

文章编号:1007-2780(2026)02-0280-14

基于高频融合注意力的轻量级图像超分辨率重建

郑龙光^{1,2}, 朱 波^{1*}, 李国梁^{1,2}, 龙家军^{1,2}

(1. 中国科学院 西安光学精密机械研究所, 陕西 西安 710119;

2. 中国科学院大学, 北京 100049)

摘要:针对现有图像超分辨率算法计算成本高昂、局部高频纹理的感知有限以及关键信息挖掘不充分的问题,本文提出了一种基于高频融合注意力的轻量化图像超分辨率重建网络。首先,在网络结构中设计了双路径增强模块,用于将特征显式解耦为高频和低频两个并行分支。低频分支用于捕获平滑结构,高频分支应用拉普拉斯算子初始化的可学习高通滤波器来准确提取边缘和纹理。其次,提出了空间细化模块,利用方向分离的卷积对局部特征进行增强,输出具有丰富空间信息的查询(Query)。最后,设计了一种高效的非对称频率感知注意力机制,用于获得局部信息和全局信息之间的有效交互。该机制以空间细化模块输出的局部信息为查询,在全局范围内动态检索互补信息,有效解决了局部窗口对长程依赖的限制。在Urban100数据集上的实验表明,算法在 $\times 4$ 放大尺度下相较于其他模型PSNR平均提升0.53 dB,SSIM平均提升0.014 7,并且运行时间仅为SwinIR的14%。在图像重建视觉质量方面,模型可以生成更清晰的边缘和更逼真的纹理,充分验证了模型在保持轻量化的同时,在图像重建过程中的有效性和先进性。

关键词:图像处理;超分辨率重建;注意力机制;特征融合;轻量级

中图分类号:TP391.4 **文献标识码:**A **doi:**10.37188/CJLCD.2025-0247 **CSTR:**32172.14.CJLCD.2025-0247

Lightweight image super-resolution reconstruction based on high frequency fusion attention

ZHENG Longguang^{1,2}, ZHU Bo^{1*}, LI Guoliang^{1,2}, LONG Jiajun^{1,2}

(1. Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences,
Xi'an 710119, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Existing image super-resolution algorithms face issues such as high computational cost, limited perception of local high-frequency textures, and insufficient extraction of key information. To address these challenges, this paper proposes a lightweight image super-resolution reconstruction network based on high-frequency fusion attention. First, a Dual-Path Enhancement module is designed within the network architecture to explicitly decouple features into parallel high-frequency and low-frequency branches. The low-frequency branch captures smooth structures, while the high-frequency branch employs learnable high-pass filters to accurately extract edges and textures. Second, a spatial refinement module is introduced,

收稿日期:2025-12-05;修订日期:2026-01-14.

基金项目:国家自然科学基金(No.52505045)

Supported by National Natural Science Foundation of China (No.52505045)

*通信联系人, E-mail:irico_zb@qq.com

utilizing directionally separated convolutions to enhance local features and generate queries rich in spatial information. Finally, an efficient asymmetric frequency-aware attention mechanism is designed to achieve effective interaction between local and global information. This mechanism uses the local information output from the spatial refinement module as a query to dynamically retrieve complementary information across the entire image, effectively overcoming the limitation of local windows on long-range dependencies. Experiments on the Urban100 dataset demonstrate that at $\times 4$ magnification scales, the proposed algorithm achieves an average PSNR improvement of 0.53 dB and an average SSIM improvement of 0.014 7 compared to other models, while running at only 14% of SwinIR's computational time. Regarding visual quality in image reconstruction, the model generates sharper edges and more realistic textures, fully validating its effectiveness and advanced capabilities in image reconstruction while maintaining lightweight efficiency.

