先验 GAN 的 CBCT 牙齿图像超分辨率方法

宋全博, 李扬科, 范业莹, 陆书一, 周元峰*

(山东大学软件学院 济南 250101)
(yfzhou@sdu.edu.cn)

摘 要: 针对高分辨率的锥形束计算机断层扫描图像难以获取的问题, 提出一种基于先验生成对抗网络的锥形束计算机断层扫描图像超分辨率方法, 其通过微型断层扫描图像作为参照, 对超分辨率网络进行弱监督训练. 首先训练一个生成对抗网络, 用于生成高质量的单颗牙齿的微型计算机断层扫描图像, 并将其嵌入到一个 U 型的主干网络中作为先验解码器; 然后利用低分辨率的多颗牙齿的锥形束计算机断层扫描图像对主干网络进行训练, 先定位到锥形束计算机断层扫描影像中的每一颗牙, 再分别提高分辨率; 最后解决锥形束计算机断层扫描与微型计算机断层扫描之间的域差距, 通过设计基于小波变换噪声提取的域适应退化模块, 间接优化生成器生成更符合微型计算机断层扫描信息分布的图像. 在锥形束计算机断层扫描数据集上进行实验的结果表明, 与现有的超分辨率方法相比, 所提方法的峰值信噪比提高了 0.79~6.02 dB, 感知相似度评价指标降低了 0.01~0.72, 并且在牙齿部分获得了更好的视觉效果, 具有较强的竞争力.

关键词: 超分辨率; 锥形束计算机断层扫描图像; 先验生成对抗网络; 弱监督学习
中图法分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2023.19756

CBCT Tooth Images Super-Resolution Method Based on GAN Prior

Song Quanbo, Li Yangke, Fan Yeying, Lu Shuyi, and Zhou Yuanfeng*

(*School of Software, Shandong University, Jinan* 250101)

Abstract: Aiming at the problem that high-resolution cone-beam computed tomography images are difficult to obtain, a cone-beam computed tomography image super-resolution method based on prior generative adversarial network is proposed, which uses the micro computed tomography image as a reference to perform weakly supervised training on the super-resolution network. Firstly, a generative adversarial network is trained to generate high-quality micro computed tomography images of a single tooth, which is embedded into a U-shaped backbone network as a prior decoder. Then, low-resolution cone-beam computed tomography images of multiple teeth are used to train the backbone network, and the network is first located to each tooth in the cone-beam computed tomography image. Finally, the domain gap between cone-beam computed tomography and micro computed tomography is solved. By designing a domain adaptive degradation module based on wavelet transform noise extraction, the generator is indirectly optimized to generate images more in line with the information distribution of micro computed tomography. The experimental results on the cone-beam computed tomography dataset show that compared with the existing super-resolution methods, the proposed method improves the peak signal-to-noise ra-

收稿日期: 2022-06-08; 修回日期: 2022-09-13. **基金项目**: 国家重点研发计划战略性科技创新合作项目(2021YFE0203800); 国家自然科学基金联合基金浙江两化融合项目(U1909210); 国家自然科学基金(62172257). **宋全博**(1996—), 男, 硕士研究生, 主要研究方向为图像处理及医学影像处理; **李扬科**(1997—), 男, 硕士研究生, 主要研究方向为计算机视觉、人机交互; **范业莹**(1998—), 女, 硕士研究生, 主要研究方向为图形图像处理; **陆书一**(1994—), 男, 博士研究生, 主要研究方向为图像处理及医学影像处理; **周元峰**(1980—), 男, 博士, 教授, 博士生导师, CCF 高级会员, 论文通信作者, 主要研究方向为几何造型、计算几何、图像处理及医学影像处理.

tio by 0.79−6.02 dB, reduces the learned perceptual image patch similarity by 0.01−0.72, and has better visual effect in the tooth super-resolution, which has strong competitiveness.

Key words: super-resolution; cone-beam computed tomography image; prior generative adversarial network; weakly supervised learning

近年来, 锥形束计算机断层扫描(cone-beam computed tomography, CBCT)已经成为临床诊断中的重要辅助工具, 其具有扫描时间短、X 射线剂量小、拍摄成本低等优点. 然而, 由于 CBCT 使用的射线散射和截断投影, 其图像往往伴随着噪声和伪影且清晰度相对较低. 因此, 提高 CBCT 图像的质量具有重要的意义.

图像超分辨率技术可以在不改变成像设备的情况下, 根据低分辨率图像重建出理想的高分辨率图像. 由于该技术在成本、实时性以及便利性等方面具有显著的优势, 因此已经成为数字图像处理技术的重要研究内容. 传统的超分辨率方法通常采用退化模型和手工设计的先验知识来对图像进行放大, 使得其在实际使用过程中的表现不尽如人意. 近年来, 深度学习在各种计算机视觉任务中表现出了优异的性能, 已有许多基于深度神经网络的超分辨率恢复方法, 它们表现出比传统方法更好的性能. 目前, 已有的基于深度学习的超分辨率方法[1-4]获得了令人满意的效果. 虽然超分辨率方法在自然图像处理领域已经取得了较大的进展, 但是其在医学图像处理领域的应用依然存在着较大的局限性. 由于各种设备产生的医学图像包含大量的、复杂的噪声和伪影, 因此 CBCT 图像超分辨率仍然是一个具有挑战性的问题.

