

融合双域特征的 CBCT-CT 生成对抗网络^①

王兴凡¹, 李 曙²

¹(吉首大学 通信与电子工程学院, 吉首 416000)

²(广州医科大学 生物医学工程学院, 广州 511436)

通信作者: 李 曙, E-mail: shuli@gzhmu.edu.cn



摘 要: 锥形束计算机断层扫描 (cone beam computed tomography, CBCT) 因其与现代直线加速器系统的集成而被广泛用于图像引导放射治疗. 然而, 由于其图像质量不如 CT, 这给实现最佳治疗计划带来了重大挑战. 本研究提出一个名为 DDFGAN (dual-domain feature fusion generative adversarial network) 的新模型, 旨在改善 CBCT 图像质量, 使其接近 CT 水平. 该模型采用双分支架构: 第 1 分支通过引入 RFB 模块来提取空间域中的多尺度特征; 第 2 个分支则设计了一个专门针对 CBCT 到 CT 合成的频率域特征提取模块. 通过将这两个分支的特征融合, DDFGAN 显著提升了 CBCT 的成像质量. 此外, 本模型引入几何一致性损失, 将传统的双向生成网络转变为单向生成网络, 这不仅更符合临床应用需求, 还大幅减少了训练时间. 实验结果显示, DDFGAN 在生成少伪影的合成 CT 图像方面优于其他 4 种比较方法, 且其合成图像的 HU 值也更接近于 CT 图像, 显著提高了自适应放射治疗的准确性.

关键词: 生成对抗网络; 锥形束计算机断层扫描; 合成 CT; 图像生成; 无监督网络

引用格式: 王兴凡, 李曙. 融合双域特征的 CBCT-CT 生成对抗网络. 计算机系统应用, 2025, 34(6): 33-40. <http://www.c-s-a.org.cn/1003-3254/9857.html>

CBCT-CT Generative Adversarial Network with Dual-domain Feature Fusion

WANG Xing-Fan¹, LI Shu²

¹(School of Communication and Electronic Engineering, Jishou University, Jishou 416000, China)

²(School of Biomedical Engineering, Guangzhou Medical University, Guangzhou 511436, China)

Abstract: CBCT is widely used in image-guided radiation therapy due to its integration with modern linear accelerator systems. However, its inferior image quality compared to CT poses significant challenges in achieving optimal treatment planning. This study proposes a new model named DDFGAN (dual-domain feature fusion generative adversarial network), aiming at increasing the image quality of CBCT to that of CT as much as possible. The model adopts a dual-branch architecture: the first branch extracts multi-scale features in the spatial domain through the introduction of an RFB module; the second branch designs a frequency domain feature extraction module specifically for CBCT to CT synthesis. By fusing features from both branches, DDFGAN significantly enhances the imaging quality of CBCT. Additionally, the model incorporates a geometric consistency loss, transforming the traditional bidirectional generative network into a unidirectional one, which not only aligns more with clinical application requirements but also substantially reduces training time. Experimental results show that DDFGAN outperforms the other four comparative methods in generating synthetic CT images with fewer artifacts, and the HU values of synthetic images are closer to those of CT images, significantly improving the accuracy of adaptive radiation therapy.

Key words: generative adversarial network (GAN); cone beam computed tomography (CBCT); synthetic CT; image generation; unsupervised network

① 基金项目: 国家自然科学基金 (42164006); 湖南省自然科学基金 (2022JJ30474)

收稿时间: 2024-11-08; 修改时间: 2024-11-29; 采用时间: 2024-12-16; csa 在线出版时间: 2025-03-24

CNKI 网络首发时间: 2025-03-25

脑肿瘤患者在治疗过程中,软组织区域的体积和位置可能会发生变化.自适应放射治疗是一种能够更好地适应这些变化的治疗方法,能够有效应对放疗过程中患者软组织区域显著变化带来的挑战.自适应放疗通常使用锥形束计算机断层扫描 (cone beam computed tomography, CBCT) 来监测每次治疗前的生理变化,并根据这些观察结果调整治疗计划,从而确保治疗的准确性^[1].然而, CBCT 图像经常会受到伪影的影响,例如条纹、阴影和散射污染,导致 HU 值严重不准确,如图 1 所示.这些问题限制了将 CBCT 用于定量应用的可行性^[2].此外, CBCT 因其与现代直线加速器系统的集成而被广泛用于图像引导放射治疗.但由于其图像质量不如 CT,这给实现最佳治疗计划带来了重大挑战.

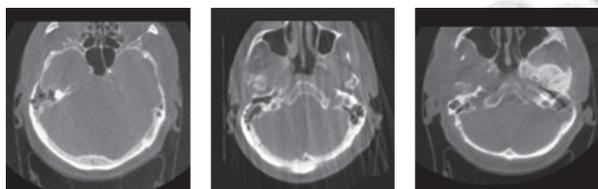


图 1 受伪影影响的 CBCT 图像

基于 CBCT 图像进行自适应放射治疗的研究主要聚焦于两个方向:一是通过伪影校正提高 CBCT 图像质量,二是根据 CBCT 图像生成与计划 CT 图像相近质量的合成 CT 图像.

