

文章编号:1671-6833(2024)06-0048-08

基于双特征提取和注意力机制的图像超分辨率重建

薄阳瑜, 武永亮, 王学军

(石家庄铁道大学 信息科学与技术学院, 河北 石家庄 050043)

摘要: 针对图像超分辨率重建过程中忽略图像高频特征, 导致特征提取不充分, 重建图像纹理细节模糊的问题, 提出了一种基于双特征提取和注意力机制的图像超分辨率重建方法。首先, 该方法采用双分支网络进行特征提取, 以解决图像重建过程中高频特征和多尺度特征无法有效提取和一致融合的问题; 其次, 为了使网络提取到更加精确的高频特征, 提出了局部空间注意力模块, 并与通道注意力模块结合构建残差融合注意力模块, 提高网络对高频特征的定位能力; 最后, 设计了空洞金字塔模块, 扩大网络感受野, 使网络多尺度提取特征。在 4 个基准数据集上的测试结果表明: 尤其是超分辨率倍数为 4 时, 所提方法较目前若干主流模型中的最佳峰值信噪比分别提升了 0.16, 0.08, 0.03, 0.20 dB, 所提方法在视觉效果和定量分析方面均有较好提升。

关键词: 图像超分辨率重建; 局部空间注意力; 残差融合注意力; 空洞金字塔; 双分支网络

中图分类号: TP751; TP391.41; TP183

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2024.03.009

图像超分辨率重建 (super-resolution, SR) 是一项重要的底层计算机视觉任务, 旨在将低分辨率 (low-resolution, LR) 图像转换为高分辨率 (high-resolution, HR) 图像, 这一技术在图像识别、遥感探测和视频监控等领域中具有重要应用价值。传统的 SR 方法通过填补图像放大后像素空缺的位置来提高图像的分辨率, 往往忽略了图像的详细纹理信息。因此, 需要更加高效的方法来解决这个问题。

近年来, 基于深度学习的 SR 方法重建出图像的纹理细节更加丰富^[1]。2015 年, Dong 等^[2]提出了首个基于 CNN 的超分辨率重建网络 SRCNN (super-resolution convolutional neural network), 大大提高了图像重建质量。Dong 等^[3]对 SRCNN 进行改进, 提出了 FSRCNN (fast super-resolution convolutional neural network), FSRCNN 是第 1 个使用反卷积层从 LR 特征图中重建 HR 图像的 SR 方法, 但是网络仍然过于简单, 重建出的纹理细节不够丰富。Kim 等^[4]提出了 VDSR (very deep convolutional networks), 其网络架构为 20 层的残差网络, 但 VDSR 将图像插值放大后进行训练, 增加了网络训练难度。

Lim 等^[5]提出了 EDSR (enhanced deep residual network for single image super-resolution), 该方法移除批处理归一化层以降低内存占用, 加深网络结构来提取更加丰富的特征。Lai 等^[6]提出了 LapSRN (deep laplacian pyramid networks for fast and accurate super-resolution), 该网络利用金字塔结构逐级生成高分辨率图像。为了进一步提升 SR 的性能, Kim 等^[7]提出了 DRCN (deeply-recursive convolutional network), 使用相同的模块来递归地学习高级特征, 但是递归学习也会导致梯度消失和梯度爆炸。Tai 等^[8]提出了 DRRN (deep recursive residual network), 网络使用残差块作为递归模块的一部分, 并结合局部残差和全局残差来提高网络的特征利用能力。Zhang 等^[9]提出了 RCAN (residual channel attention network), 在残差块中引入注意力机制, 提高网络对低频、高频特征的判别能力。Ahn 等^[10]提出了 CARN (cascading residual network), 采用全局和局部级联结构提取中间特征, 但网络将提取到的中间特征集中到 1×1 卷积中, 阻碍了信息的传递。

一些方法不断加深网络, 在一定程度上提升了

收稿日期: 2023-10-25; 修订日期: 2023-11-23

基金项目: 国家自然科学基金资助项目 (62106157); 河北省自然科学基金资助项目 (F2021210002); 山西省重点实验室开放基金 (CICIP2022001)

通信作者: 王学军 (1989—), 男, 河北衡水人, 石家庄铁道大学讲师, 博士, 主要从事机器学习、稀疏表示、图像处理研究, E-mail: wangxuejun@stdu.edu.cn。

引用本文: 薄阳瑜, 武永亮, 王学军. 基于双特征提取和注意力机制的图像超分辨率重建[J]. 郑州大学学报(工学版), 2024, 45(6): 48-55, 64. (BO Y Y, WU Y L, WANG X J. Image super-resolution reconstruction network based on double feature extraction and attention mechanism[J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(6): 48-55, 64.)

图像的重建效果,但仍然存在特征利用率低导致重建图像信息模糊的问题。与以往方法不同的是,本文充分考虑了图像特征提取以及融合对网络性能的影响,提出了一种基于双特征提取和注意力机制的图像超分辨率重建方法。特别地,本文提出了局部空间注意力模块(local spatial attention block, LSAB),并将其与 Hu 等^[11]提出的通道注意力模块(channel attention block, CAB)相结合,增强网络在通道维度和空间维度上的细节信息感知能力,使网络获取更加有效的轮廓、纹理信息。此外,本文构建空洞金字塔模块(atrous pyramid block, APB)以提高网络的多尺度特征提取能力。

