

基于反投影回归网络的医学图像超分辨率重建

郑秋富, 金燕*, 姜智伟, 胡奕标

(浙江工业大学信息工程学院 杭州 310023)
(jy@zjut.edu.cn)

摘要: 针对使用深度学习进行医学图像超分辨率重建时存在的欠定性问题, 提出了一种基于反投影回归网络的医学图像超分辨率重建算法. 该网络除了学习低分辨率到高分辨率图像的原始映射之外, 还学习一个对偶回归映射来预测退化核并重建出低分辨率图像, 形成一个闭环以提供额外的约束条件, 通过该约束减小低分辨率到高分辨率图像的映射空间, 从而缓解图像重建时的欠定性问题; 并在上采样和下采样过程中引入了反投影机制, 通过误差反馈来减少上、下采样时丢失的特征信息. 在 PyTorch 环境中, 与 EDSR, DBPN, RCAN 等算法在 Mayo 和 TCGA-KICH 数据集进行了对比实验. 结果表明, 所提算法的评价指标优于对比算法, 如 Mayo 数据集上 2 倍放大因子下, 所提算法的峰值信噪比高于对比算法 0.12~1.41 dB.

关键词: 医学图像重建; 对偶回归; 误差反馈; U-Net; 卷积神经网络

中图分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2024.20113

Medical Image Super-Resolution Reconstruction Based on Back Projection Regression Network

Zheng Qiufu, Jin Yan*, Jiang Zhiwei, and Hu Yibiao

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023)

Abstract: To address the ill-posed problem when using deep learning for medical image super-resolution reconstruction, a medical image super-resolution reconstruction algorithm based on back projection regression network is proposed. In addition to learning the primal mapping from low resolution images to high resolution images, the back projection regression network also learns a dual regression mapping to estimate the degradation kernel and reconstruct low resolution images, forming a closed loop to provide an extra constraint. The constraint reduces the mapping space from low resolution images to high resolution images so as to alleviate the ill-posed problem of super-resolution reconstruction. Besides, the back projection mechanism is introduced in the process of upsampling and downsampling to reduce the lost feature information by error feedback. In the PyTorch environment, the comparison experiments with EDSR, DBPN, RCAN and other algorithms were carried out on Mayo and TCGA-KICH data sets. The results show that the evaluation index of the proposed algorithm is better than the comparison algorithm. For instance, the peak signal to noise ratio of the proposed algorithm is 0.12–1.41 dB higher than the comparison algorithm at 2 times magnification factor on Mayo data set.

Key words: medical image reconstruction; dual regression; error feedback; U-Net; convolutional neural network

收稿日期: 2022-12-29; 修回日期: 2023-07-06. 基金项目: 国家自然科学基金(U22A2040). 郑秋富(1997—), 男, 硕士研究生, 主要研究方向为图像处理; 金燕(1964—), 女, 硕士, 副教授, 硕士生导师, 论文通信作者, 主要研究方向为图像处理、机器视觉和自动化测量; 姜智伟(1994—), 男, 硕士研究生, 主要研究方向为图像处理; 胡奕标(1995—), 男, 硕士研究生, 主要研究方向为图像处理.

在现代医疗诊断体系中, 医学图像能够提供患者体内的病理结构信息, 辅助医生对患者的病情进行分析和诊断. 医学图像因其包含的信息具有重大的诊断意义, 在临床上愈加不可或缺. 然而, 在医学图像的成像过程中, 受到图像采集设备、医疗成像技术、患者所能承受的放射剂量等各种因素的限制, 获得的医学图像的分辨率往往不高^[1]. 低清晰度的图像不利于医生对疾病的诊断和治疗.

图像超分辨率(super resolution, SR)重建旨在从低分辨率(low resolution, LR)图像恢复出高分辨率(high resolution, HR)图像^[2]. 采用 SR 重建技术对 LR 医学图像进行重建, 获取的高清图像有助于医生更清楚地看到生物组织结构及其早期病变. 该方法在不需改进硬件设备或增大放射剂量的情况下, 就能显著地提高医学图像的分辨率. 因此, 近年来医学图像 SR 重建技术成为了研究热点.

目前 SR 重建方法可以分为基于插值、基于重构和基于学习 3 大类. 基于插值的方法是根据周围邻域已知像素点的像素值, 运用插值公式计算出中心像素点的值. 经典的插值方法有最近邻域插值^[3]、双线性插值^[4]和双三次线性插值^[5]等. 这类算法的计算复杂度低, 但重建出的图像存在边缘模糊等问题. 基于重构的方法以 LR 图像为约束前提, 结合能量有限性和非负性等先验知识, 重建出 HR 图像. 典型的算法包括: 迭代反向投影法^[6]、凸集投影法^[7]和最大后验概率法^[8]等. 这类方法改善了插值法中的边缘模糊问题, 但其计算量大, 且当先验知识较少时重建效果较差. 基于学习的方法是近几年来学者研究的热点, 其基本思想是通过神经网络^[9]学习出从 LR 图像到 HR 图像的非线性映射函数; 对于待重建的 LR 图像, 用训练得到的映射函数去指导重建, 从而得到 HR 图像. Dong 等^[10]首次将深度学习应用到 SR 重建中, 提出了基于卷积神经网络(convolutional neural network, CNN)的图像 SR 算法, 该算法构建了一个 3 层的神经网络, 端到端地学习从 LR 图像到 HR 图像的映射关系, 用于图像重建, 得到了优于传统算法的重建效果. 此后, 许多基于 CNN 的方法被提出, 用于提升 SR 重建性能. He 等^[11]提出了深度残差网络, 通过恒等连接来克服网络退化问题. Haris 等^[12]提出深度反向投影网络(deep back-projection networks, DBPN), 通过迭代具有误差反馈的上、下采样层, 重建出的 HR 图像具有丰富的纹理细节. Zhang 等^[13]提出了深度残差通道注意力网络(residual channel attention networks, RCAN)通过引入通道注意力机制, 网络

侧重学习高频信息, 重建出的图像更清晰.

