

胰腺病变图像分割模型研究

黄梦滨, 顾春华

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2025年3月7日; 录用日期: 2025年3月31日; 发布日期: 2025年4月9日

摘要

胰腺病变分割是医学图像分析中的一项重要任务, 但由于胰腺的大小、形状不规则以及位置不固定, 导致该任务充满挑战。胰腺病变的早期检测对于患者的治疗和预后至关重要, 因此受到了广泛关注。本文提出了一种改进的U-Net网络模型, 通过引入MedSAM编码器和高效通道注意力(ECA)模块, 显著提升了胰腺病变分割的精度和鲁棒性。MedSAM编码器对输入图像进行预处理, 提取多层次、多尺度的特征。MedSAM编码器能够有效捕捉胰腺的复杂结构和病变区域的细节信息, 为后续分割提供高质量的特征表达。其次, 在上采样过程中引入高效通道注意力(ECA)模块, 增强通道间的依赖关系, 防止降维导致的通道信息丢失。在MSD_Pancreas数据集上的实验结果表明, 本文方法在Dice相似系数(DSC)和豪斯多夫距离(HD95)两个评价指标上均取得了显著提升。具体而言, 本文方法的DSC分数达到82.7%, HD95值为10.2 mm, 相比基线U-Net分别提高了4.4%和2.3 mm。实验结果表明, 本文方法在胰腺病变分割任务中显著优于基线模型, 尤其是在处理复杂病变和边界模糊区域时表现突出。

关键词

胰腺病变分割, 注意力机制, MedSAM, 高效通道注意力

Research on Image Segmentation Model of Pancreatic Lesions

Mengbin Huang, Chunhua Gu

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Mar. 7th, 2025; accepted: Mar. 31st, 2025; published: Apr. 9th, 2025

Abstract

Pancreatic lesion segmentation is an important task in medical image analysis, but it is challenging due to the irregular size and shape of the pancreas as well as its location. Early detection of

pancreatic lesions is crucial for patient treatment and prognosis, and thus has received much attention. In this paper, we propose an improved U-Net network model that significantly improves the accuracy and robustness of pancreatic lesion segmentation by introducing the MedSAM encoder and the Efficient Channel Attention (ECA) module. The MedSAM encoder preprocesses the input image and extracts multi-level and multi-scale features. The MedSAM encoder is able to efficiently capture the complex structure of the pancreas and the lesion region's detail information, providing high-quality feature representation for subsequent segmentation. Second, the Efficient Channel Attention (ECA) module is introduced in the up-sampling process to enhance the inter-channel dependencies and prevent the loss of channel information due to dimensionality reduction. The experimental results on the MSD_Pancreas dataset show that this paper's method achieves significant improvement in both the evaluation metrics of Dice Similarity Coefficient (DSC) and Hausdorff Distance (HD95). Specifically, the DSC score of this paper's method reaches 82.7% and the HD95 value is 10.2 mm, which are improved by 4.4% and 2.3 mm, respectively, compared with the baseline U-Net. The experimental results show that this paper's method significantly outperforms the baseline model in the task of pancreatic lesion segmentation, especially when dealing with complex lesions and regions with fuzzy boundaries.

Keywords

Segmentation of Pancreatic Lesions, Attention, Medsam, ECA

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

胰腺癌是一种恶性程度高、预后极差的消化系统肿瘤,据国外最新数据显示,胰腺癌总体 5 年生存率仅为 13%,而在中国人群中,仅有 8.5%,是所有恶性肿瘤中最低的[1]。手术是目前胰腺癌唯一可能获得治愈的手段,因此早期诊断和精准治疗对提高患者生存率至关重要。医学影像技术,如计算机断层扫描(CT)和磁共振成像(MRI),在胰腺病变的诊断和分期中发挥着关键作用。然而,由于胰腺解剖结构复杂、与周围组织对比度低,且病变形态多样,医生手动分割胰腺病变耗时且易受主观因素影响。

