

轮廓指导的层级混合多任务全卷积网络*

何克磊, 史颖欢, 高阳

(计算机软件新技术国家重点实验室(南京大学), 江苏 南京 210023)

通讯作者: 高阳, E-mail: gaoy@nju.edu.cn



摘要: 传统的深度多任务网络通常在不同任务之间共享网络的大部分层(即特征表示层).由于这样做会忽视不同任务各自的特殊性,所以往往会制约其适应数据的能力.提出了一种层级混合的多任务全卷积网络 HFFCN,以解决 CT 图像中的前列腺分割问题.特别地,使用一个多任务框架来解决这个问题.这个框架包括一个分割前列腺的主任务和一个回归前列腺边界的辅助任务.这里,第 2 个任务主要是用来精确地描述在 CT 图像中模糊的前列腺边界.因此,HFFCN 架构是一个双分支的结构,包含一个编码主干和两个解码分支.不同于传统的多任务网络,提出了一个信息共享模块,用以在两个解码分支之间共享信息.这使得 HFFCN 可以学习任务的通用层级信息,同时保留一些不同任务各自的特征表示.在一个包含有 313 个病人的 313 张计划阶段图片的 CT 图像数据集上做了详细的实验.实验结果证明了 HFFCN 网络可以超越现有其他先进的分割方法或者传统的多任务学习模型.

关键词: 全卷积网络;深度学习;多任务学习;前列腺分割

中图法分类号: TP391

中文引用格式: 何克磊,史颖欢,高阳.轮廓指导的层级混合多任务全卷积网络.软件学报,2020,31(5):1573-1584. <http://www.jos.org.cn/1000-9825/5991.htm>

英文引用格式: He KL, Shi YH, Gao Y. Contour-guided hierarchically-fused multi-task fully convolutional network. Ruan Jian Xue Bao/Journal of Software, 2020,31(5):1573-1584 (in Chinese). <http://www.jos.org.cn/1000-9825/5991.htm>

Contour-guided Hierarchically-fused Multi-task Fully Convolutional Network

HE Ke-Lei, SHI Ying-Huan, GAO Yang

(State Key Laboratory for Novel Software Technology (Nanjing University), Nanjing 210023, China)

Abstract: Conventional multi-task deep networks typically share most of the layers (i.e., layers for feature representations) across all tasks, which may limit their data fitting ability, as specificities of different tasks are inevitably ignored. This study proposes a hierarchically-fused multi-task fully-convolutional network, called HFFCN, to tackle the challenging task of prostate segmentation in CT images. Specifically, prostate segmentation is formulated into a multi-task learning framework, which includes a main task to segment prostate, and a supplementary task to regress prostate boundary. Here, the second task is applied to accurately delineating the boundary of the prostate, which is very unclear in CT images. Accordingly, the HFFCN uses a two-branch structure consisting of a shared encoding path and two complementary decoding paths. In contrast to the conventional multi-task networks, an information sharing (IS) module is also proposed to communicate at each level between the two decoding branches, by which the HFFCN endows the ability to learn hierarchically the complementary feature representations for different tasks, and also simultaneously preserve the specificities of learned feature representations for different tasks. The HFFCN is comprehensively evaluated on a large CT image dataset, including 313 images acquired from 313 patients. The experimental results demonstrate that the proposed HFFCN outperforms both the state-of-the-art segmentation methods and the conventional multi-task learning methods.

Key words: fully convolutional network; deep learning; multi-task learning; prostate segmentation

* 基金项目: 国家自然科学基金(61673203); 江苏省重点研发计划专项基金(BE2018610)

Foundation item: National Natural Science Foundation of China (61673203); Special Fund of Key Research and Development Program of Jiangsu Province (BE2018610)

收稿时间: 2019-03-07; 修改时间: 2019-09-28; 采用时间: 2019-12-07; jos 在线出版时间: 2020-04-07

前列腺癌是目前美国男性中第二致命的癌症类型^[1],治疗前列腺癌最常用的方法之一是外照射放射治疗(external beam radiation therapy,简称 EBRT).精确分割前列腺,是外照射放射治疗靶区计划阶段非常重要的一步,其目的是最大限度地提高送达肿瘤组织中的辐射剂量,同时避免损伤周边健康器官.然而,即使对于有经验的放射肿瘤医师,精确地手工描绘前列腺轮廓仍是一件具有挑战性并耗时的工作.因此,开发一种自动化方法(例如基于机器学习的方法^[2-4])进行前列腺的分割具有很大的临床价值.

近年来,全卷积网络(fully convolutional network,简称 FCN)^[5]及其衍生网络^[6-8]已被成功应用于医学图像分割.由于全卷积网络 FCN 能够提取基于任务的特征,以及像素级尺度(pixel-wise)上的语义信息及其周围局部细节,这类方法在分割任务中显示出了目前最高的性能.然而,将 FCN 直接应用于前列腺分割仍然是一项具有挑战性的工作,这是由于:(1) 在 CT 图像中,组织的对比度较低,前列腺器官边缘很难被区分出;(2) 在 CT 图像中,不同病人的前列腺器官的形状和外观差别很大.如图 1 所示,图片展示了 CT 图像的低软组织对比度(对比左列中的图像)和前列腺在病例间大的形貌尺寸变化(对比右列中的图像).

