

基于生成对抗网络与噪声分布的 图像超分辨率重建方法

王 晔¹, 孙志宽¹, 李 征^{1, 2}

(1. 四川大学计算机学院(软件学院), 成都 610065; 2. 四川大学天府工程数值模拟与软件创新中心, 成都 610207)

摘 要: 现有的图像超分辨率重建方法都较少考虑真实低分辨率图像中包含的噪声信息, 因此会影响图像的重建质量. 受真实图像去噪算法的启发, 本文引入一个噪声分布收集网络来收集低分辨率图像的噪声分布信息, 并采用生成对抗网络的模型设计, 提高含噪声图像的重建质量. 噪声分布信息会分别输入到超分辨率重建网络和判别网络, 在重建过程中去除噪声的同时保证有用高频信息的恢复, 另外由于判别网络的能力对整个模型的性能有着重要影响, 选择使用 U-Net 网络来获得更好的梯度信息反馈. 与经典图像超分辨率重建算法的对比以及消融实验表明, 使用噪声收集网络和 U-Net 判别网络后, 本文模型在噪声低分辨率图像重建任务中获得了更好的性能.

关键词: 图像超分辨率; 生成对抗网络; 真实图像; 噪声分布

中图分类号: TP391 **文献标识码:** A **DOI:** 10.19907/j.0490-6756.2023.032001

An image super-resolution reconstruction method based on Generative Adversarial Network and noise distribution

WANG Ye¹, SUN Zhi-Kuan¹, LI Zheng^{1, 2}

(1. College of Computer Science(College of Software), Sichuan University, Chengdu 610065, China;

2. Tianfu Engineering-Oriented Numerical Simulation & Software Innovation Center,
Sichuan University, Chengdu 610207, China)

Abstract: Existing image super-resolution reconstruction methods take less into account the noise information contained in real low-resolution images, which will affect the quality of image reconstruction. Inspired by the real image denoising algorithm, this paper introduces a noise distribution collection network to collect noise distribution information of low-resolution images, and adopts a model design of Generative Adversarial Network to improve the reconstruction quality of noisy images. The noise distribution information will be input to the super-resolution reconstruction network and the discriminant network respectively. During the reconstruction process, the noise is removed during while ensuring the recovery of useful high-frequency information, because the ability of the discriminant network has an important impact on the performance of the entire model, the U-Net network is selected to obtain better gradient information feedback. Comparison with the classical image super-resolution reconstruction methods and ablation experiments, the results show that the proposed model obtains better performance

收稿日期: 2022-06-20

基金项目: 国家重点研发计划项目(2020YFA0714003); 国家重大项目(GJXM92579); 四川省科技厅重点研发项目(2021YFQ0059)

作者简介: 王晔(1997-), 女, 山西晋城人, 硕士研究生, 主要研究领域为计算机视觉与图像处理. E-mail: yesical230@163.com

通讯作者: 李征. E-mail: lizheng@scu.edu.cn

in the noisy low-resolution image reconstruction task after using the noise collection network and the U-Net discriminant network.

Keywords: Image super-resolution; Generative Adversarial Network; Real images; Noise distribution

1 引言

随着科技的发展,作为重要信息载体的图像拥有了越来越高的分辨率.更高的像素密度展示了更多的纹理细节,蕴涵更丰富的信息,在很多场景中高分辨率图像都极具价值,比如 CT 图像、核磁共振图像、监控视频和卫星遥感图像等.但高分辨率的图像对成像设备和环境因素要求很高.在实际生活中,人们往往获取不到期望的高质量图片.因此,在软件层面上图像超分辨率重建技术满足了这样的需求^[1-3].

图像超分辨率重建技术需要为低分辨率图像恢复出更多的纹理信息,也就是拥有更多高频细节.真实场景中的自然低分辨率图像是通过相机镜头捕捉到的,经过了复杂的退化过程,常常包含不少的噪声信息,对图像重建过程并无用处.并且由于其图像也是高频信息,可能会在重建过程中被放大,从而生成伪影降低图像的重建质量.如果将低分辨率图像先进行去噪,然后再进行图像的超分辨率重建,则会导致部分有用的高频信息在去噪过程中被过滤丢失,尤其是一些纹理细节^[4].图像的先验信息被破坏,对后续的重建过程会有影响,使得图像重建质量一般.一些面向真实场景的图像超分辨率重建算法都是基于模糊核估计,在考虑噪声影响时只是添加一些人工合成的噪声信息.因此本文主要的贡献是:(1) 提出一个噪声收集网络用于收集图像真实的噪声分布信息,并能够拓展到已有的算法模型中;(2) 提出一个新的图像超分辨率网络模型,适用于含噪声的低分辨率图像重建.

2 相关工作

图像超分辨率重建算法的出现最早可追溯到 1960 年.早期的传统方法主要分为三大类:基于插值的方法、基于重建的方法以及基于学习的方法.基于插值的常见方法有双三次插值、双线性插值以及最邻近插值等.先将低分辨率图像放大到想要的尺寸,再将待插值的像素信息利用数学模型计算出来,获得高分辨率图像,运行速度快.但由于是线性运算,获得的高分辨率图像会出现局部平滑,效果一般.基于重建的方法主要基于频率域和空间域,

考虑图像的先验知识和退化过程进行建模,根据模型来逆向重建,经典方法有最大后验概率法^[5]、迭代反投影法^[6]以及凸集投影法^[7].基于学习的方法就不再人为定义图像的先验知识,通过低分辨率图像和高分辨率图像构成的样本数据学习映射关系来构建模型,主要方法有领域嵌入法^[8]、基于稀疏表示的算法^[9]以及锚定邻域回归法^[10]等.

