

基于对抗训练的U-Net神经网络在稀疏投影CT图像增强的应用

黄锦威¹, 萧文鹏², 朱思婷³, 丘皓怡¹, 陈星宇², 刘深泉¹

1. 华南理工大学数学学院, 广东 广州 510640; 2. 华南理工大学计算机科学与工程学院, 广东 广州 510006; 3. 华南理工大学电力学院, 广东 广州 510641

【摘要】目的:针对稀疏投影的CT重建图像附带噪声和伪影的特性,使用神经网络模型对稀疏投影得到的低质量CT重建图像进行图像增强。**方法:**在残差编码-解码卷积神经网络基础上提出一种基于对抗训练的U-Net神经网络模型,并使用公开数据集TCGA-CESC癌症CT影像进行模型训练和测试。评价模型处理效果的指标包括峰值信噪比(PSNR)、结构相似性(SSIM)和均方根误差(RMSE)。**结果:**在对180次探测的CT重建图像的测试中,模型处理后的图像相比未处理图像,PSNR、SSIM和RMSE指标平均值分别提升15.10%、37.89%和38.20%。在PSNR和SSIM指标平均值意义下,模型处理后的图像优于1800次探测的未处理CT重建图像。**结论:**本研究提出的神经网络模型能够减少伪影和噪点,对稀疏投影CT图像增强有一定效果。

【关键词】CT图像增强; U-Net神经网络; 对抗训练; 稀疏投影

【中图分类号】TP391; R318

【文献标志码】A

【文章编号】1005-202X(2020)05-0612-07

Application of adversarial training-based U-Net neural network in the enhancement of CT images obtained by sparse projection

HUANG Jinwei¹, XIAO Wenpeng², ZHU Siting³, QIU Haoyi¹, CHEN Xingyu², LIU Shenquan¹

1. School of Mathematics, South China University of Technology, Guangzhou 5106401, China; 2. School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China; 3. School of Electric Power Engineering, South China University of Technology, Guangzhou 510641, China

Abstract: Objective For the existence of noise and artifacts in reconstructed CT images obtained by sparse projection, a neural network model is proposed for the enhancement of low-quality CT images obtained by sparse projection. **Methods** Based on residual encoder-decoder convolutional neural network, an adversarial training-based U-Net neural network was proposed. Model training and testing were carried out using the CT images of cancer from TCGA-CESC dataset. The main indicators for the evaluation of the model processing results included peak signal-to-noise ratio (PSNR), structural similarity (SSIM) and root-mean-square error (RMSE). **Results** In the test on the reconstructed CT images obtained by 180 detections, the average values of the PSNR, SSIM and RMSE of the images processed by the proposed model were increased by 15.10%, 37.89% and 38.20%, respectively, compared with those of the unprocessed images. Moreover, the average values of PSNR and SSIM indicated that the processed images were superior to the unprocessed images obtained by 1800 detections. **Conclusion** The proposed neural network model can reduce artifacts and noise, and has certain effects on the enhancement of CT images obtained by sparse projection.

Keywords: CT image enhancement; U-Net neural network; adversarial training; sparse projection

前言

X射线CT技术是一种无损检测技术,广泛应用

于医疗检查和工业物品检测等领域^[1]。由于X射线的辐射性,所以在医疗诊断中,对特殊受检者要适当降低X射线剂量或减少探测次数,从而降低检测过程中辐射的影响,但低剂量的X射线投影过程容易受到噪声干扰,而稀疏投影会使探测得到的信息量减少,导致重建的CT图像带有伪影,影响进一步的医疗诊断。

传统的CT图像降噪常用滤波函数对图像进行

【收稿日期】2019-12-09

【基金项目】“攀登计划”广东大学生科技创新培育专项资金(pdjh2018-b0043)

