

基于组-信息蒸馏残差网络的轻量级图像超分辨率重建

王云涛¹ 赵 蔺² 刘李漫¹ 陶文兵²

摘要 目前, 基于深度学习的超分辨算法已经取得了很好性能, 但这些方法通常具有较大内存消耗和较高计算复杂度, 很难应用到低算力或便携式设备上. 为了解决这个问题, 设计一种轻量级的组-信息蒸馏残差网络 (Group-information distillation residual network, G-IDRN) 用于快速且精确的单图像超分辨率任务. 具体地, 提出一个更加有效的组-信息蒸馏模块 (Group-information distillation block, G-IDB) 作为网络特征提取基本块. 同时, 引入密集快捷连接, 对多个基本块进行组合, 构建组-信息蒸馏残差组 (Group-information distillation residual group, G-IDRG), 捕获多层次信息和有效重用特征. 另外, 还提出一个轻量的非对称残差 Non-local 模块, 对长距离依赖关系进行建模, 进一步提升超分性能. 最后, 设计一个高频损失函数, 去解决像素损失带来图像细节平滑的问题. 大量实验结果表明, 该算法相较于其他先进方法, 可以在图像超分辨率性能和模型复杂度之间取得更好平衡, 其在公开测试数据集 B100 上, 4 倍超分速率达到 56 FPS, 比残差注意力网络快 15 倍.

关键词 残差网络, 超分辨率, 特征蒸馏, 高频损失

引用格式 王云涛, 赵蔺, 刘李漫, 陶文兵. 基于组-信息蒸馏残差网络的轻量级图像超分辨率重建. 自动化学报, 2024, 50(10): 2063-2078

DOI 10.16383/j.aas.c211089

CSTR 32138.14.j.aas.c211089

G-IDRN: A Group-information Distillation Residual Network for Lightweight Image Super-resolution

WANG Yun-Tao¹ ZHAO Lin² LIU Li-Man¹ TAO Wen-Bing²

Abstract Recently, most super-resolution algorithms based on deep learning have achieved satisfactory results. However, these methods generally consume large memory and have high computational complexity, and are difficult to apply to low computing power or portable devices. To address this problem, this paper introduces a lightweight group-information distillation residual network (G-IDRN) for fast and accurate single image super-resolution. Specially, we propose a more effective group-information distillation block (G-IDB) as the basic block for feature extraction. Simultaneously, we introduce dense shortcut to combine them to construct a group-information distillation residual group (G-IDRG), which is used to capture multi-level information and effectively reuse the learned features. Moreover, a lightweight asymmetric residual Non-local block is proposed to model the long-range dependencies and further improve the performance of super-resolution. Finally, a high-frequency loss function is designed to alleviate the problem of smoothing image details caused by pixel-wise loss. Extensive experiments show the proposed algorithm achieves a better trade-off between image super-resolution performance and model complexity against other state-of-the-art super-resolution methods and gets 56 FPS on the public test dataset B100 with a scale factor of 4 times, which is 15 times faster than the residual channel attention network.

Key words Residual network, super-resolution, feature distillation, high-frequency loss

Citation Wang Yun-Tao, Zhao Lin, Liu Li-Man, Tao Wen-Bing. G-IDRN: A group-information distillation residual network for lightweight image super-resolution. *Acta Automatica Sinica*, 2024, 50(10): 2063-2078

收稿日期 2021-11-17 录用日期 2022-06-16

Manuscript received November 17, 2021; accepted June 16, 2022

国家自然科学基金 (61976227, 62176096), 湖北省自然科学基金 (2019CFB622) 资助

Supported by National Natural Science Foundation of China (61976227, 62176096) and Natural Science Foundation of Hubei Province (2019CFB622)

本文责任编辑 张军平

Recommended by Associate Editor ZHANG Jun-Ping

1. 中南民族大学生物医学工程学院 武汉 430074 2. 华中科技大学人工智能与自动化学院 武汉 430074

1. School of Biomedical Engineering, South-central Minzu University, Wuhan 430074 2. School of Artificial Intelligence and

单图像超分辨率或称超分 (Single image super-resolution, SISR 或 SR) 技术旨在从低分辨率 (Low-resolution, LR) 图像中重建出它的高分辨率 (High-resolution, HR) 图像. 它在真实场景中具有广泛的应用, 如医学成像^[1]、视频监控^[2]、高清显示和成像^[3]、图像压缩^[4]、卫星遥感成像^[5] 等. 然而, 数字设备在收集图像过程中, 常受到拍摄物周围环境、

Automation, Huazhong University of Science and Technology, Wuhan 430074

相机硬件和人为因素影响,导致捕获到的图像不清晰,这严重影响后续视觉任务性能.因此,针对如何从低分辨率图像复原出可靠的高清图像进行深入研究(即对图像超分辨率技术的研究)具有重要意义.

通常,由于一张 LR 图像可以对应多张 HR 图像,因此,超分辨率本质上是一个非适定性问题.为了解决这个问题,有许多超分辨率方法被提出,包括早期的传统方法^[6-8]和最近的基于学习的方法^[9-13].传统方法包括基于插值方法和基于模型方法 2 种.基于插值方式(如双线性插值和双三次插值)^[7]操作简单、有效,但应用场景范围和超分辨率效果都有限.基于模型方法(如文献^[14]和文献^[8])分别利用稀疏先验信息和自相似先验信息进行图像超分辨率重建,但这些方法极依赖先验信息,且处理优化过程非常耗时.

最近,研究者们提出基于卷积神经网络的方法来解决图像 SR 问题.超分辨率卷积神经网络(Super-resolution convolutional neural network, SRCNN)^[15]首次将深度学习用于图像超分辨率,提出一个 3 层网络学习从 LR 图像到 HR 图像的映射.但是, SRCNN 在图像块输入前,使用双三次插值方式放大到目标尺寸,这种前上采样方式会引入额外的参数量和计算量代价. Dong 等^[16]提出后上采样方式的快速 SR 卷积神经网络(Fast super-resolution convolutional neural network, FSRCNN),即在网络最后一层使用上采样放大到目标尺寸,因此,可以直接将低分辨率图像输入到网络中,在节省参数量和计算量的同时,保证了图像超分辨率效果.之后,一些 SR 方法主要集中在设计更深或更宽的模型,以进一步提高 SR 的性能,例如更深的 SR 网络(Very deep super-resolution network, VDSR)^[17]、深度递归卷积网络(Deeply-recursive convolutional network, DRCN)^[18]、增强深度 SR 网络(Enhanced deep super-resolution network, ED-SR)^[19]和残差通道注意力网络(Residual channel attention network, RCAN)^[20].尽管这些方法取得了令人满意的结果,但是通过增大网络深度提升图像超分辨率性能会增大网络计算成本,不适用于诸如手机和相机等便携式设备.

为了减少计算成本或内存消耗,研究者提出级联残差网络(Cascading residual network, CARN)^[21],它是一种级联网络体系结构,但该方法获得的超分性能较差.信息蒸馏网络(Information distillation network, IDN)^[22]将当前信息与局部信息进行融合.随后,信息多蒸馏网络(Information multi-distillation network, IMDN)^[11]设计一个信息多蒸馏模块,以进一步提高 IDN 的性能.残差特征蒸

馏网络(Residual feature distillation network, RF-DN)^[12]在 IMDN 基础上提出一种更轻便、更灵活的残差蒸馏网络.但是这些方法不够轻量,并且图像 SR 性能仍可以进一步提高.为了构建更快、更轻量 SR 模型,本文改进了 RFDN 中残差特征蒸馏模块(Residual feature distillation block, RFDB),借助组卷积在降低模型复杂度方面的优势,提出一种新颖的组-信息蒸馏模块(Group-information distillation block, G-IDB).结合密集快捷(Dense shortcut, DS)^[23]连接方式,构建一个轻量级的组-信息蒸馏残差组(Group-information distillation residual group, G-IDRG),通过 DS 连接对学习到的多层级特征进行重利用.考虑到图像超分辨率是图到图的转换,图像中存在大量重复或相似的模式,对这种相似性关系进行关注能够进一步提高 SR 的性能,因此本文借鉴非对称 Non-local 模块(Asymmetric non-local block, ANB)^[24]中非局部注意力机制思想,提出一个适用于轻量 SR 的非对称残差 Non-local 模块(Asymmetric non-local residual block, ANRB),对长距离依赖关系进行建模,捕获全局上下文信息.基于上述分析,本文构建一个性能更好的轻量级组-信息蒸馏残差网络(Group-information distillation residual network, G-IDRN),以实现快速、准确的图像 SR.如图 1 所示,本文方法在模型复杂度和性能之间取得了更好平衡.

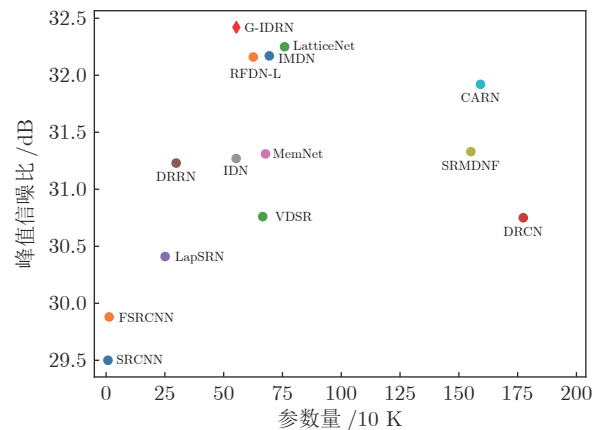


图 1 Urban100 中图像放大 2 倍时,参数量和峰值信噪比的对比结果

Fig. 1 Comparison results of the number of parameters and the peak single-to-noise ratio for 2 times images on Urban100

在图像 SR 任务中,通常使用平均绝对误差(L_1 损失)和均方误差(L_2 损失)来衡量超分辨率图像和真实 HR 图像之间在像素上的差异.但仅使用逐像素损失将导致超分辨率图像经常缺少高频细节,如图 2 所示.因此,本文提出一种简单而有效的高频损

失, 以缓解超分辨率图像过度平滑问题. 如图 2(a) ~ 图 2(c) 所示, 使用逐像素损失方法生成的 SR 图像纹理过于平滑, 而使用本文基于高频损失方法在视觉上获得了令人满意的结果, 如图 2(d) 所示. 具体地, 本文首先使用高斯滤波器从真实 HR 图像和预测的 SR 图像中提取高频信息, 然后采用两者之间的平均绝对误差作为高频损失, 使得网络更加关注细节纹理区域.

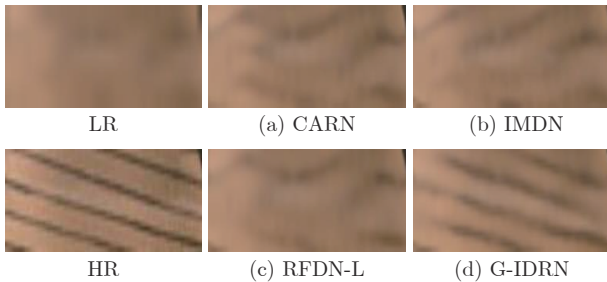


图 2 Urban100 中 Img024 放大 4 倍时, 不同 SR 方法的重建结果

Fig. 2 The reconstruction results of various SR methods for 4 times Img024 on Urban100

本文针对单图像的轻量级超分辨率展开研究, 主要贡献如下:

- 1) 提出一种轻量级的组-信息特征残差组, 可以更好地利用多层次特征信息, 重建图像的高频细节信息.
- 2) 提出一种轻量级非对称残差 Non-local 模块来捕获全局上下文信息, 进一步提高了 SISR 的性能.
- 3) 设计一个简单而有效的高频损失函数, 来缓解超分辨率图像的过度光滑问题. 在多个基准数据集上实验表明了该方法在 SISR 任务中的优越性和有效性.