Key words: image processing; super-resolution reconstruction; attention mechanism; feature fusion; lightweight

1 引言

单图像超分辨率(Single Image Super-Resolution, SISR)是计算机视觉领域的一个重要课题,其研究目标是从单张低分辨率图像中恢复出高分辨率图像,在医学影像^[1]、视频监控^[2-3]、人脸识别^[4-8]等诸多领域具有广泛的应用价值。例如,在医学影像领域,超分辨率重建技术可以帮助我们更好地观察细胞结构,应用于病理诊断。在卫星遥感领域,超分辨率重建可以提高土地覆盖分类和目标检测的准确性。但是从低分辨率图像合成清晰稳定的高分辨率图像仍然是一个热点问题。

图像超分辨率重建传统方法主要分为3类:插值、重建、学习^[9]。近年来,深度卷积神经网络(Convolutional Neural Networks, CNN)凭借其强大的特征自适应提取与表征能力,在图像超分辨率等计算机视觉任务中占据了主导地位。Dong等^[10]提出了SRCNN(Super Resolution Convolutional Neural Network)方法,在特征提取、非线性映射和重建阶段构建了一个端到端的3层卷积网络架构,将基于卷积神经网络的方法引入到图像超分辨率重建中。该方法能够大幅提高重建性能,并大幅降低时间成本,充分证明了卷积神经网络方法在超分辨率重建中的可行性和优势。Kim^[11]构造了一个具有20层深层网络的级联架构,并证明了深度特征提取在图像重建中的有效性。Lim等^[12]提出了增强型深度超分辨率网络(EDSR)。他们去除了残差块中的批归一化,并添加了额外

的卷积层以进一步改进该方法。Chen等^[13]提出了多尺度注意力残差网络,在多尺度特征提取单元加入通道注意力机制,提升了模型的计算效率。为了开发轻量级模型,Hui等^[14]引入了一种选择性融合策略,构建了一个轻量级网络,该网络采用信息蒸馏模块的级联设计。然而,受限于卷积运算固有的局部感受野,此类方法在捕捉图像全局上下文信息方面存在明显不足。伴随网络层数的增加,模型参数量与计算复杂度的急剧上升导致推理效率降低,极大地限制了其在实际场景中的部署与应用。

Transformer在自然语言处理方面取得了巨大成功,被广泛应用于计算机视觉领域^[15]。Liang等^[16]提出的SwinIR模型使用基于移位窗口的局部自注意力,有效实现了相邻局部特征区域间的信息交流。Lu等^[17]设计了一种轻量级Transformer网络,利用高效注意力机制对图像中的远距离相似区域进行处理,提高了推理效率。为使模型轻量化,Zhang等^[18]提出将移位卷积与分组多尺度自注意力相结合,实现了特征的多尺度提取。Wang等^[19]构建了包含全向自注意力的轻量级网络,协同处理空间与通道维度信息,增强了特征表达。He等^[4]提出多种注意力融合机制,加强了模型对于空间特征信息的理解能力。上述方法取得了显著进展,但往往难以在局部纹理感知的特征提取与全局建模的计算效率之间取得最佳平衡。

为解决上述问题,本文提出一种基于高频融合注意力的轻量级图像超分辨率重建网络(High

Frequency Fusion Attention Network, HFANet)。该网络以高频融合注意力模块(High Frequency Fusion Block, HFAB)为核心,旨在从解耦特征中有效捕获长程依赖关系,并协同局部窗口自注意力(Local Window Self Attention, LWSA)模块实现对图像局部细节的重建。实验表明, HFANet在维持网络轻量化的同时,显著提升了图像的重建质量,并在参数量降低的情况下,减少了网络的推理时间。本文的主要贡献总结如下:(1)针对现有网络对高频纹理信息感知不足的问题,提出了一个双路增强模块(Dual-Path Enhancement Module, DPE),用于分离和增强图像中的高频和低频特征。其中,低频分支使用标准的 3×3 卷积来捕获图像中的平滑结构和轮廓信息,高频分支则引入由拉普拉斯算子初始化的卷积层以提取边缘纹理信息。最后借助通道注意力机制对两个分支的响应进行动态加权融合。网络可以根据全局内容自适应地平衡结构信息和精细细节信息。(2)为克服卷积运算局部感受野的局限性,构建了一种高效的非对称频率感知注意力(Asymmetric Frequency-aware Attention, AFA)机制,利用空间细化模块(Spatial Refinement Module, SRM)整合输出的富含纹理特征作为查询(Query),来自 HFAB 模块的原始输入特征直接作为键(Key)和值(Value)。该设计增强了网络对长程依赖关系的捕捉能力,同时降低了网络复杂度。在多个数据集上的实验结果验证了所提方法的先进性,本文所提模型在图像超分辨率性能和推理时间上取得了较好的平衡。

