

# 结合频域先验和特征增强的心脏图像分割方法

陈柯炎<sup>1</sup>, 刘巧红<sup>2\*</sup>, 韩啸翔<sup>1</sup>, 林元杰<sup>1</sup>, 张维坤<sup>1</sup>

<sup>1</sup>上海理工大学健康科学与工程学院, 上海 200093;

<sup>2</sup>上海健康医学院医疗器械学院, 上海 201318

**摘要** 针对心脏磁共振图像中的心脏子结构之间灰度差异小导致的边界不清、右心室区域形状大小多变等影响分割精度的问题, 提出一种结合频域先验知识和特征融合增强的心脏磁共振图像分割网络。所提模型是一个由频域先验引导子网络和特征融合增强子网络组合而成的 D 形结构网络。首先, 通过傅里叶变换将原始图像从空间域转换为频域, 提取出高频的边缘特征, 并将频域先验引导的子网络的低级特征与特征融合增强子网络的对应阶段进行特征拼接融合, 以提高边缘识别的能力; 其次, 在特征融合增强子网络的跳转连接处引入具有局部和全局注意力机制的特征融合模块, 提取上下文信息并获得丰富的纹理细节; 最后, 在网络底部引入 Transformer 模块, 进一步提取长距离语义信息, 增强模型表达能力, 提高分割精度。在 ACDC 数据集上的实验结果表明, 与现有方法相比, 所提方法在客观指标和视觉效果上均取得最佳的效果, 良好的心脏分割结果能为后续图像分析和临床诊断提供参考依据。

**关键词** 图像分割; 心脏磁共振图像; 频域先验; 注意力机制; 傅里叶变换

中图分类号 TP391

文献标志码 A

DOI: 10.3788/LOP231800

## Cardiac Image Segmentation by Combining Frequency Domain Prior and Feature Enhancement

Chen Keyan<sup>1</sup>, Liu Qiaohong<sup>2\*</sup>, Han Xiaoxiang<sup>1</sup>, Lin Yuanjie<sup>1</sup>, Zhang Weikun<sup>1</sup>

<sup>1</sup>School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China;

<sup>2</sup>College of Medical Instruments, Shanghai University of Medicine and Health Sciences, Shanghai 201318, China

**Abstract** A segmentation network of heart magnetic resonance image that combines prior knowledge in the frequency domain and feature fusion enhancement is proposed to address the issues of unclear boundaries caused due to the small grayscale differences among the heart substructures in heart magnetic resonance images and the varying shapes and sizes of the right ventricular region, affecting segmentation accuracy. The proposed model is a D-shaped structured network comprising a frequency domain prior guidance and feature fusion enhancer subnetworks. First, the original image is transformed from the spatial domain to the frequency domain using Fourier transform, extracting high-frequency edge features and combining the low-level features of the frequency domain prior-guided subnetwork with the corresponding stages of the feature fusion enhancement subnetwork for improving the edge recognition ability. Second, a feature fusion module with local and global attention mechanisms is introduced at the jump connection of the feature fusion enhancer network to extract contextual information and obtain rich texture details. Finally, the Transformer module is introduced at the bottom of the network to further extract long-distance semantic information, enhance the expression ability of the model, and improve segmentation accuracy. Experimental results on the ACDC dataset reveal that compared to existing methods, the proposed method achieves the best results in objective indicators and visual effects. Good cardiac segmentation results can provide reference for subsequent image analysis and clinical diagnosis.

**Key words** image segmentation; cardiac magnetic resonance image; frequency domain prior; attention mechanism; Fourier transform

收稿日期: 2023-07-31; 修回日期: 2023-10-18; 录用日期: 2023-10-25; 网络首发日期: 2023-11-15

基金项目: 国家自然科学基金(61801288)

通信作者: hqllqh@163.com

# 1 引言

心血管疾病(CVD)成为了人类健康的头号杀手,我国 CVD 在城乡居民死亡构成比中占据首位<sup>[1]</sup>。《中国心血管健康与疾病报告 2022》<sup>[2]</sup>指出,我国 CVD 的发病率和死亡率呈不断上升的趋势,2020 年 CVD 导致的死亡占农村、城市死因的 48% 和 45.86%,每 5 例死亡中就有 2 例死于 CVD。CVD 作为目前重大的公共卫生问题之一,不仅严重地威胁着居民的身体健康,而且给家庭和社会带来了沉重的经济负担。因此,CVD 的早期诊断和治疗具有重要意义。心脏磁共振图像(MRI)作为一种非侵入式的成像技术,有助于 CVD 的及早发现与预防,其不仅可以定量评估心脏的结构和功能,还可以清晰地反映心脏运动的规律,计算射血分数、心肌质量和厚度、心室体积等诊断指标,帮助临床医生更好地分析和诊断 CVD。精确地分割出左心室、右心室和心肌等心脏子结构是对心脏磁共振图像进行后续分析的关键环节,然而由于心脏结构复杂,各个子结构之间以及子结构与周围组织之间对比度低、不同病人的器官差异性大、右心室形状大小多变等现象的存在,心脏磁共振图像分割是一项具有挑战性的任务。

近年来,深度学习技术得到迅猛的发展,由于其普遍具有自我学习能力强、泛化性能好等特点,已被广泛应用在各个领域<sup>[3-7]</sup>。Tran 等<sup>[8]</sup>首次将语义分割的经典模型全卷积神经网络(FCN)<sup>[9]</sup>用在短轴心脏 MR 图像上,分割出了左心室、心肌和右心室,但由于该模型在提取深度特征时丢失较多的空间信息,导致分割的边缘轮廓较为粗糙。U-Net<sup>[10]</sup>作为一种经典的基于编码器解码器结构的 U 形网络模型,得益于其逐步下采样、上采样操作,以及跳转连接对图像分层处理的巧妙设计,在不同类型的医学图像分割任务中备受关注。Cui 等<sup>[11]</sup>提出一种基于多尺度和深度监督的 U-Net 模型,该模型以注意力机制为指导实现多切面 MR 图像中的心脏分割,同时采用 Focal\_Tversky 损失函数解决数据不平衡的问题。Wang 等<sup>[12]</sup>基于 U-Net 结合多尺度 Inception 模块、全尺寸跳转连接和预激活模块提出 MMNet,有效地改善了 U-Net 的心脏左心室分割效果。然而基于卷积操作的模型存在着一些局限性,例如:卷积运算由于固有的局部性不能很好地处理远距离的依赖关系;跳转连接进行细节传输时过于粗糙,易产生低级特征与高级特征之间的语义鸿沟等。近来由于以自注意力机制获取全局特征的 Transformer<sup>[13]</sup>在自然语言处理领域的出色表现,其被逐渐引入医学图像分割任务上<sup>[14]</sup>,如 Chen 等<sup>[15]</sup>首次将 Transformer 引入 U-Net 结构中,提出基于 Transformer 和 CNN 的混合架构分割模型 TransUNet。它利用 Transformer 的自注意力机制在低分辨率图像中获取长距离依赖的优势,在下采样和上采样之间加入 Transformer 模块,使