为了有效提取融合高频特征和多尺度特征,本文设计了双分支网络。第1个分支通过注意力机制使网络提取到更加精细的高频特征,第2个分支使用空洞金字塔模块扩大网络感受野,提取多尺度特征,最后将提取到的特征进行上采样并进一步融合。考虑到注意力机制可以有效改善网络性能,提出了 LSAB,提高网络对高频特征的定位能力。将 LSAB 与 CAB 相结合构建残差融合注意力模块(residual fusion attention block, RFAB),使网络在通道维度和空间维度上更加关注高频特征。本文提出的 APB 模块各分支由扩张率不同的空洞卷积构成,扩大了网络感受野,使得网络以多尺度提取特征。

1 本文方法

为了提高图像的重建质量,本文提出了一种基于双特征提取和注意力机制的图像超分辨率重建方法。整体网络结构如图1所示,主要由3个模块组成:浅层特征提取模块、双分支特征提取模块、融合重建模块。

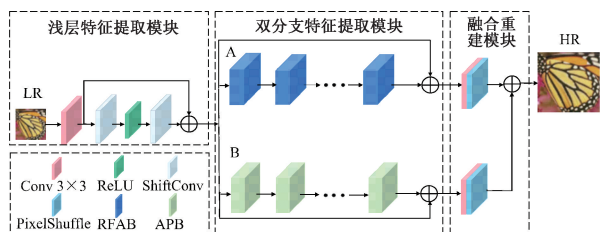


图1 整体网络结构图

Figure 1 Structure of the overall network

首先,低分辨率图像 I_{LR} 由 3×3 卷积进行浅层特征提取,为了对提取到的特征进一步细化,将低频信息进一步过滤,提取到的特征将由 Wu 等^[12]提出的移位卷积进一步处理。该过程如式(1)所示:

$$\begin{cases} H_0 = f_0(I_{LR}); \\ H_1 = H_0 + f_{\text{shift}}(\delta(f_{\text{shift}}(H_0))). \end{cases} \quad (1)$$

式中: $f_0(\cdot)$ 表示 3×3 卷积特征提取函数; $f_{\text{shift}}(\cdot)$ 表示移位卷积函数; $\delta(\cdot)$ 表示 ReLU 激活函数; I_{LR} 和 H_1 分别表示该模块的输入和输出。

细化后的浅层特征 H_1 将进入双分支特征提取模块,双分支网络由 A、B 这 2 个分支组成。 H_1 将同时进入 2 个分支,其中 A 分支由若干个残差融合注意力模块和一个长短跳跃连接组成, B 分支由若干个级联的空洞金字塔模块和一个长短跳跃连接构成。该过程如式(2)所示:

$$\begin{cases} H_2 = f_A(H_1) + H_1; \\ H_3 = f_B(H_1) + H_1. \end{cases} \quad (2)$$

式中: $f_A(\cdot)$ 和 $f_B(\cdot)$ 分别表示 A、B 这 2 个分支的深层特征提取函数。

其次,提取到的深层特征 H_2 和 H_3 分别进入图像重建模块。

最后,将重建后的两组特征进行融合生成最后的高分辨率图像 I_{SR} 。该过程如式(3)所示:

$$I_{SR} = f_{\text{up}}(H_2) + f_{\text{up}}(H_3). \quad (3)$$

1.1 局部空间注意力模块 (LSAB)

虽然 CAB 可以有效改善网络性能,但是只能捕捉图像不同通道之间的特征相关性。本文提出了 LSAB,对输入的每个特征赋予不同的权重,使网络更加关注空间特征信息,保证空间信息完整性的同时提高网络对高频特征的定位能力。该模块共分为压缩和扩张 2 个阶段,其详细结构如图2所示。

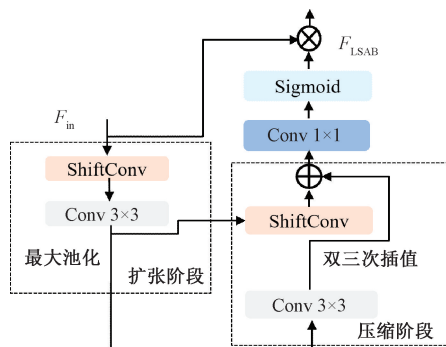


图2 局部空间注意力模块

Figure 2 Local spatial attention block

压缩阶段:输入的特征图首先通过移位卷积进行处理,移位卷积由一组移位操作和一个 1×1 卷积构成,在不引入额外参数的情况下,通过移位操作扩大感受野,并且由 1×1 卷积整合信息,其详细结构如图3所示。

其次,使用 3×3 卷积进一步提取特征,通过最大池化进行下采样,使特征图内的高频区域更加突出。该过程如式(4)所示:

$$F_{\text{seq}} = f_{\downarrow}(f_{3 \times 3}(f_{\text{shift}}(F_{\text{in}}))). \quad (4)$$

式中: F_{in} 和 F_{seq} 分别为压缩阶段的输入和输出; $f_{shift}(\cdot)$ 和 $f_{3 \times 3}(\cdot)$ 分别表示移位卷积和 3×3 卷积的特征提取函数; $f_{\downarrow}(\cdot)$ 表示最大池化下采样。

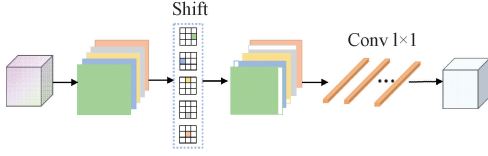


图3 移位卷积结构图

Figure 3 Structure of shift-convolution

扩张阶段:通过 3×3 卷积、双三次插值上采样和移位卷积对特征进一步处理,与压缩阶段相呼应。为了充分利用已提取到的特征,避免下采样特征丢失,将压缩阶段下采样之前提取的特征输入到扩张阶段的移位卷积层,并与上采样得到的特征图逐像素相加,然后经过一个 1×1 卷积调整通道得到该模块的输出,该过程如式(5)所示:

$$\begin{cases} F_{res} = f_{\uparrow}(f_{3 \times 3}(F_{seq})) + f_{shift}(F_{out}); \\ F_{ext} = f_{1 \times 1}(F_{res}). \end{cases} \quad (5)$$