2 设计原理及方法

本文提出的 HFANet 整体网络由 3 个模块构成:浅层特征提取、深层特征提取和图像重建,其

整体架构如图 1 所示。浅层特征提取由一个 3×3 的卷积层完成,它将低分辨率输入图像从原始输入空间映射到一个高维特征空间,首先通过公式(1)获得浅层特征 F_0 :

$$F_0 = H_{SF}(I_{LR}), \quad (1)$$

式中, H_{SF} 表示浅层特征提取函数。 F_0 被传输至深层特征提取模块。该模块由 N 个级联的残差组(Residual Group, RG)堆叠而成,旨在提取图像丰富上下文信息的深层特征。每个 RG 内部集成了高频融合注意力模块(HFAB)、局部窗口自注意力模块(LWSA)及 3×3 卷积层,这种设计能够实现长程依赖建模与局部信息的自适应增强。为稳定训练过程并促进梯度传播,网络引入了双重残差连接策略:利用长跳跃连接直接将浅层特征 F_0 传递至深层网络的末端,同时每个 RG 内部采用短跳跃连接。深层特征提取的输出 F_{deep} 定义如下:

$$F_{deep} = H_{DF}(F_0), \quad (2)$$

其中, H_{DF} 代表由 N 个残差组构成的深层特征提取函数。最终,图像重建模块接收深层特征 F_{deep} 输入到重建模块中。在该模块中,首先通过 3×3 的特征融合卷积层对特征进行融合,通过 PixelShuffle 层进行快速上采样,将特征图的分辨率提升至目标尺寸。上采样后的特征与通过双线性插值直接从输入 LR 图像得到的基准特征 I_{base} 相加,最终生成高分辨率图像 I_{HR} 。该过程可表示为:

$$I_{HR} = H_{REC}(F_{deep}) + F_{Up}(I_{LR}), \quad (3)$$

其中: F_{Up} 代表上采样算子的函数; H_{REC} 代表图像重建函数,该函数包含一个 3×3 卷积和一个像素重排操作。

在深度特征提取部分中,高频融合注意力模块(HFAB)为 RG 的核心功能单元。鉴于该模块在提升网络特征提取能力的重要性,下文将重点

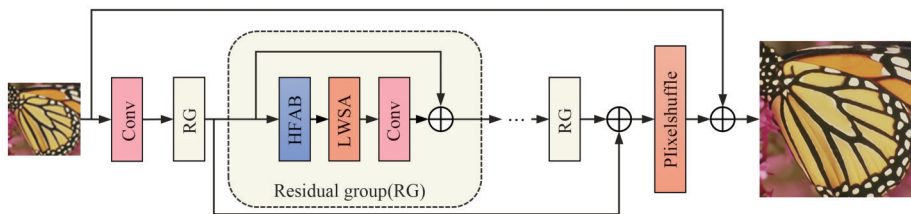


图 1 整体网络架构图

Fig. 1 Overall structure of the network

阐述HFAB的内部结构及运作机制。

HFAB的结构如图2所示。该模块旨在实现对图像特征从解耦增强到深度交互的完整处理流程。对于输入特征 X_{in} ,数据依次经过DPE模块、SRM模块、AFA机制和ConvFFN。

