

# 基于深度学习的单帧图像超分辨率重建综述

吴 靖<sup>1,2</sup>, 叶晓晶<sup>1,2</sup>, 黄 峰<sup>1,2</sup>, 陈丽琼<sup>1,2</sup>, 王志锋<sup>1,2</sup>, 刘文犀<sup>2,3</sup>

(1. 福州大学机械工程及自动化学院, 福建福州 350116; 2. 福州大学先进技术创新研究院, 福建福州 350116;  
3. 福州大学计算机与大数据学院, 福建福州 350116)

**摘要:** 图像超分辨率重建是计算机视觉中的基本图像处理技术之一, 不仅可以提高图像分辨率改善图像质量, 还可以辅助其他计算机视觉任务. 近年来, 随着人工智能浪潮的兴起, 基于深度学习的图像超分辨率重建也取得了显著进展. 本文在简述图像超分辨率重建方法的基础上, 全面综述了基于深度学习的单帧图像超分辨率重建的技术架构及研究历程, 包括数据集构建方式、网络模型基本框架以及用于图像质量评估的主、客观评价指标, 重点介绍了根据网络结构及图像重建效果划分的基于卷积神经网络的方法、基于生成对抗网络的方法以及基于Transformer的方法, 并对相关网络模型加以评述和对比, 最后依据网络模型和超分辨率重建挑战赛相关内容, 展望了图像超分辨率重建未来的发展趋势.

**关键词:** 超分辨率重建; 深度学习; 单帧图像; 卷积神经网络; 生成对抗网络; Transformer; 挑战赛  
**中图分类号:** TP751; TP391      **文献标识码:** A      **文章编号:** 0372-2112(2022)09-2265-30  
**电子学报 URL:** <http://www.ejournal.org.cn>      **DOI:** 10.12263/DZXB.20220091

## A Review of Single Image Super-Resolution Reconstruction Based on Deep Learning

WU Jing<sup>1,2</sup>, YE Xiao-jing<sup>1,2</sup>, HUANG Feng<sup>1,2</sup>, CHEN Li-qiong<sup>1,2</sup>, WANG Zhi-feng<sup>1,2</sup>, LIU Wen-xi<sup>2,3</sup>

(1. School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou, Fujian 350116, China;  
2. Advanced Technology Innovation Institute, Fuzhou University, Fuzhou, Fujian 350116, China;  
3. College of Computer and Data Science, Fuzhou University, Fuzhou, Fujian 350116, China)

**Abstract:** Image super-resolution reconstruction is one of the basic image processing techniques in computer vision, which can not only improve image resolution and image quality, but also assist other computer vision tasks. In recent years, with the rise of artificial intelligence, deep-learning-based image super-resolution reconstruction has also made remarkable progress. Based on a brief description of the image super-resolution reconstruction methodology, this paper comprehensively reviews the technical architecture and research process of deep-learning-based single image super-resolution reconstruction, including the method of datasets construction, the basic framework of the network model, the subjective and objective evaluation metrics for image quality evaluation. The methods based on convolutional neural networks, generative adversarial networks and Transformer, which are divided according to network structure and image reconstruction effect are mainly introduced, and related network models are reviewed and compared. Finally, the future development trend of image super-resolution reconstruction is prospected according to the related content of network model and super-resolution reconstruction challenges.

**Key words:** super-resolution reconstruction; deep learning; single image; convolutional neural network; generative adversarial network; transformer; challenges

## 1 引言

图像超分辨率重建(Super-resolution Reconstruction, SR)是指将同一场景的一张或多张低分辨率(Low

Resolution, LR)退化图像恢复成对应的一张或多张高分辨率(High Resolution, HR)清晰图像的技术,是计算机视觉和图像处理领域的重要技术之一. 图像SR不仅

可以提高图像的感知质量,还有助于提升目标检测、图像去噪等其他计算机视觉任务的性能<sup>[1-3]</sup>。相比于设计更复杂的光学成像系统来提升图像质量,图像SR技术能够在达到相同效果的同时大大降低成本,也能突破衍射极限对光学成像系统的限制,获取更高分辨率的重建图像,因此在视频监控、医疗成像、卫星遥感等领域有着广泛的应用<sup>[4-8]</sup>。

根据低分辨率图像在网络模型中输入数量的不同,可将图像SR技术分为单帧图像超分辨率重建(Single Image Super-resolution Reconstruction, SISR)以及多帧图像超分辨率重建(Multi-Image Super-resolution Reconstruction, MISR)。其中,SISR可大致分为3类:基于插值的方法、基于重构的方法和基于学习的方法<sup>[9]</sup>。基于学习的方法按照学习程度的不同,又可分为基于浅层学习的方法和基于深度学习的方法<sup>[10]</sup>。MISR主要可分为频域法和空域法<sup>[11]</sup>。相较于MISR需要多张同一场景具有亚像素位移的LR图像作为输入,且图像间亚像素位移的不可预知性给充分利用图像的混叠信息带来了一定难度,SISR只需输入一张LR图像即可重建出图像的纹理细节,具有较高的实用价值,因此是目前图像超分辨率重建领域的主要研究方向。

早在20世纪60年代,Harris<sup>[12]</sup>和Goodman<sup>[13]</sup>就分别提出单帧图像超分辨率重建的相关方法,并称为Harris-Goodman频谱外推法。随后,Tsai等人<sup>[14]</sup>于1984年提出用于多帧图像超分辨率重建的频域处理法。自此,研究者开始关注并研究图像超分辨率重建技术,各种基于插值的方法<sup>[15,16]</sup>、基于重构的方法<sup>[17-21]</sup>也相继被提出。随着机器学习在计算机视觉和图像处理领域的发展,Freeman等人<sup>[22]</sup>将机器学习应用于图像超分辨率重建领域,并于2000年首次提出了基于学习的图像超分辨率重建方法。此后各种基于浅层学习的方法<sup>[23-26]</sup>也陆续被提出。然而这些传统的方法大多是通过提取对图像轮廓等纹理细节表达能力有限的图像底层特征来重建高分辨率图像,故在很大程度上限制了图像的重建效果。

近年来,随着深度学习技术的快速发展,研究人员对基于深度学习的图像超分辨率重建技术展开积极的探索和研究。相较于传统方法,基于深度学习的方法能够从数据集中提取到更具表达能力的图像特征,自适应地学习低分辨率与高分辨率图像之间的映射关系,不仅有效地克服了图像获取过程中出现的模糊、噪声等退化因素的影响,同时也在各种标准数据集上取得了更好的重建效果并展现出更优的网络性能。

随着图像SR研究成果的逐年增多,综述文献的归纳整理变得极为重要。早期的图像SR综述文献<sup>[27-29]</sup>主要针对传统SR方法的算法原理及其研究成果进行总

结归纳,本文不再赘述。自深度学习应用于图像SR领域以来,相关SR综述文献开始侧重于基于深度学习的图像SR内容的阐述。大多数文献<sup>[30-33]</sup>从网络结构设计、上采样方式和损失函数等方面对SR研究成果进行论述,并总结分析不同网络模型的相关内容。部分文献<sup>[11,34,35]</sup>从有、无监督学习等角度出发,阐述分析SR中具有代表性的研究成果。唐艳秋等人<sup>[36]</sup>从模型类型、网络结构、信息传递方式等方面对各种SR算法进行详细评述,并对比分析不同算法的优缺点。而Wang等人<sup>[37]</sup>同样从有、无监督SR的角度出发,详细综述了SR最新进展,并介绍了一些特定领域的应用,同时对比部分网络模型的精度、大小和计算代价等内容。Anwar等人<sup>[38]</sup>则根据网络模型的结构差异,提出了一种新的分类方法,将现有算法分为线性、残差、多分支、递归、渐进等9种类型,并对模型之间的网络复杂性、内存占用等加以比较。Chen等人<sup>[39]</sup>针对真实世界的单帧图像超分辨率重建(Real-world SISR, RSISR)进行全面综述,并总结出四大类RSISR方法,对RSISR技术的进一步发展和应用具有重要意义。

其中,部分文献<sup>[30-32,35,38,39]</sup>仅介绍SR中常用的数据集和图像质量评价指标。然而数据集和评价指标对SR网络的训练和重建图像的评价具有重要作用,仅对常用的方法进行介绍是不够的。同时,只有少量文献<sup>[37,38]</sup>简单提及了SR的相关挑战赛,而挑战赛恰恰是SR发展趋势的一种体现。随着SR方法的逐年更新迭代,前期的综述文献已无法涵盖最新的研究成果,也无法使读者了解到更多的数据集信息及相关图像质量评价指标,且仅从网络模型的发展情况分析SR的发展趋势是不全面的。因此本文在前人的基础上,首先以网络模型的设计、训练、测试为逻辑思路介绍图像SR的相关知识,完善并丰富数据集构建方式、网络模型基本框架以及图像质量评价指标等相关内容;其次根据学习模式的不同将现有方法划分为监督式SR和无监督式SR,并根据模型的网络结构及设计策略,重点对监督式SR典型及最新的研究成果加以评述,力求系统和全面地介绍基于深度学习的SISR方法;最后从数据集构建方式、网络模型研究进展及SR挑战赛等角度分析基于深度学习的图像SR未来的发展趋势,以促进基于深度学习的SISR技术今后的发展及应用。

## 2 图像超分辨率重建相关知识

图像SR旨在从低分辨率退化图像中恢复出相应的高分辨率图像。通常,低分辨率图像满足式(1)所示的退化过程:

$$I_{LR} = D(I_{HR}; \theta) \quad (1)$$

其中, $I_{LR}$ 和 $I_{HR}$ 分别表示LR图像和HR图像, $D$ 表示退化

函数,  $\theta$  表示退化过程的参数. 根据  $\theta$  是否已知, 可将图像 SR 分为退化已知的非盲超分辨率重建方法和退化未知的盲超分辨率重建方法, 其中盲超分辨率重建方法主要应用于真实世界的图像超分辨率重建.

图像超分辨率重建是图像退化过程的逆过程, 可利用低分辨率图像中的信息重建出对应的高分辨率图像, 即

$$I_{\text{HR}} = F(I_{\text{LR}}; \beta) \quad (2)$$

其中,  $F$  表示超分辨率重建模型,  $\beta$  表示超分辨率重建模型的参数.

## 2.1 数据集构建方式

数据集作为 SR 网络的主要数据来源, 可用于训练、验证及测试, 其中训练数据集作为网络端到端学习的重要数据来源, 对网络性能的提升有着重要的作用. 一个高质量、多数量、大范围的图像数据集能够在很大程度上提升网络性能. 根据 SISR 网络模型是否使用匹配的低分辨率-高分辨率(LR-HR)图像对进行训练, 可以将其分为监督式 SISR 和无监督式 SISR. 监督式 SISR 需要用匹配的 LR-HR 图像对训练网络, 因此训练数据集的构建对监督式 SISR 至关重要.

现有的数据集主要分为两种类型. 一种是只采集 HR 图像的数据集, 如 DIV2K<sup>[40]</sup>, DIV8K<sup>[41]</sup> 等数据集, 对于此类数据集可采用不同的退化方式获取相应的 LR 图像, 从而构造匹配的 LR-HR 训练图像对, 以这种方式获得的训练数据集一般被称为合成数据集. 另一种则是直接采集同一场景不同分辨率的图像, 从而获取 LR-HR 图像对的数据集, 如 RealSR<sup>[42]</sup>, DRealSR<sup>[43]</sup> 等数据集, 这类训练数据集一般被称为真实数据集. 目前主要有 3 种方法用于真实数据集的构建, 包括基于焦距调整的方法、基于硬件分箱的方法和基于光束分束器的方法<sup>[39]</sup>. 相较于合成数据集, 真实数据集通常具有更真实的退化过程, 适用于真实场景图像的超分辨率重建. 但真实数据集的构建存在一定难度, 如真实数据集直接采集到的 LR-HR 图像对通常是不匹配的, 需要进行严格的配准操作才能得到具有相同视场且可用于训练的 LR-HR 图像对, 且真实世界图像的退化核会随着景深的变化而变化, 通常是不均匀的, 往往需要根据实际情况采取不同的策略再进行超分辨率重建.

由于难以获取同一场景下成对的 LR 图像和 HR 图像, 所以现有数据集大多只采集 HR 图像, 再通过不同的退化方式得到相应的 LR 图像, 以构造合成数据集用于训练 SISR 网络模型. 然而 LR 图像的实际退化过程未知且复杂, 易受到模糊、噪声、下采样、图片压缩等因素的影响, 难以对其进行准确的定义. 因此, 在不同的 SISR 网络模型中使用的退化方式没有一个统一的标准. 根据现有 SISR 网络模型构造合成数据集时所采取

退化方式的不同, 可总结出以下几种退化模型.

### (1) 简单退化模型

简单退化模型通过对 HR 图像进行简单的下采样操作得到相应的 LR 图像, 如式(3)所示:

$$I_{\text{LR}} = (I_{\text{HR}}) \downarrow_s, s \in \theta \quad (3)$$

其中,  $\downarrow_s$  表示尺度因子为  $s$  的下采样运算. 以往的 SISR 网络模型大多采用理想的双三次下采样进行退化以获取 LR 图像. 然而简单退化模型获得的 LR 图像与实际的退化过程存在较大差异, 不仅难以应用于真实场景的图像 SR, 也难以处理与双三次下采样具有不同退化空间的图像.

### (2) 一般退化模型

一般退化模型对下采样、模糊和噪声等退化因素加以考虑, 与简单退化模型相比, 退化过程更接近实际场景, 如式(4)所示:

$$I_{\text{LR}} = (I_{\text{HR}} \otimes k) \downarrow_s + n_\delta, \{k, s, \delta\} \in \theta \quad (4)$$

其中,  $k$  表示模糊核,  $\otimes$  表示卷积操作,  $n$  表示噪声, 通常设置为标准差为  $\delta$  的加性高斯白噪声.

SRMD (SR network for Multiple Degradations)<sup>[44]</sup> 等网络模型验证了一般退化模型的有效性, 但其仍与图像的真实退化过程存在一定差异, 且退化范围无法有效覆盖实际场景中的各种退化, 因此大规模退化模型应运而生.

### (3) 大规模退化模型

大规模退化模型是在一般退化模型或其变体的基础上对各种退化因素进行扩展, 考虑更真实和更复杂的退化过程, 从而获取具有更准确及更大范围退化空间的 LR 图像. 大规模退化模型旨在通过更准确的模糊核估计等方式扩大图像退化空间来模拟图像的真实退化过程以获取相应的 LR 图像, 因此适用于真实图像的 SR.

SFTMD (Spatial Feature Transform for Multiple Degradations)<sup>[45]</sup>, DAN (Deep Alternating Network)<sup>[46]</sup> 等网络模型通过有效的模糊核估计, 使其设计的大规模退化模型更有利于真实图像的 SR. BSRGAN (Blind SRGAN)<sup>[47]</sup> 等网络模型在一般退化模型的基础上设计了一种更加复杂且实用的大规模退化模型, 对更复杂的退化模糊、下采样和噪声等退化因素加以考虑, 从而构造合成数据集用于训练. 而 Real-ESRGAN 网络模型<sup>[48]</sup> 则是在经典退化模型(式(5))的基础上, 对模糊、下采样、噪声和 JPEG 压缩等退化元素加以考虑, 通过“ $n$  阶”退化过程(式(6))建模以扩大退化空间, 每个退化过程采用不同参数的经典退化模型, 从而合成具有更加真实退化过程的 LR 图像, 极大地提升了网络重建质量.

$$I_{\text{LR}} = \left[ (I_{\text{HR}} \otimes k) \downarrow_s + n_\delta \right]_{\text{JPEG}} \quad (5)$$

$$I_{LR} = D^n(I_{HR}) = (D_n \circ \dots \circ D_2 \circ D_1)(I_{HR}) \quad (6)$$

#### (4) 无监督式退化模型

无监督式退化模型通过生成对抗网络(Generative Adversarial Networks, GAN)等无监督的方式模拟图像的退化过程,获取相应的LR图像.相比简单退化模型和一般退化模型,无监督式退化模型能够利用生成对抗网络的对抗博弈性使网络更好地模拟图像真实的退化过程,因此主要应用于真实场景的图像SR.

KernelGAN(Kernel estimation using an internal-GAN)<sup>[49]</sup>,DSGAN(Down-Sample GAN)<sup>[50]</sup>等网络模型都是通过GAN以无监督的方式得到与原始HR图像有相同分布的LR图像,从而构造合成数据集进行真实图像的SR.FCA(Frequency Consistent Adaptation)<sup>[51]</sup>则是一种频率一致性自适应方法,通过所提出的自适应生成器以无监督的方式估计图像的退化过程,从而得到与真实场景图像具有频率一致性的LR图像,用于SR网络的训练.

目前,已有很多可用于图像SR的数据集,这些数据集在图像质量、数量、范围和分辨率等方面都存在一定差异,可以为不同的图像SR任务提供数据支持.表1对图像SR中常用的数据集进行总结<sup>[40-43, 47, 49, 52-68]</sup>,以便了解数据集的相关内容并选择合适的数据集用于SR网络的训练、验证和测试.

## 2.2 网络模型基本框架

以不同退化方式或采集方式得到合成数据集或真实数据集后,即可对网络模型进行相应的训练.虽然现有SISR的网络模型之间差异较大,但本质上可以将它们看成是网络模型框架、网络设计策略和网络学习策略等模块的不同组合<sup>[37]</sup>,从而简化复杂的网络结构.

网络模型框架是SISR网络模型中最基本的模块,根据上采样层在网络模型中位置的不同,可将模型基本框架划分为4种类型:预上采样、后上采样、渐进式上

采样、迭代式上下采样,如图1所示.上采样指的是将原始的LR图像转换为HR图像的操作,它作为SR中必不可少的环节,在网络模型框架中占有重要地位.

### 2.2.1 预上采样模型框架

预上采样模型框架中的上采样层位于网络前端的图像预处理环节,如图1(a)所示.该框架通常使用传统的基于插值的上采样方法,如线性插值、双三次插值等,最常用的是双三次插值的上采样方法.

早期的SRCNN<sup>[69,70]</sup>,VDSR<sup>[55]</sup>,DRCN<sup>[71]</sup>等网络模型都是使用预上采样模型框架,先将LR图像上采样为所需尺寸的HR图像,再将其输入卷积神经网络进行SR,以恢复HR图像的更多细节.预上采样模型框架的结构简单,能进行任意尺度因子图像的SR.但它先对LR图像进行上采样后再输入网络进行训练的操作,使网络模型的计算在高维空间中进行,显著增加了计算复杂度,时间及空间成本也随之增加,因此在近年网络模型中的使用逐渐减少.

### 2.2.2 后上采样模型框架

为了避免在高维空间中计算带来的影响,提高网络计算效率,后上采样模型框架将上采样层放置于网络末端,如图1(b)所示,直接将LR图像输入卷积神经网络中,在低维空间形成映射,最后在网络末端进行上采样后输出重建的HR图像.

