

审定成绩：_____

重庆邮电大学

毕业设计（论文）

中文题目	面向手术导航的脊柱医学图像分割 方法研究与实现
英文题目	Research and Implementation of Spinal Medical Image Segmentation Method for Surgical Navigation
学院名称	计算机科学与技术学院/人工智能学院
学生姓名	苟金和
专 业	智能科学与技术/人工智能
班 级	04812101
学 号	2021211996
指导教师	舒禹程 副教授
答 辩 组 负 责 人	

2025 年 4 月 9 日

重庆邮电大学教务处制。

计算机科学与技术/人工智能学院本科毕业设计(论文)诚信承诺书

本人郑重承诺：

我向学院呈交的论文《面向手术导航的脊柱医学图像分割方法研究与实现》，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。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品成果。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标明并致谢。本人完全意识到本声明的法律结果由本人承担。

年级 2021

专业 智能科学与技术/人工智能

班级 04812101

承诺人签名 葛金和


2025 年 4 月 9 日

学位论文版权使用授权书

本人完全了解重庆邮电大学有权保留、使用学位论文纸质版和电子版的规定，即学校有权向国家有关部门或机构送交论文，允许论文被查阅和借阅等。本人授权重庆邮电大学可以公布本学位论文的全部或部分内容，可编入有关数据库或信息系统进行检索、分析或评价，可以采用影印、缩印、扫描或拷贝等复制手段保存、汇编本学位论文。

（注：保密的学位论文在解密后适用本授权书。）

学生签名： 葛金和

指导老师签名： 

日期： 2025 年 4 月 9 日

日期： 2025 年 4 月 9 日

摘要

脊柱是生命的支柱，是人体的重要支撑结构。然而，脊柱疾病现已被认为是“全球最易被忽视的十大健康问题”之一。“2025 年脊柱健康学术大会”指出，在我国，中老年人群脊柱疾病的患病率高达 97%，成年人颈椎病和腰椎病的发病率分别为 30% 和 80%。随着社会发展和生活压力的增加，脊柱疾病呈现年轻化趋势。根据国家卫健委的数据，中国腰椎疾病患者已超过 2 亿人，其中 25~39 岁人群中椎间盘突出检出率达 13.93%，居各年龄段首位。2020 年，全国腰痛患者约 6.19 亿，预计到 2050 年，这一数字将增至 8.43 亿。

由于脊柱手术属于高精度的深部手术，传统徒手操作往往导致切口大、出血多、翻修率高等问题。因此，提高脊柱手术的精准度和安全性成为亟待解决的难题。本研究面向手术导航系统，提出了一种基于深度学习的脊柱医学图像分割方法，以辅助术中实时导航，提高手术精度。

本研究采用 Dense-3D U-Net 深度学习模型，对脊柱 CT 图像进行自动分割。本研究含有诸多创新点：在面对直接 `resize` 归一化图像大小导致图像各个通道以不同比例放缩造成比例失真的难题时，创新性地提出了先降采样到相同体素，再通过 `padding` 补齐图像的方法，使得图像在能够保证大小一致的同时不会比例失调。在面对图像 `loss` 无法收敛的难题时，创新型地提出了以 `BCELoss` 和 `DiceLoss` 加权形成的 `FusionLoss`，解决了单一 `Loss` 无法难收敛的问题。随后，利用 Dense-3D U-Net 进行端到端训练，并优化网络结构与损失函数提升模型性能。实验结果表明，该方法能够精准分割脊柱结构，并在不同病人数据上表现出良好的泛化能力。

该分割方法被集成到手术导航系统中，结合点云配准技术，使术前或术中影像数据与手术床上的病人解剖结构精准对齐。医生可通过 VR 可视化系统实时获取脊柱解剖信息，精确定位手术区域，从而降低手术难度，提高普通主治医师的手术可操作性，并减少手术创伤和风险。该系统的应用有望推动脊柱手术的微创化发展，提高临床手术质量，具有重要的临床价值和社会效益。

关键词：脊柱手术导航、医学图像分割、深度学习、Dense-3D U-Net

Abstract

The spine is hailed as the pillar of life and is an important supporting structure of the human body. However, spinal diseases have been listed as one of the "Top Ten Neglected Health Issues Globally". The "2025 Academic Conference on Spinal Health" pointed out that in China, the prevalence rate of spinal diseases among the elderly and middle-aged population is as high as 97%, while the incidence rates of cervical spondylosis and lumbar spondylosis among adults are 30% and 80% respectively. With the development of society and the increase in life pressure, spinal diseases are showing a trend of becoming more common among the young. According to data from the National Health Commission of China, the number of patients with lumbar spine diseases in China has exceeded 200 million, among which the detection rate of intervertebral disc protrusion in the 25-39 age group is 13.93%, ranking first among all age groups. In 2020, there were approximately 619 million patients with low back pain in China, and it is estimated that this number will increase to 843 million by 2050.

Due to the fact that spinal surgeries are high-precision deep surgeries, traditional manual operations often lead to large incisions, excessive bleeding, and high revision rates. Therefore, improving the accuracy and safety of spinal surgeries has become an urgent problem to be solved. This study, targeting the surgical navigation system, proposes a deep learning-based method for spinal medical image segmentation to assist in real-time intraoperative navigation and improve surgical accuracy.

This study adopts the Dense-3D U-Net deep learning model to automatically segment spinal CT images. This study contains many innovative points: When facing the problem of image size normalization through direct resizing causing distortion due to different scaling ratios of each channel, an innovative method was proposed to first downsample to the same voxel and then pad the image to ensure that the image size is consistent without proportion imbalance. When facing the problem of image loss not converging, an innovative method was proposed to form a FusionLoss by weighting BCELoss and DiceLoss, solving the problem that a single Loss cannot converge. Subsequently, Dense-3D U-Net was used for end-to-end training, and the network structure and loss function were optimized to improve model performance. Experimental results show that this method can accurately segment spinal structures and exhibit good generalization ability on different patient data.

This segmentation method was integrated into the surgical navigation system, combined with point cloud registration technology, enabling precise alignment of preoperative or intraoperative image data with the patient's anatomical structure on the surgical table. Doctors can obtain precise spinal anatomical information through the VR visualization system, accurately locate the surgical area, thereby reducing the difficulty of surgery, improving the operability of general attending physicians, and reducing surgical trauma and risks. The application of this system is expected to promote the minimally invasive development of spinal surgeries, improve clinical surgical quality, and have important clinical value and social benefits.

Keywords: Spinal surgery navigation, medical image segmentation, deep learning, Dense-3D U-Net

目录

摘要	I
第 1 章 引言	1
1.1 研究背景和意义	1
1.2 国内外研究现状	3
1.2.1 传统医学图像分割方法	3
1.2.2 基于深度学习医学图像分割方法	4
1.3 主要内容和工作安排	13
第 2 章 深度学习理论知识	15
2.1 卷积神经网络	15
2.2 损失函数	17
2.3 本章小结	19
第 3 章 脊柱 CT 图像预处理和模型训练	错误！未定义书签。
3.1 数据集分析	24
3.2 数据预处理	23
3.3 模型的选择	21
3.4 模型的训练	24
3.5 对比试验	27
3.6 本章小结	29
第 4 章 总体系统设计	30
4.1 利用 vtk 渲染 3Dnii 图像	30
4.2 模型的推理	32
4.3 利用 pyqt5 设计可视化界面	33
4.4 本章小结	36
第 5 章 总结与展望	38
5.1 主要工作与创新点	38
5.2 后续研究工作展望	39

参考文献	40
致谢	43

第 1 章 引言

1.1 研究背景和意义

随着人民的生活的水平的提升，和老龄化社会的到来，我国的居民对医疗的服务的需求一直在增长，使得医疗的器械的市场有扩张。根据《World Preview 2018, Outlook to 2025》的预测，2025 年，全球医疗器械市场规模达到 4519 亿美元，同比增长 5.63%。在这期间，中国的医疗器械的市场规模约为 6365 亿元，同比增长高达 20.00%，表现出来了巨大的市场的活力。

在高值的医用耗材细分的市场中，骨科的植入类产品是位居前列的，尤其是脊柱的类植入物，因其的技术壁垒特别高、临床需求特别大，成为了近年来的增长最为迅速的领域的之一。2024 年，我国的骨科植入的市场规模约为 598 亿元，其中脊柱的产品占比 28.23%，市场的规模达 86.99 亿元，预计在人口的老龄化、手术的渗透率提升及医保的支持政策等多重因素推动下，未来仍将保持 15%左右的年均的增长率。

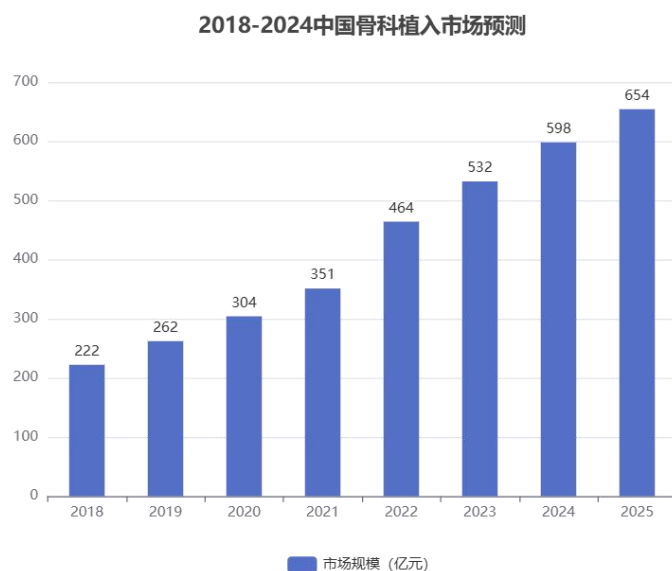


图 1.1 2010-2023 中国骨科植入市场规模及预测

脊柱是人体重要结构，由颈椎、胸椎、腰椎、骶椎及尾椎形成，承载着维持身体姿势、保护脊髓及支持运动的重要功能。常见的脊柱的疾病有退变性脊柱病

变、脊柱的骨折、先天的或特发的脊柱畸形、脊柱的肿瘤和感染等，这些疾病的高发性、复杂性和治疗难度，给患者和社会医疗体系带来了难处。据国家的相关数据表明，椎间盘的疾病、颈椎严重疾病和骨质的疏松性的脊柱骨折等慢性疾病的患病率一直上升，并表现出年轻化情况，青少年脊柱的侧弯等结构性病变也越来越受到社会关注。

为了好好的应对脊柱的疾病的诊疗的挑战，手术的导航系统正在逐渐变成临床手术的重要辅助手段，他通过结合医学的图像、导航的算法及精密的设备实现对手术的路径的可视化和正确定位，明显的增加了脊柱的外科的手术的精度与安全。图像的分割是手术导航系统的最核心的最重要的操作，他的准确性真的能够直接的影响到三维的建模、手术过程中的术中的路径的规划和术后的评估的环节的效果。

从现在的情况全面的详细考虑之下，虽然但是传统的图像的处理的方法和基于手工的特征的算法在脊柱图像的分割的任务中有所全面而又完备的表现出来，但面对复杂的临床数据，仍然会表现出来类似的情况像精度很低、适应性很差、鲁棒性很弱的问题。

这些年来，伴随着深度学习的技术的前进，医学的图像的分割在现在看来是已经取得了突破性的进展，使用的卷积神经网络和他改进结构的分割模型在特别多的多个脊柱任务中表现出来了良好的性能。但是，由于脊柱的解剖结构的变化性、病灶部分的模糊边界和 CT 图像中的类器官相像性，怎么在保证分割的精度的同时增加模型的泛化能力和临床可用性，仍然是当前研究的关键的难点。