随着卷积神经网络(Convolution Neural Network, CNN)的蓬勃发展,2014 年 Dong 等人首次提出了基于 CNN 的图像超分辨率算法 SRCNN^[9],网络类型为端到端,效果比传统算法更好.之后他们又提出了 FSRCNN^[12],速度更快,复杂度更低,但使用在特征重建时的反卷积操作使得结果出现了棋盘格.针对这一问题,Shi 等人于 2016 年提出采用亚像素卷积的神经网络结构^[13],但在网络深度上具有局限性.同年 Kim 等人受残差网络的启发提出了 VDSR^[14],使用全局残差大大加深了网络结构.同年,Ledig 等人提出了 SRResNet^[15],将全局残差与局部残差相结合. NTIRE2017 挑战赛^[16]中获得超分辨率冠军的 EDSR 算法^[17]开创性地移除了批量归一化. Tong 等人于同年提出 SRDenseNet^[18],使用了密集稠密连接.在 2018 年 Zhang 等人结合密集连接与残差连接提出了 RDN^[19],同年又提出了基于注意力机制的 RCAN^[20],500 层的网络深度获得了更好的重建效果.2019 年基于反馈机制, Li 等人提出了 SRFBN^[21],训练策略采用课程学习.同年 Muqeeet 等人提出 HRAN^[22],使用了残差注意力.2020 年 Hyun 等人提出 VarSR^[23],从一张低分辨率图像中恢复出多张高分辨率图像.2021 年, Qiao 等人提出基于傅里叶域注意力的网络模型^[24],应用于显微镜图像,拓展了图像超分辨率的应用领域.

生成对抗网络(Generative Adversarial Network, GAN)是由 Goodfellow 等人^[25]在 2014 年提出的,使用两个神经网络,分别为生成器和判别器,被广泛应用于图像生成和语音生成等领域. Ledig 等人于 2017 年提出了 SRGAN^[15],这是第一次在图像超分辨率领域使用 GAN 的网络结构.2018 年, Wang 等人提出了改进后的 ESR-

GAN^[26], 提高模型训练的稳定性。

图像超分辨率算法在人工数据集上的性能已接近瓶颈, 然而由于对图像退化过程的简化, 大多数算法在真实图像数据集上的表现不尽如人意, 因此面向真实场景的超分辨率算法获得了越来越多的关注。2018 年 Schocher 等人提出了 ZSSR^[27], 仅利用图像本身的内部信息进行重建, 是超分辨率重建领域首个基于 CNN 的无监督学习算法。2019 年 Fritsche 等人提出 FSSR^[28], 利用图像中的低频区域和高频区域。2021 年 Wang 等人提出 DASR^[29], 利用对比学习机制学习退化过程。

3 方 法

本文提出了一个基于生成对抗网络与噪声分布的图像超分辨率算法模型, 使用生成对抗的网络设计结构, 在超分辨率重建网络之前添加一个噪声分布收集网络 (Noise Distribution Collection Network, NDCN), 共同组成生成网络部分, 判别网络使用 U-Net 网络结构, 整体流程如图 1 所示。首先使用 NDCN 网络获取输入带噪声低分辨率图像中的噪声分布信息, 然后再将其与低分辨率图像一同输入到图像超分辨率重建网络, 重建出的图片与噪声分布信息一同输入到判别网络。

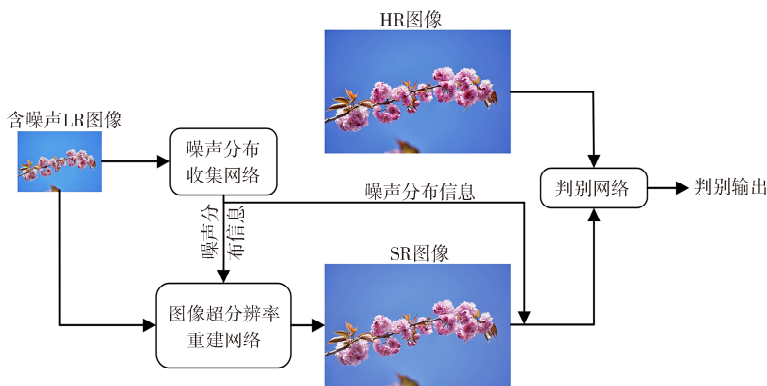


图 1 网络流程示意图
Fig. 1 Diagram of network process

3.1 NDCN 网络

Guo 等人在 2019 年提出了 CBDNet^[30] 用于真实噪声图像去噪。该网络模型由两部分组成, 分别为噪声估计子网络和非盲去噪子网络。受益于其中噪声估计子网络的设计, 本文提出了一个 NDCN 网络。CBDNet 中噪声估计子网络的结构非常简单, 使用了五层全卷积网络, 并且没有池化操作和批量归一化。为了能够收集到更深层次的特征信息, 更关注属于高频信息的噪声, 并且更加准确, 本文使用编码块-解码块的网络结构和自注意力机制, 另外使用跳跃连接使得深层特征与浅层特征可以实现多层次特征融合。第一层卷积后将特征输入到网络末端, 由于自注意力模块计算量较大, 仅在编码块的末尾添加一个自注意力模块。为了降低网络模型的计算量, 在编码块中进行降采样操作, 使用最大池化来使特征图尺寸减少为原来的一半。之后在解码块中恢复特征图尺寸, 采用的是反卷积上采样的方式, 这样便减少了全连接 CNN 中的计算开销。

