



(12) 发明专利

(10) 授权公告号 CN 119784591 B

(45) 授权公告日 2025. 07. 04

(21) 申请号 202411836612.4

G06N 3/045 (2023.01)

(22) 申请日 2024.12.13

G06N 3/0475 (2023.01)

G06N 3/094 (2023.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 119784591 A

(56) 对比文件

CN 115346045 A, 2022.11.15

CN 118172248 A, 2024.06.11

(43) 申请公布日 2025.04.08

(73) 专利权人 太原理工大学

审查员 白硕

地址 030024 山西省太原市迎泽大街79号

(72) 发明人 张军 梁铮 邓红霞 张瑞欣

杨雨晴 王培泉 李骐骋

(74) 专利代理机构 太原荣信德知识产权代理事

务所(特殊普通合伙) 14119

专利代理师 杜江涛

(51) Int. Cl.

G06T 3/4053 (2024.01)

G06T 3/4046 (2024.01)

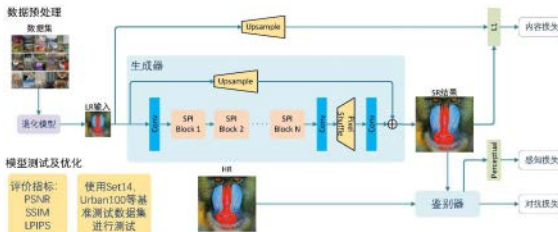
权利要求书1页 说明书6页 附图3页

(54) 发明名称

基于超像素注意力机制的真实世界图像超分重建方法

(57) 摘要

本发明属于深度学习技术领域,具体涉及一种基于超像素注意力机制的真实世界图像超分重建方法,包括以下步骤:构建数据集;构建图像超分模型SPGAN;利用所属训练集在所构建的SPGAN模型上进行训练;获取待超分重建的真实世界低分辨率图像,并通过训练完毕的SPGAN模型进行真实世界图像超分辨率重建;建立评价指标用于对SPGAN模型的评估。本发明是引入了超像素模块SPI, SPI模块包括超像素分割模块,超像素交叉注意力模块,超像素内部注意力模块,局部注意力模块四部分整体使用残差结构连接,增强了模型对边缘像素的感知能力,减少重建图像的模糊伪影等问题,提升了模型超分辨率重建的效果。



1. 基于超像素注意力机制的真实世界图像超分重建方法,其特征在于,包括下列步骤:

S1、以GAN作为基础结构,在生成器网络中融入超像素注意力机制,构建真实世界图像超分模型SPGAN;所述真实世界图像超分模型SPGAN进行超分重建的方法为:将通过二阶段门控退化后的低分辨率图像传入到生成器,图像通过N个超像素块SPI进行特征捕获,并通过Pixel Shuffle上采样操作提高重建质量,同时结合整体残差结构得到最后的SR图像;最后将此SR图像与其对应的HR图像一并传入到鉴别器当中判别图像的真实性,通过二者的对抗训练,提高图像的超分重建效果;

S2、引入二阶段门控退化模型,通过调整模糊核、大小、噪声类型和进行JPEG压缩得到真实世界退化后图像;

S3、结合超像素注意力设计超像素模块,从感知上将相似的像素分组在一起;所述S3中结合超像素注意力设计超像素模块的方法为:包括超像素分割,超像素交叉注意力,超像素内部注意力,局部注意力模块,整体采用残差结构简化训练;超像素内部注意力采用两个级联的线性层替代自注意力操作中的两次矩阵乘法操作,使得模型有更低的复杂度和更好的性能;并联多个超像素注意力模块可以更好的强化模型对局部和全局特征的关注;

S4、结合L1损失、感知损失、全变分正则化和GAN损失训练SPGAN,并采用指数移动平均EMA,以获得更稳定的训练和更好的表现。

2. 根据权利要求1所述的基于超像素注意力机制的真实世界图像超分重建方法,其特征在于,所述S1中真实世界图像超分模型SPGAN的主体框架是一个GAN结构,生成器用于退化后的低分辨率图像LR的超分过程,鉴别器负责判断超分辨率SR图像和高分辨率HR图像的相对真实性。