原有的卷积神经网络能关注到全局上下文信息,在心脏磁共振图像的分割上取得了巨大的成功。Cao 等<sup>[16]</sup>参考 U-Net 模型的编码器解码器结构,提出完全基于 Transformer 的分割模型 Swin-UNet,该模型能够天然地提取到局部和全局语义特征信息,在心脏磁共振图像的分割中表现出良好的分割精度和鲁棒泛化能力。这些模型混合了 Transformer 在获取全局上下文信息以及卷积捕捉局部特征的优势,在心脏磁共振图像的分割上普遍取得优于 CNN 结构的性能,因此受到广泛关注。

研究表明在神经网络中引入先验信息,一定程度上能够增加分割网络的可解释性和语义可靠性,获得更加准确鲁棒的分割结果<sup>[17]</sup>。Hu 等<sup>[18]</sup>首先利用卷积神经网络进行心内膜和心外膜轮廓的粗分割,进一步构建心外膜轮廓搜索的限制区域,再结合非最大梯度抑制和区域约束动态规划进行心外膜轮廓的细化。张建伟等<sup>[17]</sup>通过在神经网络中引入基于图像空间位置的先验信息引导模型聚焦于病灶区域,相较于传统的 U-Net 模型,在肝肿瘤的分割精度上有着大幅度的提升。Zhu 等<sup>[19]</sup>针对左心室的分割,开发了一种将 V-Net 与形状先验相结合的方法,得到的精确分割结果可用于计算左心室的功能参数,具有较好的临床应用价值。总的来说,先验信息可以提供分割目标的形状、轮廓、灰度、纹理、位置等信息的分布规律,为分割目标增加了约束,可有效解决较小目标或多目标轮廓交叉等问题,在增强模型可解释性的同时提高网络分割精度。

针对心脏磁共振图像中的心脏子结构之间灰度差异小导致的边界不清、右心室区域形状大小多变等影响分割精度的问题,本文提出一种结合频域先验知识和特征融合增强的心脏磁共振图像分割网络,取名为 DNet。该网络由频域先验引导子网络和特征融合增强子网络两部分组成,两个子网络均以编码器解码器结构为基础架构进行设计。将频域先验引导子网络中提取的频域先验引入特征融合增强子网络,增加目标区域的约束信息,提高对多目标交叉边缘的识别能力。在特征融合增强子网络的跳转连接处设计特征融合增强模块,增强网络上下文信息感知能力的同时保留丰富的细节信息,从而提高模型的表达能力。所提 DNet 的主要优势如下:1)结合频域先验知识和特征融合增强的心脏磁共振图像分割网络,由基于编码器解码器结构的频域先验引导子网络和特征融合增强子网络组成,充分利用频域先验引导子网络提取的频域信息作为特征融合增强子网络的边缘信息引导,从而有效地解决心脏子结构多目标分割中边界不清的问题。2)在频域先验引导子网络中设计了高频信息提取策略,生成边缘引导先验。在原始频域图像中叠加圆形掩模,有效地去除图像中的低频成分,保留利于网络边缘信息引导的高频成分,将其作为子网络的输入,进一步提取边缘特征,并与特征融合增强分支进行加权求和来

进行先验引导,增强分割效果。3)在特征融合增强子网络的跳转连接处设计了一种新的特征融合增强模块,该模块主要由全局自注意力和局部自注意力组成,充分融合了低层特征与高层特征,不仅构建了上下文依赖关系,而且丰富了局部细节,增强了网络特征提取能力,可获得更加精准的分割结果。