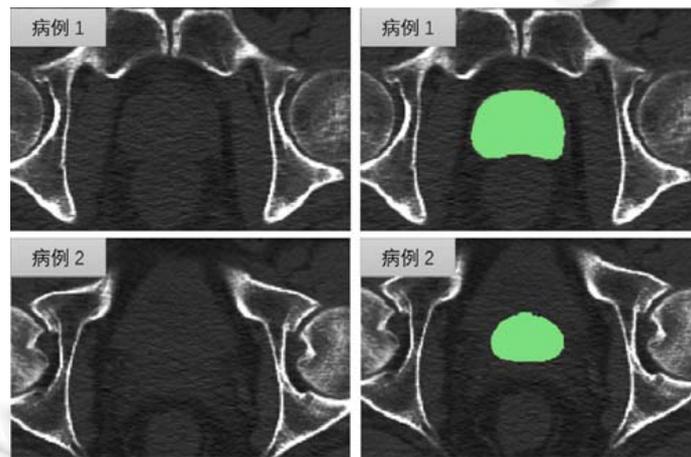


Fig.1 Prostate (left column) and its segmentation (right column) of two patients

图 1 两个典型病例的前列腺(左列)及其分割结果(右列)

为了应对这些挑战,一种可行的方法是在多任务学习框架中同时利用轮廓感知策略(contour-aware strategy)^[3,4,9],在分割过程中自动定义器官边缘.例如,Chen 等人^[9]提出了一种用于腺体分割的轮廓感知网络,其中除了主任务外,该网络将腺体轮廓的描绘(即轮廓与非轮廓体素的二元分类)纳入作为补充任务.然而,由于在 CT 图像中前列腺边缘不清晰,这种方法并不适用于前列腺分割.因此,将描绘前列腺边缘视作二元分类任务实际上并不可靠.此外,现有的大多数深度多任务网络^[9-11]的实现方式非常相似.也就是说,这些网络通常包括:(1) 一个具有共享权重的主干(backbone),以学习所有任务的共同特征;(2) 针对不同任务分支的任务特定顶层(如图 2 中左图所示).直观上,由于几乎所有网络参数在所有任务中被共享,因此这类网络结构过分地假设不同任务能够由相同的特征表示来处理.在现实应用中,这种过强的约束显然不够灵活,因为它不可避免地忽略了不同任务的特异性.

本文中,我们关注于多任务网络的性能表现.为了自动分割 CT 图像中的前列腺器官,我们提出了一个多任务全卷积网络,称为 HFFCN.具体来讲,我们将手动描绘的前列腺边缘和邻接体素作为有价值的任务相关信息,用来定位 CT 图像中不清楚的前列腺边缘.接下来,不同于将描绘器官边缘作为一个二分类问题^[9],我们将前列腺边缘形态表征的描绘(视为一个回归问题)作为前列腺分割(主任务)的补充任务.据此,我们提出的 HFFCN 由一个降采样主干(编码器)和两个用来推断主任务和补充任务热图(heatmap)的上采样分支(解码器)组成.如图 3 所示,我们网络的每个上采样分支由 3 个级联块组成,每个级联块由几个卷积层和一个变换大小的操作(转置卷积层)组成.因此,两个不同上采样分支中的同层块利用来自输入图像的共同感受野进行学习.为了鼓励两个分

支学习到的特征表示是互补的,同时不强制它们相同,两个不同采样分支中的同层块通过一个新的信息共享(information sharing,简称 IS)模块进行通信.通过在多层插入这些 IS 块,两个互补的子网络能够层级地分享面向任务的信息.显然,这里我们使用了一个更宽松的框架,这种情况下,在处理两个任务时,我们通过权衡卷积多任务网络和两个极其分散的网络的权重(如图 2 中的对比)来遵循基本的传输学习规则.因此,模型的灵活性也得以增强.

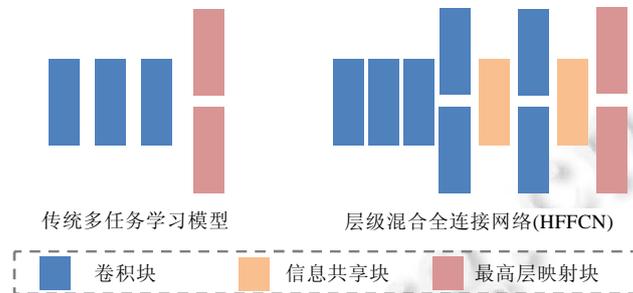


Fig.2 Comparison of the conventional multi-task network and the proposed hierarchically fused multi-task FCN

图 2 传统多任务网络和所提层级混合多任务 FCN 对比

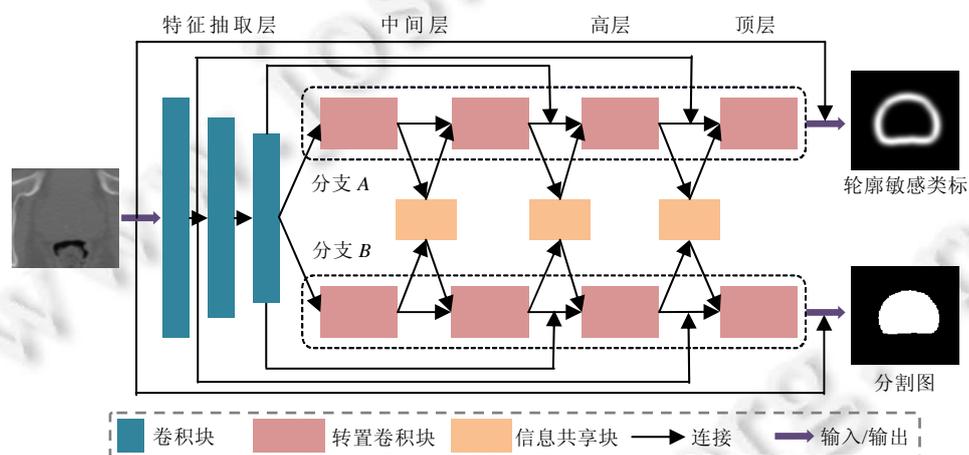


Fig.3 Architecture of HFFCN-3, which contains three IS blocks

图 3 HFFCN-3 的详细结构,其包含了 3 个 IS 块

本文的贡献主要有 3 点.