因此在目前看来，这一篇文章聚焦于面向手术导航的脊柱医学图像分割方法研究与实现，用这一方法来实现对脊柱的各阶段及其病灶的区域的高精细自动分割作为为目标，结合目前的先进的深度学习模型、医学的图像的预处理的技术和骨科的手术的应用的需求，探索出来了一套适和在临床的环境的实用型图像给出一个分割的方案，为后续的术中的导航、三维的建模和智能的辅助的决策给出了坚实的技术支撑，助力脊柱的外科的微创化、精准化的发展。



图 1.2 多模态脊柱数据集

1.2 国内外研究现状

1.2.1 传统的医学图像分割方法

传统的图像的分割的方法主要是基于灰度的信息进行深入的处理，包括了阈值法、边缘的检测法以及区域的分割法等。其中，阈值的分割的方法的核心是分析脊柱的图像的灰度的特征，选取出合适的阈值，并且以此区分前景（脊柱部分）和背景（非脊柱部分），从而实现脊柱的图像的分割。Barbieri 等人在之前提出了一种由用户引导的 MRI 的椎体的分割方法，结合 Otsu 阈值的分割和区域和生长的策略，并借助专家选定的关键点，使分割结果更加精准可靠。Wang 等人提出了一种融合深度堆叠稀疏自编码器（SSAE）与结构化回归森林（SRF）的椎骨自动定位与识别方法，开发了 SRF 的两阶段渐进式优化策略，以减少模型复杂性并提高椎骨定位的精度。

边缘检测法的核心思想是通过像素值显著变化的区域来界定不同图像部分，实现脊柱影像的精确分割。常见的边缘检测算子包括 Roberts 梯度算子、Canny 算子、Sobel 算子和 Prewitt 算子，通常结合其他方法共同完成分割任务。黄昊等人提出基于边缘与区域的水平集方法用于脊柱 CT 影像分割，该方法结合边缘特征与区域信息，提高了分割的精确度和稳定性。

区域分割方法通过分析图像的区域特征，将脊柱影像划分为多个相似部分，以实现目标识别与分割。主要方法包括区域生长法和分裂合并法。区域生长法从单个像素或小范围区域开始，逐步扩展形成完整的目标区域，而分裂合并法则采用自顶向下或自底向上的策略，对图像进行递归划分。

傅华等人结合边缘检测与区域生长技术，利用霍夫变换确定椎管 ROI 的初始位置，并通过三维区域生长降低因椎管弯曲导致的分割误差。姚姝琳等人提出了

一种融合分水岭算法与图搜索的混合方法，以优化椎体模型的分割效果，并进一步探索了统计建模方法，以实现椎骨的自动分割。内格尔等人将数学形态学与分水岭算法结合，应用于椎骨的标记与自动分割，以提升处理的智能化水平。

此外，刘恒等人提出了一种改进的水平集算法，该方法能够动态调整局部和全局信息的权重，从而增强分割的精确度。李志中等人针对区域生长法对噪声敏感的问题，结合中值滤波与窗口傅里叶滤波，优化区域生长策略，并通过形态学处理进一步提升分割效果。

尽管传统分割方法在脊柱医学图像分析中取得了一定成果，但仍然存在局限性，如特征提取能力有限、过分割或欠分割现象明显，导致分割结果与真实标签之间存在误差。因此，结合深度学习的自动化分割方法逐渐成为研究热点，以提高分割精度和稳定性。

1.2.2 基于深度学习医学图像分割方法

相较于传统机器学习方法，深度学习在 CT 图像处理方面展现出了更强的能力，能够提取更加抽象、高层次的图像信息，这是传统方法难以实现的。传统机器学习通常依赖人工设计特征，如边缘、纹理或形状特征，这不仅需要专业知识，还可能因人为主观因素影响模型的性能。此外，传统方法在面对复杂的医学影像数据时，往往难以兼顾全局信息和局部细节，从而影响最终的分割或分类效果。

深度学习的优势在于其端到端的特征学习能力，无需人为干预即可从海量数据中自动提取有效特征，避免了手动设计特征带来的局限性。同时，卷积神经网络（CNN）利用局部感受野和权重共享机制，实现了特征的复用，使得模型在不同分辨率和尺度下都能有效学习关键结构，提高了对目标区域的识别能力。此外，深度学习还具备强大的迁移学习能力，可以通过在大规模数据集上预训练模型，并在小样本数据集上进行微调，从而提升模型的泛化能力和稳定性。

在医学影像处理中，深度学习的这些优势带来了显著的突破。其端到端学习方式提高了模型的自动化程度，使得脊柱 CT 图像分割的准确率和鲁棒性得到了极大提升。同时，深度学习模型能够有效地适应不同扫描条件、患者个体差异以

及复杂的病灶区域，降低了人为干预的影响，为脊柱疾病的精准诊断和手术导航提供了更可靠的技术支持。

Sekuboyina A、Kukačka J、Kirschke J S 使用全卷积神经网络进行脊柱分割[1]，创造性的使用二维全卷积网络（2D FCN）来提供脊柱的低分辨率定位信息。使用二维-三维全卷积网络（2D-3D FCN），生成高分辨率的二值分割结果。将两个网络的预测结果融合，能够很好地对脊柱体积进行分割。

肖霖、张笠等提出将全局注意力细化(GAR)模块和注意力空间金字塔池化(A-ASPP)模块添加至 U-Net 网络[2]，运用全局与上下文双注意力 U-Net 网络，达到了腰椎的精确度为 90.58%、敏感度为 89.51%、戴斯相似系数为 90.20%。并提出一种轻量化 3D 卷积操作方法，结合改进的椎骨体素数据预处理方法，将提出的 DAU-net 扩展至 3D 模型对脊柱 L1-L5 区域进行多分类的细分割任务，有效提升椎骨分割精度的同时解决 3D 卷积模型参数量大，训练效率低下的问题。

为解决脊柱 CT 图像目前存在的椎骨结构复杂、分割精度不足等问题，刘晶提出一种基于 3D U-Net 框架的脊柱 CT 图像改进分割网络[3]，通过融合 SE 残差单元、椎骨边缘分割模型与改进混合通道-空间注意力机制，在 VerSe 19、VerSe 20 与 CTSpine1K 脊柱数据集上进行分割训练与测试。多次测试实验结果表明，该项目模型在保证分割精度和分割效率有效提高的同时具有较好的泛化性与鲁棒性，在 Dice 相似系数、豪斯多夫距离与平均表面距离上相较于其他先进网络分割精度更高。该模型在现有脊柱分割的网络中具有更强的分割性能，可为放射科医生提供有效临床信息。

张恒在通过研究椎弓根螺钉植入手术中的多个关键技术[4]，基于传统医疗资源，对术前脊椎进行分割，对术前 3D CT 与术中 2D X-ray 影像在统一的三维空间中配准，将两者在统一的三维空间中建立映射关系，通过对术中螺钉的空间位置进行准确定位，在 3d widget 中将空间关系通过多视角、多维度展现出来。

刘侠、匡瑞雪、王波、岳雨婷针对脊柱 CT 图像的分割问题[5]，给出了一个改进的基于隐马尔可夫随机场框架的算法。由隐马尔可夫随机场建立分类的先验分布，将添加了记忆器的模拟退火算法与期望最大化算法进行融合，并结合最大

后验估计准则来进行参数估计并得到最终结果。通过实验证明了所提出的算法较现有的马尔可夫随机场算法在脊柱分割的精确性上有较大程度的提高。

李贤提出一种基于多图谱的深度学习的分割方法,借鉴于 U-Net、V-Net、3D-Unet, 以及最新提出的密集卷积神经网络 DenseNet 的概念, 构建一个新的适应于 CT 分割的卷积神经网络——致密 V-Net[6], 来分割脊柱 CT, 利用多图谱来训练神经网络, 通过训练得到的参数模型来分割目标脊柱 CT, 结果远远优于基于多图谱配准的分割方法的分割结果、文献中已有的脊柱 CT 分割方法的分割结果。

杨玉聃,张俊华,刘云凤针对现有三维脊柱分割方法精度较低的问题,提出一种以三维循环残差卷积为基础的 U 型网络对脊柱 CT 图像进行分割[7]。在网络前端引入三维坐标注意力机制找出网络关注感兴趣的区域,使用三维循环残差模块代替普通卷积模块,使得网络在有效累积特征的同时缓解梯度消失问题。加入高效密集连接混合卷积模块减少底层细小特征信息的丢失,提出双特征残差注意力机制代替跳跃连接进行高低层级间的语义融合,通过聚合不同层级特征对全局上下文进行建模,提升分割性能。

何坤金 ,费雨珊,寇万福,高瑞芳,郝博,蒋俊锋,陈正鸣,钱可涵提出构建含有语义参数信息的分割模板,对脊椎骨进行由粗到细的分割和语义参数测量[8]。首先基于生成的几何平均化骨骼,引入形态学,对其进行分割并获取骨架线比例,构建分割模板;其次,标注语义特征点,计算语义参数,生成含有语义参数信息的分割模板;然后,基于分割模板,利用骨架线与测地环,结合脊椎整体形态与外表面,实现对目标骨骼的区域分割,为脊柱融合器类型的选择提供指导;最后将模板与目标骨骼进行配准、映射,实现目标骨骼语义参数的自动测量,辅助确定脊柱融合器尺寸。

在计算机辅助脊柱图像分析和疾病诊断应用中,从 X 射线脊柱图像中自动分割脊柱和椎骨是一个关键且具有挑战性的问题。申学泉、张勇为进一步提升脊柱图像分割精度,提出一种基于 VGG-Net 改进的模型[9]。首先,将 VGG16 网络去掉了后面的全连接层,用作 U-Net 的特征提取网络;其次,为了增强图像的细节信息,在特征提取网络引入小波分解模块;最后,在上采样网络中设计了一种逐像素相减的自空间注意力模块(SUB-SSAM)机制,进一步提高网络模型识别关键特征的能力。实验

结果表明,改进后的模型相较于原 VGG-Net 模型在平均交并比(mIoU)上提高了 2.39%、召回率提高了 0.96%、准确率提高了 1.31%,训练的该网络模型可以定位到每一块椎骨,准确分割椎体区域。

徐皓在 U-Net 网络的基础上融合了 Transformer 结构提出了 RUnT 椎骨分割模型[10],针对椎骨高度相似,分割模型在识别与分类时产生语义混乱引起的分割精度较低问题,提出基于自注意力机制的全局 Transformer 结构和局部 Transformer 结构,实验结果表明在降低高分辨率椎体 CT 图像的计算复杂度同时通过多体积尺度的自注意力特征图实现椎骨全局位置信息的获取,该项目提出的 RUnT 模型在椎骨分割上相较于纯卷积神经网络在 Dice 指标获得 88.4 精度平均提高了 3%,相较于和 HD 距离指标达到 3.88 最短值平均提高 21.5%。

魏子晨创新的提出了:(1)结合椎骨边界灰度先验信息,提出了基于边界信息增强的椎骨分割方法;(2)结合具有相似性的椎骨分布在 X 光图像中央区域的特点,提出了基于注意力机制的椎骨关键点检测模型;(3)针对倾斜椎骨普遍存在的问题,引入可变形卷积模块更好地捕获椎骨形态学特性,实现更为优越的椎骨分割和关键点检测。[11]充分利用椎骨形态学先验信息及空间位置信息,从而使算法在椎骨分割和关键点检测中取得理想效果。

刑立宁提出了一种二阶段分割模型,首先进行腰椎体 CT 图像粗分割[12],然后进行腰椎体边缘的细化分割。模型在 3D U-Net 基础上引入了残差模块和注意力模块,并改进了损失函数,以提高网络的分割性能。提出的基于深度学习的三维 CT 图像机会性 OP 筛查模型,通过对分割模型和分类模型的改进,改善了目前筛查模型存在的分割准确率较低、分类特征单一和感兴趣区域提取不当的问题,提升了筛查模型的分类性能。