式中: F_{out} 和 F_{ext} 分别为压缩阶段 3×3 卷积和扩张阶段的输出; $f_{\uparrow}(\cdot)$ 表示双三次插值上采样; $f_{1 \times 1}(\cdot)$ 和 $f_{3 \times 3}(\cdot)$ 分别表示 1×1 和 3×3 卷积函数; $f_{shift}(\cdot)$ 表示移位卷积函数。

最后,通过 Sigmoid 函数得到局部空间注意力权重,对输入的特征图进行调整,该过程如式(6)所示:

$$F_{LSAB} = F_{in} \cdot \delta(F_{ext}). \quad (6)$$

式中: $\delta(\cdot)$ 表示激活函数; F_{in} 和 F_{LSAB} 分别表示 LSAB 的输入和输出。

1.2 残差融合注意力模块(RFAB)

大多数图像超分辨率重建方法在特征提取时忽略了通道维度和空间维度上特征分布的差异性,导致网络无法充分提取图像的高频特征。针对该问题,本文将 CAB 与 LSAB 相结合,构建 RFAB 模块进行深层特征提取。通过 CAB 使网络根据通道信息量的多少选择特征,且 LSAB 能够提高网络对高频特征的定位能力,从而使网络选取更加精细的高频特征。其详细结构如图4所示。

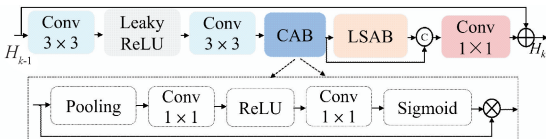


图4 残差融合注意力模块

Figure 4 Residual fusion attention block

第 k 个 RFAB 的输出结果可以表示为

$$H_k = F_k(H_{k-1}). \quad (7)$$

式中: H_{k-1} 和 H_k 分别表示第 k 个 RFAB 的输入和输出; $F_k(\cdot)$ 表示第 k 个 RFAB 函数,具体操作如下。

首先, H_{k-1} 经过 2 个 3×3 卷积层、1 个 Leaky ReLU 激活层得到 H'_{k-1} ,该过程如式(8)所示:

$$H'_{k-1} = f_{3 \times 3}(\delta(f_{3 \times 3}(H_{k-1}))). \quad (8)$$

式中: $f_{3 \times 3}(\cdot)$ 代表 3×3 卷积层函数; $\delta(\cdot)$ 为 Leaky ReLU 激活函数。

其次, H'_{k-1} 通过 CAB 提取到特征,使网络判别出不同通道特征的重要程度,忽略不相关的信息。提取到的特征之后进入 LSAB 得到 H''_{k-1} ,使网络更加准确地定位到高频特征所在的空间位置。

最后,将 LSAB 与 CAB 提取到的特征进行拼接,使用 1×1 卷积调整通道,得到输出 H_k ,该过程如式(9)所示:

$$\begin{cases} H''_{k-1} = f_{LSAB}(f_{CAB}(H'_{k-1})); \\ H'''_{k-1} = \text{concat}(f_{CAB}(H'_{k-1}), H''_{k-1}); \\ H_k = f_{1 \times 1}(H'''_{k-1}) + H_{k-1}. \end{cases} \quad (9)$$

式中: $f_{1 \times 1}(\cdot)$ 表示 1×1 卷积函数; $f_{CAB}(\cdot)$ 和 $f_{LSAB}(\cdot)$ 分别代表 CAB 和 LSAB 所用的功能函数。

1.3 空洞金字塔模块(APB)

为了扩大网络感受野,使网络多尺度提取图像特征,本文设计了 APB,详细结构如图5所示。该模块由不同类型的卷积构成,共分为 3 个分支。首先通过 2 个移位卷积和 1 个 ReLU 激活函数,使网络获得更大的感受野;其次采用并行的方式使用扩张率为 1,2,3 的 3×3 卷积多尺度提取特征;最后将每个分支提取到的特征进行拼接。为了减少参数量并且保证重建质量,本文在每个分支开端增加了 1×1 卷积进行降维。

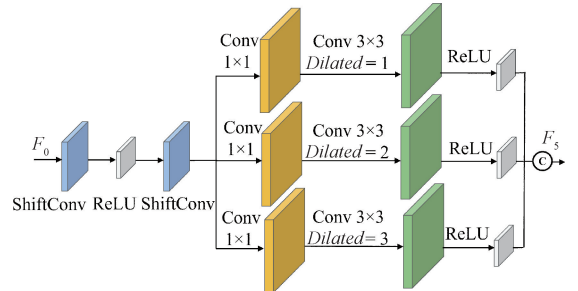


图5 空洞金字塔模块

Figure 5 Atrous pyramid block

该过程如式(10)所示:

$$\begin{cases} F_1 = f_{shift}(\delta(f_{shift}(F_0))); \\ F_2 = f_{3 \times 3}^1(f_{1 \times 1}(F_1)); \\ F_3 = f_{3 \times 3}^2(f_{1 \times 1}(F_1)); \\ F_4 = f_{3 \times 3}^3(f_{1 \times 1}(F_1)); \\ F_5 = \text{concat}[F_2, F_3, F_4] + F_0. \end{cases} \quad (10)$$

