

深度学习在医疗影像处理与分析中的研究进展与挑战综述

摘要：近年来，深度学习在医疗影像处理与分析领域取得了显著进展，为临床诊断、治疗和疾病监测提供了强有力的技术支持。医疗影像（如 CT、MRI、PET、超声等）在医学诊断中的重要性日益凸显，但其数据复杂性、标注样本稀缺等问题对传统处理方法提出了挑战。深度学习凭借其在自动特征提取和非线性建模方面的优势，已成为肿瘤检测、器官与病灶自动分割等任务的重要技术手段。本文系统梳理了深度学习在医疗影像处理领域的研究进展，重点分析了卷积神经网络（CNN）和生成对抗网络（GAN）在肿瘤检测与分割中的应用，以及 U-Net、注意力机制和多模态融合技术在器官分割和病灶检测中的发展与挑战。此外，探讨了当前医疗影像处理领域面临的挑战并展望了深度学习在医疗影像处理中的未来发展方向。

关键词：医疗影像处理、深度学习、器官分割、肿瘤检测、多模态影像融合

Abstract: In recent years, deep learning has made significant progress in the field of medical image processing and analysis, providing strong technical support for clinical diagnosis, treatment, and disease monitoring. Medical imaging modalities such as CT, MRI, PET, and ultrasound have become increasingly important in medical diagnostics, yet their complexity and the scarcity of labeled data pose challenges to traditional methods. Leveraging its advantages in automatic feature extraction and nonlinear modeling, deep learning has become a crucial tool for tasks such as tumor detection and automated organ and lesion segmentation. This paper systematically reviews the research progress of deep learning in medical image processing, with a focus on the application of convolutional neural networks (CNNs) and generative adversarial networks (GANs) in tumor detection and segmentation, as well as the development and challenges of U-Net, attention mechanisms, and multi-modal fusion technologies in organ segmentation and lesion detection. Additionally, this review discusses the current challenges in medical image processing and offers perspectives on the future directions of deep learning in this field.

Keywords: medical image processing, deep learning, organ segmentation, tumor detection, multi-modal image fusion.

引言

在现代医学中，在疾病筛查、诊断、治疗监测等方面，影像技术被广泛应用于关键的临床工具。基于不同的物理成像原理，X-射线、电脑断层扫描(CT)、磁共振成像(MRI)、正电子发射断层扫描(PET)和超声等技术为临床医生提供了丰富

的生理和病理资料。然而，在图像数据激增的背景下，传统依靠人工判读的图像分析方法受限于医生的主观和诊断准确性，面临着严峻的挑战。

近年来，深度学习（Deep Learning, DL）的快速发展显著推动了医学影像处理与分析的自动化和精确化。尤其是在图像分割、分类、检测等任务中，卷积神经网络（CNN）、生成对抗网络（GAN）和 U-NET 等模型显示出显著优势。在这些技术的帮助下，医学影像的分析从医生经验依赖的主观判读转向提高诊断效率和精确度的高效计算机辅助分析。

深度学习是一种多层神经网络构架，旨在通过对复杂或非线性特征进行高级抽象，捕捉数据的潜在模式。作为智慧医疗的重要应用，深度学习不仅有效缓解了影像医生短缺和数据量激增的矛盾，还在大规模生物医学影像数据处理中展现出独特的特征提取能力。自深度学习在计算机视觉和图像识别领域取得突破以来，其在医学影像领域的应用已迅速扩展，广泛用于肿瘤检测、病灶识别等任务。基于深度学习的医学影像分析技术正逐步实现智能化和自动化，并在现代临床诊断中发挥出重要作用。

本综述旨在对医学影像处理中深度学习技术的最新研究进展进行系统总结，重点分析其在肿瘤检测与分割、自动分割器官与病灶、多模态影像融合等领域的应用成果。本文还将探讨深度学习在医学影像处理中的核心挑战，包括数据获取和标记难题、多模态数据融合的复杂性和模型解释性问题，并提出多模态数据融合、提升模型的解释性和可信度、利用迁移学习、数据增强和合成图像技术扩充数据集等未来发展方向。

