

基于深度学习的单图像超分辨率重建综述^①



邢苏霄, 陈金玲, 李锡超, 陈 彤

(南京烽火天地通信科技有限公司, 南京 210019)
通信作者: 陈金玲, E-mail: chinling.chen@outlook.com

摘 要: 图像超分辨率重建是用于提高图像质量的一项重要技术, 得益于深度学习在计算机视觉领域的成功应用和快速发展, 单图像超分辨率重建的效果得到了显著提升. 因此, 本文针对基于深度学习的单图像超分辨率重建方法展开深入研究, 首先综合介绍了用于该领域的基准数据集、性能评价指标、损失函数等相关知识, 然后对有监督学习和无监督学习下单图像超分辨率重建技术的最新算法进行分类讨论, 并且比较分析了不同模型之间的异同点与优缺点, 最后对该领域面临的问题和未来的发展方向进行了总结与展望.

关键词: 超分辨率重建; 深度学习; 有监督学习; 无监督学习

引用格式: 邢苏霄, 陈金玲, 李锡超, 陈彤. 基于深度学习的单图像超分辨率重建综述. 计算机系统应用, 2022, 31(7): 23-34. <http://www.c-s-a.org.cn/1003-3254/8566.html>

Survey on Single Image Super-resolution Reconstruction Based on Deep Learning

XING Su-Xiao, CHEN Jin-Ling, LI Xi-Chao, CHEN Tong
(Nanjing Fiberhome Tiandi Communication Technology Co. Ltd., Nanjing 210019, China)

Abstract: Image super-resolution reconstruction is an important technique to improve image quality. Thanks to the successful application and rapid development of deep learning in the field of computer vision, significant improvement in single image super-resolution (SISR) reconstruction has been achieved. In response, this study explores SISR reconstruction methods based on deep learning in depth. Relevant background knowledge such as benchmark data sets, performance evaluation indexes, and the loss function used in this field are outlined. Then, the latest algorithms for SISR reconstruction techniques with supervised and unsupervised learning are discussed respectively, and the differences and similarities among different models as well as their advantages and disadvantages are compared. Finally, the existing problems in this field are summarized, and future trends are proposed.

Key words: super-resolution reconstruction; deep learning; supervised learning; unsupervised learning

图像超分辨率重建 (super-resolution reconstruction, SR) 是计算机视觉和图像处理领域中一个极具挑战性的热门研究课题. 大多数成像设备会受到硬件、环境等多种因素的干扰, 使得图像的分辨率并不能满足实际应用的需要. 为了提高图像空间分辨率, 研究学者们开始尝试在不改变成像设备的前提下, 利用图像处理、机器学习等算法将低分辨率 (low resolution, LR) 图像重建为高分辨率 (high resolution, HR) 图像^[1]. 正是

因为深度学习能自适应地学习 LR 图像与 HR 图像之间的非线性映射关系这种特性, 使得基于深度学习的图像 SR 算法明显优于传统方法, 因此也成为单图像超分辨率重建 (single image super-resolution, SISR) 的主流研究方法. 同时, SR 技术在医学图像的分析与识别^[2]、人脸超分辨率^[3]、视频监控与安全^[4]、遥感图像^[5]等实际场景中有着广泛的应用.

近年来, 大量介绍 SISR 方法的文献^[6,7] 以及与超

^① 基金项目: 国家自然科学基金 (62172093)

收稿时间: 2021-10-12; 修改时间: 2021-11-08; 采用时间: 2021-11-12; csa 在线出版时间: 2022-05-31

分辨率相关的综述类文章^[8,9]层出不穷,这些文献中都 对近年来 SR 领域提出的各种新型算法、设计的新 型网络框架、新的应用场景等进行了研究与总结. 与 之不同的是,本文按照学习方法的差异性将基于深度 学习的 SISR 问题分为两大类:有监督的 SISR 方法和 无监督的 SISR 方法.先全面介绍了与 SISR 问题相关 的储备知识,然后详细论述了以上两类方法的最新理 论和研究进展,同时比较了不同算法之间的异同点和 优缺点,最后对该领域存在的问题和未来的发展方向 进行了总结和展望.

1 SISR 方法的理论基础

本节将从问题定义、基准数据集、评价指标和损 失函数等方面对 SISR 方法的理论基础展开详细的调 研,并对其中涉及到的相关基础知识和理论实现进行 论述.

1.1 问题定义

日常生活中,由于成像设备、成像环境等多种退 化因素的影响,人们获取到的图像通常是 LR 图像,其 退化过程如图 1 所示.设 I_x 为由退化模型输出的 LR 图 像, I_y 为原始的 HR 图像.

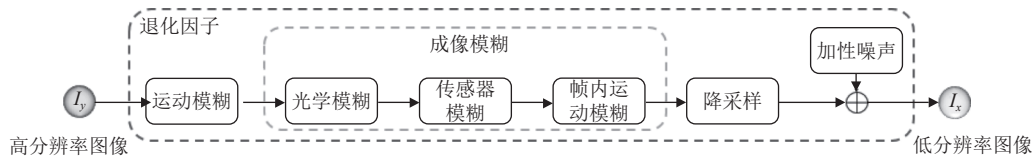


图 1 图像退化过程

那么, LR 与 HR 图像之间的关系则可以表示为:

$$I_x = D(I_y, \delta) \quad (1)$$

其中, $D(\cdot)$ 表示退化函数, δ 表示退化因子.

图像退化过程中的退化函数、模糊因子、噪声等 因素通常都是未知的,而能够被直接获取的只有 LR 图 像 I_x . 因此,在不改变硬件设备的条件下,要获取更高空 间分辨率的图像只能通过 LR 图像进行重建.那么其重 建过程可以表述为:

$$\hat{I}_y = F(I_x, \theta) \quad (2)$$

其中, $F(\cdot)$ 表示超分辨率重建模型, \hat{I}_y 为重建后的 HR 图 像, θ 为模型中的参数.由图 1 可知,图像退化是由运动 模糊、成像模糊、加性噪声等多种因素共同影响的, 则图像的退化过程^[10]可以由符号化后的式(3)进行 表述:

$$I_x = D(I_y, \delta) = (I_y \otimes k) \downarrow s + n_\sigma, \{k, s, \sigma\} \subseteq \delta \quad (3)$$

其中, $I_y \otimes k$ 表示原始高分辨率图像与模糊核的卷积过 程, $\downarrow s$ 表示下采样因子, n_σ 表示加性噪声.而基于深度 学习的 SISR 的实现过程就是模型不断迭代学习的过 程,在找到低分辨率图像与高分辨率图像之间最佳非 线性映射关系之后,损失函数达到最小值.重建过程的 损失函数可以由式(4)进行表示:

$$\hat{\phi} = [\min L(\hat{I}_y, I_y)]_\phi + \lambda \psi(\phi) \quad (4)$$

其中, $L(\hat{I}_y, I_y)$ 表示由超分辨率模型重建后的 HR 图像 与真实的 HR 图像之间的损失函数, λ 为折衷参数, $\psi(\phi)$ 表示正则化项.

