

基于特征复用的轻量级图像超分辨率重建

林子杰, 郭 恒, 葛芳艳

(成都信息工程大学 计算机学院, 四川 成都 610225)

摘要: 增加图像超分辨率重建网络模型的深度与参数量被认为是获得优秀图像重建性能的关键, 但极大增加了其复杂性和计算成本。为此, 研究一种轻量级图像超分辨率重建算法, 通过多级跳跃连接实现特征复用, 有效提升了特征信息的利用率和传播速度。首先, 为了提升卷积的特征提取效率, 设计了一种结合深度可分离卷积的快速卷积层, 极大减少了特征提取所需的卷积核数量; 其次, 引入解耦合全连接注意力模块, 以极少的参数量增强了网络模型的全局信息感知能力; 最后, 通过对比感知的特征融合机制以更好地保留图像高频信息, 使重建图像具有更好的视觉效果。实验表明, 在仅有 636 KB 参数的情况下, 所提模型优于目前大多数轻量级图像超分辨率模型, 证明了研究方法的有效性, 为深度神经网络模型的轻量化提供了新的思路。

关键词: 图像超分辨率重建; 轻量级; 特征复用; 解耦合全连接注意力; 对比感知的特征融合

DOI: 10.11907/rjdk.241276

开放科学(资源服务)标识码(OSID):

中图分类号: TP391

文献标识码: A

文章编号: 1672-7800(2025)005-0214-07



Lightweight Image Super-Resolution Reconstruction Based on Feature Reuse Method

LIN Zijie, GUO Heng, GE Fangyan

(School of Computer Science, Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: Increasing the depth and parameter count of the image super-resolution reconstruction network model is considered the key to achieving excellent image reconstruction performance, but it greatly increases its complexity and computational cost. Therefore, a lightweight image super-resolution reconstruction algorithm was studied, which achieves feature reuse through multi-level skip connections, effectively improving the utilization and propagation speed of feature information. Firstly, in order to improve the feature extraction efficiency of convolution, a fast convolutional layer combining depthwise separable convolution was designed, greatly reducing the number of convolution kernels required for feature extraction; Secondly, the introduction of a decoupled fully connected attention module enhances the global information perception capability of the network model with minimal parameter count; Finally, the fusion mechanism of perceived features is compared to better preserve high-frequency information in the image, resulting in a reconstructed image with better visual effects. Experiments have shown that with only 636 KB parameters, the proposed model outperforms most current lightweight image super-resolution models, demonstrating the effectiveness of the research method and providing new ideas for the lightweighting of deep neural network models.

Key Words: image super-resolution reconstruction; lightweight; feature reuse; decoupled full-connected attention; contrast aware feature fusion

0 引言

图像超分辨率重建(Super-resolution Reconstruction,

SR)技术是计算机视觉和数字图像处理领域的热门研究方向,旨在通过增强图像的纹理和边缘细节,将低分辨率(Low-Resolution, LR)图像转换至更高分辨率(High-Resolution, HR)版本,从而提升图像的视觉效果。同时,图像

收稿日期: 2024-03-20

扫描二维码阅读全文:

基金项目: 四川省科技计划项目(2023YFQ0072)

作者简介: 林子杰(1998-),男,成都信息工程大学计算机学院硕士研究生,研究方向为计算机视觉;郭恒(1999-),男,成都信息工程大学计算机学院硕士研究生,研究方向为计算机视觉;葛芳艳(1998-),女,成都信息工程大学计算机学院硕士研究生,研究方向为计算机视觉。



超分辨率重建技术能显著提高语义分割、目标检测、图像识别等后续处理的精度,在医学成像、遥感测绘、文物保护、工程控制等社会公共领域亦展现出广泛的应用价值^[1-7]。

传统图像超分辨率重建算法受限于有限的特征表示、固定的模型结构及对噪声的敏感性,难以实现高质量的图像重建效果^[8]。然而,随着深度学习技术的快速发展,这一问题得到了有效改善。深度神经网络凭借其出色的特征学习和表征能力,能更有效学习低分辨率图像到高分辨率图像之间的复杂映射关系,从而提高了重建图像的质量。

尽管基于深度学习的方法在图像超分辨率重建领域取得了革命性的突破,但也同样面临着不少挑战^[9]。为了实现优异的图像重建效果,深度神经网络往往需要依赖大规模的数据集进行模型训练。此外,更复杂的网络结构意味着更高的计算消耗,网络模型训练和推理时间成本也会随之上升,因此如何在有限的资源条件下提升图像超分辨率重建的计算效率,保证重建图像质量便成为了目前亟待解决的问题。

为此,本文提出一种高效特征复用与对比感知网络(Efficient Feature Reuse and Contrast Aware Network, EFR-CAN)。首先,通过跳跃连接实现特征复用,不仅能有效减少特征信息的损失,还能加快特征在网络中的传递速度;其次,为优化卷积特征提取效率,设计了一种结合普通卷积与深度可分离卷积的快速卷积层,并引入解耦合全连接(Decoupled Full-Connected, DFC)注意力模块增强网络模型对图像结构等特征的感知能力;最后,利用对比感知的特征融合策略加权图像高频信息,以显著提升重建图像的视觉质量。