后上采样模型框架在网络末端的上采样层通常使用的是基于学习的上采样方法,如转置卷积(又称反卷积)、亚像素卷积等,以实现端到端的自动学习.此外,元上采样(meta-upscale)<sup>[72]</sup>等特殊的上采样方法可用于任意尺度因子(1~4倍,步长为0.1)的SR.在后上采样模型框架的影响下,FSRCNN<sup>[56]</sup>,ESPCN<sup>[73]</sup>,BTSRN<sup>[74]</sup>和RNAN<sup>[75]</sup>等网络模型实现了网络加速并取得了较好的网络性能.

后上采样模型框架在低维空间计算的方式,能够在维持或提升网络性能的同时,降低网络计算量和空

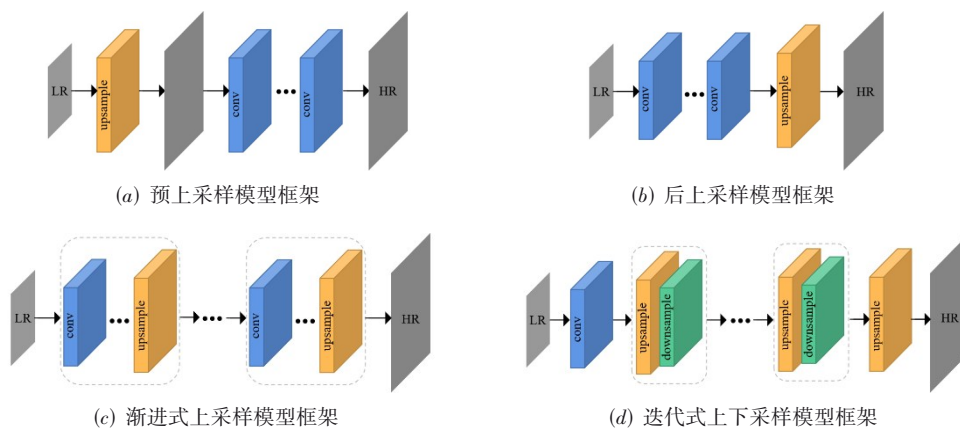


图1 图像超分辨率重建网络模型基本框架

表 1 图像超分辨率重建常用数据集概述

数据集名称	图像数量/张	合成/真实数据集 (训练集/验证集/测试集)	图像格式	图像内容
T91 <sup>[52]</sup>	91	合成数据集(训练集)	PNG	包括动植物、人、车等局部纹理图像
Timofte <sup>[53]</sup>	110	合成数据集(训练集)	PNG	包括 T91, Set5 和 Set14 三个数据集的图像
291-images <sup>[54,55]</sup>	291	合成数据集(训练集)	PNG, JPG	包括 T91 和 BSD 两个数据集的图像
General-100 <sup>[56]</sup>	100	合成数据集(训练集)	BMP	包括建筑纹理、动植物、人、食物、日用品等图像
MSCOCO <sup>[57]</sup>	328 000	合成数据集(训练集)	JPG	包括 91 种易识别的物体类型在自然环境中的复杂日常场景的图像
ImageNet <sup>[58]</sup>	3 200 000	合成数据集(训练集)	JPG	包括动植物、交通工具、家具、乐器、地质构造、工具等图像
DIV2K <sup>[40]</sup>	1 000	合成数据集(800 张训练集、100 张验证集、100 张测试集)	PNG	包括人、手工艺品、环境(城市、村庄)、动植物、自然风景等图像
Flickr2K <sup>[59]</sup>	2 650	合成数据集(训练集)	PNG	包括人、车、动植物、食物、建筑和风景等图像
DF2K <sup>[60]</sup>	3 450	合成数据集(训练集)	PNG	包括 DIV2K 和 Flickr2K 两个数据集的图像
DPED <sup>[61]</sup>	5 827	真实数据集(5 614 张训练集、113 张验证集、100 张测试集)	PNG	包括各种道路交通场景(建筑、植物、道路等)等图像
OutdoorScene(OST) <sup>[62]</sup>	10 624	合成数据集(10 324 张训练集、300 张测试集,即 OST300)	PNG	包括动物、建筑、草、山、植物、天空和水等 7 类纹理丰富的图像
DIV8K <sup>[41]</sup>	1 504	真实数据集(1 304 张训练集、100 张验证集、100 张测试集)	PNG	包括人、动物、建筑和风景等各种场景和物体的图像
RealSR <sup>[42]</sup>	595	真实数据集(550 张训练集、45 张测试集)	PNG	包括建筑、风景、动植物、海报、物体等纹理丰富的室内外场景图像
DRealSR <sup>[43]</sup>	884(×2) 783(×3) 840(×4)	真实数据集(对于尺度因子 2, 3, 4 分别用 83, 84, 93 张图像进行测试,其余用于训练)	PNG	包括广告海报、植物、办公室、建筑物等室内外场景的图像
Set5 <sup>[63]</sup>	5	合成数据集(测试集)	PNG	包括婴儿、鸟、蝴蝶、头部和女士等 5 张图像
Set14 <sup>[64]</sup>	14	合成数据集(测试集)	PNG	包括动植物、风景(船和桥)、PPT 和人等图像
Berkeley Segmentation Dataset(BSD)/BSD500 <sup>[65]</sup>	500	合成数据集(测试集)	JPG	包括动物、建筑、食物、风景、人和植物等图像. BSD100 和 BSD300 分别是 BSD500 中常用的 100 或 300 张图片
Urban100 <sup>[66]</sup>	100	合成数据集(测试集)	PNG	包括不同类型的建筑图像
Manga109 <sup>[67]</sup>	109	合成数据集(测试集)	PNG	包括来自日本漫画书的图像
PIRM <sup>[68]</sup>	200	合成数据集(100 张验证集、100 张测试集)	PNG	包括人、物、环境、植物、自然风景等图像
DIV2KRK <sup>[49]</sup>	100	合成数据集(测试集)	PNG	包括对 DIV2K 验证集的 100 张图像进行更复杂退化操作的图像
DIV2K4D <sup>[47]</sup>	400	合成数据集(测试集)	PNG	包括对 100 张 DIV2K 验证集采取四种退化的图像

间复杂度,并提高网络计算效率.但对大尺度因子的学习存在一定难度,且无法满足单一模型的多尺度因子图像 SR 的需求,对不同尺度因子的图像需要训练不同的网络模型.

### 2.2.3 渐进式上采样模型框架

渐进式上采样模型框架如图 1(c)所示,是以级联的方式连接卷积神经网络,并通过多个上采样层逐步重建得到最终的 HR 图像. LapSRN<sup>[76]</sup>是典型的采用渐进式上采样模型框架的网络模型,它将网络结构分成三级,每级进行两倍的上采样操作,通过逐级上采样

实现两倍、四倍及八倍的超分辨率重建结果. MS-LapSRN<sup>[77]</sup>, LP-KPN<sup>[42]</sup>和 E-ProSRNet<sup>[78]</sup>等网络模型也采用这种框架,实现了单一模型的多尺度因子图像 SR.

渐进式上采样模型框架采用逐步上采样的方式将困难的大尺度因子 SR 任务分解为多个简单的小尺度因子 SR 任务,极大地降低了学习难度,且在不引入过多时间和空间成本的情况下,能够满足单一模型的多尺度因子 SR 的需求.但存在模型结构设计复杂、训练稳定性差等问题.

### 2.2.4 迭代式上下采样模型框架

迭代式上下采样模型框架如图1(d)所示,该框架在网络中交替使用上、下采样层,再通过迭代反向投影不断改进重建图像细节,从而得到最终的重建图像。

DBPN<sup>[79]</sup>(Deep Back-Projection Networks)网络模型是首个采用该框架的方法,它利用迭代的上下采样层的误差反馈机制来指导网络重建,获得最终的HR图像。相较于单向前馈神经网络直接学习输入图像到目标空间非线性映射的方法,DBPN网络模型将学习过程分成多个阶段,并为每个阶段的投影误差提供误差反馈机制,使模型具有自校正的过程用于修正重建细节,从而获得更好的重建结果。同时DBPN顺应深度学习发展趋势,将其扩展为多个变体以提升网络性能<sup>[80]</sup>。此外,DSRN<sup>[81]</sup>,SRFBN<sup>[82]</sup>等网络模型也在网络中交替使用上、下采样层,并通过不同的反馈机制改善HR图像细节。

相比于其他模型框架,迭代式上下采样模型框架能够更好地挖掘LR-HR图像对之间的深层关系,从而获得更多图像细节,构建更高质量的重建图像。但迭代式上下采样模型框架的网络结构较为复杂、发展还不成熟,仍需进一步探索。

### 2.3 图像质量评价指标

对于完成训练的网络模型,可通过不同的图像质量评价指标评估重建图像质量,验证网络模型有效性,测试网络模型性能。根据评价主体不同,可以将SR的图像质量评价指标分为主观评价指标和客观评价指标。

#### 2.3.1 主观评价指标

主观评价指标是由评价人员根据自己的主观感受对图像质量进行评价的一种方式。根据是否有真实HR图像作为标准参考图像,可以将其分为绝对主观评价指标和相对主观评价指标。

绝对主观评价指标,如平均意见排名(Mean Opinion Rank, MOR)<sup>[83]</sup>,是在无标准参考图像的情况下,评价人员根据自己的主观视觉感受及设定好的评价尺度对几种SR方法的重建图像质量进行排名从而计算得出的。

相对主观评价指标,如平均意见得分(Mean Opinion Score, MOS)<sup>[84]</sup>,是在有标准参考图像的情况下,评价人员将不同SR方法获得的重建结果与标准的参考图像进行对比,并将图像进行组内对比,最后根据评价尺度对这组图像进行评分。

表2所示是主观评价指标两种方法的评价尺度。可以发现,这两种主观评价指标的评价尺度都是根据评价人员的主观感受进行衡量的,评价结果符合人类

视觉感受,因此主观评价指标是最直接、最有效的评价方法。但主观评价指标易受评价人员的主观感受及各种因素的影响,有较大的不确定性,可重复性、实时性也较差,且评价过程需要耗费大量的时间、人力、物力、财力等,在实际使用过程中存在一定困难,因此难以被广泛应用。

表2 主观评价指标的评价尺度

级别	绝对主观评价指标	相对主观评价指标
1	很好(第1名)	一组中最好(5分)
2	较好(第2名)	比该组的平均水平好(4分)
3	一般(第3名)	该组中平均水平(3分)
4	较差(第4名)	比该组的平均水平差(2分)
5	很差(第5名)	一组中最差(1分)

#### 2.3.2 客观评价指标

客观评价指标是指通过不同的数学模型和算法来评估图像质量的方法<sup>[85,86]</sup>,具有简单、高效、可重复性强等优点,因此SR中通常使用客观评价指标对重建图像进行质量评价。客观评价指标根据是否需要真实的HR图像作为参考图像,可以大致分为全参考型和无参考型两种。

全参考型的客观评价指标是将重建的HR图像与真实HR图像进行比较计算得出的,一般用于监督式SR的图像评估。该评价指标包括峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、结构相似度(Structural SIMilarity, SSIM)、信息保真度准则(Information Fidelity Criterion, IFC)、学习感知图像块相似度(Learned Perceptual Image Patch Similarity, LPIPS)等,其中PSNR和SSIM是最常用的客观评价指标。

无参考型的客观评价指标无可参考的真实HR图像,因此常用于无监督式SR的图像评估。该评价指标包括自然图像质量评价(Natural Image Quality Evaluator, NIQE)、基于感知的图像质量评价(Perception-based Image Quality Evaluator, PIQUE)、无参考质量指标(No-Reference Quality Metric, NRQM)、感知指数(Perception Index, PI)等。

##### (1) 峰值信噪比

PSNR是指通过计算重建HR图像与真实HR图像对应像素点之间的误差,从而客观地评估重建图像失真程度的指标。PSNR值主要由均方误差(Mean Square Error, MSE)决定,MSE表达式如式(7)所示:

$$MSE = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (I_{HR}(i,j) - I_{SR}(i,j))^2 \quad (7)$$

其中, $M$ 和 $N$ 分别表示图像长、宽的像素数; $I_{HR}(i,j)$ 和 $I_{SR}(i,j)$ 分别表示真实HR图像与重建HR图像在空间位置 $(i,j)$ 处的像素值。

PSNR表达式如式(8)所示,单位为分贝(dB):

$$\text{PSNR} = 10 \times \log_{10} \left( \frac{\text{MAX}^2}{\text{MSE}} \right) = 20 \times \log_{10} \left( \frac{\text{MAX}}{\sqrt{\text{MSE}}} \right) \quad (8)$$

其中,MAX为 $I_{\text{HR}}(i,j)$ 图像的最大像素值,对于8比特(bit)精度的图像,MAX取值为255.

由式(7)和式(8)可以看出,最小化MSE损失函数(即 $L_2$ 损失函数)相当于最大化PSNR评价指标.PSNR值的取值范围为 $[0, +\infty)$ ,其值越大,则表示重建HR图像与真实HR图像之间的像素误差越小,重建HR图像相对于真实HR图像的失真越少,重建图像的质量越好.PSNR通过逐像素计算的方法简单、高效,是SR领域最常用的图像质量评价指标.但PSNR仅从数学角度计算图像之间的差异,未从本质上考虑人类视觉系统(Human Visual System, HVS)特性,故PSNR计算结果反映的图像质量情况与人类主观视觉感受的图像质量情况存在一定差异,不能完全、准确地反映重建图像的感知质量.

## (2) 结构相似度

SSIM是由Wang等人<sup>[87]</sup>于2004年提出的,从亮度、对比度和结构三个方面来衡量参考图像与失真图像之间结构相似性的方法.SSIM主要由图像间的亮度、对比度和结构三部分信息组成,且三者之间是相对独立的,即亮度和/或对比度的信息变化不会影响图像的结构信息.

SSIM的表达式如式(9)所示:

$$\text{SSIM} = [l(I_{\text{HR}}, I_{\text{SR}})]^\alpha \cdot [c(I_{\text{HR}}, I_{\text{SR}})]^\beta \cdot [s(I_{\text{HR}}, I_{\text{SR}})]^\gamma \quad (9)$$

其中, $\alpha, \beta, \gamma$ 为权重参数,分别用于调整 $l(I_{\text{HR}}, I_{\text{SR}})$ ,  $c(I_{\text{HR}}, I_{\text{SR}})$ ,  $s(I_{\text{HR}}, I_{\text{SR}})$ 三个分量的相对重要性,且 $\alpha > 0, \beta > 0, \gamma > 0$ . $l(I_{\text{HR}}, I_{\text{SR}})$ ,  $c(I_{\text{HR}}, I_{\text{SR}})$ ,  $s(I_{\text{HR}}, I_{\text{SR}})$ 分别表示亮度、对比度、结构三个分量,它们的表达式分别如式(10)~(12)所示:

$$l(I_{\text{HR}}, I_{\text{SR}}) = \frac{2\mu_{I_{\text{HR}}}\mu_{I_{\text{SR}}} + C_1}{\mu_{I_{\text{HR}}}^2 + \mu_{I_{\text{SR}}}^2 + C_1} \quad (10)$$

$$c(I_{\text{HR}}, I_{\text{SR}}) = \frac{2\sigma_{I_{\text{HR}}}\sigma_{I_{\text{SR}}} + C_2}{\sigma_{I_{\text{HR}}}^2 + \sigma_{I_{\text{SR}}}^2 + C_2} \quad (11)$$

$$s(I_{\text{HR}}, I_{\text{SR}}) = \frac{\sigma_{I_{\text{HR}}I_{\text{SR}}} + C_3}{\sigma_{I_{\text{HR}}}\sigma_{I_{\text{SR}}} + C_3} \quad (12)$$

其中, $\mu_{I_{\text{HR}}}, \mu_{I_{\text{SR}}}$ 分别表示 $I_{\text{HR}}$ 和 $I_{\text{SR}}$ 的均值; $\sigma_{I_{\text{HR}}}, \sigma_{I_{\text{SR}}}$ 分别表示 $I_{\text{HR}}$ 和 $I_{\text{SR}}$ 的标准差; $\sigma_{I_{\text{HR}}I_{\text{SR}}}$ 表示 $I_{\text{HR}}$ 和 $I_{\text{SR}}$ 的协方差; $C_1, C_2$ 和 $C_3$ 是为了避免计算中出现不稳定而添加的小常数.

特别地,当 $\alpha = \beta = \gamma = 1$ ,且 $C_3 = C_2/2$ 时,SSIM可以表示为式(13),该形式是SR图像质量评价时最常使用的形式:

$$\text{SSIM} = \frac{(2\mu_{I_{\text{HR}}}\mu_{I_{\text{SR}}} + C_1)(2\sigma_{I_{\text{HR}}I_{\text{SR}}} + C_2)}{(\mu_{I_{\text{HR}}}^2 + \mu_{I_{\text{SR}}}^2 + C_1)(\sigma_{I_{\text{HR}}}^2 + \sigma_{I_{\text{SR}}}^2 + C_2)} \quad (13)$$

SSIM值的取值区间为 $[0, 1]$ ,其值越大,表明图像质量越好.由于考虑了HVS特性,SSIM的计算结果相比PSNR指标能够更有效地反映重建HR图像的感知质量,与人类感知结果具有较好的一致性,因此SSIM被广泛用于SR图像的质量评价.在SR质量评价中通常使用PSNR和SSIM共同衡量重建图像质量.

## (3) 信息保真度准则

2005年,Sheikh等人<sup>[88]</sup>提出了一种基于自然场景统计信息对图像质量进行评价的方法,即IFC评价指标,通过结合自然场景模型和失真模型的统计模型,量化参考图像与测试图像之间的相互信息,从而量化图像感知质量,并以此衡量重建图像质量的优劣.

IFC的表达式如式(14)所示,该式量化了参考图像和失真图像之间共享的统计信息:

$$\text{IFC} = \sum_{k \in \text{subbands}} I(C^{N_k, k}; D^{N_k, k} | S^{N_k, k}) \quad (14)$$

其中, $C^{N_k, k}, D^{N_k, k}, S^{N_k, k}$ 分别表示第 $k$ 个子带RF $C^k, D^k, S^k$ 的 $N_k$ 系数.

IFC是一个保真指标,而非失真指标,其值的取值范围为 $[0, +\infty)$ ,IFC值越大,图像保真度越高,重建图像质量越好.IFC在信息提取过程中使用感知质量建模,因此IFC计算结果与人类视觉感知结果具有较好的一致性.

## (4) 学习感知图像块相似度

2018年,Zhang等人<sup>[89]</sup>提出LPIPS感知评价指标,利用预训练的深度卷积神经网络提取参考图像与失真图像的特征,计算图像在深度特征空间上的 $L_2$ 距离,并评估图像间的感知相似度.

LPIPS的表达式如式(15)所示:

$$\text{LPIPS} = d(I_{\text{HR}}, I_{\text{SR}}) = \sum_l \frac{1}{H_l W_l} \sum_{h, w} \left\| w_l \odot (\hat{y}_{I_{\text{HR}}}^l - \hat{y}_{I_{\text{SR}}}^l) \right\|_2^2 \quad (15)$$

其中, $l$ 表示深度卷积神经网络的第 $l$ 层; $\hat{y}_{I_{\text{HR}}}^l, \hat{y}_{I_{\text{SR}}}^l$ 分别表示 $I_{\text{HR}}$ 和 $I_{\text{SR}}$ 从第 $l$ 层中提取特征,并在通道维度上进行单元归一化的结果,且 $\hat{y}_{I_{\text{HR}}}^l, \hat{y}_{I_{\text{SR}}}^l \in \mathbb{R}^{H_l \times W_l \times C_l}$ ;  $w_l$ 是用于缩放激活通道的向量, $w_l \in \mathbb{R}^{C_l}$ ,当 $w_l = 1 \forall l$ 时,相当于计算余弦距离.

