

X2S-Net: 基于双平面 X 线片的脊柱三维重建

王骞, 张俊华*, 王泽彤, 李博

(云南大学信息学院, 云南 昆明 650504)

摘要: 脊柱的三维模型在治疗脊柱侧弯等脊柱疾病时发挥着重要的作用,但传统的脊椎三维重建方法存在耗时长、主观性强、辐射大等问题。为应对这些挑战,提出一种基于双平面 X 线片的脊柱三维重建网络 X2S-Net。利用患者的正位和左侧位 X 线片作为输入,通过双视角平行编码器、三维重建模块以及分割监督模块后重建出对应位置的脊柱体素模型,实现了从 X 线片到可视化三维模型的端到端生成。X2S-Net 在特征提取阶段,使用了针对双平面 X 线片特点而设计的平行特征编码器,用于提取脊柱的空间信息,并设计多尺度通道注意力机制用于提取特征。在三维模型阶段,X2S-Net 结合传统图像分割任务设计了分割监督模块以提高三维重建效果。实验结果表明,X2S-Net 能够充分利用双平面 X 线片的输入信息对脊柱进行三维重建,各数据集的平均 Hausdorff 距离达到了 6.95 mm, Dice 系数达到了 92.01%。

关键词: 脊柱侧弯; 三维重建; 深度学习; 注意力机制; 计算机断层扫描

源代码链接: <https://github.com/Xaska/X2S-Net>

中图分类号: TP391

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0068706

X2S-Net: Three-Dimensional Reconstruction of Spine Based on Biplanar X-Rays

WANG Qian, ZHANG Junhua*, WANG Zetong, LI Bo

(School of Information, Yunnan University, Kunming 650504, Yunnan, China)

【Abstract】 Three-dimensional models of the spine play an important role in the treatment of spinal disorders such as scoliosis. However, traditional methods for spinal three-dimensional reconstruction suffer from issues such as long processing times, subjectivity, and high radiation exposure. We propose a spinal three-dimensional reconstruction network based on biplanar X-ray images called X2S-Net. The network takes the anteroposterior and lateral X-ray images of the patient as input and reconstructs the corresponding voxel model of the spine using a parallel encoder, three-dimensional reconstruction module, and segmentation supervision module, achieving end-to-end generation from X-ray images to visualize three-dimensional models. In the feature extraction stage, X2S-Net employs a parallel feature encoder designed for the characteristics of biplanar X-ray images to extract spatial information of the spine and incorporates a multiscale channel attention mechanism for feature extraction. In the three-dimensional modeling stage, X2S-Net combines traditional image segmentation tasks with a segmentation supervision module to improve the three-dimensional reconstruction results. The experimental results demonstrate that this method effectively utilizes input information from biplanar X-ray images for three-dimensional reconstruction of the spine, achieving an average Hausdorff distance of 6.95 mm and a Dice coefficient of 92.01% across the datasets.

【Key words】 scoliosis; three-dimensional reconstruction; deep learning; attention mechanism; Computed Tomography(CT)

0 引言

青少年特发性脊柱侧弯(AIS)是一种常见于 10~16 岁青少年的脊柱发育畸形^[1],该类型占全部脊柱侧弯畸形的约 80%^[2]。对于 AIS 的诊断以及治疗,可视化的三维模型能够给予医生的诊断以及治疗方案以直接指导^[3],医生可以通过重建的脊柱三维模型观察患者的脊柱曲度程度和症状的严重程

度,进而制定合适的治疗方案。同时对于曲度较大的严重病患,专业医生还可以通过三维模型制作定制化矫形器^[4]。近年来,使用三维重建技术对内腔、骨骼等人体组织结构进行清晰、全面地重现已成为研究热点^[5-7]。

通常而言,传统的获取脊柱三维模型的渠道主要是计算机断层扫描(CT),此类方法可以分为 3 个步骤:1)通过 CT 成像提取脊柱结构和曲度的详细

收稿日期: 2023-10-30 修回日期: 2023-12-29

基金项目: 国家自然科学基金(62063034,61841112)。

通信作者 E-mail: * 3038547763@qq.com

信息;2)专业医生对二维影像进行去噪、增强和分割,最终从二维影像中提取出脊柱的轮廓和结构;3)通过特定程序与算法将处理后的切片图像堆叠在一起,形成一个三维结构^[8]。此类传统重建方法存在的问题主要是对病人进行 CT 导致的高辐射量和经济负担,以及人工处理步骤所带来的专业技术要求和时间成本^[9]。

X 线片相较于 CT 而言拥有更低的辐射剂量,但其无法提供直观的三维空间信息,往往会给医生的诊断和处理带来较大的挑战^[10]。因此,若能采用 X 线片直接进行三维重建,便可大大降低患者在治疗过程中接受的辐射剂量,同时提供给医生可视化的三维模型用以诊断^[11]。

得益于如全身骨骼三维建模成像系统(EOS)等双平面 X 线片设备的发展,通过深度学习技术从双平面 X 射线图像中重建三维骨骼结构已经成为一个关键的研究领域^[12-14]。KASTEN 等^[15]设计了一个编码和解码的卷积神经网络(CNN),通过跳跃连接将二维图像转化为三维图像,以获得三维输出。CHEN 等^[16]利用循环神经网络去除 CT 图像的伪影,与残差编码器结合重建三维图像。DING 等^[17]通过使用近端前向-后向分裂算法对图像进行预处理,训练深度学习模型进行 CT 图像重建。XIE 等^[18]研究了 CT 图像的伪影,将其从稀疏编码图像中去除,以增强重建图像的质量。MA 等^[19]使用结构相似性指数作为 CT 图像重建的损失函数进行 CT 图像的噪声减少。FENG 等^[20]利用泊松-高斯混合噪声模拟 CT 图像噪声,使用 U-Net 架构重建 CT 图像以提高重建质量。为了增强重建的 CT 图像质量,KANG 等^[21]提出了一种 CNN 算法,将小波变换系数与方向分量相结合,通过小波变换获取 88 种伪影的方向成分。ZHENG 等^[22]使用 CNN 对