本文结构如下: 第 1 节对相关工作进行介绍; 第 2 节对本文方法进行细致描述; 第 3 节进行大量消融实验, 对实验结果进行定量和定性对比和分析; 第 4 节对全文进行总结.

1 相关工作

图像 SR 方法大致可分为传统方法^[8, 25-26]和深度学习方法^[10-12]两类. 本节主要总结深度学习方法的相关工作.

1.1 单图像超分辨率

文献 [15] 是将深度学习直接应用于图像 SR 的首批研究之一, 提出一个由 3 层卷积组成的超分辨率网络 SRCNN. 该网络包含图像块提取表征、特征非线性映射和重建 3 个部分. 利用卷积网络学习出

低分辨率到高清图像之间的映射函数, 有效提升了图像超分辨率效果. 受这项开拓性工作的启发, VD-SR 和 DRCN 基于残差学习堆叠了 16 个以上卷积层, 以进一步提高性能. 为了进一步释放深层卷积网络的性能, EDSR 将修改后的残差块集成到 SR 框架中, 从而形成一个非常宽和深的网络. 记忆网络 (Memory network, MemNet)^[27] 通过堆叠密集块以形成深度模型, 利用所有卷积层的层次化信息. Li 等^[28] 提出一种反馈机制生成有效的高级特征表示. Qiu 等^[29] 通过逐渐恢复过程, 处理纹理. 尽管这些方法实现了显著的性能, 但它们在内存消耗和计算复杂性方面的成本很高, 这限制了在资源受限设备中的应用.

因此, 最近一些研究提出了快速和轻量级的 SISR 算法, 来解决图像 SR 问题. 这些算法可以大致分为基于知识蒸馏的方法^[11-12, 22]、基于神经网络架构搜索的方法^[30-31]和基于模型的方法^[21, 32]三类. 知识蒸馏的目的是将知识从教师网络转移到学生网络. IDN 方法提出一种信息蒸馏网络, 以通过分离当前特征图, 来更好地利用层次化信息. 基于 IDN 方法, Hui 等^[11] 通过构造级联的信息多蒸馏块, 提出一种信息多蒸馏网络. RFDN 利用多个特征蒸馏连接学习更有表征性的特征表示. 而 Chu 等^[30-31] 应用神经网络架构搜索进行图像 SR, 由于策略限制, 这些方法很难达到较好性能. 此外, 文献 [21] 提出一种基于残差网络的级联机制以促进超分的性能. 晶格网络 (Lattice network, LatticeNet)^[32] 提出了一种晶格块, 应用两个蝶形结构来组合两个残差块. 这些研究表明, 轻量级 SR 网络可以在性能和模型复杂性之间保持更好平衡.

1.2 注意力机制

注意力机制是一项重要的技术, 已广泛用于各种视觉任务中, 如分类、目标检测和图像分割等. Hu 等^[33] 首次将通道注意力用于图像分类任务中, 通过对通道关系进行建模以增强网络的表达能力. Non-local^[34] 通过计算像素位置的响应作为图像所有像素位置特征的加权总和来捕获长距离依赖关系. 在图像 SR 域中, RCAN 和 Liu 等^[35] 通过考虑通道或空间维度上的注意力机制, 提高性能. Dai 等^[10] 提出一种二阶注意力机制, 以增强特征表达和相关性学习. Mei 等^[36] 通过研究跨尺度特征相关性, 提出了跨尺度非局部注意力模块. Niu 等^[37] 对层、通道和位置之间的整体相互依赖性进行建模. 基于注意力模型的有效性, 本文也将注意力机制嵌入到所提的框架中, 以改善高级特征表示. 然而, 这些注意力模型均为非轻量级, 不适用于轻量级超分网络

中. 同时, 考虑到图像中存在重复或相似模式, 本文设计轻量级的注意力机制, 来捕获长距离的依赖关系, 构建全局注意力图.

1.3 感知优化

在图像 SR 任务中, 模型优化的目标函数大多数使用预测图像和真实 HR 图像之间像素的距离作为其损失项, 如 L_1 损失和 L_2 损失. 但研究发现, 仅使用这种损失函数会导致生成模糊和过度平滑的超分辨率图像. 因此, 一些研究人员提出了多种损失函数来指导模型优化, 如在 SR 中引入内容损失^[38], 以优化特征重建错误. Sajjadi 等^[3] 使用纹理损失, 在视觉上产生更令人满意的结果; Ledig 等^[39] 使用对抗损失产生更接近自然图像的输出; Yuan 等^[40] 使用周期一致性损失避免生成对抗网络的崩溃问题, 并有助于最小化分布差异. 然而, 这些损失通常与超分生成对抗网络中的对抗性损失结合使用, 单独使用将导致性能极大下降^[3, 39]. 本文使用像素损失, 在保证 SR 结果与真实 HR 图一致性的同时, 设计一种新颖的高频损失函数学习高频信息, 恢复高保真的纹理细节.

2 组-信息蒸馏残差网络

首先, 对所提网络的整体结构进行介绍; 其次, 分别阐述组-信息蒸馏残差组和非对称残差 Non-local 模块; 最后, 描述网络的复合损失, 包括重建损失和本文的高频损失.

2.1 网络结构

如图 3 所示, 本文网络基于 RFDN 的组-信息

蒸馏残差网络, 由浅层特征提取、深层特征提取和最终的图像超分辨率重建 3 个部分组成. 在图像 SR 任务中, 各像素点的回归值与全局上下文信息密切相关. 因此, 设计一个轻量级的非对称残差 Non-local 模块.

1) 浅层特征提取. 依据大多数超分方法, 网络的第 1 层采用标准的 3×3 卷积, 提取输入 RGB 图像的浅层信息, 用于之后进一步特征提取. 浅层特征提取过程描述如下:

$$F_0 = H_{3 \times 3}(I_{LR}) \quad (1)$$

式中, I_{LR} 表示低分辨率输入图像, $H_{3 \times 3}(\cdot)$ 表示 3×3 卷积操作, F_0 表示提取到的浅层特征.

2) 深层特征提取. 首先, 使用 3 个 G-IDRG 模块 (见第 2.2 节) 进行特征提取; 然后, 采用 DS 连接融合 3 个模块的特征用于后续的图像重建. 深层特征提取的过程如下:

$$F_D = H_{DS}\{F_{RG_1}, F_{RG_2}, F_{RG_3}\} \quad (2)$$

式中, F_{RG_i} ($i = 1, 2, 3$) 表示经过第 i 个 G-IDRG 获得的特征图, $H_{DS}\{\cdot\}$ 表示 DS 连接操作 (见第 2.2 节). 深层特征提取部分的输出特征为 F_D .

3) 上采样超分重建. 深层提取后, RFDN 使用一个全局跳跃连接, 并在跳跃连接的前/后使用 2 层卷积. 然而 Yu 等^[41] 指出, 这 2 层不带激活的卷积相当于线性变换, 冗余的卷积层带来较大计算成本. 他们认为这些层的效果可以吸收到残差结构里, 通过去除冗余卷积层, 实验结果显示效果并未下降. 因此, 本文去除了全局连接和 2 层卷积, 同时使用一个重建支路, 即使用 5×5 卷积层直接从原始输入提取粗尺度特征, 再直接进行重建, 见式 (3). 特

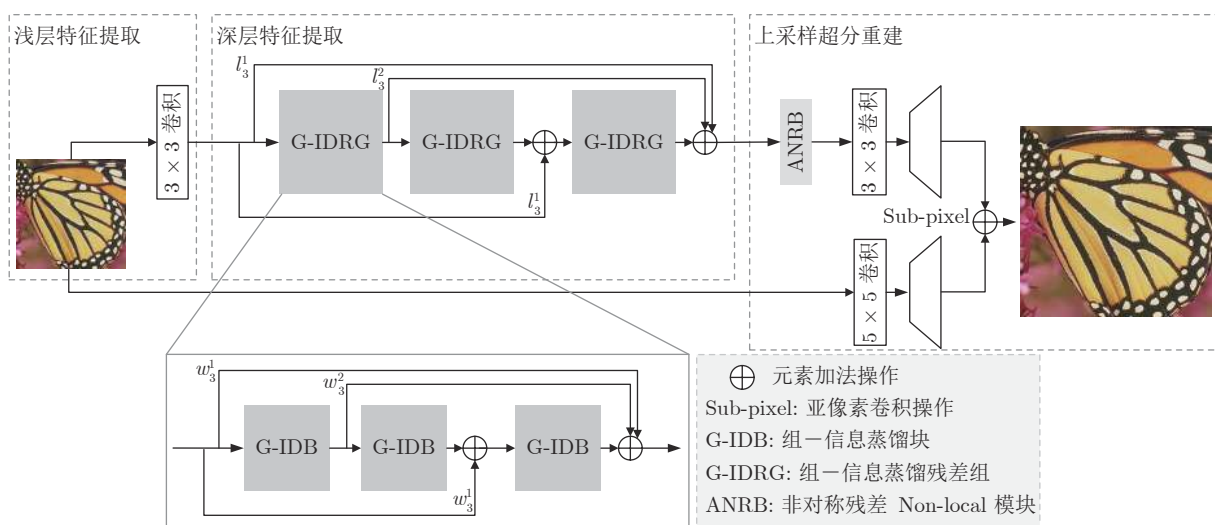


图 3 组-信息蒸馏残差网络整体架构

Fig.3 The architecture of the group-information distillation residual network

别地, 本文称这种改进的重建方式为双路重建策略. 同时, 由于在图像 SR 任务中, 相似模式建模非常重要, 本文提出一个轻量级的非对称残差 Non-local 模块 (见第 2.3 节), 来建模全局上下文相似信息, 并进一步细化学习到的特征. 过程如下:

$$\begin{cases} I_{SR_1} = H_{Up}(H_{ANRB}(F_D)) \\ I_{SR_2} = H_{Up}(H_{5 \times 5}(I_{LR})) \end{cases} \quad (3)$$

式中, $H_{ANRB}(\cdot)$ 表示 ANRB 操作, $H_{5 \times 5}(\cdot)$ 表示 5×5 卷积操作, I_{SR_1} 表示经过深层特征提取步骤和 ANRB 再进行重建获得的精细高分辨率图像, I_{SR_2} 表示经过一层卷积然后直接进行重建获得的粗糙高分辨率图像, $H_{Up}(\cdot)$ 表示利用可学习的卷积和非参数亚像素卷积操作^[42]重建 HR 图像.

最后, I_{SR_1} 与 I_{SR_2} 相加, 得到最终的 SR 输出:

$$I_{SR} = I_{SR_1} + I_{SR_2} \quad (4)$$

类似于经典的 SR 方法^[11-12, 19], 本文同样使用 L_1 损失优化本文设计网络, 并提出一个高频损失函数 (见第 2.4 节), 使网络更关注高频信息的学习.

2.2 组-信息蒸馏残差组

组-信息蒸馏残差组由多个 DS 连接的组-信息蒸馏模块组成.

