

## • 综 述 •

# 探析 U-Net 的改进机制及其在医学图像分割的应用



周涛<sup>1,2</sup>, 侯森宝<sup>1</sup>, 陆惠玲<sup>3</sup>, 赵雅楠<sup>1</sup>, 党培<sup>1</sup> 综述 董雅丽<sup>1</sup> 审校

1. 北方民族大学 计算机科学与工程学院 (银川 750021)

2. 北方民族大学 图像图形智能处理国家民委重点实验室 (银川 750021)

3. 宁夏医科大学 理学院 (银川 750004)

**【摘要】** U-Net 网络在医学图像分割任务中取得了很好的成绩。近年来, 众多学者针对 U-Net 结构不断地进行研究和扩展, 比如编、解码器的改进和跳跃连接的改进。本文针对基于 U-Net 网络结构改进的医学图像分割技术从以下角度进行总结: 首先, 阐述 U-Net 网络在医学图像分割领域中的应用; 然后, 总结 U-Net 的七大改进机制: 密集连接机制、残差连接机制、多尺度机制、集成机制、膨胀机制、注意力机制以及 Transformer 机制; 最后, 探讨 U-Net 结构改进的思路和方法, 为相关研究提供参考, 对 U-Net 的进一步发展具有一定的积极意义。

**【关键词】** U-Net; 医学图像; 语义分割; 改进机制; 网络结构

## Exploring and analyzing the improvement mechanism of U-Net and its application in medical image segmentation

ZHOU Tao<sup>1,2</sup>, HOU Senbao<sup>1</sup>, LU Huiling<sup>3</sup>, ZHAO Yanan<sup>1</sup>, DANG Pei<sup>1</sup>, DONG Yali<sup>1</sup>

1. School of Computer Science and Engineering, North Minzu University, Yinchuan 750021, P. R. China

2. Key Laboratory of Image and Graphics Intelligent Processing of State Ethnic Affairs Commission, North Minzu University, Yinchuan 750021, P. R. China

3. School of Science, Ningxia Medical University, Yinchuan 750004, P. R. China

Corresponding author: HOU Senbao, Email: hsb378093739@163.com

**【Abstract】** Remarkable results have been realized by the U-Net network in the task of medical image segmentation. In recent years, many scholars have been researching the network and expanding its structure, such as improvement of encoder and decoder and improvement of skip connection. Based on the optimization of U-Net structure and its medical image segmentation techniques, this paper elucidates in the following: First, the paper elaborates on the application of U-Net in the field of medical image segmentation; Then, the paper summarizes the seven improvement mechanism of U-Net: dense connection mechanism, residual connection mechanism, multi-scale mechanism, ensemble mechanism, dilated mechanism, attention mechanism, and transformer mechanism; Finally, the paper states the ideas and methods on the U-Net structure improvement in a bid to provide a reference for later researches, which plays a significant part in advancing U-Net.

**【Key words】** U-Net; Medical image; Semantic segmentation; Improvement mechanism; Network structure

## 引言

随着现代医学技术的快速发展, 医学图像已经成为医生进行疾病术前诊断规划、手术方案制定、

术中检测与术后评估等许多医疗活动的重要信息来源, 通过医学图像医生可以更直观、全面地了解病变的结构和特征, 提高诊断效率和正确性。医学图像分割能够从特定组织图像中提取关键信息, 是

DOI: 10.7507/1001-5515.202111010

基金项目: 国家自然科学基金项目 (62062003); 宁夏回族自治区重点研发计划项目 (引才专项) (2020BEB04022); 北方民族大学引进人才科研启动项目 (2020KYQD08); 北方民族大学研究生创新项目 (YCX21089); 宁夏自然科学基金项目 (2022AAC03149)

通信作者: 侯森宝, Email: hsb378093739@163.com



实现医学图像可视化的关键步骤,分割后的图像被提供给医生用于定量分析、诊断、解剖结构的描绘、病理改变组织的定位、制订治疗计划等不同任务。但医学图像信息量巨大,临床上手工勾画医学图像目标区域是一件费时费力的工作,给临床医生的工作增加了很大负担,故医学图像分割仍然是一个难题。在过去的几年里,深度学习在医学图像处理的应用中引起了人们极大的兴趣。迄今为止,最成功的图像分析方法是卷积神经网络(convolutional neural network, CNN),CNN<sup>[1]</sup>的突破对于医学图像分割领域非常重要。CNN经典模型有LeNet<sup>[2]</sup>、AlexNet<sup>[3]</sup>和U-Net<sup>[4]</sup>,其中U-Net取得了突出的成就<sup>[5]</sup>,成为医学图像分割的热门技术,在深度卷积神经网络的发展中具有非常重要的意义。为了解决医学图像分割中的问题,许多研究人员针对U-Net网络结构进行了研究。Millemari等<sup>[6]</sup>提出的V-Net用于三维(three-dimensional, 3D)医学图像分割,在输入3D图像按照通道拆分的同时,在每一层加入残差结构,以快速和精确的方式实现磁共振成像中前列腺体积的分割。Zhou等<sup>[7]</sup>提出U-Net++网络用于医学图像分割,网络中增加新的跳跃连接,以便在每个深度之间传输更多信息。Oktay等<sup>[8]</sup>提出Attention U-Net网络用于胰腺分割,将注意力模块引入U-Net中的跳跃连接使得图像信息更加清晰,提高分割的准确性。Ma等<sup>[9]</sup>提出多尺度水平跳跃连接U-Net(multi-scale hybrid skip U-Net, MHSU-Net)网络用于肺分割,在跳跃连接中加入上下文模块,该架构提高了分割模型的特征提取能力,有效地减少了分割过程中的特征丢失。Peng等<sup>[10]</sup>提出局部上下文感知网络(locally context-perception Net, LCP-Net),通过并行膨胀卷积提高了特征编码器上下文信息的利用率,有利于获得特征图的丰富信息。Chen等<sup>[11]</sup>提出了交叉尺度残差网络(cross-scale residual network, CSR-Net)用于舟骨骨折分割,利用不同层的特征通过跨尺度残差连接实现融合,实现不同层特征之间的尺度和通道转换,可以在不同的比例特征之间建立紧密的联系。由此可见,U-Net在医学图像分割中已经有了广泛的应用。

本文从网络结构出发对U-Net进行总结,先阐述U-Net在医学图像分割领域中的应用,再综述U-Net的七大改进机制(密集连接机制、残差连接机制、多尺度机制、集成机制、膨胀机制、注意力机制和Transformer机制),最后总结与展望U-Net未来的发展方向。

## 1 U-Net网络在医学图像分割中的应用

U-Net是一种端到端神经网络架构<sup>[12]</sup>,以其“U”形而闻名,U-Net网络以交并比(Intersection over Union, IOU)值达到了92%的优势,在ISBI 2014年胶质母细胞瘤分割挑战赛中名列前茅。在2016年,Cicek等<sup>[13]</sup>进一步将U-Net的二维网络推广到三维网络,并在MICCAI 2016中利用3D-Unet从稀疏注释中分割密集体积图像。由于其卓越的性能和优良的体系结构,U-Net迅速成为医学图像分割中的重要网络模型。

### 1.1 U-Net基本原理

U-Net采用的结构如图1所示,网络首先对输入图像进行4组卷积和下采样操作来获得图像特征信息,然后通过右侧的4组反卷积和上采样将图像放大。U-Net相同通道数的下采样和上采样操作之间增加了跳跃连接(如图中水平黄色箭头所示),能帮助解码器更好地修复目标细节。左侧卷积块的输出拼接到右侧相同深度反卷积块的输入,反卷积块的输入特征图尺寸没有变化,但通道数增大了一倍。

### 1.2 U-Net在医学图像分割领域中的应用

U-Net网络由于其结构性能卓越,在医学图像分割领域被广泛应用,目前在脑肿瘤、肺癌、肝肿瘤、乳腺癌等重大疾病的临床辅助诊断方面取得良好效果。如表1所示,从疾病分类、网络名称、时间、模型维度、数据集五方面对U-Net在医学图像分割中的应用进行归纳和总结。

脑肿瘤分为原发性和继发性肿瘤,由于脑病变结构的高可变性,组织边界模糊,以及数据有限和类不平衡性,使脑肿瘤的分割面临巨大的挑战。针对这些挑战,一些专家学者提出以下解决方法:Kumar等<sup>[14]</sup>提出分类器-分割器网络(Classifier-Segmenter network, CSNet)用于自动执行急性脑卒中诊断任务,以辅助和加快医务人员的决策过程,并进一步提高了模型的分割精度;Wang等<sup>[15]</sup>提出一种跨水平连接跨层连通U-Net(cross-level connected U-shaped network, CLCU-Net)用于从多模态磁共振图像中自动准确分割脑肿瘤,充分利用不同大小脑肿瘤的不同尺度信息,以获得更有效的特征连接,提取更多的判别信息。

肺癌的早期发现有助于尽早治疗,计算机断层扫描(computed tomography, CT)是诊断肺部疾病的首选方法。然而,肺癌患者的病灶区域通常和正常组织粘连,现有的方法不能准确地进行肺部病灶

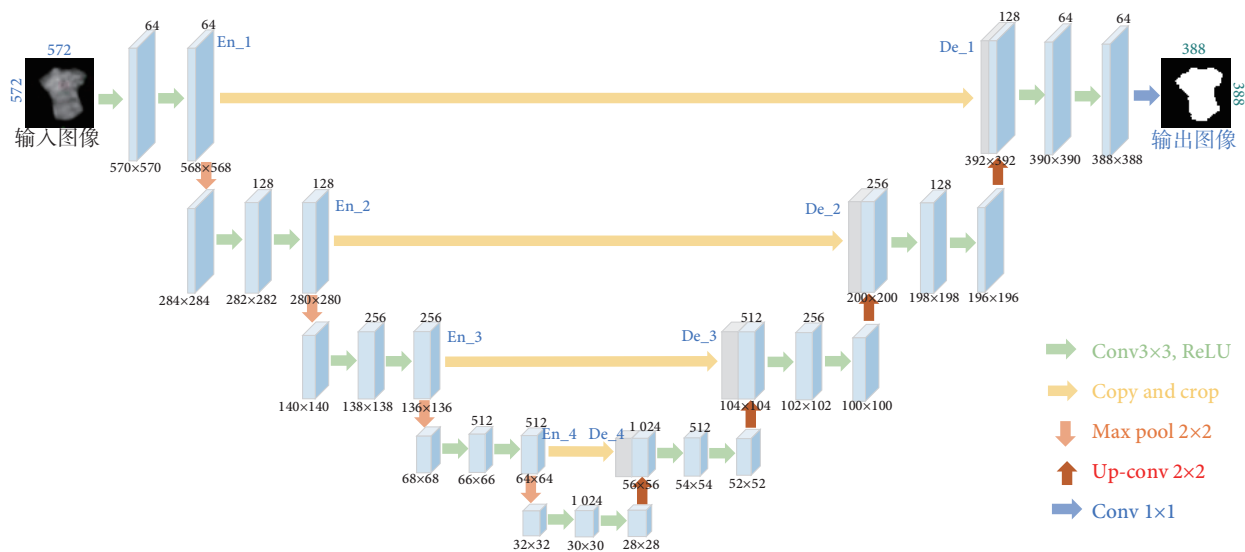


图1 U-Net 基本架构

Fig.1 U-Net basic architecture

区域的分割。一些专家学者提出以下解决方案：Zhou 等<sup>[16]</sup> 提出注意力并行 U-Net 网络 (attention mechanism parallel U-Net, APU-Net)，该网络结合正电子发射计算机断层扫描 (positron emission tomography, PET)、CT 和 PET/CT 三种模态的肺部肿瘤特征，并使用混合注意力机制聚焦病灶区域，提高了肺部肿瘤分割的精度；Zhao 等<sup>[17]</sup> 提出一种新的级联两阶段 U-Net 模型 (Distraction-Sensitive U-Net, DSU-Net)，第一阶段对整个输入 CT 体积进行全局分割，并预测潜在的干扰区域，其中包含假阴性区域和假阳性区域，第二阶段将分散区域信息转化为局部分割用于体积块，以进一步分割肺部肿瘤区域。

肝脏具有丰富的血流供应，与人体的重要血管关系密切且肝脏恶性肿瘤发病隐匿、生长快速，因此治疗甚为困难。在临床实践中，医生通常采用手工分割的方法从腹部 CT 图像中获取肝脏区域，但是该方法耗时、劳动强度大、操作繁琐。为了更准确地分割肝脏区域，一些专家学者提出以下解决方案：Xie 等<sup>[18]</sup> 提出动态自适应残差网络 (dynamic adaptive residual network, DAR-Net)，此网络去除不相关像素并采用条件随机场对肝脏边界和纹理进行优化，提升了对肝脏病灶的分割精度；Jiang 等<sup>[19]</sup> 提出注意力与混合连接网络 (Attention Hybrid Connection Network, AHCNet)，该网络结合了 UNet++、注意力机制和跳跃连接，实现更快的网络收敛和准确的语义分割，并进一步证明了所提出的方法对肝脏病灶分割具有良好的临床价值。

乳腺癌是乳腺上皮细胞在多种致癌因子的作

表1 基于 U-Net 的医学图像分割应用

Tab.1 Medical image segmentation application based on U-Net

疾病分类	网络名称	时间	模型维度	数据集
脑肿瘤	CSNet <sup>[14]</sup>	2021	2D	SISS
	CLCU-Net <sup>[15]</sup>	2021	2D	BraTS 2018
肺癌	APU-Net <sup>[16]</sup>	2022	2D	90例临床患者
	DSU-Net <sup>[17]</sup>	2022	2D	MICCAI2019
肝肿瘤	DAR-net <sup>[18]</sup>	2021	3D	3DIRCADb
	AHCNet <sup>[19]</sup>	2019	3D	3DIRCADb
乳腺癌	U-Net MALF <sup>[20]</sup>	2021	2D	DDSM
	DDTNet <sup>[21]</sup>	2022	2D	BCa-lym

用下，发生增殖失控的病变。乳腺 X 线摄影是早期发现和诊断乳腺癌的标准筛查方式，但是在发现致密和极度致密乳房的癌症病灶方面仍然无效。采用计算机视觉分割方法可以帮助放射科医生进行辅助诊断，例如：Belh 等<sup>[20]</sup> 提出结合残差块和混合注意力损失函数的改进 U-Net 乳腺肿瘤分割方法，利用残差卷积模块和扩展残差卷积模块代替编码路径的卷积模块，在不增加计算代价的情况下提取超声乳腺肿瘤更详细的特征；Zhang 等<sup>[21]</sup> 提出密集双任务网络 (dense dual-task network, DDTNet)，采用有效的特征融合策略引入癌细胞位置信息的多尺度特征，以实现精确的乳腺癌病灶分割。