每个 CT 切片进行处理,利用 Radon 逆操作符进行 CT 图像重建。GE 等^[23]分析了 CT 图像的逆投影,将正交投影转换为 CT 图像域,以在正弦图像和 CT 图像域中进行特征提取,用于 CT 图像重建。SHEN 等^[24]设计了一个基于 CNN 的转换模块,将从二维编码提取的特征转换为三维编码,用于 CT 图像重建。ZHANG 等^[25]将半二次分裂算法嵌入到 CNN 中,以增强 CT 图像重建的质量。SHIODE 等^[27]从 CT 图像生成数字重建放射图像进行 CNN 训练,实现了腕关节的重建。

得益于深度学习技术的运用,上述方法避免了传统三维重建方法的弊端,但是针对脊柱三维重建而言仍然存在一些问题:1)重建目标往往是 CT 数据而非具体的体素模型,后续仍然需要图像分割等技术对输出结果进行处理后才能得到三维模型;2)特征提取上没有充分利用双平面 X 线片输入的空间信息。基于上述问题,本文提出 X2S-Net 对脊柱进行基于 X 线片的三维重建,并通过实验室采集的数据集进行验证。

1 研究方法

1.1 本文网络结构

针对单张图像存在的位置关系模糊,本文选择正交双视角 X 线片作为重建网络的输入,即正位(PA)与左侧位(LAT)。这一输入使得网络可以有效提取不同位置信息,使其更加准确地重建三维模型;为了更有效地提取 X 线片图像中所需的对应脊柱信息,本文选择分割网络与重建网络同时进行,使得进入重建网络的特征信息关注脊柱而忽略其他软组织等噪声信息。由此,本文提出了一种用于脊柱三维重建的 X2S-Net 网络,其主要结构如图 1 所示。

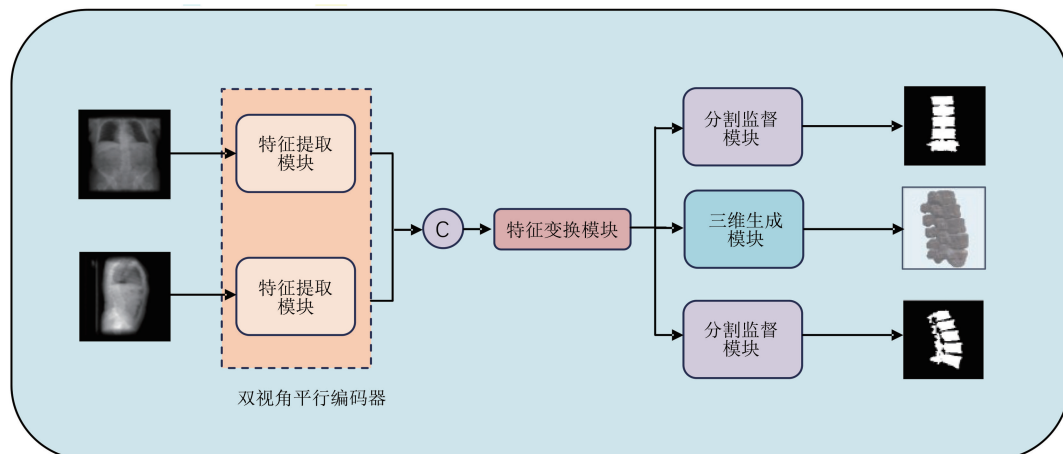


图 1 脊柱三维重建总体框架

Fig.1 Spinal three-dimensional reconstruction overall framework

本模型包括以下 3 个关键组成部分:

1) 双视角平行编码器: 为了有效提取脊柱的空间位置信息, 模型采用来自 2 个不同视角的图像作为输入数据。这些图像经过 2 个对称的编码器处理, 以获取脊柱的特征信息和空间位置信息, 获得的特征信息随后用于下游任务的输入。

2) 分割监督模块: 为了获得脊柱准确的空间位置信息并同时减少 X 线片图像中的噪声干扰, 本文将分割任务融入三维重建过程。这意味着分割任务对三维重建提供了监督。在这个模块中, 同样采用双平面图像进行分割。

3) 三维生成模块: 在经过特征编码器提取并得到分割标签监督后的特征信息后, 脊柱的位置信息和空间结构信息非常丰富, 然后采用三维反卷积来重构三维模型。

这个模型的目标是从多个视角的图像中提取脊柱的特征和位置信息, 同时在三维空间内进行准确

的重建。

1.2 双视角平行编码器

为了有效地提取脊柱三维空间信息, 本文设计了特征编码器, 该编码器采用平行的双线性结构, 主要有 3 个部分组成: 1) 预训练的 ConvNeXt 网络; 2) 多尺度动态通道注意力(MDCA); 3) 多尺度空间注意力同步融合 (MSASF) 模块。

1.2.1 多尺度动态通道注意力模块

本文提出了 MDCA 模块, 旨在进一步提升多尺度特征的表达能力, 该模块结构如图 2 中的多尺度动态通道注意力模块所示(彩色效果见《计算机工程》官网 HTML 版, 下同)。其中, k_4 表示 kernel_size 为 4, s_4 表示 stride 为 4, B 表示 Batch size。动态通道注意力(DCA)^[27]通过学习通道注意力映射, 将输入特征的每个通道与相应的权重相关联, 其结构如图 3 所示。这些权重反映了每个通道对当前任务的贡献, 因此在特征表示过程中自动选择和加权