式中: $f_{\text{shift}}(\cdot)$ 表示移位卷积函数; $f_{3 \times 3}^i(\cdot)$ 表示扩张率为 i 的空洞卷积函数; $f_{1 \times 1}(\cdot)$ 表示 1×1 卷积函数; $\delta(\cdot)$ 表示 ReLU 激活函数。

1.4 损失函数

本文实验采用 L_1 ^[13] 损失函数, 给定 N 对训练样本, 通过模型训练来研究其端到端的映射关系, 不断优化调整降低损失函数。损失函数定义为

$$L_1 = \frac{1}{N} \| I_{\text{HR}} - I_{\text{SR}} \|_1。$$

(11)

式中: N 表示样本对的数量; I_{HR} 和 I_{SR} 分别表示原始 HR 图像和生成的 SR 图像。

2 实验与结果分析

2.1 数据集

本文实验使用 DIV2K 数据集作为训练数据集。在图像处理过程中, 对图像进行旋转、翻转并裁剪成 64×64 大小的图像块, 以便模型的训练和优化。为了验证该模型性能, 采用 Set5、Set14、BSD100、Urban100^[14] 4 个基准数据集作为测试数据集。这些数据集被广泛应用于图像超分辨率重建领域, 能够有效评估模型的重建效果。为了证明本文方法的有效性, 采用峰值信噪比 ($PSNR$) 和结构相似度 ($SSIM$) 对实验结果进行客观评价^[15]。

2.2 实验环境

本文方法实验平台的 CPU 为 Intel(R) Xeon(R) Platinum 8255C, GPU 为 RTX 3080, Cuda11.1。模型的训练环境为 Python3.7, 编程框架为 Pytorch1.8.1。采用 Adam 优化器训练网络以及优化参数, $\beta_1 = 0.9$, $\beta_2 = 0.999$, 设定 $batchsize$ 为 8, 初始学习率为 0.000 1, 每个尺度因子的 $epoch$ 均为 500, 每训练 250 个回合, 学习率下降 50%。

2.3 消融实验

2.3.1 LSAB 对图像重建质量的影响

CAB 只能在通道维度上根据通道所包含的信息量多少进行选择, 无法准确定位图像的高频特征, 本文设计了 LSAB 解决此问题。为了研究 LSAB 对图像重建的影响, 去除 RFAB 中的 LSAB 进行对比。在 Set5 数据集上的重建结果如表 1 所示, 与原始网络模型相比, 去除 LSAB 的网络后, $PSNR$ 和 $SSIM$ 分别下降了 0.18 dB 和 0.003 3, 验证了 LSAB 对高频特征提取有一定的作用。

2.3.2 APB 对图像重建质量的影响

此外, 为了研究 APB 对图像重建质量的影响, 将该模块从网络中去除, 如表 1 所示, 去除 APB 后 $PSNR$ 和 $SSIM$ 都有所下降, 说明 APB 使网络提取到

的特征更加全面, 从而可以提高网络重建性能。

表 1 LSAB 和 APB 对图像重建质量的影响

Table 1 The effects of LSAB and APB on image reconstruction quality

LSAB	APB	$PSNR/\text{dB}$	$SSIM$
无	有	32.07	0.893 5
有	无	31.03	0.876 6
有	有	32.25	0.896 8

2.3.3 空洞卷积对图像重建质量的影响

本文提出 APB 的目的是扩大网络感受野以提取图像的多尺度信息, APB 由扩张率 d 为 1, 2, 3 的 3×3 卷积构成。图 6 为空洞卷积的原理图, 通过填充 0 扩张标准卷积核, 来获得不同大小的感受野^[16]。

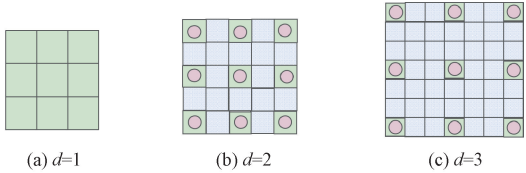


图 6 空洞卷积原理图

Figure 6 Schematic diagram of dilated convolution

为了进一步验证空洞卷积对图像重建的有效性, 本文将空洞卷积构成的金字塔结构与标准卷积构成的金字塔结构进行超分辨率重建对比, 在 Set5 数据集上的对比结果如表 2 所示。由表 2 可以看出, APB 中使用空洞卷积相比于使用标准卷积, $PSNR$ 和 $SSIM$ 分别提高了 0.05 dB 和 0.000 1, 验证了空洞卷积的有效性。

表 2 空洞卷积对图像重建质量的影响

Table 2 The effects of dilated convolution on image reconstruction quality

卷积类型	$PSNR/\text{dB}$	$SSIM$
标准卷积	32.17	0.894 8
空洞卷积	32.22	0.894 9

2.3.4 图像重建模块对图像重建质量的影响

本文对 2 个分支网络所提取到的特征分别使用亚像素卷积进行重建, 为了验证其有效性, 本文设置了消融实验, 将 A、B 这 2 个分支融合后进行一次重建, 在 Set5 数据集上的重建结果如表 3 所示。由表 3 可知, 重建 2 次得到的 $PSNR$ 和 $SSIM$ 分别提高了 0.06 dB 和 0.000 3。说明对 2 个分支的特征分别上采样进行重建有利于提高重建效果。

2.3.5 网络深度对图像重建质量的影响

本文提出的网络模型包含 A 和 B 分支, 每个分支分别由 10 个 RFAB 和 10 个 APB 构成。为了验证采用不同数目的 RFAB 和 APB 对于图像重建质