【作者简介】黄锦威,研究方向:数学物理,E-mail: 1191572861@qq.com

【通信作者】刘深泉,教授,E-mail: mashqliu@scut.edu.cn

全局滤波平滑,但这种无差别的平滑处理可能丢失CT图像的部分细节信息。而近年来基于神经网络模型的图像处理技术得到广泛研究应用,如图像分类^[2]、物体检测^[3]、图像超分辨率化^[4-6]、图像风格化^[7-8]等,在图像数据集测试中取得优于传统方法的效果。Chen等^[9]提出残差编码-解码卷积神经网络(Residual Encoder-Decoder Convolutional Neural Network, RED-CNN)模型用于低剂量CT图像增强,本研究基于RED-CNN模型,提出一种基于对抗训练的U-Net神经网络,用于稀疏投影CT图像增强。

1 CT重建算法

CT图像的重建算法主要分为两类:一类是基于Radon变换^[10-11]的解析类重建算法,另一类是以解方程为主要思想的代数重建算法^[12]。

解析类重建算法又可分为滤波反投影算法^[11,13]和直接傅里叶变换重建算法^[14]。其中,滤波反投影算法基于反投影法,在其基础上增加卷积滤波过程,具有重建速度快、空间和密度分辨率高等优点,在实际中得到广泛应用;直接傅里叶变换重建算法由于其频域插值计算复杂度较高,在实际应用中常采用重建质量高的滤波反投影算法^[15]。代数重建法是利用投影数据构建一组线性方程,通过求解线性方程组进行图像重建。但是代数重建法计算量大、重建时间长,因此在一定程度上限制了其应用^[16]。

由于滤波反投影算法在实际应用中占有重要地位,本研究将对平行束投影的滤波反投影算法的基本原理以及该算法对稀疏投影的重建图像附带伪影的现象进行介绍。

反投影算法的基本思想是:待重建图像中某个位置的像素值,可以通过计算所有经过这一点的投影值的平均值近似得到^[13]。以3×3大小的图像投影为例,原图像中有9个像素 X_1, X_2, \dots, X_9 ,分别从0°、45°、90°和135°这4个方向进行投影,得到投影值 P_1, P_2, \dots, P_{16} (图1)。投影值 $P_i (1 \leq i \leq 16)$ 和像素值 $X_j (1 \leq j \leq 9)$ 的关系如方程组(1)所示。

$$\begin{aligned} P_1 &= X_1 + X_4 + X_7, P_2 = X_2 + X_5 + X_8, \\ P_3 &= X_3 + X_6 + X_9, P_4 = X_1, \\ P_5 &= X_2 + X_4, P_6 = X_3 + X_5 + X_7, P_7 = X_6 + X_8, \\ P_8 &= X_9, P_9 = X_1 + X_2 + X_3, \\ P_{10} &= X_4 + X_5 + X_6, P_{11} = X_7 + X_8 + X_9, \\ P_{12} &= X_3, P_{13} = X_2 + X_6, P_{14} = X_1 + X_5 + X_9, \\ P_{15} &= X_4 + X_8, P_{16} = X_7 \end{aligned} \quad (1)$$

基于方程组(1)得到反投影计算式(2)。

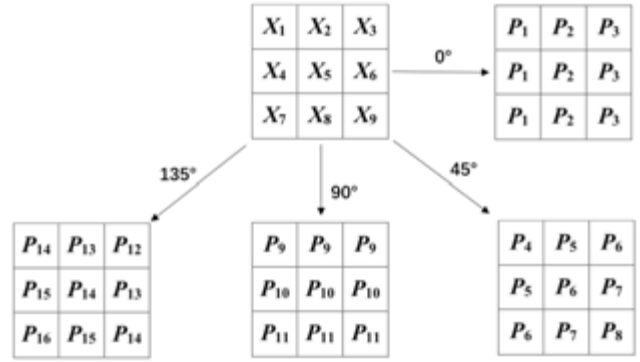


图1 3×3图像在4个不同角度的投影

Fig.1 Projection of a 3×3 image at 4 different angles

$$\begin{aligned} X_1 &= \frac{1}{4} (P_1 + P_4 + P_9 + P_{14}), \\ X_2 &= \frac{1}{4} (P_2 + P_5 + P_{10} + P_{13}), \\ X_3 &= \frac{1}{4} (P_3 + P_6 + P_{11} + P_{12}), \\ X_4 &= \frac{1}{4} (P_1 + P_5 + P_{10} + P_{15}), \\ X_5 &= \frac{1}{4} (P_2 + P_6 + P_{11} + P_{14}), \\ X_6 &= \frac{1}{4} (P_3 + P_7 + P_{12} + P_{13}), \\ X_7 &= \frac{1}{4} (P_1 + P_6 + P_{11} + P_{16}), \\ X_8 &= \frac{1}{4} (P_2 + P_7 + P_{12} + P_{15}), \\ X_9 &= \frac{1}{4} (P_3 + P_8 + P_{13} + P_{14}) \end{aligned} \quad (2)$$

