

ARU - Net: 基于残差注意力机制的胸腔积液图像分割模型*

杨靖祎¹ 陈隆鑫² 杨建凯³ 史朝霞⁴ 底涛¹ 刘晓云⁵

(¹ 河北医科大学第二医院数据中心 石家庄 050000 ² 河北医科大学第二医院信息中心 石家庄 050000

³ 河北医科大学第二医院神经外科 石家庄 050000 ⁴ 河北医科大学第二医院医学影像科 石家庄 050000

⁵ 河北医科大学第二医院神经内科 石家庄 050000)

[摘要] **目的/意义** 解决传统胸腔积液分割方法严重依赖先验知识、流程烦琐、耗时费力、性能不佳等问题, 提高效率和准确率。**方法/过程** 根据胸部 CT 图像的积液特征, 提出一种基于残差注意力机制的胸腔积液分割模型 ARU - Net。以 U - Net 模型为主干网络, 在编码和解码阶段引入残差注意力单元, 有效获取图像上下文信息, 提高对特征的利用率。**结果/结论** 在测试集上的 DICE 相似系数达到 88.76%, 与 U - Net 和 ResU - Net 相比在分割完整性和准确性方面具有显著优势, 能够满足临床需求。

[关键词] 胸腔积液; 深度学习; 图像分割; 残差单元; 注意力机制

[中图分类号] R - 058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2024.04.014

ARU - Net: A Pleural Effusion Imaging Segmentation Model Based on Residual Attention Mechanism

YANG Jingyi¹, CHEN Longxin², YANG Jiankai³, SHI Zhaoxia⁴, DI Tao¹, LIU Xiaoyun⁵

¹Data Center, The Second Hospital of Hebei Medical University, Shijiazhuang 050000, China; ²Information Center, The Second Hospital of Hebei Medical University, Shijiazhuang 050000, China; ³Department of Neurosurgery, The Second Hospital of Hebei Medical University, Shijiazhuang 050000, China; ⁴Department of Medical Imaging, The Second Hospital of Hebei Medical University, Shijiazhuang 050000, China; ⁵Department of Neurology, The Second Hospital of Hebei Medical University, Shijiazhuang 050000, China

[Abstract] **Purpose/Significance** The traditional methods for segmenting pleural effusion heavily rely on prior knowledge, are cumbersome in process, time-consuming, and often exhibit poor performance. There is a need to enhance efficiency and accuracy in addressing these issues. **Method/Process** Based on the characteristics of pleural effusion in chest CT images, the paper proposes a pleural effusion segmentation model called ARU - Net, which is based on the residual attention mechanism. The ARU - Net model utilizes the U - Net architecture as its backbone network. It introduces residual attention units in both the encoding and decoding stages to effectively capture contextual information from the images, thereby improving the utilization of features. **Result/Conclusion** The DICE similarity coefficient on the test set reaches 88.76% for ARU - Net, and shows significant advantages in segmentation integrity and accuracy compared to U - Net and ResU - Net, which can meet clinical requirements.

[Keywords] pleural effusion; deep learning; image segmentation; residual unit; attention mechanism

[修回日期] 2024 - 03 - 22

[作者简介] 杨靖祎, 硕士, 工程师; 通信作者: 底涛, 刘晓云。

[基金项目] 河北省医学重点科技研究计划 (项目编号: 20210030); 河北医科大学临床医学创新研究团队项目 (项目编号: 2022LCTD - A7); 河北省卫生健康委员会医学科学研究课题 (项目编号: 20221086)。

1 引言

胸腔积液是一类以胸膜内病理性液体实际有效聚积为特点的临床症候群^[1]。其病因复杂, 主要涉及全身多体系、多种组织脏器的各种炎症、创伤以及恶性肿瘤等^[2]。当积液大于 500 毫升时, 患者会出现明显气短和呼吸困难。目前, 医生可利用计算机断层扫描 (computed tomography, CT) 技术观察积液分布情况, 凭经验或通过记录积液最深处的度量值来预估积液量。

