

深圳大学

本科毕业论文（设计）

题目： 基于 FCN 的腹部器官 MRI 医学影像图像分割

姓名： 王家辉

专业： 计算机科学与技术

学院： 计算机与软件学院

学号： 2015150073

指导老师： 吴惠思

职称： 副教授

2019 年 04 月 16 日

深圳大学本科毕业论文（设计）诚信声明

本人郑重声明：所呈交的毕业论文（设计），题目《基于 FCN 的腹部器官 MRI 医学影像图像分割》是本人在指导教师的指导下，独立进行研究工作所取得的成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式注明。除此之外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。本人完全意识到本声明的法律结果。

毕业论文（设计）作者签名：

日期： 年 月 日

目录

摘要（关键词）	1
1 绪论.....	2
1.1 背景介绍.....	2
1.2 本课题的主要工作.....	5
2 相关研究.....	6
2.1 传统腹部器官分割方法.....	6
2.2 MRI 图像分割.....	7
2.3 多尺度方法.....	7
2.4 大核卷积.....	8
2.5 本章小结.....	9
3 基于 FCN 的腹部器官 MRI 图像分割网络方法.....	10
3.1 数据预处理.....	10
3.2 RSTN 网络.....	10
3.3 GCN++网络.....	11
3.4 UGS-Net 网络.....	13
3.5 损失函数.....	16
3.6 本章小结.....	16
4 实验设计与结果分析.....	17
4.1 实验数据和实验平台.....	17
4.2 实验设计.....	17
4.3 实验结果.....	18
4.4 实验分析.....	24

4.5 讨论.....	26
4.6 本章小结.....	28
5 总结与展望.....	29
5.1 工作总结.....	29
5.2 工作展望.....	30
参考文献.....	31
致谢.....	35
Abstract (Key words)	36

基于 FCN 的腹部器官 MRI 医学影像图像分割

计算机与软件学院（计算机科学与技术） 王家辉

学号：2015150073

【摘要】全卷积神经网络（FCN）已被广泛应用于计算机视觉和医学图像分析领域。其中，腹部 MRI 图像分割是一项基本的分割任务。尽管全卷积神经网络的应用广泛，但大多数方法只能同时处理一个器官的分割，将多目标分类应用在器官分割的研究较少，而临床实践中对分割速度的提升抱有很高期望。该研究主要有四大难点：多器官、多模态、解剖差异大、样本数量少。基于上述问题，本文提出了一种基于全卷积神经网络的注意力机制器官分割模型 UGS-Net，用于训练端到端的腹部 MRI 多模态影像。该模型使用 R2 模块运用多尺度分割的方法解决轮廓边缘分割精度不高的问题。将 SE 模块与 R2 模块组合，运用注意力机制理论解决难正样本的分割精度问题。本课题使用 CHAOS（Combined（CT-MR）Healthy Abdominal Organ Segmentation）竞赛提供的数据集。实验评估表明，本文方法同时分割四个器官（肝脏，右肾，左肾，脾脏）的 DSC 分别达到了 91.6%，88.6%，87.5%，79.7%，具有较高的分割精度，自适应性，和鲁棒性。

【关键词】全卷积神经网络；腹部 MRI 图像；多目标分类；多尺度；注意力机制

1 绪论

1.1 背景介绍

全卷积神经网络（FCN）在语义分割领域得到广泛的应用（图 1）。它采用了逐像素分类，使用端到端的卷积网络。在上采样的过程中使用了反卷积网络，使得压缩后的特征重新缩放为原始图像的大小。支持任意大小图像的输入。运用了跳层连接结合了深浅不同层次的特征，改善粗糙程度。其缺点是分割的效果不够精细，对细节不够敏感，并且没有考虑像素和像素之间的位置关系，没有考虑空间规整（spatial regularization）。

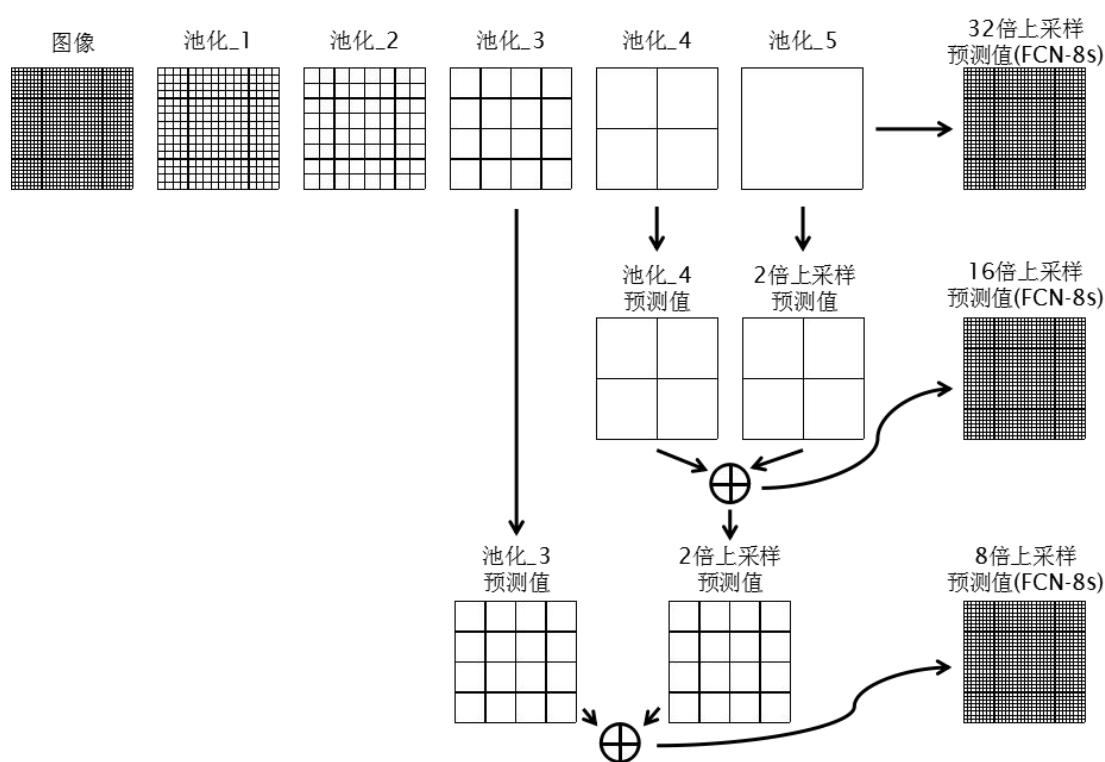


图 1 FCN 模型结构图

FCN 模型解决的逐像素分类问题主要应用于语义分割和实例分割，与医学影像分割的任务目标一致。医学影像分割在医学图像分析中是一项基本的任务，它可以对器官或结构进行自动划分，计算机辅助诊断[1][2][3]，视觉增强，干预，在影像中获取定量指数等任务，最终应用于器官手术指导，病变部位定位，3D 影像重建等任务。

用于分割的医学图像通常是由 3D 图像组成，可分为三个维度轴位方向（axial view, z 轴）、冠状方向（coronal view, x 轴）、矢状方向（sagittal view, y 轴）。目前有基于 2D 影像分割和基于 3D 影像分割[4][5][6][7]。医学影像分割技术已经广泛应用于多重器官的分割，如胰脏[8]、肝脏[9][10]、脾脏[11]、肾脏[12][13]、肺部[14]等。

腹部 MRI 图像分割是一项基本的分割任务。CHAOS 数据集提供的 MRI 图像分为 T1-DUAL (同相位和反相位) 和 T2-SPIR (T2 加权频谱预饱和和反转恢复序列), 具有多模态的特点。两种加权模式的差异较大, 在腹部器官的图像中有如下特征:

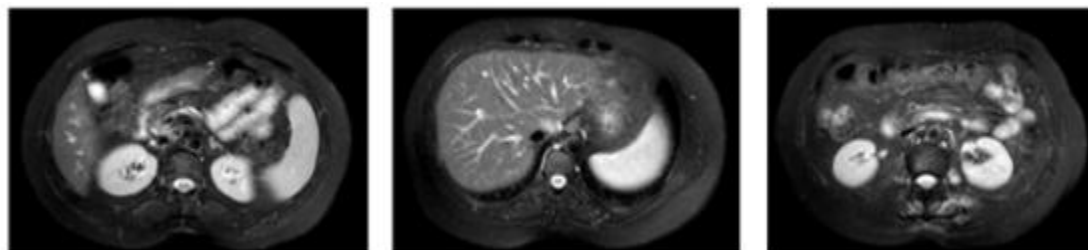


图 2 T2-SPIR 序列腹部 MRI 图像样本

1) SPIR (频谱预饱和和反转恢复序列) 代表一个混合成像序列, 并使用 T2 加权对比度机制。对于脂肪质子的选择性抑制, 预饱和脉冲分别应用于每个切片选择梯度。此序列需要校准的灵敏调整和非常均匀的磁场。上述 SPIR 的特点使其成为研究肝脏的首选序列, 因为肝脏实质可以很好地分析与抑制实质内的脂肪含量。此外, 腹部器官的边界外观在视觉上变得更清晰, 因为它们周围的脂肪组织受到抑制。T2 加权也可以在肝脏内的血管中导航, 因为它们看起来高度紧张。相邻的腹部器官和组织以其较高的信号值变得更加分离。SPIR 序列的一个更重要的贡献是它对运动的灵敏度较低。此功能可最大限度地减少腹部研究中对图像质量产生不利影响的伪影。

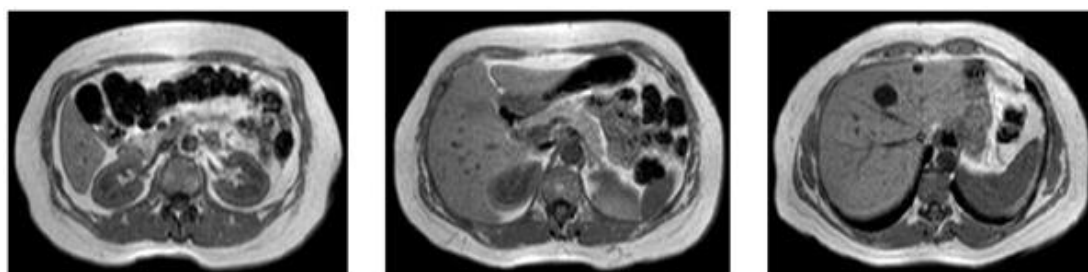


图 3 T1-DUAL (相位) 序列的腹部 MRI 图像样本

2) T1-DUAL (同相位和反相位) 是一个脂肪抑制序列, 它使用的是脂肪和水质子的 T1 倍的差异。信号被获得两次: 第一, 当水和脂肪质子处于阶段时, 第二, 当它们处于相位时 (而令人兴奋的质子正在返回到它们的第一个位置)。对于 1.5 Tesla 设备, 在阶段时间, 水和脂肪质子在同一个方向是 4.6 毫秒, 并且出相时间, 脂肪和水质子是相反方向, 是 2.3 毫秒。通过使用此信息确定 TE (回声时间) 值, 可以通过减去相应的脂肪和水信号频率来实现脂肪抑制。这个序列是非常有用的, 以了解脂肪含量的病变。由于 T1-DUAL 是一个 T1 加权序列, 它是非常有效的识别血液和组织是丰富的蛋白质。此序列还有助于确定肝脏润滑的水平。在超相图像中, 由于器官边界上的脂肪和水量突然变化, 取消了获得的信号, 器官的边界似乎是黑色的。T1-DUAL 的此属性有时用于边界划定算法。

分割 MRI 图像的腹部器官有助于病灶的分析。比如肝脏的影像诊断，可以识别肝脓肿。从形态和边缘角度分析，为圆形、椭圆形且边缘锐利。T1 加权像的信号特点为：低信号，不均匀，脓肿壁信号略高于脓腔而低于肝实质，厚薄均匀，外侧可见低信号水肿带。



图 4 T1 加权肝脓肿分析

全卷积神经网络（FCN）已被广泛应用于医学图像分析领域，许多研究在基于 FCN 的 U-Net 上做了不少改进，取得了良好的表现。尽管全卷积神经网络的应用广泛，但大多数方法提出的神经网络只能同时处理一个器官的分割，一个模型只能针对一个器官，或者一个阶段只能针对一个器官，另一阶段针对器官中的病变部位（图 5）。由于医学图像人工标注的难度较高，风险较大，可训练样本数量少等原因，将多目标分类应用在器官分割的研究较少，而临床实践中对分割速度的提升抱有很高期望。