## 2 改进机制

近年来，U-Net 得到了长远的发展，从 U-Net 与其他深度模型耦合的角度出发，本文详细总结了基于 U-Net 的七大改进机制。如图 2 所示，七大机制具体包括：密集连接机制、残差连接机制、

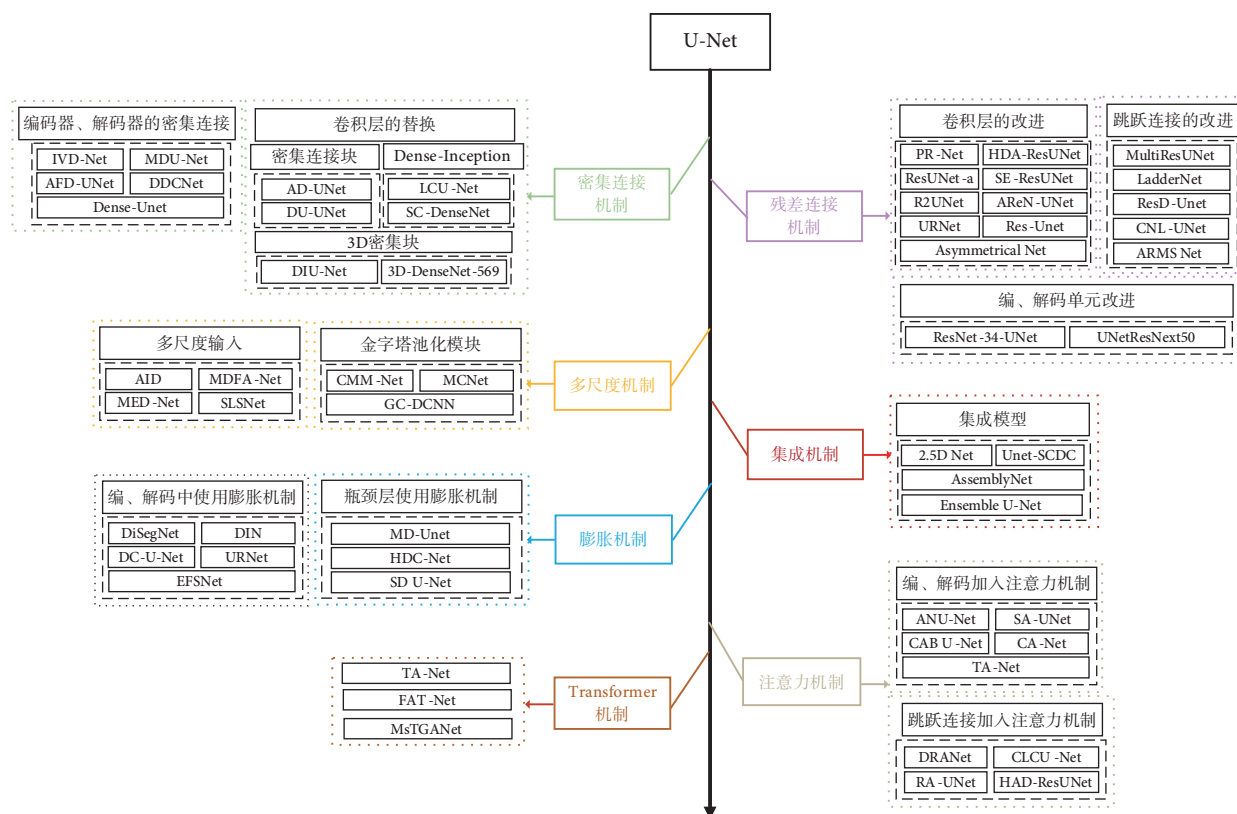


图2 改进机制总框图

Fig.2 General block diagram of the improvement mechanism

多尺度机制、集成机制、膨胀机制、注意力机制和Transformer 机制。

## 2.1 密集连接机制

U-Net 中的密集连接机制就是将卷积层或者编、解码器部分使用密集连接<sup>[22]</sup>。图3给出基于U-Net 的密集连接机制典型代表 DenseUNet<sup>[23]</sup>, 该网络由三部分组成: 用于特征提取的下采样过程、用于增加网络深度的密集块和用于图像重建的上采样过程。DenseNet<sup>[24]</sup> 通过增加网络深度有效提取面部特征, 而 U-Net 则通过跳跃连接保留重要的面部细节, 引入密集块来提高网络拟合能力, 从而较好地解决了医学图像细节的丢失。U-Net 中密集连接机制的特点是: 任何两层之间都存在直接连接, 这意味着每一层网络的输入是所有网络输出的并集之前的层, 以及这一层学习到的特征图是作为输入直接传递给它后面的所有层, 不仅可以减轻梯度消失问题, 增强了特征传播, 鼓励特征重用, 并大大减少了参数量。

密集连接神经网络 (Densely Connected Convolutional Networks, DenseNet) 主要包含卷积层、密集块、过渡层和分类器, 任一卷积层的输入包含前面所有卷积层的输出, 使得特征充分复用,

并且高低层次的特征融合使得网络具有较强的抗过拟合性能。近年来许多学者在 U-Net 中引入密集连接机制: Li 等<sup>[25]</sup> 结合 U-Net 和密集跳跃连接的嵌套分割网络 (attention-based nested U-Net, ANU-Net) 获得不同语义层次的全分辨率特征图; Tang 等<sup>[26]</sup> 提出基于双重注意力的密集网络 (Dual Attention-based Dense SU-net, DA-DSUnet), 用密集块替换了 U-Net 中的卷积层使得分割精度和参数效率均为更优。本节将详细归纳 U-Net 的密集连接机制, 从卷积层的替换以及编、解码器中加入密集连接两个方面进行总结。

**2.1.1 卷积层的替换** 卷积层的替换是指 U-Net 将原卷积层替换为密集块。在密集块中, 任意两层之间存在直接连接, 意味着网络每一层的输入都是所有前一层输出的并集, 该层学习到的特征图信息作为输入直接传递给它后面的所有层。本节从原卷积层替换成密集连接块、3D 密集块与 Dense-Inception 模块三部分进行介绍。如图4所示, 箭头按照时间先后顺序的方式从左到右依次呈现, 每种背景颜色分别代表着一种密集块。

第一, 将传统的密集块引入到 U-Net 中, 提高了传输效率, 使网络提取的特征信息更加清晰, 从



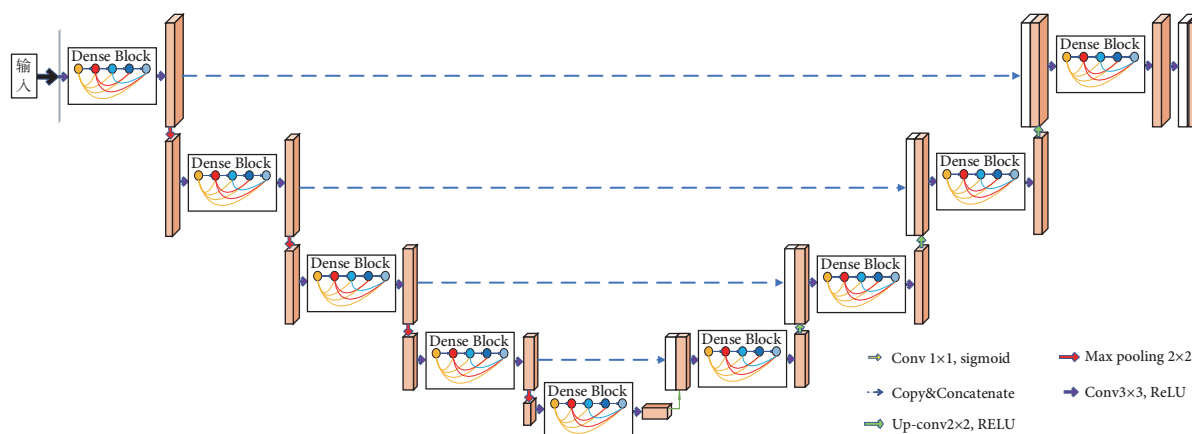


图3 基于 U-Net 的密集连接机制典型代表

Fig.3 Representative of dense connection mechanism based on U-Net

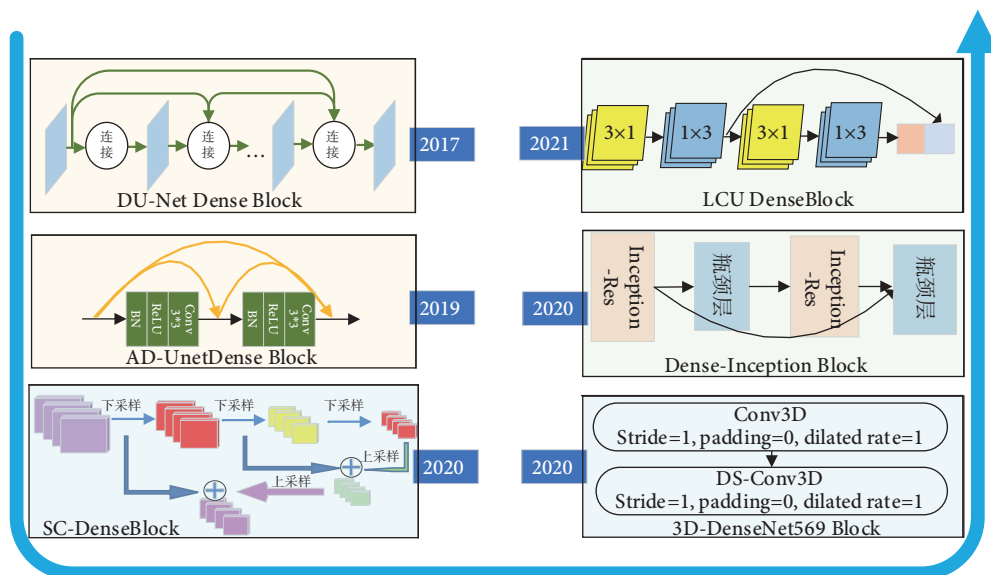


图4 密集连接机制中卷积层的替换

Fig.4 Replacement of convolutional layer in dense connection mechanism

而较好地解决了精细图像细节的丢失问题,提高了网络分割的准确性。将卷积层替换为密集块主要方法有 DU-Net 和 AD-Unet 两种模型: Manal 等<sup>[27]</sup>提出密集块 U-Net (Denseblock U-Net, DU-Net), 此网络将卷积层替换为密集块,该模型优于其他相关的深度学习模型,并改善了梯度流动性; Luo 等<sup>[28]</sup>提出注意力密集网络 (Attention-Dense-UNet, AD-Unet), 该网络在卷积层的基础上加入密集连接,较好地解决了精细血管细节的丢失问题。

第二, 3D 密集块可以提取图像中多变和复杂的空间信息,有助于网络提取时空特征信息,从而提高网络参数的利用率。将卷积层替换为 3D 密集块主要方法有 SC-DenseNet 和 3D-DenseNet-569 两种模型: Ke 等<sup>[29]</sup>提出自约束密集网络 (self-constrained DenseNet, SC-DenseNet), 该模型保留

原 U-Net 中首、尾卷积层,将其余卷积层替换为 3D 密集连接模块,将 3D 密集块下采样与上采样的图像做加和,实现了输入图像的自动检测和分割; Nasser 等<sup>[30]</sup>提出 3D-DenseNet-569,用卷积层与 3D 密集块组合代替了原 U-Net 的两次卷积,增强了在深层网络中收敛到最优解的能力。

第三, Dense-Inception 模块是将密集连接与 Inception 结构组合起来用于 U-Net 中,使网络更深更宽,减少梯度消失现象和冗余计算。将卷积层替换为 Dense-Inception 的主要方法有 DIU-Net 和 LCU-Net 两种模型: Zhang 等<sup>[31]</sup>提出密集 Inception U-Net (Dense-Inception U-Net, DIU-Net), 此网络用 Dense-Inception 模块替换瓶颈层的卷积层,模块内部使用 Res-Inception 模块,外部使用密集连接,将该模块连接到 U-Net 结构体系

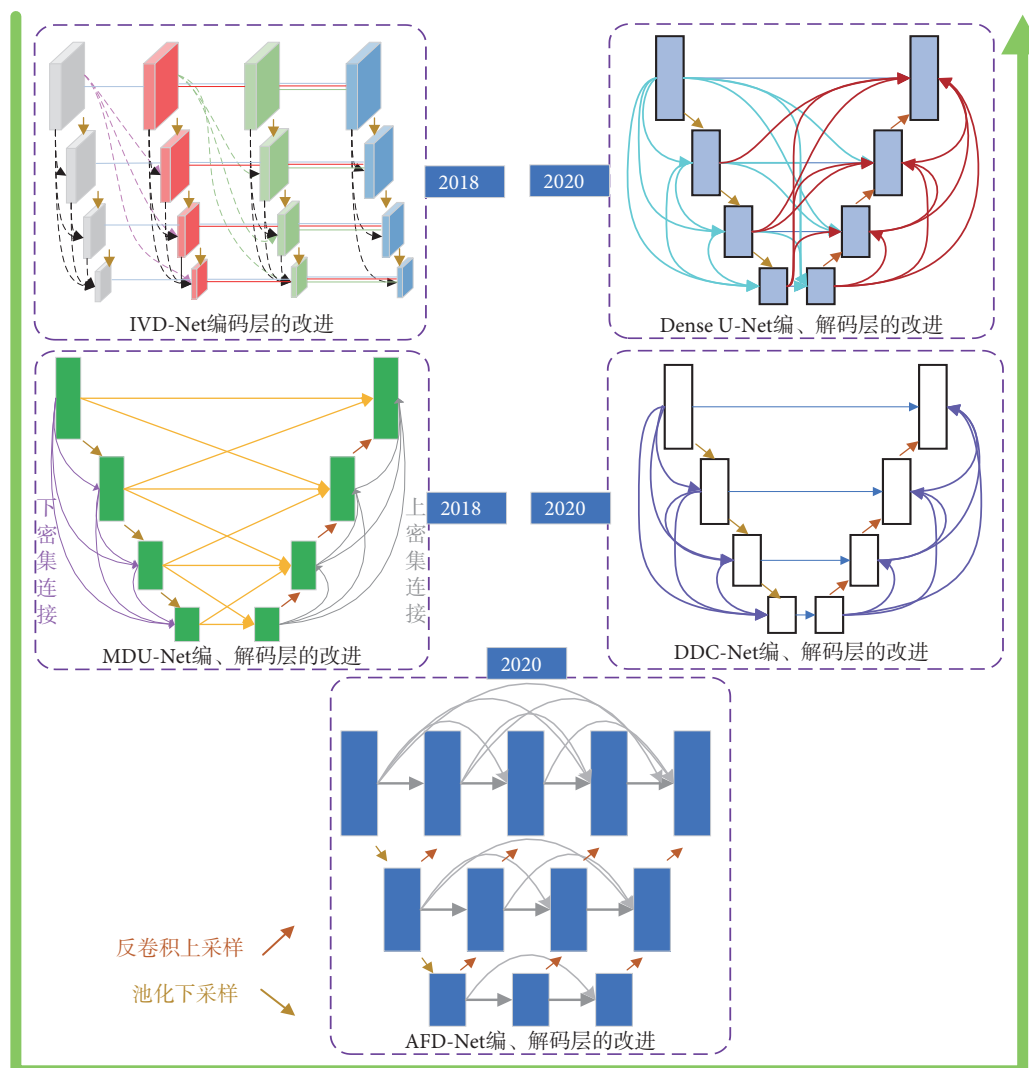


图5 编、解码器中加入密集连接

Fig.5 Dense connection built into encoder and decoder

中,避免了网络训练期间梯度消失或冗余计算;Zhang 等<sup>[32]</sup>提出低消耗 U-Net (low-cost U-Net, LCU-Net),此网络用 Inception 模块替换卷积层,解决了 U-Net 单一感受野设置的局限性和相对较高的存储成本。

