

## 图像超分辨重建算法综述

史振威, 雷 森

(北京航空航天大学宇航学院图像处理中心, 北京, 100191)

**摘 要:** 图像超分辨重建是一种提升图像分辨率的图像处理技术, 而超分辨问题是一个难解的欠定问题, 近些年来研究人员主要采用基于学习的方法, 从大量数据中学习图像先验信息, 以实现对解空间的约束。本文介绍了近20年来主流的图像超分辨重建算法, 主要分为基于传统特征的方法和基于深度学习的方法。对于传统的超分辨重建算法, 阐述了基于邻域嵌入的方法、基于稀疏表示的方法以及基于局部线性回归的方法。对于基于深度学习的超分辨重建算法, 详细总结了网络模型结构设计、上采样方式、损失函数形式以及复杂条件下的算法设计4个方面。此外, 本文简要分析了超分辨重建技术在视频超分辨、遥感图像超分辨以及在视觉高层任务方面的应用。最后, 本文展望了图像超分辨重建技术的未来发展方向。

**关键词:** 图像超分辨; 邻域嵌入; 稀疏表示; 局部线性回归; 深度学习

**中图分类号:** TP391.41      **文献标志码:** A

### Review of Image Super-Resolution Reconstruction

SHI Zhenwei, LEI Sen

(Image Processing Center, School of Astronautics, Beihang University, Beijing, 100191, China)

**Abstract:** Image super-resolution reconstruction is an image processing technology, which recovers high-resolution images from low-resolution images. While, the super-resolution problem is under-determined. In recent years, researchers have proposed learning-based methods to learn image prior information from a large amount of data, in order to constrain the super-resolution solution space. This paper introduces the mainstream image super-resolution reconstruction algorithms in the past two decades, which are divided into two categories: traditional features based methods and deep learning based methods. For the traditional super-resolution reconstruction algorithms, this paper mainly presents the methods based on neighborhood embedding, the methods based on sparse representation, and the methods based on local linear regression. For the deep learning based methods, the super-resolution model design, the up-sampling method and the loss function form are provided. In addition, this paper introduces the application of super-resolution reconstruction technology in video super-resolution, remote-sensing image super-resolution, and high-level vision tasks. Finally, the future development directions of image super-resolution reconstruction technology are provided.

**Key words:** image super-resolution reconstruction; neighborhood embedding; sparse representation; local

---

**基金项目:** 国家重点研发计划(2017YFC1405605)资助项目; 国家自然科学基金(61671037)资助项目; 北京市自然科学基金(4192034)资助项目。

**收稿日期:** 2019-11-30; **修订日期:** 2019-12-25

linear regression; deep learning

## 引言

图像超分辨率重建是一种利用低分辨率图像恢复得到高分辨率图像的图像处理技术<sup>[1-3]</sup>。该技术能在不改变物理成像设备的限制下,获取更高分辨率的图像,在很多领域得到了应用,如医学成像<sup>[4-6]</sup>、安全监控<sup>[7-8]</sup>、遥感图像处理<sup>[9-11]</sup>等。除了能够提升图像质量,图像超分辨率重建技术也可以改善许多计算机视觉任务<sup>[12-15]</sup>。

图像超分辨率最早由 Harris<sup>[16]</sup>在 20 世纪 60 年代提出,旨在通过一幅低分辨率图像重构出高分辨率图像。Tsai<sup>[17]</sup>于 1984 年利用多幅低分辨图像实现对高分辨率图像的恢复。随着机器学习技术的研究和发展,Freeman 等<sup>[18-19]</sup>于 2000 年首次将机器学习的方法应用至图像超分辨率领域。随后大量基于机器学习的超分辨方法涌现出来,如基于邻域嵌入的方法<sup>[20]</sup>,基于稀疏表示的方法<sup>[21-25]</sup>和基于局部线性回归的方法<sup>[26-28]</sup>等。但这些方法大多是利用图像的底层特征进行超分辨率重建,特征的表达能力有限,很大程度上限制了重建效果。

深度学习方法能够自适应地从训练集中学习得到深层特征,在最近几年被广泛应用于图像超分辨率领域。Dong 等<sup>[29-30]</sup>于 2014 年利用卷积神经网络(Convolutional neural networks, CNNs)直接学习低分辨率图像到高分辨率图像之间的非线性映射关系,相较于传统方法,重建效果得到了较大提升。此后,研究人员提出了大量基于深度学习的超分辨网络模型,如利用残差学习和残差模块构造深层的超分辨率网络模型<sup>[31-33]</sup>,采用递归结构<sup>[34-36]</sup>和密集连接<sup>[37-38]</sup>,以及将注意力机制与超分辨网络相结合<sup>[39-40]</sup>等。这些方法大都通过优化逐像素损失实现网络模型参数的学习,易使重构图像产生过模糊的现象,从而降低图像的感知质量<sup>[41-42]</sup>。因此,研究人员将生成对抗网络(Generative adversarial networks, GANs)<sup>[43]</sup>引入图像超分辨率领域,用以改善重建结果的感知质量<sup>[44-45]</sup>。此外,由于上述方法通常利用固定的下采样方式得到低分辨率图像,如双立方卷积下采样或高斯模糊下采样,很难适应于复杂的真实场景,针对这一问题,一些工作则使用更加复杂的退化模型<sup>[46]</sup>或使用网络去模拟退化过程<sup>[47-50]</sup>。同时部分研究人员通过模拟数据的获取过程,提出了适应于真实应用场景的超分辨数据集<sup>[51-53]</sup>,用以提升网络模型在真实场景的适用性。

本文首先介绍了关于图像超分辨率重建问题背景,然后详细总结了基于传统特征的超分辨算法和基于深度学习的超分辨算法,接着阐述了图像超分辨率重建数据集与评价指标。此外,简要分析了图像超分辨率技术在视频超分辨、遥感图像超分辨以及在视觉高层任务方面的应用。最后,本文总结了图像超分辨率算法的当前研究现状以及未来发展方向。

## 1 图像超分辨率重建问题背景

图像超分辨率重建是指利用一幅或多幅低分辨率图像重建得到相同场景的高分辨率图像。根据低分辨率图像数目的多少可以将超分辨问题分为单幅图像超分辨率(Single image super-resolution, SISR)和多幅图像超分辨率(Multi-image super-resolution, MISR)。相对于多幅图像超分辨,单幅图像超分辨的数据集获取更加容易,并且单幅图像超分辨率技术通常是多幅图像超分辨率技术的研究基础,所以在近 20 年来,单幅图像超分辨是图像超分辨率重建领域的研究热点,于是本文主要关注于单幅图像超分辨率问题。

本文令  $y$  表示观测得到的低分辨率图像,  $x$  表示其对应的高分辨率图像,则退化过程可表示为

$$y = \Phi(x; \delta) \quad (1)$$

式中: $\Phi$ 表示退化函数; $\delta$ 表示退化参数。在真实场景中,通常只能获得低分辨率图像 $y$ ,而无法得到关于退化函数及其参数的信息,研究人员需要利用 $y$ 重构出高分辨率图像,即

$$\hat{x} = F(y; \theta) \quad (2)$$

式中: $F$ 为超分辨模型; $\theta$ 为模型参数。

因为实际的退化过程未知且复杂,并容易受到各种因素的影响,如散焦、运动模糊、传感器噪声、压缩噪声等<sup>[1-2]</sup>,对退化过程进行准确的定义非常困难,所以绝大多数研究人员通常将式(1)的退化过程描述成如下的形式

$$y = (x \otimes k) \downarrow_s + n \quad (3)$$

式中: $k$ 表示模糊核; $x \otimes k$ 是指高分辨率图像与模糊核之间的卷积操作; $\downarrow_s$ 是指放缩尺度为 $s$ 的下采样操作; $n$ 为加性的高斯白噪声。

由于图像超分辨问题是一个难解的欠定问题,近年来研究人员主要采用基于学习的方法,从大量数据中学习图像先验信息去对解空间进行约束。具体即是利用式(3)的退化模型和高分辨率图像仿真得到对应的低分辨率图像数据,然后利用高低分辨率图像数据集,对超分辨模型 $F$ 的模型参数 $\theta$ 进行学习,即

$$\hat{\theta} = \arg \min_{\theta} \mathcal{L}(x, \hat{x}) + \lambda \Psi(\theta) \quad (4)$$

式中: $\mathcal{L}(x, \hat{x})$ 为数据保真项,用以度量生成的高分辨图像 $\hat{x}$ 和真实的参考图像 $x$ 之间的差异; $\Psi(\theta)$ 为正则化项(或称之为先验项); $\lambda$ 为保真项和正则化项之间的权衡。

## 2 基于传统特征的超分辨重建算法

20世纪之前,研究者主要采用基于插值的方法来解决单幅图像超分辨问题,例如双线性插值<sup>[54]</sup>以及基于双线性的改进插值方法<sup>[55-56]</sup>。这些方法主要是基于图像的连续性假设,由于没有引入额外信息,其重建结果的边缘与轮廓通常比较模糊,纹理也无法得到较好的恢复。进入21世纪,研究人员提出了一系列基于学习的超分辨重建算法,如基于邻域嵌入的方法<sup>[20]</sup>,基于稀疏表示的方法<sup>[21-25]</sup>和基于局部线性回归的方法<sup>[26-28]</sup>等。这些方法大都假设低分辨率图像块和高分辨率图像块分布在不同的子空间中,并且具有相似的局部几何流行结构,利用提前构建的高低分辨率图像块字典以更好地实现图像恢复。

最早基于学习的图像超分辨方法是由Freeman等于2000年提出<sup>[18-19]</sup>。该方法需要先收集高分辨率图像,并采用模糊和下采样操作仿真得到其对应的低分辨率图像,然后截取适宜大小的高低分辨率图像块构建得到训练集。在预测阶段,通过对测试图像进行滑框操作,在训练集中找到与测试图像块近似的低分辨率图像块,从而得到其对应的高分辨率图像块。整个重建过程基于两个前提假设:(1)相似的低分辨率图像块对应着相似的高分辨率图像块;(2)相邻图像块的边缘应保持一致,从而保证重构图像的连续性。为此,Freeman使用了Markov网络得到这两个条件下的最佳选择。Freeman所提出的方法对后续的基于传统特征的图像超分辨算法产生了很大的影响,其大都延续了“构建低频/高频图像对+滑窗重构”的基本流程。