伪影校正的方法通常包括防散射栅格、主调制光束滤波器和格子状光束阻挡器等技术.这些方法的目标是减少 CBCT 图像中的散射伪影,提高图像的清晰度.然而,这些技术尽管能够有效减轻伪影,但通常伴随着降低系统效率的风险^[3].例如,放散射栅格可以减少散射光的影响,但也可能导致信号强度减弱,从而降低图像的对比度和分辨率.此外,主调制光束滤波器和格子状光束阻挡器等技术虽然在校正伪影方面有较好的效果,但它们的复杂性可能增加图像采集时间,并降低整体的图像质量^[4,5].

另一方面,基于深度学习的方法在去噪^[6]、超分辨率^[7]和减少伪影^[8]等医学图像分析任务中展现了巨大的潜力并且取得了显著成功.在此背景下,基于 CBCT 图像生成合成 CT 图像的方法也日益成熟.该方法通过训练算法学习 CBCT 图像和计划 CT 图像之间的映射关系,生成与原始计划 CT 相匹配的图像,以达到临床上所需的精确度和图像质量.这种技术在不牺牲系统效率的情况下,有效提升 CBCT 图像的实用价值.

1 相关工作

在基于 CBCT 到 CT 图像合成的研究中,现有的合成方法主要可以分为有监督方法和无监督方法.

1.1 有监督方法

有监督的方法通过使用配准后的 CBCT 和 CT 图像对来进行模型的训练, Xie 等人^[9]使用两步配准法生成配对数据,然后训练卷积神经网络从 CBCT 图像生成合成 CT 图像.然而,配准过程中会有不可避免的误差,这会导致训练的模型学习到不正确的映射,从而在合成 CT 图像的过程中造成解剖结构的不正确变化.

1.2 无监督方法

无监督方法无需配准的图像对进行训练. Liang 等人^[10]使用循环一致生成对抗网络 (cycle-consistent generative adversarial network, CycleGAN) 从 CBCT 图像生成了合成 CT 图像,但 CycleGAN 依赖于循环一致性损失,无法充分保留细微的解剖结构,导致在某些复杂区域出现不准确的生成结果. Deng 等人^[11]提出一种具有扩张卷积一致性生成对抗网络的循环-残差连接方法,改善了 CBCT 的图像质量.然而,扩张卷积的设计可能在某些解剖区域引入非自然的平滑效应,影响局部细节的保真度. Zhang 等人^[12]提出了一种多模态无监督表征分离学习框架,利用屏息 CBCT 图像作为指导,生成了更好的合成 CT 图像.然而,该框架完全在图像域中运行,无法完全消除固有的噪声和伪影.

2 网络结构

本文提出了一种融合双域特征的生成对抗网络,用于 CBCT 到 CT 图像的合成,其核心目标是训练生成器学习 CBCT 和 CT 图像之间的映射关系,从而生成高质量的合成 CT 图像.具体来说,设定两个不同的域: X 域和 Y 域, X 域代表 CBCT 图像, Y 域代表目标 CT 图像.生成器 G 的任务是将输入的 X 域图像 (即 CBCT 图像) 转换为 Y 域图像 (即合成 CT 图像),而判别器 D 则负责判断输入的图像是否来自目标域 Y,以确保生成器生成的图像尽可能接近真实 CT 图像.

2.1 生成器

如图 2 所示,生成器由编码器 E、解码器 D 和特征提取网络组成.输入图像首先经过编码器编码,然后通过特征提取网络提取图像特征信息,然后通过解码器解码后,生成器将输出最终图像.

