

一种改进 FCN 的肝脏肿瘤 CT 图像分割方法

段 杰^{1,2}, 崔志明^{1,2}, 沈 艺^{1,2}, 冯 威¹, 吴宏杰¹, 冯雨晴^{1,2}

(1. 苏州科技大学电子与信息工程学院, 江苏 苏州 215009;
2. 苏州市虚拟现实智能交互及应用技术重点实验室, 江苏 苏州 215009)

摘 要: 精准的医学图像分割是辅助疾病诊断和手术规划的必要步骤。由于腹部器官边界模糊、对比度不高, 肝脏肿瘤的自动分割一直是一个难题。针对传统全卷积神经网络(FCN)实现端到端分割精度不佳等问题, 提出了一种卷积型多尺度融合 FCN 的 CT 图像肝脏肿瘤分割方法。首先, 通过提高对比度、增强和去噪的方式对原始的 CT 图像数据集进行预处理; 然后使用处理后的数据集对所设计好的 FCN 网络进行训练; 最终得出能够精确分割肝脏肿瘤的网络模型。实验效果采用多种评价指标进行分割结果的评估, 并且与多种常见的分割网络进行对比。实验结果表明本文方法可以精准分割 CT 图像中各种形状和大小的肝脏肿瘤, 分割效果良好, 能够为临床的诊断提供可靠的依据。

关 键 词: 全卷积神经网络; 肝脏肿瘤分割; 深度学习; 图像分割; 卷积多尺度融合
中图分类号: TP 391 **DOI:** 10.11996/JGJ.2095-302X.2020010100
文献标识码: A **文章编号:** 2095-302X(2020)01-0100-08

A CT image segmentation method for liver tumor by an improved FCN

DUAN Jie^{1,2}, CUI Zhi-ming^{1,2}, SHEN Yi^{1,2}, FENG Wei¹, WU Hong-jie¹, FENG Yu-qing^{1,2}

(1. College of Electronics and Information Engineering, Suzhou University of Science and Technology, Suzhou Jiangsu 215009, China;
2. Suzhou Key Laboratory of Virtual Reality Intelligent Interaction and Application Technology, Suzhou Jiangsu 215009, China)

Abstract: Accurate medical image segmentation is a necessary step in assisting disease diagnosis and surgical planning. The automatic segmentation of liver tumors has always been a difficult problem due to the blurred borders and the low contrast of abdominal organs. Aiming at the problem that the traditional full convolutional network (FCN) achieves low accuracy in end-to-end segmentation, this paper proposes a CT image liver tumor segmentation method based on convolutional multi-scale fusion FCN. Firstly, the original CT image dataset is preprocessed by improving the contrast, enhancement and denoising. Secondly, the designed FCN network is trained using the processed dataset. Finally, a network model capable of accurately segmenting the liver tumor is obtained. The experiment adopts a variety of evaluation indicators to evaluate the effectiveness of segmentation results and makes comparison with a variety of common segmentation networks. The experimental results show that the method proposed in this paper can accurately segment liver tumors of various shapes and sizes in CT images, and the segmentation effect is good, which can provide a reliable support for clinical diagnosis.

收稿日期: 2019-07-26; 定稿日期: 2019-09-20

基金项目: 国家自然科学基金项目(61876217); 苏州市科技发展计划项目(SYG201707)

第一作者: 段 杰(1994-), 男, 江苏泰州人, 硕士研究生。主要研究方向为计算机视觉、数据挖掘、图像分割及目标检测。

E-mail: 983861019@qq.com

通信作者: 崔志明(1961-), 男, 江苏苏州人, 教授, 博士, 博士生导师。主要研究方向为机器学习、智能信息处理、计算机视觉及数据挖掘。

E-mail: zmcui@mail.usts.edu.cn

Keywords: full convolutional network; liver tumor segmentation; deep learning; image segmentation; convolution multi-scale fusion

肝脏肿瘤是世界上最常见的肿瘤之一, 其死亡率位居全球第二, 是导致人类死亡的重要因素。随着肿瘤放射技术的飞速发展, 放射治疗进入了以图像引导和自适应放疗为代表的精准放疗阶段^[1-2]。精准放疗需要准确勾画出放疗的靶区(肿瘤), 从而引导治疗和后续的放射计划。但现阶段, 在临床医学中精准的靶区勾画需要由有经验的医师手工操作完成, 其准确性和效率完全依赖医师的临床经验。这项工作不仅耗时, 而且重现性很差。

因此, 基于计算机视觉及人工智能技术实现肝脏肿瘤自动精准分割便成为了解决上述问题的关键。常见的肝脏肿瘤自动分割方法可分为图像和统计模型 2 类。单纯基于图像分割的方法是指运用图像灰度、梯度或纹理等可以直接从图像中获取信息进行分割, 主要方法包括聚类^[3]、阈值、插值^[4]、区域生长^[5-6]、水平集^[7]、图割^[8-10]和活动轮廓模型^[11-13]等。