3. 根据权利要求1所述的基于超像素注意力机制的真实世界图像超分重建方法,其特征在于,所述S2中引入二阶段门控退化模型的方法为:二阶段门控退化模型包括两个相同的阶段,其中每个阶段包括模糊核、调整大小、添加噪声和JPEG压缩四个部分;数学模型表示为: $I^{LR} = D_g^2(I^{HR}) = (\sigma D_k^1 \cdot \sigma D_s^1 \cdot \sigma D_n^1 \cdot \sigma D_j^1 \cdot \sigma D_k^2 \cdot \sigma D_s^2 \cdot \sigma D_n^2 \cdot \sigma D_j^2)(I^{HR})$ 其中 D_g^2 表示二阶门控退化模型,其中的 D^1 和 D^2 分别表示第一阶段和第二阶段的退化操作, D_k 表示模糊核退化, D_s 表示下采样退化, D_n 表示噪声退化, D_j 表示JPEG压缩退化; σ 表示门控概率,用于控制对应的退化操作是否发挥作用;通过上述四个部分融入了门控操作机制,实现对退化操作强度的灵活调控,以更加精准地模拟出贴合实际的退化现象。

4. 根据权利要求1所述的基于超像素注意力机制的真实世界图像超分重建方法,其特征在于,所述S4中结合L1损失、感知损失、全变分正则化和GAN损失训练SPGAN的方法为:

首先,采用L1损失来计算超分图像与真实图像之间的平均误差,损失公式定义如下:

$$MAE = \frac{1}{n} \sum_{i=1}^n |f(x) - y_i| \text{ 其中 } n \text{ 为样本总数, } f(x) \text{ 为超分后图像, } y_i \text{ 为真实世界图像, L1 损失函数通过计算 } f(x) \text{ 和 } y_i \text{ 的差值绝对值来迭代优化模型;}$$

数通过计算 $f(x)$ 和 y_i 的差值绝对值来迭代优化模型;

训练一个面向PSNR的模型,训练得到的模型以SPNet命名;

然后,使用训练好的面向PSNR的模型作为生成器的初始化,并结合L1损失、感知损失、全变分正则化和GAN损失训练SPGAN。

基于超像素注意力机制的真实世界图像超分重建方法

技术领域

[0001] 本发明属于深度学习技术领域,具体涉及一种基于超像素注意力机制的真实世界图像超分重建方法。

背景技术

[0002] 真实世界图像超分辨率(RSISR)是计算机视觉中的一个重要研究课题,在现实生活中存在大量针对真实世界图像进行采集的设备,而由于设备性能等因素的限制,所采集到的真实世界图像往往是质量不足的低分辨率图像(LR),存在着细节丢失、色彩不准确、压缩伪影等问题,严重影响了医学影像、卫星遥感、视频监控和交通检测等计算机视觉任务的效果。因此,研究如何从真实世界中获取的低分辨率图像(LR)恢复出高分辨率图像(HR)具有非常重要的现实意义。

[0003] 真实世界图像超分辨率(RSISR)与普通的单幅图像超分辨率(SISR)不同,真实世界中图像的退化过程非常复杂,包含如模糊、噪声、失真等,并且退化后低分辨率图像可能无法保留真实世界图片的纹理等细节特征。同时,真实世界图像超分辨率任务中很难采集到对齐得很好的低分辨率和高分辨率图像对(LR-HR),这使得一些基于图像对的监督学习方法很难获得较好的超分效果,泛化效果也比较差。

[0004] 最近,基于退化建模的方法在一些真实世界超分辨率任务上取得了较大进展。利用人工设计的退化模型对HR图像进行退化,生成出对应的LR图像,然后使用HR-LR图像对进行模型训练。这种方法的核心在于模拟真实世界中图像从高分辨率到低分辨率的退化过程,包括模糊、压缩、噪声等因素的影响。通过这种方式,可以生成大量的合成数据集,用于训练和优化超分辨率模型。这些合成数据集不仅能够提供足够的训练样本,而且能够模拟不同的退化场景,增强模型的泛化能力。

[0005] 在实际应用中,退化建模方法能够显著提高超分辨率重建的性能。例如,通过混合退化操作和设置不同的退化参数,可以生成具有大范围退化作用的LR图像。这种方法的优势在于,它能够轻松地得到非常多成对的退化LR图像,不需要费力地收集,也不需要忍受不成对训练数据的不对准问题。此外,通过改变退化参数,可以容易地得到多种退化模型,从而提高模型在不同场景下的适应性和鲁棒性。

