

单幅图像超分辨重建的深度学习方法综述

张 德, 林青宇, 郭茂祖

北京建筑大学 电气与信息工程学院&建筑大数据智能处理方法研究北京市重点实验室, 北京 100044

摘 要: 图像超分辨重建(Super-Resolution, SR)是指利用信号处理和机器学习等方法, 从单幅或者多幅低分辨率图像(Low Resolution, LR)中重建对应的高分辨率图像(High Resolution, HR)的技术。由于多幅LR图像之间亚像素位移的不可预知性, 单幅图像超分辨重建(Single Image Super-Resolution, SISR)逐渐成为超分辨研究的主要方向。近年来, 深度学习得到迅速发展, 并广泛应用到图像处理领域。因此, 针对单幅图像超分辨重建所使用的深度学习相关算法和网络模型进行系统的总结。介绍图像超分辨问题的设置和评价指标; 讨论和比较单幅图像超分辨重建的深度学习算法, 主要从网络结构设计、损失函数和上采样方式三方面进行论述; 介绍常用的标准数据集, 并选用基于不同网络模型的几种典型算法进行实验对比分析; 展望图像超分辨技术未来的研究趋势和发展方向。

关键词: 图像超分辨; 深度学习; 卷积神经网络(CNN); 图像处理

文献标志码: A **中图分类号:** TP391 **doi:** 10.3778/j.issn.1002-8331.2106-0055

Review of Single Image Super-Resolution Based on Deep Learning

ZHANG De, LIN Qingyu, GUO Maozu

School of Electrical and Information Engineering & Beijing Key Laboratory of Intelligent Processing for Building Big Data, Beijing University of Civil Engineering and Architecture, Beijing 100044, China

Abstract: Image Super-Resolution (SR) is a technique that uses signal processing and machine learning to reconstruct High Resolution (HR) images from single or multiple Low Resolution (LR) images. Due to the unpredictability of sub-pixel displacement between multiple LR images, Single Image Super-Resolution (SISR) has gradually become the main direction of super-resolution research. In recent years, deep learning method has been developed rapidly and widely used in the field of image processing. Therefore, the deep learning algorithms and network models used in SISR are summarized systematically. Firstly, the setting and evaluation index of SR are introduced. Then, the deep learning algorithms of SISR reconstruction are discussed and compared, mainly from three aspects of network structure design, loss function and upsampling method. Next, the commonly used standard data sets are introduced, and several typical algorithms based on different network models are selected for experimental comparative analysis. Finally, the future research trend and development direction of SR technology are prospected.

Key words: image super-resolution; deep learning; Convolutional Neural Network (CNN); image processing

图像超分辨重建(Super-Resolution, SR)是计算机视觉和图像处理领域的重要研究方向之一。在不改善成像设备硬件条件的前提下, 通过信号处理和软件方法等提高图像分辨率, 即为SR技术。SR不仅有重要的学术研究价值, 并且在很多场合有实际应用价值, 例如医学影像分析^[1-2]、卫星遥感成像^[3]和视频安全监控^[4]等方面。

单幅图像超分辨重建(Single Image Super-Resolution, SISR)仅利用一幅低分辨率(Low Resolution, LR)图像

为输入, 重建具有丰富图像细节和清晰纹理的高分辨率(High Resolution, HR)图像, 实用价值很高。早在20世纪60年代, SISR技术由Harris和Goodman提出, 并通过图像复原的方法来实现^[5-6]。然后, Brown和Wadaka等使用长椭圆波函数方法^[7]和叠加正弦模板方法^[8]等其他复原技术进行超分辨重建, 但是实际应用效果并不好。

因此, 研究学者们陆续提出各种SISR技术, 主要包括基于插值的方法, 如双线性插值(Bilinear)、双立方插

基金项目: 国家自然科学基金(61871020); 北京市教育委员会科技计划重点项目(KZ201810016019)。

作者简介: 张德(1979—), 男, 副教授, CCF会员, 研究方向为模式识别与图像处理, E-mail: zhangde2000@163.com; 林青宇(1994—), 男, 硕士研究生, CCF会员, 研究方向为计算机视觉; 郭茂祖(1966—), 男, 教授, CCF高级会员, 研究方向为机器学习和人工智能。

收稿日期: 2021-06-03 **修回日期:** 2021-08-18 **文章编号:** 1002-8331(2021)22-0028-14

值(Bicubic)^[9]、曲率插值^[10]和自适应图像插值^[11-13]等;基于重建的方法,如迭代反向投影法^[14-16]、凸集投影法^[17-19]和最大后验概率法^[20-22]等;基于样例学习的方法,如邻域嵌入法^[23-25]、稀疏表示法^[26-28]和回归方法^[29-31]等。但由于超分辨重建是一个难解的欠定问题,传统方法难以重建图像高频细节信息^[32-33]。

随着机器学习在计算机视觉领域的发展,深度学习被广泛应用到图像超分辨重建任务中,并取得了很好的重建效果。如今,基于深度学习的SISR方法已逐渐成为主流,图1给出了SISR深度学习研究方法的发展时间轴线。从开始基于卷积神经网络(Convolutional Neural Network, CNN)的算法,到后来使用不断提出的新型网络结构,包括深度残差网络(Deep Residual Network, ResNet)、递归神经网络(Recursive Neural Network, RNN)、密集连接网络(Dense Connected Convolutional Network, DenseNet)和生成对抗网络(Generative Adversarial Network, GAN)等,都表现出不错的性能,使重建生成的HR图像更加逼真。

近年来,已有研究学者对图像超分辨重建算法进行了总结^[46-49],包括基于传统特征的方法和基于深度学习的方法。文献[46-47]针对这两大类方法都介绍得比较详细。其中,史振威等^[46]总结了近20年来的主流算法,还介绍了在视频超分辨、遥感图像超分辨和复杂条件下超分辨算法的设计和应用。唐艳秋等^[47]首先对传统方法进行了完整的分类介绍,然后对深度学习方法从不同的网络结构模型方面进行了分析和阐述。刘颖等^[48]将超分辨技术根据输入输出分为多图像超分辨、视频超分辨和单幅图像超分辨,然后重点介绍了SISR算法,对深度学习方法从有监督学习和无监督学习两方面进行分类叙述,并选取了几种典型的深度学习算法进行实验和比较。南方哲等^[49]主要进行了基于深度学习的SISR算法综述,把超分辨深度学习模型分为基于插值预处理、基于原始图像处理、基于分层特征和基于高频细节四类进行介绍,然后选取了几种来自不同类别的算法进行实验和比较。

本文针对SISR的深度学习方法进行全面的分析和总结,重点阐述并总结了近年来主流的基于深度学习的图像超分辨重建方法,并以图形和表格的方式对典型的基础网络结构进行展示和比较。同时选取了多种

不同的经典深度学习算法进行实验复现和比较,并考虑时间的跨度,包括了近两年内的算法。

本文首先介绍图像超分辨相关基础内容,包括问题设置概述和质量评价指标。然后,详细阐述为解决单幅图像超分辨问题和提高重建质量而采用的深度学习方法,包括网络模型设计、损失函数和上采样方式三方面,这也是SISR深度学习技术必需的三个组成部分。而后,选取来自不同网络模型的主要算法在标准数据集上进行实验,给出不同算法之间更直观的对比并进行讨论。最后,对基于深度学习的SISR方法进行总结和展望。

1 图像超分辨问题概述

图像超分辨的目的是从LR图像中恢复出相应的HR图像,而LR图像通常是由HR图像退化得到。用 I_x 表示LR图像, I_y 表示HR图像,则退化过程如式(1)所示:

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

其中, D 为退化函数, δ 表示退化过程中的相关参数,如噪声类型和缩放因子等。现实中,退化过程信息 D 和 δ 是未知的,仅有已知信息为低分辨率图像 I_x 。图形超分辨问题的目的为消除退化效果,恢复出高分辨率甚至真实场景图像,即

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

其中, θ 为超分辨模型 F 中的参数, \hat{I}_y 表示重建得到的HR图像。

现实中退化过程未知且复杂,影响因素众多,难以确定实际问题中准确的图像退化过程。因此大多数研究都是基于单个下采样作为退化模型,常用下采样操作为具有抗锯齿的双三次插值。也有一些研究将退化模型设置为如下形式:

$$I_x = (I_y \otimes k) \downarrow n_s \quad (3)$$

其中, \downarrow 表示缩放因子为 s 的下采样运算, \otimes 是HR图像 I_y 和卷积核算子 k 之间的卷积操作,变量 n_s 表示具有标准差的加性高斯白噪声。与单一采用下采样退化过程相比,式(3)的退化过程更接近于真实场景,并已被证实对SR更有效^[50]。

至此,SR问题可以理解为:

$$\hat{\theta} = \arg_{\theta} \min L(\hat{I}_y, I_y) + \lambda \varphi(\theta) \quad (4)$$