胸腔积液的高精度分割测量可帮助医生鉴别积液良恶性, 对临床治疗具有重要意义。因此, 国内外众多专家学者提出了胸腔积液分割方法。2009 年 Donohue R 等^[3]提出一种新型编码方法实现条件区域生长, 从而分割积液的渗出形状; 高俊刚等^[4]于 2015 年提出游离胸腔积液简易定量法。2016 年 Mansoor A 等^[5]提出基于空间上下文信息的 CT 图像胸腔积液自动分割法, 利用多图谱分割技术生成概率性胸腔积液图。2017 年丁杰^[6]利用移动立方体算法进行三维重构, 提取积液模型再测量体积, 但其本质是对二维数据的聚类, 需人工矫正错误聚类。

虽然上述方法取得一定效果, 但大多是以先验知识为基础, 提取积液形状、所在位置以及纹理等特征, 再进行分割, 其流程烦琐、算法复杂, 效率较低, 无法满足临床快速测量胸腔积液的需求。近年来, 随着计算机运算能力显著提升, 深度学习技术在医学图像分割、分类等领域取得成功。2015 年 Ronneberger O 等^[7]提出全卷积神经网络模型 U-Net, 以绝对优势取得国际生物医学成像研讨会 (international symposium on biomedical imaging, ISBI) 比赛细胞图像分割冠军。2021 年刘蕊等^[8]提出基于多维度特征提取网络 RDD-UNet 模型, 通过长短跳跃连接的组合形式融合多尺度语义特征进行肝脏分割。2022 年王正旭等^[9]设计 3D-UNet 模型实现对阿尔茨海默病神经成像数据库 (Alzheimer's disease neuroimaging initiative, ADNI) 数据集脑磁共振成像 (magnetic resonance imaging, MRI) 图像的高效海马体分割, 帮助临床医生快速诊断阿尔茨

海默病。深度学习相较于传统方法流程简单, 分割更加高效。目前缺乏公开的胸腔积液数据集, 导致利用深度学习进行积液分割的研究较少。因此, 本研究收集就诊于河北医科大学第二医院患者的胸部 CT 图像, 以 U-Net 作为主干网络, 提出一种融合残差和注意力机制的分割模型, 实现对胸腔积液的高效分割。

2 基于深度学习的胸腔积液分割模型

2.1 分割模型结构

2.1.1 基本结构 U-Net 模型由编码器和解码器组成, 编码器部分利用卷积和池化逐步提取图像的特征获得高层抽象信息, 解码器利用转置卷积或者上采样恢复图像的细节和空间维度信息, 编码器和解码器在通道维度进行特征融合, 其在医学图像分割领域取得巨大成功^[7]。U-Net 网络编码器虽然通过一系列卷积层和下采样层提取深度特征, 但下采样亦导致空间信息丢失; 编码器在网络浅层计算, 提取低级特征, 解码器在网络深层计算, 提取更高级别特征, 融合的两组特征之间可能存在语义鸿沟; U-Net 网络不够深, 导致其对特征表达不够准确等问题^[10-12]。

2.1.2 残差网络 (residual network, Res-Net) 在传统卷积神经网络中引入了残差块/残差单元的概念^[13]。残差单元由跳跃连接和恒等映射组成, 使输入信号可以从任意低层传播至高层, 实现特征重用, 有效解决梯度问题与网络退化问题。

2.1.3 注意力机制 人类在处理信息时, 会有选择地关注某一局部重要信息。受这一行为启发, 专家学者提出注意力机制, 将人类注意力行为应用在机器上, 使机器学会感知数据中重要和不重要的部分。将该机制与图像分割网络结合, 使网络更加专注于感兴趣区域, 抑制不相关部分, 提高模型分割精度。常见的注意力模块有通道注意力、空间注意力、门控注意力和自注意力等^[14]。

2.2 ARU-Net 模型

2.2.1 模型设计 传统 U-Net 的编码为局部操

作，无法整合全局信息，下采样时易丢失空间信息；解码同样为局部操作，上采样时很难从编码器有效获取全局信息，导致其对生物学图像分割极为不利^[15]。编码器在网络的浅层进行计算提取低级特征，解码器在网络的深层计算提取高级特征，融合的两组特征之间可能存在语义鸿沟。U-Net 网络不够深，导致其对特征的表达不够准确^[10-12]。积液只存在于胸腔中，本研究使用的数据集中，积液占图像面积最大比例不足20%，平均占比不足5%，前景和背景类别极度不平衡，显著增加了胸腔积液分割难度。本研究针对传统 U-Net 存在的不足，以及胸腔积液图像特点，结合残差网络^[13]和残差单

