

深度学习下的医学影像分割算法综述

彭 璟^{1,2}, 罗浩宇^{1,2}, 赵淦森^{1,2}, 林成创^{1,2}, 易序晟^{1,2}, 陈少洁^{1,2}

1. 华南师范大学 计算机学院, 广州 510663

2. 广州市云计算安全与测评技术重点实验室, 广州 510663

摘要: 医学影像分割是计算机视觉在医学影像处理中的一个重要应用领域, 其目标是从医学影像中分割出目标区域, 为后续的疾病诊断和治疗提供有效的帮助。近年来深度学习技术在图像处理方面取得了巨大进展, 基于深度学习的医学影像分割算法逐渐成为该领域研究的重点和热点。叙述了计算机视觉下的医学影像分割任务及其难点, 重点综述了基于深度学习的医学影像分割算法, 对当前具有代表性的相关方法进行了分类和总结, 介绍了医学影像分割算法常用的评价指标和数据集。对该技术的发展进行了总结和展望。

关键词: 深度学习; 计算机视觉; 医学影像; 图像分割

文献标志码: A **中图分类号:** TP391 **doi:** 10.3778/j.issn.1002-8331.2010-0335

Survey of Medical Image Segmentation Algorithm in Deep Learning

PENG Jing^{1,2}, LUO Haoyu^{1,2}, ZHAO Gansen^{1,2}, LIN Chengchuang^{1,2}, YI Xusheng^{1,2}, CHEN Shaojie^{1,2}

1. School of Computer Science, South China Normal University, Guangzhou 510663, China

2. Guangzhou Key Lab on Cloud Computing Security and Assessment Technology, Guangzhou 510663, China

Abstract: Medical image segmentation is an important application area of computer vision in the medical image processing, its goal is to segment the target area from medical images and provide effective help for subsequent diagnosis and treatment of diseases. Since deep learning technology has made great progress in the image processing, medical image segmentation algorithm based on deep learning has gradually become the focus and hotspot of research in this field. This paper gives a description on the tasks and difficulties of medical image segmentation. Then, it details the deep learning-based medical image segmentation algorithm, classifies and summarizes the current representative methods. Moreover, this paper presents the frequently-used algorithm evaluation indicators and datasets in the field of medical image segmentation. The development of medical image segmentation technology is summarized and forecasted.

Key words: deep learning; computer vision; medical image; image segmentation

图像分割与图像分类、目标检测同为图像处理技术在计算机视觉应用下的重要研究分支。图像分割, 具体可分为语义分割与非语义分割两大类。语义分割, 也称像素级分类, 是将属于同一对象类的图像部分聚集在一起的任务。相反, 非语义分割仅基于单个对象类的一般特征将像素聚集在一起^[1]。图像分割技术通常用于定位图像中目标和边界的位置, 为图像内容的进一步理解打下良好的基础。

传统的图像分割技术包括基于阈值的方法、基于边

缘的方法、基于区域的方法、基于特定理论的方法等^[2]。受限于计算机的硬件设备限制, 传统的图像分割技术仅能对灰度图像进行处理, 后期才逐渐发展到可以对 RGB 图像进行处理的阶段, 但在分割的精度和效率上仍难以达到实际应用的要求。

近些年来, 随着集电路上可容纳的电子元件数目不断翻倍, GPU 的运算性能越来越强, 受益于此的深度学习技术也因此迅猛发展, 结合深度学习技术的图像分割算法逐渐成为该领域的主流算法, 被广泛应用于移

基金项目: 国家重点研发计划(2018YFB1404402); 广东省重点研发计划(2019B010137003); 唯链基金会资金(SCNU-2018-01); 广东省科技计划(2018A07071702, 201804010314, 2016B030305006, 201222412)。

作者简介: 彭璟(1997—), 男, 硕士研究生, CCF 会员, 研究领域为计算机视觉、图像分割; 罗浩宇(1989—), 男, 博士, 研究员, 研究领域为服务计算、深度学习; 赵淦森(1977—), 通信作者, 男, 博士, 教授, 博士生导师, 研究领域为区块链、医疗人工智能; 林成创(1990—), 男, 博士, 研究领域为计算机视觉、医疗人工智能; 易序晟(1996—), 男, 硕士研究生, 研究领域为深度学习; 陈少洁(1998—), 女, 硕士研究生, 研究领域为图像处理。

收稿日期: 2020-10-23 **修回日期:** 2020-12-09 **文章编号:** 1002-8331(2021)03-0044-14

动机器人、自动驾驶、无人机、医学诊断等实际场景中。

目前有部分综述研究对基于深度学习的图像分割技术进行了总结。例如文献[3-4],对现有的基于深度学习的图像分割方法做出了综述,为后来的研究者提供了很好的帮助,但是文献[3-4]主要集中于图像分割整个领域,缺乏对某个实际应用场景的单独综述性讲解。文献[5]概述了基于深度学习的医学影像分割算法及其网络架构、训练技巧和难点,但对深度学习技术应用于其中的最新研究成果有失详尽。

因此,针对计算机视觉结合深度学习技术在医学影像分割领域的应用,还缺乏全面的综述性研究,本文撰写的主要目的是尝试对当前深度学习下的医学影像分割算法进行综述,为相关研究提供参考。本文首先叙述了计算机视觉下医学影像分割研究的任务和难点,随后根据深度学习下医学影像分割算法的发展历程和所采用的骨干网络,从基于全卷积神经网络的方法、基于U-Net网络及其变体的方法和基于特定设计思想的方法三个分类进行了归纳和总结,重点阐述了每种算法的网络架构并分析了存在的不足。其次,介绍了医学影像分割算法常用的评价指标和数据集。最后,指出了医学影像分割领域未来的研究方向。

1 医学影像分割的任务与难点

医学影像分割是医学影像分析中的重要领域,也是计算机辅助诊断、监视、干预和治疗所必需的一环,其关键任务是对医学影像中感兴趣的对象(例如器官或病变)进行分割,可以为疾病的精准识别、详细分析、合理诊断、预测与预防等方面提供非常重要的意义和价值。受深度学习技术在多个领域成功应用的推动,医学影像分割的研究人员将基于深度学习的方法应用于大脑^[6-8]、肝脏^[9]、胰腺^[10]、前列腺^[11]和多器官^[12]分割等方面。相比传统方法,这些基于深度学习的方法,在医学影像分割任务中取得了卓越的性能。

与自然图像相比,医学影像除了具有分辨率低、对比度低、目标分散等自身特性外,对分割算法结果的准确性和稳定性也有着更高的要求。自然图像分割更重视像素点的多分类信息,往往对多个类别进行分割,对分割细节要求不高。而对于医学影像分割而言,一般不需要进行多分类,只需要进行病灶或器官的区分即可,但医学影像的分割细致程度要求较高,不正确或是不稳定的分割将会直接影响病人后续的诊断和治疗,从而失去了对医学影像分割的本来意义。在医学影像分割任务中,目前主要存在以下几个难点:

(1)标注数据少^[13]。造成该问题的一个重要的原因是收集标注困难,手工标注医学影像是一个费时费力的过程,标注质量的好坏很大程度上取决于专家的临床经验和耐心程度,而这个标注过程在实际的临床实践中可

能并不需要。

(2)传感器噪声或伪影^[14]。现代医学影像最基本的成像模态有XR(X光)、US(超声)、CT(电子计算机断层扫描)和MRI(磁共振成像)等,用于成像的医学设备会存在物理噪声和图像重建误差,而医学影像模态和成像参数设定的差别则会造成不同大小的伪影。在此基础上,dicom(医学数字成像和通信)的历史标准不统一、医学设备性能不一致等问题,给医学影像分割任务带来了更大的挑战。

(3)分割目标形态差异大^[15]。患者之间存在胖或瘦、高或矮、成年人或小孩等体型差异,且病变的大小、形状和位置可能存在巨大差异,因此解剖结构上会有差异。不同的分割部位也存在差异,以血管和肿瘤的分割为例,目标都非常小,而且极其不规则,因此不同模态、不同分割部位往往需要不同的算法,有时还需要考虑到先验知识的加入。

(4)组织边界信息弱^[16]。人体内部的一些器官都是具有相似特征的软组织,它们相互接触且边界信息非常弱,而胰腺肿瘤、肝肿瘤、肾脏肿瘤等边界不清楚的肿瘤往往还非常小,导致很难被识别到。