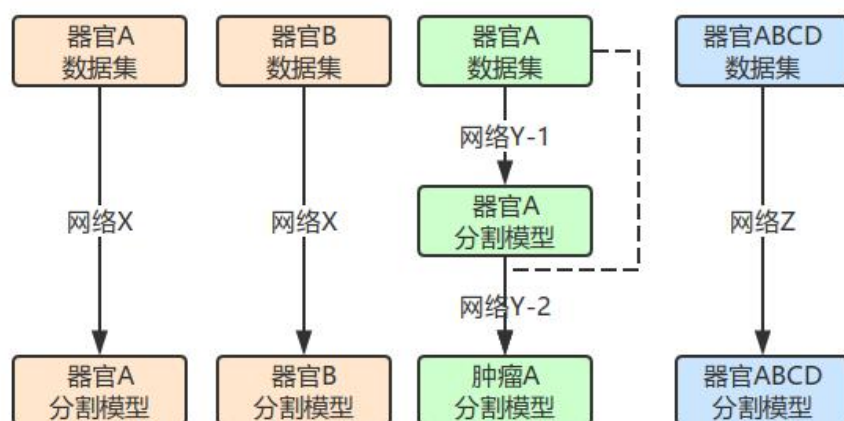


图 5 器官神经网络模型类型示意图

本文提出了一种基于全卷积神经网络的注意力机制器官分割模型 UGS-Net，用于训练端到端的腹部 MRI 多模态影像。该模型使用 R2 模块运用多尺度分割的方法解决轮廓边缘分割精度不高的问题。将 SE 模块（Squeeze-and-Excitation）与 R2 模块组合，运用注意力机制理论解决难正样本的分割精度问题。本文做了以下努力：详细地介绍模型中的细节，例如网络结构、参数和损失函数的方法。对模块的组合形式做了多次修改，探究不同特征

融合形式（逐项相加，逐项相乘，拼接）对分割性能的影响进行全面的分析，例如使用 4 折交叉验证。本工作对模型训练较长时间（训练至少 35 代）。提供的 UGS-Net 模型源码是公开的。

1.2 本课题的主要工作

第一章陈述腹部器官 MRI 图像的分割方法，介绍国内外现状研究，阐释了多模态 MRI 图像的特点以及分割难点，最后阐述了本文的内容安排。

第二章针对本课题工作，概括性地介绍了传统腹部器官分割方法、MRI 图像分割、多尺度方法、大核卷积优化四个方面的相关研究。

第三章介绍并分析了基于 FCN 的单器官图像分割模型 RSTN[15]，提出了基于 GCN 的多目标 MRI 图像分割的轻量化的 GCN++模型，和基于 FCN 的多目标 MRI 图像分割模型 UGS-Net。通过调节下采样的次数，改变不同层次间特征融合的方式，增加边界细化的流程，组合各种先进的模块，显著提升了分割精度，泛化能力强。

第四章对 RSTN、GCN++、UGS-Net 三种模型进行了实验设计与结果分析，并且对卷积核的选择、损失函数的设计、模块的性能进行了深入讨论。

第五章总结陈述本课题，回顾本文工作的特色之处与创新点，同时梳理研究的不足，对如何进一步提升 MRI 影像分割性能提出研究方向。

2 相关研究

2.1 传统腹部器官分割方法

传统腹部器官分割方法有基于阈值、边缘、区域生长等方法。除此之外新的研究还提出：1) 基于超像素的多目标腹部分割方法[16]结合多目标分类的方法，提高分类分割的精度。分类约束条件为超像素的有向邻接图生成腹部器官的 3D 区域。2) Graph cuts 分割算法[17]，运用自适应形状约束, 实现从腹部 CT 图像中提取肝脏、肾脏以及脾等组织器官的目标, 首先使用基于多图谱配准的分割方法得到形状先验，然后由形状先验的轮廓计算符号距离图, 将其通过形状比较函数加入到 Graph cuts 能量函数中; 接着根据配准分割过程中产生的目标概率图选择能量函数中每条边的形状约束项系数, 实现自适应形状约束。最后使用最大流最小割算法优化 Graph cuts 能量函数, 得到分割结果。该方法可以有效处理 CT 图像器官组织对比度低, 灰度不均, 组织边界模糊等分割问题。3) 针对器官区域小的问题, 有研究提出了用重新缩放图像大小[4][18]，使之与目标区域相匹配的方法，或者初阶段设置遮罩，后续阶段分割子区域[19]的方法。

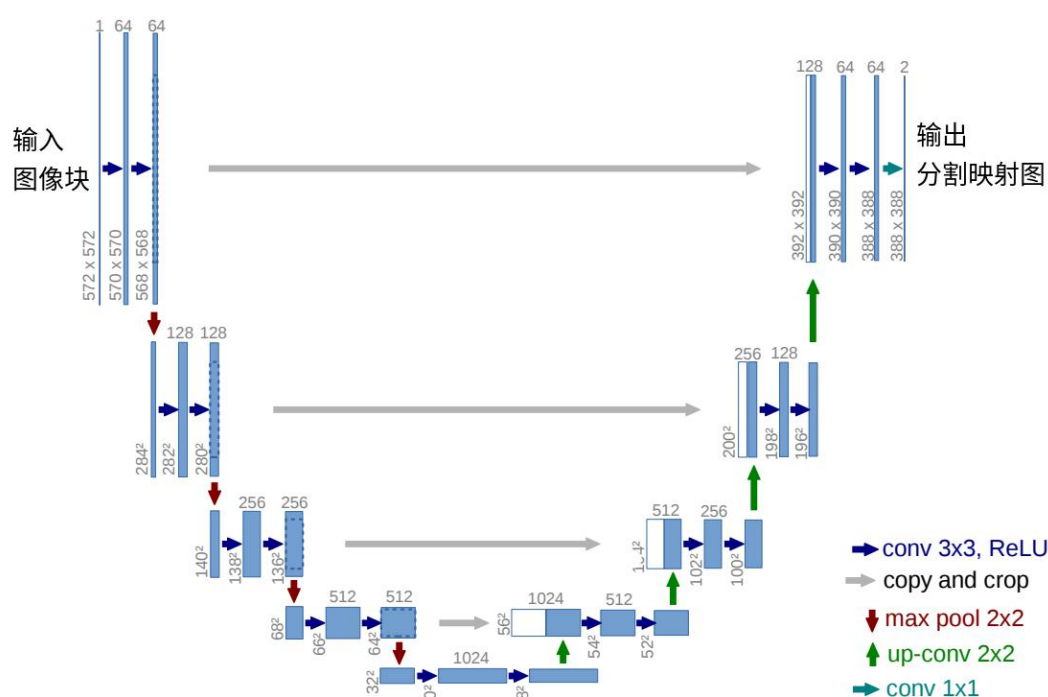


图 6 U-Net 网络结构图

2.2 MRI 图像分割

MRI 图像分割研究中，U-net[20]网络结构应用广泛。MRI 图像具有分辨率低，样本数量少的特点，所以高级特征和低级特征均具有很高的的重要性，而 U-net[20]网络结构（图 5）擅长处理高级特征和低级特征的融合问题。它基于全卷积神经网络（FCN），结构简单，高低层次特征相互融合，没有全连接层，优势在于可以在原图的分辨率上做特征融合。V-Net[6]基于 CNN 模型，在描述前列腺的 MRI 体积上进行端到端训练，并学会一次预测整个体积的分割。使用了新的损失函数，在训练期间根据 *Dice* 系数进行优化，以达到处理前景体素和背景体素之间存在强烈不平衡的情况。为了应对可用于训练的有限数量的数据集，使用随机非线性变换和直方图匹配来增加数据，取得良好的分割效果。

2.3 多尺度方法

由于腹部器官大小形状各异，其分割性能的提升依赖于多尺度。近年以来，在各种视觉任务中涌现出众多骨干网络（backbone networks）[21][22][23][24][25][26]，它们致力于加强多尺度表示法展现出先进的性能。从网络设计的角度，基于卷积神经网络（CNN）发展的网络均具备多尺度分割的能力，它们都遵循由粗到细的分割（coarse to fine）。AlexNet[27]使用了 ReLU 激活函数，将 Dropout 用在最后几层全连接中，相比过去的传统方法，对图像进行连续滤波，性能相比之前显著性提高。缺点在于网络结构不够深。后来涌现出深层次的网络，如 VGGNet[24]和 GoogLeNet [25]。VGGNet[24]将网络深度提升到 16 层和 19 层，特点在于仅使用 3×3 的卷积核堆叠在不同的层次上。相比 GoogLeNet [25]，VGGNet[24]由于其更深的网络，在多尺度的表现更好。缺点在于训练的速度特别慢，而且就带宽而言网络架构权重本身非常大。而且两者最后输出的结构为线性结构，这意味着它们接收的特征信息不够灵活，可处理的对象范围较小。尽管 GoogLeNet [25]使用不同大小的卷积核和过滤器并行处理特征（图 7），用于增强多尺度的表现，但受限于计算资源，其架构形式仍不够丰富，拓展性不强。

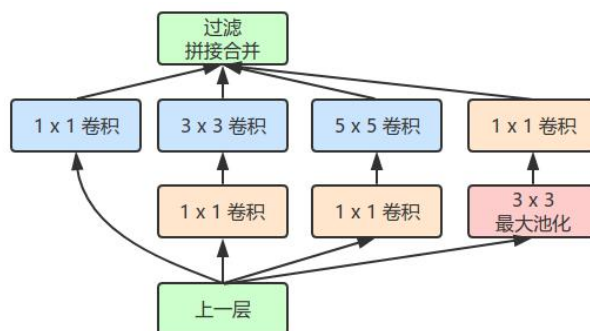


图 7 降低维度的 Inception 模块示意图

ResNet[21]相比于之前的模型，训练的卷积神经网络深度更高，进一步提升精确度。特点在于建立前后层之间的连接（skip connection, shortcuts）（图 8），这样有助于反向传播中梯度更新的敏感性，降低梯度弥散的副作用，从而更深的 CNN 网络得以训练。短连接使得多尺度方法可以在更多等效的特征尺度上运用。DenseNet[22] 提出了一种密集连接层的网络，模块的输入来自前面所有层输出的叠加，丰富了跳转连接的形式。最近提出的 DLA 方法[26]采取了更加复杂的树形结构进行更加深入地融合增强。

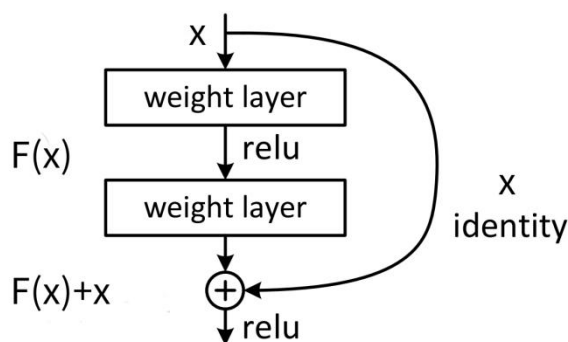


图 8 跳层连接示意图

多尺度方法可应用于多种视觉图像处理任务，例如图像分类[27][28]、目标检测[29][30]、语义分割[31][32][33]等等。

2.4 大核卷积

在器官分割任务中，性能和计算资源的权衡不可忽视。GCN 提出一种具有非常大的内核卷积的编码器 - 解码器架构[34]，是基于 ResNet [21]的模型改进。语义分割需要分割对象的分割和分类。由于完全连接层不能存在于分段架构中，因此采用大内核进行卷积。虽然 ResNet [21]更深层次的网络使其具有非常大的感知领域，但研究表明网络倾向于从更小的区域（有效的接收领域）收集信息。较大的内核计算量很大并且具有许多参数。因此， $k \times k$ 卷积用 $1 \times k + k \times 1$ 和 $k \times 1$ 和 $1 \times k$ 卷积的和来近似。该模块在本文中称为全球卷积网络（GCN）。从架构上分析（图 9），ResNet [21]（没有任何扩散的卷积）形成架构的编码器部分，而 GCN 和反卷积形成解码器。还使用称为边界细化（BR）的简单残余块。GCN 的伪大核卷积在参数量有限的情况下能获取较大的感受野，计算资源充足的情况下使用真大核卷积性能更加稳定。