虽然上述基于学习的方法在图像 SR 重建任务中表现出良好的性能, 但仍然存在不足之处. 首先, 学习从 LR 图像到 HR 图像的映射是一个典型的欠定性问题^[14]. 因为一幅 LR 图像对应多幅不同的 HR 图像, 即存在无限多幅 HR 图像能被下采样到同一幅 LR 图像, 这导致 LR→HR 可能的映射函数空间非常大; 而让神经网络在如此大的函数空间中学习到一个好的解是非常困难的, 从而限制了神经网络的 SR 重建性能. 许多基于学习的 SR 重建方法, 如 DBPN^[12]和 RCAN^[13], 虽然通过增加网络模型的容量来提高 SR 重建性能, 但仍然受到映射函数空间过大的影响. 此外, 网络模型通过下采样和上采样进行特征图的压缩和恢复时会丢失特征信息, 从而增加重建误差^[15].

为了解决上述 2 个问题, 本文提出了一种基于反投影回归网络(back projection regression networks, BPRN)的医学图像 SR 重建算法. 该算法除了学习 LR→HR 的原始映射之外, 还学习一个 HR→LR 的对偶回归映射, 重建出 LR 图像. 如果 LR→HR 的映射是最优的, 则重建出的 HR 图像可以被下采样到与输入相同的 LR 图像. 通过在 LR 数据上引入这个额外约束来减小可能函数的空间, 从而提高网络的 SR 重建性能. 此外, 在网络的上采样和下采样结构中引入反投影机制, 通过误差反馈来减少重建误差. 实验结果表明, 相比于其他算法, 本文算法的 SR 重建性能更好, 且参数量更少.

1 相关工作

1.1 医学图像 SR 重建

近年来, 随着深度学习研究的不断深入, 利用 CNN 进行医学图像 SR 重建的算法受到研究者追捧. Li 等^[16]针对部分医学图像中存在噪声污染的问题, 提出了一种基于抗噪声深度 CNN 的 SR 算法, 有效地抑制了图像重建过程中的噪声干扰. Hong 等^[17]将 CNN 应用于正电子发射断层扫描图像的重建, 通过将外部图像信息作为模型训练的先验知识, 减小了 SR 图像估计过程的难度. 网络模型训练时需要成对的 LR 和 HR 图像, 而医学图像中配对的数据集较少. 针对该问题, Song 等^[18]提出了用 CNN 提取的高维特征和空间信息等作为输入, 使用模拟的配对数据集来训练模型. 为了保证特征信息在较深的网络中也能得到有效传递,

Xia 等^[19]用多个跳跃式连接单元来传递从 LR 磁共振图像中提取的特征信息. 现有的许多模型对不同通道上的层次特征平等对待, 这阻碍了模型对不同层次特征的甄别和针对性处理, 从而降低了模型的表达能力. 对此, Zhao 等^[20]提出了一种用于单张磁共振图像 SR 重建的通道分离网络, 该网络有残差分支和密集分支 2 个分支, 前者能够促进特征重用, 后者有利于探寻新特征.

1.2 对偶学习

对偶学习方法最早在机器翻译领域中被提出, 它包含原始模型和对偶模型, 同时学习 2 个相反的映射, 两者相互提供反馈, 从而提高语言翻译的性能. 例如, 先用原始模型学习中文翻译成英文的映射, 再用对偶模型学习英文翻译成中文的映射, 形成的闭环过程称为对偶学习. 最近, 该学习策略也被用于无配对训练数据的图像翻译中. Yi 等^[21]提出用原始的生成对抗网络 (generative adversarial network, GAN) 学习将图像从 U 域翻译到 V 域, 再用 1 个对偶的 GAN 学习相反的任务. 此外, Zhu 等^[22]

在 GAN 的训练过程中引入循环一致性损失, 以避免 GAN 方法的模式崩溃问题. 受对偶学习的启发, 本文用附加的回归网络来减小 SR 重建的映射空间.

2 本文算法

2.1 整体网络模型

U-Net^[23]自 2015 年起被广泛应用于医学图像领域中, 尤其在医学图像重建和分割中表现出了优越的性能^[24]. U-Net 是一种编解码器结构^[25], 它既能利用上下文信息也能利用全局位置信息. 此外, U-Net 通过镜像操作来保留更多的边缘信息, 提高网络对输入信息的利用效率, 因而它在小规模的数据集上也能训练得非常好, 这恰好解决了医学图像数据集稀缺的问题. 因此, 本文提出的 BPRN 以 U-Net 为基础, 并进行了一定的改进. 本文所提网络的总体结构如图 1 所示, 它包括原始网络和回归网络 2 部分, 前者学习 LR→HR 的原始映射, 后者学习 HR→LR 的回归映射.

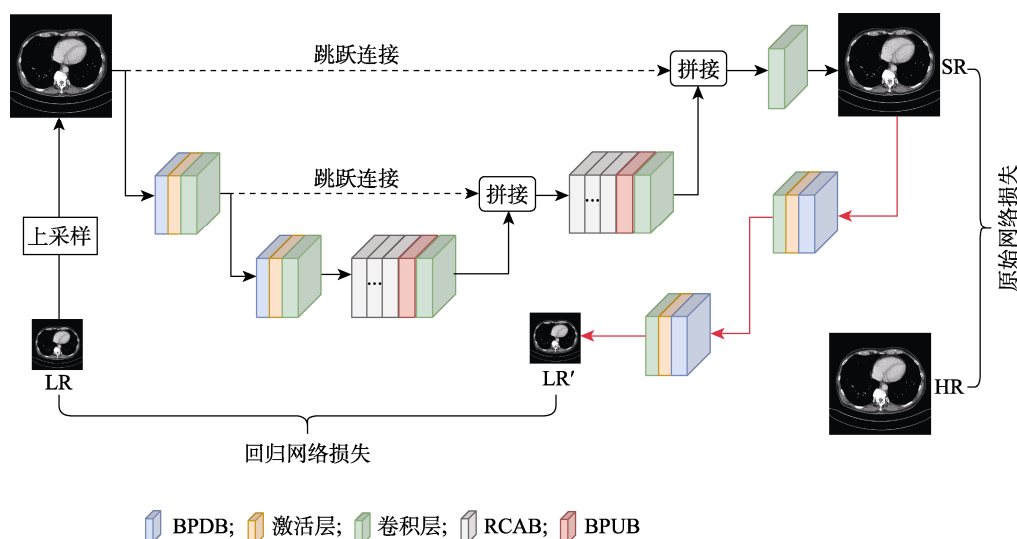


图 1 BPRN 的结构示意图

原始网络用于学习 LR→HR 的原始映射, 重建出 HR 图像. 为了减少 U-Net 通过下采样和上采样进行特征图的压缩和恢复时所丢失的特征信息, 将 U-Net 中用于下采样的最大池化层和用于上采样的反卷积层, 分别替换为具有误差纠正的反投影下采样模块 (back projection downsample block, BPDB) 和反投影上采样模块 (back projection upsample block, BPUB). 图像 SR 重建可视为一个尽可能多地恢复高频信息的过程, 高频信息恢复越多重建出的图像越清晰^[13]. 因此, 为了使网络在