**2.1.2 编、解码器中加入密集连接** 编、解码器加入密集连接是在 U-Net 的编码器或者解码器中加入密集连接,具体拓扑结构如图 5 所示,按照箭头的时间年份从左到右呈现了编、解码器中加入密集连接的简略图,可以更好地构建编、解码器不同模块之间的关系。

Jose 等<sup>[33]</sup>提出椎间盘分割网络 (intervertebral disc network, IVD-Net),该模型使用密集机制将编码器逐层连接,每个编码器处理不同的图像模式,使得模型可以自由地了解不同的模式应该在哪里以及如何进行处理和组合。Zhang 等<sup>[34]</sup>提出多尺

度密集连接 U-Net (multi-scale densely connected U-Net, MDU-Net),该网络的编、解码之间使用密集连接,直接融合相邻的高低层不同尺度的特征图,改进了编码器、解码器与跳跃连接,减少了来自密集连接的过拟合。Wang 等<sup>[35]</sup>提出自适应全密集连接网络 (adaptive fully dense UNet, AFD-UNet),此网络在 Unet++ 的基础上,通过水平密集连接将 Unet++ 每层的特征密集连接起来,自适应地有效利用浅层和深层特征。Shi 等<sup>[36]</sup>提出双密集上下文感知网络 (dual dense context-aware network, DDC-Net),在编码器和解码器中引入密集连接融合来自较高层和较低层的多分辨率特征图,以加强当前层中的特征传播,从而改善编码器和解码器之间的信息流。Mohammad 等<sup>[37]</sup>提出 Dense-Unet,此网络将编码器的每层逐层向下层连接与解码器的每层逐层向上连接形成密集连接的效果,从而可以利用

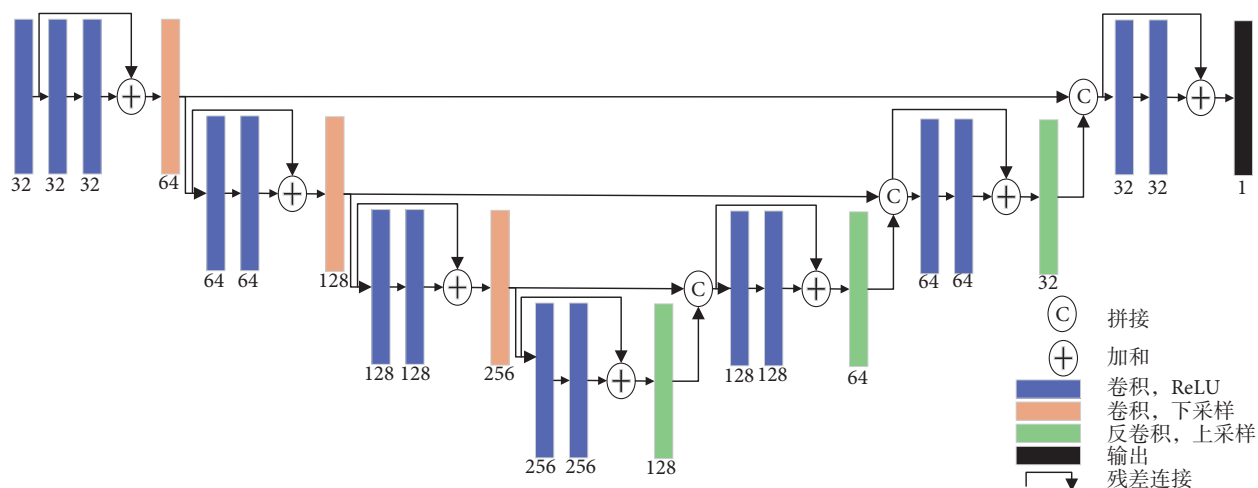


图6 基于 U-Net 的残差连接机制典型代表 ResUnet

Fig.6 Typical representation of the U-Net based residual connection mechanism: ResUnet

不同级别的图像组合。

## 2.2 残差连接机制

U-Net 中的残差连接机制 (residual neural network, ResNet)<sup>[38]</sup> 就是将 U-Net 的卷积层替换为残差块, 跳跃连接使用残差连接路径与编、解码器用残差网络来替换。图 6 描述了 ResUnet<sup>[39]</sup> 架构的细节, 在 ResUnet 架构中进行反卷积操作提高图像分辨率, 并从低分辨率特征中恢复高质量图像。跳跃连接对于 ResUnet 来说是不可或缺的, 它可以提高网络的表示能力, 加快梯度反向传播, 防止训练的不稳定性。基于 U-Net 中残差连接机制的两个特点是: ① 残差结构会简化网络的训练; ② 残差单元中的跳跃连接以及 U-Net 网络的编码和解码之间的跳跃连接可以促进信息传播, 避免梯度消失现象。

残差单元是残差网络的基本组成部分<sup>[40]</sup>, 残差单元由卷积层 (convolution, Conv), 批量归一化层 (batch normalization, BN)、线性整流函数 (rectified linear unit, ReLU) 激活函数和跳跃连接构成。残差连接的特点是保证反向传播参数的更新, 避免反向传播导致的梯度消失问题, 使得优化深层模型更为简单。近年来许多学者将残差连接机制引入 U-Net 中: Lu 等<sup>[41]</sup> 提出环状残差网络 (ringed residual U-Net, RRA-UNet), 此网路添加环形残差模块提取网络深层的附加特征, 可有效地解决胰腺的边界模糊问题; Gu 等<sup>[42]</sup> 提出上下文编码器网络 (context encoder network, CE-Net), 使用 ResNet-34 残差块替换原始 U-Net 编码器块作为固定特征提取器。本节对 U-Net 中加入残差连接机制进行总结, 具体分为: 卷积层的改进, 跳跃连接

处的改进和编、解码单元的改进。

**2.2.1 卷积层的改进** 通过将 U-Net 的卷积层改成残差单元, 可以对网络进行有效的训练, 很好地解决退化问题, 加深网络层数并提高模型性能。本节从基本残差单元、预激活残差单元、双路径残差单元、SE 残差单元、递归残差单元和注意力残差单元等方面进行总结, 如图 7 所示。

第一, 残差单元。残差单元包括相邻堆叠的卷积层进行残差连接, 由卷积层、批量归一化层和 ReLU 函数组成, 许多专家学者将残差单元加入 U-Net 中, 以提高网络架构的性能。Hu 等<sup>[43]</sup> 提出基于补丁的可变形配准网络 (patch-based deformable registration network, PR-Net), 该网络是由残差单元组成, 用于缩短整个网络, 并减轻梯度消失问题; Feng 等<sup>[44]</sup> 提出基于 U-Net 的残差网络 (U-Net based residual network, URNet), 网络在瓶颈结构中嵌入了残差单元, 显著改善了图像去雾效果。Liu 等<sup>[45]</sup> 提出 Res-Unet 结构, 编、解码部分包含两个残差单元, 此网络加深了网络中的层数, 对毛细血管分割具有良好的性能。Wang 等<sup>[46]</sup> 提出混合注意力残差网络 (hybrid dilation and attention residual U-Net, HDA-ResUNet), 将每层卷积中都加入了残差连接来获取高级特征, 与 U-Net 相比, 此模型具有更少的参数和更好的分割性能。

第二, 预激活残差单元。预激活残差单元是通过调整 BN 和卷积层位置实现预激活方式使得网络结构达到最优。研究数据表明, 预激活残差单元比基本残差单元更容易训练和泛化。Waldner 等<sup>[47]</sup> 提出 ResUNet-a, 此架构中编码器和解码器的构建模块由预激活残差单元组成, 有助于缓解梯度消失

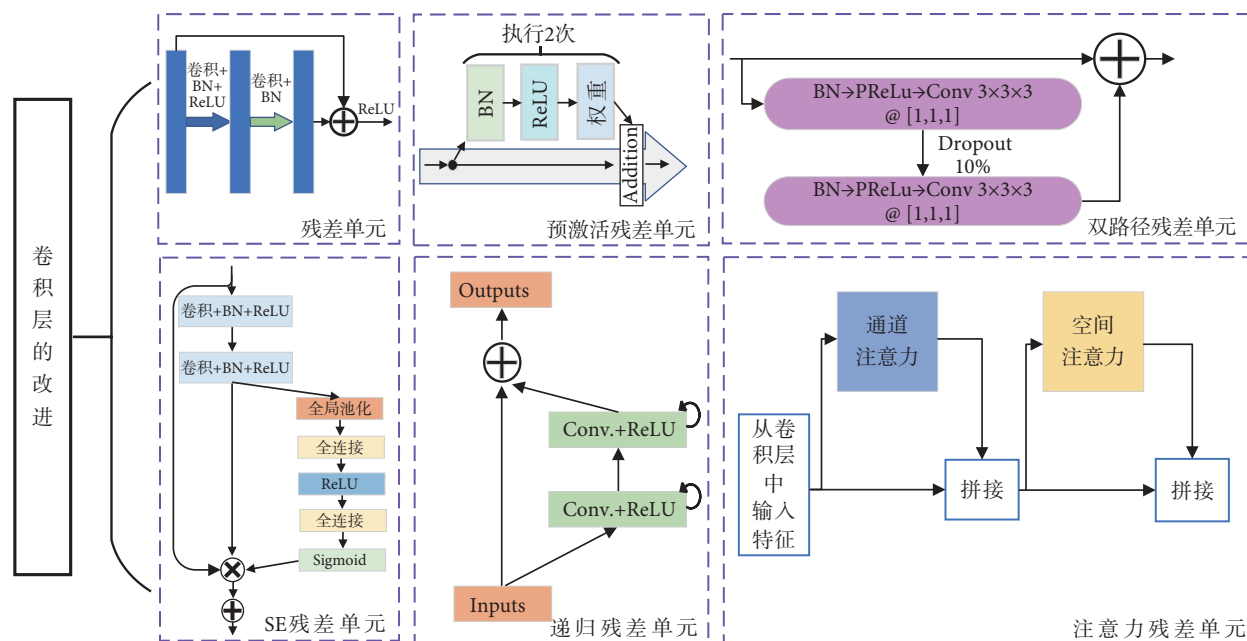


图7 卷积层替换为残差单元

Fig.7 Replacement of convolutional layers with residual units

和爆炸的问题,以准确检测边界并丢弃无关的边界。

第三,双通道残差单元。双通道残差单元能共享公共特征,重复利用已有特征,同时该架构保留了残差连接路径使得神经网络足够灵活而能够学习新的特征。Albert等<sup>[48]</sup>提出Asymmetrical Net,此架构在编码器中使用双通道残差单元,允许编码器在原始和对称模式之间提取关键特征,这种策略可避免表征的瓶颈,同时保留所包含的参数数量。

第四,挤压激励(Squeeze-and-Excitation, SE)残差单元。SE残差单元是由残差路径和SE模块组成,网络在进行粗略定位之后,加入SE残差单元,由粗到细的策略大幅提高了分割精度,可以获得更好的特征表示。Cao等<sup>[49]</sup>提出SE-ResUNet,在编码器部分使用SE残差单元来得到清晰的图像特征,在分割小体积器官方面,此架构比其他的网络架构拥有更好的分割性能。

第五,递归残差单元。主要将每次的卷积递归一次做感兴趣点积累的方法,递归残差卷积执行两次之后再输出给其他模块,具有递归残差卷积层的特征累积确保了分割任务中更好的特征表示。Alom等<sup>[50]</sup>提出R2UNet,在编、解码单元中使用递归卷积层(recurrent convolutional layer, RCL)和带有残差单元的RCL代替常规卷积层,有助于开发更有效、更深入的模型。

第六,注意力残差单元。注意力残差单元是将通道注意力与空间注意力先后依次加入残差单元当中,加入注意力模型能够忽略无关信息而更多地

关注重点信息,再与原特征层进行拼接。Aamer等<sup>[51]</sup>提出注意力残差网(Attention Residual UNet, AReN-UNet),该架构将残差注意力模块加入UNet++中的卷积块,可以提高网络的泛化和表示能力,有效地增加了分割结果。

