

新型架构下的密集网络在肺部影像的分割研究

陈 亮^{1,2}, 高文根^{1,2}, 张 晨^{1,2}, 陈 东^{1,2}

1. 安徽工程大学 电气工程学院, 安徽 芜湖 241000

2. 安徽工程大学 高端装备先进感知与智能控制教育部重点实验室, 安徽 芜湖 241000

摘 要:在医学图像分割领域中,肺实质的分割对肺结节检测有着至关重要的作用,在考虑到模型参数量的情况下追求更高的精度一直是研究热点之一;为此提出了新的三层密集卷积神经网络 DA-UNet,首先用密集卷积模块代替在传统 U-Net 使用的普通 3×3 卷积,利用密集卷积特征重用特点,加强了网络的特征提取能力。再者在没有太过影响分割网络精确度的前提下加以修剪,减少了上下采样次数,减少不必要的算力消耗。此外,使用了注意力门(Attention gate),加强了跳跃连接中高底层信息融合效果,并且使用空洞空间金字塔池化(Atrous spatial pyramid pooling),模型加入了不同尺度的特征信息,进一步加强图像中任务相关的区域特征,有效减小噪声干扰,提高网络分割精度。通过实验证明:三次上下采样改进模型的参数量只有传统四次上下采样的 75.2% 左右,但是分割效果没有太大的影响,用 LUNA 竞赛肺部影像数据集进行了分割验证,实验结果在测试集上的准确率达到 0.991,而 *IoU* 则为 0.961,比起传统 U-Net 的评价指标 *IoU* 提升了 2.9%;在泛化实验的肝脏图像中,DA-UNet 的 *IoU* 稳定在 0.929 左右,而 U-Net 稳定在 0.838 左右。这些结果证明了改进的 U-Net 有更佳的分割效果。

关键词:U-Net;密集网络;肺实质分割;空洞空间金字塔池化;注意力门;DA-UNet;评价指标

中图分类号:TP391.4 **文献标识码:**A **doi:**10.16055/j.issn.1672-058X.2023.0006.007

Segmentation of Lung Images with Dense Network Models Under a New Architecture

CHEN Liang^{1,2}, GAO Wengen^{1,2}, ZHANG Chen^{1,2}, CHEN Dong^{1,2}

1. School of Electrical Engineering, Anhui Polytechnic University, Anhui Wuhu 241000, China

2. Key Laboratory of Advanced Sensing and Intelligent Control for High-end Equipment, Ministry of Education, Anhui Polytechnic University, Anhui Wuhu 241000, China

Abstract: In the field of medical image segmentation, the segmentation of lung parenchyma plays a crucial role in lung nodule detection. Pursuing higher accuracy while considering the number of model parameters has been one of the research hotspots. In this regard, a new three-layer dense convolutional neural network, DA-UNet, was proposed. Firstly, a dense convolution module was used instead of the normal 3×3 convolution used in traditional U-Net, which enhances the feature extraction capability of the network by taking advantage of the reuse feature of dense convolution. Furthermore, the network was pruned without affecting the accuracy of the segmentation network too much, reducing the number of up and down samples, and reducing unnecessary computational power consumption. In addition, the Attention Gate was used to enhance the effect of information fusion between high and low levels in the jumping connection, and atrous spatial pyramid pooling was used. The model incorporated feature information at different scales to further enhance task-related regional features in the images, effectively reducing noise interference and improving network segmentation

收稿日期:2022-09-29 **修回日期:**2022-10-29 **文章编号:**1672-058X(2023)06-0053-08

基金项目:国家自然科学基金项目(61976005)。

作者简介:陈亮(1996—),男,安徽马鞍山人,硕士,从事图像分割研究。

通讯作者:高文根(1973—),男,安徽芜湖人,教授,硕士生导师,从事新能源系统控制与优化、智能控制与新型检测技术研究。

Email: ahpuchina@ahpu.edu.cn.

引用格式:陈亮,高文根,张晨,等. 新型架构下的密集网络在肺部影像的分割研究[J]. 重庆工商大学学报(自然科学版), 2023, 40(6): 53—60.

CHEN Liang, GAO Wengen, ZHANG Chen, et al. Segmentation of lung images with dense network models under a new architecture[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2023, 40(6): 53—60.

accuracy. It was demonstrated experimentally that the number of parameters of the three-time up-and-down sampling improved model was only about 75.2% of those of the traditional four-time up-and-down sampling, but the segmentation effect was not much affected. The LUNA competition lung image data set was used for segmentation verification. The experimental results showed that the accuracy reached 0.991, and the IoU was 0.961. Compared with the traditional U-Net, the evaluation index IoU was improved by 2.9%. In the liver images of the generalization experiment, the IoU of DA-U-Net was stable at about 0.929, while that of U-Net was stable at about 0.838. These results prove that the improved U-Net has better segmentation results.