(5)维度信息丰富^[17]。自然图像是二维的,医学影像绝大多数都是三维的,直接将自然图像分割算法迁移到医学影像分割中无法充分的利用维度信息,虽然有不少系统性的工作,但三维的图像分割技术远远没有二维那么趋于成熟。

2 深度学习下的医学影像分割算法分类

早期的图像分割算法建立在传统方法上,例如边缘检测滤波器等数学方法。然后,依靠手工提取特征的机器学习方法在很长一段时间内成为了一种占据主导地位的技术,设计和提取特征的复杂性制约了此种技术的发展。与人工规则构造特征的方法相比,基于深度学习的方法能够刻画出数据更丰富的内在信息,从而逐渐成为了图像分割领域的首选方法。本章按照深度学习下的医学影像分割算法的发展历程和所采用的骨干网络,将其划分为基于全卷积神经网络、基于U-Net网络及其变体和基于特定设计思想3类医学影像分割算法,分别介绍了3类分割方法的基本思想、代表性网络架构以及优缺点等。

2.1 基于全卷积神经网络的方法

在CNN卷积神经网络模型用于图像分类时,末尾的全连接层会将原始图像中的二维矩阵信息压缩,导致图像的空间信息丢失,这对卷积神经网络模型用于图像分割会产生很大影响。全卷积神经网络^[18]的问世开创了卷积神经网络用于图像分割的先河,其网络架构如图1所示。它的基本思想是将传统卷积神经网络模型中的全连接层替换成卷积层,接着使用反卷积操作在最后输

Huang 等人于 2020 年提出 U-Net3+^[38], 它的网络结构如图 4 所示。针对 U-Net++ 没有从多尺度中提取足够的信息这一不足之处, U-Net3+ 利用全尺度的跳跃连接和深度监督来改善该问题。全尺度的跳跃连接把来自不同尺度特征图中的高级语义与低级语义结合, 而深度监督则从多尺度聚合的特征图中学习层次表示。此外, U-Net3+ 还进一步提出了一种混合损失函数, 并设计了一个分类引导模块来增强器官边界和减少非器官图像的过度分割, 从而获得更准确的分割结果。

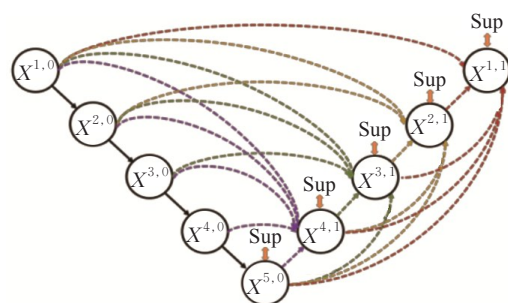


图4 U-Net3+网络架构

U-Net++ 和 U-Net3+ 网络架构的第一个优势就是精度的提升, 这个它是整合了不同层次的特征所带来的, 第二个是灵活的网络结构配合深度监督, 让参数量巨大的深度神经网络在可接受的精度范围内大幅度地缩减参数量。但是因为多次跳跃连接操作, 同样的数据在网络中会存在多个复制, 模型训练时的显存占用相对较高, 需要采用一些显存优化技术来更好地训练模型。

2.2.3 融合残差思想的 U-Net 算法

神经网络因为宽度和深度的增加, 就会面临梯度消失或梯度爆炸引起的网络退化问题, 为此 He 等人^[39]提出了残差网络 (ResNet)。如图 5 所示, 残差块的输入通过残差路径直接叠加到残差块的输出之中, 残差块会尝试去学习并拟合残差以保证增加的网络层数不会削弱网络的表达性能。

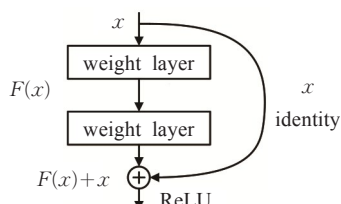
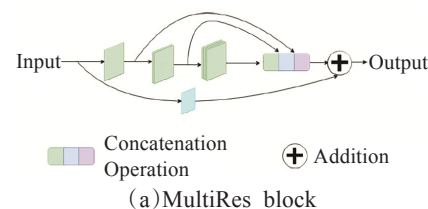


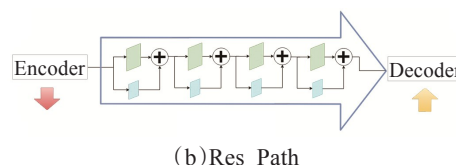
图5 ResNet 残差学习块

2019 年 Ibtehaz 等人^[40]借鉴了 ResNet 提出了 MultiResUNet 网络, MultiResUNet 运用残差思想改造了 U-Net 中的卷积块和跳跃连接。如图 6(a) 所示, MultiResNet 使用一系列 3×3 卷积核来模拟 5×5 卷积核和 7×7 卷积核的感受野, 卷积块的输入经过 1×1 卷积核后经由残差路径直接与卷积后的输入叠加, 作者称之为 MultiRes block, 在减少网络计算量的同时可以提取不同尺度的空间特征。作者同时提出了 Res Path 来减少跳跃连接

过程中所丢失的空间信息, 如图 6(b) 所示, Res Path 由一系列的 3×3 卷积、 1×1 卷积和残差路径组成, 编码器的输入特征图经过 Res Path 与解码器特征图连接, 一方面减少了语义鸿沟, 另一方面增强了网络的学习能力。



(a) MultiRes block



(b) Res Path

图6 MultiResUNet 的 MultiRes block 和 Res Path

ResUNet 由 Zhang 等人^[41]提出, 该模型在 U-Net 网络中融合了残差思想, 被设计用于航拍图像的道路检测, 2019 年 Jha 等人^[42]在其基础上增加了 SE (Squeeze-and-Excitation) 模块^[43]和空洞空间金字塔池化 (ASPP) 模块^[44], 提出了用于结肠息肉分割的 ResUNet++ 网络。SE 模块通过全局平均池化操作将图像每个通道的特征压缩为 1 个值, 再由全连接层将通道的激励映射到 $[0, 1]$ 范围, 最后将其与输入图像的每个通道相乘, 能有效地建模通道间的相关性。ASPP 模块的设计来自于 He 等人提出的空间金字塔池化网络, 该模块将不同大小的空洞卷积核堆叠, 可以有效地提取同一特征图上的不同尺度信息。虽然 ResNet 残差学习块结构确实有助于网络精度的提升, 但同时也增加了网络的训练时间。

2.2.4 基于循环神经网络的 U-Net 算法

2018 年 Alom 等人^[45]提出了 R2U-Net 网络架构, 该网络架构整合了 U-Net、ResNet、RCNN^[46]的结构, 在血管、肺部、视网膜等多个医学影像分割任务上都获得了很好的实验结果。RCNN 将卷积神经网络和循环神经网络结合, 能让神经网络记忆序列化的输入信息, 在隐式地增加了原始 CNN 深度的同时, 增强了模型捕获特征长期依赖的能力。R2U-Net 中的循环残差卷积单元 (RRCU) 如图 7 所示, 通过将原始 U-Net 网络架构中的每个卷积单元替换成 RRCU, 融合了编码-解码、残差连接、循环卷积的设计思想, 使得 R2U-Net 网络能够提取

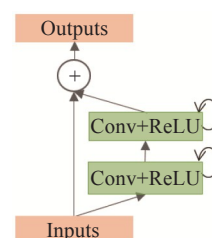


图7 R2U-Net 循环残差卷积单元

到更好的特征,在拥有相同模型参数量的情况下可以获得更好的表现。

BCDU-Net由Azad等人^[47]于2019年提出,是在U-Net网络中应用循环神经网络的另一种策略。LSTM^[48]是一种特殊的RNN,主要是为了解决长序列训练过程中的梯度消失和梯度爆炸问题。ConvLSTM^[49]是CNN和LSTM的结合体,BCDU-Net通过在跳跃连接路径中加入双向的ConvLSTM,以非线性方式合并编码和解码阶段对应的特征图,以产生更精细的分割结果。作者还在U-Net最后一个编码阶段加入了密集连接,以产生更多样化的特征来增强模型的表达能力。

尽管循环神经网络能更好地捕捉到序列数据中的语义信息,但因其状态计算固有的特性难以并行化拓展,并且基于RNN的医学影像算法更适合于分割跨越多个切片的病变或器官,对于单一切片的分割任务相比其他算法可能并不具有优势。