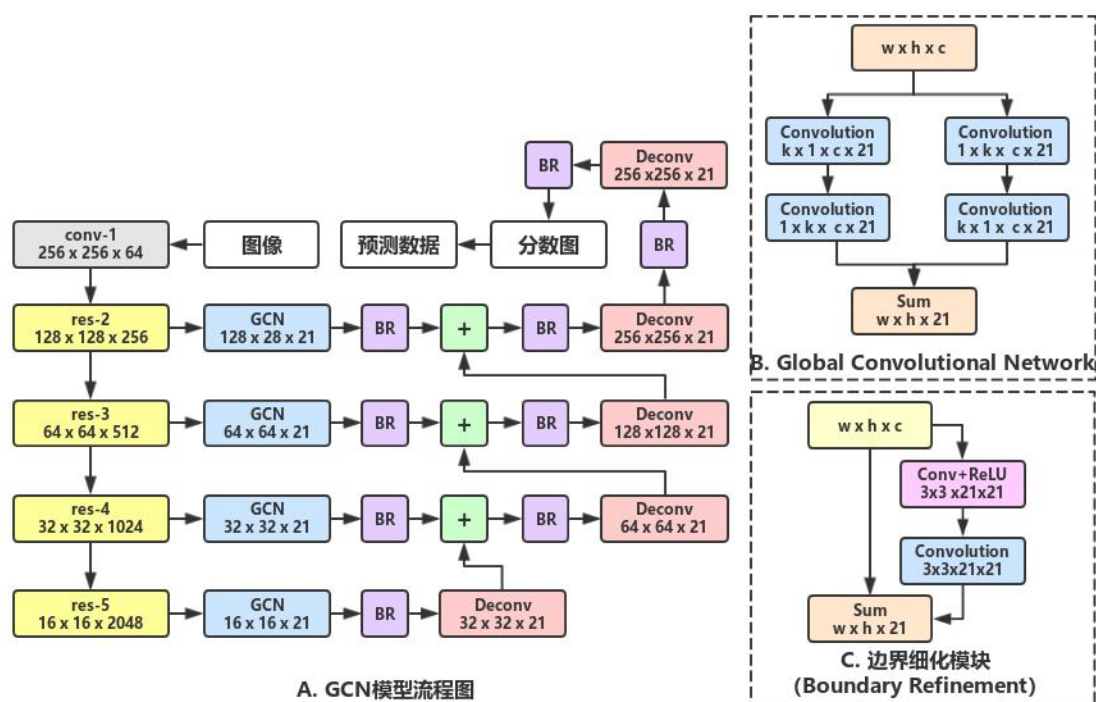


图 9 GCN 模型结构图

2.5 本章小结

腹部器官 MRI 图像的分割历经传统方法到机器学习方法的演进。传统方法有基于阈值、边缘、区域生长等分割方法以及基于超像素的方法以及 Graph cuts 分割算法。他们主要是基于数学知识和先验知识建立方法，具有较高的可解释性。现代方法更加青睐机器学习，基于 FCN 的 U-net 网络架构在医学影像分割领域中广泛应用，多尺度理论、大核卷积在 U-net 架构上的使用可以显著提高分割性，但是医学影像诊断是一项复杂的工作，端到端的网络架构设计容易让诊断医师对分割的可解释性造成困惑，因此采取新的策略提升分割的可解释性是深度学习分割 MRI 图像的一个挑战。

3 基于 FCN 的腹部器官 MRI 图像分割网络方法

本文分析了基于 FCN 的小器官分割模型（RSTN[15]），提出了基于 GCN 的多器官 MRI 图像分割模型（GCN++）以及基于 FCN 的多器官 MRI 图像分割模型（UGS-Net）。RSTN 模型支持轴位方向、冠状位方向、矢状位方向的分割，为了保证一致性，三种模型均对轴位方向（Z 轴）的图像分割。RSTN 模型（运用多尺度理论）仅分割肝脏，GCN++ 模型（运用多尺度理论，特征融合优化）和 UGS-Net 模型（运用多尺度理论，特征融合优化，注意力机制）同时分割肝脏、右肾、左肾、脾脏 4 种器官。

3.1 数据预处理

部分 MRI 图像的分辨率超过 256 x 256，从中心截取 256 x 256 的图像，使得模型允许多批次训练。

RSTN 网络模型：对强度范围做截取，范围为-100~2000，并对其做归一化，最小值为 0，最大值为 1。2%最高强度和 2%最低强度的训练数据被剔除，以减少异常值的影响。

GCN++网络模型：对强度范围做截取，做针对性截取，T1-DUAL 加权的影像范围为 0~1200，T2-SPIR 加权的影像范围为 0~1800。为了增强模型的鲁棒性，提升对难正样本的准确率，对原图 50%的概率在强度值范围内对强度值翻转（ $HU_{new} = range_{low} + range_{high} - HU$ ）。该翻转操作的可行性基于以下事实：MRI 图像中，T1 加权的肝脏强度高于脾脏。而在 T2-SPIR，脾脏的强度高于肝脏。

UGS-Net 网络模型：在 GCN++模型对数据预处理的基础上，增加了对数据拓展的测试，将平均切片 Z 轴间距 7.84 毫米的原图通过三次插值拓展切片数量，真实标注通过近邻插值拓展切片数量，原图切片数量由 1917 变为 7978。最终得到的平均切片 Z 轴间距为 2.0 毫米。

3.2 RSTN 网络

RSTN 网络[15]是一种基于 FCN 的图像分割方法，创新点在于提出了一种递归显著性变换网络，运用多尺度理论。特点在于信息模块的显著性转换，该模块反复将前一次迭代中的分段概率映射转换为空间权重，并将这些权重应用到当前迭代中。这给研究提供两份优点：1) 在训练阶段，它允许在处理不同输入尺度的深度网络上进行联合优化。2) 在测试阶段，它在迭代过程中传播多阶段的视觉信息，以提高分割的准确性。主要由两个 FCN8s 模型和显著性转换模块构成（图 10）。其中 FCN8s 模型使用预训练的模型。该模型

之前被应用在 CT 腹部影像的多器官分割，分割效果良好。

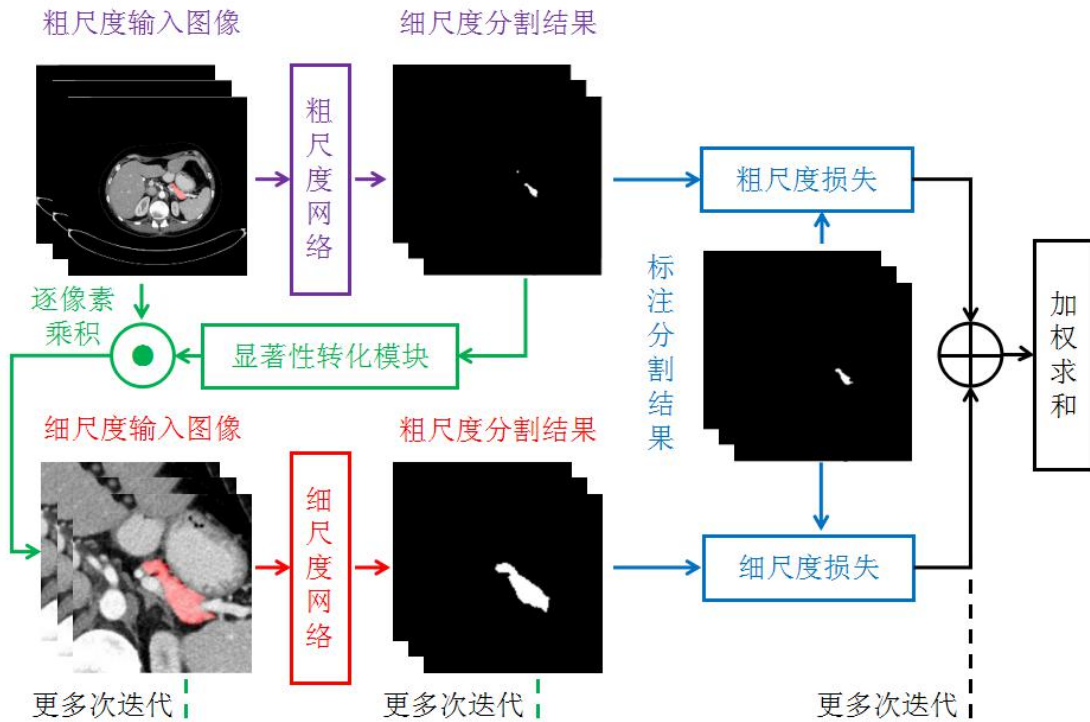


图 10 RSTN 模型[15]流程图

该模型分为三个阶段：S、I、J 三个阶段，每个阶段，细精度 FCN8s 模块仅分析真实标注周围的矩形区域。S 阶段，真实标注为细精度 FCN8s 模型提供位置范围，分析内容为原图；从 I 阶段和 J 阶段开始，粗精度 FCN8s 模块为细精度 FCN8s 模块提供位置范围。I 阶段中，细精度 FCN8s 模块分析内容为原图和显著性结果的结合，显著性结果由粗精度 FCN8s 模型的输出和显著性模块结合得出。对 J 阶段中，若粗精度 FCN8s 模块框选的范围大小为 0，则以真实标注范围为准。S、I、J 阶段分别跑 2、6、8 代，共迭代 16 次，训练时长约 40 分钟。

3.3 GCN++网络

相比于 RSTN 模型，GCN 有以下特点：GCN 是单阶段训练；参数量是 RSTN 的 1/3，能够在显存小于 4GB 的机器上运行，并且便于后期增加训练参数，调整网络结构，验证多尺度理论和特征融合理论在该模型上的运用对性能的提升；能够缓解器官形态复杂的问题，在截面面积非连通的场景下，分割效果较好。在 GCN 网络结构（图 9）的基础上，提出了改进版 GCN++ 模型，结构图如（图 11）所示。

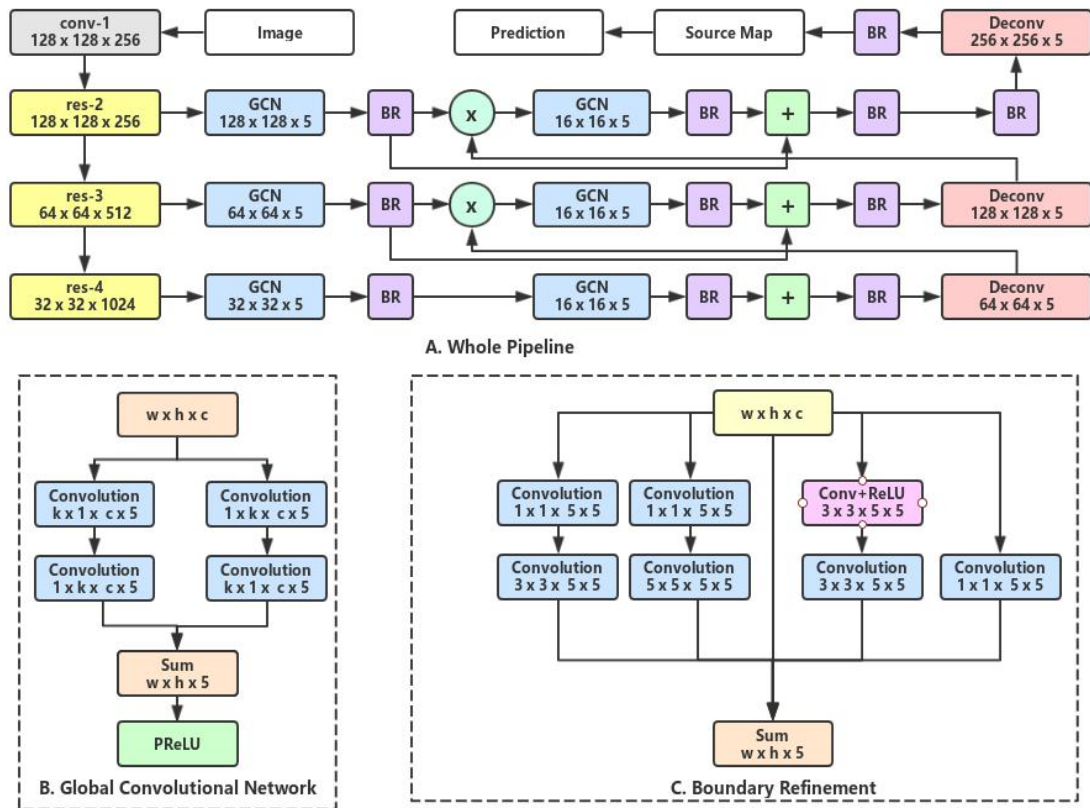


图 11 GCN++ 模型结构图

层级结构：针对数据集的特点做了以下改进：删除第四层，减少计算量的同时精确度没有太大变化；拓展了网络的宽度，增加 GCN 模块和 BR 模块，对图像做充分处理。

GCN 模块：增加了 GCN 模块的数量，目的是对边界做进一步的细化。改造 GCN 模块，在最后添加 PReLU 模块。该模块的作用是让机器自动学习参数，将负值收敛至 0 附近的一个合理区间，与 ReLU 模块相比较，添加该模块后有助于减小分割边界的误差。