Keywords: U-Net; dense network; lung parenchyma segmentation; atrous spatial pyramid pooling; attention gate; DA-U-Net; evaluation index

1 引言

随着计算机视觉的发展,对图片的分类、分割和检测的能力也越来越强。语义分割就是对目标图像进行理解与推断,将分割任务转化为对像素点的分类。而对肺部图像的分割就是对肺实质和肺部背景的像素点分类,突出目标肺实质区域。

肺实质的分割对肺结节检测有着至关重要的作用,尽早检测出肺结节的性质,对预防和医治肺癌都有着举足轻重的影响^[1]。除了医生手动分割的方法外,还有传统机器学习方法和深度学习方法。在深度学习研究中,全卷积网络 FCN(Fully Convolutional Networks)可以说是开山之作,它是由 Long 等提出,开创了语义分割的解编码结构^[2],该模型通过上采样和下采样对图像的每一个像素点进行分类。不过这种操作会在池化过程中丢失部分特征信息,而 U-Net 的提出比 FCN 稍晚一点,它独特的解编码和跳跃连接结构,使它在医学图像分割上有着出色的表现^[3]。这也使得研究者对这种 U 型网络有着持续的研究。由于医学图像的数据集规模一般都比较小,难以发挥深度学习网络模型在医学图像分割领域中的优势,所以进一步改进这种 U 型网络使其能适应像医学图像这种小规模数据集便成了研究者们研究的方向之一。文献[4]中 Res-U-Net 用于视网膜血管的分割,由 Xiao 等在 2018 年提出,该模型本质上就是经典的 Resnet 和 Attention 结合,残差连接在梯度传播上解决叠加层的局限性,而 Attention 部分属于硬注意力机制,只针对视网膜提高了关注度,而血管部分还是有所缺失。文献[5]中的 Attention U-Net 用于胰脏图像的分割,由 Oktay 等在 2018 年中期提出,该模型本质上就是加入软注意力机制,加大了局部 ROI 的权重,有效抑制无关区域的干扰。Attention U-Net 中注意力机制与 Res-U-Net 中的不同,使得 Attention U-Net 有着更广的应用场景,但是该模型的复杂度要更高一些。文献[6]中的 U-Net++ 用于肝脏等医学影像的分割,由 Zhou 等在 2019 年提出,为了解决梯度消失问题利用了密集跳跃连接,深度监控有效辨别粗细粒度特征。但是该模型并没有提取到多尺度的信息,多尺度的原始特征往往因为过度矫正而无法充分利用。

上述各种改进 U 型模型都为四层(四次下采样),而本文提出了一种新的三层卷积神经网络模型 DA-U-Net,用于肺实质的分割,其中充分利用密集连接网络,且使用补偿模块的空洞空间金字塔和注意门^[7-9],兼顾边缘局部和全局的特征信息,使得分割精度进一步提升,且新型三层分割网络对比于其他改进网络参数量较少,减少了不必要的算力消耗,在实验数据集上有着更佳的分割效果。

2 相关工作

2.1 架构设计依据

在医学分割领域,U-Net 一直是分割效果不错的模型之一。经过一些调研,发现 U-Net 作者也只是凭经验设计的网络层数,并没有具体的依据和说明。所以说这些网络层数也不是一成不变的,为了保证深度神经网络的性能增益,可以根据现实需要增减,提出了两个新的约束在深度结构设计中:

第一个约束条件是对于每一卷积层,都要保证其学习更复杂模式的能力,也就是学习能力。

由图 1 可以看出,图 1(a)中生成 pattern AB,而图 1(b)中卷积层不能很好地学习到 pattern AB 和它们的关系,因为卷积核的尺寸小于 A 特征与 B 特征响应的距离,卷积层就不能学习到较为复杂的特征。解决的办法有 3 种:

- (1) 使用更大的卷积核;
- (2) 通过下采样降低响应距离;
- (3) 使用空洞卷积。

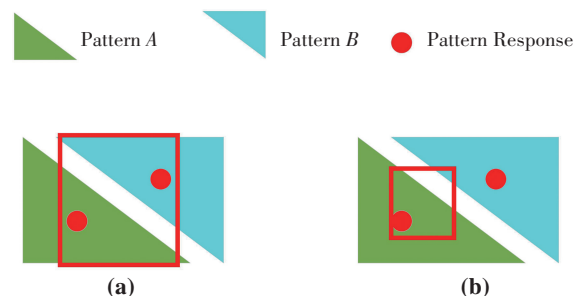


图 1 卷积图

Fig. 1 Convolution graph

为了定量测量卷积层学习能力的大小,定义了 c-

value,第一个约束数学公式如下:

$$c\text{-value} = \frac{\text{Real Filter Size}}{\text{Receptive Field Size}} \quad (1)$$

$$\frac{2^l k}{\sum_{i=1}^l 2^{i-1}(k-1)a_i} \geq t, l=1,2,\dots,n \quad (2)$$

如式(1)所示,在卷积中 Real Filter Size 是实际卷积核的尺寸,Receptive Field Size 是局部感受野的尺寸,这种关系用数学表达式显示就如式(2)所示, l 为下采样次数, k 为卷积核的尺寸。 $2^l k$ 是第 l 个 stage 的 Real Filter Size, $\sum_{i=1}^l 2^{i-1}(k-1)a_i$ 是第 l 个 stage 的最后一层的 Receptive Field Size, t 是 $c\text{-value}$ 的最小值, $t=1/6$ 对于不同任务的所有卷积层是一个很好的 $c\text{-value}$ 下界。

