

DOI:10.19651/j.cnki.emt.2106123

基于多尺度上下文信息融合的条件生成对抗神经网络用于低剂量 PET 图像去噪^{*}

杨昆^{1,2} 杜瑀¹ 钱武侠¹ 薛林雁¹ 刘琨^{1,2} 卢闫晔³(1. 河北大学 质量技术监督学院 保定 071002; 2. 计量仪器与系统国家地方联合工程研究中心 保定 071002;
3. 北京大学 生物医学工程系 北京 100871)

摘要: PET 成像在信号采集过程中存在辐射暴露风险,在保证成像质量的前提下需要尽可能降低示踪剂的使用剂量。这将导致 PET 图像伪影及信噪比低等问题。本文提出一种基于多尺度上下文信息融合的条件生成对抗网络,用于优化 PET 图像重建,以提高 PET 图像的图像质量。通过对作为生成器的编码器-解码器网络的跳过连接进行重新设计,提出一种基于多尺度上下文信息融合的跳过连接,使解码器能够在解码过程中获取来自编码器更加丰富的语义特征。采用使生成器网络学习低剂量 PET 图像的噪声分布的策略,降低了网络的学习难度。在低剂量 PET 数据集上对所提出的网络进行评估,峰值信噪比为 29.948 ± 4.062 , 结构相似性系数为 0.926 ± 0.030 , 标准均方根误差为 0.395 ± 0.211 。相比于传统去噪算法 Non-Local Mean 和 Block-Matching 3D 以及 2 种深度学习方法 RED-CNN 和以 U-Net 为生成器的条件生成对抗网络,本文所提出的网络均取得了更加优越的性能表现。

关键词: 图像去噪;低剂量 PET 图像;条件生成对抗网络;深度学习

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 520

Conditional generative adversarial network based on multi-scale contextual information fusion for low-dose PET image denoising

Yang Kun^{1,2} Du Yu¹ Qian Wuxia¹ Xue Linyan¹ Liu Kun^{1,2} Lu Yanye³

(1. College of Quality and Technical Supervision, Hebei University, Baoding 071002, China; 2. National & Local Joint Engineering Research Center of Metrology Instrument and System, Baoding 071002, China; 3. Department of Biomedical Engineering, Peking University, Beijing 100871, China)

Abstract: There is a risk of radiation exposure in the signal acquisition process of PET images. Under the premise of ensuring the image quality, it is necessary to reduce the tracer dosage as much as possible. But this will cause artifacts in the PET image and lower signal-to-noise ratio. Propose a conditional generative adversarial network (cGAN) based on multi-scale contextual information fusion, which is used to optimize PET image reconstruction to improve the image quality of PET images. By redesigning the skip connection of encoder-decoder network as generator, proposes a skip connection based on multi-scale context information fusion, which enables the decoder to obtain richer semantic features from the encoder during the decoding process. Adopt the strategy of making the generator network learn the noise distribution of low-dose PET images, which reduces the learning difficulty of the network. The proposed network is evaluated in the low-dose PET dataset. The peak signal-to-noise ratio was 29.948 ± 4.062 , the structural similarity index was 0.926 ± 0.030 , and the normalized root mean square error was 0.395 ± 0.211 . Compared with the traditional denoising algorithm Non-Local Mean (NLM) and Block-Matching 3D (BM3D) and two deep learning methods RED-CNN and cGAN (U-Net is the generator), the proposed network is superior to the above four methods.

Keywords: image denoising; low dose PET image; conditional generative adversarial network; deep learning

0 引言

正电子发射断层扫描 (positron emission tomography,

PET) 是一种分子医学成像技术,可以在分子水平上将人体代谢过程可视化,已经成为临床试验和医疗实践中不可或缺的工具^[1]。为了获得高质量的用于诊断的 PET 图像,根

收稿日期:2021-03-24

* 基金项目:河北省自然科学基金(H2019201378)、基于深度学习的低剂量 PET 图像增强算法研究(XZJJ201917)项目资助

据患者的体重,通常给患者注射 5~10 mCi 剂量的示踪剂^[2]。来自示踪剂的辐射会增加人体的损伤风险,在实际临床应用中,常规的 PET 检查存在一定的局限性,如对于儿童或者孕妇等特殊群体,常规的 PET 检查将对其造成一定的辐射伤害,因此临床中一般不建议这几类特殊人群进行 PET 检查。为了降低辐射暴露损害风险,最直接有效的方法就是降低注射示踪剂的剂量,但是示踪剂剂量的降低会导致噪声和伪影的增加,降低 PET 图像的质量,影响医生对患者的诊断。因此利用低剂量 PET 图像来生成高质量的全剂量 PET 图像,在确保 PET 图像质量的同时减少辐射暴露是非常有意义的。

目前,研究者已经提出了一系列去噪方法用于自然图像和低剂量 PET 图像中^[3-7],如 Wang 等^[3]提出将先验知识与图像重建结合,将成像的物理信息直接引入到重建过程中,来生成高质量的全剂量 PET 图像。但是该方法存在计算量过大,以及需要获取其物理投影模型的问题,具有一定的局限性。除此之外也有许多基于图像重建后处理的方法被提出,如 Kang 等^[4]提出的基于随机森林回归的方法, Wang 等^[5]提出的基于半监督三倍字典学习的方法,以及 Wang 等^[6]提出的基于映射的稀疏表示的方法。但是这些方法都存在一个相同的问题,容易生成过于平滑的图像,这对医生的诊断产生一定的影响。

