

基于生成对抗网络和噪声水平估计的低剂量CT图像降噪方法

张 雄* 杨琳琳 上官宏 韩泽芳 韩兴隆 王安红 崔学英

(太原科技大学电子信息工程学院 太原 030024)

摘 要: 生成对抗网络(GAN)用于低剂量CT(LDCT)图像降噪具有一定的性能优势, 成为近年CT图像降噪领域新的研究热点。不同剂量的LDCT图像中噪声和伪影分布的强度发生变化时, GAN网络降噪性能不稳定, 网络泛化能力较差。为了克服这一缺陷, 该文首先设计了一个编解码结构的噪声水平估计子网, 用于生成不同剂量LDCT图像对应的噪声图, 并用原始输入图像与之相减来初步抑制噪声; 其次, 在主干降噪网络中, 采用GAN框架, 并将生成器设计为多路编码的U-Net结构, 通过博弈对抗实现网络结构优化, 进一步抑制CT图像噪声; 最后, 设计了多种损失函数来约束不同功能模块的参数优化, 进一步保障了LDCT图像降噪网络的性能。实验结果表明, 与目前流行算法相比, 所提出的降噪网络能够在保留LDCT图像原有重要信息的基础上, 取得较好的降噪效果。

关键词: 图像降噪; 生成对抗网络; 低剂量CT; U-Net; 噪声水平

中图分类号: TN911.73; TP391

文献标识码: A

文章编号: 1009-5896(2021)08-2404-10

DOI: [10.11999/JEIT200591](https://doi.org/10.11999/JEIT200591)

A Low-Dose CT Image Denoising Method Based on Generative Adversarial Network and Noise Level Estimation

ZHANG Xiong YANG Linlin SHANGGUAN Hong HAN Zefang

HAN Xinglong WANG Anhong CUI Xueying

(School of Electronic Information Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China)

Abstract: Generative Adversarial Network (GAN) for Low-Dose CT (LDCT) image noise reduction has certain performance advantages, and has become a new research hot field of CT image noise reduction in recent years. When the intensity of noise and artifact distribution changes in LDCT images of different doses, the noise reduction performance of GAN network is unstable, and the generalization ability of the network is low. In order to overcome these shortcomings, this paper first designs a noise level estimation subnet with a encoder-decoder structure to generate the noise maps corresponding to LDCT images with different doses, which is subtracted from the original input image to initially suppress the noise; Secondly, the backbone of the noise reduction network is designed as a multi-coded U-Net structure that is optimized through game confrontation to suppress further CT image noise; Finally, a variety of loss functions are designed to constrain the parameter optimization of each function modules, thus to guarantee further the performance of the LDCT image noise reduction network. Experimental results show that compared with current popular algorithms, the noise reduction network proposed in this paper can achieve a better noise reduction on the basis of retaining the original important information of LDCT images.

Key words: Image denoising; Generative Adversarial Network (GAN); Low-Dose CT (LDCT); U-Net; Noise level

收稿日期: 2020-07-17; 改回日期: 2021-02-03; 网络出版: 2021-03-01

*通信作者: 张雄 zx@tyust.edu.cn

基金项目: 国家青年科学基金(62001321), 山西省高等学校科技创新项目(2019L0642), 山西省自然科学基金(201901D111261)

Foundation Items: The Natural Science for Youth Foundation (62001321), The Scientific and Technological Innovation Programs of Higher Educations Institutions in Shanxi (2019L0642), The Natural Science Foundation of Shanxi Province (201901D111261)

1 引言

近年, X射线计算机断层成像(Computed Tomography, CT)技术已成为医学诊疗中不可或缺的成像手段^[1]。重复CT扫描检查会使患者暴露于过量的X射线辐射伤害中, 进而导致由辐射诱发的癌变、代谢异常、白血病和其他遗传疾病等发生的概率增加^[2]。因此, 在临床上倡导使用ALARA(As Low As Reasonably Achievable)原则^[3], 即在保证CT图像质量满足诊断需求的同时, 将X射线对人体的辐射伤害降至最低。但是, 降低辐射剂量常导致重建图像出现斑驳噪声和不稳定的条形伪影, 进而影响临床诊断的准确性。自从Naidich等人^[4]首次提出低剂量CT(Low-Dose CT, LDCT)的概念以来, 针对如何有效地抑制LDCT图像中噪声与伪影的问题, 国内外学者不断优化成像算法的设计方案, 已在系统建模、图像重建和后处理等方面做了大量卓越的研究工作。其中, LDCT图像后处理方法不依赖于投影数据, 直接对现有CT图像进行处理, 可移植性强, 不需要对现有设备进行改进或更换, 便于使用和推广。传统后处理算法, 并没有完全解决输出图像过平滑或引入新噪声的问题, 且算法鲁棒性和泛化能力有待提高。

近年, 深度卷积神经网络(Convolutional Neural Networks, CNN)快速发展, 为医学成像领域提供了新的研究思路并展现出巨大潜力。从网络框架的构建和目标函数的优化两个角度出发, 学者提出了一些基于CNN的LDCT图像降噪方法, 研究结果表明不同的算法框架(例如2D CNN^[5], 3D CNN^[6], 残差编码器/解码器CNN^[7]和级联CNN^[8]等)和目标函数(例如均方误差^[5,8]、对抗损失^[9]和感知损失^[10]等)可能会对网络性能产生不同影响。降噪框架的设计决定了网络的复杂度, 而损失函数则会影响网络的训练精度。得益于CNN强大的特征学习与特征映射能力, 基于CNN的LDCT图像降噪方案可以取得一定的效果, 然而CNN的特征提取能力受数据集、硬件资源及运行时间等制约, 只有当数据集较大、网络层数足够深且迭代次数足够多的情况下网络才能取得比较好的降噪效果。实际应用中, 很难构建足够大的能够满足CNN充足训练的数据集, 数据集较小常会引起训练不充分、网络欠拟合等问题。为了解决这个问题, 学者做了许多努力和尝试。Goodfellow等人^[11]提出的生成对抗网络(Generative Adversarial Networks, GAN)就是其中的典型代表。GAN利用了生成器和判别器的博弈学习机制, 在图像生成^[12]、超分辨率重构^[13]、图像分割^[14]以及图像降噪^[15]等领域

都有相对优异的表现。He等人^[16]提出的条件生成对抗网络(conditional GAN, cGAN)在一定程度上解决了GAN网络生成数据不稳定的问题, 该网络在生成器与判别器中都加入了额外的约束条件, 能够生成满足约束条件的图像。Isola等人^[17]在cGAN的基础上提出了pix2pix模型, 该模型已成功应用于实现不同形式图像(如渐变图像及含内容缺失的图像等)之间的相互转换。此外, 针对传统GAN降噪框架模型泛化能力较差的问题, 在自然图像盲降噪领域有一些成功尝试。例如, Zhang等人^[18]通过对噪声水平随空间位置变化而产生的强度变化进行估计, 提出了一种采用单个网络实现对图像中噪声水平不同的位置进行区别处理的方法。