2.2.2 跳跃连接的改进 跳跃连接的改进是将残差机制应用在U-Net中的跳跃连接处,网络左、右侧分别表示U-Net中的编、解码器,中间为残差跳跃连接(residual path, Respath),可以提高U-Net网络的分割效果,减少参数量,具体如图8所示。Ibtehaz等<sup>[52]</sup>提出MultiResUNet,此架构残差跳跃连接是由一系列卷积层组成,使得网络学习更为容易,可以更好地勾画模糊边界。Zhuang等<sup>[53]</sup>提出LadderNet,在每一级中通过残差跳跃连接将每对相邻编码器和解码器分支进行连接,残差跳跃连接由两次卷积、BN、ReLU与残差连接组成,可以捕获更复杂的特征并产生更高的精度。Yuan等<sup>[54]</sup>提出ResD-UNet,此架构的残差跳跃连接由BN、ReLU、卷积组成,解决了网络加深时出现的梯度消失的问题,能够准确分割肺部CT图像的肺动脉。Shuvo等<sup>[55]</sup>提出分类和定位器网络(Classifier and Localizer UNet, CNL-UNet),残差跳跃连接是由两次不同的卷积与两次残差连接组成,使用改进的跳跃连接来减少原跳跃连接之间的语义差距,有助于模型提高精度。Wang等<sup>[56]</sup>提出自适应接受域多尺度网络(adaptive receptive field multi-scale network, ARMS Net),使用两次卷积与残差连接操



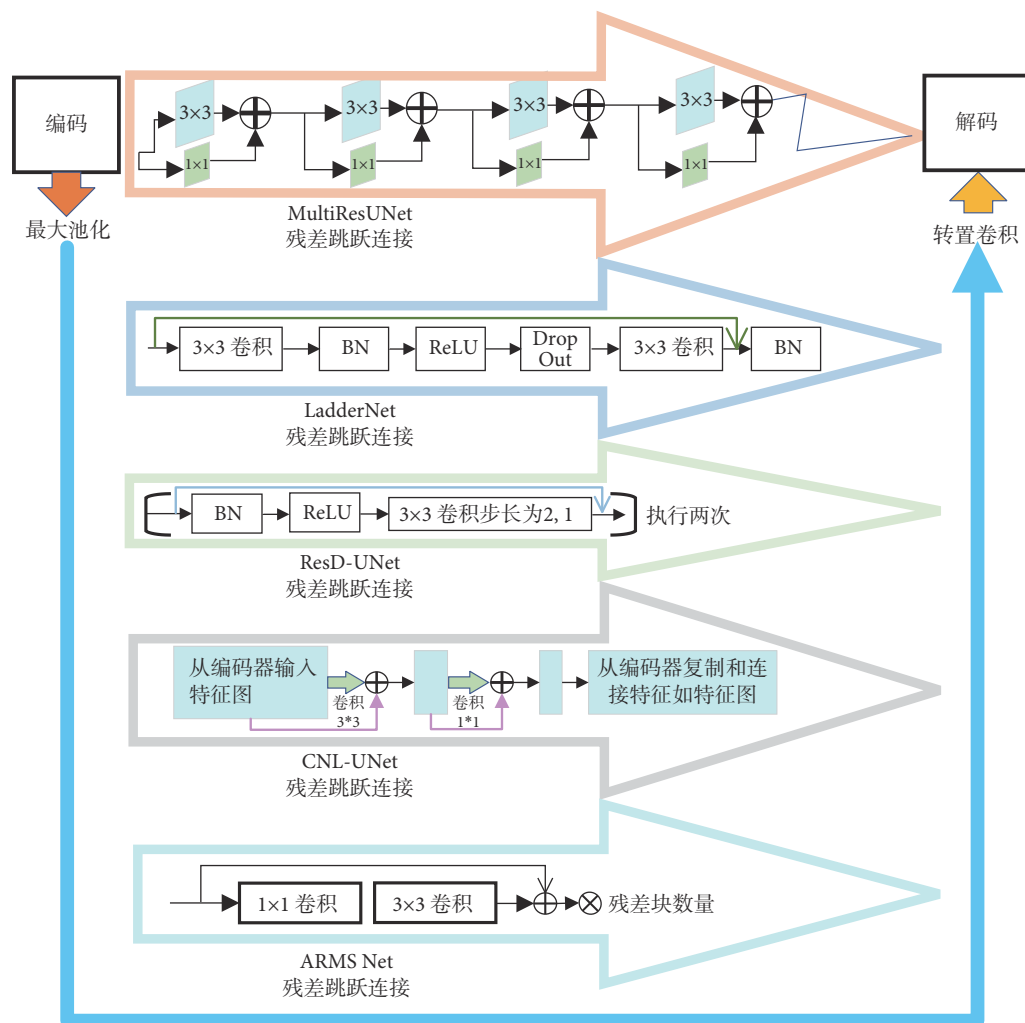


图8 残差连接机制中跳跃连接的改进

Fig.8 Improvement of skip connection in residual connection mechanism

作构成了残差跳跃连接,来代替 U-Net 中的跳跃连接,可以在不缩放或裁剪图像尺寸的情况下自适应地校正感受野,以保留更多的图像细节。

**2.2.3 编、解码单元的改进** 在残差连接机制中,编、解码单元的改进是指将 U-Net 中的编码器替换为残差网络。

Yu 等<sup>[57]</sup>开发了一种改进的 U-Net 架构,采用 ResNet-34 模型作为 U-Net 架构的编码器部分,可以加快模型训练速度,提高鲁棒性。Hari 等<sup>[58]</sup>提出 UNetResNext50,将 U-Net 的编码器替换为 ResNet-50,从 ResNet-50 模型接收到的结果被送到 U-Net 的扩展路径,将其向上采样到原始尺寸,对于磁共振图像中的脑肿瘤自动检测和分割非常有效。

## 2.3 多尺度机制

所谓多尺度,实际上就是对图像的不同粒度的采样,在一种尺度下难以获取图像信息时在另外某种尺度下容易发现或者提取图像信息,在不同尺度

下可以观察到不同的特征,从而完成不同的任务,所以多尺度技术常用于提取图像。将多尺度方法加入 U-Net 中可以通过对不同层次的上下文信息进行编码来生成区分性特征,即低层关注局部外观特征,高层关注全局表示。多尺度特征图捕捉目标图像的每个空间位置的局部邻域的丰富上下文信息,并且在 U-Net 中不同尺度的特征图的空间位置之间存在对应关系。

如图 9 所示,多尺度 U-Net<sup>[59]</sup> (multi-scale U-Net, MU-Net) 由多个 U-Net 组成,其中单个 U-Net 去除不同尺度的噪声,然后基于从粗到精的策略产生性能改进。此网络是由图像金字塔中具有不同空间尺度的 K 个 U-Net 组成,通过联合卷积输入图像,在较低级别 U-Net 中重建输出图像,上采样与当前级别 U-Net 的输入图像一起使用。这种多尺度学习方法能够实现从粗到细的图像重建,从低频到高频逐步构建目标输出。因此, MU-Net 在保留图像细节的同时有效地去除了图像噪声。

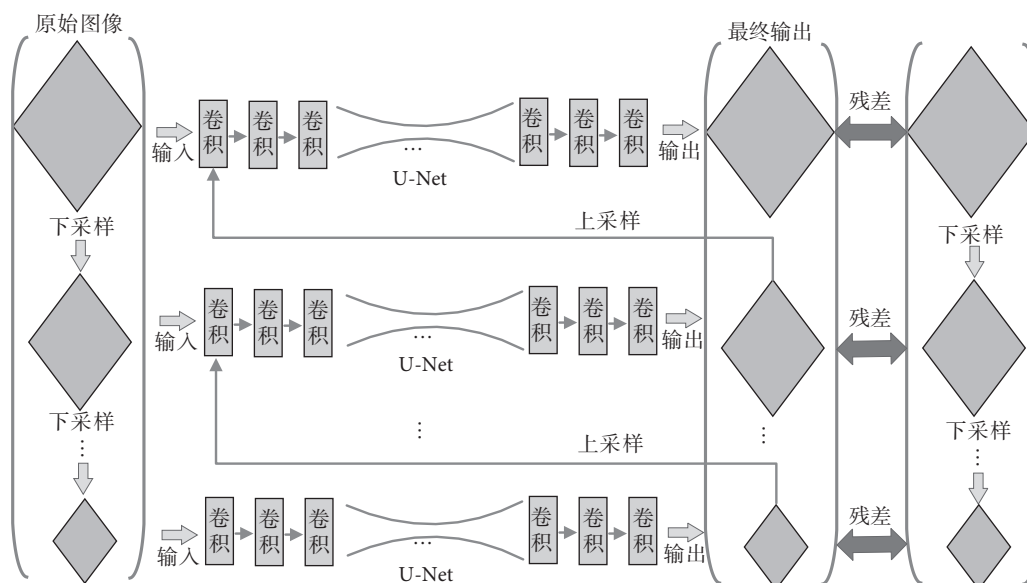


图9 基于 U-Net 的多尺度机制典型代表 MU-Net

Fig.9 Typically represents of multi-scale mechanism based on U-Net: MU-Net

在医学图像器官分割场景中,由于器官尺度差异较为严重,有的器官占据整个输入图像,有的器官只覆盖了几个像素点,不同尺度的器官需要不同尺寸的感受野才能完整地对其特征进行提取。在 U-Net 网络中多尺度机制有两个特点:一是基于 U-Net 使用多尺度输入,提升网络对不同尺度物体的识别能力;二是通过在 U-Net 嵌入金字塔池化模块,提升网络对多尺度特征的提取和学习能力。

**2.3.1 多尺度输入** 在 U-Net 各阶段的输入特征图中,不同阶段得到的特征图的尺寸不同,多尺度输入就是使用多个尺度的图像输入,然后将其结果进行融合,使得最终的输出融合了不同尺寸感受野的特征,从而提升 U-Net 对多尺度目标的识别能力,在 U-Net 中多尺度输入的方法主要有: AID、MDFA-Net、SLSNet、MED-Net。

Cui 等<sup>[60]</sup> 提出注意力输入深度监督网络 (attention input deep-supervised, AID), 在编码层使用了多尺度输入结构,在跳跃连接处使用了多尺度注意力模块,有效提高了心脏分割的准确性。Li 等<sup>[61]</sup> 提出多尺度双路径特征聚合网络 (multiscale dual-path feature aggregation network, MDFA-Net), 网络的输入层中采用多尺度输入来保持原始全局特征,该方法在心脏分割的问题中取得了良好的性能。Sarker 等<sup>[62]</sup> 提出皮损分割网络 (skin lesion segmentation network, SLSNet), 图像经过多尺度输入,将其输入在编码器的每一层内,用于聚集输入皮肤图像的从粗到细的特征,并减少伪影的影响。Kose 等<sup>[63]</sup> 提出多尺度编码器-解码器网络 (multi-

scale encoder-decoder network, MED-Net), 网络将图像以不同尺度的图像大小输入到编码器之中,对其进行多尺度融合之后,提取中间层特征输入到解码器当中, MED-Net 实现了更高的量化指标。

**2.3.2 金字塔池化模块** U-Net 中使用金字塔池化模块是一种充分利用全局信息的方式,捕获多尺度全局上下文信息并生成具有丰富表示的特征,这对于解决遮挡问题 and 提高分割性能至关重要。金字塔池化模块比起普通的单一池化更能保留全局上下文信息,在 U-Net 中金字塔池化模块主要有 CMM-Net、MCNet、GC-DCNN。

Almasni 等<sup>[64]</sup> 提出上下文多尺度多层次网络 (contextual multi-scale multi-level network, CMM-Net), 其主要思想是在网络编码器卷积之后反复利用金字塔池化模块融合多个空间尺度的全局上下文特征,有效地用作各种医学图像分割的工具。Wang 等<sup>[65]</sup> 提出多路径连接网络 (multi-path connected network, MCNet), 它将金字塔池化生成的多条路径集成到编码阶段以保留语义信息和空间细节从而增强网络编码器的细节表示能力,具有更强的特征提取能力。Lan 等<sup>[66]</sup> 提出基于全局上下文的膨胀卷积神经网络 (global context based dilated convolutional neural network, GC-DCNN), 该网络在解码器中使用空间金字塔池化模块产生多级全局上下文特征,并将它们与原始提取的特征连接起来,形成具有更丰富表示信息的最终特征。

## 2.4 集成机制

U-Net 中集成机制是由一组网络集成而成,这

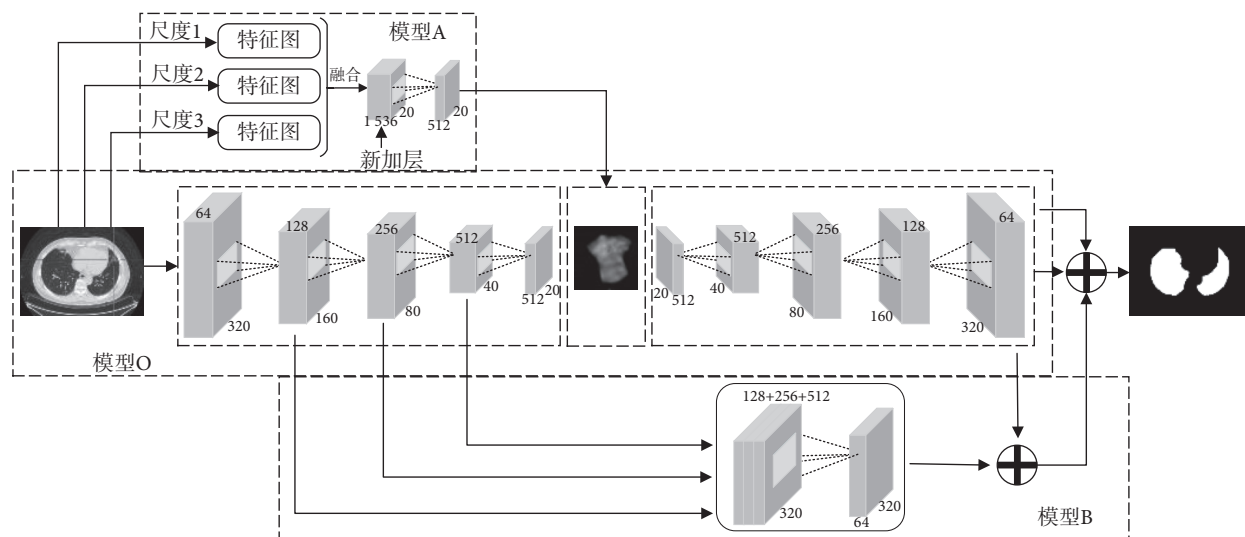


图 10 基于 U-Net 的集成机制典型代表

Fig.10 Typical representation of integration mechanism based on U-Net

些神经网络可以并行处理相同的输入数据后将它们的输出组合起来完成分割。集成机制的特点是建立多个同质且存在差异的个体分类器来解决同一个问题,然后将所有个体分类器的预测结果通过策略组合得到集成机制的最终结果,目的是更好地提高模型的鲁棒性和泛化能力。

