



计算机科学

COMPUTER SCIENCE

面向轻量化医学图像分割网络的神经结构搜索

张福昌, 仲国强, 毛玉旭

引用本文

张福昌, 仲国强, 毛玉旭. 面向轻量化医学图像分割网络的神经结构搜索[J]. 计算机科学, 2022, 49(10): 183-190.

ZHANG Fu-chang, ZHONG Guo-qiang, MAO Yu-xu. [Neural Architecture Search for Light-weight Medical Image Segmentation Network](#)[J]. Computer Science, 2022, 49(10): 183-190.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[时序知识图谱表示学习](#)

Temporal Knowledge Graph Representation Learning

计算机科学, 2022, 49(9): 162-171. <https://doi.org/10.11896/jsjcx.220500204>

[基于 Key-Value 关联记忆网络的知识图谱问答方法](#)

Key-Value Relational Memory Networks for Question Answering over Knowledge Graph

计算机科学, 2022, 49(9): 202-207. <https://doi.org/10.11896/jsjcx.220300277>

[基于安全多方计算和差分隐私的联邦学习方案](#)

Federated Learning Scheme Based on Secure Multi-party Computation and Differential Privacy

计算机科学, 2022, 49(9): 297-305. <https://doi.org/10.11896/jsjcx.210800108>

[基于非局部注意力生成对抗网络的视频异常事件检测方法](#)

Non-local Attention Based Generative Adversarial Network for Video Abnormal Event Detection

计算机科学, 2022, 49(8): 172-177. <https://doi.org/10.11896/jsjcx.210600061>

[基于深度学习的社交网络舆情信息抽取方法综述](#)

Survey of Social Network Public Opinion Information Extraction Based on Deep Learning

计算机科学, 2022, 49(8): 279-293. <https://doi.org/10.11896/jsjcx.220300099>

面向轻量化医学图像分割网络的神经结构搜索

张福昌 仲国强 毛玉旭

中国海洋大学信息科学与工程学部 山东 青岛 266100

(zfc_1527@163.com)

摘要 现有的性能优异的医学图像分割模型大都由领域专家手动设计,设计过程往往需要大量专业知识和反复实验。此外,过度复杂的分割模型不仅对硬件资源有较高要求,且分割效率较低。为此,提出了用于自动构建轻量化医学图像分割网络的神经结构搜索方法 Auto-LW-MISN(Automatically Light-Weight Medical Image Segmentation Network)。通过构建轻量级搜索空间、设计适用于医学图像分割的搜索超网络、设计添加复杂性约束的可微分搜索策略,建立用于自动搜索轻量化医学图像分割网络的神经结构搜索框架。在显微镜细胞图像、肝脏 CT 图像和前列腺 MR 图像等数据集上进行实验,结果表明,Auto-LW-MISN 能够针对不同模态的医学图像自动构建轻量化的分割模型,其分割精度相比 U-net, Attention U-net, Unet++ 和 NAS-Unet 等方法均有提高。

关键词: 深度学习;可微分神经结构搜索;轻量化卷积神经网络;自动化网络结构设计;医学图像分割

中图法分类号 TP391

Neural Architecture Search for Light-weight Medical Image Segmentation Network

ZHANG Fu-chang, ZHONG Guo-qiang and MAO Yu-xu

College of Information Science and Engineering, Ocean University of China, Qingdao, Shandong 266100, China

Abstract Most of the existing medical image segmentation models with excellent performance are manually designed by domain experts. The design process usually requires a lot of professional knowledge and repeated experiments. In addition, the over complex segmentation model not only has high requirements for hardware resources, but also has low segmentation efficiency. A neural architecture search method named Auto-LW-MISN(Automatically Light-weight Medical Image Segmentation Network) is proposed for automatic construction of light-weight medical image segmentation network. In this paper, by constructing a light-weight search space, designing a search super network for medical image segmentation, and designing a differentiable search strategy with complexity constraints, a neural architecture search framework for automatic search of light-weight medical image segmentation network is established. Experimental results on microscope cell images, liver CT images and prostate MR images show that Auto-LW-MISN can automatically construct light-weight segmentation models for different modes of medical images, and its segmentation accuracy is improved compared with U-net, Attention U-net, Unet++ and NAS-Unet.

Keywords Deep learning, Differentiable neural architecture search, Light-weight convolutional neural networks, Automatic network architecture design, Medical image segmentation

1 引言

医学图像分割有助于医生快速定位器官组织或病变位置,为后续开展疾病诊断和医学研究提供了重要依据。然而,随着临床医学影像数据量的快速增加,医生在阅片和处理图像时需要耗费大量时间和精力。因此,当前国内外均十分重视计算机辅助医学影像处理的研究^[1-3]。近年来,随着深度

学习在自然语言处理、目标检测等领域取得重大突破,基于卷积神经网络(CNNs)的医学图像分割逐渐成为深度学习与医学交叉融合过程中的重要研究方向。

为了将深度学习技术应用于医学图像分割任务,研究人员提出了很多基于 CNNs 的医学图像分割模型。2016 年,文献[4]提出了 3D FCN(Fully Convolutional Network)模型,该模型被用于脊柱 MR 医学影像分割并取得了优异效果。

到稿日期:2021-08-05 返修日期:2022-03-11

基金项目:国家重点研发计划(2018AAA0100400);装备预研教育部联合基金(6141A020337);山东省自然科学基金(ZR2020MF131);福建省医疗数据挖掘与应用工程技术研究中心开放课题(MDM2018007);青岛市科技计划(21-1-4-ny-19-nsh)

This work was supported by the National Key Research and Development Program of China(2018AAA0100400), Joint Fund of the Equipments Pre-Research and Ministry of Education of China(6141A020337), Natural Science Foundation of Shandong Province(ZR2020MF131), Open Fund of Engineering Research Center for Medical Data Mining and Application of Fujian Province(MDM2018007) and Science and Technology Program of Qingdao(21-1-4-ny-19-nsh).