AIM2020 有效超分比赛的冠军方案提出一种残差特征蒸馏块, 如图 4(a) 所示. RFDB 实质为一个渐进的细化特征模块, 其包含多个蒸馏步. 单个蒸馏步分为输入同时经过 1×1 卷积层和浅层残差块 (Shallow residual block, SRB) 左/右 2 个部分. 左部分将通道压缩后保留下来, 右部分细化特征后,

作为下一个蒸馏步的输入. 然后, 使用一层 1×1 卷积进行特征微调, 再通过一个对比度通道注意力 (Contrast-aware channel attention, CCA) 模块生成注意力特征图. 虽然 RFDB 取得了很好效果, 但是仍然存在大量冗余. 首先, 每个蒸馏步的左部分先通过 1×1 卷积层进行通道压缩, 然后在 Concat 融合操作中再次使用 1×1 卷积层进行通道压缩, 2 次通道压缩操作是减少通道以降低复杂度和通道蒸馏, 作用是相同的, 可以考虑将之融合为一次操作, 减少卷积层的使用, 使模型更简洁、有效.

1) 组-信息蒸馏模块. RFDB 中使用的均为标准卷积, 而对于轻量级超分任务, 引入更加轻量的卷积操作是很有必要的. Howard 等^[43]提出一种深度可分离卷积 (由点卷积和深度卷积组成) 用来提取特征, 相比常规的卷积操作, 其参数数量和计算成本较低. 然而, 深度卷积也带来网络运行时间变慢、内存消耗过大问题. Xie 等^[44]提出深度卷积更一般形式的卷积——分组卷积 (Group conv, GC), 当分组数等于通道数时, 分组卷积退化为深度卷积:

$$\begin{cases} N_{Conv} = C_{in} \times K^2 \times C_{out} \\ F_{Conv} = N_{Conv} \times W \times H \end{cases} \quad (5)$$

$$\begin{cases} N_{GC} = C_{in} \times K^2 \times \frac{C_{out}}{g} \\ F_{GC} = N_{GC} \times W \times H \end{cases} \quad (6)$$

式中, N_{Conv} 、 F_{Conv} 和 N_{GC} 、 F_{GC} 分别表示标准卷积和分组卷积的参数数量、计算量; C_{in} 、 C_{out} 和 K 分别表示输入通道数、输出通道数和卷积核大小; W 和 H 分别表示特征图的宽和高; $g = C/m$ 为 GC 的

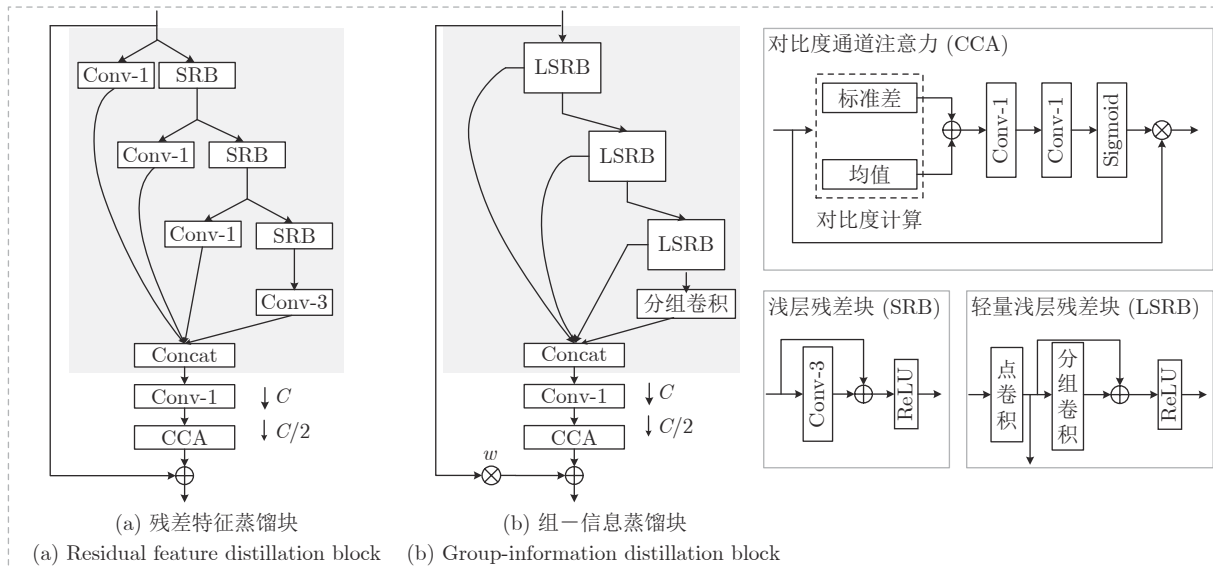


图 4 G-IDB 对 RFDB 的改进图

Fig. 4 G-IDB improvements to RFDB

分组数, 其中 C 为网络的基础通道数, m 表示每组通道个数. 一般对于 3×3 大小的卷积核, GC 的计算量和参数量为标准卷积的 $1/g$ 倍, 而内存消耗和计算成本与 g 成正比. 灵活调整 g , 可以较好地平衡参数量和内存消耗.

结合深度可分离卷积和分组卷积, 本文设计一种分组可分离卷积, 同时结合 RFDB 中的浅层残差块, 提出一个轻量 SRB (Lightweight shallow residual block, LSRB) 结构 (如图 4 所示), 以此设计一个组-信息蒸馏模块, 如图 4(b) 所示. 与深度可分离卷积先逐通道卷积、再对深度上加权组合不同, 首先, LSRB 使用点卷积对通道方向上加权求和. 然后, 使用分组卷积. 其中, 点卷积的输出保留作为后续 Concat 融合的输入, LSRB 的输出作为下一个蒸馏步的输入. 最后, 使用一个分组卷积层对学习到的特征进行微调. 考虑到当特征图增加到一定水平时, 会导致训练时数值不稳定^[45], 因此在最后的残差连接 (Residual shortcut, RS) 中引入可学习的权重 w .

2) 组-信息蒸馏残差组. 在计算机视觉任务中, 残差连接是一项重要技术. 受益于 RS 连接, 大量 SR 方法获得了超分性能的极大提高. 虽然 RS 连接能够将前层信息传输到后层, 但其缺乏灵活性, 只是简单地将输出与之前的输入相加. 为此, Huang 等^[46] 提出密集连接方式, 将浅层信息馈送到后面的每一层, 但是这种方式极其消耗内存. 受到用于图像分类中密集快捷连接的启发, 它结合了 RS 的高效性和密集连接的性能性, 本文将 DS 连接运用到本文 SR 模型中. 如图 3 所示, 考虑到本文轻量级思想, 采用 DS 连接最基本结构 (即对 3 个 G-IDB 使用 DS 连接) 构成一个组-信息蒸馏残差组模块, 其中 w_i^j 为可学习的权值, 表示第 i 个 G-IDB 的输入和第 j 个 G-IDB 输出带权重的跨模块 DS 连接. 同时, 为了减少模型的复杂度, 本文没有使用 DS 连接中的归一化技术. 另外, 对由 G-IDB 组成的 3 个 G-IDRG 使用 DS 连接, 为了区别, 将权重值 w_i^j 用 l_i^j 表示.

2.3 非对称残差 Non-local 模块

目前, 大多数 SR 方法是通过堆叠多个卷积层为网络提供一个较大的感受野, 这种方式存在网络深和效率低等问题, 且忽略了自然图像中存在大量的重复、相似结构或模式, 无法对图像中远距离特征的相似性进行有效建模. 而 Non-local 机制是一种注意力模型, 通过对全局像素之间的联系进行建模, 从而生成对应像素的注意力特征图, 可以有效

捕获长距离特征间的依赖关系. 然而, 传统 Non-local 模块在输入特征图尺寸较大时, 会带来很大计算成本, 无法直接应用在超分辨率网络中. 受到语义分割中 ANB 的启发, 本文提出一种改进的轻量级非对称残差 Non-local 模块, 可有效提升其在轻量的图像超分辨率任务中的适用性, 如图 5 所示.

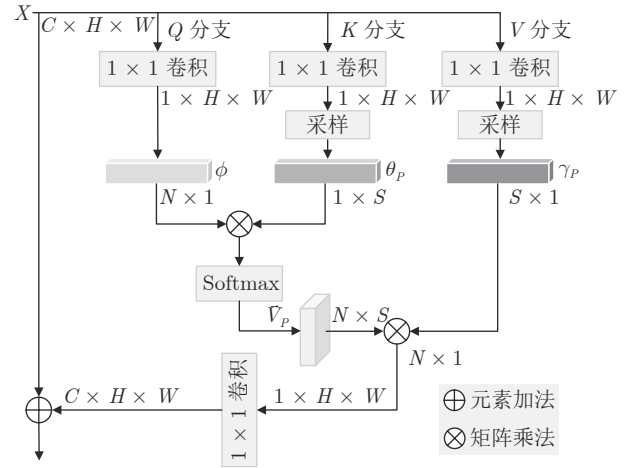


图 5 非对称残差 Non-local 模块

Fig. 5 The asymmetric Non-local residual block

具体地, 对于输入特征图 $X \in \mathbf{R}^{C \times H \times W}$, 其中 $H \times W$ 和 C 分别为输入特征图 X 的空间尺寸和通道数. 首先, 考虑 ANB 均在通道数为 C 的特征图上进行操作, 这会极大增加计算成本, 同时多通道特征图存在冗余. 为此, 本文使用 3 个 1×1 卷积, 将多通道特征 X 压缩为单通道特征 X_ϕ 、 X_θ 和 X_γ . 然后, 与 ANB 类似, 使用金字塔池化采样算法^[47] 从 K 和 V 分支采样 S 个 ($S \ll N = H \times W$) 具有代表性的特征点, 在不牺牲性能的前提下, 极大地减少了计算成本和内存消耗. 具体地, 首先进行 4 个平均池化操作, 分别得到 1×1 、 3×3 、 6×6 和 8×8 空间尺寸特征, 然后对每个特征进行平铺展开再拼接操作, 得到采样后的特征, 其长度 $S = 110$. Non-local 注意力如下:

$$X_\phi = f_\phi(X), \quad X_\theta = f_\theta(X), \quad X_\gamma = f_\gamma(X) \quad (7)$$

$$\theta_P = P_\phi(X_\phi), \quad \gamma_P = P_\gamma(X_\gamma) \quad (8)$$

$$Y = \text{Softmax}(X_\phi^T \otimes \theta_P) \otimes \gamma_P \quad (9)$$

式中, f_ϕ 、 f_θ 和 f_γ 表示通道降维的 1×1 卷积操作; P_ϕ 和 P_γ 分别表示对特征图 θ_P 和 γ_P 的金字塔池化采样方式; Y 表示包含全局语义信息的特征图.

由于注意力权重矩阵 Y 中包含了 Softmax 操作, 如果将 $\phi \otimes \theta_P$ 所得矩阵值的和归一化到 1 会导致 Y 元素存在大量 0; 如果直接与原始输入 X 点相

乘, 会导致所得值过小, 而产生梯度消失问题. 因此, 本文使用元素加法操作 (即残差连接方式) 生成最终的注意力加权特征图 X_{weighted} :

$$X_{\text{weighted}} = H_{1 \times 1}(Y) + X \quad (10)$$

式中, $H_{1 \times 1}(\cdot)$ 表示通道升维的 1×1 卷积操作, 作用是将单通道特征图 Y 转变为 C 通道特征图, 以便于后续的残差连接操作. 受益于通道压缩和采样操作, 非对称残差 Non-local 模块足够轻量且能够获取全局信息, 实现快速且精确的图像超分辨率.