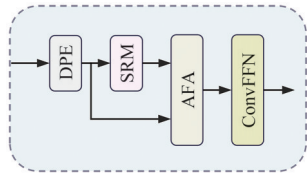


图2 高频融合注意力模块结构

Fig. 2 Structure of the HFAB

图2可以清晰地说明输入图像特征 X_{in} 从双路增强模块(DPE)和空间细化模块(SRM)到非对称频率感知注意力(AFA)的完整处理过程。该模块可以有效生成更多信息的局部特征,并使用它们来指导全局上下文中的有针对性的信息聚合,使网络具有强大的特征表示学习能力。

2.1 双路径增强模块

为了避免不同属性信息在处理过程中的相互干扰,双路增强模块(DPE)显式地将图像特征解耦为高、低频两个分量并进行针对性增强(图3)。对于输入特征 $X \in R^{H \times W \times C}$,该模块首先通过两条并行支路进行特征提取:

$$F_{low} = H_{Conv}(X), \quad (4)$$

$$F_{high} = H_{LaplaceConv}(X), \quad (5)$$

其中:低频支路 H_{Conv} 采用标准的 3×3 卷积捕获平滑结构;高频支路 $H_{LaplaceConv}$ 则是一个由拉普拉斯算子初始化的可学习卷积层,以精准提取边缘与纹理细节。

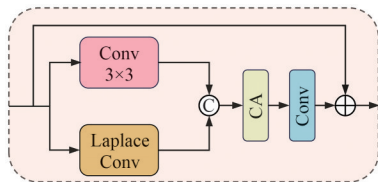


图3 双路径增强模块结构

Fig. 3 Structure of the DPE

两条支路输出的特征在通道维度上进行拼接,形成一个通道数为 $2C$ 的特征图 F_{cat} 。随后,该特征图被送入一个通道注意力模块进行动态

加权,并最终通过一个 1×1 卷积融合,并与DPE的原始输入特征进行残差相加,得到最终的增强特征 F_{out} 。该过程可表示为:

$$F_{cat} = \text{Concat}(F_{low}, F_{high}), \quad (6)$$

$$F_{out} = H_{Conv1 \times 1}(H_{CA}(F_{cat})) + X, \quad (7)$$

其中: $H_{Conv1 \times 1}$ 表示 1×1 卷积, H_{CA} 表示通道注意力模块动态加权。DPE模块通过频率解耦和自适应融合为网络提供了一种特征增强机制,使模型能够以输入驱动的方式调整结构信息和细节信息的权重,增强特征的代表质量。

2.2 空间细化模块

为了进一步挖掘特征中非刚性的、不规则的局部纹理,本文设计了空间细化模块(SRM),该模块(图4)利用方向分离的可变形卷积对DPE的输出特征进行空间层面的细化。

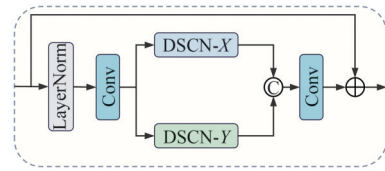


图4 空间细化模块结构

Fig. 4 Structure of the SRM

对于输入特征 X ,首先通过层归一化和 1×1 卷积进行预处理,并将其在通道维度上分成 X_1, X_2 两部分。其中 X_1, X_2 分别送入 X 方向和 Y 方向的卷积分支。将它们的拼接结果送入另一个 1×1 卷积,并通过与输入 X 的残差连接得到输出 F_{out} 。其过程可表示为:

$$[X_1, X_2] = \text{Split}[H_{Conv1 \times 1}(\text{LN}(X))], \quad (8)$$

$$Y_{cat} = \text{Concat}(H_{DSCN-X}(X_1), H_{DSCN-Y}(X_2)), \quad (9)$$

$$F_{out} = H_{Conv1 \times 1}(Y_{cat}) + X, \quad (10)$$

其中: H_{DSCN-X} 和 H_{DSCN-Y} 分别表示 X 和 Y 方向的可变形卷积, $H_{Conv1 \times 1}$ 表示 1×1 卷积,SRM模块为后续的注意力机制提供了蕴含丰富且精细空间细节的特征。