通信作者:仲国强(gqzhong@ouc.edu.cn)

同年,文献[3]针对细胞分割问题提出了 U-net 模型,该模型因为引入了跳跃连接来减少信息损失的设计和独特的编码器-解码器结构,在许多医学图像分割任务上表现优异,逐渐成为了最受欢迎的医学图像分割模型之一。随后,文献[5]在 U-net 模型的解码器中加入注意力门控(Attention Gates)机制,提出了 Attention U-net 模型,该模型在腹部器官分割任务上取得了优异的结果。借鉴残差网络[6]和循环卷积神经网络(Recurrent Convolutional Neural Network, RCNN)的思想,文献[7]在 U-net 模型的基础上构建了在视网膜血管分割和皮肤癌分割任务上表现性能良好的 R2U-net 模型。

上述模型的提出在一定程度上促进了深度学习与医学领域的融合。然而,目前主要有两方面的问题制约着深度学习在医学领域的进一步发展:1)当前性能优异的医学图像分割网络大都由经验丰富的领域专家手工设计,不仅设计过程繁琐,而且难以针对不同模态的医学图像设计出所需的神经网络模型;2)现有医学图像分割模型通常参数量较大且计算复杂度较高,不仅对医疗检测设备的计算模块和内存资源需求高,而且难以达到实时分割的效果。因此,探索一种自动构建轻量化医学图像分割模型的方法是十分必要的。

神经结构搜索(Neural Architecture Search, NAS)方法极大地降低了网络结构设计问题中的人工成本,拓展了轻量化网络结构设计的可能性。近年来, DARTS[8], DenseNAS[9], SGAS[10], FBNetV2[11], FBNetV3[12]等 NAS 方法在图像分类等任务上搜索得到的网络结构的性能已经达到甚至超过人工设计的网络结构。但是,目前基于神经结构搜索的医学图像分割研究较少。Weng 等[13]基于神经结构搜索方法自动构建了类似 U-net 结构的医学图像分割网络 NAS-Unet,并将其成功应用于神经超声影像、肝脏影像等分割任务。Mortazi 等[14]基于强化学习中的策略梯度方法,在心脏 MR 图像上自动构建了性能优异的分割模型。现有的基于神经结构搜索的医学图像分割研究在分割精度以及模型复杂度等方面有较大的提升空间。

本文提出了基于神经结构搜索的自动构建轻量化医学图像分割网络,通过设计轻量化搜索空间、搜索超网络以及添加复杂性约束的可微分搜索策略, Auto-LW-MISN 可以针对不同模态数据集实现轻量化医学图像分割网络的自动设计,为深度学习辅助医疗诊断提供帮助。

2 本文方法

2.1 轻量化搜索空间

搜索空间的设计是神经结构搜索中的关键步骤,其对模型精度、参数量和计算量(FLOPs)等性能指标有直接影响。最近,有研究人员提出将称为“Cell”的网络模块作为搜索空间,并堆叠若干个模块构建完整的深度神经网络[8]。本文结合已有的研究工作,设计了一个轻量化搜索空间,搜索空间中包含最新的轻量化操作,旨在从神经结构搜索的源头控制搜索得到的神经结构的复杂度。具体地,本文设计的搜索空间拓扑结构如图1所示,该结构可抽象为一个由若干节点和边构成的有向无环图。其中,编号为 C_1 和 C_2 的两个框表示输入节点,编号为 0-3 的框表示中间节点,节点可看作网络层。图1中有向边表示用于特征图变换的候选操作集合,如恒等映射、 3×3 空洞卷积等。从搜索空间的拓扑结构来看,网络层间的变换操作集合是影响搜索空间复杂度的决定性因素。

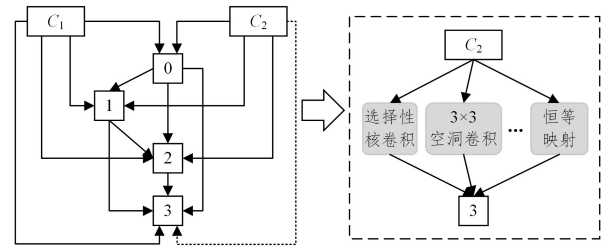


图1 搜索空间拓扑结构示意图

Fig. 1 Topology diagram of search space

为了设计轻量化的搜索空间,本文在选取搜索空间中的候选操作时,综合考虑变换操作的性能、参数量以及计算量等影响因素。首先,限制操作的感受野大小为 3×3 。当前方法的搜索空间通常包含 5×5 甚至 7×7 的卷积操作,相比 3×3 卷积,它们的参数量和浮点计算量更大。此外,沿宽度和深度维度堆叠 3×3 卷积可达到与 5×5 卷积相似的效果。其次,确保搜索空间中任一操作无法由其他操作近似或等效代替。文献[15]已证明,现有可微分搜索方法的候选操作集存在多重共线性,即多个操作输出的特征图存在高度线性相关现象,例如平均池化和最大池化操作。本文为了减弱候选操作集中不同操作之间的相互制约,在构建搜索空间时避免出现具有多重共线性的操作。最终,本文构建的搜索空间包括8个高精度、轻量化的操作,如表1中第一列所示,分别是恒等映射、 3×3 空洞卷积[16]、 3×3 深度可分离卷积、“幻影”卷积[17]、选择性核卷积[18]、 3×3 混洗分组卷积[19]、 3×3 最大池化和 Zero 操作(表示节点间无连接)。

表1 本文搜索空间与 DARTS 搜索空间的复杂度对比

Table 1 Comparison between search space designed in this paper and that used in DARTS

Auto-LW-MISN(本文)	Params/kB	FLOPs/MB	DARTS	Params/kB	FLOPs/MB
恒等映射	—	—	恒等映射	—	—
3×3 深度可分离卷积	2.81	2.68	5×5 深度可分离卷积	2.81	2.68
3×3 空洞卷积	1.41	1.34	3×3 深度可分离卷积	3.86	3.07
“幻影”卷积	0.74	0.67	3×3 空洞卷积	1.41	1.34
选择性核卷积	2.44	0.65	5×5 空洞卷积	1.93	1.64
3×3 混洗分组卷积	0.30	0.20	3×3 最大池化	—	—
3×3 最大池化	—	—	3×3 平均池化	—	—
Zero	—	—	Zero	—	—
平均值	1.54	1.11	平均值	2.50	2.18