- (1) 我们通过将轮廓感知纳入网络来处理前列腺分割任务,并通过多任务学习来解决它.网络由一个考虑到前列腺边缘和上下文的,对特别设计的轮廓敏感类标的回归来引导.
- (2) 我们提出一种层级混合多任务结构,该结构能够更好地以层级方式学习两个任务的互补特征.具体来讲,我们提出以信息贡献块来实现两个任务层级特征的交互.
- (3) 我们在 313 位患者 CT 图像数据集上评估我们提出的方法.与当前用来分割前列腺的其他先进方法相比,我们的方法具有更高的性能.

本文第 1 节从前列腺 CT 图像分割和深度多任务网络两个方面对现有的相关工作做一些回顾.第 2 节详细说明本文所提的层级混合全卷积网络的结构以及学习方法.第 3 节对实验相关的数据集、实验方法、评价标准以及实验结果进行说明.最后总结全文,讨论本文工作的一些不足和未来可以进行的研究方向.

1 相关工作

我们将在两个方面回顾本文的相关工作:(1) 前列腺 CT 图像分割的相关方法;(2) 深度多任务网络的相关方法.

1.1 前列腺CT图像分割

有很多方法用于前列腺分割^[2-4],这些方法中,形变模型(deformable model)是一系列模拟某个特定类别对象的可变性的算法^[12].由于这些方法利用统计数据对某个特定类别对象的形状表征进行形变,因此它们通常能够在医学图像中很好地描述器官轮廓,尤其是在 CT 图像中.因此,最先进的基于 CT 图像的前列腺分割方法通常基于形变模型.例如,Gao 等人^[3]提出使用形变模型来对器官位移图(displacement map)进行回归.Shao 等人^[4]提出了一种用于前列腺和直肠分割的结合形变模型并基于 landmark 的轮廓检测方法.值得注意的是,这些先进的方法通常会设计专门的策略用来区分不清晰的前列腺边缘,并利用多任务学习来解决该任务.这体现出上述两种方法的重要性.但是,以往的方法大多是在手工描绘的特征上开发出的.近年来,由于神经网络(deep neural networks)的面向任务学习能力,神经网络被广泛应用于医学图像分析领域^[6,7].神经网络也在 CT 图像前列腺分割任务中取得了很好的性能表现.例如,Balagopal 等人^[13]提出了一种极联了 2D-UNet 和 3D-UNet 的分割方法分割前列腺、膀胱、直肠、左股骨头、右股骨头这 5 种盆腔器官.Shahedi 等人^[14]提出了一种半自动的前列腺分割模型,他们首先使用一些位置先验找到前列腺的区域,然后学习一个包含 86 个坐标点的前列腺外轮廓表示模型,用以提升前列腺的分割效果,取得了较好的性能.我们提出的网络结构也归类于 FCN,用来解决具有挑战性的 CT 图像分割任务.

1.2 深度多任务神经网络

在神经网络中,结合多任务学习的有效性已经在多种计算机视觉应用中得以证明,例如,面部 landmark 检测^[10]、语音合成^[11]和人体姿势估计^[15].并且,很多研究已经将这种范式应用于医学图像分析领域.例如,Chen 等人^[9]提出一种用于分割腺体的结合轮廓感知的多任务深度网络.Liu 等人^[16]提出一种通过多任务多通道网络能够同时预测脑部疾病和回归临床评分.它们的网络结构可以被概况为前文所述方式,即使用一个基础网络对所有目标域进行特征共享.但是由于没有保留任务特定特征,这种方法的效力具有局限性.为了解决这个问题,我们放宽了这个约束,来探索更灵活的网络结构.其核心假设是,通过分离不同任务的信号路径使得自域信息被保留,并通过插入网络间模块来实现学习到的特征被共享.

2 层级混合多任务深度网络

我们提出一种新的多任务学习网络 HFFCN 来分割 CT 图像中的前列腺器官.为此,我们设置了两个精确分割前列腺的任务:(1) 分割前列腺;(2) 描绘前列腺边缘.这里,引入第 2 个任务是为了向网络提供关键的形态指导,这样有助于区分前列腺边缘.我们提出的 HFFCN 结构如图 3 所示,其中,“3”表示模型中含有 3 个信息共享块.其中,分支 a 为轮廓感知子网络,分支 b 为分割子网络.如图所示,在两个分支的每一对同级模块(即使用尺寸变换操作,如转置卷积之前)的输出之后,网络会插入一个信息共享块用以学习任务间的共同特征,并在下一级模块之前将共享特征分别发送给两个分支的模块.

2.1 前列腺器官检测

基本上,相对于整个盆腔 CT 图像,前列腺位于一个较小的区域.因此,我们首先采用一种区域定位网络来取出完全包含整个前列腺区域的子图.与半自动化方法不同^[17],我们提出的框架能够通过一个具有 U-Net^[6]结构的网络来自动定位前列腺区域.而后,基于粗预测分割图,前列腺的中心可以被确定.我们根据前列腺中心取 $128 \times 128 \times 128$ 大小的区域,以确保完全覆盖前列腺器官.接下来,为了更准确并高效地分割前列腺器官,在随后的网络(HFFCN)中,我们使用从被识别的区域中截取的局部块,而不是原始 CT 图像.

2.2 轮廓感知子网络

由于在传统 FCN 中使用了降采样操作,位于前列腺边缘的体素的空间相关性被破坏,因此导致最终的分割结果不能够分出完整的前列腺轮廓.为了解决这个问题,一些研究提出可以在基于学习的方法中结合器官轮廓作为补充信息,该想法的有效性已被证明^[3,4,9].例如,Chen^[9]提出了一种以腺体轮廓作为指导的深度多任务网络,从而改进了分割效果.之前的方法通常将器官边缘的描绘表示为二元分类任务,它们简单地将位于器官边缘的体素视为正采样,其他位置的体素视为负采样.但是,前列腺器官的边缘是难以分辨出的,这使得以上方法不能很好地适用.位于前列腺边缘周围的体素实际上包含了轮廓体素丰富的上下文信息.此外,手工描绘的前列腺轮廓通常是不可靠的.因此,我们认为靠近器官边缘的体素也是任务相关的.在本工作中,我们将这个问题解决为回归问题,把靠近轮廓的像素点作为亚真值(sub-ground-truth).图 4(其中,颜色从深到浅预示类标的强度值从低到高)说明了两个任务之间的不同:(1) 传统的二元分类任务;(2) 提出的轮廓点和其周围像素点(称为轮廓敏感类标)的回归.

