



DOI:10.13364/j.issn.1672-6510.20230129

基于对比学习的深度残差网络图像超分辨率方法

陈亚瑞, 徐肖阳

(天津科技大学人工智能学院, 天津 300457)

摘要: 传统的基于对比学习的图像超分辨率方法, 一般将原始图像作为正样本, 将退化图像或其他类图像作为负样本, 存在对纹理细节恢复差的问题。本文提出基于对比学习的深度残差网络图像超分辨率 (depth residual image super-resolution based on contrast learning, CEDSR) 方法, 针对残差超分辨率模型, 采用对高分辨率图像锐化后的图像作为正样本, 对高分辨率图像轻微模糊的图像作为负样本, 利用正负样本下的对比损失提升对纹理细节的恢复增强。增强锐化后的正样本图像携带更丰富的纹理信息, 基于不同函数生成的模糊负样本图像刻画了纹理模糊特征, 正负样本构建的对比损失有利于图像超分辨率图像对纹理细节的恢复。本文模型在 4 个标准数据集 DIV2K、Set14、BSDS100 和 Urban100 上与经典算法进行实验对比, 定性和定量实验结果均表明本文模型可以获得效果更好的超分辨率图像。

关键词: 图像超分辨率; 对比学习; 残差网络

中图分类号: TP319

文献标志码: A

文章编号: 1672-6510(2024)03-0072-09

Depth Residual Image Super-Resolution Based on Contrast Learning

CHEN Yarui, XU Xiaoyang

(College of Artificial Intelligence, Tianjin University of Science & Technology, Tianjin 300457, China)

Abstract: The traditional image super-resolution method based on contrast learning generally takes the original image as the positive sample and the degraded image or other types of images as the negative sample, which has the problem of poor texture detail restoration. In this article, we propose a depth residual image super-resolution based on contrast learning (CEDSR) method. In our proposed method, for the residual super-resolution model, the sharpened images of high-resolution images are used as positive samples, and the slightly blurred images of high-resolution images are used as negative samples. The contrast loss lifting under positive and negative samples is used to restore and enhance the texture details. The positive sample images after enhancement and sharpening carry more abundant texture information, and the fuzzy negative sample images generated based on different functions depict the texture fuzzy features. The contrast loss of positive and negative samples is conducive to the restoration of texture details in super-resolution images. Our proposed model was compared with the classical algorithms on DIV2K, Set14, BSDS100 and Urban100 standard data sets. The qualitative and quantitative experimental results show that the model has better super-resolution image.

Key words: image super-resolution; contrast learning; residual network

引文格式:

陈亚瑞, 徐肖阳. 基于对比学习的深度残差网络图像超分辨率方法[J]. 天津科技大学学报, 2024, 39(3): 72-80.

CHEN Y R, XU X Y. Depth residual image super-resolution based on contrast learning[J]. Journal of Tianjin university of science & technology, 2024, 39(3): 72-80.

单图像超分辨率 (single image super-resolution, SISR) 是指从低分辨率 (low resolution, LR) 图像恢复

成高分辨率 (high resolution, HR) 图像的过程。图像超分辨率是计算机视觉中提高图像分辨率的一类重

收稿日期: 2023-06-11; 修回日期: 2023-08-18

作者简介: 陈亚瑞(1982—), 女, 河北邢台人, 教授, yrchen@tust.edu.cn

要的图像处理技术, 受到了全球范围的高度关注, 它在现实世界中有着广泛的应用。图像超分辨率重建在医学成像、卫星图像处理、安防监控以及人脸识别等领域发挥了非常重要的作用。基于传统的插值法是早期应用的一种图像超分辨率重建方法, 通过图像像素之间的插值实现图像的放大或缩小。

传统的插值法可以分为双线性插值法^[1]、双三次插值法^[2]、重采样法^[3]等。传统的插值法可以处理任意大小的图像, 并且精度高、计算速度快, 但是生成图像的边缘或纹理细节往往较为模糊。随着深度学习技术的快速发展, 基于深度学习的图像超分辨率方法得到了积极探索, 各种深度学习方法已被应用到图像超分辨率任务中。基于深度学习的图像超分辨率神经网络主要包括卷积神经网络^[4](convolutional neural network, CNN)、生成对抗网络^[5](generative adversarial network, GAN)、残差网络^[6]等。

1 图像超分辨率方法

1.1 基于卷积神经网络的图像超分辨率

卷积神经网络通过学习输入图像的特征, 获得更好的图像特征表达, 因此被广泛应用于计算机视觉任务。Dong 等^[7]将卷积神经网络应用到图像超分辨率重建任务中, 提出了一种基于卷积神经网络的图像超分辨率模型(SRCNN)。相对于传统的图像超分辨率方法, SRCNN 模型能够通过卷积神经网络学习到更多的图像特征信息, 从而能够还原更多的图像细节, 得到更加清晰、自然的图像。

SRCNN 需要大量的数据集和计算资源训练模型, Dong 等^[8]改进了 SRCNN 模型, 提出了一种快速超分辨率卷积神经网络模型(FSRCNN), FSRCNN 网络的输入无须将 LR 图像放大, 仅将 LR 图像作为输入, 网络的最后一层将反卷积作为上采样层, 大大减少时间损耗。Kim 等^[9]提出了超分辨率重建极深的卷积神经网络模型(VDSR), 相比于 SRCNN 模型, VDSR 模型使用了更深的神经网络, 可以提取图像更多的特征信息, 使重建后图像的纹理细节更加丰富。

