

基于梯度指导的生成对抗网络内镜图像去模糊重建

时永刚* 张岳 周治国 李祎 夏卓岩

(北京理工大学信息与电子学院 北京 100081)

摘要: 胃肠镜检查是目前临床上检查和诊断消化道疾病最重要的途径, 内窥镜图像的运动模糊会对医生诊断和机器辅助诊断造成干扰。现有的去模糊网络由于缺乏对结构信息的关注, 在处理内窥镜图像时普遍存在着伪影和结构变形的问题。为解决这一问题, 提高胃镜图像质量, 该文提出一种基于梯度指导的生成对抗网络, 网络以多尺度残差网络(Res2net)结构作为基础模块, 包含图像信息支路和梯度支路两个相互交互的支路, 通过梯度支路指导图像去模糊重建, 从而更好地保留图像结构信息, 消除伪影、缓解结构变形; 设计了类轻量化预处理网络来纠正过度模糊, 提高训练效率。在传统胃镜和胶囊胃镜数据集上分别进行了实验, 实验结果表明, 该算法的峰值信噪比(PSNR)和结构相似度(SSIM)指标均优于对比算法, 且复原后的视觉效果更佳, 无明显伪影和结构变形。

关键词: 胃镜图像; 去模糊; 生成对抗网络; 梯度指导

中图分类号: R331; TN911.73

文献标识码: A

文章编号: 1009-5896(2022)01-0070-08

DOI: [10.11999/JEIT210920](https://doi.org/10.11999/JEIT210920)

Deblurring and Restoration of Gastroscopy Image Based on Gradient-guidance Generative Adversarial Networks

SHI Yonggang ZHANG Yue ZHOU Zhiguo LI Yi XIA Zhuoyan

(School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China)

Abstract: Gastrointestinal endoscopy plays a critical role in examination and diagnosis upper gastrointestinal diseases. The motion blur of endoscopic images can interfere with doctor's judgment and machine-assisted diagnosis. Due to the lack of attention to structural information in existing deblurring networks, artifacts and structural distortions occur easily when processing endoscopic images. In order to solve this problem and improve the image quality of gastroscopy, a gradient-guided generative adversarial network is proposed in this paper. The network uses the Res2net structure as the backbone, including two interactive branches, the image branch with its intensity and the gradient one. The gradient branch guides the deblurring and reconstruction of the image which in the other branch. Thus more structure information of the image can be kept, with less artifacts and alleviating structural deformation. A quasi-lightweight preprocessing network is designed to correct excessive blur and improve training efficiency. Experiments are performed on the traditional gastroscopy and the capsule gastroscopy datasets. The test results show that the Peak Signal to Noise Ratio(PSNR) and Structural SIMilarity(SSIM) indicators of the algorithm are better than those of the comparison algorithms, and the visual effect of the restored image is evidently improved, without obvious artifacts and structural deformation.

Key words: Gastroscopy image; Deblurring; Generative adversarial network; Gradient guidance

1 引言

根据世界卫生组织统计, 2020年有3624820例新增消化道癌症病例, 有2248042名患者死于消化道癌症, 消化系统的癌症死亡人数占到癌症总死亡

人数的22.6%^[1]。内窥镜检查是现在主要的消化道疾病诊断方法, 传统内窥镜具有直接、准确、价格低的优点, 是目前最常用的胃部检查方法。随着医药科技的发展, 新兴的胶囊胃镜技术因其更佳检测体验, 已逐渐得到普及。

然而, 无论是使用传统的胃镜还是胶囊胃镜, 在图像采集过程中都必然会出现运动模糊造成的图像质量下降的问题。特别是胶囊胃镜, 在数据采集过程中医生无法实时控制采集设备。其运动受到胃

收稿日期: 2021-09-01; 改回日期: 2021-12-26; 网络出版: 2022-01-04

*通信作者: 时永刚 ygshi@bit.edu.cn

基金项目: 国家自然科学基金(60971133,61271112)

Foundation Items: The National Natural Science Foundation of China (60971133, 61271112)

肠蠕动的影 响。当胃肠蠕动快时, 拍摄数量减少锐度降低, 图像模糊程度比普通胃镜更严重。模糊图像会给计算机辅助诊断带来巨大的困难, 进而干扰医生对病变部位的判断。因此, 通过内窥镜的去模糊算法, 提高了内窥镜的成像质量, 对提高临床诊断准确率具有重要意义。