1. 基于深度学习的肿瘤检测与分割

1.1 卷积神经网络（CNN）在肿瘤检测中的应用

近年来，卷积神经网络（CNN）在医学影像处理，尤其是肿瘤检测中取得了显著突破。CNN 通过多层卷积和池化操作，可以自动学习影像中的空间特征，尤其适合复杂医学影像。相对于传统的特征提取方法而言，CNN 对复杂图像的处理能力有了明显的提升，不需要预先设计好的特征提取规则。特别是 3D CNN 技术

已在三维 CT 和 MRI 的医学影像分析中获得成功，大幅提高了三维图像中肿瘤和器官的检测精度。此外，迁移学习利用在 ImageNet 等大规模数据集上预训练的 CNN 模型，可以增强小规模医学影像数据集的泛化能力。

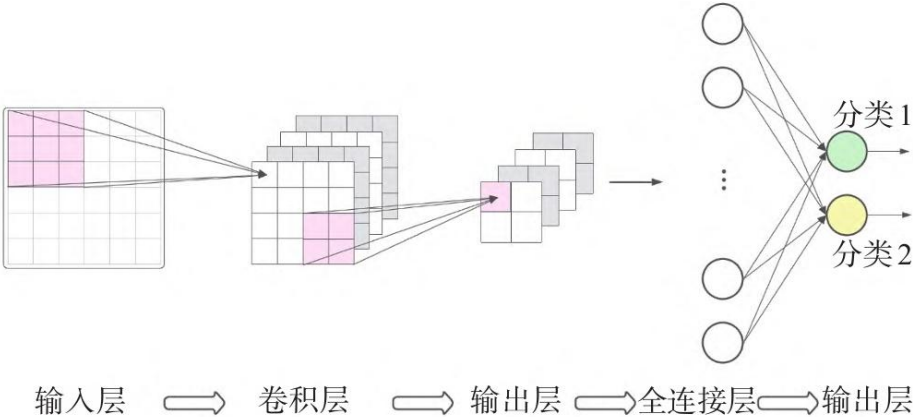


图1 卷积神经网络示意图

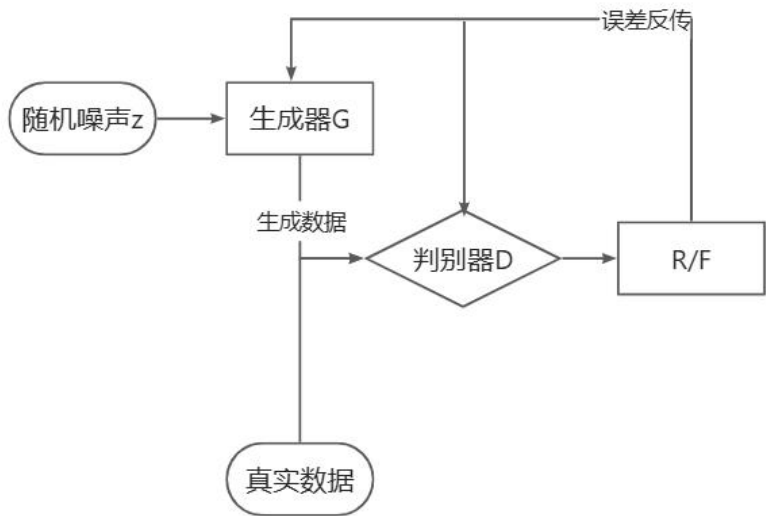
引用自文献[15]

Resnet 和 DenseNet 等网络被广泛应用于肺结节、脑肿瘤和乳腺癌的检测当中。ResNet 在每层中通过增加残差模块, 有效地缓解了深层网络训练中的梯度消失问题，从而使网络在 CT 或 MRI 图像中的细节特征学习起来更加有效。而 DenseNet 则通过密集的连接策略将各个层的特征映射传递给后续的所有层，从而使前面层的特征得到最大的利用，而这种结构在医学图像分析任务中特别适用于多尺度的特征提取。

然而，高效的 CNN 训练通常依赖大量数据，以避免模型过拟合。由于医学影像数据的样本量通常较少，深度学习的需求难以完全满足。为了解决这一问题，广泛采用了数据增强技术（如旋转、缩放、平移和对比度调整），以生成更多的训练样本，提升模型的鲁棒性。此外，迁移学习在医学影像处理中同样得到广泛应用。通过将已在 ImageNet 等大规模数据集上预训练的网络迁移至肿瘤检测任务中，CNN 模型可以在医学影像分析中展现出较强的泛化能力。