1.2 基于生成对抗网络的图像超分辨率

基于卷积神经网络的图像超分辨率都是以峰值信噪比(peak signal-to-noise ratio, PSNR)作为一项重要的评价指标。随着网络不断地训练, PSNR 值可能会越来越高, 但是生成的图像在视觉感知质量上较差, 无法恢复图像真实的纹理和细节。为了生成符合

人类感官质量的图像, Ledig 等^[10]提出了基于生成对抗网络的图像超分辨率模型(SRGAN)。SRGAN 模型的感知损失函数能够更好地评估图像的视觉质量, 还可以更好地保留图像的细节信息。尽管 SRGAN 生成的图像在视觉上有较好的效果, 但是随着网络的加深, 可能会使图像出现伪影。Wang 等^[11]对 SRGAN 模型深入研究, 提出了一种增强的生成对抗网络图像超分辨率模型(ESRGAN)。ESRGAN 模型在 SRGAN 模型的基础上添加了残差密集块, 并且将原来的绝对判别器改为相对判别器, 即判别器收到生成器生成的图像时, 不仅只是预测这幅图像是否真实的概率, 而且预测这幅图像相对于假的图像更真实的概率。SRGAN 模型是用激活后的特征计算感知损失, 而 ESRGAN 模型是在激活前计算感知损失, 激活前图像特征更加密集, 用激活前的特征计算感知损失会得到更多的图像信息。与 SRGAN 模型相比, ESRGAN 模型重建后的图像在视觉上看起来更加自然。

1.3 基于残差网络的图像超分辨率

随着神经网络层数的增加, 网络更加难以训练, He 等^[12]提出一种残差网络, 残差网络通过堆叠多个残差块实现, 残差块可以有效地传递图像间的信息。传统的神经网络训练过程中容易出现梯度消失的问题, 而残差网络则能够通过跨层连接保留网络中的梯度信息, 从而解决梯度消失的问题。残差网络中的跨层连接使网络在训练过程中可以更快收敛, 在训练轮数相同时可以实现更好的模型生成效果。近年来, 基于残差网络的图像超分辨率模型已经被广泛应用。Lim 等^[13]提出基于增强深度残差网络的图像超分辨率(EDSR)。EDSR 和 SRGAN 这两种模型主要采用的是残差网络, 其残差块结构如图 1 所示。

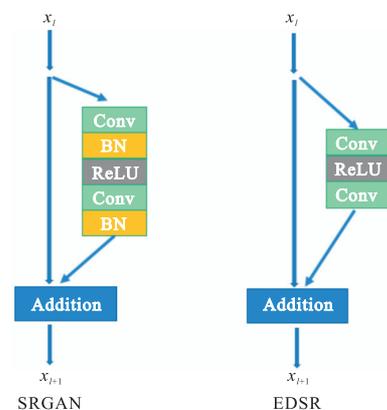


图 1 残差块结构图

Fig. 1 Residual block structure diagram

EDSR 模型通过删减多余的批归一化层,有效减少了对残差图像的处理,提升了网络的训练效率。

1.4 其他方法

目前,零样本学习、元学习、对比学习在图像超分辨率中也有所应用。2018年,Shocher等^[14]提出了一种无监督学习的图像超分辨率模型(ZSSR)。传统的图像超分辨率任务需要大量的数据作为支撑,ZSSR模型中的零样本学习为超分辨率任务提供了一种新的解决方案,并证明了单张图像内部块之间的信息熵小于多张图像块之间的信息熵,即图像内部块比多张图像块拥有更强的预测能力。2020年,Soh等^[15]提出MZSR(meta-transfer learning for zero-shot super-resolution)模型,该模型是在ZSSR模型的基础上引入元学习。由于ZSSR模型只关注LR图像本身的内部信息,并没有利用外部信息,因此MZSR模型在ZSSR模型的基础上加入了元学习,可以同时利用内部信息和外部信息,并且在样本较少的情况下经过几次迭代就能得到一个较好的实验结果。

当前,基于对比学习的下游任务方法主要集中于将模糊图像作为负样本,而让原始图像作为正样本。Wu等^[16]将退化图像(输入模糊图像)视为负样本,提出了一种具有对比正则化的新图像去模糊方法。Wang等^[17]将数据集中的其他示例作为图像超分辨率的负样本,这种方法进行了初步尝试,并证明了对比损失在下游任务的有效性。但是,简单地将退化图像或者数据集中的其他图像作为负样本,这些负样本与重建图像不同,并且容易区分,即它们相距太远而不会产生对比损失,忽略了为重建图像构建有效正负样本的重要性。因此,本文提出一种有效的数据增强方法,即HR图像通过锐化构建带有丰富纹理信息的正样本,HR图像进行轻微模糊处理构建负样本。基于深度学习的方法导致超分辨率重建结果趋于平滑,通过对比学习构建正负样本的方式,能得到更多纹理信息,从而使生成图像远离平滑结果。

EDSR模型将原始图像下采样得到的LR图像作为输入,并没有充分利用原始图像的高频信息(例如图像的纹理细节等),从而导致生成图像的纹理效果不理想。对比学习框架所构建的正样本能充分利用图像所包含的高频信息,在对比损失的约束下,使生成图像具有更好的纹理细节。因此,本文在EDSR模型的基础上引入对比学习框架,构建有效的正负样本。本文创新点如下:

(1) 移除EDSR网络中多余的卷积层。由于残差

模块中已经具备卷积层的作用,所以去除冗余卷积层会使计算更快,且不会影响图像的整体质量。

(2) 构建对比学习的正负样本对,对HR图像进行模糊操作得到负样本,将HR图像锐化后的样本作为正样本。

(3) 基于改进后的EDSR网络与构建好的正负样本对,通过对比损失进一步约束超分辨率(super-resolution, SR)图像向正样本靠近,保证生成的图像能够具有更好的纹理细节。

2 本文方法

本文提出基于对比学习的深度残差网络图像超分辨率(depth residual image super-resolution based on contrast learning, CEDSR)方法。在EDSR模型中引入对比学习框架,通过对原始图像进行数据增强,将HR图像锐化后得到的图像作为正样本,将HR图像轻微模糊后得到的图像作为负样本,通过特征提取后,在潜在空间对比损失的约束下,生成纹理细节更好的超分辨率图像。首先将详细介绍本文提出的网络结构,然后介绍在对比学习下的正负样本的生成方法,最后通过实验表明本文的对比学习方法能进一步提高图像超分辨率的性能。

2.1 CEDSR网络结构

CEDSR网络结构如图2所示,主要由两部分组成,分别为生成网络和对比网络。生成网络如图2左边红色虚线框所示,对比网络如图2右边绿色虚线框所示。

生成网络包括特征提取模块、残差模块和重建模块。特征提取模块使用1组卷积层提取图像的低频信息,其中卷积核大小为 5×5 。残差模块由若干个残差块组成,每个残差块由若干个包含残差连接的卷积层、ReLU层和残差缩放组成。卷积层用于对图像进行特征增强。由于较多的残差块会导致网络训练不稳定,因此采取残差缩放的方法,保证网络更加稳定。重建模块由卷积层和像素重组组成,使用1组卷积层将增强后的特征图像转换为高分辨率图像。

对比网络包括两个模块:生成样本模块和特征提取模块。生成样本模块包括生成正样本和负样本,将HR图像锐化后生成带有丰富纹理信息的图像作为正样本,将HR图像进行轻微模糊后的图像作为负样本。本文的特征提取模块使用预先训练的VGG网络^[18]处理图像特征。VGG网络由16个卷积层和3个

全连接层组成,使用 3×3 的卷积核,同时采用了深度相同、尺寸不同的卷积层结构进行图像的特征提取。正样本提取后的特征为 f ,负样本提取后的特征为 p ,SR 图像提取后的特征为 n ,在对比损失 L_{CL} 的约束下,希望超分辨率图像的特征 n 向正样本特征 f 靠近,同时远离负样本特征 p 。

LR 图像通过生成网络生成 SR 图像,生成的 SR

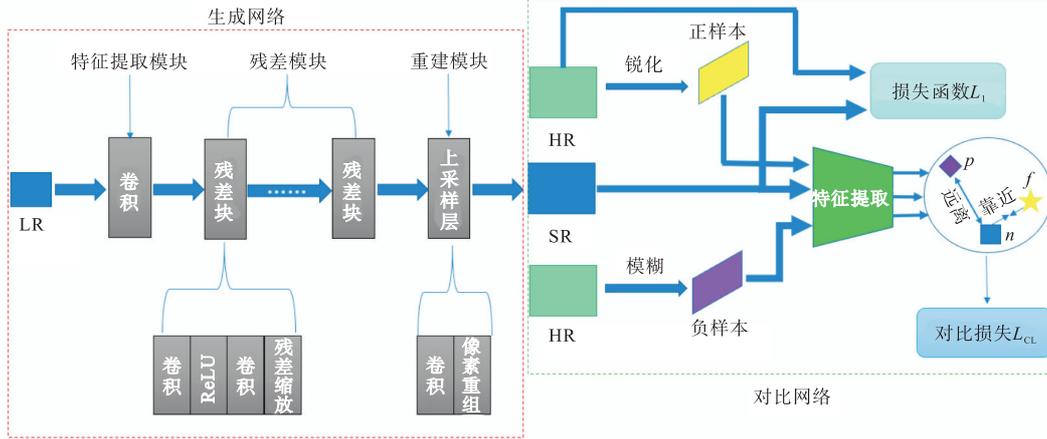


图2 CEDSR网络结构

Fig. 2 CEDSR network structure

2.2 对比学习下的正负样本构建

对比学习是一种无监督学习方式,其主要思想就是通过最大化目标样本与正样本之间的相似度、最小化目标样本与负样本之间的相似度,让模型学会区分不同的数据样本,从而实现无监督学习。