2.3 非对称频率感知注意力

非对称频率感知注意力(AFA)通过非对称的输入处理,实现了基于局部输入的全局信息检索,该模块结构如图5所示。经过DPE和SRM双重处理后的特征,代表了当前图像区域最需要被重建的、信息量最丰富的局部细节,因此将其作为查

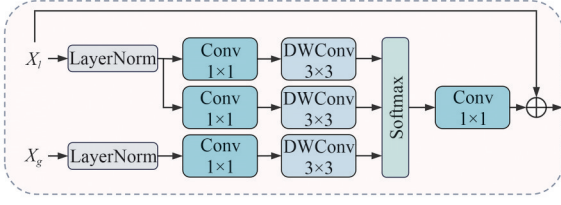


图5 非对称频率感知注意力模块结构

Fig. 5 Structure of the AFA

询(Query, Q)。而HFAB模块的原始输入特征则作为未经深度处理的全局上下文信息,构成了一个内容丰富的特征数据库,用于生成键(Key, K)和值(Value, V)。给定由SRM模块输出的局部特征 F_{local} 和 HFAB 的原始输入全局特征 F_{global} , Q 、 K 、 V 的生成过程如下:

$$Q = H_{\text{DWConv}}^Q \left(H_{\text{Conv}1 \times 1}^Q \left(\text{LN}(F_{\text{local}}) \right) \right), \quad (11)$$

$$K = H_{\text{DWConv}}^K \left(H_{\text{Conv}1 \times 1}^K \left(\text{LN}(F_{\text{global}}) \right) \right), \quad (12)$$

$$V = H_{\text{DWConv}}^V \left(H_{\text{Conv}1 \times 1}^V \left(\text{LN}(F_{\text{global}}) \right) \right), \quad (13)$$

其中, H^Q 、 H^K 、 H^V 均是由层归一化、 1×1 卷积和深度可分离卷积构成的投影函数。注意力计算遵循标准的缩放点积注意力范式:

$$\text{Attention}(Q, K, V) = \text{Softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V, \quad (14)$$

其输出经过最终的 1×1 卷积进行投影。

2.4 卷积前馈网络

在非对称频率感知注意力之后堆叠了一个卷积前馈网络(ConvFFN)用于非线性特征处理。该模块遵循了 Restormer^[20] 中的设计,由两个线性变换层(通过 1×1 卷积实现),夹在 GELU 非线性激活函数之间。在 GELU 之后,添加一个 3×3 深度可分离卷积,以进一步聚合通道中的局部空间信息。最后,整个模块通过残差连接添加到输入中,以细化通道中的特征。

在 HFAB 模块进行长程和跨特征交互后,本文添加一个局部窗口自注意力(LWSA)模块,进一步细化和聚合局部空间邻域中的特征表示。该模块采用了 Swin Transformer 中高效的窗口自注意力模块,作为 HFAB 的补充模块,进行全局-局部建模。

2.5 损失函数

在图像超分辨率重建领域中,为了与该领域一系列先进算法保持一致,也为了进一步突出模型

的先进性,在优化策略方面,本文采用 L1 损失函数作为网络训练的目标优化函数。

L1 损失被广泛证明能够更好地保留图像的边缘锐利度和纹理细节,生成视觉上更清晰、更自然的结果。L1 损失函数的表达式如式(15)所示:

$$\text{Loss}_{\text{L1}} = \frac{1}{M} \sum_{i=1}^M (I_{\text{HR}}^i - I_{\text{SR}}^i), \quad (15)$$

其中: I_{HR} 为高分辨率参考图像, I_{SR} 为网络生成的超分辨率图像。

3 实验结果及分析

3.1 实验数据与评价指标

本文采用被广泛使用的 DIV2K^[21] 数据集作为训练与验证集。DIV2K 包含 800 张来自不同场景的图像,并通过双三次插值以 $\times 2$ 、 $\times 3$ 和 $\times 4$ 的比例对原始高分辨率(HR)图像进行下采样来提取相应的低分辨率(LR)图像。为了评估不同训练数据条件下的性能和泛化性能,本文还在 5 个标准测试数据集上进行了评估,分别是 Set5^[22]、Set14^[23]、BSD100^[24]、Urban100^[25] 和 Manga109^[26]。