元中的挤压激励模块（squeeze and excitation layer, SE Layer)^[16]设计基于残差注意力机制的模型（attention ResU-Net, ARU-Net）。一是在编码器中引入注意力残差单元，通过跳跃连接将浅层特征与深层特征拼接，以保留空间特征，减少空间信息的丢失。二是 SE 模块对特征图的每个通道进行重要性预测分配权值，以提高模型前景区域的关注度，降低对背景区域的关注，从而降低积液图像前景和背景类别极度不平衡带来的分割难度。三是编码器和解码器之间的桥接阶段引入注意力门（attention gate, AG）模块^[17]，抑制噪声和无关区域，减少上采样阶段产生的信息丢失，见图1。

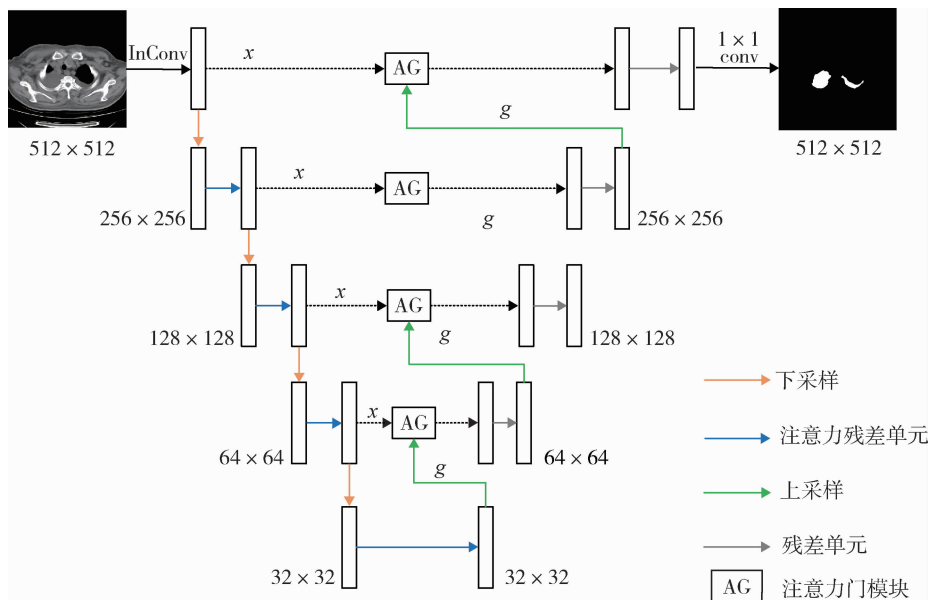


图1 ARU-Net 网络结构

2.2.2 编码阶段 首先经过一个输入卷积块，再依次经过4个下采样层和注意力残差单元。输入卷积块由3×3的卷积层、批归一化（batch normalization, BN）和修正线性单元（rectified linear unit, ReLU）组成。注意力残差单元由卷积块、SE Layer 和跳跃连接组成，见图2。卷积块中包含两个连续的3×3的卷积核，步长设置为1。卷积层中间添加BN层，以使网络的输入特征图保持同分布并有效防止梯度消失或爆炸。SE Layer 由池化层、全连接层和激活函数组成，包括挤压（Squeeze）和激励（Excitation）两部分操作。Squeeze 部分通

过平均池化和全连接对上层网络输出的特征图在通道维度上编码为一个全局特征，获得一个全局感受野。Excitation 操作采用包含两个全连接层的瓶颈（bottleneck）结构，不仅学习各通道之间的关系，而且实现了数据的降维、提升了模型的泛化能力。Excitation 部分最后采用 Sigmoid 形式的门机制（gate mechanism），学习各通道的激活值，与注意力残差单元中卷积块的输出特征图相乘，从而提高模型对通道特征的辨别能力，提升模型对通道特征的敏感性。输入图像经过编码器后，特征图为32×32。