本文从低分辨率的 CBCT 图像中恢复具有逼真纹理的高分辨率的 CBCT 图像, 其难点是对医学图像进行超分辨率处理时, 难以获得足够高的高分辨率的图像对模型的训练过程进行监督; 而如果使用不同种类的医学图像, 如生成微型断层扫描(micro computed tomography, Micro-CT)进行监督, 则会面临数据位置不对应、数据存在域差距等问题; 与 Micro-CT 相比, 由于 CBCT 图像提取时使用了更少的剂量, 因此图像容易被噪声污染. 在使用深度学习方法提取 CBCT 图像特征的过程中, 容易被这些噪声干扰, 若直接使用同一模型同时对 CBCT 和 Micro-CT 进行超分辨率处理, 则模型可能误将噪声提取为特征, 进而使输出的图像中存在不真实的细节. 此外, 由于 Micro-CT 图像需要的剂量过高, 只能提取脱离人体的单颗牙齿图像. 使用 Micro-CT 进行监督时, 需要对 Micro-CT 中的单颗牙和 CBCT 中的每颗牙齿进行匹配. 因此, 在 CBCT 与 Micro-CT 这 2 种不同类型的图像相互转换时, 需要解决域差距问题, 即需要处理 CBCT 图像与 Micro-CT 图像之间不同的噪声和伪影, 并匹配二者在图像中的空间位置.

为了解决高分辨率数据不足的问题, 本文使用弱监督学习方法, 提出一种基于先验生成对抗网络(generative adversarial network, GAN)的 CBCT 牙齿图像超分辨率方法. 首先预先训练一个 GAN 生成 Micro-CT 图像, 并将其嵌入到主干网络中, 作为 CBCT 超分辨率之前的解码器, 主干网络将输入的退化图像映射到所需的隐向量空间, 以便先验 GAN 能够重现所需的高分辨率 CBCT 图像; 为了匹配训练过程中的 Micro-CT 数据和 CBCT 数据, 使用一个牙齿检测模块定位 CBCT 图像中的单颗牙齿, 再对每颗牙齿分别进行超分辨率优化; 由于使用的扫描设备不同, CBCT 图像与 Micro-CT 图像之间往往存在不同的噪声、伪影等信息, 为解决此域差距, 因此本文使用小波变换提取出 CBCT 的高频噪声信息, 在构造低分辨率-高分辨率图像对时加入提取的高频噪声, 促使预训练的生成器能够构造超分辨率图像时忽视 CBCT 噪声的分布信息, 生成更符合 Micro-CT 信息分布的图像.

本文在 CBCT 数据集上进行充分的实验, 以验证本文方法的有效性, 结果显示, 与现有的 CBCT 超分辨率方法相比, 本文方法构造超分辨率图像具有较强的竞争力.

1 相关工作

1.1 图像超分辨率

目前, 基于卷积神经网络(convolutional neural network, CNN)的方法已成为超分辨率域的主流. Dong 等[1]提出基于 CNN 获得超分辨率结果的超分辨率方法, 使用 3 层 CNN 学习低分辨率和高分辨率图像对之间的映射函数; Lim 等[5]提出多尺度深度超分辨率系统, 重构不同放大因子的高分辨率图像; Zhang 等[6]使用残差连接和通道注意力机制提升高频

信息的恢复效果; Xu 等[7]设计双 CNN, 利用原始图像模拟相机成像的射线信息. 为了得到更符合人眼认知的评价标准, Dai 等[8]提出一种感知损失, 即测量特征空间中 2 幅图像, 而不是像素空间的误差.

仅使用像素级约束超分辨率训练通常会导致输出的图像不符合人眼的感知, 同时产生过度的平滑[9-10]. 为了缓解这个问题, 文献[11-13]采用 GAN 近似自然图像流形, 得到更真实的结果. 然而, 由于生成器需要同时学习真实图像分布的特征和输入图像分布的特征, 因此在输出中仍然可以观察到非自然的伪影. 本文通过将预训练 GAN 模型嵌入主干网络, 生成更真实、符合认知的超分辨率 CBCT 图像.