迟晓帆基于深度学习理论基础,开展多方位 X 光影像脊柱、椎体部位语义分割工作[13],具体工作为:第一,开展分割任务时,为保证深度学习模型在训练和测试时具备充分的数据来源,与三甲医院展开合作,收集并构造包含冠状位、矢状位与左右 Bending 位配套的脊柱 X 光影像集。与此同时,提出了两步走的脊椎自动分割框架。首先对四方位影像中的脊柱、骶骨区域实现初次分割,锁定影像中的脊柱区域。其次,在目标区域内结合图像形态学算法实现各椎体的精细化分割。第二,提出了一种

应用于四方位 X 光影像中脊柱、骶骨区域初步分割的循环残差跳连网络 RRSC-Net, 该网络使用循环残差路径弥补了跳连处特征融合带来的信息损失, 多尺度跳跃连接结构与 Inception 模块的使用, 加强了网络对多尺度信息的探索, 空间和通道注意力模块的加入抑制了无关信息的干扰。实验结果表明, 该网络在冠状位、矢状位与左右 Bending 位 X 光影像中脊柱、骶骨区域分割 IOU 值分别为 92.89%、93.30%、94.10%与 93.74%, 实现了目标的精确分割。

陈越鹏基于深度学习提出了一种新的方法, 以实现从正交双平面 X 光片自动重建三维椎骨[14]。首先, 在脊柱 CT 数据上利用 ITK 生成数字重建放射影像, 并以 3D 质心信息对 2D 脊柱 DRR 裁剪及标注。其次, 利用质心标注信息, 应用自注意力机制融合不同视角的互补信息, 将融合后特征转换成 3D 特征。最后, 以全卷积网络模型对 3D 特征进行监督学习, 实现 2D 图像到 3D 形态重建任务。基于上述研究和临床医生对于脊柱 3D 重建应用的需求, 设计并实现了基于脊柱双平面 X 光片的 3D 重建系统。从系统架构和功能模块等方面进行了概要设计, 并在此基础上对所有功能模块进行详细设计从而实现 3D 重建系统。目前, 该系统已准备在合作医院进行试运行以进一步明确其临床应用价值。

吕娟提出一种基于空时特征提取的双支路超声心动视频分割网络[15]。提出一个分割支路来提取视频各帧图像的空间语义特征和一个光流支路来学习相邻帧图像间的空时关联, 并设计了时间一致性约束对两个支路进行联合训练, 实现了网络对视频空时信息的有效利用, 提高了模型对心动视频分割的准确性。

覃军焱开发了基于 U-Net 的脊柱侧弯 Cobb 角测量系统[16]。Cobb 角是指脊柱侧弯的角度, 是脊柱侧弯严重程度的参考标准之一。目前临床中测量 Cobb 角是由专业医生根据经验在脊柱 X 光片上画出两条参考线, 然后使用量角器测量角度, 效率低下且容易产生误差, 不同医生的诊断经验不一样, 测量相同图像可能会产生相差较大的结果。为了解决上述问题, 该项目提出了一种基于 U-Net 模型的脊柱侧弯 Cobb 角全自动测量方法, 基于深度学习模型对脊柱 X 光图像进行分割, 对分割结果进行轮廓识别和边界确定后计算出脊柱侧弯的最大 Cobb 角度。

孟晋提出了基于图像分割的机器人辅助脊柱手术路径自动规划方法。以路径点轨迹和椎骨的三维模型为参考依据,提出了一种在手术过程中末端器械的姿态计算方法,确保骨钻在手术磨削过程中的稳定性。为了避免在手术磨削过程损伤椎管区域,对所规划出的移动轨迹进行了优化调整,提出了安全性约束方法。采用 CT 数据进行仿真实验完成数据分析,并利用脊柱手术机器人系统完成实体实验。基于医院采集的病人 CT 数据完成椎弓根置钉手术路径规划,针对规划结果分析有效性和可行性。设计集成了脊柱手术机器人系统,通过图像配准和机器注册将图像数据、规划路径和真实椎骨进行映射配准。在真实椎骨上利用机器人系统分别完成了两种手术实验,验证了路径规划算法的可行性。

张宸翰提出了一种分别基于 2D 和 3D 卷积的深度学习脊柱分割算法[17]进行研究,并针对当前医学图像分割模型的不足进行改进,最终分别得到分割精度和模型参数性能表现优异的 2D 和 3D 椎骨分割模型。而且针对医学图像分割任务中数据标注困难的问题,通过构建基于半监督学习的椎骨分割框架,在使用少量标注数据的情况下就可以获取较好分割结果。

刘宇波在脊柱扫描路径规划算法[18]中,搭建了 FCN、U-Net 和 Deeplab 三种卷积神经网络,对比了三种网络对深度摄像头拍摄的彩色图分割脊柱区域的效果,同时结合深度图的多视角信息,使用四通道法和双特征提取网络结构方法对深度摄像头拍摄的深度图、彩色图进行整合,提高分割的准确率。最后使用基于人体轮廓的分割结果判定方法和最小二乘曲线拟合方法对分割结果进行后处理,得到较为精确的脊柱扫描曲线。并且借鉴移动最小二乘形变方法提出以脊柱线为控制曲线对人体轮廓进行变形的数据增强方法,拟合多种脊柱侧凸情况,增加数据集的多样性,从而提高网络的泛化能力。最后以医学超声仪器 Sonix RP、六自由度机器人、深度摄像头 Kinect、工控计算机为基础搭建了一套脊柱三维超声自动扫描系统,在机械自动控制程序中嵌入改进的脊柱扫描路径规划算法,实现了脊柱的精确扫描与成像。

匡瑞雪提出了一种高性能的脊柱分割办法[19],首先对原图像进行增强,通过锐化脊柱椎体的边缘、消减噪声、增加对比度等方式使 CT 图像中的椎体更加明确。在增强后的图像的基础上,对脊柱 CT 图像应用隐马尔可夫随机场进行建模,其中标号场采用 MLL 模型,特征场上采用高斯混合模型。之后利用最大后验准则引出能量

函数,图像处理被转化为求解数学上的最优解问题。最后利用期望最大化算法进行图像分割,考虑到该算法容易陷入局部极值,应用模拟退火算法得到全局最优解,但是模拟退火算法的处理速度不甚理想,所以加入振动点和条件迭代算法的思想。

涂勇成以某医院提供的脊柱侧凸患者的 X 射线、CT 以及 MRI 影像为研究对象,进行了如下方面的研究[20]:(1)提出了一种针对 MRI 和 CT 脊柱侧凸影像的基于传统图像处理算法的 Cobb 角自动测量算法。首先采用增强的分水岭分割算法对 MRI 或 CT 脊柱影像中脊椎进行分割并提取各脊椎中心点,然后使用六次多项式对中心点集进行曲线拟合得出脊柱曲线。最后通过计算脊柱曲线中二阶导为 0 的点切线之间夹角自动计算出 Cobb 角。经实验验证,算法可达到 94.2%的脊椎分割精度与 $\pm 4^\circ$ 的 Cobb 角测量误差,解决了临床脊柱侧凸影像 Cobb 角测量以及现有 Cobb 角计算机辅助测量算法中需手动定义脊柱上下端椎的问题。

Deng Y, Gu F, Zeng D 提出了一种有效的 U-Net 和 BiSeNet 补充网络用于脊柱分割[21],以充分发挥它们在特征提取方面的优势。该网络由残差 U-Net、空间特征提取器和特征融合提取器组成。残差 U-Net 采用基于 ResNet34 的 U-Net 框架,旨在充分提取 MRI 中脊椎骨和椎间盘的特征。空间特征提取器采用条带池化 (SP) 块来替代 BiSeNet 网络中的空间提取路径,旨在捕捉脊椎骨和椎间盘之间的长距离关系以及同一类别的相互连接。特征融合提取器采用注意力细化模块 (ARM) 来整合来自残差 U-Net 和空间特征提取器的特征,旨在通过注意力机制引导网络通过学习椎骨与椎间盘之间的区别特征来实现这一目标,并充分利用这两条路径的互补特征。对 172 名受试者的 MRI 数据进行的实验表明,该网络取得了令人瞩目的性能,在所有脊柱结构中的平均 Dice 相似系数达到 82.802%。所提出的方法在脊柱疾病的诊断和治疗方面具有很大的价值。

Saeed M U, Bin W, Sheng J 提出了一种新颖的三维多特征注意力 (MFA) 模型用于脊柱分割[22]。标准的 MobileNetv3 经过修改,添加了 RCBAM (反向卷积块注意力模块) 模块以及 FPP (特征金字塔池化) 以增强特征。每个修改后的 MobileNetv3 都在三维图像的轴位、冠状位和矢状位视图上分别进行单独训练。将特征进行拼接以形成三维特征图,并将其传递给解码器部分进行脊柱分割。

Xie H, Huang Z, Leung F H F 为了编码关于其结构属性和空间关系的丰富先验知识,提出了一种新颖的结构亲和注意力基于转换器编码器 (SATR) [23]来分割脊柱。它利用分层架构输出多尺度特征表示。同时,对脊柱结构信息的约束增强了网络的特征可用性,从而提高了分割精度。对比实验验证了 SATR 在脊柱分割方面的性能优于其他最先进的候选方法,这使得它能够方便地取代智能脊柱侧弯评估的骨干网络。

Da Mutton R, Zanier O, Theiler S 为了实现对每层切片上单独标记椎体这一耗时过程的自动化,提出了一种全自动的流水线方案[24],该方案能够在计算机断层扫描 (CT) 图像上实现自动分割,并可作为进一步虚拟现实应用和影像组学分析的基础。首先训练了一个 YOLOv8m 来单独检测每一个脊椎。然后在裁剪后的图像上开发了一个 2D-U-Net,并在另外两个不同的公共数据集上进行了外部验证,达到了极佳的效果。

对椎体进行准确标注对于脊柱 X 射线图像的自动分析至关重要。然而,由于椎体结构复杂 (包括尺寸小和形状多变),对其进行人工标注是一项费力且成本高昂的工作。为了解决这一挑战并加快标注过程,Chen Y, Mo Y, Readie A 提出了一个名为 VertXNet 的集成管道[25]。该管道目前结合了两种分割机制,即使用 U-Net 进行的语义分割和使用 Mask R-CNN 进行的实例分割,以自动分割和标注颈椎和腰椎侧位脊柱 X 射线图像中的椎体。VertXNet 通过采用基于规则的策略 (称为集成规则) 有效地结合 U-Net 和 Mask R-CNN 的分割结果来增强其效果。它通过识别特定的参考椎体实例来确定椎体标签。

脊椎的精确定位和清晰界定对于诊断和治疗脊柱疾病至关重要。为了实现这一目标, Deng S, Yang Y, Wang J 提出了一种基于增强型 U-Net 架构的高效 X 射线全脊柱脊椎实例分割方法[26]。该方法进行了多项关键改进:采用 ConvNeXt 编码器以有效捕捉复杂特征,并在跳跃连接中引入 IFE 特征提取以聚焦于纹理丰富且边缘清晰的线索。在瓶颈处使用 CBAM 注意力机制来整合粗粒度和细粒度语义信息。解码器采用残差结构结合跳跃连接来实现多尺度上下文信息和特征融合。该方法通过在前弓和侧位脊柱分割实验中的验证得到了证实,展示了强大的特征提取能力和精确的语义分割能力。它有效地处理了各种脊柱疾病,包括脊柱

侧弯、椎体楔形变、腰椎滑脱和腰椎裂。这种分割基础使得脊椎参数能够实现快速校准，并能计算出相关指标，为医学成像领域的进步提供了有价值的参考和指导。

柯布角是衡量脊柱侧弯严重程度的金标准，并用于评估脊柱侧弯的严重程度。柯布角测量的准确性往往依赖于医生的主观经验，而且整个过程非常耗时。在本研究中，Peng L, Hu Y, Zhang K 提出了一个新的深度神经网络——基于 UNet 3+ 的 ATT UNet 3+[27]。该方法在网络的上采样过程中融入了一种新颖的混合注意力机制。这种机制能够对融合的多尺度信息进行适当的重新加权，并有助于对最终输出结果进行有效的监督。所提出的神经网络在 155 个 X 射线正位片上进行了训练、测试和验证。深度学习网络与当今更有效的神经网络进行了比较。ATT UNet 3+ 在分割评估结果中表现最佳。关于最终的柯布角计算，最长距离椭圆点 (LDEP) 方法与专家测量值之间的绝对平均误差为 1.6° 。ATT UNet 3+ 提供了一种用于 X 射线图像中脊柱分割的潜在工具，这能够提高医生处理脊柱侧弯病理图像时的效率和准确性。