综上所述, 基于GAN框架的降噪网络仍然存在网络训练过程不稳定、泛化能力较差和特征提取不充分的问题。在训练过程中, 本文采用了多种损失函数对各个子网络的输出结果进行针对性约束, 用来保障网络训练的稳定性; 在降噪网络结构中, 增加了一个噪声水平估计子网络, 用来估计不同剂量LDCT图像的噪声伪影特征图, 从而提高降噪网络的泛化能力; 在GAN网络主通道中, 本文设计了多路编码的U-Net结构, 用以提高生成器的多尺度特征表达能力, 进而解决降噪网络特征提取不充分的问题。

2 本文方法

假设 $\mathbf{X} \in R^{w \times h}$ 表示大小为 $w \times h$ 的给定LDCT图像, $\mathbf{Y} \in R^{w \times h}$ 表示相应的常规剂量CT(Normal-Dose CT, NDCT)图像。一般地, 它们之间的关系可以表示为

$$\mathbf{X} = T(\mathbf{Y}) \quad (1)$$

其中, $T: R^{w \times h} \rightarrow R^{w \times h}$ 表示以非线性的方式将NDCT图像 \mathbf{Y} 退化为与之对应的LDCT图像 \mathbf{X} 的降质过程。降噪模型的目的是学习 $\mathbf{X} \rightarrow \mathbf{Y}$ 的映射关系。若 \mathbf{A} 表示噪声和伪影, 则

$$\mathbf{A} = \mathbf{X} - \mathbf{Y} \quad (2)$$

实际中, 希望设计一种可以有效抑制噪声伪影的方法, 使所获取图像的质量尽可能接近NDCT。本文以生成对抗思想作为指导, 将LDCT图像伪影抑制问题分为两个子问题: 伪影噪声水平估计与伪影噪声抑制。具体地, 提出了一种基于GAN的LDCT图像降噪网络, 如图1所示, 其结构包括噪声水平估计子网络和主干降噪网络两部分。一方面, 为了提高网络对噪声和伪影分布存在明显差异的不同剂量LDCT的处理能力, 将噪声水平估计子网络设计为编解码结构, 编码部分用来检测噪声水平; 解码部分用来生成能够反映LDCT噪声水平的噪声图

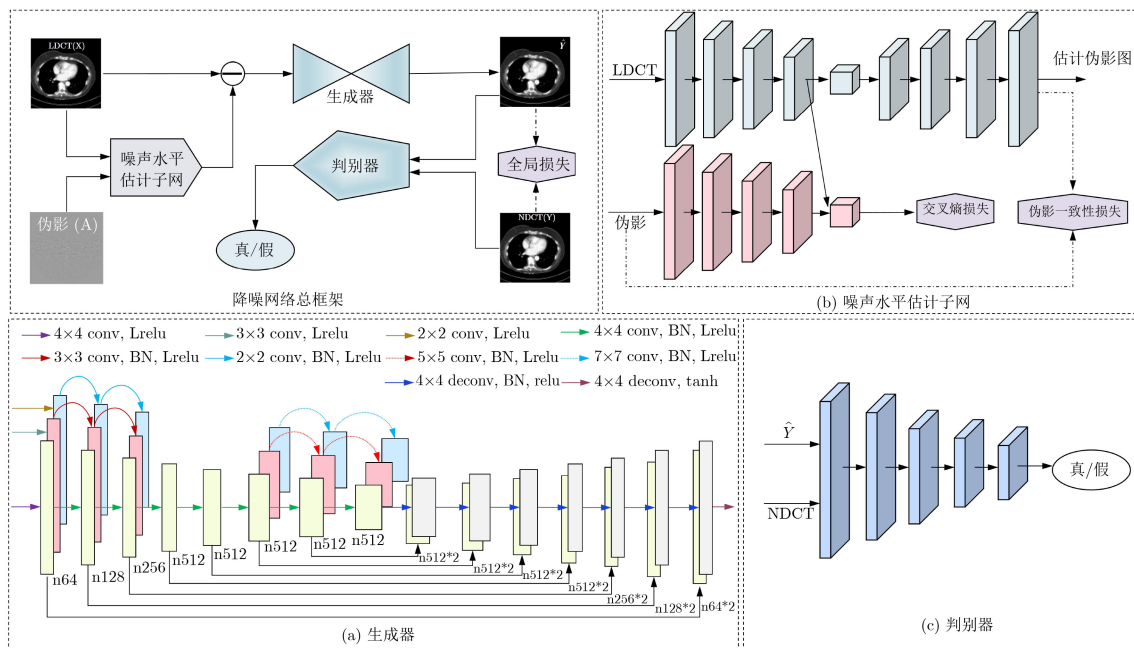


图1 本文降噪网络整体框架

像。另一方面,为了提高降噪图像的质量,将主干降噪网络设计为能够进行内部自我优化的GAN框架,生成器G采用多路编码U-Net结构;判别器D采用Markovian判别器^[17]结构。总的来说,采用噪声水平估计子网络检测不同剂量LDCT被噪声污染的程度,将LDCT与生成的伪影图相减获得的初步降噪图输入主干降噪网络,采用多路编码U-Net生成子网络进一步抑制噪声伪影,两者相辅相成,最终生成满足需求的降噪图像。此外,对损失函数进行了针对性设计:最小二乘损失及像素级L1范数用来保证网络的稳定性;噪声等级一致性损失用来优化噪声水平估计子网络;伪影一致性损失用来优化上采样网络。

2.1 降噪主干网络

(1)生成器G:传统U-Net网络的编码器采用8层卷积操作(卷积核均为 4×4)来提取输入图像特征,该操作虽然在诸如图像识别等任务上实现了一定的效果,但是并不能充分提取图像的特征。目前主流方法通过加深网络结构来解决这个问题,但随着网络的加深,模型复杂度也不断提升。众所周知,卷积神经网络的浅层主要提取视觉特征,而深层主要提取图像语义信息。考虑到这一因素,本文对传统U-Net网络进行了改进,提出了一种多路编码U-Net结构(如图1(a)所示)。该设计不仅能够提高网络对LDCT的特征提取能力,而且能够解决传统神经网络基本卷积单元的冗余问题,可以在无损性能的前提下,简化模型复杂度、提升计算速度。具体地,在U-Net网络编码端,前3个卷积层分3路