近年来,随着深度学习的兴起,尤其是卷积神经网络(convolution neural networks, CNN)的出现,其凭借自身强大的非线性特征提取的能力,在计算机视觉和医学成像领域得到了广泛的研究,在低剂量 PET 图像去噪领域, CNN 和生成对抗网络(generative adversarial networks, GAN)已经取得了显著的效果。

在 CNN 方面, Xu 等^[8]首次提出利用一种编码器解码器残差深度神经网络来实现在超低剂量(1/200 剂量)下重建标准剂量图像。Chen 等^[9]将低剂量的 PET 图像和多模态的磁共振(magnetic resonance, MR)序列作为输入训练一个深度卷积编码器-解码器网络,实现超低剂量(1/100 剂量)的 PET 图像到标准剂量 PET 图像的重建。一些研究者没有将深度学习作为一种后处理工具,而是选择将深度学习网络嵌入到迭代重建框架中,如 Gong 等^[10]将残差网络与 U-Net 网络相结合得到一个用于 PET 图像降噪的编码器-解码器网络,并将该网络引入迭代重建框架之中,来生成高信噪比的 PET 图像。Gong 等^[11]通过将深度神经网络 U-Net 与期望最大化算法(expectation-maximum, EM)的迭代更新步骤相结合,形成了一个基于 EM 算法的展开神经网络框架,用于 PET 图像重建,获得更好的 PET 图像。

在 GAN 方面, Ouyang 等^[12]通过预训练的方式训练了一个特定的网络作为 GAN 的感知损失函数,对生成的 PET 图像进行约束,能够从超低剂量(1/100 剂量)的 PET 图像中重建出高质量的全剂量 PET 图像。Wang 等^[13]研

究者则设计了一种基于三维的条件生成对抗网络(condition generative adversarial networks, cGAN),从对应的低剂量 PET 图像中生成高质量的全剂量 PET 图像。

虽然上述方法对于低剂量 PET 图像去噪都取得良好的效果,但是仍存在一定的局限性。对于涉及多模态信息融合的方法, PET 图像与 MR 或者 CT 不匹配可能会导致伪影和噪声的增加。目前用于低剂量 PET 图像去噪的深度学习算法大多数是基于编码器-解码器网络(如 U-Net^[14])和将编码器-解码器网络作为生成器的 GAN。传统的编码器-解码器网络中的跳过连接只是将来自编码器的高分辨率特征与其对应的解码器的语义信息特征融合起来,以提高最终生成图像的质量。这些结构的跳过连接仅仅考虑到编码器与解码器相对应尺度的特征融合,而忽略了相对于当前层次的更浅层或者更深层的编码器特征也将对解码过程产生影响,即编码器的多尺度上下文信息融合将有助于解码器在解码过程中生成更加精确的图像。

针对上述问题,本文以 cGAN 为主干网络,重新设计了作为生成器的编码器-解码器网络中的跳过连接,以此来改善传统编码器-解码器网络在解码过程无法获取更多来自编码器的语义信息的问题,提出了一种基于多尺度上下文信息融合的 cGAN 用于低剂量 PET 图像去噪工作,以提高在低剂量 PET 图像上的去噪能力,为低剂量 PET 图像用于临床实践,降低 PET 检查的辐射暴露风险做出一定的贡献。本文对 cGAN 和所设计的跳过连接传统编码器-解码器网络框架进行消融实验,以证明本文所设计的跳过连接能够使解码器在解码过程获取来自编码器更加丰富的语义信息和使用 cGAN 能够使生成的图片更加清晰,保留更多的细节信息。并将其与传统去噪算法 NLM^[15]和 BM3D^[16]、深度学习方法 RED-CNN^[17]和以 U-Net^[14]作为生成器的 cGAN(简称 cGAN-U-Net)进行对比,本文所提出的方法均优于上述 4 种,取得了更好的效果。

1 方 法

1.1 概 述

传统的编码器-解码器网络(如 U-Net^[14]),如图 1 所示,该网络通过卷积和池化的方式逐渐将图像折叠成高维特征,再通过多次上采样的方式将高维特征重构成高质量的全剂量 PET 图像,在解码过程中,通过跳过连接,将来自编码器的高分辨率特征和对应的解码器的语义特征相融合,进一步提高网络的性能。编码器-解码器网络凭借自身优越的性能以及在训练过程中对训练数据量要求小的优点,在低剂量 PET 图像去噪领域,得到了广泛的应用。

传统编码器-解码器网络的跳过连接只考虑了对应尺度的编码器和解码器特征的融合,忽略了其他尺度的编码器特征也将对当前尺度的解码过程产生影响。针对这一问题,本文重新设计了编码器-解码器网络的跳过连接,提出了一种基于多尺度上下文信息融合的 cGAN 用于低剂量