以上示例仅采用4个不同角度进行投影,当图像像素比较多时,为得到高质量的重建图像,需要选取更密集的投影角度,一般是取值跨度180°的多个均匀分布的投影角度。由反投影计算的原理可知,该算法的重建结果是近似的。从图2可以看出由于中心像素的像素值1的影响,使得周围的像素值在重建后变大了。这种在反投影重建中由一个位置的像素值影响其周围位置像素值产生的差异称为星状伪影。产生星状伪影的原因是反投影重建是将一个投影值均匀地反投影到其投影路径的每个像素上,因此当该投影路径上各点的像素值并非均匀分布时,其像素值较大的位置的值会被“分摊”到像素值较小的位置上^[17]。星状伪影可以通过增加投影角度或对投影数据进行滤波得到减轻或者去除,由于滤波反投影算法在反投影算法上加入了卷积滤波过程,因而重建效果更好。常用的滤波函数有Hamming、Hanning、R-L、S-L、Cosine等。但在稀疏投影的情况下,即使对投影数据进行了滤波,其反投影结果依然会带有明显伪影和噪点。

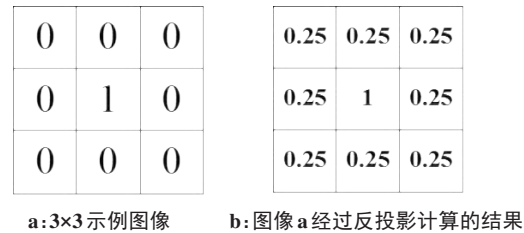


图2 3×3 图像与反投影计算结果比较

Fig.2 Comparison of a 3×3 image and the result of its backprojection calculation

2 基于对抗训练的 U-Net 神经网络

U-Net 模型由 Ronneberger 等^[18]首次提出并用于解决医学图像分割的任务,通过低层信息往高层补充,使模型能够处理医学图像的细节。近年来,对抗生成网络 GAN^[19]被广泛应用于图像的生成与转换,通过内部生成器与鉴别器的对抗、博弈,在对抗过程

中生成器不断提高生成更真实目标图像的能力,鉴别器则不断增强自己的鉴别能力,以区分输入图像的真伪。鉴别器直接区分“模糊的、有噪声的 CT 图像与清晰、高质量的 CT 图像”的能力对于增强模型的性能有引导作用。

2.1 模型与原理

基于对抗训练的 U-Net 神经网络模型结构如图 3 所示,模型以稀疏投影数据经过滤波反投影算法重建得到的带噪声和伪影的 CT 图像 Input(I)作为输入,经过增强模型 G(一个浅层的 U-Net)处理,输出增强后的 CT 图像 Output(O)。将输出图像 O 与对应的目标图像 GT 分别输入鉴别器 Decoder(D),鉴别器 D 对输入进行判断,如果输入是 GT,则 D 倾向于判断为 True;如果输入是图像 O,则 D 倾向于判断为 False。鉴别器 D 由连续的 4 层卷积操作与最后的全连接层 FC 构成,结构如图 4 所示。