1 相关工作

1.1 基于深度学习的图像超分辨率重建

Dong等^[10]提出基于深度学习的图像超分辨率卷积神经网络(Super-Resolution Convolutional Neural Network, SRCNN),利用双三次插值对低分辨率图像进行预处理,通过卷积进行特征提取,实现了图像从低分辨率到高分辨率的重建。虽然该方法相较于传统算法SRCNN具有显著优势,但在泛化能力方面仍存在一定的局限性。为此, Kim等^[11]提出了一种非常深的超分辨率卷积神经网络(Very Deep Super-Resolution, VDSR),通过增加网络深度学习更复杂和抽象的特征表示,进而提升模型泛化性能。同时, VDSR还引入了残差学习及跳跃连接,有效缓解了训练过程中的梯度消失问题,不仅提升了训练稳定性,还加速了网络收敛过程。

尽管SRCNN和VDSR证明了基于深度学习的图像超分辨率重建算法的优势,但都存在一个共同问题,即首先

对低分辨率图像进行上采样,其次在高分辨率上进行特征提取的策略会导致网络模型计算复杂度增加。为了降低网络计算量, Dong等^[12]提出一种快速的图像超分辨率卷积神经网络(Fast Super-Resolution Convolutional Neural Network, FSRCNN),创新之处在于将可学习的转置卷积层置于网络结构的末端,使特征提取过程在较低的分辨率上进行,从而显著降低计算消耗。与此同时,为了提高上采样的计算效率并优化图像重建效果, Shi等^[13]设计了一种高效的亚像素卷积算法。该算法通过在通道维度上进行特征重排,将相邻像素的值组合成更高分辨率的像素,不仅能显著提升计算效率,还能有效减少生成图像中的伪影效应。

在网络模型轻量化方面, Ahn等^[14]提出级联残差网络(Cascading Residual Network, CARN),通过级联架构设计加速特征信息在网络中的传递速度,使网络训练和推理更高效。Kim等^[15]则提出深度递归卷积神经网络(Deeply-Recursive Convolutional Network, DRCN),通过递归多次应用相同的卷积层来增加网络深度,在提高性能的同时不引入新参数也能实现网络模型轻量化。与此同时,为了解决卷积神经网络缺乏全局信息感知能力的问题, Zhang等^[16]提出高效的远程注意力网络(Efficient Long-range Attention Network, ELAN),引入一种高效的组间多尺度自注意力模块,在不重叠的特征组间使用不同的窗口大小计算注意力。Hui等^[17]提出信息多重蒸馏网络(Information Multi-Distillation Network, IMDN),通过多重特征分割实现信息蒸馏,并使用注意力机制融合特征,在保证网络轻量化的同时提升图像超分辨率重建的质量和效率。

上述网络结构和模块设计为轻量级图像超分辨率重建的应用与部署提供了切实可行的方案,推动了图像超分辨率重建技术的进步与发展。

1.2 注意力机制

在注意力机制中, Dosovitskiy等^[18]提出针对图像识别任务的视觉Transformer模型,将图像分割为16×16大小的切片序列,通过动态计算输入序列不同部分之间的权重以灵活捕捉长距离依赖关系。需要注意的是,自注意力机制突破了传统方法(固定窗口大小或卷积核大小的限制),不仅在自然语言处理领域有着强大的性能表现,还推动了计算机视觉领域的发展^[19]。同样,自注意力机制在图像超分辨率重建领域也有着广泛的应用, Lu等^[19]实现了卷积神经网络与Transformer结合的混合神经网络。Liang等^[20]构建了纯Transformer的图像超分辨率重建网络模型Swin-IR。

尽管这些基于自注意力机制的网络模型在图像超分辨率重建任务中展现了强大的图像重建能力,但其计算复杂度高和存储需求大的问题仍不容忽视,计算内积、归一化操作、大量的注意力权重存储和处理等操作,使得其部署在资源受限的设备上十分困难^[21]。虽然,已有MobileV-

iT^[22]、Swin Transformer^[23]等算法尝试降低计算消耗,但问题仍未得到很好的解决。为此,Tang等^[24]提出了一种解耦合全连接注意力算法,在横向和纵向分别使用全连接特征提取来捕获图像的远程依赖关系,以处理长距离依赖的同时有效降低网络模型整体的计算消耗,为其在轻量级图像超分辨率重建中的应用提供可能。

2 网络设计

本文针对单图像超分辨率重建任务,提出一种高效特

征复用与对比感知增强网络,主要包含浅层特征提取模块(Shallow Feature Extraction Module, SFEM)、深层特征提取模块(Deep Feature Extraction Module, DFEM)及图像重建模块(Image Reconstruction Module, IRM),整体结构如图1所示。在浅层特征提取阶段,初步提取输入的低分辨率图像特征;在深度特征提取阶段,进一步学习更复杂、抽象的特征表达;在图像重建阶段,采用对比感知的特征融合算法,将浅层与深层进行融合后进行上采样;最后阶段,将上采样结果与双三次插值的低分辨率图像相加得到重建的高分辨率重建图像。