2017年Mahesh等人^[2]提出一种卷帘(Rolling Shutter, RS)盲去运动模糊的模型, 用以解决互补金属氧化物半导体(Complementary Metal Oxide Semiconductor, CMOS)传感器的卷帘快门引起的运动模糊。2018年, Tao等人^[3]提出尺度递归网络(Scale-Recurrent Network-DeblurNet, SRN-DeblurNet), 它类似于构造重建金字塔并逐步优化以获得最终的重建清晰图片。Kupyn等人^[4]提出的去模糊生成对抗网络(Deblurring Generative Adversarial Networks, DeblurGAN)是一种结合条件生成对抗网络和内容损失的端到端运动去模糊学习模型。两年之后, Kupyn等人^[5]又在DeblurGAN的基础上提出了改进型去模糊生成对抗网络(Deblurring Generative Adversarial Networks-v2, DeblurGAN-v2), 改进了生成器的网络结构和鉴别器, 并使该算法便于使用现有成熟的基础网络, 不仅提高了去模糊的图像质量, 而且能够以低成本设计模型。与普通图像不同, 消化道内镜图像的纹理细节较为丰富, 且图像结构较为多样, 现有的去模糊方法如SRN-DeblurNet, DeblurGAN-v2在处理这一类图像时因为过度关注细节特征, 全局特征关注不够, 总会伴有一定程度的结构变形甚至出现伪影, 虚构生成了实际不存在的重影轮廓。在图像超分辨率领域, 已经被证明可以用梯度轮廓作为先验^[6-8]指导超分辨率图像更好地保留结构信息, 生成更加清晰的高分辨率图像。其中, 清华大学Ma等人^[8]提出了基于生成对抗网络的保留结构信息的超分辨率网络(Structure-Preserving Super Resolution, SPSR), 证明了在梯度信息指导下通过生成对抗网络生成结构不变形且纹理细节丰富的图像的可行性。2021年Tran等人^[9]提出一种方法来编码任意的清晰-模糊图像对数据集的模糊核空间, 并利用这种编码来解决一些特定的任务, 如图像去模糊和模糊合成。Chi等人^[10]提出一种自监督元辅助学习, 通过整合外部和内部学习来提高去模糊的性能, 并提出一个元辅助训练方案, 进一步优化作为基础学习的预训练模型。Chen等人^[11]提出一个非盲目去模糊网络(Non-Blind Deblurring Network, NBDN)来恢复夜间模糊的图像, 为降低个别不符合反模糊模型的像素所带来的副作用, 设计了置信

度估计单元(Certainty Estimation Unit, CEU), 确保这些像素在去卷积步骤中的贡献较小, 并通过共轭梯度(Conjugate Gradient, CG)方法进行优化。Dong等人^[12]提出一种在最大后验(Maximum A-Posteriori, MAP)框架内联合学习空间变异数据(spatially-variant data)和正则化项的方法, 用于非盲目图像去模糊化。

本文提出基于梯度指导的生成对抗去模糊网络和类U型网络(U-shaped Network, U-Net)预训练模型。预训练模型使用3层类U-Net结构, 解码器部分引入残差连接, 预先对剧烈模糊图像进行处理, 大大缩短了网络的训练周期。去模糊网络以Res2net^[13]结构作为基础模块增大感受野, 减轻冗余特征带来的运算负担, 网络同时包含图像信息和梯度信息两个相互交互的支路, 通过梯度支路指导真实图像的复原重建, 更好地保留图像的结构信息, 消除伪影缓解结构变形, 从而复原生成更加清晰的胃镜图像。

2 模型

2.1 梯度指导的去模糊生成对抗网络

在去模糊网络中, 模糊图像 I^B 作为输入, 根据清晰的标签图像 I^{SI} 生成清晰图像 I^{SO} 。生成器 G , 其参数表示为 θ_G , 由 G 得到 $I^{SO} = G(I^B; \theta_G)$, 输出 I^{SO} 在分布上应尽可能地与输入 I^{SI} 相似。整个去模糊网络结构如图1所示。生成器由两个分支组成, 其中一个保留结构信息的图像分支, 另一个是图像的梯度分支。图像信息分支以 I^B 作为输入, 在梯度分支的指导下恢复出清晰图像 I^{SO} 。

生成对抗网络(Generative Adversarial Networks, GAN)因其对抗学习的特性具有强大的图像生成能力, 对于生成图像的细节纹理质量有很大提升^[4], 因此在图像信息支路中本文使用了类似增强型超分辨率生成对抗网络(Enhanced Super-Resolution Generative Adversarial Networks, ESRGAN)^[14]的网络结构, ESRGAN具有良好的细节生成能力, 并且能够解决一定程度的伪影问题。ESRGAN算法的生成器使用密集残差网络(Residual in Residual Dense Block, RRDB)的基础模块, 将残差块以稠密连接网络(DenseNet)^[15]方式进行连接, 这种方式增加了训练参数, 并保留了更多的有效特征。通常在对胃镜图像去模糊测试, 会有过多的特征冗余, 容易导致伪影现象的出现, 虚构生成不存在的重影轮廓。为改善去模糊效果, 本文引入基于Res2net的基础模块来代替RRDB, 这样既增大了感受野, 同时又减少了冗余特征带来的运算负担和重影轮廓。

梯度支路采用了梯度强度作为梯度图,从图像到图像之间的转换空间来学习两种模式之间的映射。由于梯度图中大部分区域数值接近于0,而轮廓边缘具有较大的数值,卷积神经网络能更多地关注轮廓的空间关系。在梯度支路中与基于生成对抗网络的保留结构信息的超分辨率网络(Structure-Preserving Super Resolution,SPSR)不同,本文没有使用梯度模块加卷积层的结构,而是直接使用Res2net模块,减少了中间的冗余特征,网络更容易捕获关键部位结构依赖性,由梯度图指导生成清晰图像。