量的影响,本次实验在 2 个测试数据集上设置了 3 组实验(模块数量分别为 9,10,11)进行 4 倍的图像超分辨率实验,如表 4 所示。由表 4 可知,随着网络深度的增加, $PSNR$ 和 $SSIM$ 都有提高的趋势。当 2 个模块数量均为 10 时,模型的性能达到最优;当 2 个模块数量均为 9 时,模型的性能指标明显降低;当 2 个模块数量均为 11 时,模型的性能指标没有明显升高,甚至降低。考虑到计算成本,本文深层特征提取采用 10 个 RFAB 和 10 个 APB 模块。

表 3 不同重建次数对图像重建质量的影响
Table 3 The effects of different reconstruction times on image reconstruction quality

重建次数	$PSNR/dB$	$SSIM$
1	31.95	0.891 7
2	32.01	0.892 0

表 4 网络深度对图像重建质量的影响
Table 4 The effects of depths of network on image reconstruction quality

模块数量	$PSNR/dB$		$SSIM$	
	Set5	Urban100	Set5	Urban100
9	32.04	25.91	0.892 5	0.780 4
10	32.09	25.93	0.893 5	0.780 8
11	32.09	25.94	0.893 2	0.780 5

为了进一步验证 A 分支中 RFAB 数量对重建质量的影响,本文将单独使用 A 分支网络进行重建。如图 7 所示,可以看出第 3 个 RFAB 后输出的特征图包含更多的高频特征,婴儿脸部的纹理细节更加明显,因此适当加深网络结构深度,有助于提高图像重建效果。

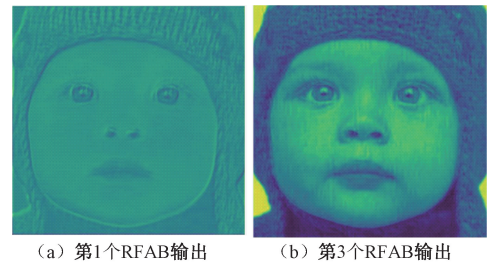


图 7 输出特征图对比

Figure 7 Comparison of output feature maps

2.4 客观评价

为了验证本文方法的有效性,将本文方法与 SRCNN^[2]、FSRCNN^[3]、VDSR^[4]、LapSRN^[6]、DRCN^[7]、DRRN^[8]、CARN^[10]、LAPAR-A^[17]、MemNet^[18]、AWSRN^[19]、IDN^[20]、IMDN^[21]、CBPN^[22]、RFDN^[23]、RLFN^[24] 15 种方法进行比较,在 Set5、Set14、BSD100、Urban100 上分别进行放大因子为 2 倍、3 倍、4 倍超分辨率重建对比,如表 5 所示。

相比于其他主流方法,本文提出的方法在 4 个公开测试数据集上的 $PSNR$ 和 $SSIM$ 都取得了最优或次优值。具体来说,对于测试数据集 Set5、Set14、BSD100、Urban100,当放大倍数为 2 时,本文方法的 $PSNR$ 比 RLFN 算法的次优值分别提升了 0.03, 0.03, 0.04, 0.15 dB;当放大倍数为 3 时,本文方法在测试数据集 Set14 和 Urban100 上的 $PSNR$ 相比于 RFDN 方法的次优 $PSNR$ 分别提高了 0.01, 0.11 dB;当放大倍数为 4 时,本文方法在 4 个数据集上的 $PSNR$ 比 RLFN 方法和 CBPN 方法的次优值分别提升了 0.16, 0.08, 0.03, 0.20 dB。此外,当放大倍数为 2, 3, 4 时,本文方法在 4 个测试数据集上均取得了最优的 $SSIM$,说明双分支网络充分考虑图像的多尺度特征以及高频特征,更有利于挖掘图像的内部信息。本文方法性能提升主要得益于 APB 结构、LSAB 结构以及 RFAB 结构。APB 能够充分提取图像不同尺度的特征,使网络获取图像丰富的特征信息;LSAB 能够更加准确地定位到高频特征信息;RFAB 分别在通道和空间维度上提高网络的高频特征选择能力,引导网络聚焦可靠信息;且本文采用的双分支网络能够有效融合多尺度特征和高频特征,提高网络的特征利用率,这些因素都是本文方法性能显著提高的原因。

同时,本文与其他方法进行了模型参数量和计算量的对比。从表 5 可以看出,本文方法的性能和参数量之间具有很好的平衡。相比于 DRCN 和 CARN,本文方法在提升 $PSNR$ 和 $SSIM$ 的同时,参数量约为其他方法的一半。对比 RFDN、RLFN、IMDN,虽然本文方法的参数量有所增加,但是本文方法的 $PSNR$ 和 $SSIM$ 显著提升,这表明本文方法能够在较小参数量的情况下使网络获得更好的性能。

2.5 实验结果分析

为了进一步展示不同超分辨率方法的重建效果,图 8 展示了在不同测试数据集上的重建效果。从图 8 可以看出,Bicubic 重建后的图像整体虚化,轮廓最为模糊。对于 BSD100 中的 img86000 图像,其他方法中放大后的叶片有轻微的模糊感,背景轮廓虚化,而本文方法重建后的图像叶片更加清晰;本文方法重建的 Urban100 中的 img080 图像结构更加规则,与 HR 图像更接近;Set14 中的 comic 图像,人物的服装细节较多,RFDN 和 RLFN 重建图像放大后的边缘轮廓不清晰并且细节缺失,本文模型重建图像放大后有较为清晰的纹理细节,给人的视觉效果要比其他模型更好。相比于一些主流方法,本文方法不仅取得了较高的 $PSNR$ 和 $SSIM$,而且重建图