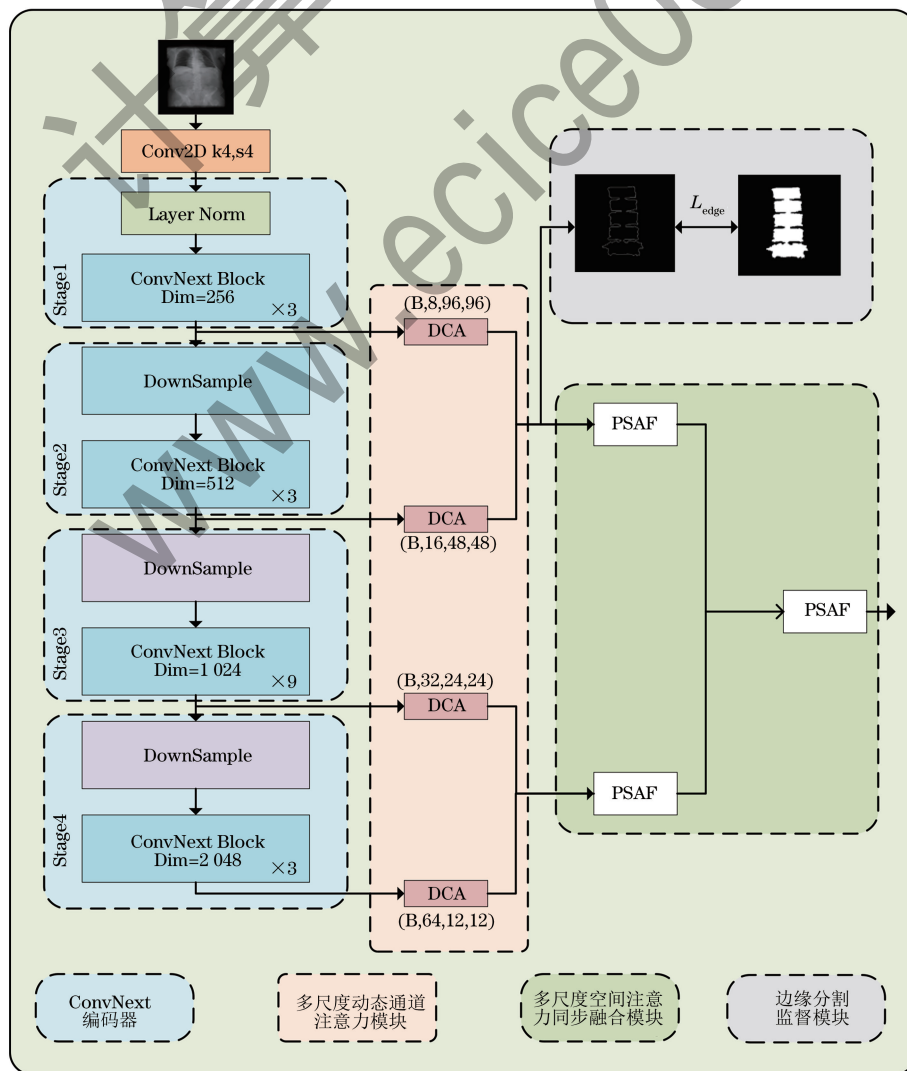


图 2 特征编码模块结构

Fig.2 Feature encoding module structure

不同通道的信息。通常, DCA 模块由 2 个关键组件组成: 通道注意力映射和特征重加权。通道注意力映射使用可学习的参数计算每个通道的注意力权重, 这些权重可以根据输入特征的上下文信息进行动态调整, 以适应不同输入样本。在特征重加权阶段, 原始输入特征与计算得到的注意力权重相乘, 从而获得加权调整后的特征表示。引入 DCA 模块有助于模型更好地学习特征的相关性和重要性, 从而提高模型在各种任务中的性能。它在计算机视觉、自然语言处理和语音识别等领域具有广泛的应用, 已被证明在提高模型性能和解决特征冗余等问题方面非常有效。

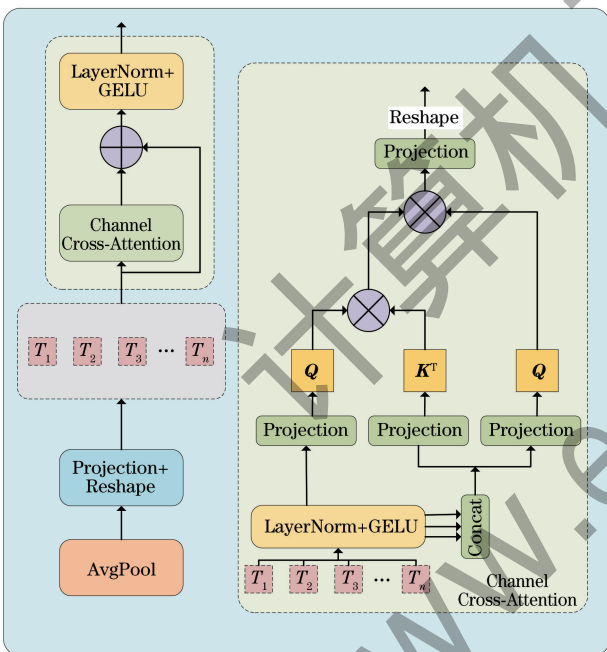


图 3 动态通道注意力机制

Fig.3 Dynamic channel attention mechanism

本文设计了 4 个尺度的 DCA 机制, 如图 3 所示, 这 4 个尺度分别为 96×96 、 48×48 、 24×24 和 12×12 。通过 DCA 对特征提取模块得到的多个不同尺度的特征图进行通道增强, 有效保留了原有不同尺度特征的结构信息。采用 MDCA 机制不仅能保留原有特征的结构信息, 还能进一步提取各个尺度的特征表达能力, 同时减少维度, 降低计算复杂度。

1.2.2 多尺度空间注意力同步融合模块

感受野较大的特征和感受野较小的特征在计算机视觉和深度学习领域具有不同的应用和优势。在多尺度特征提取方面, 以 96×96 和 48×48 的特征信息为例, 它们拥有相对较小的感受野, 能够有效捕获更丰富的局部特征, 并提高抵抗噪声干扰的能力。相比之下, 24×24 和 12×12 的特征信息感受野相

对较大, 能够捕捉更多的全局特征, 为模型提供更丰富的上下文信息, 有助于提升模型对于变化和遮挡等复杂环境的鲁棒性。为了有效整合多尺度特征, 本文提出了 MSASF 模块, 使模型能够动态地调整全局和局部特征的权重, 高效地融合不同尺度的特征信息, 其结构如图 2 中的多尺度空间注意力同步融合模块所示。

MSASF 模块由并行空间注意力融合 (PSAF) 模块组成, 如图 4 所示, 首先将 96×96 与 48×48 以及 24×24 与 12×12 的特征信息通过 PSAF (空间注意力^[28] 与 1×1 卷积并行运算), 然后再通过 PSAF 模块将融合后的特征进一步融合。PSAF 模块采用了空间注意力机制与 1×1 卷积的并行计算。图 4 右侧空间注意力机制模块首先对输入的特征图在通道维度上进行全局池化与平均池化操作, 分别计算出每个通道在空间维度上的最大特征和平均特征, 最终对这些特征经过一系列的线性变换和卷积等操作生成用于表达空间位置信息的特征图。该模块通过空间注意力机制提升了模型对空间结构的表达能力, 并通过并行卷积方式来减小特征信息的损失, 最后将这些特征进行合并。