为了证明本文设计的搜索空间更轻量化,与现在流行的 DARTS^[8] 搜索空间进行复杂度对比,比较二者参数量和计算量的差异。以一对节点间的候选操作为例,假设输入特征图的分辨率为 32×32 ,卷积核大小为 3×3 ,输入输出通道数皆为 32,卷积操作步长均设为 1。此外,混洗分组卷积分组数为 4,且对空洞卷积也进行分组(分组数与输入通道数相同)。为了公平比较,对有参操作的参数量和和计算量取均值,计算式如式(1)、式(2)所示:

$$Params := \frac{1}{N} \sum_{o \in O} p_o, \text{ s. t. } p_o > 0 \quad (1)$$

$$FLOPs := \frac{1}{N} \sum_{o \in O} f_o, \text{ s. t. } f_o > 0 \quad (2)$$

其中, O 表示候选操作集, p_o 和 f_o 分别表示操作 o 的参数量和和浮点计算量, N 表示有参操作数量。二者的对比结果如表 1 所列。可见,本文设计的搜索空间比 DARTS 搜索空间的参数量和浮点计算量更少。

2.2 医学图像分割搜索超网络

U-net^[3] 模型是流行的图像分割模型之一,本文从基于

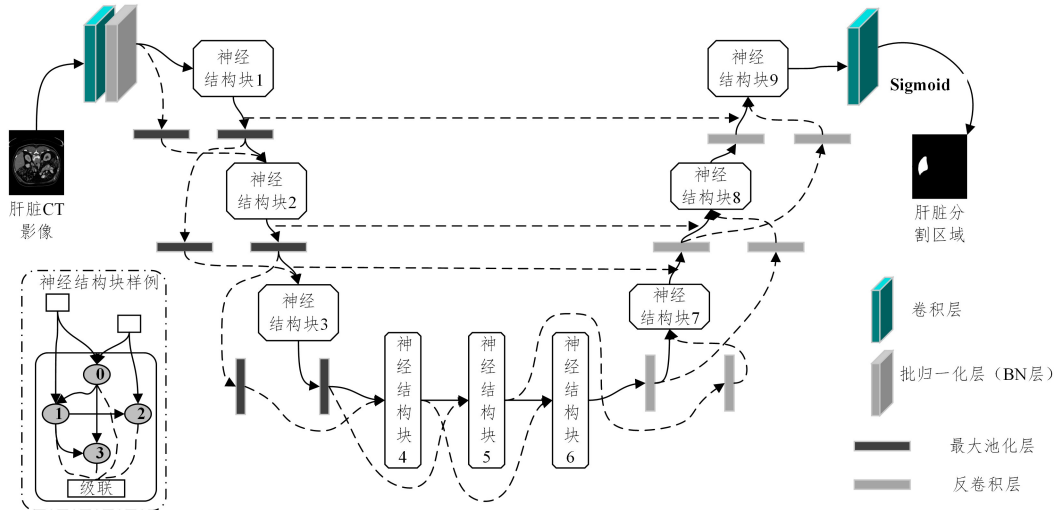


图 2 医学图像分割搜索超网络示意图

Fig. 2 Schematic diagram of medical image segmentation search hypernetwork

此外,为了降低编码器中降采样带来的特征信息损失,使用横向连接的方式连接通道维度级联编码器与解码器中对应位置的特征图。同时,在整个超网络中,除第一个神经结构块以外,其他神经结构块的输入均包含之前两个神经结构块的输出,这种设计不仅可以减少信息损失,还能降低反向传播过程中梯度爆炸或消失的可能性,使模型训练更稳定。

本文构建的搜索超网络融合了 U-net 结构和传统搜索超网络的设计思想,实现了神经结构搜索与医学图像分割任务的结合,进而可以通过自动化方式搜索用于医学图像分割的轻量化神经结构块。

2.3 添加复杂性约束的可微分搜索策略

如 2.1 节中的图 1 所示,本文设计的搜索空间结构中,信息只能由索引小的节点流向索引大的节点,输入节点的特征图经候选操作变换后可以流向任意中间节点。中间节点 j 由其所有前驱节点特征图经候选操作变换后在通道维度逐像素相加计算得到,如式(3)所示:

神经结构搜索的网络自动设计角度出发,构建了一个新的 U 型搜索超网络(见图 2),用于搜索轻量化、高性能的医学图像分割神经结构块。

具体地,搜索超网络主干由多个神经结构块组成,对于特征复杂度和影像数量等存在差异的不同模态数据集,搜索超网络的深度设置不同(详见实验设置部分)。为了限制最终分割网络的模型复杂度,使用 2.1 节设计的轻量化搜索空间构建神经结构块,整个搜索超网络是一个对称结构。左侧为编码器,其中每个神经结构块后跟随一个接受域为 2×2 的最大池化层,每经过一个最大池化层,特征图尺寸降采样为原来的一半,编码器能够提取医学图像的高层语义信息。为了加深网络深度,增强模型特征的学习能力,中间部分顺序堆叠 3 个神经结构块。右侧为解码器,解码器中神经结构块数量与编码器相同,并且特征图输入到神经结构块之前,都会经过一个反卷积层将特征图尺寸扩大为原来的两倍,旨在将抽象的特征逐步还原到原图尺寸,最后通过 Sigmoid 函数输出分割结果。

$$x^{(j)} = \sum_{i < j} O^{(i,j)}(x^{(i)}) \quad (3)$$

其中, $x^{(i)}$ 和 $x^{(j)}$ 分别表示索引为 i 和 j 的节点的特征图, $O^{(i,j)}$ 表示节点 i 到节点 j 的候选操作集。为了使整个搜索空间连续,使用 Softmax 函数将相互独立的各个离散操作进行联结,为每个操作赋予一个可学习的结构权重参数 α , 如式(4)所示:

$$O^{(i,j)} = \sum_{o \in O} \frac{\exp(\alpha_o^{(i,j)})}{\sum_{o' \in O} \exp(\alpha_{o'}^{(i,j)})} o(x), \text{ s. t. } i < j \quad (4)$$