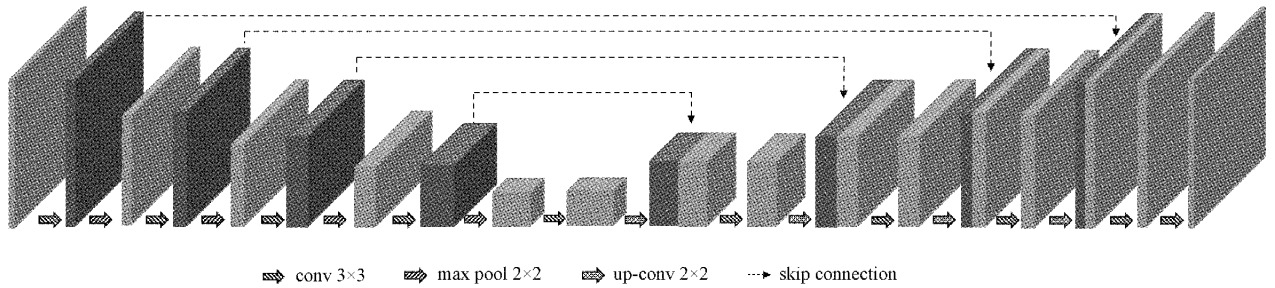


图 1 传统编码器-解码器网络框架

PET 图像去噪。与其他常规的深度学习去噪算法不同,本文没有采用让生成器直接学习生成高质量的 PET 图像的策略,而是选择让生成器去学习低剂量 PET 图像中的噪声分布,通过原始的低剂量 PET 图像减去其自身的噪声分布来获取高质量的 PET 图像,将高质量的 PET 图像送入鉴别器去鉴别。本文设计了一种多阶损失函数应用于所提出的框架,提出的总体网络框架如图 2 所示。

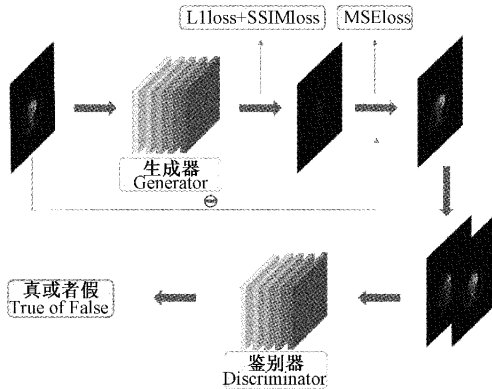


图 2 总体网络的功能框架

1.2 基于多尺度融合的跳过连接

本文设计了一种基于多尺度上下文信息融合的跳过

连接,其通过融合编码器不同尺度的特征信息来形成一个包含编码器所有尺度特征信息的语义特征,为解码器在解码过程中提供来自编码器的更加丰富和有效的语义信息,从而生成更加精确的图像。

如图 3 所示,在所设计的生成器网络中,由于编码器不同尺度的高分辨率特征有着不同的分辨率,通过使用反卷积的方式将其上采样到同一分辨率,通过串联的方式将其拼接在一起,形成一个张量。该张量通过一个卷积核为 3,步长为 1 的卷积层来得到一个包含编码器多尺度上下文信息的特征图 F_{ME} :

$$F'_{E_n} = ConvT^{(k,s)}(F_{E_n}) \quad (1)$$

$$F_{ME} = Conv^{(3,1)}(F_{E_1} \oplus F'_{E_2} \oplus F'_{E_3} \oplus F'_{E_4}) \quad (2)$$

其中, F_{E_n} 为不同尺度的编码器特征; F'_{E_n} 为通过反卷积上采样得到的分辨率一致的不同尺度的编码器特征; $ConvT^{(k,s)}$ 为卷积核为 k ,步幅为 s 的反卷积; n 代表不同尺度。在本文中, n 取 1、2、3、4(其中排序按照编码器网络由浅到深层),与其对应 k, s 的取值分别为 1、2、4、8 和 1、2、4、8。

在解码过程中,得到的包含编码器多尺度上下文信息的特征图 F_{ME} 与解码器不同尺度的语义特征相融合时,同样会出现相融合特征之间的分辨率不一致的问题。为了

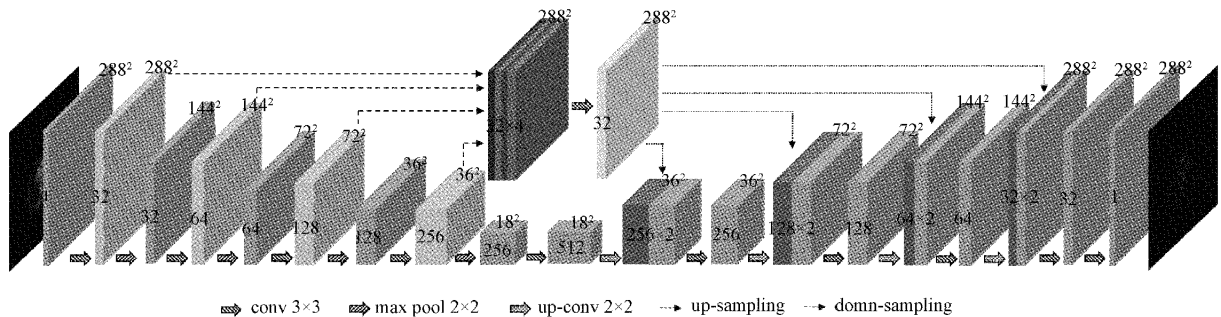


图 3 生成器的结构框架

解决这一问题,将得到的特征图 F_{ME} 分别通过不同的卷积层来降低其自身的分辨率,从而得到一系列与解码器不同尺度的特征相匹配的语义特征 F_{C_n} ,最后通过串联的方式将得到的语义特征 F_{C_n} 与解码器不同尺度的语义特征 F_{D_n} 相融合。基于多尺度上下文信息融合的跳过连接(MSSC)