其中, $L(\hat{I}_y, I_y)$ 为高分辨率图像 \hat{I}_y 和真实图像 I_y 的损失函

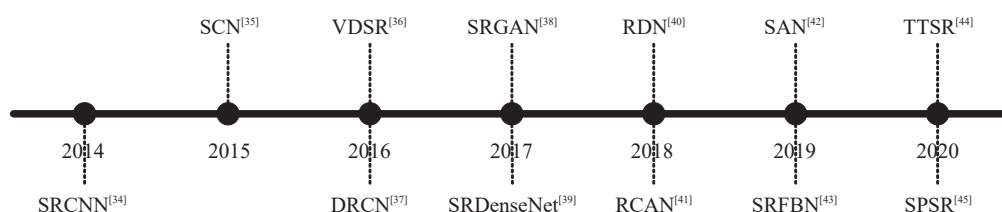


图1 SISR深度学习主要方法的发展时间轴线

Fig.1 Timeline of deep learning based SISR progress

数, φ (θ 为正则化项, λ 为权衡参数。损失函数为测量重建图像的误差, 将在第3章中详细介绍。

2 图像超分辨率质量评价

通过超分辨率技术生成的高分辨率图像, 与真实高分辨率图像的逼真度越高, 则认为图像超分辨率算法越好。一般来说, 主要通过两种客观的量化指标进行评价, 包括峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 和结构相似度 (Structural Similarity, SSIM)。

PSNR 是一种最常用的图像重构质量评价方式, 其主要由 MSE 决定, PSNR 定义如下:

$$PSNR = 10 \lg \frac{a_{\max}^2}{MSE} \quad (5)$$

其中, a_{\max} 表示参考图像中的最大像素值, MSE 如式(6)所示:

$$MSE = \frac{1}{N} \sum_{i=1}^N (a_i - \hat{a}_i)^2 \quad (6)$$

其中, a_i, \hat{a}_i 分别表示真实参考图像和重构图像第 i 个像素。由式(5)可以看出, 最小化 MSE 损失函数相当于最大化 PSNR 指标。但 MSE 度量的是同位置的像素级差异, 缺乏对图像整体的可视化感知。因此在很多情况下, 重建生成的图像虽然具有较高的 PSNR 评价指标, 但却不能给人以较好的主观感受。不过, 因为无法对图像感知质量进行准确的定量分析, PSNR 仍然是图像超分辨率领域最常用的评价方式。

考虑到人类的视觉系统是从所见场景中提取结构信息进行观察, Wang 等^[51]提出使用结构相似度 (SSIM) 评价生成图像的质量。根据亮度 (luminance)、对比度 (contrast) 和结构 (structure) 三方面相互独立的比较, SSIM 可用来衡量图像之间的结构相似度。

$$SSIM(I_y, \hat{I}_y) = \frac{(2\mu_{I_y}\mu_{\hat{I}_y} + C_1)(\sigma_{I_y\hat{I}_y} + C_2)}{(\mu_{I_y}^2 + \mu_{\hat{I}_y}^2 + C_1)(\sigma_{I_y}^2 + \sigma_{\hat{I}_y}^2 + C_2)} \quad (7)$$

其中, μ_{I_y} 和 $\mu_{\hat{I}_y}$ 分别表示 I_y 和 \hat{I}_y 的均值, σ_{I_y} 和 $\sigma_{\hat{I}_y}$ 为 I_y 和 \hat{I}_y 的方差, $\sigma_{I_y\hat{I}_y}$ 为 I_y 和 \hat{I}_y 的协方差, C_1 和 C_2 为常数。因为 SSIM 可以较好评估图像的感知质量^[52], 所以在图像超分辨率领域得到了广泛的使用。

另外, 有时也经常用到主观评价指标平均意见分数 (Mean Opinion Score, MOS), 给出图像的主观感受质量。由观察人员给重建生成的图像打分, 分数范围在 1~5 之间, 5 分为最高分。计算所有分数的平均分, 即为 MOS 指标值。

3 基于深度学习的单幅图像超分辨率重建

近年来深度学习在单幅图像超分辨率重建 (SISR) 领域得到了广泛的应用, 并取得了很好的重建效果。各种基于深度学习的 SISR 方法层出不穷, 本文从网络结构

设计、损失函数和上采样框架三方面进行阐述。表 1 对经典网络模型从以上三方面进行总结。

3.1 网络结构设计

网络结构设计是深度学习方法最重要的组成部分, 并且结构的多样性很强, 难以准确地划分类别。这里, 从网络结构设计的基本原理和策略出发, 大概分为线性网络、残差网络、递归网络、密集连接、注意力机制和生成对抗网络等几个方面进行介绍和比较。其中, 把早期应用于 SISR 的 CNN 模型归类为线性网络, 模型较为简单, 但效果也优于传统方法。随着深度学习技术的不断发展, 网络结构的深度和宽度不断增加, 感受野变大, 可以带来更强的特征学习能力, 明显提升超分辨率重建效果。不过, 更深的网络结构增加了模型复杂度, 导致网络不易收敛、性能退化和过度拟合等现象。残差学习网络和递归网络结构的出现可以帮助解决这些问题。而后出现的密集连接可以进一步减轻网络梯度消失的问题。结合注意力机制的网络模型能够进一步提高图像重建质量评价指标, 而生成对抗网络则有助于提升重建图像的视觉感知质量。

3.1.1 线性网络

SRCNN (Super-Resolution CNN) 是第一个只使用卷积神经网络来进行图像超分辨率重建的成功尝试^[39], 是深度学习在图像超分辨率重建领域的开山之作, 并激发了后续在这个方向上的研究。SRCNN 结构简单明了, 它只由卷积层组成, 除最后一层外, 每一层后面都有一个 ReLU 作为激活函数。第一层为特征提取层, 从输入图像中提取特征图块。第二层为非线性映射层, 它将特征映射转换为高维特征向量。最后一层重建特征图片来作为最终的高分辨率图像输出。

而后 Dong 等^[53]进一步提出 FSRCNN (Fast Super-Resolution CNN), 其是对 SRCNN 的改进 (对比如图 2 所示)。FSRCNN 虽然是 4 个卷积层和 1 个反卷积层的简单网络, 但是其处理速度提升明显, 图片质量也有所提升。其改进主要体现在: 一是最后使用了一个反卷积层放大尺寸, 从而不需要对图像进行预处理。二是改变特征维数, 使用更小的卷积核和更多的映射层, 在映射前减少输入特征的维数, 从而提高计算效率。

ESPCN (Efficient Sub-Pixel Convolutional Neural network) 也是一种快速高效的超分辨率方法^[54]。ESPCN 主要使用亚像素卷积 (Sub-Pixel Convolution), 作为重建生成的 HR 图像中的一个子块。这个变换实际上没有卷积操作, 图像从 LR 到 HR 放大过程的插值函数被隐含在前面的卷积层中, 只是在最后一层对图像大小做变换, 因此提高了算法的效率。

早期应用在图像超分辨率领域的线性网络虽然模型相对简单, 但重建效果优于传统方法, 并起到开创性目的。

表1 SISR深度学习主要方法简述
Table 1 Description of some typical deep learning based SISR methods

网络名称	发表时间	网络模型	损失函数	上采样方法	上采样框架	适用场景	改进点
SRCNN ^[34]	2014,ECCV	线性网络	L2	双三次插值	预上采样	实景图像	首次提出将深度学习引入超分辨率领域,提高重建质量
SCN ^[35]	2015,ICCV	线性网络	L2	双三次插值	预上采样	实景图像	在深度卷积网络中加入稀疏表示先验信息
FSRCNN ^[53]	2016,ECCV	线性网络	L2	转置卷积	后上采样	实景图像	较SRCNN提升算法处理速度
ESPCN ^[54]	2016,CVPR	线性网络	L2	亚像素卷积	预上采样	实景图像	较SRCNN提高算法效率
VDSR ^[36]	2016,CVPR	残差网络	MSE	双三次插值	后上采样	实景图像	运用残差学习加深网络层数,同时可处理多尺度SR问题
LapSRN ^[55]	2017,CVPR	残差网络	Char损失函数	亚像素卷积	逐步上采样	实景图像	通过逐步上采样、逐级残差和自身损失函数监督,提升算法速度以及重建效果
EDSR ^[56]	2017,CVPRW	残差网络	L1	亚像素卷积	后上采样	实景图像	去除BN加快运行速度,加深网络深度
MSRN ^[57]	2018,ECCV	残差网络	L1	亚像素卷积	后上采样	实景图像	运用多尺度残差网络提升算法效率
DRCN ^[37]	2016,CVPR	残差、递归网络	L2	双三次插值	预上采样	实景图像	运用递归网络增大网络感受野,从而提升网络性能
DRRN ^[58]	2017,CVPR	残差、递归网络	L2	双三次插值	预上采样	实景图像	运用局部多尺度和全局残差学习,减少参数并提升重建性能
SRFBN ^[43]	2019,CVPR	残差、递归、稠密网络	L1	转置卷积	后上采样	实景图像	引入反馈机制的循环神经网络,从而提升重建效果
SRDenseNet ^[39]	2017,ICCV	密集网络	L2	亚像素卷积	后上采样	实景图像	运用密集连接加强特征传播,减少参数数量,减轻梯度消失
RDN ^[40]	2018,CVPR	密集、残差网络	L1	亚像素卷积	后上采样	实景图像	将残差、密集学习结合,客观指标和主观效果有所提升
SeINet ^[59]	2017,CVPRW	残差网络、注意力机制	L2	双三次插值	后上采样	实景图像	使用选择单元允许特定值向下传递,提升重建效果
RCAN ^[41]	2018,ECCV	残差网络、注意力机制	L1	亚像素卷积	后上采样	实景图像	将通道注意力机制引入网络,捕捉更重要的通道信息,从而提升重建效果
SAN ^[42]	2019,CVPR	残差网络、注意力机制	L1	亚像素卷积	后上采样	实景图像	使用二阶通道注意力机制,更好的长距离空间内容信息
TTSR ^[44]	2020,CVPR	残差网络、注意力机制	L1、对抗、感知损失	亚像素卷积	后上采样	实景图像	使用硬注意力机制进行纹理迁移,使用软注意力机制进行纹理合成
SRGAN ^[38]	2017,CVPR	生成对抗网络	对抗、感知损失	亚像素卷积	后上采样	实景图像	最早使用GAN网络进行图像超分辨,在主观视觉效果上提升明显
ESRGAN ^[60]	2018,ECCVW	生成对抗网络	L1、对抗、感知损失	亚像素卷积	后上采样	实景图像	相较于SRGAN,生成器去掉BN、判别器对损失函数做出改进,提升重建效果
SPSR ^[45]	2020,CVPR	生成对抗网络	L1、对抗、感知、梯度损失	亚像素卷积	后上采样	实景图像	基于GAN网络的方法,考虑保持图像内部的几何结构