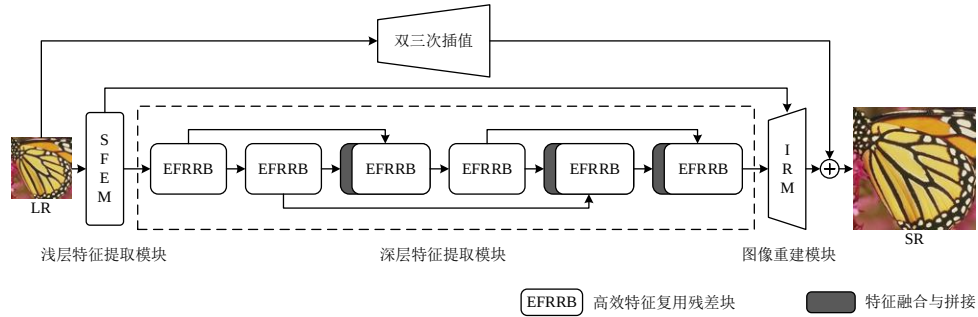


Fig. 1 Network structure

图1 网络结构

2.1 浅层特征提取模块

浅层特征提取模块仅包含一层核大小为 3×3 的卷积层,目的是将输入的低分辨率图像映射到一个更高维的特征空间中,将图像内容拓展为更丰富的特征表示。定义输入的低分辨率图像为 $I_{LR} \in \mathbb{R}^{H \times W \times C_{in}}$,其中 H 、 W 、 C_{in} 分别为图像的高、宽和通道数,输出的浅层特征为 $F_{shallow} \in \mathbb{R}^{H \times W \times C}$, C 为浅层特征通道数,具体计算过程为:

$$F_{shallow} = H_{conv3 \times 3}(I_{LR}) \quad (1)$$

式中: $H_{conv3 \times 3}()$ 表示 3×3 的普通卷积。

2.2 深层特征提取模块

深层特征提取模块包含6个高效特征复用残差块(Efficient Feature Reuse Residual Block, EFRRB),通过多级跳跃连接和特征拼接与融合模块实现特征复用。具体为,假设第 i 个模块输出的中间特征为 $F_i \in \mathbb{R}^{H \times W \times C}$,最后一个高效特征复用残差块的输出为 $F_{deep} \in \mathbb{R}^{H \times W \times C}$,计算过程为:

$$F_1 = H_{EFRRB_1}(F_{shallow}) \quad (2)$$

$$F_2 = H_{EFRRB_2}(F_1) \quad (3)$$

$$F_3 = H_{EFRRB_3}(H_{fusion}(F_1, F_2)) \quad (4)$$

$$F_4 = H_{EFRRB_4}(F_3) \quad (5)$$

$$F_5 = H_{EFRRB_5}(H_{fusion}(F_2, F_4)) \quad (6)$$

$$F_{deep} = H_{EFRRB_6}(H_{fusion}(F_4, F_5)) \quad (7)$$

式中: $H_{EFRRB_i}()$ 为第 i 个高效特征复用残差块的特征提取过程; $H_{fusion}()$ 表示特征拼接与融合模块的计算流程。

2.2.1 特征拼接与融合模块

首先,通过特征拼接与融合模块将前序模块传入的两个特征在通道维度上进行拼接;其次,利用卷积对拼接特

征进行通道压缩与特征融合。具体为,假设前序模块传入的两个特征为 $Feat_1 \in \mathbb{R}^{H \times W \times C}$ 、 $Feat_2 \in \mathbb{R}^{H \times W \times C}$,拼接融合的特征为 $Feat_{out} \in \mathbb{R}^{H \times W \times C}$,具体计算过程为:

$$Feat_{out} = H_{conv1 \times 1}(H_{concat}(Feat_1, Feat_2)) \quad (8)$$

式中: $H_{concat}()$ 表示通道维度的特征拼接; $H_{conv1 \times 1}()$ 表示 1×1 的普通卷积。

2.2.2 高效特征复用残差块

图2为高效特征复用残差块的结构,左侧为特征提取模块,右侧为解耦合全连接注意力模块,整体采用了残差学习的结构提升网络稳定性。其中,特征提取模块使用快速卷积(Fast Convolution, FastConv)替代普通卷积,从而提升特征提取效率。

高效特征复用残差块流程为:首先,采用不同尺度的

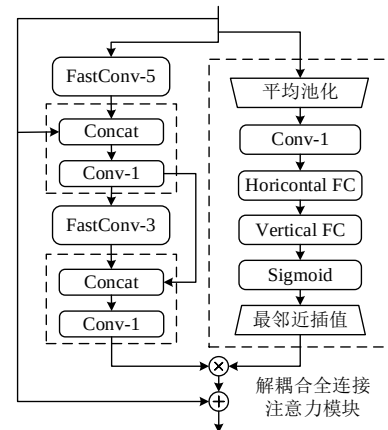


Fig. 2 Architecture of EFRRB

图2 高效特征复用残差块结构

快速卷积捕捉图像中不同大小和结构的特征信息;其次,使用跳跃连接及特征拼接与融合模块实现特征复用;再次,将快速卷积与解耦合注意力模块两部分提取的特征相乘,得到包含局部信息与远程依赖关系的聚合特征;最后,将聚合特征与输入相加得到输出特征。

