

中图分类号:

单位代号: 10280

密 级:

学 号: 18721753

上海大学



硕士学位论文

SHANGHAI UNIVERSITY

MASTER'S DISSERTATION

题 目	基于多维度编解码网络的 肝脏与肿瘤分割方法研究
-----	----------------------------

作 者 李小龙

学科专业 计算机应用技术

导 师 韩越兴

完成日期 2021.05.01

姓 名：李小龙

学号：18721753

论文题目：基于多维度编解码网络的肝脏与肿瘤分割方法研究

上海大学

本论文经答辩委员会全体委员审查, 确认符合上海大学硕士学位论文质量要求。

答辩委员会签名:

主任:

委员:

导 师:

答辩日期:

姓 名：李小龙

学号：18721753

论文题目：基于多维度编解码网络的肝脏与肿瘤分割方法研究

原创性声明

本人声明：所呈交的论文是本人在导师指导下进行的研究工作。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已发表或撰写过的研究成果。参与同一工作的其他同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

签 名：_____日期：_____

本论文使用授权说明

本人完全了解上海大学有关保留、使用学位论文的规定，即：学校有权保留论文及送交论文复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容。

(保密的论文在解密后应遵守此规定)

签 名：_____ 导师签名：_____ 日期：_____

上海大学学术硕士学位论文

基于多维度编解码网络的 肝脏与肿瘤分割方法研究

姓 名：李小龙

导 师：韩越兴

学科专业：计算机应用技术

上海大学计算机工程与科学学院

2021 年 5 月

A Dissertation Submitted to Shanghai University for the
Degree of Master in Engineering

**Research on Liver and Tumor
Segmentation Method Based on
Multi-dimensional Codec Network**

MA Candidate: Xiaolong Li

Supervisor: Yuexing Han

Major: Computer Application Technology

School of Computer Engineering and Science

College, Shanghai University

May, 2021

摘要

肝癌作为全球第六大肿瘤致死病因，是一种常见的癌症。当肝细胞出现失控生长时就会导致原发性肝癌。在临床实践中，肝肿瘤的分类与分期通常取决于肿瘤的大小、位置和形状，可以通过 CT 和 MRI 等影像检查技术对肝癌进行诊断。准确地测量肝脏与肝肿瘤的位置、大小和形状能够帮助医生对病人病情做出更全面的评估和规划。然而在 CT 图像中，肝脏与肝肿瘤的对比度较低，形状大小因人而异，准确地测量肝脏与肝肿瘤的大小、位置和形状是肝肿瘤分类和分期的关键所在。

经典的图像处理方法和机器学习方法，如阈值法、边缘检测法和聚类等方法通常依赖人工设计与提取的特征，难以对肝脏与肝肿瘤进行准确地自动化分割。而诸如 FCN、U-Net 等深度学习方法由于网络维度的限制，难以探索三维 CT 图像中的空间特征信息。本文针对腹部扫描 CT 图像中肝脏与肝肿瘤的特点，结合了多种图像处理与深度学习技术提出了两种不同的自动化肝脏与肿瘤分割方法。在第一个方法中，本文基于 FCN 提出了一种新的 2.5D 全卷积网络结构，能够在减少网络参数量和计算资源消耗的同时，有效地探索 CT 图像中的空间特征信息，提升肝脏与肝肿瘤分割准确度；针对医学图像的特点和常见的损失函数缺少优化网络探索边界特征能力的问题，设计了一种新的边界损失函数，融合了图像轮廓的距离、面积和边界信息，能够有效地优化深度学习网络，探索更多的图像边界和轮廓特征；针对深度学习网络的分割结果仍然存在假正例与假反例的情况，本文结合 3D 密集条件随机场和 3D 最大连通区域对网络的肝脏与肿瘤分割结果进行后处理，进一步提升图像的分割精度。在第二个方法中，为了进一步优化网络结构，针对编码解码网络存在忽视局部特征之间相关性和依赖性的问题，本文将 2D、2.5D、3D 网络与注意力机制整合起来，提出了一种基于双路注意力的编解码网络分割框架，该框架融合了双路自注意力机制模块、密集网络块、残差网络块和双路径网络块，包含九种不同的网络结构，能够有效地完成肝脏与肿瘤 CT 图像的自动化分割任务。

本文从不同的角度在 LiTS 和 3DIRCADb 数据集上对提出的方法进行了实验验证，并与近年来的优秀模型进行比较。实验结果表明，本文提出的方法具有较高的分割精度，验证了方法的有效性，为临床治疗提供了一定的参考价值，为智能医疗的早日实现提供了经验基础。

关键词：医学图像分割，肝脏与肿瘤，深度学习，边界损失函数，注意力机制

ABSTRACT

Liver cancer is common cancer as the sixth leading cause of cancer death in the world. Primary liver cancer occurs when liver cells grow out of control. Classification and staging of liver tumors usually depend on the tumor's location, size, and shape. And liver cancer can be diagnosed by imaging techniques such as CT and MRI in clinical practice. Accurately measuring the size, location, and shape of the liver and liver tumors can help doctors make a more comprehensive assessment and plan for a patient's condition. However, in CT images, the contrast between liver and liver tumors is low, and the shape and size of liver tumors vary from person to person. Therefore, accurate measurement of the size, location and shape of liver and liver tumors is a difficult problem for liver tumor classification and staging.

Classical image processing methods and machine learning methods, such as threshold method, edge detection method and clustering method, usually rely on the design and extraction of artificial features. So it is difficult to perform accurate automatic segmentation of liver and liver tumors. Due to the limitation of network dimension, deep learning methods such as FCN and U-Net are difficult to explore the spatial feature information in 3D CT images. Therefore, in view of the characteristics of liver and liver tumors in abdominal CT scan images, this paper combined multiple image processing techniques and deep learning techniques to propose two different automated liver and tumor segmentation methods. In the first method, this paper proposes a new 2.5D network structure based on FCN, which can effectively explore the spatial feature information in CT images, reduce the number of network parameters and computational resource consumption and improve the segmentation accuracy of liver and liver tumors. To solve the problem some loss functions cannot optimize the network to explore boundary features in the medical images, we integrate the distance, area, and boundary information of

the image contour to a boundary loss function, which can effectively optimize the deep learning network to explore more image boundary and contour features. For the segmentation result of the deep learning network, there are still false-positive examples and false-negative examples. This paper combined 3D dense conditional random field and 3D maximum connected region to post-process the segmentation results of liver and tumor, and further optimize the segmentation results of CT image. In the second method, in order to further optimize the network structure, aiming at the problem of ignoring the correlation and dependency of local features in the encoding and decoding network. This paper integrates 2D, 2.5D and 3D networks with the attention mechanism, and proposes a segmentation framework of codec network based on dual path attention. This framework integrates the dual-path attention mechanism module, the dense network block, the residual network block and the dual-path network block, including nine different networks, which can effectively complete the automatic segmentation task of liver and tumors.

The proposed method is validated on LiTS and 3DIRCADb datasets from different perspectives and compared with the excellent models in recent years. The experimental results show that the proposed method has a high segmentation accuracy, which verifies the effectiveness of the method, provides a certain reference value for clinical treatment and provides an empirical basis for the realization of intelligent medical treatment.

Keywords: Medical Image Segmentation, Liver and Tumor, Deep Learning, Boundary Loss Function, Attention Mechanism

目录

第一章 绪论	1
1.1 课题研究背景与意义	1
1.2 国内外研究概况	3
1.2.1 基于经典图像处理技术的分割方法概述	3
1.2.2 基于机器学习的分割方法概述	5
1.2.3 基于深度学习的分割方法概述	6
1.3 论文主要工作	9
1.4 论文组织结构	10
第二章 相关理论和技术概述	12
2.1 全卷积神经网络概述	12
2.2 本文相关的卷积神经网络	15
2.2.1 经典的卷积网络结构	15
2.2.2 经典的卷积网络基本模块	17
2.3 用于网络优化的损失函数	20
2.3.1 损失函数概述	20
2.3.2 本文相关的损失函数	20
2.4 计算机视觉中的注意力机制	22
2.5 常用医学图像分割评价指标	25
2.6 本章小节	26
第三章 基于边界损失的2.5D网络分割方法研究	27
3.1 方法概述	27
3.2 CT图像预处理	28
3.2.1 图像CT值截断与归一化	29
3.2.2 图像增强	30

3.2.3 图像扩充	32
3.3 基于FCN改进的2.5D全卷积神经网络设计	33
3.4 用于优化2.5D网络的边界损失函数设计	35
3.5 图像后处理	37
3.5.1 密集条件随机场	37
3.5.2 最大连通区域	39
3.6 实验结果与分析	40
3.6.1 实验环境	40
3.6.2 分析与讨论	41
3.7 本章小节	47
 第四章 基于双路注意力的编解码网络分割方法研究	 48
4.1 方法概述	48
4.2 编解码网络框架设计	49
4.3 基于双路注意力机制的编解码网络框架设计	51
4.3.1 改进的双路注意力模块设计	51
4.3.2 基于双路注意力的2D编解码网络设计	54
4.3.3 基于双路注意力的3D编解码网络设计	56
4.3.4 基于双路注意力的2.5D编解码网络设计	58
4.4 实验结果与分析	60
4.4.1 实验环境	60
4.4.2 结果与分析	60
4.5 本章小节	68
 第五章 结论与展望	 69
5.1 结论	69
5.2 展望	70
 参考文献	 72

作者在攻读硕士学位期间公开发表的论文	82
作者在攻读硕士学位期间所参与的项目	83
致谢	84

第一章 绪论

1.1 课题研究背景与意义

肝脏是人体最重要的组织器官之一，位于肺部和右锁骨的下方，在维持人体的生命活动中主要负责消化食物、储存营养、过滤血细胞、分解有毒物质和抵抗入侵的病原体。从解剖学的角度来看，正常的肝脏分为两叶：左叶和右叶，其中右叶又分为方叶和尾叶。在细胞的更新过程中，部分肝脏细胞可能出现失控增殖和扩散至身体其他部位的情况，这种异常的细胞行为被称为细胞癌变 [1]。当大量的肝部癌变细胞不受人体免疫系统控制生长时，就会形成原发性恶性肝脏肿瘤，即肝癌 [1]。肝癌是导致男性死亡的第二大癌症，也是导致女性死亡的第六大癌症。在全球范围内，每年约有75万人被诊断为肝癌，其中约有69万人不治死亡，而在确诊的患者中，男性罹患肝癌的概率是女性的两倍左右 [2]。肝癌还可以由病毒性肝炎转化发展而来。在中国，由于慢性乙型肝炎的高患病率，使得肝细胞癌(HCC)尤为常见，约占原发性肝癌的90% [3]。HCC也是中国第四大常见的恶性肿瘤，对中国人民群众的生命健康构成了严重的威胁 [3]。在临床治疗中，肝癌患者的存活率通常依赖于癌症的诊断时期，因此越早发现，越早治疗，病人的存活率也就越高 [1]。

在临床实践中，原发性恶性肝脏肿瘤一般通过血液检验和影像学检查进行诊断。常见的影像学检查技术包括X射线、计算机断层扫描(Computed Tomography, CT)、磁共振成像(Magnetic Resonance Imaging, MRI)和超声等 [4]。其中，CT扫描技术因其扫描速度快、对人体危害小和扫描图像清晰等优点被广泛应用于肝脏相关的疾病检测。具体而言，CT扫描技术利用X射线和探测器环绕人体的某个部位进行断层扫描，并使用计算机对这些扫描信号进行相关的处理，从而获得不同角度的人体矢状面、冠状面和轴向图像，最后将人体的器官、骨骼和软组织清晰地呈现出来 [5]。此外，为了增强CT扫描技术对肝脏疾病的诊断作用，医生在进行影像检查的过程中通常需要向病人的静脉中注入造影剂，以增强患者正常器官与病变区域之间的对比度，这种CT扫描方式也被称为对比度增强CT扫描 [6]。对比度增强CT扫

描通常有助于鉴别急性感染、慢性炎症、纤维化和肝硬化造成的恶性病灶 [7]。

原发性恶性肝脏肿瘤的分期通常取决于肝脏肿瘤的位置、大小和形状 [7, 8]。在以往的肝脏肿瘤诊断中，医生需要手动测量CT中肿瘤的大小和体积，然后才能结合临床知识和治疗经验判断肿瘤所处的分期。但是手工测量的方式通常具有精度不足的缺点，同时过度依赖临床医生的主观治疗经验，可能导致误诊情况的出现。随着计算机技术的快速发展，一些计算机辅助诊断系统(Computer-aided Diagnosis, CAD)能够帮助医生将肝脏肿瘤从CT扫描图像中分割出来，并对肿瘤进行定量和定性的分析，从而节省了医生大量的时间和精力 [9, 10, 11]。然而，目前大多数计算机辅助诊断系统通常是半手动的，需要在医生参与的情况下才能正常工作，分割结果难免掺杂医生的主观意识。因此研究一种能够自动检测和准确提取CT中肝脏与肿瘤的方法对肝癌的临床治疗具有重要的实际意义和价值。

近年来大量的研究人员尝试使用不同的方法实现CT中肝脏与肿瘤的自动化分割。尽管研究者们提出了各种不同的方法对CT中的肝脏与肿瘤进行自动化分割，但是肝脏与肿瘤的精确分割仍然是一个难以攻克的问题。其中，CT图像组织与器官的灰度强度过于相似且边界不够清晰，病变组织密度低且灰度分布不均是导致许多方法分割精度难以提升的重要原因 [12]。图1.1展示了两个不同患者肝脏与肿瘤CT切片图，肝脏和肿瘤染色结果图和两份CT的三维渲染效果图。从图1.1中可以看出不同患者携带的肝脏肿瘤的位置、大小和形状数量各不相同，这是造成许多自动化分割方法对肿瘤分割精度不足的主要原因。其次，肝脏和肿瘤的CT值不但跨度大，而且具有大量重叠区域，造影剂的使用也会导致CT图像中的噪声增多，这些问题进一步增加了肝脏与肿瘤的区分难度。此外，不同CT的切片间分辨率(0.45 mm - 6.0 mm)和切片内分辨率(0.45 mm - 0.98 mm)差异巨大，也使得一般的分割方法难以进行。

自动化分割CT图像中的肝脏与肿瘤具有如此多的困难，而在临床应用与医学影像领域又迫切需要一种高精度的自动化分割方法帮助医生诊断患者病情，提高工作效率，深度学习技术的发展为解决这一问题提供了契机。因此本文尝试基于多种深度学习技术来探索更优秀的肝脏和肿瘤图像分割方法，提升肝脏与肿瘤CT图像

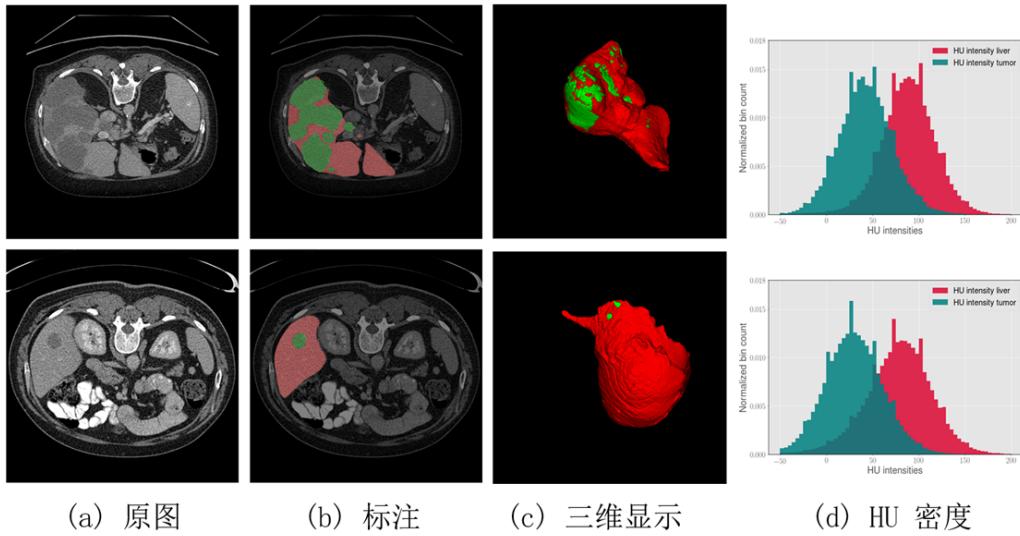


图 1.1 CT图像中的肝脏与肿瘤, (a)为肝脏与肿瘤的CT切片图像; (b)和(c)是与(a)对应的标注图像和三维渲染结果, 其中红色部分代表健康的肝脏区域, 绿色部分代表病变的肿瘤区域; (d)则是与(a)对应的CT值分布直方图, 其中红色区域表示肝脏, 绿色区域表示肿瘤, 阴影区域表示二者的CT值存在大量的重叠。

的分割精度, 为智能医疗的发展添砖加瓦。

1.2 国内外研究概况

图像分割是计算机视觉领域中一个传统而又具有挑战性的研究课题。图像分割的结果蕴含着重要的视觉信息, 工业界的许多任务往往都需要利用这些视觉信息提升工作效率, 如汽车自动驾驶 [13]、机器人导航 [14]和医学图像分割 [15]等。医学图像分割是图像分割问题的细化, 是最重要的图像分割问题之一。高效的医学图像分割方法能够突显人体器官和病理结构的变化, 从而达到辅助医生诊断病情和规划治疗方案的目的 [16]。本节将国内外的医学图像分割方法分为三类: 基于经典图像技术的分割方法、基于机器学习的分割方法和基于深度学习的分割方法, 并对这些方法进行了简要回顾和总结。

1.2.1 基于经典图像处理技术的分割方法概述

早在20世纪70年代, 一些研究人员就在不断地尝试使用经典图像处理技术对医

学图像进行分割。迄今为止，在医学图像分割领域产生了许多基于经典图像处理技术的肝脏及肿瘤分割方法。这些方法通常根据图像的色彩、灰度、空间纹理、边界等特征进行分割，如阈值分割法、边缘检测分割法、区域分割法和活动轮廓分割法等 [17]。阈值分割法的优点是算法简单，执行速度快，缺点是无法保证图像分割结果的准确性。Otsu阈值法是经典的阈值分割算法，能够基于给定的像素阈值将图像分割成前景和背景两个区域 [18]。1995年，Frank等人 [19]提出了一种最优阈值法分割肝脏与肿瘤CT图像。边缘检测算子通过检测不同分割区域的边界，从而达到分割图像的目的，常见的的边缘检测算子包括Roberts、Sobel和Canny等。上世纪90年代，Staib等人 [20]提出了一种平面曲线拟合区域边界的方法，能够利用拟合的区域边界将图像分割成不同区域。尽管基于边缘检测的分割方法能够快速准确地检测出图像的边缘，但是由于图像边缘的连续性难以确定，因此边缘检测只能检测出不连续的边界线，无法实现完整的图像分割。基于区域的图像分割算法原理是假设图像一定区域内的灰度值具有相似性，进而根据灰度值的相似性将图像分割成多个不同的区域。2008年，Wong等人 [21]提出了一种基于知识约束的二维区域生长方法来分割CT图像中的肝脏肿瘤。区域生长方法的优点是可以通过不断调整起始种子点的选取提高分割的准确度，其缺点则是手动选取合适的种子点需要耗费大量的时间和精力，因此无法完成自动化分割医学图像的任务。Snake模型是一种常见的活动轮廓模型，通过一条可变形的参数曲线和与之对应的能量函数寻找目标轮廓，从而达到分割目标的目的。2010年，Jiang等人 [22]提出了一种改进的GVF Snake模型，能够将肝脏的初始轮廓扩展到边界凹陷区域，获取精确的肝脏轮廓，进而完成肝脏肿瘤的分割任务。尽管Snake模型能够准确地捕获目标轮廓，但它对初始轮廓位置的选取非常敏感，要求初始轮廓尽可能的接近真实轮廓，然而当图像边缘模糊，目标比较复杂，其初始轮廓通常难以确定。

基于图像处理技术的医学图像分割方法种类繁多，往往只适用于特征简单的图像分割问题，同时这些方法通常过度依赖人工特征的设计与选取，在面对复杂场景的分割任务时，存在分割精度不足的缺点。这些问题导致基于经典图像处理技术的分割方法难以有效地完成肝脏与肿瘤的自动化分割任务。

1.2.2 基于机器学习的分割方法概述

随着机器学习的发展，越来越多的研究人员尝试使用机器学习方法来实现医学图像的自动化分割，如聚类分割法、支持向量机(Support Vector Machine, SVM)分类法和水平集分类法等 [17]。聚类分割方法的原理是根据特定的规则将具有相似性质的像素点分为同一类，从而将图像分割成多个不同的区域。如Wells等人 [23]提出了一种自适应聚类算法，能够从MRI图像中分割出脑肿瘤。但是聚类方法对离群点和噪声点非常敏感，面对具有大量噪声的医学图像通常存在分割精度不足的问题。SVM是一种定义在特征空间上的间隔最大线性分类器。2014年，Vorontsov等人 [24]提出了一种基于SVM分类器的分割方法，能够实现肝脏肿瘤的有效分割。与该方法类似，Kuo等人 [25]使用纹理特征向量训练SVM，并利用训练好的SVM分类器对肝脏肿瘤图像进行分类和分割。不足的是，SVM分类法对参数变化非常敏感，面对复杂的医学图像数据，难以基于固定参数实现有效地分割。水平集方法是一种借助水平集函数追踪界面的数值技术，也是一种广泛应用的医学图像分割方法。在2011年，Jimenez等人 [26]提出了一种基于自适应曲线曲率技术的三维水平集方法，实现对CT图像中肝脏与肿瘤的分割。此外，为了提升机器学习方法对医学图像的分割效果，一些研究人员尝试将多种机器学习方法的优点结合起来，提出了许多更有效的分割方法。如Le等人 [27]将快速行进算法与单层前馈神经网络(Single Hidden Layer Feedforward Neural Network, SLFN)结合起来，先使用快速行进算法生成初始的肿瘤区域，然后再使用单层前馈神经网络对区域结果进行细化分割，从而实现肝脏肿瘤的分割。Huang等人 [28]使用极限学习机(Extreme Learning Machine, ELM) [29]作为随机特征子空间集成分类器对CT图像的体素进行分类，接着在基分类器分类的结果上使用投票算法进一步优化分类的结果，从而完成肝脏肿瘤的自动检测与分割任务。相较于经典图像技术分割方法，尽管这些机器学习分割方法有较大的改进和提升，但它们的分割过程步骤繁多，当面对复杂的医学图像时，仍然难以获得有效的分割精度。

1.2.3 基于深度学习的分割方法概述

近年来，深度学习作为机器学习的一个分支领域，在计算机视觉领域发展迅速。在训练样本足够多的条件下，深度学习能够自动地学习数据特征进行图像的分类、检测和分割 [30, 31]。越来越多的研究者利用深度学习算法来解决医学图像的分割问题，在医学图像分割领域中产生了巨大的影响。这是因为深度学习网络在探索图像语义特征时对数据的噪声、模糊和对比度并不敏感，因此能够对医学图像产生优秀的分割效果。目前，在医学图像分割任务中根据使用标注数据的数量将深度学习分为监督学习、弱监督学习和无监督学习 [32]。监督学习方法能够利用标注数据训练模型；无监督学习方法在训练时不需要标注数据，但也因此增加了模型学习的难度；弱监督学习方法则是一种介于监督学习和无监督学习之间的方法，因为它的模型训练只需要少量的标注数据和部分无标注数据。

对于医学图像分割任务而言，监督学习是最流行最有效的方法，监督学习方法的优点是在准确标注的数据上训练模型，能够获得较高的分割精度。而医学图像分割任务通常又要求网络具有较高的稳定性和分割精度，这点是无监督学习和弱监督学习方法无法达到的。为此，许多优秀的深度学习网络结构被一些研究人员提了出来，如全卷积网络(Fully Convolutional Network, FCN) [33]、U-Net [34]和Deeplab [35]等。在这些网络结构中，U-Net的编码解码结构对于医学图像的分割效果最好，其编码器用于提取图像特征，解码器用于将提取到的图像特征恢复到原始图像大小并输出最终分割结果。尽管编码解码的网络结构在医学图像分割中是实用的，但它降低了模型的可解释性。在实践中，由于CT、MRI等医学图像数据大多以三维体的形式存在，使用三维卷积核可以更好地探索高维数据的空间相关性。受此启发，Cicek等人 [36]对U-Net的编码解码结构进行了扩展，提出一种直接处理三维医学图像数据的3D U-Net网络结构。但是由于计算资源的限制，3D U-Net仅包含3个下采样层，不能有效提取图像的深层特征，导致对三维医学图像的分割精度有限。此外，Milletari等人 [37]提出了类似的网络结构3D V-Net。与3D U-Net相比，3D V-Net利用残差连接结构设计了一个更深的网络(4次下采样)，残差连接可以避免梯度

消失和加速网络收敛，从而易于设计更深层次的网络结构，增强网络的特征表达能力，使得V-Net获得了更高的分割精度。但同时，这些3D网络需要训练大量参数，存在计算成本高和显存占用大等问题。

递归神经网络(Recurrent Neural Network, RNN) [38]可以用来处理序列问题，一些研究人员注意到了这点并利用RNN构建图像序列相关的时间依赖关系，进而解决医学图像的分割问题。长短时记忆网络(Long Short-Term Memory, LSTM) [39]是目前最流行的RNN之一，通过引入自循环，可以在网络中保持较长时间的梯度流动。如Alom等人 [40]提出的一种ResUNet和RNN结合的医学图像分割方法，实现了递归残差卷积层的特征积累，增强了网络在医学图像分割任务中的特征表达能力。

在使用深度学习的医学图像分割任务中，通过训练多个级联的二维或三维网络能够有效地提高分割精度。级联的网络分割模型大致可以分为三种类型：粗细分割、检测分割和混合分割 [32]。粗细分割网络框架通常使用两个级联的二维网络进行分割，其中第一个网络进行粗分割，第二个网络在粗分割的基础上实现细化分割。例如，Christ等人 [41]提出了一种用于肝脏和肝肿瘤分割的级联网络，该网络首先使用第一个FCN对肝脏进行分割，然后将第一个FCN分割的肝脏结果作为第二个FCN的输入进行肿瘤的分割。此外，Yuan等人 [42]也使用了类似的方法对肝脏和肝肿瘤图像进行分割。在此工作中，他们首先训练一个简单的卷积-反卷积神经网络(Convolutional-deconvolutional Neural Networks, CDNNs)模型(19层的FCN)，对每一份CT图像数据进行快速但粗糙的肝脏分割，然后将另一个CDNNs(29层的FCN)应用于肝脏区域，进行细粒度的肝脏分割，最后将直方图均衡化增强的肝脏分割区域作为第三个CDNNs(29层的FCN)的额外输入分割肿瘤。