第二个约束条件是顶层的感受野应该不大于图像区域。卷积层的加深,感受野慢慢到达输入边界,是为了学习新的目标特征信息。但是如果感受野超过了输入图像的尺寸,容易产生过拟合,导致分割效果变差。第二个约束数学公式如下:

$$\sum_i 2^{i-1}(k-1)a_i \leq z \quad (3)$$

如式(3)所示,左边的式子是最顶层卷积层的 Receptive Field, $2^{i-1}(k-1)$ 是第 i 个 stage 每一层的 Receptive Field 的增量, $2^{i-1}(k-1)a_i$ 是第 i 个 stage 的 Receptive Field 的总增量,右边的式子图像尺寸为 z 。由上述可见,本文提出的 DA-UNet 网络结构,三次上下采样加上空洞空间金字塔 ASPP,符合两个约束条件。其中的关键就是感受野的大小,输入图像中任何一个单元的感受野以外的地方都不会影响该单元的值,因此有必要仔细控制感受野,以确保它覆盖又不超过整个相关的图像区域。

2.2 密集模块 (Dense Block)

为了追求更高的准确率,自 AlexNet 以来,神经网络倾向于更深,更复杂的结构设计。比如上述的各种 U-Net 改进模型。这种卷积神经网络对肺实质的分割,就是将肺实质和肺部背景的像素点分割开来,之后对图像像素点的灰度值进行配对,凸显出完整的肺实质区域。本文所使用的密集网络主要包括密集模块和过渡层。

如图 2 所示,原本 U-Net 中的普通卷积在 L 层的输出为

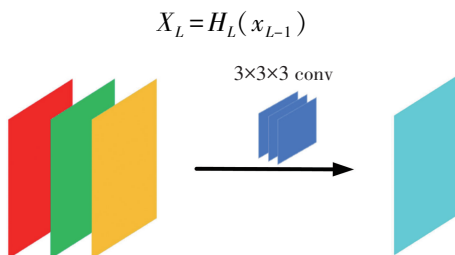


图 2 标准卷积

Fig. 2 Standard convolution

如图 3 所示,为了进一步改善层之间的信息流,提出了一种密集的连接模式,在 Dense Block 中的非线性组合函数在 L 层的输出为

$$X_L = H_L([x_0, x_1, x_2, \dots, x_{L-1}])$$

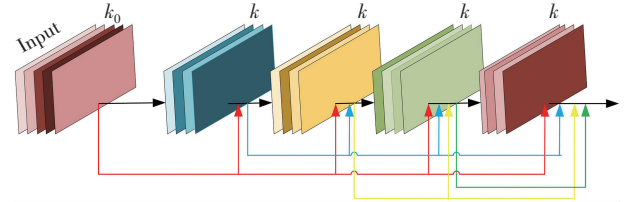


图 3 密集模块

Fig. 3 Dense module

传统的卷积是对前一层的 feature maps 进行处理,而 Dense Block 里的密集跳跃连接结构是对前面所有层 feature maps 的处理,在 L 层的输出的特征图数量公式:

$$K_L = K_0 - (L-1)K$$

在密集模块中,早期卷积层通过信道级联连接到所有后续层。这意味着密集块中每个层的输入都是之前所有层连接在一起的输出。基本上,每个层都会根据以前层获得的“集体知识”学习其他特征图。该策略通过特征重用提高了网络的表示能力。在早期层中学习的特征将被传递,从而避免了冗余特征的学习,增强了信息流动。在 Dense Block 中由于后面层的输入会非常大,内部可以采用 Dense neck 层来减少计算量,它起到的作用是降低特征数量,提升计算效率。主要是在原有结构上增加 1×1 conv,即 BN+Relu+ 1×1 conv+BN+Relu+ 3×3 conv,称为 Dense net-B 结构。

如图 4 所示,Transition 层是连接两个 Dense Block 的过渡层,在此层中可以压缩图像,和传统 U-Net 的下采样 Max pooling 有一样的效果,结构为 BN+Relu+ 1×1 conv+ 2×2 Avg pooling。

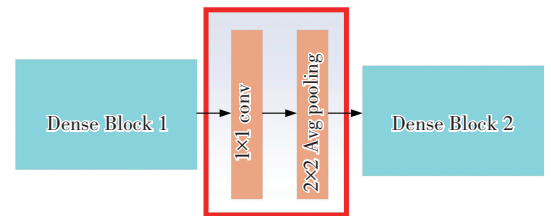


图 4 过渡层

Fig. 4 Transition layer

2.3 空洞空间金字塔池化 (ASPP)

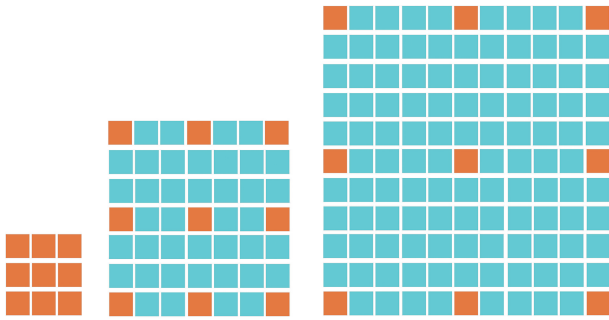
空洞空间金字塔池化,是由 Chen 等在 DeepLabv3+ 语义分割网络中首次提出,在 ASPP 中最为关键的是使用了空洞卷积和深度可分离卷积的概念。深度可分离卷积就是把普通卷积的乘法运算,因式分解成先乘再

相加的运算,这样可以大大减少模型所需的参数量和计算量。它的工作原理是将以往普通卷积操作同时考虑通道和区域,改成卷积先只考虑区域,然后再考虑通道。实现了通道和区域的分离。