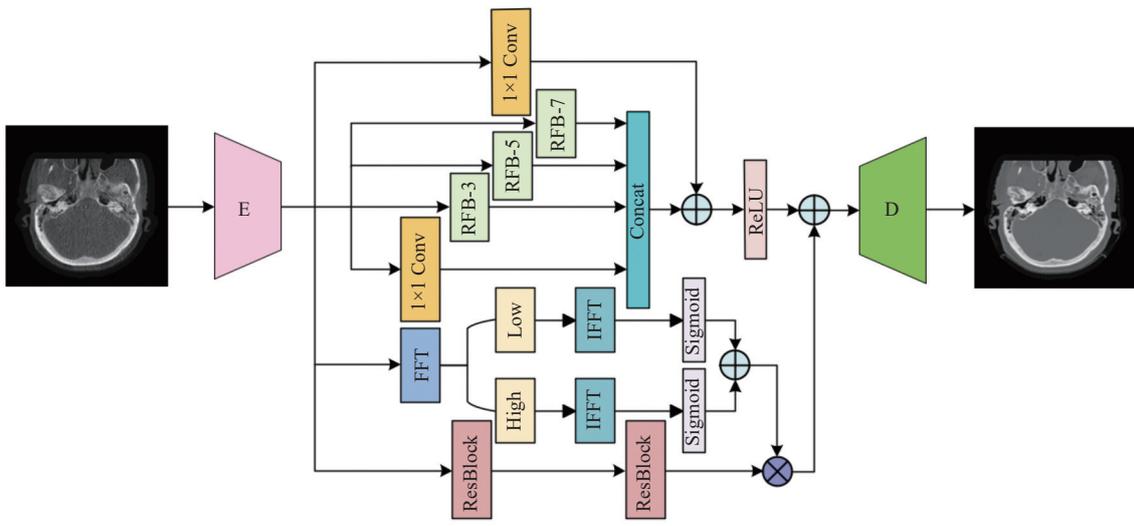


图2 生成器的网络架构图

在大多数的生成对抗网络的生成器中, 通常仅能提取单一尺度的特征图像, 这可能会导致某些语义特征的丢失. 为了解决这一问题, 本文引入了RFB^[13]作为主干模块, 构成特征提取网络的第1路分支, 用以提取多尺度的特征信息. RFB模块的设计灵感来源于人类视觉系统中的视皮层处理机制, 旨在通过模拟不同感受野的特征来增强对不同尺度目标的响应能力. 具体来说, 输入的特征图首先通过3个不同尺度的RFB模块, 这些模块对应不同的感受野大小, 以捕捉多尺度的空间信息. 图2中展示了3种规格的RFB模块: RFB-3、RFB-5和RFB-7, 分别对应3×3、5×5和7×7的卷积核大小. 这样的设计使得RFB模块能够同时捕捉局部细节和全局语义特征, 从而提升模型的综合判别能力. 接着, 经过1×1卷积操作之后, 将多尺度特征进行整合, 在保留多尺度信息的同时, 避免了直接拼接导致的通道冗余问题. 随后, 这些特征图与另一路经过1×1卷积处理的特征图相加, 进行特征融合, 还有效强化了多尺度特征之间的协同表达能力. 最后, 融合后的特征图经过ReLU激活函数激活, 完成第1分支的多尺度特征提取过程. 通过引入RFB模块, 模型在生成高质量图像时的鲁棒性和细节表现均得到了显著增强.

RFB的具体结构如图3所示(以RFB-3为例). 输入特征图 x 首先通过一个1×1的卷积层进行通道压缩, 这一步骤的目的是在减少计算开销的同时保持特征的有效性. 然后, 特征图依次通过一个1×3和一个3×1的卷积核, 以增强横向和纵向的特征提取能力, 分别强化

对横向和纵向边缘信息的捕捉. 接下来, 特征图通过一个膨胀率为3的3×3膨胀卷积层, 这一层扩大了感受野, 同时保证了局部特征的连贯性, 从而在不显著增加参数量的前提下增强全局上下文信息的感知能力. 这样的设计不仅增加了模型的感受野, 还增强了多尺度特征的提取能力.

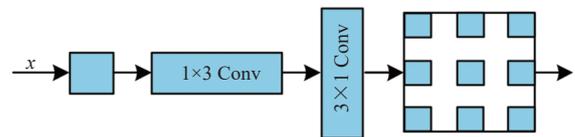


图3 RFB-3的结构图

受到Li等人^[14]提出的FGDM的启发, 我们注意到CBCT和CT图像的差异主要集中在中频带, 这为CBCT到CT图像的转换提供了新思路. 因此, 本文设计了第2路分支, 用于捕获CBCT和CT图像之间的相似频域信息, 以生成更高质量的合成CT图像. 具体来说, 本文首先对特征图进行傅里叶变换, 将图像特征从空间域转换到频域中, 以便更有效地对其频谱特征进行操作. 在频域中, 本文设计了两个滤波器, 分别用于提取图像的低频和高频信息. 低频部分主要捕捉图像中的大尺度结构和整体强度变化, 而高频部分则着重于细节信息和边缘特征. 通过这两个滤波器的分离, 能够更加精确地提取和处理CBCT和CT之间的差异信息, 特别是保留相似的频域成分. 接着, 对低频和高频部分分别进行逆傅里叶变换, 将其恢复为空间域的图像特征. 恢复后的特征图分别代表了不同频率下的

空间信息, 随后经过 Sigmoid 激活函数将它们整合相加, 从而得到一个增强的相似区域特征图. 这一步骤的设计旨在对相似区域进行强化, 以便更好地提取关键的细节特征. 随后, 将频域增强的特征图与通过两个 ResBlock^[15]提取的特征相乘, 提取出 CBCT 和 CT 之间相似区域的特征图.

最后, 将来自两路分支的特征图相加, 送入解码器以生成最终的图像. 第 1 路分支的多尺度空间信息与第 2 路分支的频域相似特征相结合, 有助于网络更全面地掌握图像的空间和频率特征, 最终得到更加逼真的合成 CT 图像. 这样, 通过频域增强和多尺度融合的设计, 成功提升了 CBCT 到 CT 合成过程中的图像质量和细节保留能力, 使生成的 CT 图像在结构一致性和细节表现上均得到了显著改善.

ResBlock 的结构图如图 4 所示, 其内部包括两个卷积层, 卷积层之间配有归一化和非线性激活函数. 通过这种残差连接方式, ResBlock 在增加网络深度和模型复杂性的同时, 仍能保持良好的可训练性和稳定的梯度传播, 并实现较好的泛化性能.

2.2 判别器

本文采用了 Zhu 等人^[16]提出的 PatchGAN 作为网

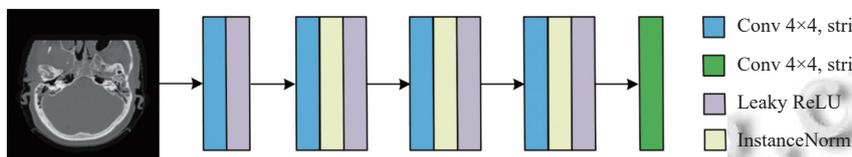


图 5 判别器的网络架构图

2.3 损失函数

为了促进网络生成更高质量的图像并使整个网络更加稳定, 总的损失函数包括对抗性损失^[17]和几何一致性损失^[18]组成.