Naga Karthik E, Valošek J, Smith A C 的研究了纳入 2002 年 7 月至 2023 年 2 月期间获取的磁共振成像 (MRI) 数据。这些数据包含了使用不同扫描仪制造商、具有不同图像分辨率（各向同性和各向异性）和方位（轴位和冠状位）所获取的 T2 加权 MRI 扫描图像。患者存在不同的病变病因（创伤性、缺血性和出血性）以及病变位置分布在颈椎、胸椎和腰椎。该团队训练了一个深度学习模型 SCISeg[28]（它是开源的，并可通过脊髓工具箱（版本 6.2 及以上）访问），该模型通过主动学习实现了对髓内脊髓损伤和脊髓自动分割的自动分割。所提出模型的分割结果与来自另外三种开源方法（PropSeg、DeepSeg 和对比无关方法，均属于脊髓工具箱的一部分）的分割结果进行了视觉和定量比较。使用 Wilcoxon 符号秩检验比较了从手动参考标准病变掩码中导出的脊髓损伤（病变体积、病变长度和最大轴向损伤比率）的定量 MRI 生物标志物与通过 SCISeg 分割自动获得的生物标志物之间的差异。

Saeed M U, Bin W, Sheng J 提出了一种名为 MSFF（多尺度特征融合）的新型深度学习模型[29]，专门用于使用 MRI 图像进行脊柱骨折的自动分割。MSFF 架

构包含五个关键模块：特征融合模块（FFM）、挤压与激励（SEM）、空洞空间金字塔池化（ASPP）、残差卷积块注意力模块（RCBAM）、残差边界细化注意力块（RBRAB）和局部位置残差注意力块（LPRAB）。这些模块共同实现了多尺度特征融合、空间特征提取、通道级特征增强、分割边界细化以及对感兴趣区域的聚焦注意力。随后，采用解码器网络来预测骨折脊柱区域。实验结果表明，所提出的 MSFF 方法在应对相关挑战方面表现出了卓越的准确性，并且优于现有的分割方法。

Bédard S, Karthik E N, Tsagkas C 提出了一种基于深度学习的方法，能够生成在 MRI 对比度下稳定的脊髓软分割[30]。使用健康参与者（Spine Generic Public Database）的脊髓图像，该项目展示了该方法的性能。该项目还比较了该方法与现有的方法在检测脊髓萎缩方面的性能差异。结果显示，该项目在检测细微萎缩方面具有更高的灵敏度。此外，该项目在不同对比度下产生的分割结果更加稳定，且不受部分体积效应的影响。

1.3 主要内容和工作安排

本文针对手术导航场景下脊柱医学图像精确分割的关键需求，开展了面向高分辨率脊柱 CT 图像的自动分割方法研究，旨在提升椎体结构分割的准确性与鲁棒性，为后续的术前三维建模与术中导航系统构建提供坚实基础。结合当前深度学习在医学图像分割中的最新进展，本文设计并实现了一套系统的脊柱图像分割方案。

首先，基于 VerSe 公共脊柱数据集，完成了针对性的数据预处理，包括统一图像空间朝向（LAS）、ROI 裁剪、各向同性重采样（ 1mm^3 ）、图像归一化、异常值 clamp 处理、噪声初步去除及多种数据增强操作，以保证网络训练的稳定性与模型泛化能力。

其次，本文引入 DenseUNet 网络结构，充分利用其密集连接机制以提升特征表达能力，同时结合深层语义特征与浅层边缘信息，增强对复杂椎体边界的识别能力。在网络训练阶段，采用基于 Dice loss 和交叉熵的复合损失函数，以优化小

体积结构的分割效果。实验结果表明，本方法在评价指标(Dice Similarity Coefficient)上均取得优异表现，模型在保持较高分辨率输出的同时，具备良好的分割准确率与边缘保留能力。

最后，本文所构建的脊柱分割模型已在本课题组后续的脊柱手术智能导航系统中获得初步应用，验证了其在术前三维重建、术中导航路径规划等任务中的可行性与有效性。整体而言，本研究为脊柱医学图像分割提供了一种高效、稳定的解决思路，在推动智能化手术辅助系统发展方面具有一定的理论价值与应用前景。

全文共分为 4 章，内容结构安排如下：

第 1 章为引言，引入课题的研究背景及意义；

第 2 章是深度学习理论背景；

第 3 章是模型的设计与训练；

第 4 章作为论文的结束语，总结毕业设计工作，提出可以在今后继续深入研究的方向。

第 2 章 深度学习理论知识

近年来，随着计算能力的显著提升和大规模数据集的广泛应用，深度学习（Deep Learning）作为人工智能领域的核心技术之一，在图像识别、自然语言处理、语音识别、医学图像分析等诸多领域取得了突破性进展。深度学习本质上是多层神经网络在特征学习和表示建模方面的一种有效方法，其通过构建多层非线性变换结构，能够自动从原始数据中提取出具有判别性的高阶特征，避免了传统方法中对特征工程的高度依赖。

深度神经网络（Deep Neural Networks, DNN）特别是卷积神经网络在图像处理任务中表现出卓越的性能。CNN 通过局部感受野、权重共享和池化操作等机制，显著减少了模型参数数量的同时保留了图像的空间层次结构，使其成为视觉任务中最常用的网络结构之一。此外，近年来提出的一系列改进结构如 U-Net、DenseNet、ResNet 等，进一步提升了模型的表达能力与训练效率，广泛应用于语义分割、目标检测等任务中。

在医学图像分析领域，深度学习已成为主流研究方向。尤其是在医学图像分割任务中，深度学习模型能够充分学习组织器官的空间结构与边界特征，在处理低对比度、形态复杂、背景噪声较多的医学图像时，表现出远超传统方法的鲁棒性与准确性。然而，医学图像数据通常存在样本稀缺、标签获取成本高、数据分布不均等挑战，这也促使研究者不断探索更高效的网络结构、损失函数设计及数据增强策略，以提升模型在实际应用场景中的泛化能力。

2.1 卷积神经网络

卷积神经网络是一种专为处理具有网格结构数据（如图像）而设计的深度学习模型，广泛应用于图像识别、目标检测、医学图像分割等任务。CNN 能够自动从数据中提取有用的特征，无需依赖人工设计，极大地提升了模型的表达能力和泛化能力。

在 CNN 中，最核心的操作是卷积运算。卷积层的本质是通过一组可学习的卷积核（也称为滤波器）对输入特征图进行局部感受，每一个卷积核都会在图像上滑动（即“卷积”），在每一个滑动位置上计算局部区域与卷积核之间的加权和。这一过程可以看作是将原图像进行“特征投影”，比如识别边缘、纹理、方向等特征。例如，一个 3×3 的卷积核在图像上滑动时，每次提取图像中一个 3×3 的局部区域，与卷积核中的权重进行逐元素相乘并求和，形成新的特征图。卷积操作通过权值共享机制大大减少了模型参数，提高了计算效率，同时保留了局部空间结构特征。

卷积操作之后，通常会引入激活函数（Activation Function）来增加模型的非线性表达能力。最常用的激活函数是 ReLU（Rectified Linear Unit），其定义是 $\text{ReLU}(x) = \max(0, x)$ 。ReLU 的作用是将输入中的负值压为零，保留正值不变，从而引入非线性特征。这种非线性变换对于深度神经网络至关重要，因为如果没有激活函数，多个卷积层叠加的效果仍然是线性的组合，模型无法学习复杂的高阶特征。ReLU 同时具有计算简单、收敛速度快、避免梯度消失等优势，是深度网络中的标准选择。

在多个卷积层之后，CNN 中还会使用池化操作（Pooling）来进一步处理特征图。池化层的主要作用是降低空间维度、减少参数数量，同时增强模型对局部平移、旋转、缩放等形变的鲁棒性。最常见的是最大池化（Max Pooling）操作，它以一个滑动窗口在特征图上滑动，在每个窗口区域内取最大值，从而提取最显著的特征。例如， 2×2 的最大池化将每 2×2 区域压缩为一个最大值，实现空间下采样。池化不会引入新的参数，而是起到了数据压缩和特征选择的作用。

综上所述，卷积神经网络通过卷积层提取特征、激活函数增强非线性能力、池化层压缩信息，从而构建出一种高效且强大的特征学习模型。在训练过程中，利用反向传播算法不断调整卷积核和全连接层的参数，使得模型可以自动学习到最优的图像表示方式。这一整套机制使得 CNN 在视觉任务中表现卓越，成为现代计算机视觉领域的主流技术基础。

2.2 损失函数

损失函数是深度学习模型训练过程中用来衡量预测结果与真实标签之间差距的工具。它本质上是一个评价指标，通过计算模型输出与实际答案之间的误差，指导模型参数的更新方向和幅度。损失函数的值越小，表示模型预测得越接近真实值；值越大，说明模型还需要进一步学习。整个训练过程的目标就是不断最小化损失函数，以提升模型在训练集和验证集上的表现。

交叉熵损失是最常见的分类损失函数之一，常用于图像分类和分割任务中。它将模型的输出概率与真实标签的 `one-hot` 编码进行比较，从概率角度量化预测的准确程度。如果模型将正确类别的概率预测得越接近 1，交叉熵损失就越小。这个损失函数特别适合多类分类问题，对分类精度要求高，但在类别不平衡或小目标分割任务中可能表现不佳。

F1 Loss 是基于分类任务中的 F1 分数构造的损失函数。F1 分数是精准率和召回率的调和平均，用于衡量模型在正负样本不平衡情况下的综合性能。F1 Loss 的目标是最大化 F1 分数，因此它更关注模型在“预测全面性”和“预测准确性”之间的平衡，能有效缓解模型只预测主类或常见类的问题。相比于交叉熵损失，F1 Loss 更强调整体判断的质量，适合用于医疗图像、文本分类等类别分布严重不均衡的场景。

Dice Loss 常用于图像分割任务，尤其在分割目标较小或前景像素稀疏时效果显著。它以 Dice 系数为基础，衡量预测区域与真实区域之间的重叠程度。如果预测和真实区域重合度越高，Dice Loss 就越小。不同于交叉熵关注每一个像素的分类结果，Dice Loss 更重视整体区域的结构一致性，具有更强的鲁棒性和泛化能力。

这些损失函数各有优劣，实际应用中常将它们组合使用，例如通过加权的方式融合交叉熵损失和 Dice Loss，形成一种更稳定、更全面的训练目标。这种方式可以同时优化像素级别的精度和区域级别的重叠率，从而提升模型的整体分割效果。本项目所使用的 FusionLoss 正是这种融合策略的体现，它结合了多个损失函数的优势，在多个层面引导模型学习，有助于获得更高质量的分割结果。

本项目使用的 FusionLoss 是将 Dice Loss 和 Binary Cross Entropy Loss (BCE Loss) 按照 0.7 和 0.3 的权重进行加权融合，即 $\text{FusionLoss} = 0.7 \times \text{DiceLoss} + 0.3 \times \text{BCE Loss}$ 。这种设计背后的核心思想是结合两种损失函数的优势，从不同角度优化模型的性能，特别适用于图像分割等任务，尤其在目标边界模糊、前景区域稀疏或类别不平衡的场景中效果更优。