其中, $\alpha_o^{(i,j)}$ 表示节点 i 到节点 j 间候选操作 o 的结构参数, $\alpha_o^{(i,j)}$ 的值代表特征图从节点 i 到节点 j 的变换过程中操作 o 的贡献度大小。使用 Softmax 函数后,搜索空间中的离散候选操作变为连续可微,进而可以使用策略梯度方法优化结构参数。

本文将参数量和计算量建模为两个连续函数,因此可以在添加模型参数量和计算量作为复杂性约束条件时保持目标函数的可微性。具体地,一对节点间的参数量和计算量期望

值可分别表示为:

$$E_{[\text{params}]^{(i,j)}} = \sum_{o \in O} (\alpha_o^{(i,j)} \times P(o^{(i,j)})) \quad (5)$$

$$E_{[\text{flops}]^{(i,j)}} = \sum_{o \in O} (\alpha_o^{(i,j)} \times F(o^{(i,j)})) \quad (6)$$

其中, $P(\cdot)$ 和 $F(\cdot)$ 分别表示操作的参数和浮点计算函数。将参数量和和计算量转换为式(5)和式(6)的形式, 便能够通过策略梯度指导结构参数更新, 如式(7)所示:

$$\begin{cases} \frac{\partial E_{[\text{params}]^{(i,j)}}}{\partial \alpha_o^{(i,j)}} = P(o^{(i,j)}) \\ \frac{\partial E_{[\text{flops}]^{(i,j)}}}{\partial \alpha_o^{(i,j)}} = F(o^{(i,j)}) \end{cases} \quad (7)$$

一个结构块的参数和计算量由其内部所有变换操作的参数量和和计算量求和得到, 如式(8)和式(9)所示:

$$E_{[\text{params}_n]} = \sum_{i \in n} \sum_{j > i} E_{[\text{params}]^{(i,j)}} \quad (8)$$

$$E_{[\text{flops}_n]} = \sum_{i \in n} \sum_{j > i} E_{[\text{flops}]^{(i,j)}} \quad (9)$$

在结构搜索阶段, 整个超网络包含 N 个相同神经结构块, 它的参数和计算量期望是所有神经结构块的总和。

$$E_{[\text{params}]} = \sum_{n \in N} E_{[\text{params}_n]} \quad (10)$$

$$E_{[\text{flops}]} = \sum_{n \in N} E_{[\text{flops}_n]} \quad (11)$$

为了搜索得到在医学图像分割任务上性能优异的轻量化神经结构块, 本文综合考虑多个指标, 提出式(12)所示的损失函数。

$$L = BCE + Dice + (\lambda_1 E_{[\text{params}]} + \lambda_2 E_{[\text{flops}]}) \quad (12)$$

其中, BCE(Binary Cross Entropy)表示二值交叉熵损失, 计算式如式(13)所示:

$$BCE = - \sum_i (y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)) \quad (13)$$

其中, y_i 是位置 i 处的真实标签, 本文对样本标签做了预处理, 所有位置上的像素值只能取 0 或 1; \hat{y}_i 表示模型在位置 i 处的预测值。

式(12)中, λ_1 和 λ_2 是权衡准确率、参数量以及计算量的尺度因子。Dice 表示集合相似性度量函数, 对于医学图像分割任务, 该函数可用于衡量模型输出的分割区域与真实标签中的标注区域之间的差异大小, 其计算式如式(14)所示:

$$Dice = 1 - \frac{2 \times |\hat{Y} \cap Y| + smooth}{|\hat{Y}| + |Y| + smooth} \quad (14)$$

其中, $|\hat{Y} \cap Y|$ 表示模型预测区域与实际标注区域间的交集, 分子系数为 2 的原因是计算分母时存在重复区域, 为了避免分母为 0, 加入一个接近 0 的浮点数 $smooth$ 。为方便计算, 本文在计算 Dice 时对预测特征图进行了离散化处理。具体地, 特征图中小于或等于 0.5 的预测值统一置为 0, 大于 0.5 的预测值置为 1。实际上, $|\hat{Y} \cap Y|$ 指样本标签值和预测特征图中同为 1 的区域。

搜索阶段结束时, 基于学得的结构参数 α 对神经结构块进行剪枝处理, 为每个中间节点仅保留 α 值最大的两个操作。最后, 在训练数据集上对剪枝完成的网络进行重训练, 以微调网络中可训练的参数, 提高模型的分割精度。

3 实验与结果分析

为了验证本文提出的 Auto-LW-MISN 方法在自动构建

轻量化医学图像分割网络任务上的有效性, 在显微镜细胞图像、肝脏 CT 影像和前列腺 MR 影像 3 种模态的真实医学图像数据上进行了实验。

3.1 显微镜细胞图像数据集的介绍及实验设置

本文使用的显微镜细胞图像数据集是 2018 年 kaggle 数据科学碗(Data Science Bowl, DSB)比赛中使用的数据集^[20]。其包含 670 张图像, 既包含彩色图像, 也包含黑白图像, 且每张图像都对应有 1 个真实标签。为了方便训练, 将黑白图像均转换为三通道相同的 RGB 图像。同时, 将所有图像大小统一为 256 像素 \times 256 像素, 将 670 张图像按照 7:3 的比例划分为训练集和验证集。