2.3.1 对抗性损失

生成对抗网络由生成器和判别器组成, 生成器的目标是生成看起来与真实数据相似的假数据, 而判别器的目标是区分生成数据和真实数据. 对抗性损失 \mathcal{L}_{adv} 可以表达为生成器和判别器之间的博弈过程.

生成器希望生成的数据能够欺骗判别器, 使其无法区分生成的数据和真实的数据. 生成器的目标是 minimized 判别器的判断, 即让判别器尽可能认为生成的数据是真实的.

判别器则试图最大化生成数据与真实数据的区分

网络的判别器. 在传统的 GAN 中, 判别器的主要目标是对输入图像整体进行判断, 决定其是真实图像还是生成器生成的图像. 这类判别器通常以完整图像作为输入, 并输出一个标量值, 代表图像为真实的概率. 这种方法重点关注图像整体的真实性.

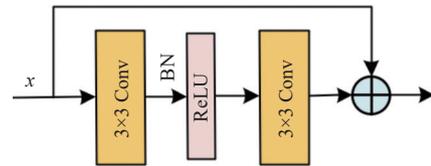


图 4 ResBlock 的结构图

然而, PatchGAN 是一种更加细粒度的判别器, 专注于图像的局部区域. 与传统的判别器不同, PatchGAN 将输入图像划分为若干小块, 并独立评估每个块的真实性和. 每个小块都分别被判断为真实或虚假, 最终的输出是所有小块判断结果的平均值或其他形式的合并结果.

如图 5 所示, DDFGAN 的判别器采用了 PatchGAN 的结构, 包括 4 个 4×4 卷积层, 步长为 2, 最后一层为 4×4 卷积层, 步长为 1, 并使用 Sigmoid 激活函数进行处理. 这样, PatchGAN 能够更加精确地评估图像的局部细节, 从而提高生成图像的质量.

能力, 目标是将生成数据分类为假数据, 并将真实数据分类为真数据.

通过这种对抗的方式, 可以有效提高生成器生成数据的质量, 使其更加逼真和接近真实数据的分布. 本文的对抗性损失可以表达为:

$$\mathcal{L}_{adv} = \mathbb{E}_{y \sim P_Y} [\log D(y)] + \mathbb{E}_{x \sim P_X} [\log (1 - D(G(x)))] \quad (1)$$

其中, G 是生成器, D 是判别器.

2.3.2 几何一致性损失

由于本文研究的目的是将 CBCT 图像生成 CT 图像的单向生成任务, 因此本文引入了几何一致性损失来实现这一目的. 几何一致性损失可以保持生成图像与参考图像在轮廓、形状等几何特征上的一致性, 从而提高生成图像的可信度和实用性.

具体来说, 将 $x \in X$ 和 $\bar{x} = f(x)$ 分别输入生成器 G 和

\tilde{G} 中. 几何一致性约束要求输出 $y' = G(x)$ 和 $\tilde{y}' = \tilde{G}(\tilde{x})$ 也应该与 x 和 \tilde{x} 一样满足 $\tilde{y}' \approx f(y')$, 此公式也可以转化为 $y' \approx f^{-1}(\tilde{y}')$. 本文的几何一致性损失可以表达为:

$$\mathcal{L}_{\text{geo}} = \mathbb{E}_{x \sim P_X} \left[\left\| G(x) - f^{-1}(\tilde{G}(f(x))) \right\|_1 \right] + \mathbb{E}_{\tilde{x} \sim P_{\tilde{X}}} \left[\left\| \tilde{G}(f(x)) - f(G(x)) \right\|_1 \right] \quad (2)$$

其中, $f^{-1}(\cdot)$ 是一个逆时针旋转 90° 的函数.

本文将对抗性损失和几何一致性损失相结合, 能够实现单向无监督的域映射, DDFGAN 完整的损失函数可以表达为:

$$\mathcal{L}_{\text{DDFGAN}} = \mathcal{L}_{\text{adv}}^G + \mathcal{L}_{\text{adv}}^{\tilde{G}} + \lambda \mathcal{L}_{\text{geo}} \quad (3)$$

其中, λ 是几何一致性损失的权重系数, 可根据不同任务微调, 文中使用 $\lambda = 2.0$.

3 实验分析

3.1 基线模型

本文选取 CycleGAN^[16]、UNIT^[19]、FastCUT^[20]和 GcGAN^[18]作为基线模型. CycleGAN 通过引入循环一致性损失, 确保两个域之间的映射符合一致性约束. UNIT 提出了共享潜在空间的假设, 结合可变分自编码器来实现图像转换. FastCUT 在 CycleGAN 的基础上, 通过对比损失函数^[21], 实现了更快速、轻量化的图像转换. GcGAN 通过几何一致性损失, 实现了单向的图

像转换.