近年来,深度学习技术在医学图像分割领域取得了显著进展,尤其是基于卷积神经网络(CNN)的算法在胰腺病变分割中展现出巨大潜力。开发自动化、高精度的胰腺病变图像分割模型具有以下重要意义:

- 1) 大幅减少医生工作量,缩短诊断时间。
- 2) 捕捉人眼难以识别的特征,减少漏诊和误诊。
- 3) 精准的分割结果为手术规划和放疗靶区勾画提供可靠依据。
- 4) 有助于大规模数据分析,推动胰腺癌的早期诊断和个性化治疗研究。

胰腺位于腹腔深处,周围被其他器官(如胃、肝脏、脾脏、肠道)包围,导致其在影像中难以清晰显示,此外,胰腺的形状、大小和位置在不同个体间差异较大,且与周围组织的对比度较低,边界模糊,导致难以分割。MedSAM 编码器能够从复杂的医学图像中提取多层次、多尺度的特征[2],因此,本实验将 MedSAM 编码器与传统的 U-Net 网络结合,同时在上采样过程中加入高效通道注意力模块,提升了模型对胰腺病变区域的细节捕捉和精确分割能力。

2. 相关工作

近年来,随着深度学习方法的兴起,基于深度学习方法对胰腺病变分割得到广泛的研究。注意力机

制源于对人类视觉的研究,能够使模型关注序列中关键的信息,为序列的不同部分分配不同的权重。SENet [3]将注意力机制应用于通道,通过动态调整不同通道的特征权重显式建模通道之间的依赖关系,自适应地重新校准通道特征响应。CBAM [4]将通道注意力与空间注意力相结合,将通道注意力模块和空间注意力模块的输出特征逐元素相乘,得到最终的注意力增强特征。这个增强的特征将用作后续网络层的输入,以在保留关键信息的同时,抑制噪声和无关信息。ECA-Net [5]提出了一种高效的通道注意力机制,避免了 SENet 中的全连接层。U-Net [6]是一种经典的医学图像分割网络,其编码器-解码器结构和跳跃连接能够有效捕捉多尺度特征。针对胰腺病变分割,研究者提出了多种 U-Net 与注意力机制结合的变体,如 Attention U-Net [7]在编解码跳跃连接操作之间加一个新型注意力门,以此来实现自动学习关注不同形状和大小的分割区域从而增强对病变区域的关注,ST-TransUNet [8]基于多尺度密集连接和极坐标余弦编码进行胰腺肿瘤分割,聚焦胰腺肿瘤的分割边缘,避免特征感知的冗余。DAU-Net [9]提出一种金字塔注意力特征融合模块,引入通道和空间两个独立注意力机制,提供多尺度输入信息并行采样,解决了连续下采样导致特征空间信息损失的问题。

3. 数据集和实验方法

3.1. 数据集

本研究采用 Medical Segmentation Decathlon (MSD)中的胰腺分割数据集,其中包含 282 例胰腺 3D CT 图像,图像尺寸为 $512 \times 512 \times D$ (D 为切片数量)。数据来自多个医疗机构,具有较高的多样性和泛化性。所有数据均由专业医生进行标注,保证了标注的准确性。部分数据集切片可视化如图 1 所示,其中胰腺使用红色标注,病变区域使用绿色标注。