## 2 心脏磁共振图像分割网络框架

### 2.1 DNet 结构

所提结合频域先验知识和特征融合增强的心脏磁共振图像分割网络 DNet 主要用于分割左心室、心肌和右心室等目标区域,其总体架构如图 1 所示。

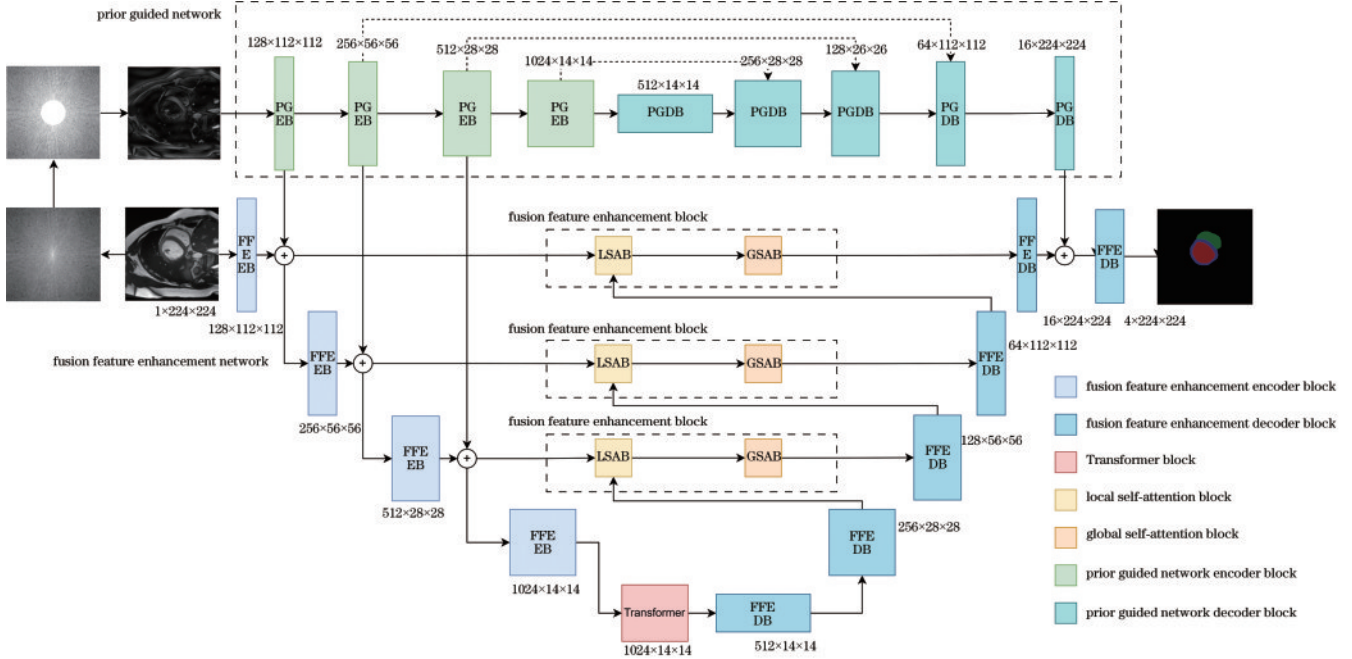


图 1 DNet 模型结构

Fig. 1 DNet model structure

所提 DNet 的上半部分是频域先验引导子网络,其基本架构基于编码器解码器结构设计。将原始心脏 MR 图像经过傅里叶变换生成频域图像后,通过先验信息提取策略保留图像中的高频成分作为网络的输入。该网络包含 4 个编码器块和 4 个解码器块,每个编码器块执行 3 个卷积操作,卷积核大小分别为  $1 \times 1$ 、 $3 \times 3$  和  $1 \times 1$ ,每次卷积后进行一个批量标准化(BN)操作,目的是减小数据偏移误差加快模型收敛,然后利用 ReLU 激活函数增强模型的非线性能力。4 个解码器均采用  $2 \times 2$  的双线性插值对输入特征图进行上采样处理,如图 2 所示,逐渐将特征图恢复到原始大小。

下半部分的特征融合增强子网络为基于 U 形结构的编码器解码器架构,其中,4 个编码器块的操作及输出特征图尺寸与频域先验引导子网络相同。将编码器的输出通过矩阵变换为一个序列数据,输入 Transformer 模块中捕获序列数据长距离依赖关系,再输入解码器。Transformer 主要由多头自注意力组成,能够计算输入序列中每个元素与其他元素的关系,并根据这些关系学习更加丰富的上下文信息,理解分布较远的器官之间的相对位置关系,从而在医学图像分割这种复杂的视觉任务中提高分割的准确性。之后经过 1 个卷积块得到特征图并输入解码器。4 个解码器同样采用  $2 \times 2$  的双线性插值,逐渐上采样恢复到原图的尺度,最后再通过 1 个卷积块输出分割预测图。在特征融合增强子网络的上采样和下采样之间且具有相同分辨率的编码器块和解码器块的跳转连接处引入新的特征融合增强模块,该模块由一个局部自注意力块和一个全局自注意力块组成,同时利用局部特征和全局长距离依赖关系学习图像的多区域特征,获取更加丰富的语义信息,提升分割效果。

将频域先验引导子网络的前 3 个编码器块提取的高频边缘特征与特征融合增强子网络相同分辨率的编码器块的特征图通过加权求和的方式进行融合。其中,频

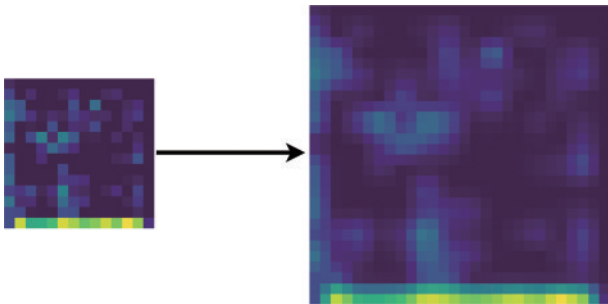


图 2 上采样过程。(a)采样前;(b)采样后

Fig. 2 Upsampling process. (a) Before sampling; (b) after sampling



域先验引导子网络的特征图权重设置为 0.3, 特征融合增强子网络的特征图权重设为 0.7。频域先验引导子网络更加关注图像的边缘信息, 利用该网络的频域先验信息引导特征融合增强子网络的边缘提取, 实现对边缘的准确识别和定位, 从而提升心脏子结构多目标分割的精度, 解决分割时边界不清的问题。

## 2.2 特征融合增强模块

在特征融合增强子网络的跳转连接处设计了特征融合增强模块, 其主要由局部自注意力模块和全局自注意力模块两部分组成。

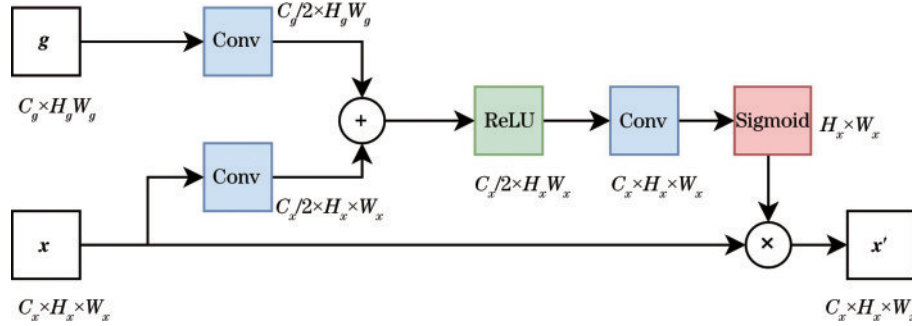


图 3 局部自注意力模块

Fig. 3 Local self-attention block

首先将来自编码器块的输入特征图  $\mathbf{x} \in \mathbf{R}^{C_x \times H_x \times W_x}$  和来自下一次解码器块的特征图  $\mathbf{g} \in \mathbf{R}^{C_g \times H_g \times W_g}$  各自进行  $1 \times 1$  的卷积运算, 使其具有相同的通道数, 并通过上采样累加。接下来依次通过 ReLU 函数、 $1 \times 1$  的卷积运算和 Sigmoid 激活函数, 得到注意力系数  $\alpha$ 。最终与输入特征图  $\mathbf{x}$  相乘, 输出具有局部注意力加权的特征块, 局部注意力的数学模型为

$$\alpha = \text{Sigmoid} \left\{ \text{Conv} \left\{ \text{ReLU} \left[ \text{Conv}(\mathbf{x}) + \text{Conv}(\mathbf{g}) \right] \right\} \right\}, \quad (1)$$

$$F(\mathbf{x}) = \mathbf{x}\alpha. \quad (2)$$

### 2.2.2 全局自注意力模块

由于心脏 MR 图像的心室和心肌区域较近, 各类别边界之间比较相似, 有效地处理全局信息能够对分割精度提升带来帮助<sup>[20]</sup>, 因此, 所提方法在局部自注意力模块后进一步利用全局自注意力模块<sup>[21]</sup>构建像素间的长距离依赖关系, 增强全局信息提取能力。全局自注意力模块由重复的两个交叉注意力模块组成, 两者共享权重, 其结构如图 4 所示。