分别采用大小为 2×2 、 3×3 和 4×4 的卷积核对图像进行卷积操作,将卷积后的特征映射图级联并入第4个卷积层。前3层网络均采用尺寸较小的卷积核,有助于提取更为准确的LDCT图像信息。编码器的第4、第5层采用大小为 4×4 的卷积核进行卷积操作。在编码器的后3个卷积层即第6、第7、第8层分3路分别采用大小为 4×4 、 5×5 和 7×7 的较大卷积核以更广的视野来提取LDCT图像的特征,该设计能更好地捕获图像的全局信息。通过在编码端分多路分尺度对LDCT图像进行不同的卷积操作,一方面可以提取图像的多尺度特征,为后面的卷积层提供更加丰富的特征信息;另一方面能够适应卷积神经网络深浅层提取不同特征的特点,增加网络的感受野,在去除伪影和噪声这些细节信息方面具有较大优势。编码完成后,将编码端所提取的不同尺度的特征,通过旁路连接操作并入解码端的对应卷积层,经8层反卷积(卷积核均为 4×4)操作后恢复出降噪后的图像。

(2)判别器D:与G相对抗的D对生成图像的质量也起着非常重要的作用,鉴于CT图像包含丰富的细节信息,图像敏感度较高,在设计降噪网络时,需同时考虑伪影噪声的去除效果和图像原有细节的保留程度。本文判别器如图1(c)所示,前4层采用conv+BN+LReLU操作来提取图像不同层次的特征,其中卷积核大小为 4×4 、滑动步长为2。第5层通过conv+sigmoid函数来区分输入图像的真假,与生成器交替迭代训练,更好约束生成器,使得最终生成高质量的降噪图像。

2.2 噪声水平估计子网络

一般地,不同剂量LDCT图像受噪声和伪影污染的程度不同,通过训练噪声水平估计子网络,可以将不同LDCT图像的噪声水平估计出来。与传统降噪网络相比,这种设计使得网络泛化能力更强。如图1(b)所示,本文所设计的噪声水平估计子网采用编解码结构,分为噪声水平检测支路和上采样支路,其中,噪声水平检测支路包含两个通道的噪声水平检测模块,两个检测模块结构相同,均为编码网络,采用一个5层网络设计,第1层和最后1层均使用大小为 4×4 、滑动步长为2的卷积核进行卷积操作,中间3层为多尺度Inception模块^[19],Inception结构能够在增加感受野的同时,不增加网络的计算复杂度。将LDCT图像和用来约束检测模块的理想伪影图像分两路输入两个通道,在每个通道内经过前4层操作后,分别提取出LDCT图像的噪声水平特征图A和伪影图像的噪声水平特征图B。在噪声等级一致性损失函数约束下,通过不断迭代优化,使得A和B越来越接近。在解码端,将特征图A输入一个卷积核大小为 1×1 的卷积层,进行卷积操作后,送入上采样网络,进行4次上采样操作后生成一幅代表噪声水平的残差图像,该残差图像的质量由伪影一致性损失来约束。

2.3 损失函数

为了提高GAN降噪网络的性能,本文在网络的不同模块采用了与其功能相对应的损失函数。具体地,为了约束噪声水平估计子网的参数优化,本文利用噪声估计子网中的编码器模块估计输入图像的噪声等级,并利用噪声等级检测的交叉熵损失来设计噪声等级一致性损失,其定义为

$$L_C = E_X [\lg F(\mathbf{X})] + E_{\mathbf{X},\mathbf{Y}} [\lg (1 - F(\mathbf{A}))] \quad (3)$$

其中, E 表示求期望, F 表示噪声水平估计子网中的噪声检测网络, \mathbf{X} 表示输入的LDCT图像, \mathbf{A} 表示理想的伪影图像, \mathbf{Y} 表示NDCT图像,该损失函数的作用是使网络提取的LDCT噪声水平特征图尽可能接近真实伪影图像的噪声水平特征图。为了保证最后生成的伪影图像尽可能接近真实的伪影图像,本文采用了与伪影相关的损失函数来保障噪声水平估计子网的性能

$$L_A = E_{\mathbf{X},\mathbf{Y}} \left[\left\| \mathbf{A} - \hat{\mathbf{A}} \right\|_{L_1} \right] \quad (4)$$

$\hat{\mathbf{A}}$ 为噪声水平估计子网估计的伪影图像。为了保证最终生成图像的质量,定义了基于 L_1 距离的全局损失

$$L_G = E_{\mathbf{X},\mathbf{Y}} \left[\left\| \mathbf{Y} - \hat{\mathbf{Y}} \right\|_{L_1} \right] \quad (5)$$

\mathbf{Y} 为NDCT图像, $\hat{\mathbf{Y}}$ 为生成的降噪图像。此外,针对GAN网络存在训练不稳定以及模型崩溃等问题,本文降噪网络使用了最小二乘损失函数来代替原始GAN的损失函数。LSGAN能够生成比常规GAN更高质量的图像,在学习过程中表现更稳定。本文降噪网络整体目标函数为

$$\min_D V(D) = \frac{1}{2} E_{\mathbf{X},\mathbf{Y}} \left((D(\mathbf{X}, \mathbf{Y}) - 1)^2 \right) + \frac{1}{2} E_{\mathbf{X},\mathbf{Y}} \left((D(\mathbf{X}, \hat{\mathbf{Y}}))^2 \right) \quad (6)$$

$$\min_G V(G) = \frac{1}{2} E_{\mathbf{X},\mathbf{Z}} \left((D(\mathbf{X}, \hat{\mathbf{Y}}) - 1)^2 \right) + \lambda_1 L_G + \lambda_2 L_C + \lambda_3 L_A \quad (7)$$

其中, λ_1 , λ_2 和 λ_3 为超参数。

3 实验

为了评估LDCT图像降噪网络的性能,本文分别在模拟和真实LDCT图像数据集上进行了实验。(1)模拟数据集为Mayo数据集^[20],该数据集包括10名匿名患者的配对CT图像(即每一幅NDCT都有与之对应的LDCT),其中LDCT是通过模拟1/4标准剂量情况下的噪声污染而获得的。实验中,从数据集中选取了1811幅CT图像作为训练集,将剩余567幅CT图像作为测试集。(2)真实数据集为piglet数据集^[21],该数据集的LDCT是通过使用GE扫描仪(Discovery CT750 HD),将源电位和切片厚度分别设为100 kVp和0.625 mm时,通过调整管电流(或电压),经不同强度X射线扫描而获得的。以管电流为300 mA时的射线剂量为正常剂量,将管电流分别降低50%, 25%, 10%和5%便可获得4种不同剂量的LDCT。X射线辐射剂量不同,重建CT图像被噪声和伪影污染的程度不同。实验中,从piglet数据集中选择了2260幅CT图像(其中,从每种剂量的CT图像中选取565幅)作为训练集,测试集由剩余CT图像构成,本文采用了4种不同剂量水平的LDCT图像测试集,每个测试集包含100幅同一种剂量的LDCT。