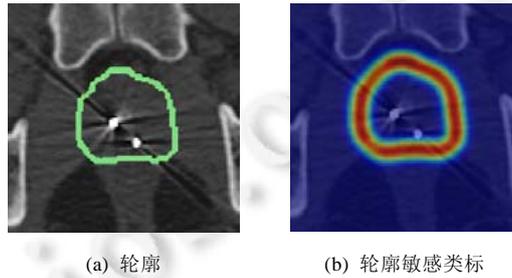


Fig.4 Comparison of contour and contour sensitive label of prostate
图 4 前列腺的轮廓和轮廓敏感类标对比

受到文献[18]中所提出想法的启发,我们将每个轮廓点以及其周围像素点表示为核(kernel)为 σ (此处 $\sigma=5$) 的高斯分布.我们将损失层的输入表示为 (X,y) ,这里, y 表示前列腺轮廓的真值.接着,关于输入 (X,y) 和权重 θ 的轮廓描绘的目标可表示为

$$\arg \min_{\theta} \sum_{(X,y)} \sum_{i,j} \|S(y_{ij}) - \phi(X, \theta)\|^2 \tag{1}$$

这里, $\phi(X, \theta)$ 是轮廓描绘子网络的输出. $S(y_{ij})$ 是代表位置 (i,j) 的高斯滤波 y 的集合,可以表示为

$$S(y_{ij}) = \sum_{k \in K, l \in L} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{((i,j)-(k,l))^2}{2\sigma^2}}.$$

这里, $(k,l) \in (K,L)$ 表示距离点 (i,j) 小于 σ 的邻域集.

2.3 信息共享块

两个子网络的信息被提出的 IS 块来结合以及学习,如图 5 所示.为了帮助确认某个块的输入信息来源,我们定义了两个连接:(1) 私有连接,即子网络之间的连接;(2) 公有连接,即子网络和 IS 块之间的连接.私有信息和公有信息也依次定义.为了方便计算,按照 Szegedy 等人^[19]最初提出的 inception 方法,我们使用连结操作(concatenation operation)来合并两种不同的信息来源.接下来是几个用于信息提取的卷积层.事实上,其他诸如求和等操作也可以应用于此.这些由 IS 块生成的特征被定义为公有信息,这些信息被送入每个子网络中.于是,接收此公有信息(将卷积核数量表示为 c_{public})的块将其与其之前块中的私有信息(将卷积核数量表示为 $c_{private}$)连结起来.为了更好地衡量公有信息和私有信息的影响,我们定义了二者之间的比率 r :

$$r = \frac{c_{public}}{c_{private} + c_{public}}.$$

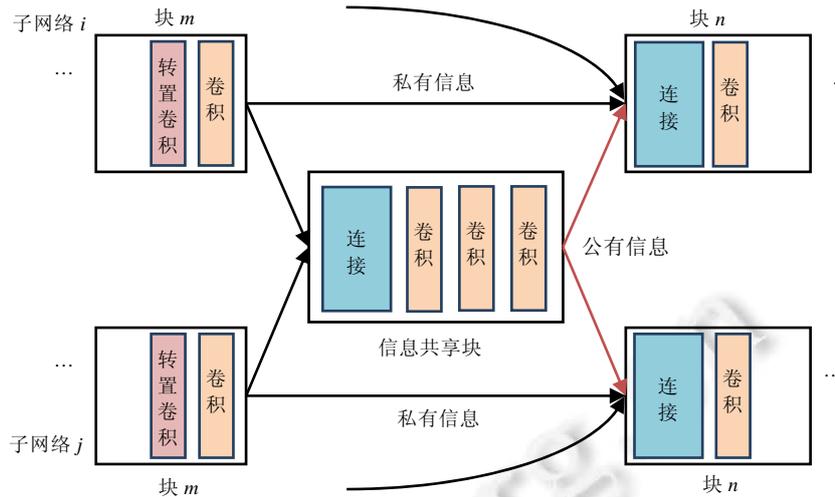


Fig.5 Illustration of information sharing block

图5 信息共享块示意

显然, r 值越大, 表示越多公有信息被送入子网络. 目前工作中, 连接操作的输入通道通常设置为等比率, 即 $r=0.5$, 或者任意数量. 本工作中, 我们将 r 作为一个超参数, 并通过调节比率 r 来探索自学习和信息共享的重要性. 实验表明, 我们方法的性能会受到 r 值的影响.

2.4 网络结构

所提网络将从前列腺区域取得到的图像局部块(image patch)作为输入. 首先, 输入的图像局部块被送入信息提取主干. 我们没有使用如 ResNet^[20]这样的预训练网络, 因为它们在我们的实验中没有取得比具有初始化训练的网络更好的效果. 然后, 提取的特征被送入到两个分支(图3中分支 a 和分支 b). 每个分支由4个块构成, 并由一个特定的任务目标监督. 为了进一步整合前一层的学習信息, 我们在每个分支中采用了 U-net^[6]中的跳过连接. 在每一层, 提出的 IS 块连结来自两个分支的相同层特征, 并在下一层之前将混合特征送回它们.

2.5 瓶颈层

显然, U-Net 结构和 IS 块利用了很多连结操作. 在连结操作后, 卷积核映射会增多, 这将给计算资源带来巨大的负担. 除此之外, 冗余的卷积核映射也增加了. 按照文献[19]中提出并已被其他工作^[20,21]成功证明的思路, 一个 1×1 大小步幅为1的卷积层, 称为瓶颈层, 能够更好地保留相同信息, 从而减少冗余的特征图. 本工作中, 我们采用此思路, 并在每个连结层后设置一个瓶颈层.