BR 模块：增加了 BR 模块的数量，目的是对边界做进一步的细化；进一步复杂化 BR 模块，受 GoogLeNet[25]的 Inception 模块的启发，加入 3×3 的卷积核与 5×5 的卷积核，该操作能充分利用特征性进行多尺度的分割，提升分割精度。

Deconv（上采样）模块：采用双线性插值 (Bilinear Interpolation) 扩展分辨率为原来的两倍。

跨阶段连接：跨阶段连接相加有助于防止梯度弥散。在同一层次的处理上，经过了两次 GCN 模块和 BR 模块的处理，该两次跳层连接使用了逐像素相加的特征融合方式。这样能缩短各阶段处理的结果与最终结果之间的过程的路径，使得分割效果提升的更快。而是采用操作后的特征放置的位置较为靠前，有助于该层次模块充分利用高层次特征的信息。

特征融合操作：该模型充分利用扩展层次特征，用逐像素乘积的方式融合不同层次的特征，相比于逐像素相加，该方式对难正样本的分割精度提升明显。

3.4 UGS-Net 网络

RSTN 模型和 GCN++模型仍有不完善的地方，它们均没有在 MRI 影像的原尺寸上进行分割。因此基于 FCN 编码器-解码器的思想，提出了 UGS-Net 模型。该模型的目标是运用多尺度理论、与上述模型不同的特征融合方式和注意力机制，探究分割精度的极限。

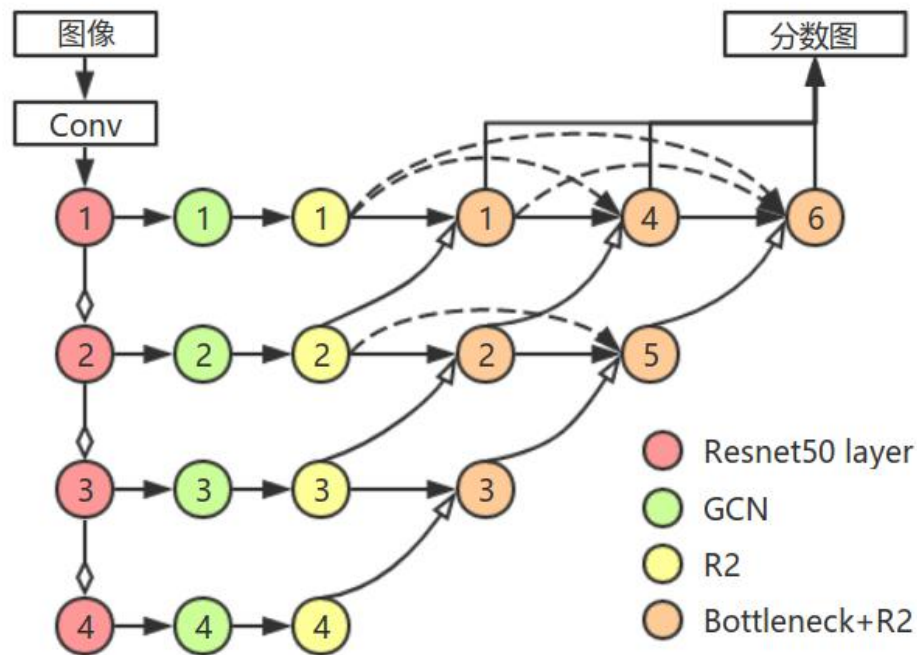


图 12 UGS-Net 模型结构图

网络结构：为了最大化利用显存性能，在网络结构上受 UNet++[36]的启发，在 FCN 编码器-解码器的基础上做修改，使得 UGS-Net 得以在 256 x 256 的分辨率上直接对 MRI 图像分割。UGS-Net 网络结构如图 12 所示。特点在于：1) 该网络一共分 4 层，比 GCN++多一层，层次越深越容易提取难正样本的特征。2) 前期的正向传播与 GCN 模型基本一致，使用 GCN 模块（图 11）减少层数的同时，用大的感受野试图保留全面的信息。3) Boundary Refine 模块被更换为基于多尺度理论的 R2 模块（受 Res2Net[35]模块启发）。4) 添加短连接与长连接跳跃结合的方案。与 FCN 的加操作（Summation）不同的是，UGS-Net 是叠操作（Concatenation），叠加后在 Bottleneck 模块中用 1x1 的卷积核压缩层数。

图 13 展示 R2 模块结构。R2 模块结构基于 Res2Net 的残差结构，目标是优化多尺度分割，同时能内嵌 SE 模块（图 14）结合使用，运用注意力机制提升分割表现。在 Res2Net 的基础上修改最后一组卷积核大小为 5 x 5，并在第一组添加 1 x 1 的卷积核。由于有 4 个器官，故分 5 个类别，网络层数训练期间分为 4 组共 20 层。最后合并 4 组为 1 组共 5 层作为最终的概率图。表 1 展示各模块的细节以及输入输出。

表 1 UGS-Net 模型细节

模块	模块细节	输入维数	输出维数
Conv	Conv2d(3, 64, kernel_size=1, stride=1) BatchNorm2d(64) ReLU	3 x 256 x 256	64 x 256 x 256
Resnet50.layer1		64 x 256 x 256	256 x 256 x 256
Resnet50.layer2		256 x 256 x 256	512 x 128 x 128
Resnet50.layer3		512 x 128 x 128	1024 x 64 x 64
Resnet50.layer4		1024 x 64 x 64	2048 x 32 x 32
GCN_1		256 x 256 x 256	20 x 256 x 256
GCN_2		512 x 128 x 128	20 x 128 x 128
GCN_3		1024 x 64 x 64	20 x 64 x 64
GCN_4		2048 x 32 x 32	20 x 32 x 32
R2_1		20 x 256 x 256	20 x 256 x 256
R2_2		20 x 128 x 128	20 x 128 x 128
R2_3		20 x 64 x 64	20 x 64 x 64
R2_4		20 x 32 x 32	20 x 32 x 32
BottleneckR2_1		40 x 32 x 32	20 x 32 x 32
BottleneckR2_2		40 x 64 x 64	20 x 64 x 64
BottleneckR2_3		40 x 128 x 128	20 x 128 x 128
BottleneckR2_4		60 x 32 x 32	20 x 32 x 32
BottleneckR2_5		60 x 64 x 64	20 x 64 x 64
BottleneckR2_6		80 x 32 x 32	20 x 32 x 32
Bottleneck	Conv2d(in_c, out_c, Kernel_size=1, stride=1) BatchNorm2d(out_channel)		
Deconv	Conv2d(in_c, in_c, Kernel_size=1, stride=1) BatchNorm2d(in_c) interpolate(x, scale_factor=2, Mode='bilinear', align_corners=True)		

R2 模块：R2 模块作为取代 Boundary Refine 的残差模块。首先经过一层 1×1 大小的卷积核，对输入数据进行微调。输入层数共 20 层，分为 4 组（x1, x2, x3, x4），每组各 5 层。每一组的边界细化是层层递进的，组号越大处理的信息越多，分割的效果也更加细化。x1 组经过 1×1 的卷积核得到 y1，x2 层经过 3×3 的卷积核得到 y2。y2 在经过 1×1 的卷积核后与 x3 作拼接合并（concatenate），经过 3×3 的卷积核得到 y3。y4 的计算方式与 y3 相近，区别在于 3×3 的卷积核换成 5×5 的大卷积核。最后对 y1, y2, y3, y4 重新拼

接，经过一层 1×1 卷积模块后被 SE 模块处理。

SE 模块：SE 模块（Squeeze-and-Excitation）被设置于 R2 模块的尾部，目的是使用计算得出的权重加强对特征的敏感性，重新衡量每层的权重。它以残差模块的输出作为输入，首先经过一层平均池化，接着经过两个全连接层，层数从 20 压缩为 5（压缩比为 $1/4$ ），再拓展为 20，最后经过 Sigmoid 操作将每层的数据转化为大于 0 的概率值，20 个概率值分别与 20 层特征图相乘得到新的 20 层特征图。

上采样模块：先经过一层 1×1 的卷积核，接着进行归一化权重（BatchNorm2d），最后使用双线性插值将分辨率提升为原来的两倍，与上一层的分辨率大小相适应。

Bottleneck 模块：目的是将多层特征融合为单层层数大小的特征。首先将多组每组 20 层的特征进行拼接，接下来的操作与上采样模块类似，区别在于去除了最后一步的分辨率扩展操作。

跨阶段连接：跨阶段连接在增加操作的同时也不影响网络的性能（引用）。跨阶段连接有一个 1×1 卷积核大小的开关，该开关系数为 0 时相当于关闭该连接。有利于后续测试的剪枝，使得测试的模型比训练的模型要小。

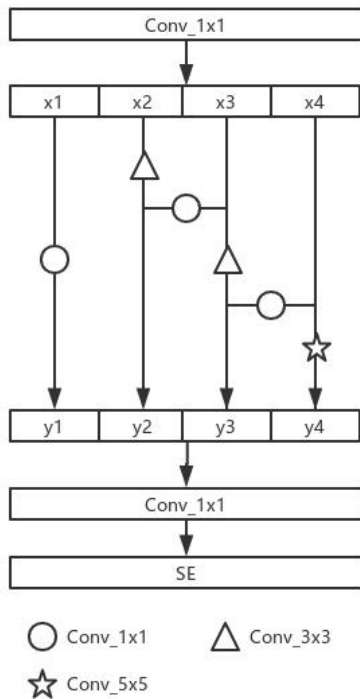


图 13 R2 模块结构图

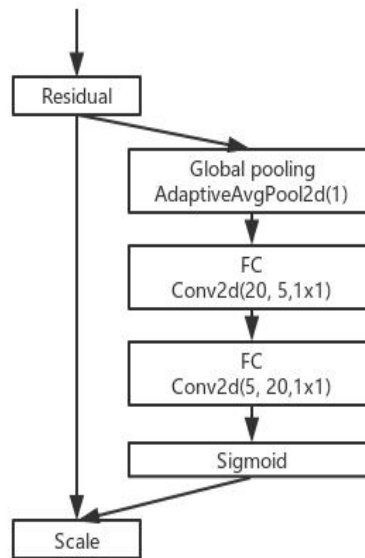


图 14 SE 模块结构图

3.5 损失函数

训练函数方面，RSTN 模型使用 *Dice* 损失对网络进行优化。RSTN 模型只能以单批次的模式训练，单批次包含 3 张相邻或者相同的图片。

$$Loss_{DSC} = \frac{2|A \cap B|}{|A| + |B|}$$

其中 A 是人工标注图像上分割器官区域的像素集合， B 是网络模型预测图像上分割器官区域的像素集合，分子表示 2 倍相交区域点的个数，分母表示两个集合各自的数量之和。

GCN++模型和 UGS-Net 模型使用交叉熵损失（cross entropy）来优化网络。两者的权重比均为为[1,4,8,8,4]（主观选择），分别对应背景、肝脏、右肾、左肾、脾脏。以下为 softmax 函数公式的解释，不包括加权操作。 z 为某像素对应各个类别的概率值， K 为类别总数， z 经过归一化处理后得到 q 。

$$q_i = \frac{\exp(z_i)}{\sum_{j=1}^K \exp(z_j)}$$

交叉熵损失公式如下，其中 i 表示多类别中的某一类。

$$Loss = -\sum_{i=1}^K p_i \log q_i$$

$$p_i = \begin{cases} 1, & \text{if } (i = y) \\ 0, & \text{if } (i \neq y) \end{cases}$$

3.6 本章小结

本章主要介绍三种神经网络，RSTN 模型（运用多尺度理论）由于训练时间有限的原因仅分割肝脏。通过粗略分析该模型，提炼出 RSTN 模型的不足之处（参数多，训练久，精度不高）等问题，研究运用多尺度理论，特征融合优化，注意力机制，先后提出 GCN++模型（运用多尺度理论，特征融合优化）和 UGS-Net 模型（运用多尺度理论，特征融合优化和注意力机制）同时分割肝脏、右肾、左肾、脾脏 4 种器官。针对 GCN++、UGS-Net 详细介绍了各模块的设计与作用。