本文采用峰值信噪比(Peak-Signal to Noise Ratio, PSNR)、结构相似性度(Structure SIMilarity, SSIM)和视觉信息保真度(Visual Information Fidelity, VIF)这3种客观评价指标来定量评估降噪图像的质量。其中,PSNR值越大,表明降噪图像中需要被保留的生理信息与需要被抑制的噪声信息之比越大,图像有用信息保留越完整,降噪图像质量越高。SSIM值越大,表明降噪图像的结构与NDCT图像的结构越接近。VIF用于测量降噪

LDCT图像和NDCT图像之间的视觉感知一致性, VIF值越大, 降噪后的LDCT图像视觉效果越好。此外, 本文使用Python语言在TensorFlow平台进行实验, 并使用NVIDIA GTX 1070 Ti GPU来训练和测试网络, 为验证降噪网络的有效性, 还与当前3种典型算法(分别为: 传统方法BM3D, CNN类方法的典型代表RED-CNN以及基于GAN的pix2pix)进行了对比。本文中的所有网络都经过了200次epochs迭代训练, 训练和测试图像尺寸均为 512×512 。实验中本文采用基于动量的Adam优化器^[22]训练降噪网络, 其中, $\beta_1 = 0.5$, $\beta_2 = 0.999$, 学习率为0.0002。考虑到参数选择对算法的性能起着极其重要的作用, 针对目标函数中涉及的3个超参数 λ_1 , λ_2 和 λ_3 , 本文在Mayo数据集上进行了实验: 将3个超参数的比例关系设置为 $\lambda_1:\lambda_2:\lambda_3 = 2:2:1$, 改变 λ_1 的取值(λ_2 和 λ_3 的取值随之变化)对网络进行训练, 测试结果的平均PSNR和SSIM值随 λ_1 取值的变化情况如图2(a)所示。从图中可以看出, 随着横坐标 λ_1 的增大, PSNR和SSIM值变化曲线呈现先增后减的变化规律, 当 $\lambda_1=20$ 时, 测试结果取得了最佳平均PSNR和SSIM值, 因此, 本文实验中将参数设置为 $\lambda_1=20$, $\lambda_2=20$, $\lambda_3=10$ 。

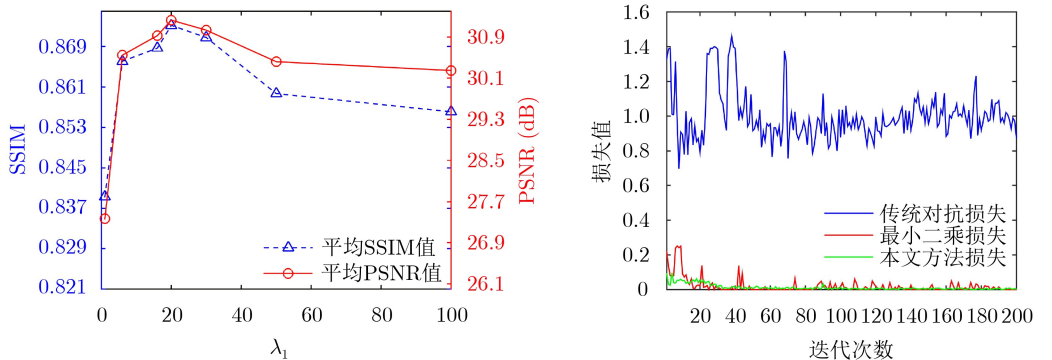
考虑到损失函数的选择对网络训练的稳定性起着极其重要的作用, 本文做了一组损失函数消融实验, 即分别采用传统对抗损失、最小二乘损失、最小二乘损失与像素级L1损失来训练降噪网络, 并绘制了如图2(b)所示的损失函数收敛曲线图。可以直观地看出, 3条损失函数变化曲线中绿色曲线所示情况下网络性能最稳定、且损失曲线收敛最快。传统GAN的对抗损失函数的变化曲线(如图2(b)蓝色曲线所示)存在明显波动, 随迭代次数增加, 网络并不能收敛于稳定的值; 与传统对抗损失函数相比, 采用最小二乘损失函数来对网络进行约束, 能够一定程度上缓解网络训练

过程不稳定的问题, 但迭代后期仍然存在损失值小幅度波动的现象(如图2(b)红色曲线所示); 将最小二乘损失与像素级L1损失函数相结合, 共同约束降噪网络的训练过程, 可以获得损失值逐渐收敛的变化曲线。因此, 在训练过程中, 采用最小二乘损失与像素级L1损失能够增加网络的稳定性。

3.1 定性分析

本文方法与对比方法对腹部与胸部LDCT图像的降噪结果分别如图3、图4、图5所示。从整体上看, 4种方法均可实现不同程度的伪影抑制, 比较不同的降噪结果图、伪影图与差值图(不同降噪结果与NDCT之差)可以发现, 经传统BM3D处理后, CT图像中仍然存在大量伪影和噪声。虽然RED-CNN和pix2pix对LDCT的伪影抑制效果优于BM3D, 但它们的降噪图像中均出现了不同程度的图像失真与边缘模糊, 如图3(c)和图3(d)所示。观察图4中红色箭头示意区域, 可以看出, 与其他3种方法相比, 本文方法在伪影噪声抑制和边缘细节保留方面的表现均是最佳的; 对比将LDCT与不同降噪图像及NDCT相减所获得的差值图像(如图5(a2)—图5(e2)所示), 可以发现, 本文方法所得结果图与理想伪影图像更接近, 图中伪影噪声的强度也是4种方法所得结果图中最大的; 观察图5(b3)—图5(e3)可以发现, LDCT与NDCT相减后, 可以得到一个有强伪影的差值图像, 4种降噪图与NDCT相减后所得到的差值图中也出现了不同强度的伪影, 4幅图中图5(b3)中残留伪影最多, 图5(e3)中残留伪影最少。综上, 从视觉效果来看, 本文方法的降噪效果更加优异。

图6展示了4种方法在piglet数据集上对不同剂量LDCT图像的降噪效果, 第1列为不同剂量的LDCT图像, 第2到4列分别为采用BM3D, RED-CNN和pix2pix与本文方法降噪效果图, 其中左下角为图6中红色虚线框所示ROI区域的放大图像。从