### 2.1 基于邻域嵌入的方法

在Freeman等所提出的方法中,重建得到的高分辨率图像块只来自于训练集中的一个最近邻的样本,模型的表达能力十分有限,从而很大程度上限制了重构图像的质量。针对这一问题,受流形学习思想的启发<sup>[57-58]</sup>,Chang等于2004年提出了基于邻域嵌入的超分辨算法<sup>[20]</sup>。该工作主要基于“高分辨率图

像流形和低分辨率图像流形具有相似的局部几何结构”这个前提假设。具体地,对于测试图像中的每一个图像块,在训练集中找到与其对应的  $K$  个近邻的低频图像块,然后通过求解式(5)的约束最小二乘问题,得到对该图像块的重构权重。

$$\begin{aligned} \min & \left\| \mathbf{y}_i^q - \sum_{x_s^p \in N_q} \omega_{qp} \mathbf{y}_s^p \right\|^2 \\ \text{s.t.} & \sum_{x_s^p \in N_q} \omega_{qp} = 1 \end{aligned} \quad (5)$$

式中:  $\mathbf{y}_i^q$  为测试图像块;  $\mathbf{y}_s^p$  为寻找出的近邻样本集合  $N_q$  中的低频图像块;  $\omega_{qp}$  为对应的权重。

利用优化好的重构权重,并基于局部相似结构这一个前提假设,就可以计算得到高分辨率图像块  $\hat{\mathbf{x}}_i^q = \sum_{x_s^p \in N_q} \omega_{qp} \mathbf{x}_s^p$ 。文献[20]基于的假设在后续大量的图像超分辨研究中都得到采用。

## 2.2 基于稀疏表示的方法

相比于基于邻域嵌入的超分辨重建算法,基于稀疏表示的方法能够为每一个测试图像块自适应地确定相关的字典元素个数,从而无需事先确定近邻个数  $K$ ,重建结果的边缘和纹理更加清晰。基于稀疏表示的超分辨方法最早由 Yang 等受到压缩感知理论的启发,于 2008 年提出<sup>[21]</sup>。这一理论指出对于稀疏信号,可以通过远低于奈奎斯特采样定理所要求的频率对信号进行采样,并能实现对原始信号的精确重构<sup>[59-60]</sup>。在文献[21]中, Yang 等首先从训练集中进行随机采样,构建高低分辨率图像块字典  $D_h$  和  $D_l$ ,然后使用字典  $D_l$  对测试图像块进行稀疏重构,并利用得到的重构系数和字典  $D_h$  获取测试图像块所对应的重建结果。该方法的核心思想在于引入稀疏先验,对病态的超分辨问题进行约束,此时超分辨率重建问题被视为  $L_1$  范数的优化问题为

$$\min \|\alpha\|_1 \quad \text{s.t.} \quad \|ED_l\alpha - E\mathbf{y}\|_2^2 \quad (6)$$

式中:  $\mathbf{y}$  为测试图像块;  $\alpha$  为重构系数;  $E$  为特征提取操作。

最后,该方法利用后向投影保证能够从重建结果中生成尽可能相似的低分辨率图像

$$\mathbf{X}^* = \arg \min \|\mathbf{X} - \mathbf{X}_0\| \quad \text{s.t.} \quad D\mathbf{H}\mathbf{X} = \mathbf{Y} \quad (7)$$

式中:  $\mathbf{X}_0$  为使用稀疏重建得到的结果;  $\mathbf{Y}$  为测试图像;  $H$  和  $D$  分别为模糊操作和下采样操作。

值得注意的是,在文献[21]中,高、低分辨率图像字典是从训练集图片中随机采样得到,字典元素的总数是十万个,重构图像时需要耗费大量时间。针对该问题, Yang 等提出字典联合学习方法<sup>[22]</sup>,旨在从数据中联合学习得到紧凑的高低分辨率图像字典(设置的字典大小为 1 024),从而极大地提升了算法的运行效率。在文献[23]中, Yang 等进一步提出了耦合字典学习,让高低分辨率字典的学习过程相互影响高度耦合,最终使得低分辨率图像块及其对应的高分辨率图像块都能分别被低分辨率字典和高分辨率字典很好地重构。此外, Yang 等提出了以下两个措施对推断过程进行加速:(1)选择性地对测试图中的图像块进行恢复,即对于平滑区域直接用双线性插值进行拉伸,对于边缘、轮廓、角点等富有变化的区域才进行稀疏学习,从而减少需要利用字典进行重构的图像块数目;(2)利用文献[25]方法,构建前馈神经网络模型学习低分辨率输入图像块特征和对应的稀疏编码系数之间的映射关系,在测试阶段中无需再使用字典迭代求解  $L_1$  问题,而是直接利用该神经网络对稀疏系数进行回归,从而极大地减少了运算量,速度提升了 10 倍左右<sup>[23]</sup>。

## 2.3 基于局部线性回归的方法

对于绝大多数基于邻域嵌入和稀疏表示的超分辨方法在进行重建的过程中通常要求解带  $L_1$  约

束的最小二乘问题,这将占用很大的计算开销。针对这一问题, Timofte等<sup>[26-27]</sup>于2013年提出了锚近邻回归(Anchored neighborhood regression, ANR)的方法,将之前的 $L_1$ 约束问题转化为带 $L_2$ 约束的最小二乘优化问题,从而较大幅度地提升了重构的速度。ANR的优化表达式为

$$\min_{\beta} \left\| y_F - D_l \beta \right\|_2^2 + \lambda \left\| \beta \right\|_2 \quad (8)$$

式中: $y_F$ 为输入图像的特征; $D_l$ 为低分辨率图像字典; $\beta$ 为重构系数。式(8)的闭式解为

$$\beta = (D_l^T D_l + \lambda I)^{-1} D_l^T y_F \quad (9)$$

利用 $\beta$ 和高分辨率图像字典 $D_h$ ,可以获取得到高分辨图像块为 $x = D_h \beta$ ,并进一步结合式(14)可以得到

$$x = D_h (D_l^T D_l + \lambda I)^{-1} D_l^T y_F \quad (10)$$

式中投影矩阵 $P$ 为 $D_h (D_l^T D_l + \lambda I)^{-1} D_l^T$ ,可以线下计算得到。

在测试阶段,只需要利用提前计算得到的 $P$ 和 $y_F$ 就可以直接得到重建图像块 $x$ ,不需要再迭代求解。因此在保持精度不损失的情况下,能够使重构的速度大大地提高(相比于文献[22]速度提升近百倍)。

## 2.4 小结

本节主要描述了基于传统特征的超分辨率重建算法,包括基于邻域嵌入的方法、基于稀疏表示的方法和基于局部线性回归的方法。这些方法都假设高分辨率图像流形和低分辨率图像流形具有相似的局部几何结构,通过寻找测试图像块在训练集中近邻的低分辨率图像块,并利用其对应的高分辨率图像块以完成对测试图像的重构。其中基于邻域嵌入和基于稀疏表示的方法需要在测试阶段求解一系列带 $L_1$ 约束的最小二乘问题,计算开销巨大,而基于局部线性回归的方法则利用投影矩阵直接求解,从而极大地提升了重构速度。然而,这些方法都是基于图像的底层特征实现对高分辨率图像的恢复,由于底层特征对高清图像的边缘、轮廓以及纹理的表达能力有限,从而很大程度上限制了重建效果。

## 3 基于深度学习的超分辨重建算法

深度学习方法能够自适应地从训练集中学习得到更具表达能力的特征,在最近几年被广泛应用于图像超分辨领域中。卷积神经网络是深度学习中最为人熟知的网络结构。Krizhevsky等利用卷积神经网络在ImageNet分类问题上取得了当时最好的成绩,相比于基于传统特征的分类方法,准确率有了显著的提升,从而使得卷积神经网络受到了研究人员极大的关注<sup>[61]</sup>。此后,深度学习开始成为计算机视觉、机器学习、自然语言处理等领域的研究热点。近年来,深度学习在图像分类<sup>[62-63]</sup>、物体检测<sup>[64-65]</sup>、图像分割<sup>[66-67]</sup>等计算机视觉高层任务上取得了长足发展,同时在图像生成<sup>[68-69]</sup>、图像修复<sup>[70-71]</sup>、图像去雾<sup>[72-73]</sup>等不同的计算机视觉底层任务上也被广泛应用,性能指标很大程度上超越了传统算法。

2014年,Dong等<sup>[29-30]</sup>率先将深度学习应用于自然图像超分辨重建任务中,提出了超分辨卷积网络(Super-resolution convolutional neural networks, SRCNN)模型。Dong等将自然图像超分辨视为回归任务,构建了一个3层卷积神经网络直接学习低分辨率图像到高分辨率图像之间的非线性映射。在训练阶段,Dong等采用带动量的随机梯度下降算法优化 $L_2$ 损失函数,寻找最优的网络模型参数。训练完成后,利用得到的网络模型便可直接实现对高分辨率图像的恢复。在SRCNN提出后的几年,随着深度学习技术的发展及其在计算机视觉各类任务的广泛应用,研究人员提出了一系列基于深度学习的超分辨重建算法。

### 3.1 网络模型结构设计

网络模型结构设计是基于深度学习的超分辨率重建算法的核心部分,利用更具表达能力的模型恢复得到的高分辨率图像通常具有更优的图像质量,很多研究者从不同的角度出发,针对图像超分辨率问题,对网络模型进行改进。