4 实验设计与结果分析

4.1 实验数据和实验平台

采用 CHAOS 竞赛 (<https://chaos.grand-challenge.org>) 提供的临床获得的 60 例患者 (40 例 T1 加权/20 例 T2 加权) 的全腹部 MRI 扫描作为实验数据, 其中每个都是例行执行扫描腹部使用不同的射频脉冲和梯度组合。该数据不包括附加说明的有关器官 (即肝脏、肾脏、脾脏) 边界的任何肿瘤或病变。数据集是通过 1.5T 飞利浦 MRI 获取的, 该 MRI 生成的 12 位 DICOM 图像的分辨率为 $256 \times 256 \sim 320 \times 320$ 。ISDs (切片间距) 在 5.5-9 毫米 (平均 7.84 毫米) 之间变化, x-y 间距在 1.36 至 1.36 毫米 (平均 1.36 毫米) 之间, 切片数量在 26 至 50 之间 (平均 36 毫米)。60 份数据样本被分为四等份用于交叉验证, 45 份 (30 份 T1 加权/15 份 T2 加权) 作为训练数据, 其余 15 份 (10 份 T1 加权/5 份 T2 加权) 作为保留用于模型测试。由于 T1w 和 T2w 扫描是从相同的患者那里获得的, 所以我们确保了来自同一患者的扫描都用于训练或测试。为了关注网络和培训策略的改进, 没有采用数据增强步骤 (平移、旋转、翻转)。针对 UGS-Net 模型添加了数据增广 (增加切片数量) 的训练测试。本课题研究使用 NVIDIA GTX 1080TI GPU (12 GB 内存) 和 CUDA 9.0 进行实验。训练和测试是在易学智能平台 (www.easyaiforum.cn) 提供的工作站上进行的, 使用 Intel E5-2665/i5-7500 CPU, 10G 内存。

RSTN 的方法代码位于 (https://github.com/twni2016/OrganSegRSTN_PyTorch)。GCN++ 和 UGS-Net 的方法代码在 Python 3.6 (www.python.org) 中实现, 源码位于 (https://github.com/armyja/CHAOS_GCN)。anaconda 是 (<https://www.anaconda.com>) 用于托管虚拟 Python 环境和相应的库配置)。使用 PyTorch 0.4.1 版本 (www.pytorch.org) 建立网络架构并进行训练。

4.2 实验设计

RSTN 模型方面, 批尺寸 (batch size) 设置为 1。采用随机梯度下降法 (SGD) 对网络进行优化, 初始学习率 lr 为 3×10^{-5} , 动量参数 (momentum) 为 0.99, 权重衰减项 (weight decay) 为 0.0005。粗 FCN8s 模块中, 有 bias 属性的参数的学习率为 2 倍 lr , 无 bias 属性的参数的学习率为 4 倍 lr ; 细 FCN8s 模块中, 有 bias 属性的参数的学习率为 10 倍 lr , 无 bias 属性的参数的学习率为 20 倍 lr , 文献[15]的 RSTN 模型数据与处理将训练数据分为三个维度 (X, Y, Z)。训练时仅选择 Z 维度进行训练。模型的输入为 $3 \times 256 \times 256$ 大小的 MRI 图像, 是 3 张连续影像的组合或 3 张一样的组合。模型的输出是 2 份 $3 \times 256 \times 256$ 大小的概率图 (粗 FCN8s 模块与细 FCN8s 模块)。损失函数使用 Dice 系数公

式。在 $loss$ 损失的权衡方面，粗 FCN8s 模块的损失占 0.25，细 FCN8s 模块的损失占 0.75。

GCN++模型方面，在 pytorch 深度学习框架下对该方法进行验证。批尺寸（batch size）设置为 4。迭代次数 14000 次，约 38 代，训练时长约为 40 分钟。训练时采用 Adagrad 优化器对网络进行优化，初始学习率 lr 为 1×10^{-3} 。训练时仅选择 Z 维度进行训练。GCN++模型的输入为 $3 \times 256 \times 256$ 大小的 MRI 图像，是 3 张同样图像的组合。模型的输出是 1 份 $5 \times 256 \times 256$ 大小的概率图（背景、肝脏、右肾、左肾、脾脏）由于是多目标腹部器官分割，损失函数改为交叉熵函数（cross entropy）。权重系数数组为[1,4,8,8,4]。做了 4 次交叉验证。此外添加特征融合方式对分割效果的对照研究，在第 1 次交叉验证的数据集上进行。

UGS-Net 模型方面，除了模型结构不同，训练参数与 GCN++保持一致。批尺寸（batch size）设置为 4。迭代次数 14000 次，约 38 代，训练时长约为 140 分钟。训练时采用 Adagrad 优化器对网络进行优化，初始学习率 lr 为 1×10^{-3} 。损失函数仍为交叉熵函数（cross entropy）。权重系数数组为[1,4,8,8,4]。做 4 次交叉验证。此外，1）第 1 次交叉验证添加了 Res2Net 的残差模块与 R2 模块作对照。2）第 1 次交叉验证添加了数据增广作为对照组，增广操作将训练集的平均切片间距 7.84 毫米的原图通过三次插值拓展切片数量，真实标注通过最近邻插值拓展切片数量，原图切片数量由 1917 变为 7978。

实验结果以 DSC（Dice similarity coefficient）作为评价不同方法分割性能的主要指标。训练时长、模型参数量作为辅助评估指标。

4.3 实验结果

实验结果分为以下几个部分：1）探究不同特征融合操作对 GCN++模型分割效果的影响。2）探究数据增广对 UGS-Net 模型分割效果的影响。3）SE 模块对 UGS-Net 模型分割效果的影响。4）对比第 1 次交叉验证中 RSTN、GCN++、UGS-Net 三者分割肝脏器官的效果。5）GCN++、UGS-Net 分割肝脏、右肾、左肾、脾脏的效果。以上 5 个部分的前 4 部分是在第 1 次交叉验证中进行的，第 5 部分是 4 次交叉验证的综合结果。

1) 不同特征融合操作对 GCN++模型分割效果的影响

对（图 11）中逐像素相乘的形式（GCN++_P）做替换，添加逐像素相加（GCN++_A）、拼接（GCN++_C）作为两组对照组测试。测试结果如表 2 所示。

表 2 GCN++特征融合操作比较,评价指标为迭代 14000 次后的 DSC

器官\模型	GCN++_P	GCN++_A	GCN++_C
肝脏	0.9136	0.9078	0.9105
右肾	0.8842	0.9021	0.8919
左肾	0.9009	0.8943	0.8992
脾脏	0.7892	0.7876	0.7750

GCN++_P 在肝脏、左肾、脾脏都取得不到 1% 的涨点, 右肾涨点-1.8%。虽然可解释性不强, 但总体而言精度得到提升。

2) 数据增广对 UGS-Net 模型分割效果的影响

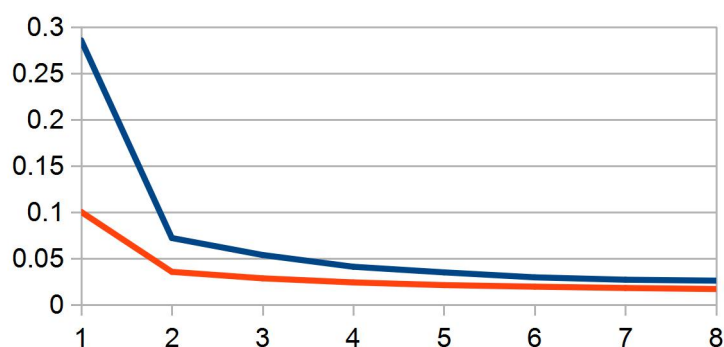


图 15 数据增广 loss 值测试

图 15 展现第 1 次交叉验证 UGS-Net 模型有无数据增广（蓝色曲线为有数据增广）的 loss 值训练代数（Epoch）变化。可见无数据增广的 loss 收敛的更快。

表 3 UGS-Net 数据增广 DSC 比较

器官\类别	无数据增广	有数据增广
肝脏	0.9191	0.9010
右肾	0.9113	0.8971
左肾	0.9219	0.9079
脾脏	0.8285	0.7975

表 3 展示 UGS-Net 第 1 次交叉验证有无数据增广迭代 12000 次后的机关分割 DSC 比较。无数据增广脾脏高 3%，其他器官平均高 1%。

上述两份数据说明了数据增广对网络的负优化。MRI 图像数据是通过三次插值生成的，生成的图像数据特点为边界模糊。而真实标注是通过近邻插值生成的。生成的标注数据轮廓与真实的器官轮廓存在一定差异。器官容积较小时，有误差的标注数据对分割提升作用不大。

3) SE 模块对 UGS-Net 模型分割效果的影响

SE 模块对 UGS-Net 模型分割效果的影响如表 4 所示。SE 模块对肝脏、肾脏能带来平均 0.8% 的涨点。

表 4 UGS-Net SE 模块 DSC 比较

器官\类别	有 SE 模块	无 SE 模块
肝脏	0.9191	0.9101
右肾	0.9113	0.9052
左肾	0.9219	0.9141
脾脏	0.8285	0.8291

图 16 为 UGS-Net 有无 SE 模块分割效果对比图，其中第 1 列是有 SE 模块的分割效果，第 2 列是无 SE 模块的分割效果，第 3 列是真实标注，第 4 列是 MRI 图像。紫色代表肝脏区域，蓝色表示右肾，浅粉色表示左肾，黄色表示脾脏。肝脏和脾脏的分割效果方面，有 SE 模块的组别表现更佳。



图 16 UGS-Net 有无 SE 模块分割效果对比图

4) RSTN、GCN++、UGS-Net 分割肝脏器官效果

分割效果评估分为直观比较和 DSC 指标评估。

直观比较：图 17，图 18，图 19 分别展示了 T1 加权同相位（In phase），T1 加权反相位（Out phase），和 T2-SPIR 加权的分割效果图，共分两栏，每栏的第一列白色区域是 RSTN 模型分割的肝脏区域。第 2 列和第 3 列，分别表示 GCN++ 模型和 UGS-Net 模型的分割结果。第四列是 MRI 图像与人工标注区域的混合。紫色代表肝脏区域，蓝色表示右肾，浅粉色表示左肾，黄色表示脾脏。

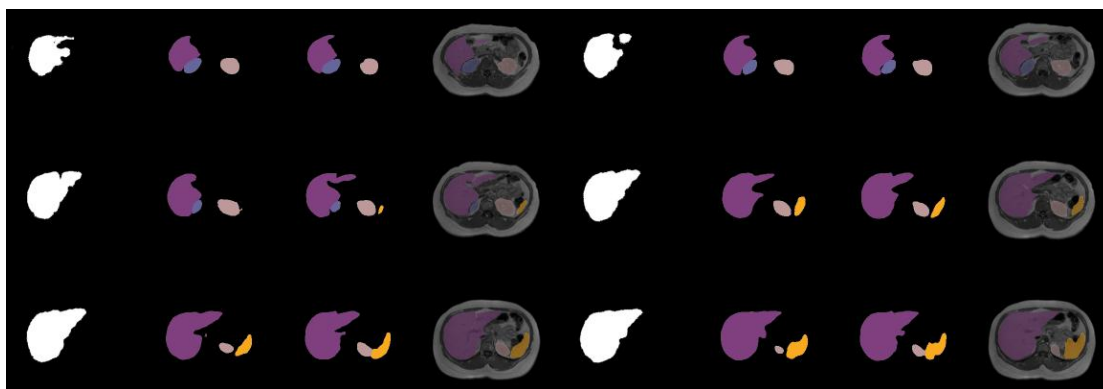


图 17 T1 加权同相位分割效果图

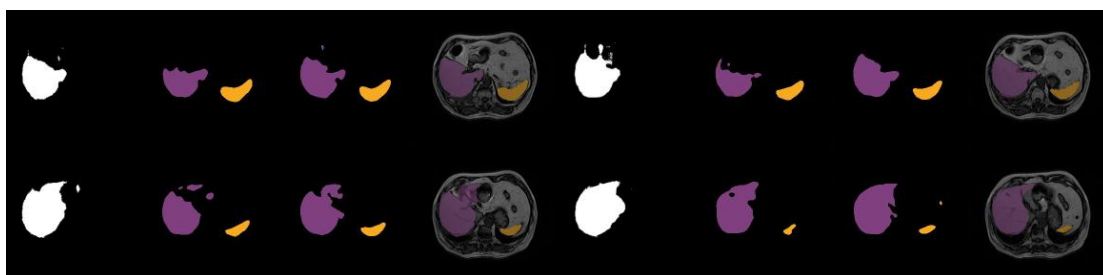


图 18 T1 加权反相位分割效果图

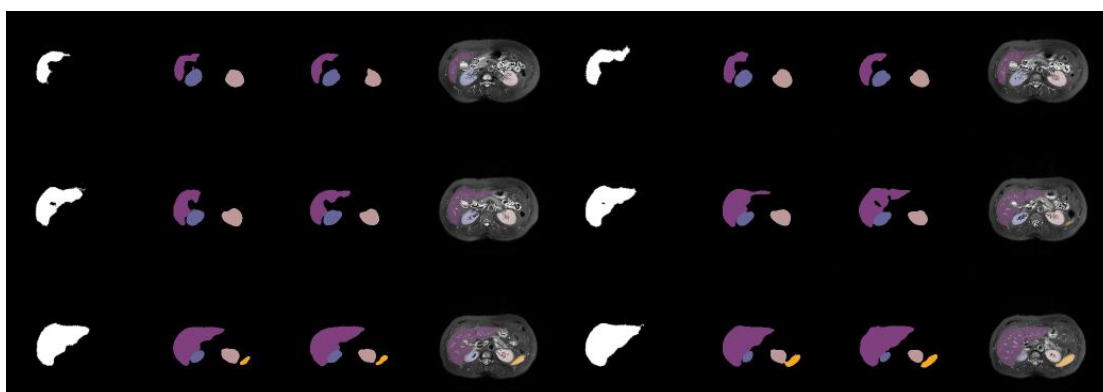


图 19 T2 加权分割效果图

可以看出，RSTN 模型倾向于覆盖较大的范围，有过拟合倾向，对复杂的轮廓细节处理的不够好。GCN++模型显得比较保守，容易受图像噪音的影响。UGS-Net 对轮廓的刻画效果更加精细。整体效果上， $RSTN < GCN++ < UGS-Net$ 。