检测分割方法一般先使用Region-CNN (R-CNN) [43]或You-Only-Look-Once (YOLO) [44]等网络模型定位分割目标，然后使用另一个网络基于定位的结果进行细化分割。例如，Al-Antari等人 [45]提出了一种类似的方法用于乳腺肿块的检测、分割和分类。在该工作中，作者首先使用区域检测网络YOLO对分割目标定位，接着将检测到的目标送入一个新的全分辨率卷积网络(Full Resolution Convolutional Network, FrCN)进行分割，最后使用一个更深的卷积神经网络用于识别肿瘤的良性

与恶性。

对于三维医学图像数据而言，二维卷积神经网络无法学习三维数据的空间特征信息，三维卷积神经网络往往又需要较高的计算成本和显存占用。为此，一些研究者提出了混合分割的方法，通过伪三维网络或者混合二维图像特征和三维图像特征减少模型的参数量，降低计算机显存的占用。例如，Oda等人 [46]提出了一种基于三平面的级联网络分割方法，能够有效地将腹部动脉区域从医学CT图像中分割出来。Vu等人 [47]也提出了一种伪三维网络方法，在该工作中他们将多张相邻切片作为输入以输出中心切片的预测结果。这些伪三维网络方法虽然能够从三维图像数据中分割出要提取的目标，但由于只利用了图像的局部特征信息，其分割精度的提高非常有限。与伪三维网络方法相比，混合的二维与三维级联网络更为流行。例如，Li等人 [17]提出了一种混合密集连接网络(H-DenseUNet)用于肝脏和肝脏肿瘤的分割。该方法首先利用简单的残差网络获得粗糙的肝脏分割掩膜，接着利用二维密集网络提取二维图像特征和三维密集网络提取三维图像特征，最后设计一种混合特征融合层，将二维与三维特征融合并输出最终的分割结果。虽然H-DenseUNet相对于三维网络来说降低了模型的复杂性，但是模型中的三维卷积仍含有大量需要训练的参数。

对于神经网络来说，注意力机制模块可以根据输入数据所属类别的重要性有选择地为它们赋予不同的权重。近年来，深度学习与视觉注意机制相结合的研究大多集中在利用标注图像实现网络的注意力机制，通过设计一个特殊的网络层来识别图像中的关键特征，然后让网络只关注图像中感兴趣的区域 [32]。视觉注意力机制通常分为空间注意力机制和通道注意力机制 [48]。空间注意力网络模块能够计算图像空间域内每个像素的重要性，提取图像中的关键特征信息。2018年，Jaderberg等人 [49]提出了一种基于空间注意力机制对胰腺图像分类的空间变换网络 (Spatial Transformer Network, STNet)，该网络将原始图像的空间信息转换到另一个空间，并保留图像中的关键信息，从而提升图像分类的效果。通道注意力网络模块可以实现对图像通道特征校准，有选择地增强有效的特征信息，抑制无用的特征信息。2020年，Chen等人 [50]使用SE模块 [51]实现了通道注意力网络FED-Net，能够用于肝脏肿瘤

的有效分割。

除了网络结构和网络功能块的设计之外，选择合适的损失函数也是提高网络性能的重要方法。在图像分割任务中，交叉熵是最常用的损失函数之一，该损失函数将预测的像素类别与目标真实的像素类别进行比较，进而优化网络的参数。2015年，Ronneberger 等人[34]提出一种新的类交叉熵损失函数，通过向交叉熵损失函数中添加一个距离函数能够有效提升网络对类间距离的学习能力。相似性系数是一种常用的医学图像分割性能评价指标，能够衡量预测图像和真实图像之间的重叠率。2016年，Milletari等人 [37]使用相似性系数作为损失函数优化网络的参数分割医学图像，有效地改善了数据类别不平衡造成的学习问题。虽然相似性系数损失函数能够在一定程度上解决数据类别不平衡的问题，但是面对数据类别极度不平衡的情况，相似性系数损失函数却难以起到应有的作用。而在2017年，Sudre等人 [52]提出的一种广义相似性系数损失函数可以平衡不同类别数据对损失的贡献，从而解决了在类别极度不平衡情况下网络的学习问题。

尽管目前有着各种各样的网络模型能够对肝脏和肝脏肿瘤进行自动化分割，但是这些方法仍然存在不足，如二维网络缺少探索三维数据的空间特征信息的能力，三维网络又需要训练大量的参数和占用高额的计算资源。其次，现有方法缺少损失函数对不同维度网络优化效果的研究，一些损失函数不能有效地优化网络学习医学图像中的轮廓和边界特征。但同时，研究发现注意力机制模块能够有效地对网络探索的高级抽象特征进行整合，增强网络的特征表达能力，提高深度学习方法在医学图像分割中的效率。

1.3 论文主要工作

为了解决肝脏与肿瘤自动化分割的问题，本文针对CT图像的特点，结合多种深度学习技术优化深度学习网络，以实现肝脏与肿瘤的自动化分割。本文主要工作和创新如下：

(1) 基于FCN提出了一种新的2.5D网络结构，在减少网络参数量和计算资源消耗的同时，有效地探索三维数据中的空间特征信息，提升肝脏与肿瘤CT图像分割

的准确度。

(2) 针对医学图像的特点，设计了一种新的边界损失函数，它融合了图像的距离、面积和边界特征信息，能够有效地优化深度学习网络学习更多的图像边界和轮廓特征信息。

(3) 为了进一步优化网络结构，本文结合双路注意力模块、密集网络块、残差网络块和双路径网络块设计了一种基于双路注意力的编解码网络分割框架，包含九种不同的网络结构，能够有效地完成肝脏与肿瘤的自动化分割任务。

(4) 针对网络预测的结果，本文提出了一种肝脏与肝肿瘤后处理方法，用于进一步优化肝脏与肿瘤的分割结果，减少深度学习网络分割结果中的假正例和假反例，提升分割的准确度。

1.4 论文组织结构

本论文是以作者攻读硕士学位期间承担的课题工作为基础，针对肝脏与肝肿瘤图像的特点，研究多种深度学习技术结合的肝脏与肝肿瘤CT图像自动化分割方法。本文的组织结构如下：

第二章：与本研究相关的理论和技术概况。本章首先简单描述了全卷积神经网络的发展历史以及相关组成结构。接着详细介绍与本文相关的两种经典卷积网络结构：FCN和U-Net，同时阐述了三种用于网络构建的基础模块设计原理：残差块、密集块和双路径块。然后对医学图像分割任务中常见的损失函数进行了简单划分，并详细说明了与本文工作相关的四种损失函数：交叉熵损失函数、相似性系数损失函数、联合损失函数和轮廓损失函数。此外，还对计算视觉中的注意力机制进行简单介绍，并详细描述了本文用到的双路注意力模块实现原理。最后对本文实验用到的评估度量给出了相关介绍。

第三章：基于边界损失函数的2.5D网络分割方法研究。本章首先介绍了基于边界损失函数的2.5D网络分割方法流程，接着详细阐述了本文针对肝脏与肿瘤CT图像特点提出的预处理方法，然后重点描述了本文设计的2.5D网络结构与边界损失函数实现原理。此外，还对本文提出的用于网络预测结果优化的后处理方法进行了

详细说明。最后详细介绍了本章的实验环境和数据集，对本章的方法进行了实验验证，同时对实验结果进行了说明和分析。

第四章：基于双路注意力的编解码网络分割方法研究。本章首先介绍了整个双路注意力编解码网络框架的分割流程。接着阐述了无注意力机制的编解码网络框架设计原理，然后详细描述基于双路注意力的编解码网络框架，并从2D网络、2.5D网络和3D网络的角度分别阐述了9种不同网络结构的设计思想。最后对本文提出的双路注意力编解码网络进行了实验验证，并对实验结果进行了分析和讨论。

第五章：结论与展望。本章对本文所提出的方法做出了总结和展望。

第二章 相关理论和技术概述

近年来深度学习发展迅速，在计算机视觉、自然语言处理、自动驾驶和情感分析等一系列领域取得了突破性的成就。各种深度学习技术的发展为医学图像处理开辟了新的道路。本文的研究目的是通过吸收现有技术方法的优点，探索更优秀的深度学习方法以实现自动化分割CT图像中的肝脏与肿瘤。因此，本章回顾了与本文工作相关的深度学习理论和技术，为后续章节的方法和实验介绍提供铺垫。

2.1 全卷积神经网络概述

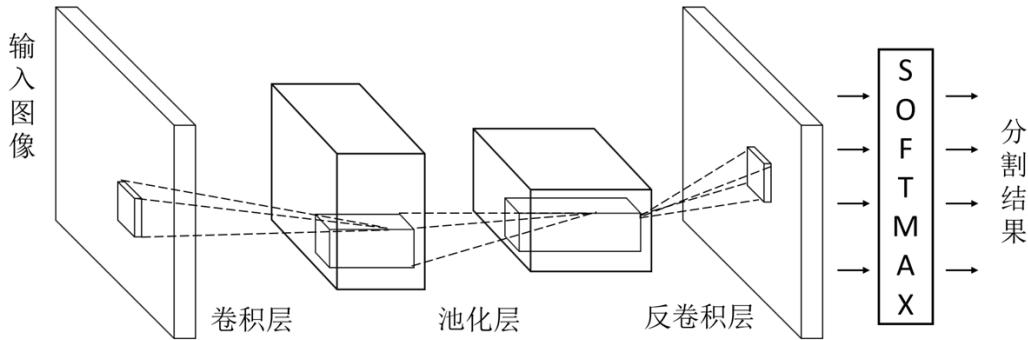


图 2.1 全卷积神经网络结构示意图

计算机领域的人工神经网络是通过模仿动物神经网络系统中的神经元连接演变而来。全卷积神经网络则是对人工神经网络的进一步延伸，通过卷积操作从图像中提取更深层的信息特征。与人工神经网络相比，全卷积神经网络的层与层之间借助卷积核的存在能够实现网络内部的参数共享，从而极大地减少需要训练的网络参数。此外，相比于将每一层的神经元都与邻近层神经元相连接的人工神经网络，全卷积神经网络在最后一层不需要将图像展开成向量，能够保留更多的图像空间特征信息。如图2.1所示，全卷积神经网络通常包含卷积层、池化层、激活层以及反卷积层。

卷积层通过卷积核对输入的数据进行卷积运算提取图像的局部特征 [53]。卷积核也被称作卷积滤波器，卷积滤波器是一种数值矩阵，卷积层则是由这样一组权值

不同的卷积滤波器组合而成。如图2.2所示，(a)是一个 5×5 大小的输入图像特征，(b)是一个 3×3 大小的卷积核，假设卷积步长为1，将卷积核(b)按照步长不断地在输入图像特征(a)上滑动计算，则得到 3×3 大小的输出图像特征(c)。卷积层具有两个重要的特性：局部连接和权值共享。在全连接网络中，网络每一层的节点都与邻近层的所有结点相连接，事实上图像中只有相邻的像素点具有较强的相关性，因此只需要学习图像的局部特征，之后再对低维的局部图像特征进行整合即可学习到更多抽象的高维图像特征。而卷积操作恰恰能够通过局部连接完成对图像局部特征的学习，同时还能保留图像原有的空间信息。卷积操作的另一个优点是卷积过程权值共享，当卷积核在图像特征上进行滑动时，其权值不变，因此能够在图像不同的位置检测相同的特征，从而达到减少网络需要训练的参数量，起到提高网络训练速度的效果。

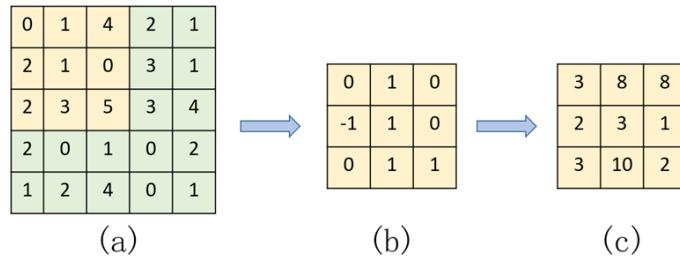


图 2.2 卷积过程示意图

池化层主要作用是对输入图像特征降维和增大图像卷积的感受野，此外对图像特征进行池化还能减少图像特征的冗余 [53]。常见的池化方法有最大池化和平均池化。由于池化降维过程会损失图像的特征质量，因此通常将池化步长设置为2，将池化窗口大小设置为 2×2 。如图2.3所示，(a)是输入的图像特征；(b)是输出的池化结果，最大池化过程是将池化窗口内的最大值填入输出特征图的对应位置，平均池化过程则是将池化窗口内所有值的平均值填入输出特征图的对应位置。

激活层的主要作用是利用激活函数对卷积层输出的线性特征进行非线性映射，从而增加网络的非线性拟合能力 [54]。常见的激活函数有Sigmoid、Tanh、Relu、Softmax等。如图2.4(a)和(b)所示，Sigmoid和Tanh激活函数都是将卷积层的特征输出映射到 $[0, 1]$ 区间内，其梯度函数曲线中间大，两端趋于0，因此在反向传播过

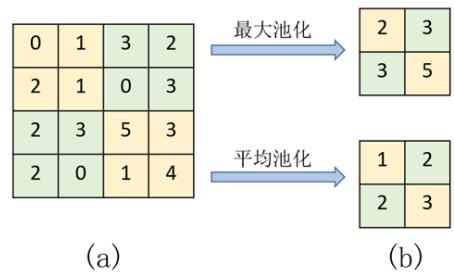


图 2.3 池化过程示意图

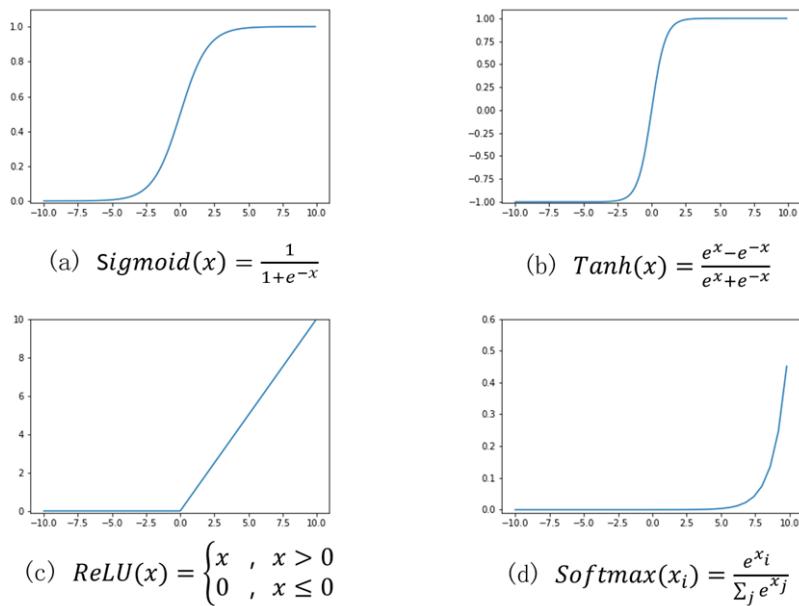


图 2.4 激活函数

程中容易造成梯度消失。如图2.4(c)所示, *ReLU*激活函数具有非饱和性, 由两段线性函数组合而成, 由于梯度简单, 在反向传播过程中梯度为0或者为1, 因此不会出现梯度消失的情况。但是*ReLU*函数在原点的不连续性使得网络训练具有不稳定性, 容易造成部分神经元提前失活。如图2.4(d)所示, *Softmax*激活函数通常用于卷积神经网络的最后一层输出分类结果。

反卷积层的作用类似于上采样和上池化，它主要用来恢复图像特征的分辨率，从而保证输出特征与输入图像具有相同的尺度大小。在卷积神经网络中反卷积又被称为转置卷积，是卷积的逆过程 [55, 56]。如图2.5所示，在反卷积过程中利用反卷积核在图像特征上滑动，能够输出分辨率更高的图像特征。

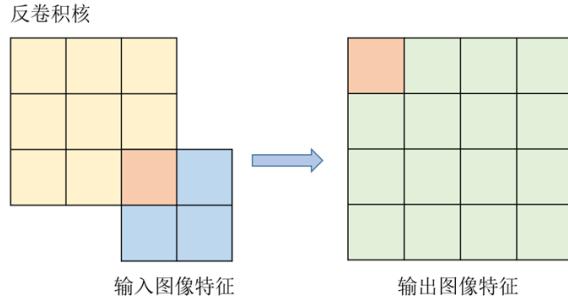


图 2.5 反卷积过程示意图

2.2 本文相关的卷积神经网络

近年来，随着深度学习的发展，计算机视觉领域出现了许多优秀的语义分割网络结构，如FCN [57]、U-Net [34]、SegNet [58]、Mask R-CNN [43]和PSPNet [59]等。与这些网络结构同期出现的还有各种用于构建网络的基本模块，如Inception Block [60, 61, 62]、残差块(Residual Block) [63]、密集块(Dense Block) [64]、残差密集块(Residual Dense Block) [65]和双路径块(Dual Path Block) [66]等。本节详细介绍了与本文工作相关的两种卷积网络结构：FCN和U-Net，以及三种用于构建网络结构的基本模块：残差块、密集块和双路径块。

2.2.1 经典的卷积网络结构

2014年出现的FCN [57]和U-Net [34]是众多全卷积神经网络中经典的两个网络，为后续许多网络结构设计提供了借鉴思路。如图2.6所示，FCN作为图像语义分割的深度学习方法，首次使用卷积层取代了网络末端的全连接层，从而使得网络能够输出与输入图像的每个像素点对应的类别预测值。由于卷积神经网络的卷积与池化操作会不断缩小输出特征的尺寸。因此FCN作者在网络内部引入反卷积操作增大特征尺寸，使得预测概率特征图与输入图像具有相同的高度和宽度，从而实现图像的像素级语义分割。尽管FCN能够实现图像语义分割，但由于在恢复特征尺寸过程中，直接对卷积和池化输出的图像特征运用反卷积操作，使得最终的分割结果相对粗糙。为此，FCN作者在网络的内部使用了跳跃融合操作对网络进行优化以探索更

多有效的特征信息，实现图像的精细分割。跳跃融合操作是一种特征对应位置相加的方法，通过上采样将不同层次的池化输出特征统一尺寸，并进行相加，从而达到增强特征位置信息的目的。尽管通过跳跃融合操作能够一定程度上改善网络下采样和上采样过程中丢失位置信息的问题，但FCN没有充分考虑像素与像素之间的联系，缺少对图像特征信息的整合能力，从而影响了最终分割结果的准确性。

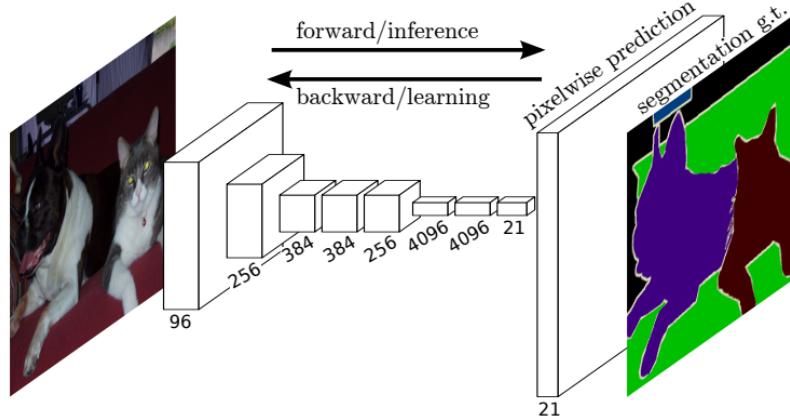


图 2.6 全卷积网络结构示意图 [57]

U-Net [34]是继FCN之后出现的另一个经典的全卷积神经网络，其特点是编码与解码的网络部分具有对称性。如图2.7所示，U-Net网络由两部分组成：探索特征上下文信息的编码网络和恢复特征空间位置信息和尺寸的解码网络。此外，U-Net在设计时使用了跳跃连接整合了浅层卷积特征与深层卷积特征，使得网络输出的最终结果由浅层特征与深层特征共同决定。现实中人体的组织器官虽然复杂，但其组成结构却是相对固定的。在临床实践中，CT扫描成像时往往只关注人体的某一部分，相比于自然图像，医学CT图像所包含的内容是简单固定的 [5]。对医学图像分割而言，网络探索的浅层特征主要是纹理特征，能够为图像像素点提供分类依据，而网络探索的深层抽象特征则能够为图像的精确分割提供位置信息。U-Net的编解码网络结构恰恰能够有效的整合来自浅层的图像纹理特征和深层的图像抽象特征信息，从而适用于医学图像分割。但由于U-Net构建网络的基本模块是简单的卷积块，尽管U-Net使用跳跃连接整合了浅层特征和深层特征信息，但仍然存在一定的梯度消失问题，限制了网络深度的增加。

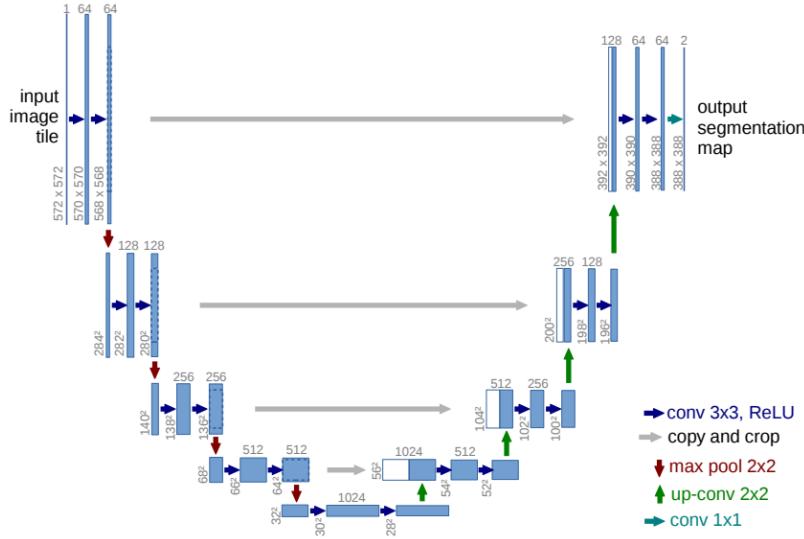


图 2.7 U-Net网络结构示意图 [34]

2.2.2 经典的卷积网络基本模块

在设计深度卷积神经网络时，通常认为网络层次设计的越深其效果也就越好。但事实是，当网络层数达到一定深度后，深度卷积神经网络会面临两个问题：梯度弥散(爆炸)和网络退化 [67]。梯度弥散(爆炸)主要是由于反向传播过程中连续多层梯度减小(增大)的情况导致的。网络退化则是当网络层数的不断加深，会出现浅层网络比深层网络效果更好的现象。为了解决这些问题，一些研究人员提出了使用不同的基本模块构建网络结构以提升模型性能，如残差块(Residual Block) [63]、密集块(Dense Block) [64]和双路径块(Dual Path Block) [66]等。

残差网络 [63]由何凯明团队在2015年提出。在残差网络中，通过对卷积的输出特征使用跳跃连接能够有效地避免梯度消失和网络退化的现象，跳跃连接也被称为残差连接，使用跳跃连接的卷积结构也被称为残差块。假设网络的输入特征为 x ，经过多次卷积的输出特征表示为 $F(x)$ ，于是可以定义 $H(x) = F(x) + x$ ，此时残差映射可以表示为 $F(x) = H(x) - x$ ，其中 $F(\cdot)$ 和 $H(\cdot)$ 表示非线性转换函数。在残差网络中通过学习残差 $F(x) = 0$ ，而不是直接学习 $H(x) = x$ 能够提升网络学习的效率。此外，整个网络需要训练的参数量也不会增加，还能有效避免梯度消失和网络退化的现象。

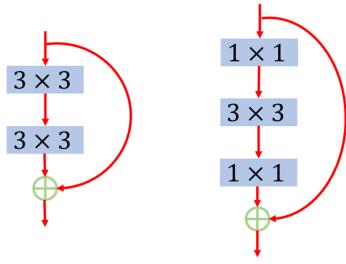


图 2.8 残差块结构示意图

图2.8描述了残差块的两种不同实现方式，左侧的残差块通过连续使用两次 3×3 的卷积再进行相应的残差连接；右侧的残差块则是先使用一次 1×1 卷积对输入特征进行降维，再使用一次 3×3 卷积探索特征，然后又使用一次 1×1 卷积对特征升维，最后对输出的结果做相应的残差连接。相比于左侧的残差块，右侧的残差块能够有效地减少参数量和计算量。

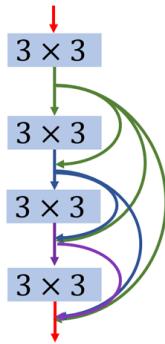


图 2.9 密集块结构示意图

残差网络的效果是如此之好，得到了许多研究者的认同。2017年，受到残差网络的启发，黄高等人提出了一种密集连接网络 [64]。在密集连接网络结构中，每一层的输出都将作为下一层额外的输入。假设 x_{l-1} 是第 $l-1$ 层的输出， $G(\cdot)$ 表示一种非线性转换函数，在一般的卷积神经网络中，第 l 层的输出可以表示为 $x_l = G(x_{l-1})$ 。在上述的残差网络中，由于第 l 层的输出会加上来自上一层的输入，因此其可以表示为 $x_l = G(x_{l-1}) + x_{l-1}$ 。而在密集连接网络中，由于每一层的输出都作为下一层的额外输入，因此第 l 层的输出表示为： $x_l = G([x_0, x_1, \dots, x_{l-1}])$ 。值得注意的是，密集连接网络在传递特征时并没有对特征进行求和，而是使用更为直接的特征拼接

操作。这种拼接的方式在网络不断变深的过程中会导致特征的通道数大幅增加。为此，作者采用了一种折中的方法，只在少量卷积块间进行特征的传递，从而实现一种密集连接网络块，也称作密集块，如图2.9所示。