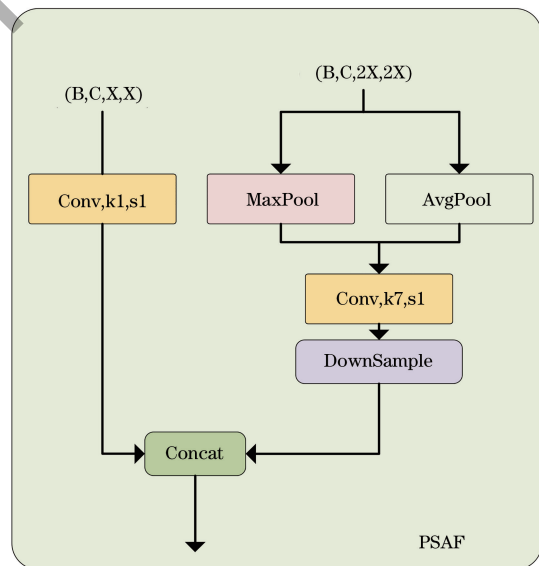


图 4 并行空间注意力融合模块

Fig.4 Parallel spatial attention fusion module

1.3 双平面 X 射线分割监督模块

与 X 射线图像相比, 二值图像更容易获取脊柱的位置关系。如果模型能够成功分割出 X 射线图像中的脊柱, 则表明该模型能够有效地提取出脊柱的位置特征, 并且能够根据不同平面的 X 射线图像来获取脊柱的三维空间位置信息。因此, 本研究引入了双平面 X 射线分割监督模块 (Edge), 该模块的主要功能是基于分割标签监督模型提取特征信息,

以帮助模型更准确地捕捉有关脊柱的关键特征。首先利用分割标签并采用阈值检测算法,以确定分割边缘图的位置,随后结合预测边缘计算边缘损失。边缘预测主要通过前两层特征的融合,并进行上采样预测。这种边缘监督模块的引入有助于提高模型对脊柱结构边缘特征的捕捉,从而改善整体重建效果。如图 5 所示,这一模块能够分别利用双平面二值图像的监督信息来提升特征表达的准确性。

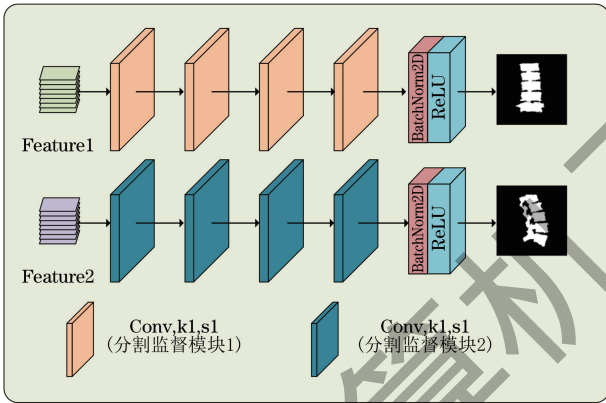


图 5 分割监督模块

Fig.5 Segmentation supervision module

1.4 三维解码器

为了将双平面 X 射线图像精确地还原成三维模型,三维解码器模块主要由 5 层三维转置卷积、BatchNorm3D 与 Sigmoid 组成,旨在将模型提取的特征信息解码为一个分辨率为 $64 \times 64 \times 64$ 像素三维脊柱模型。这个模型采用 Concat 操作首先将来自双平面的特征信息融合,接着通过三维解码器将这些特征信息重建成一个真实世界的三维脊柱模型。该模块的结构如图 6 所示。

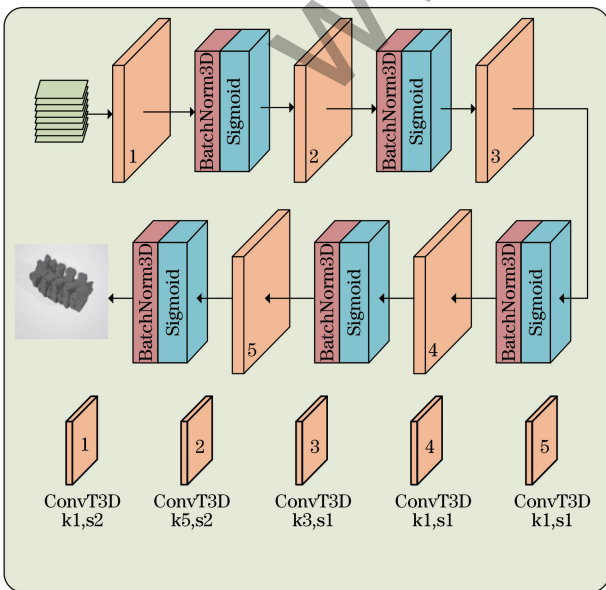


图 6 三维生成模块

Fig.6 Three-dimensional generation module

三维生成模块的设计是为了克服从二维 X 射线图像到三维脊柱模型的挑战。在这个过程中需要综合 2 个平面的信息以获得更全面的脊柱特征,而反卷积操作则负责逐步扩展和还原这些特征,以创建一个高分辨率的三维模型。这个模块的有效性在于它的能力,它可以将二维信息显著提升到更高的维度,并产生一个连贯且准确的三维脊柱模型。

1.5 损失函数

本文采用双平面 X 射线进行三维脊柱重建,以获取脊柱的空间位置关系。为了有效地分割脊柱,引入了边缘检测计算。

损失函数包含 3 个部分:1)分割损失;2)边缘损失;3)三维重建损失。

分割损失与边缘损失主要采用二元交叉熵 (BCE) 损失与 Dice 损失函数的组合。BCE 与 Dice 损失函数的数学公式如式(1)、式(2)所示:

$$B_{BCE}(y, \hat{y}) = y \times \log_a(\hat{y}) + (1 - y) \times \log_a(1 - \hat{y}) \quad (1)$$

$$D_{Dice}(y, \hat{y}) = 1 - \frac{(2 \times |y \cap \hat{y}|)}{(|y| + |\hat{y}|)} \quad (2)$$