(a) 随 λ_1 值增大测试结果的平均PSNR和SSIM变化曲线

(b) 随迭代次数增加不同损失函数的变化曲线

图2 参数选择对算法性能的影响

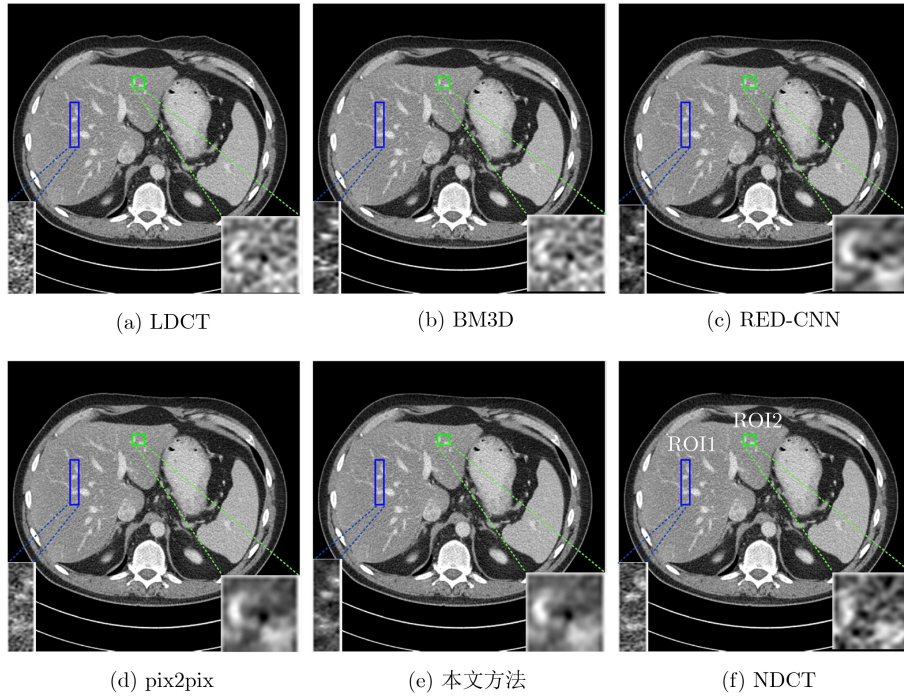


图 3 4种降噪方法对腹部LDCT图像的降噪结果示意图(显示窗为[40,400]HU)

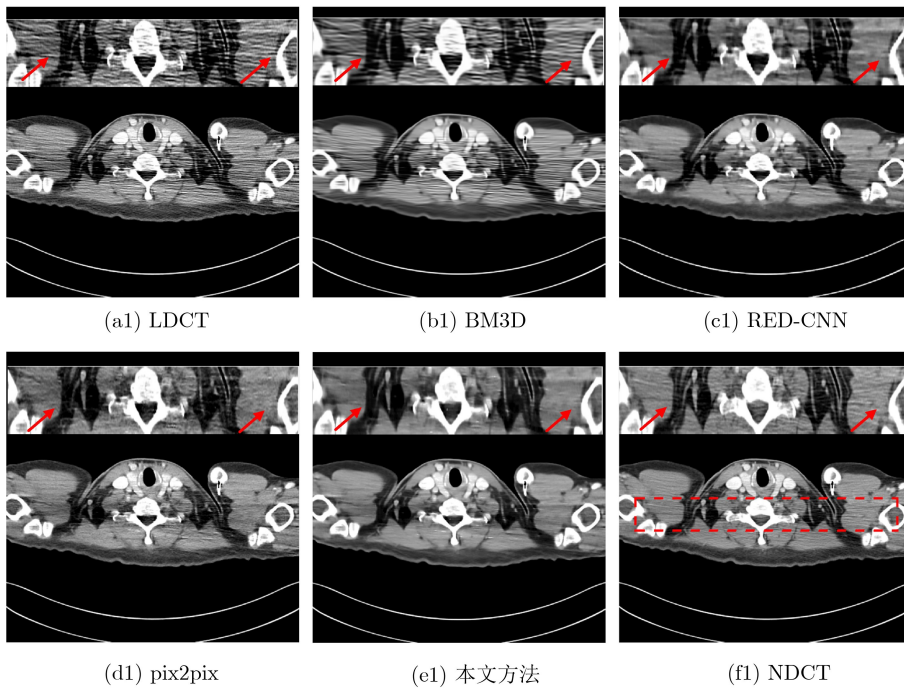


图 4 4种降噪方法对胸部LDCT图像的降噪结果示意图(显示窗为[40,400]HU)

第1列中可以观察到，在不同剂量LDCT图像中噪声呈现不同的分布特征，X射线剂量越低，噪声分布越密集，LDCT图像被噪声污染的程度越严重，当管电流变为15 mAs时，LDCT图像质量严重退化，很难辨认图像中具有临床诊断意义的组织和细微结构。观察第2和第3列降噪结果图可以发现，BM3D与RED-CNN虽然能在一定程度上对噪声进行抑制，但是当管电流分别降至30 mAs与15 mAs

时，BM3D与RED-CNN的降噪图像质量变得非常差，除了有部分噪声残留，还丢失了许多细节。将pix2pix、本文方法的降噪图像与NDCT进行比较可以发现，与前面两种方法的实验结果相比，GAN类方法对噪声抑制效果更加明显，且并没有在降噪的同时使图像变得模糊。与第6列NDCT图像进行对比可见，本文方法对不同剂量LDCT的降噪图像更接近于NDCT。此外，图中第1行—第4行分别给

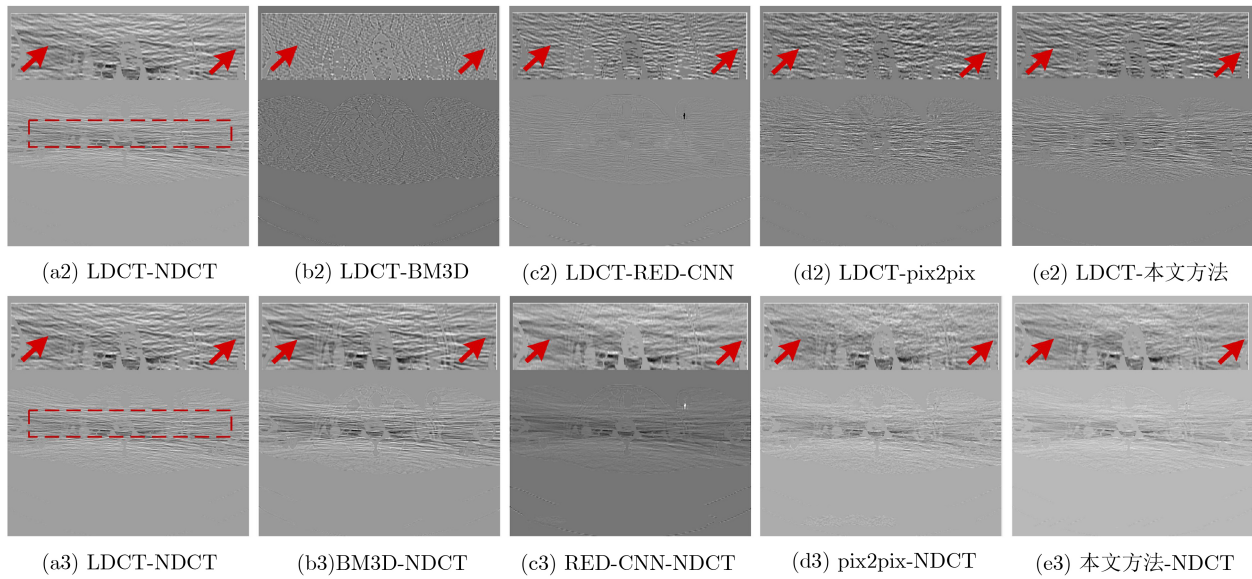


图5 4种降噪方法胸部LDCT降噪结果的伪影图与差值图(显示窗为[40,400]HU)