LPIPS是根据人类感知进行训练的,因此LPIPS值可以较好地反映人们对图像的主观感受,LPIPS值越小,图像的感知质量越好.

## (5) 自然图像质量评价

Mittal等人<sup>[90]</sup>于2013年提出了基于空域特征的完全无参考型的评价指标NIQE,通过从失真图像提取的自然场景统计特征和从自然图像提取的感知质量特征的多元高斯模型(MultiVariate Gaussian model, MVG)拟合参数之间的距离来评估失真图像的质量.

NIQE 表达式如式(16)所示:

$$\begin{aligned} \text{NIQE} &= D(v_1, v_2, \Sigma_1, \Sigma_2) \\ &= \sqrt{\left( (v_1 - v_2)^T \left( \frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (v_1 - v_2) \right)} \end{aligned} \quad (16)$$

其中,  $v_1$  和  $\Sigma_1$  分别表示自然图像 MVG 模型的均值向量和协方差矩阵,  $v_2$  和  $\Sigma_2$  分别表示失真图像 MVG 模型的均值向量和协方差矩阵. NIQE 的值越小, 表示图像的质量越好.

### (6) 基于感知的图像质量评价

Venkatanath 等人<sup>[91]</sup>于 2015 年提出了一种无参考感知的图像质量评价指标 PIQUE, 基于测试图像的局部块级别特征, 估计给定测试图像中存在的失真量. PIQUE 考虑了人类视觉注意力等人类感知图像质量原则, 通过对输入图像进行预处理, 提取自然场景统计特征, 再对图像进行块级别分析以识别不同失真块的类型, 并分配不同分数, 最后合并块级别分数以确定整体图像质量.

PIQUE 表达式如式(17)所示:

$$\text{PIQUE} = \frac{\left( \sum_{k=1}^{N_{SA}} D_{sk} \right) + C_1}{(N_{SA} + C_1)} \quad (17)$$

其中,  $N_{SA}$  表示给定图像中空间活动块的数量,  $C_1$  是防止数值不稳定添加的一个正常数,  $D_{sk}$  是失真块所分配的失真量.

PIQUE 值的取值范围为  $[0, 1]$ , 其值越小, 图像质量越好. 当 PIQUE 值接近于 0 (0~0.3) 时, 表示图像质量良好; 当 PIQUE 值接近于 1 (0.5~1.0) 时, 表示图像质量较差; 若 PIQUE 值介于 0.3 到 0.5 之间, 则可将其视为平均质量图像. PIQUE 的计算考虑了 HVS 特性, 因此该评价指标的评价结果接近人类感知质量的评价结果.

### (7) 无参考质量指标

Ma 等人<sup>[92]</sup>于 2017 年提出了一种无参考型评价指标 NRQM, 在一些文献中也将其简称为 Ma. NRQM 以重建的 HR 图像作为输入, 将频率域和空间域中计算的统计量作为输入图像的特征, 在单独的集成回归树中训练特征, 并利用线性回归模型对大量的视觉感知分数进行学习, 得到图像质量分数以评估输入图像的质量.

NRQM 的表达式如式(18)所示:

$$\text{NRQM} = \hat{y} = \sum_n \lambda_n \cdot \hat{y}_n \quad (18)$$

其中,  $\hat{y}$  表示最终的质量分数, 由 3 种类型特征的预测质量分数组成;  $\lambda_n$  表示不同类型特征的权重;  $\hat{y}_n$  表示不同类型特征的预测质量分数, 通过对  $T$  个回归树的输出进行平均得到,  $\hat{y}_n = \frac{1}{T} \sum_i^T p_i(x_n | \Theta)$  ( $x_n$  表示不同类型的低

级特征,  $n=1, 2, 3$ ;  $p_i$  表示森林中第  $i$  棵决策树的概率,  $t=1, 2, \dots, T$ ). NRQM 值越大, 表明重建图像的质量越好.

### (8) 感知指数

PI 是 Blau 等人<sup>[68]</sup>结合 NIQE 和 Ma 两个无参考客观评价指标提出的感知指标, 该指标联合量化了准确性与感知质量, 与主观评价指标具有高度的相关性.

PI 的表达式如式(19)所示:

$$\text{PI} = \frac{1}{2}((10 - \text{Ma}) + \text{NIQE}) \quad (19)$$

PI 值越小, 表明重建图像的感知质量越好.

## 3 监督式单帧图像超分辨率重建

目前大多数单帧图像超分辨率重建技术都是基于深度学习的方法进行研究的, 并取得了较为丰富的研究成果. 从早期基于卷积神经网络的超分辨率重建方法, 如 SRCNN<sup>[69]</sup>等, 到很有前景的基于生成对抗网络的超分辨率重建方法, 如 SRGAN<sup>[93]</sup>等, 再到近来大火于低级 (low level) 视觉任务界的基于 Transformer 的超分辨率重建方法, 如 IPT<sup>[94]</sup>等, 基于深度学习的方法在图像重建效果上实现了很大的提升. 图 2 给出了部分代表性 SR 算法的发展历程. 按照学习模式的不同, 可以将基于深度学习的 SISR 方法划分为监督式 SISR (图中黑色) 和无监督式 SISR (图中蓝色); 根据网络结构及图像重建效果的不同, 将监督式 SISR 进一步划分为基于卷积神经网络 (Convolutional Neural Network, CNN) 的方法 (图中轴线上方不加粗)、基于生成对抗网络 (GAN) 的方法 (图中轴线下方) 和基于 Transformer 的方法 (图中轴线上方加粗). 本节首先从监督式 SISR 角度出发, 对代表性的算法进行评述.

### 3.1 基于卷积神经网络的单帧图像超分辨率重建

卷积神经网络 (CNN) 是一种通过传统梯度下降法训练并学习图像特征的前馈神经网络, 局部连接、权重共享等特性使其相较于其他神经网络能够更好地学习与表达图像特征, 因此被广泛应用于图像处理领域<sup>[95, 96]</sup>. 基于卷积神经网络的单帧图像超分辨率重建方法主要采用卷积神经网络的结构, 以 PSNR 最大化为目标, 旨在使网络获得更真实的细节、更好的保真度, 即 PSNR, SSIM 等客观指标的提升. 虽然基于 CNN 的方法在评价指标方面表现较好, 但其重建图像往往过于平滑, 无法带来很好的感知效果.

现有 SISR 网络模型是在不同模型框架的基础上应用不同的网络设计策略及学习策略构建的. 根据网络模型中设计策略的不同, 可以进一步将近年来较为经典以及新提出的网络模型分为以下几类: 基于浅层卷积神经网络 (Shallow Convolutional Neural Network, S-CNN) 的方法、基于残差网络 (Residual Network, ResNet)

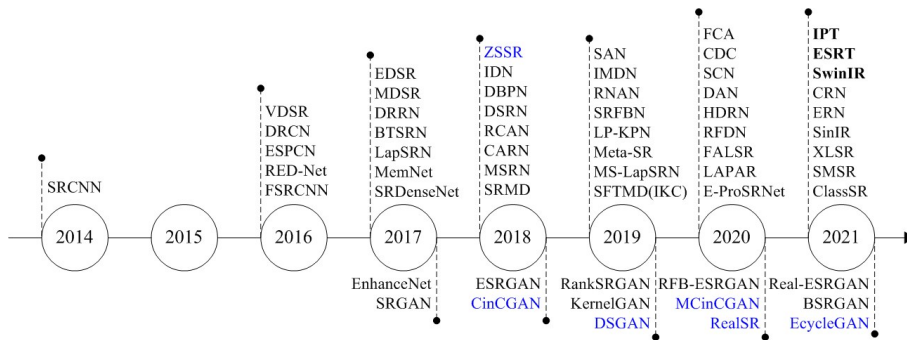


图2 基于深度学习的SISR部分网络模型的发展时间轴

的方法、基于递归神经网络(Recurrent Neural Network, RNN)的方法、基于密集卷积网络(Dense convolutional Network, DenseNet)的方法、基于注意力机制(Attentional Mechanism, AM)的方法和基于轻量化网络(Lightweight Network, LN)的方法。

### 3.1.1 基于浅层卷积神经网络的方法

2014年,Dong等人<sup>[69]</sup>首次将深度学习应用于图像SR领域,提出了第一个基于卷积神经网络的SISR网络模型SRCNN(Super-Resolution CNN)。受基于稀疏编码的图像SR方法的启发对网络结构进行设计,得到由图像块特征的提取与表示层、特征的非线性映射层及重建层等简单的三层卷积神经网络构成的SRCNN网络(图3),实现LR图像到HR图像之间的端到端映射。虽然SRCNN相较于传统的SR方法在速度和重建质量上都有所提升,但预上采样的模型框架给它带来了计算复杂、训练收敛速度慢等问题,同时SRCNN还存在结构简单、难以充分利用图像上下文信息等不足之处。

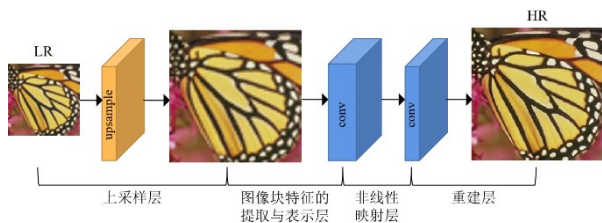


图3 SRCNN网络结构

针对SRCNN计算成本高、难以实时应用的问题,Dong等人<sup>[56]</sup>又于2016年对其进行改进及加速,提出了FSRCNN(Fast SRCNN)网络模型。FSRCNN由特征提取层、收缩层、非线性映射层、扩张层以及反卷积层构成,它的主要改进是将SRCNN中的预上采样模型框架替换为后上采样模型框架,通过网络末端的反卷积层(图4)实现上采样,以解决计算复杂度高等问题。另外,还通过改变特征维数、共享映射层参数等操作,提高网络计算效率,提升重建图像质量。

为降低计算复杂度、提升网络计算效率,Shi等

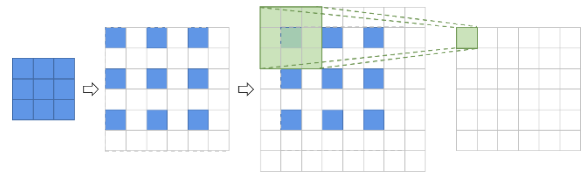


图4 反卷积层卷积过程

人<sup>[73]</sup>也于2016年提出了另一种快速、高效的SR网络模型ESPCN(Efficient Sub-Pixel CNN),与FSRCNN一样采用后上采样模型框架,但ESPCN使用的是亚像素卷积层(图5)对图像进行上采样。ESPCN网络由包含两个卷积层的隐藏层和一个亚像素卷积层构成。它先从隐藏层中提取LR输入图像的特征,再从亚像素卷积层中获取重建的HR图像,从而实现快速高效的端到端学习,得到比FSRCNN网络模型更好的重建效果。

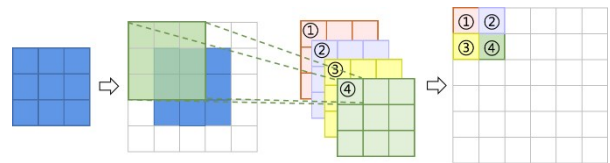


图5 亚像素卷积层卷积过程

尽管早期SRCNN,FSRCNN,ESPCN等浅层卷积神经网络模型的网络层数不超过5层,网络结构相对简单,没有使用过多的网络设计策略,但其重建效果相较于传统的图像SR方法得到了一定提升,对基于深度学习的图像SR的发展具有开创性作用。

### 3.1.2 基于残差网络的方法

为了提取更多图像特征,提升网络模型性能,最直接的方式是通过增加网络的深度或宽度来增多网络参数量。然而单纯地加深、加宽网络,随之出现的是梯度消失、梯度爆炸和网络退化等问题。对于梯度问题,通常使用批归一化或正则化等操作就能够很大程度地解决,但退化问题却仍然存在。对此,He等人<sup>[97]</sup>提出了残差网络用于解决深层网络带来的梯度和退化问题,其结构如图6所示,即在普通网络(plain network)中加入快捷连接(shortcut connections)/跳跃连接(skip connec-

tions)操作,使普通网络经过残差学习成为对应的残差网络.

ResNet具有全局残差学习和局部残差学习两种残差学习方式,主要区别在于全局残差学习是对网络模型的输入和输出之间进行快捷连接,局部残差学习则是对网络模型内部不同深度的层之间进行快捷连接. ResNet不仅能够有效提取图像细节信息,还能解决过深的网络层带来的梯度及退化问题,同时大量减少参数数量和训练时间,因此在深层网络中被广泛应用.

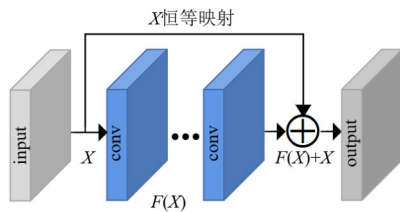


图6 残差网络应用于SR

受 ImageNet 分类比赛中深度卷积神经网络 VGG-net 的启发, Kim 等人<sup>[55]</sup>首次将 ResNet(图7(a))应用于图像SR中,于2016年提出了具有20个权重层的深度SISR网络模型VDSR(Very Deep CNN for SR). 考虑到低、高分辨率图像之间的低频信息在很大程度上是相似的,他们利用残差学习的思想,在VDSR中学习低、高分辨率图像之间高频信息的残差,从而减少训练时间,提高训练速度. 另外,他们还将ResNet与提高学习率、自适应梯度裁剪等策略相结合,使VDSR深度网络模型的训练过程更加稳定. 此后,ResNet被广泛应用于图像SR网络模型中.

同年, Mao 等人<sup>[98]</sup>提出了深度全卷积编解码网络RED-Net(Residual Encoder-Decoder Network),不仅可以进行图像SR,还能够解决图像去噪等其他图像恢复任务. 受高速公路网络(highway network)和深度ResNet的启发,RED-Net模型在对称的卷积层和反卷积层之间添加了跳跃连接,以解决深层网络带来的各种问题. 卷积层用于提取输入图像特征,反卷积层则利用跳跃连接直接与卷积层所提取的特征结合,从而更好地恢复出图像细节信息,也使得训练深层网络变得更加容易.

2017年, Lim 等人<sup>[99]</sup>在SRResNet<sup>[93]</sup>结构(图7(b))的基础上进行优化,利用ResNet的思想构建增强型深度SR网络模型EDSR(Enhanced Deep SR),创新性地去除了SRResNet结构中的批归一化(Batch Normalization, BN)层(图7(c)). BN层在SR中对图像特征进行归一化的操作,会破坏图像信息,影响图像质量,因此去除BN层不仅可以改善图像质量,还能够训练期间节省约40%的内存,实现同等计算资源条件下更深层网络的构建. 同时, Lim 等人采取残差缩放等方法解决深度网

络训练不稳定的问题,从而实现深层网络重建图像质量的显著提升. 为解决EDSR网络模型只能处理特定单尺度因子SR的问题, Lim 等人对EDSR进行扩展,又提出了多尺度深度超分辨率网络模型MDSR(Multi-scale Deep SR)用于单一模型的多尺度因子SR,而MDSR不仅能够大量减少参数数量与训练时间,还能实现与EDSR相当的网络性能.

2018年, Li 等人<sup>[100]</sup>指出一些网络模型性能的提高不是来自模型结构的改变,而是使用了一些未知的训练技巧,同时,大多数网络模型通过加深网络层数提升网络性能的方法难以充分利用LR图像特征,且无法使用单一模型处理多尺度任务. 为此, Li 等人在不使用任何训练技巧的情况下,提出了多尺度残差网络MSRN(Multi-Scale Residual Network)用于实现单一模型的多尺度SR任务,旨在通过网络结构的改变来提升网络性能. MSRN引入多尺度残差块(Multi-Scale Residual Block, MSRB)作为网络的基本构建模块,如图7(d)所示,MSRB将ResNet与不同尺度的卷积核结合,以获取不同尺度的图像特征,得到局部多尺度特征,最后将其与全局特征融合,充分利用LR图像特征,得到最终的重建图像.

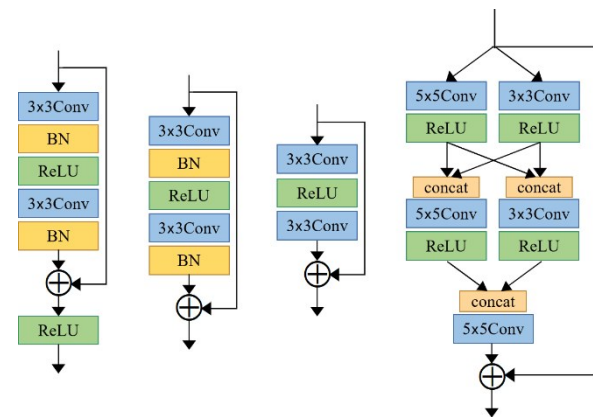


图7 不同网络模型的残差块

2021年, Lan 等人<sup>[101]</sup>指出大多数基于CNN的网络模型没有充分利用底层特征以致网络性能相对较差,因此提出了两个能够有效提取图像特征的网络模型用于图像SR. 一个是包含多个局部共享组的级联残差网络CRN(Cascading Residual Network),该网络通过级联机制促进特征融合和梯度传播,以更有效地提取图像特征. 另一个是具有双全局路径结构的增强残差网络ERN(Enhanced Residual Network),该网络通过双全局路径从原始输入中捕获长距离空间特征,以实现更强大的特征表达. 通过结构的改进,CRN和ERN网络模型能够以更少的参数数量实现与EDSR相当甚至更好的

网络性能.

ResNet 不仅能够通过局部或全局残差学习有效提取 SR 网络模型中低分辨率输入图像的特征信息,还可以解决过深网络带来的各种训练及梯度问题,因此许多 SR 网络模型都将残差学习的思想应用于网络结构中,以实现网络性能的提升,后期的很多 SR 网络模型也将残差学习思想与其他网络设计策略相结合,以获取更好的超分辨率重建效果.

3.1.3 基于递归神经网络的方法

RNN 结构通常由输入状态  $x$ 、输出状态  $y$  和循环状态  $s$  组成,如图 8(a)所示,可以将其按照时间顺序展开,以上一时刻( $t-1$ )的输出与当前时刻( $t$ )的输入同时作为当前网络的输入,从而得到当前时刻( $t$ )的输出,再不断迭代上述过程得到最终的输出<sup>[102]</sup>. 由于 RNN 的网络层主要用于记忆数据,而不是分层次处理,且每次迭代后新的图像信息都会被添加到每一层中,所以 RNN 在无限次的网络更新迭代后可以获得无限的记忆深度<sup>[103]</sup>. RNN 在 SR 中是以递归的方式多次应用相同的模块,如图 8(b)所示,其内部的模块具有参数共享的特性,使网络模型能够在不引入过多参数的情况下学习更高层次的特征,从而提升网络性能.