单图像超分辨率任务的目标是将 LR 图像转换为清晰、逼真且分辨率高的 HR 图像。由于 LR 图像通常是从 HR 图像经过双三次下采样得到的,仅包含低频信息,而图像超分辨率模型专注于学习恢复丢失的高频分量(例如边缘和纹理信息)。考虑到这一点,本文采用有效数据增强生成正样本和负样本,通过在 HR 图像上应用高通核生成锐化图像作为正样本集 P 。对于图像 I^{HR} ,将其正样本集表示为

$$P = \text{Sharpen}(I^{HR}) \quad (1)$$

其中 Sharpen 为锐化函数。图像锐化用于加强图像中的边缘和细节信息,由于边缘和细节信息对应图像的高频信息,通常对图像进行高通滤波锐化操作。

$$H(u, v) = \begin{cases} 0, & D(u, v) \leq D_0 \\ 1, & D(u, v) > D_0 \end{cases} \quad (2)$$

$$D(u, v) = \sqrt{u^2 + v^2} \quad (3)$$

其中: D_0 表示截止频率, u 表示图像在 x 方向上的频率, v 表示图像在 y 方向上的频率。图像的高频区域

图像与 HR 图像构建损失函数 L_1 ,该损失函数是衡量 SR 图像与原始图像 HR 之间差异的指标。生成的 SR 图像与构建的正负样本对通过特征提取后,在潜在空间对比损失的约束下,使 SR 图像更接近于正样本,远离负样本,使生成网络能够生成纹理细节更好的 SR 图像,从而进一步提高图像超分辨率的效果。

代表像素值剧烈变化(如图像中的纹理细节),而低频区域则描述图像的整体结构。高频信息超过了截止频率就进行保留,低频信息低于截止频率就舍弃。

现有的基于对比学习的图像超分辨率方法仅将原始图像作为正样本,这样并未充分利用图像超分辨率方法重建图像所需要的高频信息。图像超分辨率重建的目的是获得纹理效果好的图像,因此可以使用一些信息丰富的正样本作为重建结果而引入更多的高频细节。锐化图像如图 3 所示,相比于原始图像,锐化图像的纹理细节更加突出。

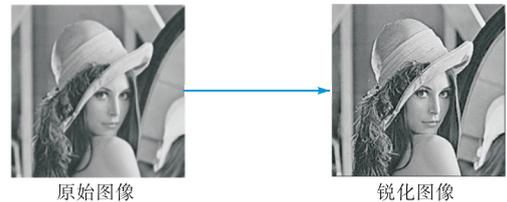


图3 锐化图像

Fig. 3 Sharpened image

现有的将对比学习引入图像超分辨率方法简单地将数据集中的退化图像(如输入 LR 图像)作为负样本,当与目标重建图像进行对比时,这些负样本容易区分,与目标样本距离太远,从而不会产生对比损失。本文将 HR 图像生成轻微模糊图像作为负样本集

N , 因为它们与原始图像相似, 所以能够迫使重建的 SR 图像变得更接近于原始图像。图像 I^{HR} 对应的负样本集为

$$N = \text{Blur}(I^{HR}) \quad (4)$$

其中 Blur 表示随机高斯核的模糊函数 Gaussian Blur , 其卷积核大小为 3×3 , 标准差为 0。轻微模糊图像如图 4 所示, 轻微模糊图像较好地保留了原始图像的信息。

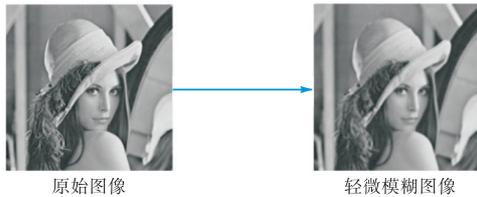


图 4 轻微模糊图像

Fig. 4 Slightly blurred image

2.3 损失函数

在对比学习方法中引入图像超分辨率方法时, 将原始图像分别进行锐化和轻微模糊后的图像作为正负样本。为了充分利用这些样本, 利用特征提取网络中的特征进行对比。对于 HR 图像, 其通过锐化和轻微模糊生成的正负样本集分别记为 P 和 N 。

对于样本 I^{HR} 图像, 其对应的目标超分辨率图像为 SR, 其正样本集为 $P = \{p_i | i=1, \dots, K_p\}$, 负样本集为 $N = \{N_j | j=1, \dots, K_n\}$, 其中正样本图像、负样本图像、SR 图像通过特征提取后的特征表示分别记为 f_i 、 p 和 n 。对于样本 I^{HR} 图像, 其对比损失定义为

$$L_i = \frac{1}{K_p} \sum_{i=1}^{K_p} -\log \frac{\exp(s(f_i, p_j)/\tau)}{\exp(s(f_i, p_j)/\tau) + \sum_{j=1}^{K_n} \exp(s(p_j, n_k)/\tau)} \quad (5)$$

其中: τ 为超参数, s 为相似性函数。图像在三维空间的特征大小为 $C \times H \times W$, 其中 C 为图像的通道数, H 为图像的高度, W 为图像的宽度。本文采用像素余弦相似度衡量特征之间的相似度 (s), s 的定义为

$$s(f^x, f^y) = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W \frac{f_{hw}^x \cdot f_{hw}^y}{\|f_{hw}^x\| \cdot \|f_{hw}^y\|} \quad (6)$$

其中 f^x 、 f^y 分别表示两幅图像的特征, 两幅图像的特征越相似, s 的值越接近于 1。对比损失为

$$L_{CL} = \frac{1}{N} \sum_{i=1}^N L_i \quad (7)$$

其中: N 代表总样本数。本文使用等式中的 L_1 损失和对比损失 L_{CL} 训练超分辨率网络。总损失函数 L_{SR} 定

义为

$$L_{SR} = L_1 + \lambda L_{CL} \quad (8)$$

其中: λ 为缩放参数, 将 $\lambda=1$ 作为默认值。 L_1 为常见的重构损失, 即

$$L_1 = |I^{SR} - I^{HR}| \quad (9)$$

3 实验

设计对比实验证明本文 CEDSR 模型能够提升生成图像的纹理细节效果, 通过计算 CEDSR 模型与 SRCNN、FSRCNN、EDSR、SRGAN 模型的峰值信噪比和结构相似性^[19]进行精度对比。展示 CEDSR 模型与 SRCNN、FSRCNN、EDSR、SRGAN 模型生成图像纹理细节的对比。