2.6 多任务学习

我们通过多任务学习来解决这两个任务. 这两个任务能够提供不同类型的信息提示来引导 HFFCN 的训练. 最终的目标函数如下:

$$L = \lambda_1 \mathcal{L}_{cls} + \lambda_2 \mathcal{L}_{reg} + R(\Theta).$$

这里, \mathcal{L}_{cls} 表示分割前列腺的分类任务的损失, \mathcal{L}_{reg} 表示描绘前列腺轮廓的回归任务的损失. 权重 $\Theta = \{\theta_{cls}, \theta_{reg}\}$ 由 $R(\cdot)$ 正则化, 这是深度神经网络的常用方法. 我们使用 Logsoftmax 损失来指导分割任务. 令 $\hat{p}_n = F(X, \Theta)$ 表示网络输出的类标可能性. 其中, F 代表网络的映射函数, Θ 表示整个网络的参数集合, X 表示网络的输入, l_n 表示真实类标, 因此损失 \mathcal{L}_{cls} 为

$$\mathcal{L}_{cls} = -\frac{1}{N} \sum_{n=1}^N \log(\hat{p}_n, l_n)$$

公式(1)中的损失被用于回归.回归任务通常会较大损失值,因此,我们用两个技巧来解决这个问题:首先,我们通过修复分割子网络和 IS 块中的所有参数,仅在冷启动阶段训练分支 a 的参数,这样能够更好地初始化不平衡的多任务学习的权重;其次,我们对两个损失设置了不平衡权重.本工作中,我们设置权重因子 $\lambda_1=1$, $\lambda_2=0.01$.

3 实验结果

3.1 数据集

本实验在一个由 313 位患者在放射治疗中的靶区规划阶段的 CT 图像组成的数据集上进行评估.该数据集包括 313 张图像,图像尺寸为 $512 \times 512 \times (61 \sim 508)$.这些图像的平面内分辨率为 $0.932 \text{mm} \sim 1.365 \text{mm}$,切片厚度为 $1 \text{mm} \sim 3 \text{mm}$.

分割该数据集非常具有挑战性,这是由于以下 3 个原因:(1) CT 图像中的组织对比度非常低,使得前列腺非常难以被区分出来;(2) 不同个体之间器官形状和表征的差异很大;(3) 该数据集是来自放射治疗中的靶区规划阶段,每一位病人仅有 1 张图像,因此数据集内部个体差异非常大.

我们在常用的开源框架 Caffe^[22]上实现了所提方法并进行了实验.实验在带有一块英伟达 1080Ti GPU 的工作站上完成.我们随机将数据集拆分为训练集(70%)验证集(10%)和测试集(20%).由于图像获取自不同制造商的不同型号机器,因此在图像预处理阶段,我们首先将所有图像采样到相同分辨率 $1:1:1 \text{ mm}^3$.对于每张图像,我们通过基于阈值的截取方法,保留 CT 图像中仅包含病人身体的部分,这样能够减少背景噪声.接下来,我们对每张图像做归一化处理,以此将强度值归一化到区间 $[-1, 1]$.我们通过使用滑动窗口策略来截取二维图像,将取得的图像输入到所提出的网络中.在训练阶段,我们随机截取前列腺区域的局部块.在测试阶段,我们密集地截取以前列腺为中心的重叠局部块.本工作中,局部块尺寸为 $64 \times 64 \times 3$,并且批尺寸设置为 128.这里,我们输入一个连续的切片续来预测中间切片的类标.这个策略结合了待预测切片的三维空间关系,这有助于提高预测性能.所提网络中,我们每一个卷积层之后使用限制性线性单元(ReLU)作为激活函数.为了更好地分析所提网络的效果,我们构建了不同结构的网络.我们构建了 3 种类型的 HFFCN,其中,IS 块分别从顶层插入到中间层.不同类型的网络均由一个 IS 块(称为 HFFCN-1)、两个 IS 块(称为 HFFCN-2)和 3 个 IS 块(称为 HFFCN-3)构成.这些网络通过随机梯度下降(SGD)算法进行优化.我们使用 6×10^5 次迭代来训练网络.学习速率从 0.01 降低到 0.000 1,步长为 2×10^5 .在测试阶段,通过训练好的网络,将基于局部块预测与最大投票算法结合起来生成最终的前列腺分割结果.

3.2 评价指标

我们利用两个常用的指标,即基于个体覆盖范围的指标(DSC)和平均表面距离指标(ASD)来衡量提出方法的分割性能.在实验中,我们比较平均 DSC 和平均 ASD,这两个指标定义为

$$DSC = \frac{2Vol_{gt} \cap Vol_{seg}}{Vol_{gt} + Vol_{seg}},$$

$$ASD = \frac{1}{2} \{ \text{mean} \min_{a \in Vol_{gt}, b \in Vol_{seg}} d(a, b) + \text{mean} \min_{a \in Vol_{seg}, b \in Vol_{gt}} d(a, b) \}.$$

其中, Vol_{seg} 表示预测个体的体素的集合, Vol_{gt} 表示真值个体的体集合, $d(a, b)$ 表示点 $\{a, b\}$ 间的欧几里得距离.

3.3 模型简化测试

(1) IS 块的效能

我们研究了具有不同 IS 块数配置的 HFFCN 的性能.为了进一步研究整个网络共享特征的效能,我们构建了一个配有 6 个 IS 块的网络(称为 HFFCN-6).为了实现该目的,我们将特征提取路径分成两个独立的路径,它们类似于推理路径,然后在其中每一层中插入 IS 块.表 1 列出了这 4 种竞争方法的 DSC 和 ASD 性能,每个 HFFCN 网络仅有最优值被包括在表中(全局的最优值已用粗体标出).为了公平地比较,所有这些网络由相同的超参数

训练.从表 1 可以看出,HFFCN-3 在这些配置中的性能最佳.这表明通过增加 IS 块的数量,能够提高分割性能.例如,HFFCN-3 的 ASD 为 0.907mm,性能明显高于 HFFCN-1(0.971mm).这验证了我们的假设,即层级共享策略有利于特征混合.值得一提的是,HFFCN-6 的表现没有 HFFCN-3 好.这说明了共享底层特征的重要性,这是由于图像底层特征大多为一些基础特征.该重要性也在其他现有工作中得到部分的证明^[23].