#### (1) 深度结构

通常情况下,更深的网络模型往往具有更强的特征学习能力,而最早提出来的SRCNN超分辨率模型只是一个浅层模型,如何构建适用于超分辨率领域的深度模型值得探讨。对此, Kim等<sup>[31]</sup>提出一个深度的超分辨率网络模型VDSR(Very deep super-resolution),采用了20个卷积层。Kim等通过实验验证了采用更深网络模型能够带来更好的重建效果,但也会对收敛速度产生很大影响。因此为了加速收敛,VD-SR模型直接对高低分辨率图像的残差进行学习。同时在训练过程中采用较高的学习速率(SRCNN的 $10^4$ 倍),并采用梯度截断的学习策略使训练过程更加稳定。随着残差网络ResNet<sup>[63]</sup>的出现,使得深度网络能够得到更好的训练。Lim等<sup>[32]</sup>利用残差模块构建更深的超分辨率重建模型EDSR(Enhanced deep super-resolution),其中的残差模块得到了有效的改进,最主要的改变是去掉了批归一化层,从而在训练过程中能够节省近40%的显存占用,为更深的网络模型训练创造了可能。EDSR模型采用了32个残差模块,共包含69个卷积层,重建效果得到了明显提升。Zhang等<sup>[33]</sup>则进一步将残差模块和残差学习相结合,提出了Residual in Residual网络结构,从而构建了更深的网络模型RCAN(Residual channel attention networks),含有500多个卷积层。

#### (2) 递归结构

提高网络模型深度往往会引入更多的参数,为了避免过拟合的发生,需要采用更多的训练图像,同时较大的网络模型也不易于存储。为了解决这些问题,一些研究人员提出采用递归结构,通过对某些卷积层进行复用,以达到控制模型参数的情况下,提升图像的重建效果。Kim等<sup>[34]</sup>利用递归模块构建得到DRCN(Deep-recursive convolutional networks)模型,在增大网络感受野的同时,能够重复利用网络参数,并通过递归监督学习和跳线连接来降低模型训练的难度,从而以较少的模型参数实现较好的重建效果。在DRCN基础上,文献[39]提出的DRRN(Deep recursive residual network)模型进一步将残差结构与递归模块相结合,通过在全局和局部方式采用残差学习,有效地降低了深度网络的训练难度。在文献[40]的MemNet(Persistent memory network)网络中,Tai等利用递归单元去学习当前状态的多层表达以作为短时记忆,并通过构建若干个记忆模块,将其输出作为长时记忆输入单元中,以解决网络模型层数加深所带来的长时依赖问题。

#### (3) 密集连接

受图像分类模型网络DenseNet<sup>[74]</sup>的启发,近年来出现了一些基于密集连接的超分辨率网络。不同于传统的连接方式,密集连接能够充分地利用网络模型各个层次的特征,使得超分辨率模型能获得更加丰富的特征表达,从而提升高分辨率图像的重建效果。Zhang等<sup>[37]</sup>将残差模块和密集模块组合,设计得到残差密集模块,并以此为基本单元构建得到RDN(Residual dense networks)超分辨率模型。通过密集连接,RDN能够得到足够的局部特征,同时进一步利用局部特征融合,从中自适应地学习得到更加有效的特征。此外,该方法利用全局特征融合,将网络模型从低到高不同层次的特征进行充分融合和利用。D-DBPN(Dense deep back-projection networks)网络<sup>[38]</sup>则利用误差反馈机制设计一系列上采样和下采样卷积层,彼此之间通过密集连接进行特征信息的流通,在较大的放大尺度上(如 $\times 8$ )取得了很好的重建效果。

#### (4) 结合注意力机制

对于大多数基于深度学习的超分辨方法,在网络模型学习到的某层特征中,不同位置以及不同通道上的神经元激活值都具有相同的权重,而结合注意力机制的超分辨模型则能挑选出对于超分辨任务的更重要的激活值,赋予更多权重,从而提升重建效果。Choi等<sup>[39]</sup>提出选择性单元,实现单元映射和基于sigmoid函数的选择模块之间的逐像素点乘操作。通过将选择性单元放置在卷积层之间,使得网络模型能够自适应地选择特征图上更有利于重建结果的响应值。不同于文献<sup>[39]</sup>,RCAN网络<sup>[37]</sup>从通道注意力这一角度出发,利用残差通道注意力模块,使网络能够对不同特征通道赋予不同的权重,并构建非常深的网络,从而使得超分辨重建的效果得到较大提升。Dai等<sup>[40]</sup>则提出二阶注意力模型SAN(Second-order attention network),该模型的关键是二阶通道注意力模块,其利用二阶特征统计值自适应地实现对通道特征的调整,从而学习得到更具表达力的特征表达和特征关系。

#### (5) 基于生成对抗网络的方法

很多超分辨模型采用逐像素损失函数(如 $L_2$ 和 $L_1$ 损失函数)作为优化函数,其重建结果的峰值信噪比(Peak signal to noise ratio, PSNR)虽然较高,但是往往会出现过模糊的现象,视觉感知质量不佳<sup>[41-42]</sup>。为了提升重建结果的感知质量,研究人员将GAN引入到图像超分辨重建邻域。GAN最早是由Ledig等<sup>[43]</sup>提出,旨在通过生成器和判别器的对抗训练,使得网络模型能够逐步实现对训练数据分布的学习,从而达到生成数据的目的。Ledig等<sup>[42]</sup>提出了超分辨对抗网络(Super-resolution generative adversarial network, SRGAN)模型,其中生成器的输入是低分辨率图像,输出是高分辨图像,而判别器则需要判断输入图像是真实图像,亦或是生成器得到生成图像。需要注意的是,SRGAN模型将对抗损失和感知损失相结合作为最终的损失函数。相比于前人方法,虽然SRGAN重建结果的PSNR值相对较低,但其视觉效果更好,在图像细节上更加逼真。受文献<sup>[75]</sup>启发,EhanceNet在感知损失的基础上添加了纹理匹配损失,该项反映的是图像特征Gram矩阵之间的欧式距离,使得网络生成的结果具有更加逼真的纹理结构<sup>[44]</sup>。加强SRGAN网络ESRGAN(Enhanced SRGAN)<sup>[45]</sup>则是在SRGAN进行了3点改进:一是使用改进的生成器基础模块;二是借鉴Relativistic GAN<sup>[76]</sup>的思想,使判别器预测相对真实度而不是绝对值;三是对感知损失进行了改进,有助于保持亮度的一致性和恢复更加真实的纹理。

### 3.2 上采样方式

上采样操作是基于深度学习的超分辨网络模型中重要环节,大致可以分为基于插值的上采样和基于学习的上采样两大类。

#### (1) 基于插值的上采样

基于插值的上采样早期传统做法可细分为最近邻插值、双线性插值和双三次插值。该方法主要基于图像连续性假设,由于只依赖于图像自身,没有引入新的信息,所以往往会带来模糊效应、噪声放大等负面效果。早期的一些基于深度学习的网络模型,如SRCNN, VDSR等通常会先对低分辨率图像进行双立方插值操作,然后学习插值后的图像和参考图像之间的非线性映射。由于网络模型是在高维空间进行操作,所以重建过程中消耗的时间和空间都相对较大。

#### (2) 基于学习的上采样

为了克服上述问题,研究人员将上采样操作放置在网络模型之中,通过端到端的方式自动学习。主要分为两种:第一种为转置卷积,也称反卷积<sup>[77]</sup>,它是先对输入图像进行填零扩充操作,然后再进行卷积操作得到分辨率变大的输出图像;第二种是亚像素卷积<sup>[78]</sup>,它是对输入图像使用卷积操作,得到 $r^2$ 个特征图,然后通过周期筛选的方式得到放大 $r$ 倍的输出图像。相对于转置卷积,亚像素层能够获得得到更大的感受野,从而使用更多的上下文信息用于细节重建。需要注意的是,不同于基于插值的上采

样,基于学习的上采样层通常放置在网络模型的末端。也就是说超分辨率网络模型在低维空间进行特征提取后,再采用转置卷积或亚像素卷积完成对高分辨率图像的重构,从而避免了直接在高维空间的大量计算。

### 3.3 损失函数形式

在图像超分辨重建领域,损失函数是用来定义超分辨模型所生成的高分辨率图像与真实的参考图像之间的差异,在基于深度学习的优化过程中起着至关重要的作用。在早期阶段,研究人员主要使用逐像素均方误差(Mean-square error, MSE)损失函数。但后续的研究表明,该损失函数并不能很好地反映重建图像的感知质量<sup>[41]</sup>,一系列其他的损失函数被陆续提出来,如感知损失、纹理损失以及对抗损失函数。当前很多基于生成对抗网络的超分辨模型通常是把这些损失函数结合起来进行优化<sup>[42]</sup>。

#### (1) 逐像素损失

逐像素损失主要是度量两个图像之间的逐像素之间的差异,主要包括 $L_2$ 损失(均方误差)和 $L_1$ 损失(平均绝对误差),具体定义如下

$$\mathcal{L}_{\text{pixel}_{12}}(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{N} \sum_{i=1}^N \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_2^2 \quad (11)$$

$$\mathcal{L}_{\text{pixel}_{11}}(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{N} \sum_{i=1}^N \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_1 \quad (12)$$

文献[36]表明,相比于 $L_2$ 损失函数, $L_1$ 损失函数能够改善收敛和提高重建性能。利用 $L_2$ 或 $L_1$ 损失函数进行优化的网络模型通常能够获取相对较高的PSNR值,重建结果中目标的边缘和轮廓较为清晰,但由于没有考虑高层的语义感知信息,所以结果中的纹理往往会出现过模糊的现象,感知质量相对较差。

#### (2) 感知损失

为了评测不同图像间感知质量的差异,研究人员将感知距离引入到了超分辨率领域。具体地,使用已经充分训练好的自然图像分类模型,如VGG<sup>[79]</sup>和ResNet<sup>[63]</sup>等,用以提取不同图像的特征,然后计算特征空间上的欧式距离,其定义如下

$$\mathcal{L}_{\text{percep}}(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{h_l \omega_l k_l} \sum_{i,j,k} (\phi_{i,j,k}^{(l)}(\mathbf{x}) - \phi_{i,j,k}^{(l)}(\hat{\mathbf{x}}))^2 \quad (13)$$