式中: y 是真实的二值图像; \hat{y} 是预测的二值图像。

在三维重建损失方面,主要采用 MAE 损失函数,其数学公式如式(3)所示:

$$M_{MAE} = (1/N) \times \sum |V_{true} - V_{pred}| \quad (3)$$

式中: N 是样本数量(通常是体素或点云数量); V_{true} 是真实值(例如,真实的深度值或点云坐标); V_{pred} 是模型的预测值。

2 实验与结果分析

2.1 数据集获取

本研究所使用的输入与输出数据均采集自云南省第一人民医院的 42 例 CT 数据(其中包含 1 例正常数据与 41 例侧弯数据)。首先,通过 TIGRE 算法^[29]对 CT 数据进行其对应 PA 与 LAT 角度下的 X 线片数字重建以获得网络的双视角输入。同时,还采集了与 X 线片相对应的三维脊柱模型用作本研究中生成模型的标签。

鉴于正常脊柱与脊柱侧弯的角度不同,以及考虑到计算能力和内存的限制,将脊柱以 5 个相邻的椎骨为一组进行重建,并将侧弯与正常数据集分开处理。以此为基准一共建立了 8 个数据集,每个数据集包含 1 000 双视角 X 线片输入与对应的椎骨体素模型。具体而言,“T1~T5”涵盖了最初的 5 个胸椎,“T5~T9”代表了接下来的 5 个胸椎,从第 6 个

到第 9 个,“T9~L1”包含了第 9 个到第 12 个胸椎以及第 1 腰椎,“L1~L5”包含了第 1 个到第 5 个腰椎。对于脊柱侧弯数据集,使用了以下标记:“ST1~T5”表示第 1 个到第 5 个胸椎,“ST5~T9”包含了第 6 个到第 9 个胸椎,“ST9~L1”包括第 9 个到第 12 个胸椎以及第 1 腰椎,“SL1~L5”代表了第 1 个到第 5 个腰椎。

2.2 实验环境

在本实验中,使用了 Python 3.9 作为编程语言,以及 PyTorch 1.9 作为深度学习工具包。实验所依赖的硬件设备配置如下:采用了一块 GPU,具有 24 GB 显存,型号为 NVIDIA GeForce RTX 3090,CPU 为 12th Gen Intel® Core™ i7-12700KF,主频为 3.61 GHz,内存规格为 16 GB×4。在模型的训练过程中,本文采用了以下超参数设置:初始学习率为 0.001,并采用指数衰减策略进行学习率的调整,这有助于模型在训练过程中更好地收敛。为了防止模型过拟合,应用了 L2 正则化技术。在优化器方面,采用了随机梯度下降(SGD)作为优化器,以最小化损失函数更新模型参数。这一硬件和超参数配置的选择是为了确保深度学习模型能够在高性能硬件上高效地运行,并以合适的学习策略获得最佳结果。

2.3 评价指标

近年来,对 3D 模型的评估越来越受到关注,最常用的评估指标包括 Hausdorff 距离(HD)、平面表面距离(ASD)、表面重叠(SO)、体积 Dice(VD)系数以及准确率(ACC)。

HD 用于衡量两个点集的最大距离,在 3D 模型评估中,它用来衡量模型之间的形状差异。其数学表达式如式(4)所示:

$$H_{HD}(A, B) = \max(\min(\|a - b\|), a \in A, b \in B) \quad (4)$$

式中: a 为点集 A 中的点; b 为点集 B 中的点。

ASD 用于评估 3D 模型的表面重建质量的指标,它测量了每个点到模型表面的平均距离,以衡量模型的精度,ASD 值越小,表明模型的表面重建质量越好。其数学表达式如式(5)~式(7)所示:

$$d(v, \rho(A)) = \min_{\rho \in \rho(A)} \|v - \rho A\| \quad (5)$$

$$P(X) = \sum_{\rho X \in \rho(X)} d(v, \rho(X)) \quad (6)$$

$$A_{ASD}(A, B) = \frac{1}{|\rho(A)| + |\rho(B)|} [(P(B) + P(A))] \quad (7)$$

SO 用于衡量 2 个模型表面之间的重叠程度,

它测量了 2 个模型表面之间的共享区域与它们的总表面积之间的比例,SO 值越高,表明 2 个模型之间的表面重叠程度越高。其数学表达式如式(8)所示:

$$S_{SO} = (A \cap B) / \min(A, B) \quad (8)$$

VD 是模型之间在三维空间中的模型相似度。将模型的体积表示为 M 。其计算公式如式(9)所示:

$$V_{VD} = \frac{2(M_1 \cap M_2)}{|M_1| + |M_2|} \quad (9)$$

Dice 系数通常用于衡量 2 个集合之间的相似度,在 3D 模型评估中,它通常用来衡量一个重建的 3D 模型与真实地面标签之间的相似度。较高的 Dice 系数表示模型的预测与真实标签之间的重叠度较高。

ACC 通常用于评估分类或分割任务的性能。在 3D 模型的情况下,ACC 可以用来衡量重建的 3D 模型与真实地面标签之间的分类或分割准确度。较高的 ACC 值表示模型的分类或分割性能较好。

3 消融实验结果与分析

3.1 消融实验

为了评估不同模块对重建性能的影响,本文进行了消融实验,建立了 7 种不同的配置。具体配置如下:

1) CN-MDCA-MSASF-Edge:这是最终模型,由双平面图像编码、MDCA、MSASF 模块以及边缘注意力模块组成。

2) CN-MDCA-MSASF:在 1) 的基础上去除了边缘注意力模块。

3) CN-MDCA:在 2) 的基础上进一步去除了 MSASF。

4) CN:在 3) 的基础上去除了动态通道注意力模块。

5) CN-PA:在 4) 的基础上仅使用单个特征提取模块,并且只使用正位 X 线片作为输入。

6) CN-LAT:在 4) 的基础上仅使用单个特征提取模块,并且只使用左侧位 X 线片作为输入。

7) CN-PA-LAT:在 4) 的基础上仅使用单个特征提取模块。

这些配置的制定旨在深入研究各模块对模型性能的贡献,以更全面地理解模型的工作原理,进一步优化并提供指导。对实验结果使用了 8 种不同的评估指标进行综合评估。具体而言,图 7 展示了 7 种不同配置在 ASD 和 Dice 系数 2 个指标