1.2 生成对抗网络（GAN）在肿瘤分割中的应用

生成对抗网络（GAN）作为一种具有生成和判别双重机制的深度学习模型，在医学图像分割，尤其是肿瘤分割任务中展现出巨大的应用潜力。GAN 由生成器和判别器两部分构成，通过生成对抗机制，生成器逐步学习逼真的图像特征，以达到与真实图像相似的效果。此过程不仅使 GAN 具备了生成高质量图像数据的能力，还通过调整生成器的架构和训练策略，使其在图像分割任务中表现优异。对于肿瘤分割任务，GAN 生成器能够生成与肿瘤区域高度匹配的分割图，而判别器则通过反馈机制优化生成器，以提高分割的准确性。



生成对抗网络结构示意图

在实际应用中，GAN 能够通过无监督或半监督的方式处理标注数据有限的医学影像任务。如 CycleGan、Pix2Pix 等变体模型，将未标示的 CT 或核磁共振图像通过实现图像到图像的转换，转化为带有标注信息的分割图，显著降低了对大规模人工标注数据的依赖。同时，GAN 的生成对抗训练机制能捕捉肿瘤的边缘信息，进一步提高分割结果的细节精度。相比于传统分割方法，基于 GAN 的分割方法在保持边缘清晰度和细节准确性方面表现更优，适用于处理边界模糊的肿瘤影像数据。

此外，GAN 在多模态影像融合的肿瘤分割中也具有一定的优势。融合 CT、MRI 等影像数据时，将多模态数据输入到 GAN 模型中，能实现不同模态信息的高效整合，进一步提升分割效果。基于 GAN 的肿瘤分割方法不仅在精确性和稳定性方面

优于传统方法，还能为后续的肿瘤检测、病灶识别和治疗规划提供更可靠的数据支持。

2. 医学影像中的器官与病灶自动分割

2.1 基于 U-Net 的器官分割

U-Net 是一种基于编码器-解码器结构的深度学习网络，由 Ronneberger 等人在 2015 年提出，专为医学图像分割任务设计。U-Net 的架构包括对称的编码路径和解码路径，以实现图像特征的提取和分割结果的生成。编码路径通过一系列卷积和最大池化操作逐步提取图像的深层特征，并在下采样过程中保留图像的主要结构信息；解码路径则通过上采样和拼接操作逐步恢复图像的分辨率，并融合编码路径的低层细节特征，最终生成具有边界细致的分割图像。得益于这种设计，U-Net 在保留空间信息的同时，能够捕捉丰富的语义特征，因而特别适用于结构复杂且边缘模糊的医学图像处理任务。

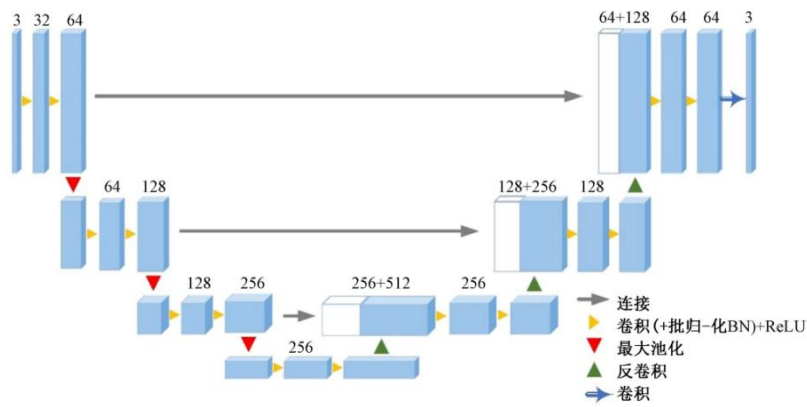


图 2 3D U-Net 结构^[10]

引用自文献[9]