重建过程中恢复更多的高频信息, 本文在 U-Net 中的上采样层前加入 N 个残差通道注意力模块 (residual channel attention block, RCAB)^[13]. RCAB 的具体结构如图 2 所示, 它将通道注意力集成到残差结构中. RCAB 的引入促使网络专注于学习高频信息, 丰富重建图像的纹理细节信息, 从而提升网络的重建效果. 回归网络用于学习 HR→LR 的回归映射, 重建出 LR 图像. 由于回归网络仅学习下采样映射, 这比原始网络学习上采样映射要简单得多, 因此回归网络仅由 BPDB, LeakyReLU 激活

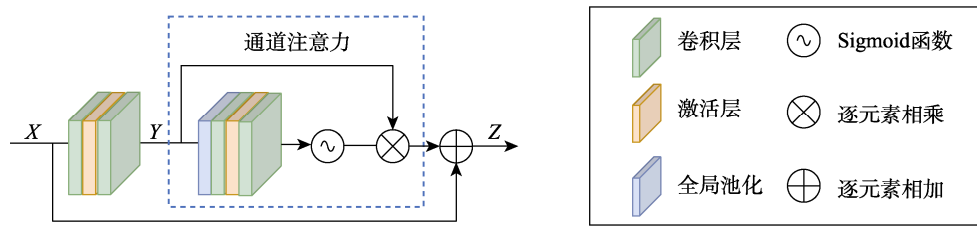


图 2 RCAB 的结构示意图

层和卷积层组成, 这样设计既能保证回归网络的性能, 又可减少不必要的计算成本。

2.2 对偶回归原理

LR 图像通常被认为是现实世界中 HR 图像的某种退化, 图像退化的数学模型可表示为

$$Y = DBX + N_{\zeta} \quad (1)$$

其中, X 表示真实的 HR 图像; Y 表示退化后的 LR 图像; D 表示下采样操作; B 表示模糊操作; N_{ζ} 表示可加的带标准差 ζ 的噪声。由图像退化的数学模型可知, 图像 SR 重建过程实质上是图像退化的逆过程, 而这是一个欠定性问题。因为重建 HR 图像实际上是求式(1)的解, 而由于退化矩阵 DB 的行数小于列数, 即方程个数小于未知数的个数, 这导致同一幅 LR 图像有多幅不同的 HR 图像与之对应, 但真实的 HR 图像是唯一的。

同一幅 LR 图像对应多幅不同的 HR 图像, 使得 LR→HR 可能的映射空间非常大^[26]。虽然, 现有的基于深度学习的方法通过增加网络模型的容量来更好地学习出 LR→HR 的映射, 但这些方法仍然受到映射函数空间过大的困扰, 导致网络训练过程非常困难。为了解决这个问题, 本文提出了对偶回归方案, 通过在 LR 数据上引入一个额外的约束来减小 LR→HR 的映射空间。如图 1 所示, BPRN 包括原始网络和回归网络。除了用原始网络学习 LR→HR 的映射, 还用回归网络学习逆映射 SR→LR。

设 $x \in X$ 为 LR 图像, $y \in Y$ 为 HR 图像。原始网络学习得到一个映射 $P: X \rightarrow Y$, 使预测值 $P(x)$ 近似于 x 所对应的 HR 图像 y 。回归网络学习得到一个映射 $R: Y \rightarrow X$, 使预测值 $R(y)$ 接近输入的 LR 图像 x 。理论上, 若 $P(x)$ 是准确无误的 HR 图像, 则其下采样后的图像 $R(P(x))$ 应非常近似于输入的 LR 图像 x 。本文利用这种约束减小 LR→HR 的映射空间, 从而使原始网络学习出更优的映射, 重建出效果更佳的 HR 图像。本文提出的网络模型联合地学习 2 个映射。假定数据集共有 N 对配对样本, x_i 和 y_i

分别是第 i 对的 LR 图像和 HR 图像。模型训练的损失函数可以表示为

$$L_{PR} = \sum_{i=1}^N L_P(P(x_i), y_i) + \lambda L_R(R(P(x_i)), x_i).$$

其中, $L_P(\cdot)$ 和 $L_R(\cdot)$ 分别是原始网络损失和回归网络损失, 且都基于 L_1 范数; λ 是回归网络损失函数的权重系数。

由于图像下采样过程不可避免地出现信息丢失问题, 导致在计算回归网络损失时难免存在偏差, 从而影响模型的重建效果。因此为了减小直接使用下采样从 SR 图像中获取 LR 图像时的信息损失, 本文提出使用具有误差反馈机制的回归网络来学习 SR→LR 的逆映射。回归网络中的 BPDB 能够在压缩图像的同时尽可能地减少信息的丢失, 从而缓解计算回归网络损失时存在的偏差问题。

2.3 反投影采样模块

在 BPRN 的原始网络中, 分别通过下采样和上采样逐步进行特征图的压缩和恢复, 这个过程会丢失特征信息并增加重建误差。此外, 在 BPRN 中, 渐进式下采样过程所丢失的信息会导致回归损失出现偏差, 不利于模型的训练。针对下采样和上采样过程中重建误差过大的问题, Haris 等^[12]运用反向投影的思想, 提出了深度反向投影网络, 该网络通过误差的反馈修正, 有效减少了重建误差, 获得了更好的结果。

反向投影思想基于这样一个假设: 高质量的 SR 图像经过下采样后得到的估计 LR 图像应尽可能地接近原始 LR 图像^[27]。受 DBPN^[12]中上投影单元和下投影单元的启发, 本文在投影单元的基础上进行改进, 构建了 BPDB 和 BPUB。具体地, 反投影采样模块在投影单元的基础上添加了一个 1×1 的卷积层, 该卷积层用于自适应地调整采样后的残差结果和特征图之间的比例。相比于 DBPN^[12]中投影单元只是简单地将残差结果和特征图相加, 本文提出的反投影采样模块能够学习出残差结果和特征图间的相互依赖关系, 并自适应地调整两者的权重, 从而提升反投影采样模块

的表征能力. 改进后的反投影采样模块能够学习出更复杂的特征用以图像重建, 进一步提高网络的重建性能.