为了改善池化过程造成的空间信息损失,利用多尺度<sup>[67]</sup>和多层级学习的方式<sup>[68]</sup>,构建两个差异化的同质型分割模型,并与 U-Net 集成在一起形成新的分割模型。构建的多模型集成算法的典型代表如图 10 所示,中间虚线框为 U-Net 模型 O,左上角虚线框为模型 A 的编码器结构,右下角虚线框为模型 B 的多尺度高中层语义信息融合结构图。

U-Net 分割算法通常对某些目标类的分割效果好,但对其他部分类的分割效果较差,无法满足所有类别的分割要求,导致模型泛化性能不佳,而其他分割模型可能会与该模型互补,因此可以将多个模型的分割结果集成起来,提升单模型分割算法效果。集成机制构建并结合多个分割网络可有效地规避单个分割网络的缺陷。U-Net 中集成的方法可以使用许多并行的模型,然后结合它们产生更精确的分割结果,比较适合用于对分割图像精度要求较高的场合特别是在医学图像分割领域,具有较大的实用价值。在 U-Net 中使用集成机制的主要有: 2.5D Net、AssemblyNet、Ensemble U-Net、UNet-SCDC。

Mei 等<sup>[69]</sup>提出 2.5D Net 来处理不同的图像信息,使用多个 U-Net 集成到一起来提高分割结果的鲁棒性,获得更稳健的结果。Coupe 等<sup>[70]</sup>提出

AssemblyNet, 此模型引入两个 U-Net 集成模型用于知识共享,每个 U-Net 都用不同的训练和验证集进行训练,以利用所有可用的训练数据,使训练和执行速度更快。Li 等<sup>[71]</sup>提出 Ensemble U-Net, 采用集成 U-Net 网络作为基础网络架构,应用膨胀空间金字塔池化 (Atrous Spatial Pyramid Pooling, ASPP) 结构来合并 U-Net 的边缘特征和高级特征,对其结构和参数进行优化和改进,以更好地提取图像特征。Lei 等<sup>[72]</sup>提出跳跃连接和密集连接网络 (skip connection and dense convolution U-Net, UNet-SCDC), 该网络集成两个模块: 基于跳跃连接和密集卷积的 U-Net 分割模块和双重区分模块 (dual discrimination, DD), 当 UNet-SCDC 模块使用密集的膨胀卷积块来生成保留细粒度信息的深度表示时, DD 模块使用两个 CNN 架构组合起来的鉴别器来共同决定鉴别器的输入是真或否。

## 2.5 膨胀机制

膨胀机制就是在 U-Net 的编、解码器中使用膨胀卷积,膨胀卷积<sup>[73]</sup>是在标准卷积核中注入空洞,扩大的幅度称为扩张率,将小的卷积核尺寸变大同时保持卷积的参数数量不变。膨胀机制的特点是在不使用大卷积核、不增加参数量、不增加卷积深度的基础上扩大感受野,获取更大范围内独立的特征信息,可以提高对大目标的分割准确率。

在 U-Net 卷积块中,用膨胀卷积代替标准卷积,改进网络提取的语义信息。膨胀卷积与普通卷积相比除了卷积核的大小以外,还有膨胀率参数,主要用来表示膨胀的大小,从而达到一次卷积看到的空间范围变化大的目的,可以使模型获得更大的



感受野,减少特征信息的丢失,有助于提高模型的性能。本文从在编、解码中使用膨胀机制,与瓶颈层使用膨胀机制两个方面来详细阐述 U-Net 改进中的膨胀机制。

**2.5.1 编、解码器中使用膨胀机制** 在编、解码器中加入膨胀机制是指在 U-Net 中,将编、解码器的原始卷积替换为膨胀卷积,主要方法有: DiSegNet、DIN、DC-U-Net、URNet、EFSNet。Xu 等<sup>[74]</sup> 提出膨胀 SegNet (Dilated SegNet, DiSegNet),在编码器与解码器体系结构中引入了膨胀空间金字塔池化,以不同的膨胀速率提取更多的局部上下文信息,有助于恢复解码器部分的心脏器官边界,提高心脏边界的语义准确预测和详细分割。Li 等<sup>[75]</sup> 提出膨胀初始网络 (dilated-inception net, DIN) 来提取和聚焦多尺度特征用于右心室分割,该网络在 U-Net 的基础上将编、解码器的每层卷积全部替换成膨胀卷积,在右心室分割挑战的基准数据库上表现优于许多最先进的模型。Chen 等<sup>[76]</sup> 提出膨胀卷积的 U-Net (dilated convolution U-Net, DC-U-Net),此模型引入了膨胀卷积来替换卷积块的结构,使腔体结构可以在不增加参数的情况下扩大模型的感受野,提高模型的非线性表达能力。Feng 等<sup>[77]</sup> 提出基于 U-Net 的残差网络 (U-Net based residual network, URNet) 用于图像去噪,编码器模块旨在通过标准卷积和膨胀卷积的混合卷积来提取图像的浅层特征图,以提取更多细节的图像特征。Hu 等<sup>[78]</sup> 提出高效快速语义分割网络 (efficient fast semantic network, EFSNet),使用膨胀卷积替换原卷积,可以在不增加额外参数的情况下扩大感受野,并且提高分割精度,降低分段的噪声。

**2.5.2 瓶颈层使用膨胀机制** 瓶颈层使用膨胀卷积是将 U-Net 底层的卷积块换成膨胀卷积。原始的卷积在很多情况都不能增大感受野,不能全面地捕获上下文信息,因此需要将瓶颈层中的卷积替换成膨胀卷积。主要方法有: HDC-Net、SD U-Net、MD-Unet。Fang 等<sup>[79]</sup> 为了填充图像中缺失或遮罩的区域提出了混合膨胀卷积网络 (Hybrid Dilated Convolution network, HDC-Net)。此架构在瓶颈层中集成 3 次不同扩张率的膨胀卷积,这样可以有效地扩大网络的感受野,进一步提高训练的稳定性。Rad 等<sup>[80]</sup> 提出堆叠膨胀卷积 (stacked dilated U-Net, SD U-Net) 用于识别发育中人类胚胎。该模型在瓶颈层中添加一个由 5 个膨胀卷积层组成的堆栈,采用了五次不同膨胀率的膨胀卷积进一步增加了感受野,在准确性方面比现有技术高出 2.5%。

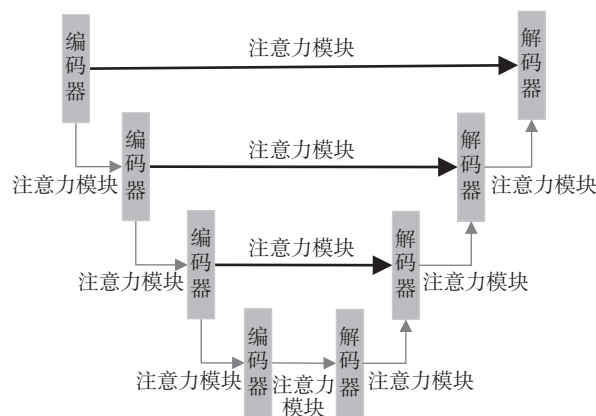


图 11 U-Net 注意力机制图

Fig.11 U-Net attention mechanism diagram

Ge 等<sup>[81]</sup> 提出多输入膨胀 U-Net (multi-input dilated U-Net, MD-Unet) 用于分割膀胱癌。其架构在 U-Net 瓶颈层中将最大池化与膨胀卷积相结合,间歇性地使用膨胀卷积,以减少肿瘤信息丢失,提高网络分割性能。

## 2.6 注意力机制

注意力机制<sup>[82]</sup> 会对输入的上下文表示进行一次基于权重的筛选,通过这种加权的方式让 U-Net 能学到空间上或者时序上的结构关系。U-Net 中的注意力机制是在编、解码器中加入注意力模块,以及在跳跃连接中加入注意力模块,如图 11 所示。在 U-Net 中将注意力模块用在跳跃连接上,可以选择聚焦位置,产生更具分辨性的特征表示,并且不同模块的特征随着网络的加深会产生适应性改变。编、解码器中加入注意力机制,可以更有针对性地对图像进行分割。注意力机制的特点是帮助 U-Net 更好地学到多种内容模态之间的相互关系,从而更好地表示这些信息,克服其无法解释从而难以设计的缺陷。

注意力机制的基本思想是忽略无关信息而关注重点信息,通过神经网络计算出梯度并通过前向传播和后向反馈来学习获得注意力权重。注意力机制分为空间注意力和通道注意力,因此模型可以在空间和通道上学到“在哪里”和“是什么”。本文对编、解码中加入注意力机制与跳跃连接处加入注意力机制进行了总结。

**2.6.1 编码器、解码器中加入注意力机制** 编、解码器使用注意力机制是在每次卷积或者反卷积之后使用注意力机制,本文从编码器、编解码器之间以及解码器加入注意力机制进行总结,如图 12 所示。

第一,在编码器中使用注意力机制是指将注意



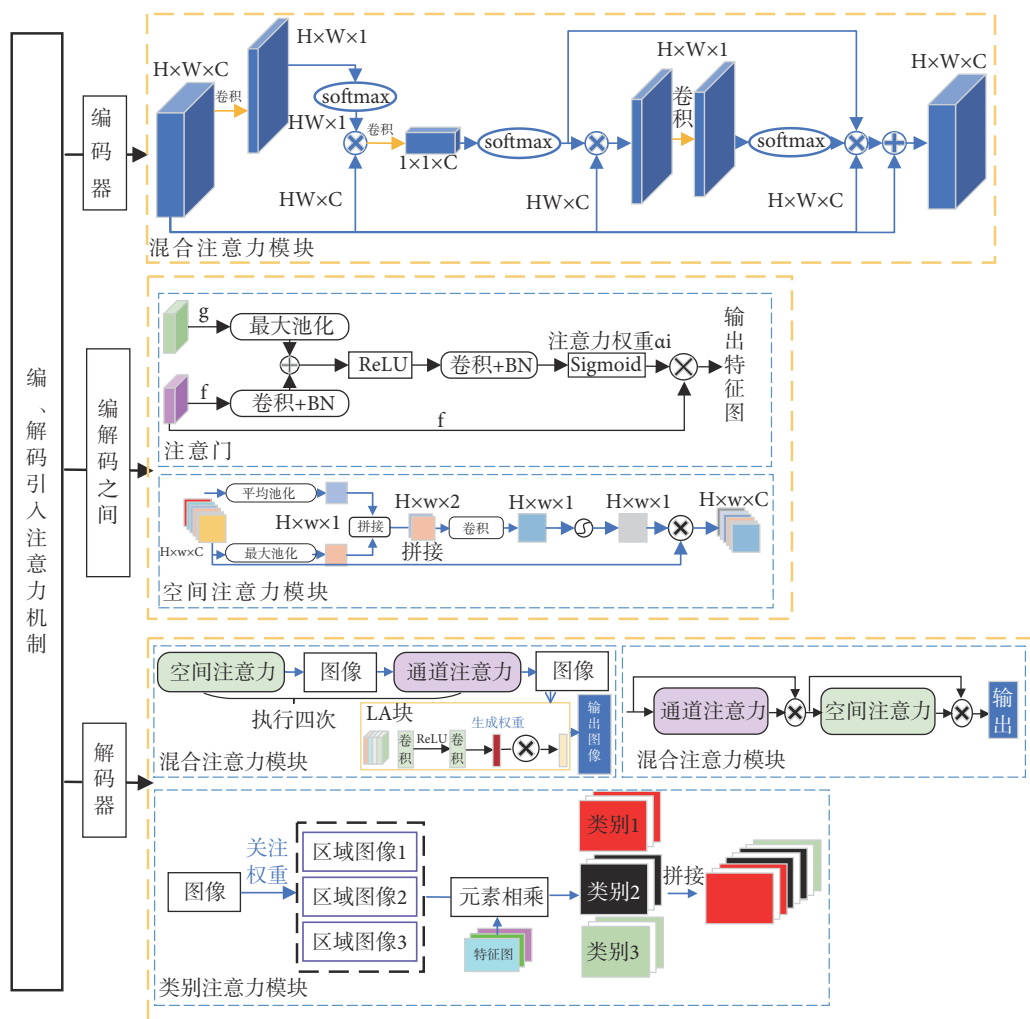


图 12 编、解码引入注意力机制

Fig.12 Introduction of attention mechanism in encoder and decoder

力机制加入 U-Net 的特征提取层,在编码阶段通过使用混合注意块来有效地增强图像特征并抑制斑点噪声。Lan 等<sup>[83]</sup>提出基于混合注意力的残差 U-Net (mixed-attention based residual U-Net, MARU),该网络在编码器中采用轻量级的混合注意块,通过对通道和空间注意的分离和再融合策略,有效地增强了图像特征,抑制了编码阶段的散斑噪声。

第二,编解码器之间使用注意力机制是指在下采样与上采样之间使用注意力机制和瓶颈层中使用注意力机制,这样可以从编码的特征中选择更有用的特征,并将其发送到上层解码器。Li 等<sup>[25]</sup>提出基于注意力的嵌套 U-Net (attention-based nested U-Net, ANU-Net),此网络在上采样与下采样之间引入注意门,使得在不同层次提取的特征可以与相关特征图进行选择合并,同时抑制与分割任务无关的背景区域。Guo 等<sup>[84]</sup>提出空间注意力 U-Net (Spatial Attention U-Net, SA-UNet),此模型在瓶颈

层中添加空间注意力模块,可以帮助网络聚焦于重要特征,抑制不必要的特征,提高网络的表示能力。

第三,解码器中使用注意力机制是指在 U-Net 解码器中加入如图 12 所示的注意力模块,以自动突出相关的特征信息,同时抑制不必要的图像特征。Gu 等<sup>[85]</sup>提出综合注意力网络 (comprehensive attention network, CA-Net),CA-Net 中的大部分注意力块都在解码器中,使用解码器中的注意力块来突出所有候选特征中的相关特征。Pang 等<sup>[86]</sup>提出肿瘤注意力网络 (Tumor Attention Networks, TA-Net),该架构的解码器连续执行通道注意力和空间注意力,可以通过进一步利用不同特征图之间的通道和空间关系更好地生成注意力图。Ding 等<sup>[87]</sup>提出类别注意力提升网络 (category attention boosting network, CAB U-Net),在解码器中利用类别注意力提升模块对特征图上的不同区域给予不同的关注,生成不同的类别图,再将类别图拼接在一起。