DSC 指标评估：表 5 展示第 1 次交叉验证各模型的肝脏分割性能，RSTN_C 模型为粗 FCN8s 分割结果，RSTN_F 模型为细 FCN8s 分割结果，对 15 例样本（5 位患者）的分割表现，评估指标为 DSC。最难的肝脏样本 DSC，GCN++表现最好，为 0.8102，UGS-Net 紧随其后，为 0.7961。最容易的肝脏样本 DSC，UGS-Net 最高达到 95%。

表 5 第 1 次交叉验证各模型的肝脏分割性能

序号.	RSTN_C	RSTN_F	GCN++	UGS-Net
1	0.9043	0.9332	0.9390	0.9498
2	0.9259	0.9394	0.9323	0.9381
3	0.9298	0.9447	0.9468	0.9479
4	0.8763	0.9128	0.9288	0.9350
5	0.8901	0.9230	0.9091	0.9302
6	0.8473	0.9173	0.8889	0.8952
7	0.8685	0.9007	0.8953	0.8990
8	0.8757	0.9067	0.8796	0.8952
9	0.8978	0.9227	0.8830	0.8827
10	0.8973	0.9208	0.9420	0.9526
11	0.9144	0.9334	0.9274	0.9346
12	0.4273	0.4431	0.8102	0.7961
13	0.9264	0.9401	0.9535	0.9569
14	0.9283	0.9481	0.9449	0.9535
15	0.9255	0.9452	0.9234	0.9321
平均值	0.8690	0.8954	0.9136	0.9199

图 20 数据来自于表 5。RSTN_C 模型为粗 FCN8s 分割结果，RSTN_F 模型为细 FCN8s 分割结果。可以看出 RSTN_F 的分割效果比较稳定，但 RSTN 的异常样本处理水平并不高。UGS-Net 在最大值和中位值方面表现最佳。

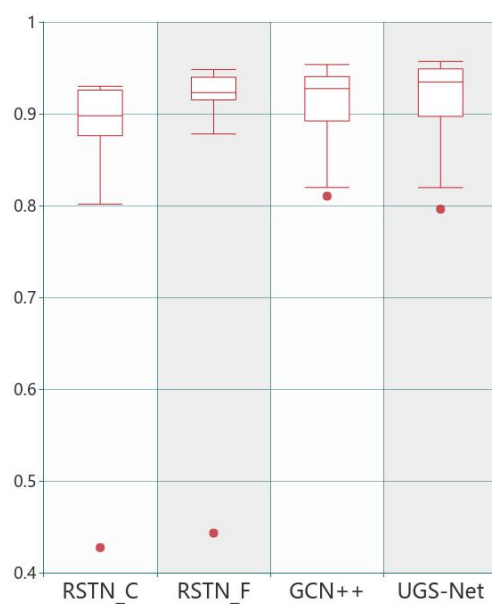


图 20 第 1 次交叉验证各模型肝脏的 DSC 箱体图

5) 4 次交叉验证中，GCN++、UGS-Net 分割肝脏、右肾、左肾、脾脏效果

效果评估分为 loss 曲线评估和 DSC 评估。

Loss 曲线评估：图 21 表示第 1 次交叉验证 UGS-Net 模型（蓝色）和 GCN++模型（红色）的 loss 值随训练代数（Epoch）变化。刨除异常值得影响，UGS-Net 的 loss 收敛的更快。

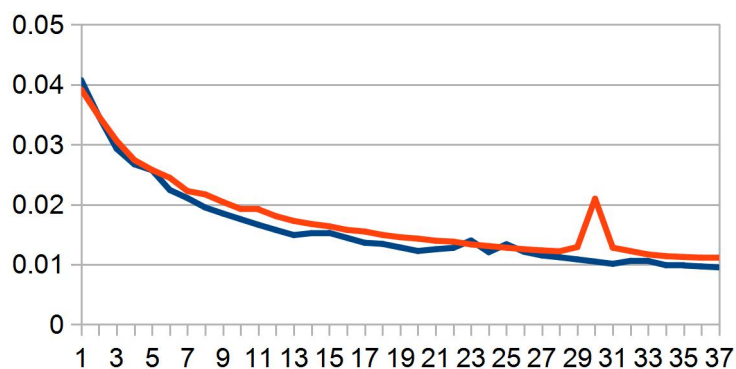


图 21 loss 值随训练代数（Epoch）变化

图 22 可以看出 UGS-Net 的 loss 值曲线走势更加密集，训练稳定性更高。

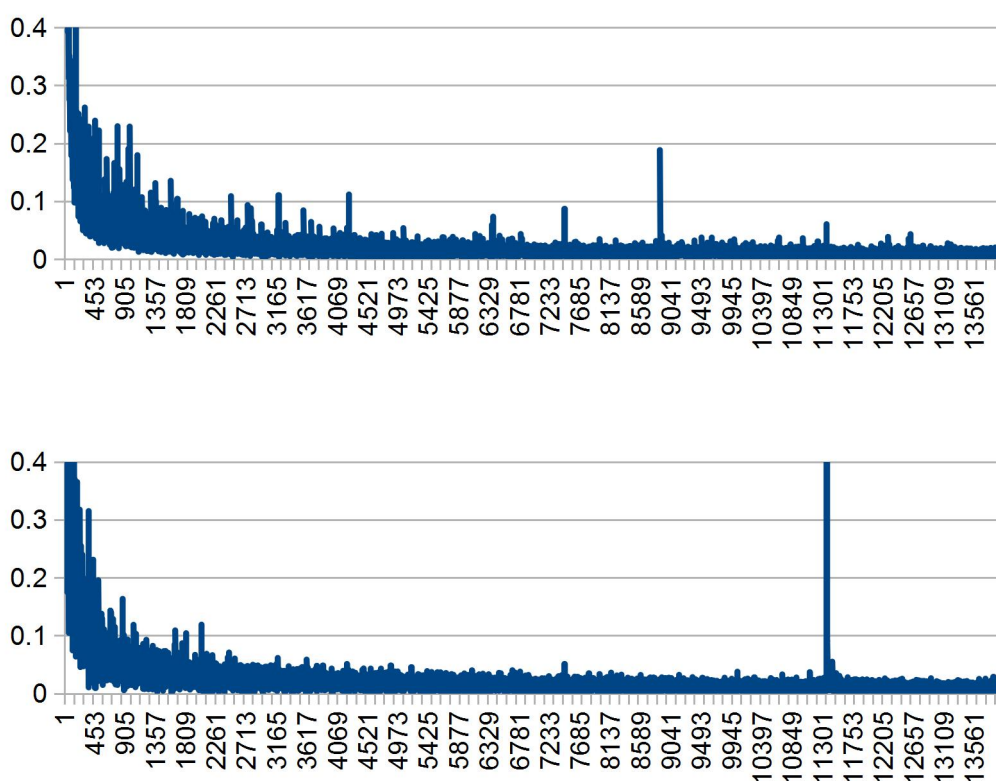


图 22 第 1 次交叉验证 UGS-Net 模型（上）和 GCN++模型（下）loss 值随迭代次数（Iteration）变化

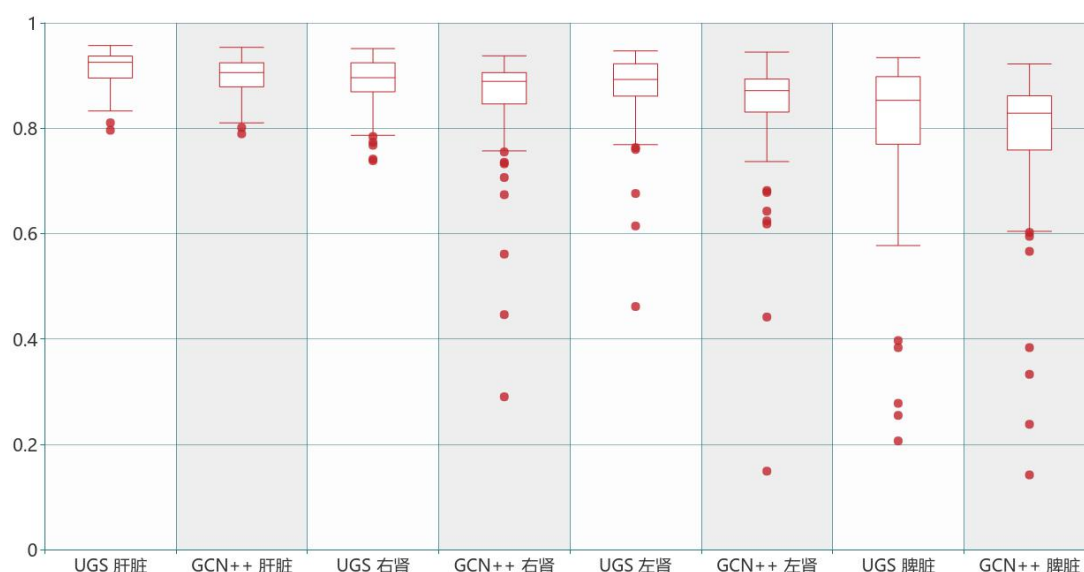


图 23 UGS-Net 模型和 GCN++模型对比, 评估指标为 DSC

DSC 评估: 图 23 为 UGS-Net 模型和 GCN++模型的腹部器官分割 *Dice* 系数比较, 采用 4 次交叉验证。表 6 表示 UGS-Net 模型。异常值用红点表示。可以看出, 从最大值和中位值的角度分析, UGS-Net 的表现优于 GCN++。从异常值的角度分析, UGS-Net 的异常值更为密集, 稳定性效果更佳。

表 6 UGS-Net (U) 和 GCN++ (G) 4 次交叉验证分割水平, 性能指标为 DSC

组号\器官	肝脏-U	肝脏-G	右肾-U	右肾-G	左肾-U	左肾-G	脾脏-U	脾脏-G
1	0.9199	0.9136	0.9118	0.8842	0.9234	0.9009	0.8329	0.7892
2	0.911	0.8846	0.8575	0.842	0.8628	0.8459	0.7435	0.7152
3	0.9181	0.8927	0.8913	0.8092	0.8111	0.7882	0.8239	0.7977
4	0.9136	0.9007	0.8842	0.8608	0.9009	0.7929	0.7892	0.7933
平均值	0.9156	0.8979	0.8862	0.8491	0.8746	0.832	0.7974	0.7738

综上所述, 1) 不同特征融合操作对 GCN++模型分割效果方面, 逐像素相乘表现稍好。2) 数据增广对 UGS-Net 模型分割效果起到负面影响。3) SE 模块对 UGS-Net 模型的帮助很大, 可以对抗过拟合。4) 对比第 1 次交叉验证中 RSTN、GCN++、UGS-Net 三者分割肝脏器官的效果, $RSTN < GCN++ < UGS-Net$ 。5) GCN++、UGS-Net 分割肝脏、右肾、左肾、脾脏的效果, UGS-Net 比 GCN++水平更高, 稳定性更强。以上实验充分说明了模型设计的有效性和网络模型的可行性。