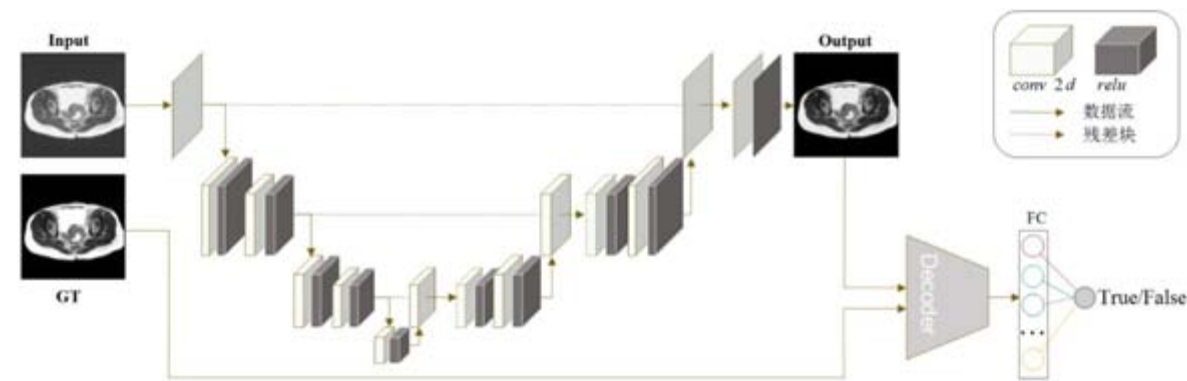


图3 基于对抗训练的 U-Net 神经网络结构

Fig.3 Structure of U-Net neural network based on adversarial training

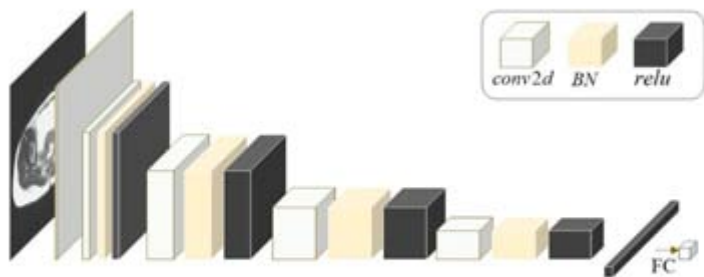


图4 鉴别器结构模型

Fig.4 Structure of the decoder

2.2 目标函数选取

2.2.1 增强网络的目标函数 模型采用逐像素的均方差 (Mean Square Error):

$$L_{con} = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W \|O_{h,w} - GT_{h,w}\|_2^2 \tag{3}$$

其中, H 、 W 分别指图像的高度与宽度。

定义 G 的对抗损失为:

$$L_{adv} = E_{O \sim P_{G(I)}} [1 - \log(D(O))] \tag{4}$$

G 的损失函数为:

$$L_G = L_{con} + \lambda L_{adv} \tag{5}$$

其中, λ 是一个超参数,在本模型实验中取 $\lambda=10$ 。

2.2.2 鉴别器的目标函数 鉴别器 D 的目标函数定义式(6)。

$$L_D = E_{GT \sim P_{GT}}[1 - \log(D(GT))] +$$
$$E_{O \sim P_{G(I)}}[\log(D(O))] \tag{6}$$

3 实验分析

3.1 数据集选取和处理

模型的训练与测试使用公开数据集 TCGA-CESC 癌症 CT 影像中的数据, 其中包含 2 900 组来自人体不同部位的 CT 图片数据。筛选出大小适宜的图片数据 1 942 组, 从中随机选择 72 组数据作为测试集, 其余 1 870 组数据作为训练集。

将每组 CT 图像数据读取为二维矩阵 $GT \in R^{H \times W}$ 。假设 CT 扫描在不同角度进行探测时设备参数条件相同, 利用 Radon 变换模拟 X 射线投影过程, 获取图像在 180 个角度(相隔 1°进行一次投影)的投影数据, 经过滤波反投影重建, 得到有噪声和伪影的 CT 图像矩阵 $I \in R^{H \times W}$ 。将每个 GT 与其对应的 I 称为一组数据。由于整体训练样本偏少, 因此将 1 870 组训练数据进行裁剪, 每组裁剪得到不同位置的 5 组图像数据, 将裁剪得到的 9 350 组数据作为训练集用于模型的训练。

3.2 评价指标

本文采用 3 种典型的图像质量评估指标。

(1) 峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 指的是峰值信号的能量与噪声的能量的均值之比。

$$PSNR = 10\log_{10}\left(\frac{MAX^2}{MSE}\right) = 20\log_{10}\left(\frac{MAX}{\sqrt{MSE}}\right) \tag{7}$$