上的实验结果,可以看出,本文的模型取得最小的 ASD 与最大的 Dice 系数,均为最佳的指标;图 8 呈现了 7 种不同配置在 SO 与 HD 指标上的差异表

现,本文模型均取得最好的结果;图 9 提供了 7 种不同配置在 ACC 指标上的实验结果,本文的模型准确率最高,且均超过 0.95。

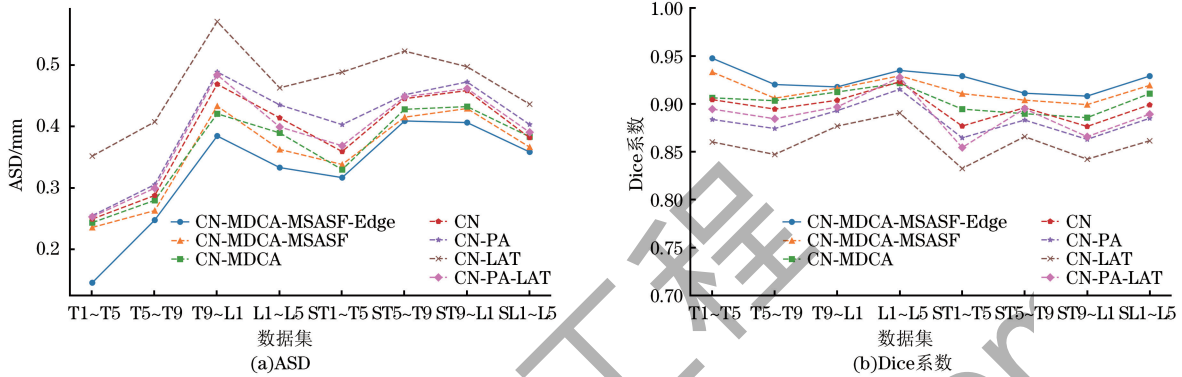


图 7 不同配置在 ASD 和 Dice 指标上的实验结果

Fig. 7 Experimental results of ASD and Dice indexes in different configurations

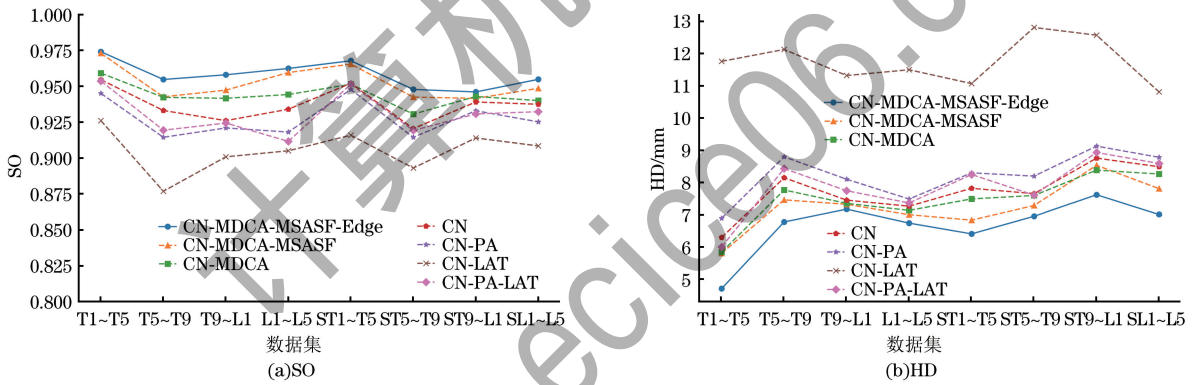


图 8 不同配置在 SO 和 HD 指标上的实验结果

Fig. 8 Experimental results of SO and HD indexes in different configurations

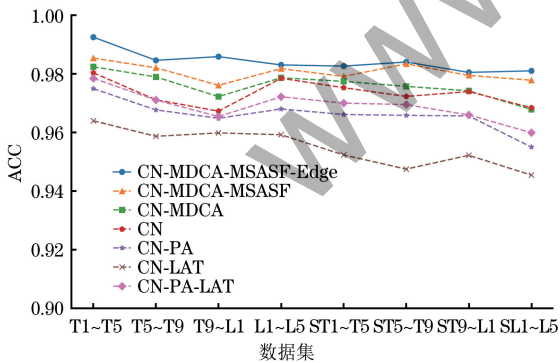


图 9 不同配置在 ACC 指标上的实验结果

Fig. 9 Experimental results of ACC indexes in different configurations

本文设计了几组对比实验,对比结果如下:

1)配置 4 与配置 7 相比,表现出双编码器分别提取双平面图像的单一特征,相较于单编码器,其重建效果更好。这表明双平面图像提供了不同的三维空间位置信息。

2)配置 4 与配置 5、6、7 相比,验证了双平面重建比单平面重建效果更好,双平面图像能够获取更

丰富的脊柱位置信息与空间信息。同时,配置 5 效果优于配置 6 也从侧面体现了 PA 图像所蕴含的空间位置信息相较于 LAT 图像更加充分。配置 4 与配置 7 相比的效果提升主要来自于 2 个特征模块所带来的参数的增加。

3)配置 3 与配置 4 相比,增加了 MDCA 模块,对 ConvNeXt 编码器提取的特征进一步增强,提高了三维模型的重建精度。

4)配置 2 与配置 3 相比,增加了 MSASF 模块,将提取的多尺度特征分类融合,减少了跨尺度融合引起的特征混淆,从数据结果图中可以看出,模型的精度进一步提高。

5)配置 1 与配置 2 相比,增加了边缘注意力模块,增强了浅层特征提取边缘的能力,进一步提高了三维重建的精度。

3.2 不同数据集性能分析

为了评估本文模型在脊柱三维重建方面的重建效果,本文进行了一系列实验。在这些实验中,以 128×128 的双平面 X 射线图像作为模型的输入数

据,并采用相同的数据增强策略来处理训练数据。本研究的核心网络采用了预训练权重进行迁移学习,最终选择了验证集中具有最高准确性的模型进行测试。

使用双平面 X 射线对 5 节椎体进行重建,分别采用 8 个数据集进行验证(第 2.1 节中介绍的 8 个数据集)。评估指标包括 Hausdorff 距离(HD)、平面表面距离(ASD)、表面重叠(SO)、体积 Dice 系数(VD)以及准确率(ACC)5 个指标。