2.2.5 集成注意力机制的U-Net算法

2018年Oktay等人^[50]提出了Attention U-Net网络架构。注意力机制借鉴了人类的注意力思维方式,最初被应用于基于RNN循环神经网络模型的图像分类^[51]、自然语言处理^[52]、图像说明^[53]等深度学习任务中并取得了显著成果,随后Yin等人^[54]对于在CNN卷积神经网络模型中使用注意力机制做了探索性工作,如何在CNN中使用注意力机制也成为了研究的热点。

如图8所示,Attention U-Net在对扩展路径每个阶段上的特征图与收缩路径中对应特征图进行拼接之前,使用了一个注意力门抑制无关区域中的特征激活来提高模型的分割准确性,在达到高分割精度的同时而无需额外的定位模块。与U-Net相比,Attention U-Net在胰腺和腹部多器官分割数据集上,提升了分割的精度同时减少了模型训练和推理的时间。

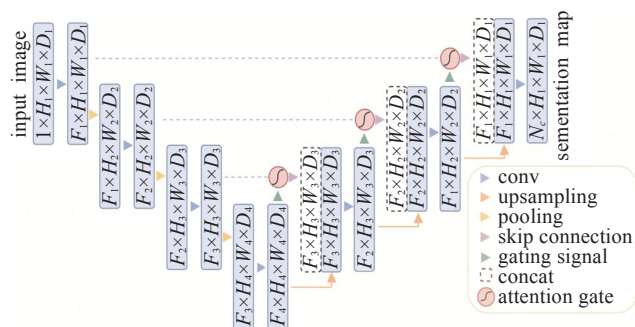


图8 Attention U-Net网络架构

2020年Li等人^[55]通过在U-Net++网络的跳跃连接中嵌入注意力门提出了ANU-Net网络架构。如图9所示,ANU-Net网络中的注意力门的两个输入分别为上采样信号 g 和编码器特征 f , g 作为门控信号可以从 f 中进行选择,经过卷积、批归范化和激活等运算后将得到的注意力图 α 与 f 相乘从而产生注意力门的输出,再

与解码器对应阶段的特征进行拼接。同时,为了充分利用U-Net++网络中的全分辨率特征信息,ANU-Net参考Dice loss、Focal loss^[56]和二进制交叉熵损失,设计了一个新颖的混合损失函数,以缓解数据不平衡问题并使模型将注意力倾向于相对更难分割的样本。

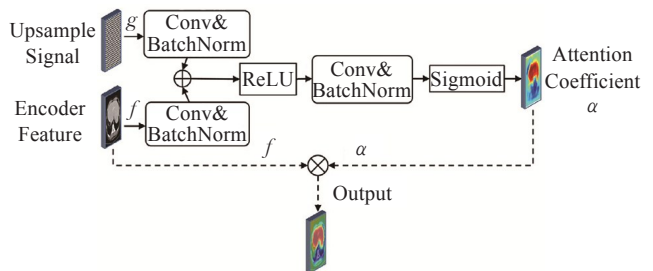


图9 ANU-Net网络中的注意力门

由于注意力门为特征图的每一个元素都重新赋予了一个权重,因此可以灵活地捕捉全局特征和局部特征的联系,同时增加了模型的可解释性,但潜在的问题是有可能破坏网络深层的特征信息,影响模型的学习能力。

2.2.6 面向3D影像的U-Net算法

3D U-Net^[57]网络架构是原始U-Net网络架构的一个简单扩展,由U-Net的研究团队在2016年提出并应用于三维图像分割。因为电脑屏幕上只能展示二维的切片,所以直接在三维层面上标注分割标签比较困难。与此同时,相邻的二维切片往往包含了近似的图片信息。基于上述两个事实,作者提出了只需要稀疏标注的二维图像进行训练的3D U-Net网络架构。3D U-Net通过将U-Net原来的2D卷积、池化、上采样操作替换成对应的3D操作,并加入Batch normalization层^[58]实现了对三维医学影像的直接分割。

如图10所示,应用3D U-Net网络架构进行医学影像分割有两种方式。图10(a)的输入是三维医学影像的稀疏标注,只标注了其中的一部分二维切片,3D U-Net经过训练可以输出三维医学影像的密集分割结果。图10(b)假定需要分割的三维医学影像有类似的代表性训练集,经过训练的3D U-Net在不带标注的三维医学影像上计算并输出分割图。

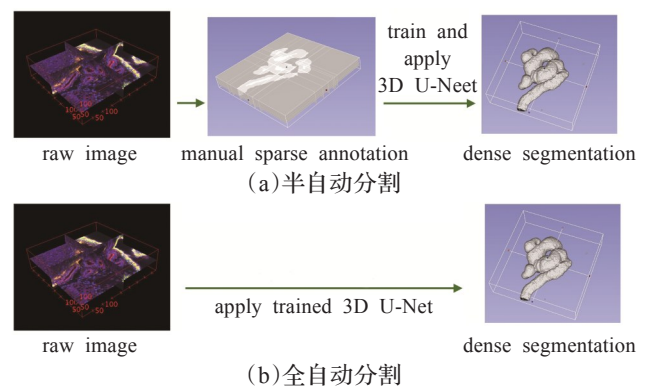


图10 3D U-Net应用场景

2016年 Milletari 等人^[59]提出了 V-Net 网络架构,是原始 U-Net 网络架构的另一种 3D 实现。V-Net 相比 3D U-Net 最大的亮点在于吸收了 ResNet 的思想,在网络拓展路径和收缩路径的每个阶段中都引入残差学习的机制。同时,受 Springenberg 等人^[60]研究的启发,V-Net 以步长为 2 的 $2 \times 2 \times 2$ 卷积核取代拓展路径里的池化操作来降低特征图的分辨率。为了解决医学影像中分割目标和背景类别不平衡问题,V-Net 还设计了新的 Dice 目标函数。

V-Net 和 3D U-Net 都是针对三维医学影像所直接构建的端到端的深度卷积神经网络,目的是运用 3D 卷积从三维进行编码,以良好的分割某些在二维没有明显表征的病理。3D 分割算法在利用医学影像的三维组织连通性方面具有优势,但相较 2D 分割算法其参数量更多,训练和推理过程对设备的算力要求更高。

2.2.7 自适应数据集的 U-Net 算法

医学影像分割领域每年有大量的新方法被提出,但在某个器官或病理分割任务中表现优秀的网络,往往无法良好地应用在其他器官或病理的分割任务上,一个重要的原因是由于医学数据集之间的数据规模、图像大小和灰度表示等方面差别很大,导致了模型推广的失败。

nnU-Net 是 Isensee 等人^[61]于 2018 年提出的一个基于 U-Net 和 3D U-Net 的医学影像分割算法框架。nnU-Net 没有设计新的网络架构,仅对 U-Net 和 3D U-Net 在网络的细节上进行了修改,着重于网络训练技巧的优化。nnU-Net 对模型的输入数据进行裁剪、重采样、标准化预处理和数据增强后,基于数据集的属性自动设置 batch size、patch size 等超参数,分别在 U-Net、3D U-Net 和两个 3D U-Net 级联模型中彼此独立地进行五折交叉验证训练,得到的 5 个网络则被用于在模型测试时进行集成推理。

Perslev 等人^[62]于 2019 年提出了一种基于多视图数据增强的 MPUNet 分割模型,训练完成后无需进行超参数调整即可准确地完成 13 个医学影像分割任务。如图 11 所示,通过从医学 3D 影像的多个视图进行各向同性采样,获取到大量与训练相关的解剖学图像,进行数据增强后输入到 6 个 2D U-Net 网络中进行分割训练,最后对 6 个网络的分割结果进行交叉验证以完成模型的融合。MPUNet 既考虑了医学影像的 3D 性质,又保持了 2D 模型的分割效率,且只进行了很少的预处理和后处理,就能很好地适应不同大小、形状和空间分布的目标分割任务。

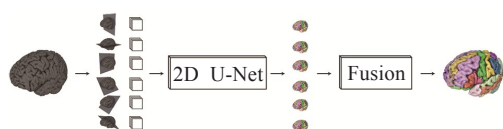


图 11 MPUNet 模型概述

nnU-Net 和 MPUNet 通过简单地使用 U-Net 与 3D U-Net 网络架构,可以动态地适应不同的医学影像数据集,在大部分医学影像分割任务中取得了非常好的成绩,证实了该类算法框架的鲁棒性,但目前来说该网络的训练时间过长,离最终的临床应用目标有所偏离。