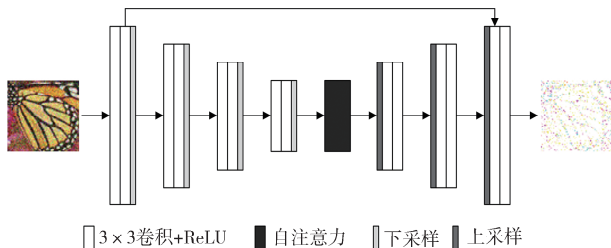


图 2 网络结构示意图
Fig. 2 Diagram of network structure

编码块由两个卷积层和 ReLU 激活函数组成。卷积核大小设置为 3×3 , 步长设置为 1, 特征图尺度减小的同时增加通道数, 解码块中的卷积核和编码块中设置相同。提取深层特征的同时压缩通道数, 最终 NDCN 网络输出与输入尺寸相同的特征图, 其中包含有噪声分布信息, 并与含有噪声的 LR 图像一同输入到重建网络中, 指导生成更好的图像重建结果。

自注意力模块可表示为:

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

其中, Q 表示查询向量矩阵; K 表示键向量矩阵; V 表示值向量矩阵; d_k 表示缩放因子. 自注意力机制使用不同的参数矩阵对 Q, K, V 执行多种线性变换, 计算 Q 和 K 之间的相似性, 重复多次后将结果拼接起来.

3.2 图像超分辨率重建网络

重建网络使用经典的密集连接和残差连接的方式. 在特征映射阶段使用引入了混合注意力模块 (Mixed Attentions Module, MAM) 的密集残差注意力块 (Residual Dense Attention Block, RD-

AB). 混合注意力包括空间注意力和频率通道注意力, 实现对不同的特征信息进行有针对性地处理, 在增强有用信息的同时, 抑制无用信息. 多个 RD-AB 组成一个密集残差注意力组 (Residual Dense Attention Group, RDAG), 其中用密集连接的方式将不同层级的特征图按通道相加, 结果输入到下一个 RDAG 模块中, 在减少计算量的同时保证多层级网络的特征复用. 另外为了减少模型的参数量, 在网络中使用 1×1 的卷积, 压缩通道数.

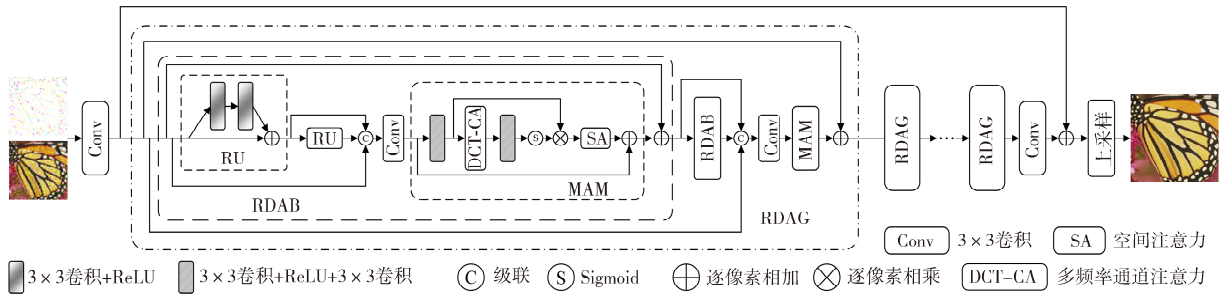


图 3 图像超分辨率重建网络整体结构图
Fig. 3 Diagram of image super-resolution reconstruction network

3.2.1 MAM 模块 在 CNN 中, 通道之间的信息相似性导致存在冗余, 并且直接使用全局平均池化处理会导致通道多样性信息丢失, 而空间注意力使网络更注重高频信息, 在关键区域增强特征表达, 因此本文在超分辨率重建网络中使用多频率通道注意力和空间注意力结合的 MAM 模块, 在通道中考虑不同位置空间关系的同时保持通道的多样性信息.

多频率通道注意力在 FcaNet^[31] 中提出, 在频域中应用全局平均池化, 采用 DCT 变换多频谱. 二维 DCT 可见式 (2), DCT 将二维输入变换到频率域见式 (3).

$$\varphi_{i,j} = \cos\left(\frac{\pi h}{H}\left(\frac{1}{2} + i\right)\right) \cos\left(\frac{\pi w}{W}\left(\frac{1}{2} + j\right)\right) \quad (2)$$

$$f_{h,w} = \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} \varphi_{i,j} x_{i,j}$$

$$\text{s. t. } h \in \{0, 1, \dots, H-1\}, w \in \{0, 1, \dots, W-1\} \quad (3)$$

其中, $\cos(\cdot)$ 表示余弦基函数, 不同的 h, w 产生不同的频率组合; $x_{i,j}$ 为输入信号; $f_{h,w}$ 为二维频谱.

在多频率通道注意力机制中, 会计算出所有频率分量的结果. 在其中挑选出表现性能最优的 K 个分量, K 值可根据特征图通道被划分出的组数

来设置. 在通道数为 C 的特征图输入后, 根据通道数划分出 n 个子特征图, 分配到每个子特征图中的通道数为 $C' = C/n$, 那么 n 个频率组合为 $[F^0, F^1, \dots, F^{n-1}]$. 根据式 (2) 计算得到相应的频率响应. 多频谱向量 F 即由这些频率组合压缩而成, 可见式 (4).

$$F = \text{compress}(X) = \text{concat}([F^0, F^1, \dots, F^{n-1}]) \quad (4)$$

其中, X 表示特征图; $\text{compress}(\cdot)$ 表示压缩操作; 将 $H \times W \times C$ 压缩为 $1 \times 1 \times C$; concat 表示向量串联.