1.2　CBCT 图像处理

CBCT 使用带有平板探测器的锥形 X 射线束成像. 与传统 CT 不同, CBCT 通过投射不同角度的锥束生成三维图像, 并利用投影几何重建三维图像. 由于 CBCT 扫描系统存储成本更小、剂量更低, 因此其广泛应用于数字化口腔医疗诊断中; 但与 CT 相比, CBCT 图像更容易受到散射和噪声的影响. 因此, 提高 CBCT 图像的质量对后续的医疗工作至关重要. 目前, 已有多种从 CBCT 二维投影图像中去除散射光子的方法. Wang 等[12]使用反散射网格的硬件抑制噪声; Xu 等[13]和 Stankovic 等[14]使用反卷积映射的估计提升 CBCT 图像的质量; Sisniega 等[15]利用 PlanCT 作为先验信息来提高 CBCT 图像的质量: 该类方法适用于直方图匹配[16]、超分辨率[17]和三维图像重建[18-20], 虽然它们不需要依赖高清的 CBCT 进行监督训练, 但依赖同一患者的 CBCT 和 PlanCT 体积对的精确空间对齐. 在改进后的 CBCT 图像中, 成对图像之间的不匹配会导致边缘模糊、变形以及部分解剖结构消失等错误. 本文提出牙齿检测模块对牙齿进行定位, 并且通过 GAN 的映射匹配牙齿的低分辨率-高分辨率特征, 对这类方法的匹配进行改进.

1.3　先验 GAN

先验 GAN 是用预先训练的 GAN 模型将给定的图像反转为隐向量. (1) 一种主流的思路是通过优化隐向量来优化生成器输出的结果. PULSE 方法[21]通过在输入和输出之间施加像素级约束来迭代地优化隐向量; mGANPrior 方法[22]通过优化多个隐向量来增加模型的表达性. DGP 方法[23]通过进一步精细化生成器和判别器, 减少训练图像和测试图像分布之间的差距. 然而, 由于隐向量是低维的, 可能无法真实地保留反演映射得到的空间信息, 因此这些方法往往会产生与真实图像不相似的结果. (2) 另一种常用的思路是预训练一个额外的编码器, 将图像空间投影回隐向量空间. Pan 等[24]将预先训练好的 GAN 嵌入到主网络中作为解码器, 使用人工退化的人脸图像对整个网络进行微调, 使网络从严重退化的野外人脸图像中生成高质量的人脸图像; Chan 等[25]将预先训练过的 GAN 作为编码器-“隐式银行”-解码器架构中的生成隐式库, 通过逐层的信息编码与解码重建逼真的高分辨率图像. 这些工作利用 GAN 作为存储先验, 从构造符合高清图像分布的隐向量空间并通过网络从中选取最优向量重建目标图像; 但由于这些工作构造的隐向量空间与输入图像不存在域差距, 因此不适用于优化不同领域的医学图像. 本文针对医学图像超分辨率的工作, 对这类方法进行域适应等方面的改进.

2　本文方法

2.1　整体框架

针对 CBCT 超分辨率问题, 本文提出一种基于先验 GAN 的 CBCT 牙齿图像超分辨率方法, 其整体框架如图 1 所示. 其中, 上行为网络的预训练部分, 下行为主干网络. 该方法包括域适应退化模块、编码器模块、先验生成器模块、牙齿检测模块和整体上采样模块. 域适应退化模块构造低分辨率-高分辨率图像对, 使模型在预训练过程中能够适应 CBCT 与 Micro-CT 之间的域差距; 编码器模块从图像中提取隐向量 $\boldsymbol{z}$; 先验生成器模块通过预训练, 能够捕获 Micro-CT 中丰富的纹理和形状先验知识, 将编码器提取的隐向量进行解码, 得到高分辨率图像; 牙齿检测模块定位 CBCT 图像中的牙齿位置, 使得后续的模块可以对每颗牙齿分别进行超分辨率处理; 整体上采样模块对整幅 CBCT 图像进行超分辨率处理, 而后将先验生成器输出的单颗牙齿图像块粘贴至 CBCT 图像上的对应位置, 最终得到高分辨率且单颗牙质量接近 Micro-CT 的 CBCT 图像.

2.2　域适应退化模块

在超分辨率工作中, 需要通过图像退化构造低分辨率-高分辨率图像对, 常见的退化方式使用降采样方法. 为了优化模型适应 CBCT 与 Micro-CT 的域差距, 本文在预训练过程中设计了域适应退化模块, 如图 2 所示.