在显微镜细胞图像数据集上, 构建由 7 个神经结构块组成的搜索超网络, 其中编码器和解码器各包含两个, 中间部分包含 3 个。在搜索过程中, 使用 Cutout 方法^[21]对图像进行数据增强, Cutout Length 设置为 16, 总共迭代 25 个 epochs, Batch Size 大小为 2, 结构参数和网络参数的初始学习率分别设置为 0.0006 和 0.005。为了使训练过程更加稳定, 在每个 epoch 后使用余弦退火技术对学习率进行调整。此外, 使用 Adam 和 SGD 优化器分别对结构参数和网络参数进行优化, Adam 和 SGD 中权重衰减分别设置为 0.001 和 0.0003。为了提高学得的结构鲁棒性, 训练中使用操作级的 Dropout 技术, 并设置 Drop path prob 为 0.3, 损失函数中的尺度因子 λ_1 和 λ_2 分别设置为 0.015 和 0.005。

超网络训练结束后, 采用 DARTS 的离散处理方法对神经结构块进行剪枝^[8], 并继承超网络的拓扑结构来构建轻量化分割网络。在模型重训练过程中(即结构评估阶段), Drop path prob 设置为 0.15, 初始学习率调整为 0.006, Batch Size 设置为 4, 使用权重衰减为 0.0002 的 Adam 进行参数优化, 总共被重训练 30 个 epochs。由于重训练阶段不需要对模型进行复杂性约束, 因此 λ_1 和 λ_2 均设置为 0。

在该数据集上, 将基于本文方法自动构建的分割网络与 U-net^[3], Attention U-net^[5], Unet++^[22] 等目前最优的手动设计的医学图像分割模型进行了对比, 同时也与当前基于神经结构搜索的医学图像分割方法进行了比较, 如 NAS-Unet^[13] 等, 以验证本文方法的有效性, 对比结果将在 3.5 节中展示。

3.2 肝脏 CT 图像数据集的介绍及实验设置

本文使用的肝脏 CT 影像数据集选自 2019 年生物医学影像国际研讨会中公开的数据集 CHAOS^[23], 该数据集包含 MR 影像和 CT 影像两种模态的数据。其中, MR 影像用于分割肾、脾等器官, CT 影像主要用于分割肝脏器官。本文选用 CT 影像数据进行实验, 总共包含 2874 个格式为 Dicom 文件的数据, 图像尺寸大小为 512 像素 \times 512 像素。使用 pydicom 包对原始数据进行处理, 并将读取的 CT 图像的像素范围控制在 0-1 之间, 图像大小均缩小为 256 像素 \times 256 像素, 同时对处理后的图像进行归一化, 均值和方差分别设置为 0.3667 和 0.3533。此外, 为了拟合 CT 图像的分辨率, 图像的掩码(真实标签)尺寸也调整为 256 像素 \times 256 像素。预处理后的肝脏 CT 图像如图 3 所示。

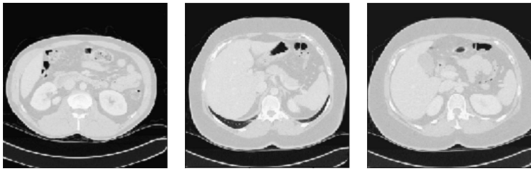


图3 肝脏 CT 影像图像示例

Fig. 3 Examples of liver CT images

在肝脏 CT 图像数据集上,构建由 9 个神经结构块组成的搜索超网络,其中编码器、解码器和中间部分各包含 3 个结构块。在结构搜索阶段,所有图像按 1:1 的比例划分为训练集和验证集。搜索过程中共迭代 35 个 epochs,结构参数和网络参数的初始学习率分别设置为 0.0006 和 0.008。损失函数中的 λ_1 和 λ_2 尺度因子分别设置为 0.025 和 0.015。其他超参数设置以及评估阶段的参数设置均与显微镜细胞图像数据集一致(详见 3.1 节)。

在该数据集上,将基于本文方法自动构建的分割网络与 U-net^[3], Attention U-net^[5], Unet++^[22] 等目前最优的手动设计的医学图像分割模型进行了对比,同时也与当前基于神经结构搜索的医学图像分割方法进行了比较,如 NAS-Unet^[13], AutoDeepLab^[24] 等,实验结果将在 3.5 节中展示。

3.3 前列腺 MR 图像数据集的介绍及实验设置

本文使用的前列腺 MR 图像数据集是全球核磁共振医学图像前列腺分割(Prostate MR Image Segmentation)大赛中的公开数据集 PROMISE12^[25],该数据集共包含大约 1377 张图像,主要用于解决医学图像中的前列腺分割难点问题。首先,使用 SimpleITK 包将 .mhd 格式的原始数据处理为用像素值表示的图像数据。此外,将所有图像尺寸统一调整为 256 像素×256 像素。同时,由于 MR 图像受检测设备等因素的影响,前列腺区域噪声比较明显,因此进一步使用 SimpleITK 包对图像进行去噪处理。预处理后的前列腺 MR 图像如图 4 所示。



图4 前列腺 MR 影像图像示例

Fig. 4 Examples of prostate MR images

在前列腺 MR 图像数据集上,构建由 8 个神经结构块组成的搜索超网络,其中编码器和解码器各包含 3 个,中间部分包含 2 个。在结构搜索阶段,所有图像按 3:2 的比例划分为训练集和验证集。搜索过程中共迭代 30 个 epochs,损失函数中的 λ_1 和 λ_2 尺度因子分别设置为 0.02 和 0.01。其他超参数设置以及评估阶段的参数设置均与显微镜细胞图像数据集一致(详见 3.1 节)。

在该数据集上,将基于本文方法自动构建的分割网络与 U-net^[3], Attention U-net^[5], Unet++^[22] 等目前最优的手动设计的医学图像分割模型进行了对比,同时也与当前基于神经结构搜索的医学图像分割方法进行了比较,如 NAS-Unet^[13], AutoDeepLab^[24] 等,实验结果将在 3.5 节中展示。

3.4 评价指标

实验采用的评价指标包括平均交并比(Mean Intersection over Union, MIoU)、Dice 相似系数(Dice Similarity Coefficient, DSC)、模型参数数量和计算量,前两者是图像分割领域的常用评价指标,用于评估模型对目标区域的分割精度,而模型参数数量和计算量可以用于衡量模型的复杂度。