1.2 常用数据集

数据作为深度学习算法的建模必不可少的基本要 素之一,其数据量的多少以及数据质量的好坏将直接 影响着目标模型的效果.现如今, SISR 领域有着各式 各样的数据集用于模型的训练与测试,这些数据集在 图像的数量、图像的质量、图像的分辨率、图像的多样 性以及图像的成像环境等方面都有所不同.其中,大 多数数据集仅包含 HR 图像,不包含用于模型训练和 测试所需的不同放大倍数下的 LR-HR 图像对,需要通 过双三次插值算法实现图像对的构造.

本节将对常用的 SISR 数据集进行调研,汇总结果 如表 1 所示.本节列出了各类数据集在图像数量、图 像格式、图像类型、图像平均分辨率以及图像中所包 含的类别信息.其中, Set5^[11]、Set14^[12]、Urban100^[13]、 BSDS100^[14]、Manga109^[15] 常作为模型重建效果评估的 基准数据集.同时,研究者们也常常倾向于多个数据集 共同参与模型的训练与测试,例如 T91^[16] 与 BSDS300^[14] 的结合, DIV2K^[17] 与 Flickr2K^[18] 的结合.其次,一些其 他视觉领域的数据集也参与到 SR 模型的训练过程 中,如 ImageNet^[19]、MS-COCO^[20]、VOC2012^[21]、 CelebA-HQ^[22]、L20^[23] 等.

表1 SISR 领域基准数据集

数据集	图像数量	图像格式	图像类型	用法	图像类别信息
BSDS100 ^[14]	100	PNG	非成对	验证	包含动物、人、建筑物等
BSDS200 ^[24]	200	PNG	非成对	训练/验证	包含动物、人、建筑物等
BSDS300 ^[14]	300	JPG	非成对	训练/验证	包含动物、人、建筑物等
BSDS500 ^[25]	500	JPG	非成对	训练/验证	包含动物、人、建筑物等
General-100 ^[26]	100	PNG	非成对	训练	包含动物、食物、植物等
CelebA-HQ ^[22]	30 000	PNG	非成对	训练	高清人脸图像
DIV2K ^[17]	1 000	PNG	成对	训练/验证	包含人、动物、风景等
Manga109 ^[15]	109	PNG	非成对	验证	漫画图像
Urban100 ^[13]	100	PNG	非成对	验证	包含城市建筑物等
Set14 ^[12]	14	PNG	非成对	验证	包含动物、植物等
Set5 ^[11]	5	PNG	非成对	验证	婴儿、鸟、蝴蝶、女性、头部
T91 ^[16]	91	PNG	非成对	训练	包含水果、花等
L20 ^[23]	20	PNG	非成对	验证	包含植物、建筑物、动物等
PIRM ^[27]	200	PNG	非成对	验证	包含自然风景、人等
OutdoorScene ^[28]	10 624	PNG	非成对	训练/验证	包含动物、植物、自然景观等
Sun-Hays 80 ^[29]	80	PNG	成对	训练/验证	包含自然风光等
Flickr2K ^[18]	2 650	PNG	非成对	训练	包含人物、动物、风景等

1.3 评价指标

根据测量方法的不同,可以将图像质量评价(image quality assessment, IQA)方法^[30]分为:通过人类视觉对图像质量进行感知的主观评价方法和通过数值计算进行定量分析的客观评价方法.其中,主观评价方法相较于客观方法更加符合实际应用,但其需要消耗更多的人力和时间资源,所以基于客观评价方法进行定量评估是目前超分辨率重建领域的主流评价指标,本节将对以上两类评价方法进行简要论述.

(1) 峰值信噪比(peak signal-to-noise ratio, *PSNR*)^[31]是目前应用最为广泛的用于图像重建质量评价的指标之一.在SR任务中,*PSNR*是由最大像素值与两幅图像之间的均方误差来定义的.*PSNR*值越大表明图像的质量越好,其计算公式如式(5)所示.其中, MAX_I 表示图像中最大像素值, MSE 表示两个图像之间对应像素差值平方的均值,单通道和多通道图像的*MSE*计算原理如式(6)、式(7)所示.

$$PSNR = 10 \log \left(\frac{MAX_I^2}{MSE} \right) \quad (5)$$

单通道的*MSE*:

$$MSE = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (I_{ij} - \hat{I}_{ij})^2 \quad (6)$$

多通道的*MSE*:

$$MSE = \frac{1}{C \times M \times N} \sum_{k=1}^C \sum_{i=1}^M \sum_{j=1}^N (I_{ijk} - \hat{I}_{ijk})^2 \quad (7)$$

(2) 结构相似性评价(structural similarity index measure, *SSIM*)^[32]是用于比较图像间亮度、对比度、结构细节的相似度指标.与*PSNR*相比,*SSIM*与人类视觉系统对图像结构提取方面的适应性更加契合,*SSIM*的取值范围为[0, 1].图片的结构相似度越大,表示其失真程度越小,也说明图像的质量越好,*SSIM*的原理实现如式(11)所示.

$$L(X, Y) = \frac{2\mu_X\mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1} \quad (8)$$

$$C(X, Y) = \frac{2\sigma_X\sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2} \quad (9)$$

$$S(X, Y) = \frac{\sigma_{XY} + C_3}{\sigma_X\sigma_Y + C_3} \quad (10)$$

$$SSIM(X, Y) = L(X, Y)^\alpha \times C(X, Y)^\beta \times S(X, Y)^\gamma \quad (11)$$

其中, μ_X 和 μ_Y 分别表示图像*X*、*Y*的像素均值, σ_X 和 σ_Y 分别表示*X*、*Y*的像素标准差, σ_{XY} 表示图像*X*、*Y*的协方差. C_1 、 C_2 、 C_3 是常数, $C_1 = (K_1 \times L)^2$, $C_2 = (K_2 \times L)^2$, $C_3 = C_2/2$,一般有 $K_1 = 0.01$, $K_2 = 0.03$, $L = 255$.权重参数 α 、 β 、 γ 能够分别控制亮度、对比度和结构细节对结构相似度的重要性.