2.2.8 基于神经网络架构搜索的 U-Net 算法

深度神经网络模型被广泛应用在计算机视觉任务中并取得了很大成功,然而神经网络架构的轻量化仍然是一个巨大的挑战。针对这一问题,模型压缩技术引起了研究人员的广泛关注,相关的研究包括手工设计轻量化网络模型^[63-64]、知识蒸馏^[65]、深度压缩^[66]和神经网络架构搜索^[67](NAS)等。其中由于 NAS 可模块化应用的特点,被较多地运用在 U-Net 算法的改进上。NAS 是一种自动化机器学习技术,通过定义搜索空间、搜索策略和性能评估策略,旨在让网络自动发现运行效率更高的轻量化架构。早期的 NAS 研究,搜索空间为整个网络架构,搜索策略包括强化学习、进化算法和贝叶斯优化等方式,因此需要非常高的算力支撑。随后的 NAS 研究主要集中于对神经网络结构块 Cell^[68],用权重共享、梯度下降等方式进行搜索来加速 NAS 的过程。

2019 年 Weng 等人^[69]首次将 NAS 应用于医学影像分割,通过用 NAS 搜索到的 Cell 替换 U-Net 网络中的对应模块,得到了 NAS-UNet 网络架构。如图 12 所示,Cell 的内部结构为一个有向无环图,每一个 Cell 的输入为前两个 Cell 的输出,图的边代表了搜索空间,分别为下采样操作集、上采样操作集与普通卷积操作集。NAS 的搜索过程转化成了一个有向无环图里面选择子图的过程,而子图权重共享的方式能有效地加速 NAS,完成搜索后将各个中间节点的输出叠加作为 Cell 的输出。以 Cell 堆叠构成的 NAS-UNet 网络,参数量仅为 U-Net 网络的 6%,却在多个医学影像数据集上取得了更高的分割精度。

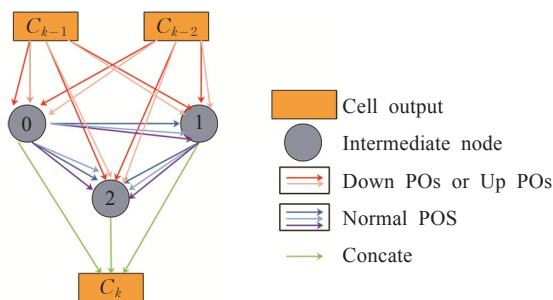


图 12 NAS-UNet 中 Cell 的结构图

Zhu 等人^[70]于 2019 年提出的 V-NAS 网络架构,进一步将 NAS 的搜索空间限定于预先定义的 3 种卷积操作,分别为 2D、3D 和伪 3D 卷积,对应了 3 种不同的卷积内核。作者利用 NAS 对 V-Net 中所有的卷积操作以梯度下降的方式进行搜索,让模型自动找到最佳的卷积操作组合,相比 V-Net 或是其他单一卷积操作的网络,更好地平衡了模型参数量和实际分割表现。

虽然基于神经网络架构搜索的U-Net算法在一些分割任务上性能表现突出,但目前NAS搜索的Cell大多是相对简单的网络操作与激活函数的排列组合,且无法解释特定网络架构表现良好的原因。

2.3 基于特定设计思想的方法

U-Net网络架构自发表以来,成为了大多数医学影像分割算法的基线模型,启发了大量研究者去思考U型分割网络,与此同时也有部分研究者针对医学影像分割中的实际难点,在融合了特定设计思想的基础上,探索并提出了一些不同于U-Net的新型网络架构,本节从相关医学影像分割算法的设计思想出发,对其中的研究工作做了介绍,旨在为研究者拓宽思路。

2.3.1 基于多任务学习的算法

当前大多数机器学习任务都是单任务学习,对于复杂的学习问题一般分解为简单且相互独立的子问题来求解,然后再合并结果得到最初复杂问题的结果,这样做忽略了问题之间的关联信息,削弱了模型的泛化效果。多任务学习是一种推导迁移学习方法,主任务使用相关任务的训练信号来提升主任务的泛化效果^[71],针对医学影像分割任务而言能帮助缓解数据标注稀疏而引发的模型过拟合问题。

2018年Mehta等人^[73]提出了Y-Net网络架构,在乳腺活检图像的分割任务中,加入乳腺癌图片的分类任务。如图13所示,Y-Net在U-Net的基础上,引入了残差网络的残差连接以帮助改善分割效果,同时添加了第二个分支用于乳腺癌图片的分类。首先以分割为目标对Y-Net网络进行预训练,然后附加第二条分支共同训练分割和分类任务。

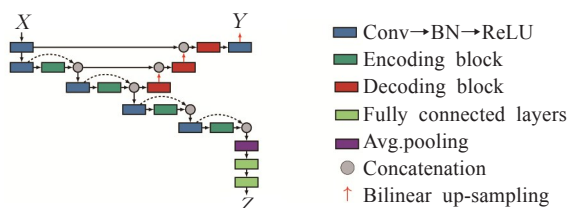


图13 Y-Net网络架构

Murugesan等人^[73]提出了一个适用于医学影像分割的多任务学习模块Conv-MCD,该模块很容易集成到现有的基于深度学习的分割网络模型中。Conv-MCD模块的三个并行学习分支,分别负责分割预测、轮廓提取和距离图估计,分割预测和轮廓提取属于分类任务,而距离图估计属于回归任务。Conv-MCD利用多任务学习能有效地处理医学影像的类平衡问题并减轻结构信息的丢失,并且没有引入额外的标注成本。

一般情况下,多任务学习能作为一种正则化方式约束模型从而缓解过拟合问题,但是不合理的任务设计可能导致模型的训练被某些任务主导,异常任务带来的负面影响则会降低模型的性能。

2.3.2 基于多模态融合的算法

在医学影像学研究中,通常会结合使用不同的核磁共振成像模式来克服单一成像技术的局限性。以大脑的分割研究为例,T1加权的影像能使灰质组织和白质组织产生良好的对比度,而T2加权的影像能有助于组织异常病变的可视化,因此考虑多种核磁共振成像方式对于获得准确的诊断结果至关重要。

2019年Dolz等人^[74]在前人的工作基础上,提出了HyperDenseNet网络架构,通过融合多模态图像进行医学影像分割。如图14所示,HyperDenseNet将密集连接的概念拓展到多模态,网络的输入为T1加权和T2加权的三维核磁共振影像,每个成像模式都有一条路径,密集连接不仅存在于同一路径的各层之间,而且可以跨越不同路径,以前馈方式将前一层直接连接到后续所有层,减少模型过拟合的风险。

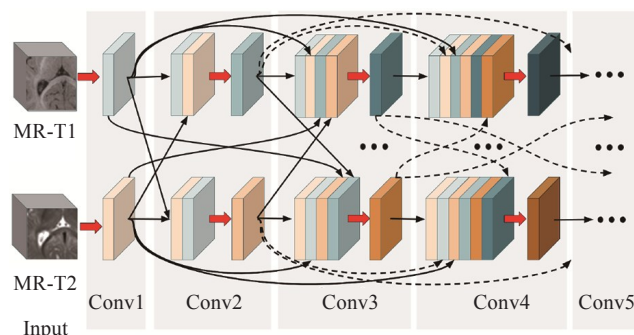


图14 HyperDenseNet网络架构

Kumar等人^[75]在2019年为融合来自多模态图像的互补信息,提出了一个用于肺癌PET和CT图像分割的多模态融合网络。通过两条不同的编码路径,得到CT图和PET图2D切片的相关图像特征,然后经过共同学习中中间层导出多模态融合图PET-CT,以加权不同位置特定于模态的特征,最后由重建组件集来自编码器不同尺寸的特征图从而转化为分割图。

HyperDenseNet等基于多模态的医学影像分割算法考虑了不同模态数据之间的互补信息,从而有助于网络更好地实现病变或器官的分割,但多模态的融合效果和分割的精度仍然有待提升。

2.3.3 基于多阶段级联的算法

现有的医学影像分割算法大多数都是单阶段算法,通过训练最小化损失函数来直接将器官或病变等目标从图像中分割出来,而由于医学影像目标形态差异大、组织边界信息弱等特点,会导致相对较小的目标在其边界附近出现不准确的分割。

2018年Roth等人^[76]提出了一种多阶段级联的3D U-Net,使模型更多地关注分割目标的边界区域,输出更精细的分割图。如图15所示,作者采用了3D U-Net作为模型的骨干网络,第一阶段3D U-Net使用基于形态学方法分割出的候选区域C1进行训练,以输出粗粒度