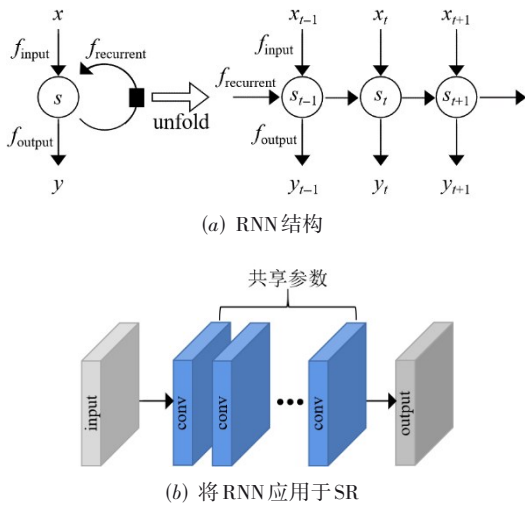


图 8 递归神经网络

为了控制深度网络中的参数量,避免增加网络深度带来的过拟合等问题, Kim 等人<sup>[71]</sup>首次将 RNN 应用于图像 SR 中,并结合残差学习的思想,于 2016 年提出了多达 16 个递归层的深度递归卷积网络 DRCN (Deeply-Recursive Convolutional Network),通过对部分卷积层采用递归学习的方式,实现在不引入过多参数的情况下提升网络性能. DRCN 由用于特征提取的嵌入网络、用于特征非线性映射的推理网络和重建网络等三个子网络构成,是将全局残差学习、单权重递归学习

及多目标优化结合的 SR 方法. Tai 等人<sup>[104]</sup>在 DRCN 的基础上,进一步结合 ResNet 和 RNN,提出了多达 52 个卷积层的深度递归残差网络 DRRN (Deep Recursive Residual Network),通过更深层次的网络结构提升网络模型的性能,同时结合多路径模式的局部、全局残差学习以及多权重的递归学习,控制参数量并稳定网络.

Han 等人<sup>[81]</sup>认为许多深层 SR 网络结构可以表示为具有各种递归函数的单状态递归神经网络的有限展开,并从 RNN 的角度理解深层结构,如图 9(a)~(c)所示. 基于此,他们提出了双状态递归网络 DSRN (Dual-State Recurrent Network),其 RNN 结构如图 9(d)所示. 与使用相同空间分辨率的单状态模型不同,DSRN 能够在不同的空间分辨率中运行,在 LR 和 HR 空间采用两个循环状态,通过网络中的延迟反馈机制,在 LR-HR 之间交换循环信号,充分利用 LR 和 HR 空间的特征,得到最终的重建图像.

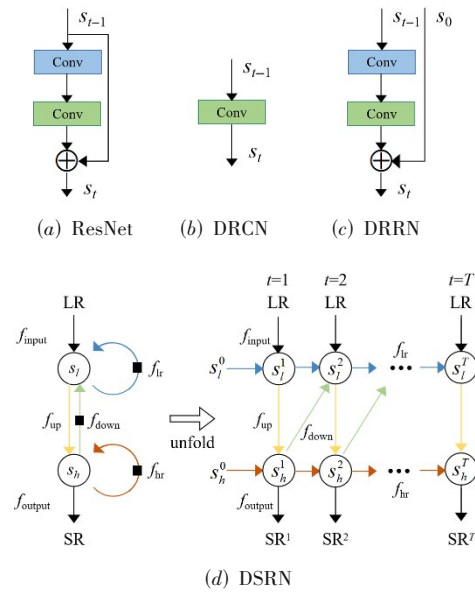


图 9 不同网络模型的 RNN 展开

2019 年, Li 等人<sup>[82]</sup>同样利用反馈机制,提出了一种图像超分辨率反馈网络 SRFBN (SR Feedback Network),将高阶信息细化为低阶表示,并逐步生成最终的 HR 图像,实现较少参数量情况下的网络性能提升. 图 10 所示是 SRFBN 网络模型的反馈机制,反馈方式通过使用带约束的 RNN 中的隐藏状态来实现. 同时, SRFBN 网络模型还引入了课程学习 (curriculum learning) 策略,通过将逐步增加重建难度的目标 HR 图像依次送入网络进行连续迭代,使网络能够逐步学习复杂的退化模型,从而更好地适应复杂的任务.

RNN 参数共享的特性使其能够在不引入过多参数的情况下学习更高层次的特征,但仍然无法避免深层

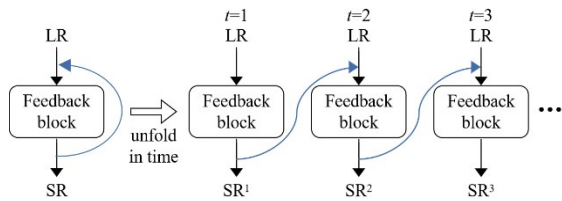


图 10 SRFBN 反馈机制

网络带来的梯度和训练等问题,因此 RNN 通常会与 ResNet、多监督学习、课程学习等网络设计及学习策略相结合,以缓解梯度及训练问题,实现网络性能的提升.

### 3.1.4 基于密集卷积网络的方法

2017年,Huang 等人<sup>[105]</sup>提出 DenseNet,并将其概括成一种简单的连接模式,即为了确保网络各层之间的最大信息流,直接连接具有相同特征图大小的任意两个层,并扩展到所有层的连接. DenseNet 在 SR 中应用的结构图如图 11 所示,对于网络中的每一层,该层前面所有层的特征图都作为该层的输入,而该层的特征图将成为后续所有层的输入之一.

与 ResNet 使用求和的方式将浅层特征传递到后续层再组合起来不同,DenseNet 是通过连接的方式来组合它们. 与传统的连接方式也不同,DenseNet 采用密集连接的方式进行组合能够充分利用层间信息. DenseNet 通过加强层间特征传播,鼓励层间特征重用,使网络模型中各层的特征被充分利用,从而大大减少参数量,改进网络的信息流动和梯度,有效缓解梯度消失的问题,使网络更容易训练.

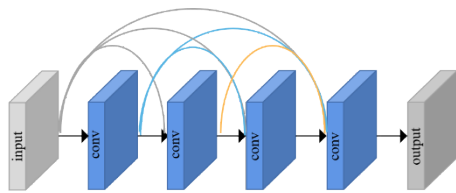


图 11 密集卷积网络应用于 SR

2017年,Tong 等人<sup>[106]</sup>首次将 DenseNet 应用于 SR 中,提出了 SRDenseNet (Super-Resolution DenseNet) 网络模型,通过密集跳跃连接将低、高层特征有效融合,再利用反卷积层进一步提升重建图像的细节信息. 同年,Tai 等人<sup>[107]</sup>提出了深度持久记忆网络 MemNet (deep persistent Memory Network),使用 DenseNet 中的密集连接操作来加强特征传播,弥补信息丢失,进一步增强高频信号. MemNet 由特征提取网络、多个堆叠的记忆块以及重构网络组成,其中最主要的结构是由用于模拟非线性函数的递归单元和用于自适应学习不同记忆权重的门单元组成的记忆块(图 12). 它通过自适应的学习过程来挖掘持久记忆,从而构建深度

网络的长期依赖关系.

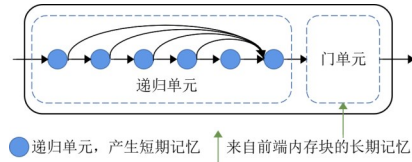


图 12 记忆块结构

单一的网络设计策略通常难以获取较大的网络性能提升,因此多策略的有效结合对网络性能的提升至关重要. 2019年,Shamsolmoali 等人<sup>[108]</sup>提出基于扩张卷积神经网络的新模型,通过将 DenseNet 与扩张卷积进行适当的结合,得到性能与效率有效权衡的网络模型. 次年,Pan 等人<sup>[109]</sup>提出基于密集残差网络的网络模型,利用基于高斯过程的神经结构搜索(GP-NAS)和异构模型集成等策略在真实图像 SR 中取得优异性能,并得到高保真度的重建图像. 而 Jiang 等人<sup>[110]</sup>提出的分层密集残差网络 HDRN (Hierarchical Dense Recursive Network) 同样在 DenseNet 的基础上利用分层残差块和全局融合模块实现整个网络由粗到细的特征重建,从而得到准确的重建效果.

DenseNet 在 SR 网络中通过密集连接的方式使网络模型中各层的特征被充分利用,不仅大大降低了网络参数量,而且有效缓解了梯度消失的现象,使得深度网络易于训练. 但单一的网络设计策略往往难以实现较大的网络性能提升,将 DenseNet 与不同网络设计及学习策略进行有效组合,才能使基于 DenseNet 的 SISR 网络模型实现更好的重建效果.

### 3.1.5 基于注意力机制的方法

注意力机制是根据图像特征的重要性程度分配不同的权重,使网络以高权重聚焦重要信息,以低权重忽略无关信息,从而改善图像细节,提升图像重建质量的机制,具有较好的灵活性和鲁棒性<sup>[111]</sup>. 目前主流的注意力机制包括通道注意力、空间注意力和自注意力. 2018年,Hu 等人<sup>[112]</sup>提出的 SENet (Squeeze-and-Excitation Network) 将通道注意力机制引入深度神经网络中,通过引入“挤压-激励”块(SE block)显式建模通道之间的相互依赖,提高网络的特征学习能力,其结构如图 13 所示.

前述基于 CNN 的网络模型在通道中平等对待 LR 输入图像包含的丰富低频信息,影响了网络的表征能

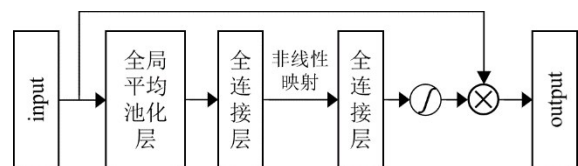


图 13 通道注意力机制

力,因此 Zhang 等人<sup>[113]</sup>于 2018 年首次将注意力机制应用于 SR 中,提出残差通道注意力网络 RCAN(Residual Channel Attention Network). RCAN 中的通道注意力机制能够根据通道之间的依赖关系自适应地重新调整每个通道的特征,从而学习到更多有用的通道特征,提高网络表征能力.此外,RCAN 中还使用了残差中的残差(Residual In Residual, RIR)结构,通过长、短跳跃连接构建深度可训练网络.

2019 年, Dai 等人<sup>[114]</sup>指出现有的基于 CNN 的网络模型大多通过设计更宽或更深层次的网络结构来提升性能,忽视了对中间层特征相关性的探索,从而限制了网络的表征能力.而 RCAN 中引入的经典通道注意力机制通过全局平均池化利用特征的一阶统计量,却忽略了高于一阶的统计量,从而阻碍了网络的判别能力,且研究表明二阶统计量较一阶统计量更有助于特征判别性的表示.为此, Dai 等人提出了二阶注意力网络 SAN(Second-order Attention Network)并引入二阶通道注意力(Second-Order Channel Attention, SOCA)机制(图 14),通过协方差归一化获取特征的二阶统计量来学习特征的相关性,使网络关注更多特别的特征,提高判别学习能力,从而实现更强大的特征相关学习和特征表达能力.受 SOCA 机制的影响, SAN 网络模型在具有纹理等更高阶信息的图像上表现更佳.

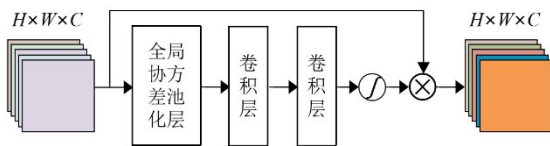


图 14 二阶通道注意力机制

2020 年, Wei 等人<sup>[43]</sup>没有通过统一处理图像中的所有像素/区域/组件或者侧重处理边缘或纹理来训练 SR 网络模型,而是受 Harris 角点检测的启发,根据图像所传达信息的重要性,将图像分为平面、边缘和角点三个低层次部分,并利用沙漏超分辨率网络 HGSR(Hour-Glass SR),分别构建与平面、边缘和角点相关的三个组件注意力块来探索不同组件的重要性,从而提出组件分治 CDC(Component Divide-and-Conquer)网络模型,旨在以分而治之的方式解决真实世界的 SR.另外, Wei 等人还提出了一种梯度加权损失函数,根据图像重建难度适应模型训练,以解决图像中不同区域在各个方向梯度不同的问题.

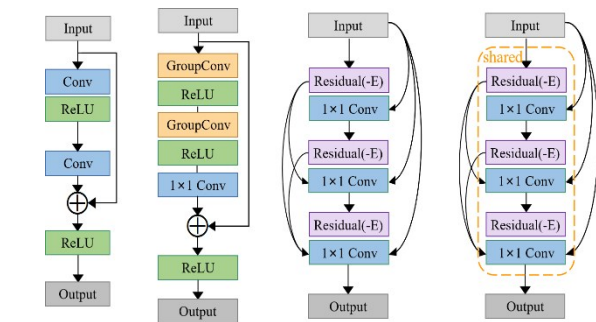
相比于其他类型的基于 CNN 的网络模型,基于注意力机制的网络模型通常会区别对待图像中的重要和不重要区域,并通过设置高权重加强对图像重要区域的特征提取,从而获取更有效的图像信息.与基于 RNN, DenseNet 等网络设计策略的网络模型相同,基于

注意力机制的网络模型也需要与残差学习、课程学习等其他网络设计及学习策略相结合,才能取得更好的网络性能提升.

### 3.1.6 基于轻量化网络的方法

轻量化网络指的是通过设计更加紧凑的网络结构或者在原始网络结构的基础上使用一些轻量化策略来减少网络参数量,提升网络速度,并保持或提升原有网络性能的一种高效网络. LN 旨在将 SR 算法应用于实际,实现真正意义上的轻量化及移动设备端的 SR 任务部署,是对性能与效率的一种权衡.目前,轻量化网络中常用的轻量化策略有很多,如使用扩张卷积(dilated convolution)、群卷积(group convolution)或者深度可分卷积(depthwise separable convolution)等先进卷积来替代传统的卷积操作,从而实现更高效的性能,同时还有网络剪枝(network pruning)、知识蒸馏(knowledge distillation)、神经架构搜索(neural architecture search)和自适应推理(adaptive inference)等轻量化策略可用于轻量化网络的构建.

为实现 SR 算法的实际应用,2018 年, Ahn 等人<sup>[115]</sup>提出了一种精确、高效的深度级联残差网络 CARN(Cascading Residual Network)及其轻量化的移动变体 CARN-M(CARN-Mobile). CARN 网络模型在 ResNet 的基础上将残差块(图 15(a))替换为级联块(图 15(c)),并通过局部和全局残差学习传递图像信息,该网络注重网络性能的提升. CARN-M 网络模型则在 CARN 的基础上结合有效的残差块(图 15(b))和将级联块进行参数共享的递归块(图 15(d))来优化参数量及操作数量,以实现性能与速度的权衡并获得高效的 SR 网络模型,使其能够应用于移动设备.



(a) 残差块 (b) 有效残差块 (c) 级联块 (d) 递归块  
图 15 CARN 不同模块结构对比

同年, Hui 等人<sup>[116]</sup>为减少网络运行时间,利用群卷积和信息蒸馏块等轻量化策略提出了一种由特征提取块、信息蒸馏块和重建块三部分组成的紧凑且深层的信息蒸馏网络 IDN(Information Distillation Network). 由增强单元和压缩单元组合而成的信息蒸馏块(图 16)作为 IDN 网络的重要结构,能够逐步提取丰富而有效的

图像特征. 其中,增强单元利用通道分离策略保留局部信息并处理后续信息,主要用于增强LR输入图像的轮廓区域,而压缩单元则由 $1\times 1$ 卷积层构成,主要用于降维及提取相关图像信息.IDN网络模型在群卷积、知识蒸馏和每层过滤器数量设置相对较少等各种策略的影响下,网络速度有了很大的提升.

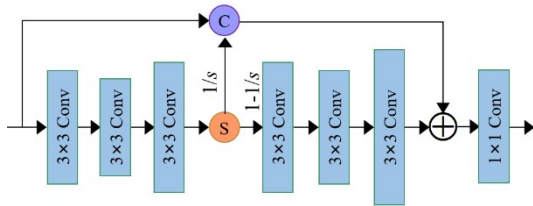


图 16 IDN 网络模型的信息蒸馏块

2019年,Hui等人<sup>[117]</sup>在IDN的基础上对信息蒸馏块加以改进,设计出信息多蒸馏块(图17)用于构建轻量化的信息多蒸馏网络IMDN(Information Multi-Distillation Network),并利用自适应裁剪策略解决任意尺度因子的SR问题.信息多蒸馏块由渐进细化模块(Progressive Refinement Module,PRM)、对比感知通道注意层以及能够减少特征通道数量的 $1\times 1$ 卷积构成,用于逐步提取更加细腻、更具真实感的图像特征.在各种策略的作用下,IMDN在客观评价指标和推理时间方面都表现良好,并取得了AIM 2019受限超分辨率重建挑战赛的冠军.次年,Liu等人<sup>[118]</sup>在IMDN基础上加以改进,提出了与通道分离策略具有相同操作的特征蒸馏连接,用于构建残差特征蒸馏网络RFDN(Residual Feature Distillation Network).RFDN在信息蒸馏网络的作用下实现了更轻量化和更灵活的图像SR,并获得了AIM 2020高效超分辨率重建挑战赛的冠军.

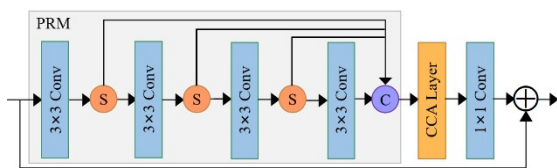


图 17 IMDN 网络模型的信息多蒸馏块

为减少网络参数数量和运算量,Chu等人<sup>[119]</sup>引入一种融合微观和宏观搜索的、新型的弹性神经架构搜索(Neural Architecture Search,NAS)方法,微观搜索空间用于提取特征单元块,宏观搜索空间则使用密集连接将特征单元块连接起来,从而构建能够实现快速、准确和轻量化的三种FALS(R,Accurate and Lightweight SR)网络模型,其中FALS-R在视觉效果方面表现最好,FALS-A具有最低的参数量,FALS-C的网络性能则表现得比轻量化的CARN网络模型更好.而Li等人<sup>[120]</sup>提出的线性组合像素自适应回归网络模型

LAPAR(Linearly-Assembled Pixel-Adaptive Regression)将直接学习LR图像到HR图像的映射问题转化为基于多个预定义过滤器字典的线性回归任务,同时根据特征通道数量( $C$ )和局部融合模块数量( $M$ )提出了LAPAR-A( $C32-M4$ ),LAPAR-B( $C24-M3$ )和LAPAR-C( $C16-M2$ )三种模型以评估网络的可扩展性.LAPAR三种网络模型在两倍尺度因子的情况下均表现出比FALS-R三种网络模型更高效的重建效果,且在保证运行速度的同时,该模型在图像去噪、JPEG去块等其他低级视觉任务中也表现良好.