[0006] 随着深度学习技术的发展,基于深度学习的真实世界超分辨率技术也得到了积极的探索与发展。目前已经提出了各种基于深度学习的真实世界图像超分辨率方法,且在公开数据集上取得了很好的重建效果。这些方法通过学习图像的映射关系,能够有效学习图像中的高频信息,包括结构、纹理等,从而在低分辨率图像中恢复出丰富的细节和结构。因此,基于退化建模的超分辨率技术在实际应用中具有广阔的前景,尤其是在医学诊断、遥感图像、计算机视觉研究等领域。

发明内容

[0007] 针对上述的技术问题,本发明提供了一种基于超像素注意力机制的真实世界图像

超分重建方法,可以生成出更接近于真实世界退化特点的低分辨率退化图像,从而在真实世界场景下获得更加优异的超分性能。同时提出了基于超像素的真实世界图像超分方法,融合GAN的网络架构,利用超像素模块有效加强模型对图像边缘像素的关注,进一步提升超分结果的精细度。

[0008] 为了解决上述技术问题,本发明采用的技术方案为:

[0009] 基于超像素注意力机制的真实世界图像超分重建方法,包括下列步骤:

[0010] S1、以GAN作为基础结构,在生成器网络中融入超像素注意力机制,构建真实世界图像超分模型SPGAN;

[0011] S2、引入二阶段门控退化模型,通过调整模糊核、大小、噪声类型和进行JPEG压缩得到真实世界退化后图像;

[0012] S3、结合超像素注意力设计超像素模块,从感知上将相似的像素分组在一起;

[0013] S4、结合L1损失、感知损失、全变分正则化和GAN损失训练SPGAN,并采用指数移动平均EMA,以获得更稳定的训练和更好的表现。

[0014] 所述S1中真实世界图像超分模型SPGAN的主体框架是一个GAN结构,生成器用于退化后的低分辨率图像LR的超分过程,鉴别器负责判断超分辨率SR图像和高分辨率HR图像的相对真实性。

[0015] 所述真实世界图像超分模型SPGAN进行超分重建的方法为:将通过二阶段门控退化后的低分辨率图像传入到生成器,图像通过N个超像素块SPI进行特征捕获,并通过Pixel Shuffle上采样操作提高重建质量,同时结合整体残差结构得到最后的SR图像;最后将此SR图像与其对应的HR图像一并传入到鉴别器当中判别图像的真实性,通过二者的对抗训练,提高图像的超分重建效果。

[0016] 所述S2中引入二阶段门控退化模型的方法为:二阶段门控退化模型包括两个相同的阶段,其中每个阶段包括模糊核、调整大小、添加噪声和JPEG压缩四个部分;数学模型可表示为:

$$[0017] \quad I^{LR} = D_g^2(I^{HR}) = (\sigma D_k^1 \cdot \sigma D_s^1 \cdot \sigma D_n^1 \cdot \sigma D_j^1 \cdot \sigma D_k^2 \cdot \sigma D_s^2 \cdot \sigma D_n^2 \cdot \sigma D_j^2)(I^{HR})$$

[0018] 其中 D_g^2 表示二阶段门控退化模型,其中的 D^1 和 D^2 分别表示第一阶段和第二阶段的退化操作, D_k 表示模糊核退化, D_s 表示下采样退化, D_n 表示噪声退化, D_j 表示JPEG压缩退化; σ 表示门控概率,用于控制对应的退化操作是否发挥作用;通过上述四部分融入了门控操作机制,实现对退化操作强度的灵活调控,以更加精准地模拟出贴合实际的退化现象。

[0019] 所述S3中结合超像素注意力设计超像素模块的方法为:包括超像素分割,超像素交叉注意力,超像素内部注意力,局部注意力模块,整体采用残差结构简化训练;超像素内部注意力采用两个级联的线性层替代自注意力操作中的两次矩阵乘法操作,使得模型有更低的复杂度和更好的性能;并联多个超像素注意力模块可以更好的强化模型对局部和全局特征的关注。

[0020] 所述S4中结合L1损失、感知损失、全变分正则化和GAN损失训练SPGAN的方法为:

[0021] 首先,采用L1损失来计算超分图像与真实图像之间的平均误差,损失公式定义如下:

$$[0022] \quad MAE = \frac{1}{n} \sum_{i=1}^n |f(x) - y_i|$$

[0023] 其中n为样本总数, $f(x)$ 为超分后图像, y_i 为真实世界图像, L1损失函数通过计算 $f(x)$ 和 y_i 的差值绝对值来迭代优化模型;

[0024] 训练一个面向PSNR的模型, 训练得到的模型以SPNet命名;

[0025] 然后, 我们使用训练好的面向PSNR的模型作为生成器的初始化, 并结合L1损失、感知损失、全变分正则化和GAN损失训练SPGAN。

[0026] 本发明与现有技术相比, 具有的有益效果是:

[0027] 本发明提出了一种基于超像素注意力机制的真实世界图像超分重建方法SPGAN。提出了一种二阶段门控退化模型, 使得经过模型退化后的低分辨率图像更接近真实世界退化图像, 一定程度上减少了因模型退化后图像与真实世界退化后图像差距较大而导致的超分效果不理想情况, 进一步提升了真实世界图像超分效果, 同时, 结合GAN网络结构, 将超像素模块应用于生成器网络, 使得模型更能关注到图像的边缘像素, 一定程度上解决了GAN网络导致的图像模糊伪影等问题, 提升了超分图像的精细程度。本发明中的预训练模型SPNet网络在Set14数据集上进行量化比较, 当噪声等级为15时, 取得了23.74的PSNR值和0.6165的SSIM值, 当噪声等级为25时, 取得了23.03的PSNR值和0.578的SSIM值; 在BSD100数据集上进行量化比较, 当噪声等级为15时, 取得了23.83的PSNR值和0.6176的SSIM值, 当噪声等级为25时, 取得了23.31的PSNR值和0.5864的SSIM值。

附图说明

[0028] 为了更清楚地说明本发明的实施方式或现有技术中的技术方案, 下面将对实施方式或现有技术描述中所需要使用的附图作简单地介绍。显而易见地, 下面描述中的附图仅仅是示例性的, 对于本领域普通技术人员来讲, 在不付出创造性劳动的前提下, 还可以根据提供的附图引申获得其它的实施附图。

[0029] 本说明书所绘示的结构、比例、大小等, 均仅用以配合说明书所揭示的内容, 以供熟悉此技术的人士了解与阅读, 并非用以限定本发明可实施的限定条件, 故不具技术上的实质意义, 任何结构的修饰、比例关系的改变或大小的调整, 在不影响本发明所能产生的功效及所能达成的目的下, 均应仍落在本发明所揭示的技术内容能涵盖的范围内。

[0030] 图1为本发明的SPGAN整体架构图;

[0031] 图2为本发明的SPGAN模型中超像素块的结构图;

[0032] 图3为本发明的SPNet与优秀真实世界超分辨率方法对合成图像的实验结果对比图;

[0033] 图4为本发明的SPGAN与优秀真实世界超分辨率方法对真实世界图像的实验结果对比图。

具体实施方式

[0034] 为使本发明实施例的目的、技术方案和优点更加清楚, 下面将对本发明实施例中的技术方案进行清楚、完整地描述, 显然, 所描述的实施例仅仅是本申请一部分实施例, 而

不是全部的实施例,这些描述只是为进一步说明本发明的特征和优点,而不是对本发明权利要求的限制;基于本申请中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0035] 下面结合附图和实施例,对本发明的具体实施方式做进一步详细描述。以下实施例用于说明本发明,但不用来限制本发明的范围。

[0036] 本发明在pytorch深度学习框架下实现,本发明提供了基于超像素注意力机制的真实世界图像超分重建方法,具体包括以下步骤:

[0037] 1、数据准备

[0038] 本发明的数据样本为DF2K图像超分辨率数据集:

[0039] DF2K数据集由DIV2K数据集和Flickr2K数据集合并得到。DIV2K数据集包含多种不同退化类型的低分辨率图像1000张,其中包括800张训练图像,100张验证图像,100张测试图像;Flickr2K数据集包含2650张2K分辨率的图像,图像内容包括人物,动物,风景等多种类型,涵盖了丰富的自然场景和日常物体。