最后使用全连接层和通道信息对多频谱向量进行整合, 使用 Sigmoid 激活函数计算注意力权重, 可见式 (5).

$$fca(X) = \sigma(fc(F)) H_{\text{scale}}(X) \quad (5)$$

其中, $fc(\cdot)$ 表示全连接层; $\sigma(\cdot)$ 表示激活函数; $H_{\text{scale}}(\cdot)$ 表示按通道相乘.

在多频谱通道注意力后加入空间注意力共同构成 MAM 模块, 对特征自适应性实现再校准, 多响应包含丰富信息的通道和空间特征, 同时抑制不重要的通道和空间特征.

3.2.2 RDAB 模块 对于图像超分辨率重建任务来说, 目标图像一般具有较少的高频信息和较多的低频信息. 网络模型要在此情况下重建出尽可能

多的高频信息,而低频信息基本变化不大,因此可以使用全局残差的方法. 低频信息直接映射到网络末端,残差学习部分使用高频信息.

批归一化处理(Batch Normalization, BN)可以加快模型收敛速度,同时一定程度缓解出现在深层网络的“梯度弥散”,因此几乎是卷积神经网络的标配处理技巧. 但在图像超分辨率重建过程中,BN 层会归一化图像的色彩分布,使得对比度信息受损,从而影响图像重建的质量. 另外 BN 层的计算会耗费大量资源,去掉后反而可以利用空间堆叠更多网络. 因此本文中的残差结构同样选择去除 BN 层. 残差单元计算过程见式(6).

$$f_{\text{RU}}(x) = x + \text{ReLU}(\text{Conv}(\text{ReLU}(\text{Conv}(x)))) \quad (6)$$

其中, x 表示输入; ReLU 表示激活函数; Conv 表示卷积.

RDAB 模块中包括残差单元和 MAM 模块. 输入的特征图通过残差单元学习高频信息,同时使用短跳跃连接将特征图输入到 Concat 层. 在 Concat 层中对特征图和不同的残差单元输出结果进行通道拼接. 使用该连接方式既能保留之前的输出结果,又能复用在深层次网络中学习的特征,提高了

网络的特征表达. 在 Concat 层后使用 1×1 卷积核进行通道压缩,再输入到 MAM 模块中计算通道和空间位置的相应权重,使用残差连接将 RDAB 模块的输入与 MAM 模块的输出相加,最终得到 RDAB 模块的输出.

3.2.3 RDAG 结构 为了构建更深的残差网络,不同于对残差网络层直接进行堆叠,RDAG 结构使用了残差内嵌残差的方式,更有利于模型训练^[17]. 短跳跃连接使得网络可以直接映射跳过大量低频信息,集中残差学习高频信息. 用公式表达网络中第 m 个 RDAG 可见式(7).

$$G_m = H_m(G_{m-1}) = H_m(H_{m-1}(\cdots H_1(G_0)\cdots)) \quad (7)$$

其中, H_m 表示第 m 个 RDAG 的过程; G_m 表示第 m 个 RDAG 的输出; G_{m-1} 表示第 m 个 RDAG 的输入; G_0 表示浅层特征.

3.3 判别网络

GAN 网络模型中使用的判别网络为 VGG19,在提取真实场景的图像特征时表现一般. 为了提高判别网络的能力,本文选择使用 U-Net 网络作为判别网络优化模型性能,使得重建出的高分辨率图像增强了细节信息.

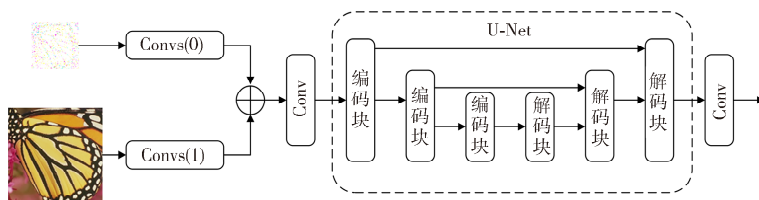


图 4 判别网络结构图
Fig. 4 Diagram of discriminant network

U-Net 网络是编码-解码的架构. 为了充分融合网络的上下文信息,从解码块到编码块使用跳跃连接的方式,对不同尺寸的特征图进行多层次的特征拼接,并且由全卷积层组成,可以处理任意形状和大小的图片. 编码块的下采样用最大池化的方式,使用了两个卷积层再加一个最大池化层,每个卷积层后加一个激活函数,每次降采样后使得通道数加倍. 解码块对特征图使用反卷积进行上采样,得到更大分辨率的特征图后拼接由编码块跳跃传递过来的特征图,结合了全局信息和局部信息,融合了深层特征和浅层特征,获得更多细节信息. 由于重建出的高分辨率图像与噪声分布信息图像的尺寸不同,使用了卷积层来调整大小为统一的 32×32 ,融合后输入到 U-Net 网络中.

3.4 损失函数

3.4.1 内容损失 对于图像的像素损失和噪声损失,均使用 L_1 损失函数. 噪声损失是通过噪声分布信息和低分辨率图像中的噪声来计算,分别表示为式(8)和式(9). 计算损失时对低分辨率图像使用高通滤波器,减少被低频信息影响.

$$L_1^G = \frac{1}{m} \sum_{i=1}^m \|I_{\text{SR}}^i - I_{\text{HR}}^i\|_1 \quad (8)$$

$$L_n^G = \frac{1}{m} \sum_{i=1}^m \|N^i - H(I_{\text{LR}}^i)\|_1 \quad (9)$$

其中, m 表示批处理图像的大小; N^i 表示第 i 个图像的噪声分布信息; $H(\cdot)$ 表示高通滤波; I_{LR}^i 表示第 i 个低分辨率图像.