定义输入特征为 $Feat_{in} \in \mathbb{R}^{H \times W \times C}$, 核大小为 5×5 、 3×3 的快速卷积提取的中间特征分别为 $Feat_1 \in \mathbb{R}^{H \times W \times C}$ 和 $Feat_2 \in \mathbb{R}^{H \times W \times C}$, 解耦合注意力模块的提取的特征为 $Feat_{dfc} \in \mathbb{R}^{H \times W \times C}$, 聚合特征为 $Feat_{total} \in \mathbb{R}^{H \times W \times C}$, 输出的特征为 $Feat_{out} \in \mathbb{R}^{H \times W \times C}$, 具体计算过程为:

$$Feat_1 = H_{fastconv5}(Feat_{in}) \quad (9)$$

$$Feat_2 = H_{fastconv3}(H_{fusion}(Feat_{in}, Feat_1)) \quad (10)$$

$$Feat_{total} = H_{fusion}(Feat_1, Feat_2) \times Feat_{dfc} \quad (11)$$

$$Feat_{out} = Feat_{total} + Feat_{in} \quad (12)$$

式中: $H_{fastconv5}(\cdot)$ 表示核大小为 5×5 的快速卷积; $H_{fastconv3}(\cdot)$ 表示核大小为 3×3 的快速卷积。

(1) 快速卷积层。由于普通卷积重叠计算的特性,会在不同位置和尺度上重复提取某些特征,导致出现大量冗余信息。为此,本文提出一种全新设计的快速卷积层来替换普通卷积(见图3)。首先,使用普通卷积层进行特征提取和通道压缩;其次,采用深度可分离卷积提取冗余特征以保证网络的泛化性能;最后,进行特征拼接与融合。

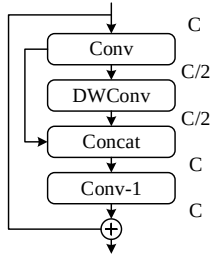


Fig. 3 Architecture of FastConv

图3 快速卷积层结构

(2) 解耦合全连接注意力模块(见图2右)。首先,使用步长为2的平均池化层对输入特征进行下采样,并添加归纳偏置;其次,分别采用核大小为 5×1 和 1×5 的深度可分离卷积层,表示纵向与横向全连接层的特征提取过程,并使用 Sigmoid 激活函数提升训练稳定性^[24];最后,通过最邻近插值对特征进行上采样,将捕获的远程依赖关系与原始特征图对齐。

定义原始特征输入为 $Feat_{src} \in \mathbb{R}^{H \times W \times C}$, 下采样的特征为 $Feat_{down} \in \mathbb{R}^{H/2 \times W/2 \times C}$, 纵向和横向特征提取后的中间特征分别为 $Feat_{vertical} \in \mathbb{R}^{H/2 \times W/2 \times C}$ 、 $Feat_{horizon} \in \mathbb{R}^{H/2 \times W/2 \times C}$, 输出为 $Feat_{target} \in \mathbb{R}^{H \times W \times C}$, 具体计算过程为:

$$Feat_{down} = H_{conv1 \times 1}(H_{avgpool}(Feat_{src})) \quad (13)$$

$$Feat_{vertical} = H_{dconv5 \times 1}(Feat_{down}) \quad (14)$$

$$Feat_{horizon} = H_{dconv1 \times 5}(Feat_{vertical}) \quad (15)$$

$$Feat_{target} = H_{nearest}(H_{sigmoid}(Feat_{horizon})) \quad (16)$$

式中: $H_{avgpool}(\cdot)$ 为平均池化; $H_{dconv5 \times 1}(\cdot)$ 为 5×1 为纵向

深度可分离卷积; $H_{dconv1 \times 5}(\cdot)$ 为 1×5 横向深度可分离卷积; $H_{nearest}(\cdot)$ 为最邻近插值上采样; $H_{sigmoid}(\cdot)$ 为 Sigmoid 激活函数。

2.3 图像重建模块

图像重建模块由对比感知特征融合模块和上采样模块两部分组成,如图4所示。定义图像的放大尺度为 s , 通过双三次插值算法放大的低分辨率图片为 $I_{bicubic} \in \mathbb{R}^{H \times sW \times C_{in}}$, 对比感知融合的特征为 $Feat_{fusion} \in \mathbb{R}^{H \times W \times C}$, 重建的高分辨率图像为 $I_{SR} \in \mathbb{R}^{sH \times sW \times C_{in}}$, 具体计算过程为:

$$Feat_{fusion} = H_{contrast}(F_{deep}) \times F_{deep} + F_{shallow} \quad (17)$$