近年来, 随着计算性能的大幅提高, 人工智能和神经网络迅猛发展, 基于各种神经网络图像分割技术也应用到医学图像处理领域。卷积神经网络(convolutional neural networks, CNN)是计算机视觉和图像处理方面的一个非常重要的模型, 在图像分类^[14-16]、人脸识别^[17-18]、目标检测^[19-21]、动作识别^[22-23]等方面具有广泛应用, 并且取得了巨大的成功。在实际应用中 CNN 表现出优秀的特征提取能力, 其中, 全卷积神经网络(fully convolutional networks, FCN)作为 CNN 的改进网络在图像分割领域得到广泛应用^[24]。与图像分类不同, 语义分割需要确定每个像素的类别以实现精确的分割。FCN 是将 CNN 最后的全连接层换成反卷积层, 从而实现像素到像素级别的分类。FCN 及其衍生物在图像分割上的应用不断拓展, 其编码器与 VGG-16 中的 13 个卷积层相同, 解码器将编码器提取的特征映射到与输入分辨率相同的编码器从大到小提取特征, 解码器从小到大逐步将提取的特征放大到输入图像大小。但传统的 FCN 网络边缘分割差、精度不高, 达不到医学图像分割的要求。

针对上述问题, 本文设计了一种卷积型多尺度融合改进的 FCN, 并将其应用于肝脏肿瘤 CT 图像的分割。

1 全卷积神经网络

FCN 是 SHELHAMER 等^[25]在现用于分类的 CNN 结构基础上改造而成的, 即将 CNN 原有的全连接层改成卷积层。FCN 是由卷积和反卷积组成, 通过卷积和反卷积的过程, 基于端到端学习, 完成了对图像每个像素的分类, 从而实现对整个输入图片的分割。

卷积层(convolution layers)在 FCN 中主要承担着特征提取的功能。卷积输出的图像大小为 $\frac{W-F+2p}{S}+1$, 其中, W 为输入图片的维度; F 为卷积核的维度; S 为步长; p 为 padding 填充的像素数。卷积产生的特征图通过激活函数来增加神经网络模型的非线性。卷积和激活函数运算为

$$x_j^l = f \left(\sum_{i \in M_j} x_i^{l-1} + b_j^l \right) \quad (1)$$

其中, 括号内的部分为卷积层提取特征的主要运算; l 为网络的当前层; x_i^{l-1} 为感受野中第 i 个像素点的值; w_{ij}^l 为 x_j^l 的权重; b_j^l 为偏置项; $f()$ 为激活函数。

反卷积是一种卷积操作, 其是池化的逆过程。在 FCN 中, 池化操作缩小了输入图片的尺寸, 然而在图像分割过程中需要对每一个像素进行分类, 并且最终得到与输入图片维度相同的分割图像, 因此需要将生成的热度图(heat map)恢复成原图像的维度。反卷积通过反向训练, 能够达到输出重构输入的效果, 使输出图像恢复到与输入图像相同大小的维度。

FCN 通过端到端的网络结构来实现图像的语义分割, 端到端的网络可以缩减人工预处理和后续处理, 尽可能使模型从原始输入到最终输出, 网络自己学习特征, 提取特征也融入到算法当中。网络模型可以根据数据自动调节, 从而增加模型的整体契合度, 并且端到端网络学习中的各项成本都比非端到端的网络结构要低。FCN 的网络结构如图 1 所示。

2 一种卷积型多尺度融合 FCN 肝脏肿瘤图像分割方法

2.1 卷积多尺度融合

为了使输出与输入具有相同的维度, FCN 将最

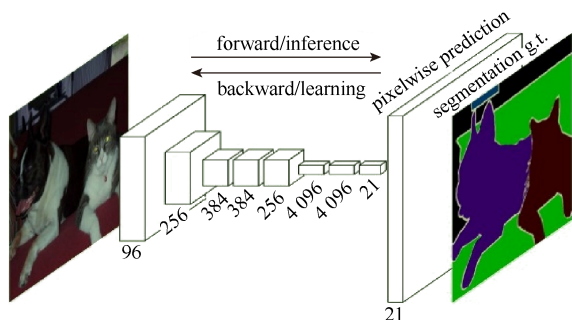


图1 FCN 的网络结构图

后一层卷积层输出的 feature map 进行 32 倍的上采样(upsampling), 上采样输出的是输入图片的分割结果。文献[25]提出的 FCN 中具有 5 个 2×2 的池化层, 每层池化后图像的分辨率均会降低二分之一, 因此最后一层卷积层输出的 feature map 分辨率比较低, 会丢失一些细节特征。直接对最后输出的 feature map 进行 32 倍上采样, 得到的图像分割精度不高, 分割边缘较粗糙, 达不到理想的分割效果, 所以文献[25]提出了多尺度特征融合的方法,