如图1所示,梯度分支和图像信息分支有多层级的特征信息交互,这些特征流作为一个强先验来提升梯度分支的性能,在这种情况下,可大大减少

参数。在实验中,本文将梯度分支的最后一层产生的特征图传递到真实图像分支。同时,通过 1×1 卷积层生成输出层梯度图。通过这种双支路交互,图像信息分支最终能够携带丰富的结构信息,确保复原后的图像不会产生明显的结构变形。判别器部分使用了相对平均判别器,网络结构为如图2所示的VGG-128^[16]。

2.2 类U型网络预训练模型

本文在使用梯度指导生成对抗网络(Gradient Guidance GAN,GGGAN)对模糊胃镜图像进行实验时发现,GGGAN具有良好的图像生成效果,但在复原重度模糊图像时复原图像会出现一定程度的纹理扭曲,且训练收敛时间较长,这一现象源于GAN网络自身的性质^[17],很难通过内部的结构调

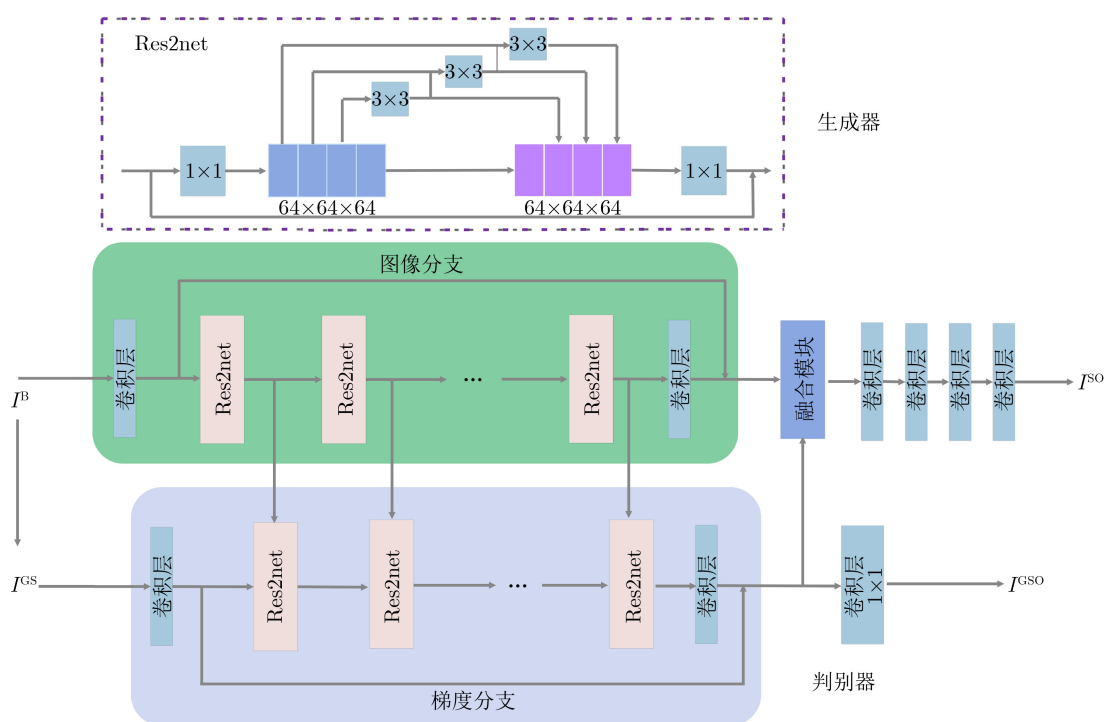


图1 梯度指导生成对抗网络生成器结构图

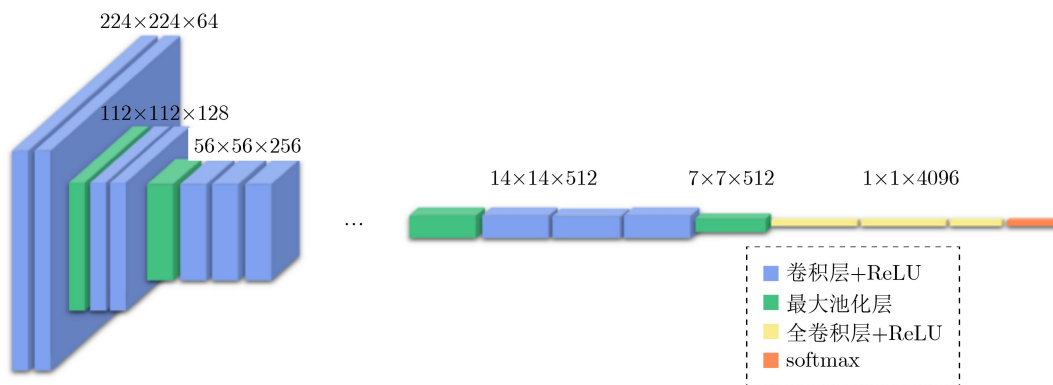


图2 梯度指导生成对抗网络判别器结构图

整避免。这一情况会影响图像的整体视觉效果，降低峰值信噪比(Peak Signal to Noise Ratio, PSNR)这一评价指标，对于图像复原效果并不理想。

为解决这一问题，本文在GGGAN网络之前设计了一个预处理模块，通过这一模块对实验图像进行预先修正处理，可以在一定程度上减小输入图像的模糊程度，避免因输入图像模糊程度太大带来的纹理扭曲，缩短训练时间。