全局注意力块主要关注的是输入特征图的上下文信息, 包括每个像素自身的特征、其周围像素的特征, 以及其在整个特征图的相对位置。具体来说, 对于每个像素, 全局注意力块通过一个新颖的交叉注意力模块获取该像素交叉路径上的邻近像素的上下文信息。这种上下文信息包括每个像素周围的特征以及其在整个特征图中的相对位置。再通过进一步的递归操作, 每个像素最终可以得到所有像素的远程依赖。而这种

### 2.2.1 局部自注意力模块

局部注意力块主要关注的是输入特征图的特定区域的空间信息, 通过上下文信息的门控信号  $\mathbf{g}$  和输入特征图  $\mathbf{x}$  确定应该关注的区域, 然后将其映射到一个新的维度空间, 并将输出压缩到通道域生成注意力系数, 这个注意力系数可以看作是一个掩码, 突出应该关注的局部区域, 同时降低对不相关区域的敏感性。局部自注意力模块接收来自对应编码器块的特征图和下一层解码器的特征图作为输入, 其结构如图 3 所示。

远程依赖实际上就是一种全局上下文信息, 它可以帮助模块理解每个像素与整个图像中其他像素之间的关系<sup>[19]</sup>。

在图 4 所示的交叉注意力模块中, 输入的特征图  $\mathbf{g} \in \mathbf{R}^{C_g \times H_g \times W_g}$  首先通过 3 个  $1 \times 1$  卷积得到  $\mathbf{Q}, \mathbf{K}, \mathbf{V}$  等 3 个参数, 其中,  $\{\mathbf{Q}, \mathbf{K}\} \in \mathbf{R}^{\frac{C_g}{8} \times H_g \times W_g}$ 、 $\mathbf{V} \in \mathbf{R}^{C_g \times H_g \times W_g}$ 。再通过相关程度计算和 Softmax 激活函数得到注意力图  $\mathbf{A} \in \mathbf{R}^{(W_g+H_g-1) \times (H_g \times W_g)}$ , 其中的相关程度计算公式为

$$d_{i,u} = \mathbf{Q}_u \mathbf{\Omega}_{i,u}^T, \quad (3)$$

式中:  $\mathbf{Q}_u \in \mathbf{R}^{\frac{C_g}{8}}$  表示在  $\mathbf{Q}$  的空间维度上每个位置  $u$  对应的向量;  $\mathbf{\Omega}_u \in \mathbf{R}^{(W_g+H_g-1) \times \frac{C_g}{8}}$  是  $\mathbf{K}$  中与位置  $u$  相同行或者列中提取的特征向量,  $\mathbf{\Omega}_{i,u} \in \mathbf{R}^{\frac{C_g}{8}}$  则表示  $\mathbf{\Omega}_u$  的第  $i$  个元素。

在得到了注意力图  $\mathbf{A}$  后, 进一步通过聚集操作捕捉长距离上下文信息, 聚集操作计算公式为

$$\mathbf{g}'_u = \sum_{i=0}^{W_g+H_g-1} A_{i,u} \mathbf{\Phi}_{i,u} + \mathbf{g}_u, \quad (4)$$

式中:  $\mathbf{g}'_u$  为特征图  $\mathbf{g}' \in \mathbf{R}^{C_g \times H_g \times W_g}$  在位置  $u$  的特征向量;  $A_{i,u}$  表示注意力图  $\mathbf{A}$  中通道  $i$ 、位置  $u$  的标量值;  $\mathbf{\Phi}_u$  是  $\mathbf{V}$  中的一个特征向量, 位于位置为  $u$  的同一行或列中,  $\mathbf{\Phi}_{i,u}$  是集合  $\mathbf{\Phi}_u$  中的第  $i$  个特征向量。

由此, 交叉注意力模块捕获了位置  $u$  在水平和垂直方向的信息, 进一步循环操作一次, 每个位置都聚合所有像素的信息, 使得分割更具备鲁棒性。

## 2.3 先验引导信息提取策略

提出一种先验引导信息提取策略用于频域先验的

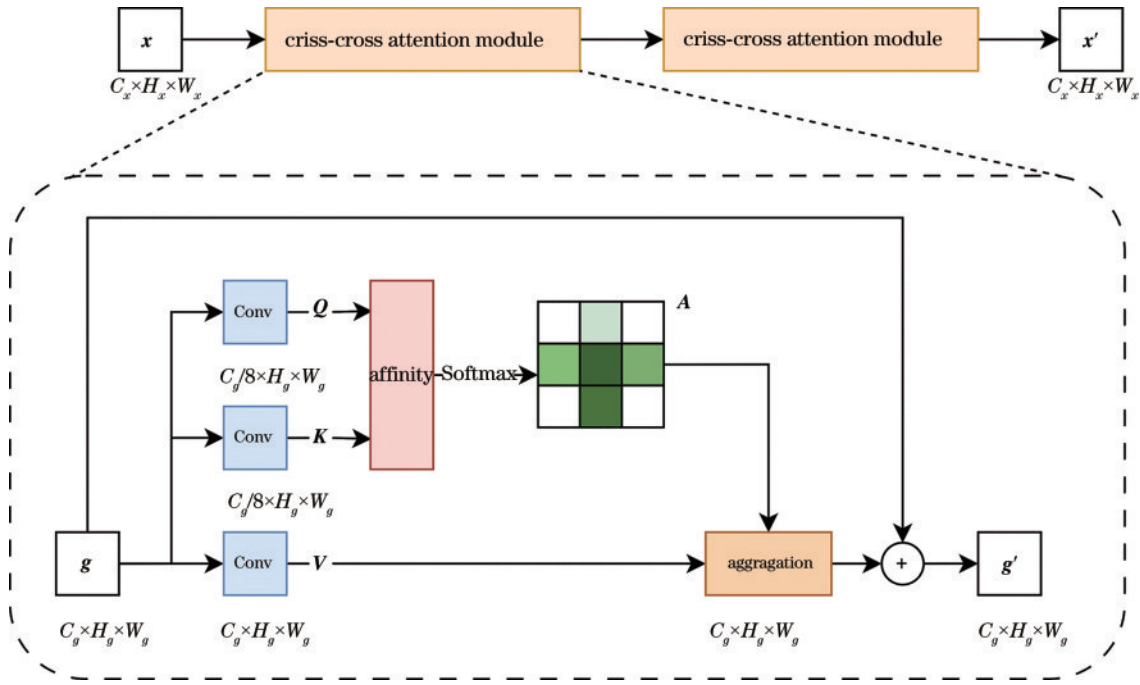


图 4 全局自注意力模块

Fig. 4 Global self-attention block

生成,为特征融合增强子网络的特征提取进行引导,为目标区域增加约束信息,提高多目标交叉边缘的识别能力,从而获得更为精确的分割边缘。具体地,在训练分割网络模型时,使用这些高频边缘信息作为先验知识,可以帮助模型更好地理解和学习图像的结构和边缘特征,从而提高分割的准确度。在图像处理中,频域图的高频信息对应的通常是图像的细节信息,包括边缘、纹理等。而边缘信息在图像分割中非常重要,可以