从高阶递归神经网络的角度分析来看，密集网络是一种特殊的高阶递归神经网络，而残差网络则是一种简化的密集网络。残差网络注重于特征的复用，但缺少不断挖掘新特征的能力，密度连接网络更注重于特征的发掘，但同时也会产生许多特征的冗余。因此，颜水成等人 [66]结合残差连接与密集连接的优点提出了一种新的双路径网络，从高阶递归神经网络的角度详细分析了残差网络和密集网络之间的联系以及它们效果优异的原因。

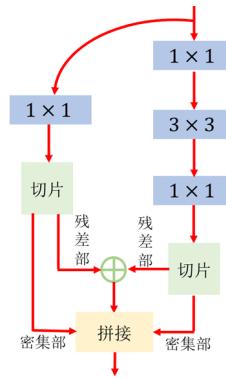


图 2.10 双路径块结构示意图

双路径网络由大量的双路径块构成，双路径块如图2.10所示。在初始阶段，输入图像特征被同时送入两条不同的通道进行不同的操作，左侧的通道主要用来保存来自上一阶段的输入特征，右侧的通道则是用来探索新的特征。在每一条通道中，图像特征会经过不同的卷积，并对各自的卷积输出特征进行切片，将切片后的特征分为残差部和密集部，残差部进行相应的残差连接，密集部则与残差连接的结果进行拼接操作，拼接的结果将作为下一层网络的输入。

2.3 用于网络优化的损失函数

2.3.1 损失函数概述

在训练深度卷积神经网络的过程中，损失函数能够用来评估真实值与预测值之间的差异，优化网络参数，提升网络拟合能力。常见的损失函数可以分为四类 [68]：基于分布的损失函数、基于区域的损失函数、基于边界的损失函数以及基于联合的损失函数。基于分布的损失函数通过最小化两个数据分布之间的差异来优化网络参数，常见的有交叉熵损失函数 [69]、TopK损失函数 [70]和Focal损失函数 [71]等。交叉熵损失的概念来自香农信息论，TopK损失函数和Focal损失函数则是交叉熵损失函数的变体，用于解决困难样本的分类问题。基于区域的损失函数通过最小化真实值与预测值之间差异性来优化网络模型，常见的有相似性系数损失函数 [37]、IoU损失函数 [72]、Tversky损失函数 [73]和广义相似性系数损失 [52]等。IoU损失由评估度量IoU变化而来，Tversky损失和广义相似性系数损失则是由相似性系数损失演变而来，其中Tversky损失主要用于解决召回率和准确率之间的平衡问题，广义相似性系数损失则是用于解决相似性系数损失的多类别扩展问题。基于边界的损失函数通过最小化标注图像与预测图像之间的距离来优化网络，常见的有哈夫曼距离损失函数 [74]和轮廓损失函数 [75]。哈夫曼距离损失函数由评估度量哈夫曼距离变化而来，轮廓损失函数则是由主动轮廓模型演变而来，主要用于解决医学图像分割问题。联合损失函数则是通过对上述损失组合加权实现的 [76]，将多种损失函数相结合用于优化网络往往能够获得更好的优化效果，如相似性系数损失函数与TopK损失函数相结合，相似性系数损失函数与交叉熵损失函数相结合。

2.3.2 本文相关的损失函数

交叉熵的概念起源于香农信息论，常用于衡量两个概率分布之间的差异性，也可以作为损失函数优化深度学习网络的参数 [69]。为了方便表示交叉熵损失函数，本文以二值交叉熵损失函数为例。假设 y 为图像的真实值， \hat{y} 是图像的预测值，二值

交叉熵损失函数则可以表示为：

$$L_b(y, \hat{y}) = -[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})] \quad (2.1)$$

一般情况下，交叉熵损失函数可以优化网络解决大多数的分割问题，但当分割目标存在类别不平衡的情况时，交叉熵则难以优化网络对样本进行有效的分类。

相似性系数是一种集合相似性度量函数，常用于计算两个样本的相似性，还可以用作衡量模型分割性能的指标。在2016年，Milletari等人 [37]证明相似性系数可以作为损失函数来优化网络参数，并比一般的损失函数更适用于医学图像分割任务。因为医学图像中正常器官与病灶类别占比差异巨大，而相似性系数损失函数恰恰能够有效地解决分割问题中类别不均衡的情况。具体而言，相似性系数损失函数通过衡量标注图像与预测图像之间重叠度的大小来优化网络，其取值范围为[0, 1]，其中1表示二者完全重叠，0表示完全不重叠，其他值表示二者部分重叠。为了表示相似性系数，假设标注图像和预测图像分别为 $T, P \in [0, 1]$ ，0和1分别表示背景像素和前景像素取值， n 表示图像像素空间 N 的索引，相似性系数则可以表示为：

$$d(T, P) = 2 \times \frac{\sum_{n=1}^N (T_n + P_n)}{\sum_{n=1}^N (T_n + P_n)} \quad (2.2)$$

其中 T_n, P_n 表示像素索引 n 处的像素值，此时，相似性系数损失函数则表示为：

$$L_d(T, P) = 1 - d(T, P) \quad (2.3)$$

尽管相似性系数损失函数能够解决分割目标类别不平衡的问题，但它缺少对图像轮廓和边界特征的探索，因此对网络的优化效果有限。

轮廓损失函数由主动轮廓模型演变而来。2019年，Chen等人 [75]将面积和尺度信息整合起来作为一种轮廓损失函数，能够使网络学习到更多的图像轮廓特征，从而提升网络对图像的分割效果。轮廓损失函数可以表示为：

$$L_a = l + \lambda \cdot r \quad (2.4)$$

其中 l 是轮廓的长度， r 是轮廓的内部面积， λ 为轮廓内部面积 r 的权重。与相似性系数损失函数相比，轮廓损失函数考虑了分割目标的轮廓特征，能够探索部分图像轮

廓特征。但轮廓损失对分割目标边界特征的探索是不充分的，因此在医学图像分割任务中并非最优的选择。

联合损失函数通过吸取不同损失函数的优点，将它们组合在一起，从而产生更强大的网络优化效果 [76]。如相似性系数损失函数与交叉熵损失函数进行加权相加可以获得一种新的联合损失函数，其可以表示为：

$$L_c = \omega_1 \cdot L_b(T, P) + \omega_2 \cdot L_d(T, P) \quad (2.5)$$

其中 ω_1 , ω_2 分别是交叉熵损失函数和相似性系数损失函数的权重。

2.4 计算机视觉中的注意力机制

深度学习中的注意力机制模拟了人类视觉能够选择性处理信息的特点，使网络模型能够关注更重要的特征信息并忽略无关的特征信息，是一种增强网络特征表达能力的优秀方法 [77]。由于注意力机制缺少严格的数学定义，因此在实际应用中常作为一个模块融入深度学习网络中。注意力机制根据实现方式不同可以分为：硬注意力和软注意力 [78, 79]。硬注意力是一种不可微的注意力机制，它主要关注图像像素点与像素点之间的联系，通常基于强化学习来实现硬注意力机制。软注意力则是一种可微的注意力机制，它主要关注空间特征和通道特征上的依赖关系，通过梯度下降的方式可以实现软注意力机制。此外，软注意力根据关注领域的不同还可以划分为：空间域注意力，通道域注意力，混合域注意力和时间域注意力等 [48]。

2018年，Fu等人 [80]提出一种新的双路注意力模块，它属于软注意力机制，能够捕获图像空间维度和通道维度上的特征依赖信息。如图2.11所示，双路注意力模块包含了两个子注意力模块：位置注意力模块和通道注意力模块。双路注意力模块可以轻易地融入卷积神经网络中，将来自卷积神经网络提取的特征分别送入位置注意力模块和通道注意力模块，用于探索更多的特征相关依赖信息，之后再将两个子注意力机制模块探索的特征融合输出，从而达到提升卷积神经网络特征表达能力的效果。

双路注意力模块中的位置注意力模块能够捕捉更多的图像局部特征信息，增

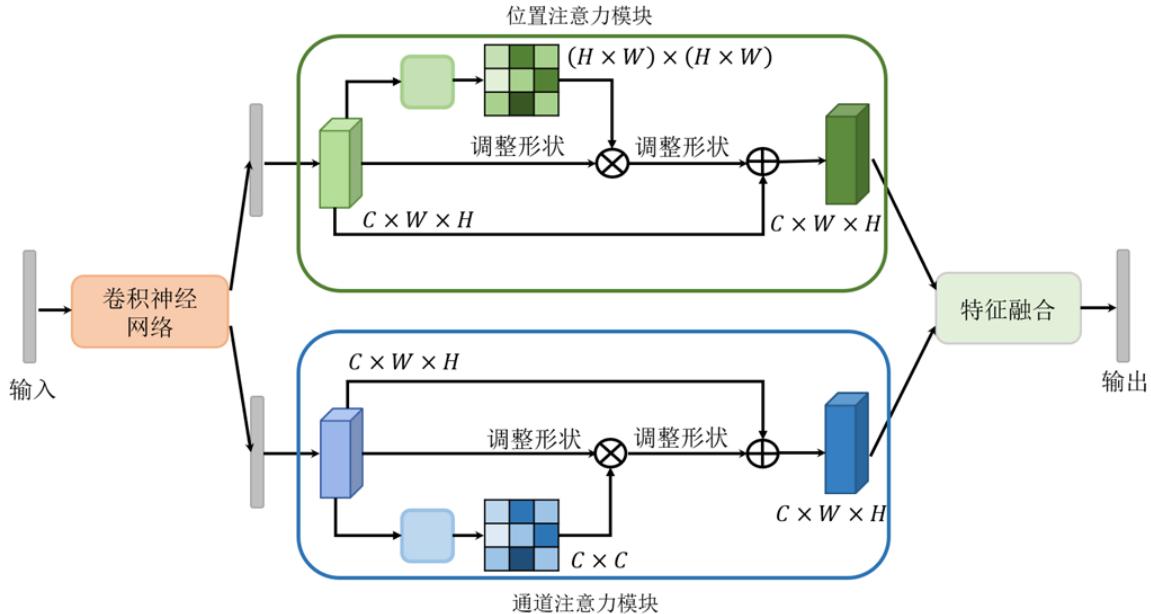


图 2.11 双路注意力模块示意图 [80]

强卷积神经网络对局部特征的表达能力。文献 [80] 中提到的位置注意力模块，如图 2.12 所示， \mathbf{A} 是来自卷积神经网络的局部特征，满足 $\mathbf{A} \in \mathbb{R}^{C \times H \times W}$ ，在位置注意力机制模块中，首先对局部特征 \mathbf{A} 做了三次不同的卷积，从而获得了局部特征 \mathbf{B} 、 \mathbf{C} 和 \mathbf{D} ，满足 $\{\mathbf{B}, \mathbf{C}, \mathbf{D}\} \in \mathbb{R}^{C \times H \times W}$ 。接着调整 \mathbf{B} 和 \mathbf{C} 的矩阵形状并获得特征矩阵 \mathbf{E} 和 \mathbf{F} ，满足 $\{\mathbf{E}, \mathbf{F}\} \in \mathbb{R}^{C \times N}$ ， $N = H \times W$ ，而后对 \mathbf{E} 转置以获得特征矩阵 \mathbf{G} ，其中 $\mathbf{G} \in \mathbb{R}^{N \times C}$ 。然后对 \mathbf{G} 和 \mathbf{F} 做矩阵乘并得到一个 $N \times N$ 的特征矩阵，对该矩阵应用 *Softmax* 操作即可得到位置注意力矩阵 \mathbf{S} ， $\mathbf{S} \in \mathbb{R}^{N \times N}$ ：

$$s_{ji} = \frac{\exp(\mathbf{G}_i \cdot \mathbf{F}_j)}{\sum_{i=1}^N \exp(\mathbf{G}_i \cdot \mathbf{F}_j)} \quad (2.6)$$

其中 \mathbf{G}_i 表示特征矩阵 \mathbf{G} 的第 i 行， \mathbf{F}_j 表示特征矩阵 \mathbf{F} 的第 j 列， s_{ji} 为位置注意力矩阵 \mathbf{S} 第 i 行、第 j 列的特征值。 s_{ji} 能够衡量第 j 个位置的特征对第 i 个位置的特征的影响，两个位置的特征越相似，其相关性也就越强。同时，调整局部特征 \mathbf{D} 的矩阵形状并生成一个 $C \times N$ 的特征矩阵，然后将该特征矩阵与位置注意力矩阵 \mathbf{S} 做矩阵乘则能够获得与局部特征矩阵 \mathbf{A} 具有相同尺寸的位置注意力加权矩阵 \mathbf{H} 。最后，将位置注意力加权矩阵 \mathbf{H} 与局部特征 \mathbf{A} 相加，从而获得了经过位置注意力矩阵加权的特征矩阵 \mathbf{I} ，满足 $\mathbf{I} \in \mathbb{R}^{C \times H \times W}$ 。特征矩阵 \mathbf{I} 继承了来自 \mathbf{A} 的全局上下文信息，同时又拥有

来自位置注意力矩阵针对位置加权的局部特征信息，从而提升了特征类间的紧凑性和语义的一致性。

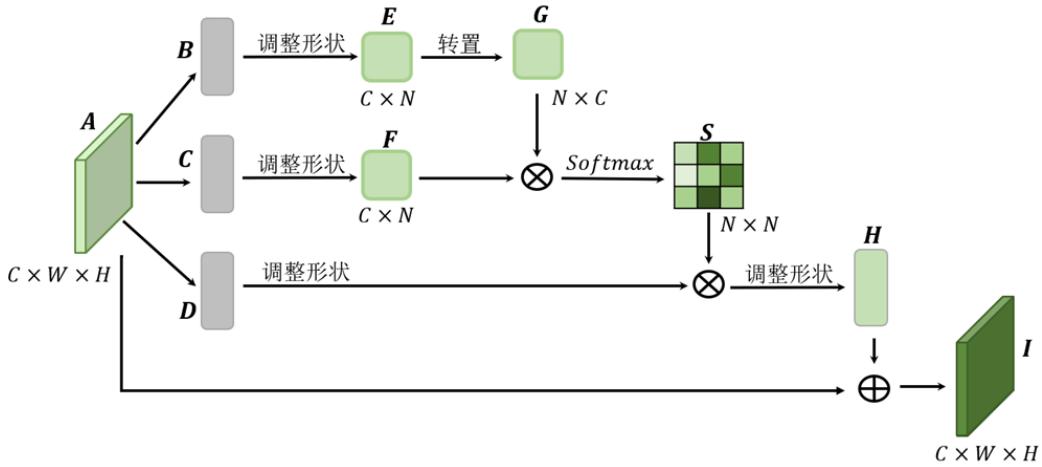


图 2.12 位置注意力模块示意图 [80]

双路注意力模块中的通道注意力模块能够捕捉通道特征之间的相关性和依赖性，增强网络特征表达的效果。文献 [80]中提到的通道注意力模块，如图2.13所示，与位置注意力模块不同，通道注意力模块利用卷积神经网络的局部特征 \mathbf{A} 直接计算通道注意力加权矩阵 \mathbf{X} ，其中 $\mathbf{A} \in \mathbb{R}^{C \times H \times W}$ ， $\mathbf{X} \in \mathbb{R}^{C \times C}$ 。具体而言，首先对局部特征 \mathbf{A} 调整形状以获得特征矩阵 \mathbf{P} ，满足 $\mathbf{P} \in \mathbb{R}^{C \times N}$ ，再对局部特征 \mathbf{A} 调整形状并转置以获得特征矩阵 \mathbf{Q} ，满足 $\mathbf{Q} \in \mathbb{R}^{N \times C}$ ，而后对 \mathbf{P} 和 \mathbf{Q} 做矩阵乘，并对矩阵乘的结果应用 $Softmax$ 操作，从而能够获得通道注意力矩阵 \mathbf{X} ：

$$x_{ji} = \frac{\exp(P_i \cdot Q_j)}{\sum_{i=1}^C \exp(P_i \cdot Q_j)} \quad (2.7)$$

其中 P_i 表示特征矩阵 \mathbf{P} 的第*i*行， Q_j 表示特征矩阵 \mathbf{Q} 的第*j*列， x_{ji} 为位置注意力矩阵 \mathbf{X} 第*j*行、第*i*列的特征值。 x_{ji} 能够衡量第*j*个通道的特征对第*i*个通道的特征的影响，两个通道的特征越相似，其依赖性也就越强。之后将局部特征矩阵 \mathbf{A} 调整为 $C \times N$ 的特征矩阵，并将该 $C \times N$ 的矩阵与通道注意力矩阵 \mathbf{X} 做矩阵乘操作，从而获得大小为 $C \times H \times W$ 的矩阵，最后将该 $C \times H \times W$ 的矩阵与局部特征矩阵 \mathbf{A} 相加，从而获得了经过通道注意力矩阵加权的特征矩阵 \mathbf{I} ，其中 $\mathbf{I} \in \mathbb{R}^{C \times H \times W}$ 。特征矩阵 \mathbf{I} 的每个通道特征继承了特征矩阵 \mathbf{A} 的通道上下文信息和经过通道注意力矩阵加权的通道特

征信息，增强了通道特征之间的语义依赖关系，有效地提升了网络对不同特征的辨识能力。

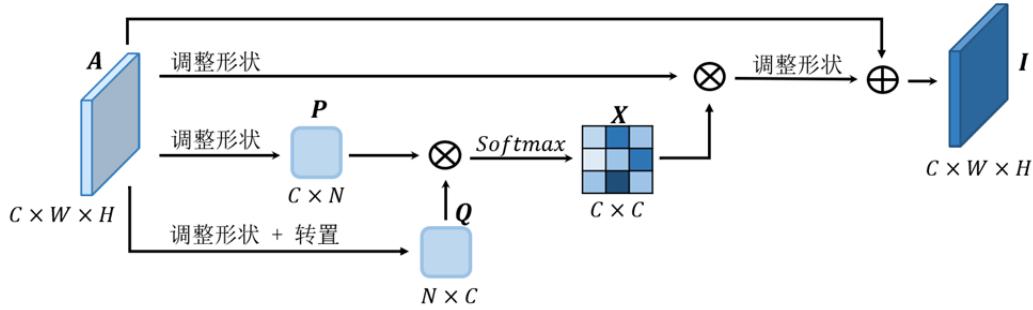


图 2.13 通道注意力模块示意图 [80]

2.5 常用医学图像分割评价指标

由于医学图像是三维数据，准确率和召回率等常见的评价指标并不能很好地评估模型分割性能的优劣。为此，本文使用了LiTS挑战赛的评价标准来评估模型分割的效果。LiTS挑战赛的主要评价指标有：Dice Per Case(DSC)、Dice Global(DG)、Volumetric Overlap Error(VOE)、Root-Mean-Square Error(RSME)、Relative Voxel Difference(RVD) 和Average Symmetric Surface Distance(ASD) [81]。假设A是标注图像，B是预测图像，则这些评价指标的表示和相关描述如下：

(1) *Dice*:

$$Dice(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (2.8)$$

*Dice*分数的本质是F1分数 [81]，其取值范围在[0,1]之间，其中最大值1表示分割表现最优。

(2) *VOE*:

$$VOE(A, B) = 1 - \frac{|A \cap B|}{|A \cup B|} \quad (2.9)$$

*VOE*由*Jaccard*系数变化而来，即 $VOE(A, B) = 1 - Jaccard(A, B)$ ，常用来衡

量目标的体积重叠误差。

(3) *RSME*:

$$RMSE(A, B) = \sqrt{\frac{1}{N} \sum_{i,j}^N (A_{i,j} - B_{i,j})^2} \quad (2.10)$$

*RMSE*是均方根误差，即预测值和真实值均方误差的平方根。其中*i, j*是像素点坐标，*N*是像素点总数。

(4) *RVD*:

$$RVD(A, B) = \frac{|B| - |A|}{|A|} \quad (2.11)$$

*RVD*是一种相对性度量指标，它能够衡量体素间相对误差。

(5) *ASD*:

假设*S(A)*是*A*表面体素点，即三维空间中像素点的集合，并将任意体素点*v*到*S(A)*最短距离定义为：

$$d(v, S(A)) = \arg \min_{s_A \in S_B} \|v - s_A\| \quad (2.12)$$

其中 $\|\cdot\|$ 表示欧氏距离，因此对称位置的平均表面距离*ASD*可以表示为：

$$ASD(A, B) = \frac{1}{|S(A)| + |S(B)|} \left(\sum_{s_A \in S(A)} d(s_A, (S_B)) + \sum_{s_B \in S(B)} d(s_B, (S_A)) \right) \quad (2.13)$$

2.6 本章小节

本章首先介绍了卷积神经网络的发展历史以及全卷积神经网络的组成结构。接着简要回顾了与本文工作相关的两种经典网络结构：FCN和U-Net，并详细阐述了本文用到的三种网络组成模块的实现原理：残差块、密集块以及双路径块。然后对用于网络优化的损失函数进行了简要的回顾和划分，并详细介绍了本文用到的四种损失函数：交叉熵损失函数、相似性系数损失函数、轮廓损失函数和联合损失函数。之后对计算机视觉中的注意力机制进行了简单的介绍，并对与本文工作相关的双路径注意力模块的原理做了详细的描述。最后对本文所使用的医学图像评价指标进行了简单的介绍。

第三章 基于边界损失的2.5D网络分割方法研究

近年来，深度学习的热潮席卷了各个学科领域，FCN和U-Net等网络不断地被用于分割医学图像中的肝脏与肿瘤。然而这些网络都是基于2D的网络结构，面对3D的肝脏与肿瘤CT图像数据，缺少探索三维空间特征信息的能力。此外，医学图像内部的肝脏与肿瘤往往具有一定的轮廓和边界，损失函数本该优化网络参数学习这些重要的特征，提升网络的分割效果，但现有的损失函数在设计时却缺少对这些重要特征信息的考虑，难以充分优化网络获取更高的分割准确度。为此，本文基于FCN设计了一种2.5D网络结构，能够有效地探索CT图像中的三维空间上下文信息。同时，本文针对CT图像中肝脏与肿瘤的特点设计了一种边界损失函数用于2.5D网络的优化，能够有效提升网络对医学图像中轮廓和边界特征的探索，提升分割精度。本章所述方法已在期刊*Algorithms*上发表。

3.1 方法概述

在人体的腹部扫描CT中，分布着肝、肾、胰腺以及脾胃等器官。分割肝脏与肿瘤CT图像的目的是将肝脏以及肝部的肿瘤从图像中提取出来，并将分割结果定量地呈现给医生，从而让医生能够更方便地诊断患者的病情。在以往的深度学习分割方法中 [82]，通常先将三维的肝脏与肿瘤CT图像分成一张张的二维切片图像，再输入深度卷积网络获得肝脏与肿瘤的分割结果。这种分割方法是简单直观的，但却忽视了一个事实，即肝肿瘤必然位于肝脏之中。因此，一个更有效的分割思路是先进行肝脏分割，在完成肝脏分割之后，利用肝脏分割结果作为掩膜覆盖在原图上以去除非肝部的组织结构，再进行肝肿瘤的分割，从而能够减少肝肿瘤分割的假正例，提升网络对CT图像中肿瘤的分割精度。先分割较大的目标区域，再分割目标区域中感兴趣区域的分步分割策略在深度学习中也常被称作级联分割 [81]。

本文基于级联分割的策略提出了一种基于FCN的2.5D网络分割方法，该方法流程图3.1所示。具体而言，在第一阶段肝脏分割任务中，本文先对原始数据进行预处理，以增强肝脏与肿瘤CT图像的清晰度，提升网络的学习效果。接着本文先使

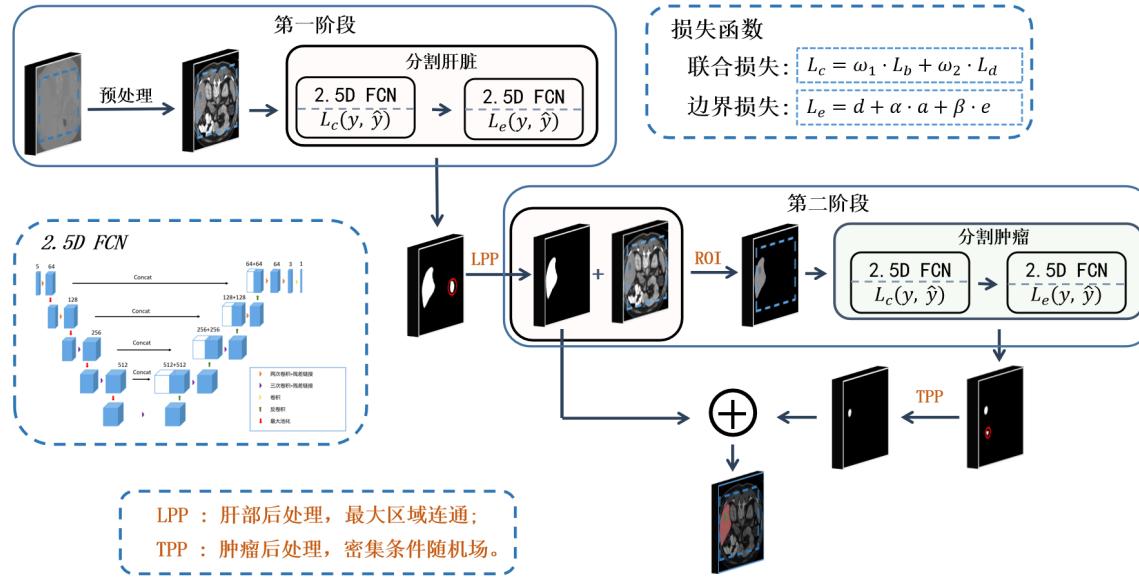


图 3.1 基于FCN的2.5D级联网络分割方法流程图，整个方法分为两个阶段，第一阶段分割肝脏，第二阶段分割肿瘤，每一阶段都采用2.5D FCN网络分割，并使用联合损失函数和边界损失函数优化此网络。

用联合损失函数优化第一阶段的2.5D网络，再使用边界损失函数优化该2.5D网络。然后本文使用第一阶段训练好的2.5D网络预测分割CT图像中的肝脏，并使用最大连通区域方法对预测的结果进行后处理，减少假正例，提升肝脏分割的精度。在第二阶段肿瘤分割任务中，本文首先利用第一阶段产生的只含有肝脏的CT图像数据和联合损失函数训练优化第二阶段2.5D网络的参数，再使用边界损失函数对该网络进行进一步的优化。在第二阶段的2.5D网络完成CT图像中肿瘤的预测分割后，本文使用了密集条件随机场对网络预测结果进行后处理，细化网络分割的结果，提升肿瘤分割的精度。最后，本文将肝脏与肿瘤的分割结果合并，作为最终的分割结果输出。