分割图,粗粒度分割图经过前景扩张后生成候选区域C2,用于第二阶段3D U-Net的训练从而输出最终的精细分割图。

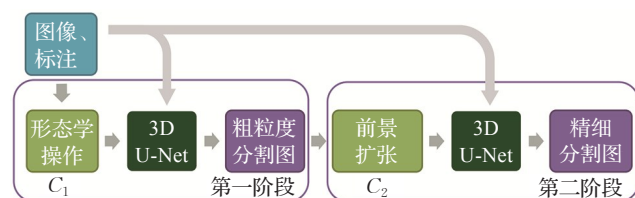


图15 多阶段级联的3D U-Net训练流程图

Wang 等人^[77]提出的级联各向异性卷积神经网络,实现了对脑肿瘤的多阶段分割。如图16所示,该分割框架由三个级联的卷积神经网络组成,作者称之为W-Net、T-Net和E-Net。W-Net从患者的3D影像中提取出整个肿瘤的边界框,基于边界框对输入图像进行裁剪后作为T-Net的输入,以得到肿瘤核心区域的边界框,进而输入到E-Net分割出肿瘤的核心区域。网络通过将 $3 \times 3 \times 3$ 的卷积核分解为 $3 \times 3 \times 1$ 和 $1 \times 1 \times 3$ 的切片内核,利用了各向异性卷积结合多视图融合的方法,以解决肿瘤的过度分割问题。

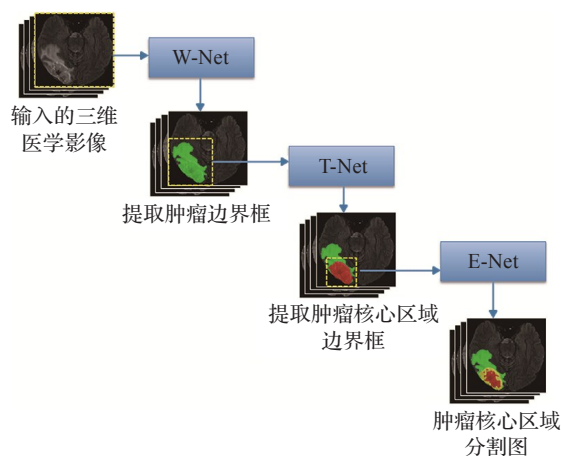


图16 多阶段级联的脑肿瘤分割框架

这类由粗到细的多阶段级联网络,相比单阶段算法虽然在检测小目标方面更具优势,但增加了额外的计算成本,在检测的实时性上需要改善。

2.3.4 基于特征增强的算法

U-Net编码-解码的体系结构推动了医学影像分割的发展,但相似的低级特征在多个尺度上被多次提取,导致了特征的冗余使用。其次,无法有效地将最佳的特征表示与每个分割类别相关联。

2020年Sinha 等人^[78]提出的MS-Dual-Guided网络,自适应地集成了局部特征和全局依赖以进行特征增强。如图17所示,MS-Dual-Guided中的位置注意力模块(PAM)和通道注意力模块(CAM)分别模拟空间和通道维度中的语义依赖性。位置注意力模块中的前两个

分支计算位置间的相关性矩阵,再与第三条分支相乘得到空间注意力图来指导输入。通道注意力模块与位置注意力模块类似,但其输入不经过卷积层,以保持通道间的相对特征,最后将两个注意力模块的输出汇总,以获得更好的像素级预测结果。

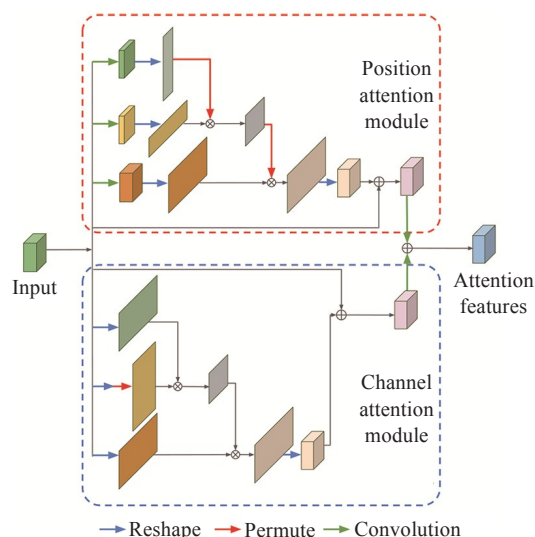


图17 MS-Dual-Guided中的注意力模块

针对卷积和池化操作可能导致的空间特征丢失问题,2019年Gu 等人^[79]提出了一种上下文编码网络CE-Net,来保留特征的空间信息并捕获更多高级特征。如图18所示,CE-Net主要包含三个组件:特征编码器模块、上下文提取模块和特征解码器模块。特征编码器模块由预训练的ResNet组成,上下文提取模块由密集空洞卷积块(DAC)和残差多核池化块(RMP)组成。空洞卷积块能够提取各个尺寸的目标特征,残差多核池化块使用四个不同大小的池化核并联以检测不同大小的目标。特征解码器模块用于恢复特征编码器中的高分辨率特征,输出与输入图像大小相同的分割图。

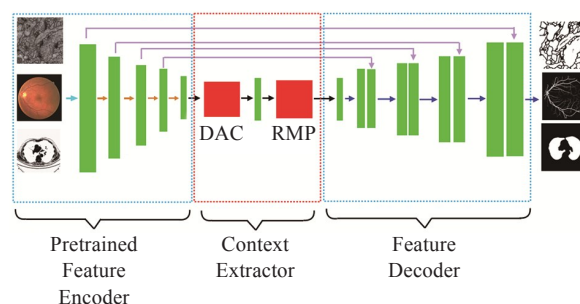


图18 CE-Net网络架构

基于特征增强的医学影像分割算法较好地保留了输入图片的原始特征,并且通过特定的网络设计最大化提取了特征信息,存在的缺点是人工痕迹太明显,且无法良好地泛化到不同类型的分割任务中。

从网络架构的主要思想、关键技术、优缺点及其主要功能对上述算法框架的综合总结如表1所示。

表1 医学影像分割领域的代表性算法

网络架构	主要思想	优点	缺点	关键技术	主要功能
FCN、U-Net	编码解码	将传统分类网络改造成了分割网络	上采样结果比较粗糙,对图像细节不敏感	反卷积	使模型输出的分割图和原图大小一致
				跳跃连接	改善上采样粗糙的像素定位,细化分割结果
U-Net++, U-Net3+	密集连接	提升了分割精度,缩减了参数量	训练数据冗余,占用较多显存	深度监督	用来弥补编码器和解码器特征图之间的语义鸿沟,改善梯度流
MultiResUNet、ResUNet++	残差学习	减缓了梯度爆炸和梯度消失现象	增加了网络的训练时间	残差路径	将 stage 的输入和输出叠加,提高训练效果
R2U-Net、BCDU-Net	记忆长期依赖	更好地捕捉到序列数据中的信息	并行化拓展困难,适合分割跨越多个切片的目标	循环神经网络	使网络能够提取到更好的特征,在拥有相同模型参数数量的情况下获得更好的表现
Attention U-Net、ANU-Net	注意力机制	灵活地捕捉了全局特征与局部特征的联系,增加了模型的可解释性	可能破坏网络深层的特征信息,影响模型的学习能力	注意力门	抑制无关区域中的特征激活来提高模型灵敏度和准确性,在达到高分割精度的同时而无需额外的定位模块
3D U-Net、V-Net	三维分割	实现了针对三维医学影像的直接分割	参数量过多、分割速度慢	3D卷积、池化、上采样	将2D运算转换为3D运算
nnU-Net、MPUNet	数据集自适应	可以动态适应不同的医学影像数据集	训练时间过长	训练技巧优化	更好地学习到数据本身的特征,加强算法的鲁棒性
NAS-UNet、V-NAS	自动化机器学习	平衡了模型的参数量和分割性能	仅是简单操作的排列组合,可解释性差	神经网络架构搜索	让网络自动发现运行效率更高的轻量化架构
Y-Net	多任务学习	能够约束模型从而缓解过拟合问题	容易被其他任务主导,从而降低模型的性能	权重共享	利用相关任务提升主任务模型的泛化效果
HyperDenseNet	多模态融合	最大程度地利用了现有数据	多模态融合效果和分割精度仍然有待提升	特征融合	克服单一成像技术的局限性,使用不同模态的数据进行互补
Cascaded 3D U-Net	多阶段分割	在检测小目标方面更具优势	增加了额外的计算成本,实时性较差	网络级联	产生候选区域后精细化模型的分割结果
MS-Dual-Guided Net、CE-Net	特征增强	最大化地提取了信息	人工痕迹明显,无法泛化到不同的分割任务	特征重编码	捕获全局依赖,还原空间信息,最大化地利用现有特征