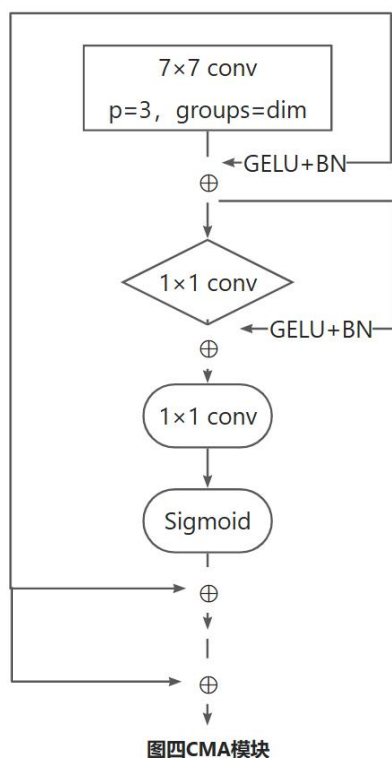
U-Net 的一个显著优势是跳跃连接（skip connections）。模型将编码路径中的低层特征直接传递至解码路径的对应层，从而实现跨层信息的融合。这种信息传递机制保证了解码过程中保留细节特征，提升分割精度。在医学图像分割中，U-Net 利用这种特征融合方式，有效分割出复杂的器官和病灶边界。其在细胞核分割、肝脏分割等多项医学图像分割任务中展现出优异的性能。

由于传统 U-Net 无法直接处理三维影像，如 MRI 和 CT 数据，研究者提出了 3D U-Net，将二维卷积替换为三维卷积，使其能捕捉三维空间中的信息，显著提升了在三维医学图像分割中的表现。此外，Attention U-Net 等变体在 U-Net 的基础上加入了注意力机制，提升了模型对重要特征的聚焦能力，有效抑制了背景噪声的影响。这些改进进一步扩展了 U-Net 在各种医学影像分割任务中的适用性，使其在小病灶和复杂结构分割中表现更为出色。

2.2 基于注意力机制的分割模型

注意力机制（Attention Mechanism）近年来在医学图像分割任务中展现出显著的优势，尤其在提升分割精度方面取得了关键突破。传统的卷积神经网络（CNN）由于受限于局部感受野的卷积操作，通常只能捕捉到图像的局部特征，难以处理全局上下文信息。这种局限性在医学图像中表现得尤为明显，例如，病变区域与正常组织的边界往往模糊，且图像中的病灶可能分布在较大范围内，这使得仅依赖局部信息的卷积网络难以准确区分病变区域。通过建模特征之间的全局依赖关系，注意力机制能够在更大范围内增强对关键区域的关注，从而帮助模型捕捉全局信息，弥补卷积网络在远距离信息捕捉方面的不足。

具体而言，卷积-自注意力模块（Convolutional Self-Attention Module, CMA）通过结合卷积层的局部特征提取能力和自注意力机制的全局特征建模能力，有效优化了特征表示。卷积操作用于细粒度的局部特征提取，而自注意力机制则通过建模全局依赖关系增强了对图像中远距离特征的捕捉。CMA 模块在编码与解码阶段进行信息传递和特征融合，通过残差连接和跨通道交互，进一步提升了模型的特征表达能力，尤其在边界复杂或局部细节模糊的医学图像处理上表现突出。CMA 模块能够自适应地为重要通道赋予更高权重，抑制无关信息干扰，从而显著提高分割的准确性。



基于注意力机制的分割模型在多个医学图像数据集上的实验结果表明，引入注意力机制显著提升了模型的分割性能。例如，朱王令等人提出的融合注意力机制与边缘预测的分割网络，采用卷积注意力模块（Convolutional Block Attention Module, CBAM）强化边缘信息提取，并通过激活函数的变分形式引入边缘先验信息，有效改善了边缘模糊病灶区域的分割效果。实验结果显示，基于注意力机制的分割网络在边缘保留、全局信息整合以及抗噪性能方面优于传统方法，能够有效应对高噪声、低对比度等挑战，进一步推动了医学图像分割领域的发展。

2.3 多模态影像数据的分割技术

多模态医学影像分割技术是近年来计算机视觉和医学影像分析领域的一个重要研究方向。这种技术旨在提取不同模态数据中的有效信息，精确分割和识别图像。该技术整合了不同影像模态的优势，特别是 CT 与 MRI 在器官分割任务中的互补性，使其在临床应用中受到广泛关注。CT（计算机断层扫描）凭借高分辨率的结构成像能力，能清晰展示人体器官的解剖结构，特别在骨骼和硬组织成像方面具有显著优势。而 MRI（磁共振成像）则能提供更丰富的软组织对比信

息，通过多种模态，适用于软组织病变的检测。两种影像模态的结合使得 CT 和核磁共振可以共同提供更全面的解剖学信息和病灶特征，从而在器官分割与肿瘤检测等任务中，显著提升了分割效果和诊断精度。