$$I_{SR} = H_{upsample}(Feat_{fusion}) + I_{bicubic} \quad (18)$$

式中: $H_{contrast}(\cdot)$ 表示对比感知的特征计算过程; $H_{upsample}(\cdot)$ 表示上采样过程。

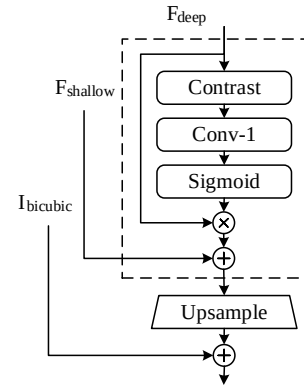


Fig. 4 Structure of image reconstruction module

图4 图像重建模块结构

(1) 特征融合模块。该模块采用对比感知的特征融合机制,其中对比感知信息为通道均值和方差之和。具体而言,定义输入特征图为 $X = [x_1, x_2, \dots, x_C]$, 高、宽和通道数分别为 H 、 W 、 C , 那么第 i 个通道 x_i 的均值 $mean_i$ 、标准差 std_i 及对比感知信息为 $contrast_i$ 的计算过程为:

$$mean_i = \frac{1}{HW} \sum_{(j,k) \in x_i} x_i^{j,k} \quad (19)$$

$$std_i = \sqrt{\frac{1}{HW} \sum_{(j,k) \in x_i} (x_i^{j,k} - mean_i)^2} \quad (20)$$

$$contrast_i = mean_i + std_i \quad (21)$$

式中: $x_i^{j,k}$ 表示第 i 个通道中第 j 行 k 列的特征值。

在得到对比感知信息后,首先使用一个核大小为 1×1 的卷积层来增加归纳偏置;其次使用 Sigmoid 激活函数提升训练稳定性,将结果与输入特征相乘得到加权的聚合特征;最后将聚合特征与浅层特征相加得到最终的特征输出。

(2) 上采样模块。由图5可见,该模块采用亚像素卷积算法,能在放大图像的同时保持图像的清晰度和细节信息。在放大尺度为4的情况下,使用两次亚像素卷积进行上采样,可保证其在高尺度下的重建图像质量。

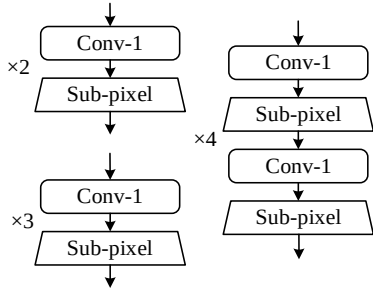


Fig. 5 Architecture of upsample module
图5 上采样模块结构

3 实验结果与分析

3.1 实验设置

本文训练阶段采用DIV2K数据集,测试阶段则采用Set5、Set14、Urban100及BSD100这4个图像超分辨率重建常用的基准数据集^[25-29]。对比模型选取SRCNN^[10]、FS-RCNN^[12]、VDSR^[11]、DRCN^[15]、CARN^[14]、IMDN^[17]等6个图像超分辨率领域经典的轻量级卷积神经网络,这些模型参数量从8 KB~1 774KB不等。每次训练迭代时,从低分辨率图片中随机选取16个48×48大小的三通道的RGB图像切片作为模型输入,图像增强策略为随机旋转90°、180°、270°或水平翻转,优化器选择ADAM算法,参数为 $\beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 10^{-8}$,训练初始学习率为 1×10^{-4} ,每经过 2×10^5 次迭代学习率下降50%。

本文实验环境为Ubuntu 16.04 LTS操作系统,PyTorch的版本为1.8.0,GPU为NVIDIA GeForce 2080Ti 10 GB。采用两个常用的评价图像相似度指标,分别为峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)和结构相似性(Structural Similarity, SSIM)定量评估网络模型的图像超分辨率性能。峰值信噪比与结构相似性计算公式为:

$$MSE = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n [HR(i, j) - SR(i, j)]^2 \quad (22)$$

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) = 20 \cdot \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \quad (23)$$

$$SSIM(x, y) = [l(x, y)]^{\alpha} [c(x, y)]^{\beta} [s(x, y)]^{\gamma} \quad (24)$$

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (25)$$

式中: $HR(i, j)$ 表示原始高分辨率图像第*i*行第*j*列的像素值; $SR(i, j)$ 表示网络模型输出的高分辨率图像第*i*行第*j*列的像素值; MAX_I 为表示像素点颜色的最大数值,本文实验涉及的图像像素值最大均为255; $l(x, y)$ 用于比较两者之间的亮度; $c(x, y)$ 用于比较两者之间的对比度; $s(x, y)$ 用于比较两者之间的结构; $\mu_x, \mu_y, \sigma_x, \sigma_y$ 分别表示

与*y*的均值和方差; σ_{xy} 为*x*与*y*的协方差; C 为常量。

3.3 实验结果

本文实验在Set5、Set14、Urban100、BSD100数据集上进行测试,以评估不同方法在2倍、3倍和4倍尺度上的图像重建性能,实验结果如表1所示(最优结果用粗体突出,次优结果以下划线标注)^[26-29]。