2021年,Wang等人<sup>[121]</sup>提出了稀疏掩码超分辨率重建网络模型SMSR(Sparse Mask SR),通过研究图像的稀疏性减少网络的冗余计算,提高网络的推理效率.相较于独立关注空间及通道维度冗余计算的自适应推理和网络剪枝方法,SMSR提供了一个统一的框架来考虑空间和通道维度的冗余计算,通过将空间掩码与通道掩码结合,使用空间掩码学习识别图像的重要区域,使用通道掩码学习标记图像的不重要区域,因此精确地删除网络中冗余计算的部分,使网络能够有效地降低计算成本,以获得更好的效率,同时保持相当的性能.此外,SMSR在移动设备上的应用也有显著的加速.

轻量化网络能够在维持或提升网络性能的同时,降低网络参数量,提高网络速度,实现更高效的SR网络,使SR算法能够部署在现实应用中.尽管出现了越来越多的轻量化网络并取得了一定的成果,但大多数网络主要关注参数量和浮点运算次数(Floating Point Operations,FLOPs),然而FLOPs越少并不意味着网络效率越好,相反,网络激活的数量是网络效率更准确的衡量标准<sup>[122]</sup>.因此,对于轻量化网络效率的衡量不能只关注参数量和FLOPs,而应该从各个角度进行全面的分析<sup>[123]</sup>,进而实现高效的轻量化网络.

### 3.2 基于生成对抗网络的单帧图像超分辨率重建

生成对抗网络是Goodfellow等人<sup>[124]</sup>于2014年提出的一个通过对抗过程评估生成模型的新框架.如图18所示,该框架包含两个子模块:一个是生成器(Generator,G),用于捕获数据分布;另一个是判别器(Discriminator,D),用于判断输入数据的“真伪”.估计输入数据是来自训练数据而非来自G的概率,通过对抗博弈的方法训练两个网络,不断完善G和D,直至D“难辨真假”则完成训练.

基于GAN的方法主要采用生成对抗网络的结构,以感知驱动的方式训练网络,旨在使重建图像获得更好的感知质量、更逼真的视觉效果,在视觉效果上更接近真实图像,但该方法在客观评价指标方面表现不佳,对图像细节的恢复存在误差.

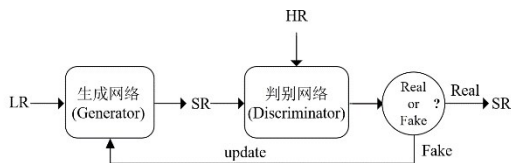


图 18 生成对抗网络

2017年, Ledig 等人<sup>[93]</sup>首次将 GAN 应用于图像 SR 领域中, 提出了 SRGAN (Super-Resolution Generative Adversarial Network) 网络模型. SRGAN 包含生成网络和对抗网络, 生成网络将输入的 LR 图像进行 SR 生成重建后的 HR 图像, 而判别网络则判断输入的是重建的 HR 图像还是原始的 HR 图像. 若判断错误则两者继续相互迭代训练, 直至判别网络将输入的重建 HR 图像当成是原始的 HR 图像则完成训练. 相较于之前基于 CNN 的网络模型, SRGAN 不仅使用 GAN 改进网络结构, 同时也将感知损失应用于 SR 中, 采用感知损失与对抗损失组合的损失函数, 从而使重建图像的细节更加丰富、图像更具真实感, 在感知质量上得到了很好的提升. 但 GAN 框架的“欺骗性”使它的 PSNR 和 SSIM 等客观评价指标相对较低.

为生成纹理逼真、图像自然的 SR 图像, Sajjadi 等人<sup>[125]</sup>提出了一种结合感知损失的自动纹理合成的增强型网络 EnhanceNet. EnhanceNet 同样采用了 GAN 的结构, 生成网络部分是一个用于纹理合成的前馈全卷积神经网络, 通过 GAN 与感知损失的结合, 能够实现高放大倍数的真实纹理, 而判别网络部分则遵循常见的设计模式. 同时为证明不同损失函数对重建图像质量的影响, Sajjadi 等人还对各种损失函数的重建结果进行比较, 验证了以 MSE 为损失函数的重建图像虽然有最高的 PSNR 和 SSIM 等客观评价指标, 但其重建结果较为平滑, 缺乏高频细节, 而将感知损失、对抗损失以及纹理损失相结合的损失函数所生成的重建图像虽然客观评价指标较低, 但却能产生具有真实纹理及逼真视觉效果的重建图像.

受 SRGAN 的启发, Wang 等人<sup>[126]</sup>于 2018 年提出了 ESRGAN (Enhanced SRGAN) 网络模型, 主要在网络结构、损失函数等方面对 SRGAN 进行改进. 在网络结构方面, 生成网络去除了残差块中的 BN 层, 并引入了残差中的残差密集块 (Residual-in-Residual Dense Block, RDB) 结构, 判别网络则用相对判别器代替原有的标准判别器. 在损失函数方面, 对感知损失进行改进, 通过使用激活层前的特征, 增强图像的代表能力, 并引入网络插值<sup>[127]</sup>的方法使重建图像从平滑的重建结果向感知质量较好的重建结果转移. 得益于这些改进, ESRGAN 网络模型的网络性能及重建图像都得到了很好的提升.

ESRGAN 通过对 SRGAN 的改进, 实现了优异的感知重建效果, 故此以后基于 GAN 的网络模型大多都是在 ESRGAN 网络模型上进行改进创新. 如 BSRGAN<sup>[47]</sup>, RFB-ESRGAN<sup>[128]</sup> (NTIRE 2020 感知极端 SR 挑战赛的冠军模型) 和 Real-ESRGAN<sup>[48]</sup> 等网络模型的结构都是基于 ESRGAN 结构进行改进, 并取得了优异的重建效果.

Zhang 等人<sup>[129]</sup>指出许多与人类主观评价高度相关的感知质量评价指标 (无参考型客观评价指标) 通常是不可微的且无法作为损失函数优化网络模型, 故提出由标准的 SRGAN 和 Ranker 组成的 RankSRGAN 网络模型. 其中 Ranker 是一个通用和可微的模型, 可以通过学习排名的方法模拟任意感知指标的行为, 并作为损失函数优化网络. RankSRGAN 网络模型在公共 SR 数据集上使用不同的 SR 方法生成 SR 图像, 再将成对的图像对根据感知质量得分进行排名后构建排名 (rank) 数据集用于网络模型的训练. 因此 RankSRGAN 网络模型能够结合不同 SR 方法的优点, 在感知质量方面产生更好的结果, 并恢复出比 SRGAN 和 ESRGAN 网络模型更真实的纹理.

基于 GAN 的方法利用 GAN 结构的对抗性使网络重建效果更具真实感, 对不关注细节的整体图像具有较好的应用效果, 但该方法存在大量的网络参数, 使得网络训练不稳定, 推理速度也因此延缓. 对于基于 GAN 的方法要注重图像细节的重建, 同时采用合适的策略构造轻量化网络并使其训练稳定.

### 3.3 基于 Transformer 的单帧图像超分辨率重建

Transformer 是由 Google 的 Vaswani 等人<sup>[130]</sup>于 2017 年提出的一种用于自然语言处理 (Natural Language Processing, NLP) 的网络架构, 其模型架构如图 19 所示. 它摒弃了 RNN 和 CNN, 是一个完全基于自注意力机制来获取输入和输出之间全局依赖关系的转换模型. 相较于 CNN 通过堆叠卷积层扩大感受野以获取全文信息, RNN 通过递推捕捉全局联系, 却难以捕捉长距离依赖, 而自注意力机制能够更好地捕捉全局联系, 解决了长距离依赖的问题, 同时能够支持并行化计算, 加快训练速度, 提升网络效率.

Transformer 最初是为了 NLP 任务中的序列建模而设计的, 而后随着 ViT (Vision Transformer)<sup>[131]</sup>, DETR (Detection Transformer)<sup>[132]</sup>, ViViT (Video Vision Transformer)<sup>[133]</sup> 等网络模型的相继提出, Transformer 逐渐被应用于计算机视觉领域, 并取得比 CNN、非极大抑制和 3D 卷积等更好的效果<sup>[134]</sup>.

基于 Transformer 的方法主要采用 Transformer 的结构, 将自注意力机制引入网络中. 由于 Transformer 强大的图像表征能力和各式各样的结构, 此类网络能够得

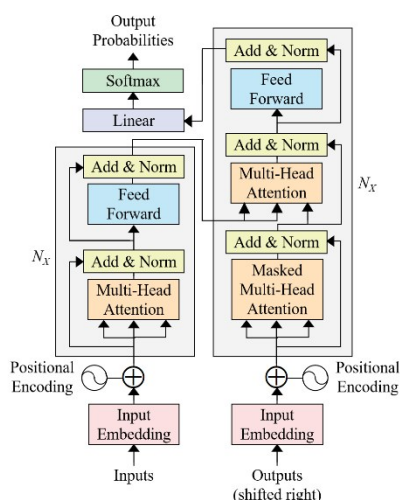


图 19 Transformer 模型架构

到较 CNN 更好的重建结果和评价指标。当前基于 Transformer 的 SR 方法主要有两种类型：一种是完全使用 Transformer 结构作为网络架构的纯 Transformer 网络模型；另一种是将 Transformer 作为主干网络与 CNN 相结合的混合 Transformer 网络模型。

2021 年, Chen 等人<sup>[94]</sup>将 Transformer 架构应用于计算机视觉领域, 联合提出了一种用于处理 SR、去噪和去雨等多种低级计算机视觉任务的预训练网络模型 IPT (Image Processing Transformer)。该网络属于纯 Transformer 模型, 以端到端的方式进行学习, 网络结构主要由用于从输入退化图像提取特征的头(heads)、用于从输入数据中重建丢失信息的编-解码器(encoder-decoder)Transformer 和用于输出重建图像的尾(tails)三部分构成, 其中编-解码器 Transformer 与原始 Transformer<sup>[130]</sup>中的结构相似, 不同之处在于该网络利用了特定任务的嵌入作为解码器的附加输入。为了最大限度地挖掘 Transformer 的潜力, 作者采用包含 1 000 个类别的 ImageNet 数据集构造了大量的退化图像数据对, 并利用这些数据对对 IPT 模型进行训练。同时, 为了使 IPT 模型更好地适应不同的图像处理任务并将其应用到未知任务上, 作者还引入了对比学习(contrastive learning)来学习通用的特征。最后, 经过微调后的预训练 IPT 模型可以有效地用于所需的任务, 且在不同任务上的表现超过了大多数现有的方法。

2021 年, Liang 等人<sup>[135]</sup>基于 Swin Transformer<sup>[136]</sup>的结构, 提出了一种用于图像恢复的网络模型 SwinIR (Image Restoration Using Swin Transformer)。该网络属于混合 Transformer 模型, 网络结构主要由浅层特征提取、深层特征提取和图像重建三个模块组成。浅层特征提取模块采用卷积层提取浅层特征, 并利用残差连接将浅层特征直接传递给重建模块, 以保留图像的低频信息;

深层特征提取模块主要由多个 RSTB (Residual Swin Transformer Block) 和一个用于特征增强的卷积层组成, 每个 RSTB 中利用 STL (Swin Transformer Layer) 进行局部注意和跨窗口交互; 而图像重建模块则通过融合浅层和深层特征, 实现高质量图像的重建。SwinIR 网络模型整合了 CNN 和 Transformer 的优势, 既能够利用 CNN 处理大尺度因子的图像 SR 问题, 也能够利用 Transformer 解决长距离依赖的问题, 从而在图像 SR、图像去噪和 JPEG 压缩伪影减少等低级计算机视觉任务上表现出良好的性能。

同年, Lu 等人<sup>[137]</sup>指出 Vision Transformer 计算成本高、GPU 内存占用大等问题导致网络无法设计得过深, 为此, 提出了一种新的高效网络 ESRT (Efficient SR Transformer), 以研究在轻量级 SR 任务中使用 Transformer 的可行性。ESRT 是一个混合 Transformer 的网络模型, 网络结构主要由浅层特征提取、轻量级 CNN 骨干 (Lightweight CNN Backbone, LCB)、轻量级 Transformer 骨干 (Lightweight Transformer Backbone, LTB) 和图像重建四个部分组成。其中, LCB 通过动态调整映射图的大小, 能够以较低的计算成本提取深层图像特征, 可用于解决 Transformer 在小数据集上特征提取能力差的问题。而由一系列高效 Transformer (Efficient Transformer, ET) 组成的 LTB, 主要用于获取图像中相似块的长期依赖关系, 同时利用 ET 解决其他 Vision Transformer 参数大和 GPU 内存消耗大的问题。通过这些改进, ESRT 能够有效地增强图像中相似块的特征表达能力和长期依赖性, 从而获得更好的性能, 验证了 Transformer 在轻量级 SR 任务中的可行性。

无论是纯 Transformer 网络模型还是混合 Transformer 网络模型, 目前基于 Transformer 的方法均表现出比基于 CNN 的方法更好的网络性能, 但基于 Transformer 的方法目前仍处于发展阶段, 因此主要注重网络模型重建质量的提升, 对于实际的应用考虑较少, 后续可结合实际应用考虑更具实用性的轻量化 Transformer 网络模型。

#### 4 无监督式单帧图像超分辨率重建

无监督式 SISR 旨在通过从未配对的 LR-HR 图像中学习函数来解决 SR 问题, 从而得到更能够处理真实世界场景下 SR 问题的网络模型。目前, 大多数 SR 网络模型都是采用有监督学习的方法进行网络训练, 但监督式 SR 仍然存在一些难以解决的问题, 如真实数据集构造过程存在一定难度, 合成数据集采用的退化过程具有一定的人工先验, 无法完全符合图像真实退化过程等。而无监督式 SR 可以对未配对的 LR-HR 图像对进行训练, 相比于需要使用成对的 LR-HR 图像对进行网

络训练的监督式 SR,其在真实世界的图像 SR 任务上更具优势,因此逐渐受到研究人员的关注,如 AIM 2019 和 NTIRE 2020 等真实世界图像超分辨率重建挑战赛就旨在以弱监督或无监督的学习方法来实现 SR. 本章主要将无监督式 SISR 分为“零样本(zeroshot)”图像超分辨率重建和弱监督式图像超分辨率重建两类进行介绍.

#### 4.1 “零样本”单帧图像超分辨率重建

2018 年,Shocher 等人<sup>[138]</sup>认为现实图像是具有模糊、噪声、伪影和压缩等性质的非理想图像,监督式的 SR 方法无法很好地解决非理想图像的 SR 问题,因此不适用于真实世界图像的 SR. 为解决此问题,Shocher 等人提出了第一个无监督式的基于 CNN 的网络模型 ZSSR(Zero-Shot SR). ZSSR 网络不依赖模型的预训练,而是利用单个图像内部的信息,在测试时直接训练一个小型的特定图像的 CNN,仅从输入的 LR 测试图像中提取内部信息,从而实现各种大小、尺度因子的非理想图像的 SR. 图 20 为监督式 SR 与 ZSSR 的对比,相较于监督式 SR 需要花费较多时间对网络进行预训练从而得到图像的先验知识,ZSSR 能够在测试时直接进行网络的训练与测试,同时在非理想的图像上表现出比 EDSR 网络模型更好的性能. 但 ZSSR 存在测试时间较长以及对每张图像都需要训练一个特定网络模型的缺点.

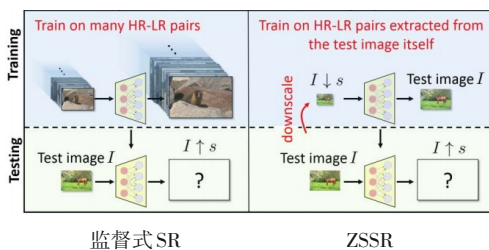


图 20 监督式 SR 与 ZSSR 对比<sup>[138]</sup>

#### 4.2 弱监督式单帧图像超分辨率重建

对于未配对的 LR-HR 图像,弱监督式 SR 主要是采用不同的方法,在不引入预定义退化的情况下进行 SR. 根据所用方法的不同,大致可以将其分为两种类型:一种是运用无监督式的退化模型学习图像退化过程以构造合成数据集进行 SR,另一种则是在网络中通过不同的循环结构以实现弱监督式 SR.

为解决监督式 SR 方法无法应用于现实世界场景图像的问题,Bulat 等人<sup>[139]</sup>于 2018 年提出了一个弱监督式的两阶段网络模型. 首先利用一个 GAN 模型以无监督的方法学习 HR 图像的退化和下采样过程,得到相应的 LR 图像,再对前一个 GAN 模型的 LR 输出图像进行重建,训练一个由 LR 图像到 HR 图像的 GAN 模型,从

而实现真实世界图像的 SR,有效地提高真实世界 LR 图像的质量. 此外,DSGAN<sup>[50]</sup>和 RealSR<sup>[60]</sup>等网络模型同样以无监督的方式获取相应 LR 图像并构造合成数据集,解决 LR-HR 图像不匹配的问题,再基于改进后的 ESRGAN 网络模型进行图像的重建,取得了很好的重建效果,并分别在 AIM 2019 和 NTIRE 2020 真实图像超分辨率重建挑战赛上获得了冠军.

2018 年,Yuan 等人<sup>[140]</sup>受 CycleGAN<sup>[141]</sup>的启发,以生成对抗网络为基本结构,提出了一种无监督式的循环生成对抗网络模型 CinCGAN(Cycle-in-Cycle GAN). 该模型包含 2 个 CycleGAN:第一个 CycleGAN 通过将有噪声和模糊的 LR 输入图像映射到一个双三次下采样的 clean LR 空间,从而得到去噪和去模糊后的 LR 图像;第二个 CycleGAN 则用于学习从第一个 CycleGAN 中输出的 LR 图像到 HR 图像的映射,即通过现有的 SR 网络模型来重建中间结果到所需的放大倍数,再以端到端的方式对这两个模块同时进行组合和微调,从而得到 HR 重建图像的输出. 此外,为进一步提高 CinCGAN 的网络性能,Zhang 等人<sup>[142]</sup>又于 2020 年采用渐进式多循环策略及模糊损失对 CinCGAN 加以改进,提出了一个多循环网络模型 MCinCGAN(Multiple Cycle-in-Cycle GAN),用于处理多种尺度因子、退化未知且复杂和 LR-HR 图像不匹配的 SR 问题.

Wu 等人<sup>[143]</sup>认为导致无监督式 SR 模型产生不满意重建效果的关键在于不真实的低频信息和不准确的高频纹理,因此利用数据约束策略从损失函数、训练数据和后处理等方面对 CycleGAN 的结构进行改进,提出了无监督的 ECycleGAN(Enhanced CycleGAN)网络模型. 他们通过引入新的图像约束损失函数来补偿无监督学习中像素级监督的缺失,并限制判别器的数据内容,促进其抑制高频纹理或伪影,此外还引入了模型平均策略用于后处理. 由于这些改进,ECycleGAN 的训练变得更加稳定,同时图像重建能力也得到了有效提升.