BCE Loss 是一种像素级别的损失函数，它逐个像素地评估预测概率与真实标签之间的差异。对于每个像素，它都鼓励模型输出一个接近真实标签的概率值，因此在分类精度方面具有很强的约束力。BCE Loss 适合精细控制每一个像素的预测值，帮助模型在局部层面提高分辨率和分类正确率。

Dice Loss 则是一种基于区域重叠的指标，它关注的是整体前景区域的预测质量，强调预测结果与真实标签之间的重叠程度。它对前景像素稀少的情況特别敏感，能有效避免模型只预测背景、忽略前景的问题。在目标面积较小、边缘模糊的图像中，Dice Loss 能显著提升模型对目标轮廓和形状的识别能力。

将两者融合使用的原因，是因为单独使用 BCE Loss 容易在类别不平衡时失效，尤其在图像中前景像素远少于背景像素时，模型可能只学习到“预测全部为背景”也能得到较低的 BCE Loss。而单独使用 Dice Loss 虽然对前景敏感，但它对像素级的误差反馈较弱，不足以驱动模型在早期阶段收敛得很快。因此，融合两者可以让模型既关注每一个像素的准确性，也注重前景区域的结构完整性。

具体采用 0.7 和 0.3 的加权比例，是为了强化 BCE Loss 的像素级监督作用（占比更大），在训练初期帮助模型快速收敛；同时保留一定的 Dice Loss 结构级别的监督，提升模型对整体轮廓的识别能力。这种权重分配在实践中往往可以达到较好的平衡效果，提高最终模型在分割任务中的精度与鲁棒性。

2.3 目前主流的分割网络 UNet

U-Net 最初是为医学图像分割任务（如器官、肿瘤、血管等结构的自动分割）而设计的，尤其适用于训练数据较少的场景。其核心设计理念源于医学图像处理对精确像素级分割的需求，网络结构兼顾局部特征提取与全局上下文信息的融合，通过编码器-解码器（Encoder-Decoder）结构配合跳跃连接（Skip Connection）机制，有效提升了分割的准确性与边界细节的恢复能力。

U-Net 的整体结构呈现出典型的对称“U”形：左侧为编码路径，右侧为解码路径，两侧通过跳跃连接相互耦合。编码路径由多个卷积层和池化层堆叠而成，通过多次下采样操作提取图像的高层语义特征，并逐步降低空间分辨率，以获得更丰富的上下文信息。每个卷积层后通常接 **Relu** 激活函数和最大池化操作，随着网络深度的增加，卷积核的数量逐渐提升，使得模型具备提取复杂图像结构的能力。

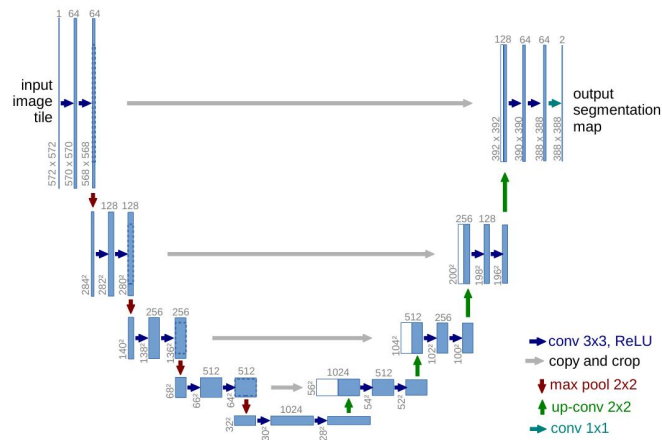


图 3.4 UNet 网络结构

位于网络中心的瓶颈层具有最小的空间分辨率，但同时拥有最大的通道数，代表了最深层次的抽象语义信息。这一层整合了整个图像的全局特征，为后续的解码过程提供了丰富的语义支撑。

解码路径通过转置卷积（Transposed Convolution）或上采样操作逐步恢复图像的空间分辨率。每一阶段的上采样后都伴随着两个标准卷积操作，用于精细还原

图像细节。为了缓解深层信息对浅层结构的模糊影响，U-Net 在解码器的每一层引入跳跃连接，将编码器对应层的特征图与当前解码层的特征图进行通道维度上的拼接。这种结构融合了编码器提取的空间边缘信息和解码器的语义信息²⁰，有助于模型精准还原目标边界和复杂结构。

跳跃连接是 U-Net 架构中的关键创新之一，它直接传递浅层的低级特征至对称位置的解码层，避免了池化操作带来的空间信息丢失，尤其有助于图像边缘和纹理的恢复，使得分割结果更加精细、可靠。整体而言，U-Net 能够在保持全局理解的同时兼顾局部细节，为医学图像分割任务提供了强大且高效的建模

2.4 本章小结

本章围绕深度学习的基础理论展开论述，系统介绍了深度神经网络的基本概念、结构组成及其在图像处理中的关键技术要点，重点分析了卷积神经网络在图像特征提取与空间建模中的优势。针对医学图像处理中的复杂性与挑战性，概述了当前主流的深度学习网络结构及其在医学图像分割任务中的典型应用，为后续章节中脊柱 CT 图像的自动分割方法设计与实现提供了理论支撑和方法基础。

第 3 章 基于改进降采样稠密网络的脊柱 CT 图像分割

3.1 利用稠密型网络 densenet 进行图像分割

3.1.1 densenet

为了增强特征提取能力,本项目在编码器部分引入了 DenseNet 结构。DenseNet 是一种高效的深度神经网络结构,其核心思想是通过密集连接实现跨层特征复用。具体而言,在同一 Dense Block 内部,每一层都会接收前面所有层的特征图作为输入,并将自身的输出传递给后续所有层,从而构建出强耦合的特征流动路径。

这种密集连接方式一方面有效缓解了梯度消失问题,增强了特征的传递与共享;另一方面也提升了网络的参数利用效率,使得模型在保持较小参数量的同时具备强大的表达能力。在图像分割任务中,DenseNet 结构能够捕捉更丰富、更多尺度的图像上下文信息,提升模型对目标区域特别是边界细节的感知能力。

在实际设计中,Dense Block 通常由若干由 Batch Normalization、ReLU 和 3×3 卷积组成的卷积单元构成,层与层之间通过特征图的拼接操作进行连接。同时,为避免特征图通道数爆炸式增长,在每个 Dense Block 之间通常引入 Transition Layer,通过 1×1 卷积降维并配合平均池化实现特征压缩与尺寸下采样。通过将 DenseNet 与 U-Net 结构融合,形成 DenseUNet 架构,能够在保持深层语义理解的同时大幅增强低层细节信息的表达能力,从而显著提升分割的精度与鲁棒性。

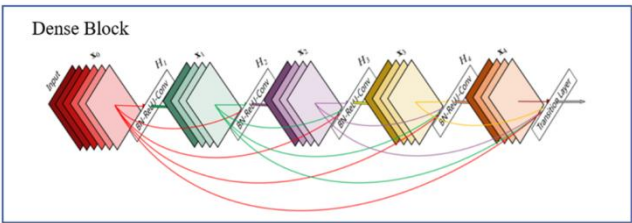


图 3.1 DenseNet 网络结构

3.1. 2Dense-3DUNet

本项目所采用的网络结构为 Dense-3D U-Net，该模型在 U-Net 架构的基础上进行了两方面的改进：一是将二维卷积操作替换为三维卷积，以适应医学图像中体数据的三维结构特征；二是在编码器和解码器各层引入 DenseNet 结构，即每一层均采用 Dense Block 构建，形成 Dense U-Net 结构。

Dense-3D U-Net 结合了 U-Net 的跳跃连接机制与 DenseNet 的密集连接优势。一方面，U-Net 的编码-解码对称结构配合跳跃连接，使模型能够在上采样过程中融合浅层的空间细节信息与深层的语义特征，提升分割精度；另一方面，Dense Block 中跨层级的特征融合机制增强了特征传播与复用能力，缓解梯度消失问题，并提高了参数效率与收敛速度。

通过将三维卷积与 dense 连接机制相结合，该模型不仅具备强大的空间建模能力，还能充分整合多层次的上下文语义信息，在保持高精度分割性能的同时显著提升了网络的表达能力与泛化性能，尤其适用于三维医学影像中器官边界复杂、目标尺度变化显著的分割任务。

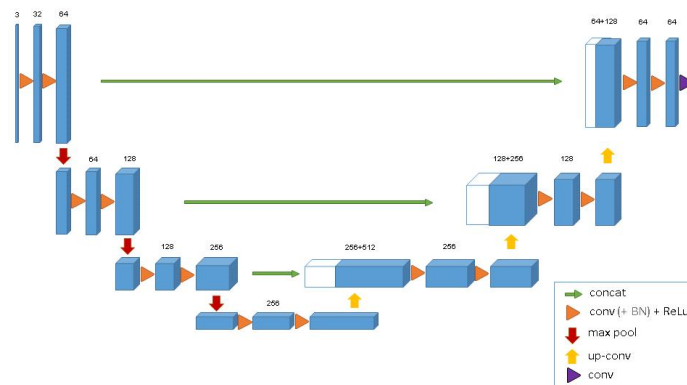


图 3.2 3D-UNet 网络结构

3.2 利用改进的降采样模块进行预处理

本项目所采用的原始 CT 图像在空间分辨率与图像尺寸方面存在显著差异。一方面，不同医院或成像设备的扫描协议存在不一致性，导致图像在轴向（Axial）、冠状（Coronal）和矢状（Sagittal）三个方向上的体素大小存在各向异性。这种体素尺度的不统一会干扰模型对空间结构的建模能力，进而影响分割结果的准确性。另一方面，原始 CT 图像通常呈长方体形态，若直接缩放至统一尺寸，可能会引入各维度压缩比例不一致的问题，造成结构性失真，削弱模型对关键解剖结构的感知能力。

为应对上述问题，本项目在数据预处理阶段引入了标准化处理流程。首先，采用线性插值方法对所有图像进行重采样，将体素尺寸统一为各向同性的 $1\text{mm} \times 1\text{mm} \times 1\text{mm}$ ，从而消除空间尺度差异，增强模型对空间特征的学习能力。其次，为去除异常灰度值对训练的干扰，采用灰度值裁剪（clamping）操作，将图像中的 Hounsfield 单位（HU）限制在 0 至 1000 之间，以排除金属伪影、气体及其他极端值带来的噪声，提高图像质量的一致性和稳定性。

在完成重采样与 HU 值裁剪后，为保证模型输入维度的一致性，将图像统一裁剪并填充至固定尺寸 $128 \times 128 \times 128$ 。该操作以感兴趣区域（ROI）为中心，通过零填充（zero-padding）补齐边缘，确保模型在训练过程中能够聚焦于关键解剖结构区域，同时满足深度学习网络对输入尺寸一致性的要求。

此外，为提升模型的泛化能力，本项目还在训练阶段引入了数据增强策略，包括随机旋转、翻转（镜像）等空间变换操作，从而增加样本多样性，减缓模型过拟合现象，进一步提高分割精度与鲁棒性。上述一系列预处理与增强策略为后续模型训练提供了统一、规范、高质量的输入数据基础，显著提升了分割模型在复杂场景下的性能表现。

3.3 模型的训练

3.3.1 数据集分析

VerSe 数据集是专门用于脊柱图像分割和定位的公开数据集，常用于医学图像分割和三维重建任务。该数据集包含大量脊柱 CT 扫描图像，并附标签信息，有 128 张训练集，116 张验证集和 110 张测试集。图像采用 NIfTI 格式 (.nii.gz)，由头文件和图像数据两部分组成。头文件包含图像的元数据，如维度、体素和空间朝向等信息。图像维度通常为 (512, 512, 128)，表示 CT 扫描在 X、Y 和 Z 轴上的体素数量。体素大小存储在头文件的 spacing 字段中，VerSe 数据集采用医学影像中的 RAS 朝向。

VerSe	Patients	Scans	Scan split	Vertebrae (Cer/Tho/Lum)
2019	141	160	80/40/40	1725 (220/884/621)
2020	300	319	113/103/103	4141 (581/2255/1305)
Total	355	374	141/120/113	4505 (611/2387/1507)