空洞卷积如图 5 所示,普通的卷积过程(dilation rate = 1),这时,3×3 卷积后的感受野为 3,而空洞卷积的感受野的公式为

$$k' = (d-1) \times (k-1) + k \quad (3)$$

由式(3)可知,在图 5 中扩张率=1 时,等同于普通的 3×3 卷积,而扩张率=3 时,感受野为 7,扩张率=5 时,感受野为 11。由此可以看出在相同的特征图下,更大的扩张率,可以得到更加密集的特征信息,但是也不能一味地加大膨胀系数,越大的空洞卷积所带来的网格效应也就越明显,所以必须按照实际分割的特征图大小和影像包含信息的实际情况设置空洞空间金字塔池化里扩张卷积的膨胀系数。



(a) 扩张率=1 (b) 扩张率=3 (c) 扩张率=5

图 5 空洞卷积

Fig. 5 Atrous convolution

空洞空间金字塔如图 6 所示,它的工作流程是输入图像经过五次处理后再叠加在一起。第一次处理是,输入图像经过 1×1×64 的卷积操作,加上 BN 操作后得到了第一个特征图,第二次到第四次的处理是,输入图像经过 3×3×64 的卷积操作,不过这里的卷积操作是深度可分离卷积,可以在扩大感受野的同时减少参数量和计算量。由式(3)可知,膨胀系数分别是 4,8,12 的 3×3 空洞卷积,当膨胀系数为最大的 12 时,这时卷积的感受野是 25,而经过三次下采样的输入图像只有 32×32,此时的感受野接近又不超过整个相关的图像区域。其网络结构可表示为 conv(3×3)+BN+conv(1×1)+BN+ReLU。值得注意的一点是,为了使输出的特征图大小一致,其中膨胀系数 rate 等于 padding 系数,由于边缘效应的影响(大的膨胀率下 padding 随之增加)所以第五次的图像处理为全局平均池化,减轻边缘效应的影响,得到有全局特征的图片。

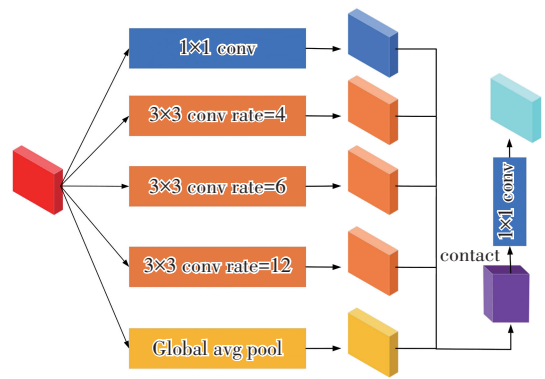


图 6 空洞空间金字塔池化

Fig. 6 Atrous spatial pyramid pooling

2.4 反卷积(Transposed Convolution)

在语义分割模型中,在输入图片经过池化下采样之后,需要进行与正常卷积相反的转换,将尺寸变小的图片恢复到原本输入图片一样的大小,这个过程也就是上采样。当输入图像在编码时,是将图片逐步缩小,获得不同层次的 feature maps,解码时是将缩小后的图片利用反卷积操作再逐步放大。正向卷积是输入图像根据卷积核的大小,按照一定比例补 0 之后旋转卷积核,进行卷积操作。用矩阵公式来说,假设输入图片的像素矩阵为 X ,输出矩阵为 Y ,那么矩阵运算公式为

$$Y = CX$$

而反向卷积就是正向卷积的逆运算,用矩阵运算来说的话,就是通过 C 和 Y 得到 X ,这是简单的逆运算过程,所以具体的计算公式为:

$$X = C^T Y$$

3 整体架构

3.1 算法流程

本模型算法框架保留 U-Net 独特的解编码和跳跃连接操作结构,并在其结构上进行模型修剪和改进,把原本的四层上下采样,削减成三次。DA-UNet 如图 7 所示,其中左侧为编码路径,包含三层六个 Dense Block 密集卷积模块和三个 Transition 过渡层,也就是下采样层,目的是为了通过逐步缩小特征图尺寸来学习更多语义分割信息,右侧是解码路径,也有三层六个 Dense Block 密集卷积模块和三个反卷积上采样模块,目的是为了恢复由下采样导致的逐步缩小特征图,并且使用了注意力解码模式(AG)为了抑制图片中的背景像素对分割效果的影响,突出包含关键信息的目标区域。此外,“U”形网络底部的加入空洞空间金字塔(ASPP),为了解决因肺部影像中肺实质与其余组织器官的相似性,导致实际分割效果不理想,特别是图像边界的分割上的问题。