2.4 损失函数

图像 SR 领域最常用的图像目标优化函数是平均绝对误差 (即 L_1 损失) 和均方误差 (即 L_2 损失). 与一些经典超分算法^[11-12, 19-20] 一样, 本文采用 L_1 损失作为主要的重建损失去衡量 SR 图像与真实高分辨率 HR 图像之间的差异. L_1 损失定义如下:

$$\mathcal{L}_1 = \frac{1}{N} \sum_{i=1}^N \|I_{HR}^i - I_{SR}^i\|_1 \quad (11)$$

式中, I_{SR}^i 和 I_{HR}^i 表示由 G-IDRN 生成的第 i 张 SR 图像和相对应的第 i 张 HR 图像.

然而, 仅使用 L_1 或 L_2 损失容易造成 SR 图像缺乏高频细节, 其过于平滑的纹理难以达到视觉上令人满意的结果如图 6 所示. 利用性能级的 SR 方法 RCAN 对 Set14 数据集中 barbara.png 图像放大 3 倍进行超分生成 SR 图像. 与真实 HR 图像对比可以看出, 重建的 SR 图像在细节纹理区域过于平滑. 通过应用高频提取算法对图像进行高频特征提取, 可以清楚地看到 SR 与 HR 的差别. 另外, 对上述 SR 图像的低频和高频部分与对应 HR 的低频和高频部分分别计算峰值信噪比 (Peak single-to-noise ratio, PSNR), 得到 SR 图的 PSNR 值为 26.05 dB, 低频 PSNR 值为 36.72 dB, 高频 PSNR 值为 7.30 dB, 这表明 SR 图仅恢复了极少量的高频信息.

因此, 本文使用简单但有效的高频损失来解决这个问题. 首先使用高频提取算法提取出 SR 图像和 HR 图像的细节纹理. 然后, 采用平均绝对误差衡量 SR 图像和 HR 图像的细节差异. 过程如下:

$$\mathcal{L}_{hf} = \frac{1}{N} \sum_{i=1}^N \|H_{hp}(I_{HR}^i) - H_{hp}(I_{SR}^i)\|_1 \quad (12)$$

式中, $H_{hp}(\cdot)$ 表示高频提取算法. 本文利用高通滤波器提取 SR 与 HR 的细节信息, 因此本文网络训练的目标函数为 $\mathcal{L} = \alpha\mathcal{L}_1 + \beta\mathcal{L}_{hf}$, 其中 α 和 β 为调节 2 种损失的权重.

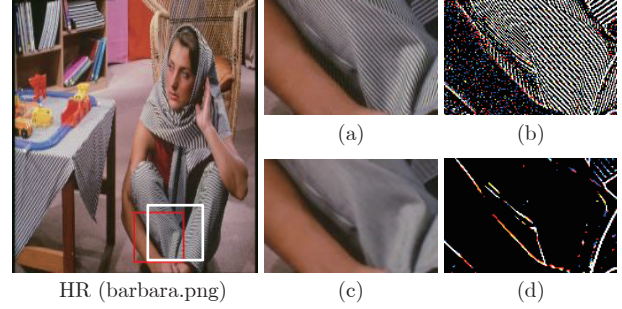


图 6 Set14 中 barbara.png 放大 3 倍的高频提取图像 (a) 裁剪的 HR 图像; (b) HR 图像的高频提取图; (c) 裁剪的 SR 图像; (d) SR 图像的高频提取图

Fig. 6 High-frequency extraction images for 3 times barbara.png on Set14 ((a) Cropped HR image; (b) High-frequency extraction image of HR image; (c) Cropped SR image; (d) High-frequency extraction image of SR image)

3 实验结果和分析

3.1 数据集和指标

DIV2K^[48] 是高质量、高分辨率的图像数据, 包含 1000 张 2 K 分辨率 RGB 图像, 图像内容包含动植物、风景等多个场景, 其中 800 张图像用于训练, 100 张图像用于验证, 100 张图像作为测试集. 根据文献^[11-12, 22], 本文使用 800 张 HR 训练图像训练所有模型, 对应的 LR 图像由 HR 使用双三次下采样生成, 下采样倍数为 2、3 和 4. 为了测试本文模型的性能和泛化性, 选择在 Set5^[49]、Set14^[50]、B100^[51]、Urban100^[25] 和 Manga109^[52] 5 个公开的单图像、超分辨率基准数据集上进行图像质量评价. 在由 RGB 空间变换到 YCbCr 空间的 Y 单通道上分别评估峰值信噪比^[53] 和结构相似度 (Structural similarity, SSIM)^[54] 指标. 用 PSNR 比较 SR 结果图和 HR 图像间的像素差异, 用 SSIM 评价 SR 的结构相似度 (如亮度、对比度和结构). 2 个定量指标的数值越高, 说明对真实图像的恢复保真度越好.

3.2 执行细节

与 EDSR 一样, 使用随机的水平翻转和 90 度旋转进行数据增强. 在训练过程中, 随机从 LR 图像中提取 32 张 64×64 像素的 RGB 图像. 采用 Adam 优化器, 其中一阶矩和二阶矩指数衰减率设置为 $\beta_1 = 0.900$, $\beta_2 = 0.999$, 数值稳定常数 $\epsilon = 10^{-8}$. 初始学习率设置为 5×10^{-4} , 学习率每经过 2×10^5 次迭代反向传播后, 更新为原来的一半. 每轮训练包含 1000 次迭代, 共训练 1000 轮. 根据 IMDN, Leaky ReLU 的超参数设置为 0.05. 设置组卷积的分组数

$g = C/m$, 其中 $C = 72$, $m = 8$. 损失函数的权重设置为 $\alpha = 0.4$, $\beta = 0.6$. 消融实验中的所有模型都从零开始训练, 以节省训练时间. 本文算法模型选择在 NVIDIA GTX1080Ti GPU 上进行训练和测试, 在 Pytorch 框架上实施.

本文所有参数量和计算量 (即浮点运算量 (Floating point operations, FLOPs)) 均由 AIM-2020 Efficient Super-resolution 竞赛官方代码计算所得.

3.3 消融实验分析

为了验证所提网络不同模块的有效性, 本文进行了消融实验. 由于本文方法是 RFDN 的改进模型, 因此将 RFDN 作为所有实验的基线模型. 以 RFDN 框架为基本框架, 分别以 RFDB 和 G-IDB 为基本块, 表 1 给出了双路重建策略、DS 连接和 ANRB 在 Set5 测试集上、放大 2 倍的消融实验结果, 表中“ \times ”表示不使用当前模块或策略, “ \checkmark ”表示使用当前模块或策略.

1) 双路重建策略. 如第 2.1 节所述, 本文去除 RFDN 中的全局跳跃连接和 2 层冗余卷积, 代替为使用 1 个 5×5 卷积的重建支流去实现这种效果, 并称为双路重建策略. 由表 1 中第 1、2 行对比可看出, 双路重建策略在 Set5 数据集上, 获得了较大的性能提升 ($\uparrow 0.038$ dB), 同时参数更少 ($\downarrow 19.8$ K), 其中“ \uparrow ”和“ \downarrow ”分别表示相对表中 RFDB 基本块的第 1 行指标的提升和下降, 并使用 256×256 像素图像计算 RFDN 和带双路重建的 RFDN 模型, 计算量分别为 34.53 G FLOPs 和 33.23 G FLOPs. 在 SR 中, 多数方法强调高频信息的学习, 但低频信息也是极其重要的. 因为人眼对低频信息更敏感, 通常通过一张粗糙图像就能很好地辨识. 使用高斯低通滤波器对原始真实 HR 图像进行低频信息提取, 如图 7 所示. 从整体上看, 低频信息图像基本和原始 HR 图一样, 表明使用大感受野的卷积核从低分

辨率图像提取信息, 然后进行超分重建, 有效保留了人眼敏感的低频信息, 使得最终的超分性能得到提升.

2) DS 连接. 为研究 DS 连接的有效性, 分别对 RFDN 中前 3 个 RFDB 模块和后 3 个 RFDB 模块使用 DS 连接, 并移除原始融合 6 个模块的 Concat 操作. 由表 1 的第 1 行和第 3 行对比可以看出, 将 Concat 操作替换为 DS 后, 在 Set5 数据集上, PSNR 仅减少了 0.002 dB, 但参数量却减少了 14 K. 这表明轻量级的 DS 成功达到了 Concat 的性能, 并实现了残差连接的高效性. 由表 1 中 RFDB 基本块的最后 2 行对比可以看出, DS 与双路重建策略、ANRB 结合, 反而比只使用双路重建策略和 ANRB 性能更优, 这表明 DS 连接能够隐式促进双路重建策略和 ANRB.

3) ANRB. 表 1 第 4 行给出了在 RFDN 增加非对称 Non-local 残差块的实验结果. 相比于原始网络 (RFDB 基本块的第 1 行), 增加了 ANRB 模块后, 参数量增加了 0.3 K (通道数为 48), PSNR 提升了 0.023 dB. 当 ANRB 与双路重建策略同时使用时, 性能有明显提升, 同时参数量大幅降低. 根据 Zhao 等^[47]提出的金字塔池化采样算法规则 (即 4 层不同金字塔池化层应包含 4 种不同尺度特征), 采用 4 个池化操作获得 4 个不同输出特征尺寸分别为 1×1 、 2×2 、 3×3 和 6×6 , 由式 (8) 计算采样特征点长度 $S = 50$. 为了探究采样点数对性能的影响, 首先, 以文献 [47] 的采样方式 (即设置 $S = 50$) 为基础. 然后将采样点数量扩大 2 倍和 4 倍. 当扩大 2 倍时, 考虑池化输出特征尺寸与 S 的对应关系, 本文设置 $S = 110$, 对应的池化输出特征尺寸分别为 1×1 、 3×3 、 6×6 和 8×8 ; 类似地, 当扩大 4 倍时, 设置 $S = 222$, 对应的池化输出特征尺寸分别为 1×1 、 6×6 、 8×8 和 11×11 . 在 Set5 和 Manga109 基准测试数据集的实验结果 (SR 放大 2 倍,

表 1 消融实验结果
Table 1 Ablation experiment results

基本块	双路重建策略	DS 连接	ANRB	PSNR (dB)	参数量 (K)	增幅 PSNR (dB) 参数量 (K)
RFDB	\times	\times	\times	37.893	534.0	0 0
	\checkmark	\times	\times	37.931	514.2	$\uparrow 0.038$ $\downarrow 19.8$
	\times	\checkmark	\times	37.891	520.2	$\downarrow 0.002$ $\downarrow 13.8$
	\times	\times	\checkmark	37.916	534.3	$\uparrow 0.023$ $\uparrow 0.3$
	\checkmark	\times	\checkmark	37.934	514.4	$\uparrow 0.041$ $\downarrow 19.6$
	\checkmark	\checkmark	\checkmark	37.940	500.5	$\uparrow 0.047$ $\downarrow 33.5$
G-IDB	\times	\times	\times	37.955	449.4	$\uparrow 0.062$ $\downarrow 84.6$
	\checkmark	\checkmark	\checkmark	37.965	383.2	$\uparrow 0.072$ $\downarrow 150.8$

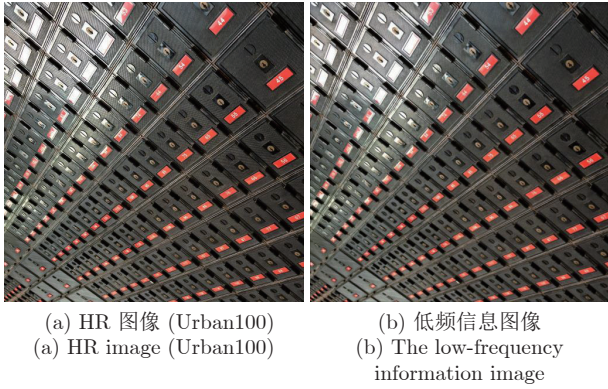


图 7 HR 图像和对应使用高斯低通滤波器提取的低频信息图

Fig.7 HR images and their low-frequency information images extracted by the Gaussian low-pass filter

性能评价指标为 PSNR 和 SSIM) 如表 2 所示, 表中 $S = \infty$ 表示不进行金字塔池化采样, 并分别给出了输入低分辨率图像尺寸在 128×128 像素和 180×180 像素下的内存消耗. 可以看出, 当未使用金字塔池化采样方法时, 随图像尺寸增加, 内存消耗呈指数增大, 如在 Manga109 数据集上, 其输入 LR 图像的平均尺寸为 413×585 像素, 使用 11 GB 1080-Ti 时, 内存仍然溢出, 这很难应用于现实设备; 当进行成倍采样一定的特征点数 (如 $S = 50, 110, 222$) 时, 内存消耗随 S 的上升而小幅增大, 而性能并不是随着 S 的增加而增加; 在相对较大特征点数 (如 $S = 222$) 时, 性能反而较低. 这可能是由于当 S 过大时, 更容易采样到干扰信息. 因此在本文实验中, 取 $S = 110$, 即 4 个平均池化后输出核尺寸为 1、3、6 和 8.