3.1 实验设置

实验在 NVIDIA GeForce RTX 3090 GPU 上进行训练。为了保证结果的公平性, 采用 TensorFlow 封装好的方法计算 PSNR 和 SSIM 指标。Adam 优化器具有 0.0002 的学习率, 可以同时处理 16 个批次的数据, 迭代次数为 12000 次。

训练数据从 BSDS200^[20]、General-100^[8]、DIV2K^[21]和 Flickr2K^[22]中随机采样高分辨率图像, 训练数据集详细信息见表 1。

表 1 训练数据

Tab. 1 Training data

数据集	图像数量/张	图像格式	平均分辨率
BSDS200	200	JPG	432 像素 × 370 像素
DIV2K	800	PNG	2 040 像素 × 1 848 像素
Flickr2K	2 650	PNG	2 040 像素 × 1 356 像素
General-100	100	BMP	435 像素 × 381 像素

测试数据为 DIV2K、Set14^[23]、BSDS100^[20]和 Urban100^[24]数据集, 测试数据集详细信息见表 2。

表 2 测试数据

Tab. 2 Test data

数据集	图像数量/张	图像格式	平均分辨率
DIV2K	100	PNG	2 040 像素 × 1 848 像素
BSDS100	100	JPG	432 像素 × 370 像素
Set14	14	PNG	492 像素 × 446 像素
Urban100	100	PNG	984 像素 × 797 像素

3.2 图像质量评估标准

为了评估超分辨率方法重建图像的质量, 运用图像质量评估标准 (image quality assessment, IQA) 对图像质量进行评价。评价方法主要以峰值信噪比和结构相似性为主, 通过公式计算出数值, 比较生成图像的质量, 该方法可以对图像进行定量分析。

(1) 峰值信噪比

峰值信噪比 (PSNR, 用符号 R_{PSNR} 表示) 是常用于图像压缩等领域的图像重建质量的评估方法, 通过图像之间的最大像素值和均方误差 (MSE, 用符号 E_{MSE} 表示) 进行定义。在图像超分辨率重建任务中, 它表示重建后的图像 x 与真实图像 y 之间的相似程度, PSNR 值越大, 说明图像的质量越好。

$$E_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N [y(i) - x(i)]^2 \quad (10)$$

$$R_{\text{PSNR}} = 10 \cdot \log \left(\frac{L^2}{E_{\text{MSE}}} \right) \quad (11)$$

其中: N 为图像像素的个数; y 为真实图像; x 为超分辨率重建后的图像; L 为图像像素的最大值, 一般为 255。

(2) 结构相似性

结构相似性 (SSIM, 用符号 S_{SSIM} 表示) 是一种基于结构信息退化的图像质量评估框架。SSIM 考虑了图像的结构信息、对比度与亮度, 反映图像结构的相似性。SSIM 值越大, 说明图像的质量越好。

给定两张图像 x 和 y , 其结构相似性为

$$S_{\text{SSIM}}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (12)$$

$$C_1 = (k_1 l)^2 \quad (13)$$

$$C_2 = (k_2 l)^2 \quad (14)$$

其中: μ_x 、 μ_y 分别为 x 、 y 的平均值, σ_x^2 、 σ_y^2 分别为 x 、 y 的方差, σ_{xy} 为 x 与 y 之间的协方差, l 为像素, $k_1 = 0.01$, $k_2 = 0.03$, C_1 、 C_2 为用于避免不稳定性的常数。结构相似性的范围为 $0 \sim 1$, 结构相似性越大说明图像越接近。

3.3 对比实验

使用 PyTorch 实现本文提出的方法, 为了公平比较, 使用它们的官方代码进行训练, 对比 PSNR 和 SSIM 两个指标。

设计实验证明去除冗余卷积层能够提升生成图像的速度, 并且图像质量并没有下降, 实验结果见表 3。由表 3 可知, 去除冗余卷积层后, 图像的 PSNR 指标有微弱的提升, 图像生成的时间有一定的缩短。

设计不同的 λ 考察模型对生成图像指标的影响, 找到一个最佳的参数 λ , 实验结果见表 4。由表 4 可知, 本文在 Set14 和 DIV2K 数据集上, 设计 4 组不同参数 λ 的实验, 当 $\lambda = 0$ 时, 即没有对比损失, 此时的 PSNR 和 SSIM 指标值在 4 组中最低; 随着 λ 的增大, PSNR 和 SSIM 指标都在增大; 当 $\lambda = 1$ 时, 图像

评价指标值最高; 当参数 λ 超过 1 时, 指标有变小的趋势。本文选择参数 $\lambda = 1$ 训练网络。

表 3 对比实验

Tab. 3 Contrast experiment

测试集	有无冗余卷积层	PSNR	图像生成时间/s
DIV2K	有	23.85	2.021
DIV2K	无	23.87	1.924

表 4 参数 λ 的选定Tab. 4 Selection of parameter λ

测试集	指标	$\lambda = 0$	$\lambda = 0.5$	$\lambda = 1$	$\lambda = 1.5$
Set14	PSNR	21.49	21.58	21.72	21.64
DIV2K	SSIM	0.735	0.746	0.756	0.747