表 5 不同 SR 算法在测试数据集上的指标值对比

Table 5 Comparison of index values of different SR algorithms on test datasets

放大因子	模型	参数量/ kB	浮点运算次 数/GFLOPs	PSNR/dB				SSIM			
				Set5	Set14	BSD100	Urban100	Set5	Set14	BSD100	Urban100
2 倍	SRCNN	57	52. 7	36. 66	32. 42	31. 36	29. 50	0. 954 2	0. 960 3	0. 887 9	0. 894 6
	FSRCNN	12	6. 0	36. 98	32. 62	31. 50	29. 85	0. 955 6	0. 908 7	0. 890 4	0. 900 9
	VDSR	665	612. 6	37. 53	33. 05	31. 90	30. 77	0. 958 7	0. 912 7	0. 896 0	0. 914 1
	DRCN	1 774	9 788. 7	37. 63	33. 04	31. 85	30. 75	0. 958 8	0. 911 8	0. 894 2	0. 913 3
	LapSRN	251	29. 9	37. 52	32. 99	31. 80	30. 41	0. 959 1	0. 912 4	0. 859 2	0. 910 3
	AWSRN	397	91. 2	37. 75	33. 31	32. 00	31. 39	0. 959 6	0. 915 1	0. 897 4	0. 920 7
	DRRN	297	6 796. 9	37. 74	32. 23	32. 05	31. 23	0. 959 1	0. 913 6	0. 897 3	0. 918 8
	CARN	1 592	222. 8	37. 76	33. 52	32. 09	31. 92	0. 959 0	0. 916 6	0. 897 8	0. 925 6
	LAPAR-A	548	171. 0	38. 01	33. 62	32. 19	32. 10	0. 960 5	0. 918 3	0. 899 9	0. 928 3
	IDN	553	127. 7	37. 83	33. 30	32. 08	31. 27	0. 960 0	0. 914 8	0. 898 5	0. 919 6
	MemNet	678	623. 9	37. 78	32. 28	32. 08	31. 33	0. 959 7	0. 914 2	0. 897 8	0. 919 5
	IMDN	684	158. 8	38. 00	33. 63	32. 19	32. 17	0. 960 5	0. 917 7	0. 899 6	0. 928 3
	CBPN	1 036	240. 7	37. 90	33. 60	32. 17	32. 14	0. 959 0	0. 917 1	0. 898 9	0. 927 9
	RFDN	534	123. 0	38. 05	33. 68	32. 16	32. 12	0. 960 6	0. 918 4	0. 899 4	0. 927 8
	RLFN	527	50. 26	38. 07	33. 72	32. 22	32. 33	0. 960 7	0. 918 7	0. 900 0	0. 929 9
	本文模型	710	88. 49	38. 10	33. 75	32. 26	32. 48	0. 961 1	0. 919 6	0. 901 0	0. 931 9
3 倍	SRCNN	57	52. 7	37. 75	29. 30	28. 41	26. 24	0. 909 0	0. 821 5	0. 786 3	0. 798 9
	FSRCNN	13	6. 0	33. 18	29. 37	28. 53	26. 43	0. 914 0	0. 824 0	0. 791 0	0. 808 0
	VDSR	666	612. 6	33. 66	29. 77	28. 82	27. 14	0. 921 3	0. 831 4	0. 797 6	0. 827 9
	DRCN	1 774	9 788. 7	33. 82	29. 76	28. 80	27. 15	0. 922 6	0. 831 1	0. 796 3	0. 827 6
	LapSRN	502	149. 4	33. 81	29. 79	28. 82	27. 07	0. 922 0	0. 832 5	0. 798 0	0. 827 5
	AWSRN	477	48. 6	34. 02	30. 09	28. 92	27. 57	0. 924 0	0. 837 6	0. 800 9	0. 839 1
	DRRN	297	6 796. 9	34. 03	29. 96	28. 95	27. 53	0. 924 4	0. 834 9	0. 800 4	0. 837 3
	CARN	1 592	118. 8	34. 29	30. 29	29. 06	28. 06	0. 925 5	0. 840 7	0. 803 4	0. 849 3
	IDN	553	57. 0	34. 11	29. 99	28. 95	27. 42	0. 925 3	0. 835 4	0. 801 3	0. 835 9
	LAPAR-A	544	114. 0	34. 36	30. 34	29. 11	28. 15	0. 926 7	0. 842 1	0. 805 4	0. 852 3
	MemNet	678	623. 9	34. 09	30. 00	28. 96	27. 56	0. 924 8	0. 835 0	0. 800 1	0. 837 6
	IMDN	703	71. 5	34. 36	30. 32	29. 09	28. 17	0. 927 0	0. 841 7	0. 804 6	0. 851 9
	RFDN	541	55. 4	34. 41	30. 34	29. 09	28. 21	0. 927 3	0. 842 0	0. 805 0	0. 852 5
	本文模型	718	89. 31	34. 37	30. 35	29. 07	28. 32	0. 927 6	0. 842 8	0. 805 4	0. 855 8
4 倍	SRCNN	57	52. 7	30. 48	27. 50	26. 90	24. 52	0. 862 6	0. 710 1	0. 751 3	0. 722 1
	FSRCNN	12	6. 0	30. 72	27. 61	26. 98	24. 62	0. 866 0	0. 715 0	0. 755 0	0. 728 0
	VDSR	665	612. 6	31. 35	28. 01	27. 29	25. 18	0. 883 8	0. 725 1	0. 767 4	0. 752 4
	DRCN	1 774	9788. 7	31. 53	28. 02	27. 23	25. 14	0. 885 4	0. 723 3	0. 767 0	0. 751 0
	LapSRN	813	149. 4	31. 54	28. 09	27. 32	25. 21	0. 885 2	0. 727 5	0. 770 0	0. 756 2
	AWSRN	588	37. 7	31. 77	28. 35	27. 41	25. 56	0. 889 3	0. 730 4	0. 776 1	0. 767 8
	DRRN	297	6796. 9	31. 68	28. 21	27. 38	25. 44	0. 888 8	0. 728 4	0. 772 0	0. 763 8
	CARN	1 592	90. 9	32. 13	28. 60	27. 58	26. 07	0. 893 7	0. 734 9	0. 780 6	0. 783 7
	IDN	553	32. 3	31. 82	28. 25	27. 41	25. 41	0. 890 3	0. 729 7	0. 773 0	0. 763 2
	LAPAR-A	659	94. 0	32. 15	28. 61	27. 61	26. 14	0. 894 4	0. 736 6	0. 781 8	0. 787 1
	MemNet	678	623. 9	31. 74	28. 26	27. 40	25. 50	0. 889 3	0. 728 1	0. 772 3	0. 763 0
	IMDN	715	40. 9	32. 21	28. 58	27. 56	26. 04	0. 894 8	0. 735 3	0. 781 1	0. 783 8
	CBPN	1 197	97. 9	32. 21	28. 63	27. 58	26. 14	0. 894 4	0. 735 6	0. 781 3	0. 786 9
	RFDN	550	31. 6	32. 24	28. 61	27. 57	26. 11	0. 895 2	0. 736 0	0. 781 9	0. 785 8
	RLFN	527	51. 96	32. 24	28. 62	27. 60	26. 17	0. 895 2	0. 736 4	0. 781 3	0. 787 7
	本文模型	730	90. 45	32. 40	28. 71	27. 64	26. 37	0. 897 7	0. 738 8	0. 784 7	0. 794 8