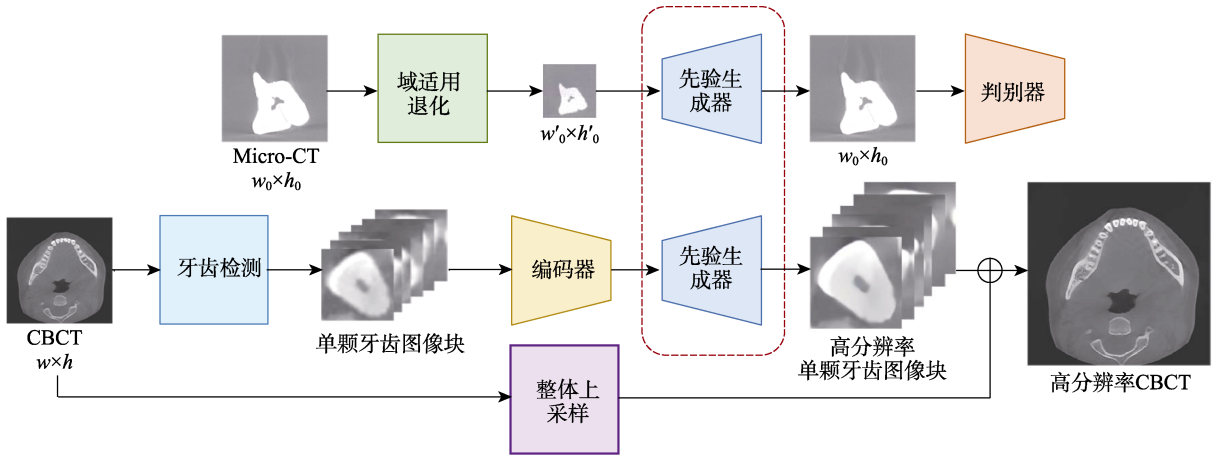


图 1 本文方法整体架构

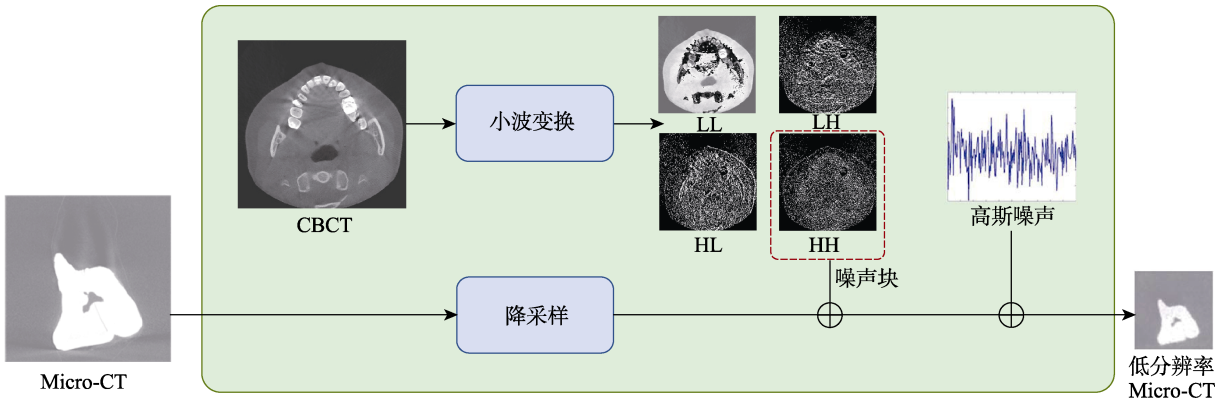


图 2 域适应退化模块

使用退化模型将高清的 Micro-CT 图像合成退化后的 Micro-CT 图像. 本文使用的退化模型表达式为

$$I_{\mathrm{L}}=\left(I_{\mathrm{H}}\otimes k_Q\right)\downarrow_r+n^* \tag{1}$$

其中，$\otimes$表示高清 Micro-CT 图像 I_{H} 与带有参数为 Q 的模糊核 k_Q 之间的滤波操作；$\downarrow_r$ 表示具有尺度因子 r 的线性降采样操作；n^* 表示域适应噪声.

根据文献[26]的先验知识，CBCT 图像中的噪声往往集中于高频部分. 为了更好地解决 CBCT 与 Micro-CT 图像的域差距，本文使用小波变换提取 CBCT 的高频噪声，并裁剪出牙齿位置的噪声块，输入预训练模型中噪声的构成为

$$n^*=H_{\mathrm{wavelet}}\left(I_{\mathrm{CBCT}}\right)+n_\delta \tag{2}$$

其中，H_{wavelet} 表示小波变换操作；n_δ 表示噪声水平为 δ 的加性高斯白噪声. 域适应退化模块输出了含有 CBCT 域噪声的下采样 Micro-CT 图像，该图像与原始的高清 Micro-CT 图像构成训练的图像对，促使生成器在训练过程中自动忽视与 CBCT 噪声有着相似分布的信息，专注生成 Micro-CT 图像分布有关的信息.

2.3 编码器模块

编码器模块包括 3 个卷积核为 3×3 的卷积层和 LeakyReLU 激活函数. 本文使用编码器生成牙齿图像 x_0 对应的隐向量

$$\boldsymbol{z}=E_\phi\left(x_0\right) \tag{3}$$

其中，E 表示编码器；ϕ 表示编码器的网络参数；$\boldsymbol{z}$ 表示隐向量空间内的隐向量. 编码器近似于生成器的一种反演，将图像映射至隐向量空间，得到反演的隐向量 $\boldsymbol{z}$，如图 3 所示. 下文中，生成器可通过隐空间中最匹配的隐向量恢复高分辨率的图像.

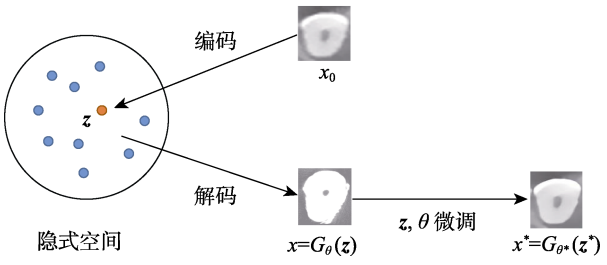


图 3 单颗牙齿图像的编码与解码

2.4　先验生成器模块

先验生成器模块包括4个卷积核为3×3的卷积层和 LeakyReLU 激活函数. 本文利用这种预先训练的 GAN 从 Micro-CT 图像中捕获丰富的纹理和形状先验信息.