图 3.3 verse 数据集的划分

选取 VerSe 数据集的原因在于该数据集是公开的，并被多篇学术论文广泛使用，具有较高的权威性和通用性。由于 VerSe 数据集包含大量带有标注的脊柱 CT 图像，且标注精细且规范，能够为模型提供高质量的训练数据。此外，VerSe 数据集已成为脊柱图像分割任务的常用基准数据集，使用该数据集进行模型训练和评估，可以将模型的分割效果与已有文献中的结果进行对比，从而客观衡量模型在精度、召回率和推理速度等方面的表现。这有助于验证模型的有效性，并发现改进方向。

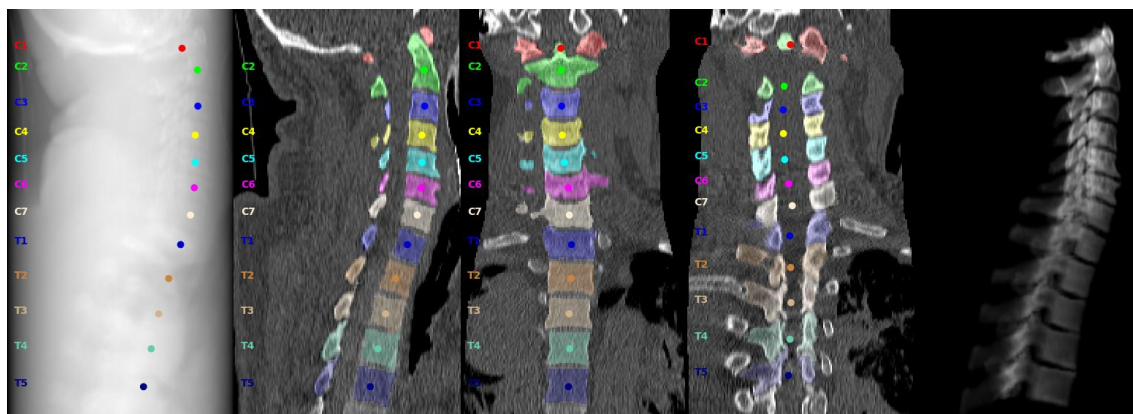


图 3.4 verse 切片展示

3.3.2 改进 loss 函数对模型 进行训练

模型训练采用有监督的端到端方式进行，利用带标注的三维医学图像数据对 Dense-3D U-Net 网络进行优化。损失函数设计为融合损失 Fusion Loss ($\text{FusionLoss} = 0.3 \times \text{BCELoss} + 0.7 \times \text{DiceLoss}$)，其中 Dice Loss 权重为 0.7，BCE Loss 权重为 0.3。该设计兼顾了全局区域重叠程度与像素级分类精度，特别适用于前景背景比例悬殊的医学图像分割任务。

优化器选择 Adam，初始学习率设为 1×10^{-4} ，并引入权重衰减策略抑制过拟合。在训练过程中，当验证集性能在连续若干 epoch 内无明显提升时，自动降低学习率以提升模型的收敛稳定性和泛化能力。

训练在 GPU 平台上完成，批量大小为 4，总训练轮数为 200。在前向传播阶段，模型接收统一尺寸的三维图像作为输入，经过稠密连接的编码-解码结构逐层提取多尺度空间语义特征，并借助跳跃连接恢复空间细节信息，最终输出与输入尺寸一致的概率分割图。反向传播根据 FusionLoss 来计算梯度并更新网络参数。

从训练曲线来看，损失值在前 10 个 epoch 内下降最为显著，反映出模型快速学习到初始判别能力；随后逐渐趋于平稳，并在第 125 个 epoch 左右收敛。准确率方面，在第 50 个 epoch 即达到 72%，至第 175 个 epoch 提升至 80%。训练过程中每轮均在验证集上评估模型性能，并监控 Dice 系数与验证损失的变化，

$$Dice = \frac{2|x \cap y|}{|x| + |y|}$$

最终保存验证集上性能最优的模型用于后续推理与测试。

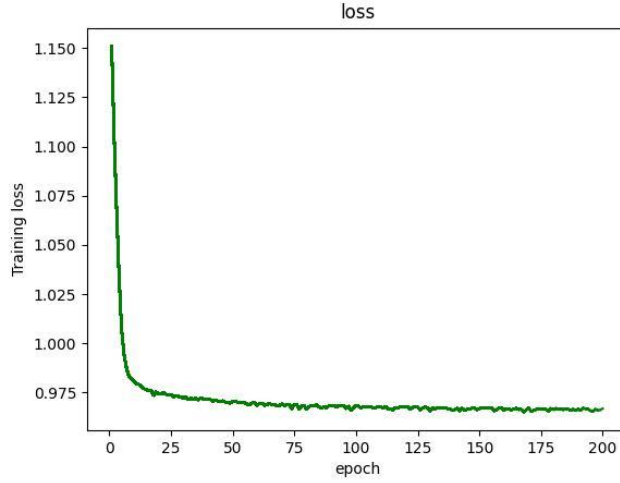


图 3.5 训练 loss 图像 （ $FusionLoss = 0.3 \times BCELoss + 0.7 \times DiceLoss$ ）

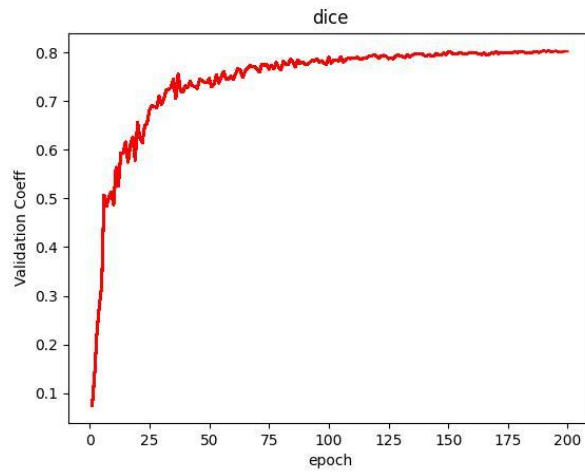


图 3.6 验证集 Dice 系数 （ $Dice = \frac{2|x \cap y|}{|x| + |y|}$ ）

3.4 对比试验

为了验证所提出的三维分割模型在脊柱医学图像分割任务中的有效性，本文设计了一系列对比实验。实验选取了当前常用的三维医学图像分割模型作为基线，包括 DiceNet3D、VNet、UNet 和 Dice3DUNet 等，并在相同的数据集与训练策略下进行评估。

在基准实验中，Dice3DUNet 已表现出良好的分割性能。为了进一步探究本文所设计的改进型降采样模块对模型性能的影响，我们分别在 Dice3DUNet 模型上进行了两组实验：一组为去除降采样模块的原始结构，另一组为加入改进型降采样模块的结构。对比结果表明，在 Dice 系数这一评价指标上，引入本文提出的降采样策略显著提高了模型的分割准确率。

此外，相比于 DiceNet3D、VNet 和传统 UNet 模型，加入改进降采样模块后的 Dice3DUNet 不仅在整体分割性能上具有明显优势，同时在边界细节的刻画和小目标结构识别方面也更为精确。这说明所提出的降采样策略有效缓解了传统模型在信息损失与空间细节保持之间的矛盾，提高了深层特征的表达能力。

综上所述，本文提出的降采样模块在提升三维分割模型性能方面具有明显优势，尤其在 Dice3DUNet 架构中表现最为显著，为后续高精度医学图像分析提供了可靠支撑。

Model Name	Average Dice Score (%)
DenseNet3D	70.2
UNet3D	76.1
Dense3DUNet	78.3
Dense3DUNet & DownsamplingBlock	80.7

表 3.1 不同模型的对比试验（verse 数据集）

本实验设计了三组网络结构对比，分别为传统的 Unet、改进后的 Dense3D UNet，以及在 Dense3D UNet 基础上进一步引入 Downsampling Block 的变体模型，旨在评估不同结构对医学图像分割性能的影响。通过对比三种模型在测试集上的分割结果可以看出，Unet 模型尽管能够完成基本的分割任务，但在复杂结构区域和边界细节处表现有限，分割轮廓较为模糊，部分关键区域存在漏分或误分现象。

相比之下，Dense3D UNet 在结构设计中引入了密集连接，有效提升了特征的复用能力和深层语义信息的表达能力，分割结果在准确性和连续性方面均优于 Unet，尤其在复杂器官边界处呈现出更好的保留能力。然而，仍存在一定程度的信息损失，特别是在纵向切片方向上，分割结果较横向切片略显粗糙。

进一步地，在 Dense3D UNet 的基础上引入 Downsampling Block 后，模型的分割性能获得了显著提升。该模块通过合理的下采样操作增强了对多尺度空间特征的建模能力，同时减缓了由于图像预处理中的重采样所带来的信息损失问题。实验显示，该结构能够更有效地保留图像的细节信息，所生成的分割图像在边缘曲线的平滑性和整体结构的连贯性上均优于前两种模型，视觉效果更加自然，医学解剖结构的还原度也更高。

此外，三个模型在不同切片方向的表现也体现出显著差异。Unet 与 Dense3D UNet 在纵向方向上的分割效果普遍低于横向方向，主要原因在于图像预处理阶段对纵向尺寸的重采样存在比例失调，导致大量结构细节在输入模型之前就已丢失。而引入 Downsampling Block 的模型在纵向切片中的表现明显优于前两者，验证了该模块对结构信息保留和跨方向一致性的重要促进作用。

从目前的实验的情况来看，引入了 Downsampling Block 的 Dense3D UNet 结构在本实验中取得出了最好的分割成果，不仅仅加强了模型的特征提取功能，也在一定程度上放缓了由图像预处理带来的信息丢失问题，具有了较高的实际应用的价值。

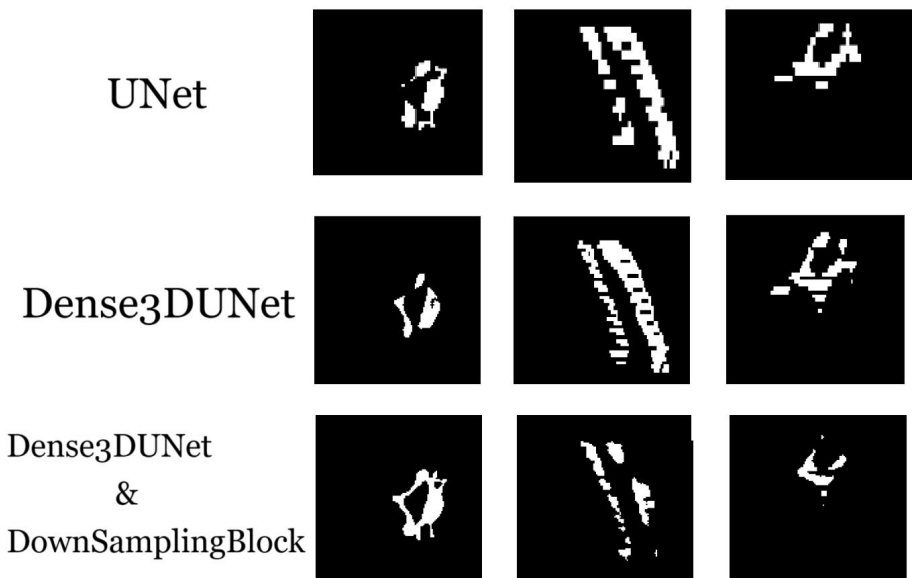


图 3.7 不同方法的切片展示

3.5 本章小结

本章节的研究围绕脊柱医学图像分割的活动，对于数据预处理、模型训练与推理部署三个阶段进行了完整的设计与优化。在预处理阶段，本文提出了“重采样—等比例缩放—零填充”的三步处理流程，确保所有输入图像在尺寸标准化的同时保持原始空间比例与结构特征的一致性，从而为模型的学习稳定性与空间理解能力奠定基础。在训练阶段，构建了基于 Dense 3D UNet 的网络结构，并结合创新设计的 FusionLoss 融合损失函数，有效缓解了传统训练过程中 loss 值难以收敛的问题，提升了模型在复杂结构区域（如脊柱边缘）的分割精度。同时，通过合理设置学习率策略、批大小、数据增强与正则化手段，实现了训练过程的稳定收敛与性能最优化。在推理阶段，模型保持了较高的运行效率，能够对三维医学图像进行快速而精确的脊柱结构预测，满足实际临床前图像处理需求，为后续脊柱手术导航等应用场景提供了可行的技术支撑。