Table 1 Quantitative results

表1 定量结果

方法	尺度	参数量(K)	Set5 ^[26]	Set14 ^[27]	Ur-ban100 ^[28]	BSD100 ^[29]
			PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
SRCNN ^[10]	×2	8	36.66/ 0.954 2	32.45/ 0.906 7	29.50/ 0.894 6	31.36/ 0.887 9
FS-RCNN ^[12]		13	37.00/ 0.955 8	32.63/ 0.908 8	29.88/ 0.902 0	31.53/ 0.892 0
VDSR ^[11]		666	37.53/ 0.958 7	33.03/ 0.914 2	30.76/ 0.914 0	31.90/ 0.896 0
DRCN ^[15]		1774	37.63/ 0.958 8	33.04/ 0.911 8	30.75/ 0.913 3	31.85/ 0.894 2
IMDN ^[17]	×3	694	<u>38.00/</u> <u>0.960 5</u>	<u>33.63/</u> <u>0.917 7</u>	<u>32.17/</u> 0.928 2	<u>32.19/</u> 0.899 6
CARN ^[14]		1592	37.76/ 0.959 0	33.52/ 0.916 6	31.92/ <u>0.928 3</u>	32.09/ 0.897 8
Ours		636	38.05/ 0.960 7	33.67/ 0.917 9	<u>32.16/</u> 0.928 4	<u>32.25/</u> <u>0.899 3</u>
SRCNN ^[10]		8	32.75/ 0.909 0	29.30/ 0.821 5	26.24/ 0.798 9	28.41/ 0.786 3
FS-RCNN ^[12]	×4	13	33.18/ 0.914 0	29.37/ 0.824 0	26.43/ 0.808 0	28.53/ 0.791 0
VDSR ^[11]		666	33.66/ 0.921 3	29.77/ 0.831 4	27.14/ 0.827 9	28.82/ 0.797 6
DRCN ^[15]		1774	34.03/ 0.924 4	29.96/ 0.834 9	27.53/ 0.837 8	28.95/ 0.800 4
IMDN ^[17]		694	<u>34.36/</u> 0.927 0	<u>30.32/</u> <u>0.841 7</u>	<u>28.17/</u> <u>0.851 9</u>	29.05/ <u>0.804 6</u>
CARN ^[14]	×5	1592	34.29/ <u>0.925 5</u>	30.28/ 0.840 7	28.06/ 0.849 3	<u>29.06/</u> 0.803 4
Ours		636	34.37/ 0.927 0	<u>30.28/</u> 0.841 9	28.25/ 0.851 7	29.10/ 0.804 5
SRCNN ^[10]		8	30.48/ 0.862 8	27.50/ 0.751 3	24.52/ 0.722 1	26.90/ 0.710 1
FS-RCNN ^[12]		13	30.72/ 0.866 0	27.61/ 0.755 0	24.62/ 0.728 0	26.98/ 0.715 0
VDSR ^[11]	×6	666	31.35/ 0.883 8	28.01/ 0.767 4	25.18/ 0.752 4	27.29/ 0.725 1
DRCN ^[15]		1774	31.53/ 0.885 4	28.02/ 0.767 0	25.14/ 0.751 0	27.23/ 0.723 3
IMDN ^[17]		694	<u>32.21/</u> <u>0.894 8</u>	28.53/ <u>0.781 1</u>	26.04/ 0.783 1	27.55/ <u>0.735 3</u>
CARN ^[14]		1592	32.13/ 0.893 7	<u>28.60/</u> 0.780 6	<u>26.07/</u> <u>0.783 7</u>	<u>27.58/</u> 0.734 9
Ours	636	32.28/ 0.895 3	28.61/ 0.781 5	26.10/ 0.784 9	<u>27.56/</u> 0.735 8	

由定量结果可知, SRCNN、FSRCNN 作为最早的轻量级图像超分辨率重建方法, 虽然参数量极低, 但是过于简单的网络结构会导致模型的特征学习和表达能力不佳, 难以获得较高的图像重建质量。VDSR 虽通过增加卷积层数来提升特征学习和表达能力, 但其先上采样后特征提取的策略不仅增加了计算成本, 还限制了卷积核所能提取的局部信息, 导致重建效果不如人意。尽管 DRCN 通过递归方式复用卷积层以减少参数量, 但其参数量仍高达 1 774 KB, 且在 PSNR 指标和 SSIM 指标上均未能超越本文方法。CARN 虽通过级联方式实现特征复用, 但复杂的连接方式和低效的特征提取算法使其参数规模庞大。本文方法得益于快速的卷积结构和优化的特征复用策略, 不仅在重建质量上更胜一筹, 还在推理速度上表现更为出色。实验结果表明, IMDN 的定量结果与本文方法最为接近, 但本文方法参数量更少。

为了进一步验证本文方法的性能, 还与 FSRCNN^[12]、VDSR^[11]、DRCN^[15]、CARN^[14] 和 IMDN^[17] 进行定性比较。图 6 为针对 Urban100 数据集第 95 张图片的重建结果, 图像放大尺度为 4 倍。图 7 为针对 BSD100 数据集第 95 张图片的重建结果, 重建图像放大尺度为 3 倍。从定性结果可知, FSRCNN、VDSR、DRCN 在应对复杂线条的高放大尺度重建时, 均出现了不同程度的伪影效应和图像失真, 尤其是 FSRCNN 的伪影效应最为明显, 影响了重建图像的整体质量。相对而言, CARN、IMDN 及所提方法在抑制伪影和失真方面表现更出色, 重建图像更清晰自然。在细节方面, 本文方法由于采用了解耦合全连接注意力与对比感知的特征融合策略, 对图像整体细节特征的捕捉能力更优秀。