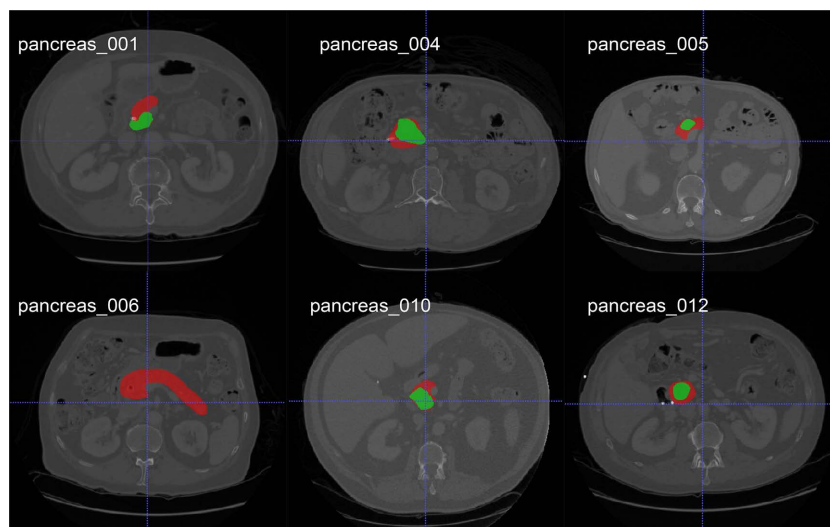


Figure 1. MSD_Pancreas datasets

图 1. MSD_Pancreas 数据集

实验时将数据集按照 7:2:1 的比例将数据集划分为训练集、测试集和验证集,并对数据集进行预处理操作,如随机旋转、翻转等操作,从而提高模型的泛化能力。

3.2. 模型架构

本研究使用的模型结构如图 2 所示。采用 MedSAM 中的编码器对胰腺图像处理后,得到初步数据

特征, MedSAM 编码器从图像中提取多层次的特征, 包括低层次的边缘、纹理信息和高层次的语义信息。随后对特征图进行下采样, 一共 4 个阶段, 将每个阶段得到的特征图分别作为网络结构第二、三、四、五层的输入特征图。每个上采样阶段加入高效通道注意力模块, 减小降维带来的特征损失, 并与对应层的输入特征图连接, 本实验采用简单的跳跃连接。最终, 通过一个 1×1 的卷积层将特征图映射到目标类别的数量, 生成最终的分割结果。