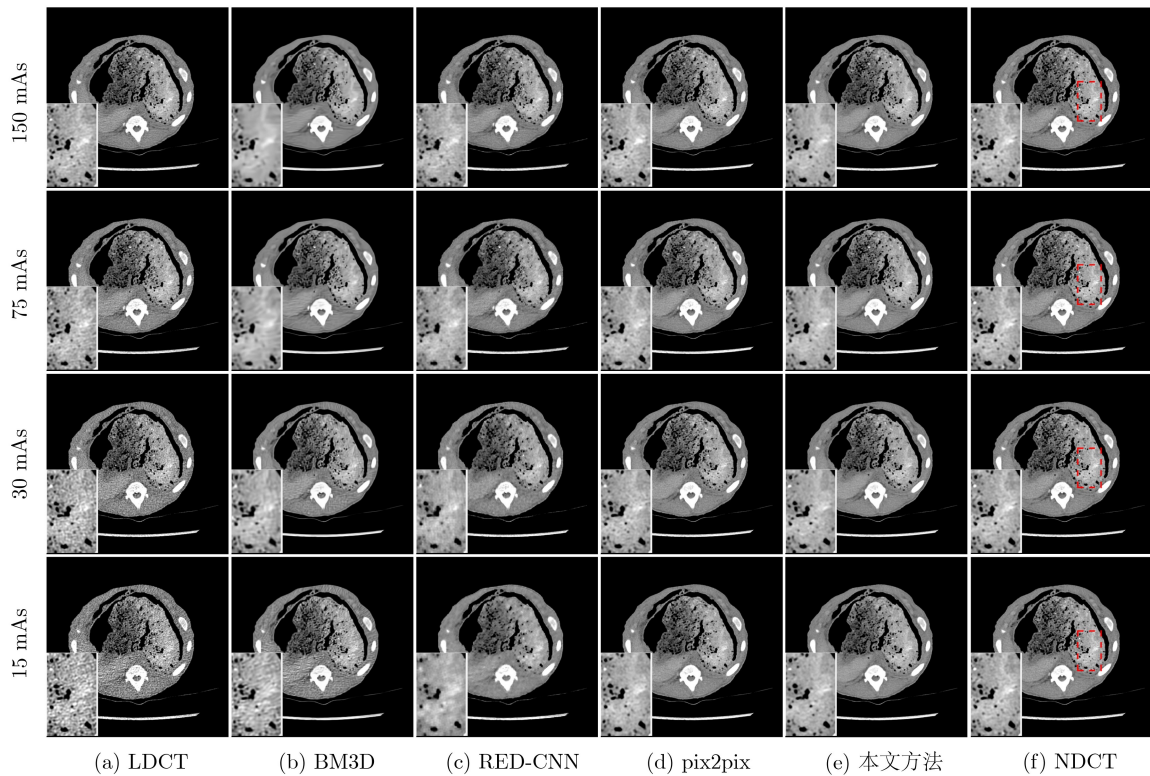


图6 4种方法在piglet数据集上对不同剂量LDCT图像的降噪结果(显示窗为[40,400]HU)

出了不同降噪方法对不同剂量LDCT的降噪图像显示。可以发现,无论采用哪种剂量的LDCT进行实验,本文方法的降噪效果最为明显,特别地,当噪声比较严重(30 mAs与15 mAs的管电流)时,本文降噪网络的优越性体现得更加充分,进一步证明本文提出的噪声水平估计子网络能够对不同剂量LDCT中的噪声伪影进行有效估计,降噪网络具有更好的泛化能力。

3.2 定量分析

为了对4种方法的降噪性能进行更全面的评估,本文还分析了4种降噪图像的量化表现。图7给出了图3所示降噪图像的2个ROI的SSIM、PSNR与VIF值。在同一ROI内分析4种降噪图像的量化表现,可以发现,除了ROI2中VIF值略低于BM3D,在ROI2其他量化指标与ROI1中,本文方法降噪图像指标值均是最大的。分析同一降噪图像的不同

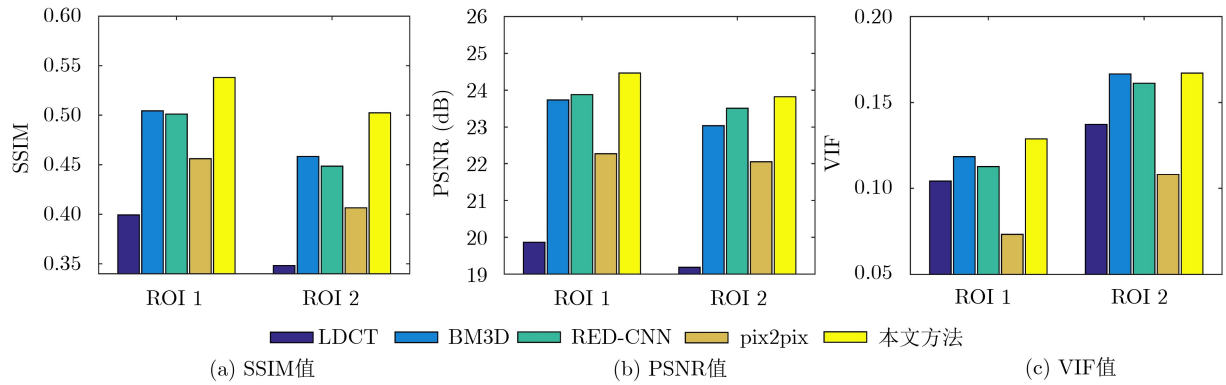


图7 图3所示降噪图像局部ROI的量化指标值

ROI可以发现SSIM方面, 4种降噪图像ROI1均比ROI2的数值更高在PSNR方面, 4种降噪图像ROI1均比ROI2的数值更高; 在VIF方面, 4种降噪图像ROI1的值均低于ROI2的值。总的来说, 从量化角度分析, 本文方法降噪性能优于其他3种方法。