(3) 平均意见得分(mean opinion score, *MOS*)^[33]是属于IQA中一种常用的图像质量评价主观方法,也是在图像感知质量评价方面最可靠的方法.该方法要求参与评测的人员对被测图像从图像的清晰度、对比

度、颜色、外观细节等方面进行视觉感知并打分,最后的评估得分是每一个评级份数的算术平均分,分数范围由坏至好分别为1到5分。但是这种方法会因为评测标准的不同或人为偏见等不可控因素的影响而存在一些缺陷和局限性。

(4) 基于学习的 IQA 方法是为了减少人为因素的干扰,更好地评估图像的视觉感知质量,在大型理想参考图像数据集上通过学习的方式对图像进行质量评估的一种算法。2017年, Kim 等人^[34]提出了基于 CNN 的全参考图像质量评估模型 DeepQA,该方法结合失真图像、主观评分和客观误差图进行模型的训练并对图像的视觉相似性进行预测。Ma 等人^[35]和 Talebi 等人^[36]分别提出了 MEON 和 NIMA 的无参考 IQA 神经网络,该方法不需要真实的参考图像,直接从视觉感知分数中进行学习并预测图像的质量分数。基于 CNN 的方法能够在视觉感知评估方面表现出更好的性能,但需要更多的数据资源。

除了上述几种常见的方法外,还有一些通过特定任务(图像分类、人脸识别、图像分割等)来间接衡量 SISR 性能的方法。同时,基于相位一致性和图像的梯度大小来计算图像的特征相似性^[37]的方法也被用来评估图像的质量。Blau 等人^[38]在2018年的论文中提出,图像的感知质量和图像的失真相互矛盾的,随着图像失真度的降低,图像的视觉感知质量也会随之降低。所以, PSNR 和 SSIM 这两种定量分析的方法仍是目前图像 SR 领域应用最为广泛的 IQA 方法。

1.4 损失函数

损失函数是深度学习模型迭代优化过程中必不可少的重要元素之一。在 SISR 任务中,损失函数的选择至关重要,其通过计算误差来衡量 HR 图像的重建质量。而长期的研究发现,多种损失函数的共同指导更能在多方面指导 HR 图像的重建过程。SR 领域中最为常见的损失是像素损失,但由于像素并不能对图像的重建质量进行绝对评估,所以内容损失、对抗损失、先验损失等各类损失函数应运而生,以下将介绍几种常用的损失函数。

(1) 像素损失 (pixel loss): 该损失是最为简单常用的一类损失函数,其主要以像素为基础衡量两幅图像之间的差异,包括 L1、L2 损失^[39]以及 L1 损失的一种可区分的变体 Charbonnier 损失^[40],其具体表达式分别如式(12)、式(13)、式(14)所示。这类损失能够提高

模型的收敛速度,但是缺乏对图像内容以及语义信息的考虑,常会出现重建图像过度平滑的现象。

$$\mathcal{L}_{L_1}(I_{SR}, I_y) = \frac{1}{hwc} \sum_{i,j,k} |I_{SR}^{i,j,k} - I_y^{i,j,k}| \quad (12)$$

$$\mathcal{L}_{L_2}(I_{SR}, I_y) = \frac{1}{hwc} \sum_{i,j,k} (I_{SR}^{i,j,k} - I_y^{i,j,k})^2 \quad (13)$$

$$\mathcal{L}_{Char}(I_{SR}, I_y) = \frac{1}{hwc} \sum_{i,j,k} \sqrt{(I_{SR}^{i,j,k} - I_y^{i,j,k})^2 + \varepsilon^2} \quad (14)$$

(2) 内容损失 (content loss)^[41]: 该损失也常被称为感知损失,如式(15)所示。其主要是从图像内容理解和感知层面对图像质量进行评价,通常使用预先训练好的图像分类网络对两幅图像之间的语义差异进行评估,从而提高重建图像的视觉相似性和真实度。

$$\mathcal{L}_{Cont}(I_{SR}, I_y, \phi) = \frac{1}{hwc} \sum_{i,j,k} (\phi_{(l)}^{i,j,k}(I_{SR}) - \phi_{(l)}^{i,j,k}(I_y)) \quad (15)$$

(3) 纹理损失 (texture loss)^[42]: 该损失主要源自于风格迁移领域,其常被称为风格常见损失,如式(17)所示。通常使用重建图像与原始图像不同通道之间的相关性来度量两幅图像在颜色、纹理、对比度等方面的相似性。

$$G_{i,j}^{(l)}(I) = \text{vec}(\phi_{(l)}^i(I)) \cdot \text{vec}(\phi_{(l)}^j(I)) \quad (16)$$

$$\mathcal{L}_{texture}(I_{SR}, I_y, \phi) = \frac{1}{c_l^2} \sqrt{\sum_{i,j,k} (G_{(l)}^{i,j,k}(I_{SR}) - G_{(l)}^{i,j,k}(I_y))^2} \quad (17)$$

(4) 对抗损失 (adversarial loss)^[43]: 该损失的提出主要是由于对抗生成网络 (generative adversarial networks, GANs), GANs 由生成器和判别器组成,如式(18)所示。SISR 任务中的生成器为超分模块,而判别器主要用高层的语义信息来鉴别生成图像的真实性,从而使得重建后的 HR 图像更加真实。

$$\mathcal{L}_{Adversarial}(I_x, G, D) = \sum_{n=1}^N -\log D(G(I_x)) \quad (18)$$

(5) 先验损失 (prior loss)^[44]: 该损失主要是将稀疏先验、梯度先验和边缘先验等知识引入到 SISR 模型中,如式(19)、式(20)所示。该损失将已知的先验作为约束条件放入损失函数中,从而能够优化图像的一些特定信息,使得重建的效果能够朝着预期的方向发展,重建后的 HR 图像也将包含更多的纹理细节。

$$\mathcal{L}_{TV}(I_{SR}) = \frac{1}{hwc} \sum_{i,j,k} \sqrt{(I_{SR}^{i,j+1,k} - I_y^{i,j,k})^2 + (I_{SR}^{i,j,k+1} - I_y^{i,j,k})^2} \quad (19)$$

$$\mathcal{L}_{\text{Edge}}(I_{SR}, I_y, E) = \frac{1}{hwc} \sum_{i,j,k} |E(I_{SR}^{i,j,k}) - E(I_y^{i,j,k})| \quad (20)$$

2 有监督的 SISR 深度学习方法

目前,大多数基于深度学习方法的 SISR 模型均侧重于使用 LR 图像和 HR 图像进行模型训练,也即这类模型都属于有监督的深度学习模型.虽然不同模型之间的差异性较大,但是他们均是由模型框架、网络设计、学习策略等组件组合而成,以适用于各种 SR 任务.因此,本节将从网络的类型和网络的框架两个层面对有监督的 SISR 方法进行分类,并对其中的代表性算法进行介绍.