对于 MIoU 指标,给定图像的真实标签掩码 Y 和预测掩码 \hat{Y} ,首先计算每个类上的 IoU 值,计算式如式(15)所示:

$$IoU = \frac{TP}{TP + FP + FN} \quad (15)$$

假设正样本为分割区域,负样本为图像中的其他区域。式(15)中, TP 表示实际上是正样本且被预测为正样本的区域(真正例), FP 表示实际上是负样本但被预测为正样本的区域(假正例), FN 表示实际上是正样本但被预测为负样本的区域(假反例)。然后,对所有类上的 IoU 值取平均得到该图像的 $MIoU$ 值。由于本文实验中所用医学图像数据仅包含一类需要分割的组织或器官,因此对于一张图像来说 IoU 等于 $MIoU$,结果中展示的 $MIoU$ 值是验证集中所有图像 IoU 的均值。

Dice 相似系数(DSC)用于衡量真实标签掩码与预测掩码之间的相似度,计算式如式(16)所示:

$$DSC = \frac{2 \times TP}{FP + 2 \times TP + FN} \quad (16)$$

3.5 实验结果及可视化

本节主要展示与分析在显微镜细胞图像、肝脏 CT 图像和前列腺 MR 图像 3 个数据集上的实验结果。

表 2 列出了本文方法构建的轻量化医学图像分割模型与其他模型在显微镜细胞图像数据集上的对比结果,对比模型主要是近年来手动设计的医学图像分割网络。由于 U-net, Attention U-net 等模型在原文献中并未展示在该数据集上的实验结果,因此基于这些模型的公开源代码在该数据集上进行了复现实验,并且实验过程中都使用与本文方法相同的方式进行数据预处理。同时,为了使对比结果更具说服力,表中展示的结果是重复 4 次实验的结果的平均值。由表 2 所列的对比结果可以看出,通过 Auto-LW-MISN 方法自动构建的模型在细胞核分割任务上表现优异,不仅 MIoU 和 DSC 指标均超过了其他对比方法,模型复杂度较手动设计的网络也显著降低。

表2 不同方法在显微镜细胞图像数据集上的对比结果

Table 2 Comparison results of different methods on microscope cell image dataset

方法	MIoU/%	DSC/%	参数量/MB	计算量/GB
U-net	80.57	87.77	9.81	20.17
FCN	84.03	90.17	18.64	25.50
Attention U-net	78.70	86.54	34.88	66.23
UNet++	83.79	90.12	9.16	35.00
R2U-net	77.50	84.87	39.09	153.18
DARTS_Based	79.37	86.02	7.01	16.69
NAS-Unet	83.15	89.18	10.95	17.43
Ours	85.75	92.11	0.57	15.76

表 3 和表 4 分别列出了通过 Auto-LW-MISN 方法自动构建的分割模型在肝脏 CT 图像和前列腺 MR 图像数据集上

的表现。实验中进行对比的模型既包括 U-net, Attention U-net 等手动设计的分割网络, 也包括基于神经结构搜索方法自动构建的图像分割网络, 如 AutoDeepLab, NAS-Unet 等。为了公平比较, 所有模型在训练过程中均使用相同的数据预处理方法(实验设置中已介绍)。通过表 3 和表 4 列出的对比结果可以发现, 本文方法自动构建的模型在肝脏 CT 图像数据集和前列腺 MR 图像数据集上均取得了最佳结果, 特别是在前列腺 MR 数据集上, 本文方法的 MIoU 和 DSC 指标超过 U-net, Attention U-net 等模型近 20%。

表 3 不同方法在肝脏 CT 图像数据集上的对比结果

Table 3 Comparison results of different methods on liver CT image dataset

方法	MIoU/%	DSC/%	参数量/MB	计算量/GB
U-net	88.00	91.96	9.81	20.17
Attention U-net	91.60	94.64	34.88	66.23
R2U-net	91.63	95.13	39.09	153.18
UNet++	90.20	93.53	9.16	35.00
AutoDeepLab	92.01	95.89	15.43	36.18
NAS-Unet	92.32	96.54	12.33	27.15
DARTS_Based	88.02	90.97	7.11	18.17
Ours	93.17	97.86	0.56	15.03

表 4 不同方法在前列腺 MR 图像数据集上的对比结果

Table 4 Comparison results of different methods on prostate MR image dataset

方法	MIoU/%	DSC/%	参数量/MB	计算量/GB
U-net	50.06	57.92	9.81	20.17
Attention U-net	56.22	61.41	34.88	66.23
R2U-net	50.59	57.96	39.09	153.18
UNet++	51.66	58.65	9.16	35.00
AutoDeepLab	58.18	68.20	15.43	36.18
NAS-Unet	64.82	72.43	12.33	27.15
DARTS_Based	50.16	57.06	6.69	17.65
Ours	66.88	72.71	0.61	15.98

综合上述实验结果可以发现, 本文提出的 Auto-LW-MISN 方法构建的医学图像分割模型在多种模态的数据集上均表现优异, 这说明 Auto-LW-MISN 方法在医学图像分割任务上具有普遍适用性。首先, 从分割精度来看, 本文方法 Auto-LW-MISN 在实验的 3 个医学图像数据集(显微镜细胞图像、肝脏 CT 图像和前列腺 MR 图像)上均取得了最佳效果, 超过了人工设计和自动构建的医学图像分割模型, 这表明 Auto-LW-MISN 方法能够针对不同模态的医学图像数据集自动构建高精度的神经网络模型。其次, 从模型复杂度来看, 使用本文方法自动构建的模型显著优于其他分割网络。具体地, 在 3 个数据集上的模型参数量均不超过 1MB, 浮点计算量都控制在 15GB 左右, 参数量仅是 U-net 的 1/16、NAS-Unet 的 1/20, 浮点计算量也远低于其他对比模型。如此低复杂度的模型完全可以嵌入到普通医疗检测设备中, 这进一步证明了本文设计的轻量化搜索空间和添加复杂性约束的搜索策略对搜索轻量化神经结构有着至关重要的作用, 同时也说明 Auto-LW-MISN 方法自动构建的医学图像分割模型可以兼顾模型精度和模型复杂度。