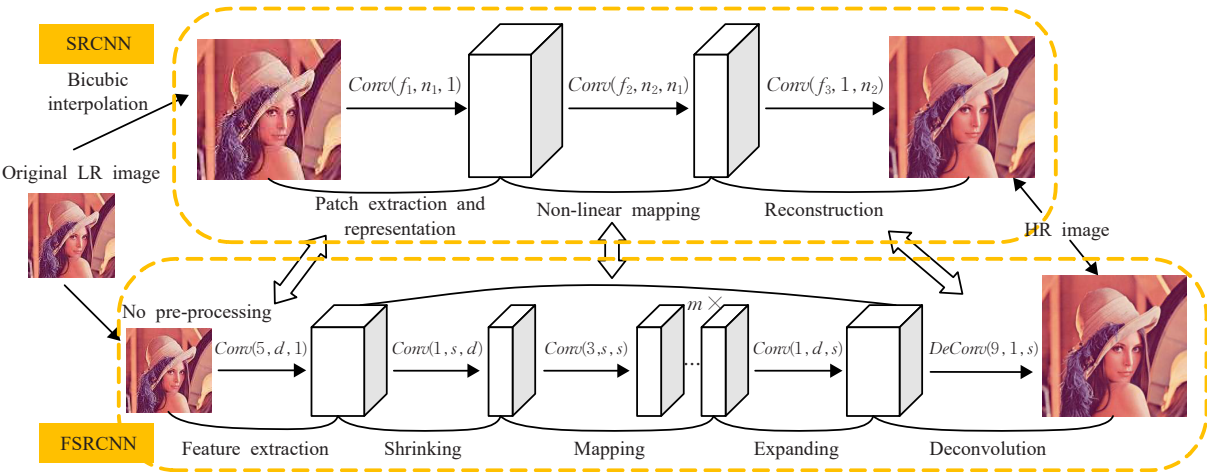


图2 SRCNN和FSRCNN网络结构比较
Fig.2 Comparison of SRCNN and FSRCNN structure

3.1.2 残差网络

与线性网络的线性模型相比,残差网络使用残差学习,在网络设计中使用了跳过连接方式进行残差单元中的直接映射^[61],有效避免了梯度消失,同时可以加快训练速度。因此残差学习的应用可以使更深的网络得到更好的训练,从而使得图像超分辨率网络模型设计变得非常深。同时残差网络能利用更多的上下文信息,学习能力和层次表征能力变得更强。因此,残差网络在图像超分辨率中应用广泛。

Kim等^[36]提出的VDSR(Deep Super-Resolution convolutional network)网络模型有20个卷积层,相较于SRCNN等线性网络层数更深。Kim等通过实验证明层数越深,图像重建效果越好,但对收敛速度也产生了很大影响。为了解决收敛速度问题,VDSR网络模型采用全局残差学习,即高分辨率图像与低分辨率图像进行残差学习,同时提高学习速率,使得训练过程更加稳定。因此残差网络可以使网络模型较深的网络得到更好的训练,提升重建效果。Lim等^[56]提出了用于图像超分辨率的改进残差网络模型EDSR(Enhanced Deep Super-Resolution network),利用改进残差块,搭建更深的网络模型。EDSR主要将原始残差块中的批归一化层(Batch Normalization, BN)去除,节省显存空间,使更深的网络有更好的可训练性,可以拥有69个卷积层,重建效果明显提升。

同时,许多多尺度残差网络也被提出。Li等^[57]提出了多尺度残差网络模型(Multi-Scale Residual Network, MSRN)。MSRN使用多尺度残差块(Multi-Scale Residual Block, MSRB)来获取不同尺度图像特征,联合 3×3 和 5×5 卷积核帮助网络获取多尺度的信息,并采用局部残差学习法可以让神经网络更高效。将每个MSRB的输出组合起来进行全局特征融合,充分利用图像中的特征信息。实验证明尽管该模型在评价指标PSNR和SSIM方面比EDSR模型稍差,但是MSRN模型参数数量远低于EDSR,训练时间短,运行效率高。图3给出了EDSR和MSRN模型的改进残差块结构,并与He等^[61]提出的原始结构进行了比较。另外,应自炉等^[62]使用多种程度的卷积核提取LR层的特征信息,将每个卷积核下的残差块局部融合后再进行全局融合。这使得每层信息不会丢失,从而使得网络梯度一直保持存在的状态,不会发生梯度消失现象,同时重建效果较SRCNN等线性网络得到提升。不同于文献[62],李现国等^[63]提出的多尺度残差网络在残差块中用两个 3×3 卷积核代替一个 5×5 的卷积核来提取特征,减少了模型参数,增加了网络的非线性特性,并且使得图像特征信息在每个维度下得到增加,从而获得更丰富的图像信息,达到更好的重建效果。

3.1.3 递归网络

随着网络模型深度的增加,也会引入更多的参数。

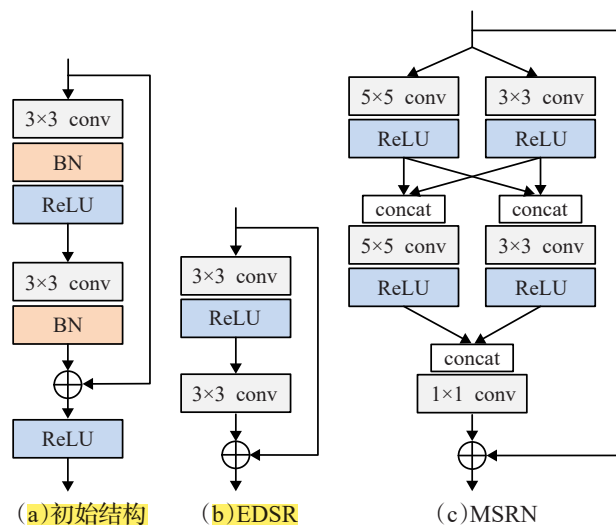


图3 EDSR、MSRN和初始残差网络结构比较

Fig.3 Comparison of EDSR,MSRN and original residual network structure

为了避免发生过拟合,通常会扩大训练集,但过大的网络模型不易存储。因此将递归学习引入到图像超分辨率领域中。递归网络为递归链接的卷积层或单元的网络结构,从而在控制模型参数的情况下,提高图像重建效果。

Kim等^[37]提出的DRCN(Deep Recursive Convolutional Network)多次应用叠加相同的卷积层。这种技术的一个优点是在增大网络感受野的同时,对于更多的递归,可以有效地重复利用权重参数,并能够优化随机梯度下降算法,缩短训练时间。用递归和跳跃连接的方式减低模型的训练难度,从而达到更好的训练效果。在此基础上,Tai等^[58]提出了DRRN(Deep Recursive Residual Network)网络模型,和DRCN仅使用全局残差学习不同,DRRN加入了多路径方式的局部残差学习,并在残差块中共享使用多个权重层。DRRN运用局部多尺度和全局残差学习的方式,在参数数量更少的情况下获得更好的重建性能。而程德强等^[64]提出的多通道递归残差网络针对单一通道的递归网络可能会丢失部分特征信息的问题,提出多通道网络,提高对细节信息重建能力。