可以表示为:

$$F_{C_n} = Conv^{(k,s)}(F_{ME}) \quad (3)$$

$$MSSC = F_{C_n} \oplus F_{D_n} \quad (4)$$

其中, $Conv^{(k,s)}$ 卷积核为 k ,步幅为 s 的卷积; n 代表不同尺度。在本文中, n 取 1、2、3、4(其中排序按照解码器

网络由浅到深层),与其对应 k, s 的取值分别为 8、4、2、1 和 8、4、2、1。

1.3 基于 patch 的鉴别器

对于所设计网络的鉴别器,采取 PatchGAN^[18]中所采用的鉴别器策略,该策略不是通过比较整幅图像中的所有像素来判断目标的真伪,而是试图去判断在整幅图像中 $N \times N$ 大小的 patch 的真伪。在产生高质量的结果的前提下, N 的取值可以比图像的尺寸小很多,从而使得 PatchGAN 的参数更少,运行速度更快,以及使其可以应用在任意大的图像上。

PatchGAN 的结构如图 4 所示,将生成的高质量 PET 图片与其对应的低剂量 PET 图片送入 3 层卷积核为 4,步长为 2,padding 为 1 的卷积层中,逐步增大其感知野,然后再使其通过一层卷积核为 4,步长为 1 的卷积层来映射为 1 维输出。

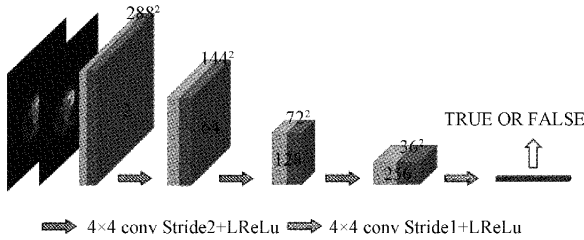


图 4 鉴别器的结构框架

1.4 多阶损失函数

与传统 GAN 不同,cGAN^[19] 还需要一个条件变量来对网络的训练进行约束。条件变量的限制,可以使网络最终生成的结果与目标更加相近,cGAN 的损失函数可以表示为:

$$\mathcal{L}_{cGAN}(D, G) = \mathbb{E}_{x, y}[\log D(x, y)] + \mathbb{E}_{x, z}[\log(1 - D(x, G(x, z)))] \quad (5)$$

其中,生成器 G 试图使式(5)损失函数最小化,而与之对抗的鉴别器 D 试图使其最大化。仅使用这个对抗性损失函数可以使生成的图像变得更加清晰,但是这并不意味着生成的图片与真实图片存在良好的相似性。为了提高低频信息准确性, $L1$ 距离以及结构相似性(SSIM)^[20] 被引入网络。其中 $L1$ 距离损失函数和结构相似性(SSIM)损失函数可以分别表示为式(6)和(7)。

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x, y, z}[\|y - G(x, z)\|_1] \quad (6)$$

$$\mathcal{L}_{SSIM}(G) = 1 - [l(G(x, z), y)^\alpha \cdot c(G(x, z), y)^\beta \cdot s(G(x, z), y)^\gamma] \quad (7)$$

其中, $l(G(x, z), y)^\alpha$ 是亮度比较; $c(G(x, z), y)^\beta$ 是对比度比较; $s(G(x, z), y)^\gamma$ 是结构比较,并且 $\alpha > 0, \beta > 0, \gamma > 0$ 。由于仅考虑到生成的噪声与真实噪声分布之间的区别,而忽略了最终生成的 PET 图像和真实全剂量 PET 图像之间的区别,因此在最终生成的 PET 图像和真实全剂量 PET 图像之间引入 MSE 损失函数来对其进行约束,

MSE 损失函数可以表示为:

$$\mathcal{L}_{MSE}(G) = \frac{1}{N} \sum_{i=1}^N \|Y_i - X_i\|_2 \quad (8)$$

最终的多阶损失函数可以表示为:

$$G^* = \underset{G}{\operatorname{argmin}} \max_D \mathcal{L}_{cGAN}(D, G) + \lambda_1 \mathcal{L}_{L1}(G) + \lambda_2 \mathcal{L}_{SSIM}(G) + \lambda_3 \mathcal{L}_{MSE}(G) \quad (9)$$

其中, $\mathcal{L}_{cGAN}(D, G)$ 和 $\mathcal{L}_{MSE}(G)$ 是作用于低剂量 PET 图像和真实全剂量 PET 图像之间, $\mathcal{L}_{L1}(G)$ 和 $\mathcal{L}_{SSIM}(G)$ 是作用于生成的噪声分布和真实噪声分布之间,并且 λ_1, λ_2 和 λ_3 的取值均为 1 000。