Table 1 Quantitative comparison of HFFCNs on DSC and ASD with different number of IS blocks in 313-patient CT image dataset

表 1 含有不同数量信息共享块的 HFFCN 在 313 个病人的 CT 图像数据集上基于 DSC 和 ASD 指标的定量比较

方法	DSC (%)	ASD (mm)
HFFCN-1	87.1	0.971
HFFCN-2	87.5	0.933
HFFCN-3	88.1	0.907
HFFCN-6	86.1	1.115

(2) 输入比率 r 的影响

我们还比较了在不同输入比率 r 下的 HFFCN 的性能.为了公平比较,我们在连结所有 HFFCN 之后设置了数目相等的特征通道(本节设为 80).在这组实验中,我们使用离散集 {0.1,0.3,0.5,0.7,0.9} 来改变输入比率 r 的值,并在图 6 中记录不同方法所得的平均 DSC 值.图 6 显示,在大多数情况下,3 种不同 HFFCN 所得的 DSC 值随着 r 值的增加先下降,后上升.此外我们发现,在 HFFCN-1 中, $r=0.9$ 时网络取得最优的性能;在 HFFCN-2 中,达到最优性能时 $r=0.7$;而在 HFFCN-3 中,达到最优值时的 r 值下降为 0.3.我们认为,这可能由于通过增加更多 IS 块,公有信息的比例增加,而每个子网络的私有信息比例下降导致的.因此,我们可以得出结论:私有信息和公有信息应当在最佳点平衡,以此达到最佳表现.在实验中,直观的感觉是,随着信息共享块的增加,为了保证网络达到最佳性能,应使用更小的 r 值,使得公有信息和私有信息保持在一个比例.另一方面,使用更多数量的 IS 块通常比使用较少数量的 IS 块的分割性能要好,这说明保留任务特定特征的重要性.值得指出的是,对于给定 HFFCN 网络结构如何自动获得最佳 r 仍然是一个未解决的问题,我们将在未来工作中对此进行研究.

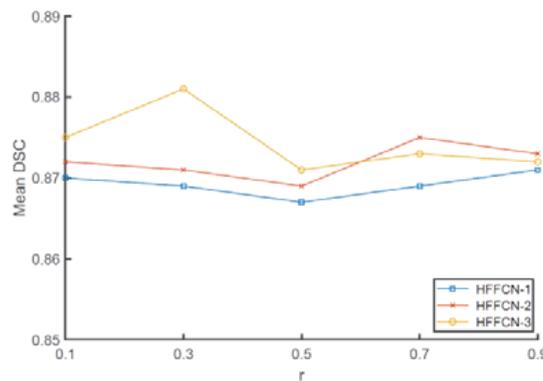


Fig.6 Comparison of performance in terms of DSC with respect to different HFFCN architectures

图 6 不同 HFFCN 结构在 DSC 指标上的性能比较

(3) 轮廓感知网络的效能

为了评估我们提出的轮廓感知方法的效能,我们将 HFFCN-1、HFFCN-2 和 HFFCN-3 与仅执行分割任务的传统的基于 U-Net 的网络进行比较.此外,我们将所提出网络与传统的多任务深度网络 OnePath 进行比较.在 OnePath 方法中,仅保留传统 U-Net 的一条主路径(具有相同的层和参数设置),接着是包含两个卷积层的顶部映射路径(每个路径对应于特定任务),实验结果见表 2.从表 2 可以看出,与 U-Net 的结果相比,将轮廓感知纳入网络(如我们在 HFFCN-1 和 HFFCN-3 中所做)能够提高分割性能.例如,就 ASD 而言,使用 HFFCN-3 和 U-Net 得到的

结果之间有较大差异($>1\text{mm}$),该差异清楚地表明在网络中考虑轮廓感知的效力.就 *DSC* 而言,与 *OnePath* 相比, *HFFCN-1* 有 1% 的改进.这说明与 *OnePath* 中使用的早期混合策略相比,在 *HFFCN-1* 中使用的后期混合策略能够取得更好的分割结果.此外, *HFFCN-3* 的总体性能最佳,因此表明了混合所有推理层的有效性.

Table 2 Evaluation of the effectiveness of contour-awareness

表 2 轮廓感知网络的性能评估

方法	<i>DSC</i>	<i>ASD</i> (mm)
U-Net	0.80	2.02
OnePath	0.86	1.01
HFFCN-1	0.87	0.97
HFFCN-3	0.88	0.91

3.4 特征可视化

为了研究 *HFFCN* 在不同层学习到的特征,我们将几个块的输出进行可视化,如图 7 所示.与在中高层共享参数的 *OnePath* 网络不同, *HFFCN* 有两个分支来处理两个特定任务.因此,我们将从 *HFFCN-3* 的两个子网络的第 4 块(中层)和第 5 块(高层)学习到的特征进行可视化.图 7 显示,在分割子网络的很大区域中出现了高激活的神经元,并且主要出现在轮廓描绘子网络的器官边缘上(我们将 3-D 的特征通过通道维平铺成 2-D 以进行展示,其中的颜色从深到浅表示神经元的活性从低到高).这表明我们的 *HFFCN* 方法能够在执行不同任务时学习针对任务的特征,这对于多任务学习问题尤其有用.