3.2 数据集

本文使用的数据来自 SynthRAD2023 挑战赛^[22,23], 并对获取的数据进行了以下预处理步骤. 首先, 删除了不完整的图像, 保留了每名患者 CT 扫描中心的 100 张切片. 其次, 为避免图像中非解剖结构的信息干扰, 本文使用规划 CT 的结构文件生成与规划 CT 图像大小相同的掩膜, 将身体轮廓外部区域设为 0, 内部区域不做处理. 第三, 统一所有切片的尺寸为 256×256 . 最后, 对 CBCT 图像和 pCT 图像进行了线性归一化处理, 将像素值归一化到 $[-1, 1]$ 区间.

3.3 实现细节

所有方法均基于 PyTorch 2.0.1 版本实现. 本文采用 Adam 优化器, 起始学习率设置为 $2E-4$, batchsize 为 8, 并通过 lambda 函数动态调整学习率. 所有训练和测试任务均在 NVIDIA RTX A6000 GPU 上进行.

3.4 视觉评估

如图 6 所示, 共展示了 3 行 6 列的图像. 图中每一行展示了一个不同的样本图像. 从左至右, 每行的图像依次为 CBCT 图像、CycleGAN、FastCUT、UNIT、GcGAN、DDFGAN 合成的 CT 图像, 以及目标 CT 图像. 图 6 中用红色箭头标注的区域特别显示了不同方法在关键结构或细节上的合成效果.

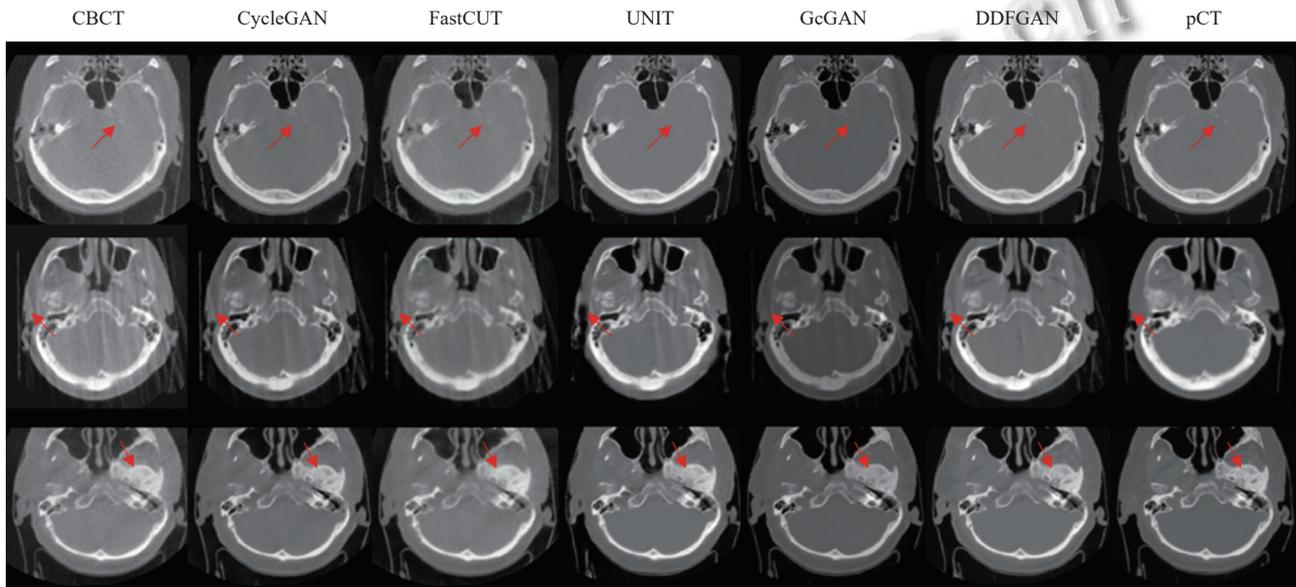


图 6 不同方法生成的 sCT 图像的视觉对比

CycleGAN 和 FastCUT 在图像纹理和整体轮廓的还原方面表现较好, 但去噪效果不理想. GcGAN 在几

何对齐上具有较好的表现, 生成的图像在整体形状上与目标 CT 图像相似, 但由于其对比度偏低, 使得红色

箭头标注区域的局部细节不够清晰. 相比之下, UNIT 和 DDFGAN 在整体对比度和亮度上更接近目标 CT 图像, 但 UNIT 在红色箭头标注的区域中, 细小纹理和边缘细节仍有所欠缺. 综合来看, DDFGAN 生成的图像在多个方面与目标图像最为接近, 不仅在轮廓形状和边缘清晰度上表现出色, 还在纹理细节、对比度和亮度方面与目标 CT 有更好的匹配.

为了比较不同方法生成的 sCT 图像之间的差异, 本文分析了一些图像的剖面数据, 并在图 7 中展示了

这些数据. 图 7 中 (a) 和 (c) 分别展示了垂直和水平方向的剖面图, 其中红色实线穿过的区域的详细数据在图 7(b) 和 (d) 中呈现. 在图 7(b) 和 (d) 中, 不同颜色的虚线代表不同方法生成的 sCT 图像在红色实线区域的 HU 值变化. 具体来说, 蓝色代表 CBCT、橙色代表 CycleGAN、绿色代表 UNIT、红色代表 FastCUT、紫色代表 GcGAN. 而棕色和黑色实线分别代表规划 CT 和 DDFGAN. 每个红框选中的范围都展示了相应区域的放大图.