其具体做法是将第 3 层和第 4 层池化层特征进行融合后进行 8 倍上采样, 图像分割精度有所提升。

本文在文献[25]的多尺度特征融合 FCN 的基础上对上采样(反卷积)过程做了一定的改进。图 2 为本文提出的卷积型多尺度融合结构和训练过程示意图。如图 2 训练过程所示, 本文将训练过程分为前向传播和反向传播 2 部分。前向传播主要通过卷积完成图片信息的提取, 池化完成特征压缩和降维操作, 通过反卷积将维度还原到原图大小, 前向传播是权重和偏置学习的过程。通过反向传播进行权重和偏置的调节, 实现参数优化的功能。传统方法直接使用 2 个池化层 pool4 和 pool3 的 feature map 进行上采样, 但是, pool3 和 pool4 保留的特征信息粗糙、边缘分割的精度不高。图 2 中, 本文在 pool3 和 pool4 池化层各增加一次卷积操作, 然后再进行融合操作。增加的卷积过程可以从 2 个池化层中提取到更深层的局部信息进行周围的信息融合, 增强模型的局部细节判断能力, 能够优化图像分割的细

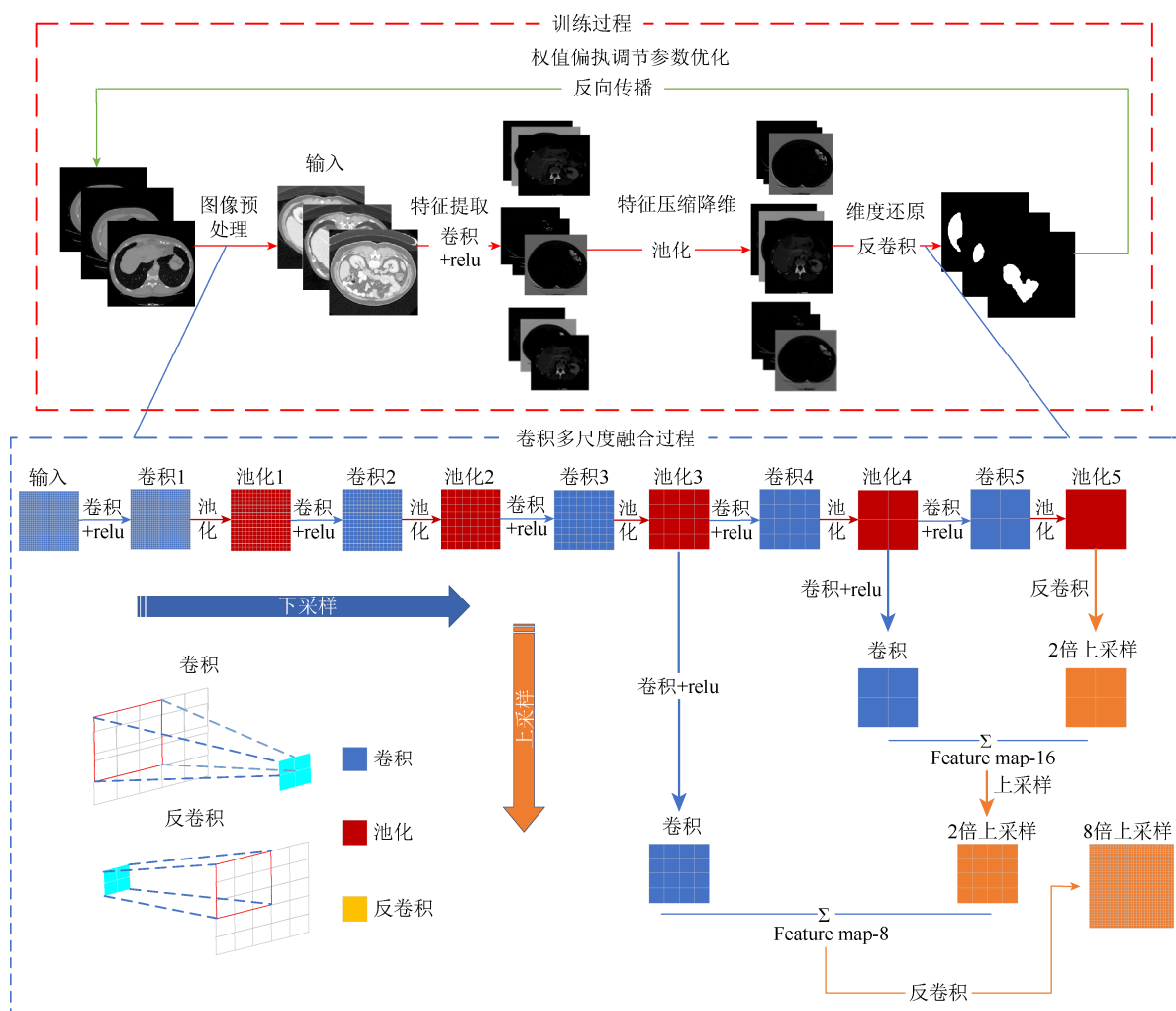


图2 训练和卷积多尺度融合示意图

节效果。采用多尺度特征融合的方法能够大大提升分割精度,使目标轮廓更加精细,从而提高分割效果。该模型共有 5 个池化层,经过每个池化层之后,输出图像依次缩小了 2, 4, 8, 16, 32 倍,每次池化之前都有若干个卷积层,最后一层输出的 feature map 被称为热度图,本模型在传统的 FCN 的基础上对第 3 和第 4 池化层多加一次卷积操作。本模型首先将热度图进行一次 2 倍上采样,同时将第 4 池化层(pool4)进行一次卷积,将其卷积到与图片分类数相同的通道数(本文中为 3),随后将两者进行相加得到 feature map-16。并将其进行一次 2 倍上采样,同时对第 3 层池化(pool3)进行一次卷积,同样将其卷积到与图片分类数相同的通道数(本文中为 3),将两者进行相加得到 feature map-8。最后将其进行一次 8 倍上采样得到最后输出的分割图。

2.2 FCN 网络结构

本文使用的网络结构包括 5 个卷积层组、5 个池化层、3 个卷积层和 3 个反卷积层。其中,第 1, 2 卷积层组每层包含 2 个卷积过程,第 3, 4, 5 卷积层组每层包含 3 个卷积过程,前 5 个卷积层组中每个卷积层组后会跟一个池化层。在卷积层组后还有 3 个卷积层和 3 个反卷积层(上采样)。每次卷积后都会使用 relu(线性整流函数)作为激活函数,该函数可以使所有输出值均为正数,同时增加网络的稀疏性,且能够防止过拟合的情况。在反卷积过程中也采用同样的函数,以保证输出为正数,防止过拟合的发生,relu 函数的数学表达式为

$$f(x) = \max(0, x) \quad (2)$$