3.2 CT图像预处理

由于原始的CT图像存在大量噪声，不同组织器官之间对比度较低，使得深度学习网络难以从中学习到有效的特征信息。为了增强肝脏与肿瘤CT图像的质量，本文结合多种不同的方法对CT图像进行预处理。如图3.2所示，本文首先对原始的

肝脏与肿瘤CT图像进行CT值截断操作并将它们转换为灰度图，接着依次使用高斯滤波器进行降噪、直方图均衡化对灰度值分布进行拉伸以及图像锐化增强图像边缘特征，最后使用双三次插值法扩充用于训练的图像数据，为后续的深度学习网络分割图像提供数据基础。

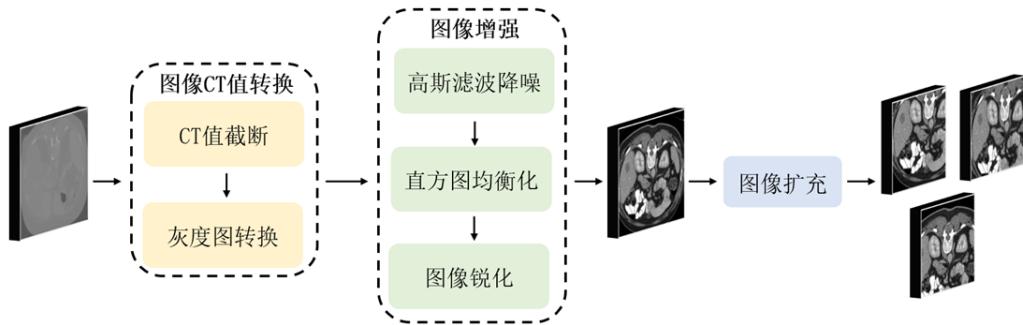


图 3.2 CT 图像预处理方法流程图，在预处理过程中首先对原始CT图像进行CT值转换，将图像CT值截断并归一化为灰度图像，接着进行图像增强操作，获得对比度更高的CT图像，最后对图像进行扩充，产生裁剪后的CT图像以适应深度学习网络的输入。

3.2.1 图像CT值截断与归一化

与由像素点构成的二维图像不同，三维CT图像由体素点构成，体素点的值被称为CT值或者密度值，其单位是亨氏(HU) [83]。CT图像的CT值通常分布在[-1024, 2048]，这是一个跨度非常大的范围，而人体腹部的组织器官CT值分布如表3.1所示，通常分布在一个较小的区间内。因此在大范围CT值分布下，图像中的许多组织器官模糊不清。

表 3.1 人体腹部组织器官CT值分布范围

器官	肝脏	脾脏	胰腺	肾脏
CT值范围	50 - 70	35 - 60	30 - 50	25 - 50

为了增强组织器官的清晰度，本文对肝脏与肿瘤CT图像进行了CT值截断操作，以获得更清晰的CT图像。如图3.3所示，(a)为原始CT切片图像，其CT值范围分布在[-2048, 1024]之间，(b)、(c)和(d)为CT值截断后的切片图像，它们的CT值范围分别分布在[-300, 300]、[-200, 200]和[-150, 300]之间。当CT值截断在[-300, 300]之间

时，CT图像中的组织器官仍然模糊不清，几乎不可辨认。而当进一步缩小截断区间时，图像中各个组织器官的清晰度和对比度明显增加。经过多组测试，最后本文选择[-150, 300]作为分割肝部的CT值截断区间。

在完成CT值截断操作之后，本文使用图像归一化操作将CT值转换为像素灰度值，从而将CT图像转换为灰度图像，其灰度范围在[0, 255]之间。由于肿瘤的CT值平均分布要低于肝脏，因此经过多组测试，本文最终选择[-135, 265]作为肿瘤分割的CT值截断区间，同时将CT值截断后的图像使用归一化操作转换为[0, 255]之间的灰度图像。

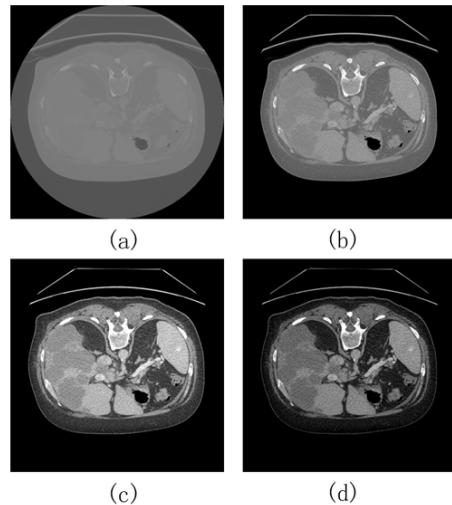


图 3.3 不同CT值分布范围下的CT切片图像，(a)为CT值范围分布在[-2048, 1024]之间的原始CT切片图像，(b)、(c)和(d)分别为CT值分布在[-300, 300]、[-200, 200]和[-150, 300]之间的CT切片图像。

3.2.2 图像增强

在完成图像CT值转换操作之后，由于CT图像在从X射线扫描、收集测量信号以及输入计算机转换为可观察图像的过程中会产生大量噪声。这些噪声降低了图像的质量，使得图像特征差异化增大，给后续的深度学习分割方法带来了阻碍。因此在使用深度学习方法进行图像分割前，有必要选择合适的方法进一步增强图像的质量，获取对比度更高的CT图像。高斯滤波是一种线性平滑滤波，本质是借助高斯函数对整张图像进行加权求平均值，能够有效地抑制图像中服从正态分布的噪声。

经过比较，本文选用 5×5 的高斯核对肝脏与肿瘤CT图像进行降噪处理，从而有效地减少了CT图像中的噪声。

在经过CT值截断和图像降噪操作后，图像的清晰度有着明显的提升。但在成像过程中，造影剂的使用导致了图像的CT值分布不均且保持在一个较小的范围内，因此有必要对图像执行灰度直方图均衡化操作以改变原始图像的灰度分布，使图像的灰度范围变得均匀以增强图像对比度。如图3.4所示，(a)是肝脏与肿瘤CT切片原图，与它对应的灰度直方图(b)中的灰度值集中分布在[0, 40]和[100, 200]之间，灰度值分布非常紧密，图像整体呈现出模糊的状态。这样的图像在分割时难以提供足够的轮廓和边界特征信息，从而阻碍了深度学习方法分割精度的提升。(c)是经过直方图均衡化的肝脏与肿瘤CT切片图像，从与(c)对应的灰度直方图(d)中可以看出，整张图像的灰度范围均匀分布在0到255之间。将图像(c)与(a)对比，可以看出图像整体变亮，轮廓和边界变得清晰可见，图像中的组织器官对比度明显增强，而这对于后续分割工作的进行是非常有利的。

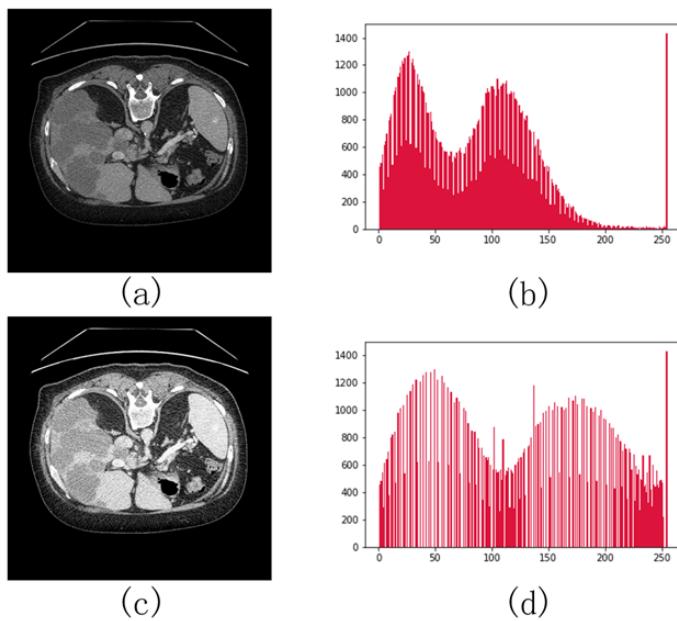


图 3.4 肝脏与肿瘤CT切片图直方图均衡化效果图

图像锐化是另一种常见的图像增强方法，通过增强高频分量能够减少图像模糊的情况，此外还能够增强图像的边缘细节、轮廓以及灰度反差。图像锐化能够

一定程度上缓解肝脏与肿瘤CT图像中的组织器官对比度不足以及边缘细节不明显的问题，从而有利于后续深度学习方法分割CT图像中的肝脏与肿瘤。本文使用了拉普拉斯算子实现图像锐化处理。拉普拉斯算子是一种线性的二阶微分算法，其旋转不变性的特点满足了不同方向的图像边缘锐化要求。如图3.5所示，(a)是肝脏与肿瘤CT切片图像；(b)为经过拉普拉斯算子增强后的肝脏与肿瘤CT切片图。从图3.5可以看出，对肝脏与肿瘤CT图像应用拉普拉斯算子可以增强图像中的灰度突变，即降低灰度变化缓慢的区域，从而达到锐化图像的效果。

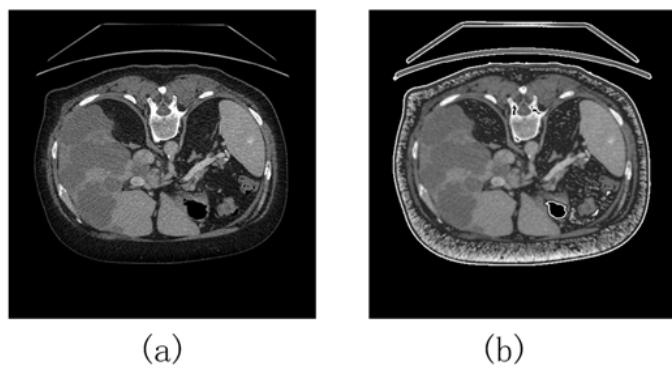


图 3.5 肝脏与肿瘤CT切片锐化效果图

3.2.3 图像扩充

深度学习是一种基于数据驱动的方法，能够从大量标注数据中自动地学习图像特征信息，从而进行图像分割。但是许多情况下难以获取足够多的训练图像，如医学图像分割，属于小样本图像处理问题。这是由于医学图像所包含的信息通常涉及到病人的隐私，这使得获取大量的医学图像数据是困难的；此外对医学图像进行标注时，通常需要拥有医学领域相关知识的人员来手动完成，标注效率极低，占用大量的人力和时间。为了满足深度学习对大量标注数据的需求，使用合适的方法对医学图像进行扩充是有必要的。

将图像按照一定的尺度比例进行缩放是一种常见的数据扩充方法。如图3.6所示，为了满足深度学习网络对海量数据的需求，本文在将图像输入网络训练前首先固定输入图像的模板尺度大小为 224×224 ，之后对该模板进行[0.8, 1.2]倍的随机缩

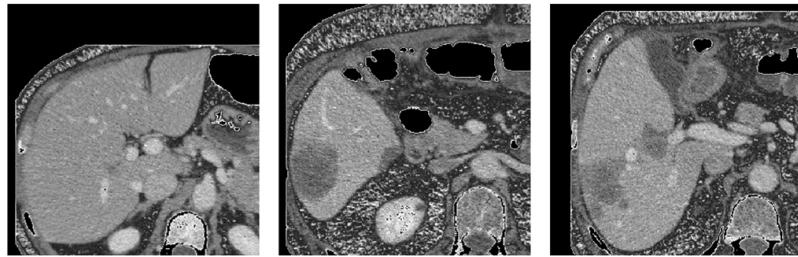


图 3.6 肝脏与肿瘤CT切片扩充示意图

放，从而得到一个新的模板尺度大小，接着按照新的模板尺度大小从原始图像裁剪局部图像数据，最后使用双三次插值法将裁剪的局部图像数据缩放到最初固定的模板尺度大小，从而完成肝脏与肿瘤CT图像数据的扩充。

3.3 基于FCN改进的2.5D全卷积神经网络设计

许多研究表明编码解码的网络结构有助于提升图像的分割效果 [33, 37]。本文借鉴了FCN编码解码网络结构的设计思路，在其基础上进行改进，提出了一种能够有效探索三维CT图像空间上下文特征信息的2.5D全卷积神经网络。通过第二章对FCN的介绍可知，FCN在设计时只使用了简单的卷积块构建网络的主体结构，因此当加深网络的深度时，模型存在梯度消失的问题。为此，本文在构建网络时引入了残差块增强梯度信息在网络内部的传播能力，从而设计出了更深的编码网络和解码网络 [63]，如图3.7所示。根据第二章的介绍，残差块的实现有两种方式，第一种残差块由多个 3×3 的卷积块构成，另一种残差块则是由两个 1×1 的卷积块和多个 3×3 卷积块构成。为了探索更多的图像特征，本文所使用的残差块内部并没有使用 1×1 的卷积块对探索的图像特征进行降维，而是使用两个或三个 3×3 的卷积块构成网络的残差块，其中构成残差块的每个卷积块都由一个 3×3 卷积层、一个归一化层和一个激活层组成。此外，FCN具有许多变体，U-Net作为一种经典的全卷积网络，因其对称的网络结构，对医学图像分割有着显著的分割效果，因此本文采用了与U-Net类似的对称性网络结构。

然而，在U-Net和其他的2D网络中 [35]，网络的输入是一张图像，输出则是一

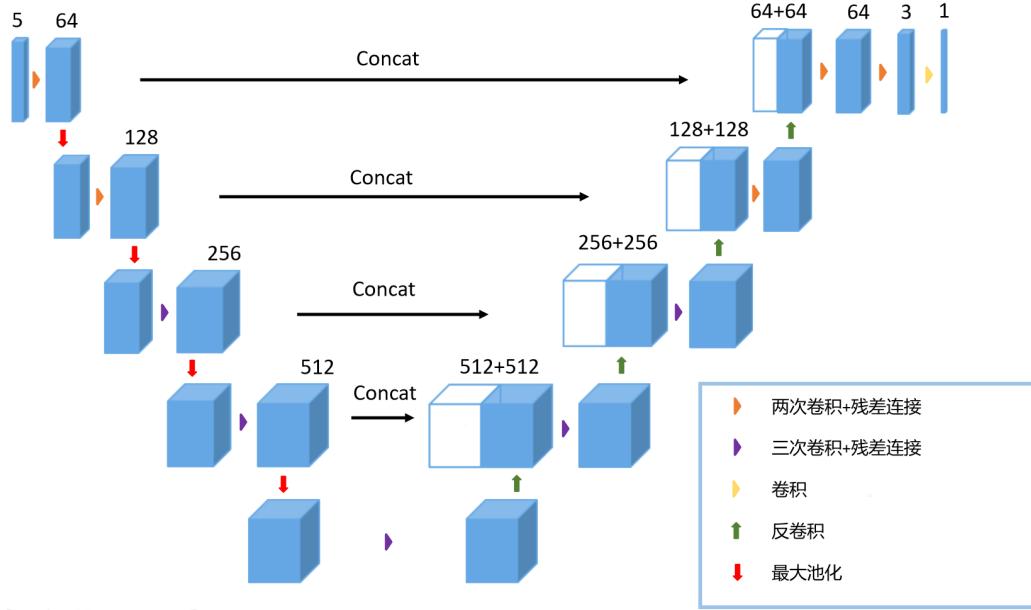


图 3.7 2.5D全卷积网络结构示意图

张与输入图像对应的概率预测图。与这些网络不同，本文设计的2.5D网络将五张级联的CT切片图像作为初始特征输入网络，在输出时却只输出五张图像中间的切片，即第三张切片所对应的概率预测图。多张级联切片作为初始特征输入网络是设计2.5D网络的关键所在，能够使网络更充分地探索三维数据的空间上下文信息。整个2.5D网络结构的表示如下：

$$I = f_{2.5D}(X, \hat{L}, \theta') \quad (3.1)$$

其中 $f_{2.5D}$ 表示2.5D全卷积网络， X 表示网络输入的初始特征，即五张级联的CT切片图像。 θ' 表示需要训练的网络参数， \hat{L} 表示用于优化网络的损失函数， I 则是网络 $f_{2.5D}$ 输出的概率预测图。如图3.7所示，2.5D网络由两个部分组成，探索特征上下文信息的编码网络和恢复特征空间位置信息和尺寸大小的解码网络。在编码网络的第一层和第二层，本文使用两个卷积块构成的残差块探索低级纹理特征，同时减少显存的占用。在编码网络的第三层至第五层，本文使用三个卷积块构成的残差块探索更多的高级抽象特征。在本文2.5D网络中，解码网络与编码网络是对称的，它们的不同之处在于解码网络使用反卷积取代了编码网络中的最大池化操作。相比于3D网络 [37]，本文设计的2.5D网络能够保证适中输入图像尺寸，同时又能减少网

络需要训练的参数量和显存的占用。

3.4 用于优化2.5D网络的边界损失函数设计

在训练深度卷积神经网络的过程中，损失函数能够评估预测值与真实值之间的差异性，优化网络的参数，提升网络的特征表达能力。因此选择合适的损失函数来优化网络参数能够有效地提升模型的分割效果。在肝脏与肿瘤CT图像自动化分割任务中，肝脏与肿瘤通常具有一定的轮廓和边界，尽管这些边界通常是模糊的，但是对于肝脏与肿瘤的分割却是非常重要的图像特征，充分利用肝脏与肿瘤CT图像中边界特征能够有效地提升网络分割的准确度。然而常见的损失函数，如交叉熵损失函数和相似性系数损失函数通常缺少优化网络学习图像中边界和轮廓特征的能力。为此本文提出了一种新的边界损失函数能够优化网络学习更多的边界和轮廓特征。经过边界损失函数优化的网络，能够探索更多图像边界特征，对肝脏的边缘分割更加平滑，有效地减少预测结果中的假正例，提升肝脏与肿瘤分割的精度。边界损失函数 L_e 可以表示为：

$$L_e = d + \alpha \cdot a + \beta \cdot e \quad (3.2)$$

其中 d 、 a 和 e 分别表示邻近体素之间的距离、轮廓的面积和轮廓的边界， α 和 β 分别是面积和边界对应的加权， d 和 a 分别可以表示为：

$$d = \sum_{i=1,j=1}^N \sqrt{(\Delta u_{x_{i,j}})^2 + (\Delta u_{y_{i,j}})^2} \quad (3.3)$$

同时，

$$a = \left| \sum_{i=1,j=1}^N u_{i,j} (1 - v_{i,j}) \right| + \left| \sum_{i=1,j=1}^N v_{i,j} (1 - u_{i,j}) \right| \quad (3.4)$$

其中 $x_{i,j}$ 和 $y_{i,j}$ 分别是像素点 (i, j) 的横纵坐标， $\Delta u_{x_{i,j}}$ 和 $\Delta u_{y_{i,j}}$ 表示对应坐标像素值相减的结果， $v_{i,j}$ 和 $u_{i,j}$ 分别表示标注值和预测值， N 为图像的像素点总数。为了表示边界 e ，假设标注图像为 A ， B 是 A 进行膨胀的结果， C 则是对 A 进行腐蚀的结果。通过 A ， B 和 C 可以得到外边界图像 $Om = A \oplus B$ 和内边界图像 $Im = A \ominus C$ 。如图3.8所示，其中 \oplus 表示异或操作， Om 和 Im 边界处的值为1，非边界处的值为0。接

着利用 Om 、 Im 以及预测图像 D 来获取预测图像的外边界损失 O 和内边界损失 I ，其中 $O = Om \times D$ ， $I = Im \times D$ 。值得注意的是， O 和 I 是与标注图像 A 具有相同大小的图像，只是它在边界处的值为网络预测的概率，非边界处的值为0， O 和 I 的示意图如3.8所示。为了利用边界损失达到优化预测图像边界的目的，需要使得外边界损失 O 尽可能小，内边界损失 I 尽可能大，而又为了保持整体损失尽可能小，因此用1减去内边界损失 I 达到取反的目的。此时，边界损失函数的 e 可以表示为：

$$e = \left| \sum_{i,j}^N O_{i,j} + \sum_{i,j}^N (1 - I_{i,j}) \right| \quad (3.5)$$

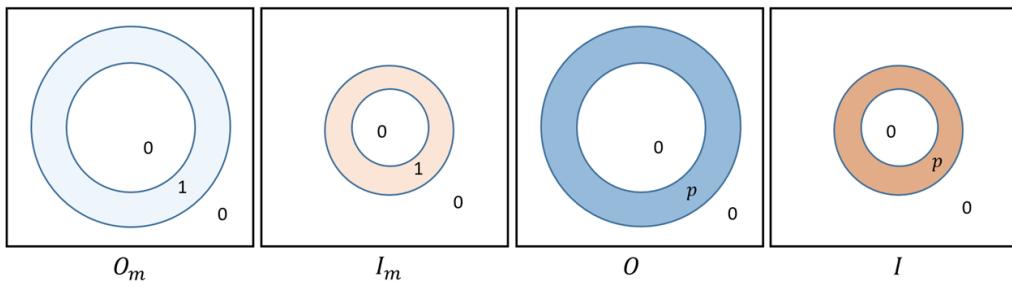


图 3.8 标注边界图像与预测边界损失图像的示意图，其中 $Om = A \oplus B$ 和 $Im = A \oplus C$ 分别为外边界图像和内边界图像，通过标注图像 A 与其膨胀图 B 和腐蚀图 C 进行异或操作获取； $O = Om \times D$ 和 $I = Im \times D$ 分别为外边界损失图像和内边界损失图像，通过标注边界图像 Om 、 Im 与预测图像 D 的对应像素点相乘获取；图中的 p 表示网络预测概率。

边界损失函数整合了分割目标邻近像素点之间的距离、轮廓的面积和轮廓的边界特征信息作为损失优化网络参数。在医学图像中，相邻的组织器官的密度值具有一定的相似性。因此，如公式(3.3)所示，最小化医学图像邻近像素点之间差异性，即最小化水平和垂直方向上邻近像素点的距离，并把该距离 d 作为损失的一部分能够增强分割目标区域内的平滑性和完整性。边界损失函数中的面积 e 本质是由相似性系数损失函数变化而来。如公式(3.4)所示，本文通过最大化目标区域内的像素值和最小化背景区域像素值来确保网络对目标区域分割的有效性。如公式(3.5)所示，本文通过标注图像获得分割目标的边界，接着计算预测图像内外边界之间距离，并将该距离作为损失来优化网络，探索更多的图像边界特征，从而达到平滑分割目标边界的效果。

3.5 图像后处理

在2.5D网络完成对肝脏与肿瘤CT图像的预测分割后，其分割结果仍存在一些假正例和假反例，具有一定的优化提升空间。如图3.9，第一列为CT标注图像；第二列为网络预测结果图像；第三列为后处理的结果图像。从图3.9可以看出，经过后处理的图像假正例有明显减少。因此有必要使用合适的后处理方法进一步优化网络的预测分割结果。在使用2.5D网络完成第一阶段的肝脏分割后，本文先使用密集条件随机场调整网络输出的概率映射图的概率，为其分配标签，优化预测结果的边缘，提升分割的准确度。接着为了进一步减少肝部假正例，本文对密集条件随机场输出的预测结果使用最大连通区域去除肝部以外的假正例，提升分割精度。而对第二阶段2.5D网络输出的肿瘤预测结果，本文首先使用密集随机场条件场优化预测的概率结果，再基于肿瘤CT值低于肝部CT值的特点，计算预测图像中每个肿瘤区域的平均CT值，将预测图像每个肿瘤CT平均值高于肝脏CT平均值的肿瘤区域作为假正例和假反例排除，提升肿瘤分割的准确度。

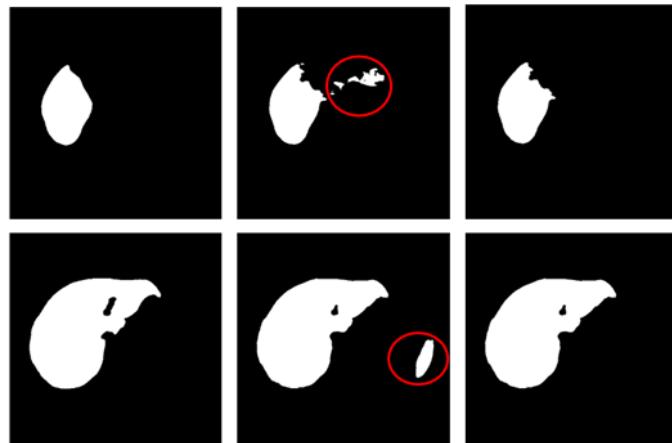


图 3.9 CT标注图像与2.5D网络预测结果图像对比，第一列为CT标注图像；第二列为2.5D网络预测结果图像，其中红色圈中为假正例；第三列为经过后处理的图像。

3.5.1 密集条件随机场

随机场本质是一组随机变量的集合，这些随机变量在按照某种分布被赋值后组

成了随机场 [84]。具有马尔可夫性质的随机场通常叫马尔可夫随机场 [85]。在马尔可夫随机场中，每一个随机变量的赋值仅与它相邻的随机变量赋值有关，而与不相邻的随机变量赋值无关。马尔可夫随机场也可以看作随机场的特例 [86]。在马尔可夫随机场上进一步加强限定条件则构成了条件随机场 [87]。假设在马尔可夫随机场中只有两种随机变量 X 和 Y ，其中 X 的赋值是给定的， Y 的赋值则是基于条件 X 的赋值得到的，该随机场就是一个条件随机场。因此，条件随机场是具有更强约束条件的马尔可夫随机场，是马尔可夫随机场的特例 [88]。