在数据集 BSD200、General-100、DIV2K 和 Flickr2K 中随机抽取数据并进行训练, 将训练好的模型在 4 个测试数据集 DIV2K、BSDS100、Set14、Urban100 上进行测试。测试时, 需要先对测试数据进行双三次下采样处理, 然后对处理后的数据进行测试, 实验结果见表 5。

表 5 定量结果实验对比

Tab. 5 Experimental comparison of quantitative results

测试集	指标	SRCNN	FSRCNN	EDSR	SRGAN	本文模型
DIV2K	PSNR	23.02	23.11	23.62	23.52	23.94
	SSIM	0.709	0.717	0.737	0.745	0.758
Set14	PSNR	21.02	21.12	21.52	21.42	21.74
	SSIM	0.639	0.651	0.669	0.673	0.672
BSDS100	PSNR	22.32	22.48	23.52	23.69	23.82
	SSIM	0.669	0.667	0.737	0.725	0.752
Urban100	PSNR	21.82	22.08	22.42	22.38	22.51
	SSIM	0.649	0.667	0.687	0.689	0.688

注: 红色标注数据为实验最优指标值。

由表 5 分析可知, 本文模型在 DIV2K 数据集上的 PSNR 和 SSIM 指标都优于其他数据集的。由于 DIV2K 数据集中图像的分辨率都是 2K, 图像原本所携带的纹理信息更丰富, 图像特征更加明显, 所以模型在训练时更容易捕获图像的高频信息, 生成的超分辨率图像质量更好。从评价指标上分析, 本文模型相比于 SRCNN、FSRCNN、EDSR、SRGAN 模型, 在 4 个数据集上测试的 PSNR 指标都是最高的。在 SSIM 指标上, 本文模型在 DIV2K 和 BSDS100 数据集上取得最优的结果, 而在 Set14 和 Urban100 数据集上 SRGAN 模型取得最优的结果, 本文模型的 SSIM 指标略低于 SRGAN 模型的。由于 SRGAN 使用了生成对抗网络结构, 其中生成器负责将低分辨率图像转化为高分辨率图像, 判别器用于评估生成图像与真实图像之间的差异, 所以通过对抗学习的训练方式, SRGAN 能够生成更具结构相似性的图像。SRGAN

模型还提出了一种新型的感知损失,该感知损失是基于 VGG 网络的特征匹配损失,它可以更好地反映出人类视觉系统对图像细节和结构的敏感性,使生成器能够产生更接近真实图像的结果,从而提高 SSIM 指标。在 SRGAN 的训练过程中,直接优化 SSIM 指标作为生成器的损失函数,生成器可以更加专注于提高结构相似性,因此在 SSIM 指标上表现优异。在 DIV2K、Set14、BSDS100、Urban100 数据集上设计本文模型与插值法 (Bicubic)、SRCNN、FSRCNN、EDSR、SRGAN 模型生成图像纹理细节效果的对比实验,实验结果如图 5—图 8 所示。为了突出图像超分辨率重建后的纹理细节效果,选取图像中纹理细节丰富的一小块进行超分辨率重建,如图 5—图 8 中黄色框所示。

图 5 是典型的建筑细节恢复。插值法生成的图像质量最差,建筑的基本纹理均恢复得不清楚。在建筑纹理清晰度上,EDSR、SRGAN 和本文模型明显优于基于卷积的 SRCNN 和 FSRCNN。但在建筑图像中间的格状花纹部分,EDSR 和 SRGAN 恢复得比较混乱,而本文 CEDSR 模型更清晰。

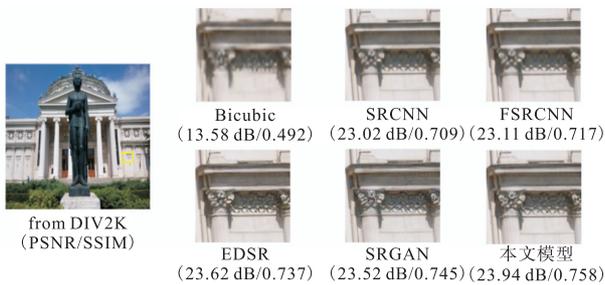


图 5 生成图像在 DIV2K 数据集上的效果展示
Fig. 5 Generated images for effect display on DIV2K data set

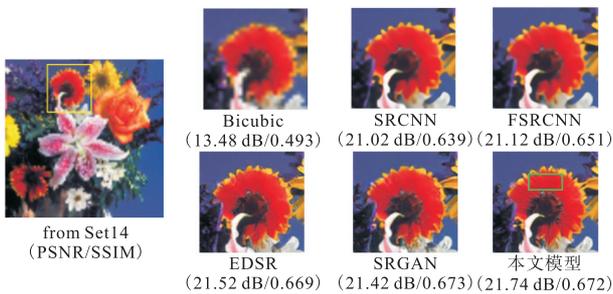


图 6 生成图像在 Set14 数据集上的效果展示
Fig. 6 Generated images for effect display on Set14 data set