表1给出了在piglet测试集上4种方法降噪结果的平均PSNR, SSIM和VIF值表现情况。将管电流分别设置为150 mAs, 75 mAs, 30 mAs和15 mAs时, 可以获得4种不同剂量的LDCT图像。从表1第1行可以看出, 剂量越低, LDCT图像质量越差。管电流从150 mAs降至75 mAs时, 图像质量急剧下降; 当管电流降至15 mAs时所获得的LDCT图像质量最差。分析4种方法对同一剂量水平的LDCT的降噪结果可以发现, 当对含有少量噪声的LDCT (例如管电流为150 mAs与75 mAs)进行降噪时, 4种方法中RED-CNN表现最差, 其次为BM3D和pix2pix, 而本文降噪网络性能最优; 当LDCT中包含大量噪声(例如管电流为15 mAs)时, 本文方法的实验结果均取得了最佳的PSNR, SSIM和VIF值; 分析每种方法对不同剂量水平的LDCT的降噪结果可以发现, BM3D对含有少量噪声(例如管电流为150 mAs)的LDCT进行降噪时获得了较高的PSNR和SSIM值, 然而当应对含有大量噪声的LDCT时, 量化值并不理想; 虽然在管电流较高的情况下, RED-CNN的实验结果并不理想, 甚至低于BM3D的实验结果, 但是随着管电流的降低, 两

者之间的差距逐渐缩小, 当管电流降至15 mAs时, 甚至超过了BM3D降噪图像的表现; pix2pix对不同剂量LDCT图像的降噪性能比较稳定, 本文方法对不同剂量LDCT降噪图像的SSIM, PSNR和VIF值表现均高于pix2pix的实验结果且均取得最大值, 这主要得益于噪声水平估计子网络强大的噪声伪影估计能力。总的来说, 本文降噪网络具有较强的鲁棒性, 对不同剂量LDCT进行降噪, 均能获得优于其他3种方法的降噪图像。

此外, 图8给出了在不同剂量水平情况下不同算法所抑制的噪声伪影图(即LDCT与不同算法降噪结果的残差图)与理想伪影图(由LDCT减去NDCT得到)的均方误差值变化曲线。从图中可以看出, 与其他算法降噪结果相比, 在4种不同剂量情况下, 本文方法所抑制的伪影噪声图与真实伪影图更加接近, 4种方法中本文方法获取的伪影噪声图与真实伪影的均方误差MSE值最小。这也说明本文网络对伪影噪声的特征提取能力更强, 进而能有效抑制更多的伪影噪声。

与基于GAN的传统降噪网络相比, 本文降噪网络在生成器部分设计了多路编码结构, 并增加了一个噪声水平估计子网络, 本文通过一组消融实验来评估这些改进对降噪网络性能影响, 具体地, 分别采用“w/o多路卷积”(网络中包括噪声水平估计子网, 主干降噪网络的生成器采用传统Unet结构)和“w/o噪声水平估计子网”(网络中不包括噪

表1 图6所示降噪图像的定量比较

	150 mAs			75 mAs			30 mAs			15 mAs		
	SSIM	PSNR(dB)	VIF	SSIM	PSNR(dB)	VIF	SSIM	PSNR(dB)	VIF	SSIM	PSNR(dB)	VIF
LDCT	0.9135	32.4367	0.3912	0.8827	29.7393	0.3087	0.8572	28.4505	0.2561	0.8232	25.6776	0.1867
BM3D	0.9346	32.7401	0.4810	0.9181	30.6341	0.4062	0.8996	30.6095	0.3547	0.8595	27.8316	0.2616
RED-CNN	0.9137	31.1185	0.3478	0.8944	28.9783	0.2960	0.9003	30.4194	0.3113	0.8930	29.1218	0.2858
pix2pix	0.9244	32.7808	0.4130	0.9084	30.8796	0.3585	0.9224	32.9193	0.3986	0.9118	31.3273	0.3537
本文方法	0.9484	34.6459	0.5030	0.9408	32.6691	0.4580	0.9453	34.1737	0.4776	0.9403	33.1630	0.4457

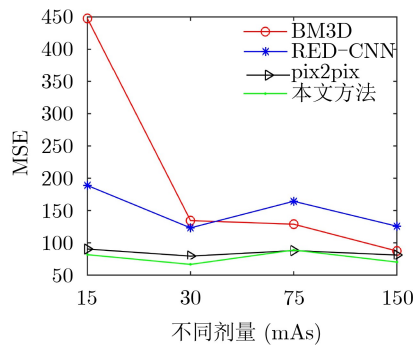


图8 不同算法所抑制的噪声伪影图与理想伪影图的均方误差值变化曲线

声水平估计子网, 主干降噪网络的生成器采用改进的多路编码Unet结构)两个网络在Mayo数据集上进行训练, 并比较其与本文网络的测试结果。表2所示为不同降噪网络测试结果的平均SSIM和PSNR值及相应的局部ROI视觉效果图汇总表, 可以看出, 本文降噪网络测试结果的平均PSNR与SSIM量化指标值是3种算法中表现最好的, “w/o多路卷积”和“w/o噪声水平估计子网”两种网络降噪效果量化指标值差别不大; 这里没有展现消融网络的视觉效果。综合测试结果的两种量化表现进行考虑, 本文网络的降噪性能均是最佳的, 这也进一步表明, 本文所做改进对改善LDCT降噪图像的质量具有积极作用。

表3给出了不同算法测试相同数据集所需要的平均测试时间。观察表中数据可以发现, RED-CNN与BM3D算法测试的时间相对较长, 由于网络结构更加复杂, 本文方法所花费时间略高于pix2pix, 执行效率差距并不大, 但本文方法的综合视觉效果和量化指标表现更好。

表2 网络结构消融对算法性能的影响

	子模块		SSIM	PSNR (dB)
	多路卷积	噪声水平估计子网		
w/o多路卷积		✓	0.8684	30.3736
w/o噪声水平估计子网	✓		0.8696	31.0011
本文方法	✓	✓	0.8732	31.2266

表3 不同算法的测试时间对比

算法	BM3D	RED-CNN	pix2pix	本文方法
测试时间(s/张)	1.2078	8.6900	0.0710	0.1085

4 结束语

CT成像技术已经成为不可或缺的医学诊断辅

助手段, 为了避免X射线剂量累加引发的辐射伤害, 低剂量CT图像的应用日益增多。针对由于减少辐射剂量引起的CT图像噪声和伪影增加问题, 本文致力于设计有效的降噪卷积神经网络模型, 以期在保持CT图像组织结构和病理信息完整的基础上尽可能抑制图像中的噪声和伪影。本文提出了一种用于低剂量CT降噪的生成对抗网络改进模型, 增加了噪声水平估计子网来提高网络对不同剂量LDCT图像的噪声估计能力, 并且引入了噪声等级一致性损失函数来约束该网络参数的优化过程, 保证了噪声估计子网的噪声估计准确度。此外, 本文提出一种多路编码U-Net结构作为生成器的主干降噪网络, 使用了不同大小卷积核提取多尺度特征的同时增大网络的感受野, 更进一步提升了网络的降噪性能。本文虽然一定程度上解决了生成对抗网络用于LDCT图像降噪中表现出来的泛化能力弱问题, 针对不同辐射剂量的LDCT图像, 如何设计性能更优的噪声估计子网来提升降噪网络的性能仍然是很有研究价值的课题。