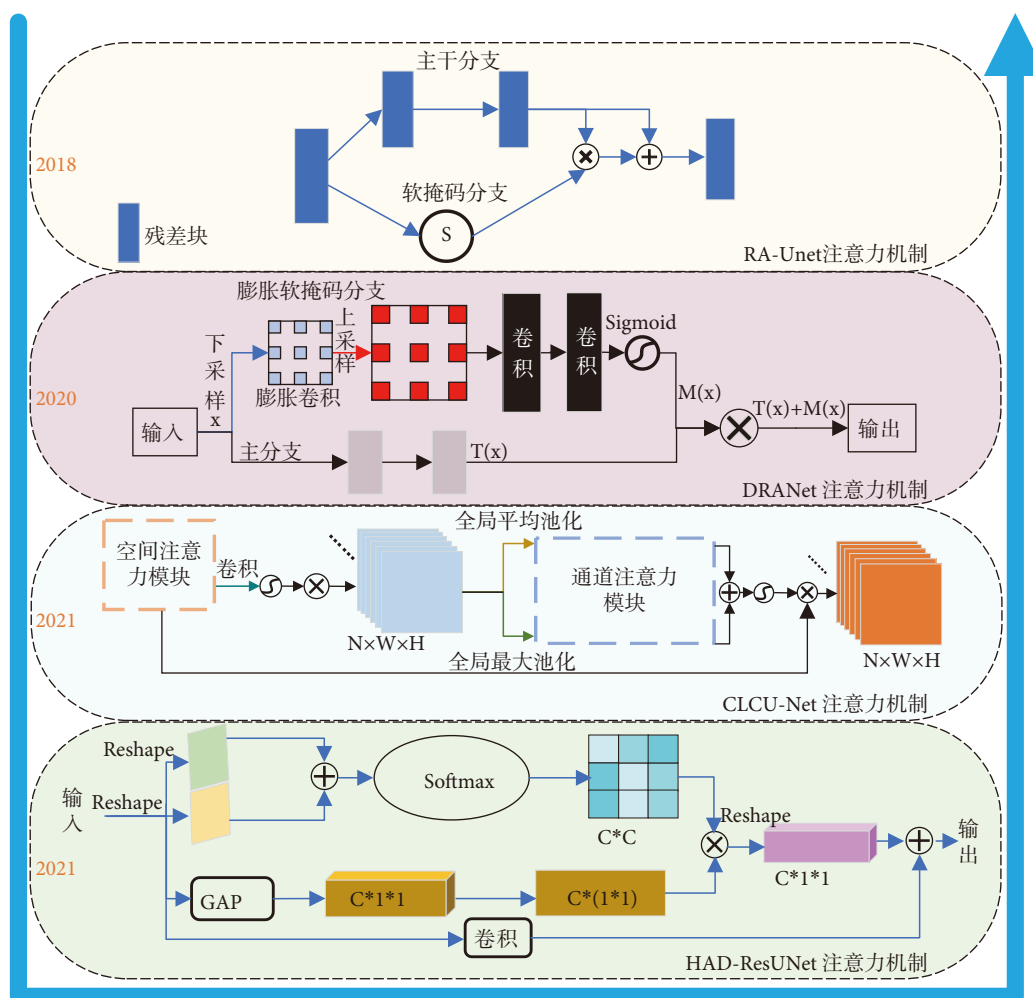


图 13 注意力机制中引入跳跃连接

Fig.13 Introduction of skip connections in the attention mechanism

此模型在大量压缩 U-Net 结构的情况下,增强了网络中的梯度流。

**2.6.2 跳跃连接中加入注意力机制** 跳跃连接加入注意力机制是在 U-Net 的跳跃连接处加入注意力模块。原始 U-Net 只是把同层下采样层的特征值直接拼接至上采样层中,改进后使用注意力模块可以对提取的特征进行重标定。如图 13 所示,按照年份的先后顺序将跳跃连接中加入注意力机制的方法进行总结。Jin 等<sup>[88]</sup>提出残差注意力 U-Net (residual attention U-Net, RAUNet)用于肝肿瘤分割。该网络中的注意力残差机制包含一个主干分支和一个软掩码分支,主干分支学习原始特征,而软掩码分支专注于减少噪声和增强好的特征,该方法在肝脏肿瘤分割中取得了较好的效果,在脑肿瘤分割中具有很高的扩展性和泛化能力。Liu 等<sup>[89]</sup>提出深度残差注意力网络 (deep residual attention network, DRANet),此模型中的注意力机制是由残差块和膨胀卷积组成的,注意力机制改进

了网络的编码器和解码器之间的特征处理,使得模型可以更好地区分两种病变类型。Wang 等<sup>[15]</sup>提出跨层连接的 CLCU-Net,在编码路径与跳跃连接路径中加入由通道注意力模块组成的分割的注意力机制,提取连通特征中的有用信息,剔除冗余信息。Wang 等<sup>[46]</sup>提出 HDA-ResUNet,此模型在跳跃连接中增加了通道注意力机制,使用混合膨胀注意力卷积层放在模型最后一层与反卷积相连,以准确有效地进行医学图像分割。

## 2.7 Transformer 机制

Transformer<sup>[90]</sup>采用编码器-解码器 (Encoder-Decoder) 架构,如图 14 所示堆叠了 6 层的编码器和解码器,是一种避免循环的模型结构。编码器每个层结构包含两个子层:多头注意力层和前馈连接层。解码器有三个子层结构:遮掩多头注意力层、多头注意力层,前馈连接层,每个子层后面都加上残差连接和正则化层。Transformer 机制的特点为善于处理空间上下文关系的优势,对医学图像

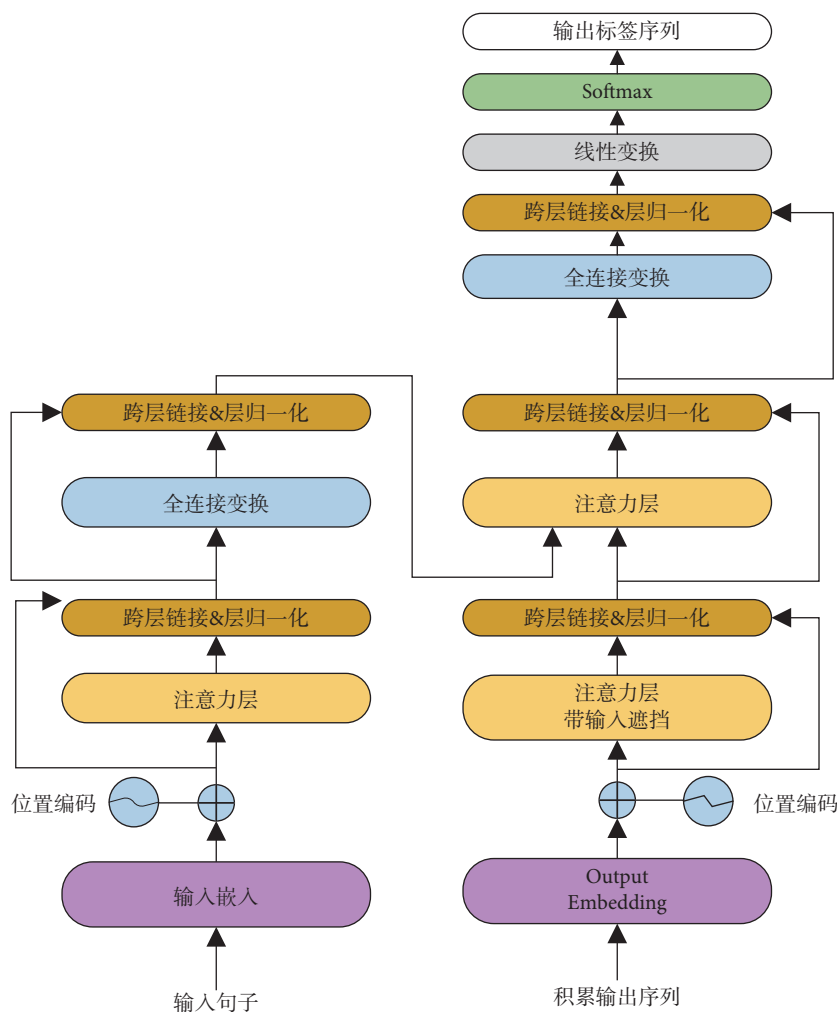


图 14 Transformer 模型架构

Fig.14 Transformer model architecture

上器官之间的关系进行挖掘,提高医学图像的分割精度。Transformer 模型也采用了编码器与解码器结构,调整流通张量的结构与数据稠密程度,提高了网络精度与效率。

U-Net 中的 Transformer 机制是将编码器、解码器与跳跃连接的操作改变为 Transformer。如图 15 所示 Chen 等<sup>[91]</sup>提出了一种基于 U-Net 的 Transformer,此模型具有跳跃连接的多级 Transformer,每个级别的转换器解码器接收三部分输入,跳跃连接的残差嵌入、更高级别的转换器解码器的输出和查询嵌入,使得更多信息特征得以重建,实现了更稳定的训练过程和更精确的异常检测和定位结果。此外,随着注意力层被分解为多级,模型计算成本和内存使用量逐渐减少。

在医学图像分割中,癌症患者病灶边缘对比度低且图像分割在单个图像块级别通常是不明确的,还需要上下文信息来达成语义共识,很容易造成漏诊。因此,如何构造更大的感受野进行上

下文建模,达到既能抽取特征信息又尽量不损失空间分辨率,一直是图像分割的难点。许多分割算法将 Transformer 与 U-Net 结合作为分割病灶的网络。

Li 等<sup>[92]</sup>提出三重注意力网络 (triple attention network, TA-Net),在编码器与解码器上分别运用了 Transformer 机制,专注于捕获各种全局特征维度和特征图的长期依赖关系。Wu 等<sup>[93]</sup>提出一种特征自适应 Transformers 网络 (feature adaptive transformers network, FAT-Net),在编码器部分集成了一个额外的 Transformer 分支,有效地捕获了远程全局上下文信息。Wang 等<sup>[94]</sup>提出多尺度 Transformer 全局注意力网络 (multi-scale transformer global attention network, MSTGANet),将 Transformer 模块插入编码器路径,旨在捕获多尺度非局部来自不同领域的具有长期依赖关系的特征编码器层,从而利用未标记数据进一步提高分割性能。

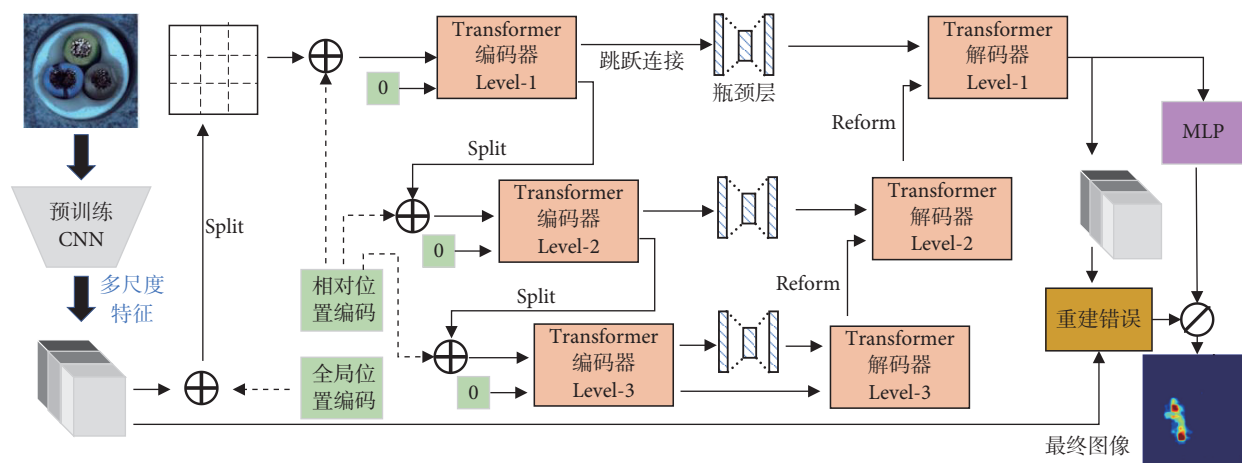


图 15 基于 U-Net 的 Transformer 机制典型代表 U-Transformer

Fig.15 Typical representation of Transformer mechanism based on U-Net: U-Transformer

### 3 总结与展望

将上文介绍的基于 U-Net 改进机制的技术脉络总结成表, 按照 U-Net 结构总结成表 2~4, 分别从结构改进、网络名称、使用方法、时间、数据集、评价参数这 6 个方面进行总结。

(1) 关于结构的改进: 包括卷积操作的改进, 编、解码器的改进和跳跃连接的改进。

(2) 关于使用方法: 将本文综述的基于 U-Net 结构改进机制 (密集连接机制、残差连接机制、多尺度机制、集成机制、膨胀机制、注意力机制和 Transformer 机制) 进行详细的总结, 所列举的方法为近五年一些专家学者所提出的。

(3) 关于数据集, 主要包括: 脑部肿瘤分割 (Brain Tumor Segmentation, BraTS) 是 MICCAI 所有比赛中历史最悠久的数据集, BraTS 2017 中的训练集有 285 个病例, 每个病例有四个模态 (t1、t2、flair、t1ce), 需要分割三个部分: 整个肿瘤 (whole tumor, WT)、增强肿瘤 (enhance tumor, ET) 与肿瘤核心 (tumor core, TC); DDSM 数据集里存放了 Cancer、Normal、Benign 与 Benign\_without\_callback 四类数据, 是一个用于筛选乳腺摄影的数字数据库, 是乳腺摄影图像分析研究社区使用的资源; LIDC-IDRI 数据集共收录了 1 018 个研究实例, 该数据集由胸部医学图像文件 (如 CT、X 线) 和对应的诊断结果病变标注组成; DRIVE 数据集是一个用于血管分割的数字视网膜图像数据集, 它由 40 张图像组成, 其中 7 张显示出轻度早期糖尿病视网膜病变迹象; STARE 数据集是用来进行视网膜血管分割的彩色眼底图数据库, 包括 20 幅眼底图像, 其中 10 幅有病变, 10 幅没有病变。由于