如图 1 所示, 输入图像 Micro-CT 降采样得到的低分辨率图像, 先验 GAN 模块生成对应的高分辨率图像为

$$x = G_\theta(\boldsymbol{z}) \tag{4}$$

其中, G 表示生成器; θ 表示生成器的参数; $\boldsymbol{z}$ 表示能够用以生成高分辨率图像的隐向量, 即预训练中的低分辨率图像.

预训练完成后, 将预训练的先验 GAN 模块嵌入主干网络中作为解码器. 如图 3 所示, CBCT 中的单颗牙齿图像块的信息 x_0 通过编码器模块的编码映射至隐向量空间. 以训练好的生成器 G 作为解码器, 将 $\boldsymbol{z}$ 映射为具有高质量纹理和形状的高分辨率图像.

主干网络中, 本文使用 CBCT 图像对编码器模块和先验生成器模块进行微调, 公式为

$$\phi^*, \theta^* = \arg\min_{\phi,\theta} L\left(G_\theta\left(E_\phi\left(x_0\right)\right), \hat{x}\right) \tag{5}$$

其中, G_θ 表示先验生成器模块; θ 表示网络参数; $\hat{x}$ 表示目标高分辨率 CBCT 图像. 由于先验生成器学习到 Micro-CT 的高分辨率特征, 因此模型能够从低分辨率的图像生成具有丰富的图像细节的高分辨率的图像.

2.5　牙齿检测模块

牙齿检测模块在 CBCT 图像中定位每颗牙齿的位置, 使得后续的模块可以对每颗牙齿分别进行超分辨率处理, 达到着重提升图像中牙齿部分质量的目的. 该模块采用预训练的 MobileNet0.25 网络[27]实现牙齿的检测, 对于每幅输入的 CBCT 图像, 通过模型的深度可分离卷积提取特征, 输出 n 个牙齿的边界框, 其中, n 表示当前 CBCT 图像中的牙齿总数, 由网络自适应生成, 每个边界框 $t_i = \left\{t_x, t_y, t_w, t_h\right\}$ 表示每颗牙齿的横、纵坐标, 以及宽和高.

在牙齿检测模块输出的结果中, 使用非极大抑制算法检测预测结果的边界框中与其他框重合的部分, 去掉其中重合度过高的边界框, 最终保留的就是 CBCT 图像中的每颗牙齿的位置. 模块输出单颗牙齿图像块, 以使用后续的模块对单颗牙齿进行超分辨率处理.

2.6　整体上采样模块

整体上采样模块基于 RRDB 网络[12], 包括 4 层 3×3 的卷积层和 LeakyReLU 激活函数. 通过结合多层的残差网络和密集连接, 提取图像中从浅层到深层的丰富特征, 获得更好的超分辨率效果, 并且训练更加稳定.

CBCT 图像中, 除了已经进行超分辨率的牙齿部分, 还包含牙床骨、软组织和背景信息. 本文将牙齿信息与其他信息分别进行超分辨率处理. 其中, 单颗牙齿图像块经过编码器模块及先验 GAN 模块进行超分辨率提升, 其他部分经过整体上采样模块进行整体的2倍超分辨率放大; 然后将先验生成器生成的高分辨率牙齿粘贴至对应的位置.

2.7　损失函数

为了兼顾图像的真实性和感知效果, 本文采用 3 种不同的损失函数: 对抗性损失 L_{adv}, 均方误差(mean square error, MSE)损失 L_{MSE} 和感知损失[8] L_{per}.

L_{adv} 基于 StyleGAN, 计算公式为

$$L_{\text{adv}} = \min_G \max_D E_{(X)} \log\left(1+\exp\left(-D\left(G\left(X_{\text{L}}\right)\right)\right)\right) \tag{6}$$

其中, X 表示高清图像; X_{L} 表示退化后的低分辨率图像; $E_{(X)}$ 表示数学期望; G 表示生成器; D 表示判别器; 限制生成的样本在生成器的得分 $G\left(X_{\text{L}}\right)$ 越高越好、在判别器的得分 $D\left(G\left(X_{\text{L}}\right)\right)$ 越低越好, 以使网络生成贴近真实的样本.

L_{MSE} 计算生成样本与真实图像之间的像素级别的损失, 计算公式为

$$L_{\text{MSE}} = \frac{1}{N}\left\|\hat{y}-y\right\|_2^2 \tag{7}$$

其中, $\hat{y}$ 表示生成器生成的高分辨率图像中的像素; y 表示高分辨率真实图像中的像素; N 表示输入图像的像素总数. 使用 L_{MSE} 能够强化图像的像素级特征, 保留原始的灰度信息.