DRCN和DRRN使用的递归结构中重复运行的部分其实是以前馈的方式共享信息,即使使用了残差连接,前面的层依然无法有效利用底层信息。因此,Li等^[43]提出了基于反馈机制的网络模型SRFBN(Super-Resolution Feedback Network),设计一个反馈块(Feedback Block, FB)使得上一次反馈的输出和整个网络的输入一起重新输入到反馈模块,不断地回传。这样回传的好处在于不会增加额外的参数并且加深网络,同时使得前面的卷积层也可以学习到后面卷积层的信息,进而提升重建效果。该模型的反馈机制如图4(c)所示,并在图4中与DRCN、DRRN模型的前馈方式进行了比较。

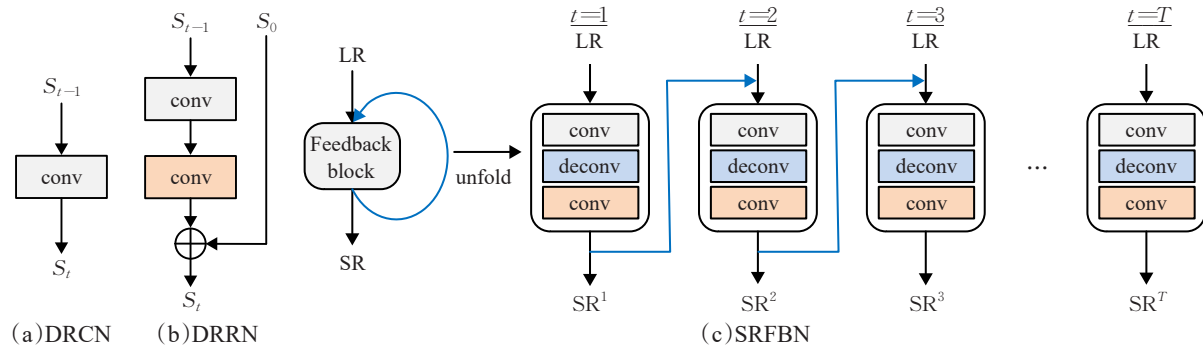


图4 DRCN,DRRN的前馈方式和SRFBN的反馈机制比较

Fig.4 Comparison of DRCN,DRRN feedforward methods and SRFBN feedback mechanism

3.1.4 密集连接

自从Huang等^[63]提出基于密集块的DenseNet以来,密集连接在视觉任务中变得越来越流行。密集连接不仅有助于减少梯度消失,增强信号传播和鼓励特征重用,而且通过采用较小的增长率(即密集块中的信道数)和级联后的压缩信道,大大减少了参数数目。对于稠密块中的每一层,所有前一层特征映射被用作输入,其自身的特征映射被用作所有后一层的输入,从而缓解了深层网络在训练过程中梯度消失而难以优化的问题。因此该方法在超分辨率领域应用丰富。这种设计的主要动机是结合网络深度上可用的分层线索,以实现高度的灵活性和丰富的特征表示。

Tong等^[39]提出了SRDenseNet(Super-Resolution DenseNet)模型,首先使用一个卷积层学习低层的特征,接着用多个密集块(Dense Block)学习高层的特征。SRDenseNet通过密集连接的方式,无论是密集块内还是全局网络中,都可以很好地将低层信息与高层信息融合,同时减少参数数量。这样的结构给整个网络带来了减轻梯度消失问题、加强特征传播、支持特征复用的效果,从而提高了重建性能。相较于SRDenseNet,Zhang等^[40]结合残差结构进一步提出RDN(Residual Dense Network)网络模型。RDN网络中的残差稠密块(Residual Dense Block,RDB)将局部残差块、密集连接层与局部特征进行融合,更好地保证了信息流的贯通。在全局网络中,与SRDenseNet的全局稠密连接不同,RDN使用了全局特征融合和全局残差学习,这样可以更好地结合残差学习以及密集链接的优点,更充分地提取多层次的特征,使训练稳定并加速训练过程,共同提高网络性能。图5比较了ResNet、DenseNet和RDN的基础块结构。

王海勇等^[66]提出了一种基于密集Inception的单图像超分辨率重建模型。该方法全局融合了简化的密集网络,核心模块引入了Inception-ResNet-A结构。因此该方法拥有较少网络参数,加快训练速度,同时得到了更好的重建效果。

3.1.5 注意力机制

在前面的网络模式中,对于超分辨率问题,所有空间

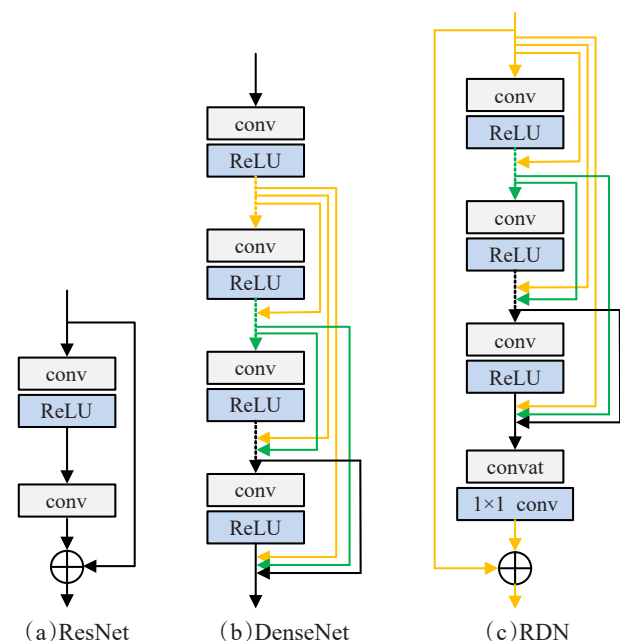


图5 ResNet、DenseNet和RDN块结构比较

Fig.5 Comparison of block structure of ResNet, DenseNet and RDN

位置和信道都有着相同的重要性(即权值相同)。在一些情况下,在给定的层中选择性地处理部分特征可能更有帮助。注意力机制则可以赋予其这样的灵活性。在结合注意力机制超分辨率模型中可以挑出更重要的激活值,给予其更大的权重,从而提升重建效果。

Choi等^[59]提出了一种新的用于图像超分辨率网络的选择单元,称为SelNet。选择单元充当卷积层之间的一扇门,只允许通过从特征图中选定的值,在每个卷积层之后都有选择单元,使得网络可以自适应地选择更有利于重建效果的响应值向下传递,提升重建效果。类似于VDSR,在SelNet中也采用了残差学习和梯度转换(梯度裁剪的一种),从而降低训练时间。

Zhang等^[41]在残差网络中加入通道注意力机制,提出RCAN(Residual Channel Attention Network)网络模型进行图像超分辨率重建。该模型使用局部嵌套残差结构(Residual in Residual)可以搭建更深的网络,并通过通道之间的依赖关系选择包含更多关键信息的通道,增

强整个网络的辨识学习能力。Dai 等^[42]使用基于协方差归一化的二阶通道注意力机制,提出 SAN(Second-order Attention Network)网络模型,通过高于一阶的特征统计量来自适应地重新缩放特征,能够更好地学习特征之间的关联度。利用协方差归一化的迭代方法来加快网络的训练速度,从而达到更优秀的网络性能和重建效果。

李彬等^[67]提出的双重注意力机制的深度网络由两种注意力机制融合而成。相较于单一的注意力机制,通过融合通道注意力机制和空间注意力机制,可以有效获取不同特征的权重值,使得网络可以根据权重精准分配计算资源,在仅仅引入极少参数的同时有效提升超分辨率重建的质量。

3.1.6 生成对抗网络

生成对抗网络^[68](Generative Adversarial Networks, GAN)采用了博弈论方法,模型由生成器和判别器两部分组成。在图像超分辨中,生成器负责生成重建图像,判别器根据自身判定条件判别生成图像与真实图像的差距,用试图欺骗判别器的方法来进一步恢复图像。由于其特殊机制的存在,通过这种方式,往往可以生成视觉效果更好的高分辨率图像。

Ledig 等^[38]提出了超分辨率生成对抗网络(Super-Resolution Generative Adversarial Network, SRGAN),首次使用 GAN 进行了单幅图像超分辨重建的尝试。如图 6 所示,SRGAN 将低分辨率图像输入至生成器 G 中进行图像重建,由判别器 D 将生成图像与真实图像对抗训练,最后输出训练后的图像。相较于之前的算法,虽然 SRGAN 在客观评价指标(如 PSNR)上相对较低,但是其无论在视觉效果,还是图像细节等直观方面,重建效果更加优秀,效果更好。这与其特殊的网络结构以及将对抗损失和感知损失相结合的损失函数有关。受 SRGAN 的启发,Wang 等^[60]又提出 ESRGAN(Enhanced SRGAN)模型,并对 SRGAN 做出改进。ESRGAN 的改进主要有以下方面:一是在生成器端去除 BN 和用残差密集块 RRDB(Residual in Residual Dense Block),这样有助于提高泛化能力,减少计算复杂度和内存使用。二是判别器端修改,使其判别不再是真伪绝对值,而是预测生成图像与真实图像的相对距离。三是对感知损失进行修改,可以更好地保持图像亮度的一致性和恢复更好的细节纹理。