参考文献

- [1] 张权. 低剂量X线CT重建若干问题研究[D]. [博士学位], 东南大学, 2015.
ZHANG Quan. A study on some problems in image reconstruction for low-dose CT system[D]. [Ph. D. dissertation], Southeast University, 2015.
- [2] HSIEH J. Computed Tomography: Principles, Design, Artifacts, and Recent Advances[M]. Bellingham, USA: SPIE Press, 2009.
- [3] BRENNER D J and HALL E J. Computed tomography-an increasing source of radiation exposure[J]. *New England Journal of Medicine*, 2007, 357(22): 2277-2284. doi: [10.1056/NEJMr072149](https://doi.org/10.1056/NEJMr072149).
- [4] NAIDICH D P, MARSHALL C H, GRIBBIN C, et al. Low-dose CT of the lungs: Preliminary observations[J]. *Radiology*, 1990, 175(3): 729-731. doi: [10.1148/radiology.175.3.2343122](https://doi.org/10.1148/radiology.175.3.2343122).
- [5] CHEN Hu, ZHANG Yi, ZHANG Weihua, et al. Low-dose CT via convolutional neural network[J]. *Biomedical Optics Express*, 2017, 8(2): 679-694. doi: [10.1364/BOE.8.000679](https://doi.org/10.1364/BOE.8.000679).
- [6] SHAN Hongming, ZHANG Yi, YANG Qingsong, et al. 3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network[J]. *IEEE Transactions on Medical Imaging*, 2018, 37(6): 1522-1534. doi: [10.1109/TMI.2018.2832217](https://doi.org/10.1109/TMI.2018.2832217).
- [7] CHEN Hu, ZHANG Yi, 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. doi: [10.1109/TMI.2017.2715284](https://doi.org/10.1109/TMI.2017.2715284).

- [8] WU Dufan, KIM K, FAKHRI G E, *et al.* A cascaded convolutional neural network for X-ray low-dose CT image denoising[EB/OL]. <https://arxiv.org/abs/1705.04267>, 2017.
- [9] WOLTERINK J M, LEINER T, VIERGEVER M A, *et al.* Generative adversarial networks for noise reduction in low-dose CT[J]. *IEEE Transactions on Medical Imaging*, 2017, 36(12): 2536–2545. doi: [10.1109/TMI.2017.2708987](https://doi.org/10.1109/TMI.2017.2708987).
- [10] YANG Qingsong, YAN Pingkun, ZHANG Yanbo, *et al.* Low-Dose CT image denoising using a generative adversarial network with wasserstein distance and perceptual loss[J]. *IEEE Transactions on Medical Imaging*, 2018, 37(6): 1348–1357. doi: [10.1109/TMI.2018.2827462](https://doi.org/10.1109/TMI.2018.2827462).
- [11] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, *et al.* Generative adversarial nets[C]. The 27th Conference and Workshop on Neural Information Processing Systems (NIPS), Cambridge, United States, 2014: 2672–2680.
- [12] BAO Jianmin, CHEN Dong, WEN Fang, *et al.* CVAE-GAN: Fine-grained image generation through asymmetric training[C]. The IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017: 2745–2754.
- [13] ZHENG Jin, MA Haocheng, and PENG Lihui. A CNN-based image reconstruction for electrical capacitance tomography[C]. 2019 IEEE International Conference on Imaging Systems and Techniques (IST), Abu Dhabi, United Arab Emirates, 2019: 1–6. doi: [10.1109/IST48021.2019.9010096](https://doi.org/10.1109/IST48021.2019.9010096).
- [14] ZHANG Zhijun and JI Xiaopeng. GAN network solves the problem of image segmentation across data[C]. 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chengdu, China, 2019. doi: [10.1109/iaeac47372.2019.8998056](https://doi.org/10.1109/iaeac47372.2019.8998056).
- [15] KIM D W, CHUNG J R, and JUNG S W. GRDN: Grouped residual dense network for real image denoising and GAN-based real-world noise modeling[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Long Beach, USA, 2019.
- [16] HE Zhang, VISHWANATH S, and VISHAL M P. Image de-raining using a conditional generative adversarial network[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, 30(11): 3943–3956. doi: [10.1109/TCSVT.2019.2920407](https://doi.org/10.1109/TCSVT.2019.2920407).
- [17] ISOLA P, ZHU Junyan, ZHOU Tinghui, *et al.* Image-to-image translation with conditional adversarial networks[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, USA, 2017: 1125–1134.
- [18] ZHANG Kai, ZUO Wangmeng, and ZHANG Lei. FFDNet: Toward a fast and flexible solution for CNN-based image denoising[J]. *IEEE Transactions on Image Processing*, 2018, 27(9): 4608–4622. doi: [10.1109/TIP.2018.2839891](https://doi.org/10.1109/TIP.2018.2839891).
- [19] SZEGEDY C, LIU Wei, JIA Yangqing, *et al.* Going deeper with convolutions[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, USA, 2015: 1–9.
- [20] AAPM. Low dose CT grand challenge[EB/OL]. <http://www.aapm.org/GrandChallenge/LowDoseCT/>, 2017.
- [21] Piglet Dataset[EB/OL]. <http://homepage.usask.ca/~xiy525/>.
- [22] DIEDERK P K and JIMMY L B. Adam: A method for stochastic optimization[C]. International Conference on Learning Representations (ICLR), Ithaca, USA, 2015.

张 雄: 男, 1973年生, 教授, 硕士生导师, 研究方向为模式识别、医学图像处理和视频目标跟踪。

杨琳琳: 女, 1992年生, 硕士生, 研究方向为医学图像处理。

上官宏: 女, 1988年生, 副教授, 硕士生导师, 研究方向为模式识别、医学图像处理。

韩泽芳: 女, 1996年生, 硕士生, 研究方向为医学图像处理。

韩兴隆: 男, 1995年生, 硕士生, 研究方向为医学图像处理。

王安红: 女, 1972年生, 教授, 博士生导师, 研究方向为图像视频编码。

崔学英: 女, 1978年生, 副教授, 硕士生导师, 研究方向为图像处理与重建。

责任编辑: 余 蓉