在评估方面,所使用的指标为峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似性(Structural Similarity, SSIM),这两种方法都是在从 RGB 色彩空间中提取的 YCbCr 色彩空间的 Y 通道上计算的。PSNR 是衡量重建图像与参考图像之间像素级差异的广泛使用的度量。一般来说,PSNR 值越大,超分辨率图像与高分辨率图像越相似,重建精度越高。但是,PSNR 只考虑了像素到像素的误差,没有考虑结构信息和人类视觉感知特性。SSIM 是一种从亮度、对比度、结构 3 个方面衡量重构图像与参考图像相似度的指标。SSIM 取值范围为 $[0, 1]$,重建图像与参考图像在结构信息上越相似,感知质量越高,SSIM 值越接近 1。

3.2 实验设置

本研究提出的网络运行于以下硬件环境: CPU 为 Intel Xeon Gold 6133, GPU 为 NVIDIA Geforce RTX 3090,采用 PyTorch 2.0.0 作为深度学习框架, CUDA 版本为 11.8,编程语言为 Python 3.8.3。

在训练阶段,使用大小为 64×64 的 LR 图像作为输入,并在训练过程中对图像进行处理,具体

操作为采用随机翻转和水平翻转对训练集进行数据增强。本文使用L1损失函数,使用 $\beta_1=0.9$ 和 $\beta_2=0.99$ 的Adam优化器,输入批次大小为16。学习率采用衰减策略进行更新,初始学习率为 2×10^{-4} ,并在迭代次数为250k,400k,450k,475k时将学习率减半,训练总迭代次数为50万。

在网络参数设置方面,网络基础特征通道设置为48。在深层特征提取阶段,本文设置了6个级联的残差组(RG),每个残差组包含HFAB、LWSA以及 3×3 卷积。多头注意力机制头数和维度分别设置为4和36。

3.3 对比实验

3.3.1 定量分析

为验证HFANet的重建性能,将其与其他主流图像超分辨率重建模型方法进行对比,分别为EDSR^[12]、CARN^[27]、IMDN^[15]、LatticeNet^[28]、LAPAR-A^[29]、HPUN-L^[30]、MSRA^[31]、CRAFT^[32]、

SwinIR-light^[16]、CAMixerSR^[33]、SRFormer^[34]、ELAN-light^[18]、LBNet^[35]、SGGLC-Net-L^[36]。本文所提模型在5个测试集上的PSNR均取得最优结果,SSIM基本取得最优结果(除SGGLC-Net-L外)。在缩放因子为2时,由表1可知,本文方法取得了全面的领先性能。与参数量较多的SGGLC-Net-L相比,本文模型在Urban100数据集上PSNR和SSIM分别提升了0.63 dB和0.005 0,在Manga109上提升了0.50 dB和0.001 1,且参数量下降245k。当缩放因子为3时,由表2可知,在BSD100测试集和Urban100测试集上,本文模型对比参数量相近的CRAFT,PSNR与SSIM分别提升0.08 dB、0.001 4和0.28 dB、0.005 1。当缩放因子为4时,由表3可知,本文提出的模型在5个测试集上均表现最优,特别是在Manga109测试集上,PSNR、SSIM相较于第二高的CRAFT提升高达0.35 dB、0.003 7。