此外, DBPN^[12]中投影单元的卷积核尺寸随着比例因子的增大而增大, 与之不同的是, 本文采用固定大小的卷积核以更好地适应 U-Net 结构. 本文用填充和步长都为 2 的 6×6 卷积层实现下采样, 用相同参数的反卷积层实现上采样, 并在每个卷积层和反卷积层之后都连接一层 LeakyReLU 激活层. BPDB 将 HR 特征图下采样为原来的一半, BPUB 将 LR 特征图上采样到原来的 2 倍. 反投影采样模块的具体结构如图 3 所示.

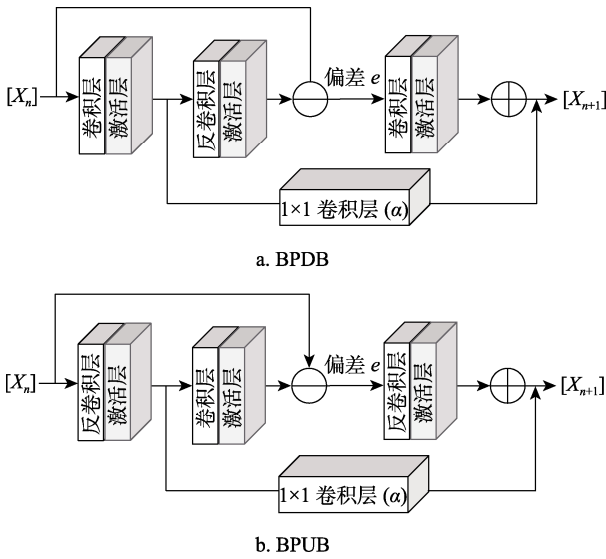


图 3 反投影采样模块

反投影采样模块中的 BPDB 下采样过程可表示为 $x_{n+1} = \alpha D x_n + D(x_n - U D x_n)$. 其中, x_n 为 BPDB 的输入; x_{n+1} 为 BPDB 的输出; D 为下采样操作; U 为上采样操作; α 是权重系数. 同理, BPUB 上采样过程可表示为 $x_{n+1} = \alpha U x_n + U(x_n - D U x_n)$.

3 实验及结果分析

3.1 实验数据及评价标准

为了验证本文算法的 SR 重建性能, 在 2 个临床医学数据集上进行实验. 第 1 个临床数据集由美国癌症基因组谱组肾脏嫌色细胞(TCGA-KICH)^[28]提供, 该数据集包含 113 名患者的 1 083 幅腺癌和腺瘤 CT 图像. 本文将 CT 图像的体素值缩放到[0,1], 并将其作为 HR 图像, HR 图像的分辨率大小为 500×400 像素; 然后在 HR 图像上进行 Bicubic 下采样得到不同缩放比例的 LR 图像, 选取 1 000 对

LR 和 HR 图像作为训练集, 剩余的 83 对图像作为测试集.

第 2 个临床数据集来自 2016 年美国国立卫生研究院医学物理学家协会梅奥医学中心(NIH-AAPM-Mayo Clinic)低剂量 CT 大型挑战赛^[29], 该数据集包含 10 位患者的 5 936 幅全剂量 CT 图像. HR 图像通过体素尺寸为 0.74×0.74×0.80 的 FBP 算法重建得到, 其分辨率大小为 512×512 像素. 通过在 HR 图像上进行 Bicubic 下采样得到不同缩放比例的 LR 图像, 选取 5800 对 LR 和 HR 图像作为训练集, 剩余的 136 对图像作为测试集.

对于实验结果的定量分析, 本文采用峰值信噪比(peak signal to noise ratio, PSNR)和结构相似性(structural similarity, SSIM)作为实验结果的评价标准. PSNR 是图像客观评价中广泛使用的一种指标, 通过计算原始图像和重建图像之间全部像素的均方误差来判断图像的失真度. PSNR 的数值越大表示重建后的图像质量越高, 其计算公式为

$$\text{PSNR} = 10 \times \lg \left(\frac{\max(I^2)}{\text{MSE}} \right).$$

其中, $\max(I^2)$ 为图像可能的最大像素值.

SSIM 是一种衡量 2 幅图像相似度的指标. 通过比较原始图像和重建图像的结构信息来判断图像失真度. SSIM 的取值范围为[0,1], 数值越大表示 2 幅图像相似度越高, 当 2 幅图像完全相同时, SSIM=1. SSIM 的计算公式为

$$\text{SSIM}(x, y) = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}.$$

其中, x 和 y 分别表示原始图像(HR)和重建图像(SR); μ 为均值; σ^2 为方差; c_1 和 c_2 是用于避免除零的 2 个常数.

3.2 实验设置

本文实验基于 Pytorch 深度学习框架进行, 所有算法模型均在单张 NVIDIA GeForce RTX 3090 24 GB 显卡上训练完成. 实验的 batchsize 均为 16, 并使用图像块学习方式将图像分割成小块输入网络, 裁剪块的尺寸为 64×64. 本文采用 Adam 算法优化损失函数, 设置 Adam 的超参数 $\beta_1=0.9$, $\beta_2=0.999$. 初始学习率为 0.000 5, 再采用余弦退火策略调整学习率, 模型迭代共 120 次后终止训练.

3.3 实验结果分析

3.3.1 不同算法的对比实验

为验证本文提出的 BPRN 的重建性能, 本文选取 EDSR^[30], DBPN^[12], RCAN^[13], RDN^[31], DRN^[26]和

CSNLTN^[2]这 6 种基于深度学习的算法作为对比算法. 本文分别在 2 个数据集 TCGA-KICH 和 Mayo 上进行训练和测试, 测出各算法在 2 倍和 4 倍放大因子下的 PSNR 均值和 SSIM 均值. 各算法在 TCGA-KICH 和 Mayo 数据集上的客观评价指标分别如表 1 和表 2 所示, 各算法在 TCGA-KICH 和 Mayo 数据集上 4 倍重建的视觉效果如图 4 和图 5 所示. 通过比较和分析能得出, 无论是在客观评价指标还是在视觉效果上, 本文 BPRN 与其他 6 种常见的算法相比均有所提升.