预处理模块的整体结构类似于U-Net^[18]，使用3层编码器-解码器结构，并在解码部分引入残差连接^[15]，扩充感受野，其结构如图3。在满足预处理需求的前提下，3层编码器-解码器结构的训练参数量较少，不会带来更多的训练负担。

2.3 GGGAN与预训练模块的结合

预训练模型和GGGAN的生成器构成的网络结构如图4所示，预处理模块放在此结构之前。

计算模糊图像预处理后输出 I_X 与清晰图像 I_Y 之间的结构相似性指标^[19]

$$SSIM_{pre}(I_X, I_Y) = \frac{2\mu_X\mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1} \cdot \frac{2\sigma_X\sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2} \cdot \frac{\sigma_{XY} + C_3}{\sigma_X\sigma_Y + C_3} \quad (1)$$

其中， μ_X, μ_Y 分别表示图像 I_X 和 I_Y 的像素值均值， σ_X, σ_Y 分别表示图像 I_X 和 I_Y 的像素值标准均方差， σ_{XY} 表示图像 I_X 和 I_Y 的协方差； C_1, C_2, C_3 为常数， $C_1 = 6.5025, C_2 = 58.5225, C_3 = 0.5C_2$ 。

将 $SSIM_{pre}$ 加入到损失函数中，参与回传，以此指导训练中参数的调整。损失函数即为预处理损失 $Loss_{pre}$ 与GGGAN损失 $Loss_{GG}$ 加权之和，即

$$Loss = \alpha \cdot Loss_{pre} + Loss_{GG} \quad (2)$$

其中

$$Loss_{pre} = 1 - SSIM_{pre} \quad (3)$$

为使得两损失函数在同一数量级，以免将另一指标在训练中淹没，选取 $\alpha = 10^{-3}$ 。

3 实验结果与分析

3.1 实验准备

3.1.1 数据集

采用内窥镜伪影检测(Endoscopy Artifact Detection, EAD)挑战赛数据集^[20]以及胶囊胃镜数据集^[21](internet-based digital video atlas of capsule endoscopy for research purposes)中的图像作为清晰图像，人为添加运动模糊产生对应的模糊图像，其中运动模糊的范围为平移50~80像素点，旋转模

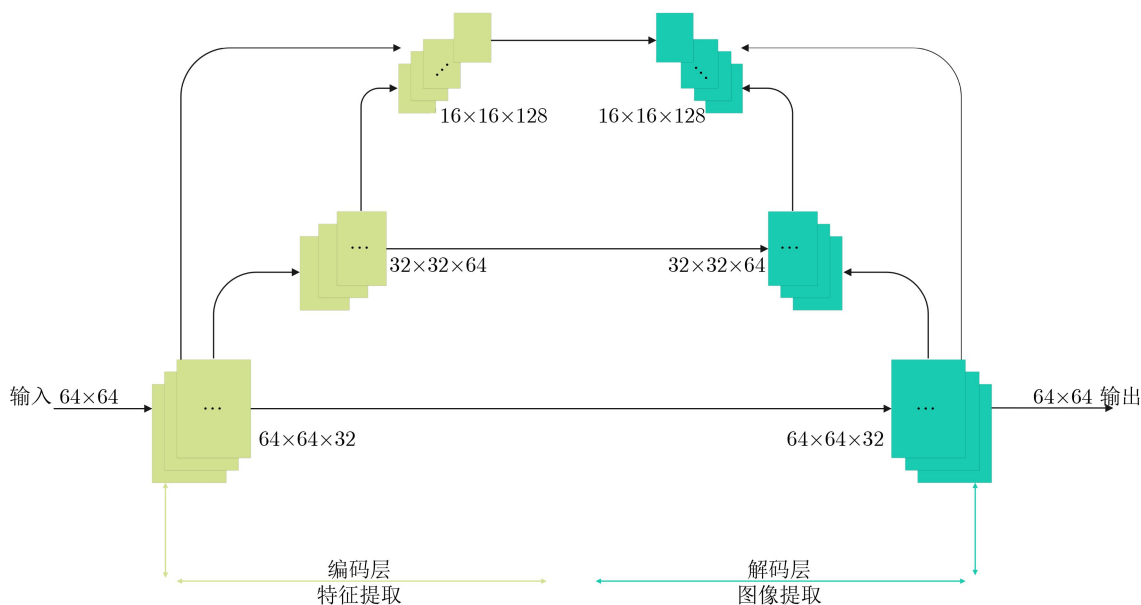


图3 预处理模块的网络结构

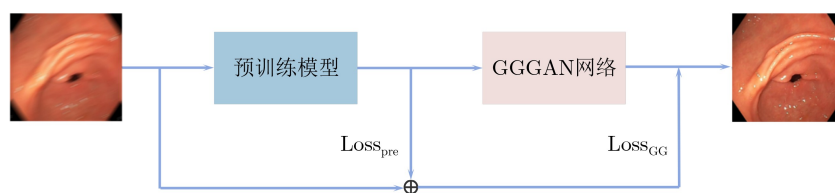


图4 本文网络整体结构

模糊核在 $-40^\circ \sim 40^\circ$ 范围内随机产生, 每个图像对的模糊核不同。传统胃镜数据集共有清晰-模糊图像对780组, 其中480对为训练集, 300对为测试集。胶囊胃镜数据集中共有清晰-模糊图像对2398组, 分为训练集1398对, 测试集1000对。