2.1 基于网络类型

根据网络设计的差异性,可以根据网络设计不同将有监督的深度学习超分辨率重建网络分为以 PSNR 为导向的 SISR 方法和以视觉感知为驱动的 SISR 方法.

2.1.1 以 PSNR 为导向的 SISR 方法

随着卷积神经网络 (convolutional neural network, CNN) 在图像分类领域的成功应用, SISR 方法也在 CNN 模型的启发下取得了显著的进步,而前期基于深度学习的 SISR 模型大多数均是以获取较高的 PSNR 为导向,也即通过像素损失来指导网络的学习过程. Dong 等人^[45]首次将传统的稀疏编码方法与深度学习相结合,提出了基于深度卷积神经网络的端到端图像超分辨率重建算法 (super-resolution convolutional neural network, SRCNN). 随后, Kim 等人^[46]加深网络层数以提取更深层次的图像特征,并且为了减少网络参数、提升特征图利用率,进一步提出了基于深层卷积的 VDSR 和 DRCN. 由于深层网络容易出现梯度爆炸或梯度消失的现象, ResNet 的思想也参与到 SR 任务中, Zhang 等人^[47]基于残差密集块构建了 RDN, 该网络实现了全局残差的学习和局部特征的融合,并在 PSNR 上取得了不错的性能. 同时,为了突破固定放大倍数的限制,实现对深入图像的任意放大, Shi 等人^[48]提出了一种基于像素重排思想的 ESPCN 方法用于超分辨率重构. 为进一步实现对 LR 图像的非整数缩放, Hu 等人^[49]提出了可通过单一模型解决任意缩放因子问题的全新模型 Meta-SR, 并在基准数据集上取得了优越性能. 为了解决 LR 与 HR 图像映射空间的不确定性, DBPN^[50]、DRN^[51]、SRFBN^[52] 等方法利用反馈机制在网络迭代学习的过程中对图像的重建特征进行约束,从而减少

可能的函数空间并在 PSNR 上显示出更好的重建性能. 为了进一步提升 PSNR, attention 机制被引入到 SR 领域, 该机制能够有效提升网络的表征能力以及加强信息的流动性, 典型的算法有 RCAN^[53]、SAN^[54]、HAN^[55] 等. 因此, CNN 与各类深度学习策略的有效结合, 使得以 PSNR 为导向的 SISR 方法取得了较为完善的重建性能.

2.1.2 以视觉感知为驱动的 SISR 方法

SR 问题也可以看作是一种图像转换的任务, 也即输入一张退化的 LR 图像, 经过模型处理后得到一张高质量的 HR 图像. 考虑到以 PSNR 为导向的重建方法常常会导致高频信息的缺失, 重建后的 HR 图像过度平滑, 与真实的 HR 图像在视觉感知上存在一定的差异. Johnson 等人^[41]在 2016 年则提出利用感知损失来衡量重建后的图像质量, 以此来提升模型的重建性能. Ledig 等人^[30]在 2016 年提出了一个以视觉感知为驱动的 SRGAN 方法, 该方法首次将生成对抗网络 (generative adversarial network, GAN)^[44]用于 SISR 领域, 并在 MOS 上表现出了卓越的性能. 随后, 大量以视觉感知为驱动的基于 GAN 的 SR 方法被提出, 例如 Wang 等人^[56]通过在 ESRGAN 中引入密集残差块 (residual-in-residual dense block, RRDB) 来提高网络的性能; Soh 等人^[57]基于自然流形鉴别器设计了 NatSR, 使得重建后的图像具有更加真实的纹理和更加自然的细节信息; Wang 等人^[29]则基于特征的空间信息感知提出了 SFTGAN, 该方法让图像的纹理信息恢复更精细、更接近于真实的 HR 图像; Ma 等人^[58]发现基于上述方法重建后的 HR 图像常常存在明显的结构变形问题, 于是提出了结构保留的 SPSR 网络, 该方法能在缓解图像结构变形的同时还能够生成丰富的问题细节; 为了更好的构建 LR 图像和高频信息之间的关系, Lugmayr 等人^[59]基于归一化流设计了能够基于低分辨率输入学习输出的条件分布的 SRFlow, 该方法能够探索超分辨率的解空间并生成多样性的图像, 且最终的生成图像在感知度量上均优于当前最优的 GAN 方法. 尽管现有的以感知为驱动的 SISR 方法的确能够有效提升图像的感知质量, 但是仍会出现伪影、几何变形等现象.

2.2 基于模型框架

根据模型框架的差异性, 可以根据框架设计不同将有监督的深度学习超分辨率重建网络分为预定义上采样网络、单次上采样网络、渐进上采样网络和迭

代上下采样网络。

2.2.1 预定义上采样网络

预定义上采样网络 (predefined up-sampling network) 的实现通常先利用传统的插值算法将输入的 LR 图像放大到具有目标尺寸的粗略 HR 图像, 然后使用深度卷积神经网络 (deep convolutional neural network, DCNN) 进行特征提取, 使得网络能自动学习 LR 图像到 HR 图像的端到端非线性映射关系, 从而实现 SR 图像的重建. Dong 等人^[26] 率先使用预定义上采样的 SR 框架提出了 SRCNN, 首次利用 CNN 实现了高质量的重建效果. 如图 2 所示, 该方法解决了 LR 图像直接映射到 HR 图像的难度, 降低了学习的难度, 同时也能实现任意缩放因子下的比传统单尺度 SR 模型更加精细的重建效果. 文献 [53,54] 在这种主流框架的基础上对学习策略、模型设计等方面进行改进, 使得基于深度学习的 SISR 模型的性能逐步得到提升. 但是, 这种预先处理至目标尺寸的上采样操作会给输入图像带来额外的噪声和模糊, 并且使得大多数计算都是在高维空间实现的, 使得网络的时间复杂度和空间复杂度要明显高于其他框架.

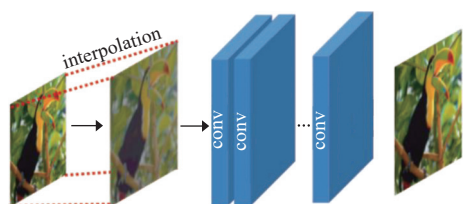


图 2 预定义上采样网络

2.2.2 单次上采样网络

单次上采样网络 (single up-sampling network) 直接在低维空间实现大部分的计算过程, 也即在不对输入的 LR 图像进行目标尺寸缩放的前提下, 利用 DCNN 进行特征提取, 并在网络的后端采用可端到端学习的

上采样层 (亚像素卷积层或者反卷积) 对非线性映射后的图像进行放大, 其网络设计如图 3 所示. 该方法能有效提高图像的空间分辨率, 且相较于预定义上采样网络能够大大降低网络的时间和空间成本, 因此, 基于这种网络的设计思想, 通过对特征提取层、上采样层、学习策略等模块的改进, 文献 [56,60,61] 的方法也取得了更好的性能. 但是, 其无法学习到 LR 与 HR 图像之间复杂的非线性映射关系, 而且网络中只进行了一步上采样操作, 这样会增加网络的学习难度, 同时针对每一个缩放因子都需要训练一个独立的 SR 模型, 所以无法实现任意缩放尺寸的 SR 能力.