其中, MSE 为均方差, 计算式如式 (3); MAX 为图像的灰度级, 一般取 MAX=255。PSNR 值越大, 说明两张图像相似度越高。

(2) 结构相似性 (Structural Similarity, SSIM) 是基于两张图像的亮度、对比度与结构这 3 个方面衡量两者相似度的方法, SSIM 值越高, 说明两张图像相似度越高^[20]。

(3) 均方根误差 (Root-Mean-Square Error, RMSE) 计算式见式 (8), 用于衡量两组数据的偏差。RMSE 值越小, 说明两张图像相似度越高。

$$RMSE = \sqrt{MSE} \tag{8}$$

3.3 实验细节

模型的训练和测试在一台 CPU 核为 i7、内存 32G、内置 1080Ti GPU 的服务器上完成, 使用的图像分辨率为 384×384。测试阶段, 无 GPU 驱动下单张图像的处理时间为 (0.353±0.010) s; 1080Ti GPU 驱动下单张图像的处理时间为 (3.406±0.350) ms。

3.4 实验结果

将模型输入图像 I 和最终输出图像 O 分别与数据集中原始图像 GT 进行比较, 并计算 PSNR、SSIM 和 RMSE, 两张图像相似度越高表明图像质量越高。

3.4.1 基于测试集的单样本分析 从测试集中随机选择一个样本, 本研究提出的模型得到的结果见图 5 和图 6。

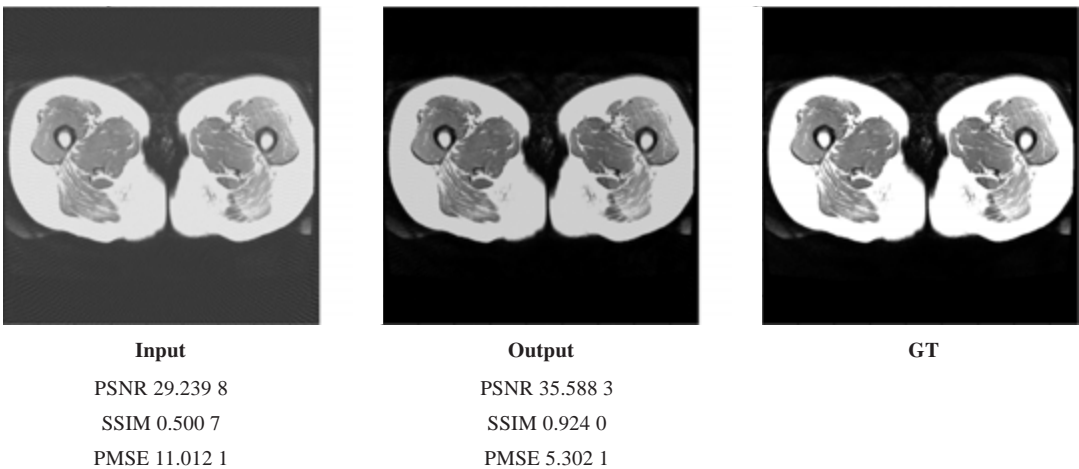


图5 模型输出结果与输入样本的比较
Fig.5 Comparison of the input and the output of one sample

由图 5 可知, 模型输出图像的 3 项评价指标均优于所选的输入样本图像。从图 6 可以明显发现模型输出图像相比输入样本图像有更少的伪影和噪声以及更清晰的边缘轮廓, 细节上更接近原始图像。

3.4.2 基于测试集的总体分析 基于测试集中数据,

统计本研究提出的模型对测试图像在 3 个指标中的表现, 结果见表 1。其中, 提升数值表示模型处理结果相对于未处理图像在相应指标上提升的数值, PSNR 和 SSIM 指标的提升数值为模型处理结果的指标值减去未处理图像的指标值, PMSE 指标的提升数