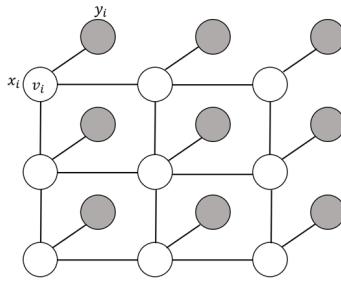


图 3.10 条件随机场模型示意图

为了使用条件随机场优化肝脏与肿瘤CT图像分割的边缘，本文基于肝脏CT图像构造了一个条件随机场。在一张肝脏CT图像中，每个体素点 i 可以看作一张图 $G = (V, E)$ 中的节点 $v \in V$ 。该节点 v 具有固定的类别标签 x_i 以及对应的观测值 y_i ，节点与节点之间的依赖关系视为 $e \in E$ ，这样也就构成了一个基于肝脏CT图像的条件随机场，如图3.10所示。基于节点与相邻节点之间依赖关系所组成的条件随机场又被称作稀疏的条件随机场。这种条件随机场由于只考虑了邻近节点之间的关系，导致其对神经网络分割的图像进行后处理的效果提升不大。因此在实际应用中，通常使用密集条件随机场来构建具有更高节点连接性的概率优化模型，密集条件随机场也被称为全连接条件随机场 [88]。

具体而言，假设 $G = (V, E)$ 是基于肝脏CT图像 I 构造的一张无向完全图，其对应的密集条件随机场为 X ，则 G 服从吉布斯分布的能量函数为：

$$E(X) = \sum_{v_i} \phi_i(x_i) + \sum_{(v_i, v_j) \in E} \phi_{ij}(x_i, x_j) \quad (3.6)$$

其中 $\phi_i(x_i) = -\log P(x_i|I)$ 是一阶势能函数，对应深度学习网络输出的类别概率映

射图。 $\phi_{ij}(x_i, x_j)$ 则是二阶势能函数，能够对相似的体素点分配相同的标签，具体表示为：

$$\phi_{ij}(x_i, x_j) = u(x_i, x_j) \left(\omega_{pos} \exp\left(-\frac{|p_i - p_j|^2}{2\sigma_{pos}^2}\right) + \omega_{bil} \exp\left(-\frac{|p_i - p_j|^2 - |I_i - I_j|^2}{2\sigma_{bil}^2 + 2\sigma_{int}^2}\right) \right) \quad (3.7)$$

其中 $u(x_i, x_j) = 1$ 为标签兼容函数，即分类错误的惩罚项， $|p_i - p_j|$ 表示体素 v_i 与 v_j 之间的空间距离， $|I_i - I_j|$ 则是体素 v_i 与 v_j 的体素值的差。 ω_{pos} 和 ω_{bil} 是二阶势能函数的权重， σ_{pos} 、 σ_{bil} 和 σ_{int} 是对应的核宽度，它们共同决定着二阶势能函数的影响力。由于密集条件随机场会将肝脏CT图像中每个体素点连接构造边集，从而导致计算机无法对如此庞大边集关系进行推理运算。因此本文采用了文章 [88]提出的平均近似算法来构建密集条件随机场，通过一组线性高斯滤波器来替代二阶势能函数执行消息传递的过程，进而调整概率预测图中的概率值，优化肝脏与肿瘤CT图像的边界分割。

3.5.2 最大连通区域

经过密集条件随机场对网络输出的概率映射图调整分配标签后，肝脏与肿瘤CT图像的分割结果中仍然存在一些假正例和假反例。根据常识，在人体的腹部CT扫描中只可能存在一个肝部区域，同时该区域是连通的。因此当分割结果中存在两个以上的连通区域，必然只有一个肝脏区域，其他连通区域是错误的分割结果。基于这样的思路，本文提出一种新的最大连通区域方法，用于减少分割结果中假正例和假反例。

对于经过密集条件随机场处理过的预测结果是一份三维的二值CT图像数据。假设图3.11是一张CT切片图像的局部数据。在图中，像素值相同且相邻的两个像素点被视作处于同一个相互连通的区域。因此，图3.11共可以划分为三个互相连通的区域，即黄、绿和蓝三个不同颜色的区域。最大连通区域方法的目标即是寻找值为1的最大连通区域，因此首先可以排除掉值为0的区域，此时剩下蓝色和黄色两个连通区域。又由于卷积神经网络的分割结果准确度一般较高，因此在剩余的区域中，面积最大的区域必然属于肝部 [17]，对每个区域计算它们的面积并进行比

1	0	0	0	0
1	1	0	0	0
1	0	1	0	0
0	0	0	0	1
0	0	0	1	1

图 3.11 二值CT切片局部图像数据示意图

较，最后保留面积最大的黄色连通区域作为最终的肝脏分割结果。后处理结果如图3.12，从图中可以看出经过最大连通区域操作处理后，可以有效的去除大部分假正例和假反例。

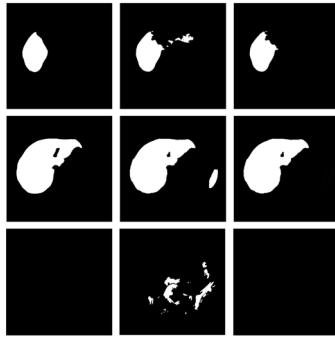


图 3.12 最大连通区域后处理示意图，第一列为CT标注图像，第二列为网络预测结果图像，第三列为后处理结果图像。

3.6 实验结果与分析

3.6.1 实验环境

本文在一台操作系统为Ubuntu 18.04的服务器上训练和测试提出的2.5D网络模型，该服务器配有一块Intel至强E5-2620 V4 2.10GHz的CPU和128G的内存，同时还配有3块Nvidia Tesla K80显卡，每块显卡具有24G显存。在开发过程中，本文使用Python语言来编写基础处理程序，同时基于Pytorch深度学习开发框架搭建本文的网络 [89]。

在训练和测试本文的网络时，本文使用了两种不同的肝脏肿瘤数据集，LiTS [90]和3DIRCADb [91]。LiTS数据集中含有131份有标注的训练CT图像和70份无标注的测试CT图像。这些CT图像来源于六所不同的医院，分别使用不同的扫描协议采集，因此这些CT图像的分辨率具有较大的差异性，其中切片内的分辨率在0.55 mm到1.0 mm之间，切片间的分辨率在0.45 mm到6.0 mm之间 [81, 17]。3DIRCADb 数据集包含了20份肝脏肿瘤数据集，其中15份CT图像数据含有肿瘤病灶，5份CT图像数据不含有肿瘤病灶。

为了对实验结果统一比较和分析，在进行每个实验前，本文都会先使用3.1小节所述的方法对数据进行预处理。其次在每个实验过程中，当网络完成对肝脏肿瘤CT图像的预测分割，本文统一使用3.5小节所述的方法对预测结果进行后处理，进一步提升分割效果。此外，在使用轮廓损失函数和边界损失函数优化网络时，本文将学习率设置为0.00005，当使用交叉熵损失函数、相似性系数损失函数和联合损失函数优化网络时，本文将学习率设置为0.001。同时，为了保证能够充分训练网络模型，当损失函数不再下降时本文将学习率乘以0.5的衰减因子，直到无法使损失进一步下降，才会停止网络训练。最后，为了防止过拟合，本文的所有实验统一使用了Adam (Adaptive Moment Estimation)优化器优化网络 [92]，并将网络参数的权重衰减系数设置为0.0005。

3.6.2 分析与讨论

为了验证本章设计的2.5D网络和边界损失函数的有效性，本文利用不同维度的网络与损失函数设计了一组消融性实验。具体而言，为了验证2.5D网络的有效性，本文修改2.5D网络的输入卷积块得到了一个新的2D网络。在实验中，2.5D网络输入图像特征的尺寸为 $5 \times 320 \times 320$ ，与2.5D网络将级联的五张切片图像作为初始特征输入不同，2D网络输入的是单张CT切片图像，输入图像特征的尺寸为 $1 \times 320 \times 320$ 。接着本文分别使用联合损失函数、轮廓损失函数和边界损失函数在131份LiTS肝脏与肿瘤训练数据集上训练优化2D网络和2.5D网络，并在70份LiTS肝脏与肿瘤测试数据集上测试训练好的模型。不同网络与损失函数组合的分割结果如表3.2所示，当

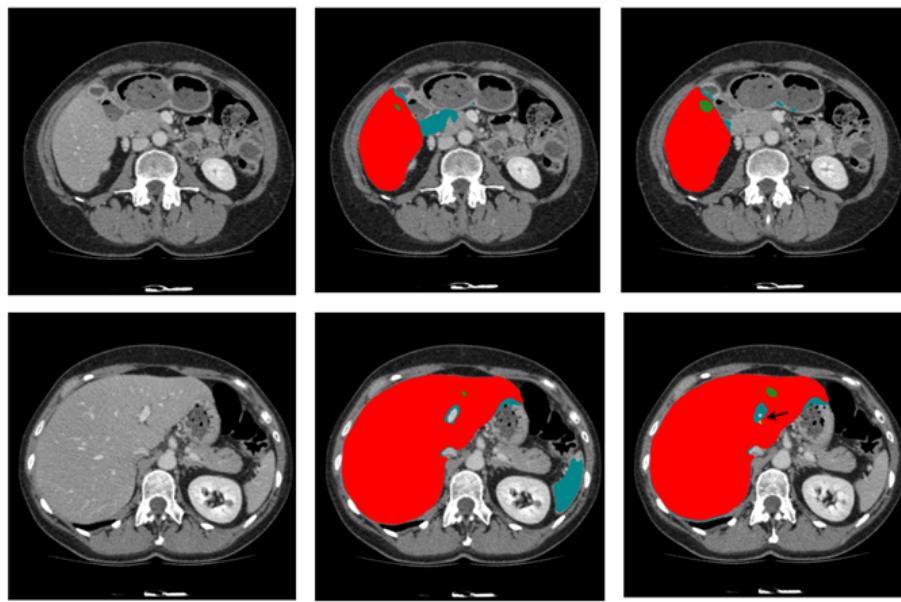


图 3.13 基于边界损失函数分别优化2D网络和2.5D网络的分割结果对比图，第一列为肝脏与肿瘤CT预处理切片图，第二列为使用边界损失函数分别2D网络的分割结果图，第三列为使用边界损失函数分别2.5D网络的分割结果图，其中肝脏分割真正例(红色)、肝脏分割假正例与假反例(蓝色)、肿瘤分割真正例(绿色)和肿瘤分割假正例与假反例(黄色，如黑色箭头所指)。

使用边界损失函数优化2D网络和2.5D网络分割肿瘤时，2D网络获得了73.7%的DPC分数和58.3%的IoU分数。与2D网络相比，2.5D网络获得了0.9%的DPC分数和1.1%的IoU分数提升。其次，在使用边界损失函数优化网络分割肝脏的过程中，2D网络获得了93.9%的DPC分数和88.5%的IoU分数，2.5D网络获得了0.4%的DPC分数和0.8%的IoU分数提升。一个实验结果如图3.13所示，显然2.5D网络更加准确地分割出了肝部和肿瘤区域。2.5D网络之所以表现比2D网络优秀，是因为2.5D网络在设计时将多张级联的切片图像作为初始特征输入网络，保留了切片间的空间依赖关系，因此能够探索更多的空间上下文特征信息，提升对CT图像的分割效果。

为了验证边界损失函数的有效性，本文使用了五种不同损失函数在131份LiTS肝脏与肿瘤训练数据集上训练优化2.5D网络，并在70份LiTS肝脏与肿瘤测试数据集上进行测试。如表3.3所示，在分割肝脏的任务中，使用交叉熵损失函数、相似性系数损失函数和联合损失函数优化2.5D网络所产生的DG分数差异并不大，相比之下，使用轮廓损失函数和边界损失函数优化2.5D网络能够获得更优异的分割结果。

表 3.2 不同网络与损失函数组合分割肝脏与肿瘤的实验结果对比 (DPC和IoU: %)

模型 + 损失	肿瘤		肝脏	
	DPC	IoU	DPC	IoU
2D 网络 + 联合损失	65.2	48.4	91.3	84.0
2D 网络 + 轮廓损失	72.4	56.8	93.2	87.3
2D 网络 + 边界损失	73.6	58.3	93.9	88.5
2.5D 网络 + 联合损失	68.2	51.8	92.6	86.3
2.5D 网络 + 轮廓损失	73.8	58.5	94.1	88.9
2.5D 网络 + 边界损失	74.5	59.4	94.3	89.3

表 3.3 不同损失函数优化2.5D网络的肝脏与肿瘤实验分割结果对比 (Dice: %)

损失	肿瘤		肝脏	
	DPC	DG	DPC	DG
交叉熵损失	62.4	67.7	91.2	92.6
相似性系数损失	65.7	68.3	92.3	93.4
联合损失	68.2	71.2	92.6	93.8
轮廓损失	73.8	75.6	94.1	95.2
边界损失	74.5	77.2	94.3	96.1

在分割肿瘤的任务中，使用相似性系数损失函数优化的2.5D网络相较于交叉熵损失函数获得了3.3%的DPC分数提升。相比于相似性系数损失函数，联合损失函数优化的2.5D网络获得了2.5%的DPC分数提升。如图3.14所示，上述的实验结果表明使用联合损失函数优于单独使用交叉熵损失函数或者相似性系数损失函数。此外，进一步使用轮廓损失函数和边界损失函数优化2.5D网络可以分别获得5.6%和6.3%的DPC分数提升。相比于轮廓损失函数，在分割肿瘤时，本文提出的边界损失函数获得了0.7%的DPC分数提升，这也表明边界损失函数能够优化2.5D网络探索更多的边界和轮廓特征，减少预测结果中假正例，提升分割精度。

图3.15展示了使用联合损失函数和边界损失函数分别优化2.5D网络分割肝脏与肿瘤的一个实验结果。在图3.15中，第二列为第一列肝脏与肿瘤CT切片的标注染

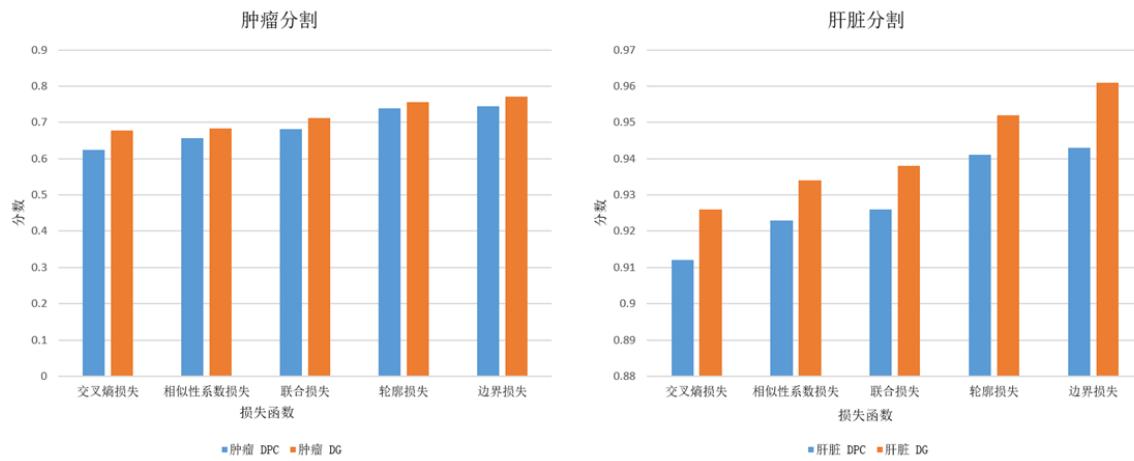


图 3.14 基于不同损失函数优化2.5D网络的实验结果柱形图，其中蓝色表示肿瘤分割结果，橙色表示肝脏分割结果。

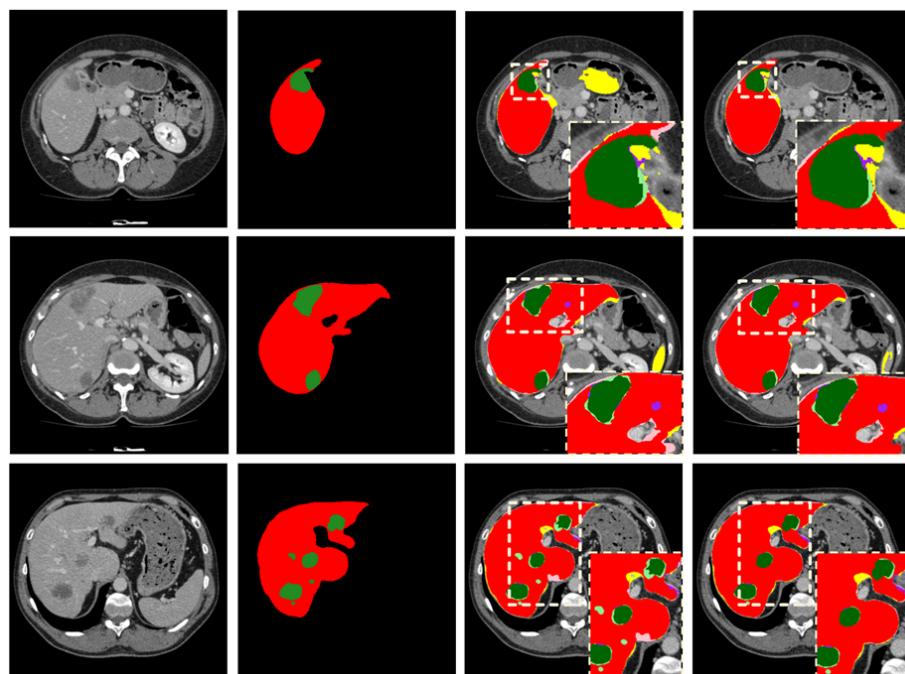


图 3.15 基于不同损失函数优化2.5D网络的肝脏与肿瘤分割结果图，第一列为肝脏与肿瘤CT切片预处理图，第二列为相应的肝脏与肿瘤CT切片标注图，第三列和第四列分别是使用联合损失函数和边界损失函数优化2.5D网络的分割结果图，其中肝脏分割真正例(红色)、肝脏分割假正例(黄色)、肝脏分割假反例(粉色)、肿瘤分割真正例(绿色)、肿瘤分割假正例(紫色)、肿瘤分割假反例(浅绿色)。

表 3.4 在3DIRCADb肝脏肿瘤数据集上肝脏分割结果对比(Dice: %)

模型	VOE	RVD	ASD	RSME	DG
2D UNet [93]	14.2 ± 5.7	- 0.05 ± 0.1	4.3 ± 3.3	8.3 ± 7.5	0.923 ± 0.03
2D FCN [81]	10.7	-1.4	1.5	24.0	0.943
2.5D ResNet [94]	11.6 ± 4.1	-0.03 ± 0.06	3.9 ± 3.9	8.1 ± 9.6	0.938 ± 0.02
Li 等人 [95]	9.2	-11.2	1.6	28.2 -	
Li 等人 [96]	-	-	-	-	0.945
我们的方法	8.5 ± 6.6	0.01 ± 0.02	1.6 ± 2.0	3.9 ± 6.5	0.961 ± 0.08

表 3.5 在3DIRCADb肝脏肿瘤数据集上肿瘤分割结果对比(Dice: %)

模型	VOE	RVD	ASD	RSME	DG
2D UNet [93]	62.5 ± 22.3	0.38 ± 1.95	11.1 ± 12.0	16.7 ± 13.8	0.51 ± 0.25
2D FCN [81]	-	-	-	-	0.56 ± 0.26
2.5D ResNet [94]	56.4 ± 13.6	-0.41 ± 0.21	6.3 ± 3.7	11.6 ± 7.6	0.60 ± 0.12
Li等人 [17]	49.7 ± 5.2	-0.33 ± 0.10	5.29 ± 6.1	11.1 ± 29.1	0.65 ± 0.02
我们的方法	57.5 ± 13.8	-0.18 ± 0.12	4.8 ± 3.7	15.9 ± 12.3	0.68 ± 0.15

色图，第三列和第四列分别为使用联合损失函数和边界损失函数优化2.5D网络的分割结果图，在第三列和第四列中，肝脏分割的真正例、假正例和假反例分别以红色、黄色和粉色表示，肿瘤分割的真正例、假正例和假反例则分别以绿色、紫色和浅绿色表示。综合比对第三列和第四列的图像，可以发现基于边界损失函数优化的2.5D网络分割结果中的假正例与假反例更少，分割结果的边界更加平滑，说明了边界损失函数能够优化网络从图像中学习到更多的边界与轮廓特征，减少了分割结果中的假正例与假反例。

为了进一步验证本章方法的有效性，本文在3DIRCADb肝脏肿瘤数据集上进行了六次交叉验证。在肝脏分割任务中，相比于2D网络U-Net [93]和FCN [81]，本文提出的2.5D网络具有更深的网络深度，同时多张级联切片作为初始特征输入的设计使得2.5D网络能够探索更多的三维数据空间上下文信息，因此本文提出的2.5D网

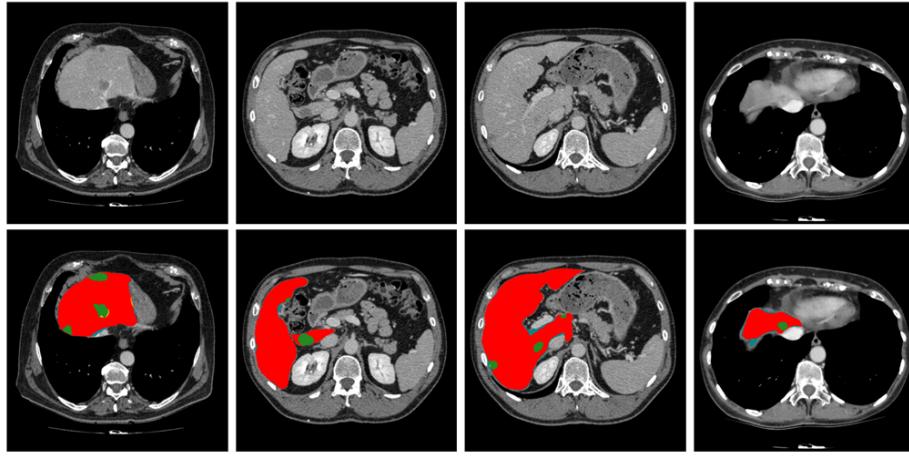


图 3.16 3DIRCADb肝脏肿与瘤数据集实验分割结果图, 第一行为肝脏与肿瘤CT预处理切片图, 第二行为肝脏与肿瘤CT分割结果图, 其中肝脏分割真正例(红色)、肝脏分割假正例与假反例(蓝色)、肿瘤分割真正例(绿色)和肿瘤分割假正例与假反例(黄色)。

络在分割效果上优于U-Net等2D网络。如表3.4所示, 本文提出的2.5D网络获得了更高的肝脏分割DG分数96.1%。在表3.5中, 相比于2D FCN [81]、3D H-DenseUNet [17]和2.5D ResNet [94], 本文提出的方法获得了最高的肿瘤分割DG分数68%。一个实验结果如图3.16所示, 从图中可以看出本文提出的2.5D网络能够较完整地分割CT图像中的肝脏, 对CT图像中的肿瘤同样具有优异的分割效果, 只存在少量假正例。

表 3.6 在LiTS肝脏与肿瘤数据集上与现有其他优秀方法的实验结果对比 (Dice: %).