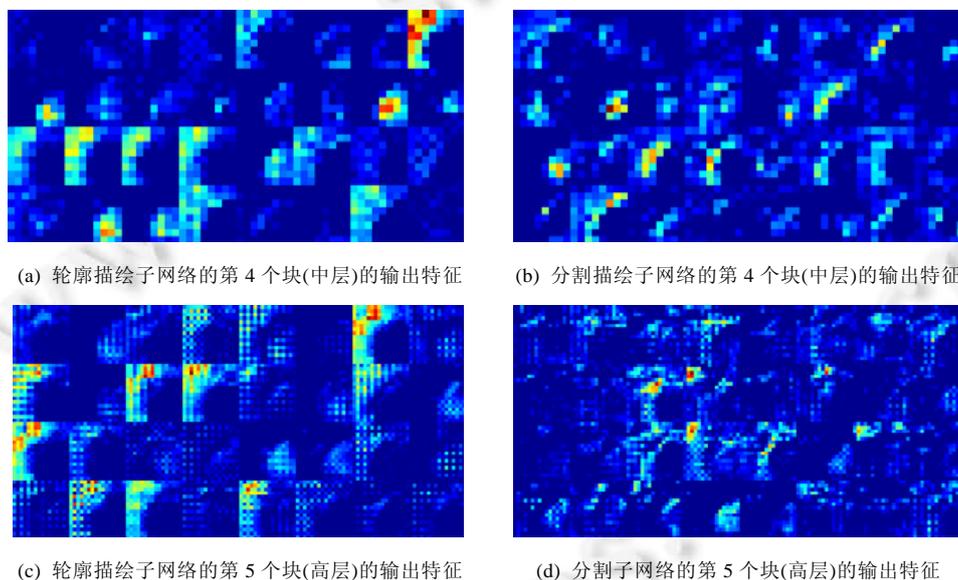


Fig.7 Visualization of mid- and high-level features in the proposed method

图 7 所提网络中层和高层特征的可视化

3.5 与其他先进算法比较

我们进一步将我们的 *HFFCN* 方法与其他先进的分割方法作对比^[3,4,24,25],它们大多数是基于可变形模型的方法.需要指出的是,一些竞争方法也利用边缘信息(例如使用回归线索^[4]或使用位移图^[3,25])进行分割,并且在过去取得了非常好的分割效果.其中, Martinez 等人^[24]提出了一种基于贝叶斯先验初始化的可变形模型方法. Shao 等人^[4]提出了一种使用基于回归森林方法生成的形状先验的可变形模型方法. Gao 等人^[3,25]提出了一种同时训练回归器和分类器的可变形模型方法.实验结果见表 3(最优值用粗体标出).通过表 3 可以观察到, *HFFCN-3* 方法在 *DSC* 和 *ASD* 方面始终取得最佳性能,这表明了所提出算法的效力.

Table 3 Quantitative comparison of *DSC* and *ASD* (unit: mm) for prostate segmentation on planning CT image of 313 patients

表 3 313 个病人的计划阶段 CT 图像上前列腺分割基于 *DSC* 和 *ASD*(单位:mm)指标的定量比较

方法	<i>DSC</i>	<i>ASD</i> (mm)
Martinez 等人 ^[24]	0.87	—
Shao 等人 ^[4]	0.88	1.86
Gao 等人 ^[25]	0.86	1.85
Gao 等人 ^[3]	0.87	1.77
HFFCN-3	0.88	0.91

3.6 可视化结果

我们给出了所提出的 HFFCN-3 以及 OnePath 网络的可视化分割结果.注意,HFFCN-3 和 OnePath 均隐舍地利用轮廓信息,因此它们均可在大多数 CT 图像中较好地描绘前列腺轮廓.因此,我们仅可视化了几个具有挑战性的样本的分割结果,如图 8 所示.深色的轮廓线表示 ground-truth,浅色的轮廓线表示所提方法的分割结果(基于 HFFCN-3),灰色的轮廓线表示 OnePath 网络的分割结果.

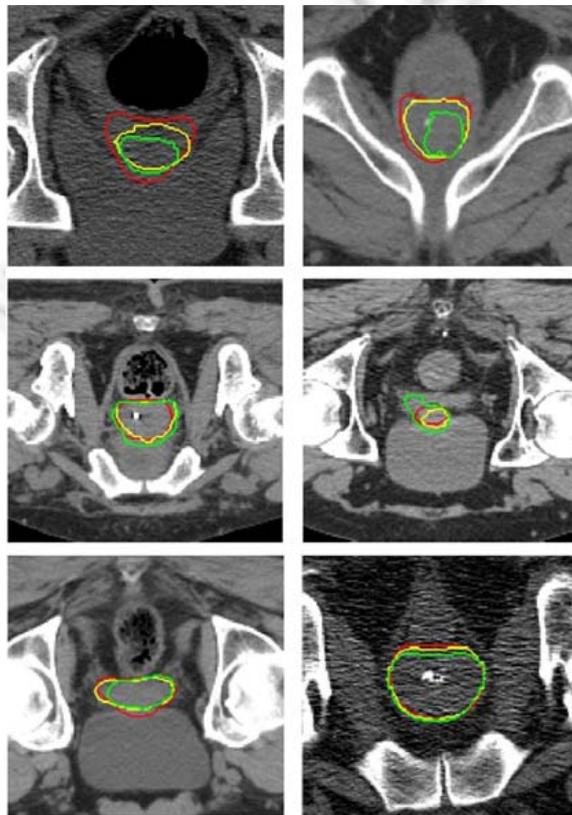


Fig.8 Visualization results of the proposed method on prostate segmentation

图 8 所提方法的前列腺分割可视化结果

通过图 8 可看出,对于这些具有挑战性的样本,我们的 HFFCN-3 方法始终取得比 OnePath 网络更好的分割结果.对于这些具有挑战性的样本,HFFCN-3 的分割结果与 ground-truth 重合度更高.与此同时,OnePath 的分割结果质量较低.这些结果显示,我们提出的 HFFCN 能够更好地在组织对比度很低的 CT 图像数据集完成分割任务.