[0040] 训练过程中,图像被随机裁剪成 256×256 的尺寸。

[0041] 2、模型构建

[0042] 构建的SPGAN模型主体框架是一个GAN结构,具体网络结构如图1所示,SPGAN模型包含生成器和鉴别器两大模块。生成器首先使用一个 3×3 的卷积层进行低分辨率图像 I^{LR} 的浅层特征,然后用n个超像素模块构成编码器。如图2所示,每个超像素模块由超像素分割模块,超像素交叉注意力模块,超像素内部注意力模块,局部注意力模块四部分组成。在超像素块构建的细节上,我们首先通过线性层将像素信息传递到超像素层,生成自注意力机制所需的查询向量 Q^s 、键向量 K^l 和值向量 V^l 。这些向量通过与相应的权重矩阵 W_q^s 、 W_k^l 和 W_v^l 相乘得到。接着,我们计算查询向量 Q^s 与键向量 K^l 之间的相似度,并利用这一相似度作为注意力权重,与值向量 V^l 相乘,得到更新后的超像素特征。这一过程可以表示为:

$$[0043] \quad \text{Attention}(Q^s, K^l, V^l) = \text{softmax}\left(\frac{Q^s (K^l)^T}{\sqrt{D}}\right) V^l$$

[0044] 其中, $\sqrt{D}$ 是缩放因子,用于防止梯度消失。更新后的超像素特征随后通过交叉注意力机制映射回像素级别,这一步骤解决了超像素之间远程信息传递的问题,确保了信息在像素间的有效流通。经过n个超像素块后,提取出的特征再经过一个 3×3 的卷积层和一个上采样Pixel Shuffle层,最后将结果与低分辨率图像 I^{LR} 进行残差融合;鉴别器采用具有更强鉴别能力的U-Net结构设计,可以提供更为详细的鉴别反馈,同时为了增加训练稳定性,采用谱归一化正则化方法来稳定训练过程。

[0045] 3、模型训练

[0046] 训练过程采用两个阶段。

[0047] 首先,我们采用L1损失来计算超分图像与真实图像之间的平均误差,损失公式定义如下:

[0048]
$$MAE = \frac{1}{n} \sum_{i=1}^n |f(x) - y_i|$$

[0049] 其中n为样本总数, $f(x)$ 为超分后图像, y_i 为真实世界图像, L1损失函数通过计算 $f(x)$ 和 y_i 的差值绝对值来迭代优化模型。

[0050] 训练一个面向PSNR的模型, 训练得到的模型以SPNet命名。

[0051] 然后, 我们使用训练好的面向PSNR的模型作为生成器的初始化, 并结合L1损失、感知损失、全变分正则化和GAN损失训练SPGAN。

[0052] 4、测试结果

[0053] 训练由低分辨率图像进行超分辨率重建得到超分重建结果的方法为: 对于真实世界图像数据集使用双立方插值将数据集中测试图像分辨率调整为 256×256 作为HR图像; 使用设计好的二阶段门控退化模型将HR图像退化至 16×16 分辨率的图像作为LR图像。首先使用训练好的SPNet模型将LR图像重建得到SR图像, 使用SSIM和PNSR作为性能评估指标进行重建效果评价; 然后通过SPGAN模型将LR图像超分重建后得到SR图像, 使用NIQE作为性能评估指标进行重建效果评价。

[0054] 5、模型评估

[0055] 利用超分结果和真实图像计算峰值信噪比 (PSNR) 和结构相似性 (SSIM) 评价指标以评估SPNet模型的性能。

[0056] 表1不同模型在Set14和BSD100数据集上对比结果表

[0057]

	Set14				BSD100			
	$\sigma=15$		$\sigma=25$		$\sigma=15$		$\sigma=25$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
BSRNet	22.08	/	19.58	/	22.14	/	19.66	/
DASR	23.26	/	21.73	/	23.14	/	21.84	/
Real-SwinIR-L	23.61	/	22.12	/	23.75	/	22.48	/
Real-ESRNet	23.86	0.6188	<u>22.69</u>	0.5719	23.9	0.6194	<u>22.97</u>	0.5797
SPNet(Ours)	<u>23.74</u>	0.6165	23.03	0.578	<u>23.83</u>	0.6176	23.31	0.5864