3.4.2 感知损失 感知损失注重的是特征层面,

特征提取使用预训练好的 VGG19 网络, 仅使用内容损失, 会使得最后获得的图像比较平滑, 缺乏高频细节. 感知损失可以弥补这个问题. 其表达式如下.

$$L_p^G = \frac{1}{m} \sum_{i=1}^m \|VGG_l(I_{SR}^i) - VGG_l(I_{HR}^i)\|_1 \quad (10)$$

其中, $VGG_l(\cdot)$ 表示 VGG 在第 l 层输出的特征图.

3.4.3 对抗损失 生成对抗网络是非常具有开创性和应用前景的网络模型, 但也存在着难以训练的弊端. 过优的判别网络可能造成生成网络的梯度消失, 而差劲的判别网络无法对生成网络的训练进行有效指导. 因此对生成对抗网络损失函数的优化方案不断涌现. WGAN^[32] 引入了 Wasserstein 距离来量化生成样本分布和真实样本分布间的差距, 定义见式(11).

$$W(p_r, p_G) = \inf_{\gamma \in \Pi(p_r, p_G)} E_{(x,y) \sim \gamma} [\|x - y\|] \quad (11)$$

其中, x 表示真实样本; y 表示生成样本; p_r 表示真实样本分布; p_G 表示生成样本分布; $\Pi(p_r, p_G)$ 为 p_r 与 p_G 所有可能的联合分布集合; $\|x - y\|$ 表示样本间距离; $E_{(x,y) \sim \gamma}[\cdot]$ 表示期望值; \inf 表示下限; $W(p_r, p_G)$ 表示联合分布集合中的期望值下限. 但该公式不能直接求解, 需要满足 Lipschitz 连续才能求解, 因此近似表示为式(12).

$$W(p_r, p_G) = \frac{1}{K} \sup_{\|f\|_L \leq K} E_{x \sim p_r} [f(x)] - E_{y \sim p_G} [f(y)] \quad (12)$$

其中, $\|f\|_L$ 表示 Lipschitz 常数; K 表示一个非负的常数; \sup 表示上限.

使用参数组 w 定义一系列函数, 同时生成网络与 $E_{x \sim p_r} [f(x)]$ 无关, 因此可获得生成网络和判别网络的损失函数分别为式(13)和式(14).

$$L_G = -E_{y \sim p_G} [f_w(y)] \quad (13)$$

$$L_D = E_{y \sim p_G} [f_w(y)] - E_{x \sim p_r} [f_w(x)] \quad (14)$$

之后在使用 Wasserstein 距离的基础上, WGAN-GP^[33] 又在判别网络中引入了梯度约束项来帮助满足 Lipschitz 连续, 定义见式(15).

$$(\|\nabla_x D(x)\|_2 - 1)^2 \quad (15)$$

本文超分辨率重建网络使用的对抗损失函数为

$$L_{adv}^G = -E_{I_G^{SR} \sim p_G} [D(I_G^{SR})] \quad (16)$$

在判别网络中添加了梯度约束项的对抗损失函数为

$$L_{adv}^G = -E_{I_G^{SR} \sim p_G} [D(I_G^{SR})] L_{adv}^D = E_{I_G^{SR} \sim p_G} [D(I_G^{SR})] - E_{I_R^{HR} \sim p_r} [D(I_R^{HR})] + \lambda E_{I_G^{SR} \sim p_G} [(\|\nabla_{I_G^{SR}} D(I_G^{SR})\|_2 - 1)^2] \quad (17)$$

本文中超分辨率重建网络的损失函数包括内容损失、感知损失以及对抗损失三部分. 损失函数表示为

$$L_G = \lambda_A L_{adv}^G + \lambda_C L_1^G + \lambda_N L_n^G + \lambda_P L_p^G \quad (18)$$

其中, λ_A 、 λ_C 、 λ_N 与 λ_P 为损失函数的权重系数. 在本文模型中, 我们设置 $\lambda_A = 5 \times 10^{-2}$, $\lambda_C = 1 \times 10^{-2}$, $\lambda_N = 1 \times 10^{-2}$, $\lambda_P = 1$.

4 实验

4.1 数据集

本文基于真实图像的噪声分布设计网络模型, 使用真实场景的图像数据集 DIV2K^[14] 和 RealSR^[34] 作为训练数据集. DIV2K 在 NTIRE2017 时被提出, 在图像超分辨率重建领域应用非常广泛, 作为 NTRIE 和 PIPM 系列竞赛的基准数据集, 其中包含了 1000 张高质量图片, 涵盖了丰富的场景, 有人物、风景以及生活场景等, 分辨率达到 2 K, 有 800 张作为训练数据, 100 张作为验证数据, 100 张作为测试数据. RealSR 数据集集中的图片采集自 Canon 5D3 和 Nikon D810 两种相机, 通过调整相机的焦距获得具有不同尺度因子的低分辨率-高分辨率图像对. 数据集有 3 个版本, 版本 1 有 234 个场景, 204 个用于训练, 30 个用于测试; 版本 2 和版本 3 都是包含 559 个场景, 其中 459 个用于训练, 100 个用于测试, 不同在于版本 2 的高质量图片和低质量图片有着相同的分辨率, 版本 3 是不同的分辨率. 测试数据集使用 AIM2019 测试集和 RealSR 测试集. AIM2019 测试集数据来自于 DIV2K, 制作低分辨率图像时使用了不同的退化过程, 包括随机的噪声信息, 与 RealSR 数据集相比有较高的噪声强度.