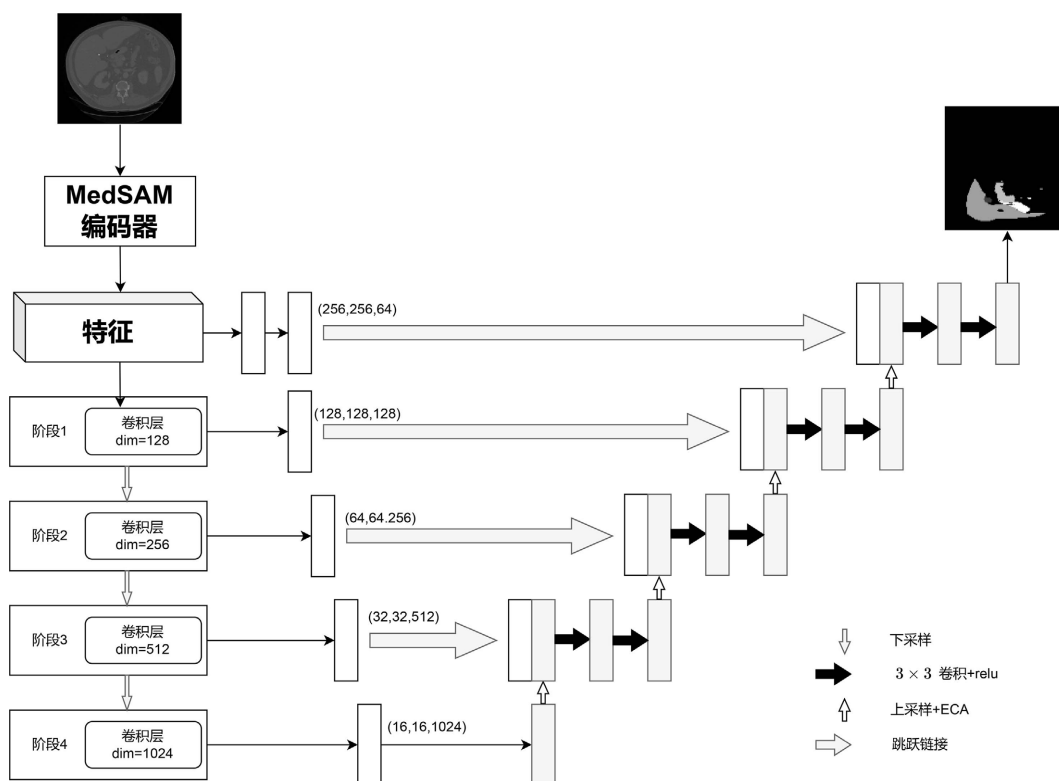


Figure 2. Model architecture
图 2. 模型架构

3.3. 高效通道注意力

高效通道注意力的核心思想是通过局部跨通道交互来计算通道权重, 如图 3 所示, 首先对输入特征图的各个通道进行全局平均池化, 然后使用一维卷积进行局部跨通道交互, 一维卷积核的大小通过自适应方法确定:

$$k = \frac{\lfloor \log_2(C) + b \rfloor}{\gamma} \quad (1)$$

其中, C 是通道数, b 和 γ 是超参数。随后, 将通道权重与输入特征图相乘, 得到加权后的特征图。

4. 实验设计

4.1. 超参数

本文实验在 Python 3.8、CUDA 12.0、Pytorch 2.2.0、Nvidia RTX 4090 显卡环境下运行。超参数设置如表 1 所示。

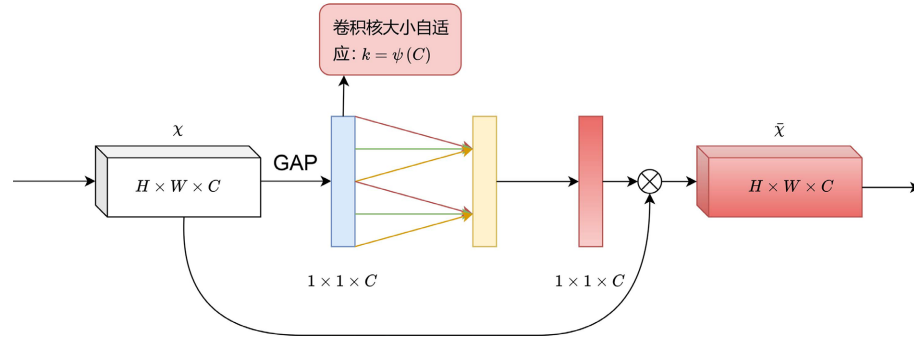


Figure 3. ECA
图 3. 高效通道注意力

Table 1. Hyper-parameter setting
表 1. 超参数设置

优化器	初始学习率	Epochs	Batch Size
AdamW	0.001	300	16

4.2. 损失函数

为了直观展示模型在胰腺病变图像分割任务中的性能表现, 我们采用 CeDice 损失函数。CeDice 损失函数是一种结合了 Cross-Entropy (交叉熵) 和 Dice Loss (Dice 损失) 的混合损失函数, 能够有效解决类别不平衡问题并提高模型的分割性能。

$$L_{Ce} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log(\hat{y}_{i,c}) \quad (2)$$

$$L_{Dice} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \quad (3)$$

$$L_{CeDice} = \alpha L_{Ce} + \beta L_{Dice} \quad (4)$$

其中, N 和 C 分别代表样本总数和类别总数, y_i 和 $\hat{y}_i$ 分别代表真实标签和预测标签, $\hat{y}_{i,c}$ 表示模型预测样本 i 属于类别 c 的概率, α 和 β 表示损失函数的权重, 默认值均为 1。

4.3. 评价指标

极度不平衡的分割在医学图像分析中十分普遍, 本文的胰腺病变分割任务便是其中之一[10]。病变区域通常远小于背景, 常规的交叉熵损失函数应用于此类任务不具备优势, 因此, 本实验采用 Dice 相似系数(DSC)以及 95%豪斯多夫距离(HD95)作为评价指标。DSC 是一种用于衡量两个样本(通常是分割结果和真实标签)之间相似度的指标。其取值范围为[0, 1], 值越接近 1, 表示分割结果与真实标签的重叠度越高。DSC 主要用于衡量分割结果与真实标签之间的重叠程度。本实验目标区域通常较小, DSC 能够有效评估小目标的分割效果。其公式如下:

$$DSC(P, G) = \frac{2 \times |P \cap G|}{|P| + |G|} \quad (5)$$

其中 P 表示预测的分割标签, G 表示真实的标签。

HD 是一种用于衡量两个点集之间最大距离的指标, 常用于评估分割结果的边界精度。 HD 值越小,

表示分割结果与真实标签的边界越接近。 HD 对分割结果中的异常值(如孤立的错误预测点)非常敏感, 能够反映分割结果的边界精度, 对分割结果的形状和位置变化敏感。其公式如下:

$$HD(P, G) = \max \left\{ \sup_{p \in P} \inf_{g \in G} d(p, g), \sup_{g \in G} \inf_{p \in P} d(g, p) \right\} \quad (6)$$

其中, $d(p, g)$ 表示点 p 和点 g 之间的欧氏距离。 HD 值越小, 表示预测边界和真实边界之间的吻合度越高。

5. 实验结果与分析

5.1. 对比实验

为了验证本文所提模型的有效性, 本研究与常用的 U-Net [6]、V-Net [11]、Attention U-Net [7]、TransUnet [12]、Res U-Net [13]以及 Deeplabv3+ [14]模型在 MSD_Pancreas 数据集上进行对比分析。这些模型均是医学图像分割领域的经典和前沿方法, 具有广泛的代表性。对比结果见表 2, 本文方法均优于对比方法, DSC 分数和 HD95 值分别为 82.7%和 10.2 mm, 相比于基线 U-Net 方法提升了 4.4%和 2.3 mm, 表明本文方法在分割精度上具有显著优势。

Table 2. Comparison of experimental results
表 2. 对比实验结果数据

方法	DSC (%)	HD95 (mm)	Parameters
U-Net	78.3	12.5	31M
V-Net	79.8	11.8	76M
Attention U-Net	80.1	11.2	34M
TransUnet	82.5	10.5	105M
Res U-Net	81.2	10.9	35M
Deeplabv3+	79.5	11.5	45M
Ours	82.7	10.2	53M

5.2. 可视化

为直观展现各模型在胰腺病变图像上的分割效果, 我们选取 pancreas_15 的第 33 张切片进行可视化展示, 如图 4 所示, 我们的模型对胰腺目标边缘的分割最接近于真实标签, 且对病变区域的细节捕捉也优于选定的对比模型。

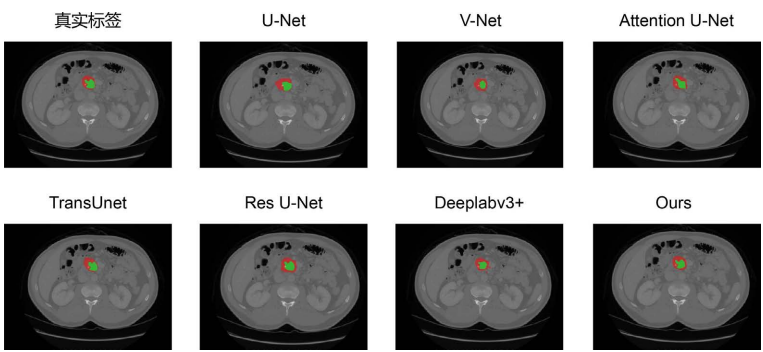


Figure 4. Visualization comparison
图 4. 可视化对比

5.3. 消融实验

为了进一步验证模型各模块的有效性, 设计了三组消融实验, 分别验证 MedSAM 编码器和高效通道注意力模块在模型中的贡献。实验结果如下表 3 所示, 当模型仅加入高效通道注意力模块时, DSC 分数相比于 HD 值提升更大; 而当模型仅加入 MedSAM 编码器模块时, HD 值相较于 DSC 分数提升更为显著。根据实验结果可知, MedSAM 编码器能够提升边界分割能力, 而高效通道注意力能够提升标签预测的精度。通过该实验可知引入 MedSAM 编码器和高效通道注意力的有效性。