[0058] 实验首先对比了SPNet和四种先进的真实世界图像超分辨率算法, 分别是BSRNet、DASR、Real-SwinIR-L和Real-ESRNet, 量化对比超分图像效果评价在Set14数据集和BSD100数据集中进行, 表中最优指标为加粗字体。

[0059] 通过表1发现, 在相同实验环境下SPNet相较于对比算法, 在Set14数据集和BSD100数据集中PSNR、SSIM结果表现优异, 在Set14数据集中, 两种噪声等级情况下, PSNR相较于BSRNet和DASR方法均有明显提升, 相较于Real-SwinIR-L也有小幅提升, 对比Real-ESRNet模型性能相差无几; SSIM与Real-ESRNet同样非常接近。在BSD100数据集中, 两种噪声等级情况下同样表现良好, PSNR仍然相较于BSRNet和DASR方法有较大提升, 相较于Real-SwinIR-L也有小幅提升; 而对比Real-ESRNet, PSNR和SSIM均非常接近。通过图3的对比结果可以发现, SPNet对于退化严重的图像相较于当前的真实世界超分辨率优秀模型的重建效果更好。

[0060] 利用超分结果和真实图像计算NIQE (Natural Image Quality Evaluator) 分数以

评估SPGAN模型的性能。

[0061] 表2不同模型在多种真实世界图像数据集上对比结果表

[0062]		Bicubic	ESRGAN	CDC	BSRGAN	SPGAN (ours)
	RealSR-Canon	6.1269	6.7715	6.1488	5.7489	5.6214
	OST300	4.4440	3.5245	4.7441	4.1662	3.2413
	ADE20K val	7.5239	3.6905	6.9219	3.9434	3.5814

[0063] 表2对比了SPGAN模型与四种先进的真实世界图像超分辨率方法,分别是Bicubic、ESRGAN、CDC、BSRGAN,量化对比超分图像效果评价在RealSR-Canon数据集、OST300数据集和ADE20Kval数据集中进行,表中最优指标为加粗字体。

[0064] 通过表2对比结果发现,SPGAN在三种数据集上的表现均优于大部分真实世界超分方法,在RealSR-Canon数据集和OST300数据集上,SPGAN模型的NIQE分数分别为5.6214和3.2413,重建效果显著优于其他真实世界超分辨率方法,而在ADE20Kval数据集上,SPGAN模型的NIQE分数也略优于其他优秀的超分辨率模型。这表明SPGAN能够更好应对真实世界中的多种退化情况,最大程度上将LR图像重建为高质量的SR图像。通过图4的对比可以发现,SPGAN不仅在提升图像分辨率方面有较大提升,同时也可以很好的保留细节信息,如边缘轮廓,纹理等,这表明SPGAN在进行真实世界图像超分辨率任务中具有优越的性能。

[0065] 上面仅对本发明的较佳实施例作了详细说明,但是本发明并不限于上述实施例,在本领域普通技术人员所具备的知识范围内,还可以在不脱离本发明宗旨的前提下作出各种变化,各种变化均应包含在本发明的保护范围之内。