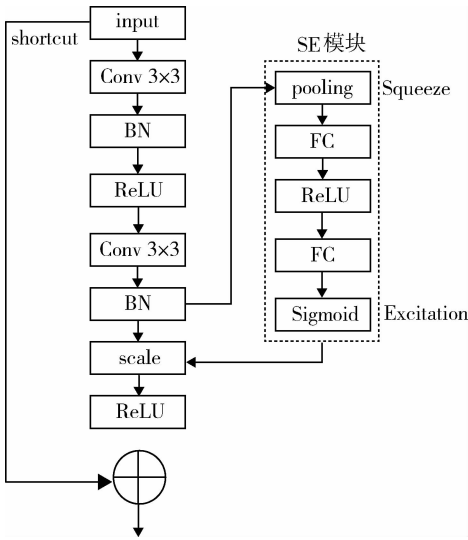


图 2 注意力残差单元

2.2.3 解码阶段 注意力门模块，见图 3。 g 为经过上采样的特征图， x 为编码器对应的输出特征图，通过两个 1×1 的卷积 W_g 和 W_x 分别对 g 和 x 进行线性变换后相加，再依次经过 1×1 的卷积、BN 和 Sigmoid 激活函数生成特征图的注意力系数 α 。注意力系数 α 与编码器的输出特征图 g 相乘获得最后的输出 X_a ，见公式 (1)。通过注意力门模块，使注意力集中到目标区域，减轻上采样阶段产生的信息丢失。模型最后通过 Sigmoid 激活函数对 CT 图像的积液和背景进行分类，从而达到分割效果。

$$X_a = f(R(w_g g + w_x x)) \times x \quad (1)$$

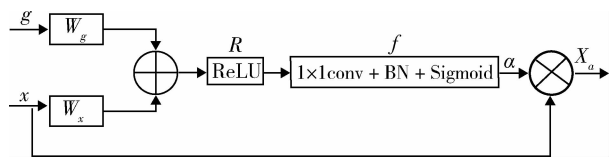


图 3 注意力门模块

2.3 积液图像分割流程

数据集来源于 2020 年就诊于河北医科大学第二医院的 26 例患者的胸部 CT 图像，分辨率为 512×512 ，由影像科医生进行标注，将标注结果生成掩膜图像。对原始图像进行降噪和均衡化等预处理。经过预处理的数据按 7:2:1 的比例划分为训练集 1、验证集和测试集，对训练集 1 进行数据扩展生成训练集 2，见图 4。

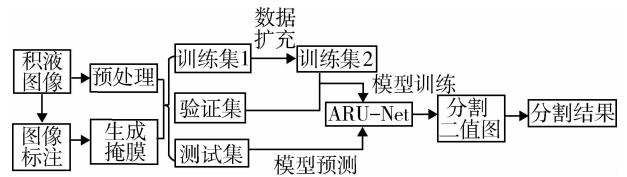


图 4 算法流程

3 实验与方法

3.1 数据预处理

胸腔积液图像边缘轮廓的辨别仅需调整到纵隔窗，其他图像信息是冗余的。因此，图像窗宽设置为 400，窗位设置为 40。由于积液图像的灰度分布不均匀、特征边界模糊等特点，使用中值滤波进行图像降噪，并采用直方图均衡化进行数据增强，以提高图像对比度，见图 5。数据经过预处理后，最终获得 1 151 张 CT 切片。医学图像数据由于其特殊性，具有规模小、标注困难等特点。因此，本研究通过平移、旋转和翻转等操作对训练集样本进行数据扩充^[18]，验证集和测试集不作数据增广。