各种方法采用的数据集不尽相同, 本文列举的数据集只提供一个数据集名称索引。

(4) 表中总结各个网络采用的评价指标以及指标数值, 评价指标主要包括 Dice scores、IOU 和 Accuracy 等, 可以为研究学者在改进 U-Net 网络模型时提供参考。

U-Net 已在深度学习领域取得突破性的进展, 本文对 U-Net 的研究背景和意义进行概述, 阐述了 U-Net 在医学图像分割中的应用; 归纳了 U-Net 的七大机制: 密集连接机制、残差连接机制、多尺度机制、集成机制、膨胀机制、注意力机制和 Transformer 机制; 最后对 U-Net 网络结构的改进进行总结, 为临床计算机辅助诊断提供参考。虽然 U-Net 在医学图像领域取得了显著性进展, 但该网络在医学图像分割领域的应用中仍然存在一些挑战和局限。

第一, 医学图像的分割与自然图像分割大相径庭, 即使同一张病灶图像, 不同经验的医学专家也可能会给出不同的诊断方案。资历深厚的专家更能迅速找出器官与病变组织, 这表明先验知识在医学诊断中占很大比重。针对这种现象, 根据医学图像中不同对象灰度分布情况与不同成像设备的特点, 融合先验知识指导分割网络模型结构的设计是未来的研究方向。

第二, 在肿瘤患者的医学图像分割中, 通常是 PET、CT 与 PET-CT 三种模态图像, 由于这三种模态图像具备局部相关性, 不同模态之间肿瘤病灶密切相关, 将网络设置成三种模态输入便会造成输入冗余度非常高。而 Transformer 在医学图像分割任务中激起了热潮, 因此如何利用 Transformer 与 U-Net 二者架构的优势, 解决输入的冗余性会成为未



表 2 U-Net 卷积层的改进

Tab.2 Improvements to the U-Net convolutional layer

网络名称	使用方法	时间	数据集	评价指标	数值 (%)
DIU-Net <sup>[31]</sup>	Dense-Inception块	2020	MICCAI2017BraTS	Dice scores	98.57
LCU-Net <sup>[32]</sup>	Dense-Inception块	2021	EMDS-5	Accuracy	96.91
AD-UNet <sup>[28]</sup>	密集块	2019	DRIVE和STARE	AUC	98.46
DU-UNet <sup>[27]</sup>	密集块	2017	DDSM	F1-score	92.19
SC-DenseNet <sup>[29]</sup>	3D密集块	2020	3 218名NPC患者	DSC	77.07
3D-DenseNet-569 <sup>[30]</sup>	3D密集块	2021	LiTS-2017	Dice	96.70
PR-Net <sup>[43]</sup>	残差单元	2020	OASIS-1	Dice	86.70
URN <sup>[44]</sup>	残差单元	2021	REVITE	SSIM	95.89
Res-UNet <sup>[45]</sup>	残差单元	2020	50幅指甲毛细血管图像	Dice	97.66
ResUNet <sup>[39]</sup>	残差单元	2021	UIH	SSIM	92.20 ± 3.3
HDA-ResUNet <sup>[46]</sup>	残差单元	2021	Lung	Dice	97.97
ResUNet-a <sup>[47]</sup>	预激活残差单元	2020	5个卫星站点提取数据	MCC	81.00
R2UNet <sup>[50]</sup>	循环残差单元	2018	CHASE_DB1	AUC	97.84
Asymmetrical Net <sup>[48]</sup>	双路径残差单元	2020	ISLES 2015	DSC	59.00 ± 31.00
AReN-UNet <sup>[51]</sup>	残差注意力	2021	DRIVE	Accuracy	96.96
DIN <sup>[75]</sup>	膨胀卷积	2019	RVSC	Accuracy	94.00
EFSNet <sup>[78]</sup>	膨胀卷积	2020	CamVid	Global Acc	84.90
HDC-Net <sup>[79]</sup>	膨胀卷积	2020	CelebA	SSIM	95.00

表 3 U-Net 编、解码器的改进

Tab.3 Improvements to the U-Net encoder and decoder

网络名称	使用方法	时间	数据集	评价指标	数值 (%)
IVD-Net <sup>[33]</sup>	密集连接	2018	16个3D多模态MRI	DSC	91.62
DenseUNet <sup>[23]</sup>	密集连接	2019	Conghua Road	Kappa	80.35
ResNet34 UNet <sup>[57]</sup>	ResNet34替换编码器	2019	RIGA	Dice	97.31
UNetResNext50 <sup>[58]</sup>	ResNet50替换编码器	2021	3 929个MRI	DICE	95.73
MCNet <sup>[65]</sup>	金字塔池化	2020	DRIVE	Accuracy	98.20
GC-DCNN <sup>[66]</sup>	金字塔池化	2020	CNDS	F1-score	94.02
DiSegNet <sup>[74]</sup>	膨胀卷积	2021	Thorax	DSC	77.00 ± 5.00
URN <sup>[44]</sup>	膨胀卷积	2021	REVITE	SSIM	95.89
RAUNet <sup>[88]</sup>	注意力机制	2019	CATA7	DICE	97.71
ANU-Net <sup>[25]</sup>	注意力机制	2020	ISBI 2019	DICE	94.79
SA-UNet <sup>[84]</sup>	注意力机制	2021	DRIVE	AUC	98.64
CA-Net <sup>[85]</sup>	注意力机制	2021	ISIC 2018	DICE	92.08
TA-Net <sup>[86]</sup>	注意力机制	2021	COLON CANCER	Accuracy	95.47
TA-Net <sup>[92]</sup>	Transformer机制	2021	DRIVE和STARE	Accuracy	97.18
FAT-Net <sup>[93]</sup>	Transformer机制	2021	ISIC	Accuracy	93.26
MSTGANet <sup>[94]</sup>	Transformer机制	2022	UCSD	DSC	86.20 ± 0.5

表 4 U-Net 跳跃连接的改进

Tab.4 Improvements to the U-Net skip connection

网络名称	使用方法	时间	数据集	评价指标	数值 (%)
MultiresUNet <sup>[52]</sup>	残差连接	2020	CAC-ClinicDB	Jaccard	91.60 ± 0.95
LadderNet <sup>[53]</sup>	残差连接	2019	CHASE_DBI	AUC	98.39
ResD-UNet <sup>[54]</sup>	残差连接	2021	4 000张肺部图像	Precision	98.00
CNL-UNet <sup>[55]</sup>	残差连接	2021	Chest X-ray	DICE	95.94
ARMS Net <sup>[56]</sup>	残差连接	2021	13 434个重叠染色体	IOU	99.45
Attention UNet <sup>[8]</sup>	注意力机制	2018	150个腹部3DCT扫描	DSC	94.30 ± 0.92
CLCU-Net <sup>[15]</sup>	注意力机制	2021	BRATS 2018	Precision	91.98
RA-Net <sup>[88]</sup>	注意力机制	2018	LITS	Jaccard	92.60
HDA-ResUNet <sup>[46]</sup>	注意力机制	2021	LITS 2017	DICE	94.90
DRANet <sup>[89]</sup>	注意力机制	2021	SISS 2015	Accuracy	94.04

来一个非常值得研究的方向。

第三, U-Net 不仅在医学图像上表现出了优异的成绩, 而且在其他应用领域也做出了巨大的贡献, 例如: 基于 U-Net 的人体活动识别 (Human Activity Recognition, HAR) 架构<sup>[95]</sup>; 光学领域中的 variant U-Net<sup>[96]</sup> 架构解决了多光谱图像问题; 建筑行业用于分割钢筋图像的 Multi-Attention U-Net<sup>[97]</sup> 方法; 基于 U-Net 的方法<sup>[98]</sup> 解决了地震数据偏移; 在服装行业 Tan 等<sup>[99]</sup> 提出了非局部试衣网络 (non-local virtual try-on network, NL-VTON) 等。将 U-Net 扩展到更多行业领域里, 并进一步提升 U-Net 的准确性与计算效率是未来值得研究的方向。

综上所述, 从深度学习的发展可以预见, U-Net 网络将会迈向更加广阔的应用, 极大可能推动医学图像分割领域不断突破和创新, 为医疗水平的提升做出巨大贡献。

## 重要声明

利益冲突声明: 本文全体作者均声明不存在利益冲突。

作者贡献声明: 侯森宝为综述主要撰写人, 完成文献资料的整理收集与分析; 赵雅楠、党培参与文献资料分析与论文修改; 周涛与陆惠玲参与论文的修改与指导; 董雅丽参与论文的指导和审校。

## 参考文献

- Shichung L, Shyhliang L, Jyhshysn L, *et al.* Artificial convolution neural network techniques and applications for lung nodule detection. *IEEE Trans Med Imaging*, 1995, 14(4): 711-718.
- LeCun Y, Bottou L, Bengio Y, *et al.* Gradient-based learning applied to document recognition. *Proc IEEE*, 1998, 86(11): 2278-2324.
- Krizhevsky A, Sutskever I, Hinton G. Imagenet classification with deep convolutional neural networks. *Commun Acn*, 2017, 60(6): 84-90.
- Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation// *Medical Image Computing and Computer-assisted Intervention*. Munich: Springer Lncs, 2015: 234-241.
- Liu Liangliang, Cheng Jianhong, Quan Quan, *et al.* A survey on u-shaped networks in medical image segmentations. *Neurocomputing*, 2020, 409: 244-258.
- Milletari F, Navab N, Ahmadi S, *et al.* V-Net: fully convolutional ceural networks for volumetric medical image segmentation// 2016 Fourth International Conference on 3D Vision. Stanford: 3DV, 2016: 565-571.
- Zhou Z, Md S, Nima T, *et al.* Unet++: redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Trans Med Imaging*, 2020, 39(6): 1856-1867.
- Oktay O, Schlemper J, Loic L F, *et al.* Attention U-Net: learning where to Look for the Pancreas (2018-5-30) [2021-11-04]. <https://arxiv.org/pdf/1804.03999.pdf>.
- Ma Hao, Zou Yanni, Liu P. MHSU-net: a more versatile neural network for medical image segmentation. *Comput Meth Prog Bio*, 2021, 208(2): 106230.
- Peng Dunlu, Xiong Shiyong, Peng Wenjia, *et al.* LCP-net: a local context-perception deep neural network for medical image segmentation. *Expert Syst Appl*, 2021, 168: 114234.
- Chen Cheng, Liu Bo, Zhou Kangneng, *et al.* CSR-net: cross-scale residual network for multi-objective scaphoid fracture segmentation, *Comput Biol Med*, 2021, 137: 104776.
- 周涛, 董雅丽, 霍兵强, 等. U-Net网络医学图像分割应用综述. *中国图象图形学报*, 2021, 26(9): 2058-2077.
- Cicek O, Abdulkadir A, Abdulkadir A, *et al.* 3D U-Net: learning dense volumetric segmentation from sparse annotation// *Medical Image Computing and Computer-assisted Intervention*. Athens: Springer Cham, 2016, 9901: 424-432.
- Kumar A, Neha U, Ghosal P, *et al.* CSNet: A new DeepNet framework for ischemic stroke lesion segmentation. *Comput Meth Prog Bio*, 2020, 193: 105524.
- Wang Y, Zhao Z, Syh B. CLCU-Net: cross-level connected U-shaped network with selective feature aggregation attention module for brain tumor segmentation. *Comput Meth Prog Bio*, 2021, 207: 106154.
- Zhou Tao, Dong Yali, Lu Huiling, *et al.* APU-Net: an attention mechanism parallel U-Net for lung tumor segmentation. *Biomed Res Int*, 2022, 2022: 5303651.
- Zhao Juntaing, Dang Meng, Chen Zhihao, *et al.* DSU-Net: distraction-sensitive U-Net for 3D lung tumor segmentation. *Eng Appl Artif Intel*, 2022, 109: 104649.
- Xie Xiwang, Zhang Weidong, Wang Huadeng, *et al.* Dynamic adaptive residual network for liver CT image segmentation. *Comput Electr Eng*, 2021, 91: 107024.
- Jiang Huiyan, Shi Tianyu, Bai Zhiqi, *et al.* AHCNet: an application of attention mechanism and hybrid connection for liver tumor segmentation in ct volumes. *IEEE Access*, 2019, 7: 24898-24909.
- Belh K, Naima K, Nabil M, *et al.* Breast cancer: one-stage automated detection, segmentation, and classification of digital mammograms using unet model based-semantic segmentation. *Biomed Signal Proces*, 2021, 66: 102481.
- Zhang Xiaoxuan, Zhu Xiongeng, Tang Kai, *et al.* DDTNet: a dense dual-task network for tumor-infiltrating lymphocyte detection and segmentation in histopathological images of breast cancer. *Med Image Anal*, 2022, 78: 102415.
- Zhou Tao, Ye Xinyu, Lu Huilin, *et al.* Dense Convolutional Network and its application in medical image analysis. *Biomed Res Int*, 2022, 2022: 2384830.
- Xu Jiangtao, Lu Kaige, Shi xingping, *et al.* DenseUnet generative adversarial network for near-infrared face image colorization. *Signal Process*, 2021, 183: 108007.
- Huang Gao, Liu Zhuang, Maaten L V, *et al.* Densely Connected Convolutional Networks// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 2261-2269.
- Li Chen, Tan Yusong, Chen Wei, *et al.* ANU-Net: Attention-based nested U-Net to exploit full resolution features for medical image segmentation. *Comput Graph*, 2020, 90: 11-20.
- Tang Pin, Zu Chen, Hong Mei, *et al.* DA-DSUnet: dual attention-based dense SU-NET for automatic head-and-neck tumor segmentation in MRI images. *Neurocomputing*, 2021, 435: 103-113.

