

基于全卷积网络的肝脏CT语义分割

徐婷宜, 朱家明, 李祥健

(扬州大学信息工程学院, 江苏 扬州 225000)

✉634687849@qq.com; qingzhu@yzu.edu.cn; 1049091648@qq.com



摘要: 肝脏CT(计算机断层扫描)图像分割为临床肝脏医疗分析提供了可靠依据。文中探索了完全卷积网络(FCN)用于肝脏CT图像中的检测和分割。FCN已被证明是用于语义分段的非常强大的工具, 它能接受任意大小的输入并通过有效地推理与学习产生相应大小的输出。该文将分类网络VGG调整为完全卷积网络, 并通过迁移学习将其转移到分割任务, 展示了由端到端, 像素到像素训练的卷积网络语义分割。此架构能将来自深层粗糙层的语义信息与来自浅层精细层的外观信息相结合, 以生成准确而精细的分割。本架构肝脏分割IOU值达到0.9, 取得较好的分割效果。

关键词: 全卷积网络; 语义分割; 深度学习

中图分类号: TP391 **文献标识码:** A

Semantic Segmentation of Liver CT based on Full Convolutional Network

XU Tingyi, ZHU Jiaming, LI Xiangjian

(College of Information Engineering, Yangzhou University, Yangzhou 225000, China)

✉634687849@qq.com; qingzhu@yzu.edu.cn; 1049091648@qq.com

Abstract: Computed Tomography (CT) image segmentation provides a reliable basis for clinical liver medical analysis. A Full Convolutional Network (FCN) is explored for detection and segmentation in liver CT images. FCN has been proven to be a very powerful tool for semantic segmentation. It can accept inputs of any size and generate corresponding output through effective reasoning and learning. This paper adjusts the classification network of Visual Geometry Group (VGG) to a fully convolutional network and transfers it to the segmentation task through transfer learning. It shows an end-to-end, pixel-to-pixel trained convolutional network semantic segmentation. This architecture can combine semantic information from deep rough layers with appearance information from shallow fine layers to generate accurate and fine segmentation. In this architecture, the liver segmentation Intersection-over-Union (IOU) value reaches 0.9, and a good segmentation effect is achieved.

Keywords: full convolutional network; semantic segmentation; deep learning

1 引言(Introduction)

计算机断层扫描(CT)是观察骨关节及软组织的一种理想的方式, 肝脏病变检测常用CT图像观察^[1]。肝脏手术需要有关肝脏大小、形状和精确位置的信息。临床诊断大多数依靠医生耗费大量时间手动检测和分割, 这就突出了计算机分析的必要性。本文中所述图像语义分割的意思是计算机把图像中所有的像素点分配给其对应的标签。采用VGG、Alex-Net等CNN网络, 以图像中每个像素点为中心提取像素补丁, 将

补丁送入分类后对应的标签中, 每个补丁中心像素得到相对应的标签, 将所有像素点执行操作, 即可得对应像素点的标签^[2]。全卷积网络用卷积层取代全连接层, 使得网络能进行像素级分割的分类, 从而解决语义分割问题^[3]。FCN可以接受任意大小的输入, 并通过有效地推理与学习产生相应的输出, 使这个体系结构的损失函数在整个图像分割结果上进行计算^[4]。网络处理的是整个图像, 因此可以通过图像的分辨率更有效地进行缩放。相比于传统的CNN图像分割, FCN分割图

像还能减少不必要的重复存储和计算卷积,使得训练更加高效。

2 肝脏分割算法构建(Construction of liver segmentation algorithm)

2.1 数据采集

IRCAD(Research Institute Against Digestive Cancer)汇集了消化道癌症研究实验室,计算机科学/机器人研究部门和微创科培训中心。本文采用IRCAD的Liver segmentation-3D-ircadb-01数据库,它由对75%的肝肿瘤的10位女性和10位男性进行3D CT扫描组成。数据库提供了一系列有关图像的信息,例如根据库尼诺(Couninaud)分割的肝脏大小(宽度、深度、高度)或肿瘤的位置。这也表明与邻近器官的接触,肝脏的非典型形状或密度,甚至图像中的伪像等问题都会成为肝脏分割的主要困难。二十组肝脏3D图像如图1所示。