## 5 图像超分辨率重建发展趋势

### 5.1 图像超分辨率重建网络模型汇总

为便于分析图像超分辨率重建的发展趋势,表 3 对前述部分网络模型相关内容进行总结对比. 表中的类型按照前述网络模型进行分类. LR 图像获取方式:Bi-cubic 表示双三次下采样,B 表示模糊核,GB 表示高斯模糊核,N 表示噪声,GN 表示高斯噪声,J 表示 JPEG 压缩. 由于数据增强<sup>[144]</sup>能够扩大数据容量,减少迭代次数,并在一定程度上提升网络性能,所以许多网络模型常通过随机翻转、旋转和缩放等操作对数据集进行数据增强.

表 3 基于深度学习的单帧图像超分辨率重建典型网络模型总结

网络名称-- 发表时间(类型)	网络框架 (上采样方法)	训练集	LR 图像获取 方式	数据 增强	测试集	损失 函数	评价指标
SRCNN--2014 (基于 S-CNN)	预上采样 (双三次插值)	91-images, ImageNet	Bicubic+GB	—	Set5, Set14	$L_2$	PSNR, SSIM, runtime
FSRCNN--2016 (基于 S-CNN)	后上采样 (反卷积)	91-images, General-100	下采样	缩放、 旋转	Set5, Set14, BSD200	$L_2$	PSNR, SSIM, IFC, runtime
ESPCN--2016 (基于 S-CNN)	后上采样 (亚像素卷积)	ImageNet	下采样+GB	—	91-images, Set5, Set14, BSD300, BSD500, super texture	$L_2$	PSNR, runtime
VDSR--2016 (基于 ResNet)	预上采样 (双三次插值)	291-images	下采样	旋转、 翻转	Set5, Set14, Urban100, B100	$L_2$	PSNR, SSIM, runtime
RED-Net--2016 (基于 ResNet)	预上采样 (双三次插值)	BSD300	下采样	旋转、 翻转	Set5, Set14, BSD100	$L_2$	PSNR, SSIM
DRCN--2016 (基于 RNN)	预上采样 (双三次插值)	91-images	下采样	—	Set5, Set14, B100, Urban100	$L_2$	PSNR, SSIM
LapSRN--2017 (基于 ResNet)	渐进式上采样 (转置卷积)	291-images	Bicubic	缩放、 旋转、 翻转	Set5, Set14, BSDS100, Urban100, Manga109	$L_{\text{Charbonnier}}$	PSNR, SSIM, IFC
EDSR--2017 (基于 ResNet)	后上采样 (亚像素卷积)	DIV2K	Bicubic	翻转、 旋转	Set5, Set14, B100, Urban100, DIV2K	$L_1$	PSNR, SSIM
DRRN--2017 (基于 RNN)	预上采样 (双三次插值)	291-images	下采样	翻转、 旋转	Set5, Set14, BSD100, Urban100	$L_2$	PSNR, SSIM, IFC
SRDenseNet--2017 (基于 DenseNet)	后上采样 (反卷积)	ImageNet	Bicubic	—	Set5, Set14, B100, Urban100	$L_2$	PSNR, SSIM
MemNet--2017 (基于 DenseNet)	预上采样 (双三次插值)	291-images	Bicubic	翻转、 旋转	Set5, Set14, BSD100, Urban100	$L_2$	PSNR, SSIM
SRGAN--2017 (基于 GAN)	后上采样 (亚像素卷积)	ImageNet	Bicubic	—	Set5, Set14, BSD100	$L_{\text{percep}},$ $L_{\text{GAN}}$	PSNR, SSIM, MOS
EnhanceNet--2017 (基于 GAN)	后上采样 (最近邻插值)	MSCOCO	Bicubic	—	Set5, Set14, BSD100, Urban100	$L_{\text{percep}},$ $L_{\text{GAN}},$ $L_{\text{texture}}$	PSNR, SSIM, IFC, MOS
DBPN--2018 (基于 ResNet)	迭代式上下采样 (反卷积)	DIV2K, Flickr2K, ImageNet	Bicubic	—	Set5, Set14, BSDS100, Urban100, Manga109	$L_2$	PSNR, SSIM
MSRN--2018 (基于 ResNet)	后上采样 (亚像素卷积)	DIV2K	Bicubic	缩放、 旋转、 翻转	Set5, Set14, BSDS100, Urban100, Manga109	$L_1$	PSNR, SSIM
RCAN--2018 (基于 AM)	后上采样 (亚像素卷积)	DIV2K	①Bicubic ②模糊下采样	旋转、 翻转	Set5, Set14, B100, Urban100, Manga109	$L_1$	PSNR, SSIM
ESRGAN--2018 (基于 GAN)	后上采样 (最近邻插值)	DIV2K, Flickr2K, OST	Bicubic	翻转、 旋转	Set5, Set14, BSD100, Urban100, PIRM	$L_{\text{percep}},$ $L_{\text{GAN}}, L_1$	PSNR, PI
LP-KPN--2019 (基于 ResNet)	渐进式上采 (shuffle upsample)	RealSR	—	旋转、 翻转	RealSR	$L_2$	PSNR, SSIM
SRFBN--2019 (基于 RNN)	后上采样 (反卷积)	DIV2K, Flickr2K	①Bicubic ②下采样+GB ③Bicubic+GN	翻转、 旋转	Set5, Set14, B100, Urban100, Manga109	$L_1$	PSNR, SSIM

续表

网络名称-- 发表时间(类型)	网络框架 (上采样方法)	训练集	LR 图像获取 方式	数据 增强	测试集	损失 函数	评价指标
Meta-SR--2019 (基于 DenseNet)	后上采样 (元上采样)	DIV2K	下采样	翻转、 旋转	Set14, B100, Manga109, DIV2K	$L_1$	PSNR, SSIM
SAN--2019 (基于 AM)	后上采样 (亚像素卷积)	DIV2K	①Bicubic ②模糊下采样	旋转、 翻转	Set5, Set14, BSD100, Urban100, Manga109	$L_1$	PSNR, SSIM
DAN--2020 (基于 ResNet)	后上采样 (亚像素卷积)	DIV2K, Flickr2K	①GB+下采样 ②GB(旋转、噪 声、归一化)+ 下采样	—	①Set5, Set14, Urban100, BSD100, Manga109 ②DIV2K, KRK	$L_1$	PSNR, SSIM
HDRN--2020 (基于 DenseNet)	后上采样 (亚像素卷积)	DIV2K	Bicubic	翻转、 旋转	Set5, Set14, BSD100, Urban100, Manga109, Jilin-1 satellite imagery, Kaggle	$L_{\text{Charbonnier}}$	PSNR, SSIM
CDC-2020 (基于 AM)	渐进式上采样 (亚像素卷积)	RealSR, DRealSR	—	—	RealSR, DRealSR	$L_{\text{CW}}(L_1)$	PSNR, SSIM, LPIPS
RealSR--2020 (无监督式)	后上采样 (最近邻插值)	①DF2K ②DPED	无监督式 退化模型	—	①DF2K ②DPED	$L_1$ , $L_{\text{percep}}$ , $L_{\text{GAN}}$	①PSNR, SSIM, LPIPS ②MOR
RFB-ESRGAN-- 2020 (基于 GAN)	后上采样 (交替使用最近邻 插值和亚像素卷 积)	DIV8K, DIV2K, Flickr2K, OST	Bicubic	翻转、 旋转	DIV8K	$L_1$ , $L_{\text{VGG}}$ , $L_{\text{GAN}}$	PSNR, SSIM, LPIPS, PI
CRN--2021 (基于 ResNet)	后上采样 (亚像素卷积)	DIV2K	Bicubic	—	Set5, Set14, B100, Urban100	$L_1$	PSNR, SSIM
SMSR--2021 (基于 LN)	后上采样 (亚像素卷积+双三 次插值)	DIV2K	下采样	旋转、 翻转	Set5, Set14, B100, Urban100, Manga109	$L_1, L_{\text{reg}}$	PSNR, SSIM, Params, FLOPs
BSRGAN--2021 (基于 GAN)	后上采样 (最近邻插值)	DIV2K, Flickr2K, WED, FFHQ	B+下采样+N	—	①合成的 DIV2K4D, ②真实的 RealSRSet	$L_1$ , $L_{\text{percep}}$ , $L_{\text{PatchGAN}}$	①PSNR, LPIPS ②NIQE, NRQM, PI
Real-ESRGAN-- 2021 (基于 GAN)	后上采样 (最近邻插值)	DIV2K, Flickr2K, OST	Bicubic+B+N+J	—	RealSR, DRealSR, OST300, DPED, ADE20K	$L_1$ , $L_{\text{percep}}$ , $L_{\text{GAN}}$	视觉效果
IPT--2021 (基于 Transformer)	—	ImageNet	Bicubic	—	Set5, Set14, B100, Urban100	$L_1$ , $L_{\text{contrastive}}$	PSNR
SwinIR--2021 (基于 Transformer)	后上采样 (亚像素卷积)	DIV2K, Flickr2K	①Bicubic ②B+下采样+N	—	Set5, Set14, Urban100, BSD100, Manga109	① $L_1$ ② $L_1$ , $L_{\text{GAN}}$ , $L_{\text{percep}}$	PSNR, SSIM, Params, Multi-Adds

对于网络模型的总结对比,除了表3所示的内容外,本文还在 Wang 等人<sup>[37]</sup>的基础上,对最新网络模型的 SR 精度(即 PSNR)、模型大小(即参数数量)和计算代价(即操作数 multi-add)进行基础测试对比,如图 21 所示. SR 精度由网络模型在 Set5, Set14, B100 和 Urban100 四个基准数据集上的平均 PSNR 值来衡量,再计算模型大小和计算成本,尺度因子为 2.

## 5.2 图像超分辨率重建挑战赛

ICCV (IEEE International Conference on Computer Vision, 即国际计算机视觉大会)、CVPR (IEEE Conference on Computer Vision and Pattern Recognition, 即国际计算机视觉与模式识别会议)和 ECCV (European Conference on Computer Vision, 即欧洲计算机视觉国际会议)是世界三大顶级的计算机视觉会议. 近年来,利用

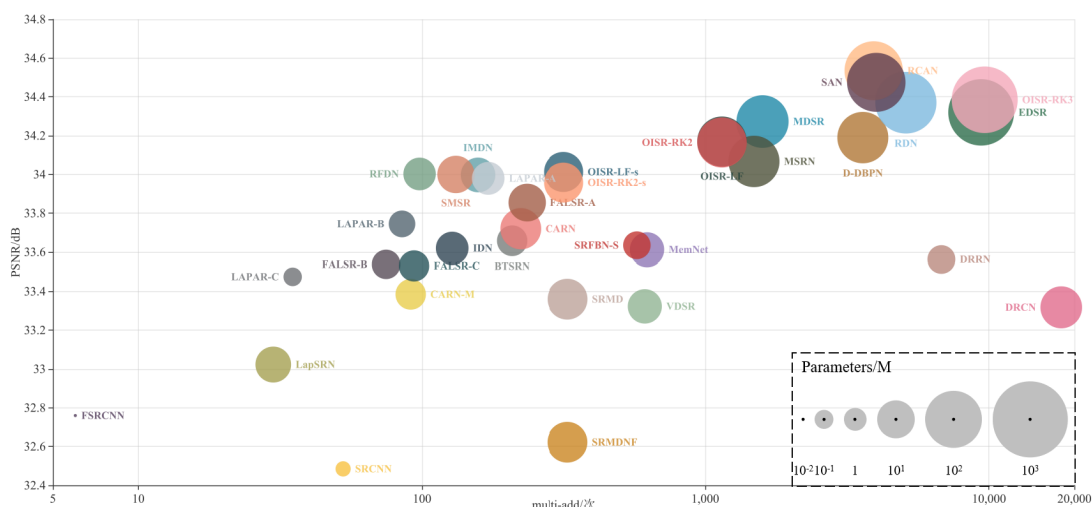


图 21 图像超分辨率重建基准测试

深度学习技术的图像SR研究取得了显著进展,相关挑战赛也层出不穷.其中最有影响力的图像SR挑战当属三大顶会所组织的NTIRE(New Trends in Image Restoration and Enhancement,即图像恢复与增强的新趋势)挑战赛、AIM(Advances in Image Manipulation,即图像处理的进展)挑战赛以及PIRM(Perceptual Image Restoration and Manipulation,即感知图像恢复与处理)挑战赛.

NTIRE挑战赛一般是与同年的CVPR会议同期开展,研究图像恢复与增强的新趋势,聚焦图像、视频的质量改进与评估,涉及图像及视频的超分辨率重建、去噪、去雾、去模糊等内容.NTIRE挑战赛自2017年开始举办至今,吸引了不少图像处理领域的研究者和爱好者的关注,是近年来计算机视觉领域非常有影响力的一场赛事.2021年,NTIRE挑战赛在图像处理挑战方面没有关于SISR的相关挑战,但在视频处理挑战方面有一场关于MISR的挑战.另外,还有一场与CVPR 2021联合举办的首届Mobile AI 2021挑战赛,旨在从移动设备端开发端到端基于深度学习的图像SR的解决方案,并在移动或边缘NPU上实现实时性.

AIM挑战赛是关于图像处理类的国际挑战赛,涵盖了图像、视频处理的多个热门研究方向,例如超分辨率重建、重光照、图像补全等.AIM挑战赛中关于图像SR的挑战主要涉及真实世界的图像SR和高效的图像SR.AIM挑战赛与NTIRE挑战赛一样在计算机视觉领域有很大的影响力,自2019年起连续举办了两年,2021年由于时间和资源的限制没有举办.

PIRM挑战赛是2018年与ECCV 2018 PIRM研讨会联合举办的第一个关于感知图像超分辨率重建的挑战.PIRM挑战赛中提出的无参考型客观评价指标PI联合量化了准确性和感知质量,且与主观评价指标具有较强的相关性,为感知SR的图像质量提供了有效的评

价指标.虽然PIRM挑战赛只举办了一届,但它为感知图像SR今后的发展奠定了坚实的基础.

为分析图像超分辨率重建的发展趋势,表4<sup>[59, 68, 83, 123, 145-153]</sup>从赛道设置、数据集使用、评价指标等方面总结了近年来上述挑战赛的相关信息.

### 5.3 图像超分辨率重建发展趋势

从基于CNN到基于GAN再到基于Transformer的单帧图像超分辨率重建的发展,表明了基于深度学习的图像超分辨率重建技术取得了阶段性的成功,开始从PSNR目标最大化的SR向感知驱动的SR再到性能显著提升的SR转换.虽然现阶段基于深度学习的图像超分辨率重建技术已有很多表现优异的成果,但仍存在很大的发展空间.根据表3对现有研究成果的总结和表4对近年相关挑战赛的相关内容总结,可以从以下几个方面对图像超分辨率重建领域未来的发展趋势进行展望.

#### (1) 性能提升

当前,影响SR网络性能的主要因素是训练数据集以及网络结构等.网络结构包括网络模型中使用的网络框架、损失函数、网络设计及学习策略、网络深度及宽度等.改善这些因素可以提升网络性能,获得重建效果更好的网络模型.

(a) 训练数据集.数据集是影响网络性能的重要因素,如何采集数据集并构建训练数据集则至关重要.目前,训练数据集可以通过合成数据集以及真实数据集的两种方式获取.

从表3和表4可以看出,近年来训练集中LR图像的获取方式已逐渐从简单的双三次合成数据集向真实数据集以及无监督式模拟退化过程的合成数据集转变.在网络结构变化较小的情况下,通过转变训练集获取方式的网络模型性能也能有所提升,如RealSR<sup>[60]</sup>网

表 4 图像超分辨率重建相关挑战赛内容总结

挑战赛	赛道设置(尺度因子)	数据集	评价指标
NTIRE 2017--单图像超分辨率重建挑战 <sup>[59]</sup>	赛道 1:经典的双三次下采样( $\times 2$ 、 $\times 3$ 、 $\times 4$ ) 赛道 2:未知下采样( $\times 2$ 、 $\times 3$ 、 $\times 4$ )	DIV2K 数据集	PSNR, SSIM
NTIRE 2018--单图像超分辨率重建挑战 <sup>[145]</sup>	赛道 1:经典的双三次( $\times 8$ ) 赛道 2:现实温和不利条件( $\times 4$ ) 赛道 3:现实困难不利条件( $\times 4$ ) 赛道 4:现实野生不利条件( $\times 4$ )	DIV2K 数据集	PSNR, SSIM
PIRM 2018--感知图像超分辨率重建的挑战 <sup>[68]</sup>	赛道:采用双三次核下采样的单幅图像 SR ( $\times 4$ )	PIRM 数据集	PI, RMSE
NTIRE 2019--真实图像超分辨率重建挑战 <sup>[146]</sup>	赛道:真实世界 SR(未知的尺度因子和退化算子)	RealSR 数据集	PSNR, SSIM
AIM 2019--真实世界图像超分辨率重建挑战 <sup>[147]</sup>	赛道 1:源域( $\times 4$ ) 赛道 2:目标域( $\times 4$ )	DIV2K, Flickr2K 数据集	PSNR, SSIM, LPIPS, MOS
AIM 2019--受限超分辨率重建挑战 <sup>[148]</sup>	赛道 1:参数( $\times 4$ ) 赛道 2:推理时间( $\times 4$ ) 赛道 3:保真度(PSNR)( $\times 4$ )	DIV2K 数据集	参数个数、平均运行时间和平均 PSNR
AIM 2019--图像极端超分辨率重建挑战 <sup>[149]</sup>	赛道 1:保真度( $\times 16$ ) 赛道 2:感知质量( $\times 16$ )	DIV8K 数据集	赛道 1:PSNR, SSIM 赛道 2:PSNR, SSIM, MOS
NTIRE 2020--真实世界图像超分辨率重建挑战 <sup>[83]</sup>	赛道 1:图像处理伪影( $\times 4$ ) 赛道 2:智能手机图像( $\times 4$ )	赛道 1:Flickr2K, DIV2K 数据集 赛道 2:DPED 数据集	赛道 1:PSNR, SSIM, LPIPS, MOS 赛道 2:NIQE, BRISQUE, PIQE, NRQM, PI, IQA-Rank, MOR
NTIRE 2020--感知极端超分辨率重建挑战 <sup>[150]</sup>	赛道:感知极端 SR( $\times 16$ )	DIV8K 数据集	PSNR, SSIM, LPIPS, PI
AIM 2020--真实图像超分辨率重建挑战 <sup>[151]</sup>	赛道 1:上采样 $\times 2$ 赛道 2:上采样 $\times 3$ 赛道 3:上采样 $\times 4$	DRealSR 数据集	PSNR, SSIM
AIM 2020--高效超分辨率重建挑战 <sup>[123]</sup>	赛道:保持 PSNR 的同时,减少运行时间、参数、FLOPs、激活和内存消耗等一个或者几个方面( $\times 4$ )	DIV2K 数据集	验证和测试 PSNR、运行时间、参数数量、FLOPs 数量、激活数量和推断期间消耗的最大 GPU 内存
NTIRE 2021--突发超分辨率重建挑战 <sup>[152]</sup>	赛道 1:合成数据集( $\times 4$ ) 赛道 2:真实世界数据集( $\times 4$ )	赛道 1:合成突 发数据集 赛道 2:BurstSR 数据集	赛道 1:PSNR, SSIM, LPIPS 赛道 2:PSNR, SSIM, LPIPS, MOR
Mobile AI 2021--基于移动 NPU 的实时量化图像超分辨率重建挑战 <sup>[153]</sup>	赛道:开发一个有效实际的解决方案用于移动任务	DIV2K 数据集	PSNR, SSIM, 运行时间

络模型在网络结构方面只对 ESRGAN 判别网络进行改进,并通过无监督式退化模型合成数据集,就得到了很大的性能提升. 因此为了获得更好的图像质量,可以通过改善训练数据集的构建方式,模拟图像真实退化过程,扩大退化空间使模型更具泛化性,以便对各种真实场景进行 SR.