表 2 ANRB 中, 不同采样特征点数的实验结果

Table 2 The experimental results for different sampled feature points in ANRB

特征点数	Set5	Manga109	128 × 128 像素	180 × 180 像素
	PSNR (dB)	PSNR (dB)	内存消耗 (MB)	内存消耗 (MB)
无 ANRB	37.888	38.396	216	419
$S = 50$	37.893	38.439	224	436
$S = 110$	37.895	38.443	232	452
$S = 222$	37.861	38.325	246	480
$S = \infty$	37.883	内存溢出	2266	8431

4) G-IDB. 为了进一步验证本文改进 RFDB 结构组-信息蒸馏模块的有效性, 将 RFDB 中 RFDB 替换成 G-IDB. 2 种网络均在同样的实验设置下重新训练, 实验结果如表 1 所示. 对比表 1 中 RFDB 基本块的第 1 行和 G-IDB 基本块的第 1 行, 去除

RFDB 中的冗余卷积后, 结合了组卷积的 G-IDB 反而以较低参数量超过了 RFDB 性能. 这可能是由于对于标准卷积, 卷积核可以看为一个通道维 + 空间维的三维滤波器, 常规的卷积操作其实就是实现通道相关性和空间相关性的联合映射, 而卷积层通道间的相关性和空间相关性是可以退耦合的, 将它们分开映射能达到更好效果^[55]. 而分组卷积经过不同卷积路径得到的特征图之间的耦合性较低, 关注的主要特征不同, 可得到互为补充的特征图, 从而更完整地表示图像. 最终, G-IDB 结合双路重建策略、DS 连接和 ANRB 在参数量减少约 151 K 情况下, PSNR 值提高了 0.072 dB.

5) 高频损失. 本文在 5 个基准数据集上进行消融实验, 以验证所提高频损失的有效性. 本文给出像素损失 \mathcal{L}_1 和高频损失 \mathcal{L}_{hf} 不同权重组合的实验结果, 放大 2 倍超分的定量结果如表 3 所示. 可以看出, 增加了高频损失 \mathcal{L}_{hf} 后, 在 5 个基准数据集上都获得了 PSNR 提升, 当 $\alpha = 0.4, \beta = 0.6$ 时, 取得最好结果. 为了更清晰地看出它们之间的差异, 将仅用 \mathcal{L}_1 (即 $\alpha = 1.0, \beta = 0$) 时各测试集上的 PSNR 值作为基线, 计算其他 4 种不同权重的组合与它的差值, 差值结果如图 8 所示. 可以看出, \mathcal{L}_{hf} 的权重系数 β 占比不是越高越好, 随着 β 的提高, 整体性能呈先上升、再下降的趋势. 同时, 对于包含大量的结构性纹理的 Urban100 数据集, 增加了高频损失的性能提升非常明显, 这表明了本文的高频损失可以使网络更加关注图像中的纹理细节. 综上所述, 本文选取网络训练的目标函数为 $\mathcal{L} = 0.4\mathcal{L}_{hf} + 0.6\mathcal{L}_1$.

表 3 使用不同损失权重系数的 PSNR 对比结果 (dB)

Table 3 Comparison results of PSNR with different loss weights (dB)

权重系数	Set5	Set14	B100	Urban100	Manga109
$\alpha = 1.0, \beta = 0$	37.907	33.423	32.063	31.830	38.483
$\alpha = 0.8, \beta = 0.2$	37.900	33.406	32.071	31.850	38.476
$\alpha = 0.6, \beta = 0.4$	37.930	33.421	32.075	31.843	38.483
$\alpha = 0.4, \beta = 0.6$	37.975	33.444	32.084	31.878	38.576
$\alpha = 0.2, \beta = 0.8$	37.901	33.467	32.084	31.860	38.462

通过消融实验结果分析, 最终采用双路重建策略、组-信息蒸馏模块、密集快捷连接和非对称 Non-local 残差块去构建网络, 使用 L_1 损失和高频损失的复合损失优化网络, 实验结果见第 3.4 节.

3.4 与先进算法的对比分析

本文在 5 个公开图像超分辨率基准测试数据集

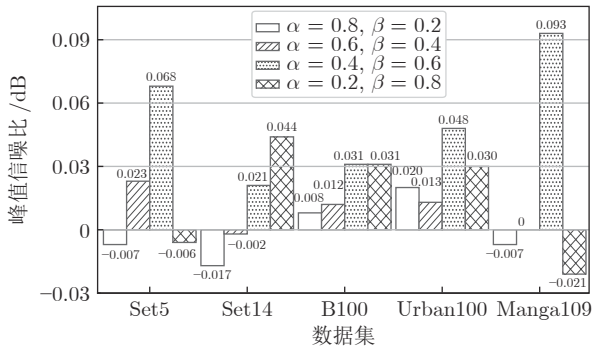


图 8 使用不同损失权重系数的 PSNR 分数差值对比结果

Fig.8 Comparison results of PSNR score differences with different loss weights

上进行实验并与其他 SR 算法进行对比, 对比算法包括 Bicubic、SRCNN、FSRCNN、VDSR、DRCN、拉普拉斯金字塔 SR 网络 (Laplacian pyramid SR network, LapSRN)^[56]、深度递归残差网络 (Deep recursive residual network, DRRN)^[57]、MemNet、IDN、微调的多退化 SR 网络 (Fine-tuning SR network for multiple degradations, SRMDNF)^[58]、CARN、稀疏掩膜网络 (Sparse mask super-resolution, SMSR)^[59]、IMDN、IMDN-联合蒸馏 SR (IMDN-joint distillation super-resolution, IMDN-JD-SR)^[60]、像素注意力网络 (Pixel attention network,

PAN)^[61]、RFDN-L^[12] 和 LatticeNet. 下面展示定量分析和定性分析结果.

1) 定量分析. 表 4 展现了在 5 个基准数据集上, 图像放大 2 倍、3 倍、4 倍时, 各算法的参数量、PSNR 和 SSIM 定量分析结果, 其中性能指标表现最好值加粗表示, 次优值加下划线表示. 为了清晰地看出本文方法的改进效果, 按照文献 [12] 的实验设置训练了 RFDN-L. IMDN (采用官方的预训练模型)、RFDN-L 和本文方法的测试基于同样测试环境. 由表 4 可知, 与其他方法相比, 本文 G-IDRN 在 5 个基准数据集的所有尺度上实现了最好性能. 为了方便看出本文方法的提升效果, 以 IMDN 的 PSNR 和 SSIM 值为基线, RFDN-L 和本文方法 G-IDRN 的 PSNR 和 SSIM 值分别与其相减, 得到的差值图如图 9 所示. 图 9 中, 圆点折线与 $y = 0$ 轴的差距表示以 IMDN 为基线, RFDN-L 在每个数据集上提升的幅度. 方点折线与圆点折线的差距表示以 RFDN-L 为基线, G-IDRN 在每个数据集上提升的幅度. 由图 9 可以看出, RFDN-L 在某些尺度的 B100 和 Urban100 数据集上性能没有提升或反而下降 (见图 9(a) 和图 9(d)), 而本文方法在所有放大倍数上, 较 IMDN 和 RFDN-L 均有提升. 从提升幅度看, 本文方法较 RFDN-L 在 PSNR 指标上获得了更大提升. 在 5 个基准数据集上, 图像放大 2 倍、3 倍和 4

表 4 在 5 个基准数据集上, 图像放大 2 倍、3 倍和 4 倍时, 各算法的参数量、PSNR 和 SSIM 定量分析结果

Table 4 Parameters, PSNR and SSIM quantitative comparisons of various algorithms for 2, 3, and 4 times images on the five benchmark datasets

方法	放大倍数	参数量 (K)	Set5	Set14	B100	Urban100	Manga109
			PSNR (dB) / SSIM	PSNR (dB) / SSIM	PSNR (dB) / SSIM	PSNR (dB) / SSIM	PSNR (dB) / SSIM
Bicubic	-	-	33.66 / 0.9299	30.24 / 0.8688	29.56 / 0.8431	26.88 / 0.8403	30.80 / 0.9339
SRCNN		8	36.66 / 0.9542	32.45 / 0.9067	31.36 / 0.8879	29.50 / 0.8946	35.60 / 0.9663
DRCN		1774	37.63 / 0.9588	33.04 / 0.9118	31.85 / 0.8942	30.75 / 0.9133	37.55 / 0.9732
LapSRN		251	37.52 / 0.9591	32.99 / 0.9124	31.80 / 0.8952	30.41 / 0.9103	37.27 / 0.9740
DRRN		298	37.74 / 0.9591	33.23 / 0.9136	32.05 / 0.8973	31.23 / 0.9188	37.88 / 0.9749
MemNet		678	37.78 / 0.9597	33.28 / 0.9142	32.08 / 0.8978	31.31 / 0.9195	37.72 / 0.9740
IDN		553	37.83 / 0.9600	33.30 / 0.9148	32.08 / 0.8985	31.27 / 0.9196	38.01 / 0.9749
SRMDNF	2 倍	1511	37.79 / 0.9601	33.32 / 0.9159	32.05 / 0.8985	31.33 / 0.9204	38.07 / 0.9761
CARN		1592	37.76 / 0.9590	33.52 / 0.9166	32.09 / 0.8978	31.92 / 0.9256	38.36 / 0.9765
SMSR		985	38.00 / 0.9601	33.64 / 0.9179	32.17 / 0.8993	32.19 / 0.9284	38.76 / 0.9771
IMDN		694	38.00 / 0.9605	33.63 / 0.9177	32.19 / 0.8997	32.17 / 0.9282	38.88 / 0.9774
IMDN-JDSR		694	38.00 / 0.9605	33.57 / 0.9176	32.16 / 0.8995	32.09 / 0.9271	- / -
PAN		261	38.00 / 0.9605	33.59 / 0.9181	32.18 / 0.8997	32.01 / 0.9273	38.70 / 0.9773
RFDN-L		626	<u>38.03</u> / 0.9606	33.65 / 0.9183	32.18 / 0.8997	32.16 / 0.9282	<u>38.88</u> / 0.9772
LatticeNet		759	<u>38.03</u> / <u>0.9607</u>	<u>33.70</u> / <u>0.9187</u>	<u>32.20</u> / <u>0.8999</u>	<u>32.25</u> / <u>0.9288</u>	- / -
G-IDRN		554	38.09 / 0.9608	33.80 / 0.9203	32.42 / 0.9003	32.42 / 0.9311	38.96 / <u>0.9773</u>