L_{per} 通过提取 VGG16 网络[28]提供的高层的特征, 监督生成图像在人眼感知方面的损失, 计算公式为

$$L_{\text{per}} = \frac{1}{N}\left\|f\left(\hat{y}\right)-\left(y\right)\right\|_2^2 \tag{8}$$

其中, f 表示 VGG16 提取的特征空间. 使用 L_{per} 能够利用卷积层抽象高层特征的能力, 从高维度强化模型生成更接近人的思维的层次感知图像.

最终的损失函数为

$$L = L_{\text{adv}} + \alpha L_{\text{MSE}} + \beta L_{\text{per}} \tag{9}$$

其中，α 和 β 表示为了平衡各损失函数设置的超参数；L_{MSE} 加强精细特征并保留原始灰度信息；L_{per} 和 L_{adv} 共同约束网络生成贴近真实的图像. 为了获得令人满意的结果，在实验中，本文将超参数设置为 $\alpha=1$ 和 $\beta=0.2$.

3 实验及结果分析

3.1 数据集与评价标准

本文与首尔大学医学院合作，收集了一系列 Micro-CT 数据集. 其中，训练数据集包含 18 套分辨率为 512×512×512 的 Micro-CT 扫描，精度为 0.1 mm；测试数据集包括 6 套分辨率为 512×512×512 的 CBCT 扫描，精度为 0.3 mm. 将 CBCT 图像和 Micro-CT 的强度分别归一化到[0,1]. 由于医疗信息属于敏感信息，因此本文的数据集不进行公开使用；所有患者资料已完全匿名，且均按照相关机构制定的指南和规定执行.

本文使用峰值信噪比(peak signal-to-noise ratio, PSNR)和感知相似度(learned perceptual image patch similarity, LPIPS)[29]作为评估指标. 其中，PSNR 计算图像最大值信号和背景噪声像素级别的均方误差，其值越大越好；LPIPS 提取抽象高层特征的能力，更符合人的思维的层次感知的评价标准，其值越小越好.

3.2 对比实验

为了验证本文方法的性能，选取传统的双三次插值(Bicubic)方法，基于深度神经网络(deep neural network, DNN)的 EDSR 方法[8]，以及基于 GAN 的 mGAN[22]，DGP[23]和 GLEAN[25]这 5 种方法，在数据集上进行对比实验. 其中，EDSR 使用 CBCT 数据集训练；PULSE，DGP 和本文方法先使用 Micro-CT 数据集预训练，再使用 CBCT 数据微调其主干网络. 6 种方法的 PSNR 和 LPIPS 比较的结果如表 1 所示. 可以看出：(1) 仅使用像素级别损失的约束下，EDSR 方法的输出感知效果不足，且在 2 倍以上的超分辨率中性能下降；在低维向量和低分辨率空间约束的引导下，GAN 反演方法的输出无法保持良好的保真度，特别是 mGAN 方法无法恢复具有足够真实性的图像；通过在优化过程中对发生器进行微调，DGP 和 GLEAN 方法在质量和保真度方面都有显著提高：上述方法输出的图像与真实高清 CBCT 图像之间仍然有细微差别. (2) 本文方法的 PSNR 略优于其他方法，并且在 LPIPS 上取得了更优的效果. 本文方法通过预训练的先验 GAN 方法，能够提取到 Micro-CT 中丰富的细节和形状特征，因此可以获得更好的感知效果；由于本文在预训练过程中进行了域适应优化，因此比其他方法更适用于 CBCT 数据集.

表 1 6 种方法的 PSNR 和 LPIPS 比较

方法	PSNR/dB↑		LPIPS↓	
	2×	4×	2×	4×
Bicubic	32.86	30.27	0.176 4	0.173 7
EDSR[8]	**37.60**	30.63	0.152 5	0.161 8
mGANPrior[22]	25.82	24.84	0.152 7	0.151 4
DGP[23]	34.72	34.22	0.122 6	0.120 6
GLEAN[25]	36.83	35.50	0.121 7	0.110 3
本文	36.98	**36.29**	**0.097 2**	**0.101 4**

注. 粗体表示最优值.

图 4 所示为在 4 倍超分辨率下，本文方法与其他方法的可视化结果，其中，最具代表性的牙齿分别用红色、蓝色、黄色和绿色标出. 可以看出，本文方法在清晰度、对比度和保真度上都具有不错的效果.

3.3 消融实验

为了验证牙齿检测模块和域适应退化模块的有效性，通过修改后模型的超分辨率性能进行消融实验. (1) 仅删除牙齿检测模块，训练与测试时直接将整个的 CBCT 图像输入到主干网络的编码器中；(2) 删除小波变换提取的噪声信息块，只使用高斯噪声进行图像退化；(3) 去除预训练的过程，直接使用 CBCT 数据对主干网络进行训练. 本文在数据集上，将进行 3 种修改后的模型与本文原模型的对比，结果如表 2 所示. 可以看出，预训练 GAN 的模块可以拟合出特征更丰富、更具有细节和真实性的牙齿图像分布，对后续的训练有较好的提升作用；如果没有提取 CBCT 域的特有噪声至训练模型，则不能有效地在 CBCT 提升的过程中解决域差异所带来的伪影等问题；牙齿检测模块保证了模型能够有针对性地提升图像中牙齿的细节特征，而不是将牙齿的特征误学习到其他部分. 总体来说，本文方法表现出优于其变体的性能，证明了预训练 GAN 和本文对 CBCT 超分辨率训练策略的有效性.