本文所提方法中前 5 个卷积层组均采用 3×3 的卷积核,该卷积核既有良好的接受场,同时参数的数量也少,能减少计算量。第 6 层卷积采用 8×8 的卷积核,第 7 层卷积选择 2×2 的卷积核,第 8 层卷积选择 1×1 的卷积核。所有卷积过程中的步长均为 1,以便更精细地提取特征,尽量不丢失任何边缘信息。在初始化卷积核方法中,本文选择“Xavier”^[22],通过从零均值和特定方差的分布绘制网络中的权重来实现网络初始化。该方法权值分布均匀且使网络收敛速度快,权值能达到优值,并且能够保持每一层梯度大小相似,防止梯度消失或梯度爆炸。对于池化层,一般的平均池化能保留更多的图像背景信息,最大池化能保留更多纹理信息,本文选择最大池化。网络的详细结构见表 1。

表 1 网络详细结构

网络层	卷积核	通道数	步长	输出维度
卷积层组 1	3×3	64	1	512×512×64
池化 1	2×2		2	256×256×64
卷积层组 2	3×3	128	1	256×256×128
池化 2	2×2		2	128×128×128
卷积层组 3	3×3	256	1	128×128×256
池化 3	2×2		2	64×64×256
卷积层组 4	3×3	512	1	64×64×512
池化 4	2×2		2	32×32×512
卷积层组 5	3×3	512	1	32×32×512
池化 5	2×2		2	16×16×512
卷积 6	8×8	4 096	1	16×16×4096
卷积 7	2×2	4 096	1	16×16×4096
卷积 8	1×1	3	1	16×16×3
反卷积 1	4×4	3	2	32×32×3
反卷积 2	4×4	3	2	64×64×3
反卷积 3	16×16	3	8	512×512×3

2.3 图像预处理方式的改进

原始的医学图像像素分布广泛,肝脏与周围器官的边缘分布不是很明显,对比度比较低。若直接将图像归一化到 0~255,则会丢失大量的重要信息。传统的 CT 图像预处理方式为像素填充值(pixel padding value)和像素值转换的方法,本文采用 contrast, enhancement 和 denoise 的路线进行数据的预处理,此方法在保留 CT 图像像素值的同时可增强图像的对比度,提高图像的亮度,降低噪声干扰。

首先,本文采用像素值线性变换对原始数据进行亮度和对比度的调节。随后统计图像中每个灰度级出现的次数,计算累积归一化直方图,重新计算像素点的像素值。最后通过拉伸像素强度分布范围,重新分配图像像素值,使一定灰度范围的像素的数量大致相等,达到对原始数据进行直方图均衡化的效果,从而提高图像质量,增强图像对比度。直方图均衡化是一个可逆的过程,且相对于其他处理方式更能够保留图片的原始特征。随后将直方图均衡化的图像使用全变差(图像梯度的范数的积分)的方式去噪声,一幅图像的噪声或干扰越多,全变差的值越高,为使全变差最小化,需去掉噪声同时,保留边缘等主要细节。如图 3 所示,由于 CT 图像的特殊性,原图的直方图与处理后的图像具有很大区别,可见处理后的图片具有更高的亮度和鲜明的对比度。