帮助算法区分不同的对象或区域。先验引导信息提取流程如图 5 所示,首先将原始图像通过傅里叶变换从空间域转为频域,其计算公式为

$$F_k = \sum_{n=0}^{N-1} f_n \exp\left(-\frac{i2\pi kn}{N}\right), k=0, 1, 2, \dots, N-1, \quad (5)$$

式中: $f_n$ 是原始离散信号在时空域中表示; $N$ 是信号长度。

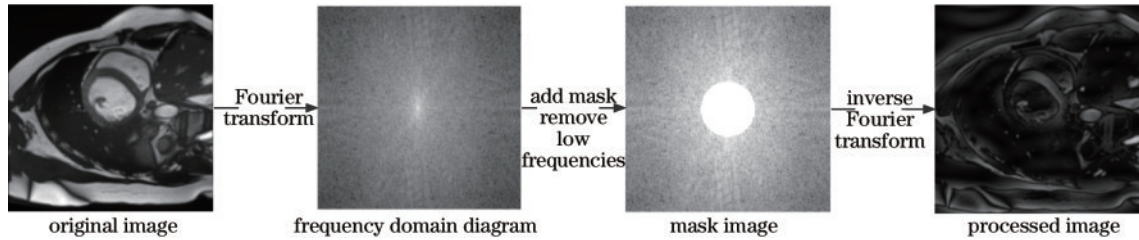


图 5 先验引导信息提取流程

Fig. 5 Prior guidance information extraction process

本文设计了一个高通滤波掩模,将傅里叶变换后的频域图像应用该掩模保留高频成分同时去除低频信号。掩模为圆形,圆形内部区域取值为 1,其余全部为 0,公式如下:

$$M(u, v) = \begin{cases} M_1(u, v) = 1, & (u - u_c)^2 + (v - v_c)^2 \leq R^2 \\ M_0(u, v) = 0, & (u - u_c)^2 + (v - v_c)^2 > R^2 \end{cases} \quad (6)$$

式中: $M_1$ 表示圆形内部区域; $M_0$ 表示其余区域; $u_c$ 和 $v_c$ 分别表示中心; $R$ 表示掩模半径。高通滤波器只允许高频信息通过,低频信息被去除,因此边缘信息可以被保留下来。经过实验计算,当半径 $R$ 取 30 pixel 时,边

缘信息保留得最好。

$$\begin{cases} F_k' = F_k(u, v) = 0 \\ F_k(u, v) \in M_1(u, v) \end{cases}, \quad (7)$$

式中: $F_k'$ 是经过掩模处理后的频域信号。

最后,通过傅里叶逆变换将经过掩模处理后的频域图像再转换回空间域,从而实现信号的重构,其数学公式为

$$f_n' = \frac{1}{N} \sum_{k=0}^{N-1} F_k' \exp\left(\frac{i2\pi kn}{N}\right), n=0, 1, 2, \dots, N-1, \quad (8)$$

式中: $f_n'$ 是经过掩模处理后的频域信号逆变换后的空间域图像。

### 3 实验结果分析与讨论

#### 3.1 实验环境及参数

实验过程中为了训练 DNet,使用配置 Intel Core i9-10920X 3.5GHz 12 核处理器、单张 12 GB 显存 NVIDIA GeForce RTX 2080 Ti 的计算设备,软件开发环境为 Python 3.9、PyTorch 1.13.0 和 CUDA 11.7。DNet 的训练中,输入图像的分辨率统一为  $224 \times 224$ ,损失函数使用交叉熵损失函数和 Dice 损失函数的加权和。模型的损失函数可以表示为

$$L_{\text{total}} = 0.3 \times L_{\text{cross\_entropy}} + 0.7 \times L_{\text{Dice}} \quad (9)$$

实验采用权重衰减自适应矩阵估计(AdamW)优化器,学习率设置为  $1 \times 10^{-3}$ ,批处理量(batch size)设置为 10,模型迭代次数为 600。

#### 3.2 数据集

所使用的心脏磁共振图像分割数据集来自 2017 年国际医学图像计算和计算机辅助干预会议举办的自动心脏诊断挑战赛(ACDC)<sup>[22]</sup>。该数据集由两台磁场强度不同的西门子磁共振扫描仪采集到的 DICOM 格式的数据组成,包括心脏收缩期和舒张期 2 个周期。该数据集共包含 100 个样本,有临床医生分割标注的 ground truth(GT),其中,70 个作为训练集,10 个作为验证集,20 个作为测试集。每个样本的标签包括左心室(LV)、右心室(RV)和心肌(MYO),每像素的空间分辨率在  $1.37 \sim 1.68 \text{ mm}^2$  之间。

#### 3.3 评价指标

为了验证所提网络的可行性和有效性,采用 4 个性能指标对模型分割结果进行评估,分别是平均相似系数(mDice)、对称位置的平均表面距离(mASD)、平均豪斯多夫距离(mHD)和平均雅卡尔指数(mJC)。

其中,mDice 是 Dice 系数的平均值,用于衡量模型分割结果和标注的分割结果之间的相似程度,Dice 系数的值越大,说明分割结果的精度越高。

$$\text{Dice}(A, B) = 2 \frac{|A \cap B|}{|A| + |B|}, \quad (10)$$

式中: $A$  表示标注的分割结果; $B$  表示模型的分割结果; $|A \cap B|$  表示  $A$  与  $B$  完全重合的部分; $|A|$  和  $|B|$  分别表示  $A$  和  $B$  的个数。

mASD 即平均对称位置的表面距离,用于测量两个点集之间的平均差异。平均对称位置的表面距离越小,分割精度也就越高。

$$\text{ASD} = \frac{1}{S(A) + S(B)} \times \left\{ \sum_{s_A \in S(A)} d[s_A, S(B)] + \sum_{s_B \in S(B)} d[s_B, S(A)] \right\}. \quad (11)$$

mHD 即豪斯多夫距离的平均值,用于衡量两组不同点集之间的最大差异,豪斯多夫距离越接近 0,说明模型分割结果的精度越高。

$$\text{HD}(A, B) =$$

$$\left\{ \max_{a \in A} \left\{ \min_{b \in B} [d(a, b)] \right\}, \max_{b \in B} \left\{ \min_{a \in A} [d(a, b)] \right\} \right\}, \quad (12)$$