模型	肿瘤		肝脏	
	DPC	DG	DPC	DG
2D UNet [93]	65.0	-	-	-
2D FCN [81]	67.0	-	-	-
3D V-Net [37]	-	-	-	93.9
Chlebus 等人 [97]	68.0	79.6	-	-
Yuan 等人 [98]	65.7	82.0	96.3	96.7
Li等人 [17]	72.2	82.4	96.1	96.5
3D Med [99]	-	-	-	94.6
本章的方法	74.5	77.2	94.3	96.1

在LiTS肝脏肿瘤数据集上，本方法和其他多种方法进行了对比实验，如表3.6所示。U-Net [93]、FCN [81]和Chlebus等人提出的网络 [97]都是2D网络，这些网络忽视了三维数据中的空间上下文特征信息，因此它们的分割准确度相对不足。在肝脏分割任务中，本文提出的2.5D网络分割效果远远超越了这些2D网络，获得了96.1%的DG分数。3D V-Net [37]和3D Med [99]本质都是类3D U-Net网络，与它们相比，本文提出的方法获得了2.2%DG分数提升。H-DenseUNet [17]与Yuan等人提出的网络 [98]同样都是3D网络，这些网络具有更深的网络深度和更高的网络维度，因此这些网络的DPC分数和DG分数要高于一般的2D网络方法。相比于这些3D网络，尽管本文提出的2.5D网络在肝脏分割分数上并不占优，但是本文提出2.5D网络参数量更少，需求的计算资源也更少。如在Batchsize同样为1的条件下，H-DenseUNet输入一份 $12 \times 224 \times 224$ 的CT图像数据需要占用24G的显存，而本方法输入一份 $5 \times 224 \times 224$ 的CT图像数据只消耗4G的显存。此外，在分割肿瘤任务中，与这些3D网络相比本文提出的方法获得了最高的DPC分数74.5%。

3.7 本章小节

本章首先从级联分割策略的角度详细地阐述本章的分割方法流程。接着简要介绍了CT值截断、高斯滤波降噪、直方图均衡化、图像锐化以及图像扩充等CT图像预处理方法。随后详细介绍了本文设计的基于FCN的2.5D全卷积神经网络和用于网络优化的边界损失函数。此外，还详细描述了用于减少假正例与假反例的后处理方法，密集条件随机场和最大连通区域方法。最后本文在LiTS和3DIRCADb肝脏与肿瘤数据集上从不同的角度对本章提出的方法进行了实验验证，实验结果表明本文提出的方法在肝脏与肿瘤分割任务中优于现有方法。

第四章 基于双路注意力的编解码网络分割方法研究

在上一章，本文基于FCN提出了一种新的2.5D全卷积神经网络，可以有效地探索三维数据中的空间上下文特征信息，同时本文针对肝脏与肿瘤CT图像的特点设计了一种用于优化网络参数的边界损失函数，能够优化网络学习更多的轮廓和边界特征。然而该网络缺少探索图像局部特征的能力，忽视了相邻像素之间的依赖关系，同时该网络深度不足，结构也稍显简单，存在特征冗余和特征探索不足的问题。因此选择合适的方法增强网络的特征表达能力是非常必要的，本文在第二章介绍的三种不同网络基础构建模块和双路注意力模块能够在一定程度上解决上述问题。然而仅使用这些基础网络模块并不足以实现肝脏与肿瘤CT图像的有效分割，因此还需借助一定的网络搭建技巧来构建更优秀的网络模型。本章基于编解码的网络结构，将三种不同的网络基础构建模块和双路注意力模块融入了编解码网络结构中，提出了一种新的基于双路注意力的编解码网络框架，包含2D、2.5D和3D三种不同维度的网络，共九种不同的网络模型，可以实现肝脏与肿瘤CT图像的有效分割。

4.1 方法概述

基于双路注意力的编解码网络方法流程如图4.1(a)所示，与第三章分割方法的过程类似，本章提出的方法同样采用了级联分割策略，分为两个阶段：分割肝脏阶段和分割肿瘤阶段，并对每个阶段的网络预测概率图进行后处理，通过合并每个阶段后处理的结果作为最终的分割结果。

具体而言，在肝脏分割阶段，本文会首先使用第三章的预处理方法增强CT图像的质量，消除图像中的噪音，提高关键组织器官的对比度，经过预处理的肝脏与肿瘤CT会被送入双路注意力编解码网络中。如图4.1(b)所示，双路注意力编解码网络分割框架实际上包含九种不同的注意力编解码网络，根据输入数据的维度不同，它们可以分为：2D FCN、2.5D FCN和3D FCN，而根据编码器主体构成的不同，它们又可以分为：Residual FCN、Dense FCN和Dual Path FCN。如图4.1(c)所

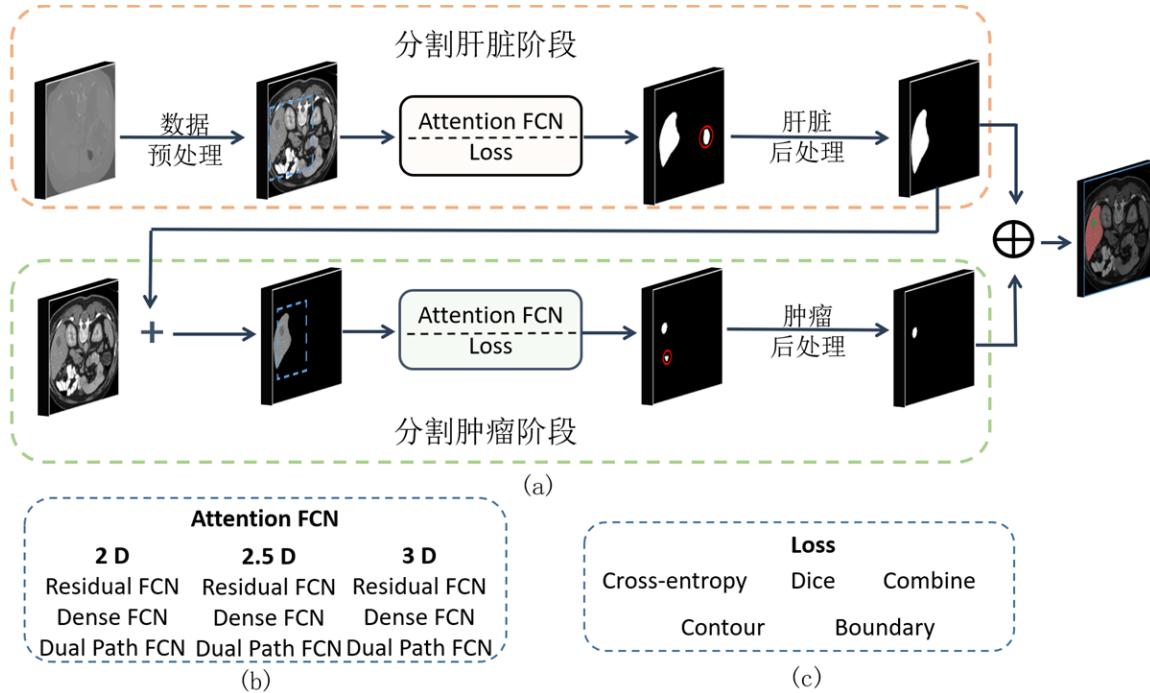


图 4.1 基于双路注意力的编解码网络级联分割方法流程图, (a)为基于双路注意力的编解码网络分割方法流程, (b)为双路注意力的编解码网络框架包含的九种不同网络, (c)为本文用于优化网络的损失函数。

示, 在训练注意力编解码网络时, 除了本文设计的Boundary loss, 本文还选择了其他四种损失函数作为对比来优化网络, 它们分别是Cross-entropy loss、Dice loss、Combined loss和Contour loss。在肿瘤分割阶段, 本文首先利用第一阶段经过预处理的原始数据和后处理的肝脏分割结果去除多余的组织器官, 获取只含有肝脏区域的CT图像数据, 接着利用这些数据训练双路注意力编解码网络, 并使用肿瘤后处理方法对网络预测结果进行优化。最后, 将经过后处理的肝脏分割掩码和肿瘤分割掩码合并, 并作为最终的分割结果输出。

4.2 编解码网络框架设计

为了更好地描述双路注意力编解码网络框架, 本节将首先介绍不含注意力的编解码网络框架。对于医学图像分割任务而言, 网络探索到的浅层纹理特征能够为图像分割提供类别依据, 深层抽象特征能够为精确的分割提供位置信息。因此, 在医

学图像分割任务中，浅层的纹理特征与深层的抽象特征同等重要。而编解码的网络结构恰恰能够有效的整合来自浅层的图像纹理特征和深层的图像抽象特征，因此与医学图像分割十分契合。此外，不同维度的网络对医学图像的分割效果不同，使用不同基础模块构建的网络探索特征的能力也不尽相同。基于上述原因，本文尝试将不同维度和不同基础模块构建的网络结构融入到同一种网络框架中，设计一种新的编解码网络分割框架，为双路注意力编解码网络框架的设计提供基础。

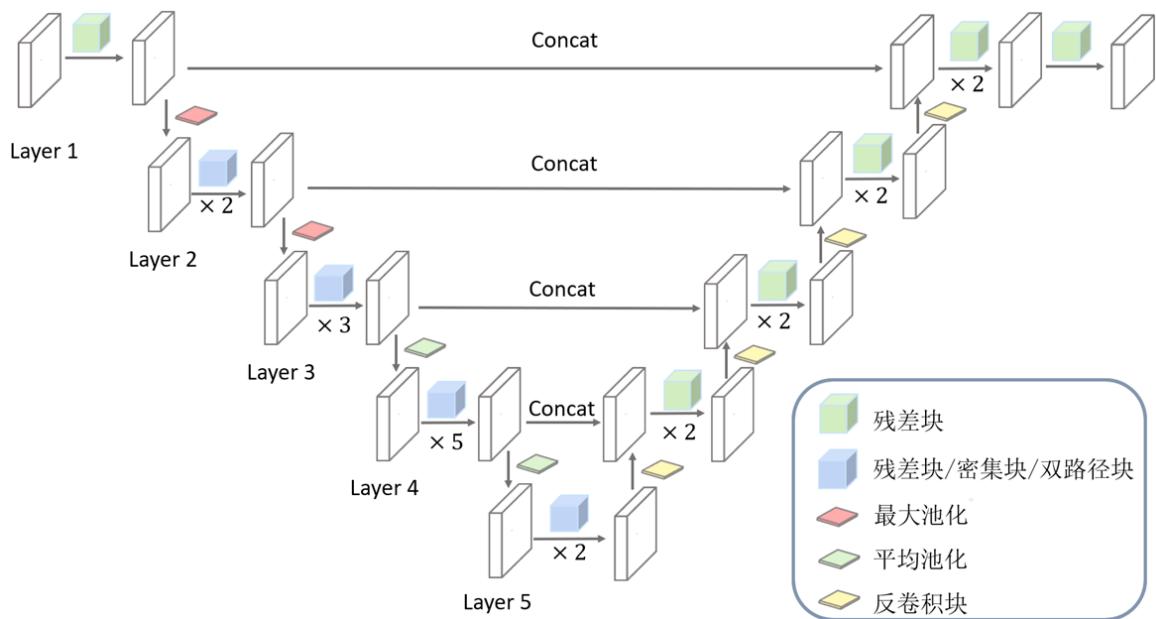


图 4.2 编解码网络框架示意图

编解码网络框架如图4.2所示，整个编解码网络框架分为五层，左边是对输入数据不断进行卷积和池化的编码器，右边则是对来自编码器的输出使用卷积和反卷积操作恢复图像特征分辨率的解码器。编码器的第一层，由一个适应不同维度输入的残差块构成，经过残差块卷积输出的特征在经过一个最大池化操作后将会进入网络的第二层。根据层次的网络不同，编码器的第二至第五层由不同数目的残差块、密集块和双路径块构成。在编码网络中，网络的浅层主要用于探索图像的纹理特征，因此使用较少的网络基础模块，而网络的深层主要用于探索图像的高级抽象特征，因此用了较多的网络基础模块。出于同样原因，本文在编码器的第一至第三层之间使用最大池化操作进行下采样，而在编码器的第三至第五层之间则是使用平均

池化进行下采样。此外，为了保证编码器中网络基础模块的一致性，对于残差块、密集块和双路径块，本文每次只选择其中的一种来构建编码器网络。在整个编解码网络框架中，解码器的主要作用是将编码器探索的高级抽象特征恢复成与输入图像具有相同尺寸的概率特征图，从而实现图像像素级语义分割。在设计解码器网络结构时，本文使用反卷积操作实现对图像特征上采样的过程，同时在每个反卷积操作之后加入两个残差块，从而保证上采样后的图像特征能够进行充分地组合和筛选。最后，为了提升网络的拟合效果，本文使用跳跃连接对编码器中的浅层特征与解码器中的深层特征重新整合，从而完成整个编解码网络结构框架的设计。

4.3 基于双路注意力机制的编解码网络框架设计

第二章描述的U-Net等编解码网络虽然能够通过融合上下文信息来提高网络的特征表达能力，同时获得丰富的全局特征信息，但它们忽视了对局部特征之间的相关性和依赖性的探索。注意力模块作为一种能够轻易地融入编解码网络之中的轻量级网络结构，可以有效提升网络的特征表达能力，探索医学图像中的局部特征相关性和依赖性。2019年，Fu等人 [80]提出了一种双路注意力模块，可以轻易地融入全卷积神经网络以解决场景语义分割的问题，其中的通道注意力模块和位置注意力模块可以有效地帮助神经网络探索更多通道特征和空间特征，从而提升场景图像分割的准确度。本文对双路注意力模块进行了修改和扩展，并将它融入了本文的编解码网络结构框架中，从而提出一种新的基于双路注意力的编解码网络分割框架。

4.3.1 改进的双路注意力模块设计

在场景语义分割问题中，分割目标往往受到光照、遮挡和尺度等因素的影响而难以辨别，而注意力模块能够增强网络的特征表达能力，使网络提取的图像特征更加具有辨识度 [80]。此外注意力模块还能更好的捕捉图像的细节纹理特征和边缘特征，从而有效地缓解上述问题。为了增强编解码网络对局部特征的探索能力，本文选用Fu等人 [80]提出的双路注意力模块对编解码网络框架进行优化。双路注意力模

块包含两个子模块：位置注意力模块和空间注意力模块，在第二章本文已经对它们进行了详细介绍，这里不再赘述。双路注意力模块通常用于2D网络的优化，为了将它融入本文提出的编解码网络框架，本文对它进行了扩展，使双路注意力模块能够同时适用于2D网络、2.5D网络和3D网络。如图4.3所示，相比于无注意力模块的编解码网络框架，本文并没有直接对编码器探索的特征进行上采样以恢复图像分辨率，而是在编码器与解码器网络的连接处融入了一个改进的双路注意力模块来探索更多的空间特征和通道特征。

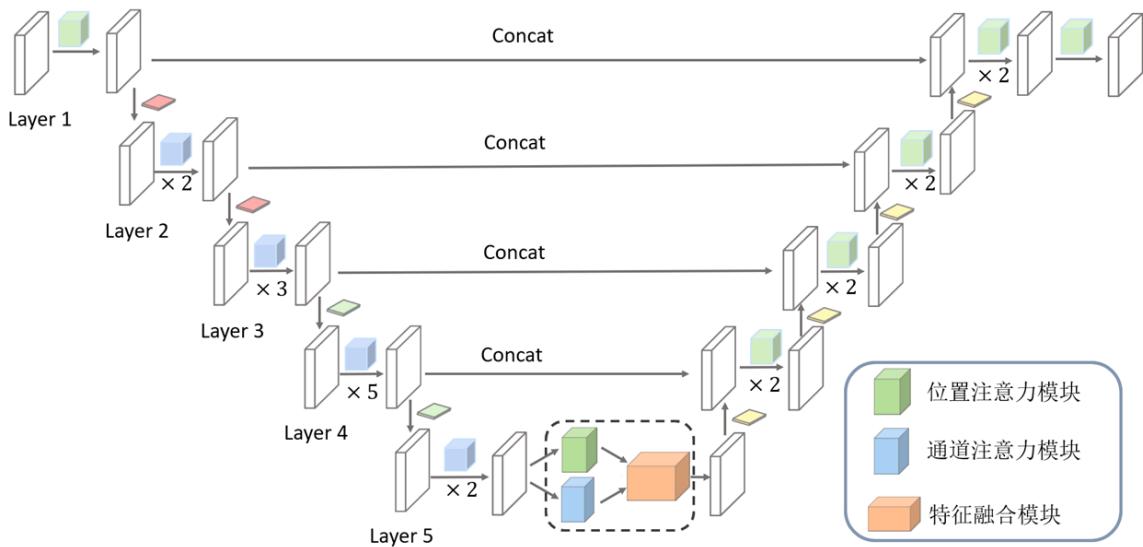


图 4.3 基于双路注意力的编解码网络框架示意图

在将双路注意力模块融入2D网络和2.5D网络时，本文使用的仍然是第二章介绍的双路注意力模块，只修改了其中的特征融合模块。2D网络与2.5D网络中的特征融合模块如图4.4所示，经过两个注意力模块增强的特征被送入特征融合模块进行重组输出。2D特征融合模块会先使用 3×3 卷积块调整上一层输入特征的通道维度，同时将 3×3 卷积输出的特征相加并保存，接着使用 1×1 卷积块对来自上一层的卷积特征进行线性组合，最后将这些重新组合的特征与上一步保存的特征拼接起来送入解码器中。特征融合模块能够对位置注意力特征和通道注意力特征进行组合和变化，从而生成更加抽象复杂的特征，提升整个网络的特征表达能力。由于2D网络和2.5D网络中的编码器网络和解码器网络的连接处具有相似性，因此2D网络和2.5D

网络中的双路注意力模块是相同的，为了与3D网络中的双路注意力模块区分，本文将2D网络与2.5D网络中的双路注意力模块称作2D双路注意力模块。

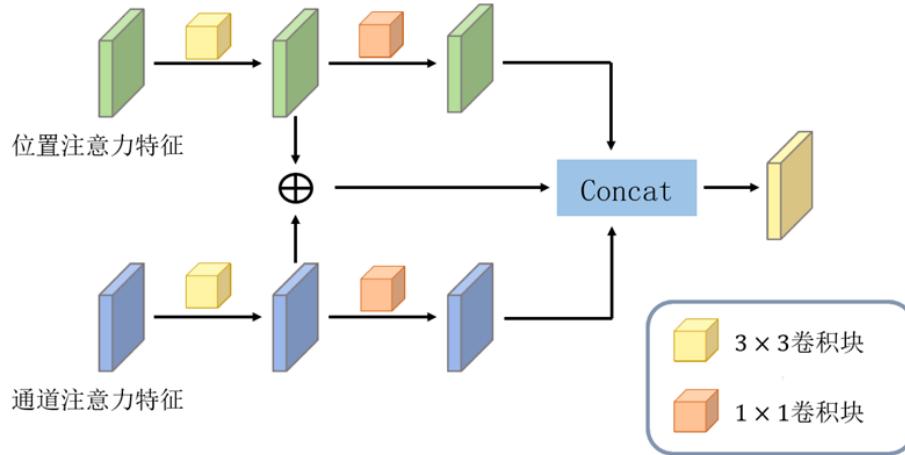


图 4.4 2D特征融合模块示意图

在将双路注意力模块融入3D网络时，本文对第二章介绍的位置注意力模块、空间注意力模块以及特征融合模块都进行了扩展，并将这种注意力模块称作3D双路注意力模块。如图4.5和4.6所示，本文将位置注意力模块、通道注意力模块和特征融合模块的中所有的2D卷积换成了3D卷积，从而匹配3D编码器网络提取的特征维度。此外，由于3D网络中编码器网络提取的特征维度为 $C \times D \times W \times H$ ，因此在两个注意力模块中进行调整形状操作时，本文实际上是将 $C \times D \times W \times H$ 的特征调整为 $C \times N$ 的特征，其中 $N = D \times W \times H$ 。

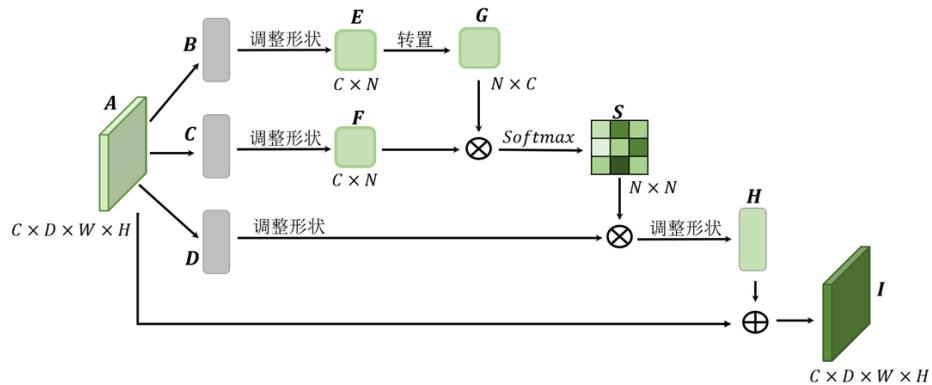


图 4.5 3D位置注意力模块示意图

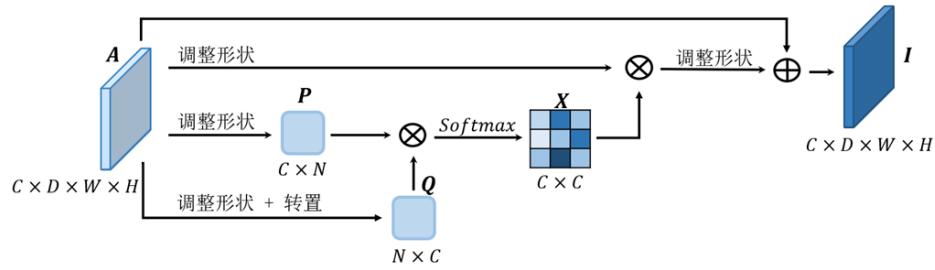


图 4.6 3D位置注意力模块示意图

4.3.2 基于双路注意力的2D编解码网络设计

在设计基于双路注意力的2D编解码网络时，本文使用了三种不同的网络基础模块来构建编码器网络，从而设计出了三种不同的2D双路注意力编解码网络。基于双路注意力的2D编解码网络属于双路注意力编解码网络框架中的2D网络，包含了三种2D网络：2D Attention Dense FCN(2D AD FCN)、2D Attention Residual FCN(2D AR FCN)和2D Attention Dual Path FCN(2D ADP FCN)，它们的网络结构模块配置如表4.1所示。

对于基于双路注意力的2D编解码网络而言，它们输入的是 $N \times 1 \times 224 \times 224$ 的肝脏与肿瘤CT切片图像，其中 N 表示每批次输入网络的数据量大小，1表示输入的特征数量，即一张图片，224表示输入图像的长度和宽度。为了将不同的2D网络融入同一个网络框架中，在设计时每种2D网络的第一层都是由一个残差块开始探索网络的低级纹理特征。网络第二至第五层则是由不同数目的残差块、密集块以及双路径块组成，用于探索图像中的高级抽象特征。接着本文将双路注意力模块(2D Attention Module)融入了编码器与解码器的连接处。在设计解码器时，为了使所有2D双路注意力编解码网络特征维度保持一致，本文没有使用难以控制特征维度的密集块和双路径块进行卷积，而是交替使用一个反卷积块和两个残差块逐步恢复特征的分辨率。最后，对于解码器的输出，我们使用单个残差块来产生与输入图像具有相同尺寸的图像特征，再使用Sigmoid激活函数输出最终的概率映射图。基于双路注意力的2D编解码网络其优点在于网络的参数量少，易于训练，需要的显存资源少。然而由于肝脏与肿瘤CT图像本质是由一张张连续的二维切片堆叠而成的

表 4.1 基于双路注意力的2D编解码网络结构配置表

2D FCN	特征维度	网络模块			模块数量
输入	(N,1,224,224)	-			-
Layer 1	(N,32,224,224)	2D Residual Block	2D Residual Block	2D Residual Block	1
	(N,32,112,112)	2D Max pool			1
Layer 2	(N,64,112,112)	2D Residual Blocks	2D Dense Blocks	2D Dual Blocks	2
	(N,64,56,56)	2D Max pool			1
Layer 3	(N,128,56,56)	2D Residual Blocks	2D Dense Blocks	2D Dual Blocks	3
	(N,128,28,28)	2D Avg pool			1
Layer 4	(N,256,28,28)	2D Residual Blocks	2D Dense Blocks	2D Dual Blocks	5
	(N,256,14,14)	2D Avg pool			1
Layer 5	(N,512,14,14)	2D Residual Blocks	2D Dense Blocks	2D Dual Blocks	2
		2D Attention			1
Up Layer 4	(N,256,28,28)	2D Deconvolution Block			1
		2D Residual Blocks			2
Up Layer 3	(N,128,56,56)	2D Deconvolution Block			1
		2D Residual Blocks			2
Up Layer 2	(N,64,112,112)	2D Deconvolution Block			1
		2D Residual Blocks			2
Up Layer 1	(N,32,224,224)	2D Deconvolution Block			1
		2D Residual Blocks			2
Conv 0	(N,1,224,224)	2D Residual Block			1
输出	(N,1,224,224)	Sigmoid Function			1

三维图像数据，相邻的切片之间具有紧密的相关性，而基于双路注意力的2D编解码网络使用的却是2D卷积核进行卷积，因此无法有效地探索数据中的三维空间上下文特征信息，对肝脏与肿瘤CT图像的分割能力也十分有限。

4.3.3 基于双路注意力的3D编解码网络设计

虽然基于双路注意力的2D编解码网络计算量更小，训练时间更短，但它忽略了相邻切片之间的依赖关系，影响了分割准确度。因此一些研究人员提出了基于3D网络的医学图像分割方法，如3D U-Net [36]和3D V-Net [37]。3D网络利用3D卷积核进行卷积，从而能够探索更多的三维图像空间中的上下文特征信息。本文设计的基于双路注意力的编解码网络框架中同样包含了3D网络结构，即基于双路注意力的3D编解码网络，包含三种不同3D网络：3D Attention Dense FCN(3D AD FCN)、3D Attention Residual FCN(3D AR FCN)和3D Attention Dual Path FCN(3D ADP FCN)。

基于双路注意力的3D编解码网络配置如表4.2所示，与上一小节的2D双路注意力编解码网络不同，除了图像固有的长度和宽度，3D网络的初始输入比2D网络多了一个维度，即深度，因此其输入维度为 $(N, 1, 12, 224, 224)$ ，其中 N 表示每批次输入网络的数据量大小，1表示输入的特征数量，12表示输入图像的深度，即1个输入特征包含了12张图像，224表示输入图像的长度和宽度。在设计3D网络中的编码器时，本文将残差块、密集块以及双路径块都扩展成了相应的三维卷积形式。此外，为了降低显存开销，同时确保网络能够探索到足够多的三维空间特征信息，本文在网络的第一层至第三层使用3D最大池化逐渐缩小特征的深度。此后，在编码器的第三至第五层将特征的深度始终保持为3。为了适应3D网络结构，在编码器与解码器的连接处，本文将双路注意力模块(3D Attention Module)扩展成了3D形式。具体而言，本文将位置注意力模块、通道注意力模块以及特征融合模块的卷积块换成了相应的三维卷积块，即完成了2D Attention Module到3D Attention Module的扩展。为了保持特征维度的一致性，在设计解码器过程中，本文保持解码器网络的第三层和第四层的网络深度不变，只对长度和宽度进行特征尺度的提升。从这里可以

表 4.2 基于双路注意力的3D编解码网络结构配置表

3D FCN	特征维度	网络模块			模块数量
输入	(N,1,12,224,224)	-			-
Layer 1	(N,32,12,224,224)	3D Residual Block	3D Residual Block	3D Residual Block	1
	(N,32,6,112,112)	3D Max pool			1
Layer 2	(N,64,6,112,112)	3D Residual Blocks	3D Dense Blocks	3D Dual Blocks	2
	(N,64,3,56,56)	3D Max pool			1
Layer 3	(N,128,3,56,56)	3D Residual Blocks	3D Dense Blocks	3D Dual Blocks	3
	(N,128,3,28,28)	3D Avg pool			1
Layer 4	(N,256,3,28,28)	3D Residual Blocks	3D Dense Blocks	3D Dual Blocks	5
	(N,256,3,14,14)	3D Avg pool			1
Layer 5	(N,512,3,14,14)	3D Residual Blocks	3D Dense Blocks	3D Dual Blocks	2
		3D Attention			1
Up Layer 4	(N,256,3,28,28)	3D Deconvolution Block			1
		3D Residual Blocks			2
Up Layer 3	(N,128,3,56,56)	3D Deconvolution Block			1
		3D Residual Blocks			2
Up Layer 2	(N,64,3,112,112)	3D Deconvolution Block			1
		3D Residual Blocks			2
Up Layer 1	(N,32,6,224,224)	3D Deconvolution Block			1
		3D Residual Blocks			2
Conv 0	(N,1,12,224,224)	3D Residual Block			1
输出	(N,1,12,224,224)	Sigmoid Function			1

看出，除了深度不同，整个3D双路注意力编解码网络的骨骼结构与2D双路注意力编解码网络的骨骼结构保持一致。

基于双路注意力的3D编解码网络符合医学图像数据的三维体积特性，因此其能够有效地探索数据中的三维空间特征信息，保证分割结果的三维连续性和体积完整性 [37]。虽然3D双路注意力编解码网络对数据中的空间特征信息具有较强的感知力，但它占用了更多的计算资源并且花费了更多的时间用于网络训练，同时由于3D网络扩展了更高的维度，使得网络参数量增大，一旦遇到规模较小的数据集，往往具有较高的过拟合风险。