3 常用数据集与评价指标

3.1 常用数据集

在医学影像分割领域的实践中,大部分情况下一个网络模型的训练需要许多已标注数据的输入,而不同的算法框架也需要一个共同的数据集来判断性能的优劣。现实中,个人收集一个较大且有标注的医学影像数据集难度很大,一方面需要大量的时间和专业领域的知识,另一方面医疗机构出于隐私保护等问题很少会开放相关医学影像数据,因此可供研究人员利用的数据集大部分是知名研究团队联合有关医疗机构开源发布的。表2从数据集的所属器官部位、开放年份、具体内容和成像模态等方面,对医学影像分割领域常用的数据集进行了划分整理。

从成像模态的角度,相机成像一般只适用于可以非创伤直接观察的器官部位,是深度学习早期应用于医学影像分割任务的主要成像方式。X光和CT对以骨质及肺脏等空腔脏器的成像效果很好,同时成像速度快,器官部位的运动伪影较小。MRI与CT比较,其主要优点

是对软组织的分辨更加清晰,更容易明确肿瘤等一类小的病灶,但成像速度慢,所以对于大脑、前列腺等相对静止的器官部位应用更多。

从器官部位的角度,视网膜血管和皮肤黑色素瘤的分割目标相对分散,但由于其成像比较准确直观,因此其分割难度相对较低,基于编码解码的FCN和U-Net算法就可以较为理想地分割出目标。前列腺和心脏的分割任务为对器官部位的提取,目标相对较大,其中前列腺的成像边界相对模糊、强度分布不均匀,而心脏是一个不停运作的器官,其内部解剖结构复杂且相位会产生变化,因而考虑到注意力机制和特征增强的算法会取得更好的分割效果。乳腺、肺、肝脏、脾脏、肾脏和大脑等器官部位主要是对肿物或病灶的分割,分割目标相对病变的器官部位来说占比较小,往往形状不规则、形变较大且会跨越多个切片,因此应用多阶段分割、多模态融合、RNN和3D分割的算法能够更精细地分割出目标,模型也会更加复杂。胰腺和多器官的分割是目前更具挑战性的问题,尽管胰腺附近的腹部器官已经可以被分

表2 医学影像分割领域常用数据集

器官部位名称	数据集名称	开放年份	具体内容	成像模态
视网膜	DRIVE	2003	包含40张视网膜图像中的血管分割图	相机
	STARE	2004	包含400张视网膜图像,标注了其中40张的血管	相机
	CHASE_DB1	2012	包含28张分辨率为1 280×960的视网膜血管分割图	相机
皮肤	Pratheepan	2012	从谷歌随机下载且标注的图像,用于人类皮肤检测研究	相机
	ISIC2016	2016	包括1 250张黑色素瘤分割图	相机
乳腺	INbreast	2012	标注了115位患者的410张乳腺肿块分割图	XR
前列腺	Promise 2012	2012	包含50个不同设备获取的前列腺分割图	MRI
	I2CVB	2015	提供了用于前列腺癌检测的前列腺分割图	MRI
心脏	SCD	2009	标注了不同心脏疾病的45个受试者的左心室分割图	MRI
	ACDC	2017	收集了150个受试者的心脏分割图	MRI
肺	LUNA16	2016	包括了888张肺结节分割图	CT
	CTVIE19	2019	提供了50个患有不同呼吸道疾病的病人肺部分割图	CT
	COVID-19-20	2020	标注了296名新型冠状病毒感染者的肺部分割图	CT
肝脏	LIVER07	2007	总共30张肝脏肿瘤分割图	CT
	LiTS2017	2017	含有200张肝脏肿瘤分割图	CT
脾脏	CHAOS	2019	标注了120位病人的脾脏分割图	MRI
肾脏	KiTS19	2019	标注了300位病人的肾脏肿瘤分割图	CT
胰腺	NIH Pancreas-CT	2015	包括了82张胰腺分割图	CT
大脑	BraTS2017	2017	提供了300多张大脑肿瘤分割图	MRI
	iSeg2017	2017	提供了6个月大的婴儿大脑解剖结构图	MRI
多器官	Computational anatomy	2017	包含240张人体躯干图,标注了19个解剖结构图	CT
	Decathlon	2018	囊括了10个器官或病变的数百张分割图	CT和MRI

割得较好,但由于胰腺没有包膜所以边界很不清楚,多器官分割的难点则在于需要分割的目标尺寸不一致、相邻器官的空间界限难以确认以及不同器官分割训练时存在的对抗性,考虑到多视图信息、模型融合的数据集自适应算法是目前解决这类问题的主流算法。

3.2 评价指标

为了公平地比较图像分割领域中的不同算法,必须有标准的、被广泛认可的指标用于评估。常用的医学影像分割算法评估标准有精确率(accuracy)、召回率(recall)、特异率(specificity)、Dice系数(Dice coefficient)和Jaccard指数(Jaccard index)。

以图19为例进行说明,A为一张医学影像的真实标注,B为分割模型的预测结果,则准确率AC、召回率SE、特异率SP、Dice系数DSC和Jaccard指数JAC分别表示为:

$$AC = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$SE = \frac{TP}{TP + FN} \quad (2)$$

$$SP = \frac{TN}{TN + FP} \quad (3)$$

$$DSC = \frac{2(A \cap B)}{A + B} \quad (4)$$

$$JAC = \frac{A \cap B}{A \cup B} \quad (5)$$

准确率是预测正确的像素占总像素的百分比,在类别不平衡的情况下,并不能作为很好的指标来衡量分割

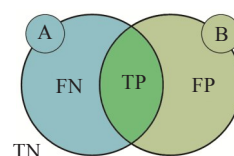


图19 医学影像分割结果示例

结果。召回率又称敏感率(sensitivity),只关注真实标注被正确预测的比例,而特异率的关注则刚好相反,这两种指标对分割目标的大小比较敏感。Dice系数是医学影像分割任务中最常用的评价指标,能较好地规避医学影像领域中普遍存在的类别不平衡问题。Jaccard指数又称交并比(IoU),它与Dice系数的关系为:

$$DSC = \frac{2JAC}{1 + JAC} \quad (6)$$

实际应用中,往往会根据需求对上述评价指标进行取舍,从多个维度证明分割算法的准确性和稳定性。

4 总结和展望

本文阐述了医学影像分割的任务及其难点,对于深度学习下的医学影像分割算法进行了综述,介绍了医学影像分割领域的研究现状、相关评价指标和数据集。总的来说,基于深度学习的医学影像分割在未来将发挥实质性的作用,但该技术的落地还存在以下亟待研究的问题:

(1)分割网络架构的轻量化。对于现阶段的医学影像分割网络架构来说,进行模型压缩以减少对硬件设备

的算力需求是需要考虑的实际问题。模型压缩现有的研究方向包括手工设计、知识蒸馏、深度压缩和神经网络架构搜索等,医学影像分割领域也有针对模型压缩的部分研究工作,如使用3D空洞卷积的多尺度脑肿瘤分割^[80]、基于知识蒸馏的脑肿瘤分割^[81]、基于权重量化的腺体细胞分割^[82]和基于神经网络架构搜索的头颈肿瘤分割^[83]等,但这些研究尚处于起步阶段,神经网络模型在嵌入式设备上的存储与计算仍然是一个未解决的难点。在保证准确率和稳定性的同时压缩模型,实现医学影像的实时分割将会是未来研究的重点。

(2)分割结果的不确定性分析。不确定性分析目的是让模型给出分割结果的同时,指出哪些是不确定的分割,需要人工介入修正。虽然有少数研究工作^[84]探索了不确定性分析在医学影像分割中的作用,但现阶段的大多数医学影像分割算法只输出确定性的分割图。允许医生能够根据模型预测的不确定性分割结果进行修改,提升分割的结果和质量,这是理论与实际场景结合的重要环节,也是医学影像分割算法值得进一步研究的问题。