彭晏飞等^[69]在 SRGAN 的基础上,迁移了 SVM 中的 Hinge 损失作为目标函数,又在原有的感知损失中加入了 Charbonnier 损失函数和 TV 正则项,使其具有更好的稳定性、鲁棒性、抗噪性。同时保持重建图像边缘和细节清晰,去掉了残差块和判别器中不必要的 BN 层,并在生成器和判别器中使用谱归一化来稳定模型训练。而汪鑫耘等^[70]在 ESRGAN 基础上对激活函数、基础网络结构和损失函数进行优化。将激活函数改为 LeakyReLU,用来避免某些神经元不激活数据现象。用 RDB 来代替 RRDB,减少参数量的同时加深了网络。采用双层特征损失,并引入渗透指数 PI 作为损失权重来改进损失函数。该方法在客观评价指标和主观视觉上较 ESRGAN 均有所提升。

3.2 损失函数

在图像超分辨领域,损失函数用来度量生成的高分辨率图像与真实图像之间的差异。早期研究人员通常采用逐像素 L2 损失,即均方误差(Mean-Square Error, MSE)。但后来发现它并不能很精确地测量图像的感知质量,因此采用其他损失函数(如内容损失^[71]、生成损失^[38]等)来更好地测量重构误差。目前,在超分辨领域中,损失函数发挥着重要的作用。

3.2.1 像素损失

像素损失主要是度量两个图像之间的像素差异,主要包括 L1 损失(平均绝对误差)和 L2 损失(均方误差)。与 L1 损失相比,L2 损失能抑制较大误差的产生,但对较小的误差往往具有更大的容错性,因此常常导致结果过于平滑。

在实际应用中,L1 损失比 L2 损失^[56]表现出更好的性能,尤其是收敛性。由于 PSNR 的定义与像素差高度相关,并且最小化像素损失直接最大化 PSNR,因此像素损失逐渐成为最广泛使用的损失函数。然而,由于像素损失实际上并不考虑图像质量(例如感知质量^[71]、纹理^[11]),往往结果缺乏高频细节,感知效果不好的过度平滑的纹理^[37-38,71]。

3.2.2 感知损失

感知损失是近年来提出的一个比较巧妙的思想。在图像超分辨领域中,由于经常使用 L1 损失,会导致重建图像较为平滑,丢失细节部分或高频部分。因此感知

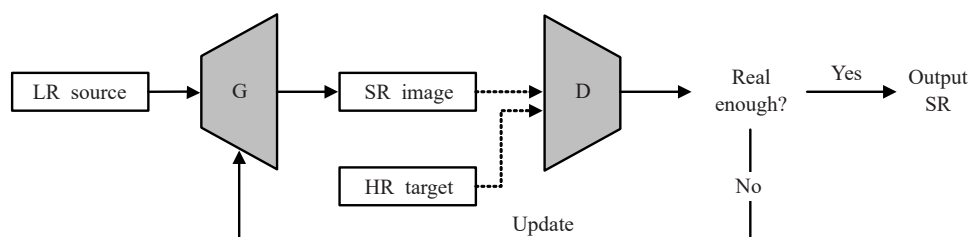


图6 SRGAN 基本结构示意图

Fig.6 Basic structure diagram for SRGAN

损失被引入到超分辨领域中^[50,72],从而更好地评价并提升的感知质量。

感知损失使用已经训练好的图像训练网络,其中ResNet^[61]和VGG^[73]是最常用的CNN预训练网络,然后提取两个图像的特征,计算两个图像在特征空间上的欧氏距离。感知损失与像素损失相比,没有迫使像素之间精确地匹配,而是在感知上与目标图像类似。因此其在视觉上产生更佳的效果^[38,58,71]。

3.2.3 纹理损失

由于重建图像应该与目标图像具有相同的纹理风格,并受Leon等^[74-75]的影响,纹理损失被引入到图像超分辨领域中。图像的纹理被视为不同特征通道之间的相关性。

纹理损失可使SR模型创建更真实的纹理,并在视觉上产生更令人满意的结果^[44,76]。然而,确定图块大小以匹配纹理仍然是经验性的问题。由于纹理统计信息是在不同纹理区域取平均值,过小的补丁会导致纹理区域出现伪影现象,而过大的补丁会导致整个图像出现伪影现象。

3.2.4 对抗损失

在基于生成对抗网络的超分辨领域中,通常超分辨模型为生成器,而判别器根据判断条件来进行判别。而损失函数则通常在像素损失或感知损失的基础上再结合对抗损失,从而通过判别器使得生成器生成的图像能有更逼真的高分辨率图片。在对抗损失计算方面,SRGAN^[38]和EnhanceNet^[76]等方法使用交叉熵损失函数。而后,Wang等^[77]和Yuan等^[78]使用最小二乘方法^[79]进行损失计算,可以生成质量更高的图像。RaGAN^[80]和WGAN-GP^[81]等变体GAN模型使用新的距离度量方式改进生成对抗学习的性能,例如FID(Fr chet Inception Distance)和Wasserstein距离。

目前GAN的训练过程还比较困难和不稳定。尽管已有一些关于如何稳定GAN训练的研究,但如何保证整合到超分辨模型中的GAN得到正确训练并发挥积极作用仍是一个问题。

3.3 上采样方法和模型框架

上采样指从低分辨率(LR)图像输出高分辨率(HR)

图像的操作。在图像超分辨深度网络模型中,依据上采样层的位置和上采样操作的基本方式,大致分为4类模型框架,包括预定义上采样模型框架、后端上采样模型框架、渐进式上采样模型框架和上下采样迭代式模型框架。所使用的上采样方法主要可分为两大类,基于插值的上采样方法和基于学习的上采样方法。

3.3.1 上采样方法

基于插值的上采样方法指通过传统的图像插值操作进行图像尺寸的放大。常用的插值方法主要有最近邻插值、双线性插值和双立方插值。基于学习的上采样方法指以端到端学习的方式进行上采样,可以通过在网络结构中加入转置卷积层或者亚像素卷积层实现。表2给出了上采样方法的分类和相应的描述。

传统基于插值的上采样方法原理比较简单,也容易实现。其中,双立方插值还经常应用在预定义上采样模型框架中,但是也增加了模型的计算复杂度。另外,插值方法还会引起噪声被放大等副作用,因此可以使用端到端学习的方式进行上采样,以避免这些问题的发生。例如,转置卷积应用在SRDenseNet^[39]等多个超分辨网络模型中,亚像素卷积应用在RDN^[40]和SRGAN^[38]等多种不同的网络模型中。亚像素卷积比转置卷积有更大的感受野,因此能够提供更多的邻域相关信息,重建图像细节部分更真实。

3.3.2 模型框架

4种模型框架如图7所示。图7(a)给出了预定义上采样框架示意图,如图可知,该框架是先使用传统的插值方法将图像尺寸放大到要求的倍数,然后再使用深度学习网络模型重建图像细节。该框架所使用的上采样方法简单,但深度神经网络的很多操作都需要在高维空间进行,占用空间大,时间效率低。图7(b)给出了后上采样框架示意图,如图所示,上采样层位于网络结构末端。这样,深度神经网络的很多操作在低维空间进行,极大提高了训练速度和算法效率。网络可以自适应地学习上采样过程,进行端到端可学习的上采样。图7(c)给出了逐步上采样框架示意图,基于该框架的深度网络模型通过多次上采样逐步获得目标尺寸的超分辨图像。这样,学习难度降低,并且在不需要过多时间和空间成

表2 上采样方法分类和简单描述
Table 2 Category and description of upsampling methods

方法分类	方法名称	方法描述
基于插值的上采样方法	最近邻插值	使用距离最近的像素点的值进行插值操作
	双线性插值	沿着图像的一个方向进行插值,然后再沿着另一个方向进行插值,使用线性插值函数。插值的结果与先进行哪个方向的插值无关
	双立方插值	使用的插值函数为三次多项式,也是进行图像两个方向的插值操作
基于学习的上采样方法	转置卷积	也称为反卷积,卷积核大小为3×3,步长为1,并使用0进行填充的卷积操作
	亚像素卷积	通过卷积操作得到多个通道的特征图,然后对特征图进行重新排列,在多个通道之间重组,从而放大图像的尺寸

本的情况下,能够满足多尺度超分辨的需求。图7(d)给出了迭代上下采样框架示意图,交替进行上采样和下采样操作,通过反向映射不断地改进超分辨重建图像的质量。最后,基于全部中间上采样层得到的高分辨率特征图,重建最终的超分辨图像。