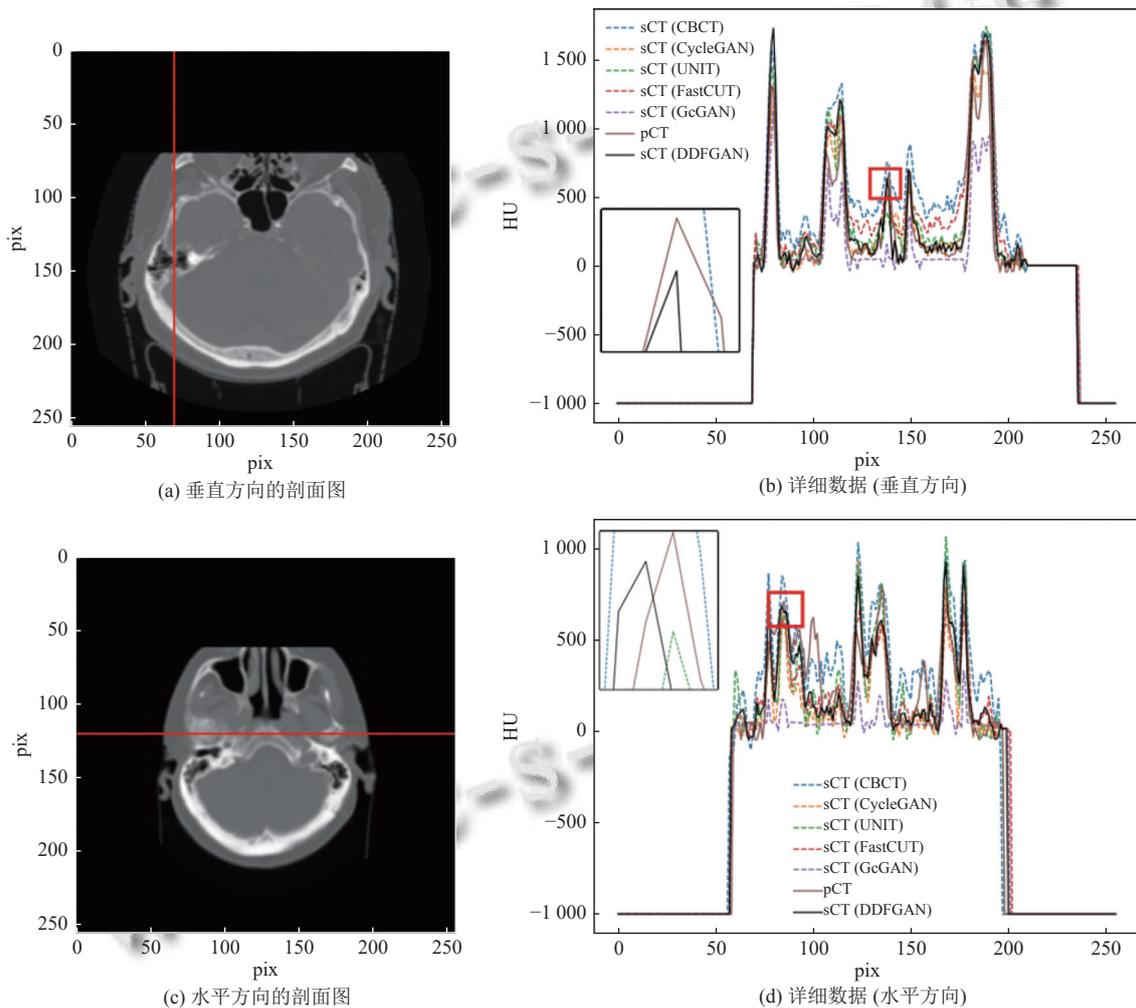


图 7 垂直和水平方向的剖面统计数据

在明显的波峰区域, 黑色和棕色实线在位置和幅度上表现出高度的一致性, 这表明本文的方法很好地保留了规划 CT 的结构. 在接近 0 值附近的波动幅度可以反映模型的噪声抑制效果, 本文的方法在该范围内更好地拟合了 pCT 的 HU 值, 这说明本文的方法在噪声抑制方面表现优异, 成功地减少了噪声对图像质量

的影响.

3.5 定量评估

图 8 和图 9 分别通过小提琴图展示了测试样本的 SSIM 和 PSNR 值的分布情况. SSIM 用于评估两幅图像在结构、亮度和对比度方面的相似性. PSNR 用于衡量图像之间的相似性. 在小提琴图中, 白色区域表示数