式中: $d(a, b)$  表示  $a$  和  $b$  之间的欧氏距离。

mJC 即平均雅卡尔指数,用于衡量两个不动点集之间的不相似度,雅卡尔指数越大,表示两个集合的重叠部分越大。

$$\text{JC}(A, B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}, \quad (13)$$

式中: $|A \cap B|$  表示  $A$  和  $B$  完全重合的部分; $|A \cup B|$  表示  $A$  和  $B$  的并集。

#### 3.4 实验结果与分析

##### 3.4.1 分割效果

为了验证 DNet 的性能和分割效果,将 DNet 分别与 FCN、U-Net、TransUNet、Swin-UNet 等网络模型在 ACDC 心脏磁共振数据集上进行比较。表 1 是不同方法在 4 种定量指标上的数值结果对比,具体给出了在左心室、心肌和右心室的 mDice 指标。从表中可以看出,所提分割模型 DNet 在各项指标上均达到了最优,表明模型有很强的学习能力,能够有效地提升心脏图像的分割精度。具体而言,DNet 的 mDice 系数达到了 91.36%,比 FCN、U-Net、TransUNet 和 Swin-UNet 分别提升 3.71 百分点、3.05 百分点、1.8 百分点和 3.8 百分点,mJC、mASD 和 mHD 指标同样也取得了最优的性能指标,这主要得益于频域先验引导策略对于边缘提取能力的加强以及特征融合增强模块对于模型表达能力的提升。

表 1 不同方法在 ACDC 数据集上分割效果对比

Table 1 Comparison of segmentation effects of different methods on ACDC dataset

Method	mDice				mJC	mASD	mHD
	Average	MYO	RV	LV			
FCN	87.65	86.57	83.51	92.87	0.78	0.59	3.06
U-Net	88.31	85.97	85.03	93.94	0.79	0.56	2.84
TransUNet	89.56	86.42	87.54	94.71	0.81	0.54	2.79
Swin-UNet	87.56	82.97	85.69	94.03	0.79	0.42	1.92
DNet	<b>91.36</b>	<b>89.07</b>	<b>89.37</b>	<b>95.64</b>	<b>0.84</b>	<b>0.34</b>	<b>1.62</b>

图 6 展示了所提 DNet 与其他模型在 ACDC 数据集上的分割结果可视化比对。其中,绿色表示右心室,蓝色表示心肌,红色表示左心室。

从图 6 可以看出,对比方法中出现了分割冗余或分割不全的情况,而所提 DNet 的分割效果更加接近于真实标签,通过先验引导模块的边缘引导以及特征融合增强的特征细化,改善了分割冗余或分割不全的情



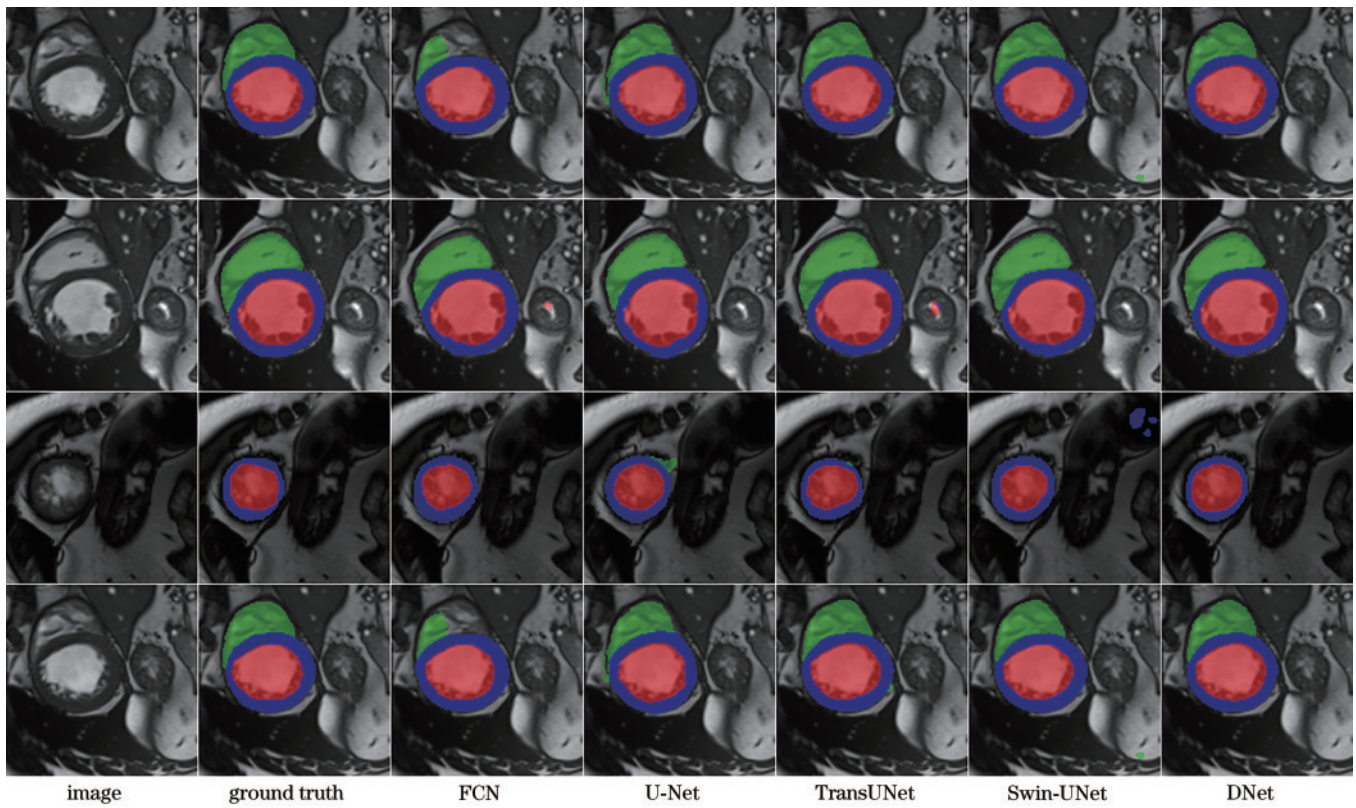


图 6 分割结果对比图

Fig. 6 Comparison of segmentation results

况,有效提取了多目标区域间的交叉边缘,使得分割形状更加清晰,边缘更加平滑,同时保留更多的细节信息。

为了验证先验引导策略的效果,设计了对比实验,结果如表 2 所示。

表 2 不同先验引导策略的分割效果对比

Table 2 Comparison of segmentation effects of different prior guided strategies

Strategy	mDice	mJC	mASD	mHD
Canny	86.70	0.78	3.54	25.43
Sobel	89.40	0.82	0.48	2.42
Fourier	<b>91.36</b>	<b>0.84</b>	<b>0.34</b>	<b>1.62</b>