式中: $h_l$ ,  $\omega_l$ 和 $k_l$ 分别表示第 $l$ 层特征图的高、宽和特征层数目。

#### (3) 纹理损失

受文献[76]启发,Sajjadi等<sup>[44]</sup>提出了纹理损失,用来描述生成图像与参考图像在纹理风格上的差异,具体定义如下

$$\mathcal{L}_{\text{percep}}(\mathbf{x}, \hat{\mathbf{x}}) = \left\| G(\phi(\mathbf{x})) - G(\phi(\hat{\mathbf{x}})) \right\|_2^2 \quad (14)$$

式中 $G(F) = FF^T \in \mathbf{R}^{n \times n}$ 为Gram矩阵。

#### (4) 对抗损失

在基于GAN的超分辨率网络模型中,通常会在逐像素损失或感知损失的基础上再结合对抗损失,旨在通过生成器和判别器的相互竞争,使得判别器能够从真实的参考图像中提取出难以学习的潜在模式,并强迫生成器进行模型调整,从而使生成器能够产生逼真的高分辨率图像。基本的损失函数如下

$$\begin{aligned} \mathcal{L}_{\text{gan,g}} &= \sum_{i=1}^N -\log(D(\mathbf{x}_i)) \\ \mathcal{L}_{\text{gan,d}} &= \sum_{i=1}^N -\log(D(\mathbf{x}_i)) - \log(1 - d(\hat{\mathbf{x}}_i)) \end{aligned} \quad (15)$$



式中: $\mathcal{L}_{\text{gan}_g}$ 和 $\mathcal{L}_{\text{gan}_d}$ 分别为生成器和判别器所对应的对抗损失函数。需要注意的是,采用不同的GAN模型,损失函数的基本形式会有所区别。

由于现阶段GAN的训练仍比较困难且不稳定,所以结合了对抗损失的超分辨模型有时会产生人工痕迹和不自然的变形。如何更好地将GAN应用于图像超分辨率重建领域仍值得深入研究。

### 3.4 复杂条件下的超分辨率算法设计

#### (1) 复杂的退化模型

通常情况下,低分辨率图像被视为高分辨率图像经过模糊、下采样和噪声干扰后所得到的退化图像。从单一低分辨率图像恢复重构出高分辨率图像,是一个难解的欠定问题,研究人员需要利用先验信息去约束解空间,如早期的稀疏先验<sup>[21]</sup>以及近几年利用深度学习直接从大量训练数据中学习到的图像先验信息。然而,Efrat等<sup>[80]</sup>通过实验指出“精确的模糊核模型要比复杂的图像先验更加重要”,当训练过程中采用的模糊核和测试时的模糊核不一致时,重建结果往往会产生明显的退化。Zhang等<sup>[46]</sup>考虑多种退化方式,包含不同的模糊核和不同的噪声水平,并通过维度延展策略将模糊核参数和噪声参数编码成退化图,与低分辨率图一同作为网络输入,于是深度学习学习的是退化图和低分辨率图到高分辨率图之间的映射。但文献[46]中的方法假设退化模型是已知的,针对真实的低分辨率图像,需要先进行网格搜索寻找退化参数,这极大地限制了该方法的适用性。对此,许多研究人员致力于盲超分辨领域的相关研究。Bulat等<sup>[47]</sup>使用High-to-Low GAN对图像的退化方式进行学习,其中生成器对高分辨率图像进行退化操作,判别器则判断输入图像是真实的低分辨率图像还是生成的图像,同时使用Low-to-High GAN完成低分辨率图像的重构。Bulat等将该方法应用于人脸的超分辨中,有效地提升了真实场景下人脸的重建效果。针对模糊核未知这个问题,Gu等<sup>[48]</sup>提出迭代核校正(Iterative kernel correction, IKC)方法用于模糊核的估计,并使用空间特征变换层将估计的模糊核参数嵌入到超分辨网络中,需要注意的是,文献[48]主要是对高斯模糊核进行估计。Zhou等<sup>[49]</sup>则利用模糊核估计算法<sup>[50]</sup>和GAN对真实的自然场景图片中退化方式进行估计和建模,构建模糊核池,并从中随机提取模糊核,从而构建高低分辨率训练图像对,完成盲超分辨模型的训练。不同于文献[48]中高斯模糊核的假设,文献[49]中模糊核的方式很大程度上取决于文献[50]中模糊核估计算法能够估计出哪些类型的模糊核。由于真实的退化过程复杂多变,如何能够更加有效地对退化模型进行估计,仍然是一个开放的问题,等待着研究人员进一步解决。

#### (2) 真实的应用场景

在超分辨以往的研究中,研究人员通常是采用人为设定的降质模型(如双三次立方下采样或高斯模糊下采样等)对高分辨图像进行处理,以得到仿真的低分辨率图像,然后再训练得到超分辨网络模型。然而,这些人工降质模型与真实图像获取过程中的退化过程往往有一定差异,因此研究人员针对真实数据获取过程,提出了一些满足真实应用场景的超分辨数据集,并进行超分辨率模型的设计。Chen等<sup>[51]</sup>从真实的相机镜头成像的角度出发,构建分辨率和视场之间的关系,并简称之为R-V退化模型。同时,利用相机对同一场景进行拍摄,提出了满足R-V退化模型的高低分辨率图像数据集City100。之前绝大多数超分辨算法都是在8 bit的RGB图像上进行操作,Zhang等<sup>[52]</sup>则考虑直接使用12~14 bit的原始数据(Raw格式),由于没有经过相机图像处理器的处理,相比于RGB图像,原始数据包含的图像高频信息更加完整。Zhang等使用变焦镜头构建得到SR-RAW数据集,并采用深度网络模型对原始数据进行超分辨率重建,其实验结果优于直接在RGB训练的模型。与Zhang等提出的方法类似,Cai等<sup>[53]</sup>通过调整数码相机焦距获得同一场景的高低分辨率图像对,在此基础上提出了一种有效的图像匹配算法完成图像对之间像素级别的对齐,并采用基于拉普拉斯金字塔核预测网络完成超分辨任务。

### 3.5 小结

相比于基于传统特征的超分辨率重建算法,基于深度学习的方法能够从大量数据中自动地学习得到适用于超分辨率任务的特征,从而极大地提升了重建的效果。本节从网络模型结构设计、上采样方式和损失函数形式3个方面对基于深度学习的超分辨率重建算法进行详细介绍,并分析在复杂条件下的超分辨率算法设计。其中,网络模型结构设计是基于深度学习的超分辨率重建算法中的核心部分,研究人员从不同的角度出发,设计并利用深度结构、递归结构、密集连接、与注意力机制相结合等不同的思路完成网络模型的结构设计。此外,为了提升重建图像的感知质量,生成对抗网络被引入到了超分辨率重建领域。

## 4 图像超分辨率重建数据集与评价指标

在图像超分辨率领域,研究人员利用基准数据集对不同算法的性能进行评测,包括 Set5<sup>[81]</sup>、Set14<sup>[82]</sup>、BSD100<sup>[83]</sup>、Urban100<sup>[84]</sup>和 Manga109<sup>[85]</sup>等,这些数据集的数量及图像内容都有一定的差异。除了这些基准数据集,还有一些数据集经常被当作训练集应用于该领域,如 ImageNet<sup>[86]</sup>、DIV2K<sup>[87]</sup>、Flickr2K<sup>[88]</sup>和 OST<sup>[89]</sup>等。需要注意的是,这些数据集的低分辨率图像通常都是采用人工设定的退化模型仿真得到(如双三次插值下采样和高斯模糊下采样等)。最近一年,很多研究者发布了在真实的应用场景下获取得到的高低分辨率图像数据集,如 City100<sup>[51]</sup>、SR-RAW<sup>[52]</sup>和 RealSR<sup>[53]</sup>等。

用于评价超分辨率重建结果的客观指标主要分为两大类:一类是全参考指标,包括 PSNR,结构相似性(Structure similarity, SSIM)<sup>[90]</sup>以及最近被提出来的学习感知图像相似度(Learned perceptual image patch similarity, LPIPS)<sup>[91]</sup>等;另一类是无参考指标,包括 PI<sup>[92]</sup>、MA<sup>[93]</sup>和 NIQE<sup>[94]</sup>等。其中 PSNR 和 SSIM 是2个最为常用的指标,被广泛应用于各种图像处理相关的底层任务中。

(1) PSNR 主要由 MSE 决定,定义如下

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2 \quad (16)$$

$$\text{PSNR} = 10 \log_{10} \left( \frac{x_{\max}^2}{\text{MSE}} \right) \quad (17)$$

式中: $x$ 为真实的参考图像; $\hat{x}$ 为重构图像; $x_{\max}$ 为参考图像中的最大像素值。需要注意的是,最小化 MSE 损失函数(亦称  $L_2$  损失函数)等效于最大化 PSNR。因为 MSE 度量的是相同位置的逐像素差异,并没有考虑到人的视觉感知,因此易导致图像感知质量较差<sup>[46]</sup>。但由于当前缺乏准确的感知质量指标,因此 PSNR 在超分辨率领域仍被广泛使用。

(2) SSIM 度量图像之间的结构相似性<sup>[90]</sup>,由图像间的亮度、对比度和结构所共同决定,其定义如下

$$\text{SSIM}(x, \hat{x}) = \frac{(2\mu_x \mu_{\hat{x}} + c_1)(\sigma_{x\hat{x}} + c_2)}{(\mu_x^2 + \mu_{\hat{x}}^2 + c_1)(\sigma_x^2 + \sigma_{\hat{x}}^2 + c_2)} \quad (18)$$

式中: $\mu_x$ 和  $\mu_{\hat{x}}$ 分别为  $x$ 和  $\hat{x}$ 的均值; $\sigma_x$ 和  $\sigma_{\hat{x}}$ 为方差; $\sigma_{x\hat{x}}$ 为  $x$ 和  $\hat{x}$ 的协方差; $c_1$ 和  $c_2$ 为常数。