4.4 实验分析

RSTN: 仅在第 1 次交叉验证中做了训练测试, 测试数据 (表 5) 和箱体图 (图 20、图

23) 展示了测试结果, 显著性模块的输出对细 FCN8s 模块的分割精确度提升 3%, 取得不错的效果。在理解模型架构的基础上, 对初始学习率 lr 、强度范围结合 CHAOS 竞赛数据集做了调试优化。其优点在于: 1) 利用了 FCN8s 输入分辨率为任意大小的特性, 使得裁剪后的图像直接导入细精度 FCN8s 模型训练。2) 细阶段缺点在于: 1) 模型偏大, checkpoint 模型大小超过 1GB; 2) 训练的模型只能针对一个器官进行分割; 3) 边界不清晰; 4) 仅允许单批次训练, 显存要求过高, 模型无法在 4GB 显存的机器上训练; 5) 对难正样本的分割精确度过低, 模型泛化能力不强。

GCN++: 箱体图 (图 20、图 23) 展示并比较了肝脏分割测试结果, 与 RSTN 的细 FCN8s 模块的分割精度相比, 虽然分割精度误差偏大, 但是 GCN++ 对难正样本的分割精度有了显著提升。平均 *Dice* 值为 0.9060, 超过前者的 0.8954。GCN++ 模型在不同层次间的特征如何融合、如何进一步细化边界方面做了探索性的尝试, 提高了平均分割精度。其优点在于: 训练的模型能分割多个器官, 泛化能力强, 训练时间短, 迭代次数多。缺点在于: 分割精度的稳定性与 RSTN 相比偏差较大。

UGS-Net: 测试结果如图 23 和图 20 所示。C 表示 RSTN 模型粗 FCN8s 模块, F 表示 RSTN 模型粗 FCN8s 模块, 相比 GCN++, UGS-Net 模型的中位数水平更高。肝脏、右肾、左肾和脾脏的分割, UGS-Net 模型比较稳定。图 17 可以看出, 小器官边界处的分割仍有提升空间。UGS-Net 尝试在原尺寸影像上分割, 借鉴 UNet++ 的网络结构, 采用大感受野的 GCN 模块、多尺度分割的 R2 模块以及基于注意力加权的 SE 模块。进一步提高了平均分割精度。其优点在于: 训练的模型能分割多个器官, 泛化能力强。缺点在于: 显存占用较高, 训练数据量大, 训练时间长。模型综合比较如表 7 所示。

表 7 不同模型各属性对比

属性\模型	RSTN	GCN++	UGS-Net
数据增广	无	无	无
强度倒转	无	有	有
训练参数	268555786	8691752	24634659
训练时间	40min	40min	140min
迭代次数	18	38	38
肝脏平均 DSC	0.8954270509785	0.897913385046568	0.915639383491634
右肾平均 DSC		0.849064113830736	0.886195485397131
左肾平均 DSC		0.831956045026205	0.874551831887752
脾脏平均 DSC		0.773821762259504	0.797368283498678

4.5 讨论

1) 特征融合

低级特征包含线条、轮廓等信息，高级特征包含更多语义信息。本课题在融合高低级特征的操作上比较简单。常见的分辨率扩张方式有两种：反卷积和双线性插值。虽然理论上反卷积的效果更好，但是实际应用中偏向于双线性插值。GCN++网络直接使用双线性插值扩大分辨率。UGS-Net 网络在双线性插值之前进行 1×1 的卷积和批次归一化。如何使用先验知识确定高低级特征的性质差异，以及使用什么样的模块来训练出一种局部算法与先验知识相匹配，是提升模型可靠性的一个值得探讨的方向。

2) 预训练模型

特征工程技术对神经网络的快速收敛的作用富有成效，预训练的模型有助于将模型突破最佳水准，Fine tuning（微调）通过降低预训练模型参数部分的学习率、提升模型新模块部分的学习率，来达到快速收敛的目的，本课题由于时间有限，没有对提出的 GCN++网络、UGS-Net 网络做微调（初始阶段模块加载了 ResNet50 的预训练参数）。然而如果迭代次数足够大，并且初始参数随机化，无预训练模型参数的性能表现将会更好。本课题的 4 次交叉实验，每次实验模型均运行了 38 代。38 代附近已经到达一个瓶颈，DSC 指标在这个区间内上下波动，波动范围在 1%DSC 左右。

3) 过拟合

小样本数量的数据通常会出现过拟合的情况。这在 UGS-Net 无 SE 模块中显得尤为突出，最大涨点可达-2%。UGS-Net 网络层数最深达到 4 层，基于注意力机制的 SE 模块对模型迭代维持高分割水准的帮助很大。

4) loss 函数的选择

本文提出的基于 FCN 的 MRI 腹部图像分割方法，使用加权交叉熵损失来训练模型，而不是使用 *Dice* 损失作为监督学习的损失函数。加权交叉熵损失有一定的主观因素，脾脏属于难正样本，其体积的大小有很大的解剖变化（图 23），这对分割准确率的提升造成困难。下一步模型的改进将针对不同的器官使用不同的损失函数综合考量。例如让 loss 值除以正样本点的数量，提升对小器官的敏感性。

5) 性能的分配

本课题同时对 4 个器官分割，提出的 UGS-Net 网络在实际运行中占用了接近 10000MB 的显存。多尺度分割的机制决定了 5 个类别在占用 4 张特征图。分割效果上，肝脏与脾脏的分割效果差距较大。试想，当显存性能足够充足时，如果神经网络可以自动判别样本的

难易程度，给予难正样本更多的特征图数量权重（相当于使用更多的卷积核训练），减少易正样本的权重，可能有助于整体性能的提升。亦或是神经网络根据不同难度的类别训练出各自的前向传播路径（类似于不同的个体回答一份心理测试问卷，从一个问题跳向另一个问题，路径不尽相同），也可能是一个潜在的涨点。

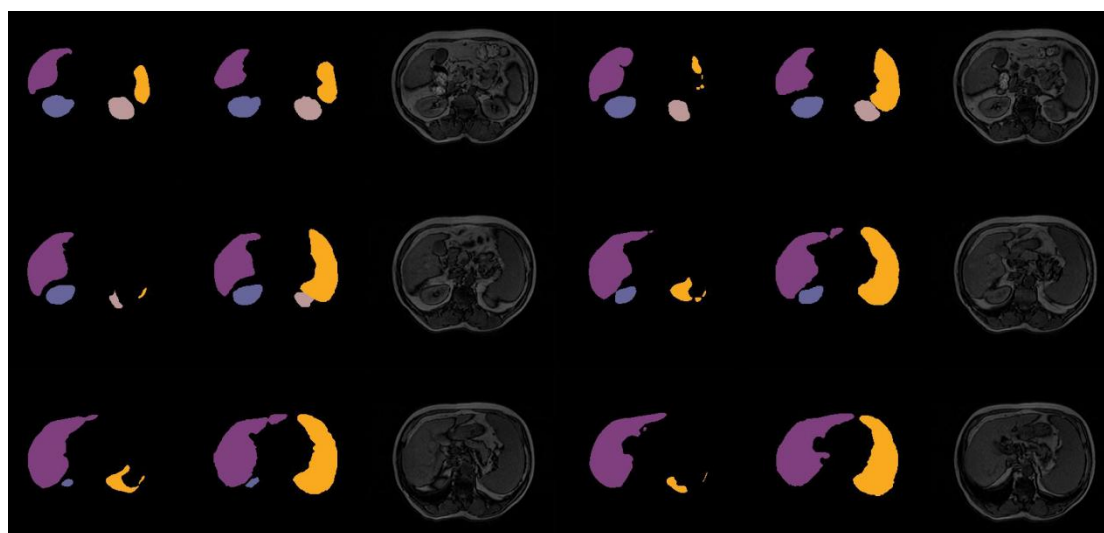


图 24 腹部器官分割难正样本比较

图 24 分为两栏，每栏的第一列是 UGS-Net 模型的分割结果，第二列是人工标注的分割图，第三列是 T2 加权的 MRI 图像。可以看出，模型对肝脏的分割效果较好，对脾脏的分割，小面积的分割效果较好（左上），对于大面积脾脏分割不敏感。由于两者的交叉熵权重是相同的，这说明在样本数量过少的情况下，难正样本对该模型分割效果并不理想。

6) 综述

提出的框架中使用了许多模块，例如伪内核（GCN），BR 模块，SE 模块，R2 模块。在图 10 和表二中，使用大卷积核的好处已经在 GCN++ 和 UGS-Net 中有所体现。RSTN 使用规则的 3×3 大小卷积核，GCN++ 和 UGS-Net 使用假的 7×7 卷积核 (1×7 和 7×1 组合) 和 5×5 大小的卷积核。结果表明，较大的卷积核具有较好的分割性能。相比 GCN++，UGS-Net 将残余模块由 BR 模块替换为 R2 模块，多尺度协同卷积提升了分割的敏感性。基于注意力机制的模型已经广泛应用于自然图像的分割中[23][37][38]。基于注意力的模块（例如 R2 和 SE）有助于该模型运用到对图像中小目标的检测任务[39]或者检测器官的病变区域[40]。由于训练时间有限，部分实验仅进行了一次交叉验证。未来将对各模型进行充分的数据增广，同时使用更多次数的交叉验证，更加完善地评估模型。

本文现有的工作，提出了在多尺度分割中深尺度使用大核进行卷积的思想。本文采用逐像素相乘的特征融合方法、运用基于注意力机制的 SE 模块，用于解决肺部难正样本分割水平的问题。本文建议使用更加复杂的损失函数解决难正样本分割精度的问题。本研究

使用 T1 加权和 T2 加权模式混合，同时对 4 个器官进行分割，从以往研究应用的角度看是比较少见的。在未来可以尝试多模态的 MRI 图像分割通过协同的方式使用多种网络进行上下文分割，通过某种方式使得模型输出元数据，用于辅助分析具有异常形状的器官体素。

4.6 本章小结

本章详细介绍了实验数据、实验平台，对 RSTN 模型、GCN++模型、UGS-Net 模型做了细致的测试，设置数据增广、SE 模块、特征融合模式等自变量，使用直观对比、Loss 值、DSC 值等指标评估模型分割性能，充分体现提出的模块以及模型在分割腹部 MRI 方面的先进性，多尺度理论、特征融合操作、注意力机制方面的改进都得到了有效验证。针对实验结果开展了特征融合、预训练模型、过拟合、loss 函数的选择、性能的分配共 5 个方面的思考，最后综述实验的有效性、可行性。

5 总结与展望

5.1 工作总结

提出了一种基于 FCN 的 GCN++模型，和一种基于全卷积神经网络的注意力机制器官分割模型 UGS-Net，运用多尺度技术和特征融合技术，用于训练端到端的腹部 MRI 多模态影像，同时分割 4 个器官（肝脏、右肾、左肾、脾脏）。本文模型在 CHAOS 竞赛（Combined (CT-MR) Healthy Abdominal Organ Segmentation）数据集上运行。

GCN++模型对 GCN 模型的修改主要有以下方面：

1) 删除第四层网络，网络架构更浅，分割速度提升，减少计算量的同时精确度没有太大变化。

2) 拓展了网络的宽度，增加 GCN 模块和 BR 模块，对 BR 模块添加更多卷积核，使用卷积和并行计算，对图像做充分处理。

3) 修改特征融合的方式，逐像素相加改为逐像素相乘。充分利用 MRI 影像的高级语义信息和低级特征信息的融合，提升了分割性能。

GCN++模型相比 GCN 模型，最终效果以第一次交叉验证对比，肝脏 91.36%，右肾 88.42%，左肾 90.09%，脾脏 78.92%。而 GCN 模型的结果为 88.95%，86.29%，85.70%，76.62%。平均涨点 2%，效果良好。

UGS-Net 的特点如下：

1) 下采样部分与 GCN++模型的区别在于添加第 4 层，支持在原分辨率图像上做分割，提升分割水平。后半部分采取类似 U-Net++的跳层连接结构，防止梯度弥散。

2) 将 GCN++的 BR 模块替换为使用 R2 模块（基于 Res2Net）运用多尺度分割的方法解决轮廓边缘分割精度不高的问题。

2) 将 SE 模块（Squeeze-and-Excitation）与 R2 模块组合，运用注意力机制理论解决难正样本的分割精度问题，提高了网络对多目标腹部器官的分割稳定性。