2 实验研究

2.1 概述

为了验证和评估所提出的方法,本文用 37 例全身 PET 数据进行了此项实验,将所得结果与传统去噪算法 NLM、BM3D 以及 2 种深度学习算法 RED-CNN、cGAN-UNet 进行对比,实验结果表明本文所提出的方法实现了更加优越的性能。

2.2 数据集

数据集来源于 37 例人体的全身 PET 数据,由北京锐视康科技发展有限公司提供,数据中不包含个人身份等隐私信息。每一例患者数据包括 326~477 张 2 D 径向切片数据,数据分辨率为 288×288 。通过对 PET 原始数据进行随机采样以降低剂量来模拟临床场景下的低剂量示踪剂注射的情况,并采用与全剂量 PET 图像重建时相同的重建参数对数据进行重建,包括所有的物理校正,如衰减、散射和随机校正,以确保二者的可比较性,从而获得对应的 1/10 剂量的 PET 图像。37 例全身 PET 数据分为 26 例和 11 例分别作为训练集和测试集。

2.3 实验细节

所有的实验是在一个装有 NVIDIA RTX2080Ti 11G * 8 的服务器上进行的。深度学习算法均是基于 PYTORCH 实现,训练的所有网络均使用 ADAM 优化器,其动量参数 $\beta_1 = 0.5, \beta_2 = 0.999$ 。所提出的网络和 cGAN-UNet 的批处理大小(BatchSize)设置为 64,RED-CNN 的批处理大小设置为 16。

对于学习率的设定 RED-CNN 的学习率设置为 2×10^{-4} 。为了使 cGAN 能够更加稳定的训练,实验采用了 two-timescale update rule(TTUR)策略^[20],在训练过程中给予生成器和鉴别器不同的学习速率,以平衡二者的训练,对所提出的网络的学习率设置为生成器学习率 1×10^{-4} ,鉴别器学习率 2×10^{-4} ,对于 cGAN-UNet,生成器学习率为 2×10^{-4} ,鉴别器学习率为 2×10^{-3} 。实验使用了学习率线性衰减策略,对于 RED-CNN 和 cGAN-UNet,训练时间均为 200 个 epoch,其中前 50 个 epoch 的学习率不变,当 50 个 epoch 之后,随着 epoch 的增加,学习率呈线性下降。对于所提出的网络,训练时间为 300 个 epoch,其中前

20 个 epoch 的学习率不变,当 20 个 epoch 之后,随着 epoch 的增加,学习率呈线性下降。

2.4 评价指标

选取峰值信噪比 (peak signal-to-noise ratio, PSNR)、结构相似性系数 (structural similarity index, SSIM)^[21] 和标准均方根误差 (Normalized root mean square error, NRMSE) 这 3 个指标来进行图像质量评估。

其中峰值信噪比、结构相似性系数和标准均方根误差可以表示为式(10)、(11)和(12)。

$$PSNR = 20 * \log_{10} \left(\frac{MAX}{MSE} \right) \tag{10}$$

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{11}$$

$$NRMSE = \sqrt{\frac{\sum_{i=1}^N \sum_{j=1}^M (x_{ij} - y_{ij})^2}{\sum_{i=1}^N \sum_{j=1}^M y_{ij}^2}} \times 100 \tag{12}$$

其中, μ_x 和 μ_y 分别代表 x 和 y 的平均值; σ_{xy} 代表 x 和 y 的协方差; σ_x 和 σ_y 分别代表 x 和 y 的方差; MAX 代表图像中数值取值的最大值; c_1 和 c_2 是 2 个变量来稳定当弱分母的除法; MSE 代表图片二者之间的绝对均方误差。对于 PSNR 和 SSIM 二种评价指标,当其值越大时,表示其图片质量越高,对于 NRMSE 来说,当其值越小,表示图片

质量越高。

2.5 消融实验

为了验证所设计的基于多尺度上下文信息融合的跳过连接的优越性能以及 cGAN 在图像去噪方面的优越性能,本文在不改变其他条件的情况下,不使用 cGAN 作为主体网络,仅使用改进的 U-Net 作为主体网络,与 UNet 和所提出的网络进行对比,来探究所设计的基于多尺度上下文信息融合的跳过连接的有效性和条件生成对抗网络在图像去噪方面的作用。

消融实验的定量结果如表 1 所示,定性结果如图 5 所示。相比于 U-Net 网络,本文所提出的改进的 U-Net 在 3 种评价指标 PSNR、SSIM、NRMSE 上均取得了更好的结果。在 PSNR 和 SSIM 上分别提升了 0.117 dB 和 0.004,在 NRMSE 上降低了 0.011,有效证明了本文所设计的基于多尺度上下文信息融合的跳过连接能够使解码器在解码过程中获取来自编码器的更加丰富的语义特征信息,从而提升网络的性能。而定性结果中所表现的差异性不大,本文认为这是由于 UNet 和改进的 UNet 均存在生成图片过于模糊的现象,从而导致丢失许多的细节信息。在接下来的实验结果中与 cGAN-UNet 的对比中,将进一步体现所设计的基于多尺度上下文信息融合的跳过连接的优越性能。