尽管该技术具有显著优势，但其应用也面临诸多挑战。首先，由于 CT 与核磁共振的成像机制不同，使二者在空间分辨率和对比度上存在差异。CT 影像通常分辨率较高，边界信息清晰；而核磁共振由于软组织对比度复杂，边缘模糊。这种差异使图像的配准和对齐变得困难起来，而分割效果则直接受到多模态配对的精确度的影响。如果出现配准上的误差，就会影响到融合特征的精确，从而使得分割结果的可靠性随之降低。

此外，融合策略也是技术上的核心难点。当前深度学习模型主要采用三类融合策略：输入级融合、特征级融合和决策级融合。在输入级融合中，不同模态的图像被直接堆叠为多通道输入进行训练，这种方法尽管能够保留原始信息，但对模态间的特征互补性利用不足，且易导致信息冗余。而特征级融合方法则通过并行特征提取网络分别提取 CT 和 MRI 的特征，随后对这些特征进行融合，此法在提高信息利用效率的同时也增加了模型的复杂性和计算开销。

3. DL 在医疗影像中面临的挑战及发展前景

经过近几年的快速发展，深度学习技术在临床医学影像辅助分析中得到了越来越多的重视，基于深度学习建立的计算机辅助分析方法已经涵盖了几乎全部的医学影像种类，涉及的疾病类别十分广泛，已经能够在临床影像的分类、分割、配准和重建等方面提供高效可靠的解决方案。然而，要实现这些技术的临床应用，仍面临诸多挑战。

获取高质量的医学影像数据并为其进行精确标注是医疗影像处理的首要挑战之一。目前，公开的医学图像数据平台较少，大多数数据为非公开或小样本数据集。尤其在罕见病的研究中，数据的稀缺性尤为明显。因此，如何有效地扩充数据集，并通过迁移学习、数据增强和合成图像等技术来缓解数据不足的问题，已成为研究的重点之一。近年来，部分研究直接将自动学习方法应用于数据增强，

例如利用卷积神经网络预测灰度图像或透射光图像中的荧光部分[95, 96], 或从未染色组织切片的图像预测染色后的图像[97], 为生成标准化数据提供了一种新的有效方法。

传统的深度学习模型多为“黑盒”结构, 其决策过程缺乏可解释性, 难以满足临床决策的可靠性需求。在医疗高风险领域, 增强模型的可解释性和可信度对于临床推广尤为重要。深度神经网络虽然具有高鉴别力, 但因可解释性不足, 广受质疑。尽管部分专家认为性能比可解释性更值得关注, 但提升可解释性有助于特征组织与模型性能提升。因此, 未来基于注意力机制的可解释模型、自适应权重分配的深度学习方法, 以及透明化设计将成为研究热点。通过这些改进, 模型不仅能够实现高精度的分割和诊断, 还将增强临床应用价值, 提高医生对模型结果的信任度。

总之, 深度学习技术在医学影像处理领域展现出巨大的发展潜力和广泛的应用空间。随着算法模型的不断优化, 深度学习在疾病诊断、分割精度、特异性和鲁棒性等方面的表现将持续提升。在医疗大数据快速发展的背景下, 深度学习技术将有效辅助医生进行更加精准的诊断和个性化治疗, 推动医疗技术智能化发展, 助力医学影像学迈向新阶段。

参考文献:

- [1]林冰洁, 王梅云. 深度学习在医学影像学中的研究现状及发展前景[J]. 山东大学学报(医学版), 2023.
- [2]施俊, 汪琳琳, 王珊珊, 陈艳霞, 王乾, 魏冬铭, 梁淑君, 彭佳林, 易佳锦, 刘盛锋, 倪东, 王明亮, 张道强, 沈定刚. 深度学习在医学影像中的应用综述[J]. 中国图象图形学报, 2020.
- [3]南嘉格列, 李锐, 王海霞, 周旭, 王毅, 倪东. 基于深度学习的肝包虫病超声图像分型研究[J]. 深圳大学学报(理工版), 2019.
- [4]左艳, 黄钢, 聂生东. 深度学习在医学影像智能处理中的应用与挑战. 中国图象图形学报 2021.
- [5]姚超, 赵基准, 马博渊, 等. 基于深度学习的宫颈癌异常细胞快速检测方法[J]. 工程科学学报, 2021.
- [6]蔡立志, 章伟, 陈敏刚, 等. 基于深度学习的细粒度皮肤癌图像分类研究[J]. 计算机应用与软件, 2023.
- [7]宫霞, 吴卫华, 张文涛, 等. 基于深度学习的肺癌患者颈部淋巴结良恶性辅助超声诊断[J]. 计算机应用与软件, 2019.
- [8]张驰名, 王庆凤, 刘志勤, 等. 基于深度迁移学习的肺结节辅助诊断方法[J]. 计算机工程, 2020.