(3) LPIPS 是近年提出的用来度量图像间的感知距离<sup>[91]</sup>,其利用一个预训练的深度学习模型,通过提取参考图和重构图像特征并计算  $L_2$  距离得到,表达式为

$$\text{LPIPS}(x, \hat{x}) = \sum_l \frac{1}{N} \left\| \omega_l \odot (\phi(x)_l - \phi(\hat{x})_l) \right\|_2^2 \quad (19)$$

式中: $\phi(x)_l$ 表示使用预训练模型对  $x$ 提取的第  $l$ 层特征; $\omega_l$ 为第  $l$ 层特征所对应的权重;运算符  $\odot$  则表

示逐像素乘积操作。

(4) 无参考图像质量指标:MA 是利用底层特征和回归模型对主观评价分数进行学习,得到回归模型参数后,用于非参考图像的质量评价<sup>[93]</sup>;NIQE 利用待评价图像的特征模型参数与预先在原始图像库上建立的模型参数之间的距离来确定图像质量<sup>[94]</sup>;PI 则进一步将 MA 和 NIQE 这两个指标结合起来综合考虑<sup>[92]</sup>,表达式为

$$PI = \frac{1}{2} ((10 - MA) + NIQE) \quad (20)$$

此外,研究人员有时会采用主观指标,如平均意见分数(Mean opinion score, MOS),评价图像的感知质量。在计算 MOS 时,评分人员需要对测试图像的感知质量进行打分,取值范围通常为 1~5 之间,计算所有测试图像的平均得分则得到最终的 MOS 值。

需要注意的是,绝大多数超分辨算法主要报道了在基准数据集上的 PSNR 和 SSIM 值,详细结果可见表 1—3,其中上标 a 表示数据来源文献[40],上标 b 表示数据来源文献[30],上标 c 表示数据来源文献[1]。从表中可以看出,基于深度学习的超分辨重建算法的实验效果大幅度地超过了基于传统特征的方法,这表明对于图像超分辨重建,相比于传统的底层特征,利用深度学习方法所得到的自适应特征更有助于高频信息的恢复和重构<sup>[99-100]</sup>。

表 1 基准数据集中不同超分辨算法的平均 PSNR 和 SSIM 对比(尺度因子为×2)

Table 1 Average PSNR and SSIM values of different super-resolution methods on some benchmark datasets (scale factor is ×2)

方法	Set5	Set14	BSD100	Urban100	Manga109
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Bicubic <sup>a</sup>	33.66/0.929 9	30.24/0.868 8	29.56/0.843 1	26.88/0.840 3	30.80/0.933 9
NE+LLE <sup>[20]b</sup>	35.77/0.949 0	31.76/0.899 3	—/—	—/—	—/—
ANR <sup>[26]b</sup>	35.83/0.949 9	31.80/0.900 4	—/—	—/—	—/—
A+ <sup>[27]b</sup>	36.54/0.954 4	32.28/0.905 6	—/—	—/—	—/—
SRCNN <sup>[30]a</sup>	36.66/0.954 2	32.45/0.906 7	31.36/0.887 9	29.50/0.894 6	35.60/0.966 3
FSRCNN <sup>[95]a</sup>	37.05/0.956 0	32.66/0.909 0	31.53/0.892 0	29.88/0.902 0	36.67/0.971 0
SCN <sup>[96]c</sup>	36.52/0.953 0	32.42/0.904 0	31.24/0.884 0	29.50/0.896 0	35.51/0.967 0
VDSR <sup>[31]a</sup>	37.53/0.959 0	33.05/0.913 0	31.90/0.896 0	30.77/0.914 0	37.22/0.975 0
LapSRN <sup>[97]a</sup>	37.52/0.959 1	33.08/0.913 0	31.08/0.895 0	30.41/0.910 1	37.27/0.974 0
DRCN <sup>[34]c</sup>	37.63/0.958 8	33.06/0.912 1	31.85/0.894 2	30.76/0.913 3	37.57/0.973 0
DRRN <sup>[35]c</sup>	37.74/0.959 1	33.23/0.913 6	32.05/0.897 3	31.23/0.918 8	37.92/0.976 0
ZSSR <sup>[98]c</sup>	37.37/0.957 0	33.00/0.910 8	31.65/0.892 0	—/—	—/—
MemNet <sup>[36]a</sup>	37.78/0.959 7	33.28/0.914 2	32.08/0.897 8	31.31/0.919 5	37.72/0.974 0
EDSR <sup>[32]a</sup>	38.11/0.960 2	33.92/0.919 5	32.32/0.901 3	32.93/0.935 1	39.10/0.977 3
MDSR <sup>[32]c</sup>	38.11/0.960 2	33.85/0.919 8	32.29/0.900 7	32.84/0.934 7	38.96/0.978 0
SRMD <sup>[46]a</sup>	37.79/0.960 1	33.32/0.915 9	32.05/0.898 5	31.33/0.920 4	38.07/0.976 1
D-DBPN <sup>[38]c</sup>	38.09/0.960 0	33.85/0.919 0	32.27/0.900 0	32.55/0.932 4	38.89/0.977 5
RDN <sup>[37]a</sup>	38.24/0.961 4	34.01/0.921 2	32.34/0.901 7	32.89/0.935 3	39.18/0.978 0
RCAN <sup>[33]a</sup>	38.27/0.961 4	34.11/0.921 6	32.41/0.902 6	33.34/0.938 4	39.43/0.978 6
SAN <sup>[40]a</sup>	38.31/0.962 0	34.07/0.921 3	32.42/0.902 8	33.10/0.937 0	39.32/0.979 2

表2 基准数据集中不同超分辨算法的平均PSNR和SSIM对比(尺度因子为×3)

Table 2 Average PSNR and SSIM values of different super-resolution methods on some benchmark dataset (scale factor is ×3)

方法	Set5	Set14	BSD100	Urban100	Manga109
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Bicubic <sup>a</sup>	30.39/0.868 2	27.55/0.774 2	27.21/0.738 5	24.46/0.734 9	26.95/0.855 6
SC <sup>[22]b</sup>	31.42/0.882 1	28.31/0.795 4	—/—	—/—	—/—
NE+LLE <sup>[20]b</sup>	31.84/0.895 6	28.60/0.807 6	—/—	—/—	—/—
ANR <sup>[26]b</sup>	31.92/0.896 8	28.65/0.809 3	—/—	—/—	—/—
A+ <sup>[27]b</sup>	32.59/0.908 8	29.13/0.818 8	—/—	—/—	—/—
SRCNN <sup>[30]a</sup>	32.75/0.909 0	29.30/0.821 5	28.41/0.786 3	26.24/0.798 9	30.48/0.911 7
FSRCNN <sup>[95]a</sup>	33.18/0.914 0	29.37/0.824 0	28.53/0.791 0	26.43/0.808 0	31.10/0.921 0
SCN <sup>[96]c</sup>	32.62/0.908 0	29.16/0.818 0	28.33/0.783 0	26.21/0.801 0	30.22/0.914 0
VDSR <sup>[31]a</sup>	33.67/0.921 0	29.78/0.832 0	28.83/0.799 0	27.14/0.829 0	32.01/0.934 0
LapSRN <sup>[97]a</sup>	33.82/0.922 7	29.87/0.832 0	28.82/0.798 0	27.07/0.828 0	32.21/0.935 0
DRCN <sup>[34]c</sup>	33.82/0.922 6	29.77/0.831 4	28.80/0.796 3	27.15/0.827 7	32.31/0.936 0
DRRN <sup>[35]c</sup>	34.03/0.924 4	29.96/0.834 9	28.95/0.800 4	27.53/0.837 7	32.74/0.939 0
ZSSR <sup>[98]c</sup>	33.42/0.918 8	29.80/0.830 4	28.67/0.794 5	—/—	—/—
MemNet <sup>[36]a</sup>	34.09/0.924 8	30.01/0.835 0	28.96/0.800 1	27.56/0.837 6	32.51/0.936 9
EDSR <sup>[32]a</sup>	34.65/0.928 0	30.52/0.846 2	29.25/0.809 3	28.80/0.865 3	34.17/0.947 6
MDSR <sup>[32]c</sup>	34.66/0.928 0	30.44/0.845 2	29.25/0.809 1	28.79/0.865 5	34.17/0.947 0
SRMD <sup>[46]a</sup>	34.12/0.925 4	30.04/0.838 2	28.97/0.802 5	27.57/0.839 8	33.00/0.940 3
RDN <sup>[37]a</sup>	34.71/0.929 6	30.57/0.846 8	29.26/0.809 3	28.80/0.865 3	34.13/0.948 4
RCAN <sup>[33]a</sup>	34.74/0.929 9	30.64/0.848 1	29.32/0.811 1	29.08/0.870 2	34.43/0.949 8
SAN <sup>[40]a</sup>	34.75/0.930 0	30.59/0.847 6	29.33/0.811 2	28.93/0.867 1	34.30/0.949 4

## 5 图像超分辨重建应用

### 5.1 视频超分辨

不同于单幅图像超分辨,视频超分辨(Video super-resolution, VSR)利用低分辨率视频序列重建得到高分辨率视频序列。早期的视频超分辨方法通常将该任务视为一个优化问题,需要反复迭代求解,计算量较大。而近年来,很多学者则利用深度学习去构建视频帧序列之间的时空依赖关系。这些方法大体可以分为两类:一是基于运动补偿的算法;二是基于递归的算法,无需显式地计算光流或运动补偿。对于基于运动补偿的方法,Caballero等<sup>[101]</sup>使用一个多分辨率空间变换模块来估计运动补偿,并将此与低分辨率图像帧序列相结合,输入到ESPCN网络模块<sup>[78]</sup>中,从而得到重建结果。Tao等<sup>[102]</sup>则将亚像素运动补偿模块嵌入卷积神经网络模型中,通过端到端的方式,同时学习运动补偿和完成超分辨率重建,从而能够更好地从多帧图像中恢复出细节信息。对于基于递归的方法,Jo等<sup>[103]</sup>使用一个端到端的卷积网络模型,通过计算图像像素的局部时空邻域关系,同时生成动态上采样滤波器和残差图,从而避免显式地计算运动补偿。Sajjadi等<sup>[104]</sup>提出了一个帧递归视频超分辨框架,通过利用先前估计出的高分辨率图像帧辅助后续高分辨率图像帧的估计。Yi等<sup>[105]</sup>则提出逐步融合网络以充分利用视频帧序列之间的时空信息,并引入非局部网络<sup>[106]</sup>去学习视频帧序列间的长距离依赖关系。