表 1 所提出网络上的消融实验的定量比较

模型	是否采用 cGAN	PSNR	SSIM	NRMSE
U-Net	否	29.309±4.154	0.919±0.033	0.436±0.258
改进的 U-Net	否	29.426±3.876	0.923±0.031	0.425±0.233
所提出的网络	是	29.948±4.062	0.926±0.030	0.395±0.211

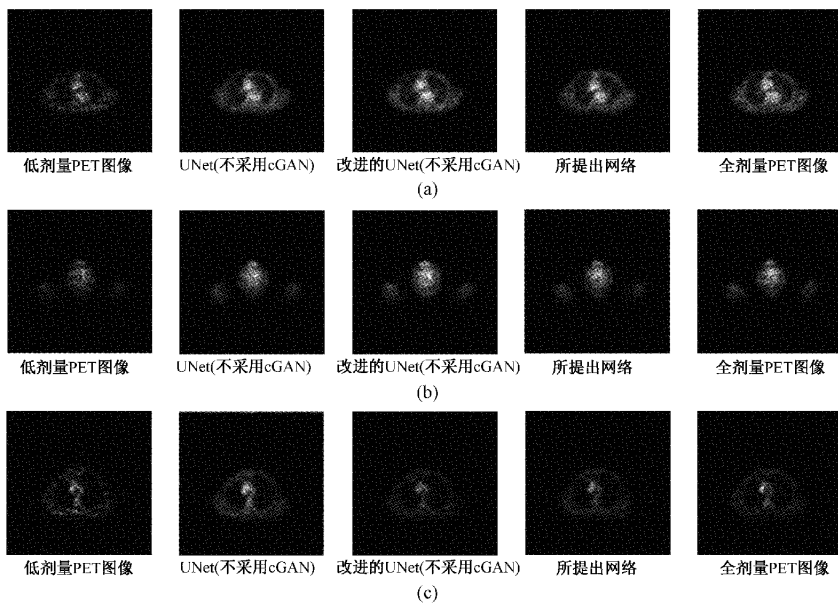


图 5 所提出网络上的消融实验的定性比较

相比于改进的 U-Net(不采用 cGAN),所提出的网络在 3 种评价指标 PSNR、SSIM、NRMSE 上均取得了更好的结果,在 PSNR 和 SSIM 上分别提升了 0.522 dB 和 0.003,在 NRMSE 上降低了 0.03。并且从定性结果中可以看出,使用 cGAN 能够进一步提升生成的 PET 图像的质量,所生成的 PET 图像拥有更好的细节信息和结构信息,而没有使用 cGAN 所得到的结果存在过于模糊的现象,导致丢失大量细节信息。如图 5(b)中,相比于改进的 U-Net(不采用 cGAN)的结果,使用 cGAN 所得到的结果,在边缘结构上更加接近于全剂量的 PET 图像,以及在其内部拥有更多细节信息。这是因为 cGAN 中加入了条件变量对网络进行了约束,以及采用的对抗损失函数,使生成器和鉴别器在不断对抗中得到优化,从而使生成的图像更加清晰。从定量结果和定性结果中,可以看出使用 cGAN 能够有效地提升网络的去噪性能,使其生成的 PET 图像更加清晰。

2.6 实验结果

本文提出的网络与传统算法 NLM、BM3D 和深度学习算法 RED-CNN 和 cGAN-UNet 在测试集上进行了对比。其定量结果如图 6 所示,具体定量比较如表 2 所示,相比于传统算法 NLM 和 BM3D 以及深度学习算法 RED-CNN 和 cGAN-UNet,实验中使用的网络在峰值信噪比(PSNR)、结构相似性(SSIM)以及标准均方根误差(NRMSE)上均取得了更好的结果。在 PSNR 评价指标上分别提升了 1.123、1.050、0.551 和 0.380 dB,在 SSIM 评价指标上分别提升了 0.008、0.016、0.001 和 0.006,在 NRMSE 评价指标上分别降低了 0.029、0.025、0.035 和 0.015。尽管在 SSIM 评价指标上相较于 RED-CNN 算法仅有微弱的提升,这是因为 RED-CNN 网络中加入了残差连接,使其能够更好的学习到 PET 图像的结构特征。但是该算法也存在明显的问题,其网络训练结果存在过于平滑和缺少细节的问题。