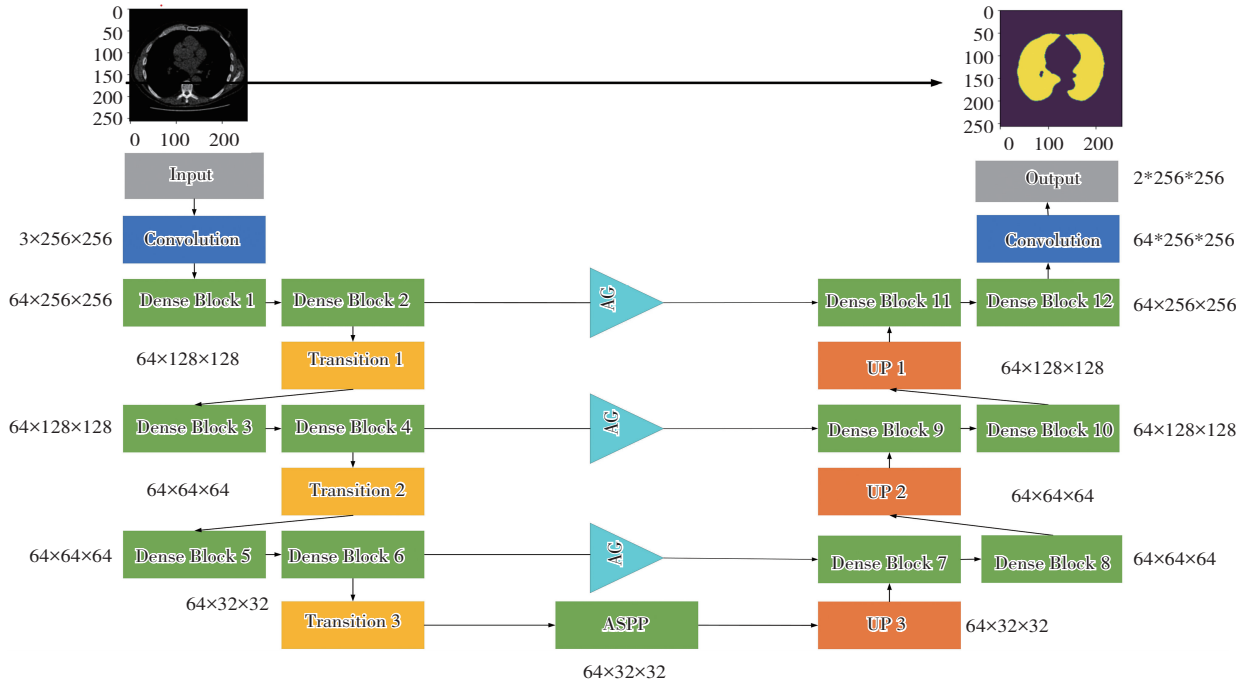


图 7 整体模型

Fig. 7 Overall model

编码阶段:该模型有三次下采样,为了提高分割模型充分提取有用特征信息和利用上下文信息的能力,并且在不过分加深网络而导致参数量急剧增加的前提下,受 Dense Net 密集检测网络的启发下,使用两个 Dense Block 卷积模块代替原本的两次 3×3 普通卷积,其内部的密集跳跃连接结构可以更有效地提取和利用有效信息。而感受野的问题,受 DeepLabv3+采用空间金字塔池模块提醒^[10],能够通过以多种空洞卷积和多个有效视野传入特征来编码多尺度上下文信息,扩大卷积核感受野。将深度可分离卷积应用于空洞空间金字塔池模块,从而形成更快、更强的编码网络。

解码阶段:DA-UNet 采用编解码结构进行语义分割,解码过程也就是图像恢复过程,对应三次下采样,解码部分也有三次上采样,上采样的方法是反卷积,反卷积是使用 2×2 的卷积,步长为 2。同样的是使用两个 Dense Block 卷积模块代替原本两次 3×3 普通卷积,而为了更好地还原图像,使用了一种新的用于医学成像的注意力门 (AG) 模型^[11-13]。

可以看出,图 8 中 g 是解码阶段的高级特征图, x^l 为编码阶段的低级特征图, W_g 是门控信号, W_x 为注意力门。

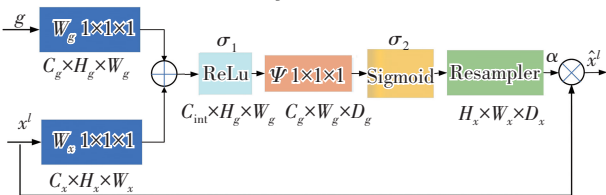


图 8 注意力门

Fig. 8 Attention gate

$$q_{attn}^l = \Psi^T(\sigma_1(W_x^T x_i^l + W_g^T g_i + b_g)) + b_\psi \quad (5)$$

如式(5)所示,首先 g, x^l 都会经过 1×1 卷积,目的

是为了保证两个特征图有同样的尺寸和通道数,然后两者相加,这时高层信息的目标区域的值会变大。再使用 ReLU 激活函数,经过 Ψ 的 1×1 卷积通道数变为 1。

之后如式(6)所示,利用 sigmoid 函数和双线性插值的重采样恢复与 x^l 相同的特征图大小,生成注意力系数表 α 。

$$\alpha = \sigma_2(q_{attn}^l(x_i^l, g_i)) \quad (6)$$

最后输入特征(x^l)通过在 AG 中计算的注意系数 α 进行相乘,此时目标区域的值就会变大,从而抑制无关区域。得到新的浅层特征图 $\hat{x}^l$ 。

3.2 损失函数(CE)