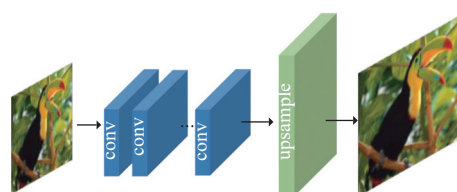


图 3 单次上采样网络

2.2.3 渐进上采样网络

渐进上采样网络 (progressive up-sampling network) 利用拉普拉斯金字塔网络在前向过程中重建出不同尺度的 SR 图像, 如图 4 所示. 该网络通过 CNN 的级联, 输入的 LR 图像在每一个上采样阶段都会被重建到一个更高分辨率的 HR 图像, 这种上采样叠加的操作使得网络能够更好地对浅层特征进行利用. LapSRN^[40]、MS-LapSRN^[62] 和渐进式 SR (ProSR)^[63] 也都采用了这种网络结构, 并取得了较单次上采样网络更优的性能, 并且 ProSR 在网络的学习过程中加强了上下文信息的流行性. 渐进式上采样网络可以解决多种尺度的 SR 任务, 面对大尺度放大需求, 该框架可以有效降低时间复杂度, 但是将学习过程拆分为多个阶段会造成网络性能的不稳定.

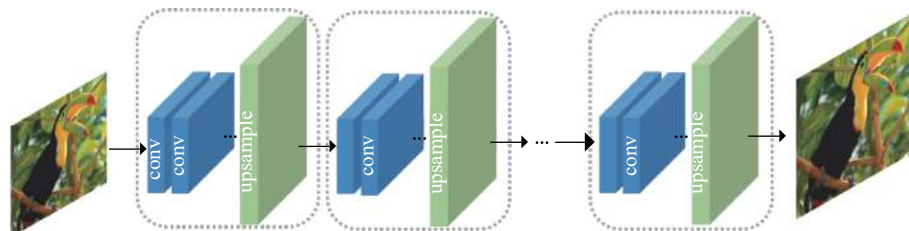


图 4 渐进上采样网络

2.2.4 迭代上下采样网络

迭代上下采样网络 (iterative up and down-sampling

network) 中加入了一种迭代式的误差修正反馈机制, 尝试利用反向传播的机制构建相互连接的上采样模块

和下采样模块,从而获取到不同层次的分辨率特征,并利用误差对重建后的 HR 图像进行加强,网络结构如图 5 所示. Haris 等人^[50]基于该网络提出了 DBPN,通过上采样层和下采样层交替连接的方式将中间所有的重建结果重建最终的 HR 图像. 随后, SRFBN^[52]采用

了迭代的上下采样反馈块,跳过连接更密集,学习到了更好的表示方法. 该框架下的模型可以更好地挖掘 LR-HR 图像对之间的深层关系,产生更深层次的特征,从而提供更高质量的重建结果,但目前反向传播模块的设计并没有明确的标准.

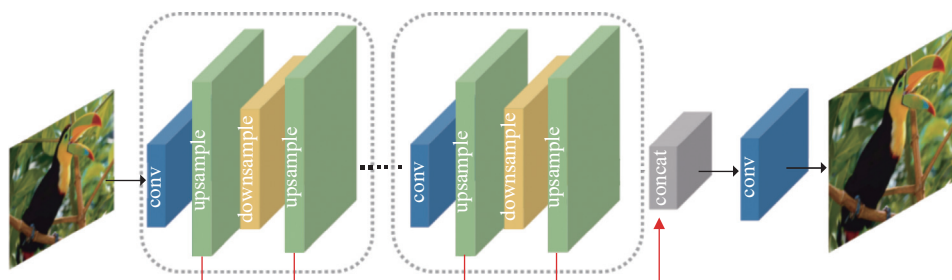


图 5 迭代上下采样网络

3 无监督的 SISR 深度学习方法

目前, SISR 网络的实现大多数都基于有监督的学习机制,也即使用成对的 LR-HR 图像进行模型训练. 但想要收集到同一个场景并且具有不同分辨率的图像需要大量的人力和时间成本,而且现有的大多数 SR 数据集都需要通过传统的插值算法生成与 HR 图像向配对的 LR 图像,而这种已知的退化过程对真实场景下的 SR 需求并不适用. 因此,基于无监督学习的 SR 方法受到研究学者的广泛关注,该方法在训练过程中不需要成对的 LR-HR 图像,模型也能学习到真实场景下的图像退化函数,最终获取到的模型更容易适应真实自然场景下的 SR 问题. 本小节将对现有的几种无监督学习下的 SISR 深度学习方法进行介绍.

3.1 弱监督学习下的 SISR 方法

基于弱监督学习的 SISR 方法在不利用插值算法构造成对的 LR-HR 数据的前提下,使用非配对的 LR-HR 图像进行网络的训练,在该学习方法下有两种网络训练方式:

一种是先学习 HR 图像退化为 LR 图像的退化核,再用该函数生成 HR 图像对应的 LR 图像,以构建成对的 LR-HR 数据集用于 SR 模型的训练. CVPR2021 年接收的一篇最新的研究无监督学习下的退化感知 SISR 算法 DASR^[64],该网络采用无监督学习的方式利用退化表达学习机制自适应地从 LR 图像中捕获到具有辨识性的特征,从而获取到准确的退化信息,最后结合对比学习的思想对低分辨率特征进行调制.

另一种是利用循环学习的思想让网络学习 LR 图像到 HR 图像的映射和 HR 图像到 LR 图像的退化函数,从而实现相互验证和相互促进. 2018 年, Yuan 等人^[65]受到 CycleGAN^[66]的启发,于是设计了一个由 4 个生成器和 2 个判别器组合而成的无监督 SISR 模型 CinCGAN. 该网络将 HR 图像和 LR 图像分别看作两个不同的作用域,并利用 CycleGAN 的循环结构来解决真实场景下 LR 图像的重建问题.