(3)稀疏标注下的弱监督学习。尽管深度学习下的医学影像分割算法不断达到更高的分割精度,但仍然离不开大规模的高质量标注数据集支持,因此有不少研究人员尝试利用未标注和稀疏标注的数据进行弱监督学习,如基于着色还原的皮肤分割^[85]、基于点注释的细胞核分割^[86]等,但完成的分割任务相对简单且算法无法扩展到其他分割任务。如何实现大量标注不完善的数据集的弱监督学习,在自然图像和医学影像未来的研究中都尤为重要。

(4)小数据集下的数据增强。克服医学影像标注稀缺的另一种手段就是数据增强,为深度模型扩充训练集。传统的方法包括几何变换、颜色变换、仿射变换和高斯噪声等,这类方法取得的效果相对有限。基于深度学习的生成对抗网络^[87]在自然图像的生成任务中表现突出,也有部分研究工作^[88]将其应用于医学影像分割模型的数据增强中,其他方法还有基于超像素的数据增强^[89]、基于无监督的数据增强^[90]等,但该类方法都存在生成的数据分布差、算法适用范围有限等缺点。因此,设计合理且泛化效果好的医学影像数据增强算法,将会是未来发展的趋势。

(5)融合先验知识的分割算法设计。医学影像分割不同于自然图像分割,即使是同一张医疗影像,不同经验的专家也可能给出不同的诊断,有经验的专家更能快速地找出器官和病变,这表明先验知识在医学诊断中占了很大比重。针对这个现象,根据医学影像中不同对象的灰度分布情况、解剖知识和空间几何关系以及不同成像设备的特点,融合先验知识指导模型结构和损失函数设计,应当是未来的研究方向。

参考文献:

- [1] THOMA M.A survey of semantic segmentation[J].arXiv: 1602.06541, 2016.
- [2] 梁新宇,罗晨,权冀川,等.基于深度学习的图像语义分割技术研究进展[J].计算机工程与应用,2020,56(2):18-28.
- [3] 邝辉宇,吴俊君.基于深度学习的图像语义分割技术研究综述[J].计算机工程与应用,2019,55(19):12-21.
- [4] LIU X,DENG Z,YANG Y.Recent progress in semantic image segmentation[J].Artificial Intelligence Review,2019,52(2):1089-1106.
- [5] HESAMIAN M H,JIA W,HE X,et al.Deep learning techniques for medical image segmentation: achievements and challenges[J].Journal of Digital Imaging,2019,32(4): 582-596.
- [6] HAVAEI M,DAVY A,WARDE-FARIEY D,et al.Brain tumor segmentation with deep neural networks[J].Medical Image Analysis,2017,35:18-31.
- [7] PEREIRA S,PINTO A,ALVES V,et al.Brain tumor segmentation using convolutional neural networks in MRI images[J].IEEE Transactions on Medical Imaging,2016,35(5):1240-1251.
- [8] MENZE B H,JAKAB A,BAUER S,et al.The multi-modal brain tumor image segmentation benchmark (BRATS)[J].IEEE Transactions on Medical Imaging,2014,34(10):1993-2024.
- [9] CHRIST P F,ELSHAER M E A,ETTLINGER F,et al. Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2016: 415-423.
- [10] ROTH H R,LU L,FARAG A,et al.Deeporgan: multi-level deep convolutional networks for automated pancreas segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2015: 556-564.
- [11] YU L,YANG X,CHEN H,et al.Volumetric convnets with mixed residual connections for automated prostate segmentation from 3D MR images[C]//Thirty-First AAAI Conference on Artificial Intelligence, 2017.
- [12] ZHOU X,TAKAYAMA R,WANG S,et al.Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting method[J].Medical Physics,2017,44(10):5221-5233.
- [13] DESPOTOVIĆ I,GOOSSENS B,PHILIPS W.MRI segmentation of the human brain: challenges, methods, and applications[J].Computational and Mathematical Methods in Medicine,2015(6):1-23.
- [14] 张翥,范虹.基于模糊C均值聚类的医学图像分割研究[J].计算机工程与应用,2014,50(4):144-151.

- [15] 李铮,白柯鑫,赵柳,等.MRI脑肿瘤图像分割研究进展及挑战[J].中国图象图形学报,2020,25(3):419-431.
- [16] RAZZAK M I,NAZ S,ZAIB A.Deep learning for medical image processing:overview,challenges and the future[M]//Classification in bioApps.Cham:Springer,2018:323-350.
- [17] 许存禄,高佳,武国德.Chan-Vese 模型下的脑肿瘤图像分割方法[J].计算机工程与应用,2010,46(9):155-158.
- [18] LONG J,SHELHAMER E,DARRELL T.Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2015:3431-3440.
- [19] KRIZHEVSKY A,SUTSKEVER I,HINTON G E.Imagenet classification with deep convolutional neural networks[C]//Advances in Neural Information Processing Systems,2012:1097-1105.
- [20] SIMONYAN K,ZISSERMAN A.Very deep convolutional networks for large-scale image recognition[J].arXiv:1409.1556,2014.
- [21] SZEGEDY C,LIU W,JIA Y,et al.Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2015:1-9.
- [22] BEN-COHEN A,DIAMANT I,KLANG E,et al.Fully convolutional network for liver segmentation and lesions detection[M]//Deep learning and data labeling for medical applications.Cham:Springer,2016:77-85.
- [23] YUAN Y,CHAO M,LO Y C.Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance[J].IEEE Transactions on Medical Imaging,2017,36(9):1876-1886.
- [24] DASGUPTA A,SINGH S.A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation[C]//2017 IEEE 14th International Symposium on Biomedical Imaging,2017:248-251.
- [25] 章琳,袁非牛,张文睿,等.全卷积神经网络研究综述[J].计算机工程与应用,2020,56(1):25-37.
- [26] 魏帼,郭浩,安居白,等.基于全卷积神经网络的SAR海面溢油图像分割方法[J].计算机应用,2019,39(S1):182-186.
- [27] 孙莹莹.基于多尺度FCN-CRF和强化学习的高分辨SAR图像语义分割[D].西安:西安电子科技大学,2018.
- [28] 谢凯浪,赵泉华,李玉.结合FCN和多特征的全极化SAR土地覆盖分类[J].测绘科学,2020,45(1):77-83.
- [29] BADRINARAYANAN V,HANDA A,CIPOLLA R.Segnet: a deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling[J].arXiv:1505.07293,2015.
- [30] 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,2017,40(4):834-848.
- [31] LIN G,MILAN A,SHEN C,et al.Refinenet: multi-path refinement networks for high-resolution semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2017:1925-1934.
- [32] FU J,LIU J,TIAN H,et al.Dual attention network for scene segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2019:3146-3154.
- [33] 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.Cham:Springer,2015:234-241.
- [34] GORDIENKO Y,GANG P,HUI J,et al.Deep learning with lung segmentation and bone shadow exclusion techniques for chest X-ray analysis of lung cancer[C]//International Conference on Computer Science, Engineering and Education Applications.Cham:Springer,2018:638-647.
- [35] IANDOLA F,MOSKEWICZ M,KARAYEV S,et al. Densenet: implementing efficient convnet descriptor pyramids[J].arXiv:1404.1869,2014.
- [36] ZHOU Z,SIDDIQUEE M M R,TAJBAKHSH N,et al. Unet++: a nested u-net architecture for medical image segmentation[M]//Deep learning in medical image analysis and multimodal learning for clinical decision support. Cham:Springer,2018:3-11.
- [37] LEE C Y,XIE S,GALLAGHER P,et al.Deeply-supervised nets[C]//Artificial Intelligence and Statistics,2015:562-570.
- [38] HUANG H,LIN L,TONG R,et al.UNet 3+: a full-scale connected UNet for medical image segmentation[C]//IEEE International Conference on Acoustics, Speech and Signal Processing,2020:1055-1059.
- [39] HE K,ZHANG X,REN S,et al.Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2016:770-778.
- [40] IBTEHAZ N,RAHMAN M S.MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation[J].Neural Networks,2020,121:74-87.
- [41] ZHANG Z,LIU Q,WANG Y.Road extraction by deep residual u-net[J].IEEE Geoscience and Remote Sensing Letters,2018,15(5):749-753.
- [42] JHA D,SMEDSRUD P H,RIEGLER M A,et al.Resunet++: an advanced architecture for medical image segmentation[C]//2019 IEEE International Symposium on