交叉熵损失 Cross Entropy Loss,也叫 Log Loss,最初的交叉熵公式如下所示。交叉熵损失函数值越小,两种分布越相似,交叉熵就是衡量分布的相似程度。

$$L(y, f(x)) = -y \log(f(x))$$

在本文中分割出的二图片和与标签图片越相似,损失函数的值越小,对于二分类问题,交叉熵损失函数如下:

$$L = -[y \log + (1 - \hat{y}) \log(1 - \hat{y})]$$

4 实验结果与分析

4.1 实验材料

本文数据集采用了两种数据集,一种公共数据集是 2016 年 LUNA 肺结节检测数据集,其中包含 267 组双肺的原图和对应的标签图片,剔除了分割有误的图片,利用了 262 组图片和对应的标签图片,按照训练集验证集 9 : 1 的比例,其中训练集是 234 组图片,验证集是 28 组图片。为了测试本模型对其他数据集的泛化能力,使用了另外一种肝脏数据集,有 398 组图像及其对应的标签图像。其中 320 组为训练集,78 组为验证集。

4.2 参数设置

本模型搭建框架是 Pytorch, 用 Jupyter notebook 编写了相关模型代码, 实验室的电脑为 NVIDIA GTX3050TI, 训练和验证的 Batchsize 为 2, 数据集 1、数据集 2 的 Epoch 设置为 80, 本模型损失函数为 Cross Entropy Loss, Adam 优化器, 学习率为 0.001, 并使用学习率衰减, 每 7 轮学习率缩小 90%, 加快模型收敛速度。

4.3 评价指标

在图像分割任务中, 分割像素的准确率 (Accuracy, f_{ACC}), 和分割图像与实际的标签图像的交并比 (IoU), 是评价分割效果的重要指标, 在本文医学肺部图像分割中也采用了这两种评价指标。计算公式为:

$$IoU = \frac{|GT \cap SR|}{|GT \cup SR|} \quad (7)$$

式(7)中, GT 是标签图像的分割结果, SR 是三层 DA-UNet 模型分割结果。

$$f_{ACC} = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

式(8)中, 像素点是肺实质像素也被模型判定为肺实质像素点的数量称为 TP , 像素点是肺实质背景像素也被模型判定为肺实质背景像素点的数量称为 TN , 像素点是肺实质背景像素却被模型判定为肺实质像素点的数量称为 FP , 像素点是肺实质像素却被模型判定为肺实质背景像素点的数量称为 FN ^[14]。

4.4 实验结果

4.4.1 对比实验

为了证明改进模块对分割效果的提升, 做了对比实验, 并且使用了两种数据集, 1 是 LUNA 肺结节检测数据集, 2 是肝脏数据集。实验结果分别如下:

表 1 数据集 1 验证集结果

Table 1 Validation set results of dataset 1

Model	f_{ACC}	IoU
U-Net	0.976	0.932
三层 DA-UNet	0.991	0.961
四层 DA-UNet	0.992	0.962
Attention-U-Net	0.982	0.958
Res-UNet	0.986	0.956
U-Net++	0.987	0.957

模型除了层数上的区别, 还有空洞空间金字塔 ASPP 的膨胀系数不一样, 因为根据第一节架构设计中可知, 要想模型的实验效果比较好, 模型里卷积的感受野要做到接近又不超过整个相关的图像区域, 二层 DA-UNet 经过两次下采样, 最小像素是 64×64 , 所以两层的膨胀系数是 6, 12, 18, 感受野最大为 37, 三层 DA-UNet 经过三次下采样, 最小像素是 32×32 , 所以三层的膨胀系数是 6, 8, 12, 感受野最大为 25, 而四层的最小像素是 16×16 , 系数为 2, 4, 6, 感受野最大为 13, 避免应感受野过大带来的网络过拟合问题。如图 9 所示, 传统 U-Net 和改进的 DA-UNet 在数据集 1 的验证集上运行 80 轮的 IoU 结果。

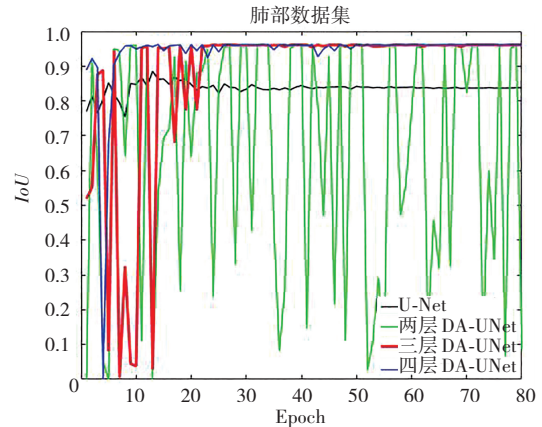


图 9 数据集 1 IoU 结果

Fig. 9 IoU results of dataset 1

由表 1 和图 9 可知, 三层和四层的 DA-UNet 的模型分割效果相差不多。本文提出三层 DA-UNet 模型为红线, 评价指标 IoU 稳定在 0.961 左右, 而黑线 U-Net 稳定在 0.932 左右。绿线是两层 DA-UNet 的分割结果, 很明显 IoU 结果不稳定, 难以收敛。可能的结果有几种, 比如下采样次数太少, 网络深度不够, 前两层的卷积没有提取到足够多的高阶信息, 其次使用空洞金字塔具有网格效应, 再者就是 Batchsize 太小, 无法一次性提取到太多的信息, 不过因为此次实验为消融对比实验, 不能考虑太多的变量, 或者实验室电脑性能有限, 本文就暂且不考虑这些问题, 待之后解决。

从表 2 可知, 当输入张量为 $1 \times 3 \times 256 \times 256$ 时, 参数量用科学计数法计数, 本模型在对比各种 U-Net 的改进版本中, 参数量都是最少的, 且三层和四层的 DA-UNet 的参数量相差 25% 左右, 所以出于算力消耗的方面考虑, 三层的 DA-UNet 模型是最佳的分割网络。