3.2 零次学习下的 SISR 方法

Shocher 等人^[67]最早将零次学习 (zero-shot learning) 应用到 SR 领域并提出完全无监督的 SISR 深度学习方法 ZSSR. 其利用网络在对输入的单一特定图像进行测试时训练 SR 网络的思想,学习图像内部的先验特征,进而预测出图像的退化函数,预估出 LR 图像与 HR 图像之间的退化关系并恢复出图像的高频细节. ZSSR 的实现不需要大型的 SR 数据集,只需要通过学习到的退化函数生成小型数据集,并在该 SR 数据集上完成网络的训练. 所以,该方法对真实场景下的图像进行重建时具有更大的优势,但 ZSSR 需要对每张特定的图像进行上千次的迭代学习,因此其推理时间要远大于其他基于深度学习的 SISR 方法. 2020 年, Soh 等人^[68]为解决推理时间过长的问题,将零次学习与元转换学习相结合应用到 SR 问题中,从而提出了只需一次梯度更新便可使重建效果可观的 MZSR 方法. 虽然,基于零次学习方法的 SR 网络在未知退化函数的条件下取得了出色的性能,并且能够有效改善自然场景下

LR 图像的重建效果,但仍然会存在噪声和伪影。

4 讨论与分析

为了更加直观地比较 SISR 的性能,本节将在 4 倍放大因子的条件下,对目前一些典型的有监督和无监督的 SISR 算法在 Set5、Set14、BSDS100 三个基准数据集上分别计算 PSNR 和 SSIM,从定量分析的角度进行对比分析。

由表 2 可知, SRCNN 作为第一个将 CNN 引入 SISR 领域的算法,其得益于 CNN 强大的特征捕获和表达能力,使得 SRCNN 对单张 LR 图像的重建能力远高于传统的双三次插值方法 Bicubic;其次,密集连

接、残差连接、特征金字塔等网络结构也在 SISR 领域取得了显著的进步,如 VDSR^[53]、RRDB^[56]、LapSRN^[40] 等算法;随后,为了解决 SR 的不适定性,提高 SISR 网络对真实自然场景图像重建的自适应性和泛化能力,研究人员逐渐将反馈回归机制、基于感知损失、先验损失、生成式建模思想、无监督学习等方法用于 SISR 领域。其中,基于归一化流的 SRFlow 重建方法取得了优异的重建性能,无监督的重建算法 ZSSR 也能有效学习到 LR 图像的超分辨率信息,其重建性能超越了经典的 SRCNN,也为基于无监督学习方法的 SISR 研究打下了坚实的基础。

表 2 ×4 放大因子下不同算法在 Set5、Set14、BSDS100 数据集下的 PSNR/SSIM 指标

类型	方法	Set5	Set14	BSDS100
有监督	Bicubic	28.42/0.8104	26.00/0.7027	25.96/0.6675
	SRCNN	30.48/0.8628	27.49/0.7503	26.90/0.7101
	VDSR	31.35/0.8838	28.01/0.7674	27.29/0.7251
	SRGAN	29.40/0.8472	26.02/0.7397	25.16/0.6688
	RRDB	32.73/0.9011	28.99/0.7917	27.85/0.7455
	DRCN	31.53/0.8838	28.02/0.7670	25.14/0.7510
	ESPCN	30.66/0.8646	27.71/0.7562	24.60/0.7360
	LapSRN	31.54/0.8855	28.19/0.7720	25.21/0.7560
	Meta-RDN	—/—	32.35/0.9019	27.75/0.7423
	DBPN	32.47/0.8980	28.82/0.7860	26.38/0.7946
	ProSR	—/—	28.94/—	27.79/—
	DRN-L	32.74/0.9020	28.98/0.7920	27.03/0.8130
	SRFBN	32.47/0.8983	28.81/0.7868	26.60/0.8015
	RCAN	32.63/0.9002	28.87/0.7889	26.82/0.8087
	SAN	32.64/0.9003	28.92/0.7888	26.79/0.8068
	HAN	32.64/0.9002	28.90/0.7890	26.85/0.8094
	NatSR	30.98/0.8606	27.42/0.7329	26.44/0.6827
	SPSR	30.400/0.8627	26.640/0.7930	25.505/0.6576
	SRFlow	36.29/0.9468	33.02/0.9065	31.74/0.8864
	SwinIR ^[69]	32.92/0.9044	29.09/0.7950	27.45/0.8254
无监督	ZSSR	31.13/0.8796	27.12/0.7211	27.12/0.7211
	MZSR(10)	30.50/0.8704	—/—	26.89/0.7168
	IKC ^[70]	—/—	28.26/0.7688	27.29/0.8014
	DAN ^[71]	—/—	28.43/0.7693	27.51/0.8078

尽管各类 SISR 算法已经取得了显著的进步,但是对结构复杂、纹理信息丰富的图像重建后还是会存在一定的重建伪影、信息丢失等问题,所以 SISR 任务仍然面临着许多挑战。

5 总结与展望

本文对深度学习方法在 SISR 领域的最新研究进展进行了详细调研与论述。SISR 技术作为计算机视觉领域的重要研究任务之一,其可以作为刑事侦查、医

学诊断、遥感图像等领域的有效辅助工具。基于深度学习的 SISR 算法层出不穷,本文主要围绕基于有监督、无监督的 SISR 技术进行讨论,也对其中涉及到的背景知识进行了具体的介绍。虽然 SISR 的性能已经取得了显著的提升,但是仍然还存在很多挑战,以下将对 SISR 研究过程中存在的问题和未来发展趋势进行总结。

(1) 评价指标

SISR 问题中常用 IQA 来衡量 SR 算法的性能, PSNR 和 SSIM 作为两个最常用的评价指标。前者容易

导致过度平滑的问题,使得差异较小的图片之间的定量分析误差较大,而后者虽然从亮度、对比度和结构细节方面模拟人类的视觉感知,但是不能完全准确表示图像的视觉感知效果,MOS虽然与视觉感知质量最为相近,但是其评估过程需要消耗大量的时间和人工成本。所以,制定通用的、准确的SR质量评价指标是目前亟待解决的问题。

(2) 真实场景图片的SR

自然场景下的图片常常会受到加性噪声、运动模糊、压缩伪影等因素的影响,使得模型面对未知的退化问题,缺少成对的真实场景下的LR-HR图像。因此,在有人工设计的数据集上训练的SISR模型对自然场景图像的重建效果并不理想。虽然,已经有很多研究方法被提出,但是仍然存在训练难度大、重建细节不够完善等缺点。因此,针对未知图像退化过程的建模依旧是SISR的重要发展方向。

(3) 网络框架的设计趋势

一个好的网络框架不仅能提升SISR的性能,还能有效减少模型的时间复杂度和空间复杂度。因此,后续可以考虑在模型的学习过程中,将全局信息和局部信息、底层特征与高层次特征相结合的思想,利用不同尺度、不同层次的语义信息加强网络对图像特征的表征能力,从而提升SISR算法的重建性能。