- Multimedia(ISM),2019:225-2255.
- [43] HU J,SHEN L,SUN G.Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2018:7132-7141.
- [44] HE K,ZHANG X,REN S,et al.Spatial pyramid pooling in deep convolutional networks for visual recognition[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2015,37(9):1904-1916.
- [45] ALOM M Z,HASAN M,YAKOPCIC C,et al.Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation[J].arXiv:1802.06955,2018.
- [46] LIANG M, HU X.Recurrent convolutional neural network for object recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015:3367-3375.
- [47] AZAD R,ASADI-AGHBOLAGHI M,FATHY M,et al. Bi-directional convLSTM U-net with densely connected convolutions[C]//Proceedings of the IEEE International Conference on Computer Vision Workshops,2019.
- [48] GREFF K,SRIVASTAVA R K,KOUTNÍK J,et al.LSTM: a search space odyssey[J].IEEE Transactions on Neural Networks and Learning Systems,2016,28(10):2222-2232.
- [49] SONG H,WANG W,ZHAO S,et al.Pyramid dilated deeper convlstm for video salient object detection[C]//Proceedings of the European Conference on Computer Vision,2018:715-731.
- [50] OKTAY O,SCHLEMPER J,FOLGOC L L,et al.Attention u-net:learning where to look for the pancreas[J].arXiv: 1804.03999,2018.
- [51] MNIH V,HEESS N,GRAVES A.Recurrent models of visual attention[C]//Advances in Neural Information Processing Systems,2014:2204-2212.
- [52] BAHDANAU D,CHO K,BENGIO Y.Neural machine translation by jointly learning to align and translate[J]. arXiv:1409.0473,2014.
- [53] XU K,BA J,KIROS R,et al.Show,attend and tell: neural image caption generation with visual attention[C]// International Conference on Machine Learning,2015: 2048-2057.
- [54] YIN W,SCHÜTZE H,XIANG B,et al.Abcnn:attention-based convolutional neural network for modeling sentence pairs[J].Transactions of the Association for Computational Linguistics,2016,4:259-272.
- [55] LI C,TAN Y,CHEN W,et al.ANU-Net:attention-based nested U-Net to exploit full resolution features for medical image segmentation[J].Computers & Graphics, 2020.
- [56] LIN T Y,GOYAL P,GIRSHICK R,et al.Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision,2017: 2980-2988.
- [57] ÇİÇEK Ö,ABDULKADIR A,LIENKAMP S S,et al. 3D U-Net:learning dense volumetric segmentation from sparse annotation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham:Springer,2016:424-432.
- [58] IOFFE S,SZEGEDY C.Batch normalization:accelerating deep network training by reducing internal covariate shift[J].arXiv:1502.03167,2015.
- [59] MILLETARI F,NAVAB N,AHMADI S A.V-net: fully convolutional neural networks for volumetric medical image segmentation[C]//2016 Fourth International Conference on 3D Vision(3DV),2016:565-571.
- [60] SPRINGENBERG J T,DOSOVITSKIY A,BROX T,et al. Striving for simplicity:the all convolutional net[J].arXiv: 1412.6806,2014.
- [61] ISENSEE F,PETERSEN J,KLEIN A,et al.Nnu-net: self-adapting framework for u-net-based medical image segmentation[J].arXiv:1809.10486,2018.
- [62] PERSLEV M,DAM E B,PAI A,et al.One network to segment them all:a general,lightweight system for accurate 3d medical image segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention.Cham:Springer,2019:30-38.
- [63] HOWARD A G,ZHU M,CHEN B,et al.Mobilenets: efficient convolutional neural networks for mobile vision applications[J].arXiv:1704.04861,2017.
- [64] ZHANG X,ZHOU X,LIN M,et al.Shufflenet:an extremely efficient convolutional neural network for mobile devices[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2018:6848-6856.
- [65] HINTON G,VINYALS O,DEAN J.Distilling the knowledge in a neural network[J].arXiv:1503.02531,2015.
- [66] HAN S,MAO H,DALLY W J.Deep compression:compressing deep neural networks with pruning,trained quantization and huffman coding[J].arXiv:1510.00149,2015.
- [67] ZOPH B,LE Q V.Neural architecture search with reinforcement learning[J].arXiv:1611.01578,2016.
- [68] ZOPH B,VASUDEVAN V,SHLENS J,et al.Learning transferable architectures for scalable image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2018:8697-8710.
- [69] WENG Y,ZHOU T,LI Y,et al.Nas-unet:neural architecture search for medical image segmentation[J].IEEE Access,2019,7:44247-44257.
- [70] ZHU Z,LIU C,YANG D,et al.V-nas:neural architecture search for volumetric medical image segmentation[C]// 2019 International Conference on 3D Vision(3DV), 2019:240-248.

- [71] 邵蔚元,郭跃飞.多任务学习及卷积神经网络在人脸识别中的应用[J].计算机工程与应用,2016,52(13):32-37.
- [72] MEHTA S, MERCAN E, BARTLETT J, et al. Y-Net: joint segmentation and classification for diagnosis of breast biopsy images[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2018: 893-901.
- [73] MURUGESAN B, SARVESWARAN K, SHANKARANA-RAYANA S M, et al. Conv-MCD: a plug-and-play multi-task module for medical image segmentation[C]//International Workshop on Machine Learning in Medical Imaging. Cham: Springer, 2019: 292-300.
- [74] DOLZ J, GOPINATH K, YUAN J, et al. Hyper Dense-Net: a hyper-densely connected CNN for multi-modal image segmentation[J]. IEEE Transactions on Medical Imaging, 2018, 38(5): 1116-1126.
- [75] KUMAR A, FULHAM M, FENG D, et al. Co-learning feature fusion maps from PET-CT images of lung cancer[J]. IEEE Transactions on Medical Imaging, 2019, 39(1): 204-217.
- [76] ROTH H R, ODA H, ZHOU X, et al. An application of cascaded 3D fully convolutional networks for medical image segmentation[J]. Computerized Medical Imaging and Graphics, 2018, 66: 90-99.
- [77] WANG G, LI W, OURSELIN S, et al. Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks[C]//International MICCAI Brainlesion Workshop. Cham: Springer, 2017: 178-190.
- [78] SINHA A, DOLZ J. Multi-scale guided attention for medical image segmentation[J]. arXiv: 1906.02849, 2019.
- [79] GU Z, CHENG J, FU H, et al. Ce-net: context encoder network for 2d medical image segmentation[J]. IEEE Transactions on Medical Imaging, 2019, 38(10): 2281-2292.
- [80] CHEN C, LIU X, DING M, et al. 3D dilated multi-fiber network for real-time brain tumor segmentation in MRI[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2019: 184-192.
- [81] LACHINOV D, SHIPUNOVA E, TURLAPOV V. Knowledge distillation for brain tumor segmentation[C]//International MICCAI Brainlesion Workshop. Cham: Springer, 2019: 324-332.
- [82] XU X, LU Q, YANG L, et al. Quantization of fully convolutional networks for accurate biomedical image segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8300-8308.
- [83] GUO D, JIN D, ZHU Z, et al. Organ at risk segmentation for head and neck cancer using stratified learning and neural architecture search[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 4223-4232.
- [84] NAIR T, PRECUP D, ARNOLD D L, et al. Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation[J]. Medical Image Analysis, 2020, 59: 101557.
- [85] TAJBAKHSH N, HU Y, CAO J, et al. Surrogate supervision for medical image analysis: effective deep learning from limited quantities of labeled data[C]//2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019: 1251-1255.
- [86] QU H, WU P, HUANG Q, et al. Weakly supervised deep nuclei segmentation using points annotation in histopathology images[C]//International Conference on Medical Imaging with Deep Learning, 2019: 390-400.
- [87] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Advances in Neural Information Processing Systems, 2014: 2672-2680.
- [88] SHIN H C, TENENHOLTZ N A, ROGERS J K, et al. Medical image synthesis for data augmentation and anonymization using generative adversarial networks[C]//International Workshop on Simulation and Synthesis in Medical Imaging. Cham: Springer, 2018: 1-11.
- [89] ZHANG Y, YANG L, ZHENG H, et al. SPDA: superpixel-based data augmentation for biomedical image segmentation[C]//International Conference on Medical Imaging with Deep Learning, 2019: 572-587.
- [90] ZHAO A, BALAKRISHNAN G, DURAND F, et al. Data augmentation using learned transformations for one-shot medical image segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 8543-8553.