由图 6 可知, SRCNN 和 FSRCNN 生成的图像质量相比于插值法有较大提升,但是花瓣的纹理细节效果较为模糊。在 EDSR 和 SRGAN 模型上,生成的花瓣图像有一定的轮廓,但不明显,而本文提出的

CEDSR 模型生成的花瓣纹理细节较为明显,如图 6 中绿色框所示。

图 7 是恢复建筑结构纹理细节,基于卷积神经网络的 SRCNN、FSRCNN 模型生成的图像整体上较为模糊。基于残差网络的 EDSR 模型和基于生成对抗网络的 SRGAN 模型能恢复窗户的整体轮廓,但窗户的纹理细节较为模糊,本文 CEDSR 模型生成的图像能够较好地还原窗户的轮廓及纹理细节。

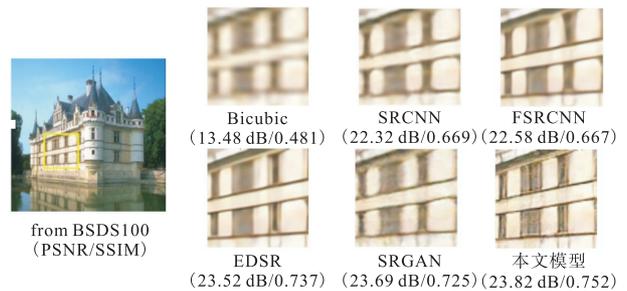


图 7 生成图像在 BSDS100 数据集上的效果展示
Fig. 7 Generated images for effect display on BSDS100 data set

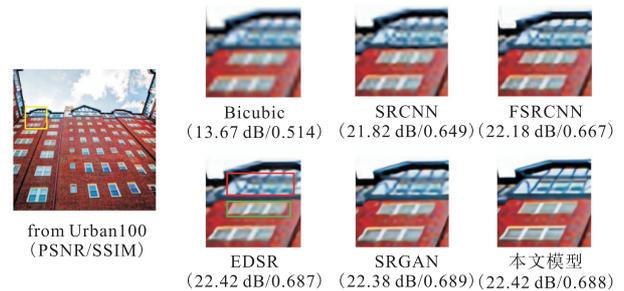


图 8 生成图像在 Urban100 数据集上的效果展示
Fig. 8 Generated images for effect display on Urban100 data set

在图 8 中,插值法、SRCNN、FSRCNN 生成图像的纹理细节较差,而 EDSR 和 SRGAN 模型能够较好地还原白色窗户的纹理细节,如图 8 绿色框所示,而蓝色顶窗的纹理细节恢复的效果较差,如图 8 红色框所示。本文 CEDSR 模型生成的图像在窗户的纹理细节方面都优于其他模型。

本文 CEDSR 模型相比于 SRCNN、FSRCNN、EDSR、SRGAN 模型,在生成图像纹理细节效果上有一定的提升,证实了本文提出的对比学习方法通过构造有效的正负样本,在特征空间对比损失的约束下,能够使生成网络学到正样本所携带的纹理信息,从而生成纹理细节更好的超分辨率图像,使重建图像的生成质量更好。

4 结 语

本文提出一种基于对比学习的图像超分辨率方法,进一步提高了现有的图像超分辨率的性能。从样本构建角度研究对比学习,为了使模型更好地重建丢失的高频信息,设计了有效的数据增强方法和简单有效的正负样本对,将原始图像锐化和轻微模糊后的图像作为正负样本,在对比损失的约束下,使生成网络能够生成纹理细节较好的超分辨率图像,从而进一步提高了超分辨率性能。但是,本文构建正负样本的形式过于单一,还可以通过多种数据增强方式构建丰富的正负样本对,例如可以使用多种高通滤波器进行图像锐化得到正样本,也可以使用不同的模糊核构建负样本,以便在现实世界的超分辨率任务中有更好的应用。实验结果表明,本文提出的对比学习方法可以获得纹理细节效果较好的超分辨率图像。

参考文献:

- [1] 王森,杨克俭. 基于双线性插值的图像缩放算法的研究与实现[J]. 自动化技术与应用,2008,27(7):44-45.
- [2] 李英民. 图像双三次插值算法的研究[D]. 兰州:兰州大学,2020.
- [3] 郝丽. 数字图像重采样检测研究[J]. 大连:大连理工大学,2009.
- [4] ALBAWI S, MOHAMMED T A, AL-ZAWI S. Understanding of a convolutional neural network[C]//IEEE. 2017 International Conference on Engineering and Technology (ICET). New York: IEEE, 2017: 8308186.
- [5] CRESWELL A, WHITE T, DUMOULIN V, et al. Generative adversarial networks: an overview[J]. IEEE Signal processing magazine, 2018, 35(1): 53-65.
- [6] TAI Y, YANG J, LIU X. Image super-resolution via deep recursive residual network[C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2017: 3147-3155.
- [7] DONG C, LOY C C, HE K, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on pattern analysis and machine intelligence, 2015, 38(2): 295-307.
- [8] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network[C]//ECCV. Computer Vision-ECCV 2016: 14th European Conference. Berlin: Springer International Publishing, 2016: 391-407.
- [9] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016: 1646-1654.
- [10] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2017: 4681-4690.
- [11] WANG X, YU K, WU S, et al. ESRGAN: enhanced super-resolution generative adversarial networks[C]//ECCV. Proceedings of the European Conference on Computer Vision (ECCV) Workshops. Berlin: Springer, 2018: 63-79.
- [12] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016: 770-778.
- [13] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. New York: IEEE, 2017: 136-144.
- [14] SHOCHER A, COHEN N, IRANI M. "zero-shot" super-resolution using deep internal learning[C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 3118-3126.
- [15] SOH J W, CHO S, CHO N I. Meta-transfer learning for zero-shot super-resolution[C]//IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2020: 3516-3525.
- [16] WU H, QU Y, LIN S, et al. Contrastive learning for compact single image dehazing[C]//IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2021: 10551-10560.
- [17] WANG Y, LIN S, QU Y, et al. Towards compact single image super-resolution via contrastive self-distillation [EB/OL]. [2023-06-01]. <https://doi.org/10.48550/arXiv.2105.11683>.
- [18] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition

- [EB/OL]. [2023-06-01]. <https://arxiv.org/pdf/1409.1556.pdf>.
- [19] 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.
- [20] ARBELAEZ P, MAIRE M, FOWLKES C, et al. Contour detection and hierarchical image segmentation[J]. IEEE Transactions on pattern analysis and machine intelligence, 2010, 33(5): 898-916.
- [21] AGUSTSSON E, TIMOFTE R. NTIRE 2017 challenge on single image super-resolution: dataset and study[C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. New York: IEEE, 2017: 126-135.
- [22] TIMOFTE R, AGUSTSSON E, VAN GOOL L, et al. NTIRE 2017 challenge on single image super-resolution: methods and results[C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. New York: IEEE, 2017: 114-125.
- [23] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations[C]//LNTCS. Curves and Surfaces: 7th International Conference. Berlin: Springer Berlin Heidelberg, 2012: 711-730.
- [24] HUANG J B, SINGH A, AHUJA N. Single image super-resolution from transformed self-exemplars[C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2015: 5197-5206.

责任编辑: 郎婧

(上接第34页)

- [4] 周景辉, 朱宏伟. 纸地膜的研究及应用概况[J]. 中国造纸, 2002, 21(5): 56-58.
- [5] 韩永俊, 陈海涛. 国内外纸地膜的发展现状及思考[J]. 农机化研究, 2008(12): 244-246.
- [6] 胡稳. 高雾度透明全纤维素薄膜的制备及在电子器件中的应用[D]. 广州: 华南理工大学, 2019.
- [7] HOU G Y, LIU Y, ZHANG D J, et al. Approaching theoretical haze of highly transparent all-cellulose composite films[J]. ACS Applied materials & interfaces, 2020, 12(28): 31998-32005.
- [8] 王楠. 纳米纤维素形貌调控及其纳米纸光学雾度和悬浮液流变性能的研究[D]. 西安: 陕西科技大学, 2020.
- [9] 张燕杰, 罗军, 葛明桥. 完全降解纸基棉质地膜的制备及性能研究[J]. 产业用纺织品, 2011, 29(11): 5-9.
- [10] 傅玉全. 纸基地膜地面覆盖栽培试验[J]. 纸和造纸, 2000(1): 42-43.
- [11] 新华网. 新方法让制造透明纸更简单[J]. 功能材料信息, 2014, 11(3): 61.
- [12] YAO Y G, TAO J S, ZOU J H, et al. Light management in plastic-paper hybrid substrate towards high-performance optoelectronics[J]. Energy & environmental science, 2016, 9(7): 2278-2285.
- [13] 李玉洁, 李冠辉, 侯高远, 等. 木质纤维聚合度对高透光率纤维素薄膜力学性能的影响[J]. 中国造纸学报, 2023, 38(1): 1-8.
- [14] 王柯, 闫巧, 马瑞花, 等. 高固含量液体石蜡纳米乳液的制备及稳定性[J]. 精细石油化工进展, 2021, 22(4): 25-29.
- [15] 全国塑料标准化技术委员会. 透明塑料透光率和雾度的测定: GB/T 2410—2008[S]. 北京: 中国标准出版社, 2008.
- [16] 张召, 路思远, 张直直, 等. 浆料打浆度对纤维素薄膜光学及力学性能的影响[J]. 中国造纸, 2022, 41(10): 24-29.
- [17] 程金兰, 翟华敏, 汤丽娟. PAE 对不同浆料薄型纸湿强度特性的影响[J]. 中国造纸, 2008, 27(4): 17-20.
- [18] 程金茹, 胡建全, 杨乾, 等. 一种热封纸袋用膜的制备研究[J]. 中国造纸, 2022, 41(9): 46-53.
- [19] 白媛媛, 类延豪, 姚春丽, 等. 环保型造纸湿强剂的研究进展[J]. 中国造纸学报, 2016, 31(4): 49-54.
- [20] 杨叶城, 张昕, 韩小帅, 等. 带像差校正功能的柔性变焦透镜[J]. 光学学报, 2022, 42(4): 165-170.
- [21] 陈港, 蒋晨颖, 刘映尧, 等. 透明纸基材料的研究与应用[J]. 中国造纸, 2018, 37(7): 65-75.
- [22] 全国造纸工业标准化技术委员会. 半透明纸: GB/T 22812—2008[S]. 北京: 中国标准出版社, 2008.
- [23] 董轩萌, 郭立稳, 董宏伟, 等. 基于傅里叶变换红外光谱的不同煤种微观结构定量分析[J]. 中国科技论文, 2022, 17(1): 55-61.

责任编辑: 周建军