由于数据采集过程中采用了随机裁切的方法, 无须再对数据集做单独的增广。随机裁切使得训练过程中同一对图像每次输入是不同的, 极大程度上避免了过拟合现象的产生, 保证了网络的泛化能力。

3.1.2 实验评价指标

为比较不同算法的去模糊效果, 本文使用两种评价指标来测量最终结果: 峰值信噪比(PSNR)和结构相似度(SSIM)。PSNR是评价图像质量最常用、应用最广泛的客观测量方法^[22]。对于灰度图像, 给定清晰的图像 I 和大小为 $m \times n$ 的模糊图像 K , 均方误差(Mean Square Error, MSE)定义为

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (4)$$

然后将PSNR(dB)定义为

$$\text{PSNR} = 10 \cdot \lg \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right) \quad (5)$$

其中, MAX_I 是图像的最大像素值。如果每个像素用8位二进制表示, 则为255。对于彩色图像, 计算红绿蓝(Red-Green-Blue, RGB)3个通道的MSE, 然后除以3得到最终的PSNR值。

SSIM指标是度量两幅图像的结构相似程度, 是一种感知模型。与MSE绝对误差不同, 它是更符合人类直观感知的图像质量评估标准, 可以抵消MSE无法测量图像结构^[23]相似性的缺点。SSIM已在损失函数部分中简要介绍, 在这里将给出其计算方法。

SSIM公式基于图像的3个关键特征度量: 亮度($l(I_X, I_Y)$)、对比度($c(I_X, I_Y)$)和结构($s(I_X, I_Y)$)

$$l(I_X, I_Y) = \frac{2\mu_X\mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1} \quad (6)$$

$$c(I_X, I_Y) = \frac{2\sigma_X\sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2} \quad (7)$$

$$s(I_X, I_Y) = \frac{\sigma_{XY} + C_3}{\sigma_X\sigma_Y + C_3} \quad (8)$$

一般采用 $C_3 = 0.5C_2$ 。 $C_1 = (k_1L)^2$, $C_2 = (k_2L)^2$ 以避免分母为0。 L 是像素值的范围上限, 取255。默认值为 $k_1 = 0.01$, $k_2 = 0.03$ 。所以

$$\text{SSIM}(I_X, I_Y) = [l(I_X, I_Y)^\alpha c(I_X, I_Y)^\beta s(I_X, I_Y)^\gamma] \quad (9)$$

以 $\alpha = \beta = \gamma = 1$ 为例

$$\text{SSIM}(I_X, I_Y) = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)} \quad (10)$$

3.1.3 实验计算资源

本文中, 所有进行实验比较的模型都是基于PyTorch框架, 并在Intel Xeon E5-2680 v4处理器和2片NVIDIA GeForce RTX 2080 Ti(总计显存为22 GB)的服务器上训练了500000个epoch。

3.2 实验方法

本实验数据中图像尺寸为 1349×1079 , 为适应运行时设备限制的内存占用空间, 将图像随机裁切为 128×128 大小进行训练, 生成器和鉴别器的初始学习率均为 10^{-4} , 在迭代周期分别为50000, 100000, 200000, 300000时进行 10^{-1} 的衰减。网络的训练过程中采用了小批次训练迭代, 批量尺寸(batch size)被设置成4。使用自适应矩估计(Adaptive moment estimation, Adam)优化算法^[24]来优化模型参数。在传统胃镜数据集和胶囊胃镜数据集上分别与SRN-DeblurNet, DeblurGAN-v2两种主流的去模糊算法及改进SPSR网络进行了对比。SRN-DeblurNet, DeblurGAN-v2为目前效果较好的开源算法, 改进SPSR网络则是将Ma等人^[8]的超分辨率网络改为去模糊网络, 用来对比算法性能。在胶囊胃镜数据集的实验中, 本文增加了迁移测试实验, 利用传统胃镜数据集训练好的模型在胶囊胃镜数据集上直接进行测试, 对比不同模型的泛化性能。

对传统胃镜数据集进行测试, 结果如表1所示。实验结果显示, 在数据指标上本文算法比其他方法在两个指标上均有所提升。与对照组中指标较高的DeblurGAN-v2相比较, 本文算法将PSNR值从26.55提升至26.89, SSIM值从0.818提升至0.849。

从图5可以看出, SRN-DeblurNet, DeblurGAN-v2方法处理后, 图像仍然存在模糊, 尤其在消化道中溃疡出血、增生息肉及肠壁边缘处较为明显。改进的SPSR方法虽然在细节生成效果上优于

表1 不同算法在传统胃镜数据集上的指标测试结果

	SRN-DeblurNet	改进SPSR	DeblurGAN-v2	本文算法
峰值信噪比(PSNR)	26.40	26.47	26.55	26.89
结构相似度(SSIM)	0.799	0.815	0.818	0.849