表 2 参数量对比结果

Table 2 Comparative results of parameter volumes

Model	参数量
两层 DA-UNet	3.012 100E6
三层 DA-UNet	4.496 549E6
四层 DA-UNet	5.980 998E6
Attention-U-Net	3.486 301 4E7
Res-UNet	8.240 669 0E7
U-Net++	3.660 781 0E7

表 3 为肝脏图片分割, 因为只是作为泛化能力验证, 所以出于计算力的考虑, 只做了两个模型实验。

表 3 数据集 1 验证集结果

Table 3 Validation set results of dataset 1

Model	f_{ACC}	IoU
U-Net	0.99	0.838
三层 DA-UNet	0.993	0.929

如图 10 所示, 在泛化实验的数据集 2 肝脏图像分割结果中, 实验经过 80 轮训练之后, 评价指标 IoU 有了明显的差距, DA-UNet 稳定在 0.929 左右, 而 U-Net 稳定在 0.838 左右。

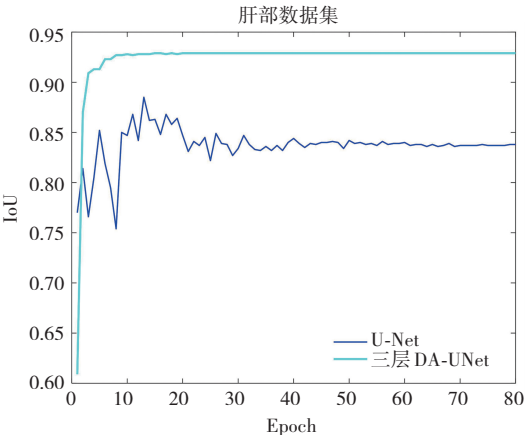


图 10 数据集 2 IoU 结果
Fig. 10 IoU results of dataset 2

4.4.2 分割效果图

如图 11 所示,这是使用两种数据集产生的分割图,A 和 B 是 LUNA 肺结节检测数据集,C 和 D 是肝脏的数据集,A 和 C 为传统 U-Net,A 中看红色框内很明显可以看出,有多处明显缺陷,特别是肺部末端都没有很好的分割,C 分割图像中红色框则分割区域不完整,所以传统 U-Net 分割效果不够优秀。而 B 和 D 为改进的 DA-UNet 网络,B 中在模糊边界上的分割效果还是有一定瑕疵不够平滑,同样 D 中也出现的同样问题。不过可以看到在整体肺实质的分割没有太大的缺陷。