表 1 展示了本研究中提出的 X2S-Net 在对 5 个椎体进行重建时的结果。可以看出,本文模型在 8 组数据集中的平均 HD 达到了 6.953 2 mm,平均 SO 达到了 95.59%,同时在其他评估指标上表现也非常出色。这表明了 X2S-Net 在三维脊柱重建方面表现出卓越的性能。这一高准确率的表现得益于模型能够通过双平面图像提取椎体的空间位置信息,以及通过分割监督模块更加准确地提取椎体的位置,从而更好地实现三维脊柱模型重建。

表 1 不同数据集的对比实验

Table 1 Comparative experiments of different data sets

Dataset	ACC	HD/mm	SO	ASD/mm	VD
T1~T5	0.992 4	4.707 9	0.974 0	0.145 3	0.947 6
T5~T9	0.984 5	6.774 4	0.954 6	0.247 2	0.920 2
T9~L1	0.985 8	7.173 7	0.957 9	0.384 2	0.917 7
L1~L5	0.983 0	6.738 6	0.962 3	0.332 7	0.934 9
ST1~T5	0.982 6	6.406 1	0.967 7	0.316 4	0.929 0
ST5~T9	0.983 9	6.950 6	0.947 7	0.408 6	0.911 1
ST9~L1	0.980 4	7.618 8	0.946 0	0.406 0	0.908 1
SL1~L5	0.980 9	7.010 4	0.954 7	0.358 1	0.929 0
Avg.	0.983 0	6.953 2	0.955 9	0.350 5	0.920 1

3.3 可视化

图 10 分别展示了 8 个数据集中通过模型进行的三维重建结果。这些结果以体素、点云和多边形文件等常见数据格式分别进行了展示。通过可视化结果可以看出,本文的重建模型 X2S-Net 在面对不

同部位以及病患情况时均能得到稳定准确的对应部位椎骨模型。X2S-Net 通过充分利用双平面 X 线片的信息,精准地还原了对应部位椎骨的形态学特征,为医疗专业人员提供了更全面、更精确的患者信息。

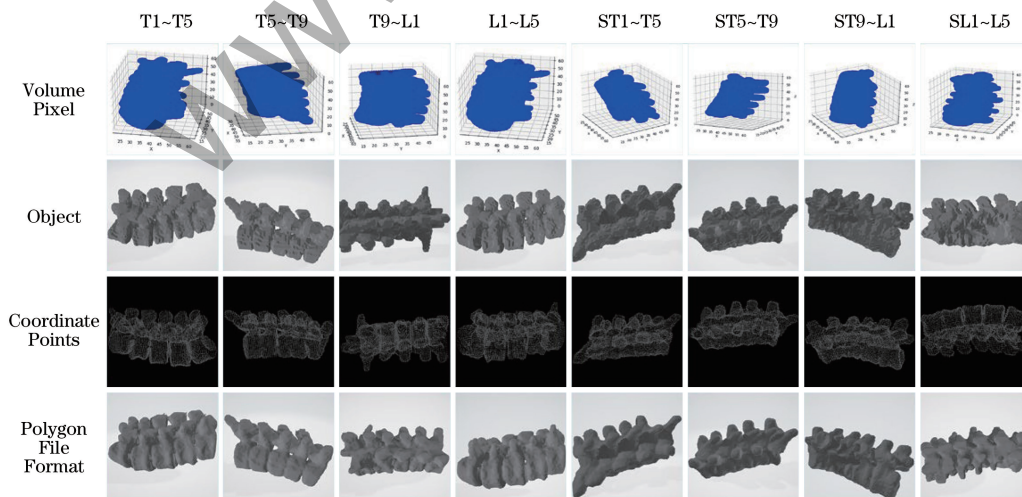


图 10 三维重建可视化结果

Fig.10 Visualization results of three-dimensional

3.4 模型效率对比

由图 10 可知,本文模型在三维脊柱重建中具有较高的重建准确性,为了进一步验证本文模型的重建效率,本文对不同模型以 Params、FLOPS 和 FPS

3 个指标来进行对比实验。实验结果如表 2 所示,提出的最终模型相较于基础模型由于参数数量的增加在推理速度上有所下降,但相较于传统的手动三维重建方法仍然维持了较好的效率。

表 2 模型效率对比实验

Table 2 Model Efficiency Comparative Experiment

模型	Params/ 10^6	FLOPS/ 10^6	FPS
CN-MDCA-MSASF-Edge	7.82	116 190.68	52.35
CN-MDCA-MSASF	7.82	116 187.93	53.46
CN-MDCA	7.79	116 182.72	55.64
CN	7.02	116 018.74	77.33
CN-PA	7.02	114 202.60	71.34
CN-LAT	7.02	114 202.60	71.34
CN-PA-LAT	6.74	116 190.68	71.96

4 结束语

本文提出一种基于双平面 X 射线的脊柱三维重建网络(X2S-Net),通过采用 X2S-Net 模型成功地展示了脊柱的三维重建具备稳定性和高准确度。经过训练的模型可以快速生成脊柱的三维重建结果,并将其以多种三维格式进行存储,提供了广泛的应用灵活性。此外,下一步将探索三维重建结果的应用,包括对脊柱应力分布的详细表征等骨骼力学方向的应用,进而帮助医疗专业人员更好地理解脊柱的负荷情况,为手术规划提供支持。同时,研究脊柱侧弯的物理矫正方法,旨在为脊柱侧弯患者提供更个性化的治疗选择。