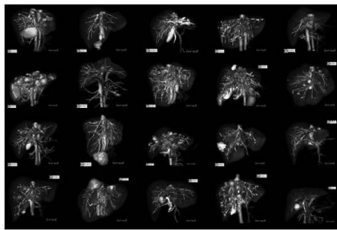


图1 肝脏3D图

Fig.1 Liver 3D picture

2.2 全卷积网络的构建

全卷积网络采用卷积神经网络实现了从图像像素到像素类别的变换。全卷积网络通过转置卷积(transposed convolution)层将中间层特征图的高和宽变换回输入图像的尺寸,从而令预测结果与输入图像在空间维(高和宽)上一一对应:给定空间维上的位置,通道维的输出即该位置对应像素的类别预测^[5]。全卷积像素预测如图2所示。

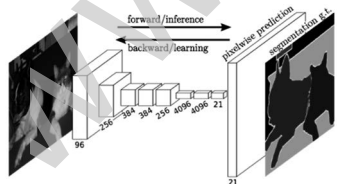


图2 FCN像素预测分割图

Fig.2 FCN pixel prediction segmentation map

本文的前置基础网络为VGG16,通过丢弃最终的分类器层并将全连接层转换为卷积层来构建全卷积网络。架构中附加一个通道维数为2的 1×1 卷积来预测每个粗略输出位置的肝脏评分,然后再附加一个反卷积层来将粗略输出提升到像素密集输出。上采样在网络中进行,通过像素级损失的反向传播进行端到端的学习。本文初始网络为FCN-8s的DAG网络,它学会了将粗糙的、高层的信息与精细的、低层的信息

结合起来^[6]。我们还探讨了添加另一个较低级别的链接层来创建FCN-4s DAG网络的附加值。这是通过与图3中的Pool3和Pool4层的链接相似的方式链接Pool2层来完成的。最初的网络架构如图3所示。

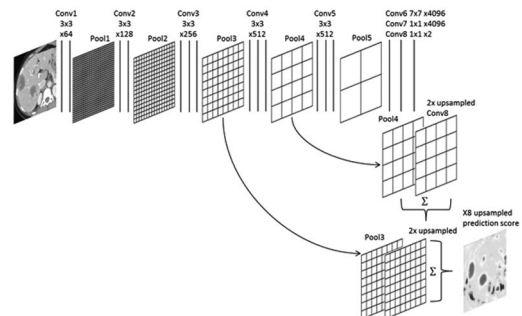


图3 网络架构图(未显示Relu激活函数)

Fig.3 Network architecture diagram (relu activation function not shown)

全卷积网络工作流程:

- (1)首先对输入的原图像实行卷积操作conv1和池化操作pool1,则原图像缩小到原来的1/2;
- (2)将步骤(1)的输出结果作为输入信息,对图像进行第二次卷积操作conv2和池化操作pool2,则图像进一步缩小到原图的1/4;
- (3)将步骤(2)的输出结果作为输入信息,对图像进行第三次卷积操作conv3和池化操作pool3,则图像进一步缩小到原图的1/8,此外,保留pool3过程中产生的feature map;
- (4)将步骤(3)的输出结果作为输入信息,对图像进行第四次卷积操作conv4和池化操作pool4,则图像进一步缩小到原图的1/16,此外,保留pool4过程中产生的feature map;
- (5)将步骤(4)的输出结果作为输入信息,对图像进行第三次卷积操作conv5和池化操作pool5,则图像进一步缩小到原图的1/32;
- (6)将步骤(5)的输出结构作为输入信息,利用conv6和conv7构建的全连接层进行进一步卷积操作,输出图像的大小依然是原图的1/32,此时得到的feature map记为heat map;
- (7)先将步骤(3)、步骤(4)中的feature map,以及步骤(6)中的heat map进行上采样操作,得到的图像记为X;
- (8)利用conv4的卷积核对X进行反卷积操作来进一步补充图像分割细节部分,得到的图像记为Y;
- (9)最后利用conv3中的卷积核对Y进行第二次的反卷积操作,得到图像Z,图像Z即为图像语义分割的结果^[7]。

2.3 数据增强

手动分割遮罩对于数据集来说只在2D之内,肝脏分割的数据集相对较小,因此数据增强是最适合的。当只有很少的

训练样本可用时,数据增强对于教导网络期望的不变性和鲁棒性是必不可少的。数据增强的方式有多种选择,例如,调整亮度、对比饱和色调等因素来降低模型对色彩的敏感度。本文通过将比例转换应用于可用的训练图像来达到数据增强的目的。

3 实验(Experiment)