表 4 在 5 个基准数据集上, 图像放大 2 倍、3 倍和 4 倍时, 各算法的参数量、PSNR 和 SSIM 定量分析结果 (续表)

Table 4 Parameters, PSNR and SSIM quantitative comparisons of various algorithms for 2, 3, and 4 times images on the five benchmark datasets (continued table)

方法	放大倍数	参数量 (K)	Set5	Set14	B100	Urban100	Manga109
			PSNR (dB) / SSIM	PSNR (dB) / SSIM	PSNR (dB) / SSIM	PSNR (dB) / SSIM	PSNR (dB) / SSIM
Bicubic		–	30.39 / 0.8682	27.55 / 0.7742	27.21 / 0.7385	24.46 / 0.7349	26.95 / 0.8556
SRCNN		8	32.75 / 0.9090	29.30 / 0.8215	28.41 / 0.7863	26.24 / 0.7989	30.48 / 0.9117
DRCN		1774	33.82 / 0.9226	29.76 / 0.8311	28.80 / 0.7963	27.15 / 0.8276	32.24 / 0.9343
LapSRN		502	33.81 / 0.9220	29.79 / 0.8325	28.82 / 0.7980	27.07 / 0.8275	32.21 / 0.9350
DRRN		298	34.03 / 0.9244	29.96 / 0.8349	28.95 / 0.8004	27.53 / 0.8378	32.71 / 0.9379
MemNet		678	34.09 / 0.9248	30.00 / 0.8350	28.96 / 0.8001	27.56 / 0.8376	32.51 / 0.9369
IDN		553	34.11 / 0.9253	29.99 / 0.8354	28.95 / 0.8013	27.42 / 0.8359	32.71 / 0.9381
SRMDNF	3 倍	1528	34.12 / 0.9254	30.04 / 0.8382	28.97 / 0.8025	27.57 / 0.8398	33.00 / 0.9403
CARN		1592	34.29 / 0.9255	30.29 / 0.8407	29.06 / 0.8034	28.06 / 0.8493	33.50 / 0.9440
SMSR		993	34.40 / 0.9270	30.33 / 0.8412	29.10 / 0.8050	28.25 / 0.8536	33.68 / 0.9445
IMDN		703	34.36 / 0.9270	30.32 / 0.8417	29.09 / 0.8047	28.16 / 0.8519	33.61 / 0.9445
IMDN-JDSR		703	34.36 / 0.9269	30.32 / 0.8413	29.08 / 0.8045	28.12 / 0.8498	– / –
PAN		261	<u>34.40</u> / 0.9271	<u>30.36</u> / <u>0.8423</u>	<u>29.11</u> / 0.8050	28.11 / 0.8511	33.61 / 0.9448
RFDN-L		633	34.39 / 0.9271	30.35 / 0.8419	<u>29.11</u> / <u>0.8054</u>	<u>28.24</u> / <u>0.8534</u>	<u>33.74</u> / <u>0.9453</u>
LatticeNet		765	<u>34.40</u> / <u>0.9272</u>	30.32 / 0.8416	29.10 / 0.8049	28.19 / 0.8513	– / –
G-IDRN		565	34.43 / 0.9277	30.41 / 0.8431	29.14 / 0.8061	28.32 / 0.8552	33.79 / 0.9456
Bicubic		–	28.42 / 0.8104	26.00 / 0.7027	25.96 / 0.6675	23.14 / 0.6577	24.89 / 0.7866
SRCNN		8	30.48 / 0.8626	27.50 / 0.7513	26.90 / 0.7101	24.52 / 0.7221	27.58 / 0.8555
DRCN		1774	31.53 / 0.8854	28.02 / 0.7670	27.23 / 0.7233	25.14 / 0.7510	28.93 / 0.8854
LapSRN		502	31.54 / 0.8852	28.09 / 0.7700	27.32 / 0.7275	25.21 / 0.7562	29.09 / 0.8900
DRRN		298	31.68 / 0.8888	28.21 / 0.7720	27.38 / 0.7284	25.44 / 0.7638	29.45 / 0.8946
MemNet		678	31.74 / 0.8893	28.26 / 0.7723	27.40 / 0.7281	25.50 / 0.7630	29.42 / 0.8942
IDN		553	31.82 / 0.8903	28.25 / 0.7730	27.41 / 0.7297	25.41 / 0.7632	29.41 / 0.8942
SRMDNF	4 倍	1552	31.96 / 0.8925	28.35 / 0.7787	27.49 / 0.7337	25.68 / 0.7731	30.09 / 0.9024
CARN		1592	32.13 / 0.8937	28.60 / 0.7806	27.58 / 0.7349	26.07 / 0.7837	30.47 / 0.9084
SMSR		1006	32.13 / 0.8937	28.60 / 0.7806	27.58 / 0.7349	26.11 / 0.7868	30.54 / 0.9084
IMDN		715	32.21 / 0.8948	28.58 / 0.7811	27.56 / 0.7354	26.04 / 0.7838	30.45 / 0.9075
IMDN-JDSR		715	32.17 / 0.8942	28.62 / 0.7814	27.55 / 0.7350	26.06 / 0.7820	– / –
PAN		272	32.13 / 0.8948	28.61 / <u>0.7822</u>	<u>27.59</u> / <u>0.7363</u>	26.11 / 0.7854	30.51 / <u>0.9095</u>
RFDN-L		643	<u>32.23</u> / <u>0.8953</u>	28.59 / 0.7814	27.57 / <u>0.7363</u>	<u>26.14</u> / <u>0.7871</u>	<u>30.61</u> / <u>0.9095</u>
LatticeNet		777	32.18 / 0.8943	28.61 / 0.7812	27.57 / 0.7355	<u>26.14</u> / 0.7844	– / –
G-IDRN		580	32.24 / 0.8958	28.64 / 0.7824	27.61 / 0.7378	26.24 / 0.7903	30.63 / 0.9096

倍时, 本文方法的 PSNR 平均提升了 15.50 倍、1.93 倍、1.90 倍; 本文方法的 SSIM 平均提升了 15.75 倍、2.50 倍、1.90 倍. 特别地, 在包含更丰富结构纹理信息的 Urban100 数据集上, 幅度差距明显变大, 表明 G-IDRN 性能明显提升, 说明本文方法能够成功重建出丰富的纹理细节.

2) 定性分析. 各方法在 Urban100 上 4 倍 SR

的定性比较如图 10 所示. 可以看出, 对于具有挑战性细节恢复的图像 Img005 和 Img020, 多数方法都存在边缘模糊和明显的伪影. IMDN 和 RFDN-L 虽然可以缓解模糊边缘、恢复更多的细节, 但产生了不同程度的虚假信息. 相比之下, 本文 G-IDRN 通过恢复更清晰、更精确的边缘, 获得了更忠实于真实的高分辨率图像. 特别是对于图像 Img020, 所有

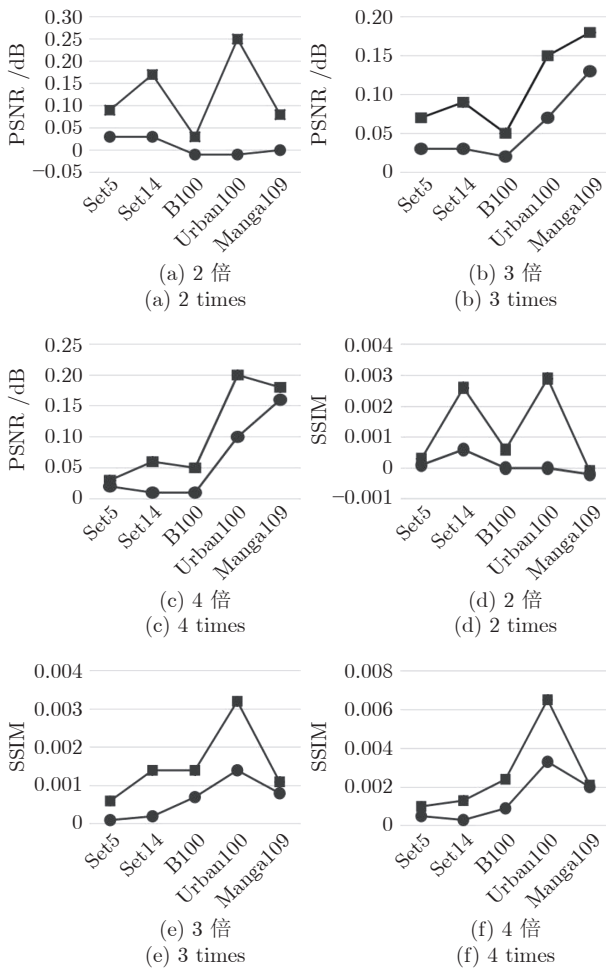


图9 PSNR和SSIM的差值图

Fig. 9 Differential results of PSNR and SIMM scores

对比方法重构的边缘纹理方向都是错误的,而本文 G-IDRN 可以充分利用学习到的特征,获得更清晰的轮廓而不会产生严重的伪影.由图 10 可以看出,与其他方法相比,本文 G-IDRN 可以更好地恢复 HR 图像中更多的信息成分,显示出令人满意的图像 SR 结果.进一步在真实场景下测试 G-IDRN 性能.本文选取真实图像数据集 RealSR^[62],采用 Bicubic、RFDN-L 和 G-IDRN 方法,图像放大 4 倍的可视化对比结果如图 11 所示.可以看出,本文 G-IDRN 实现了更好的感知质量,生成了更为清晰的线条纹理.

3.5 模型分析

1) 模型参数量.对于轻量级超分模型,网络参数是需要考虑的关键因素.表 4 展示了本文方法与其他方法在图像放大 2、3、4 倍时,参数量和图像 SR 性能的对比.为了更全面地了解模型的复杂性,图 1 可视化在 Urban100 数据集上,模型参数量和峰值信噪比的对比结果.由于 SSIM 能够更好地反

映重建图像的视觉效果,因此本节给出了在 Urban100 数据集上,图像放大 4 倍时,SSIM 和参数量的对比结果如图 12 所示.图 1 和图 12 可以看出,与其他先进的轻量级模型相比,本文 G-IDRN 在图像 SR 性能和模型复杂性之间实现了更好的平衡.

2) 模型计算量.仅通过模型参数来衡量模型的轻量化是不够的,计算成本也是一个重要指标.使用 256×256 像素的单张图像,采用 FLOPs 衡量 G-IDRN 与其他先进算法(如 CARN、IMDN 和 RFDN-L)的计算量,同时在 Set14 数据集上计算 PSNR 和 SSIM,实验结果如表 5 所示.可以看出,相较于其他方法,本文 G-IDRN 实现了更高的 PSNR 和 SSIM,且只用了更少的计算量.表明本文方法很好地平衡了计算成本和图像超分性能.