为了更直观地展示本文 Auto-LW-MISN 方法的有效性, 对部分图像分割结果进行了可视化。如图 5 所示, 第一列

表示输入神经网络模型前的原始图像; 第二列是输入图像对应的真实标签, 即领域专家手工标记的区域分割掩码; 第三列和第四列分别对应 U-net 模型和 NAS-Unet 模型的分割结果, U-net 是当前最受欢迎的基于人工设计的医学图像分割模型之一, 而 NAS-Unet 是当前鲜有的基于神经结构搜索方法自动构建医学图像分割模型的方法之一; 最后一列展示了本文方法构建的轻量化医学图像分割模型在 3 个数据集上的分割效果图。图 5 中前两行是在显微镜细胞图像数据集上的结果, 中间两行是在肝脏 CT 图像数据集上的结果, 最后两行是在前列腺 MR 图像数据集上的结果。从图 5 给出的结果比较图可看出, 本文方法的分割结果与图像真实标签更接近。这进一步说明本文方法的分割效果优于对比模型, 能够准确标记出输入图像中器官或者组织的位置区域, 这对医生进行临床诊断和医学研究具有重要价值。

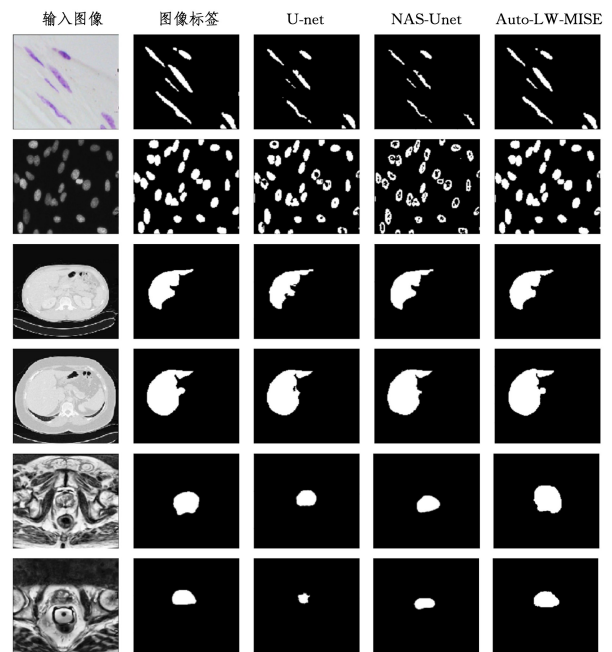


图 5 部分医学图像的分割可视化结果图

Fig. 5 Visualization results of part of medical image segmentation

3.6 不同拓扑结构的超网络对比实验

为了验证本文设计的搜索超网络在医学图像分割任务上的有效性, 针对搜索超网络设置了对比实验。对比的超网络结构如图 6 所示。具体地, 对比超网络的设计参考 DARTS^[8] 等流行可微分结构搜索方法, 其由两种类型的神经结构块构成, 分别是标准神经结构块和降维神经结构块, 两者的结构基本相同(与前述本文方法所设计的神经结构块相同), 唯一的差别是降维神经结构块的输入特征图尺寸是输出特征图的两倍, 而标准神经结构块的特征图尺寸始终保持一致。对比超网络中降维神经结构块的位置处在超网络的 1/3 和 2/3 深度位置处。对比超网络的头部是一个包含 3×3 标准卷积层和批归一化层的复合结构。对比超网络最后通过 Sigmoid 激活函数输出分割结果。除搜索超网络的拓扑结构不同之外, 对比超网络上的实验设置与本文方法的实验设置完全相同。在 3 个数据集上的实验结果如表 5 所列。

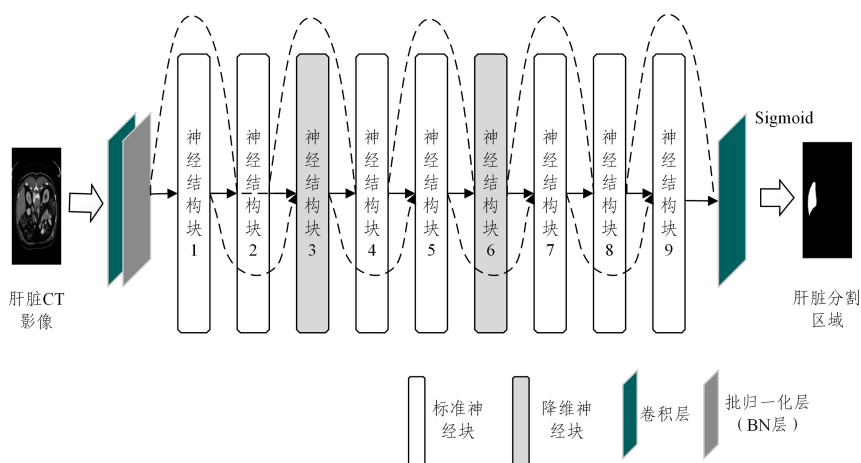


图6 对比超网络示意图

Fig. 6 Schematic diagram of comparative search hypernetwork

表5列出了本文设计的超网络和对超网络对医学图像分割的影响。分别使用上述两个超网络在本文使用的3个数据集上构建了医学图像分割模型并测试了模型的分割精度。由表5可看出,通过本文设计的超网络所构建的医学图像分割模型在3个数据集上的表现均优于对比超网络,这证明本文针对医学图像分割任务所设计的搜索超网络是非常有效的,能够用于自动构建高质量的分割模型。

表5 不同拓扑结构的超网络对医学图像分割的影响

Table 5 Influence of hypernetworks with different topologies on medical image segmentation

(单位:%)