参考文献

- [1] 叶华韵,王正飞. 青少年特发性脊柱侧凸的流行病学研究进展[J]. 系统医学, 2022, 7(20): 192-196.
YE H Y, WANG Z F. Progress of epidemiological study on adolescent idiopathic scoliosis[J]. Systems Medicine, 2022, 7(20): 192-196. (in Chinese)
- [2] 海涌,王云生. 青少年特发性脊柱侧凸:探索、发展、创新[J]. 骨科, 2023, 14(1): 1-3.
HAI Y, WANG Y S. Adolescent idiopathic scoliosis: exploration, development and innovation[J]. Orthopaedics, 2023, 14(1): 1-3. (in Chinese)
- [3] CHENG J, CASTELEIN R, CHU W, et al. Adolescent idiopathic scoliosis[J]. Nat Rev Dis Primers 2015, 1:15030.
- [4] LUD Z, LI T, YU W Q, et al. Expert consensus on the design, manufacture, materials, and clinical application of customized three-dimensional printing scoliosis orthosis[J]. Digital Medicine, 2022, 8(1): 2.
- [5] 吴海滨,徐若彤,王爱丽,等. 基于计算机视觉的人体内腔三维重建技术综述[J]. 计算机工程, 2021, 47(10): 1-15.
WU H B, XU R T, WANG A L, et al. Overview of 3D human lumen reconstruction technology based on computer vision[J]. Computer Engineering, 2021, 47(10): 1-15. (in Chinese)
- [6] 方威扬,林东鑫,寇万福,等. 医学图像三维重建系统的研究进展[J]. 中国医学物理学杂志, 2022, 39(7): 823-827.
FANG W Y, LIN D X, KOU W F, et al. Advances in medical image three-dimensional reconstruction system[J]. Chinese Journal of Medical Physics, 2022, 39(7): 823-827. (in Chinese)
- [7] 郑丽萍,李光耀,姜华. 口腔颌面疾病辅助诊断系统的设计与实现[J]. 计算机工程, 2011, 37(21): 279-281, 284.
ZHENG L P, LI G Y, JIANG H. Design and implementation of aided diagnosis system for oral and maxillofacial diseases[J]. Computer Engineering, 2011, 37(21): 279-281, 284. (in Chinese)
- [8] ILLÉS T, SOMOSKEÖY S. Comparison of scoliosis measurements based on three-dimensional vertebra vectors and conventional two-dimensional measurements; advantages in evaluation of prognosis and surgical results[J]. European Spine Journal, 2013, 22(6): 1255-1263.
- [9] HOU Y Y, LIU X X, GUO Y, et al. Strategies for effective neural circuit reconstruction after spinal cord injury: use of stem cells and biomaterials[J]. World Neurosurgery, 2022, 161: 82-89.
- [10] MCCOLLOUGH C H, PRIMAK A N, BRAUN N, et al. Strategies for reducing radiation dose in CT[J]. Radiologic Clinics of North America, 2009, 47(1): 27-40.
- [11] CHAHINE G A, RICHARD M I, HOMES-REGOJO R A, et al. Imaging of strain and lattice orientation by quick scanning X-ray microscopy combined with three-dimensional reciprocal space mapping [J]. Journal of Applied Crystallography, 2014, 47(2): 762-769.
- [12] ZHANG X L, ZHENG Y Q, BAI X M, et al. Femoral image segmentation based on two-stage convolutional network using 3D-DMFNet and 3D-ResUnet[J]. Computer Methods and Programs in Biomedicine, 2022, 226: 107110.
- [13] ČAVOJSKÁ J, PETRASCH J, MATTERN D, et al. Estimating and abstracting the 3D structure of feline bones using neural networks on X-ray (2D) images [J]. Communications Biology, 2020, 3: 337.
- [14] 王帅坤,周志勇,胡冀苏,等. 基于深度学习的肝脏 CT-MR 图像无监督配准[J]. 计算机工程, 2023, 49(1): 223-233.
WANG S K, ZHOU Z Y, HU J S, et al. Unsupervised registration for liver CT-MR images based on deep learning[J]. Computer Engineering, 2023, 49(1): 223-233. (in Chinese)
- [15] KASTEN Y, DOKTOFSKY D, KOVLER I. End-to-end convolutional neural network for 3D reconstruction of knee bones from bi-planar X-ray images[M]. Berlin, Germany: Springer, 2020.
- [16] CHEN H, ZHANG Y, KALRA M K, et al. Low-dose CT with a residual encoder-decoder convolutional neural network[J]. IEEE Transactions on Medical Imaging, 2017, 36(12): 2524-2535.
- [17] DING Q Q, CHEN G Y, ZHANG X Q, et al. Low-dose CT with deep learning regularization via proximal forward-backward splitting[J]. Physics in Medicine & Biology, 2020, 65(12): 125009.
- [18] XIE S P, ZHENG X Y, CHEN Y, et al. Artifact removal using improved GoogLeNet for sparse-view CT reconstruction [J]. Scientific Reports, 2018, 8: 6700.
- [19] MA Y J, FENG P, HE P, et al. Low-dose CT with a deep convolutional neural network blocks model using mean squared error loss and structural similar loss[C]//Proceedings of the 11th International Conference on Information Optics and Photonics. Washington D. C., USA: IEEE Press, 2019: 116-127.
- [20] FENG Z W, LI Z H, CAI A L, et al. A preliminary study on projection denoising for low-dose CT imaging using modified dual-domain U-net[C]//Proceedings of the 3rd International Conference on Artificial Intelligence and Big Data. Washington D. C., USA: IEEE Press, 2020: 223-226.
- [21] KANG E, MIN J H, YE J C. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction[J]. Medical Physics, 2017, 44(10): 360-375.
- [22] ZHENG A, GAO H, ZHANG L, et al. A dual-domain deep learning-based reconstruction method for fully 3D sparse data helical CT [J]. Physics in Medicine & Biology, 2020, 65(24): 245030.
- [23] GE Y S, SU T, ZHU J T, et al. ADAPTIVE-NET: deep

- computed tomography reconstruction network with analytical domain transformation knowledge[J]. *Quantitative Imaging in Medicine and Surgery*, 2020, 10(2): 415-427.
- [24] SHEN L Y, ZHAO W, XING L. Patient-specific reconstruction of volumetric computed tomography images from a single projection view via deep learning[J]. *Nature Biomedical Engineering*, 2019, 3: 880-888.
- [25] ZHANG H M, LIU B D, YU H Y, et al. MetaInv-net: meta inversion network for sparse view CT image reconstruction[J]. *IEEE Transactions on Medical Imaging*, 2021, 40(2): 621-634.
- [26] SHIODE R, KABASHIMA M, HIASA Y, et al. 2D-3D reconstruction of distal forearm bone from actual X-ray images of the wrist using convolutional neural networks[J]. *Scientific Reports*, 2021, 11: 15249.
- [27] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[M]. Berlin, Germany: Springer, 2018.
- [28] CHEN W C, ZHANG D, LI M, et al. STCAM: spatial-temporal and channel attention module for dynamic facial expression recognition[J]. *IEEE Transactions on Affective Computing*, 2023, 14(1): 800-810.
- [29] KASER S, BERGAUER T, BIRKFELLNER W, et al. First application of the GPU-based software framework TIGRE for proton CT image reconstruction[J]. *Physica Medica*, 2021, 84: 56-64.

编辑 索书志