4.2 评价指标

在图像超分辨率重建领域, 验证模型有效性常见的评价指标是峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 和结构相似性 (Structural Similarity, SSIM)^[35]. 模型性能越好, 获得的数值越高. 而使用的评价指标学习感知图像块相似度 (Learned Perceptual Image Patch Similarity, LPIPS)^[36] 则是数值越低, 模型性能越好.

4.2.1 PSNR PSNR 是信号最大功率和破坏性噪声功率的比值. 由于信号的动态范围比较大, 表

示时使用对数分贝单位,其定义为

$$\text{PSNR}(I_{\text{HR}}, I_{\text{SR}}) = 10 \times \log_{10} \frac{(2^n - 1)^2}{\text{MSE}(I_{\text{HR}}, I_{\text{SR}})}$$

(19)

其中, MSE 表示两个图像的均方误差; n 表示像素的位深, 一般取值为 8, 则像素的最大值为 255. MSE 越小, PSNR 越大, 则表明图像相似度越高, 在图像超分辨率重建领域则表明图像重建质量越好. PSNR 高于 38 dB 时人眼便很难分辨图像之间的差异, 低于 20 dB 时图像质量便很差, 不能被接受.

4.2.2 SSIM SSIM 是从图像组成的角度来估计两个图像之间的相似度. 结构信息、亮度和对比度组合起来共同反应失真建模. 给定图像 x 和 y , SSIM 定义表示为

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

(20)

其中, μ_x 表示 x 的平均值; σ_x^2 表示 x 的方差; μ_y 表示 y 的平均值; σ_y^2 表示 y 的方差; σ_{xy} 表示 x 和 y 的协方差; c_1 和 c_2 表示常数, 用来维持稳定. SSIM 值的范围是 $[0, 1]$, 值为 1 时, 表示两幅图像完全相同.

4.2.3 LPIPS LPIPS 用来度量生成图像和真实图像的特征差异, 与人类的感知更加一致, 需要将两个图像输入到网络中, 计算方式表示为

$$d(x, x_0) = \sum_l \frac{1}{H_l W_l} \sum_{h,w} \|w_l \odot (\hat{y}_{hw}^l - \hat{y}_{0hw}^l)\|_2^2$$

(21)

其中, x 表示真实图像; x_0 表示生成图像; $H_l W_l$ 表示第 l 层特征图大小; \hat{y}_{hw}^l 和 \hat{y}_{0hw}^l 分别表示真实图像和生成图像第 l 层提取的特征; w_l 表示通道向量, 在每个通道中计算 L_2 获得差异, 最后计算加权平均值.

4.3 实验设置

为获得更多训练数据, 对数据集进行数据增强, 分别进行 90° 、 180° 以及 270° 旋转, 对旋转后的图像再做水平方向和垂直方向的镜像反转. 由于硬件设备有限, 在训练前将图像裁剪为 32×32 的图像块, 重建后的图像块为 128×128 .

训练网络的过程中, 网络参数的迭代优化使用 Adam 算法, 其中设置 $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\beta = 10^{-8}$, 学习率初值设置为 1×10^{-4} , 在训练过程中按周期减半. 实验过程使用到的硬件环境 CPU 为 Intel i5-10400F @2.90 GHz, 32 G 的计算内存, GPU 为 NVIDIA GTX 1080Ti, 系统环境为 Windows 10, 使用深度学习框架 PyTorch 和开发环境 PyCharm.

4.4 方法对比

在图像超分辨率重建任务中, 4 倍放大的重建难度最高, 其效果也更能展现模型性能, 由于篇幅原因, 本文仅给出了不同方法在 RealSR 测试集和 AIM2019 测试集进行 4 倍放大重建的实验结果.

由表 1 可看出, 本文模型在含有噪声信息的低分辨率图像重建结果中具有最优表现. 相比基于真实场景的无监督网络模型 DASR, PSNR 增长了 0.66 dB, SSIM 增长了 0.0308, LPIPS 降低了 0.085, 在各项指标中有着不同程度的提升. 另外相比于经典的 SRGAN 方法提升较大, PSNR 增长了 2.7, SSIM 增长了 0.0652, LPIPS 降低了 0.962.

由于 AIM2019 测试集包含较多的噪声信息, 在表 1 中可看到, 较于 RealSR 测试集整体的指标表现均有下降, 但本文模型的重建表现仍为最佳. 相较于 DASR, PSNR 增长了 0.44 dB, SSIM 增长了 0.0339, LPIPS 降低了 0.118.

表 1 不同方法×4 重建时在 RealSR 和 AIM2019 测试集的实验数据表

Tab. 1 Experimental data table of different methods under ×4 super-resolution in RealSR and AIM2019 test datasets

		SRCNN	SRGAN	ESRGAN	RCAN	ZSSR	DASR	Ours
RealSR	PSNR/dB	23.65	24.17	25.91	25.95	26.00	26.21	26.87
	SSIM	0.6784	0.7312	0.7466	0.7479	0.7483	0.7656	0.7964
	LPIPS	1.325	1.134	0.415	0.395	0.386	0.257	0.172
AIM2019	PSNR/dB	20.09	20.71	21.38	21.53	21.68	21.54	21.98
	SSIM	0.4638	0.5084	0.5469	0.5837	0.6007	0.5728	0.6067
	LPIPS	0.793	0.689	0.547	0.509	0.634	0.471	0.353

除了使用 PSNR、SSIM 和 LPIPS 这样的量化指标来验证本文模型的效果,还可通过视觉效果图来进行分析. 本文给出了 AIM2019 测试集中“0846”、“0883”和“0898”的结果展示. 从图 5 中可以看出,低分辨率图像包含较明显的噪声信息,

SRCNN、SRGAN、ESRGAN、RCAN 和 ZSSR 的重建效果一般,DASR 稍好一点,但噪声问题仍存在,无法较好抑制噪声. 而本文模型重建出的结果更清晰,视觉效果与高分辨率图像更接近,较好地抑制了噪声.