Table 3. Results of ablation experiments

表 3. 消融实验结果

实验编号	MedSAM 编码器	高效通道注意力	DSC (%)	HD (mm)
1	×	√	82.1	11.8
2	√	×	82.2	10.5
3	√	√	82.7	10.2

5.4. 局限性和改进方向

现有胰腺病变数据集在病变类型方面缺乏多样性且数据量较少, 影响了模型的泛化能力; 模型在胰腺病变分割数据集上训练和推理需要大量的计算资源, 本文方法的参数量达到了 53 M, 导致了模型推理速度慢, 难以满足实时性要求高的场景。未来将进一步研究如何优化模型结构, 减少模型参数量, 并采用数据增强技术扩充数据集, 提高模型的泛化能力。

6. 总结

本文提出了一种融合 MedSAM 编码器和高效通道注意力的 U 型胰腺病变分割网络。使用 MedSAM 编码器, 旨在有效捕捉医学图像中的复杂结构和细节, 为后续分割提供高质量的特征表示。在采样过程中加入高效通道注意力模块, 旨在增强通道间的依赖关系, 使网络更关注重要特征, 抑制不重要的信息。在 MSD_Pancreas 数据集上的实验结果表明, 该模型在 Dice 相似系数(DSC)和 95%豪斯多夫距离(HD95)上均优于基线 U-Net 和其他对比模型。通过消融实验, 证明了融合 MedSAM 编码器和高效通道注意力模块的有效性。这一改进为胰腺病变分割以及医学图像分割提供了新的思路。未来工作将进一步优化模型, 提升模型的泛化能力, 并减少模型参数量。

参考文献

- [1] 杜楠, 魏妙艳, 徐近. 胰腺癌的靶向治疗进展及前沿[J]. 中国临床药理学与治疗学, 2025, 30(2): 183-192.
- [2] Ma, J., He, Y., Li, F., Han, L., You, C. and Wang, B. (2024) Segment Anything in Medical Images. *Nature Communications*, 15, Article No. 654. <https://doi.org/10.1038/s41467-024-44824-z>
- [3] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/cvpr.2018.00745>
- [4] Woo, S., Park, J., Lee, J. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Computer Vision—ECCV 2018*, Springer, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1
- [5] Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W. and Hu, Q. (2020) ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 11531-11539. <https://doi.org/10.1109/cvpr42600.2020.01155>
- [6] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation.

-
- Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference*, Munich, 5-9 October 2015, 234-241.
- [7] Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Rueckert, D., *et al.* (2018) Attention U-Net: Learning Where to Look for the Pancreas. arXiv: 1804.03999.
 - [8] 李昊轩, 窦凤岐. 基于深度学习的胰腺肿瘤分割算法研究[J]. 电脑编程技巧与维护, 2023(8): 119-121.
 - [9] 张国栋, 唐晓艺, 鞠蓉晖, 等. 融合双注意力机制的多尺度胰腺分割方法[J]. 计算机工程与设计, 2024, 45(4): 1189-1194.
 - [10] 生琳, 王朝立. 基于改进 MaxViT 的胰腺肿瘤图像分割[J]. 软件, 2024, 45(6): 13-15.
 - [11] Milletari, F., Navab, N. and Ahmadi, S. (2016) V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. 2016 *Fourth International Conference on 3D Vision (3DV)*, Stanford, 25-28 October 2016, 565-571. <https://doi.org/10.1109/3dv.2016.79>
 - [12] Chen, J.N., *et al.* (2021) TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. arXiv: 2102.04306.
 - [13] Diakogiannis, F.I., Waldner, F., Caccetta, P. and Wu, C. (2020) ResUNet-a: A Deep Learning Framework for Semantic Segmentation of Remotely Sensed Data. *ISPRS Journal of Photogrammetry and Remote Sensing*, **162**, 94-114. <https://doi.org/10.1016/j.isprsjprs.2020.01.013>
 - [14] Chen, L., Zhu, Y., Papandreou, G., Schroff, F. and Adam, H. (2018) Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Computer Vision—ECCV 2018*, Springer, 833-851. https://doi.org/10.1007/978-3-030-01234-2_49