本次训练数据集中的图像格式是DICOM格式,因此在将数据转换为TFRecord格式以供以后在TensorFlow中进行训练之前,我们必须先做一些预处理工作。在本项目范围内,我们只划分了肝、骨、肾等四类。因此,每个预处理的数据样本将是图像-形状(512, 512)的输入图像和遮罩二维数组具有与图像相同的空间形状,指示每个像素属于哪个类。训练中随机抽取数据集对数据进行训练,激活函数为relu函数,使用Adam优化方法和交叉熵损失函数对全卷积函数进行训练。二维训练样本2258张,验证样本565张,在进入网络前进行归一化处理,减去图像均值。在FCN预训练模型的基础上进行迭代,实验平台为Linux平台下的TensorFlow框架。使用批量大小为32的完整训练大概需要48小时,同时使用基本的数据增强。

3.1 Adam优化算法

Adam优化算法来源于适应性矩估计,同时具有AdaGrad(适应性梯度算法)和RMSprop(均方根传播)的优点。Adam的主要参数有学习率 α 、一阶矩估计的指数衰减率 β_1 和二阶矩估计的衰减率 β_2 。 α 控制了权重的更新比率,在迭代优化的前期,学习率较大,则前进的步长较长,这时便能以较快的速度进行梯度下降;而在迭代的后期,逐步减小学习率的值,减小步长,这样有助于算法的收敛,获得最优解。 β_1 用于计算导数的平均值, β_2 计算平方版指数加权平均数, ϵ 是固定值用来防止分母为0,本文参数设置为 $\alpha=0.0003$, $\beta_1=0.9$, $\beta_2=0.99$, $\epsilon=10e^{-8}$ 。Adam参数具有很好的解释性,通常无须调整或仅需很少的微调。

3.2 交叉熵损失

交叉熵损失的计算分为两个部分:softmax分类器与交叉熵损失。

Softmax分类器将网络最后的输出 y 通过指数转变为概率公式,如公式(1)所示。

$$\text{softmax}(y)_i = \frac{e^{y_i}}{\sum_{i=1}^n e^{y_i}} \quad (1)$$

e^{y_i} 用于计算类别 i 的网络输出类别,分母为输出指数和。

交叉熵损失是用来评估当前训练得到的概率分布与真实分布的差异情况。在深度学习中, $p(x)$ 表示真实分布, $q(x)$ 表

示预测分布,在实际实验中,交叉熵值越小,说明预测分布与真实分布越接近^[8]。交叉熵公式如式(2)。

$$H(p, q) = -\sum p(x_i) \log(q(x_i)) \quad (2)$$

3.3 交并比

IOU(Intersection-Over-Union)即交并比,是进行目标检测的一个重要算法,它具有非负性、不可同一性、对称性和满足三角不等式等优点。

IOU表示了产生的候选框(candidate bound)与原标记框(ground truth bound)的交叠率或者说重叠度。它衡量了两个边界框重叠的相对大小,一般约定0.5是阈值,IOU越高,边界框越精确。IOU的公式如式(3)。

$$IOU = \frac{(A \cap B)}{(A \cup B)} \quad (3)$$

4 实验结果分析(Analysis of results)

训练过程的可视化包括训练集和验证集的acc和loss曲线,根据曲线的不同特点进行超参数调节,可以不断优化网络。用tensorboard观察训练(train)和验证(val)的损失(loss)在训练时的变化如图所示。Train/loss不断下降, val/loss不断下降,说明网络训练正常, val/iou_loss不断上升,说明分割的精度不断提高。本文分割的IOU值为0.9。

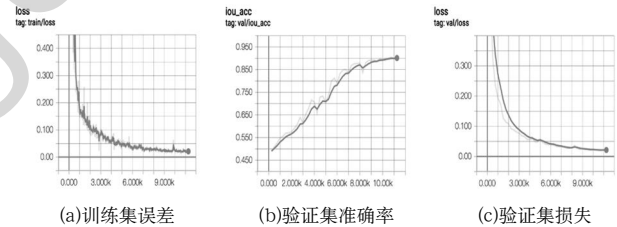


图4 曲线变化图

Fig.4 Curve change graph

肝脏CT图,手动真实分割图,本文分割图如图5所示。

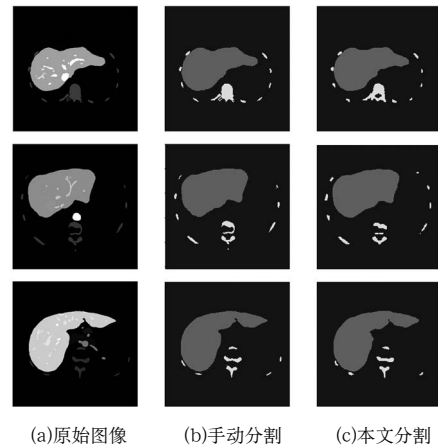


图5 肝脏分割结果

Fig.5 Liver segmentation results

(下转第16页)