第 4 章 总体系统设计

4.1 利用 vtk 渲染 3Dnii 图像

医学的图像的可视化为医学图像处理的系统中地核心的组成部分，直接的关系到了医生对病灶的、组织的结构等信息的理解与分析。本系统选用了 VTK 作为可视化渲染工具，结合了 ITK 进行了图像地读取与格式的转换，成功的实现了对三维的 NIfTI 医学图像的三维体的绘制和展示。该部分的实现不仅仅为后续模型的推理结果展示并且提供了基础，也增强了整个系统的人机的交互性和直观性。

在图像的读取阶段，系统采用 ITK 中的 `itk.imread()` 方法对 NIfTI 格式文件进行加载。NIfTI 格式是医学图像处理中常见的数据存储格式，具有良好的跨平台支持性和三维体数据结构的优势。ITK 作为专业的医学图像处理库，能够稳定、高效地完成图像数据的读取和初步解析。由于 VTK 和 ITK 使用的是不同的图像数据结构，为了实现后续在 VTK 中的渲染，需要调用 ITK 提供的桥接接口 `itk.vtk_image_from_image()`，将 ITK 图像对象转换为 VTK 图像对象。

完成数据转换后，系统进入 VTK 的渲染阶段。渲染采用 VTK 中的 `vtkGPUVolumeRayCastMapper`，该映射器是基于 GPU 的体绘制方法，可以极大地提高三维图像的渲染效率，适用于医学图像这类高分辨率、数据量庞大的三维数据。通过设置该映射器的输入数据为 VTK 格式的图像，实现了图像数据与渲染管线的连接。

为了增强图像的视觉表现效果，本系统引入了两种传输函数：不透明度传输函数和颜色传输函数。前者控制不同强度值（如 CT 值）在显示中的透明程度，使背景区域或低密度组织在可视化过程中逐渐隐藏，突出显示骨骼或病灶等高密度结构；后者则根据强度值映射对应的颜色值，在本系统中使用灰度映射（从黑色到白色线性过渡）来保持医学图像的原始风格。将这两类传输函数通过

`vtkVolumeProperty` 封装成体绘制属性对象，并与 `vtkVolume` 体绘制体元结合后，能够完整地反映三维图像的灰度变化和空间结构。

在图像渲染显示部分，系统设置了 `vtkRenderer` 作为场景渲染器，背景为黑色，以增强图像对比度；`vtkRenderWindow` 作为窗口用于实际的图像显示；`vtkRenderWindowInteractor` 提供了用户与图像交互的功能，使用户可以通过鼠标进行图像的旋转、缩放、平移等操作。调用 `Render()` 函数后，图像会被即时渲染在窗口中，而通过 `Start()` 函数可以启动交互器，支持用户对渲染图像进行实时操作。

综上所述，通过 VTK 与 ITK 的结合，本系统成功实现了 NIfTI 格式医学图像的三维可视化，为医生或科研人员提供了清晰、可交互的图像界面，也为后续模型分割结果的对比与展示奠定了坚实基础。

```
def show_3d_nifti_image(nifti_file_name):  
    # Read NIFTI file using ITK  
    itk_img = itk.imread(nifti_file_name)  
  
    # Convert ITK image to VTK image  
    vtk_img = itk.vtk_image_from_image(itk_img)  
  
    # Set up volume rendering mapper  
    volume_mapper = vtkGPUVolumeRayCastMapper()  
    volume_mapper.SetInputData(vtk_img)  
  
    # Set up opacity transfer function  
    opacity = vtkPiecewiseFunction()  
    opacity.AddPoint(0, 0.0)  
    opacity.AddPoint(50, 0.05)  
    opacity.AddPoint(1000, 0.2)  
    opacity.AddPoint(3000, 0.4)  
  
    # Set up color transfer function (grayscale mapping)  
    color = vtkColorTransferFunction()  
    color.AddRGBPoint(0, 0.0, 0.0, 0.0)  
    color.AddRGBPoint(1000, 1.0, 1.0, 1.0)
```

图 4.1 VTK 3D 渲染部分关键代码

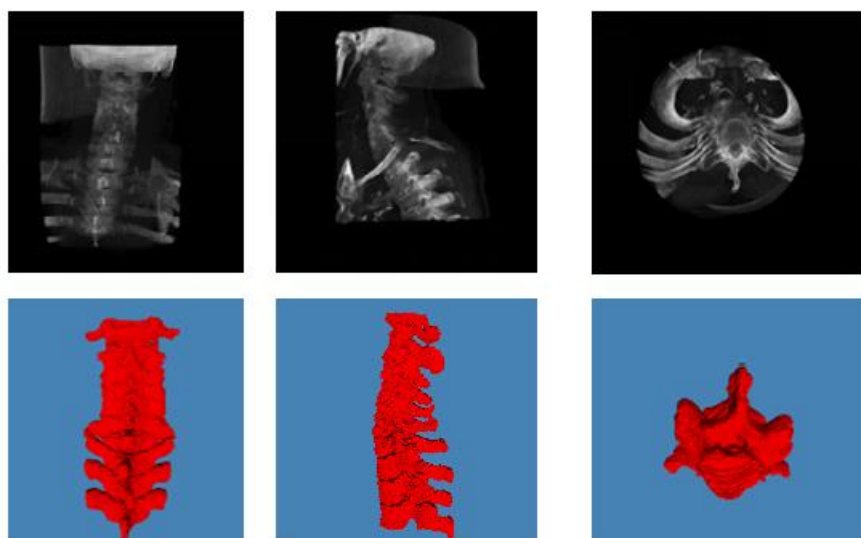


图 4.2 利用 VTK 渲染出的三视图（原图和分割后的图像）

4.2 模型的推理

在模型训练完成并保存权重之后，需要对模型在测试集上的性能进行评估和结果可视化。推理模块的主要任务包括：加载训练好的模型参数、将待预测图像输入模型进行前向传播、对预测结果进行阈值处理和可视化展示、保存预测图像以及评估模型的分割性能指标（如 Dice 系数）。本节详细介绍模型推理的实现流程与关键技术。

首先，系统从 `config.yaml` 文件中读取网络结构和数据相关配置，并根据配置初始化模型。若配置选择了 `DenseUNet3d`，则加载对应网络结构；否则默认使用 `VNet` 结构。模型参数通过 `torch.load()` 加载，并使用 `load_state_dict()` 进行恢复。推理时启用 `torch.no_grad()`，避免梯度计算以节省内存和计算资源。

推理过程中，图像通过数据加载器 `DataLoader` 分批加载，并传入模型进行前向传播，输出预测概率图。通过 `torch.sigmoid()` 函数对输出进行归一化，并以 0.5 为阈值将连续值转化为二值化掩码，从而得到最终的分割结果。模型输出、真实标签和原始图像经过处理后传入 `show_slices()` 和 `show_featuremap()` 函数进行可

可视化。具体可视化方式为：分别截取图像的第 16、26、36 层展示分割结果和真实标签，利用 matplotlib 生成对比图并保存至本地 ./result/20240423/ 目录。

为了评估模型的分割性能，本系统采用 Dice 系数（Dice Similarity Coefficient, DSC）作为衡量指标，计算方式见第 3 章相关内容。在推理过程中，系统对每一组输入图像计算其 Dice 系数，并将所有结果进行平均作为模型在测试集上的整体性能表现。

此外，推理模块还提供了两种形式的输出保存：一种是将预测后的三维掩码图像保存为 NIfTI 格式，便于后续医学图像分析和临床应用；另一种是保存各层的特征图可视化图像，有助于研究者观察模型在不同层级的语义学习能力和分割表现。

```
if __name__ == '__main__':
    files_saving_nii = './OUTPUT_nii'
    files_saving_featuremap = './OUTPUT_Featuremap'
    weight_path = 'result/jpgimage/9_11_12012018/epoch49_13-9.951_0C-9.826.pth'

    with open('./config.yaml', 'r') as infile:
        config = yaml.load(infile, Loader=yaml.FullLoader)
        device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
        # device = torch.device('cpu')
        logging.info('Using device {}'.format(device))

        if not os.path.exists(files_saving_nii):
            os.makedirs(files_saving_nii)
        if not os.path.exists(files_saving_featuremap):
            os.makedirs(files_saving_featuremap)

        if config['model'] == 'DenseNet30':
            net = DenseNet30()
        else:
            net = VNet()
        net.to(device=device)
        s = torch.load(weight_path)
        net.load_state_dict(s)
        torch.backends.cudnn.benchmark = True
```

图 4.3 模型推理部分的关键代码

4.3 利用 PyQt5 设计可视化界面

为提升医学图像分割系统的人机交互性和实用性，本系统基于 PyQt5 构建了一套功能完整、界面友好的图形用户界面（GUI）。PyQt5 是 Python 下功能强大的图形界面开发库，其结合了 Qt 框架丰富的控件和灵活的布局机制，能够高效构建响应式桌面应用。本节将重点介绍该可视化界面的功能布局设计、与模型推理模块的联动机制以及图像渲染实现方式。

界面整体采用垂直布局，包括文件选择区、图像显示区、推理控制区、进度与日志区域四个部分。用户可通过点击“选择 CT NII 文件”按钮调用 `QFileDialog`，选取本地的 NIfTI 格式 CT 图像。文件路径一旦确认，将在左侧图像区域中显示该图像的中间层切片。右侧区域则用于展示模型预测结果，以便直观对比模型输出与原始图像。

界面中“开始预测”按钮用于触发推理流程。该流程被封装在独立的 `InferenceThread` 子线程中，以避免阻塞主线程界面响应。推理线程接收输入路径与输出路径，使用 GPU 加速加载 VNet 分割模型，并在 PyTorch 无梯度模式下执行前向传播，输出二值分割结果。结果保存为 NIfTI 格式并与原图进行 Dice 系数评估，同时通过信号机制返回进度与结果。主线程接收信号后刷新进度条显示，并在文本区域输出预测完成信息和 Dice 系数。

为了增强可视化效果，系统不仅支持二维切片图像显示，还集成了基于 VTK（Visualization Toolkit）的三维体渲染模块。具体地，使用 `vtkImageData` 对 NIfTI 数据进行封装，将三维 CT 数据通过 `vtkGPUVolumeRayCastMapper` 实现体绘制。设置合适的颜色映射函数与不透明度传递函数，并通过 `vtkVolumeProperty` 控制插值方式与光照模式，从而实现对不同组织密度的清晰展示。渲染窗口嵌入至 PyQt5 界面中，实现交互式旋转、缩放等操作，提升用户交互体验。

此外，界面还配有进度条和多行文本框，实时反馈推理进展与关键日志信息，提高系统可用性与透明度。该 GUI 应用程序整体架构清晰，操作逻辑流畅，既降低了医学图像处理门槛，又为模型性能调试与评估提供了有力工具。

综上所述，本节所设计的可视化界面不仅集成了 CT 图像的加载、模型预测与结果展示等核心功能，还兼顾了线程处理与图像渲染效率，极大提升了系统的工程实用性与用户交互体验。该设计为本系统的推广与临床落地提供了良好基础。