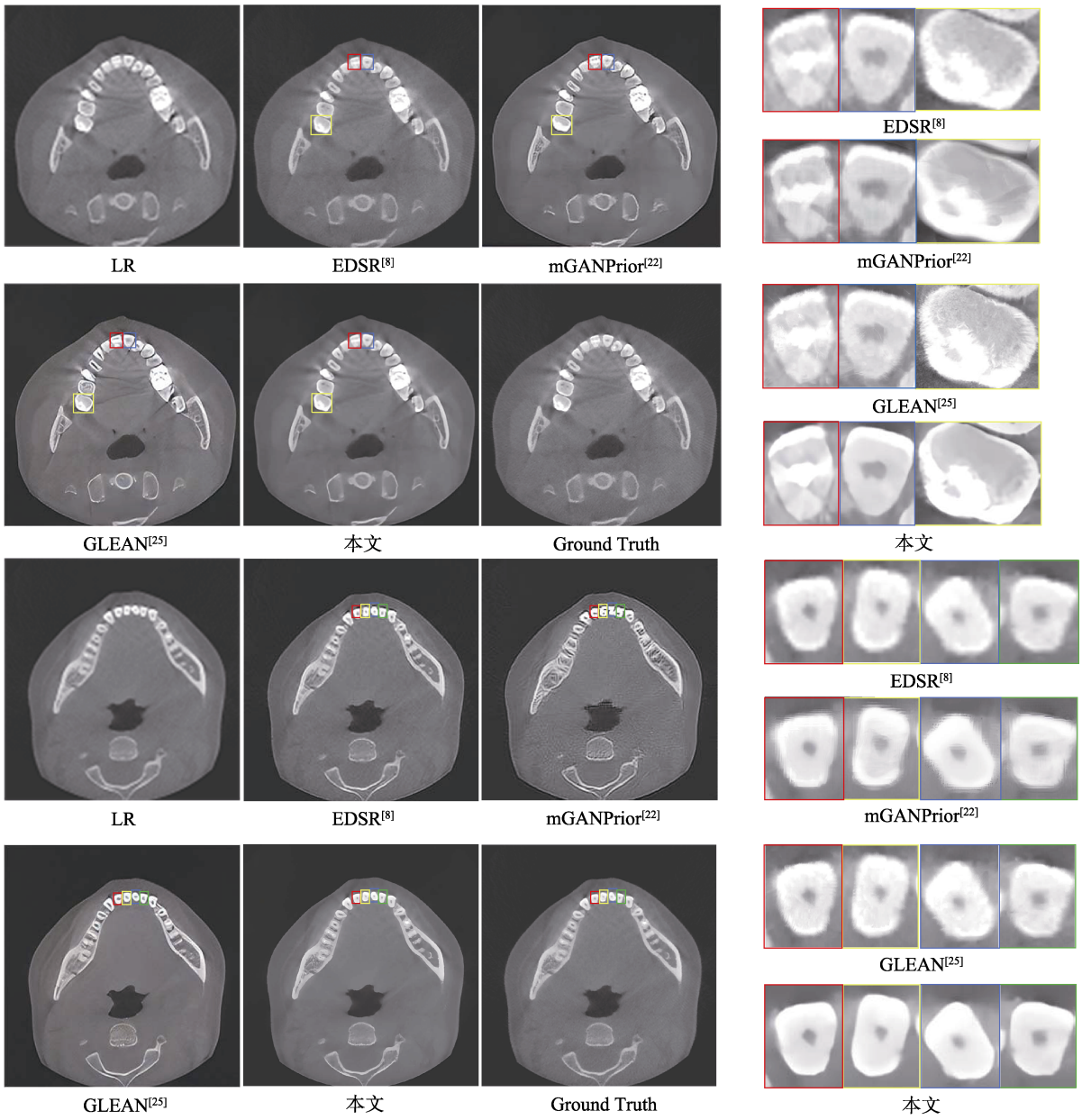


图 4　4 倍超分辨率下 4 种方法可视化结果比较

表 2　模型消融实验结果

预训练	域适应退化模块	牙齿检测模块	PSNR/dB
×	√	√	34.63
√	×	√	36.03
√	√	×	36.72
√	√	√	**36.98**

注. 粗体表示最优值.

4　结　语

本文提出一种基于先验 GAN 的 CBCT 超分辨率方法，可以将 CBCT 图像的分辨率提升到接近 Micro-CT 的水平. 该方法通过预训练的 GAN 学习到符合 Micro-CT 图像分布的隐向量空间，再将预训练的 GAN 模型嵌入到 U 型网络中，生成具有接近 Micro-CT 丰富特征的 CBCT 图像. 此外，借助牙齿检测模块定位 CBCT 图像中的每颗牙齿，通过先验生成器隐向量空间的映射，提高其分辨率至接近 Micro-CT 的水平. 本文使用小波变换提取 CBCT 图像的噪声特征，进一步优化模型的训练过程，解决了 CBCT 与 Micro-CT 之间存在的域差距对模型的影响. 实验结果表明，本文方法在 CBCT

数据集上取得了非常有竞争力的表现性能, 具体的可视化结果也进一步说明了该方法的有效性.