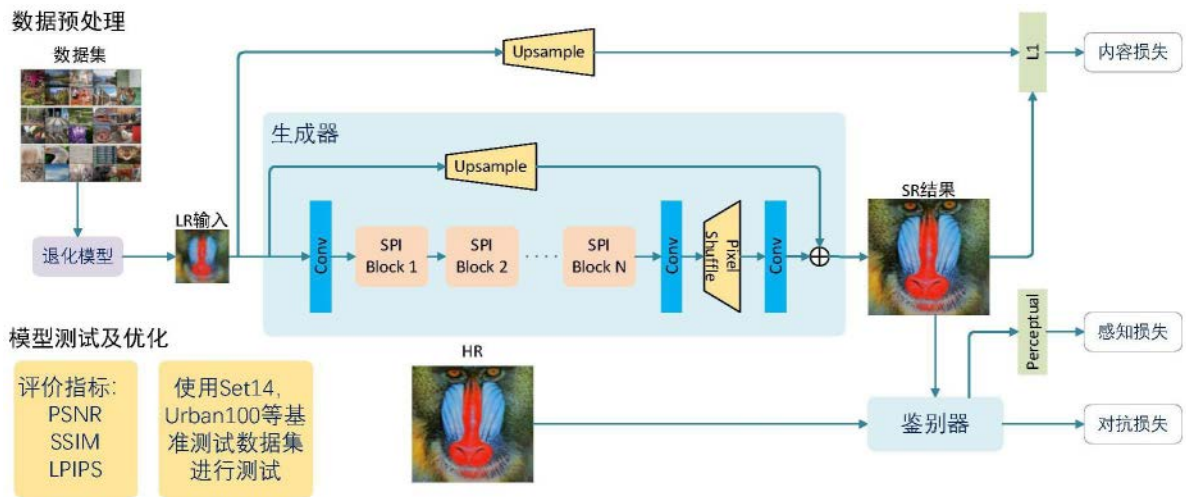


图1

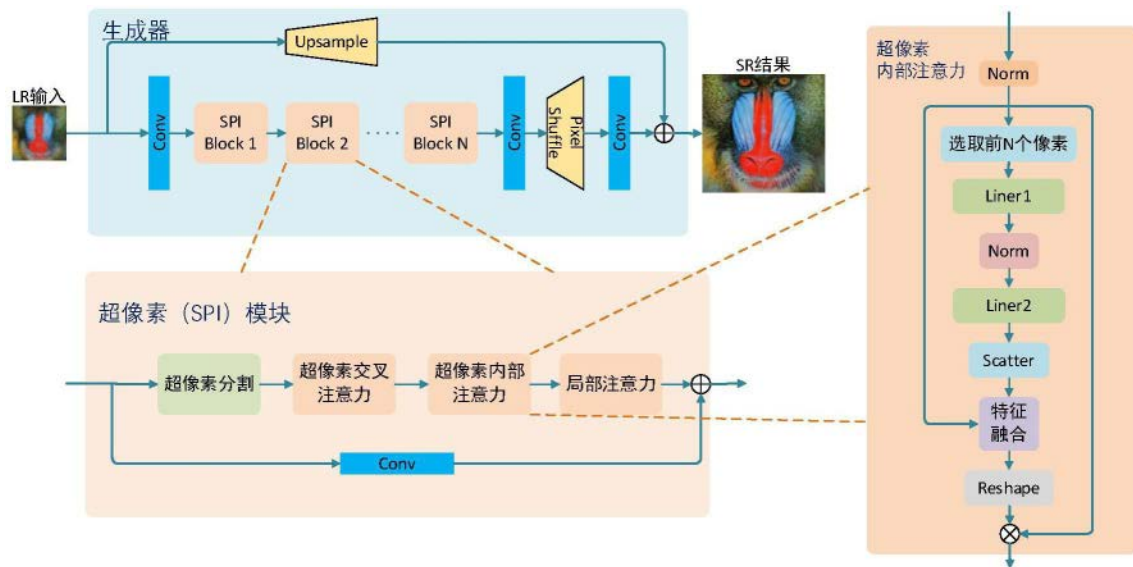


图2