表 1 不同算法在 TCGA-KICH 数据集上的性能比较

放大因子	算法	PSNR/dB	SSIM
2×	EDSR ^[30]	33.17	0.923
	DBPN ^[12]	32.87	0.913
	RCAN ^[13]	33.96	0.930
	RDN ^[31]	33.87	0.929
	DRN ^[26]	32.67	0.924
	CSNLTN ^[2]	33.50	0.929
	BPRN(本文)	34.08	0.930
4×	EDSR ^[30]	29.10	0.833
	DBPN ^[12]	28.37	0.801
	RCAN ^[13]	29.50	0.810
	RDN ^[31]	29.55	0.838
	DRN ^[26]	29.72	0.838
	CSNLTN ^[2]	29.62	0.847
	BPRN(本文)	30.61	0.847

注. 粗体表示最优值.

由表 1 可知, 在 TCGA-KICH 数据集上, 本文算法的 PSNR 均值和 SSIM 均值均高于其他 6 种算法. 具体地, 当放大因子为 2 倍时, 6 种对比算法中性

能最优及次优的分别是 RCAN 和 RDN. 而相比于 RCAN, 本文算法的 PSNR 均值提高 0.12 dB, SSIM 均值与之相等; 相比于 RDN, 本文算法的 PSNR 均值提高 0.21 dB, SSIM 均值提供高 0.001. 当放大因子为 4 倍时, 6 种对比算法中表现最佳的 2 种算法是 DRN 和 CSNLTN, 相比于 DRN, 本文算法的 PSNR 均值提高 0.89 dB, SSIM 均值提高 0.009; 相比于 CSNLTN 算法, 本文算法的 PSNR 均值提高 0.99 dB, SSIM 均值与之相等.

由表 2 可知, 在 Mayo 数据集上, 本文算法取得的 PSNR 均值和 SSIM 均值都高于其他 6 种算法. 具体地, 当放大因子为 2 倍时, 6 种对比算法中性

表 2 不同算法在 Mayo 数据集上的性能比较

放大因子	算法	PSNR/dB	SSIM
2×	EDSR ^[30]	28.82	0.898
	DBPN ^[12]	28.46	0.874
	RCAN ^[13]	28.76	0.896
	RDN ^[31]	28.78	0.896
	DRN ^[26]	30.45	0.903
	CSNLTN ^[2]	28.50	0.897
	BPRN(本文)	30.69	0.904
4×	EDSR ^[30]	27.06	0.796
	DBPN ^[12]	26.97	0.793
	RCAN ^[13]	27.15	0.767
	RDN ^[31]	26.97	0.794
	DRN ^[26]	29.23	0.819
	CSNLTN ^[2]	26.67	0.760
	BPRN(本文)	29.31	0.820

注. 粗体表示最优值.

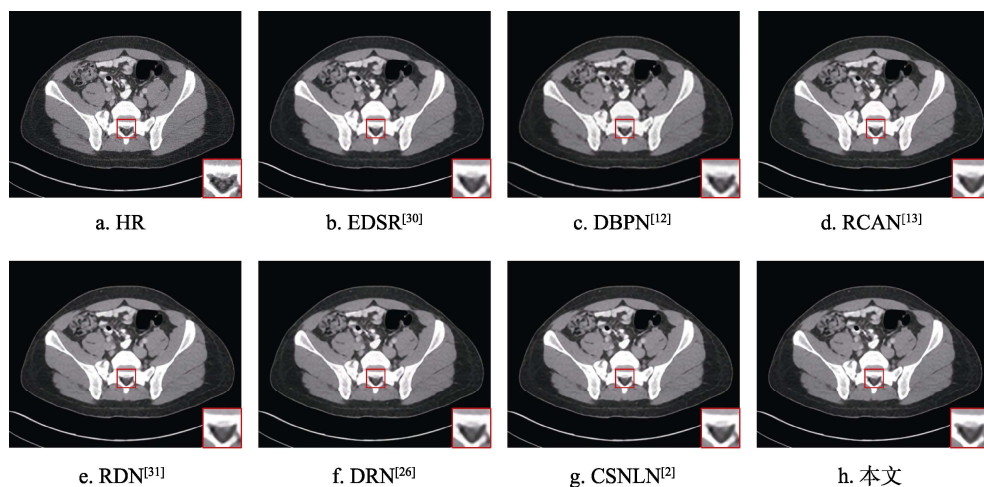


图 4 不同算法在 TCGA-KICH 测试集上的重建结果对比

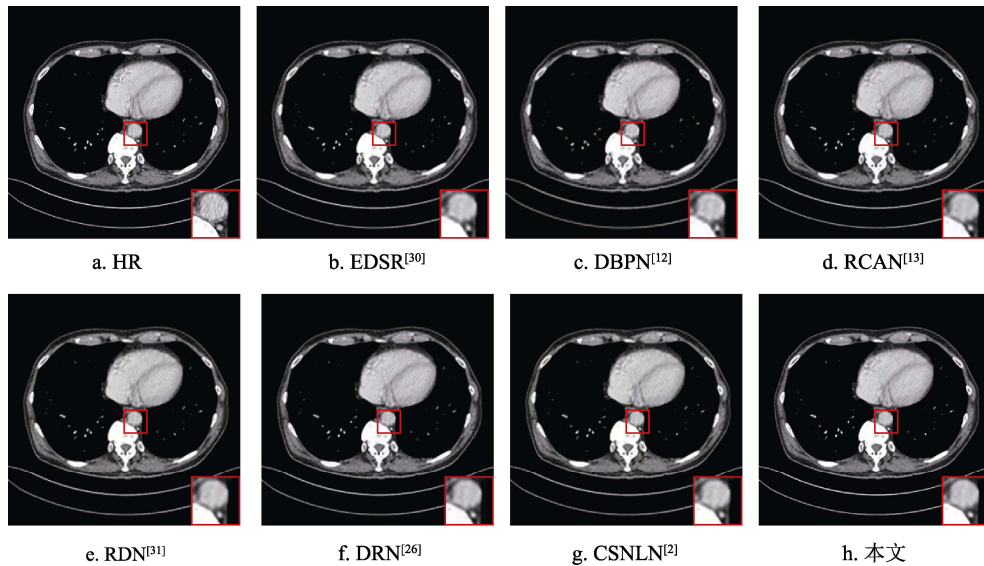


图 5 不同算法在 Mayo 测试集上的重建结果对比

提高 0.001; 相比于 EDSR, 本文算法的 PSNR 均值提高 1.87 dB, SSIM 均值提高 0.006. 当放大因子为 4 倍时, 6 种对比算法中表现最佳是 DRN, 相比于 DRN, 本文算法的 PSNR 均值提高 0.08 dB, SSIM 均值提高 0.001. 综上, 在客观评价指标方面, 相比其他 6 种主流算法, 本文算法具有更好的性能.