表3 基准数据集中不同超分辨算法的平均PSNR和SSIM对比(尺度因子为 $\times 4$ )Table 3 Average PSNR and SSIM values of different super-resolution methods on some benchmark dataset (scale factor is  $\times 4$ )

方法	Set5	Set14	BSD100	Urban100	Manga109
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Bicubic <sup>a</sup>	28.42/0.810 4	26.00/0.702 7	25.96/0.667 5	23.14/0.657 7	24.89/0.786 6
NE+LLE <sup>[20]b</sup>	29.61/0.840 2	26.81/0.733 1	—/—	—/—	—/—
ANR <sup>[26]b</sup>	29.69/0.841 9	26.85/0.735 2	—/—	—/—	—/—
A+ <sup>[27]b</sup>	30.28/0.860 3	27.32/0.749 1	—/—	—/—	—/—
SRCNN <sup>[30]a</sup>	30.48/0.862 8	27.50/0.751 3	26.90/0.710 1	24.52/0.722 1	27.58/0.855 5
FSRCNN <sup>[95]a</sup>	30.72/0.866 0	27.61/0.755 0	26.98/0.715 0	24.62/0.728 0	27.90/0.861 0
SCN <sup>[96]c</sup>	30.39/0.862 0	27.48/0.751 0	26.87/0.710 0	24.52/0.725 0	27.39/0.857 0
VDSR <sup>[31]a</sup>	31.35/0.883 0	28.02/0.768 0	27.29/0.072 6	25.18/0.754 0	28.83/0.887 0
LapSRN <sup>[97]a</sup>	31.54/0.885 0	28.19/0.772 0	27.32/0.727 0	25.21/0.756 0	29.09/0.890 0
DRCN <sup>[34]c</sup>	31.53/0.885 4	28.03/0.767 3	27.24/0.723 3	25.14/0.751 1	28.97/0.886 0
DRRN <sup>[35]c</sup>	31.68/0.888 8	28.21/0.772 0	27.38/0.728 4	25.44/0.763 8	29.46/0.896 0
ZSSR <sup>[98]c</sup>	31.13/0.879 6	28.01/0.765 1	27.12/0.721 1	—/—	—/—
MemNet <sup>[36]a</sup>	31.74/0.889 3	28.26/0.772 3	27.40/0.728 1	25.50/0.763 0	29.42/0.894 2
EDSR <sup>[32]a</sup>	32.46/0.896 8	28.80/0.787 6	27.71/0.742 0	26.64/0.803 3	31.02/0.914 8
MDSR <sup>[32]c</sup>	32.50/0.897 3	28.72/0.785 7	27.72/0.741 8	26.67/0.804 1	31.11/0.915 0
SRMD <sup>[46]a</sup>	31.96/0.892 5	28.35/0.778 7	27.49/0.733 7	25.68/0.773 1	30.09/0.902 4
D-DBPN <sup>[38]c</sup>	32.47/0.898 0	28.82/0.786 0	27.72/0.740 0	26.38/0.794 6	30.91/0.913 7
RDN <sup>[37]a</sup>	32.47/0.899 0	28.81/0.787 1	27.72/0.741 9	26.61/0.802 8	31.00/0.915 1
RCAN <sup>[33]a</sup>	32.62/0.900 1	28.86/0.788 8	27.76/0.743 5	26.82/0.808 7	31.21/0.917 2
SAN <sup>[40]a</sup>	32.64/0.900 3	28.92/0.788 8	27.78/0.743 6	26.79/0.806 8	31.18/0.916 9

## 5.2 遥感图像超分辨

遥感图像是利用遥感卫星所搭载的传感器从远处非接触的方式对地面或空间目标的电磁波特性记录下来所形成的影像。相比于对成像设备的改进以提高遥感图像的分辨率,采用图像超分辨技术是一种更加廉价方便的方式。稀疏学习方法被广泛应用于遥感图像处理领域<sup>[107-108]</sup>,包括遥感图像超分辨重建。Pan等<sup>[109]</sup>将压缩感知与结构自相似性进行结合,利用学习好的稀疏字典完成单幅遥感图像超分辨。Chavez-Roman等<sup>[110]</sup>基于稀疏表示和离散小波变换,从单幅低分辨图像重建得到高分辨图像。最近几年,深度学习被广泛应用在遥感图像超分辨领域中。Lei等<sup>[111]</sup>利用深度卷积神经网络去学习遥感图像中的多层次特征表达;Wang等<sup>[112]</sup>利用多个卷积神经网络去分别近似遥感图像中的小波多层次表达。Pan等<sup>[113]</sup>则利用残差模块和密集连接提升遥感图像超分辨率效果。需要注意的是,当前遥感图像超分辨率算法大多是从自然图像超分辨率领域中进行借鉴,如何将深度学习方法与遥感图像特性结合是遥感图像超分辨领域未来的发展方向。

## 5.3 超分辨重建在视觉高层任务中的应用

虽然图像超分辨算法是视觉底层任务,但是大量的研究人员将它与视觉高层任务联系起来,用以提升高层任务的效果。Dai等<sup>[12]</sup>通过实验验证图像超分辨对于边缘检测、语义分割、数字识别和场景识

别等任务的作用,即在输入为低分辨率图像的情况下,通过对输入图像采用超分辨率处理能够提升这些任务的性能。FSR-GAN<sup>[114]</sup>则是从特征空间的角度出发,使用GAN框架实现特征超分辨,将低分辨率图像的原始特征转换成更具判别力的特征,从而大幅提升了对低分辨率图像的检索结果。此外,很多研究人员将图像超分辨技术应用于目标检测领域,对小目标的像素空间或特征空间加以改善,从而提升小目标检测的效果。在像素空间方面,Bai等<sup>[115]</sup>将超分辨率技术应用于极小人脸检测流程中,通过使用生成对抗网络将模糊低分辨率人脸重构成清晰的高分辨率人脸,提升对极小人脸的检测效果。同时,Bai等<sup>[13]</sup>采用类似的技术,并结合多任务学习,应用于通用目标检测任务中。在特征空间方面,Li等<sup>[116]</sup>提出Perceptual GAN对小尺寸目标的特征进行变换,改善与大尺寸目标之间的特征差异。Noh等<sup>[117]</sup>则采用图像金字塔的方式,构造高低分辨率目标特征对,从而更加有效地训练超分辨率目标特征生成器。

## 6 总结和展望

本文介绍了近20年主流的图像超分辨重建算法,大体可以分为基于传统特征的超分辨算法和基于深度学习的超分辨率算法。传统的超分辨算法主要可以分为基于邻域嵌入、基于稀疏表示和基于局部线性回归的方法,其大都假设高低分辨率图像块分布在不同的流形上,且具有类似的几何局部结构。然而这些方法主要基于图像的底层特征,特征的表达能力有限。而深度学习方法能够自适应地从大量训练集中学习得到深层特征,相比于传统方法,基于深度学习的超分辨算法的重建效果有了大幅提升。此外,为了改善重建结果的感知质量,生成对抗网络被引入到超分辨率重建领域。同时,针对不同退化模型的超分辨网络被提出来。

尽管基于深度学习的超分辨重建效果已经有了很大提升,但还有很多方面值得进一步研究:

(1) 图像感知质量的提升。现有的基于生成对抗网络的超分辨率网络模型虽然能够在一定程度上提升高分辨图像重建结果的感知质量,但是现阶段生成对抗网络仍面临训练困难和不稳定的难题,重建结果中易存在不理想的人为痕迹和伪影,如何添加图像先验或约束,使得超分辨结果既拥有较好的感知质量,又能避免产生人为痕迹和伪影,这值得进一步研究;此外,目前最为常用的评价指标PSNR和SSIM无法很好地描述图像的感知质量,更为准确的评价指标也值得深入探索。

(2) 网络模型轻量化设计。为了提升重建效果,当前大多数超分辨模型都采用深度网络结构,致使模型参数以及计算量都十分庞大,从而极大地限制了真实场景中的应用(如手机端实时处理)。因此,研究者需要关注超分辨模型的轻量化设计,使得超分辨结果在不发生明显退化的同时,重建速度得到大幅度提升,以满足实际应用中的速度要求。

(3) 盲超分辨重建。现存的大多数超分辨率算法都采用固定的退化方式,然而在真实的应用场景中,图像的退化方式是未知的,从而限制了已有算法的应用效果。虽然已存在一些盲超分辨重建算法,但大多都还是对退化方式进行约束,如各向同性的高斯模糊核参数在一定范围内适用性十分有限,因此更具适用性的盲超分辨重建算法需要进一步研究。

(4) 面向真实数据集的超分辨模型研究。对于当前大多数基于深度学习的超分辨算法,在训练过程中都是利用固定的退化方式将高分辨率图像下采样得到低分辨率图像,然后利用高低分辨率图像完成对模型的训练。近一年来研究人员提出了一系列在真实场景下获取的超分辨数据集,如City100<sup>[51]</sup>、SR-RAW<sup>[52]</sup>和RealSR<sup>[53]</sup>等,相比于传统人工构造的数据集,利用面向真实数据集的超分辨率模型,将更加适用于真实的应用场景,因此值得研究者在未来深入探索。

(5) 图像超分辨对于特定任务的提升。图像超分辨算法不仅能够提升低分辨率图像的图像质量,而且有助于一些其他视觉特定任务,如目标识别、目标检测、人脸识别、视频监控等。如果目标的尺寸比较微小,或者轮廓不够清晰,可以通过结合超分辨重建算法改善对这些目标的处理与分析。然而,如