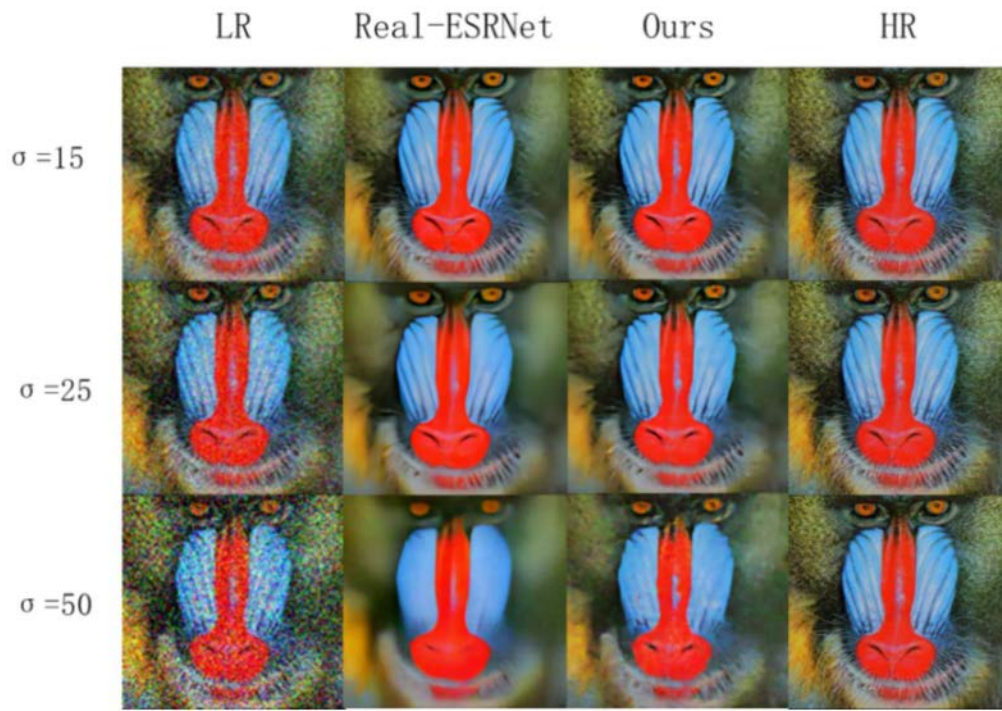


图3

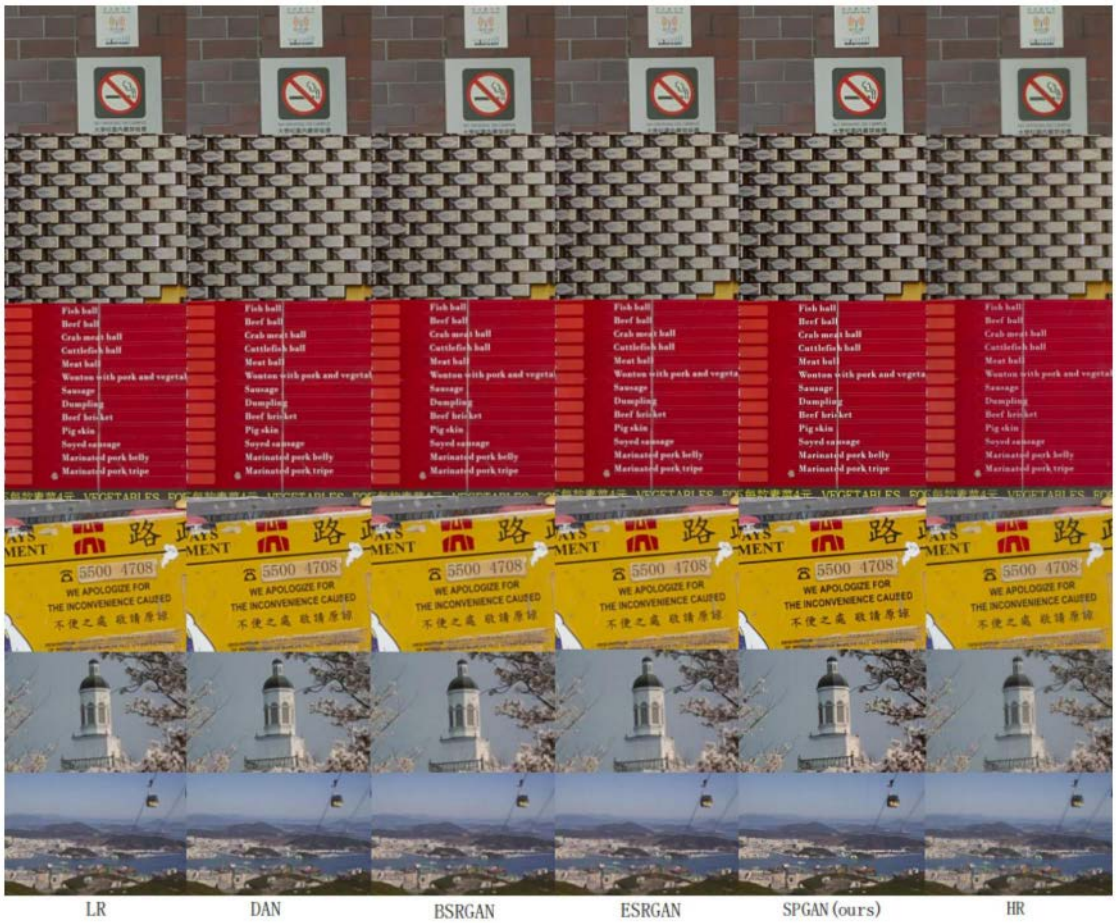


图4