4.3.4 基于双路注意力的2.5D编解码网络设计

2D网络通常难以达到医学图像分割精度的标准，3D网络往往又需要消耗较多的计算资源。为此，一些研究人员设计出了2.5D网络结构，2.5D网络结构又被称作伪3D网络，能够在减少网络参数量和计算成本的同时探索医学图像中的三维空间特征信息。本文设计了一种基于双路注意力的2.5D编解码网络结构，它包含了三种2.5D FCN：2.5D Attention Dense FCN(2.5D AD FCN)、2.5D Attention Residual FCN(2.5D AR FCN)和2.5D Attention Dual Path FCN(2.5D ADP FCN)。

如表4.3所示，2.5D双路注意力编解码网络和2D双路注意力编解码网络最大不同之处在于输入和输出图像的特征维度。基于双路注意力的2.5D编解码网络的输入图像特征维度为 $(N, 5, 224, 224)$ ，其中 N 表示每批次输入网络的数据量大小，5表示输入的特征数量，即5张图片，224表示输入图像的长度和宽度。然而，2.5D编解码网络的输出图像特征维度却是 $(N, 1, 224, 224)$ ，这是因为2.5D双路注意力编解码网络的输入是级联的五张CT切片图像，输出的却是五张切片的中间切片的概率映射图。因此2.5D双路注意力编解码网络能够探索一定的三维医学图像数据中的空间特征信息。事实上，由于2.5D双路注意力编解码网络与2D注意力编解码网络本质差别并不大，所以相比于3D双路注意力编解码网络，2.5D双路注意力编解码网络探索三维数据的空间特征信息的能力是有限的，对三维医学图像的分割能力也是有限的。因此，2.5D网络更多作为2D网络与3D网络之间的一种折中选择 [100]。

表 4.3 基于双路注意力的2.5D编解码网络结构配置表

2.5D FCN	特征维度	网络模块			模块数量
输入	(N,5,224,224)	-			-
Layer 1	(N,32,224,224)	2D Residual Block	2D Residual Block	2D Residual Block	1
	(N,32,112,112)	2D Max pool			1
Layer 2	(N,64,112,112)	2D Residual Blocks	2D Dense Blocks	2D Dual Blocks	2
	(N,64,56,56)	2D Max pool			1
Layer 3	(N,128,56,56)	2D Residual Blocks	2D Dense Blocks	2D Dual Blocks	3
	(N,128,28,28)	2D Avg pool			1
Layer 4	(N,256,28,28)	2D Residual Blocks	2D Dense Blocks	2D Dual Blocks	5
	(N,256,14,14)	2D Avg pool			1
Layer 5	(N,512,14,14)	2D Residual Blocks	2D Dense Blocks	2D Dual Blocks	2
		2D Attention			1
Up Layer 4	(N,256,28,28)	2D Deconvolution Block			1
		2D Residual Blocks			2
Up Layer 3	(N,128,56,56)	2D Deconvolution Block			1
		2D Residual Blocks			2
Up Layer 2	(N,64,112,112)	2D Deconvolution Block			1
		2D Residual Blocks			2
Up Layer 1	(N,32,224,224)	2D Deconvolution Block			1
		2D Residual Blocks			2
Conv 0	(N,1,224,224)	2D Residual Block			1
输出	(N,1,224,224)	Sigmoid Function			1

4.4 实验结果与分析

4.4.1 实验环境

本章实验所用的服务器、程序开发语言和网络搭建框架均与第三章的介绍的实验环境相同。本章实验同样使用了LiTS和3DIRCADb数据集来训练和测试本章提出的网络模型。然而在对数据校检时，本文发现3DIRCADb数据集中部分数据存在标注缺失现象，因此本文剔除了3DIRCADb数据集中的部分数据，只保留了其中的15份肝脏肿瘤数据用于实验。

为了对实验结果进行统一比较和分析，在进行每个实验前，本文都会先使用3.1小节所述的方法对数据进行预处理，同时在每个实验过程中，当网络完成对肝脏肿瘤CT图像的预测分割后，同样统一使用3.5小节所述的方法对网络预测结果进行后处理，进一步提升分割精度。其次，在使用轮廓损失函数和边界损失函数优化网络时，本文将学习率设置为0.00005，当使用交叉熵损失函数、相似性系数损失函数和联合损失函数优化网络参数时，本文将学习率设置为0.001。为了保证能够充分训练网络模型，当损失不再下降时本文会将学习率乘以0.5的衰减因子，直到无法使损失进一步下降，才会停止网络的训练。最后，为了防止过拟合，本文的所有实验统一使用了Adam (Adaptive Moment Estimation)优化器优化网络，同时将权重衰减系数设置为0.0005。

4.4.2 结果与分析

为了对比不同网络的分割性能，本文基于边界损失函数设计了一组不同网络分割肝脏与肿瘤CT图像的实验。在该实验中，本文使用LiTS数据集的131份肝脏CT图像作为训练数据，3DIRCADb数据集的15份有标注肝脏CT图像作为测试数据。此外，不同维度的网络对输入数据的维度要求不同，因此在对训练数据进行预处理后，本文将肝脏与肿瘤CT图像分别裁剪成 $1 \times 224 \times 244$ 、 $5 \times 224 \times 244$ 和 $12 \times 224 \times 244$ 以适应2D、2.5D和3D网络的输入。同时，为了最大化利用显存训练不同维度的网络，在训练时本文将2D、2.5D和3D网络的输入批量(Batchsize)分别设置30，

20和6。

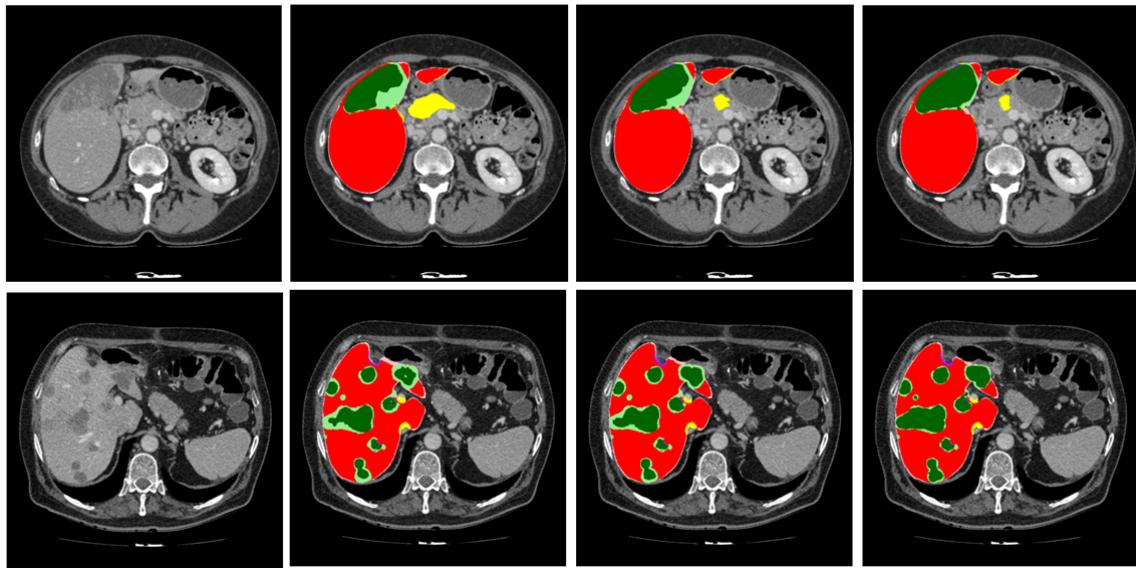


图 4.7 基于边界损失函数优化的三种不同维度的ADP FCN的分割结果对比图，第一列为肝脏与肿瘤CT切片预处理图，第二列至第四列分别为2D、2.5D和3D ADP FCN的分割结果图，其中肝脏分割真正例(红色)、肝脏分割假正例(黄色)、肝脏分割假反例(粉色)、肿瘤分割真正例(绿色)、肿瘤分割假正例(紫色)、肿瘤分割假反例(浅绿色)。

本文设计的基于边界损失函数的不同网络分割实验结果如表4.4所示。从表中数据可以看出，网络维度是影响分割效果的重要因素。在表4.4中，3D ADP FCN对肝脏和肿瘤的分割分数相比于2.5D ADP FCN分别提升了0.014和0.026。2.5D ADP FCN对肝脏和肿瘤的分割结果相较于2.5D ADP FCN分别提升了0.007和0.025。其中3D ADP FCN获得了最高的肝脏与肿瘤分割分数，分别为0.765和0.963。从整体上看，网络的维度越高其分割效果也就越好，3D网络的分割效果优于2.5D网络，而2.5D网络的分割效果则会优于2D网络。从图4.8中也可以看出当保持网络基础块不变，维度越高的网络其损失收敛点也越低。此外，如图4.7所示，对比三种不同维度的ADP FCN可以发现，维度越高的网络分割结果中假正例与假反例越少，这说明在相似网络的结构下，网络维度越高，其分割效果也就越好。而在表4.4中，3D AD FCN的肝脏和肿瘤分割分数比2.5D ADP FCN分别低了0.015和0.006，而2.5D AD FCN的肝脏分割分数又比2D ADP FCN低了0.011。此外，3D ADP FCN对肝脏与肿瘤CT图像的分割结果上要优于3D AR FCN，肝脏与肿瘤分割分数分别提升

了0.009和0.011，而3D AR FCN对肝脏与肿瘤CT图像的分割结果又要高于3D AD FCN，肝脏与肿瘤分割分数分别提升了0.032和0.012，这说明网络基本模块同样对分割的结果起着至关重要的作用。整体而言，基于双路径块构建的网络要优于基于残差块构建的网络，基于残差块构建的网络优于基于密集块构建的网络。

表 4.4 基于边界损失函数的九种不同网络对肝脏与肿瘤CT图像的实验分割结果对比

模型	肿瘤		肝脏	
	DPC	DG	DPC	DG
2D AD FCN	0.679	0.711	0.934	0.943
2D AR FCN	0.695	0.745	0.938	0.951
2D ADP FCN	0.714	0.762	0.941	0.956
2.5D AD FCN	0.703	0.737	0.938	0.952
2.5D AR FCN	0.732	0.751	0.945	0.961
2.5D ADP FCN	0.739	0.776	0.949	0.963
3D AD FCN	0.724	0.762	0.940	0.957
3D AR FCN	0.756	0.794	0.952	0.965
3D ADP FCN	0.765	0.813	0.963	0.972

图4.8展示了九种不同网络的训练损失随着迭代变化的曲线。在图4.8中，可以明显看出所有网络的训练损失随着迭代的进行最终会收敛至一个较小的区间内，在网络训练的前20个周期损失快速下降并逐渐收敛，而在后续的30个周期损失缓慢收敛并停留在一个较小区间内波动。在图中，3D ADP FCN的损失收敛点比其他网络的损失收敛点都要低，在维度相同的情况下，基于双路径块构建的网络的损失略低于基于残差块构建的网络的损失收敛点，远低于基于密集块构建的网络的损失收敛点。这说明双路径块确实吸收了残差块与密集块的优点，能够探索新的特征并对这些特征进行复用。而基于残差块构建的网络的损失收敛点同样远低于基于密集块构建的网络的损失收敛点，说明在医学图像分割任务中，残差块对浅层网络的特征复用的特性要优于密集块不断探索新特征的特性。在医学图像中，由于组织与器官的位置相对固定，因此实际上密集块探索的新特征有较多的冗余，导致其构建的网络分割效果不如基于残差块构建的网络。

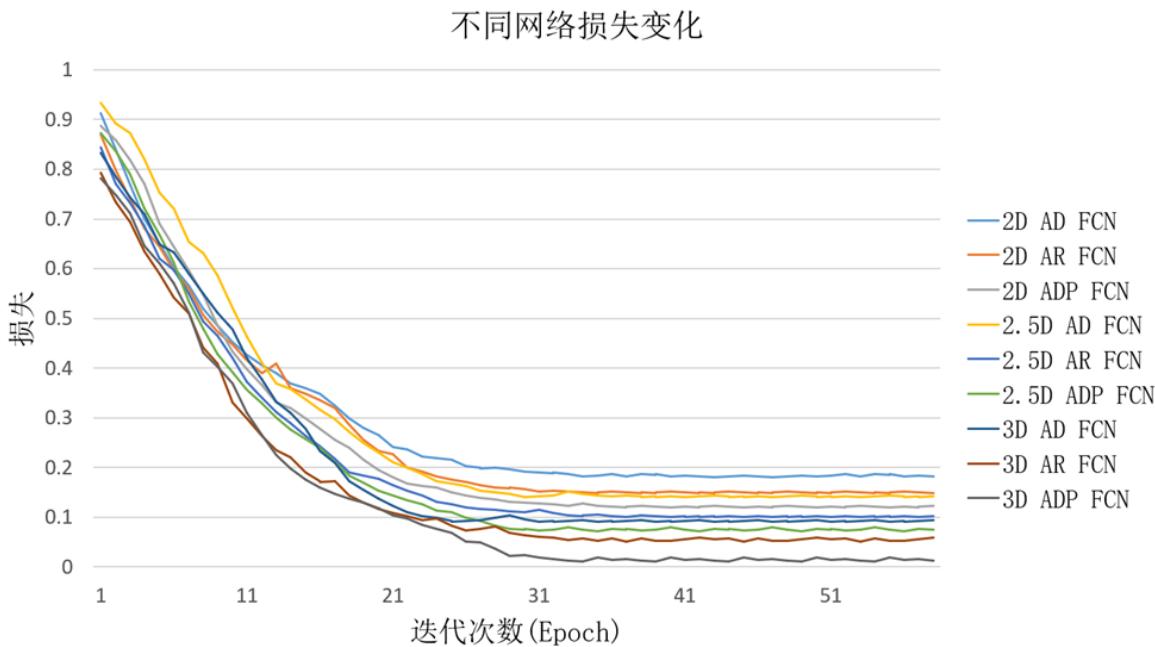


图 4.8 基于边界损失函数的九种不同网络训练损失变化曲线

为了证明双路注意力模块增强网络特征表达能力的有效性，本文使用边界损失函数优化无双路注意力模块的3D Dual Path FCN(3D DP FCN)和含有双路注意力模块的3D Attention Dual Path FCN(3D ADP FCN)，并在肝脏与肿瘤CT图像上进行了对比分割实验。在该实验中，本文使用LiTS的131份肝脏与肿瘤CT图像作为训练数据，3DIRCADb的15份有标注肝脏与肿瘤CT图像作为测试数据。同时本文将网络的学习率设置为0.00005、学习率衰减因子设置为0.5以及权重衰减系数设置为0.0005。此外，为了降低训练网络时的显存占用，避免显存溢出的情况发生，在将肝脏与肿瘤CT图像送入网络前，本文对其进行了裁剪，保持每份输入数据大小为 $12 \times 224 \times 244$ ，同时将输入批量Batchsize设置为6。

本文针对注意力机制模块设计的消融实验结果如表4.5所示。从表中可以看出在同样使用边界损失函数优化网络的情况下，3D ADP FCN获得了更高的肝脏与肿瘤分割分数，分别是0.963和0.765。在其他条件相同的情况下，3D ADP FCN仅比3D DP FCN多出一个双路注意力模块，而从图4.9也可以看出3D ADP FCN分割结果中的假正例更少（黄色，如黑色箭头所指），证明注意力模块确实能够整合编码器提取的高级语义特征，探索更多的位置特征信息与通道特征信息，提升网络的

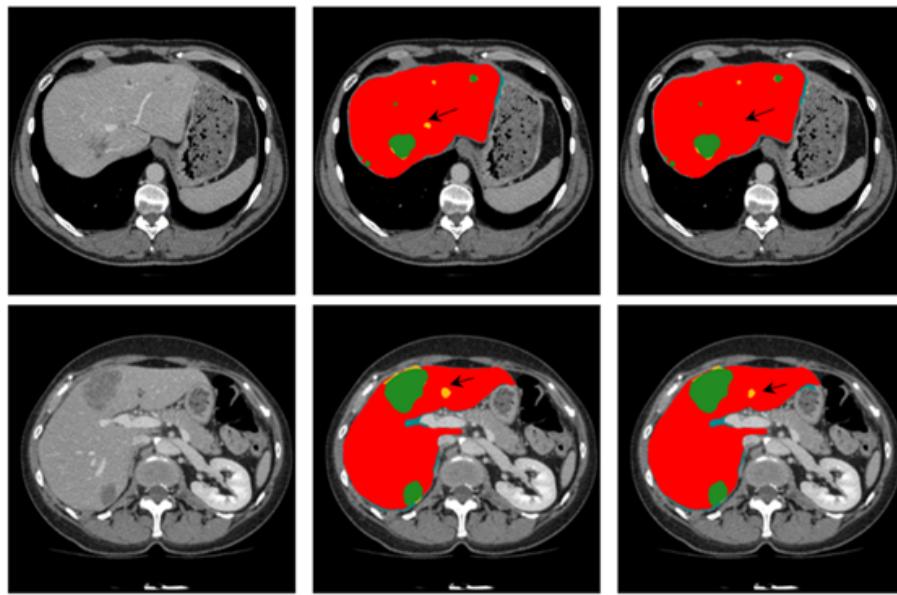


图 4.9 双路注意力模块消融实验的肝脏与肿瘤分割结果对比，第一列为肝脏与肿瘤CT切片预处理图，第二列与第三列分别为3D DP FCN和3D ADP FCN的分割结果图，其中肝脏分割真正例(红色)、肝脏分割假正例与假反例(蓝色)、肿瘤分割真正例(绿色)和肿瘤分割假正例与假反例(黄色，如黑色箭头所指)。

分割效果。

表 4.5 双路注意力模块消融实验的肝脏与肿瘤分割结果对比

模型	肿瘤		肝脏	
	DPC	DG	DPC	DG
3D DP FCN	0.739	0.782	0.945	0.961
3D ADP FCN	0.765	0.813	0.963	0.972

在表4.6和表4.7中，本文使用RVD、VOE、ASD以及RSME对双路注意力模块消融实验的分割结果进行了进一步的评估。表4.6是3D DP FCN与3D ADP FCN对肝脏分割的实验结果对比，表4.7是3D DP FCN与3D ADP FCN对肿瘤分割的实验结果对比。从这两张表中可以看出，3D ADP FCN在各个评价指标上都要优于3D DP FCN，进一步证明注意力机制模块能够有效增强网络的特征表达能力。

损失函数是深度学习分割方法的重要组成部分。为了验证本文提出的边界损失函数的有效性以及对比不同损失函数的优化效果，本文基于3D ADP FCN和不同的

表 4.6 双路注意力模块消融实验肝脏分割结果对比

模型	RVD	VOE	ASD	RSME
3D DP FCN	9.36 ± 3.34	0.29 ± 0.08	4.89 ± 2.08	7.15 ± 9.16
3D ADP FCN	5.57 ± 2.66	0.03 ± 0.02	2.28 ± 1.02	4.58 ± 7.58

表 4.7 双路注意力模块消融实验肿瘤分割结果对比

模型	RVD	VOE	ASD	RSME
3D DP FCN	30.61 ± 10.44	15.97 ± 12.04	4.18 ± 9.60	5.09 ± 10.71
3D ADP FCN	24.6 ± 13.8	-2.20 ± 15.88	0.72 ± 0.33	1.10 ± 0.49

损失函数设计了一组肝脏肿瘤分割对比实验。在该实验中，本文使用LiTS的131份肝脏与肿瘤CT图像作为训练数据，3DIRCADb的15份有标注肝脏与肿瘤CT图像作为测试数据。

不同损失函数优化3D ADP FCN的分割结果如表4.8所示。交叉熵损失函数本质是一种基于数据分布的损失函数，它通过最小化两个分布之间的差异性来优化网络。当目标数据分布不均时它对网络的优化效果是有限的。而肝脏与肿瘤CT图像中前景与背景的数据分布差异极大，因此使用交叉熵损失函数优化的网络获得了较低的0.946肝脏DG分数和0.714的肿瘤DG分数。相似性系数损失函数则是一种基于区域的损失函数，它通过最大化真实样本和预测样本之间的重叠区域来提升网络的分割表现。由于只考虑数据的重叠性，相似性系数损失函数可以在数据分布不均的情况下优化网络，其对网络的优化效果也要优于交叉熵损失函数。相比于交叉熵损失函数，相似性系数损失函数优化的网络对肝脏和肿瘤CT图像的分割分数分别提升了0.022和0.009。在深度学习中，联合多种不同损失函数往往能产生更优秀的网络优化效果，交叉熵损失函数与相似性系数损失函数能够组合出一种新的联合损失函数。从表4.8中可以看出联合损失函数优化的网络对肝脏和肿瘤的分割结果相较于交叉熵损失函数和相似性系数损失函数均有所提升。一个实验分割结果如图4.10所示，从图中可以看出不同损失函数对网络的优化效果不同，将第六列使用边界损失函数优化3D ADP FCN的分割结果与其他列的分割结果对比，可以发现边界损失优化3D ADP FCN的分割结果中，肝脏与肿瘤的假正例和假反例更少，一定

程度上证明了边界损失函数的有效性。

表 4.8 不同损失函数优化3D ADP FCN的肝脏与肿瘤分割结果对比

损失	肿瘤		肝脏	
	DPC	DG	DPC	DG
交叉熵损失	0.682	0.714	0.928	0.946
相似性系数损失	0.690	0.736	0.944	0.953
联合损失	0.724	0.758	0.952	0.961
轮廓损失	0.759	0.793	0.959	0.967
边界损失	0.765	0.813	0.963	0.972

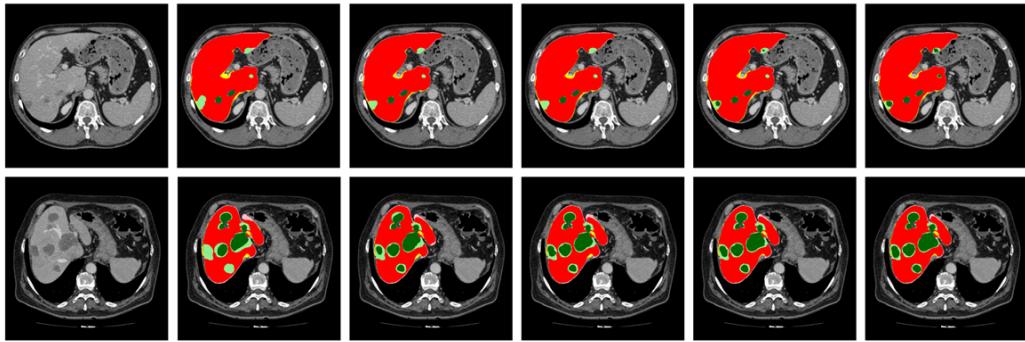


图 4.10 基于不同损失函数优化3D ADP FNC的肝脏与肿瘤分割结果对比图，第一列为肝脏与肿瘤CT切片预处理图，第二列至第六列分别为使用交叉熵损失函数、相似性系数损失函数、联合损失函数、轮廓损失函数和边界损失函数优化3DADP FNC分割肝脏与肿瘤CT图像的结果图，其中肝脏分割真正例(红色)、肝脏分割假正例(黄色)、肝脏分割假反例(粉色)、肿瘤分割真正例(绿色)、肿瘤分割假正例(紫色)、肿瘤分割假反例(浅绿色)。

然而上述的损失函数都只是基于分布和区域来优化网络，缺少对图像轮廓与边界特征的探索。轮廓损失函数可以看作改进的相似性系数损失函数。它同时考虑了长度和区域作为网络优化的目标来提升网络的分割效果。在图4.11中，轮廓损失函数优化的网络对肿瘤的分割分数有较大的提升，达到了0.759，相较于联合损失函数提升了0.035。尽管交叉熵损失函数、相似性系数损失函数，联合损失函数以及轮廓损失函数都能够有效地优化网络分割医学图像，但它们仍然有着提升的空间。事实上，由于人体的器官和组织相对固定且具有一定的边界，因此将边界特征作为网络优化的目标能够使得网络探索更多的边界特征，从而提升网络的分割能力。本文提出的边界损失函数也正是基于这种想法，将边界信息融入轮廓损失函数，从而

进一步优化网络和提升分割效果。从图4.11中可以看出基于边界损失函数优化的网络获得了最高的肝脏与肿瘤DPC分数，分别达到了0.963和0.765。

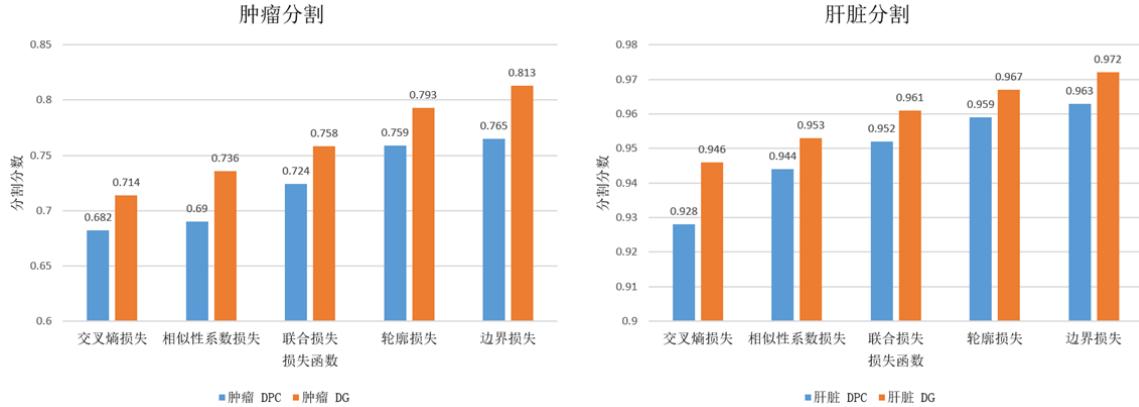


图 4.11 基于不同损失函数的网络分割肝脏与肿瘤结果

表 4.9 基于3D ADP FCN的肝脏与肿瘤CT图像分割结果与现有的方法结果对比

模型	肿瘤		肝脏	
	DPC	DG	DPC	DG
2D UNet [93]	0.65	-	-	-
2D FCN [81]	0.67	-	-	-
3D V-Net [37]	-	-	-	0.939
Med3D [99]	-	-	-	0.946
Wang 等人 [101]	-	0.702	-	0.951
H-DenseUNet [17]	0.722	0.824	0.961	0.965
3D ADP FCN	0.736	0.781	0.952	0.966