像的结构和边缘更加清晰,整体效果明显提升。

然而,对于一些纹理比较密集且不规则的图像,本文的重建效果没有显著提高。BSD100 数据集的 Image8023 中包含较多不规则的黑色纹理图案,本文重建图像中出现黑色边缘纹理缺失、轮廓模糊的现象;Urban100 数据集的 Image024 含有较多建筑物的规则结构,本文重建的图像中存在扭曲、失真且细

节不完整的情况。本文方法在重建过程中主要关注图像的高频特征和多尺度特征,并未考虑图像局部与非局部信息之间的交互性,重建出图像的轮廓虽然更加清晰,但是图像一些局部的复杂纹理区域重建效果并不显著。如何进一步提高密集且不规则的形状和结构的重建效果,仍是本文进一步研究和解决的问题。

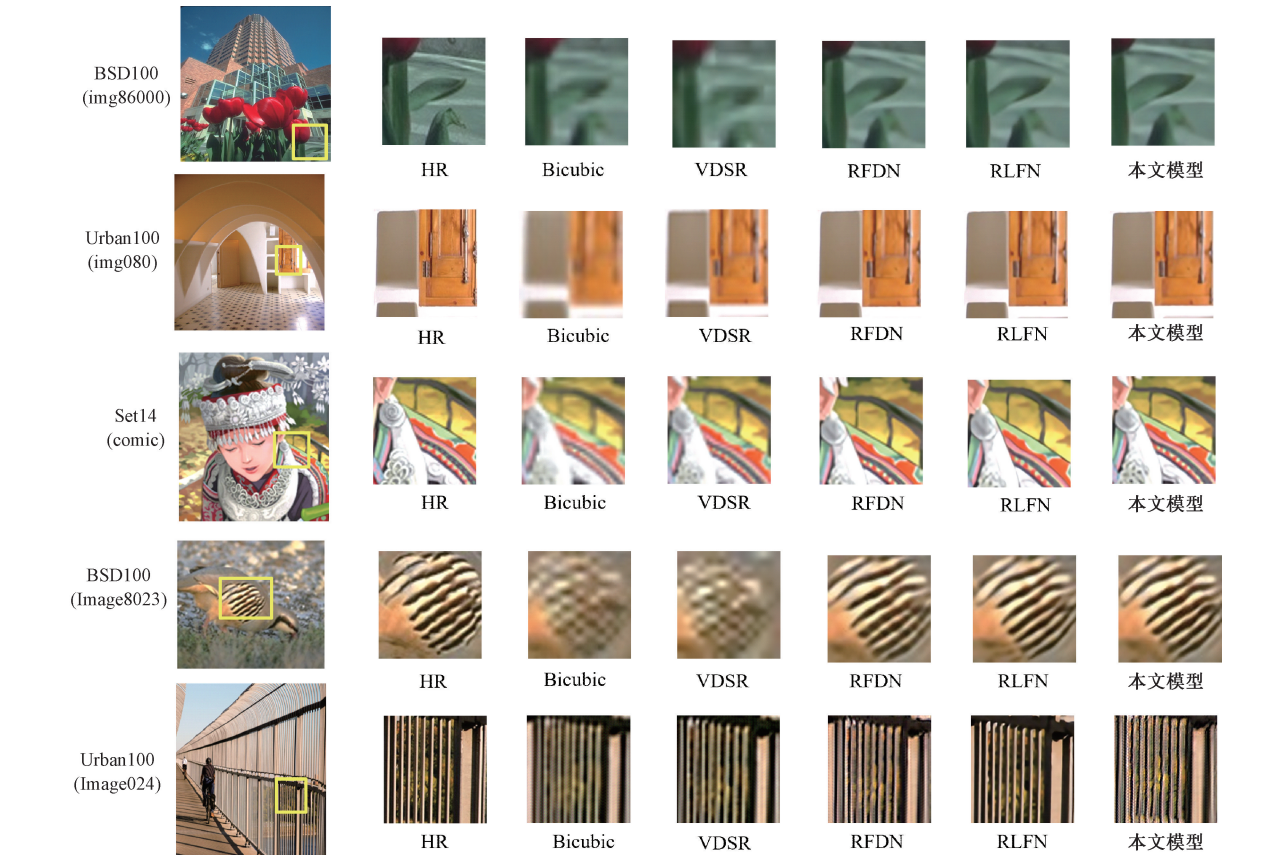


图 8 放大 4 倍时各网络视觉效果对比

Figure 8 Visual effect comparison by different networks with 4 times of scale factor

3 结 论

本文提出了一种基于双特征提取和注意力机制的图像超分辨率重建方法。使用双分支网络有效传递信息,将提取的高频特征和多尺度特征重建后进一步融合,提高图像的重建质量。通过构建残差融合注意力模块,将局部空间注意力和通道注意力相结合,增强了网络在空间维度和通道维度上对高频特征的提取能力,使网络更准确地判别和定位高频特征;使用不同类型的卷积构建空洞金字塔模块,扩大了网络感受野,提升网络多尺度特征的提取能力。本文训练数据集使用 DIV2K 数据集,PSNR、SSIM 作为评价指标,实验结果表明,本文方法重建出的图像有较好的纹理细节,PSNR 和 SSIM 都有所提升,证明了本文研究方法的有效性。但是该模型计算量略

大,后续工作将在确保算法重建精度的前提下,进一步降低参数量并减小网络结构。

参考文献:

[1] 成科扬, 荣兰, 蒋森林, 等. 基于深度学习的遥感图像超分辨率重建方法综述[J]. 郑州大学学报(工学版), 2022, 43(5): 8-16.
CHENG K Y, RONG L, JIANG S L, et al. Overview of methods for remote sensing image super-resolution reconstruction based on deep learning[J]. Journal of Zhengzhou University (Engineering Science), 2022, 43(5): 8-16.
[2] DONG C, LOY C C, HE K M, et al. Learning a deep convolutional network for image super-resolution[C]//European Conference on Computer Vision. Cham: Springer, 2014: 184-199.
[3] DONG C, LOY C C, TANG X O. Accelerating the su-

- per-resolution convolutional neural network [C]//European Conference on Computer Vision. Cham: Springer, 2016: 391–407.
- [4] 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 (CVPR). Piscataway:IEEE, 2016: 1646–1654.
- [5] 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). Piscataway:IEEE, 2017: 1132–1140.
- [6] 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). Piscataway:IEEE, 2017: 5835–5843.
- [7] 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 (CVPR). Piscataway: IEEE, 2016: 1637–1645.
- [8] 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). Piscataway:IEEE, 2017: 2790–2798.
- [9] ZHANG Y L, LI K P, LI K, et al. Image super-resolution using very deep residual channel attention networks [C]//European Conference on Computer Vision. Cham: Springer, 2018: 294–310.