- 27 Manal G, Mohamed A, Fernando C. DU-Net: Convolutional network for the detection of arterial calcifications in mammograms. *IEEE Trans Med Imaging*, 2020, 39(10): 3240-3249.
- 28 Luo Zhongming, Zhang Yu, Zhou Lei. Micro-vessel image segmentation based on the AD-UNet model. *IEEE Access*, 2019, 7: 143402-143411.
- 29 Ke Liangru, Deng Yishu, Xia Weixiong, *et al.* Development of a self-constrained 3D denseNet model in automatic detection and segmentation of nasopharyngeal carcinoma using magnetic resonance images. *Oral Oncol*, 2020, 110: 104862.
- 30 Nasser A, Amr A, AbdAllah E, *et al.* Efficient 3D deep learning model for medical image semantic segmentation. *Alex Eng J*, 2021, 60(1): 1231-1239.
- 31 Zhang Ziang, Wu Chendong, Sonya C, *et al.* DENSE-INception U-net for medical image segmentation. *Comput Meth Prog Bio*, 2020, 192: 105395.
- 32 Zhang Jinhua, Li Chen, Kosov S. LCU-Net: A novel low-cost U-Net for environmental microorganism image segmentation. *Pattern Recogn*, 2021, 115(4): 107885.
- 33 Jose D, Christian D, Ismail B. IVD-Net: Intervertebral disc localization and segmentation in MRI with a multi-modal UNet. *Lect Note Comput Sci*, 2019, 11397: 130-143.
- 34 Zhang Jiawei, Jin Yuzhen, Xu Jilan, *et al.* MDU-Net: Multi-scale Densely Connected U-Net for biomedical image segmentation (2018-12-4) [2021-11-04]. <https://arxiv.org/pdf/1812.00352.pdf>.
- 35 Wang E, Chen C, Ahmad A, *et al.* A deep learning based medical image segmentation technique in Internet-of-Medical-Things domain. *Future Gener Comp Sy*, 2020, 108: 135-144.
- 36 Shi Jiali, Zhang Rong, Guo Lijun, *et al.* Dual dense context-aware network for hippocampal segmentation. *Biomed Signal Proces*, 2020, 61(6): 102038.
- 37 Mohammad Y, Philippe J, Farida C. Dense-Unet: a light model for lung fields segmentation in chest X-ray images// 2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Montreal: IEEE, 2020: 1242-1245.
- 38 He Kaiming, Zhang Xiangyu, Ren Shaoqing, *et al.* Deep residual learning for image recognition// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 770-778.
- 39 Liu Jin, Kang Yanqin, Qiang Jun, *et al.* Low-dose CT imaging via cascaded ResUnet with spectrum loss. *Methods*, 2021, 202: 78-87.
- 40 周涛, 刘赞琛, 陆惠玲, 等. ResNet及其在医学图像处理领域的应用: 研究进展与挑战. *电子与信息学报*, 2022, 44(1): 149-167.
- 41 Lu Lin, Jian Liqiong, Luo Jun, *et al.* Pancreatic segmentation via ringed residual U-Net. *IEEE Access*, 2019, 7: 172871-172878.
- 42 Gu Zaiwang, Cheng Jun, Fu Huazhu, *et al.* CE-Net: context encoder network for 2D medical image segmentation. *IEEE Trans Med Imaging*, 2019, 38(10): 2281-2292.
- 43 Hu Shunbo, Zhang Lintao, Li Guoqiang, *et al.* Brain deformable registration using global and local label-driven deep regression learning in the first year of life. *IEEE Access*, 2020, 8: 25691-25705.
- 44 Feng Ting, Wang Chuansheng, Chen Xinwei, *et al.* URNet: a U-Net based residual network for image dehazing. *Appl Soft Comput*, 2021, 102: 106884.
- 45 Liu S, Li Yuemei, Zhou Jingjing, *et al.* Segmenting nailfold capillaries using an improved U-Net network. *Microvasc Res*, 2020, 130: 104011.
- 46 Wang Zekun, Zou Yanni, Liu P. Hybrid dilation and attention residual U-Net for medical image segmentation. *Comput Biol Med*, 2021, 134(12): 104449.
- 47 Waldner F, Diakogiannis F. Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network ScienceDirect. *Remote Sens Environ*, 2020, 245: 111741.
- 48 Albert C, Sergi V, Jose B, *et al.* Acute and sub-acute stroke lesion segmentation from multimodal MRI. *Comput Meth Prog Bio*, 2020, 194: 105521.
- 49 Cao Zheng, Yu Bohan, Lei Biwen, *et al.* Cascaded SE-ResUnet for segmentation of thoracic organs at risk. *Neurocomputing*, 2021, 453: 357-368.
- 50 Alom Z, Chris Y, Tarek T, *et al.* Nuclei segmentation with recurrent residual convolutional neural networks based U-Net (R2U-Net)// NAECON 2018 IEEE National Aerospace and Electronics Conference. Dayton: IEEE, 2018: 228-233.
- 51 Aamer A, Biswal B, Pavani G, *et al.* Robust segmentation of vascular network using deeply cascaded AREN-UNet. *Biomed Signal Proces*, 2021, 69: 102953.
- 52 Ibtehaz N, Rahman M S. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Networks*, 2020, 121: 74-87.
- 53 Zhuang J. LadderNet: multi-path networks based on U-Net for medical image segmentation (2019-8-28) [2021-11-04]. <https://arxiv.org/pdf/1810.07810.pdf>.
- 54 Yuan Hongfang, Liu Zhenhong, Shao Yajun, *et al.* ResD-Unet research and application for pulmonary artery segmentation. *IEEE Access*, 2021, 9: 67504-67511.
- 55 Shuvo B, Rifat A, Reza S, *et al.* CNL-UNet: a novel lightweight deep learning architecture for multimodal biomedical image segmentation with false output suppression. *Biomed Signal Proces*, 2021, 70: 102959.
- 56 Wang Guangjie, Liu Hui, Yi Xianpeng. ARMS net: overlapping chromosome segmentation based on adaptive receptive field multi-scale network. *Biomed Signal Proces*, 2021, 68: 102811.
- 57 Yu Shuangang, Xiao Di, Frost S, *et al.* Robust optic disc and cup segmentation with deep learning for glaucoma detection. *Comput Med Imag Grap*, 2019, 74: 61-71.
- 58 Hari R, Kalyan C, Sergey D. Automatic and accurate abnormality detection from brain MR images using a novel hybrid unetresnext-50 deep CNN model. *Biomed Signal Proces*, 2021, 66: 102477.
- 59 Lee S, Makiko N, Hidetoshi U. Mu-net: multi-scale U-net for two-photon microscopy image denoising and restoration. *Neural Networks*, 2020, 125: 92-103.
- 60 Cui Hengfei, Chang Yuwen, Jiang Lei, *et al.* Multiscale attention guided U-Net architecture for cardiac segmentation in short-axis MRI images. *Comput Meth Prog Bio*, 2021, 206(1): 106142.
- 61 Li Feiyan, Li Weisheng, Qin Sheng, *et al.* MDFA-Net: multiscale dual-path feature aggregation network for cardiac segmentation on multi-sequence cardiac MR. *Knowl Based Syst*, 2021, 215(10): 106776.
- 62 Sarker M, Hatem R, Farhan A, *et al.* SLSNet: skin lesion segmentation using a lightweight generative adversarial network. *Expert Syst Appl*, 2021, 183: 115433.
- 63 Kose K, Bozkurt A, Fox C, *et al.* Segmentation of cellular patterns in confocal images of melanocytic lesions in vivo via a multiscale encoder-decoder network. *Med Image Anal*, 2021, 67: 101841.
- 64 Almasni M, Kim D. CMM-Net: contextual multi-scale multi-level network for efficient biomedical image segmentation. *Sci Rep*, 2021, 11: 10191.
- 65 Wang Dan, Hu Guoqing, Lyu C. Multi-Path connected network for medical image segmentation. *J Vis Commun Image Rep*, 2020, 71: 102852.



- 66 Lan Meng, Zhang Yipeng, Zhang Lefei, *et al.* Global context based automatic road segmentation via dilated convolutional neural network. *Inform Sciences*, 2020, 535: 156-171.
- 67 Chen L C, Yang Yi, Wang Jiang, *et al.* Attention to scale: Scale-aware semantic image segmentation// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 3640-3649.
- 68 Chen L C, Barron J, George P, *et al.* Semantic Image Segmentation with Task-Specific Edge Detection Using CNNs and a Discriminatively Trained Domain Transform// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 4545-4554.
- 69 Mei Haochen, Lei Wenhui, Gu Ran, *et al.* Automatic segmentation of gross target volume of nasopharynx cancer using ensemble of multiscale deep neural networks with spatial attention. *Neurocomputing*, 2021, 438: 211-222.
- 70 Coupe P, Boris M, Michael C, *et al.* AssemblyNet: A large ensemble of CNNs for 3D whole brain MRI segmentation. *NeuroImage*, 2020, 219: 117026.
- 71 Li Qianwen, Jia Weikuan, Sun Meili, *et al.* A novel green apple segmentation algorithm based on ensemble U-Net under complex orchard environment. *Comput and Electron Agr*, 2021, 180(6): 105900.
- 72 Lei Baiying, Xia Zaiming, Jiang Feng, *et al.* Skin lesion segmentation via generative adversarial networks with dual discriminators. *Med Image Anal*, 2020, 64: 101716.
- 73 Fisher Y, Vladlen K, Schemmel J. Towards the second generation brainscales system (2015-11-23) [2021-11-04]. <https://arxiv.org/pdf/1511.07122v1.pdf>.
- 74 Xu Guoping, Cao Hanqiang, Udpa J, *et al.* DiSegNet: A deep dilated convolutional encoder-decoder architecture for lymph node segmentation on PET/CT images. *Comput Med Imag and Grap*, 2020, 88: 101851.
- 75 Li Jingcong, Zhu Liangyu, Gu Zhenghui, *et al.* Dilated-inception Net: Multi-scale feature aggregation for cardiac right ventricle segmentation. *IEEE Trans Biomed Eng*, 2019, 66(12): 3499-3508.
- 76 Chen Kuanbing, Xuan Y, Lin A, *et al.* Lung computed tomography image segmentation based on U-Net network fused with dilated convolution. *Comput Meth and Prog Bio*, 2021, 207: 106170.
- 77 Feng Ting, Wang Chuansheng, Chen Xinwei, *et al.* URNet: A U-Net based residual network for image dehazing. *Appl Soft Comput*, 2020, 102(12): 106884.
- 78 Hu Xuegang, Wang Haibo. Efficient fast semantic segmentation using continuous shuffle dilated convolutions. *IEEE Access*, 2020, 8: 70913-70924.
- 79 Fang Yuchun, Li Yifan, Tu Xiaokang, *et al.* Face completion with Hybrid Dilated Convolution. *Signal Process Image*, 2019, 80: 115664.
- 80 Rad R M, Saeedi P, Au J, *et al.* Multi-resolutional ensemble of stacked dilated U-Net for inner cell mass segmentation in human embryonic images// 2018 25th IEEE International Conference on Image Processing (ICIP). Athens: IEEE, 2018: 3518-3522.
- 81 Ge Ruiquan, Cai Huihuang, Yuan Xin, *et al.* MD-UNET: Multi-input dilated U-shape neural network for segmentation of bladder cancer. *Comput Biol Chem*, 2021, 93: 107510.
- 82 Xu K, Ba J, Kiros R, *et al.* Show, attend and tell: neural image caption generation with visual attention. *Comput Sci*, 2015, 37: 2048-2057.
- 83 Lan Yancheng, Zhang Xuming. Real-time ultrasound image despeckling using mixed-attention mechanism based residual UNet. *IEEE Access*, 2020, 8: 195327-195340.
- 84 Guo Changlu, Szemenyei M, Yi Yugen, *et al.* SA-UNet: Spatial attention U-Net for retinal vessel segmentation// 2020 25th International Conference on Pattern Recognition (ICPR). Milan: IEEE, 2021: 1236-1242.
- 85 Gu Ran, Wang Guotai, Huang Rui, *et al.* CA-Net: Comprehensive attention convolutional neural networks for explainable medical image segmentation. *IEEE Trans Med Imaging*, 2021, 40(2): 699-711.
- 86 Pang Shutao, Du Aanan, Orgun M, *et al.* Tumor attention networks: Better feature selection, better tumor segmentation. *Neural Networks*, 2021, 140: 203-222.
- 87 Ding Xiaofeng, Peng Yaxin, Shen Chaomin, *et al.* CAB U-Net: An end-to-end Category Attention Boosting algorithm for segmentation. *Comput Med Imag Grap*, 2020, 84(5): 101764.
- 88 Jin Qiangguo, Meng Zhaopeng, Sun Changming, *et al.* RA-UNet: A hybrid deep attention-aware network to extract liver and tumor in CT scans. *Front Bioeng Biote*, 2020, 8: 605132.
- 89 Liu Liangliang, Kurgan L, Wu Fangxiang, *et al.* Attention convolutional neural network for accurate segmentation and quantification of lesions in ischemic stroke disease. *Med Image Anal*, 2020, 65: 101791.
- 90 Ashish V, Noam S, Niki P, *et al.* Attention is all you need// 2017 NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: ACM, 2017: 6000-6010.
- 91 Chen Liyang, You Zhiyuan, Zhang Nian, *et al.* UTRAD: Anomaly detection and localization with U-Transformer. *Neural Networks*, 2022, 147: 53-62.
- 92 Li Yang, Yang Jun, Ni Jiajia, *et al.* TA-Net: Triple attention network for medical image segmentation. *Comput Biol Med*, 2021, 137: 104836.
- 93 Wu Huisi, Chen Shihuai, Chen Guilian, *et al.* FAT-Net: Feature adaptive transformers for automated skin lesion segmentation. *Med Image Anal*, 2022, 76: 102327.
- 94 Wang Meng, Zhu Weifang, Shi Fei, *et al.* MsTGANet: Automatic drusen segmentation from retinal OCT images. *IEEE Trans Med Imaging*, 2022, 41(2): 394-406.
- 95 Zhang Yong, Zhang Zhao, Zhang Yu, *et al.* Human activity recognition based on motion sensor using U-Net. *IEEE Access*, 2019, 7: 75213-75226.
- 96 Zeng Tao, Diao Changyu, Lu Dongming. U-Net-based multispectral image generation from an RGB image. *IEEE Access*, 2021, 9: 43387-43396.
- 97 Shi Jie, Wu Kunpeng, Yang Chaolin, *et al.* A method of steel bar image segmentation based on multi-attention U-Net. *IEEE Access*, 2021, 9: 13304-13313.
- 98 Wang Benfeng, Li Jiakuo, Luo Jingrui, *et al.* Intelligent deblending of seismic data based on U-Net and transfer learning. *IEEE Trans Geosci Remote*, 2021, 59(10): 8885-8894.
- 99 Tan Zelin, Bai Jing, Zhang Shaomin, *et al.* NL-VTON: a non-local virtual try-on network with feature preserving of body and clothes. *Sci Rep*, 2021, 11(1): 19950.

收稿日期: 2021-11-03 修回日期: 2022-05-27  
 本文编辑: 李蓓兰