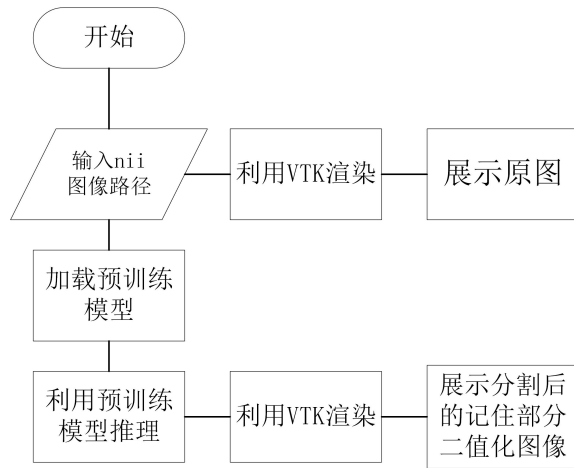


图 4.4 系统流程图



图 4.5 利用 pyqt5 设计的可视化界面

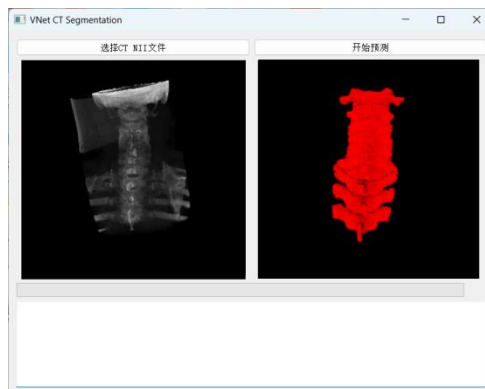


图 4.6 利用该系统预测 CT 图像的脊柱部分

4.4 本章小结

本章主要完成了系统的总体设计与关键模块实现，包括图像读取、模型推理、可视化显示与结果输出等环节。用户首先通过界面选择本地存储的 NIFTI 格式医学图像路径，系统利用 VTK 对该图像进行三维重建与渲染，直观展示脊柱的原始结构信息。随后，用户点击“预测”按钮，系统调用训练完成的深度学习模型对输入图像进行前向推理，提取脊柱区域并生成相应的分割掩码。预测完成后，系统将分割结果与原图进行融合，再次使用 VTK 进行三维可视化展示，使用户能够直观查看分割效果。同时，系统将预测结果保存为新的 NIFTI 文件，并存储在指定路径下，便于后续分析或医学应用。

通过本章的设计与实现，系统已基本具备完整的医学图像分割流程，为后续系统优化与临床适配奠定了坚实基础。

第 5 章 总结与展望

5.1 主要工作与创新点

本研究围绕脊柱术前三维医学图像的自动分割任务，基于 Dense 3D UNet 网络结构，构建了一套高效的脊柱图像分割方案。在公开的 VerSe 数据集上进行训练与验证后，模型在保持较高推理速度的同时，分割准确率达到了 80%，在保证实际应用可行性的基础上显著提升了分割精度，为脊柱图像智能处理提供了有力支撑。

为进一步提升模型性能并解决实际训练过程中的关键难题，本研究在数据预处理与损失函数设计两个方面提出了具有创新性的改进方法。首先，针对医学图像在尺寸标准化过程中易导致形态比例失调的问题，本文提出了一种“先重采样、后等比例缩放并结合零填充”的预处理策略，有效实现了在保持图像尺寸统一的同时保留其原始结构比例，提升了模型对脊柱区域的空间感知能力。其次，为应对传统损失函数在医学图像分割任务中常出现的梯度不稳定、训练难收敛等问题，本文创新性地设计了一种融合型损失函数 FusionLoss，将多种损失函数优势互补地融合于一体，以增强模型在前景-背景区分和边界学习方面的鲁棒性。实验结果表明，该损失函数在提升模型收敛速度的同时，有效降低了训练过程中的 loss 值波动，为模型性能的进一步优化提供了稳定保障。

5.2 后续研究工作展望

尽管本研究所提出的分割方法在标准尺寸图像上取得了较为理想的结果，但由于受限于显存大小的限制，训练过程中输入图像需被统一压缩至 $128 \times 128 \times 128$ 的固定尺寸。在该尺寸下进行训练虽可提升训练效率与显存利用率，但在面对原始分辨率较高的图像时，模型难以捕捉脊柱结构中更多边缘细节信息，导致推理精度有所下降。

针对上述问题，后续研究工作将考虑引入 3D Patch-based 分割策略。该方法旨在将原始尺寸的三维医学图像划分为具有重叠区域的等大小三维子块（3D Patches），在不牺牲图像分辨率的前提下，有效缓解显存压力，使模型能够在更高分辨率下捕捉细粒度结构特征。该策略不仅有助于提升模型对边缘区域细节的识别能力，增强分割精度，同时也简化了预处理流程。由于所有图像在切分后均具备一致的 Patch 尺寸，无需再对原始图像进行零填充处理，从而有效减少图像中的无效信息，提高模型训练与推理的效率与鲁棒性。该思路将在后续工作中进一步深入探索与验证，以提升模型在真实临床应用中的实用性与可靠性。

参考文献

- [1] Sekuboyina A, Kukačka J, Kirschke J S, et al. Attention-driven deep learning for pathological spine segmentation[C]//International workshop on computational methods and clinical applications in musculoskeletal imaging. Cham: Springer International Publishing, 2017: 108-119.
- [2] 肖霖,张笠,唐渔,等.基于全局与上下文双注意力 U-Net 网络于脊柱矢状位 X 线片中分割胸椎及腰椎[J].中国医学影像技术,2025,41(01):128-132.DOI:10.13929/j.issn.1003-3289.2025.01.027.
- [3] 金顺楠,周迪斌,何斌,等.基于多尺度特征融合与注意力机制的脊柱 CT 图像分割[J].计算机系统应用,2021,30(10):280-286.DOI:10.15888/j.cnki.csa.008118.
- [4] 张恒.椎弓根螺钉植入手术导航系统关键技术的研究与实现[D].中国科学院大学(中国科学院工程管理与信息技术学院),2017.
- [5] 刘侠,匡瑞雪,王波,等.基于隐马尔可夫场的脊柱 CT 图像分割算法[J].哈尔滨理工大学学报,2018,23(02):1-5.DOI:10.15938/j.jhust.2018.02.001.
- [6] 李贤.基于多图谱的医学图像分割算法研究[D].杭州电子科技大学,2018.
- [7] 杨玉聃,张俊华,刘云凤.基于三维循环残差卷积的脊柱 CT 图像分割[J].计算机工程,2024,50(04):237-246.DOI:10.19678/j.issn.1000-3428.0067751.
- [8] 何坤金,费雨珊,寇万福,等.基于脊柱融合器设计的椎骨分割及测量方法[J].计算机集成制造系统,2024,30(11):3815-3824.DOI:10.13196/j.cims.2023.0436.
- [9] 申学泉,张勇,张润杰,等.基于 VGG-Net 的 X 射线全脊柱冠状面图像分割方法[J].国外电子测量技术,2024,43(01):135-140.DOI:10.19652/j.cnki.femt.2305336.
- [10] 徐皓.基于 Transformer 和卷积神经网络椎骨 CT 图像分割[D].甘肃中医药大学,2024.DOI:10.27026/d.cnki.ggszc.2024.000527.
- [11] 魏子晨.基于脊柱 X 光图像的椎骨分割及关键点检测研究[D].北京交通大学,2023.DOI:10.26944/d.cnki.gbifu.2023.003138.
- [12] 邢立宁.基于深度学习的三维 CT 图像机会性骨质疏松症筛查模型研究[D].山东大学,2023.DOI:10.27272/d.cnki.gshdu.2023.001871.

- [13]迟晓帆.基于循环残差跳连结构的多方位脊椎 X 光影像分割[D].青岛大学,2023.DOI:10.27262/d.cnki.gqdau.2023.002033.
- [14]陈越鹏.基于脊柱双平面 X 光片的 3D 重建系统的研究与实现[D].北京邮电大学,2023.DOI:10.26969/d.cnki.gbydu.2023.000654.
- [15]吕娟.基于深度学习的医学图像分类选择与分割研究[D].哈尔滨工程大学,2022.DOI:10.27060/d.cnki.ghbcu.2022.002679.
- [16]覃军焱.基于 U-Net 的脊柱侧弯 Cobb 角测量系统研究与实现[D].桂林电子科技大学,2022.DOI:10.27049/d.cnki.gglde.2022.000660.
- [17]张辰翰.基于深度学习的脊柱 CT 图像分割[D].电子科技大学,2021.DOI:10.27005/d.cnki.gdzku.2021.003578.
- [18]刘宇波.基于卷积神经网络图像分割的脊柱三维超声自动扫描系统研究[D].华南理学,2019.DOI:10.27151/d.cnki.ghnlu.2019.000766.
- [19]匡瑞雪.基于隐马尔可夫随机场的脊柱 CT 图像分割算法研究[D].哈尔滨理工大学,2017.
- [20]涂勇成.脊柱侧凸影像 Cobb 角计算机辅助测量算法研究[D].安徽大学,2019.
- [21]Deng Y, Gu F, Zeng D, et al. An effective U-Net and BiSeNet complementary network for spine segmentation[J]. Biomedical Signal Processing and Control, 2024, 89: 105682.
- [22]Saeed M U, Bin W, Sheng J, et al. 3D MFA: An automated 3D Multi-Feature Attention based approach for spine segmentation using a multi-stage network pruning[J]. Computers in Biology and Medicine, 2025, 185: 109526.
- [23]Xie H, Huang Z, Leung F H F, et al. Satr: A structure-affinity attention-based transformer encoder for spine segmentation[C]//2024 IEEE International Symposium on Biomedical Imaging (ISBI). IEEE, 2024: 1-5.
- [24]Da Muten R, Zanier O, Theiler S, et al. Whole Spine Segmentation Using Object Detection and Semantic Segmentation[J]. Neurospine, 2024, 21(1): 57.
- [25]Chen Y, Mo Y, Readie A, et al. VertXNet: an ensemble method for vertebral body segmentation and identification from cervical and lumbar spinal X-rays[J]. Scientific Reports, 2024, 14(1): 3341.

- [26]Deng S, Yang Y, Wang J, et al. Efficient SpineUNetX for X-ray: A spine segmentation network based on ConvNeXt and UNet[J]. Journal of Visual Communication and Image Representation, 2024, 103: 104245.
- [27]Peng L, Hu Y, Zhang K, et al. Cobb angle measurement based on spine segmentation using att unet 3+[J]. IEEE Access, 2024.
- [28]Naga Karthik E, Valošek J, Smith A C, et al. Sciseg: Automatic segmentation of intramedullary lesions in spinal cord injury on t2-weighted mri scans[J]. Radiology: Artificial Intelligence, 2024, 7(1): e240005.
- [29]Saeed M U, Bin W, Sheng J, et al. MSFF: An automated multi-scale feature fusion deep learning model for spine fracture segmentation using MRI[J]. Biomedical Signal Processing and Control, 2024, 91: 105943.
- [30]Bédard S, Karthik E N, Tsagkas C, et al. Towards contrast-agnostic soft segmentation of the spinal cord[J]. Medical Image Analysis, 2025: 103473.

致谢

首先，我衷心感谢我的导师舒禹程副教授在我整个研究生学习期间给予的悉心指导和无私帮助。从课题的选择到论文的撰写，舒老师始终以严谨的学术态度、渊博的专业知识和耐心细致的指导引领我不断进步，使我受益匪浅。

同时，感谢我的朋友和室友们，感谢你们在学习和生活上的陪伴与鼓励，是你们让我在研究生阶段的生活充满温暖与力量，帮助我度过了一个又一个难关。

我还要感谢所有教导过我的老师和在学习过程中给予我帮助的同学，正是你们的支持和鼓励让我不断前行，收获成长。

最后，我要特别感谢党和国家、感谢教育部给予我宝贵的研究生学习机会，是国家良好的教育制度和稳定的发展环境让我能够安心学习、专心科研，并顺利完成学业。

再次向所有帮助、支持和关心我的人表示最诚挚的感谢！