参考文献(References):

[1] Dong C, Loy C C, He K M, *et al*. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38(2): 295-307

[2] He K M, Zhang X Y, Ren S Q, *et al*. Deep residual learning for image recognition[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2016: 770-778

[3] Huang G, Liu Z, Van Der Maaten L, *et al*. Densely connected convolutional networks[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2016

[4] He X Y, Mo Z T, Wang P S, *et al*. ODE-inspired network design for single image super-resolution[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 1732-1741

[5] Lim B, Son S, Kim H, *et al*. Enhanced deep residual networks for single image super-resolution[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Los Alamitos: IEEE Computer Society Press, 2017: 1132-1140

[6] Zhang Y L, Li K P, Li K, *et al*. Image super-resolution using very deep residual channel attention networks[C] //Proceedings of the 15th European Conference on Computer Vision. Heidelberg: Springer, 2018: 286-301

[7] Xu X Y, Ma Y R, Sun W X. Towards real scene super-resolution with raw images[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 1723-1731

[8] Dai T, Cai J R, Zhang Y B, *et al*. Second-order attention network for single image super-resolution[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 11057-11066

[9] Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution[C] //Proceedings of the 14th European Conference on Computer Vision. Heidelberg: Springer, 2016: 694-711

[10] Ledig C, Theis L, Huszár F, *et al*. Photo-realistic single image super-resolution using a generative adversarial network[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2017: 105-114

[11] Sajjadi M S M, Schölkopf B, Hirsch M. EnhanceNet: single image super-resolution through automated texture synthesis[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 4501-4510

[12] Wang X T, Yu K, Wu S X, *et al*. ESRGAN: enhanced super-resolution generative adversarial networks[C] //Proceedings of the European Conference on Computer Vision. Heidelberg: Springer, 2018: 63-79

[13] Xu X Y, Sun D Q, Pan J S, *et al*. Learning to super-resolve blurry face and text images[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 251-260

[14] Stankovic U, Ploeger L S, Van Herk M, *et al*. Optimal combination of anti-scatter grids and software correction for CBCT imaging[J]. Medical Physics, 2017, 44(9): 4437-4451

[15] Sisniega A, Zbijewski W, Badal A, *et al*. Monte Carlo study of the effects of system geometry and antiscatter grids on cone-beam CT scatter distributions[J]. Medical Physics, 2013, 40(5): Article No.051915

[16] Sun M, Star-Lack J M. Improved scatter correction using adaptive scatter kernel superposition[J]. Physics in Medicine and Biology, 2010, 55(22): 6695-6720

[17] Shi L X, Tsui T, Wei J K, *et al*. Fast shading correction for cone beam CT in radiation therapy via sparse sampling on planning CT[J]. Medical Physics, 2017, 44(5): 1796-1808

[18] Kidar H S, Azizi H. Assessing the impact of choosing different deformable registration algorithms on cone-beam CT enhancement by histogram matching[J]. Radiation Oncology, 2018, 13(1): Article No.217

[19] Oyama A, Kumagai S, Arai N, *et al*. Image quality improvement in cone-beam CT using the super-resolution technique[J]. Journal of Radiation Research, 2018, 59(4): 501-510

[20] Kida S, Nakamoto T, Nakano M, *et al*. Cone beam computed tomography image quality improvement using a deep convolutional neural network[J]. Cureus, 2018, 10(4): Article No.e2548

[21] Menon S, Damian A, Hu S J, *et al*. PULSE: self-supervised photo upsampling via latent space exploration of generative models[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 2434-2442

[22] Karras T, Laine S, Aila T. A style-based generator architecture for generative adversarial networks[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 4396-4405

[23] Gu J J, Shen Y J, Zhou B L. Image processing using multi-code GAN prior[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 3009-3018

[24] Pan X G, Zhan X H, Dai B, *et al*. Exploiting deep generative prior for versatile image restoration and manipulation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(11): 7474-7489

[25] Chan K C K, Wang X T, Xu X Y, *et al*. GLEAN: generative latent bank for large-factor image super-resolution[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 14240-14249

[26] Zhong J, Ning R, Conover D. Image denoising based on multiscale singularity detection for cone beam CT breast imaging[J]. IEEE Transactions on Medical Imaging, 2004, 23(6): 696-703

[27] Howard A G, Zhu M L, Chen B, *et al*. MobileNetV2: inverted Residuals and Linear Bottlenecks[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 4510-4520

[28] Yang T, Ren P R, Xie X S, *et al*. GAN prior embedded network for blind face restoration in the wild[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 672-681

[29] Zhang R, Isola P, Efros A A, *et al*. The unreasonable effectiveness of deep features as a perceptual metric[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 586-595