(4) 上采样方法的改进

目前使用的上采样方法会存在不能端到端学习、感受野分布不均匀、会产生棋盘效应等问题,因此会导致SISR算法低效、重建效果不稳定等问题。同时目前的上采样多为整数倍。所以,探索出一个高效地、适用于任意放大因子的上采样方法是未来值得进一步研究的方向。

参考文献

- 唐艳秋,潘泓,朱亚平,等.图像超分辨率重建研究综述.电子学报,2020,48(7):1407-1420.[doi:10.3969/j.issn.0372-2112.2020.07.022]
- 刘朋伟,高媛,秦品乐,等.基于多感受野的生成对抗网络医学MRI影像超分辨率重建.计算机应用.<http://kns.cnki.net/kcms/detail/51.1307.TP.20210708.1626.008.html>. [2021-10-07].
- 蒋文杰,罗晓曙,戴沁璇.基于对抗网络人脸超分辨率重建算法研究.计算机工程与应用,2021,57(11):219-223.[doi:10.3778/j.issn.1002-8331.2002-0370]
- Kappeler A, Yoo S, Dai QQ, et al. Video super-resolution with convolutional neural networks. IEEE Transactions on Computational Imaging, 2016, 2(2): 109-122. [doi: 10.1109/TCI.2016.2532323]
- Yuan QQ, Li Y, Li JC, et al. Remote sensing image super-resolution via regional spatially adaptive total variation model. Proceedings of 2014 IEEE Geoscience and Remote Sensing Symposium. Quebec City: IEEE, 2014. 3073-3076.
- Wang ZH, Chen J, Hoi SCH. Deep learning for image super-resolution: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(10): 3365-3387. [doi: 10.1109/TPAMI.2020.2982166]
- Bashir SMA, Wang Y, Khan M, et al. A comprehensive review of deep learning-based single image super-resolution. PeerJ Computer Science, 2021, 7: e621. [doi: 10.7717/peerj-cs.621]
- 向海燕.超分辨率图像恢复方法综述.重庆理工大学学报(自然科学),2014,28(9):72-76.
- 陈文静,唐轶.基于深度学习的图像超分辨率重建方法综述.云南民族大学学报(自然科学版),2019,28(6):597-605.
- Zhang K, Zuo WM, Zhang L. Learning a single convolutional super-resolution network for multiple degradations. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 3262-3271.
- Bevilacqua M, Roumy A, Guillemot C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. Proceedings of British Machine Vision Conference. Surrey: BMVC, 2012. 1-10.
- Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations. Proceedings of the 7th International Conference on Curves and Surfaces. Avignon: Springer, 2010. 711-730.
- Huang JB, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars. Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 5197-5206.
- Martin D, Fowlkes C, Tal D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. Proceedings of the 8th IEEE International Conference on Computer Vision. Vancouver: IEEE, 2001. 416-423.
- Ogawa T, Otsubo A, Narita R, et al. Object detection for comics using Manga109 annotations. arXiv: 1803.08670, 2018.
- Yang JC, Wright J, Huang TS, et al. Image super-resolution

- via sparse representation. *IEEE Transactions on Image Processing*, 2010, 19(11): 2861–2873. [doi: [10.1109/TIP.2010.2050625](https://doi.org/10.1109/TIP.2010.2050625)]
- 17 Agustsson E, Timofte R. NTIRE 2017 challenge on single image super-resolution: Dataset and study. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Honolulu: IEEE, 2017. 1122–1131.
- 18 Lim B, Son S, Kim H, *et al.* Enhanced deep residual networks for single image super-resolution. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Honolulu: IEEE, 2017. 1132–1140.
- 19 Deng J, Dong W, Socher R, *et al.* ImageNet: A large-scale hierarchical image database. *Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition*. Miami: IEEE, 2009. 248–255.
- 20 Lin TY, Maire M, Belongie S, *et al.* Microsoft COCO: Common objects in context. *Proceedings of the 13th European Conference on Computer Vision*. Zurich: Springer, 2014. 740–755.
- 21 Everingham M, Eslami SMA, van Gool L, *et al.* The PASCAL visual object classes challenge: A retrospective. *International Journal of Computer Vision*, 2015, 111(1): 98–136. [doi: [10.1007/s11263-014-0733-5](https://doi.org/10.1007/s11263-014-0733-5)]
- 22 Tolosana R, Vera-Rodriguez R, Fierrez J, *et al.* Deepfakes and beyond: A survey of face manipulation and fake detection. *Information Fusion*, 2020, 64: 131–148. [doi: [10.1016/j.inffus.2020.06.014](https://doi.org/10.1016/j.inffus.2020.06.014)]
- 23 Timofte R, Rothe R, van Gool L. Seven ways to improve example-based single image super resolution. *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 1865–1873.
- 24 Chao D, Loy CC, Tang XO. Accelerating the super-resolution convolutional neural network. *Proceedings of the 14th European Conference on Computer Vision*. Amsterdam: Springer, 2016. 391–407.
- 25 Arbeláez P, Maire M, Fowlkes C, *et al.* Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(5): 898–916. [doi: [10.1109/TPAMI.2010.161](https://doi.org/10.1109/TPAMI.2010.161)]
- 26 Dong C, Loy CC, He KM, *et al.* Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295–307. [doi: [10.1109/TPAMI.2015.2439281](https://doi.org/10.1109/TPAMI.2015.2439281)]
- 27 Blau Y, Mechrez R, Timofte R, *et al.* The 2018 PIRM challenge on perceptual image super-resolution. *Proceedings of European Conference on Computer Vision*. Munich: Springer, 2018. 334–355.
- 28 Wang XT, Yu K, Dong C, *et al.* Recovering realistic texture in image super-resolution by deep spatial feature transform. *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 606–615.
- 29 Sun LB, Hays J. Super-resolution from Internet-scale scene matching. *Proceedings of 2012 IEEE International Conference on Computational Photography*. Seattle: IEEE, 2012. 1–12.
- 30 Ledig C, Theis L, Huszár F, *et al.* Photo-realistic single image super-resolution using a generative adversarial network. *IEEE Computer Society. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 105–114.
- 31 Almohammad A, Ghinea G. Stego image quality and the reliability of PSNR. *Proceedings of the 2010 2nd International Conference on Image Processing Theory, Tools and Applications*. Paris: IEEE, 2010. 215–220.
- 32 Marabini R, Sorzano COS, Matej S, *et al.* 3-D reconstruction of 2-D crystals in real space. *IEEE Transactions on Image Processing*, 2004, 13(4): 549–561. [doi: [10.1109/TIP.2003.822620](https://doi.org/10.1109/TIP.2003.822620)]
- 33 Viswanathan M, Viswanathan M. Measuring speech quality for text-to-speech systems: Development and assessment of a modified mean opinion score (MOS) scale. *Computer Speech & Language*, 2005, 19(1): 55–83.
- 34 Kim J, Lee S. Deep learning of human visual sensitivity in image quality assessment framework. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 1969–1977.
- 35 Ma C, Yang CY, Yang XK, *et al.* Learning a no-reference quality metric for single-image super-resolution. *Computer Vision and Image Understanding*, 2017, 158: 1–16. [doi: [10.1016/j.cviu.2016.12.009](https://doi.org/10.1016/j.cviu.2016.12.009)]
- 36 Talebi H, Milanfar P. NIMA: Neural image assessment. *IEEE Transactions on Image Processing*, 2018, 27(8): 3998–4011. [doi: [10.1109/TIP.2018.2831899](https://doi.org/10.1109/TIP.2018.2831899)]
- 37 Zhang L, Zhang L, Mou XQ, *et al.* FSIM: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*, 2011, 20(8): 2378–2386. [doi: [10.1109/TIP.2011.2109730](https://doi.org/10.1109/TIP.2011.2109730)]
- 38 Blau Y, Michaeli T. The perception-distortion tradeoff. *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 6228–6237.