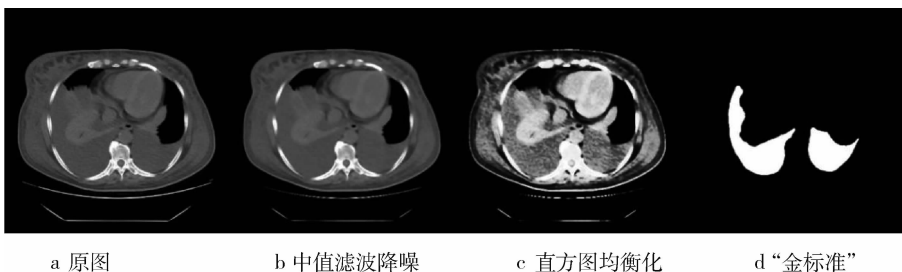


图 5 图像预处理

3.2 实验参数设置

实验在 Windows10 平台下完成，使用 PyTorch1.0 搭建 ARU - Net 网络模型，使用 Adam 优化器，训练过程中利用 Visdom 进行可视化，使用 NVIDIA GeForce RTX3060 GPU 进行加速训练。网络训练次数设置为 100，初始学习率设置为 0.001，批量大小 (batch size) 设置为 8。

3.3 评价标准

模型使用 Dice 相似系数损失函数 (dice loss)，见公式 (2)。X 代表“金标准”图像，Y 为分割预测图像， $|X \cap Y|$ 表示 X 和 Y 的交集， $|X|$ 和 $|Y|$ 分别表示 X 和 Y 的元素个数。对于二分类分割问题而言，“金标准”只有 0 和 1。因此， $|X \cap Y|$ 可以有效地将预测分割图中未在“金标准”分割图中激活的像素值进行清零。对于激活的像素，主要是惩罚低置信度的预测，较高值得到更好的 Dice 系数。Dice Loss 指导模型不断学习，使预测结果逐渐接近“金标准”。

$$d = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \tag{2}$$

4 结果分析

在训练集上经过 100 次训练后，ARU - Net 在测试集上的 DICE 系数为 88.76%。不同模型在测试集上的 DICE 系数，见表 1。ARU - Net 模型的 DICE 系数显著高于 U - Net ($P = 0.009$)，与 ResU - Net 相比亦有优势 ($P = 0.020$)。ARU - Net 的 DICE 系数高于 PSPNet^[19] 和 DeepLab V3^[20] 模型，但并无显著优势。

表 1 不同模型积液图像分割性能对比

模型	Dice 系数 (均值)	Dice 系数 (中位数)
U - Net	78.76	85.01
ResU - Net	81.39	85.20
PSPNet	86.26	90.92
DeepLab V3	85.05	88.97
ARU - Net	88.76	91.40

不同模型的可视化结果，见图 6。

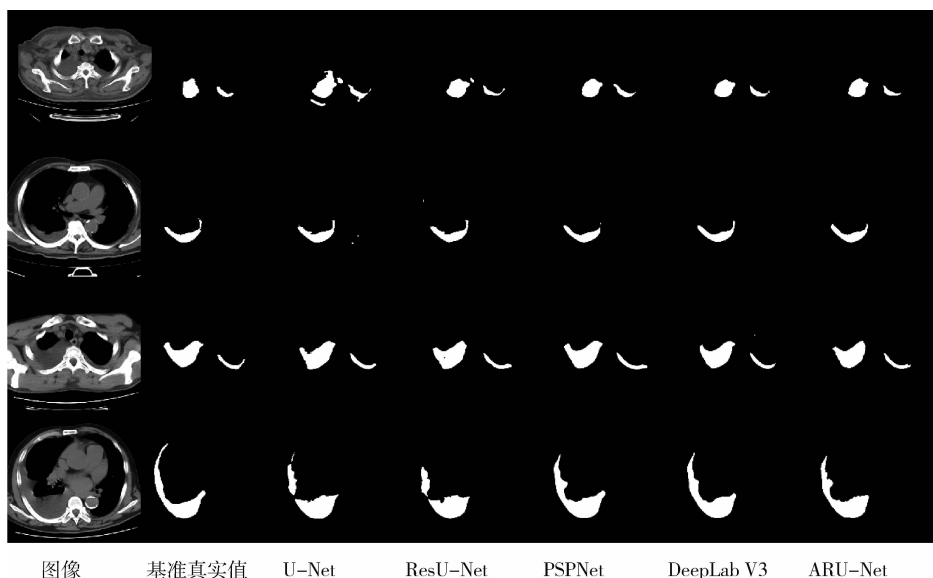


图 6 不同模型分割结果

U - Net 在分割积液时会出现显著偏差，存在冗余区域、分割结果不完整；ResU - Net 虽然改进了 U - Net 的冗余问题，但仍存在分割区域不连续问题。ARU - Net 在分割完整性、准确性上相较于 U