此外, 为验证本文 BPRN 是否显著优于其他 SOTA 算法, 本文统计了在 TCGA-KICH 和 Mayo

数据集上的对比实验结果, 进行了配对样本 t 检验. 表 3 展示了 TCGA-KICH 数据集上 2 倍放大因子下, BPRN 与 RDN 的 PSNR 值的配对样本 t 检验详细结果. 表 4 和表 5 分别展示了 TCGA-KICH 数据集和 Mayo 数据集上 BPRN 与 6 种对比算法重建图像的 PSNR 值和 SSIM 值在配对样本 t 检测中的 P 双尾值.

从表 3 中可以看出, 经配对 t 检验, BPRN 重建

表 3 BPRN 与 RDN 的 PSNR 值的配对样本 t 检验结果

算法	均值/dB	方差	泊松相关系数	df	t start	$P(T \leq t)$ 单尾	t 单尾临界	$P(T \leq t)$ 双尾	t 双尾临界
BPRN	34.08	0.41	0.61	82	3.84	0.000 2	1.68	0.000 4	2.02
RDN ^[31]	33.87	0.39							

表 4 TCGA-KICH 数据集上 BPRN 与 6 种对比算法的配对 t 检测的 P 双尾值

放大因子	对比算法	P 双尾值	
		PSNR/dB	SSIM
2×	EDSR ^[30]	1.5×10^{-17}	2.0×10^{-8}
	DBPN ^[12]	8.2×10^{-29}	3.6×10^{-45}
	RCAN ^[13]	7.3×10^{-3}	3.4×10^{-2}
	RDN ^[31]	4.0×10^{-4}	1.5×10^{-2}
	DRN ^[26]	7.6×10^{-28}	9.8×10^{-36}
	CSNLT ^[2]	8.7×10^{-16}	2.1×10^{-11}
4×	EDSR ^[30]	2.8×10^{-23}	3.3×10^{-8}
	DBPN ^[12]	2.8×10^{-34}	4.7×10^{-19}
	RCAN ^[13]	3.0×10^{-20}	5.5×10^{-18}
	RDN ^[31]	8.7×10^{-20}	1.5×10^{-5}
	DRN ^[26]	6.6×10^{-29}	1.9×10^{-36}
	CSNLT ^[2]	9.6×10^{-16}	1.2×10^{-8}

表 5 Mayo 数据集上 BPRN 与 6 种对比算法的配对 t 检测的 P 双尾值

放大因子	对比算法	P 双尾值	
		PSNR/dB	SSIM
2×	EDSR ^[30]	3.1×10^{-33}	2.3×10^{-32}
	DBPN ^[12]	1.2×10^{-34}	4.1×10^{-41}
	RCAN ^[13]	2.2×10^{-33}	7.8×10^{-33}
	RDN ^[31]	2.5×10^{-33}	8.2×10^{-33}
	DRN ^[26]	1.3×10^{-13}	2.4×10^{-25}
	CSNLT ^[2]	1.2×10^{-35}	8.1×10^{-32}
4×	EDSR ^[30]	5.4×10^{-21}	1.2×10^{-3}
	DBPN ^[12]	1.3×10^{-27}	1.9×10^{-12}
	RCAN ^[13]	2.0×10^{-20}	3.4×10^{-13}
	RDN ^[31]	2.4×10^{-21}	6.0×10^{-4}
	DRN ^[26]	1.7×10^{-7}	9.1×10^{-24}
	CSNLT ^[2]	1.3×10^{-24}	1.1×10^{-12}

结果的 PSNR 均值为 34.08 dB, RDN 算法重建结果的 PSNR 均值为 33.87 dB, 并且 P 双尾值为 0.000 4, 小于 0.05, 因而此差异是统计学显著的. 由表 4 和表 5 可知, 在 TCGA-KICH 和 Mayo 数据集上, 无论是在 2 倍放大因子还是在 4 倍放大因子情况下, BPRN 与 6 种对比算法重建图像的 PSNR 值和 SSIM 值在配对 t 检测中的 P 双尾值均小于 0.05. 因此, 本文提出的 BPRN 是显著优于 6 种对比的 SOTA 算法.

为了比较本文算法与其他算法的重建视觉效果, 本文随机选取了 TCGA-KICH 测试集和 Mayo 测试集中的 1 幅图像进行对比测试, 视觉效果对比图如图 4 和图 5 所示. 为便于观察, 本文对每幅图像同一位置的局部图进行放大, 并用红色框线标记. 从图 4 中可以看出, 其他主流算法重建出的 SR 图像中的腺体连接处有较重的模糊感, 且细节信息有所丢失, 而本文算法的重建出的 SR 图像中的腺体连接处没有造成模糊效应. 此外, 相比于其他 6 种算法, 本文算法重建出的 SR 图像具有更好的锐化纹理信息. 综上分析, 在视觉效果方面, 本文算法与其他 6 种算法相比具有更好的结果.

此外, 表 6 中展示了不同算法在 4 倍放大因子下的参数量. 由表 6 可知, 参数量由小到大排序时, 7 种算法中, 本文提出的 BPRN 的参数量排在第 3 位, 与第 2 位 DRN 仅相差 0.2M, 但 BPRN 重建出的图像比 DRN 拥有更高的 PSNR 和 SSIM 值. 虽然 CSNLN 算法的参数量是最少的, 但它的中间变量和计算量都非常大, 网络模型的训练和测试都非常耗时. 并且 BPRN 重建出的图像比 CSNLN 拥有更高的客观评价指标和更好的视觉效果, 尤其是在 Mayo 数据集上. 相比于 EDSR, DBPN, RCAN 和 RDN, BPRN 的参数量分别减少 33.1M, 0.3M, 5.6M 和 12.3M. 因此, 与 EDSR, DBPN, RCAN 和 RDN 相比, BPRN 在参数量和性能上都具有优势.