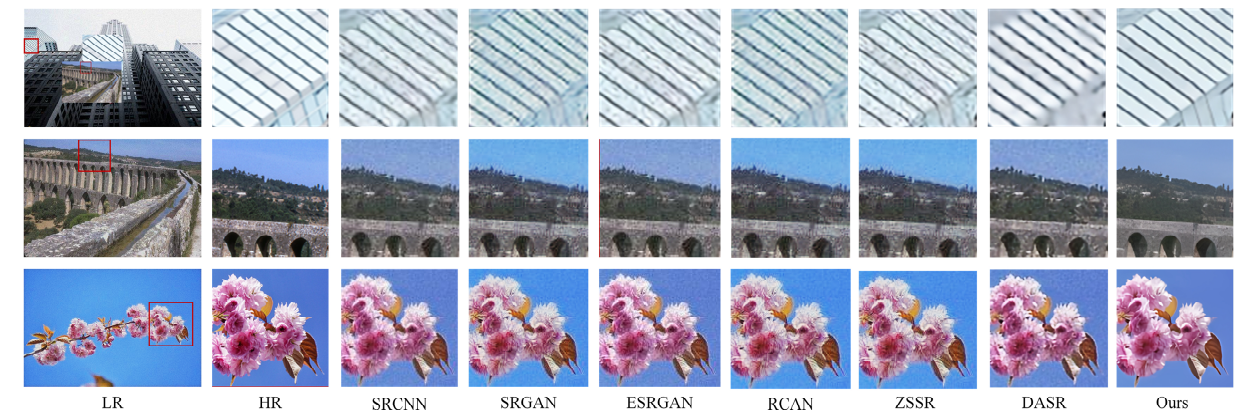


图 5 重建结果对比图
Fig. 5 Comparisons of reconstruction results

4.5 消融实验

为验证本文模型中 NDCN 网络的优势,用经典去噪方法中值滤波(Median filter)处理图像后再使用本文的重建网络. 另外对 NDCN 网络和判别网络两部分进行消融实验对比(结果如表 2),去掉噪声收集网络,并使用传统的 VGG19 判别网络,然后再分别添加噪声收集网络和使用 U-Net 判别网络.

表 2 NDCN 与 U-Net 消融实验结果					
Tab. 2 Results of NDCN and U-Net ablation					
实验设置	NDCN	U-Net	PSNR/dB	SSIM	LPIPS
With Median filter	×	√	26.09	0.7728	0.312
No NDCN+U-Net	×	×	26.14	0.7756	0.307
With NDCN	√	×	26.57	0.7828	0.223
With U-Net	×	√	26.46	0.7791	0.258
With NDCN+U-Net	√	√	26.87	0.7964	0.172

×表示去掉某个网络;√表示添加某个网络

由表 2 可知,对低分辨率图像先进行去噪再进行图像超分辨率重建的方法性能较差,说明将图像去噪与超分辨率重构融合处理的思想具有一定的优势. 另外本文网络模型与不使用噪声收集网络和 U-Net 判别网络的设置下相比,PSNR 值增长了 0.73 dB,SSIM 增长了 0.0208,LPIPS 下降了 0.135. 在与分别不使用噪声收集网络和 U-Net 判别网络的情况对比时,指标体现的效果均有不同程度的下降,说明噪声收集网络和 U-Net 判别网络

都对网络模型的效果提升做出了贡献.

5 结 论

本文基于存在噪声信息的真实场景低分辨率图像来展开图像超分辨率研究. 直接对含有噪声信息的低分辨率图像进行重建会影响图像的重建质量,而如果在重建前进行去噪处理,会使得对重建过程有用的高频信息也被过滤掉,原有信息被破坏,提高了图像重建的难度. 因此本文提出了一种新的基于噪声分布的真实图像超分辨率算法模型,使用生成对抗网络的结构设计,通过引入 NDCN 网络,能够获取低分辨率图像的噪声分布,用于指导生成网络和判别网络,并在此基础上利用 U 型网络提高判别网络的判断能力,在完成图像重建任务的同时降低噪声信息对图像质量的影响.

虽然本文网络模型在包含噪声信息的真实场景图像超分辨率重建任务中有着不错表现,但真实场景下的自然图像超分辨率重建还面临着很多问题. 比如,现在用于超分辨率重建的真实场景数据集包含成对低分辨率和高分辨率图像的数量并不多. 未来可从两个方面考虑该问题的解决方案. 一方面是从扩充数据集的角度,考虑真实低分辨率图像的多种退化过程,生成更贴近自然的低分辨率图像;另一方面是从非成对数据集训练方法的角度,只需要任意真实场景的高分辨率图像和低分辨率图像就可进行训练.