(b) 损失函数. 损失函数作为网络结构的一部分,在网络模型中也占据重要地位. 相较于之前单纯使用一种损失函数( $L_2$  损失或  $L_1$  损失)的方法来看,多种损失函数的组合已表现出明显优势,能给图像带来更好的感知效果<sup>[125]</sup>. 但目前尚未出现标准的组合型损失函

数. 如何进行损失函数的有效组合得到最佳的重建结果,仍需继续探索研究.

(c) 评价指标. 图像质量评价指标作为衡量图像质量的重要方式之一,在 SR 图像的质量评价中具有重要地位. MOS 和 MOR 等常用的主观评价指标虽然能够较为准确地衡量图像的感知质量,但昂贵的人工成本及不可复现性,使该指标并没有被广泛应用. 尽管 PSNR 和 SSIM 等客观评价指标通过计算图像间的像素差距对图像质量进行衡量的方式,与主观评价指标存在反相关的关系,无法准确反映人对重建图像的主观评价,但从表 3、表 4 中可以看出,PSNR 和 SSIM 仍然是

目前主流的SR评价指标,同时,从表中可以发现,近年来的评价指标有从全参考型的客观评价指标向无参考型客观评价指标转变的趋势.由于现有的客观评价指标与主观评价指标存在一定的差异,所以探索一种符合人类视觉感受且便于使用的客观评价指标也将是一个重要的研究方向.

目前,评价网络模型视觉效果测试图像通常是采用一些自然图像,如动物、建筑、文字等,并通过放大图像,观察动物毛发、建筑框架、文字信息等图像细节来衡量重建图像的视觉效果,该方法从整体及局部对图像质量进行评估,虽然能够较为准确地评价图像重建质量,但对于一些视觉效果较为相似的图案就难以辨别其重建质量.因此,可以通过引入西门子星图(图22(a))、USAF-1951分辨率板(图22(b))等靶标图对网络模型的性能进行评估,并通过图形的线对关系等判断空间分辨率,从而更为准确地评估网络模型的性能.

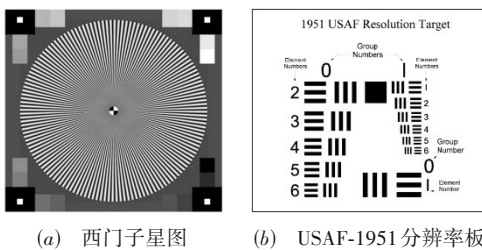


图22 不同类型靶标图

(d) 算法结合.目前,已有将SR算法与其他计算机视觉任务相结合共同处理图像问题的应用,如SCN<sup>[154]</sup>,SinIR<sup>[155]</sup>和IPT<sup>[94]</sup>等网络模型.图像的退化过程是由模糊、噪声等各种因素造成的,相较于单独训练网络处理去模糊、去噪等单个图像问题后再进行图像SR,组合处理的方式能够更有效地一次性处理各种图像问题,很大程度上节省了存储空间,提升了计算效率.因此,可以进一步研究SR算法与其他计算机视觉任务结合共同处理图像恢复问题的方式.

(e) 通用插件及框架.除了从整体的网络结构上进行改进实现特定网络模型的SR,也有不少研究人员研究通用的插件和框架旨,在从局部上对网络进行改进,从而实现网络性能的提升.如Kong等人<sup>[156]</sup>提出的一种利用数据特性加速SR网络的通用框架ClassSR,通过类模块将子图像按照难易程度分类,用不同分支处理不同难度的类,最终实现几乎所有基于学习的大图像SR方法的加速.又如Wang等人<sup>[157]</sup>利用条件卷积开发的插件模块,不仅可以处理任意尺度的SR问题,还可以处理非对称的SR问题.

## (2) 实际应用

一个好的网络模型不仅要具有良好的网络性能,且要能够将算法应用于实际才能真正起到作用.尽管现有的大多数网络具有良好的性能,但很多却无法应用于实际.因此现有的网络模型越来越重视实际的应用,追求高效、实时、轻量化.从表4所示的挑战赛发展趋势可以看出,SR算法正在从学术研究向工业产品应用转化.

(a) 注重网络效率,构建轻量化SR网络模型.从AIM 2019受限SR及AIM 2020高效SR的挑战赛中可以看出当前网络对实时性的需求,同时从评价指标的变化也可以看出实时性不仅是由参数量、运行时间以及浮点运算次数(FLOPs)等决定的,网络激活的数量等也具有一定的影响.所以衡量一个网络的效率不仅是从运行时间、参数量和FLOPs方面考虑,还要从其他方面考虑并进行全面的分析.近来,SR网络模型已实现较好的性能,但却难以实现实时性需求,因此往后的发展要注重轻量化网络模型的构建,使SR算法能够应用于实际.

(b) 注重网络质量,构建应用于真实场景的SR网络模型.从基于CNN到基于GAN方法的转变,是从PSNR目标最大化的SR向感知驱动的SR转变,同时也表明了感知SR的重要性.从表3、表4中训练数据集构建方式的转变可以看出,SR的发展逐渐从经典SR转向真实世界的SR,网络模型更加注重对现实场景SR的应用,可见能够实现真实场景图像而非合成场景图像的应用才是各项SR技术的最终落脚点.

(c) 注重网络应用,构建基于移动设备的SR网络模型.从Mobile AI 2021挑战赛中可以看出,目前超分的一个发展趋势是从硬件设备端转向移动设备端,使相关的SR网络模型能够应用于实际,实现算法的落地.如Ayazoglu等人<sup>[158]</sup>针对现有模型在移动设备端运行的限制因素对网络构建块加以改进,提出了一个极轻量化的超分辨率网络XLSR(Extremely Lightweight SR),在运行时间和模型参数等方面得到很大提升,并且能实现移动硬件端的高效运行,最终获得了Mobile AI 2021挑战赛的冠军.

(d) 特定领域的应用.目前大多数的SR网络模型主要用自然图像进行训练,虽具有一定的泛化性,但对于红外图像、遥感图像等特定场景图像的SR依旧存在局限性.如何将这此特定场景的先验知识与深度学习的框架结合起来应用于SR中是一个值得探索的方向.

## 6 总结

本文对基于深度学习的单帧图像超分辨率重建的相关知识及现有研究成果进行综述.随着单帧图像超

分辨率重建技术的发展,构建数据集的方式已逐渐从简单的双三次合成数据集向真实数据集以及大规模和无监督式合成数据集转变,旨在构建具有更大退化空间和更真实退化过程的退化模型.模型框架仍然以后上采样模型框架为主.评价指标虽有向无参考型客观评价指标转变的趋势,但目前仍以 PSNR 和 SSIM 为主.单帧图像超分辨率重建从最初的以 PSNR 最大化为目标到以感知驱动为目标再到以 Transformer 为主干的发展,是图像超分辨率重建算法的一大进步,也为后续超分辨率重建网络的发展提供了新思路,促进了图像超分辨率重建技术的进一步发展.目前,单帧图像超分辨率技术已在学术研究上取得了较好的成果,后续的研究重点会向算法的应用部署以及工业产品研发方面转移.

#### 参考文献

- [1] DAI D X, WANG Y J, CHEN Y H, et al. Is image super-resolution helpful for other vision tasks? [C]//2016 IEEE Winter Conference on Applications of Computer Vision. Lake Placid: IEEE, 2016: 1-9.
- [2] BAI Y C, ZHANG Y Q, DING M L, et al. SOD-MTGAN: A small object detection via multi-task generative adversarial network[C]//Proceedings of the European Conference on Computer Vision(ECCV 2018). Munich: Springer, 2018: 210-226.
- [3] BEI Y J, DAMIAN A, HU S J, et al. New techniques for preserving global structure and denoising with low information loss in single-image super-resolution[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). Salt Lake City: IEEE, 2018: 987-9877.
- [4] 吴秀秀,肖珊,张煜.基于配准的肺4D-CT图像超分辨率重建[J].电子学报,2015,43(2):383-386.  
WU X X, XIAO S, ZHANG Y. Registration based super-resolution reconstruction for lung 4D-CT image[J]. Acta Electronica Sinica, 2015, 43(2): 383-386. (in Chinese)
- [5] YIN Y, ROBINSON J, ZHANG Y L, et al. Joint super-resolution and alignment of tiny faces[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12693-12700.
- [6] HUANG Y W, ZHENG F, WANG D Y, et al. Super-resolution and inpainting with degraded and upgraded generative adversarial networks[C]//Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence. Yokohama: ACM, 2021: 645-651.
- [7] WANG Z Y, JIANG K, YI P, et al. Ultra-dense GAN for satellite imagery super-resolution[J]. Neurocomputing, 2020, 398: 328-337.
- [8] GAO G W, YU Y, XIE J, et al. Constructing multilayer locality-constrained matrix regression framework for noise robust face super-resolution[J]. Pattern Recognition, 2021, 110: 107539.
- [9] 李云红,王珍,张凯兵,等.基于学习的图像超分辨率重建方法综述[J].计算机工程与应用,2018,54(15):13-21.  
LI Y H, WANG Z, ZHANG K B, et al. Survey on example learning-based single image super-resolution technique[J]. Computer Engineering and Applications, 2018, 54(15): 13-21. (in Chinese)
- [10] 孙旭,李晓光,李嘉锋,等.基于深度学习的图像超分辨率复原研究进展[J].自动化学报,2017,43(5):697-709.  
SUN X, LI X G, LI J F, et al. Review on deep learning based image super-resolution restoration algorithms[J]. Acta Automatica Sinica, 2017, 43(5): 697-709. (in Chinese)
- [11] 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 Ying, ZHU Li, LIM K P, et al. Review and prospect of image super-resolution technology[J]. Journal of Frontiers of Computer Science and Technology, 2020, 14(2): 181-199. (in Chinese)
- [12] HARRIS J L. Diffraction and resolving power[J]. Journal of the Optical Society of America, 1964, 54(7): 931-936.
- [13] GOODMAN J W, COX M E. Introduction to fourier optics[J]. Physics Today, 1969, 22(4): 97-101.
- [14] TSAI R Y, HUANG T S. Multiframe image restoration and registration[J]. Advances in Computer Vision and Image Processing, 1984, 1(2): 317-339.
- [15] DUCHON C E. Lanczos filtering in one and two dimensions[J]. Journal of Applied Meteorology, 1979, 18(8): 1016-1022.
- [16] KEYS R. Cubic convolution interpolation for digital image processing[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1981, 29(6): 1153-1160.
- [17] STARK H, OSKOUI P. High-resolution image recovery from image-plane arrays, using convex projections[J]. Journal of the Optical Society of America. A, Optics and Image Science, 1989, 6(11): 1715-1726.
- [18] IRANI M, PELEG S. Improving resolution by image registration[J]. CVGIP: Graphical Models and Image Processing, 1991, 53(3): 231-239.
- [19] SCHULTZ R R, STEVENSON R L. A Bayesian approach to image expansion for improved definition[J]. IEEE Transactions on Image Processing, 1994, 3(3): 233-242.
- [20] 伍政华,孙明健,顾宗山,等.基于二阶广义方向性全变

- 分的图像超分辨率重建方法[J]. 电子学报, 2017, 45(11): 2625-2632.
- WU Z H, SUN M J, GU Z S, et al. Second-order directional total generalized variation regularization for image super-resolution[J]. *Acta Electronica Sinica*, 2017, 45(11): 2625-2632. (in Chinese)
- [21] 王相海, 赵晓阳, 毕晓昀, 等. 小波域多角度轮廓模板变换模型的单幅图像超分辨率重建[J]. 电子学报, 2018, 46(9): 2256-2262.
- WANG X H, ZHAO X Y, BI X Y, et al. Single image super-resolution reconstruction approach based on multi-angle contour templates variational calculus model in wavelet domain[J]. *Acta Electronica Sinica*, 2018, 46(9): 2256-2262. (in Chinese)
- [22] FREEMAN W T, PASZTOR E C, CARMICHAEL O T. Learning low-level vision[J]. *International Journal of Computer Vision*, 2000, 40: 25-47.
- [23] CHANG H, YEUNG D Y, XIONG Y M. Super-resolution through neighbor embedding[C]//Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington: IEEE, 2004: 275-282.
- [24] YANG J C, WRIGHT J, HUANG T, et al. Image super-resolution as sparse representation of raw image patches [C]//2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage: IEEE, 2008: 1-8.
- [25] TIMOFTE R, DE V, GOOL L V. Anchored neighborhood regression for fast example-based super-resolution [C]//2013 IEEE International Conference on Computer Vision. Sydney: IEEE, 2013: 1920-1927.
- [26] 潘宗序, 禹晶, 肖创柏, 等. 基于自适应多字典学习的单幅图像超分辨率算法[J]. 电子学报, 2015, 43(2): 209-216.
- PAN Z X, YU J, XIAO C B, et al. Single image super resolution based on adaptive multi-dictionary learning[J]. *Acta Electronica Sinica*, 2015, 43(2): 209-216. (in Chinese)
- [27] PARK S C, PARK M K, KANG M G. Super-resolution image reconstruction: A technical overview[J]. *IEEE Signal Processing Magazine*, 2003, 20(3): 21-36.
- [28] 江静, 张雪松. 图像超分辨率重建算法综述[J]. 红外技术, 2012, 34(1): 24-30.
- JIANG J, ZHANG X S. A review of super-resolution reconstruction algorithms[J]. *Infrared Technology*, 2012, 34(1): 24-30. (in Chinese)
- [29] 苏衡, 周杰, 张志浩. 超分辨率图像重建方法综述[J]. 自动化学报, 2013, 39(8): 1202-1213.
- SU H, ZHOU J, ZHANG Z H. Survey of super-resolution image reconstruction methods[J]. *Acta Automatica Sinica*, 2013, 39(8): 1202-1213. (in Chinese)
- [30] YANG W M, ZHANG X C, TIAN Y P, et al. Deep learning for single image super-resolution: A brief review[J]. *IEEE Transactions on Multimedia*, 2019, 21(12): 3106-3121.
- [31] 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. (in Chinese)
- [32] 张德, 林青宇, 郭茂祖. 单幅图像超分辨率重建的深度学习综述[J]. 计算机工程与应用, 2021, 57(22): 28-41.
- ZHANG D, LIN Q Y, GUO M Z. Review of single image super-resolution based on deep learning[J]. *Computer Engineering and Applications*, 2021, 57(22): 28-41. (in Chinese)
- [33] 黄健, 赵元元, 郭苹, 等. 深度学习的单幅图像超分辨率重建方法综述[J]. 计算机工程与应用, 2021, 57(18): 13-23.
- HUANG J, ZHAO Y Y, GUO P, et al. Survey of single image super-resolution based on deep learning[J]. *Computer Engineering and Applications*, 2021, 57(18): 13-23. (in Chinese)
- [34] YANG Z Y, SHI P, PAN D. A survey of super-resolution based on deep learning[C]//2020 International Conference on Culture-oriented Science & Technology(ICCST). Beijing: IEEE, 2020: 514-518.
- [35] 李佳星, 赵勇先, 王京华. 基于深度学习的单幅图像超分辨率重建算法综述[J]. 自动化学报, 2021, 47(10): 2341-2363.
- LI J X, ZHAO Y X, WANG J H. A review of single image super-resolution reconstruction algorithms based on deep learning[J]. *Acta Automatica Sinica*, 2021, 47(10): 2341-2363. (in Chinese)
- [36] 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. (in Chinese)
- [37] WANG Z H, CHEN J, HOI S C H. Deep learning for image super-resolution: A survey[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(10): 3365-3387.
- [38] ANWAR S, KHAN S, BARNES N. A deep journey into super-resolution[J]. *ACM Computing Surveys*, 2021, 53(3): 1-34.
- [39] CHEN H G, HE X H, QING L B, et al. Real-world single

- image super-resolution: A brief review[J]. *Information Fusion*, 2022, 79: 124-145.
- [40] AGUSTSSON E, TIMOFTE R. NTIRE 2017 challenge on single image super-resolution: Dataset and study[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu: IEEE, 2017: 1122-1131.
- [41] GU S H, LUGMAYR A, DANELLJAN M, et al. DIV8K: DIVERse 8K resolution image dataset[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Seoul: IEEE, 2019: 3512-3516.
- [42] CAI J R, ZENG H, YONG H W, et al. Toward real-world single image super-resolution: A new benchmark and a new model[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019: 3086-3095.
- [43] WEI P X, XIE Z W, LU H N, et al. Component divide-and-conquer for real-world image super-resolution[C]//Computer Vision-ECCV 2020. Glasgow: Springer, 2020: 101-117.
- [44] ZHANG K, ZUO W M, ZHANG L. Learning a single convolutional super-resolution network for multiple degradations[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 3262-3271.
- [45] GU J J, LU H N, ZUO W M, et al. Blind super-resolution with iterative kernel correction[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 1604-1613.
- [46] LUO Z X, HUANG Y, LI S, et al. Unfolding the alternating optimization for blind super resolution[EB/OL]. (2020-11-25)[2022-01-16]. <https://arxiv.org/abs/2010.02631>.
- [47] ZHANG K, LIANG J Y, VAN GOOL L, et al. Designing a practical degradation model for deep blind image super-resolution[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 4771-4780.
- [48] WANG X T, XIE L B, DONG C, et al. Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data[EB/OL]. (2021-08-17) [2022-01-16]. <https://arxiv.org/abs/2107.10833>.
- [49] BELL-KLIGLER S, SHOCHER A, IRANI M. Blind super-resolution kernel estimation using an Internal-GAN [C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems (NeurIPS). Vancouver: Curran Associates, Inc., 2019: 284-293.
- [50] FRITSCHÉ M, GU S H, TIMOFTE R. Frequency separation for real-world super-resolution[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Seoul: IEEE, 2019: 3599-3608.
- [51] JI X Z, TAO G P, CAO Y, et al. Frequency consistent adaptation for real world super resolution[EB/OL]. (2020-12-18)[2022-01-16]. <https://arxiv.org/abs/2012.10102>.
- [52] YANG J C, WRIGHT J, HUANG T S, et al. Image super-resolution via sparse representation[J]. *IEEE Transactions on Image Processing*, 2010, 19(11): 2861-2873.
- [53] TIMOFTE R, DE SMET V, VAN GOOL L. A+: Adjusted anchored neighborhood regression for fast super-resolution[C]//Computer Vision-ACCV 2014. Singapore: Springer, 2015: 111-126.
- [54] SCHULTER S, LEISTNER C, BISCHOF H. Fast and accurate image upscaling with super-resolution forests[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 3791-3799.
- [55] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 1646-1654.
- [56] DONG C, LOY C C, TANG X O. Accelerating the super-resolution convolutional neural network[C]//Computer Vision-ECCV 2016. Amsterdam: Springer, 2016: 391-407.
- [57] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common objects in context[C]//Computer Vision-ECCV 2014. Zurich: Springer, 2014: 740-755.
- [58] DENG J, DONG W, SOCHER R, et al. ImageNet: A large-scale hierarchical image database[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami: IEEE, 2009: 248-255.
- [59] TIMOFTE R, AGUSTSSON E, VAN GOOL L, et al. NTIRE 2017 challenge on single image super-resolution: Methods and results[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Honolulu: IEEE, 2017: 1110-1121.
- [60] JI X Z, CAO Y, TAI Y, et al. Real-world super-resolution via kernel estimation and noise injection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle: IEEE, 2020: 1914-1923.
- [61] IGNATOV A, KOBYSHEV N, TIMOFTE R, et al. DSLR-quality photos on mobile devices with deep convolutional networks[C]//2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 3297-3305.
- [62] WANG X T, YU K, DONG C, et al. Recovering realistic texture in image super-resolution by deep spatial feature transform[C]//2018 IEEE/CVF Conference on Computer

- Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 606-615.
- [63] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding[C]//Proceedings of the British Machine Vision Conference 2012. Guildford: BMVA Press, 2012: 135.1-135.10
- [64] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations[C]//Proceedings of the 7th International Conference on Curves and Surfaces. Avignon: Springer, 2012: 711-730.
- [65] 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 Eighth IEEE International Conference on Computer Vision(ICCV). Vancouver: IEEE, 2001: 416-423.
- [66] HUANG J B, SINGH A, AHUJA N. Single image super-resolution from transformed self-exemplars[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 5197-5206.
- [67] 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. Cancun: ACM, 2016: (2) 1-5.
- [68] BLAU Y, MECHREZ R, TIMOFTE R, et al. The 2018 PIRM challenge on perceptual image super-resolution[C]//Lecture Notes in Computer Science. Munich: Springer, 2019: 334-355.
- [69] DONG C, LOY C C, HE K M, et al. Learning a deep convolutional network for image super-resolution[C]//Computer Vision-ECCV 2014. Zurich: Springer, 2014: 184-199.
- [70] DONG C, LOY C C, HE K M, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295-307.
- [71] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 1637-1645.
- [72] HU X C, MU H Y, ZHANG X Y, et al. Meta-SR: A magnification-arbitrary network for super-resolution[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach: IEEE, 2019: 1575-1584.
- [73] SHI W Z, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 1874-1883.
- [74] FAN Y C, SHI H H, YU J H, et al. Balanced two-stage residual networks for image super-resolution[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu: IEEE, 2017: 1157-1164.
- [75] ZHANG Y, LI K, LI K, et al. Residual non-local attention networks for image restoration[C]//Proceedings of the Seventh International Conference on Learning Representations(ICLR 2019). New Orleans: ICLR Press, 2019: 1-18.
- [76] LAI W S, HUANG J B, AHUJA N, et al. Deep laplacian pyramid networks for fast and accurate super-resolution [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Honolulu: IEEE, 2017: 5835-5843.
- [77] LAI W S, HUANG J B, AHUJA N, et al. Fast and accurate image super-resolution with deep Laplacian pyramid networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(11): 2599-2613.
- [78] CHUDASAMA V, UPLA K. E-ProSRNet: An enhanced progressive single image super-resolution approach[J]. Computer Vision and Image Understanding, 2020, 200: 103038.
- [79] HARIS M, SHAKHAROVICH G, UKITA N. Deep back-projection networks for super-resolution[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 1664-1673.
- [80] HARIS M, SHAKHAROVICH G, UKITA N. Deep back-Projection Networks for single image super-resolution[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(12): 4323-4337.
- [81] HAN W, CHANG S Y, LIU D, et al. Image super-resolution via dual-state recurrent networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 1654-1663.
- [82] LI Z, YANG J L, LIU Z, et al. Feedback network for image super-resolution[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach: IEEE, 2019: 3862-3871.
- [83] LUGMAYR A, DANELLJAN M, TIMOFTE R, et al. NTIRE 2020 challenge on real-world image super-resolution: Methods and results[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). Seattle: IEEE, 2020: 2058-2076.
- [84] STREIJL R C, WINKLER S, HANDS D S. Mean opinion score(MOS) revisited: Methods and applications, limi-

- tations and alternatives[J]. *Multimedia Systems*, 2016, 22(2): 213-227.
- [85] 张凯兵, 朱丹妮, 王珍, 等. 超分辨率图像质量评价综述[J]. *计算机工程与应用*, 2019, 55(4): 31-40, 47.  
ZHANG K B, ZHU D N, WANG Z, et al. Survey of super-resolution images quality assessment[J]. *Computer Engineering and Applications*, 2019, 55(4): 31-40, 47. (in Chinese)
- [86] 高敏娟, 党宏社, 魏立力, 等. 全参考图像质量评价回顾与展望[J]. *电子学报*, 2021, 49(11): 2261-2272.  
GAO M J, DANG H S, WEI L L, et al. Review and prospect of full reference image quality assessment[J]. *Acta Electronica Sinica*, 2021, 49(11): 2261-2272. (in Chinese)
- [87] 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.
- [88] SHEIKH H R, BOVIK A C, DE VECIANA G. An information fidelity criterion for image quality assessment using natural scene statistics[J]. *IEEE Transactions on Image Processing*, 2005, 14(12): 2117-2128.
- [89] ZHANG R, ISOLA P, EFROS A A, et al. The unreasonable effectiveness of deep features as a perceptual metric[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 586-595.
- [90] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a “completely blind” image quality analyzer[J]. *IEEE Signal Processing Letters*, 2013, 20(3): 209-212.
- [91] VENKATANATH N, PRANEETH D, BH M C, et al. Blind image quality evaluation using perception based features[C]//2015 Twenty First National Conference on Communications(NCC). Mumbai: IEEE, 2015: 1-6.
- [92] MA C, YANG C Y, YANG X K, et al. Learning a no-reference quality metric for single-image super-resolution[J]. *Computer Vision and Image Understanding*, 2017, 158: 1-16.
- [93] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 105-114.
- [94] CHEN H T, WANG Y H, GUO T Y, et al. Pre-trained image processing transformer[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021: 12294-12305.
- [95] HUBEL D H, WIESEL T N. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex[J]. *The Journal of Physiology*, 1962, 160(1): 106-154.
- [96] LI Z W, LIU F, YANG W J, et al. A survey of convolutional neural networks: Analysis, applications, and prospects[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021: 1-21.
- [97] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.
- [98] MAO X, SHEN C, YANG Y. Image restoration using convolutional auto-encoders with symmetric skip connections[C]//Proceedings of the Advances in Neural Information Processing Systems 29(NIPS 2016) [C]. Barcelona: Curran Associates, Inc., 2016.
- [99] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). Honolulu: IEEE, 2017: 1132-1140.
- [100] LI J, FANG F, MEI K, et al. Multi-scale residual network for image super-resolution[C]//Proceedings of the European Conference on Computer Vision. Munich: Springer, 2018: 527-542.
- [101] LAN R S, SUN L, LIU Z B, et al. Cascading and enhanced residual networks for accurate single-image super-resolution[J]. *IEEE Transactions on Cybernetics*, 2021, 51(1): 115-125.
- [102] 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. (in Chinese)
- [103] 杨祎玥, 伏潜, 万定生. 基于深度循环神经网络的时间序列预测模型[J]. *计算机技术与发展*, 2017, 27(3): 35-38, 43.  
YANG Y Y, FU Q, WAN D S. A prediction model for time series based on deep recurrent neural network[J]. *Computer Technology and Development*, 2017, 27(3): 35-38, 43. (in Chinese)
- [104] TAI Y, YANG J, LIU X M. Image super-resolution via deep recursive residual network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 2790-2798.
- [105] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Honolulu: IEEE, 2017: 2261-2269.

- [106] TONG T, LI G, LIU X J, et al. Image super-resolution using dense skip connections[C]//2017 IEEE International Conference on Computer Vision(ICCV). Venice: IEEE, 2017: 4809-4817.
- [107] TAI Y, YANG J, LIU X M, et al. MemNet: A persistent memory network for image restoration[C]//2017 IEEE International Conference on Computer Vision(ICCV). Venice: IEEE, 2017: 4549-4557.
- [108] SHAMSOLMOALI P, ZAREAPOOR M, ZHANG J H, et al. Image super resolution by dilated dense progressive network[J]. *Image and Vision Computing*, 2019, 88: 9-18.
- [109] PAN Z, LI B, XI T, et al. Real image super resolution via heterogeneous model ensemble using GP-NAS[C]// Proceedings of the European Conference on Computer Vision(ECCV). Glasgow: Springer, 2020: 423-436.
- [110] JIANG K, WANG Z Y, YI P, et al. Hierarchical dense recursive network for image super-resolution[J]. *Pattern Recognition*, 2020, 107: 107475.
- [111] CHAUDHARI S, MITHAL V, POLATKAN G, et al. An attentive survey of attention models[J]. *ACM Transactions on Intelligent Systems and Technology*, 2021, 12 (5): 53(1-32).
- [112] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020,42(8): 2011-2023.
- [113] 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: Springer, 2018: 294-310.
- [114] DAI T, CAI J R, ZHANG Y B, et al. Second-order attention network for single image super-resolution[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach: IEEE, 2019: 11057-11066.
- [115] AHN N, KANG B, Fast SOHN K., accurate, and lightweight super-resolution with cascading residual network [C]//Proceedings of the European Conference on Computer Vision(ECCV). Munich: Springer, 2018: 256-272.
- [116] HUI Z, WANG X M, GAO X B. Fast and accurate single image super-resolution via information distillation network[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 723-731.
- [117] HUI Z, GAO X, YANG Y, et al. Lightweight image super-resolution with information multi-distillation network[C]//Proceedings of the 27th ACM International Conference on Multimedia. Nice: ACM, 2019: 2024-2032.
- [118] LIU J, TANG J, WU G. Residual feature distillation network for lightweight image super-resolution[C]//Proceedings of the European Conference on Computer Vision(ECCV). Glasgow: Springer, 2020: 41-55.
- [119] CHU X X, ZHANG B, MA H L, et al. Fast, accurate and lightweight super-resolution with neural architecture search[C]//2020 25th International Conference on Pattern Recognition(ICPR). Milan: IEEE, 2021: 59-64.
- [120] LI W B, ZHOU K, QI L, et al. LAPAR: Linearly-assembled pixel-adaptive regression network for single image super-resolution and beyond[C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver: ACM, 2020: 20343-20355.
- [121] WANG L G, DONG X Y, WANG Y Q, et al. Exploring sparsity in image super-resolution for efficient inference [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Nashville: IEEE, 2021: 4915-4924.
- [122] RADOSAVOVIC I, KOSARAJU R P, GIRSHICK R, et al. Designing network design spaces[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Seattle: IEEE, 2020: 10425-10433.
- [123] ZHANG K, DANELLJAN M, LI Y, et al. AIM 2020 challenge on efficient super-resolution: methods and results[C]//Proceedings of the European Conference on Computer Vision(ECCV). Glasgow: Springer, 2020: 5-40.
- [124] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems(NIPS 2014). Montreal: MIT Press, 2014: 2672-2680.
- [125] SAJJADI M S M, SCHÖLKOPF B, HIRSCH M. EnhanceNet: single image super-resolution through automated texture synthesis[C]//2017 IEEE International Conference on Computer Vision(ICCV). Venice: IEEE, 2017: 4501-4510.
- [126] WANG X, YU K, WU S, et al. ESRGAN: enhanced super-resolution generative adversarial networks[C]//Proceedings of the European Conference on Computer Vision(ECCV). Munich: Springer, 2018: 63-79.
- [127] WANG X T, YU K, DONG C, et al. Deep network interpolation for continuous imagery effect transition[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach: IEEE, 2019: 1692-1701.
- [128] SHANG T Z, DAI Q J, ZHU S C, et al. Perceptual ex-

- treme super resolution network with receptive field block [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). Seattle: IEEE, 2020: 1778-1787.
- [129] ZHANG W L, LIU Y H, DONG C, et al. RankSRGAN: generative adversarial networks with ranker for image super-resolution[C]//2019 IEEE/CVF International Conference on Computer Vision(ICCV). Seoul: IEEE, 2019: 3096-3105.
- [130] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017: 6000-6010.
- [131] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: Transformers for image recognition at scale[C]//Proceedings of the International Conference on Learning Representations(ICLR 2021). London: ICLR Press, 2021: 1-22.
- [132] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers[C]//Proceedings of the European Conference on Computer Vision(ECCV 2020). Glasgow: Springer, 2020: 213-229.
- [133] ARNAB A, DEGHANI M, HEIGOLD G, et al. ViViT: A video vision transformer[C]//2021 IEEE/CVF International Conference on Computer Vision(ICCV). Montreal: IEEE, 2021: 6816-6826.
- [134] BI J R, ZHU Z L, MENG Q L. Transformer in computer vision[C]//2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology(CEI). Fuzhou: IEEE, 2021: 178-188.
- [135] LIANG J Y, CAO J Z, SUN G L, et al. SwinIR: image restoration using swin transformer[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops(ICCVW). Montreal: IEEE, 2021: 1833-1844.
- [136] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//2021 IEEE/CVF International Conference on Computer Vision(ICCV). Montreal: IEEE, 2021: 9992-10002.
- [137] LU Z, LIU H, LI J, et al. Efficient transformer for single image super-resolution[EB/OL]. (2021-10-28)[2022-01-16]. <https://arxiv.org/abs/2108.11084>.
- [138] SHOCHER A, COHEN N, IRANI M. Zero-shot super-resolution using deep internal learning[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Salt Lake City: IEEE, 2018: 3118-3126.
- [139] 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 15th European Conference on Computer Vision(ECCV). Munich: Springer, 2018: 187-202.
- [140] YUAN Y, LIU S Y, ZHANG J W, et al. Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). Salt Lake City: IEEE, 2018: 814-81409.
- [141] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 2242-2251.
- [142] ZHANG Y B, LIU S Y, DONG C, et al. Multiple cycle-in-cycle generative adversarial networks for unsupervised image super-resolution[J]. IEEE Transactions on Image Processing, 2020, 29: 1101-1112.
- [143] WU S, DONG C, QIAO Y. Blind image restoration based on cycle-consistent network[J/OL]. IEEE Transactions On Multimedia, 2022. DOI: 10.1109/TMM.2021.3139209.
- [144] TIMOFTE R, ROTHE R, VAN GOOL L. Seven ways to improve example-based single image super resolution [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Las Vegas: IEEE, 2016: 1865-1873.
- [145] TIMOFTE R, GU S H, WU J Q, et al. NTIRE 2018 challenge on single image super-resolution: Methods and results[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). Salt Lake City: IEEE, 2018: 965-96511.
- [146] CAI J R, GU S H, TIMOFTE R, et al. NTIRE 2019 challenge on real image super-resolution: Methods and results[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). Long Beach: IEEE, 2019: 2211-2223.
- [147] LUGMAYR A, DANELLJAN M, TIMOFTE R, et al. AIM 2019 challenge on real-world image super-resolution: Methods and results[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop(ICCVW). Seoul: IEEE, 2019: 3575-3583.
- [148] ZHANG K, GU S H, TIMOFTE R, et al. AIM 2019 challenge on constrained super-resolution: Methods and results[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop(ICCVW). Seoul: IEEE, 2019: 3565-3574.
- [149] GU S H, DANELLJAN M, TIMOFTE R, et al. AIM

2019 challenge on image extreme super-resolution: Methods and results[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Seoul: IEEE, 2019: 3556-3564.

- [150] ZHANG K, GU S, TIMOFTE R, et al. NTIRE 2020 challenge on perceptual extreme super-resolution: Methods and results[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle: IEEE, 2020: 2045-2057.
- [151] WEI P, LU H, TIMOFTE R, et al. AIM 2020 challenge on real image super-resolution: methods and results[C]//Proceedings of the European Conference on Computer Vision (ECCV). Berlin: Springer, 2020: 392-422.
- [152] BHAT G, DANELLJAN M, TIMOFTE R, et al. NTIRE 2021 challenge on burst super-resolution: Methods and results[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Nashville: IEEE, 2021: 613-626.
- [153] IGNATOV A, TIMOFTE R, DENNA M, et al. Real-time quantized image super-resolution on mobile NPUs, mobile AI 2021 challenge: Report[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Nashville: IEEE, 2021: 2525-2534.
- [154] FAN Y C, YU J H, LIU D, et al. Scale-wise convolution for image restoration[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 10770-10777.
- [155] YOO J, CHEN Q. SinIR: Efficient general image manipulation with single image reconstruction[C]//Proceedings of the 38th International Conference on Machine Learning (ICML). Champaign: PMLR, 2021: 12040-12050.
- [156] KONG X T, ZHAO H Y, QIAO Y, et al. ClassSR: A general framework to accelerate super-resolution networks by data characteristic[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021: 12011-12020.
- [157] WANG L G, WANG Y Q, LIN Z P, et al. Learning a single network for scale-arbitrary super-resolution[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 4781-4790.
- [158] AYAZOGLU M. Extremely lightweight quantization robust real-time single-image super resolution for mobile devices[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Nashville: IEEE, 2021: 2472-2479.

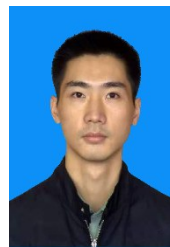
### 作者简介



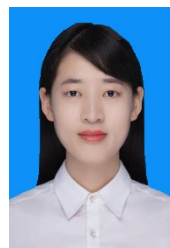
吴靖男, 1986年生. 博士. 福州大学机械工程及自动化学院副教授、硕士生导师. 主要研究方向为计算机视觉、光机电算一体化.  
E-mail: wujing@fzu.edu.cn



叶晓晶女, 1998年生. 福州大学机械工程及自动化学院硕士研究生. 主要研究方向为图像处理、图像超分辨率重建.  
E-mail: yexiaojingyxj@163.com



黄峰(通讯作者)男, 1979年生. 博士. 福州大学机械工程及自动化学院研究员、博士生导师. 主要研究方向为机器视觉、光学精密测量与成像、智能光学感知.  
E-mail: huangf@fzu.edu.cn



陈丽琼女, 1994年生. 博士. 福州大学机械工程及自动化学院讲师、硕士生导师. 主要研究方向为计算机视觉、超分辨率重建、目标检测.  
E-mail: liqiongchen@fzu.edu.cn



王志锋男, 1996年生. 福州大学机械工程及自动化学院硕士研究生. 主要研究方向为深度学习、图像超分辨率重建、图像增强处理.  
E-mail: 15880277679@139.com



刘文犀男, 1986年生. 博士. 福州大学计算机与大数据学院教授、博士生导师. 主要研究方向为计算机视觉.  
E-mail: wenxi.liu@hotmail.com