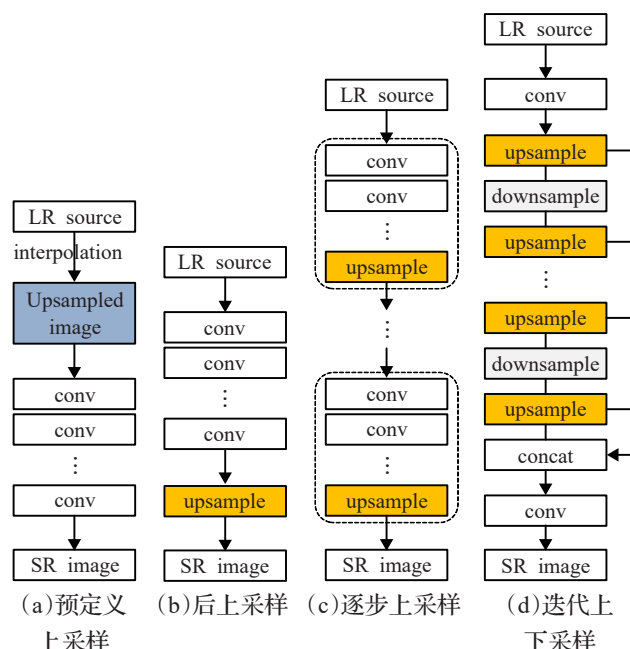


图7 上采样模型框架比较

Fig.7 Comparison of upsampling frameworks

4 数据集及实验分析

4.1 数据集介绍

目前,有很多公用数据集可用于基于深度学习的单幅图像超分辨重建任务。这些数据集包含的图像数量和图像内容各不相同,可以综合测试超分辨重建方法的性能。表3列出了常用的数据集,并进行了简单介绍。一些数据集已被划分为固定的训练集、验证集和测试集,也有一些较大的数据集经常作为训练集,如ImageNet^[89]、DIV2K^[87]和Flickr2K^[88]等。研究人员也可以根据使用需求在数据集上自行划分训练集、验证集和测试集,或者通过图像旋转等方式进行训练集扩增,或者联合多个数据集进行训练。

4.2 典型方法实验结果及分析

依据不同的网络结构,本文选取8种较为典型的深度学习单幅图像超分辨重建方法进行实验,并与传统的双立方插值方法进行比较。这8种方法包括属于线性网络结构的SRCNN^[34]和FSRCNN^[53],属于残差网络结构的VDSR^[36],属于递归网络结构的DRCN^[37],属于密集连接结构的RDN^[40],使用注意力机制的SAN模型^[42],基于生成对抗网络学习的SRGAN^[38]、USRGAN^[90]。然后选取了4个常用的数据集,包括Set5、Set14、BSD100和Urban100,并采用4倍放大因子进行了测试。训练数据集使用T91和DIV2K。

本文从Set14和Urban100各选取了1张图片,用于展示和比较不同算法的重建效果,如图8和图9所示。

表3 常用数据集介绍

Table 3 Description of widely used datasets

数据集名称	图片数量	图片格式	数据集简单描述
Set5 ^[82]	5	PNG	包含的5张图片分别来自婴儿、鸟、蝴蝶、小孩的头部的一个女士
Set14 ^[83]	14	PNG	包含的图片来自人物、动物、昆虫、花、蔬菜和喜剧演员等
Berkeley segmentation ^[84]	500	JPG	简称BSD500。包含的图片来自动物、建筑、食物、人和植物等。经常使用其中的100张或300张图片,称为BSD100、BSD300数据集
Urban100 ^[85]	100	PNG	包含的图片主要是不同类型的城市建筑物
Manga109 ^[86]	109	PNG	包含的图片全部来自日本漫画
T91 ^[26]	91	PNG	包含的图片来自车辆、花、水果和人脸等,经常作为训练集使用
General-100 ^[52]	100	BMP	包含的图片来自动物、日常用品、食物、植物和人物等,经常作为训练集使用
DIV2K ^[87]	1 000	PNG	高清图片数据集,图片来自自然环境、风景、手工艺品和人物等,其中的800张图片经常作为训练集使用
Flickr2K ^[88]	2 650	PNG	包含的图片来自人物、动物和风景等,图片数量多,经常作为较大规模训练集使用

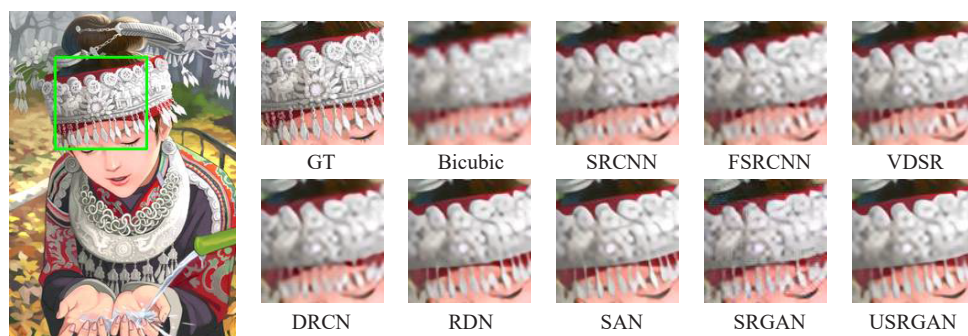


图8 Set14数据集的示例图片实验结果比较

Fig.8 Results comparison of different methods for sample image from Set14 dataset

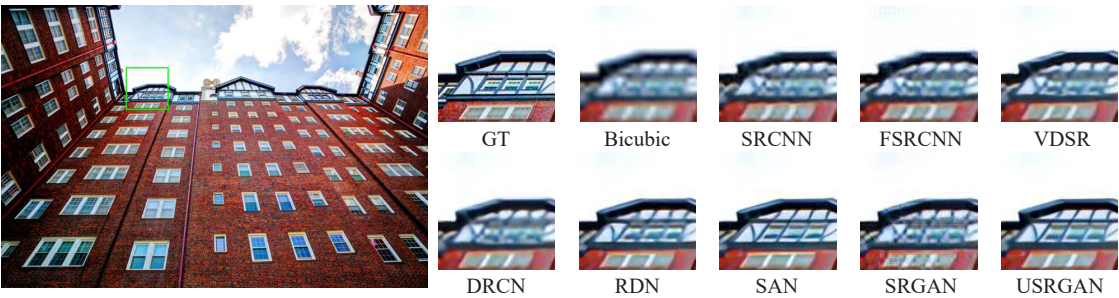


图9 Urban100数据集的示例图片实验结果比较

Fig.9 Results comparison of different methods for sample image from Urban100 dataset

表4 不同方法在4个数据集上的PSNR/SSIM评价指标比较(×4)

Table 4 Comparison of PSNR and SSIM for different methods on 4 datasets(×4)

方法名称	Set5		Set14		BSD100		Urban100	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
Bicubic	28.43	0.810 9	26.00	0.702 3	25.96	0.667 8	23.24	0.657 4
SRCNN	30.48	0.862 8	27.50	0.751 3	26.90	0.722 1	24.52	0.722 1
FSRCNN	30.70	0.865 7	27.61	0.755 0	26.98	0.715 0	24.60	0.712 8
VDSR	31.35	0.885 4	28.02	0.768 0	27.29	0.725 2	25.18	0.752 5
DRCN	31.53	0.885 4	28.03	0.767 3	27.24	0.723 3	25.14	0.751 1
RDN	32.47	0.899 0	28.81	0.787 1	27.72	0.741 9	26.61	0.802 8
SAN	32.64	0.900 3	28.92	0.788 8	27.78	0.743 6	26.79	0.806 8
SRGAN	29.40	0.847 2	26.02	0.739 7	25.16	0.668 8	24.12	0.697 4
USRGAN	32.42	0.896 5	28.83	0.765 9	27.69	0.735 1	26.44	0.802 1

其中,GT(Ground Truth)表示从原图像中截取的部分。表4给出了这些方法在4个常用数据集上的PSNR和SSIM评价指标,实验结果表现各有差异。

由图8和图9可看出,主观视觉特性较好的有RDN、SAN和USRGAN三种网络模型,从表4中可知,其相应的评价指标也较高。这几种网络模型是近几年(2018年及以后)发表的,与前些年发表的方法相比,在超分辨重建质量和效果上确实有了很大的提升。SRGAN模型的PSNR和SSIM评价指标在表4中较低,甚至低于最开始的超分辨深度学习模型SRCNN。但是从主观感知上看,观察图8和图9,SRGAN的重建性能表现优于其他早期(2018年之前)的方法。由此说明,GAN网络的生成对抗学习方式可以增强超分辨重建图像的逼真程度,但也会引起图像结构上的一些变形。因此,基于GAN网络的多种改进模型近年来不断被提出,这也是目前深度学习在图像超分辨领域的发展趋势。

5 结束语

本文主要归纳和分析了近年来单幅图像超分辨重建的深度学习方法。从早期的浅层模型SRCNN开始,到越来越深的残差网络模型VDSR、递归网络模型DRCN和密集连接网络模型RDN等,并结合注意力机制,使得图像超分辨重建质量不断得到提高。质量评价客观指标PSNR和SSIM从SRCNN开始就优于传统方法,并且重建图像的主观感受质量也比传统方法高。研