从实验结果可以看出,傅里叶域的高频提取策略的效果更好一些。为了对比不同特征融合增强方法的效果,对比了 U-Net++ 的密集连接增强、RefineNet 的特征细化增强,以及 DNet 的特征融合增强的分割效果。实验结果如表 3 所示。

表 3 不同特征融合增强策略的分割效果对比

Table 3 Comparison of segmentation effects of different feature fusion enhancement strategies

Strategy	mDice	mJC	mASD	mHD
U-Net++	89.43	0.81	0.44	2.33
RefineNet	90.19	0.83	0.54	3.26
DNet	<b>91.36</b>	<b>0.84</b>	<b>0.34</b>	<b>1.62</b>

从实验结果可以看出,特征融合增强的效果更好一些。

### 3.4.2 消融实验

为了验证先验信息引导和特征融合增强的有效性,进行了消融实验。为了观察模型中的模块对模型性能的影响,验证这些模块的有效性,消融实验通过分别去除先验引导子网络模块(PGN)和特征融合增强模块(FFEB)来进行模块性能验证。

进行消融实验时,通过逐步去除模型的某些部分进而得到 3 个新的模型: DNet-PGN-FFEB(去除了 PGN 和 FFEB)、DNet-PGN(去除了 PGN)和 DNet-FFEB(去除了 FFEB)。

接着进行性能比较,使用相同的数据集和评价指标来比较这些模型的性能。实验结果如表 4 所示,从 mDice 性能指标来看, DNet 相比于 DNet-PGN-FFEB、DNet-PGN 和 DNet-FFEB 分别提升 1.57 百分点、1.05 百分点和 1.31 百分点。此外,在 mJC、mASD 和 mHD 等 3 个指标上, DNet 同样也取得最高的数值。

表 4 消融实验结果

Table 4 Results of ablation experiments

Method	mDice	mJC	mASD	mHD
DNet-PGN-FFEB	89.79	0.82	0.62	2.51
DNet-PGN	90.31	0.82	0.48	2.54
DNet-FFEB	90.05	0.83	0.55	2.20
DNet	<b>91.36</b>	<b>0.84</b>	<b>0.34</b>	<b>1.62</b>

实验结果进一步证明了DNet中先验引导子网络和特征融合增强模块的有效性,在一定程度上有助于模型性能的提升。

为了验证频域先验引导子网络的低级特征与特征融合增强子网络的对应阶段融合的有效性,进行拼接融合前后效果验证实验。实验结果如表5所示,其中,not concatenated表示不与低级特征进行特征拼接融合,concatenated表示与低级特征进行特征拼接融合。

表5 先验特征拼接实验结果  
Table 5 Results of prior feature concatenation

Method	mDice	mJC	mASD	mHD
not concatenated	87.41	0.78	0.53	2.63
concatenated	<b>91.36</b>	<b>0.84</b>	<b>0.34</b>	<b>1.62</b>

可以看出,特征拼接后4个指标更好,结合图7拼接融合的梯度加权类激活热图可更好地说明频域先验引导子网络的低级特征与特征融合增强子网络的对应阶段进行特征拼接融合的有效性。

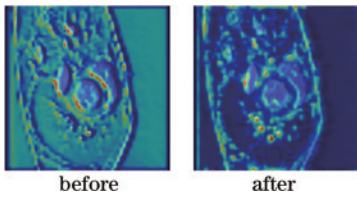


图7 拼接融合前后的特征热图

Fig. 7 Feature heatmaps before and after fusion

为了剖析局部和全局注意力,设计了一组实验来对比单独和组合使用局部和全局注意力的效果。实验结果如表6所示,其中,local attention表示只有局部注意力,global attention表示只有全局注意力,both表示结合了局部和全局注意力。

表6 局部和全局注意力单独和结合使用实验结果  
Table 6 Results of using local and global attention individually and in combination

Method	mDice	mJC	mASD	mHD
local attention	90.16	0.83	0.43	2.17
global attention	88.38	0.80	1.10	5.49
both	<b>91.36</b>	<b>0.84</b>	<b>0.34</b>	<b>1.62</b>

可以看出,结合使用局部和全局注意力在4个指标中均取得更好的表现,因此局部和全局的注意力的结合使用的效果是要好于单独使用的,图8拼接融合的梯度加权类激活热图也证明了局部和全局注意力的有效性。

在掩模形状的选取和设计时,主要以简单、常规、便于计算作为基本原则,重点考虑了圆形和方形。从

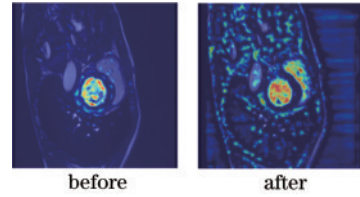


图8 拼接融合前后的特征热图

Fig. 8 Feature heatmaps before and after fusion

提高模型性能角度出发,进行了圆形掩模和方形掩模的对比实验。实验结果如表7所示,其中,round表示圆形掩模,square表示方形掩模。

表7 使用不同掩模形状的实验结果  
Table 7 Results of using different mask shapes

Shape	mDice	mJC	mASD	mHD
round	<b>91.36</b>	<b>0.84</b>	0.34	1.62
square	90.89	0.83	<b>0.30</b>	<b>1.23</b>

综合考虑,取效果稍好的圆形掩模作为原始频域图像中的覆盖区域。为了验证圆形掩模对后续的影响,设计了3组不同半径的掩模实验,实验结果如表8所示,实验结果表明,半径取30是一个适当的值。

表8 不同掩模半径的实验结果  
Table 8 Results of different mask radii

Radius	mDice	mJC	mASD	mHD
20	90.07	0.83	0.38	2.07
30	<b>91.36</b>	<b>0.84</b>	<b>0.34</b>	1.62
40	90.47	0.83	<b>0.34</b>	<b>1.39</b>

### 3.5 讨论

通过对比实验和消融实验可知,高频提取策略在引导网络模型进行分割上表现优异,原因在于其能捕捉医学图像的细节信息。结合局部和全局注意力,可以更好地捕捉图像的多维度信息,从而提升精度。在特征融合方面,频域先验引导子网络的低级特征与特征融合增强子网络的对应阶段进行特征拼接融合,效果更佳,这是由于特征拼接融合能更好地保留和利用低级和高级特征。在掩模形状和半径方面,圆形掩模更符合图像的自然特性和人眼的视觉感知。通过选择适当的掩模半径,可以保留对模型目标区域分割有利的细节信息。尽管实验结果表明了各种策略和技术在改进模型性能方面的有效性,但仍存在局限性。例如,特征拼接融合可能导致模型复杂性增加,增加了计算成本。局部和全局注意力可能导致模型过于关注某些特定的信息,而忽视了其他可能有用的信息。因此,在应用这些策略和技术时,需要根据具体的应用场景和需求来进行选择和调整。