UGS-Net 模型和 GCN++模型的分割效果对此采取 4 次交叉综合验证，UGS-Net 模型肝脏、右肾、左肾、脾脏平均 DSC 达到 91.6%，88.6%，87.5%，79.7%，效果均优于 GCN++模型的 89.79%，84.90%，83.20%，77.38%，具有较好的精确性和鲁棒性。同时通过对照实验研究了特征融合、注意力机制、数据增广对分割效果的影响，证明了特征融合、注意力机制在网络模型中的有效性。

5.2 工作展望

本文提出的分割方法仍有值得优化的地方。

1) 本文使用的交叉熵损失函数中, 不同器官对损失值的权重影响被设置为定值。经过测试, 提升对难正样本的权重对其他分类的分割精度容易造成负面效果, 通过使用动态损失函数从而优化模型尚待研究。

2) 数据增强方面, 尝试对 MRI 影像进行顺时针 90 度旋转、对角线翻转, 大大提高了模型训练初期的难度。由于时间原因, 无法测试模型最终的鲁棒性。未来可研究旋转较小角度 (正负 10°) 或者平移短距离等数据增强策略对网络模型分割效果的影响。

3) UGS-Net 不同层次间的跳层连接较多, 有剪枝优化的空间。训练时保持完整分支, 测试时通过调试剪枝, 可以在保证分割性能的情况下节省显存。考虑多阶段分割在 GCN++ 模型上的应用, 也是值得优化的方向之一。

4) 数据和特征决定机器学习的上限, 在样本数量小的情况下进行特征数据优化可以有效提升分割性能。特征工程、先验知识、图像预处理 (如降噪) 在本课题的研究中没有涉及, 上述方法对网络模型的分割效果提升有待探索。

【参考文献】

- [1] T. Brosch, L. Tang, Y. Yoo, D. Li, A. Traboulsee, and R. Tam. Deep 3D Convolutional Encoder Networks with Shortcuts for Multiscale Feature Integration Applied to Multiple Sclerosis Lesion Segmentation[C], IEEE Transactions on Medical Imaging, 35(5):1229 – 1239, 2016.
- [2] D. Wang, A. Khosla, R. Gargeya, H. Irshad, and A. Beck. Deep Learning for Identifying Metastatic Breast Cancer, arXiv preprint arXiv:1606.05718, 2016.
- [3] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P. Jodoin, and H. Larochelle. Brain Tumor Segmentation with Deep Neural Networks[C], Medical Image Analysis, 2017.
- [4] L. Yu, X. Yang, H. Chen, J. Qin, and P. Heng. Volumetric ConvNets with Mixed Residual Connections for Automated Prostate Segmentation from 3D MR Images[C], AAAI Conference on Artificial Intelligence, 2017.
- [5] J. Merkow, D. Kriegman, A. Marsden, and Z. Tu. Dense Volume-to-Volume Vascular Boundary Detection [C], International Conference on Medical Image Computing and Computer-Assisted Intervention, 2016.
- [6] Milletari F, Navab N, Ahmadi S-A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation[C]. International Conference on 3D Vision; IEEE International Conference on 3D Vision, et al. 2016 Fourth International Conference on 3D Vision: 3DV 2016: 25-28 October 2016, Stanford, California, USA : proceedings. Piscataway, NJ: IEEE, 2016: 565-571.
- [7] K. Kamnitsas, C. Ledig, V. Newcombe, J. Simpson, A. Kane, D. Menon, D. Rueckert, and B. Glocker. Efficient Multi Scale 3D CNN with Fully Connected CRF for Accurate Brain Lesion Segmentation [C], Medical Image Analysis, 36: 61 – 78, 2017.
- [8] H. Roth, L. Lu, N. Lay, A. Harrison, A. Farag, A. Sohn, and R. Summers. Spatial Aggregation of Holistically Nested Convolutional Neural Networks for Automated Pancreas Localization and Segmentation. arXiv preprint arXiv:1702.00045, 2017..
- [9] T. Heimann, B. Van Ginneken, M. Styner, Y. Arzhaeva, V. Aurich, C. Bauer, A. Beck, C. Becker, R. Beichel, G. Bekes, et al. Comparison and Evaluation of Methods for Liver Segmentation from CT Datasets[C], IEEE Transactions on Medical Imaging,

28(8):1251 – 1265, 2009.

- [10]H. Ling, S. Zhou, Y. Zheng, B. Georgescu, M. Suehling, and D. Comaniciu. Hierarchical, Learning-based Automatic Liver Segmentation[C], Computer Vision and Pattern Recognition, 2008.
- [11]M. Linguraru, J. Sandberg, Z. Li, F. Shah, and R. Summers. Automated Segmentation and Quantification of Liver and Spleen from CT Images Using Normalized Probabilistic Atlases and Enhancement Estimation. Medical Physics [C],37(2):771 – 783, 2010.
- [12]D. Lin, C. Lei, and S. Hung. Computer-Aided Kidney Segmentation on Abdominal CT Images. IEEE Transactions on Information Technology in Biomedicine[C], 10(1):59 – 65, 2006.
- [13]A. Ali, A. Farag, and A. El-Baz. Graph Cuts Framework for Kidney Segmentation with Prior Shape Constraints [C], International Conference on Medical Image Computing and Computer-Assisted Intervention, 2007.
- [14]S. Hu, E. Hoffman, and J. Reinhardt. Automatic Lung Segmentation for Accurate Quantitation of Volumetric X-ray CT Images [C], IEEE Transactions on Medical Imaging, 20(6):490 – 498, 2001.
- [15]Qihang Yu, Lingxi Xie, Yan Wang, Yuyin Zhou, Elliot K. Fishman, Alan L. Yuille, "Recurrent Saliency Transformation Network: Incorporating Multi-Stage Visual Cues for Small Organ Segmentation"[C]IEEE Conference on CVPR, Salt Lake City, Utah, USA, 2018.
- [16]吕朝晖. 基于超像素的腹部图像多器官分割算法研究[D].西北大学,2018.
- [17]谢勤岚,潘先攀.自适应形状约束 Graph cuts 算法在腹部 CT 图像分割中的应用[J].中南民族大学学报(自然科学版),2019,38(01):119-125.
- [18]F. Xia, P. Wang, L. Chen, and A. Yuille. Zoom Better to See Clearer: Human and Object Parsing with Hierarchical Auto-Zoom Net[C], European Conference on Computer Vision, 2016.
- [19]Y. Zhou, L. Xie, W. Shen, Y. Wang, E. Fishman, and A. Yuille. A Fixed-Point Model for Pancreas Segmentation in Abdominal CT Scans [C], International Conference on Medical Image Computing and Computer-Assisted Intervention, 2017.
- [20]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, 2015, Springer. pp. 234 - 241.
- [21]K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition[C], In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 770 - 778, 2016
- [22]G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely connected convolutional networks[C], In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [23]G. Lin, A. Milan, C. Shen, and I. Reid. RefineNet: Multi Path Refinement Networks with Identity Mappings for High Resolution Semantic Segmentation[C], Computer Vision and Pattern Recognition, 2017.
- [24]K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556, 2014.
- [25]C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions [C], In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1 - 9, 2015.
- [26]F. Yu, D. Wang, E. Shelhamer, and T. Darrell. Deep layer aggregation[C], In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2403 - 2412, 2018.
- [27]A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks [C], In Advances in Neural Information Processing Systems (NIPS), pages 1097 - 1105, 2012.,
- [28]Z. Xu, Y. Huo, J. Park, B. Landman, A. Milkowski, S. Grbic, et al., "Less is More: Simultaneous View Classification and Landmark Detection for Abdominal Ultrasound Images," [C], International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 711-719, 2018.
- [29]O. Pauly, B. Glocker, A. Criminisi, D. Mateus, A. M. M?ller, S. Nekolla, et al., "Fast multiple organ detection and localization in whole-body MR Dixon sequences,"[C], International Conference on Medical Image Computing and Computer-Assisted Intervention, 2011, pp. 239-247.
- [30]B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Object detectors emerge in

- deep scene cnns",arXiv preprint arXiv:1412.6856, 2014.A
- [31]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.
- [32] Milletari, F., Ahmadi, S.A., Kroll, C., Plate, A., Rozanski, V., Maiostre, J., Levin, J., Dietrich, O., Ertl-Wagner, B., B ¨ otzel, K., et al.: Hough-cnn: Deep learning for segmentation of deep brain regions in mri and ultrasound, arXiv preprint arXiv:1601.07014, 2016
- [33]Noh, H., Hong, S., Han, B.: Learning deconvolution network for semantic segmentation. [C], Proceedings of the IEEE International Conference on Computer Vision. pp. 1520 – 1528, 2015
- [34]Peng C,Zhang X Y,Yu G,et al. Large kernel matters——improve semantic segmentation by global convolutional network [C]Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu,HI,USA: IEEE,2017: 4353-4361. [DOI: 10.1109 / CVPR. 2017. 189]
- [35]Gao, S.-H., Cheng, M.-M., Zhao, K., Zhang, X.-Y., Yang, M.-H., Torr, P. Res2Net: A New Multi-scale Backbone Architecture, 2019, arXiv e-prints arXiv:1904.01169.
- [36]Zhou, Z., Mahfuzur Rahman Siddiquee, M., Tajbakhsh, N., Liang, J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation[C], Stoyanov D. et al. (eds) Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA 2018, ML-CDS 2018, arXiv e-prints arXiv:1807.10165.
- [37]L. Chen, Y. Yang, J. Wang, W. Xu, and A. Yuille. Attention to Scale: Scale-aware Semantic Image Segmentation [C], Computer Vision and Pattern Recognition, 2016.
- [38]G. Li, Y. Xie, L. Lin, and Y. Yu. Instance-Level Salient Object Segmentation [C], Computer Vision and Pattern Recognition, 2017.
- [39]S. Singh, D. Hoiem, and D. Forsyth. Learning to Localize Little Landmarks. Computer Vision and Pattern Recognition, 2016.
- [40]Y. Zhou, L. Xie, E. Fishman, and A. Yuille. Deep Supervision for Pancreatic Cyst Segmentation in Abdominal CT Scans [C], International Conference on Medical Image Computing and Computer-Assisted Intervention, 2017.

致谢

在深圳大学四年本科学习生活即将结束，回想这四年，学到了很多，包括新知识和一些为人处世的道理。在这四年的学习生活里，有收获的喜悦，也有困难面前的沮丧。本次毕业设计和论文得以圆满完成，首先要特别感谢易学智能平台对本课题的巨大帮助，使得网络模型得以快速落地应用，同时保障实验环境和实验环境的稳定性。其次要感谢我的指导教师吴惠思老师给我的悉心指导和大力支持，感谢我的学友陈展华给我的鼓励和启发，也感谢舍友刘远平、陈先钦、陈锐泽在这过程中的支持和提醒，你们的帮助让我获益良多，终生受用，在此表示我深深的谢意。

感谢我的父母以及所有的家人，这些年来对我的默默付出和支持，让我能够无忧无虑、全身心的投入到科研和学习中，在此对他们致以崇高的敬意，你们的鼓励是我前进的最大动力。

最后，感谢在工作繁忙之际抽出时间评阅论文的各位专家和老师，感谢你们为完善论文提出的中肯建议，感谢答辩委员会的各位老师对本文提出的宝贵意见。

MRI medical image segmentation of abdominal organs based on FCN

School of Computer and Software (Computer Science and Technology) Wang Jiahui

Student ID: 2015150073

【Abstract】 Fully Convolutional Network (FCN) has been widely used in computer vision and medical imaging analysis. Abdominal MRI image segmentation is a basic segmentation task. Although the Fully Convolutional Network is widely used, most methods can only deal with the segmentation of one organ at the same time, and there are few research on the application of multi-target classification in organ segmentation. However, there are high expectations for the improvement of the speed of the segmentation in clinical practice. There are four major difficulties in this study: multiple organs, multiple modality, large anatomical variability and small dataset. Based on the problems above, we propose an organ segmentation model UGS-Net based on the Fully Convolutional Network for the training of end-to-end abdominal MRI multimodal images. The model uses R2 module to solve the problem of low precision of contour edge segmentation with multi-scale theory. Combining the SE module and R2 module, the attention mechanism theory is applied to solve the segmentation accuracy problem of difficult positive samples. This topic using dataset provided by CHAOS (Combined (CT-MR) Healthy Abdominal Outraged Segmentation) competition. Experimental evaluations that the DSC of four organs (liver, right kidney, left kidney and spleen) report 91.6%, 88.6%, 87.5% and 79.7% respectively, with high segmentation accuracy, adaptability and robustness.

【Key words】 Fully Convolutional Network; Abdominal MRI image; multi-target classification; multi-scale; attention mechanism