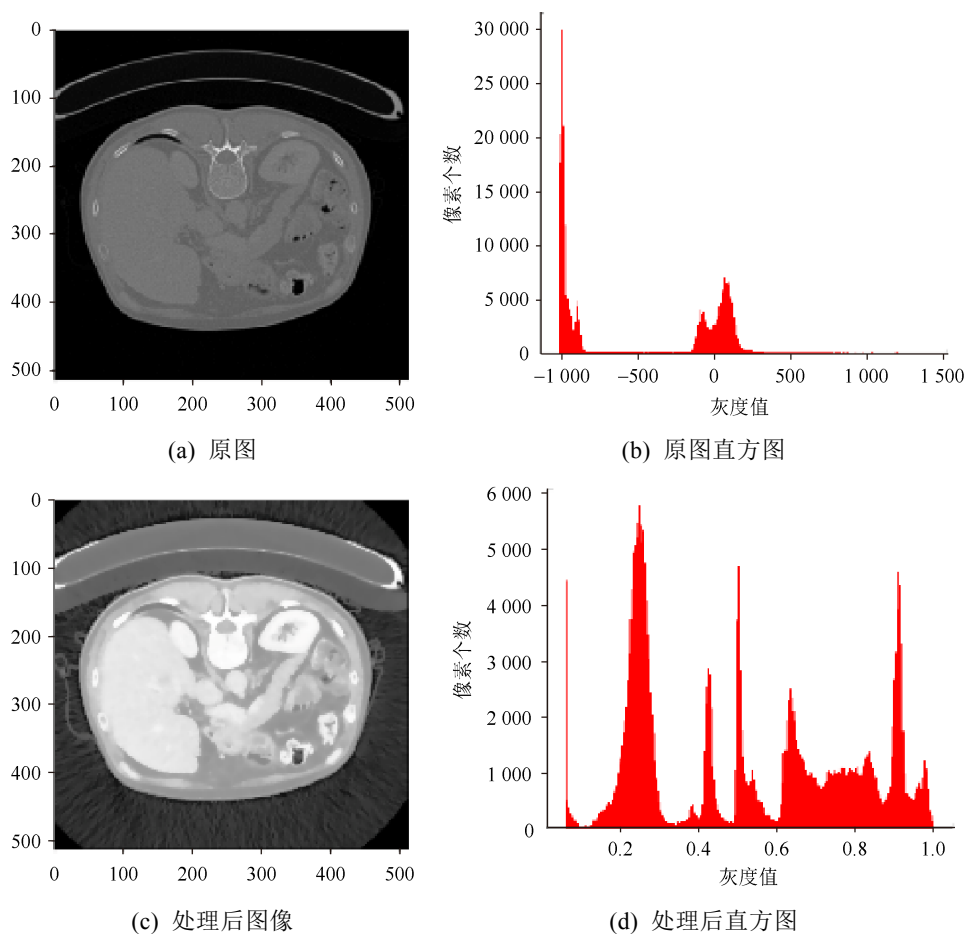


图3 图像处理前后及其直方图对比

由于医学图像的特殊性,一些医学图像的切片可能与靶器官的边界相切,相切的部分的面积很小,对比度较低,对分割精度影响很大。对于这类问题,普通的损失函数,收敛速度慢,学习效果差,不能达到预期的分割效果。所以本文采用 Dice 作为损失函数,其表达式为 $\frac{2(a \cdot b)}{|a|^2 + |b|^2}$,

可以衡量网络分割出来的图像与标签图像每一个像素的相似度,通过训练网络,可以让神经网络学习到更好的参数。本文将网络训练出的预测图片先做个标准和归一化,提高背景体和前景体的平衡度,尽量减少用于训练的损失,提高收敛的效率。由于处理后的图像的像素值不是 0 就是 1,且平方后的值不变,所以将损失函数表达式改进为 $\frac{2a \cdot b}{|a| + |b|}$,可减少不必要的计算,其中 a 为网络预测出的图像; b 为标签图像。通过训练损失函数,可以使神经网络学习到更好的参数与模型,减小预测图片和标签图片的差异,从而得到更好的分割效果。

3 实验

3.1 实验设备及数据集

实验硬件环境为 Intel(R) Xeon(R) E5-1603, 16 GB 内存,使用型号为 GTX1080ti 的 GPU 加速图像处理,卷积神经网络在 tensorflow 开源框架下实现, tensorflow 版本为 1.70。实验数据集由 3 454 张 training 图片和 1 068 张 test 图片构成, Validation 图片在 training 集中按照 7:3 的比例随机挑选作为验证集,即交叉验证。图片的维度是 512×512,为了防止训练和测试过程中过拟合情况,本文将所有数据打乱,再按照训练数据与测试数据 7:3 的比例随机分配,不仅预防过拟合的出现,也能提高实验的随机性。本文使用的数据集来自 codalab (<https://competitions.codalab.org/competitions/17094>),是肝脏分割竞赛使用的数据集,其包含 131 个序列, codalab 包含肝脏 CT 图像的三期 ground truth,格式为 enhanced CT,分辨率为 512×512。