究学者们不断把新出现的CNN网络结构和设计策略应用在SISR领域,并进行改进,以进一步提高PSNR和SSIM指标值。但是重建图像的主观视觉感受质量却没有得到同步提升,不能达到理想的效果。

而基于GAN网络结构的SISR方法,尽管PSNR和SSIM客观质量评价值较低,却可以生成主观评价质量高的HR图像,这是由于GAN网络模型方便结合感知损失函数。因此,近两年来,基于GAN网络的SISR方法迅速发展,并在解决实际问题中得到了应用。随着基于深度学习的SISR技术越来越成熟和稳定,研究人员希望超分辨重建能够在视觉高层任务中得到更多的应用。但目前仍存在问题,这也是未来的研究方向和发展趋势:

(1)算法的效率。现有方法虽然可以重建生成较高质量的HR图像,但是往往需要较长的时间,特别是在缩放因子较大(×8和×16)的情况下。同时,还需要占用很多的内存资源完成模型的训练。为了能够满足实际应用的需求,需要在保证重建质量的前提下,进一步优化网络结构,降低模型参数数量,对算法进行轻量化改进,提高时间效率。

(2)真实图像超分辨。在真实的应用场景中,图像退化模型经常是未知的,难以获得匹配的LR-HR图像数据集进行训练,直接进行盲超分辨效果也不够好。因此,可结合指定任务情景下的先验信息进行无监督的学习,这将是亟待发展的研究方向。

(3)重建结果评价。目前的重建结果评价指标仍然以PSNR和SSIM为主,作为客观评价指标有时候不能准确评估图像的重建逼真程度,与人眼视觉的主观感知有差异。而常用的主观评价MOS不能很方便地被使用,现在也有一些研究人员提出了新的评价指标,但是没有受到广泛的认可。因此,急需进一步找到更合适的图像超分辨重建评价指标体系。

(4)合适的损失函数。损失函数在整个算法模型中其实也占据着很重要的地位,而目前对于SR任务而言,尚无明确的最合适的损失函数。现在使用最多的是多种损失函数的加权组合,目的是增强重建图像的主观感知程度,实际上这是SR的根本目标。损失函数决定了算法网络模型学习的方向,而学习的目的是提高SR的结果评价指数。因此,本文认为,可以将评价指标和损失函数结合进行选择,这也是未来的一个研究趋势。

参考文献:

- [1] ISAAC J S, KULKARNI R. Super resolution techniques for medical image processing[C]//Proceedings of International Conference on Technologies for Sustainable Development, 2015:2170-2175.
- [2] 邢晓羊,魏敏,符颖.基于特征损失的医学图像超分辨率重建[J].计算机工程与应用,2018,54(20):202-207.
XING X Y, WEI M, FU Y. Super-resolution reconstruction of medical images using feature-based loss[J]. Computer Engineering and Applications, 2018, 54(20):202-207.
- [3] JIA X, LI F, FRASER D. Universal HMT based super resolution for remote sensing images[C]//Proceedings of IEEE International Conference on Image Processing, 2008:333-336.
- [4] ZHANG L, ZHANG H, SHEN H, et al. A super-resolution reconstruction algorithm for surveillance images[J]. Signal Processing, 2010, 90(3):848-859.
- [5] HARRIS J L. Diffraction and resolving power[J]. Journal of the Optical Society of America, 1964, 54(7):931-936.
- [6] GOODMAN J W. Introduction to Fourier optics[M]. 2nd ed. New York: McGrawHill, 1996.
- [7] BROWN H A. Effect of truncation on image enhancement by prelate spheroidal functions[J]. Journal of the Optical Society of America, 1969, 59(2):228-229.
- [8] WADAKA S, SATO T. Super-resolution in incoherent imaging system[J]. Journal of the Optical Society of America, 1975, 65(3):354-355.
- [9] KEYS R. Cubic convolution interpolation for digital image processing[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1981, 29(6):1153-1160.
- [10] KIM H, CHA Y, KIM S. Curvature interpolation method for image zooming[J]. IEEE Transactions on Image Processing, 2011, 20(7):1895-1903.
- [11] MISHIBA K, SUZUKI T, IKEHARA M. Edge-adaptive image interpolation using constrained least squares[C]//Proceedings of IEEE International Conference on Image Processing, 2010:2837-2840.
- [12] HAN J W, KIM J H, SULL S, et al. New edge-adaptive image interpolation using anisotropic Gaussian filters[J]. Digital Signal Processing, 2013, 23(1):110-117.
- [13] LIU Y, ZHANG Y, GUO Q, et al. Image interpolation based on weighted and blended rational function[C]//Asian Conference on Computer Vision, 2014:78-88.
- [14] IRANI M, PELEG S. Improving resolution by image registration[J]. Graphical Models and Image Processing, 1991, 53(3):231-239.
- [15] YANG X, ZHANG Y, ZHOU D, et al. An improved iterative back projection algorithm based on ringing artifacts suppression[J]. Neurocomputing, 2015, 162:171-179.
- [16] TAN J, TAO Z, CAO A, et al. An edge-preserving iterative back-projection method for image super-resolution[C]//Proceedings of International Conference on Digital Image Processing, 2016:1-8.
- [17] TEKALP A M, OZKAN M K, SEZAN M I. High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration[C]//Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 1992, 3:169-172.
- [18] PATTI A J, ALTUNBASAK Y. Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants[J]. IEEE Transactions on Image Processing, 2001, 10(1):179-186.
- [19] KIM J Y, PARK R H, YANG S. Super-resolution using POCS-based reconstruction with artifact reduction constraints[C]//Proceedings of SPIE on Visual Communications and Image Processing, 2005, 5960(4):1810-1818.
- [20] SUN D, GAO D, LU Y, et al. A high quality single-image super-resolution algorithm based on linear Bayesian MAP estimation with sparsity prior[J]. Digital Signal Processing, 2014, 35:45-52.
- [21] VRIGKAS M, NIKOU C, KONDI L P. Robust maximum a posteriori image super-resolution[J]. Journal of Electronic Imaging, 2014, 23(4):30-35.
- [22] CHANTAS G K, GALATSANOS N P, WOODS N A. Super-resolution based on fast registration and maximum a posteriori reconstruction[J]. IEEE Transactions on Image Processing, 2007, 16(7):1821-1830.
- [23] CHANG H, YEUNG D Y, XIONG Y. Super-resolution through neighbor embedding[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004:1-8.
- [24] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Neighbor embedding based single-image super-resolution using semi-nonnegative matrix factorization[C]//Proceedings of

- IEEE International Conference on Acoustics, Speech and Signal Processing, 2012: 1289-1292.
- [25] CHEN X, QI C. Low-rank neighbor embedding for single image super-resolution[J]. IEEE Signal Processing Letters, 2014, 21(1): 79-82.
- [26] 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.
- [27] DONG W, ZHANG L, SHI G, et al. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization[J]. IEEE Transactions on Image Processing, 2011, 20(7): 1838-1857.
- [28] PELEG T, ELAD M. A statistical prediction model based on sparse representations for single image super-resolution[J]. IEEE Transactions on Image Processing, 2014, 23(6): 2569-2582.
- [29] NI K S, NGUYEN T Q. Image super-resolution using support vector regression[J]. IEEE Transactions on Image Processing, 2007, 16(6): 1596-1610.
- [30] KIM K I, KWON Y. Single-image super-resolution using sparse regression and natural image prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(6): 1127-1133.
- [31] HE H, SIU W C. Single image super-resolution using Gaussian process regression[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2011: 449-456.
- [32] BAKER S, KANADE T. Limits on super-resolution and how to break them[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(9): 1167-1183.
- [33] LI X G, LI F H, ZHUO L. Research progress of joint reconstruction of high resolution and high dynamic range images[J]. Measurement and Control Technology, 2012, 31(5): 8-12.
- [34] 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, 2014: 184-189.
- [35] WANG Z W, LIU D, YANG J C, et al. Deep networks for image super-resolution with sparse prior[C]//Proceedings of IEEE International Conference on Computer Vision, 2015: 370-378.
- [36] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016: 1646-1654.
- [37] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016: 1637-1645.
- [38] LEDIG C, THEIS L, HUSZAR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2017: 105-114.
- [39] TONG T, LI G, LIU X, et al. Image super-resolution using dense skip connections[C]//Proceedings of IEEE International Conference on Computer Vision, 2017: 4809-4817.
- [40] ZHANG Y L, TIAN Y P, KONG Y, et al. Residual dense network for image super-resolution[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018: 2472-2481.
- [41] ZHANG Y L, LI K P, LI K, et al. Image super-resolution using very deep residual channel attention networks[C]//Proceedings of European Conference on Computer Vision, 2018: 294-310.
- [42] DAI T, CAI J R, ZHANG Y B, et al. Second-order attention network for single image super-resolution[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019: 1057-1066.
- [43] LI Z, YANG J L, LIU Z, et al. Feedback network for image super-resolution[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019: 3862-3871.
- [44] YANG F Z, YANG H, FU J L, et al. Learning texture transformer network for image super-resolution[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020: 5790-5799.
- [45] MA C, RAO Y M, CHENG Y, et al. Structure-preserving super resolution with gradient guidance[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020: 7766-7775.
- [46] 史振威, 雷森. 图像超分辨率重建算法综述[J]. 数据采集与处理, 2020, 35(1): 1-20.
- SHI Z W, LEI S. Review of image super-resolution reconstruction[J]. Journal of Data Acquisition and Processing, 2020, 35(1): 1-20.
- [47] 唐艳秋, 潘泓, 朱亚平, 等. 图像超分辨率重建研究综述[J]. 电子学报, 2020, 48(7): 1407-1420.
- TANG Y Q, PAN H, ZHU Y P, et al. A survey of image super-resolution reconstruction[J]. Acta Electronica Sinica, 2020, 48(7): 1407-1420.
- [48] 刘颖, 朱丽, 林庆帆, 等. 图像超分辨率技术的回顾与展望[J]. 计算机科学与探索, 2020, 14(2): 181-199.
- LIU Y, ZHU L, LIN Q F, et al. Review and prospect of image super-resolution technology[J]. Journal of Frontiers of Computer Science and Technology, 2020, 14(2): 181-199.
- [49] 南方哲, 钱育蓉, 行艳妮, 等. 基于深度学习的单图像超分辨率重建研究综述[J]. 计算机应用研究, 2020, 37(2): 321-326.
- NAN F Z, QIAN Y R, XING Y N, et al. Survey of single