## 4 结 论

针对心脏磁共振图像的多目标分割任务,提出一种基于频域先验知识和特征融合增强的分割网络模型 DNet,其由频域先验引导子网络和特征融合增强子网络两部分结合构成。在 ACDC 数据集上的实验结果表明,所提方法的 mDice、mJC、mASD、mHD 分别达到了 91.36%、0.84、0.34、1.62,均高于对比方法,表明所提频域先验特征能够增强网络的学习能力,这是由于其通过提取边缘信息,减小了噪声。特征融合模块能够优化上下采样路径,从而进一步提高网络的分割精度,局部和全局注意力机制能够获取低维更丰富的纹理细节以及多个感兴趣区域的特征信息。未来的工作将考虑通过学习医生的临床经验,引入临床先验信息,增加模型的可解释性,提高分割精度更好地服务于临床应用。

### 参 考 文 献

- [1] 马丽媛,王增武,樊静,等.《中国心血管健康与疾病报告 2021》关于中国高血压流行和防治现状[J]. 中国全科医学, 2022, 25(30): 3715-3720.  
Ma L Y, Wang Z W, Fan J, et al. Epidemiology and management of hypertension in China: 《an analysis using data from the annual report on cardiovascular health and diseases in China(2021)》[J]. Chinese General Practice, 2022, 25(30): 3715-3720.
- [2] 中国心血管健康与疾病报告编写组. 中国心血管健康与疾病报告 2022 概要[J]. 中国循环杂志, 2023, 38(6): 583-612.  
The Writing Committee of the Report on Cardiovascular Health, China D I. Report on cardiovascular health and diseases burden in China: an updated summary of 2022 [J]. Chinese Circulation Journal, 2023, 38(6): 583-612.
- [3] Wu F, Yang X, Ma Y, et al. Machine-learning guided optimization of laser pulses for direct-drive implosions[J]. High Power Laser Science and Engineering, 2022, 10: e12.
- [4] Lin J P, Haberstroh F, Karsch S, et al. Applications of object detection networks in high-power laser systems and experiments[J]. High Power Laser Science and Engineering, 2023, 11: e7.
- [5] Andreas D, Christoph E, Sunny H, et al. Data-driven science and machine learning methods in laser-plasma physics[J]. High Power Laser Science and Engineering, 2023, 11: e55.
- [6] 撒子奇,刘巧红,凌晨,等.结合 HarDNet 和反向注意力的息肉分割方法[J]. 激光与光电子学进展, 2023, 60(2): 0217003.  
Han Z Q, Liu Q H, Ling C, et al. Polyp segmentation method combining HarDNet and reverse attention[J]. Laser & Optoelectronics Progress, 2023, 60(2): 0217003.
- [7] 张欢,仇大伟,冯毅博,等. U-Net 模型改进及其在医学图像分割上的研究综述[J]. 激光与光电子学进展, 2022, 59(2): 0200005.  
Zhang H, Qiu D W, Feng Y B, et al. Improved U-net models and its applications in medical image segmentation: a review[J]. Laser & Optoelectronics Progress, 2022, 59(2): 0200005.
- [8] Tran P V. A fully convolutional neural network for cardiac segmentation in short-axis MRI[EB/OL]. (2016-04-02)[2023-04-05]. <https://arxiv.org/abs/1604.00494>.
- [9] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[C]//IEEE Transactions on Pattern Analysis and Machine Intelligence, May 24, 2016, New York: IEEE Press, 2016: 640-651.
- [10] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [11] Cui H F, Chang Y W, Jiang L, et al. Multiscale attention guided U-Net architecture for cardiac segmentation in short-axis MRI images[J]. Computer Methods and Programs in Biomedicine, 2021, 206: 106142.
- [12] Wang Z Y, Peng Y J, Li D P, et al. MMNet: a multi-scale deep learning network for the left ventricular segmentation of cardiac MRI images[J]. Applied Intelligence, 2022, 52(5): 5225-5240.
- [13] 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, December 4-9, 2017, Long Beach, California, USA. New York: ACM Press, 2017: 6000-6010.
- [14] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16 x 16 words: transformers for image recognition at scale[EB/OL]. (2021-10-22) [2022-09-16]. <https://arxiv.org/abs/2010.11929>.
- [15] 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) [2023-03-05]. <https://arxiv.org/abs/2102.04306>.
- [16] Cao H, Wang Y Y, Chen J, et al. Swin-UNet: unet-like pure transformer for medical image segmentation[M]//Karlinsky L, Michaeli T, Nishino K. Computer vision-ECCV 2022 workshops. Lecture notes in computer science. Cham: Springer, 2023, 13803: 205-218.
- [17] 张建伟,张旭斌,徐宇扬,等.融合空间先验的医学影像分割神经网络模型[J]. 计算机辅助设计与图形学学报, 2021, 33(8): 1287-1294.  
Zhang J W, Zhang X B, Xu Y Y, et al. Spatial prior-embedded neural networks for medical image segmentation[J]. Journal of Computer-Aided Design & Computer Graphics, 2021, 33(8): 1287-1294.
- [18] Hu H F, Pan N, Wang J Y, et al. Automatic segmentation of left ventricle from cardiac MRI via deep learning and region constrained dynamic programming[J].

- Neurocomputing, 2019, 347: 139-148.
- [19] Zhu F B, Li L X, Zhao J Y, et al. A new method incorporating deep learning with shape priors for left ventricular segmentation in myocardial perfusion SPECT images[J]. Computers in Biology and Medicine, 2023, 160: 106954.
- [20] 李擎, 皇甫玉彬, 李江昀, 等. UConvTrans: 全局和局部信息交互的双分支心脏图像分割[J]. 上海交通大学学报, 2023, 57(5): 570-581.
- Li Q, Huangfu Y B, Li J Y, et al. UConvTrans: dual-branch cardiac image segmentation with global and local information interaction[J]. Journal of Shanghai Jiao Tong University, 2023, 57(5): 570-581.
- [21] Huang Z L, Wang X G, Huang L C, et al. CCNet: criss-cross attention for semantic segmentation[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Republic of Korea. New York: IEEE Press, 2020: 603-612.
- [22] Bernard O, Lalande A, Zotti C, et al. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? [J]. IEEE Transactions on Medical Imaging, 2018, 37(11): 2514-2525.