数据集	本文超网络		对比超网络	
	MIoU	DSC	MIoU	DSC
显微镜细胞图像	85.75	92.11	82.16	89.73
肝脏CT图像	93.17	97.86	91.15	95.31
前列腺MR图像	66.88	72.71	63.01	69.78

结束语 针对深度学习与医学领域交叉融合过程中的实际应用问题——医学图像分割,本文提出了自动构建轻量化医学图像分割网络的方法 Auto-LW-MISN。该方法设计了轻量化搜索空间和适用于图像分割的搜索超网络,并从图像分割精度和模型轻量化两方面进行综合考虑,设计了添加复杂性约束的可微分搜索策略。实验结果表明,Auto-LW-MISN在医学图像分割任务上表现优异,不仅达到了与最优模型相当的分割精度,而且构建的分割网络的模型参数量和计算量远低于当前医学图像分割模型。后续工作中需要收集更多的医学影像数据,为轻量化医学图像分割网络的自动构建提供数据支撑。此外,下一步将设计更加有效的轻量化搜索空间和搜索策略,探索更有效的神经结构搜索方法。

参考文献

[1] LIU C, XIAO Z Y, DU N M. Application of Improved Convolutional Neural Network in Medical Image Segmentation [J]. Computer Science and Exploration, 2019, 13(9): 1593-1603.

[2] XU H, SUI L, ZHANG J W, et al. Research Progress of Convolutional Neural Network in Medical Image Segmentation [J].

Chinese Journal of Medical Physics, 2019, 36(11): 1302-1306.

[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. Springer, 2015: 234-241.

[4] KOREZ R, LIKAR B, PERNUS F, et al. Model-Based Segmentation of Vertebral Bodies from MR Images with 3D CNNs [C]// International Conference on Medical Image Computing and Computer-assisted Intervention. Springer, 2016: 433-441.

[5] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-net: Learning Where to Look for the Pancreas[J]. arXiv: 1804. 03999, 2018.

[6] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2016: 770-778.

[7] 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.

[8] LIU H, SIMONYAN K, YANG Y. DARTS: Differentiable Architecture Search[J]. arXiv: 1806. 09055, 2018.

[9] FANG J, SUN Y, ZHANG Q, et al. Densely Connected Search Space for More Flexible Neural Architecture Search [C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition. Computer Vision Foundation/IEEE, 2020: 10625-10634.

[10] LI G, QIAN G, DELGADILLO I C, et al. SGAS: Sequential Greedy Architecture Search [C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition. Computer Vision Foundation/IEEE, 2020: 1617-1627.

[11] WAN A, DAI X, ZHANG P, et al. FBNetV2: Differentiable Neural Architecture Search for Spatial and Channel Dimensions [C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition. Computer Vision Foundation/IEEE, 2020: 12962-12971.

[12] DAI X, WAN A, ZHANG P, et al. FBNetV3: Joint Architecture-Recipe Search Using Predictor Pretraining [C] // IEEE Conference on Computer Vision and Pattern Recognition. Computer

Vision Foundation/IEEE, 2021; 16276-16285.

- [13] WENG Y, ZHOU T, LI Y, et al. NAS-Unet: Neural Architecture Search for Medical Image Segmentation[J/OL]. IEEE Access, 2019, 7: 44247-44257. <https://ieeexplore.ieee.org/document/8681706>.
- [14] MORTAZI A, BAGCI U. Automatically Designing CNN Architectures for Medical Image Segmentation[C]// Machine Learning in Medical Imaging — 9th International Workshop. Springer, 2018; 98-106.
- [15] LI G, ZHANG X, WANG Z, et al. StacNAS: Towards Stable and Consistent Differentiable Neural Architecture Search[J]. arXiv: 1909.11926, 2019.
- [16] YU F, KOLTUN V. Multi-scale Context Aggregation by Dilated Convolutions[J]. arXiv:1511.07122, 2015.
- [17] HAN K, WANG Y, TIAN Q, et al. GhostNet: More Features From Cheap Operations[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition. Computer Vision Foundation/IEEE, 2020; 1577-1586.
- [18] LI X, WANG W, HU X, et al. Selective Kernel Networks[C]// IEEE Conference on Computer Vision and Pattern Recognition. Computer Vision Foundation/IEEE, 2019; 510-519.
- [19] ZHANG X, ZHOU X, LIN M, et al. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices [C]// IEEE Conference on Computer Vision and Pattern Recognition. Computer Vision Foundation/IEEE Computer Society, 2018; 6848-6856.
- [20] CAICEDO J, GOODMAN A, KARHOHS K, et al. Nucleus Segmentation Across Imaging Experiments: the 2018 Data Science Bowl[J]. Nature Methods, 2019, 16(12): 1247-1253.
- [21] DEVRIES T, TAYLOR G W. Improved Regularization of Convolutional Neural Networks with Cutout [J]. arXiv: 1708.04552, 2017.
- [22] ZHOU Z, SIDDIQUEE M M R, TAJBAKHS 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.
- [23] KAVUR A E, SELVER M A, DICLE O, et al. Chaos-combined (CT-MR) Healthy Abdominal Organ Segmentation Challenge Data[C]// IEEE International Symposium on Biomedical Imaging (ISBI). Imag. (ISBI), 2019.
- [24] LIU C, CHEN L C, SCHROFF F, et al. Auto-DeepLab: Hierarchical Neural Architecture Search for Semantic Image Segmentation[C]// IEEE Conference on Computer Vision and Pattern Recognition. Computer Vision Foundation/IEEE, 2019; 82-92.
- [25] LITJENS G, TOTTH R, VAN DE VEN W, et al. Evaluation of Prostate Segmentation Algorithms for MRI: the PROMISE12 Challenge[J]. Medical Image Analysis, 2014, 18(2): 359-373.



ZHANG Fu-chang, born in 1996, post-graduate. His main research interests include neural architecture search and image processing.



ZHONG Guo-qiang, born in 1981, professor, Ph. D supervisor, is a senior member of China Computer Federation. His main research interests include pattern recognition, machine learning and computer vision.

(责任编辑:喻黎)