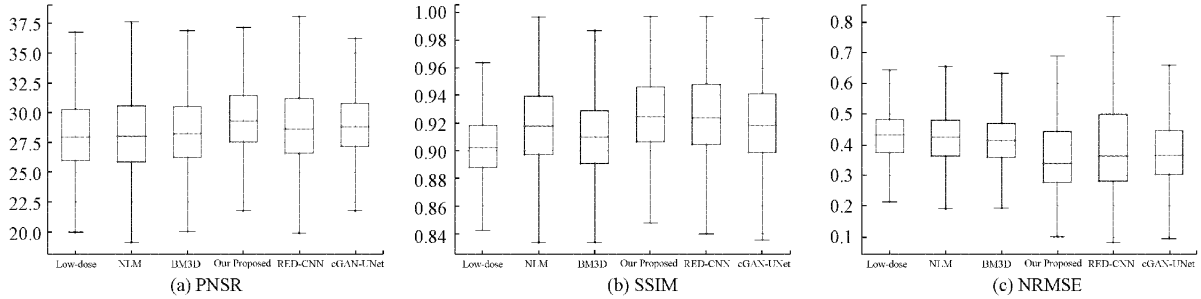


图 6 不同模型间去噪结果的量化对比

表 2 所提出网络与 NLM、BM3D、RED-CNN 和 cGAN-UNet 的定量对比

模型	PSNR	SSIM	NRMSE
Low-dose(1/10 dose)	28.616 ± 4.238	0.906 ± 0.029	0.433 ± 0.105
NLM	28.825 ± 4.603	0.918 ± 0.034	0.424 ± 0.105
BM3D	28.898 ± 4.207	0.910 ± 0.034	0.420 ± 0.108
RED-CNN	29.397 ± 4.708	0.925 ± 0.033	0.430 ± 0.244
cGAN-UNet	29.568 ± 4.193	0.920 ± 0.032	0.410 ± 0.213
本文提出的网络	29.948 ± 4.062	0.926 ± 0.030	0.395 ± 0.211

本文对实验中得到的结果进行了定性评估,如图 7 所示,传统算法 NLM 和 BM3D 的结果相较于 Ground Truth(全剂量的 PET 图像),其丢失大量的细节信息。这是由于传统算法 NLM 和 BM3D 都是通过以图像块为单位在图片中寻找与其相似的区域的方式来进行滤波,只是 BM3D 算法相较于 NLM 算法更加复杂一些。但是这 2 种算法都只能去除图片中存在的噪声和保留图片原本的结构信息,而无法补充低剂量 PET 图像相较于全剂量 PET 图像所缺失的细节信息。相比于深度学习算法 RED-CNN,本文提出的网络拥有更好的细节信息,而 RED-CNN 存在过于平滑的现象,从而丢失很多细节信息。相较于深度学习算法

cGAN-UNet,本文提出的网络的结果拥有更好的结构信息和更多的细节信息,更加接近于 Ground Truth。如在图 7(a)中,相较于 cGAN-UNet,所提出的网络在内部拥有更好的细节信息,以及更加平滑的边界。以及在图 7(c)中,相较于 cGAN-UNet,所提出的网络在图片内部拥有更好的结构信息,和在边界上更加接近 Ground Truth。能够得到这样的结果,主要是得益于所设计的基于多尺度上下文信息融合的跳过连接,其通过融合编码器所有尺度的特征信息,形成一个包含编码器所有尺度特征信息的语义特征,使解码器能够在解码过程中获取来自编码器的更加丰富的语义特征。并且凭借 cGAN 优秀的去噪能力,以及生