为了进一步证明本文提出的双路注意力编解码网络的有效性，本文选用分割效果最好的3D ADP FCN进行肝脏与肿瘤分割实验，并与现有的一些方法进行了对比。在该实验中，本文使用LiTS的131份肝脏与肿瘤CT图像作为训练数据，LiTS的70份无标注肝脏与肿瘤CT图像作为测试数据。实验结果如表4.9所示，相比于其他方法，本文的方法在肝脏与肿瘤的分割任务中获得了最高的肝脏DG分数0.966和最高的肿瘤DPC分数0.736。

4.5 本章小节

本章首先详细介绍了基于双路注意力的编解码网络框架的肝脏与肿瘤CT图像分割流程，接着重点阐述了本文设计的编解码网络框架，然后介绍了改进的双路注意力模块，并将该模块融入到本文的编解码网络框架之中，从不同的网络维度分析了本文的九种双路注意力编解码网络结构。最后本文在LiTS和3DIRCADb肝脏与肿瘤数据集上从不同的角度对本章提出的方法进行了验证，实验结果表明了本文提出的方法在肝脏与肿瘤分割任务上是有效的。

第五章 结论与展望

5.1 结论

准确地测量肝脏与肝肿瘤的位置、大小和形状能够帮助医生针对病人病情做出更全面的诊断和治疗规划。在第三章，本文首先充分研究了腹部扫描CT图像的特点，针对性地对CT图像进行预处理，然后基于FCN提出了一种新的2.5D全卷积网络，在有效减少网络的参数量和对计算机显存资源需求的同时，增强网络对三维数据空间特征信息的探索能力，提升模型的分割效果。然而该网络具有结构相对简单的缺点，同时缺少对图像局部特征之间相关性的探索能力。其次本文设计了一种边界损失函数，它融合了图像中的距离、面积和边界特征信息，能够有效地优化深度学习网络学习更多边界与轮廓特征。该损失函数的局限是对网络的参数变化敏感，在训练网络的初始阶段难以收敛，因此需要使用相关损失函数进行预训练。此外本文还提出了一种针对网络分割的肝脏与肝肿瘤CT图像的后处理方法，利用密集条件随机场和最大连通区域减少深度学习网络预测结果中的假正例和假反例，提升图像分割的准确度。该后处理方法的问题是当网络的分割效果较差时，难以产生有效的优化提升。在第四章，为了进一步优化网络结构，本文结合双路注意力机制模块、密集网络块、残差网络块和双路径网络块设计了一种基于双路注意力的编解码网络分割框架，它包含了九种不同的网络结构，能够有效地探索图像中的局部特征信息，提升对CT图像的分割准确度。然而该网络框架中的网络参数量大，含有一些冗余的卷积操作，对计算机硬件资源要求较高。最后，本文对上述方法在公开的肝脏与肿瘤数据集上进行了实验验证和分析，并将本文分割效果最好的3D ADP FCN与现有的一些方法进行了对比，实验结果表明本文提出的方法对肝脏与肿瘤CT图像具有优秀的分割能力，同时具备良好的鲁棒性和泛用性，能够为临床研究提供一定的参考价值，同时为智能医疗的早日实现提供一定的经验基础。

本文的主要研究成果和贡献如下：针对腹部扫描CT图像中肝脏与肝肿瘤的特点，提出了一种基于边界损失函数的2.5D网络，提升肝脏CT图像的分割准确度；针

对CT图像分割目标边界模糊的问题，提出一种边界损失函数，能够有效强化网络对图像边界与轮廓特征信息的学习能力；为了进一步优化网络分割的结果，提出一种基于密集条件随机场和最大连通区域的后处理方法，减少网络预测结果中的假正例和假反例；为了解决编解码网络缺少探索图像局部特征之间相关性和依赖性的问题，本文结合双路注意力模块、密集网络块、残差网络块和双路径网络块提出了一种新的双路注意力网络分割框架，有效增强了网络学习图像局部特征之间相关性和依赖性的能力，提升了网络的分割效果。

5.2 展望

本文主要结合了多种深度学习技术提出一种基于边界损失函数的2.5D全卷积网络和一种基于双路注意力的编解码网络框架分割腹部扫描CT图像中的肝脏与肝肿瘤。实验充分证明了深度学习方法是自动化分割医学图像的强有力工具。近年来，各种优秀的深度学习技术层出不穷，本文的方法虽然取得了一定的成果，但仍然存在一些有待解决的问题，需要进一步优化提升。

首先由于不同的CT图像使用不同的协议采集而来，导致每一份腹部CT扫描图像中的肝脏与肝肿瘤CT值分布范围并不一致，使用固定的CT值截断方法会造成部分图像信息丢失。因此在未来的工作中将会考虑使用变化的CT值截断方法，从而最大程度减少预处理过程中CT图像数据的信息丢失问题。其次，本文提出的边界损失函数对网络参数的变化非常敏感，必须使用相关损失函数进行预训练才能发挥应有的作用。在接下来的工作中，将考虑进一步优化边界损失函数，减轻网络训练时对参数变化的敏感度。尽管3D网络模型能够充分探索CT图像中的空间特征信息，有效地分割肝脏与肝肿瘤，但是模型需要训练大量的参数和占用高额的计算资源，因此在后续的工作中将考虑优化网络结构，减少网络中的冗余卷积，降低网络参数量和对计算资源的需求。此外，虽然级联的分割策略能够在大多数情况下更准确地分割肿瘤，但是当肝脏区域分割错误，就会影响到肿瘤的分割，因此在未来的研究中将考虑寻找更好的分割策略，从而避免肝脏分割对肿瘤分割造成的影响。最后，本文的方法只考虑分割腹部CT扫描图像中的肝脏与肝肿瘤，但是腹部CT图像

中还存在其他组织和器官，后续研究将考虑把本文的方法应用到多器官与多模态的分割任务中。

参考文献

- [1] ANWANWAN D, SINGH S, SINGH S, et al. Challenges in Liver Cancer and Possible Treatment Approaches.[J]. Biochimica et biophysica acta. Reviews on cancer, 2019: 188314.
- [2] ABRAHAM G, OBLEY A J, HUMPHREY L, et al. World Health Organization Guidelines on Treatment of Hepatitis C Virus Infection: Best Practice Advice from the American College of Physicians[J]. Annals of Internal Medicine, 2020, 174: 98 – 100.
- [3] ZHOU J, SUN H-C, WANG Z, et al. Guidelines for Diagnosis and Treatment of Primary Liver Cancer in China (2017 Edition)[J]. Liver Cancer, 2018, 7: 235 – 260.
- [4] WEI W, ZHOU B, POŁAP D, et al. A Regional Adaptive Variational PDE Model for Computed Tomography Image reconstruction[J]. Pattern Recognit., 2019, 92: 64–81.
- [5] VARMA D. Managing DICOM Images: Tips and Tricks for The Radiologist[J]. The Indian Journal of Radiology and Imaging, 2012, 22: 4 – 13.
- [6] YUAN J, WANG D, CHERIYADAT A. Factorization-Based Texture Segmentation[J]. IEEE Transactions on Image Processing, 2015, 24: 3488–3497.
- [7] DAVARPANAHA A. Diagnosis and Staging of Hepatocellular Carcinoma[J]. Iranian Journal of Radiology, 2017, 14.
- [8] WOŹNIAK M, POŁAP D, CAPIZZI G, et al. Small Lung Nodules Detection Based on Local Variance Analysis and Probabilistic Neural Network[J]. Computer methods and programs in biomedicine, 2018, 161: 173–180.
- [9] EL-REGAILY S, SALEM M, AZIZ M H A, et al. Multi-view Convolutional Neural Network for Lung Nodule False Positive Reduction[J]. Expert Syst. Appl., 2020, 162: 113017.
- [10] ATTEYA M, SALEM M A-M, HEGAZY D, et al. Image Segmentation and Particles Classification Using Texture Analysis Method[J]. Research on Biomedical Engineering, 2016, 32: 243–252.
- [11] HAN X. Automatic Liver Lesion Segmentation Using A Deep Convolutional Neural Network Method[J]. ArXiv, 2017, abs/1704.07239.

- [12] YAN J, SCHWARTZ L, ZHAO B. Segmentation of Liver Metastases on Volumetric CT Images.[J]. Medical physics, 2015, 42 11 : 6283 – 93.
- [13] LI B, LIU S, XU W, et al. Real-time Object Detection and Semantic Segmentation for Autonomous Driving[C] // International Symposium on Multispectral Image Processing and Pattern Recognition. 2018.
- [14] ZHANG Y, CHEN H, HE Y, et al. Road Segmentation for All-day Outdoor Robot Navigation[J]. Neurocomputing, 2018, 314 : 316 – 325.
- [15] JIANG F, GRIGOREV A, RHO S, et al. Medical Image Semantic Segmentation Based on Deep Learning[J]. Neural Computing and Applications, 2017, 29 : 1257 – 1265.
- [16] TJOA E, GUAN C. A Survey on Explainable Artificial Intelligence (XAI): Towards Medical XAI[J]. IEEE transactions on neural networks and learning systems, 2020, PP.
- [17] LI X, CHEN H, QI X, et al. H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation from CT Volumes[J]. IEEE Transactions on Medical Imaging, 2018.
- [18] OTSU N. A Threshold Selection Method from Gray Level Histograms[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9 : 62 – 66.
- [19] FRANK R J, GRABOWSKI T J, DAMASIO H. Voxelize Percentage Tissue Segmentation of Human Brain Magnetic Resonance Images[J]. Society for Neuroscience, 1995.
- [20] NALWA V S, BINFORD T O. On Detecting Edges[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, PAMI-8(6) : 699 – 714.
- [21] WONG D, LIU J, FENGSHOU Y, et al. A Semi-automated Method for Liver Tumor Segmentation Based on 2D Region Growing with Knowledge-based Constraints[C] // MICCAI workshop : Vol 41. 2008 : 159.
- [22] JIANG H, GAO X. Semi-Automatic Liver Segmentation Using Improved GVF Snake Model[J]. Advanced Materials Research, 2010, 121-122 : 435 – 440.
- [23] WELLS W, GRIMSON W, KIKINIS R, et al. Adaptive Segmentation of MRI Data[C] // CVRMed. 1995.

- [24] VORONTSOV E, ABI-JAOUDEH N, KADOURY S. Metastatic Liver Tumor Segmentation Using Texture-Based Omni-Directional Deformable Surface Models[C] // ABDI@MICCAI. 2014.
- [25] KUO C, CHENG S-C, LIN C, et al. Texture-based Treatment Prediction by Automatic Liver Tumor Segmentation on Computed Tomography[J]. 2017 International Conference on Computer, Information and Telecommunication Systems (CITS), 2017: 128–132.
- [26] JIMENEZ-CARRETERO D, de MANUEL L F, PASCAU J, et al. Optimal Multiresolution 3D Level-set Method for Liver Segmentation Incorporating Local Curvature Constraints[J]. 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2011: 3419–3422.
- [27] LE T-N, BAO P, HUYNH H T. Liver Tumor Segmentation from MR Images Using 3D Fast Marching Algorithm and Single Hidden Layer Feedforward Neural Network[J]. BioMed Research International, 2016, 2016.
- [28] HUANG W, YANG Y, LIN Z, et al. Random Feature Subspace Ensemble Based Extreme Learning Machine for Liver Tumor Detection and Segmentation[J]. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2014: 4675–4678.
- [29] HUANG G, ZHU Q-Y, SIEW C. Extreme Learning Machine: A New Learning Ccheme of Feedforward Neural Networks[J]. 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541), 2004, 2: 985–990 vol.2.
- [30] CHEN L-C, PAPANDREOU G, SCHROFF F, et al. Rethinking Atrous Convolution for Semantic Image Segmentation[J]. ArXiv, 2017, abs/1706.05587.
- [31] NOH H, HONG S, HAN B. Learning Deconvolution Network for Semantic Segmentation[J]. 2015 IEEE International Conference on Computer Vision (ICCV), 2015: 1520–1528.
- [32] LEI T, WANG R, WAN Y, et al. Medical Image Segmentation Using Deep Learning: A Survey[J]. ArXiv, 2020, abs/2009.13120.

- [33] SHELHAMER E, LONG J, DARRELL T. Fully Convolutional Networks for Semantic Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39 : 640–651.
- [34] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation[J]. ArXiv, 2015, abs/1505.04597.
- [35] CHEN L-C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40 : 834–848.
- [36] ÇIÇEK Ö, ABDULKADIR A, LIENKAMP S, et al. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation[J]. ArXiv, 2016, abs/1606.06650.
- [37] MILLETARI F, NAVAB N, AHMADI S-A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation[J]. 2016 Fourth International Conference on 3D Vision (3DV), 2016 : 565–571.
- [38] MIKOLOV T, KARAFIÁT M, BURGET L, et al. Recurrent Neural Network Based Language Model[C] // INTERSPEECH. 2010.
- [39] HOCHREITER S, SCHMIDHUBER J. Long Short-Term Memory[J]. Neural Computation, 1997, 9: 1735–1780.
- [40] 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, 2018, abs/1802.06955.
- [41] CHRIST P, ELSHAER M E A, ETTLINGER F, et al. Automatic Liver and Lesion Segmentation in CT Using Cascaded Fully Convolutional Neural Networks and 3D Conditional Random Fields[C] // MICCAI. 2016.
- [42] TANG W, ZOU D, YANG S, et al. DSL: Automatic Liver Segmentation with Faster R-CNN and DeepLab[C] // ICANN. 2018.
- [43] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42 : 386–397.

- [44] BOCHKOVSKIY A, WANG C-Y, LIAO H. YOLOv4: Optimal Speed and Accuracy of Object Detection[J]. ArXiv, 2020, abs/2004.10934.
- [45] AL-ANTARI M A, AL-MASNI M A, CHOI M-T, et al. A Fully Integrated Computer-aided Diagnosis System for Digital X-ray Mammograms via Deep Learning Detection, Segmentation, and Classification[J]. International journal of medical informatics, 2018, 117: 44–54.
- [46] ODA M, ROTH H, KITASAKA T, et al. Abdominal Artery Segmentation Method from CT Volumes Using Fully Convolutional Neural Network[J]. International Journal of Computer Assisted Radiology and Surgery, 2019, 14: 2069 – 2081.
- [47] VU M H, GRIMBERGEN G, NYHOLM T, et al. Evaluation of Multi-Slice Inputs to Convolutional Neural Networks for Medical Image Segmentation[J]. ArXiv, 2019, abs/1912.09287.
- [48] CHAUDHARI S, POLATKAN G, RAMANATH R, et al. An Attentive Survey of Attention Models[J]. ArXiv, 2019, abs/1904.02874.
- [49] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-Net: Learning Where to Look for The Pancreas[J]. ArXiv, 2018, abs/1804.03999.
- [50] CHEN X, ZHANG R, YAN P. Feature Fusion Encoder Decoder Network for Automatic Liver Lesion Segmentation[J]. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019: 430–433.
- [51] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42: 2011 – 2023.
- [52] SUDRE C, LI W, VERCAUTEREN T K M, et al. Generalised Dice Overlap as A Deep Learning Loss Function for Highly Unbalanced Segmentations[J]. Deep learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support : Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, held in conjunction with MICCAI 2017 Quebec City, QC,..., 2017, 2017: 240 – 248.
- [53] GOODFELLOW I, BENGIO Y, COURVILLE A C. Deep Learning[J]. Nature, 2015, 521: 436 – 444.

- [54] KALAN S M M, SOLTANOLKOTABI M, AVESTIMEHR A. Fitting ReLUs via SGD and Quantized SGD[J]. 2019 IEEE International Symposium on Information Theory (ISIT), 2019: 2469–2473.
- [55] ZEILER M D, KRISHNAN D, TAYLOR G W, et al. Deconvolutional networks[J]. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010: 2528–2535.
- [56] ZEILER M D, TAYLOR G W, FERGUS R. Adaptive Deconvolutional Networks for Mid and High Level Feature Learning[J]. 2011 International Conference on Computer Vision, 2011: 2018–2025.
- [57] LONG J, SHELHAMER E, DARRELL T. Fully Convolutional Networks for Semantic Segmentation[J]. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 3431–3440.
- [58] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39: 2481–2495.
- [59] ZHAO H, SHI J, QI X, et al. Pyramid Scene Parsing Network[J]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 6230–6239.
- [60] SZEGEDY C, LIU W, JIA Y, et al. Going Deeper with Convolutions[J]. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 1–9.
- [61] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the Inception Architecture for Computer Vision[J]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 2818–2826.
- [62] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, Inception-ResNet and The Impact of Residual Connections on Learning[C] // AAAI. 2017.
- [63] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[J]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770–778.

- [64] HUANG G, LIU Z, WEINBERGER K Q. Densely Connected Convolutional Networks[J]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 2261–2269.
- [65] ZHANG Y, TIAN Y, KONG Y, et al. Residual Dense Network for Image Super-Resolution[J]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 2472 – 2481.
- [66] CHEN Y, LI J, XIAO H, et al. Dual Path Networks[C] // NIPS. 2017.
- [67] IOFFE S, SZEGEDY C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[J]. ArXiv, 2015, abs/1502.03167.
- [68] MA J, CHEN J, NG M, et al. Loss Odyssey in Medical Image Segmentation.[J]. Medical image analysis, 2021, 71: 102035.
- [69] GIRSHICK R. Fast R-CNN[C] // Proceedings of the IEEE international conference on computer vision. 2015: 1440–1448.
- [70] WU Z, SHEN C, HENGEL A V. Bridging Category-level and Instance-level Semantic Image Segmentation[J]. ArXiv, 2016, abs/1605.06885.
- [71] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal Loss for Dense Object Detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, PP(99): 2999 – 3007.
- [72] RAHMAN M, WANG Y. Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation[C] // ISVC. 2016.
- [73] SALEHI S, ERDOĞMUŞ D, GHOLIPOUR A. Tversky Loss Function for Image Segmentation Using 3D Fully Convolutional Deep Networks[J]. ArXiv, 2017, abs/1706.05721.
- [74] KARIMI D, SALCUDEAN S E. Reducing the Hausdorff Distance in Medical Image Segmentation With Convolutional Neural Networks[J]. IEEE Transactions on Medical Imaging, 2020, 39(2): 499–513.
- [75] CHEN X, WILLIAMS B M, VALLABHANENI S, et al. Learning Active Contour Models for Medical Image Segmentation[J]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 11624 – 11632.

- [76] HELLER N, ISENSEE F, MAIER-HEIN K, et al. The State of The Art in Kidney and Kidney Tumor Segmentation in Contrast-enhanced CT Imaging: Results of The KiTS19 Challenge[J]. *Medical image analysis*, 2021, 67: 101821.
- [77] ZHANG Q, WU Y, ZHU S. Interpretable Convolutional Neural Networks[J]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 8827–8836.
- [78] CHU X, YANG W, OUYANG W, et al. Multi-context Attention for Human Pose Estimation[J]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 5669–5678.
- [79] XU K, BA J, KIROS R, et al. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention[C] // ICML. 2015.
- [80] FU J, LIU J, TIAN H, et al. Dual Attention Network for Scene Segmentation[J]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 3141–3149.
- [81] CHRIST P, ETTLINGER F, GRÜN F, et al. Automatic Liver and Tumor Segmentation of CT and MRI Volumes Using Cascaded Fully Convolutional Neural Networks[J]. ArXiv, 2017, abs/1702.05970.
- [82] LI C, TAN Y, CHEN W, et al. ANU-Net: Attention-based Nested U-Net to Exploit Full Resolution Features for Medical Image Segmentation[J]. *Comput. Graph.*, 2020, 90: 11–20.
- [83] ZORMPAS-PETRIDIS K, FAILMEZGER H, RAZA S, et al. Superpixel-Based Conditional Random Fields (SuperCRF): Incorporating Global and Local Context for Enhanced Deep Learning in Melanoma Histopathology[J]. *Frontiers in Oncology*, 2019, 9.
- [84] GÖTZ M, HEIM E, MÄRZ K, et al. A Learning-based, Fully Automatic Liver Tumor Segmentation Pipeline Based on Sparsely Annotated Training Data[C] // SPIE Medical Imaging. 2016.
- [85] ZHANG Y, BRADY M, SMITH S M. Segmentation of Brain MR Images Through A Hidden Markov Random Field Model and The Expectation-maximization Algorithm[J]. *IEEE Transactions on Medical Imaging*, 2001, 20: 45–57.

- [86] CHLEBUS G, MEINE H, MOLTZ J, et al. Neural Network-Based Automatic Liver Tumor Segmentation With Random Forest-Based Candidate Filtering[J]. ArXiv, 2017, abs/1706.00842.
- [87] ALAM F, ZHOU J, LIEW A, et al. Conditional Random Field and Deep Feature Learning for Hyperspectral Image Classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57: 1612 – 1628.
- [88] KRÄHENBÜHL P, KOLTUN V. Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials[C] // NIPS. 2011.
- [89] PASZKE A, GROSS S, MASSA F, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library[C] // NeurIPS. 2019.
- [90] BILIC P, CHRIST P, VORONTSOV E, et al. The Liver Tumor Segmentation Benchmark (LiTS)[J]. ArXiv, 2019, abs/1901.04056.
- [91] L SOLER A H. 3D Image Reconstruction for Comparison of Algorithm Database: A Patient-specific Anatomical and Medical Image Database[J]. Dircadb3, 2012.
- [92] KINGMA D P, BA J. Adam: A Method for Stochastic Optimization[J]. CoRR, 2015, abs/1412.6980.
- [93] CHLEBUS G, MEINE H, MOLTZ J H, et al. Neural Network-based Automatic Liver Tumor Segmentation with Random Forest-based Candidate Filtering[J]. arXiv preprint arXiv:1706.00842, 2017.
- [94] HAN X. Automatic liver lesion segmentation using a deep convolutional neural network method[J]. arXiv preprint arXiv:1704.07239, 2017.
- [95] LI G, CHEN X, SHI F, et al. Automatic liver segmentation Based on Shape Constraints and Deformable Graph Cut in CT Images[J]. IEEE Transactions on Image Processing, 2015, 24(12): 5315 – 5329.
- [96] LI C, WANG X, EBERL S, et al. A Likelihood and Local Constraint Level Set Model for Liver Tumor Segmentation from CT Volumes[J]. IEEE Transactions on Biomedical Engineering, 2013, 60(10): 2967 – 2977.

- [97] CHLEBUS G, SCHENK A, MOLTZ J H, et al. Automatic Liver Tumor Segmentation in CT with Fully Convolutional Neural Networks and Object-based Postprocessing[J]. *Scientific reports*, 2018, 8(1): 1–7.
- [98] YUAN Y. Hierarchical Convolutional-deconvolutional Neural Networks for Automatic Liver and Tumor Segmentation[J]. arXiv preprint arXiv:1710.04540, 2017.
- [99] CHEN S, MA K, ZHENG Y. Med3D: Transfer Learning for 3D Medical Image Analysis[J]. ArXiv, 2019, abs/1904.00625.
- [100] ZHANG Y, LIAO Q, ZHANG J. Exploring Efficient Volumetric Medical Image Segmentation Using 2.5D Method: An Empirical Study[J]. ArXiv, 2020, abs/2010.06163.
- [101] WANG L, ZHANG R, CHEN Y, et al. The Application of Panoramic Segmentation Network to Medical Image Segmentation[J]. 2020 15th IEEE International Conference on Signal Processing (ICSP), 2020, 1: 640–645.

作者在攻读硕士学位期间取得的研究成果

攻读硕士期间发表论文情况：

1. Yuexing Han, Xiaolong Li, Bing Wang and Lu Wang. Boundary Loss-Based 2.5D Fully Convolutional Neural Networks Approach for Segmentation: A Case Study of the Liver and Tumor on Computed Tomography. *Algorithms*, 2021.
2. Yuexing Han*, Bing Wang, Jie Luo, Long Li, Xiaolong Li. A Classification Method for EEG Motor Imagery Signals Based on Parallel Convolutional Neural Network. *Biomedical Signal Processing and Control* 71 (2022) 103190, 2022.

攻读硕士期间专利获取情况：

1. 基于双路注意力编解码网络的三维图像分割方法，发明人：韩越兴，李小龙，钱权，王冰。申请号：202110863600.0，申请日：2021.07.29。

攻读硕士期间软著获取情况：

1. 基于双路径网络的肝脏与肿瘤CT图像自动化分割软件V1.0，申请人：上海大学，开发人：韩越兴，李小龙。登记号：2021SR1229122，开发完成日期：2021.6.21，登记日期：2021.08.19。

作者在攻读硕士学位期间所参与的项目

1. 上海市科学技术委员会面上项目名称：小样本环境下物体自适应识别方法研究，上海市自然科学基金，编号：20ZR1419000，日期：2020.07-2023.06。
2. 国家重点研发计划：材料基因工程专用数据库平台建设与示范应用（科技部）中“陶瓷涂层材料数据自动采集、整合与应用”课题中的子课题“材料复杂显微结构高精度图像处理方法研究”，编号：2018YFB0704400, 2018YFB0704402，日期：2018.07-2022.06。
3. 国家自然科学基金（青年）：面向医疗器械软件产品家族的形式化开发与验证，项目编号：61602293，日期：2017.01-2019.12。
4. 上海市科学技术委员会的浦江人才项目：基于形状空间理论和复杂网络理论的纳米材料图像处理方法的研究，编号：17PJ1402900，日期：2017.06-2019.05。

致谢

本文是在导师韩越兴教授的悉心指导下完成的。承蒙韩老师的亲切关怀和精心指导，在繁忙的教学、科研工作之余，对我的硕士论文选题、框架结构、资料收集、行文表述、乃至格式修改都倾注了大量的心血。韩老师对学生认真负责的态度、严谨的科学的研究方法、敏锐的学术洞察力、勤勉的工作作风以及勇于创新、勇于开拓的精神是我永远学习的榜样。在此，谨向韩老师致以深深的敬意和由衷的感谢。

还要感谢我的父母，他们在生活上给予我很大的支持和鼓励，是他们一直在背后默默为我付出，让我能够专心完成学业。

最后，感谢所有关心我、支持我和帮助过我的同学、朋友、老师和亲人。在这里，我仅用一句话来表明我无法言语的心情：感谢你们！