3) 运行时间.模型的运行时间对于轻量级 SR 算法也极为重要.因此,在相同环境下,将本文 G-IDRN 与最优的非轻量级算法 EDSR、RCAN,以及轻量级算法 IMDN、RFDN-L 测试每 100 次迭代的平均训练时间(统一设置图像块尺寸为 64×64 像素,最小训练批次为 16),并在 B100 数据集上,图像放大 4 倍时,测试它们的平均运行时间,实验结果如表 6 所示.可以看出,对比 EDSR 和 RCAN,本文方法在性能下降不多的情况下,推理时间分别获得 12 倍和 15 倍的提升;对比 IMDN 和 RFDN-L,受到分组卷积的影响,本文方法较为耗时.虽然在不使用分组卷积情况下,有助于提升性能和减少训练时间,但 IDRN 的参数量急剧增加到 2 MB 以上且伴随着很大计算量和参数量,其实际推理时间较慢.本文 G-IDRN 在网络前馈过程中,由于分组卷积包含的参数量和计算量比普通卷积少,能够达到较快速度,因此本文方法能够获得最小的推理时间,这对于实际应用非常重要.

4 结束语

本文提出一种新型的轻量级组-信息蒸馏残差网络,用于快速、准确地提取图像的深度特征.在 RFDB 基础上,在保证其性能的前提下,提出更轻量的 G-IDB,从低分辨率图像中提取特征.基于 DS 连接, G-IDRG 可以有效地重利用浅层特征,有效捕获多尺度信息.此外,引入一个轻量级的非对称残差 Non-local 模块捕获全局上下文信息.特别地,为缓解像素损失造成的平滑图像细节问题,设计一个简单但有效的高频损失,来帮助优化所提模型.实验结果表明,与其他先进图像 SR 方法相比, G-IDRN 在图像 SR 性能和模型复杂度之间,取得了更好的平衡.

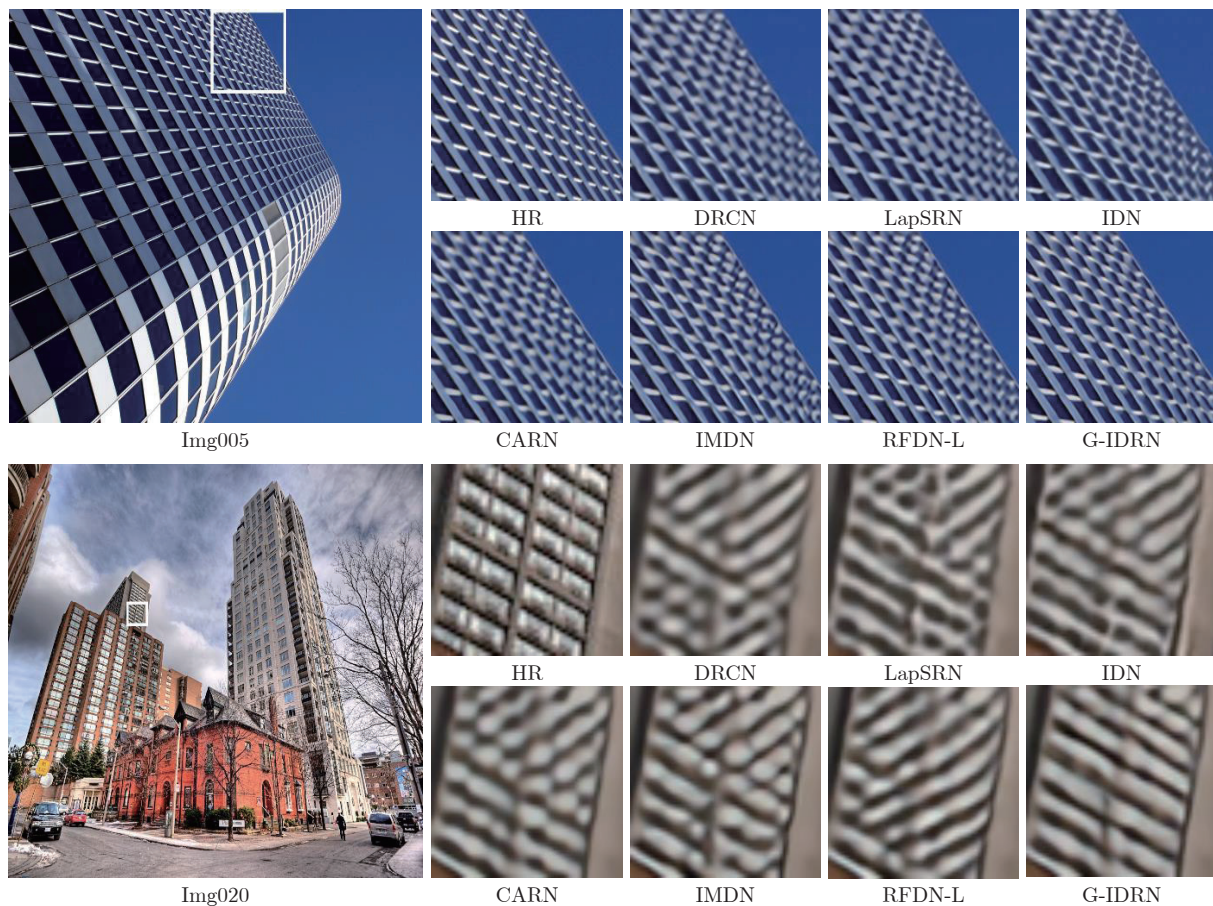


图 10 各方法在 Urban100 上 4 倍 SR 的定性比较
Fig.10 Qualitative comparisons of each method for 4 times SRs on Urban100

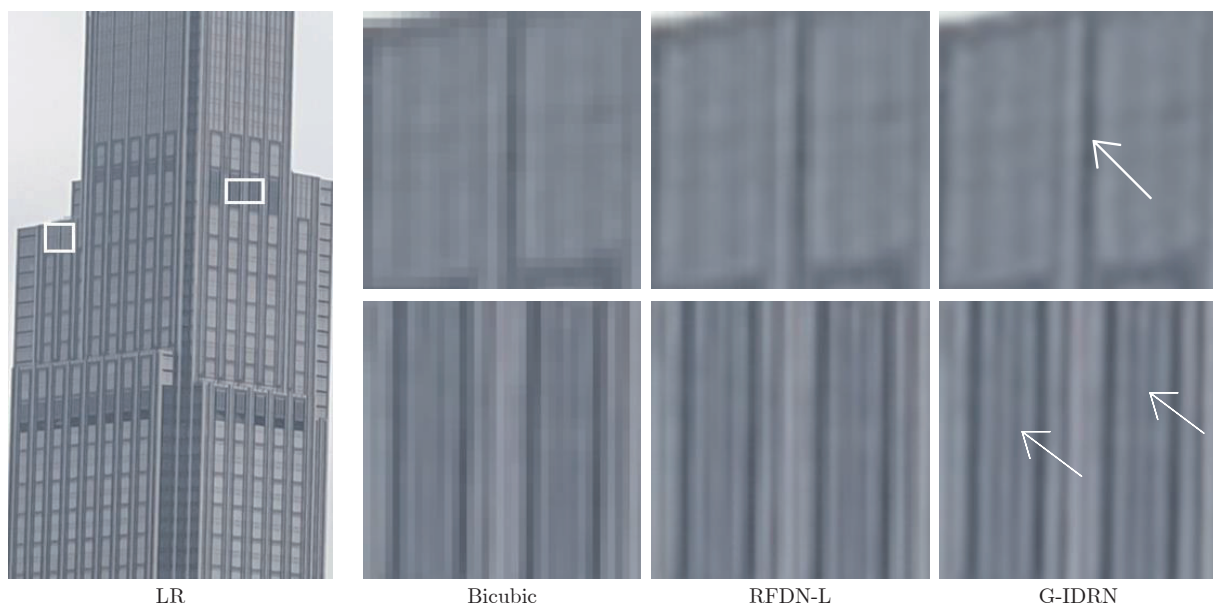


图 11 在真实图像上的可视化对比结果
Fig.11 Visual comparison on a real-world image

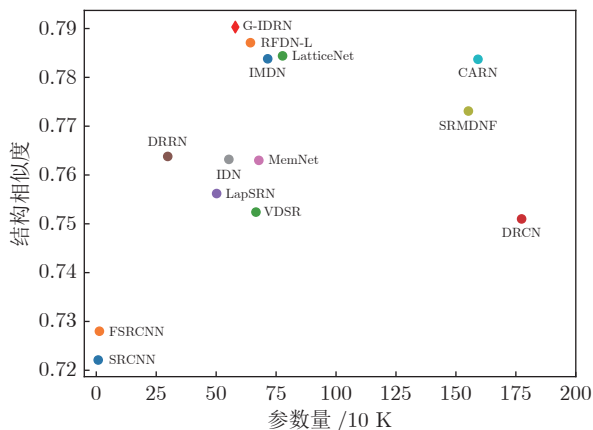


图 12 Urban100 中图像放大 4 倍时, 参数量和结构相似度的对比结果

Fig. 12 Comparison results of the number of parameters and the structural similarity for 4 times images on Urban100

表 5 Set14 中图像放大 4 倍时, SSIM、PSNR 和 FLOPs 的比较结果

Table 5 Comparison results of SSIM, PSNR and FLOPs for 4 times images on Set14

评价指标	CARN	IMDN	RFDN-L	G-IDRN
SSIM	0.7806	0.7810	0.7814	0.7826
PSNR (dB)	28.60	28.58	28.59	28.64
FLOPs (GB)	103.58	46.60	41.54	36.19

表 6 B100 中图像放大 4 倍时, 平均运行时间的比较结果

Table 6 Comparison results of average running time for 4 times images on B100

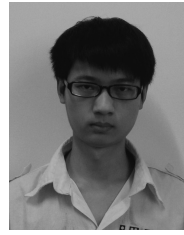
方法	PSNR (dB) / SSIM	参数量 (K)	训练时间 (s)	推理时间 (s)
EDSR	27.71 / 0.7420	43090	—	0.2178
RCAN	27.77 / 0.7436	15592	—	0.2596
IMDN	27.56 / 0.7354	715	5.4	0.0217
RFDN-L	27.57 / 0.7363	633	6.1	0.0250
G-IDRN	27.61 / 0.7378	580	12.7	0.0177
IDRN	27.64 / 0.7389	2047	8.5	0.0692