- [10] AHN N, KANG B, SOHN K A. Fast, accurate, and lightweight super-resolution with cascading residual network [C]//Computer Vision—ECCV 2018: 15th European Conference. New York:ACM, 2018: 256–272.
- [11] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7132–7141.
- [12] WU B C, WAN A, YUE X Y, et al. Shift: a zero FLOP, zero parameter alternative to spatial convolutions [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 9127–9135.
- [13] 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). Piscataway:IEEE, 2017: 4809–4817.
- [14] 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). Piscataway:IEEE, 2018: 814–822.
- [15] 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: A Publication of the IEEE Signal Processing Society, 2004, 13(4): 600–612.
- [16] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [EB/OL]. (2015–11–23) [2023–06–14]. <https://arxiv.org/abs/1511.07122>.
- [17] LI W B, ZHOU K, QI L, et al. LAPAR: linearly-assembled pixel-adaptive regression network for single image super-resolution and beyond [EB/OL]. (2021–05–21) [2023–06–14]. <https://arxiv.org/abs/2105.10422>.
- [18] 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). Piscataway:IEEE, 2017: 4549–4557.
- [19] WANG C F, LI Z, SHI J. Lightweight image super-resolution with adaptive weighted learning network [EB/OL]. (2019–04–04) [2023–06–14]. <https://arxiv.org/abs/1904.02358>.
- [20] 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. Piscataway:IEEE, 2018: 723–731.
- [21] HUI Z, GAO X B, YANG Y C, et al. Lightweight image super-resolution with information multi-distillation network [C]//Proceedings of the 27th ACM International Conference on Multimedia. New York: ACM, 2019: 2024–2032.
- [22] ZHU F Y, ZHAO Q J. Efficient single image super-resolution via hybrid residual feature learning with compact back-projection network [C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Piscataway:IEEE, 2019: 2453–2460.
- [23] LIU J, TANG J, WU G S. Residual feature distillation network for lightweight image super-resolution [C]//European Conference on Computer Vision. Cham: Springer, 2020: 41–55.
- [24] KONG F Y, LI M X, LIU S W, et al. Residual local feature network for efficient super-resolution [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2022: 765–775.