- 39 Ahn N, Kang B, Sohn KA. Fast, accurate, and lightweight super-resolution with cascading residual network. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 256–272.
- 40 Lai WS, Huang JB, Ahuja N, *et al.* Deep Laplacian pyramid networks for fast and accurate super-resolution. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 5835–5843.
- 41 Johnson J, Alahi A, Li FF. Perceptual losses for real-time style transfer and super-resolution. Proceedings of the 14th European Conference on Computer Vision. Amsterdam: Springer, 2016. 694–711.
- 42 Gatys LA, Ecker AS, Bethge M. Image style transfer using convolutional neural networks. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 2414–2423.
- 43 Goodfellow IJ, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets. Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal: NIPS, 2014. 2672–2680.
- 44 Bulat A, Tzimiropoulos G. Super-FAN: Integrated facial landmark localization and super-resolution of real-world low resolution faces in arbitrary poses with GANs. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 109–117.
- 45 Dong C, Loy CC, He KM, *et al.* Learning a deep convolutional network for image super-resolution. ECCV. Proceedings of the 13th European Conference on Computer Vision. Zurich: Springer, 2014. 184–199.
- 46 Kim J, Lee JK, Lee KM. Deeply-recursive convolutional network for image super-resolution. Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016. 1637–1645.
- 47 Zhang YL, Tian YP, Kong Y, *et al.* Residual dense network for image super-resolution. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 2472–2481.
- 48 Shi W, Caballero J, Huszár F, *et al.* Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 1874–1883.
- 49 Hu XC, Mu HY, Zhang XY, *et al.* Meta-SR: A magnification-arbitrary network for super-resolution. Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 1575–1584.
- 50 Haris M, Shakhnarovich G, Ukita N. Deep back-projection networks for super-resolution. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 1664–1673.
- 51 Guo Y, Chen J, Wang JD, *et al.* Closed-loop matters: Dual regression networks for single image super-resolution. Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 5406–5415.
- 52 Li Z, Yang JL, Liu Z, *et al.* Feedback network for image super-resolution. Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 3862–3871.
- 53 Zhang YL, Li KP, Li K, *et al.* Image super-resolution using very deep residual channel attention networks. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 294–310.
- 54 Dai T, Cai JR, Zhang YB, *et al.* Second-order attention network for single image super-resolution. Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 11057–11066.
- 55 Niu B, Wen WL, Ren WQ, *et al.* Single image super-resolution via a holistic attention network. Proceedings of the 16th European Conference on Computer Vision. Glasgow: Springer, 2020. 191–207.
- 56 Wang XT, Yu K, Wu SX, *et al.* ESRGAN: Enhanced super-resolution generative adversarial networks. Proceedings of European Conference on Computer Vision. Munich: Springer, 2019. 63–79.
- 57 Soh JW, Park GY, Jo J, *et al.* Natural and realistic single image super-resolution with explicit natural manifold discrimination. Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 8114–8123.
- 58 Ma C, Rao YM, Cheng YA, *et al.* Structure-preserving super resolution with gradient guidance. Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 7766–7775.
- 59 Lugmayr A, Danelljan M, van Gool L, *et al.* SRFlow: Learning the super-resolution space with normalizing flow. Proceedings of the 16th European Conference on Computer Vision. Glasgow: Springer, 2020. 715–732.
- 60 Tong T, Li G, Liu XJ, *et al.* Image super-resolution using dense skip connections. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice: IEEE,

2017. 4809–4817.
- 61 Han W, Chang SY, Liu D, *et al.* Image super-resolution via dual-state recurrent networks. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 1654–1663.
- 62 Lai WS, Huang JB, Ahuja N, *et al.* Fast and accurate image super-resolution with deep Laplacian pyramid networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(11): 2599–2613. [doi: [10.1109/TPAMI.2018.2865304](https://doi.org/10.1109/TPAMI.2018.2865304)]
- 63 Wang YF, Perazzi F, McWilliams B, *et al.* A fully progressive approach to single-image super-resolution. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE, 2018. 977–986.
- 64 Wei YX, Gu SH, Li YW, *et al.* Unsupervised real-world image super resolution via domain-distance aware training. Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 13380–13389.
- 65 Yuan Y, Liu SY, Zhang JW, *et al.* Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE, 2018. 814–823.
- 66 Zhu JY, Park T, Isola P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017. 2242–2251.
- 67 Shocher A, Cohen N, Irani M. Zero-shot super-resolution using deep internal learning. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 3118–3126.
- 68 Soh JW, Cho S, Cho NI. Meta-transfer learning for zero-shot super-resolution. Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 3513–3522.
- 69 Liang JY, Cao JZ, Sun GL, *et al.* SwinIR: Image restoration using swin transformer. Proceedings of 2021 IEEE/CVF International Conference on Computer Vision Workshops. Montreal: IEEE, 2021. 1833–1844.
- 70 Gu JJ, Lu HN, Zuo WM, *et al.* Blind super-resolution with iterative kernel correction. Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 1604–1613.
- 71 Luo ZX, Huang Y, Li S, *et al.* Unfolding the alternating optimization for blind super resolution. Proceedings of the 34th Conference on Neural Information Processing Systems. Vancouver: NeurIPS, 2020. 1–12.

(校对责编: 孙君艳)