表 6 不同算法的参数量比较

算法	参数量
EDSR ^[30]	43.1M
DBPN ^[12]	10.3M
RCAN ^[13]	15.6M
RDN ^[31]	22.3M
DRN ^[26]	9.8M
CSNLN ^[2]	6.6M
BPRN(本文)	10.0M

注: 粗体表示最优值.

3.3.2 消融实验

消融实验常用于复杂神经网络的性能分析, 对于添加了多种模块的复杂网络, 可以通过删减部分网络结构来验证不同模块对网络性能的影响. 首先, 为了验证回归网络和反投影采样模块对本文网络模型的影响, 本文设计了如下消融实验. Network1 表示删去 BPRN 中的回归网络, 以验证回归网络的作用. Network2 表示将 BPRN 中的 BPDB 替换为用于下采样的最大池化层, 并将 BPUB 替换为用于上采样的反卷积层, 以验证反投影采样模块的作用. 其次, 由于反投影采样模块是基于 DBPN^[12]中投影单元的改进, 为验证改进后的提升效果, 本文将 BPRN 中的反投影采样模块替换为 DBPN^[12]中投影单元, 并用 Network3 表示. 最后, 为了验证用回归网络学习 SR $\rightarrow$ LR 逆映射的必要性, 本文将 BPRN 中的回归网络替换为 Bicubic 下采样, 用 Network4 表示. 通过比较 Network1, Network2, Network3, Network4 和 BPRN 在 TCGA-KICH 数据集上的重建效果, 来验证各部分的作用. 消融实验的客观评价指标结果如表 7 所示.

表 7 消融实验结果

放大因子	模型	PSNR/dB	SSIM
2 $\times$	Network1	32.58	0.923
	Network2	32.66	0.924
	Network3	34.01	0.930
	Network4	33.84	0.929
	BPRN(本文)	34.08	0.930
4 $\times$	Network1	29.83	0.839
	Network2	29.71	0.838
	Network3	30.51	0.847
	Network4	30.28	0.846
	BPRN(本文)	30.61	0.847

注: 粗体表示最优值.

由表 7 可知, 删去回归网络的 Network1 相较于 BPRN, 其 PSNR 均值和 SSIM 均值都有所下降, 在 2 倍放大因子情况下, 其 PSNR 均值和 SSIM 均值分别降低 1.5 dB 和 0.007, 在 4 倍放大因子情况下, 其 PSNR 和 SSIM 均值分别降低 0.78 dB 和 0.008, 因而本文提出的回归网络是有效的. 替换掉反投影采样模块的 Network2 相较于 BPRN, 其 PSNR 均值和 SSIM 均值也都有所下降, 在 2 倍放大因子情况下, 其 PSNR 均值和 SSIM 均值分别降低 1.42 dB 和 0.006, 在 4 倍放大因子情况下, 其

PSNR 均值和 SSIM 均值分别降低 0.9 dB 和 0.009, 因而本文提出的反投影采样模块是有效的。

由于本文提出的反投影采样模块是基于 DBPN^[12]中投影单元的改进, 即在投影单元中添加一个 1×1 的卷积作为权重层, 用于自适应地调整采样后的残差结果和特征图之间的比例。为了验证引入权重层这一改进的有效性, 本文将 BPRN 中的反投影采样模块替换为原始的投影单元, 用 Network3 表示, 并比较 Network3 和 BPRN 性能差异。由表 7 可知, Network3 相比于 BPRN, 其 SSIM 均值未变, 但 PSNR 均值有所下降, 在 2 倍放大因子情况下, 其 PSNR 均值降低 0.07 dB, 在 4 倍放大因子情况下, 其 PSNR 均值降低 0.1 dB, 因而本文对投影单元的改进是有效的。

此外, 由表 7 可知, 用 Bicubic 下采样替换回归网络的 Network4 相比于 BPRN, 其 PSNR 均值和 SSIM 均值都有所下降, 在 2 倍放大因子情况下, 其 PSNR 均值和 SSIM 均值分别降低 0.24 dB 和 0.001, 在 4 倍放大因子情况下, 其 PSNR 和 SSIM 均值分别降低 0.33 dB 和 0.001。虽然仅用下采样也能实现 $SR \rightarrow LR$ 逆映射, 但相比于通过回归网络实现逆映射, 前者重建出的图像质量要低于后者。因此, 本文使用回归网络而非用下采样来实现 $SR \rightarrow LR$ 逆映射是有必要的。

图 6 展示了消融实验结果的局部放大图, 可以看出 BPRN 重建出的图像边界模糊感较弱, 整体的视觉效果更好。综上所述, 本文提出的回归学习策略和反投影采样模块是有效的, 对 DBPN^[12]中投影单元的改进是有意义的, 采用回归网络来学习 $SR \rightarrow LR$ 逆映射是有必要的。

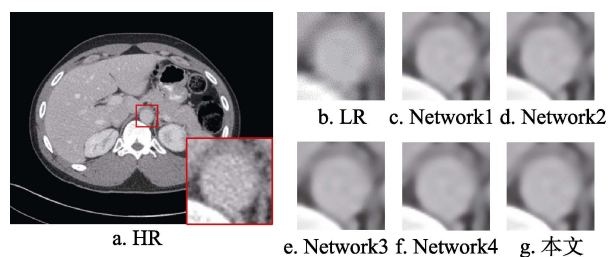


图 6 消融实验结果的局部放大图

3.3.3 超参数 λ 对 BPRN 的影响

由 BPRN 的损失函数可知: 超参数 λ 控制着 BPRN 中回归网络的贡献度。为了使回归网络发挥最佳效果, 本文通过实验来验证回归网络的权重 λ 对 BPRN 重建性能的影响, 实验结果如表 8 所示。从表 8 可以看出, 当权重 λ 从 0.001 逐渐增至 0.100

时, 回归网络的损失随之变得愈加重要, BPRN 的重建性能越来越好。但是当权重 λ 从 0.100 逐渐增至 10.000 时, 回归网络的权重过大, 挤占了原始网络的作用, 从而导致 BPRN 的重建性能下降。因此, 为了平衡 BPRN 中原始网络和回归网络的作用, 从而获得最佳的重建性能, 本文设置超参数 $\lambda=0.100$ 。