- Net 和 ResU - Net 均有显著优势，可以满足临床需求。虽然 ARU - Net 的 DICE 系数高于 PSPNet 和 DeepLab V3，但在可视化结果上差异并不明显。

5 结语

由于医学影像中的灰度分布不均匀、出现伪造影、特征边界模糊等特点,采用直方图均衡化提高图像对比度;鉴于传统图像分割方法的局限性,利用深度学习进行积液分割的研究较少。本研究探索深度学习在积液图像中的分割,提出一种融合残差注意力机制的编码器-解码器结构的分割模型 ARU-Net,在编码器中设计注意力残差单元,改进 U-Net 模型无法有效整合全局信息的问题;编码器和解码器通过注意力门模块进行特征融合,使注意力集中到目标区域,减轻上采样阶段产生的信息丢失。通过在测试集上进行验证,ARU-Net 模型能够有效应用于胸部 CT 图像积液分割,满足临床需求。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

- MISEROCCHI G. Physiology and pathophysiology of pleural fluid turnover [J]. *European respiratory journal*, 1997, 10 (1): 219 - 225.
- LIGHT R W. Pleural effusion [J]. *New England journal of medicine*, 2002, 346 (25): 1971 - 1977.
- DONOHUE R, SHEARER A, BRUZZI J, et al. Pleural effusion segmentation in thin-slice CT [C]. Dublin: Proceedings of SPIE, the International Society for Optical Engineering, 2009.
- 高俊刚,白璐,金晨望,等.基于 CT 图像对游离胸腔积液简易定量方法 [J]. *中国医学影像技术*, 2015, 31 (2): 4.
- MANSOOR A, CASAS R, LINGURARU M G. Spatial context learning approach to automatic segmentation of pleural effusion in chest computed tomography images [C]. San Diego: Medical Imaging 2016: Computer - Aided Diagnosis, 2016.
- 丁杰.计算机辅助胸腔积液的 CT 测量方法 [D]. 宁波:宁波大学,2017.
- RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [C]. Munich: Medical Image Computing and Computer - Assisted

- Intervention - MICCAI 2015: 18th International Conference, 2015.
- 刘蕊,续欣莹,谢珺.基于多维度特征提取网络的肝脏图像分割 [J]. *河北大学学报 (自然科学版)*, 2021, 41 (4): 426 - 435.
- 王正旭,赵文兵,蔡越江,等.基于三维卷积神经网络对脑 MRI 海马体的高效分割研究 [J]. *中国数字医学*, 2022, 17 (1): 8 - 14.
- IBTEHAZ N, RAHMAN M S. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation [EB/OL]. [2023 - 06 - 21]. <https://www.sciencedirect.com/science/article/pii/S0893608019302503?via%3Dihub>.
- LOU A, GUAN S, LOEW M H. DC-UNet: rethinking the U-Net architecture with dual channel efficient CNN for medical image segmentation [C]. online: Medical Imaging 2021: Image Processing, 2021.
- 马巧梅,梁昊然,郎雅琨.融合残差模块的 U-Net 肺结节检测算法 [J]. *计算机工程与设计*, 2021, 42 (4): 1058 - 1064.
- HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. Las Vegas: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- 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, 39 (5): 1297 - 1306.
- WANG Z, ZOU N, SHEN D, et al. Non-local U-Nets for biomedical image segmentation [C]. New York: 34th AAAI Conference on Artificial Intelligence, 2020.
- HU J, SHEN L, SUN G. Squeeze-and-excitation networks [EB/OL]. [2023 - 03 - 01]. <https://arxiv.org/pdf/1709.01507.pdf>.
- OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-net: learning where to look for the pancreas [EB/OL]. [2023 - 06 - 21]. <https://arxiv.org/abs/1804.03999>.
- 杨婧.面向低分辨 CT 影像的微浸润性腺癌和浸润性腺癌深度学习分类方法研究 [D]. 合肥:中国科学技术大学,2019.
- ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]. Honolulu: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation [EB/OL]. [2023 - 06 - 21]. <https://arxiv.org/abs/1706.05587>.