据的中位数,内部矩形的宽度对应上四分位数和下四分位数,而小提琴的宽度表示数据分布的概率,宽度越大表示该区域的数据分布越密集。

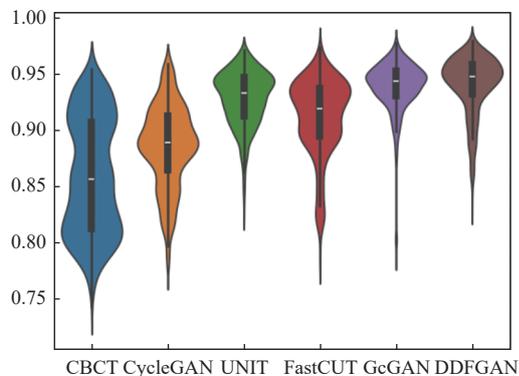


图8 不同方法的SSIM数据分布情况

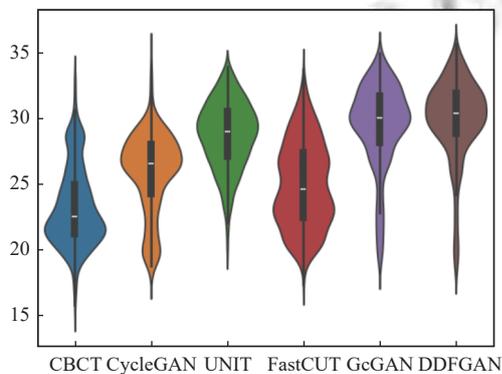


图9 不同方法的PSNR数据分布情况

在图8中, CBCT、CycleGAN和FastCUT的SSIM值分布较为宽泛,表明这些方法生成的图像一致性有待提升。与其他方法相比, DDFGAN显示出最高的SSIM分布,其数据主要集中在0.95附近,表明DDFGAN在保持生成图像的结构细节方面表现优异,具有较高的一致性和稳定性。

在图9中, CBCT、CycleGAN和FastCUT的PSNR值较低,说明其生成的图像质量不佳。相比之下, UNIT的PSNR值分布相对集中,表明其在一致性方面有较好的表现。而DDFGAN则显示出最高的PSNR分布,表明其生成的图像在质量方面优于其他方法,能够更准确地生成高质量图像。

表1列出了更详细的结果。本文的方法在所有指标上都优于其他对比方法。其中, SSIM从0.859提高到0.94(如表1第2列所示), PSNR从23.28 dB提高到29.934 dB(如表1第3列所示), FID从87.277降低至24.091(如表1第4列所示)。

表1 不同方法的指标对比

网络	SSIM↑	PSNR↑(dB)	FID↓
CBCT	0.859±0.051	23.280±3.023	87.277
CycleGAN	0.886±0.036	25.909±3.067	68.750
UNIT	0.928±0.024	28.747±2.539	51.800
FastCUT	0.912±0.035	24.958±3.219	43.720
GcGAN	0.937±0.033	29.808±2.931	30.719
DDFGAN	0.940±0.026	29.934±3.105	29.067

3.6 消融实验

本文设计了消融实验来验证所提出创新措施的有效性,实验结果如表2所示。通过对原始CycleGAN网络的生成器进行改进,形成了DDFGAN的初始网络结构,所有性能指标均得到了提升(如表2第3行所示)。这种提升源于双域融合特征网络架构的设计,该架构使模型更有效地捕获多尺度的双域特征,从而使生成的图像更接近于规划CT图像。

表2 DDFGAN消融实验的指标对比

网络	SSIM↑	PSNR↑(dB)	FID↓
CBCT	0.859±0.051	23.280±3.023	87.277
CycleGAN	0.886±0.036	25.909±3.067	68.750
DDFGAN	0.928±0.034	28.784±2.797	34.265
DDFGAN+ \mathcal{L}_{geo}	0.940±0.026	29.934±3.105	29.067

在加入了几何一致性损失后,各项指标进一步得到提升(如表2第4行所示)。这是因为几何一致性损失引导模型更加关注图像的几何结构,从而在结构一致性上实现了显著的改进,提升了生成图像的整体质量。

4 结论与展望

本文提出了一种融合双域特征的CBCT到CT的生成对抗网络。所提出的双域融合特征模块能够有效提取不同感受野下的特征,引导生成器聚焦于CBCT和CT的相似区域,使模型能够捕获更全面的信息。实验结果表明, DDFGAN可以更稳定地生成质量更高、与规划CT更接近的图像。此外,本文将CycleGAN的双向生成改为单向生成,大幅减少了计算资源的消耗。

然而,由于本文的输入是二维切片,无法充分利用医学图像数据中的三维空间信息。因此,在未来的工作中,将研究一种能够输入多张二维切片的新模型,以更好地利用空间信息,从而实现更加精准的图像生成。

参考文献

- Chan TCY, Mišić VV. Adaptive and robust radiation therapy optimization for lung cancer. *European Journal of*