参考文献:

- [1] 李千, 赵逢禹. 融合多路径与混合注意力的遥感图像超分辨率重建[J/OL]. 小型微型计算机系统, 录用定稿[2022-06-02]. <https://kns.cnki.net/kcms/detail/21.1106.TP.20220513.1247.014.html>.
- [2] 夏英, 黄秉坤. 采用改进 YOLOv3 的高分辨率遥感图像目标检测[J]. 重庆邮电大学学报: 自然科学版, 2022, 34: 383.
- [3] 朱海琦, 李宏, 李定文, 等. 基于生成对抗网络的单图像超分辨率重建[J]. 吉林大学学报: 理学版, 2021, 59: 1491.
- [4] 张意, 阚子文, 邵志敏, 等. 基于注意力机制和感知损失的遥感图像去噪[J]. 四川大学学报: 自然科学版, 2021, 58: 042001.
- [5] Schultz R R, Stevenson R L. Extraction of high-resolution frames from video sequences [J]. IEEE T Image Process, 1996, 5: 996.
- [6] Irani M, Peleg S. Improving resolution by image registration [J]. CVGIP: Graph Mod Image Process, 1991, 53: 231.
- [7] Elad M, Feuer A. Restoration of a single super-resolution image from several blurred, noisy, and undersampled measured images [J]. IEEE T Image Process, 1997, 6: 1646.
- [8] Chang H, Yeung D Y, Xiong Y. Super-resolution through neighbor embedding [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2004.
- [9] Yang J, Wright J, Huang T S, *et al.* Image super-resolution via sparse representation [J]. IEEE T Image Process, 2010, 19: 2861.
- [10] Timofte R, De Smet V, Van Gool L. Anchored neighborhood regression for fast examplebased super-resolution [C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Computer Society, 2013.
- [11] Dong C, Loy C C, He K, *et al.* Image super-resolution using deep convolutional networks [J]. IEEE T Pattern Anal, 2015, 38: 295.
- [12] Dong C, Loy C C, Tang X. Accelerating the super-resolution convolutional neural network [C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2016.
- [13] Shi W, Caballero J, Huszár F, *et al.* Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2016.
- [14] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2016.
- [15] Ledig C, Theis L, Huszár F, *et al.* Photo-realistic single image super-resolution using a generative adversarial network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2017.
- [16] Agustsson E, Timofte R. Ntire 2017 challenge on single image super-resolution: Dataset and study [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2017.
- [17] Lim B, Son S, Kim H, *et al.* Enhanced deep residual networks for single image superresolution [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2017.
- [18] Tong T, Li G, Liu X, *et al.* Image super-resolution using dense skip connections [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2017.
- [19] Zhang Y, Tian Y, Kong Y, *et al.* Residual dense network for image super-resolution [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2018.
- [20] Zhang Y, Li K, Li K, *et al.* Image super-resolution using very deep residual channel attention networks [C]//Proceedings of the European Conference on Computer Vision. Berlin, German: Springer, 2018.
- [21] Li Z, Yang J, Liu Z, *et al.* Feedback network for image super-resolution [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2019.
- [22] Muqeet A, Iqbal M T B, Bae S H. HRAN: Hybrid residual attention network for single image super-resolution [J]. IEEE Access, 2019, 7: 137020.
- [23] Hyun S, Heo J P. VarSR: Variational super-resolution network for very low resolution images [C]//Proceedings of the European Conference on Comput-

- er Vision. Berlin: Springer, 2020.
- [24] Qiao C, Li D, Guo Y, *et al.* Evaluation and development of deep neural networks for image super-resolution in optical microscopy [J]. Nat Methods, 2021, 18: 194.
- [25] Goodfellow I, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets [C]//Neural Information Processing Systems. Cambridge: MIT Press, 2014.
- [26] Wang X, Yu K, Wu S, *et al.* Esrgan: enhanced super-resolution generative adversarial networks [C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2018.
- [27] Shocher A, Cohen N, Irani M. “zero-shot” super-resolution using deep internal learning [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2018.
- [28] Fritsche M, Gu S, Timofte R. Frequency separation for real-world super-resolution [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision Workshop. Piscataway: IEEE Computer Society, 2019.
- [29] Wang L, Wang Y, Dong X, *et al.* Unsupervised degradation representation learning for blind super-resolution [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2021.
- [30] Guo S, Yan Z, Zhang K, *et al.* Toward convolutional blind denoising of real photographs [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2019.
- [31] Qin Z, Zhang P, Wu F, *et al.* Fcanet: frequency channel attention networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2021.
- [32] Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks [C]//Proceedings of the International Conference on Machine Learning. New York: PMLR, 2017.
- [33] Gulrajani I, Ahmed F, Arjovsky M, *et al.* Improved training of wasserstein GANs [J]. NeurIPS, 2017, 30: 5767.
- [34] Cai J, Zeng H, Yong H, *et al.* Toward real-world single image super-resolution; a new benchmark and a new model [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2019.
- [35] Wang Z, Bovik A C, Sheikh H R, *et al.* Image quality assessment: from error visibility to structural similarity [J]. IEEE T Image Process, 2004, 13: 600.
- [36] Zhang R, Isola P, Efros A A, *et al.* The unreasonable effectiveness of deep features as a perceptual metric [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2018.

引用本文格式:

中 文: 王晔, 孙志宽, 李征. 基于生成对抗网络与噪声分布的图像超分辨率重建方法[J]. 四川大学学报: 自然科学版, 2023, 60: 032001.

英 文: Wang Y, Sun Z K, Li Z. An image super-resolution reconstruction method based on Generative Adversarial Network and noise distribution [J]. J Sichuan Univ: Nat Sci Ed, 2023, 60: 032001.