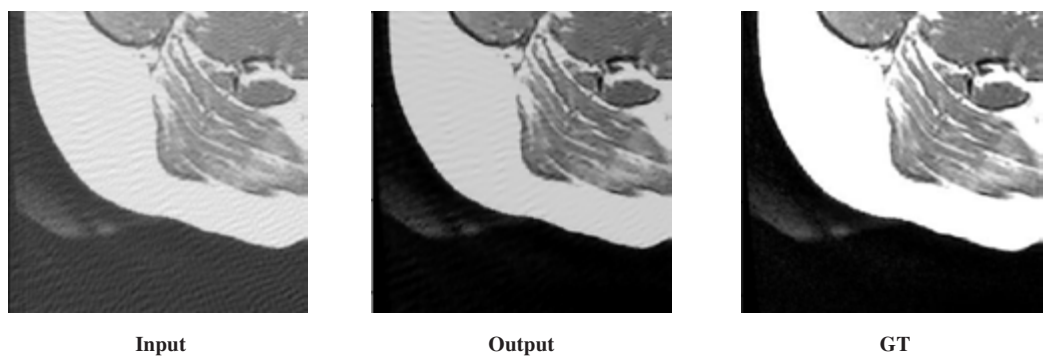


图6 图5中图片的局部放大
Fig.6 Partial magnification of Fig.5

值为未处理图像的指标值减去模型处理结果的指标值。提升百分比为指标提升数值与未处理图像相应指标的比值。

本研究提出的模型是基于RED-CNN上的创新，将本研究提出的模型与RED-CNN的处理结果进行比较,各项评价指标对比如表2所示。

表1 本文模型在测试集上的提升效果
Tab.1 Enhancement effects of the proposed model on the test set

提升效果	PSNR			SSIM			RMSE		
	最高	最低	平均	最高	最低	平均	最高	最低	平均
提升数值	9.1435	-2.3907	4.7117	0.4685	-0.0311	0.2170	7.3519	-0.5666	3.1703
提升百分比/%	29.35	-6.11	15.10	100.73	-3.18	37.89	65.10	-31.69	38.20

表2 本文模型与RED-CNN在测试集上的处理结果比较
Tab.2 Comparison of the results of the proposed network with RED-CNN on the test set

处理方法	PSNR			SSIM			RMSE		
	最高	最低	平均	最高	最低	平均	最高	最低	平均
RED-CNN	42.454 8	25.516 8	35.889 1	0.965 2	0.815 2	0.909 7	15.107 5	2.001 0	5.305 4
本文模型	42.260 3	25.751 9	35.840 7	0.963 9	0.812 1	0.910 1	14.822 9	2.054 3	5.387 0
无处理	39.144 4	25.740 1	31.129 1	0.980 5	0.465 1	0.693 1	17.294 5	1.788 2	8.557 3

从表1的提升数据可知,在平均意义下,本研究提出的模型的处理结果相比于未处理的图像有较大提升,其中SSIM和RMSE的平均提升百分比均超过35%。同时某些测试图像的模型处理结果在评价指标意义下差于未处理图像,在72组测试集中有2组数据出现负提升,图7为其中一组出现负提升的样本。经人工对比分析,发现这两幅图像在数据集TCGA-CESC中的原图有较多噪点,而模型处理后减少了图像的噪点,处理后图像更加平滑(图8),因此在基于原图比对的指标下出现了负提升现象。

从表2的对比数据可知,在3种评价指标下本研究提出的模型的结果与RED-CNN比较接近,相较于处理前的图像(Input)有明显的提升。本研究提出的模型与RED-CNN相比,PSNR得分的极差分别为

16.508 4与16.938 0, RMSE得分的极差分别为12.768 6与13.106 5, SSIM得分的极差分别为0.151 8与0.150 0,在评价指标极差意义下本研究提出的模型更稳定。

为了进一步说明模型对稀疏投影重建图像的提升效果,将本研究提出的模型对180次探测(1°/次)重建图像的处理结果与180次探测(1°/次)、360次探测(0.5°/次)、720次探测(0.25°/次)和1 800次探测(0.1°/次)的未经本模型增强的重建图像进行比较。图9中3项评价指标的数值为测试集的评价结果的均值归一化,可以看出本文模型相比180次探测的重建图像在3个指标上均有明显提升,PSNR和RMSE指标与360次探测、720次探测和1 800次探测的重建结果接近,SSIM指标则优于1 800探测的重建结果。说明本文