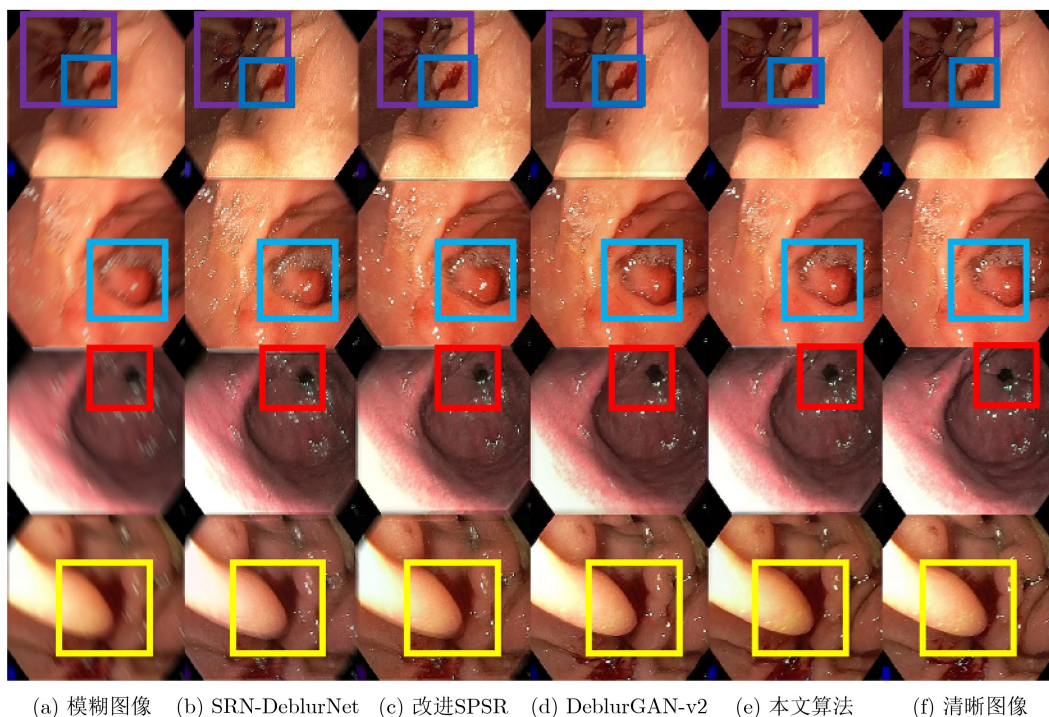


图5 不同算法在传统胃镜数据集上的图像测试结果

SRN-DeblurNet, 但是出现了重影的现象, 实际的复原效果也并不理想。本文算法处理后的胃镜图像与SRN-DeblurNet方法和DeblurGAN-v2方法相比不但提高了清晰度, 丰富了边缘和纹理信息; 还缓解了GAN网络虚构出的重影现象, 复原图像更加贴近真实的清晰图像, 在胃内粘液、出血和褶皱等细节处, 本文算法有更好的表现力。

表2为不同算法在胶囊胃镜数据集上的指标测试结果, 从表2实验数据可以看出, 本文算法在不同的数据上也能有更好的表现力。在评价指标方面, 本文算法训练后的PSNR和SSIM值分别达到了29.04和0.826, 高于SRN-DeblurNet的28.01和0.795, 也高于改进SPSR的28.70和0.771以及DeblurGAN-v2的28.81和0.813; 迁移测试中, 本文算法在传统胃镜数据集上的PSNR和SSIM值分别为28.83和0.801, 高于其他3种对比算法SRN-DeblurNet、改进SPSR和DeblurGAN-v2。

从图6标识迁移的图像是使用光学胃镜数据集训练后直接在胶囊胃镜数据集上迁移测试的结果,

标识训练的图像是使用胶囊胃镜数据集训练并测试得到的结果, 本文算法在胶囊胃镜数据集上直接训练实验及迁移测试复原的胃镜图像效果均优于SRN-DeblurNet、改进SPSR和DeblurGAN-v2 3种对比算法, 在毛细血管、局部溃疡等细节变化多的位置最为明显。

4 结束语

本文提出一种基于梯度指导生成对抗网络的内镜图像去模糊算法并对损失函数进行了改进, 通过梯度信息指导模糊图像的复原, 缓解了结构变形和虚构重影的现象; 设计了预处理网络来纠正过度模糊, 提高训练效率。通过在传统光学胃镜和胶囊内镜的数据集上进行训练和测试, 检验其性能。实验结果表明, 在性能指标方面, 本文算法的PSNR值及SSIM值较SRN-DeblurNet和DeblurGAN-v2等均有所提升, 且迁移测试效果也优于上述算法。直接视觉评估, 本文算法在细节和边缘处有更好的生成效果, 缓解了结构变形和虚构重影的现象, 视觉效

表2 不同算法在胶囊胃镜数据集上的指标测试结果

		SRN-DeblurNet	改进SPSR	DeblurGAN-v2	本文算法
峰值信噪比(PSNR)	迁移	26.93	28.33	28.79	28.83
	训练	28.01	28.70	28.81	29.04
结构相似度(SSIM)	迁移	0.777	0.767	0.783	0.801
	训练	0.795	0.771	0.813	0.826