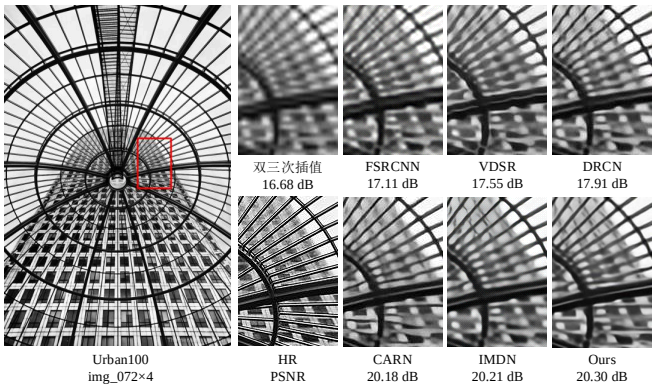


Fig. 6 Qualitative results comparison on Urban100 dataset
图 6 Urban100 数据集上的定性结果对比

3.4 消融实验

为了验证特征复用结构的有效性, 在放大尺度为 4 倍的图像超分辨率场景下, 与特征串联的线性结构进行比较, 并进行消融实验, 结果如表 2 所示。由此可知, 特征复用结构相较于特征串联而言, 能显著加速特征信息在网络中的传递, 在相同网络深度下能实现更出色的图像超分辨率重建效果。

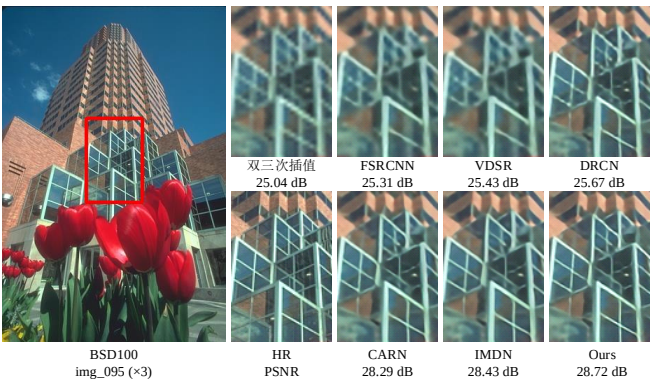


Fig. 7 Qualitative results comparison on BSD100 dataset
图 7 BSD100 数据集定性结果比较

Table 2 Ablation experimental result of feature reuse structure
表 2 特征复用结构的消融实验结果

结构	Set5 ^[26]	Set14 ^[27]	Urban100 ^[28]	BSD100 ^[29]
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
特征串联	31.70/0.889 2	28.21/0.772 5	25.35/0.769 0	27.37/0.729 2
特征复用	31.82/0.890 1	28.29/0.773 2	25.49/0.769 8	27.43/0.729 8

此外, 为了深入探究快速卷积层、解耦合自注意力模块以及对比感知特征融合策略对图像重建性能的影响, 在放大尺度为 4 倍的图像超分辨率场景进行消融实验, 实验均采用了特征复用的网络结构, 结果如表 3 所示。由此可知, 将普通卷积替换为快速卷积过后参数量从 963 KB 减少到 602 KB, 且并未造成过大的性能损失, 证明了快速卷积层能在减少网络参数量的同时, 保持较高的图像重建质量; 引入解耦合的自注意力模块后, 网络获得了捕获远程依赖和空间信息的能力, 提升了图像重建效果, 特别是在 Set5 数据集上 PSNR 值提升高达 0.45 dB, 充分证明了解耦合自注意力模块在提升图像边缘细节和整体结构重建方面的有效性; 引入对比感知特征融合策略进一步增强了网络对图像高频信息的感知能力, 图像重建性能得到了进一步提升, 尤其在 Urban10 数据集上 PSNR 提升了 0.1 dB, 表明该策略在提升网络对复杂纹理和细节的处理方面发挥了积极作用。

Table 3 Ablation experimental results
表 3 消融实验结果

快速卷积	解耦合注意力	对比感知	参数量(K)	Set5 ^[24]	Set14 ^[25]	Urban100 ^[26]	BSD100 ^[27]
				PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
×	×	×	963	31.83/0.890 5	28.30/0.772 8	25.62/0.769 9	27.45/0.729 9
				31.75/0.889 2	28.28/0.772 1	25.57/0.769 0	27.39/0.729 3
√	×	×	602	32.20/0.894 4	28.57/0.780 5	26.00/0.783 5	27.55/0.735 2
				32.28/0.895 3	28.61/0.781 5	26.10/0.784 9	27.56/0.735 8

4 结语

本文提出了一种高效特征复用与对比感知增强网络重建图像超分辨率。首先结合解耦合全连接注意力模块增强网络的远程依赖;其次引入对比感知的特征融合策略等方面对图像超分辨率重建网络进行改进,以提升特征信息在网络中的传递速度、优化卷积层的特征提取效率。