- image super resolution based on deep learning[J]. Application Research of Computers, 2020, 37(2): 321-326.
- [50] ZHANG K, ZUO W, ZHANG L. Learning a single convolutional super-resolution network for multiple degradations[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018: 3262-3271.
- [51] 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.
- [52] WANG Z, BOVIK A C. Mean squared error: let it or leave it? A new look at signal fidelity measures[J]. IEEE Signal Processing Magazine, 2009, 26(1): 98-117.
- [53] DONG C, LOY C C, TANG X O. Accelerating the super-resolution convolutional neural network[C]//Proceedings of European Conference on Computer Vision, 2016: 391-407.
- [54] SHI W Z, CABALLERO J, HUSZAR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016: 1874-1883.
- [55] LAI W S, HUANG J B, AHUJAN, et al. Deep Laplacian pyramid networks for fast and accurate super-resolution [C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2017: 624-632.
- [56] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2017: 1132-1140.
- [57] LI J C, FANG F M, MEI K F, et al. Multi-scale residual network for image super-resolution[C]//Proceedings of European Conference on Computer Vision, 2018: 527-542.
- [58] TAI Y, YANG J, LIU X M. Image super-resolution via deep recursive residual network[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2017: 2790-2798.
- [59] CHOI J, KIM M. A deep convolutional neural network with selection units for super-resolution[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2017: 1150-1156.
- [60] WANG X T, YU K, WU S H, et al. ESRGAN: enhanced super-resolution generative adversarial networks[C]//Proceedings of European Conference on Computer Vision Workshops, 2018: 63-79.
- [61] HE K, ZHANG X, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [62] 应自炉, 龙祥. 多尺度密集残差网络的单幅图像超分辨率重建[J]. 中国图象图形学报, 2019, 24(3): 410-419.
- YING Z L, LONG X. Single-image super-resolution construction based on multi-scale dense residual network[J]. Journal of Image and Graphics, 2019, 24(3): 410-419.
- [63] 李现国, 冯欣欣, 李建雄. 多尺度残差网络的单幅图像超分辨率重建[J]. 计算机工程与应用, 2021, 57(7): 215-221.
- LI X G, FENG X X, LI J X. Single image super-resolution reconstruction based on multi-scale residual network[J]. Computer Engineering and Applications, 2021, 57(7): 215-221.
- [64] 程德强, 郭昕, 陈亮亮, 等. 多通道递归残差网络的图像超分辨率重建[J]. 中国图象图形学报, 2021, 26(3): 605-618.
- CHENG D Q, GUO X, CHEN L L, et al. Image super-resolution reconstruction from multi-channel recursive residual network[J]. Journal of Image and Graphics, 2021, 26(3): 605-618.
- [65] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2017: 2261-2269.
- [66] 王海勇, 张开心, 管维正. 基于密集Inception的单图像超分辨率重建[J/OL]. 计算机应用(2021-04-15)[2021-07-13]. <https://kns.cnki.net/kcms/detail/51.1307.TP.20210415.1300.004.html>.
- WANG H Y, ZHANG K X, GUAN W Z. Single image super-resolution reconstruction based on dense Inception [J/OL]. Journal of Computer Applications (2021-04-15) [2021-07-13]. <https://kns.cnki.net/kcms/detail/51.1307.TP.20210415.1300.004.html>.
- [67] 李彬, 王平, 赵思逸. 基于双重注意力机制的图像超分辨率重建算法[J]. 图学学报, 2021, 42(2): 206-215.
- LI B, WANG P, ZHAO S Y. Image super-resolution reconstruction based on dual attention mechanism[J]. Journal of Graphics, 2021, 42(2): 206-215.
- [68] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of Advances in Neural Information Processing Systems, 2014: 2672-2680.
- [69] 彭晏飞, 高艺, 杜婷婷, 等. 生成对抗网络的单图像超分辨率重建方法[J]. 计算机科学与探索, 2020, 14(9): 1612-1620.
- PENG Y F, GAO Y, DU T T, et al. Single image super-resolution reconstruction method for generative adversarial network [J]. Journal of Frontiers of Computer Science and Technology, 2020, 14(9): 1612-1620.
- [70] 汪鑫耘, 李丹. 生成式对抗网络在超分辨率图像重建中的应用[J]. 计算机科学与探索, 2020, 14(4): 680-687.
- WANG X Y, LI D. Application of generative adversarial network in super-resolution image reconstruction[J]. Journal of Frontiers of Computer Science and Technology, 2020, 14(4): 680-687.
- [71] JUSTIN J, ALEXANDRE A, LI F F. Perceptual losses for real-time style transfer and super-resolution[C]//Proceed-

- ings of European Conference on Computer Vision, 2016: 694-711.
- [72] ALEXEY D, THOMAS B. Generating images with perceptual similarity metrics based on deep networks[C]// Proceedings of Annual Conference on Neural Information Processing Systems, 2016: 658-666.
- [73] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]// Proceedings of International Conference on Learning Representations, 2015: 234-240.
- [74] LEON A G, ALEXANDER S E, MATTHIAS B. Texture synthesis using convolutional neural networks[C]// Proceedings of Annual Conference on Neural Information Processing Systems, 2015: 262-270.
- [75] LEON A G, ALEXANDER S E, MATTHIAS B. Image style transfer using convolutional neural networks[C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016: 2414-2423.
- [76] SAJJADI M, SCHOLKOPF B, HIRSCH M, et al. EnhanceNet: single image super-resolution through automated texture synthesis[C]// Proceedings of IEEE International Conference on Computer Vision, 2017: 4501-4510.
- [77] WANG Y F, PERAZZI F, MCWILLIAMS B, et al. A fully progressive approach to single-image super-resolution[C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2018: 977-986.
- [78] YUAN Y, LIU S Y, ZHANG J W, et al. Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks[C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2018: 814-823.
- [79] MAO X D, LI Q, XIE H R, et al. Least squares generative adversarial networks[C]// Proceedings of IEEE International Conference on Computer Vision, 2017: 2813-2821.
- [80] ALEXIA J M. The relativistic discriminator: a key element missing from standard GAN[C]// Proceedings of International Conference on Learning Representations, 2019: 1240-1246.
- [81] ISHAAN G, FARUK A, MARTIN A, et al. Improved training of Wasserstein GANs[C]// Proceedings of Annual Conference on Neural Information Processing Systems, 2017: 5768-5778.
- [82] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-complexity single-image super-resolution based on non-negative neighbor embedding[C]// Proceedings of British Machine Vision Conference, 2012: 201-210.
- [83] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations[C]// Proceedings of International Conference on Curves and Surfaces, 2010: 711-730.
- [84] 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 IEEE International Conference on Computer Vision, 2001: 416-423.
- [85] HUANG J B, SINGH A, AHUJA N. Single image super-resolution from transformed self-exemplars[C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015: 5197-5206.
- [86] FUJIMOTO A, OGAWA T, YAMAMOTO K, et al. Manga109 dataset and creation of metadata[C]// Proceedings of International Workshop on Comics Analysis, Processing and Understanding, 2016: 1110-1118.
- [87] AGUSTSSON E, TIMOFTE R. NTIRE 2017 challenge on single image super-resolution: dataset and study[C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2017: 1122-1131.
- [88] TIMOFTE R, AGUSTSSON E, GOOL L C, et al. NTIRE 2017 challenge on single image super-resolution: methods and results[C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2017: 1110-1121.
- [89] RUSSAKOVSKY O, DENG J, SU H, et al. Image net large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
- [90] ZHANG K, VAN G L, TIMOFTE R. Deep unfolding network for image super-resolution[C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020: 3214-3223.