- [9]黄晓鸣,何富运,唐晓虎,王 勋,丘森辉,胡聪 U-Net 及其变体在医学图像分割中的应用研究综述[J] 中国生物医学工程学报, 2022.
- [10]魏炜,刘振宇,王硕,等. 影像组学技术研究进展及其在结直肠癌中的临床应用[J]. 中国生物医学工程学报, 2018.
- [11]巨变端倪:人工智能在医疗影像诊断领域全面爆发[J]. 中国医学计算机成像杂志, 2017.
- [12]徐莹莹,沈红斌. 基于模式识别的生物医学图像处理研究现状[J]. 电子与信息学报, 2020.
- [13]张钦和,刘爱连. MRI 功能成像及定量成像技术在宫颈癌诊疗中的应用述评[J]. 磁共振成像, 2024.
- [14]萧毅,刘士远. 客观看待人工智能在医学影像中的作用[J]. 放射学实践, 2018.
- [15]蒋佳旺,陈艳,王佳庆. 卷积神经网络与迁移学习的颅脑癌症识别方法的研究[J]. 中国医疗设备, 2020.
- [16]王霄,朱恩照,艾自胜. 卷积神经网络的原理及其在医学影像诊断中的应用[J]. 中国医学物理学杂志, 2022, 39(12):1485-1489.
- [17]王海星,杨志清,郭玲玲,郭燕青,张靓,齐昊. 基于大数据和人工智能的超声医学发展现状及问题研究[J]. 肿瘤影像学, 2020.
- [18]张颖,仇大伟,刘静. 生成对抗网络在肝脏肿瘤图像分割中的应用综述[J]. 计算机工程与应用, 2022.
- [19]朱王令,金正猛,王皓. 融合注意力机制和边缘预测的医学图像分割网络算法[J]. 南京邮电大学学报(自然科学版), 2024.
- [20]赵凡,张学典. 集成自注意力机制的医学图像分割方法[J]. 数据采集与处理, 2024
- [21]杨鸿杰,徐巧枝,于磊. 基于深度学习的多模态医学影像分割研究综述[J]. 计算机应用研究, 2022.
- [22]张驰名,王庆凤,刘志勤,等. 基于深度迁移学习的肺结节辅助诊断方法[J]. 计算机工程, 2020.
- [23]袁灵,成思航,苏童,等. 医学影像学的研究进展综述[J] 中国科学: 生命科学, 2021.
- [24]苗扬,张硕,陈俊,等. 基于卷积神经网络的下咽癌医学影像分析综述[J]. 北京工业大学学报, 2024.
- [25]Mazurowski MA, Buda M, Saha A, et al. Deep learning in radiology: an overview of the concepts and a survey of the state of the art with focus on MRI [J] . J Magn Reson Imaging, 2019, 49(4): 939-954.
- [26]Cheplygina V, de Bruijne M, Pluim JPW. Not-so-supervised: a survey of semi-supervised multi-instance, and transfer learning in medical image analysis [J] MedImageAnal, 2019.
- [27]Jia H Z, Song Y, Huang H, Cai W D and Xia Y. 2019a. HD-Net: hybrid discriminative network for prostate segmentation in MR images // Proceedings of the 22nd International Conference on Medical Image Computing and Computer-Assisted Intervention. Shenzhen: Springer: 110-118 [DOI: 10.1007/978-3-030-32245-8_13]
- [28]SHI F, WANG J, SHI J, et al. Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19[J]. IEEE Reviews in Biomedical Engineering, 2020, 14: 4-15.
- [29]LOU A, GUAN S, LOEW M. DC-UNet: Rethinking the U-Net architecture with dual channel efficient CNN for medical image segmentation[C]//Proceedings of Medical Imaging 2021: Image Processing. [S.l.]: SPIE, 2021, 11596: 758-768.

[30]RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]//Proceedings of Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference.Munich, Germany: Springer, 2015: 234-241.