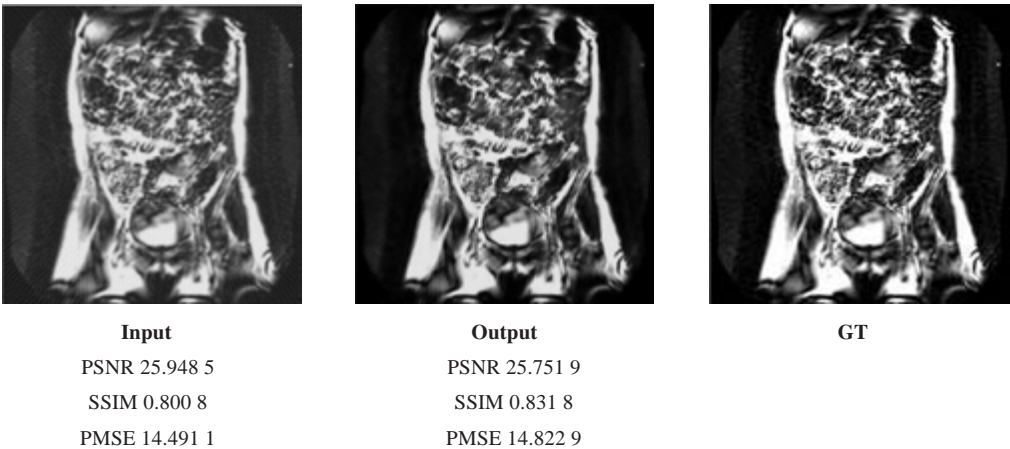


图7 出现负提升现象的一组数据的模型输出结果与输入样本比较
Fig.7 Comparison of the input with the output of one sample with negative enhancement

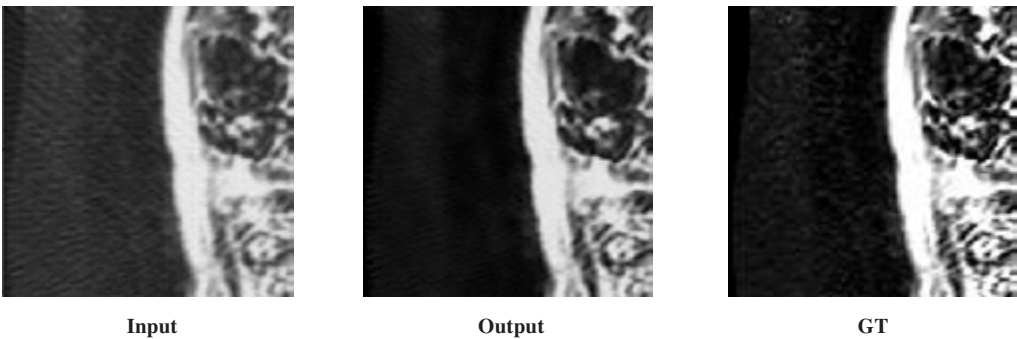


图8 图7中图像的局部放大
Fig.8 Partial magnification of Fig.7

模型对稀疏探测重建的CT图像确有提升作用。此外,360次探测、720次探测和1 800次探测的重建结果指标上接近,这是因为当探测密度增加到一定程度时,继续增加探测密度对重建图像的质量提升作用并不显著。

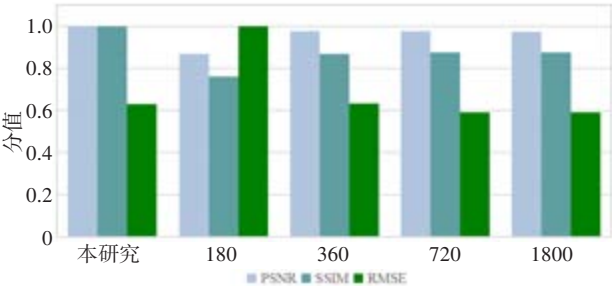


图9 本文模型结果与不同探测密度的重建图像比较
Fig.9 Comparison of the results obtained by the proposed network and reconstructed images with different detection densities