何有效地将超分辨重建算法和特定任务相结合,需要进一步研究。

#### 参考文献:

- [1] ANWAR S, KHAN S, BARNES N. A deep journey into super-resolution: A survey[J/OL]. (2019-07-17)[2019-11-30]. <https://arxiv.org/abs/1904.07523?context=cs>.
- [2] WANG Z, CHEN J, HOI S C H. Deep learning for image super-resolution: A survey[J/OL]. (2019-02-16)[2019-11-30]. <https://arxiv.org/abs/1902.06068>.
- [3] 李云红,王珍,张凯兵,等.基于学习的图像超分辨重建方法综述[J].计算机工程与应用.2018,54(15):13-21.  
LI Yunhong, WANG Zhen, ZHANG Kaibing, et al. Survey on example learning-based single image super-resolution technique [J]. Computer Engineering and Applications, 2018, 54(15): 13-21.
- [4] ISAAC J S, KULKARNI R. Super resolution techniques for medical image processing[C]// Proceedings of 2015 International Conference on Technologies for Sustainable Development (ICTSD). [S.l.]: IEEE, 2015: 1-6.
- [5] HUANG Y, SHAO L, FRANGI A F. Simultaneous super-resolution and cross-modality synthesis of 3D medical images using weakly-supervised joint convolutional sparse coding[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2017: 6070-6079.
- [6] 黄浩锋,肖南峰.基于组稀疏表示的医学图像超分辨率重建[J].计算机学报,2015(S1):151-153.  
HUANG Haofeng, XIAO Nanfeng. Super-resolution reconstruction of medical images based on group sparse representation[J]. Computer Science, 2015 (S1): 151-153.
- [7] ZHANG L, ZHANG H, SHEN H, et al. A super-resolution reconstruction algorithm for surveillance images[J]. Signal Processing, 2010, 90(3): 848-859.
- [8] RASTI P, UIBOUPIN T, ESCALERA S, et al. Convolutional neural network super resolution for face recognition in surveillance monitoring[C]//Proceedings of International Conference on Articulated Motion and Deformable Objects. Cham: Springer, 2016: 175-184.
- [9] JI H, GAO Z, MEI T, et al. Vehicle detection in remote sensing images leveraging on simultaneous super-resolution[J]. IEEE Geoscience and Remote Sensing Letters, 2019. DOI: 10.1109/LGRS.2019.2930308.
- [10] SHERMEYER J, VAN ET TEN A . The effects of super-resolution on object detection performance in satellite imagery[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. [S.l.]: IEEE, 2019.
- [11] 杨超,杨斌,黄国玉.基于多光谱图像超分辨率处理的遥感图像融合[J].激光与光电子学进展,2016,53(2):88-95.  
YANG Chao, YANG Bin, HUANG Guoyu. Remote sensing image fusion based on multispectral image super-resolution[J]. Laser & Optoelectronics Progress, 2016, 53(2): 88-95.
- [12] DAI D, WANG Y, CHEN Y, et al. Is image super-resolution helpful for other vision tasks?[C]// Proceedings of 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). [S.l.]: IEEE, 2016: 1-9.
- [13] BAI Y, ZHANG Y, DING M, et al. SOD-MTGAN: Small object detection via multi-task generative adversarial network [C]//Proceedings of the European Conference on Computer Vision (ECCV). [S.l.]: Springer, 2018: 206-221.
- [14] ZHANG H, LIU D, XIONG Z. Convolutional neural network-based video super-resolution for action recognition[C]// Proceedings of 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). [S.l.]: IEEE, 2018: 746-750.
- [15] 畅青,冯晶明,洪伟杰,等.基于图像超分辨网络的目标检测算法[J].现代计算机,2019(25):47-50.  
CHANG Qing, FENG Jingming, HONG Weijie, et al. Object detection algorithm based on image super-resolution network [J]. Modern Computer, 2019 (25): 47-50.
- [16] HARRIS J L. Diffraction and resolving power[J]. JOSA, 1964, 54(7): 931-936.
- [17] TSAI R. Multiframe image restoration and registration[J]. Advance Computer Visual and Image Processing, 1984, 1: 317-339.
- [18] FREEMAN W T, PASZTOR E C, CARMICHAEL O T. Learning low-level vision[J]. International Journal of Computer Vision, 2000, 40(1): 25-47.
- [19] FREEMAN W T, JONES T R, PASZTOR E C. Example-based super-resolution[J]. IEEE Computer Graphics and Applications, 2002 (2): 56-65.

- [20] 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 (CVPR). [S.l.]: IEEE, 2004: 275-282.
- [21] YANG J, WRIGHT J, HUANG T, et al. Image super-resolution as sparse representation of raw image patches[C]//Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2008: 1-8.
- [22] YANG J, WRIGHT J, HUANG T S, et al. Image super-resolution via sparse representation[J]. *IEEE Transactions on Image Processing*, 2010, 19(11): 2861-2873.
- [23] YANG J, WANG Z, LIN Z, et al. Coupled dictionary training for image super-resolution[J]. *IEEE Transactions on Image Processing*, 2012, 21(8): 3467-3478.
- [24] 孙玉宝, 韦志辉, 肖亮, 等. 多形态稀疏性正则化的图像超分辨率算法[J]. *电子学报*, 2010, 38(12): 2898-2903.  
SUN Yubao, WEI Zhihui, XIAO Liang, et al. Multimorphology sparsity regularized image super-resolution[J]. *Acta Electronica Sinica*, 2010, 38(12): 2898-2903.
- [25] GREGOR K, LECUN Y. Learning fast approximations of sparse coding[C]//Proceedings of the 27th International Conference on International Conference on Machine Learning. Haifa, Israel: [s.n.], 2010: 399-406.
- [26] TIMOFTE R, DE SMET V, VAN GOOL L. Anchored neighborhood regression for fast example-based super-resolution [C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2013: 1920-1927.
- [27] TIMOFTE R, DE SMET V, VAN GOOL L. A+: Adjusted anchored neighborhood regression for fast super-resolution [C]//Proceedings of Asian Conference on Computer Vision. Cham: Springer, 2014: 111-126.
- [28] SCHULTER S, LEISTNER C, BISCHOF H. Fast and accurate image upscaling with super-resolution forests[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2015: 3791-3799.
- [29] DONG C, LOY C C, HE K, et al. Learning a deep convolutional network for image super-resolution[C]//Proceedings of European Conference on Computer Vision. Cham: Springer, 2014: 184-199.
- [30] DONG C, LOY C C, HE K, et al. Image super-resolution using deep convolutional networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 38(2): 295-307.
- [31] KIM J, LEE K J, LEE M K. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2016: 1646-1654.
- [32] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. [S.l.]: IEEE, 2017: 136-144.
- [33] 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 (ECCV). Munich, Germany: [s.n.], 2018: 286-301.
- [34] KIM J, KWON LEE J, MU LEE K. Deeply-recursive convolutional network for image super-resolution[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 1637-1645.
- [35] TAI Y, YANG J, LIU X. Image super-resolution via deep recursive residual network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 3147-3155.
- [36] TAI Y, YANG J, LIU X, et al. Memnet: A persistent memory network for image restoration[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 4539-4547.
- [37] 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. Salt Lake City, USA: IEEE, 2018: 2472-2481.
- [38] HARIS M, SHAKHAROVICH G, UKITA N. Deep back-projection networks for super-resolution[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 1664-1673.
- [39] CHOI J S, KIM M. A deep convolutional neural network with selection units for super-resolution[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, HI, USA: IEEE, 2017: 154-160.
- [40] DAI T, CAI J, ZHANG Y, et al. Second-order attention network for single image super-resolution[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2019: 11065-11074.
- [41] JOHNSON J, ALAHI A, FEI-FEI L. Perceptual losses for real-time style transfer and super-resolution[C]//Proceedings of European Conference on Computer Vision. Cham: Springer, 2016: 694-711.
- [42] 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. Honolulu, HI, USA: IEEE, 2017: 4681-4690.
- [43] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]// Proceedings of Advances in Neural Information Processing Systems. Montreal, QC, Canada: [s.n.], 2014: 2672-2680.
- [44] SAJJADI M S M, SCHOLKOPF B, HIRSCH M. Enhancenet: Single image super-resolution through automated texture synthesis[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2017: 4491-4500.
- [45] WANG X, YU K, WU S, et al. Esrgan: Enhanced super-resolution generative adversarial networks[C]//Proceedings of the European Conference on Computer Vision. Munich: [s.n.], 2018.
- [46] ZHANG K, ZUO W, ZHANG L. Learning a single convolutional super-resolution network for multiple degradations[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 3262-3271.
- [47] BULAT A, YANG J, TZIMIROPOULOS G. To learn image super-resolution, use a GAN to learn how to do image degradation first[C]//Proceedings of the European Conference on Computer Vision. [S.l.]: IEEE, 2018: 185-200.
- [48] GU J, LU H, ZUO W, et al. Blind super-resolution with iterative kernel correction[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2019: 1604-1613.
- [49] ZHOU R, SUSSTRUNK S. Kernel Modeling Super-Resolution on Real Low-Resolution Images[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2019: 2433-2443.
- [50] PAN J, SUN D, PFISTER H, et al. Blind image deblurring using dark channel prior[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 1628-1636.
- [51] CHEN C, XIONG Z, TIAN X, et al. Camera lens super-resolution[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2019.
- [52] ZHANG X, CHEN Q, NG R, et al. Zoom to learn, learn to zoom[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2019: 3762-3770.
- [53] CAI J, ZENG H, YONG H, et al. Toward real-world single image super-resolution: A new benchmark and a new model[J/OL].(2019-04-01)[2019-11-30]. <https://arxiv.org/abs/1904.00523?context=cs.CV>.
- [54] KEYS R. Bicubic interpolation[J]. IEEE Transactions on Acoustics Speech and Signal Processing, 1981, 29: 1153-1160.
- [55] THURNHOFER S, MITRA S K. Edge-enhanced image zooming[J]. Optical Engineering, 1996: 35.
- [56] FEKRI F, MERSEREAU R M, SCHAFER R W. A generalized interpolative VQ method for jointly optimal quantization and interpolation of images[C]// Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing. [S.l.]: Institute of Electrical Engineers Inc, 1998, 5: V-2657.
- [57] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding[J]. Science, 2000, 290(5500): 2323-2326.
- [58] TENENBAUM J B, DE SILVA V, LANGFORD J C. A global geometric framework for nonlinear dimensionality reduction [J]. Science, 2000, 290(5500): 2319-2323.
- [59] CANDÈS E J, ROMBERG J, TAO T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information[J]. IEEE Transactions on Information Theory, 2006, 52(2): 489-509.
- [60] CANDÈS E J. Compressive sampling[C]//Proceedings of the International Congress of Mathematicians. Madrid, Spain: [s.n.], 2006, 3: 1433-1452.
- [61] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]// Proceedings of Advances in Neural Information Processing Systems. [S.l.]: ACM, 2012: 1097-1105.
- [62] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2015: 1-9.
- [63] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [64] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[C]// Proceedings of Advances in Neural Information Processing Systems. [S.l.]: IEEE, 2015: 91-99.