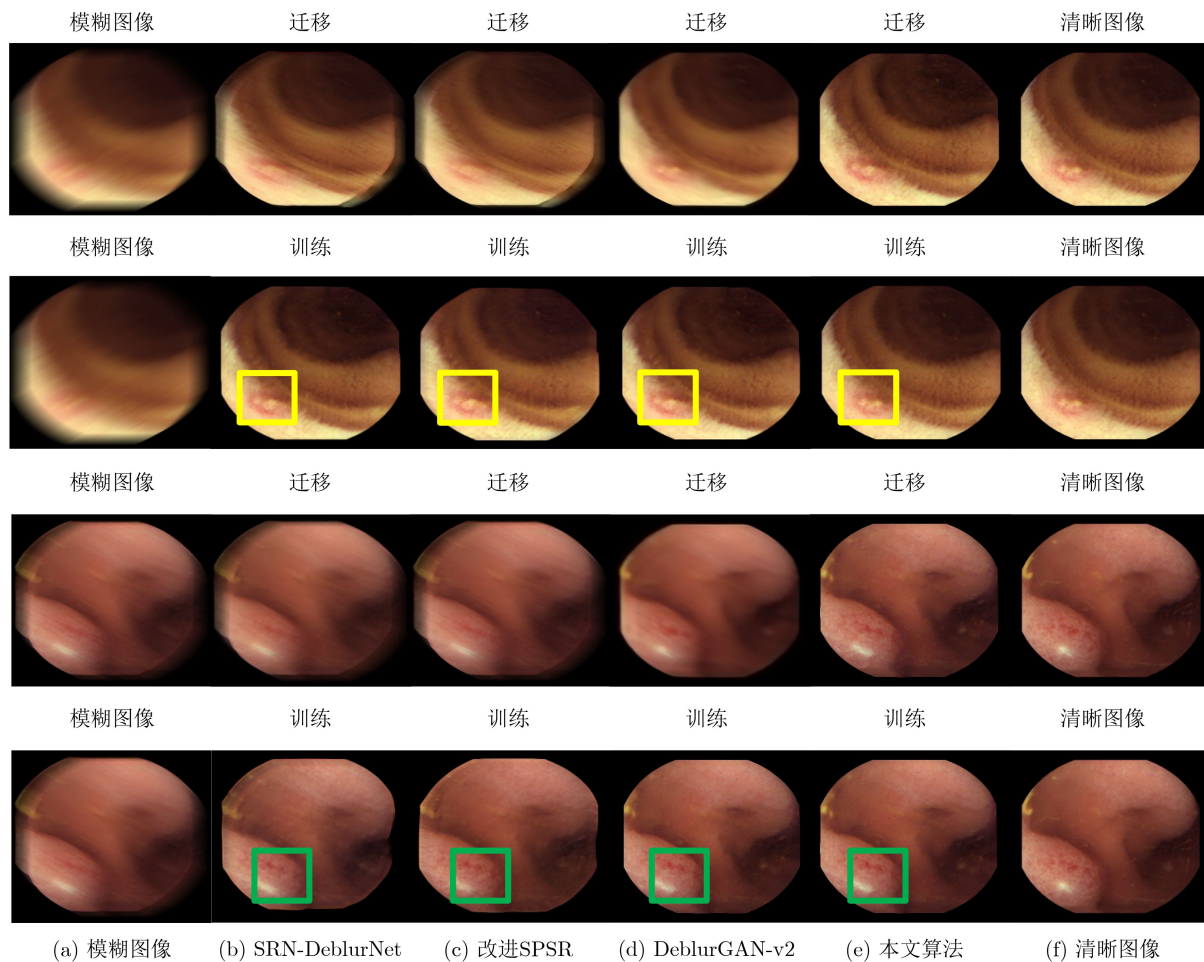


图6 不同算法在胶囊胃镜数据集上的图像测试结果

果更好。下一步工作,针对特征提取对全局信息的关注仍然不充分,可考虑采用计算机视觉中近期研究热点Transformer结构,Transformer本身具有很好的全局信息提取能力,与卷积神经网络结合,可更好地整合全局信息和局部信息,预期可获得更好的胃镜图像去模糊效果。

参考文献

- [1] All cancers source: Globocan 2020[EB/OL]. <https://gco.iarc.fr/today/data/factsheets/cancers/39-All-cancers-fact-sheet.pdf>, 2020.
- [2] MAHESH M M R, RAJAGOPALAN A N, and SEETHARAMAN G. Going unconstrained with rolling shutter deblurring[C]. 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017: 4030–4038. doi: [10.1109/ICCV.2017.432](https://doi.org/10.1109/ICCV.2017.432).
- [3] TAO Xin, GAO Hongyun, SHEN Xiaoyong, *et al.* Scale-recurrent network for deep image deblurring[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 8174–8182. doi: [10.1109/CVPR.2018.00853](https://doi.org/10.1109/CVPR.2018.00853).
- [4] KUPYN O, BUDZAN V, MYKHAILYCH M, *et al.* DeblurGAN: Blind motion deblurring using conditional adversarial networks[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 8183–8192. doi: [10.1109/CVPR.2018.00854](https://doi.org/10.1109/CVPR.2018.00854).
- [5] KUPYN O, MARTYNIUK T, WU Junru, *et al.* DeblurGAN-v2: Deblurring (orders-of-magnitude) faster and better[C]. The IEEE/CVF International Conference on Computer Vision, Seoul, Korea (South), 2019: 8877–8886. doi: [10.1109/ICCV.2019.00897](https://doi.org/10.1109/ICCV.2019.00897).
- [6] YAN Qing, XU Yi, YANG Xiaokang, *et al.* Single image superresolution based on gradient profile sharpness[J]. *IEEE Transactions on Image Processing*, 2015, 24(10): 3187–3202. doi: [10.1109/TIP.2015.2414877](https://doi.org/10.1109/TIP.2015.2414877).
- [7] ZHU Yu, ZHANG Yanning, BONEV B, *et al.* Modeling deformable gradient compositions for single-image super-resolution[C]. The IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 5417–5425. doi: [10.1109/CVPR.2015.7299180](https://doi.org/10.1109/CVPR.2015.7299180).
- [8] MA Cheng, RAO Yongming, CHENG Ye'an, *et al.* Structure-preserving super resolution with gradient guidance[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2020:

- 7766–7775. doi: [10.1109/CVPR42600.2020.00779](https://doi.org/10.1109/CVPR42600.2020.00779).
- [9] TRAN P, TRAN A T, PHUNG Q, *et al.* Explore image deblurring via encoded blur kernel space[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2021: 11951–11960. doi: [10.1109/CVPR46437.2021.01178](https://doi.org/10.1109/CVPR46437.2021.01178).
- [10] CHI Zhixiang, WANG Yang, YU Yuanhao, *et al.* Test-time fast adaptation for dynamic scene deblurring via meta-auxiliary learning[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2021: 9133–9142. doi: [10.1109/CVPR46437.2021.00902](https://doi.org/10.1109/CVPR46437.2021.00902).
- [11] CHEN Liang, ZHANG Jiawei, PAN Jinshan, *et al.* Learning a non-blind deblurring network for night blurry images[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2021: 10537–10545. doi: [10.1109/CVPR46437.2021.01040](https://doi.org/10.1109/CVPR46437.2021.01040).
- [12] DONG Jiangxin, ROTH S, and SCHIELE B. Learning spatially-variant MAP models for non-blind image deblurring[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2021: 4884–4893. doi: [10.1109/CVPR46437.2021.00485](https://doi.org/10.1109/CVPR46437.2021.00485).
- [13] GAO Shanghua, CHENG Mingming, ZHAO Kai, *et al.* Res2Net: A new multi-scale backbone architecture[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(2): 652–662. doi: [10.1109/TPAMI.2019.2938758](https://doi.org/10.1109/TPAMI.2019.2938758).
- [14] WANG Xintao, YU Ke, WU Shixiang, *et al.* ESRGAN: Enhanced super-resolution generative adversarial networks[C]. The European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 2019: 63–79. doi: [10.1007/978-3-030-11021-5_5](https://doi.org/10.1007/978-3-030-11021-5_5).
- [15] HUANG Gao, LIU Zhuang, VAN DER MAATEN L, *et al.* Densely connected convolutional networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 2261–2269. doi: [10.1109/CVPR.2017.243](https://doi.org/10.1109/CVPR.2017.243).
- [16] SIMONYAN K and ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]. 3rd International Conference on Learning Representations, San Diego, USA, 2015: 1–14. doi: [10.1109/cvpr.2016.182](https://doi.org/10.1109/cvpr.2016.182).
- [17] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, *et al.* Generative adversarial nets[C]. The 27th International Conference on Neural Information Processing Systems, Montreal, Canada, 2014: 2672–2680. doi: [10.1145/3422622](https://doi.org/10.1145/3422622).
- [18] RONNEBERGER O, FISCHER P, and BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]. 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 2015: 234–241. doi: [10.1007/978-3-319-24574-4_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [19] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Deep residual learning for image recognition[C]. The IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 770–778. doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [20] ALI S, ZHOU F, DAUL C, *et al.* Endoscopy artifact detection (EAD2019) challenge dataset[EB/OL]. <https://arxiv.org/abs/1905.03209>, 2019.
- [21] KOULAOUZIDIS A, IAKOVIDIS D K, YUNG D E, *et al.* KID Project: An internet-based digital video atlas of capsule endoscopy for research purposes[J]. *Endoscopy International Open*, 2017, 5(6): E477–E483. doi: [10.1055/s-0043-105488](https://doi.org/10.1055/s-0043-105488).
- [22] HORÉ A and ZIOU D. Image quality metrics: PSNR vs. SSIM[C]. 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 2010: 2366–2369. doi: [10.1109/ICPR.2010.579](https://doi.org/10.1109/ICPR.2010.579).
- [23] WANG Zhou, BOVIK A C, SHEIKH H R, *et al.* Image quality assessment: From error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600–612. doi: [10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861).
- [24] KINGMA D P and BA J. Adam: A method for stochastic optimization[C]. 3rd International Conference on Learning Representations, San Diego, USA, 2015.
- 时永刚: 男, 1969年生, 副教授, 研究方向为医学图像处理、目标检测识别、目标分类、图像复原和超分辨率重建。
- 张岳: 男, 1996年生, 硕士生, 研究方向为医学图像处理。
- 周治国: 男, 1977年生, 副教授, 研究方向为智能感知与导航。
- 李祎: 女, 1996年生, 硕士生, 研究方向为医学图像处理。
- 夏卓岩: 男, 1997年生, 硕士生, 研究方向目标检测。

责任编辑: 余蓉