- Operational Research, 2013, 231(3): 745–756. [doi: [10.1016/j.ejor.2013.06.003](https://doi.org/10.1016/j.ejor.2013.06.003)]
- 2 Kim S, Yoo S, Yin FF, *et al.* Kilovoltage cone-beam CT: Comparative dose and image quality evaluations in partial and full-angle scan protocols. *Medical Physics*, 2010, 37(7): 3648–3659. [doi: [10.1118/1.3438478](https://doi.org/10.1118/1.3438478)]
 - 3 Harms J, Lei Y, Wang TH, *et al.* Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography. *Medical Physics*, 2019, 46(9): 3998–4009. [doi: [10.1002/mp.13656](https://doi.org/10.1002/mp.13656)]
 - 4 Zhao W, Brunner S, Niu K, *et al.* Patient-specific scatter correction for flat-panel detector-based cone-beam CT imaging. *Physics in Medicine & Biology*, 2015, 60(3): 1339–1365. [doi: [10.1088/0031-9155/60/3/1339](https://doi.org/10.1088/0031-9155/60/3/1339)]
 - 5 Shi LX, Vedantham S, Karellas A, *et al.* Library based X-ray scatter correction for dedicated cone beam breast CT. *Medical Physics*, 2016, 43(8): 4529–4544. [doi: [10.1118/1.4955121](https://doi.org/10.1118/1.4955121)]
 - 6 Chen H, Zhang Y, Kalra MK, *et al.* Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Transactions on Medical Imaging*, 2017, 36(12): 2524–2535. [doi: [10.1109/TMI.2017.2715284](https://doi.org/10.1109/TMI.2017.2715284)]
 - 7 Hammernik K, Klatzer T, Kobler E, *et al.* Learning a variational network for reconstruction of accelerated MRI data. *Magnetic Resonance in Medicine*, 2018, 79(6): 3055–3071. [doi: [10.1002/mrm.26977](https://doi.org/10.1002/mrm.26977)]
 - 8 Zhang YB, Yu HY. Convolutional neural network based metal artifact reduction in X-ray computed tomography. *IEEE Transactions on Medical Imaging*, 2018, 37(6): 1370–1381. [doi: [10.1109/TMI.2018.2823083](https://doi.org/10.1109/TMI.2018.2823083)]
 - 9 Xie SP, Yang CY, Zhang ZJ, *et al.* Scatter artifacts removal using learning-based method for CBCT in IGRT system. *IEEE Access*, 2018, 6: 78031–78037. [doi: [10.1109/ACCESS.2018.2884704](https://doi.org/10.1109/ACCESS.2018.2884704)]
 - 10 Liang X, Chen LY, Nguyen D, *et al.* Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy. *Physics in Medicine & Biology*, 2019, 64(12): 125002. [doi: [10.1088/1361-6560/ab22f9](https://doi.org/10.1088/1361-6560/ab22f9)]
 - 11 Deng LW, Zhang MX, Wang J, *et al.* Improving cone-beam CT quality using a cycle-residual connection with a dilated convolution-consistent generative adversarial network. *Physics in Medicine & Biology*, 2022, 67(14): 145010. [doi: [10.1088/1361-6560/ac7b0a](https://doi.org/10.1088/1361-6560/ac7b0a)]
 - 12 Zhang YW, Li CP, Dai ZH, *et al.* Breath-hold CBCT-guided CBCT-to-CT synthesis via multimodal unsupervised representation disentanglement learning. *IEEE Transactions on Medical Imaging*, 2023, 42(8): 2313–2324. [doi: [10.1109/TMI.2023.3247759](https://doi.org/10.1109/TMI.2023.3247759)]
 - 13 Liu S, Huang D. Receptive field block net for accurate and fast object detection. *Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018. 404–419. [doi: [10.1007/978-3-030-01252-6_24](https://doi.org/10.1007/978-3-030-01252-6_24)]
 - 14 Li YX, Shao HC, Liang X, *et al.* Zero-shot medical image translation via frequency-guided diffusion models. *IEEE Transactions on Medical Imaging*, 2024, 43(3): 980–993. [doi: [10.1109/TMI.2023.3325703](https://doi.org/10.1109/TMI.2023.3325703)]
 - 15 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 770–778. [doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90)]
 - 16 Zhu JY, Park T, Isola P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks. *Proceedings of the 2017 IEEE International Conference on Computer Vision*. Venice: IEEE, 2017. 2242–2251. [doi: [10.1109/ICCV.2017.244](https://doi.org/10.1109/ICCV.2017.244)]
 - 17 Goodfellow IJ, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets. *Proceedings of the 28th International Conference on Neural Information Processing Systems*. Montreal: MIT Press, 2014. 2672–2680.
 - 18 Fu H, Gong MM, Wang CH, *et al.* Geometry-consistent generative adversarial networks for one-sided unsupervised domain mapping. *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019. 2422–2431. [doi: [10.1109/CVPR.2019.00253](https://doi.org/10.1109/CVPR.2019.00253)]
 - 19 Liu MY, Breuel T, Kautz J. Unsupervised image-to-image translation networks. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach: Curran Associates Inc., 2017. 700–708.
 - 20 Park T, Efros AA, Zhang R, *et al.* Contrastive learning for unpaired image-to-image translation. *Proceedings of the 16th European Conference on Computer Vision*. Glasgow: Springer, 2020. 319–345. [doi: [10.1007/978-3-030-58545-7_19](https://doi.org/10.1007/978-3-030-58545-7_19)]
 - 21 van den Oord A, Li YZ, Vinyals O. Representation learning with contrastive predictive coding. *arXiv:1807.03748*, 2018.
 - 22 Thummerer A, van der Bijl E, Galapon Jr A, *et al.* SynthRAD2023 grand challenge dataset: Generating synthetic CT for radiotherapy. *Medical Physics*, 2023, 50(7): 4664–4674. [doi: [10.1002/mp.16529](https://doi.org/10.1002/mp.16529)]
 - 23 Huijben EMC, Terpstra ML, Galapon Jr A, *et al.* Generating synthetic computed tomography for radiotherapy: SynthRAD2023 challenge report. *Medical Image Analysis*, 2024, 97: 103276. [doi: [10.1016/j.media.2024.103276](https://doi.org/10.1016/j.media.2024.103276)]

(校对责编: 张重毅)