- [65] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector[C]//Proceedings of European Conference on Computer Vision. Cham: Springer, 2016: 21-37.
- [66] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2015: 3431-3440.
- [67] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]//Proceedings of International Conference on Medical Image Computing and Computer-assisted Intervention. Cham: Springer, 2015: 234-241.
- [68] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[J/OL]. (2016-01-07)[2019-11-30]. <https://arxiv.org/abs/1511.06434>.
- [69] KARRAS T, LAINE S, AILA T. A style-based generator architecture for generative adversarial networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2019: 4401-4410.
- [70] REN W, LIU S, ZHANG H, et al. Single image dehazing via multi-scale convolutional neural networks[C]//European Conference on Computer Vision. Cham: Springer, 2016: 154-169.
- [71] LI B, PENG X, WANG Z, et al. AOD-net: All-in-one dehazing network[C]//Proceedings of the IEEE International Conference on Computer Vision. Honolulu, HI, USA: IEEE, 2017: 4770-4778.
- [72] XIE J, XU L, CHEN E. Image denoising and inpainting with deep neural networks[C]// Proceedings of Advances in Neural Information Processing Systems. Venice, Italy: IEEE, 2012: 341-349.
- [73] YEH R A, CHEN C, LIM T Y, et al. Semantic image inpainting with deep generative models[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 5485-5493.
- [74] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 4700-4708.
- [75] GATYS L A, ECKER A S, BETHGE M. Image style transfer using convolutional neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 2414-2423.
- [76] JOLICOEUR-MARTINEAU A. The relativistic discriminator: A key element missing from standard GAN[J/OL].(2018-09-10)[2019-11-30]. <https://arxiv.org/abs/1807.00734?context=stat.ML>.
- [77] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks[C]//Proceedings of European Conference on Computer Vision. Cham: Springer, 2014: 818-833.
- [78] 赵丽玲, 孙权森, 张泽林. 基于深度学习特征字典的单帧图像超分辨率重建[J]. 数据采集与处理, 2018, 33(4): 740-750.  
ZHAO Liling, SUN Quansen, ZHANG Zelin. Single image super resolution reconstruction based on deep features dictionary [J]. *Journal of Data Acquisition and Processing*, 2018, 33(4): 740-750.
- [79] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J/OL]. (2015-04-10)[2019-11-30]. <https://arxiv.org/abs/1409.1556>.
- [80] EFRAT N, GLASNER D, APARTSIN A, et al. Accurate blur models vs. image priors in single image super-resolution[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2013: 2832-2839.
- [81] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding[C]// Proceedings of British Machine Vision Conference (BMVC). Guildford, Surrey, United Kingdom: [s.n.], 2012.
- [82] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations[C]//Proceedings of International Conference on Curves and Surfaces. Berlin, Heidelberg: Springer, 2010: 711-730.
- [83] 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[C]//Proceedings of the IEEE International Conference on Computer Vision. Vancouver, BC, Canada, Canada: IEEE, 2001. DOI: 10.1109/ICCV.2001.937655.
- [84] HUANG J B, SINGH A, AHUJA N. Single image super-resolution from transformed self-exemplars[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2015: 5197-5206.
- [85] FUJIMOTO A, OGAWA T, YAMAMOTO K, et al. Manga109 dataset and creation of metadata[C]//Proceedings of the 1st International Workshop on coMics ANalysis, Processing and Understanding. [S.l.]: ACM, 2016: 2.

- [86] DENG J, DONG W, SOCHER R, et al. Imagenet: A large-scale hierarchical image database[C]// Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2009: 248-255.
- [87] 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 Workshops. Honolulu, HI, USA: IEEE, 2017: 126-135.
- [88] TIMOFTE R, AGUSTSSON E, VAN GOOL L, et al. Ntire 2017 challenge on single image super-resolution: Methods and results[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, HI, USA: IEEE, 2017: 114-125.
- [89] WANG X, YU K, DONG C, et al. Recovering realistic texture in image super-resolution by deep spatial feature transform [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 606-615.
- [90] WANG Z, 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.
- [91] 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. Salt Lake City, USA: IEEE, 2018: 586-595.
- [92] BLAU Y, MECHREZ R, TIMOFTE R, et al. The 2018 PIRM challenge on perceptual image super-resolution[C]// Proceedings of the European Conference on Computer Vision. Munich, Germany: [s.n.], 2018.
- [93] MA C, YANG C Y, YANG X, et al. Learning a no-reference quality metric for single-image super-resolution[J]. Computer Vision and Image Understanding, 2017, 158: 1-16.
- [94] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a “completely blind” image quality analyzer[J]. IEEE Signal Processing Letters, 2012, 20(3): 209-212.
- [95] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network[C]// Proceedings of European Conference on Computer Vision. Cham: Springer, 2016: 391-407.
- [96] LIU D, WANG Z, WEN B, et al. Robust single image super-resolution via deep networks with sparse prior[J]. IEEE Transactions on Image Processing, 2016, 25(7): 3194-3207.
- [97] LAI W S, HUANG J B, AHUJA N, et al. Deep laplacian pyramid networks for fast and accurate super-resolution[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 624-632.
- [98] 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. Salt Lake City, USA: IEEE, 2018: 3118-3126.
- [99] 付争方, 朱虹, 余顺园, 等. 基于灰度级映射函数建模的多曝光高动态图像重建[J]. 数据采集与处理, 2019, 34(3): 472-490. FU Zhengfang, ZHU Hong, YU Shunyuan, et al. Multi-exposure HDR image reconstruction based on gray scale mapping function modeling[J]. Journal of Data Acquisition and Processing, 2019, 34(3): 472-490.
- [100] 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. Las Vegas, NV, USA: IEEE, 2016: 1874-1883.
- [101] CABALLERO J, LEDIG C, AITKEN A, et al. Real-time video super-resolution with spatio-temporal networks and motion compensation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 4778-4787.
- [102] TAO X, GAO H, LIAO R, et al. Detail-revealing deep video super-resolution[C]//Proceedings of the IEEE International Conference on Computer Vision. Honolulu, HI, USA: IEEE, 2017: 4472-4480.
- [103] JO Y, WUG OH S, KANG J, et al. Deep video super-resolution network using dynamic upsampling filters without explicit motion compensation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 3224-3232.
- [104] SAJJADI M S M, VEMULAPALLI R, BROWN M. Frame-recurrent video super-resolution[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 6626-6634.

- [105] YI P, WANG Z, JIANG K, et al. Progressive fusion video super-resolution network via exploiting non-local spatio-temporal correlations[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2019: 3106-3115.
- [106] WANG X, GIRSHICK R, GUPTA A, et al. Non-local neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 7794-7803.
- [107] 孔繁镨, 朱成, 徐诚, 等. 空间相关性约束联合子空间追踪的高光谱图像稀疏解混[J]. 南京航空航天大学学报, 2019, 51(5): 577-585.  
KONG Fanqiang, ZHU Cheng, XU Cheng, et al. Spatial correlation constrained simultaneous subspace pursuit for sparse unmixing of hyperspectral imagery[J]. Journal of Nanjing University of Aeronautics and Astronautics, 2019, 51(5): 577-585.
- [108] KONG Fanqiang, GUO Wenjun, SHEN Qiu, et al. Recursive dictionary-based simultaneous orthogonal matching pursuit for sparse unmixing of hyperspectral data[J]. Transactions of Nanjing University of Aeronautics & Astronautics, 2017, 34(4): 456-466.
- [109] PAN Z, YU J, HUANG H, et al. Super-resolution based on compressive sensing and structural self-similarity for remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51(9): 4864-4876.
- [110] CHAVEZ-ROMAN H, PONOMARYOV V. Super resolution image generation using wavelet domain interpolation with edge extraction via a sparse representation[J]. IEEE Geoscience and Remote Sensing Letters, 2014, 11(10): 1777-1781.
- [111] LEI S, SHI Z, ZOU Z. Super-resolution for remote sensing images via local-global combined network[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(8): 1243-1247.
- [112] WANG T, SUN W, QI H, et al. Aerial image super resolution via wavelet multiscale convolutional neural networks[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(5): 769-773.
- [113] PAN Z, MA W, GUO J, et al. Super-resolution of single remote sensing image based on residual dense backprojection networks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(10): 7918-7933.
- [114] TAN W, YAN B, BARE B. Feature super-resolution: Make machine see more clearly[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 3994-4002.
- [115] BAI Y, ZHANG Y, DING M, et al. Finding tiny faces in the wild with generative adversarial network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 21-30.
- [116] LI J, LIANG X, WEI Y, et al. Perceptual generative adversarial networks for small object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 1222-1230.
- [117] NOH J, BAE W, LEE W, et al. Better to Follow, Follow to Be Better: Towards Precise Supervision of Feature Super-Resolution for Small Object Detection[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2019: 9725-9734.

#### 作者简介:



史振威(1977-),男,教授,博士生导师,研究方向:图像处理、模式识别、深度学习、遥感图像处理等, E-mail: shizhenwei@buaa.edu.cn。



雷森(1992-),男,博士研究生,研究方向:深度学习、超分辨重建、遥感图像质量提升等。

(编辑:张黄群)