References

- Isaac J S, Kulkarni R. Super resolution techniques for medical image processing. In: Proceedings of the International Conference on Technologies for Sustainable Development. Mumbai, India: IEEE, 2015. 1–6
- Rasti P, Uiboupin T, Escalera S, Anbarjafari G. Convolutional neural network super resolution for face recognition in surveillance monitoring. In: Proceedings of the International Conference on Articulated Motion and Deformable Objects. Cham, Netherlands: Springer, 2016. 175–184
- Sajjadi M S M, Scholkopf B, Hirsch M. Enhancenet: Single image super-resolution through automated texture synthesis. In: Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 4491–4500
- Tan Y, Cai J, Zhang S, Zhong W, Ye L. Image compression algorithms based on super-resolution reconstruction technology. In: Proceedings of the IEEE 4th International Conference on Image, Vision and Computing. Xiamen, China: IEEE, 2019. 162–166
- Luo Y, Zhou L, Wang S, Wang Z. Video satellite imagery super resolution via convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 2017, **14**(12): 2398–2402
- Yang Xin, Zhou Da-Ke, Fei Shu-Min. A self-adapting bilateral total aaration technology for image super-resolution reconstruction. *Journal of Computer Research and Development*, 2012, **49**(12): Article No. 2696 (杨欣, 周大可, 费树岷. 基于自适应双边全变差的图像超分辨率重建. *计算机研究与发展*, 2012, **49**(12): Article No. 2696)
- Zhang L, Wu X. An edge-guided image interpolation algorithm via directional filtering and data fusion. *IEEE Transactions on Image Processing*, 2006, **15**(8): 2226–2238
- Pan Zong-Xu, Yu Jing, Hu Shao-Xing, Sun Wei-Dong. Single image super resolution based on multi-scale structural self-similarity. *Acta Automatica Sinica*, 2014, **40**(4): 594–603 (潘宗序, 禹晶, 胡少兴, 孙卫东. 基于多尺度结构自相似性的单幅图像超分辨率算法. *自动化学报*, 2014, **40**(4): 594–603)
- Zhang Yi-Feng, Liu Yuan, Jiang Cheng, Cheng Xu. A curriculum learning approach for single image super-resolution. *Acta Automatica Sinica*, 2020, **40**(2): 274–282 (张毅锋, 刘袁, 蒋程, 程旭. 用于超分辨率重建的深度网络递进学习方法. *自动化学报*, 2020, **40**(2): 274–282)
- Dai T, Cai J, Zhang Y, Xia S T, Zhang L. Second-order attention network for single image super-resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE, 2019. 11065–11074
- Hui Z, Gao X, Yang Y, Wang X. Lightweight image super-resolution with information multi-distillation network. In: Proceedings of the 27th ACM International Conference on Multimedia. New York, USA: Association for Computing Machinery, 2019. 2024–2032
- Liu J, Tang J, Wu G. Residual feature distillation network for lightweight image super-resolution. In: Proceedings of the 20th European Conference on Computer Vision. Cham, Netherlands: Springer, 2020. 41–55
- Sun Chao-Wen, Chen Xiao. Multi-scale feature fusion back-projection network for image super-resolution. *Acta Automatica Sinica*, 2021, **47**(7): 1689–1700 (孙超文, 陈晓. 基于多尺度特征融合反投影网络的图像超分辨率重建. *自动化学报*, 2021, **47**(7): 1689–1700)
- Sun Yu-Bao, Fei Xuan, Wei Zhi-Hui, Xiao Liang. Sparsity regularized image super-resolution model via forward-backward operator splitting method. *Acta Automatica Sinica*, 2010, **36**(9): 1232–1238 (孙玉宝, 费选, 韦志辉, 肖亮. 基于前向后向算子分裂的稀疏性正则化图像超分辨率算法. *自动化学报*, 2010, **36**(9): 1232–1238)
- Dong C, Loy C C, He K, Tang X. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **38**(2): 295–307
- Dong C, Loy C C, Tang X. Accelerating the super-resolution convolutional neural network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 391–407
- Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 1646–1654
- Kim J, Lee J K, Lee K M. Deeply-recursive convolutional network for image super-resolution. In: Proceedings of the IEEE

- Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 1637–1645
- 19 Lim B, Son S, Kim H, Nah S, Mu Lee K. Enhanced deep residual networks for single image super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, USA: IEEE, 2017. 136–144
- 20 Zhang Y, Li K, Li K, Wang L, Zhong B, Fu Y. Image super-resolution using very deep residual channel attention networks. In: Proceedings of the 18th European Conference on Computer Vision. Mohini, Germany: Springer, 2018. 286–301
- 21 Ahn N, Kang B, Sohn K A. Fast, accurate, and lightweight super-resolution with cascading residual network. In: Proceedings of the 18th European Conference on Computer Vision. Mohini, Germany: Springer, 2018. 252–268
- 22 Hui Z, Wang X, Gao X. Fast and accurate single image super-resolution via information distillation network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 723–731
- 23 Zhang C, Benz P, Argaw D M, Lee S, Kim J, Rameau F, et al. Resnet or densenet? Introducing dense shortcuts to resnet. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. Waikoloa, USA: IEEE, 2021. 3550–3559
- 24 Zhu Z, Xu M, Bai S, Huang T, Bai X. Asymmetric non-local neural networks for semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE, 2019. 593–602
- 25 Huang J B, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015. 5197–5206
- 26 An Yao-Zu, Lu Yao, Zhao Hong. An adaptive-regularized image super-resolution. *Acta Automatica Sinica*, 2012, **38**(4): 601–608 (安耀祖, 陆耀, 赵红. 一种自适应正则化的图像超分辨率算法. *自动化学报*, 2012, **38**(4): 601–608)
- 27 Tai Y, Yang J, Liu X, Xu C. MemNet: A persistent memory network for image restoration. In: Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 4539–4547
- 28 Li Z, Yang J, Liu Z, Jeon G, Wu W. Feedback network for image super-resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE, 2019. 3867–3876
- 29 Qiu Y, Wang R, Tao D, Cheng J. Embedded block residual network: A recursive restoration model for single-image super-resolution. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, South Korea: IEEE, 2019. 4180–4189
- 30 Chu X, Zhang B, Ma H, Xu R, Li Q. Fast, accurate and lightweight super-resolution with neural architecture search. In: Proceedings of the 25th International Conference on Pattern Recognition. Milan, Italy: IEEE, 2021. 59–64
- 31 Chu X, Zhang B, Xu R. Multi-objective reinforced evolution in mobile neural architecture search. In: Proceedings of the 20th European Conference on Computer Vision. Glasgow, UK: Springer, 2020. 99–113
- 32 Luo X, Xie Y, Zhang Y, Qu Y, Li C, Fu Y. LatticeNet: Towards lightweight image super-resolution with lattice block. In: Proceedings of the 20th European Conference on Computer Vision. Glasgow, UK: Springer, 2020. 23–28
- 33 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City, USA: IEEE, 2018. 7132–7141
- 34 Wang X, Girshick R, Gupta A, He K. Non-local neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City, USA: IEEE, 2018. 7794–7803
- 35 Liu D, Wen B, Fan Y, Loy C C, Huang T S. Non-local recurrent network for image restoration. In: Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal, Canada: MIT Press, 2018. 1680–1689
- 36 Mei Y, Fan Y, Zhou Y, Huang L, Huang T S, Shi H. Image super-resolution with cross-scale non-local attention and exhaustive self-exemplars mining. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE, 2020. 5690–5699
- 37 Niu B, Wen W, Ren W, Zhang X, Yang L, Wang S, et al. Single image super-resolution via a holistic attention network. In: Proceedings of the 20th European Conference on Computer Vision. Glasgow, UK: Springer, 2020. 191–207
- 38 Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution. In: Proceedings of the 14th European Conference on Computer Vision. Amsterdam, Netherlands: Springer, 2016. 694–711
- 39 Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, et al. Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 4681–4690
- 40 Yuan Y, Liu S, Zhang J, Zhang Y, Dong C, Lin L. Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City, USA: IEEE, 2018. 701–710
- 41 Yu J, Fan Y, Huang T. Wide activation for efficient image and video super-resolution. In: Proceedings of the 30th British Machine Vision Conference. Cardiff, UK: BMVA Press, 2020. 1–13
- 42 Shi W, Caballero J, Huszár F, Totz J, Aitken A P, Bishop R, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 1874–1883
- 43 Howard A G, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications [Online], available: <https://arxiv.org/abs/1704.04861>, April 17, 2017
- 44 Xie S, Girshick R, Dollár P, Tu Z, He K. Aggregated residual transformations for deep neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 1492–1500
- 45 Szegedy C, Ioffe S, Vanhoucke V, Alemi A A. Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, USA: AAAI Press, 2017. 4278–4284
- 46 Huang G, Liu Z, Van Der Maaten L, Weinberger K Q. Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 4700–4708
- 47 Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 2881–2890
- 48 Timofte R, Agustsson E, Van Gool L, Yang M H, Zhang L. Ntire 2017 challenge on single image super-resolution: Methods and results. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, USA: IEEE, 2017. 114–125
- 49 Bevilacqua M, Roumy A, Guillemot C, Morel M L A. Lowcom-

- plexity single-image super-resolution based on nonnegative neighbor embedding. In: Proceedings of the British Machine Vision Conference. Surrey, UK: BMVA Press, 2012. 1–10
- 50 Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations. In: Proceedings of the International Conference on Curves and Surfaces. Berlin, Germany: Springer, 2010. 711–730
- 51 Arbelaez P, Maire M, Fowlkes C, Malik J. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **33**(5): 898–916
- 52 Matsui Y, Ito K, Aramaki Y, Fujimoto A, Ogawa T, Yamasaki T, et al. Sketch-based manga retrieval using Manga109 dataset. *Multimedia Tools and Applications*, 2017, **76**(20): 21811–21838
- 53 Gao X, Lu W, Tao D, Li X. Image quality assessment based on multi-scale geometric analysis. *IEEE Transactions on Image Processing*, 2009, **18**(7): 1409–1423
- 54 Wang Z, Bovik A C, Sheikh H R, Simoncelli E P. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, **13**(4): 600–612
- 55 Chollet F. Xception: Deep learning with depth-wise separable convolutions. In: Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 1251–1258
- 56 Lai W S, Huang J B, Ahuja N, Yang M H. Deep Laplacian pyramid networks for fast and accurate super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 624–632
- 57 Tai Y, Yang J, Liu X. Image super-resolution via deep recursive residual network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 3147–3155
- 58 Zhang K, Zuo W, Zhang L. Learning a single convolutional super-resolution network for multiple degradations. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City, USA: IEEE, 2018. 3262–3271
- 59 Wang L, Dong X, Wang Y, Ying X, Lin Z, An W, et al. Exploring sparsity in image super-resolution for efficient inference. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE, 2021. 4917–4926
- 60 Luo X, Liang Q, Liu D, Qu Y. Boosting lightweight single image super-resolution via joint-distillation. In: Proceedings of the 29th ACM International Conference on Multimedia. Virtual Event: Association for Computing Machinery, 2021. 1535–1543
- 61 Zhao H, Kong X, He J, Qiao Y, Dong C. Efficient image super-resolution using pixel attention. In: Proceedings of the European Conference on Computer Vision. Cham, Netherlands: Springer, 2020. 56–72
- 62 Cai J, Zeng H, Yong H, Cao Z, Zhang L. Toward real-world single image super-resolution: A new benchmark and a new model. In: Proceedings of IEEE/CVF International Conference on Computer Vision. Seoul, South Korea: IEEE, 2019. 3086–3095



王云涛 中南民族大学生物医学工程学院硕士研究生. 主要研究方向为图像处理, 深度学习和图像超分辨率.

E-mail: ytao-wang@scuec.edu.cn

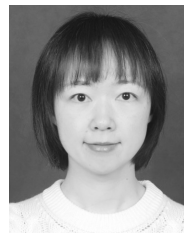
(WANG Yun-Tao Master student at the School of Biomedical Engineering, South-central Minzu University. His research interest covers image processing, deep learning, and image super-resolution.)



赵 藺 华中科技大学人工智能与自动化学院博士研究生. 主要研究方向为图像识别, 图像超分辨率和点云实例语义分割.

E-mail: linzhao@hust.edu.cn

(ZHAO Lin Ph.D. candidate at the School of Artificial Intelligence and Automation, Huazhong University of Science and Technology. His research interest covers image recognition, image super-resolution, and point cloud instance semantic segmentation.)



刘李漫 中南民族大学生物医学工程学院副教授. 主要研究方向为图像处理, 深度学习和计算机视觉. 本文通信作者.

E-mail: limanliu@mail.scuec.edu.cn

(LIU Li-Man Associate professor at the School of Biomedical Engineering, South-central Minzu University. Her research interest covers image processing, deep learning, and computer vision. Corresponding author of this paper.)



陶文兵 华中科技大学人工智能与自动化学院教授. 主要研究方向为图像分割, 目标识别和 3D 重建.

E-mail: wenbingtao@hust.edu.cn

(TAO Wen-Bing Professor at the School of Artificial Intelligence and Automation, Huazhong University of Science and Technology. His research interest covers image segmentation, target recognition, and 3D reconstruction.)