表 8 超参数 λ 对 BPRN 的影响

λ	PSNR/dB
0.001	30.43
0.010	30.60
0.100	30.62
1.000	30.33
10.000	29.39

注: 粗体表示最优值。

4 结 语

针对使用深度学习进行医学图像 SR 重建时出现的欠定性问题, 本文提出了一种 BPRN 来重建图像。BPRN 通过回归网络引入一个额外的约束来减小 $LR \rightarrow HR$ 可能的映射空间, 从而使 BPRN 学习出一个更好的映射, 提升重建性能。此外, BPRN 引入具有误差反馈纠正的上、下采样块来减小网络进行上、下采样时丢失的特征信息。实验结果表明, 相较于其他主流的基于深度学习的图像 SR 重建算法, 本文算法重建出的图像具有更高的客观评价指标和更好的视觉效果。然而, 本文 BPRN 需要在配对的医学图像数据集上进行训练学习, 而真实的医疗场景中配对的医学图像较少。因此, 在未来的工作中将改进并优化 BPRN, 使其可以在未配对的医学数据集上进行图像重建。

参考文献(References):

- [1] Lyu Q, Shan H M, Wang G. MRI super-resolution with ensemble learning and complementary priors[J]. IEEE Transactions on Computational Imaging, 2020, 6: 615-624
- [2] Mei Y Q, Fan Y C, Zhou Y Q, et al. Image super-resolution with cross-scale non-local attention and exhaustive self-exemplars mining[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 5689-5698
- [3] Blu T, Thevenaz P, Unser M. Linear interpolation revitalized[J]. IEEE Transactions on Image Processing, 2004, 13(5): 710-719
- [4] Zhou D W, Shen X L, Dong W M. Image zooming using directional cubic convolution interpolation[J]. IET Image Processing, 2012, 6(6): 627-634

- [5] Keys R. Cubic convolution interpolation for digital image processing[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1981, 29(6): 1153-1160
- [6] Irani M, Peleg S. Improving resolution by image registration[J]. CVGIP: Graphical Models and Image Processing, 1991, 53(3): 231-239
- [7] Tekalp A M, Ozkan M K, Sezan M I. High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration[C] //Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. Los Alamitos: IEEE Computer Society Press, 1992: 169-172
- [8] Liu C, Sun D Q. On bayesian adaptive video super resolution[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(2): 346-360
- [9] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444
- [10] Dong C, Loy C C, He K M, *et al.* Learning a deep convolutional network for image super-resolution[C] //Proceedings of the 13th European Conference on Computer Vision. Heidelberg: Springer, 2014: 184-199
- [11] 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
- [12] Haris M, Shakhnarovich G, Ukita N. Deep back-projection networks for super-resolution[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 1664-1673
- [13] 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: 294-310
- [14] Badran Y K, Salama G I, Mahmoud T A, *et al.* Single image super resolution based on learning features to constrain back projection[C] //Proceedings of the International Conference on Innovative Trends in Computer Engineering. Los Alamitos: IEEE Computer Society Press, 2019: 23-28
- [15] Song Zhaoyang, Zhao Xiaoqiang, Hui Yongyong, *et al.* Inverted N-type lightweight network based on back projection for image super-resolution reconstruction[J]. Journal of Computer-Aided Design & Computer Graphics, 2022, 34(6): 923-932 (in Chinese)
(宋昭漾, 赵小强, 惠永永, 等. 基于反向投影的倒 N 型轻量化网络图像超分辨率重建[J]. 计算机辅助设计与图形学学报, 2022, 34(6): 923-932)
- [16] Li H J, Gao Y X, Dong J W, *et al.* Super-resolution based on noise resistance deep convolutional network[C] //Proceedings of the 6th International Conference on Bioinformatics and Computational Biology. New York: ACM Press, 2018: 88-94
- [17] Hong X, Zan Y L, Weng F H, *et al.* Enhancing the image quality via transferred deep residual learning of coarse PET sinograms[J]. IEEE Transactions on Medical Imaging, 2018, 37(10): 2322-2332
- [18] Song T A, Chowdhury S R, Yang F, *et al.* PET image super-resolution using generative adversarial networks[J]. Neural Networks, 2020, 125: 83-91
- [19] Xia H, Cai N, Wang H H, *et al.* Brain MR image super-resolution via a deep convolutional neural network with multi-unit upsampling learning[J]. Signal, Image and Video Processing, 2021, 15(5): 931-939
- [20] Zhao X L, Zhang Y L, Zhang T, *et al.* Channel splitting network for single MR image super-resolution[J]. IEEE Transactions on Image Processing, 2019, 28(11): 5649-5662
- [21] Yi Z L, Zhang H, Tan P, *et al.* DualGAN: unsupervised dual learning for image-to-image translation[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 2868-2876
- [22] Zhu J Y, Park T, Isola P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 2242-2251
- [23] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C] //Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Heidelberg: Springer, 2015: 234-241
- [24] Jin Y, Jiang Z W, Huang M J, *et al.* Low-dose CT image restoration based on noise prior regression network[J]. The Visual Computer, 2023, 39(1): 459-471
- [25] Jin Yan, Xue Zhizhong, Jiang Zhiwei. Medical image segmentation based on recurrent residual convolution neural network[J]. Journal of Computer-Aided Design & Computer Graphics, 2022, 34(8): 1205-1215(in Chinese)
(金燕, 薛智中, 姜智伟. 基于循环残差卷积神经网络的医学图像分割算法[J]. 计算机辅助设计与图形学学报, 2022, 34(8): 1205-1215)
- [26] Guo Y, Chen J, Wang J D, *et al.* Closed-loop matters: dual regression networks for single image super-resolution[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 5406-5415
- [27] Ma L, Li S M. Enhanced back projection network based stereo image super-resolution considering parallax attention[C] //Proceedings of the IEEE International Conference on Image Processing. Los Alamitos: IEEE Computer Society Press, 2021: 1834-1838
- [28] TCGA-KICH. The cancer genome atlas kidney chromophobe[EB/OL]. [2022-12-29]. <https://wiki.cancerimagingarchive.net/display/Public/TCGA-KICH>
- [29] AAPM. Low dose CT grand challenge(2017)[EB/OL]. [2022-12-29]. <https://www.aapm.org/GrandChallenge/LowDoseCT/#>
- [30] 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
- [31] Zhang Y L, Tian Y P, Kong Y, *et al.* Residual dense network for image super-resolution[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 2472-2481