成器学习低剂量 PET 图片中噪声分布的策略,使所提出的网络拥有良好的去噪性能。

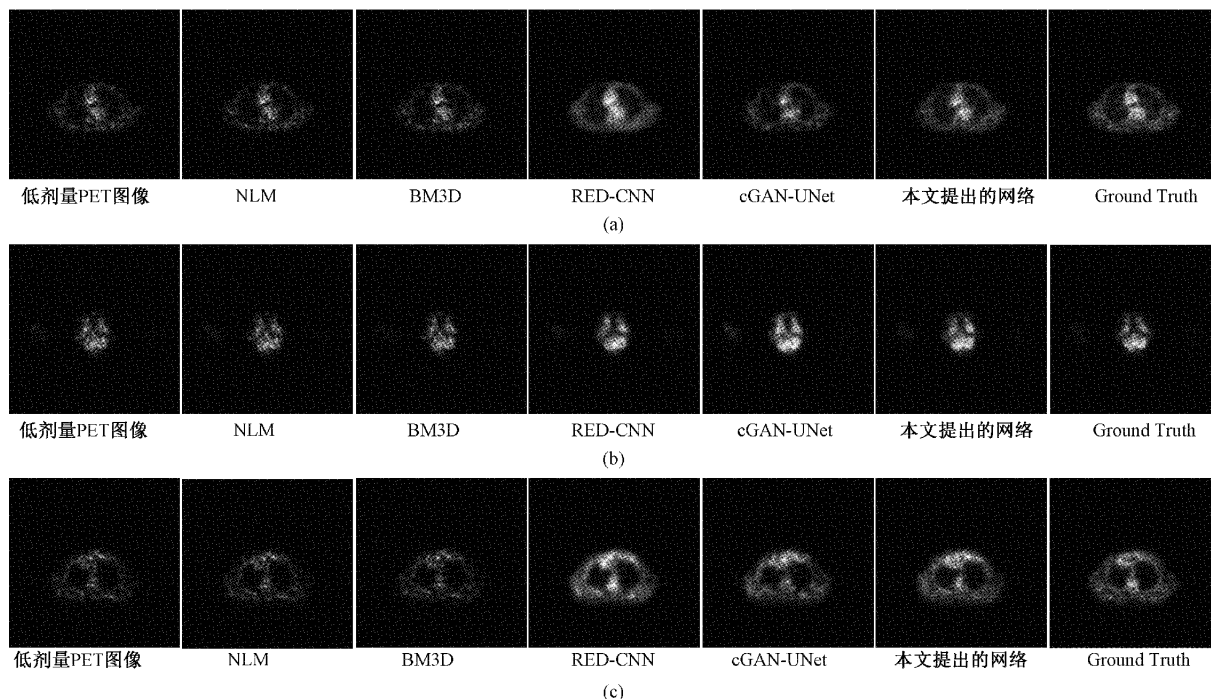


图 7 所提出的网络与 NLM、BM3D、RED-CNN 和 cGAN-UNet 的定性对比

3 结 论

本文设计一种基于多尺度上下文信息融合的条件生成对抗神经网络用于低剂量 PET 图像去噪工作,该网络重新设计了作为生成器的编码器-解码器网络中的跳过连接,其通过融合编码器所有尺度的特征信息,从而形成一个包含编码器所有尺度特征信息的语义特征,使解码器能够在解码过程中获取来自编码器的更加丰富的语义特征信息,进一步提升了网络的去噪性能。通过使用与传统深度学习去噪网络不同的训练方法,即生成器来学习低剂量的 PET 图像中噪声分布,而不是直接学习生成全剂量的 PET 图像,从而降低了学习难度。本文将所提出的网络与 2 种传统算法 NLM、BM3D 和 2 种深度学习算法 RED-CNN、cGAN-UNet 相比较,取得了更好的性能表现,能够生成更加精确的 PET 图像。

参考文献

- [1] LIU Z, CHEN H, LIU H. Deep learning based framework for direct reconstruction of PET images[C]. International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Cham, 2019: 48-56.
- [2] ZHOU L, SCHAEFFERKOETTER J D, THAM I W K, et al. Supervised learning with cycleGAN for low-dose FDG PET image denoising [J]. Medical Image Analysis, 2020, 65: 101770.
- [3] WANG C, HU Z, SHI P, et al. Low dose PET reconstruction with total variation regularization[C]. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2014: 1917-1920.
- [4] KANG J, GAO Y, SHI F, et al. Prediction of standard-dose brain PET image by using MRI and low-dose brain [^{18}F] FDG PET images [J]. Medical Physics, 2015, 42(9): 5301-5309.
- [5] WANG Y, MA G, AN L, et al. Semisupervised tripled dictionary learning for standard-dose PET image prediction using low-dose PET and multimodal MRI[J]. IEEE Transactions on Biomedical Engineering, 2016, 64(3): 569-579.
- [6] WANG Y, ZHANG P, AN L, et al. Predicting standard-dose PET image from low-dose PET and multimodal MR images using mapping-based sparse representation[J]. Physics in Medicine & Biology, 2016, 61(2): 791.
- [7] 荣霞,薛伟,朱继超.一种新的小波阈值函数在图像去噪中的应用[J].电子测量技术,2016,39(5):84-88.
- [8] XU J, GONG E, PAULY J, et al. 200x low-dose PET reconstruction using deep learning [J]. ArXiv Preprint 2017, ArXiv:1712.04119.
- [9] CHEN K T, GONG E, DE CARVALHO MACRUZ F B, et al. Ultra-low-dose ^{18}F -florbetaben amyloid PET imaging using deep learning with multi-contrast MRI

- inputs[J]. *Radiology*, 2019, 290(3): 649-656.
- [10] GONG K, GUAN J, KIM K, et al. Iterative PET image reconstruction using convolutional neural network representation [J]. *IEEE Transactions on Medical Imaging*, 2018, 38(3): 675-685.
- [11] GONG K, WU D, KIM K, et al. Emnet: An unrolled deep neural network for pet image reconstruction[C]. *Medical Imaging 2019: Physics of Medical Imaging*, International Society for Optics and Photonics, 2019, 10948: 1094853.
- [12] OUYANG J, CHEN K T, GONG E, et al. Ultra-low-dose PET reconstruction using generative adversarial network with feature matching and task-specific perceptual loss [J]. *Medical Physics*, 2019, 46(8): 3555-3564.
- [13] WANG Y, YU B, WANG L, et al. 3D conditional generative adversarial networks for high-quality PET image estimation at low dose[J]. *Neuroimage*, 2018, 174: 550-562.
- [14] RONNEBE R, RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]. *International Conference on Medical Image Computing and Computer-assisted Intervention*, Springer, Cham, 2015: 234-241.
- [15] BUADES A, COLL B, MOREL J M. A non-local algorithm for image denoising [C]. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, IEEE, 2005, 2: 60-65.
- [16] DABOV K, FOI A, KATKOVNIK V, et al. Image denoising by sparse 3-D transform-domain collaborative filtering[J]. *IEEE Transactions on Image Processing*, 2007, 16(8): 2080-2095.
- [17] 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.
- [18] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-image translation with conditional adversarial networks [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 1125-1134.
- [19] MIRZA M, OSINDERO S. Conditional generative adversarial nets [J]. *ArXiv Preprint* 2014, ArXiv:1411.1784.
- [20] MIYATO T, KATAOKA T, KOYAMA M, et al. Spectral normalization for generative adversarial networks[J]. *ArXiv Preprint* 2018, ArXiv:1802.05957.
- [21] WANG Z, SIMONCELLI E P, BOVIK A C. Multiscale structural similarity for image quality assessment[C]. *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers*, 2003, 2: 1398-1402.

作者简介

杨昆,博士,教授,主要研究方向为基于人工智能的医学图像处理技术。

E-mail: yangkun@hbu.edu.cn