4 结 论

本研究提出了基于对抗训练的U-Net神经网络模型,并基于公开数据集TCGA-CESC癌症CT影像通过实验计算证明该模型对稀疏投影重建后带伪影

的CT图像有明显的增强效果。在180次探测测试中,在均值意义下SSIM指标和RMSE指标的提升超过35%,且本研究模型对180次探测重建图像的处理结果在PSNR和SSIM指标下优于1 800次探测的重建图像。说明本研究模型对稀疏探测的条件下的重建图像进行增强,其图像质量接近甚至优于密集探测条件下重建的CT图像。基于上述研究结果,进一步可以研究模型的稳定性,以及如何通过改良模型框架和调整必要参数使模型有更好的表现。

【参考文献】

[1] 先武,李时光,王珏.最佳无损检测手段-工业CT技术的发展[J].光电工程,1995,22(4): 51-58.
XIAN W, LI S G, WANG J. Development of the optimal nondestructive testing approach-industrial CT[J]. Opto-Electronic Engineering, 1995, 22(4): 51-58.

[2] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks[C]//Neural Information Processing Systems 25. New York: Curran Associates Inc, 2012: 1106-1114.

[3] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Trans Pattern Anal Machine Intell, 2017, 39(6): 1137-1149.

[4] ROMANO Y, ISIDORO J, MILANFAR P R. Rapid and accurate image super resolution[J]. IEEE Trans Comput Imaging, 2016, 3

- (1): 110-125.
- [5] LEDIG C, THEIS L, HUSZAR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[J]. arXiv: 1609.04802, 2016.
- [6] CHEN Y, SHI F, CHRISTODOULOU A G, et al. Efficient and accurate MRI super-resolution using a generative adversarial network and 3D multi-level densely connected network[C]. Medical Image Computing and Computer Assisted Intervention-MICCAI, 2018: 91-99.
- [7] GATYS L A, ECKER A S, BETHGE M. Image style transfer using convolutional neural networks [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016. DOI: 10.1109/CVPR.2016.265.
- [8] JIMENEZ-ARREDONDO V H, CEPEDA-NEGRETE J, SANCHEZ-YANEZ R E. Multilevel color transfer on images for providing an artistic sight of the world[J]. IEEE Access, 2017, 5: 15390-15399.
- [9] CHEN H, ZHANG Y, KALRA M K, et al. Low-dose CT with a residual encoder-decoder convolutional neural network [J]. IEEE Trans Med Imaging, 2017, 36(12): 2524-2535.
- [10] BEYLKIN G. Discrete radon transform[J]. IEEE Trans Acoust Speech Signal Process, 1987, 35(2): 162-172.
- [11] DEANS S R. The radon transform and some of its applications[M]. New York: Wiley, 1984: 55-61.
- [12] GORDON R, BENDER R, HERMAN G T. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography[J]. J Theor Biol, 1970, 29(3): 471-481.
- [13] KUCHMENT P. The radon transform and medical imaging [M]. Philadelphia: Society for Industrial and Applied Mathematics, 2014: 21-48.
- [14] SHEPP L A, LOGAN B F. The Fourier reconstruction of a head section [J]. IEEE Trans Nucl Sci, 1974, 21(3): 21-43.
- [15] 毛小渊. 二维CT图像重建算法研究[D]. 南昌: 南昌航空大学, 2016. MAO X Y. The research on the two-dimensional CT image reconstruction algorithm [D]. Nanchang: Nanchang Hangkong University, 2016.
- [16] 张顺利. ART算法几种重建模型的研究和比较[J]. 航空计算技术, 2005, 35(2): 39-41. ZHANG S L. Research and comparison on several reconstruction models of ART[J]. Aeronautical Computer Technique, 2005, 35(2): 39-41.
- [17] 范慧赞. CT图像滤波反投影重建算法的研究[D]. 西安: 西北工业大学, 2007. FAN H Y. Research on filtered backprojection algorithms of CT image reconstruction [D]. Xi'an: Northwestern Polytechnical University, 2007.
- [18] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[J]. Comput Sci, 2015, arXiv: 1505. 04597.
- [19] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//International Conference on Neural Information Processing Systems. Massachusetts: MIT Press, 2014: 2672-2680.
- [20] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Trans Image Process, 2004, 13(4): 600-612.

(编辑:谭斯允)