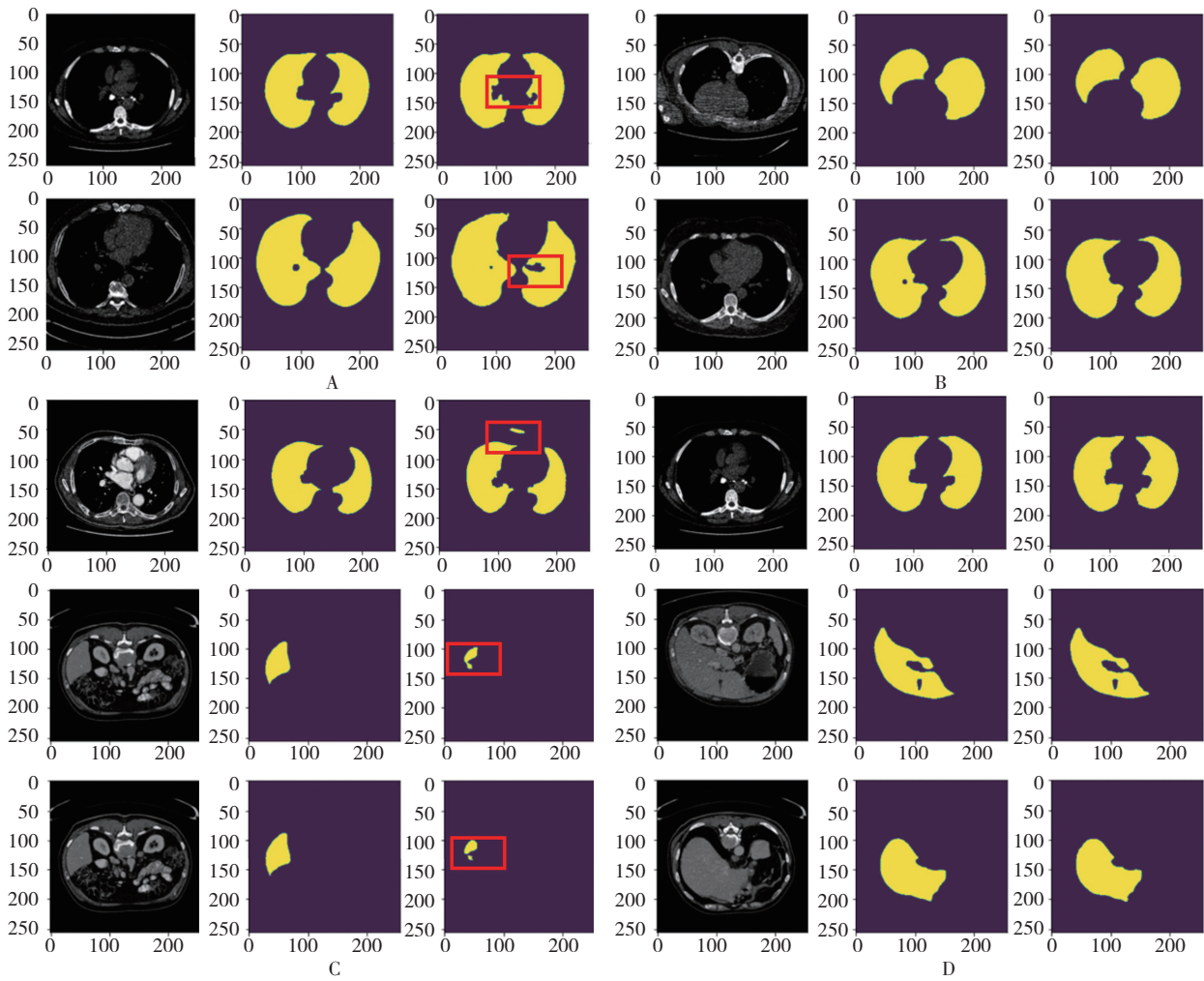


图 11 分割结果图
Fig. 11 Segmentation results

5 结 论

改进的 DA-UNet 网络三次上下采样的参数量只有传统四次上下采样的 75.2%左右,所以出于算力消耗的方面考虑,三层的 DA-UNet 模型是最佳的分割网络。把解编码网络从普通的卷积网络提升为特征重用的密

集卷积网络,以优化网络的特征提取能力,此网络不仅可以做到更深,并且因为 Dense Block 内的短的跳跃连接结构,可以解决随着网络层数增加而出现的网络退化问题。因为和跳跃连接中 U-Net 的简单复制拼接不同,注意门在解码操作之前被插入到编码路径提取的

特征图中,以提升跳跃连接中信息融合操作的效果。此外,“U”形网络底部的 Bottleneck 被空洞空间金字塔(ASPP)所取代,扩大感受野,获得上下文多尺度信息,加强了对全局信息的掌握^[15]。在肺部数据集实验中可以看到三层 DA-UNet 也就是三次下采样的 DA-UNet,运行 80 轮之后在测试集上准确率达到了 0.991,而 *IoU* 则为 0.961,比起传统 U-Net 的评价指标 *IoU* 提升了 2.9%;在泛化实验的肝脏图像中,DA-UNet 的 *IoU* 稳定在 0.929 左右,而 U-Net 稳定在 0.838 左右。从实验结果可以看出,在两种数据集中无论是收敛速度还是稳定性,改进模型都有着不错的表现。

参考文献(References):

- [1] JEMAL A, THUN M J, RIES L, et al. Annual report to the nation on the status of cancer, 1975—2005, featuring trends in lung cancer, tobacco use, and tobacco control [J]. Journal of the National Cancer Institute, 2008, 100(23): 1672—1694.
- [2] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 39(4): 640—651.
- [3] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer-assisted Intervention. Berlin: Springer, 2015: 234—241.
- [4] XIAO X, LIAN S, LUO Z, et al. Weighted ResUNet for high quality retina vessel segmentation[C]// International Conference on Information Technology in Medicine and Education (ITME). IEEE, 2018: 327—331.
- [5] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-Net: Learning where to look for the pancreas[J]. Chinese Journal of Medical Physics, 2018, 39(9): 1133—1139.
- [6] ZHOU Z W, SIDDIQUEE M M R, TAJBAKHS N, et al. UNet++: A nested U-Net architecture for medical image segmentation[C]//Proceedings of the 4th International Workshop on Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Granada: Springer, 2018: 3—11.
- [7] FU K, LIU S, LUO X, et al. Robust point cloud registration framework based on deep graph matching[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2021: 8893—8902.
- [8] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 7132—7141.
- [9] 朱凯俊. 基于改进的 FCM 算法对图像分割的研究和应用[J]. 重庆工商大学学报(自然科学版), 2022, 39(5): 24—33.
ZHU Kai-jun. Research and application of image segmentation based on improved FCM algorithm[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2022, 39(5): 24—33.
- [10] 邓瑞, 林金朝, 杨宏志. 基于深度学习的建筑物识别[J]. 重庆工商大学学报(自然科学版), 2019, 36(4): 17—22.
DENG Rui, LIN Jin-chao, YANG Hong-zhi. Building recognition based on deep learning[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2019, 36(4): 17—22.
- [11] MO Y L, SINOPOLI B. Secure control against replay attacks[C]// 47th Annual Allerton Conference. Monticello: IEEE, 2009: 911—918.
- [12] ZHEN M, WANG J, ZHOU L, et al. Joint semantic segmentation and boundary detection using iterative pyramid contexts[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2020: 13666—13675.
- [13] QI CR, SU H, MO K, et al. Pointnet: Deep learning on point sets for 3D classification and segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017: 652—660.
- [14] 钱宝鑫, 肖志勇, 宋威. 改进的卷积神经网络在肺部图像上的分割应用[J]. 计算机科学与探索, 2020, 14(8): 1358—1367.
QIAN Bao-xin, XIAO Zhi-yong, SONG Wei. Segmentation application of improved convolutional neural network on lung images[J]. Computer Science and Exploration, 2020, 14(8): 1358—1367.
- [15] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834—848.