3.2 训练参数设置

本文实验中选择使用指数衰减学习率,即根据训练下降的速度自行调节变化,防止产生振荡。指

数衰减学习率的方程为

$$lr = lr_0 \times \text{gamma}^{\left(\frac{\text{globalstep}}{\text{decaysteps}}\right)} \quad (3)$$

其中, lr 为当前学习率; lr_0 为初始学习率; gamma 为学习率衰减系数(一般在 0~1 之间); globalstep 为当前迭代次数; decaysteps 为衰减速度。本文设置的初始学习率为 $3e-5$, 学习率衰减速度为 0.90, 当前迭代次数为 0, 衰减速度为 1 727。batch_size 设置为 2, 训练的总轮数为 50 000 轮。选择的优化器为 Adamoptimizer, 此优化方法相对于传统的梯度下降法(gradient descent), 具有更快的收敛速度, 且能快速跳出局部最优点, 寻找到全局最优点。

3.3 实验结果分析

为了验证本文方法的有效性, 将本文方法与传统的 FCN, U-net 和 Segnet 网络进行了比较, 使用相同的数据及参数设置进行训练。对训练好的模型, 使用测试数据集进行验证。采用 5 种流行的评估指标即 Dice 相似函数(Dice similarity coefficient, DSC), 体积重叠误差(volumetric overlap error, VOE), 相对体积差异(relative volume difference, RVD), 平均对称距离(average surface distance, ASD) 和均方根差(root mean square error, RMSD)。

DSC 表示分割结果(S)与标签(T)之间的交叉区域的比值, 其表达式为

$$DSC(S, T) = \frac{2|S \cap T|}{|S| + |T|} \quad (4)$$

VOE 与 DSC 类似, 其将乘法换成了减法, 以此代表错误率, 即

$$VOE(S, T) = 1 - \frac{|S \cap T|}{|S \cup T|} \quad (5)$$

RVD 表示分割结果(S)与标签(T)之间的体积差异, 其表达式为

$$RVD(S, T) = \left(\frac{|S|}{|T|} - 1 \right) \times 100\% \quad (6)$$

ASD 表示分割结果(S)与标签(T)的平均表面距离, 其表达式为

$$ASD(S, T) = \frac{1}{|B_S| + |B_T|} \times \left(\sum_{x \in B_S} d(x, B_T) + \sum_{x \in B_T} d(x, B_S) \right) \quad (7)$$

其中, B_S 和 B_T 分别为分割结果和标签肝脏区域的轮廓; $d(x, B_T)$ 为任意像素 x 与 B_S 之间的最小距离, $d(x, B_T) = \min_{b_T \in B_T} \|x - b_T\|$, $\|\cdot\|$ 为欧几里得距离。

RMSD 表示分割结果(S)与标签(T)之间数值之间差异的度量, 其表达式为

$$RMSD(S, T) = \sqrt{\frac{1}{|B_S| + |B_T|} \times \left(\sum_{x \in B_S} d^2(x, B_T) + \sum_{x \in B_T} d^2(x, B_S) \right)} \quad (8)$$

表 2 为 5 种评价指标的定量分析。

从表 2 可以看出, 本文 FCN 方法, 在 5 个评价指标上除 DICE 指标外, 其他均优于传统的 FCN 网络, 且本文方法具有很大提升。与 Segnet 相比, 本文方法在 DICE 和 VOE 2 个评价指标上不及 Segnet, 但相差不大, 且其他 3 个指标均优于 Segnet。与 U-net (专业医学图像分割网络)相比, 本文方法在 RVD 和 RMSD 2 个评价指标上优于 U-net, 与其他评价指标均较接近。从表 2 可知, 虽然本文提出的方法具有优质的分割效果, 有很多评价指标优于其他网络, 且具有更好的稳定性, 但是从 ASD 和 RMSD 等评价指标来看, 分割结果与肝脏区域的标签有很大距离, 原因在于网络中实验的结果没有进一步细化, 输出边界相对粗糙。

图 4 在第 600 轮时, 本文方法与 U-net 都出现分割轮廓, 相对于传统 FCN 具有明显优势。在第 18 000 轮时, 所有网络均能显示出分割成果, 且本文方法与 U-net 最接近标签图片。在第 40 000 轮时, 所有网络均已成功分割, 传统 FCN 和 Segnet 的分