实验表明,本文方法相较于比较模型,因其快速的卷积结构和优化的特征复用策略,不仅在重建质量上更胜一筹,还在推理速度上表现更为出色。与此同时,本文提出的网络轻量级策略也为其他计算机视觉任务提供了新的思路,未来可在图像复原、图像去模糊等图像处理技术上进行探索与改进。

参考文献:

- [1] AAKERBERG A, JOHANSEN A, NASROLLAHI K, et al. Semantic segmentation guided real-world super-resolution [C]// Proceedings of the IEEE Winter Conference on Applications of Computer Vision, 2022: 449-458.
- [2] XIAOLIN F, FAN H, MING Y, et al. Small object detection in remote sensing images based on super-resolution [J]. Pattern Recognition Letters, 2022, 153: 107-112.
- [3] KABIRAJ A, PAL D, GANGULY D, et al. Number plate recognition from enhanced super-resolution using generative adversarial network [J]. Multimedia Tools and Applications, 2023, 82(9): 13837-13853.
- [4] QIU D, ZHENG L, ZHU J, et al. Multiple improved residual networks for medical image super-resolution [J]. Future Generation Computer Systems, 2021, 116: 200-208.
- [5] WANG Y, BASHIR S M A, KHAN M, et al. Remote sensing image super-resolution and object detection: benchmark and state of the art [J]. Expert Systems with Applications, 2022, 197: 116793.
- [6] LIU J, GE Y F, TIAN M. Research on super-resolution reconstruction algorithm of cultural relic images [J]. Acta Electronica Sinica, 2023, 51(1): 139-145.
刘杰, 葛一凡, 田明. 文物图像的超分辨率重建算法研究 [J]. 电子学报, 2023, 51(1): 139-145.
- [7] QIN J Z. Research on microcrack detection and image super resolution reconstruction of aircraft engine turbine blades [D]. Nanchang: Nanchang Hangkong University, 2022.
秦基展. 航空发动机涡轮叶片微裂纹检测及图像超分辨率重建研究 [D]. 南昌: 南昌航空大学, 2022.
- [8] LEPCHA D C, GOYAL B, DOGRA A, et al. Image super-resolution: a comprehensive review, recent trends, challenges and applications [J]. Information Fusion, 2023, 91: 230-260.
- [9] LIU H A, ZHENG Q X, TAO R L, et al. Review of image super-resolution based on deep learning [J]. Journal of Graphics, 2023, 44(1): 1-15.
李洪安, 郑峭雪, 陶若霖, 等. 基于深度学习的图像超分辨率研究综述 [J]. 图学学报, 2023, 44(1): 1-15.
- [10] 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.
- [11] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1646-1654.
- [12] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network [C]// Proceedings of the European Conference on Computer Vision, 2016: 391-407.
- [13] SHI W, CABALLERO J, HUSZAR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1874-1883.
- [14] AHN N, KANG B, SOHN K A. Fast, accurate, and lightweight super-resolution with cascading residual network [C]// Proceedings of the European Conference on Computer Vision, 2018: 252-268.
- [15] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1637-1645.
- [16] ZHANG X, ZENG H, GUO S, et al. Efficient long-range attention network for image super-resolution [C]// Proceedings of the European Conference on Computer Vision, 2022: 649-667.
- [17] HUI Z, GAO X, YANG Y, et al. Lightweight image super-resolution with information multi-distillation network [C]// Proceedings of the 27th ACM International Conference on Multimedia, 2019: 2024-2032.
- [18] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale [C]// Vienna: International Conference on Learning Representations, 2021.
- [19] LU Z, LI J, LIU H, et al. Transformer for single image super-resolution [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2022: 457-466.
- [20] LIANG J, CAO J, SUN G, et al. SwinIR: Image restoration using swin transformer [C]// Proceedings of the IEEE International Conference on Computer Vision, 2021: 1833-1844.
- [21] HAN K, WANG Y, CHEN H, et al. A survey on vision transformer [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(1): 87-110.
- [22] MEHTA S, RASTEGARI M. MobileViT: light-weight, general-purpose, and mobile-friendly vision transformer [DB/OL]. <https://arxiv.org/abs/2110.02178>.
- [23] LIU Z, LIN Y, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows [C]// Proceedings of the IEEE International Conference on Computer Vision, 2021: 10012-10022.
- [24] TANG Y, HAN K, GUO J, et al. Ghostnetv2: enhance cheap operation with long-range attention [J]. Advances in Neural Information Processing Systems, 2022, 35: 9969-9982.
- [25] AGUSTSSON E, TIMOFTE R. NTIRE 2017 challenge on single image super-resolution: dataset and study [C]// Honolulu: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017.
- [26] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding [C]// Surrey: Electronic Proceedings of the British Machine Vision Conference, 2012.
- [27] YANG J, WRIGHT J, HUANG T S, et al. Image super-resolution via sparse representation [J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861-2873.
- [28] HUANG J B, SINGH A, AHUJA N. Single image super-resolution from transformed self-exemplars [C]// Boston: IEEE Conference on Computer Vision and Pattern Recognition, 2015.
- [29] MARTIN D, FOWLKES C, TAL D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics [C]// Proceedings of the 8th IEEE International Conference on Computer Vision, 2001: 416-423.

(责任编辑: 刘嘉元)