4 本文总结

本文工作中,我们提出了一个针对前列腺 CT 图像分割的层级混合多任务全卷积网络(HFFCN),以更好地学习多任务学习中任务各自的特定信息.具体来讲,我们提出的富表示性的轮廓敏感类标能够使得网络更好地捕捉前列腺的边缘.同时,我们提出的信息共享模块能够使得跨任务特征在每层得以共享.除此之外,不同于先前网络的全共享策略,我们探索了共享信息与任务特定信息的比率,以此提升多任务网络的性能.我们在包含 313 位患者的 CT 图像的数据集上对 HFFCN 进行评估,评估结果说明该方法具有卓越的性能,与其他高性能方法相比,我们提出的 HFFCN 能够取得更好的分割效果.另外,我们的框架能够被简单地应用于其他轮廓敏感的分割任务上.我们提出的网络是一种通用方法,能够推广到其他处理多任务的应用.在未来工作中,我们将研究为不同的子网络设计不同的结构,以处理不同的任务.

References:

- [1] Cancer.org. 2018. <https://www.cancer.org/cancer/prostate-cancer/about/key-statistics.html>
- [2] Guo YR, Gao YZ, Shen DG. Deformable MR prostate segmentation via deep feature learning and sparse patch matching. *IEEE Trans. on Medical Imaging*, 2016,35(4):1077–1089.
- [3] Gao YZ, Shao YQ, Lian J, Wang AZ, Chen RC, Shen DG. Accurate segmentation of CT male pelvic organs via regression-based deformable models and multi-task random forests. *IEEE Trans. on Medical Imaging*, 2016,35(6):1532–1543.
- [4] Shao YQ, Gao YZ, Wang Q, Yang X, Shen DG. Locally-constrained boundary regression for segmentation of prostate and rectum in the planning CT images. *Medical Image Analysis*, 2015,26(1):345–356.
- [5] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. 2015. 3431–3440.
- [6] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *Proc. of the Int'l Conf. on Medical Image Computing and Computer-assisted Intervention*. 2015. 234–241.
- [7] Nie D, Wang L, Gao YZ, Shen DG. Fully convolutional networks for multi-modality isointense infant brain image segmentation. In: *Proc. of the IEEE 13th Int'l Symp. on Biomedical Imaging (ISBI)*. 2016. 1342–1345.
- [8] Milletari F, Navab N, Ahmadi SA. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: *Proc. of the Fourth Int'l Conf. on 3D Vision (3DV)*. 2016. 565–571.
- [9] Chen H, Qi XJ, Yu LQ, Dou Q, Qin J, Heng PA. Dcan: Deep contour-aware networks for object instance segmentation from histology images. *Medical Image Analysis*, 2017,(36):135–146.
- [10] Zhang ZP, Luo P, Loy CC, Tang XO. Facial landmark detection by deep multi-task learning. In: *Proc. of the European Conf. on Computer Vision*. 2014. 94–108.
- [11] Wu ZZ, Valentini-Botinhao C, Watts O, King S. Deep neural networks employing multi-task learning and stacked bottleneck features for speech synthesis. In: *Proc. of the IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)*. 2015. 4460–4464.
- [12] McInerney T, Terzopoulos D. Deformable models in medical image analysis: A survey. *Medical Image Analysis*, 1996,1(2): 91–108.
- [13] Balagopal A, Kazemifar S, Nguyen D, *et al.* Fully automated organ segmentation in male pelvic CT images. *Physics in Medicine & Biology*, 2018,63(24). [doi: 10.1088/1361-6560/aaf11c]
- [14] Shahedi M, Ma L, Halicek M, *et al.* A semiautomatic algorithm for three-dimensional segmentation of the prostate on CT images using shape and local texture characteristics. In: *Proc. of the Medical Imaging 2018: Image-guided Procedures, Robotic Interventions, and Modeling*. Int'l Society for Optics and Photonics, 2018.
- [15] Li SJ, Liu ZQ, Chan AB. Heterogeneous multi-task learning for human pose estimation with deep convolutional neural network. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops*. 2014. 482–489.
- [16] Liu MX, Zhang J, Adeli E, Shen DG. Deep multi-task multi-channel learning for joint classification and regression of brain status. In: *Proc. of the Int'l Conf. on Medical Image Computing and Computer-assisted Intervention*. 2017. 3–11.

- [17] Shi YH, Gao YZ, Liao S, Zhang DQ, Gao Y, Shen DG. Semi-automatic segmentation of prostate in ct images via coupled feature representation and spatial-constrained transductive lasso. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2015,37(11): 2286–2303.
- [18] Pfister T, Charles J, Zisserman A. Flowing convnets for human pose estimation in videos. In: *Proc. of the IEEE Int'l Conf. on Computer Vision (ICCV)*. 2015.
- [19] Szegedy C, Liu W, Jia YQ, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2015.
- [20] He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. 2015. 770–778.
- [21] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2017.
- [22] Jia YQ, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T. Caffe: Convolutional architecture for fast feature embedding. In: *Proc. of the 22nd ACM Int'l Conf. on Multimedia*. 2014. 675–678.
- [23] Karpathy A, Toderici G, Shetty S, Leung T, Sukthankar R, Li FF. Large-scale video classification with convolutional neural networks. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2014.
- [24] Martínez F, Romero E, Dréan G, Simon A, Haigron P, De Crevoisier R, Acosta O. Segmentation of pelvic structures for planning CT using a geometrical shape model tuned by a multi-scale edge detector. *Physics in Medicine and Biology*, 2014,59(6):1471–1484.
- [25] Gao YZ, Lian J, Shen DG. Joint learning of image regressor and classifier for deformable segmentation of ct pelvic organs. In: *Proc. of the Int'l Conf. on Medical Image Computing and Computer-assisted Intervention*. 2015. 114–122.



何克磊(1989—),男,安徽合肥人,博士,助理研究员,主要研究领域为医学图像分析,机器学习.



高阳(1972—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为强化学习,大数据,云计算.



史颖欢(1984—),男,博士,副教授,CCF 专业会员,主要研究领域为机器学习,机器视觉,医疗图像分析.