表 2 评价指标定量分析

方法	DICE (%)	VOE (%)	RVD (%)	ASD (mm)	RMSD (mm)
传统 FCN	80.03±3.04	24.30±5.04	18.24±1.25	14.64±1.32	19.39±6.16
U-net	90.82±2.05	14.28±4.07	21.22±2.56	6.79±2.12	13.62±6.18
Segnet	89.46±1.35	18.31±3.53	52.39±5.07	9.61±1.47	16.97±6.20
本文 FCN	88.46±1.03	18.36±4.33	18.12±1.222	7.64±0.71	12.52±3.45

割效果稍差。综上所述,本文方法具有优秀的分割精度的同时,还具有好的收敛性,分割效果优于传统 FCN 和 Segnet,这与数据分析出的结果是一

致的。由于实验使用的数据集来自不同病人的各种 CT 图片,且具有不同的病理特征,所以本文提出的方法具有很强的鲁棒性。

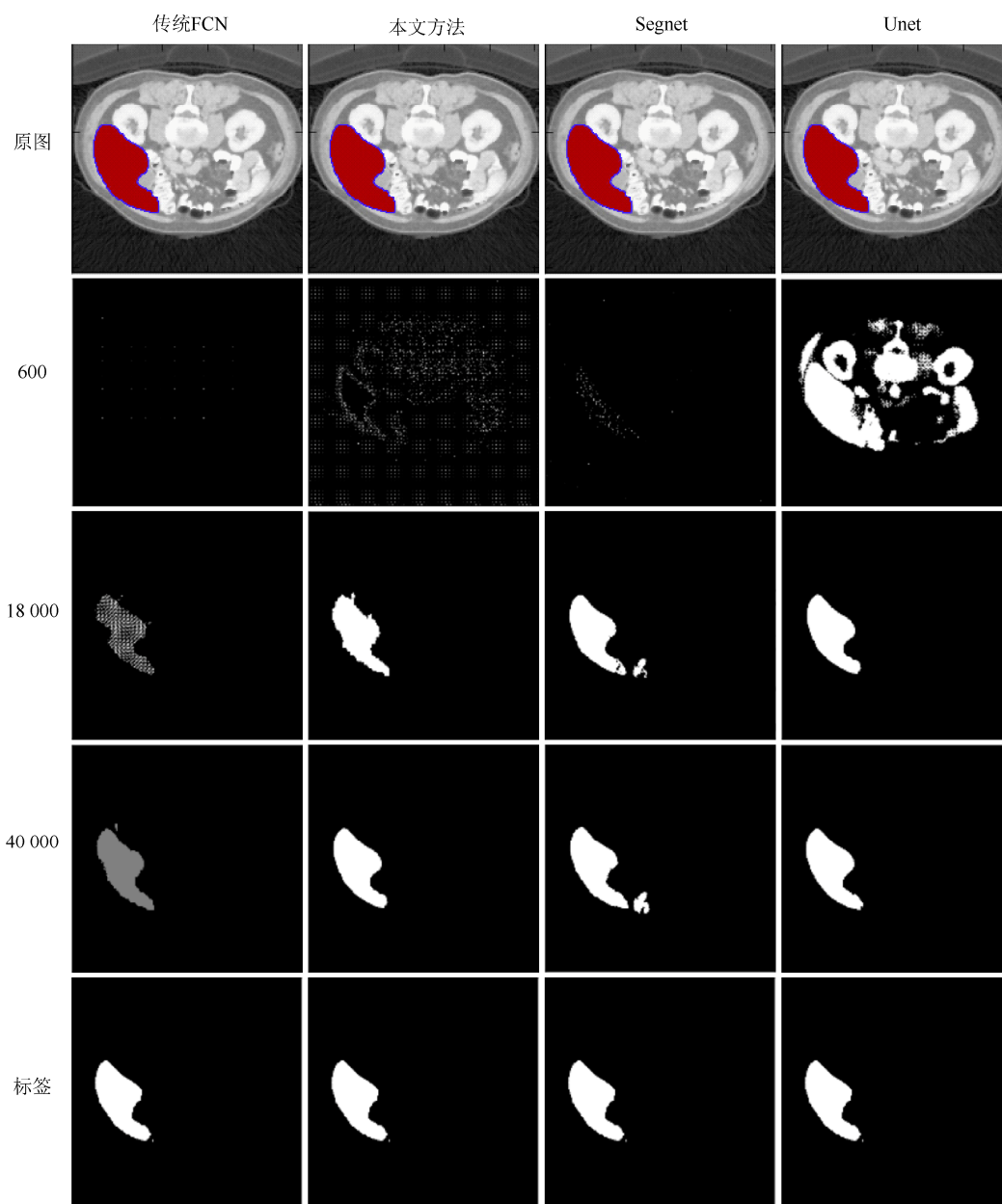


图4 分割效果图

4 结 束 语

本文提出了一种基于卷积型多尺度融合 FCN 图像分割方法,抛弃传统 FCN 只保留 pool3 和 pool4 的 feature map 进行上采样的方式,对 pool3 和 pool4 各增加一次卷积,并进行特征信息的提取,然后再进行上采样过程恢复到原始图片的维度。实验结果显示,本文提出的方法能够精确地从一堆对比度比较低的器官中分割出器官和病灶,提升了图像分割的精度。

参 考 文 献

- [1] 张建光, 黄术林, 赵继国, 等. 多发性肝癌质子放射治疗剂量学研究[J]. 中华肿瘤防治杂志, 2017, 24(16): 1162-1165.
- [2] 孟鸿宇, 江林宫, 张火俊. PET-CT 在非小细胞肺癌放射治疗中的应用进展[J]. 中国医学装备, 2015, 12(11): 67-70.
- [3] SELVER M A, KOCAOĞLU A, DEMIR G K, et al. Patient oriented and robust automatic liver segmentation for pre-evaluation of liver transplantation[J]. Computers in Biology and Medicine, 2008, 38(7): 765-784.

- [4] 吴健, 崔志明, 叶峰, 等. 基于轮廓形状的 CT 断层图像插值[J]. 计算机应用与软件, 2008, 25(11): 63-65.
- [5] LU X Q, WU J S, REN X Y, et al. The study and application of the improved region growing algorithm for liver segmentation[J]. Optik, 2014, 125(9): 2142-2147.
- [6] WEI J, LI P, YANG J, et al. A new automatic ship detection method using Σ -band polarimetric SAR imagery[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 7(4): 1383-1393.
- [7] ZHAO Y Q, YAN G X, XU X W, et al. Automatic segmentation of livers from CT series based on level set method with prior knowledge[J]. Journal of Central South University: Science and Technology, 2015, 46(4): 1310-1317.
- [8] AFIFI A, NAKAGUCHI T. Liver segmentation approach using graph cuts and iteratively estimated shape and intensity constraints[J]. Medical Image Computing and Computer-Assisted Intervention, 2012: 15(2): 395-403.
- [9] CHEN X, UDUPA J K, BAGCI U, et al. Medical image segmentation by combining graph cuts and oriented active appearance models[J]. IEEE Transactions on Image Processing, 2012, 21(4): 2035-2046.
- [10] LIAO M, ZHAO Y Q, LIU X Y, et al. Automatic liver segmentation from abdominal CT volumes using graph cuts and border marching[J]. Computer Methods and Programs in Biomedicine, 2017, 143: 1-12.
- [11] 吉宏伟, 何江萍, 杨新. 基于层次上下文活动轮廓的三维 CT 肝脏图像分割[J]. 生物医学工程学杂志, 2014, 31(2): 405-412.
- [12] PENG J L, WANG Y, KONG D X. Liver segmentation with constrained convex variational model[J]. Pattern Recognition Letters, 2014, 43: 81-88.
- [13] HAN M, LIU J M, MENG J Y, et al. Local energy information combined with improved signed distance regularization term for image target segmentation algorithm[J]. Journal of Electronics and Information Technology, 2015, 37(9): 2047-2054.
- [14] HE X T, PENG Y X. Fine-grained image classification via combining vision and language[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE Press, 2017: 5994-6002.
- [15] YE C, WU J, SHENG V S, et al. Multi-label active learning with label correlation for image classification[C]//2015 IEEE International Conference on Image Processing (ICIP). New York: IEEE Press, 2015: 3437-3441.
- [16] CAI S J, ZUO W, ZHANG L. Higher-order integration of hierarchical convolutional activations for fine-grained visual categorization[C]//2017 IEEE International Conference on Computer Vision (ICCV). New York: IEEE Press, 2017: 511-520.
- [17] WANG H, WANG Y T, ZHOU Z, et al. CosFace: large margin cosine loss for deep face recognition[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2018: 5265-5274.
- [18] DENG J K, GUO J, XUE N N, et al. ArcFace: Additive angular margin loss for deep face recognition[C]//2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE Press, 2019: 4690-4699.
- [19] SINGH B, DAVIS L S. An analysis of scale invariance in object detection – SNIP[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2018: 3578-3587.
- [20] SINGH B, LI H D, SHARMA A, et al. R-FCN-3000 at 30fps: decoupling detection and classification[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2018: 1081-1090.
- [21] CAI Z W, VASCONCELOS N. Cascade R-CNN: delving into high quality object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2018: 6154-6162.
- [22] CAO Z, SIMON T, WEI S-E, et al. Realtime multi-person 2D pose estimation using part affinity fields[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE Press, 2017: 7291-7299.
- [23] MA C Y, KADAV A, MELVIN I, et al. Attend and interact: higher-order object interactions for video understanding[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2018: 6790-6800.
- [24] MOSTAJABI M, YADOLLAHPOUR P, SHAKHNAROVICH G. Feedforward semantic segmentation with zoom-out features[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE Press, 2015: 3376-3385.
- [25] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.