表1 缩放因子为2时不同算法的性能对比

Tab. 1 Quantitative comparison of different methods at scale factor $\times 2$

Models	Params/k	Set5	Set14	BSD100	Urban100	Manga109
		(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)
EDSR	1 370	37.99/0.960 4	33.57/0.917 5	32.16/0.899 4	31.98/0.927 2	38.54/0.976 9
CARN	1 592	37.76/0.959 0	33.52/0.916 6	32.09/0.897 8	31.92/0.925 6	38.36/0.976 5
IMDN	694	38.00/0.960 5	33.63/0.917 7	32.19/0.899 6	32.17/0.928 3	38.88/0.977 4
LatticeNet	756	38.06/0.960 7	33.70/0.918 7	27.57/0.735 5	32.20/0.899 9	—/—
LAPAR-A	548	38.01/0.960 5	33.62/0.918 3	32.19/0.899 9	32.10/0.928 3	38.67/0.977 2
SGGLC-Net-L	870	38.13/0.963 1	33.88/0.921 2	32.28/0.901 0	32.61/0.932 7	39.05/0.978 0
HPUN-L	714	38.09/0.960 8	33.79/0.919 8	32.25/0.900 6	32.37/0.930 7	39.07/0.977 9
SwinIR-light	878	38.14/0.961 1	33.86/0.920 6	32.31/0.901 2	32.76/0.934 0	39.12/0.978 3
CAMixerSR	746	38.23/0.961 3	34.00/0.921 4	32.34/0.901 6	32.95/0.934 8	39.32/0.978 1
SRFormer-L	853	38.23/0.961 3	33.94/0.920 9	32.36/0.901 9	32.91/0.935 3	39.28/0.978 5
ELAN-light	582	38.17/0.961 1	33.94/0.920 7	32.30/0.901 2	32.76/0.934 0	39.11/0.978 2
MSRA	769	38.23/0.961 4	34.01/0.921 1	32.33/0.901 7	32.98/0.935 8	39.24/0.978 3
CRAFT	737	38.23/0.961 5	33.92/0.921 1	32.33/0.901 6	32.86/0.934 3	39.39/0.978 6
Ours	625	38.33/0.961 9	34.06/0.922 0	32.40/0.902 4	33.24/0.937 7	39.55/0.979 1

注:粗体数字表示最优值

实验结果表明,在客观质量评价指标上,本文模型取得了最优的性能表现。该结果也表明本文模型实现了更精确的图像重建,进一步验证了本文所提模型在图像超分辨率任务中的有效性和优越性。

3.3.2 主观结果展示

本文展示了5种具有挑战性的视觉比较结

果,以更直观地体现模型在复杂图像细节上的恢复性能。这些样本分别选自Urban100和BSD100数据集,涵盖了精细的结构化纹理、多样的光照条件以及复杂的自然纹理,所有结果均在 $\times 4$ 倍放大基础上进行测试。对于每种样本,从所有对比方法的复原结果中裁剪出同一空间位置的局部区域进行放大显示。

表 2 缩放因子为 3 时不同算法的性能对比

Tab. 2 Quantitative comparison of different methods at scale factor $\times 3$

Models	Params/k	Set5	Set14	BSD100	Urban100	Manga109
		(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)
EDSR	1 555	34.37/0.927 0	30.28/0.841 7	29.09/0.805 2	28.15/0.852 7	33.45/0.943 9
CARN	1 592	34.29/0.925 5	30.29/0.840 7	29.06/0.803 4	28.06/0.849 3	33.50/0.944 0
IMDN	703	34.36/0.927 0	30.32/0.841 7	29.09/0.804 6	28.17/0.851 9	33.61/0.944 5
LatticeNet	765	34.53/0.928 1	30.39/0.842 4	29.15/0.805 9	28.33/0.853 8	—/—
LAPAR-A	544	34.36/0.926 7	30.34/0.842 1	29.11/0.805 4	28.15/0.852 3	33.51/0.944 1
LBNet	736	34.47/0.927 7	30.38/0.841 7	29.13/0.806 1	28.42/0.855 9	33.82/0.946 0
SGGLC-Net-L	856	34.40/0.929 8	30.55/0.846 8	29.29/0.814 7	28.61/0.861 6	34.12/0.947 9
HPUN-L	723	34.56/0.928 1	30.45/0.844 5	29.18/0.807 2	28.37/0.857 2	33.90/0.946 3
SwinIR-light	886	34.62/0.928 9	30.54/0.846 3	29.20/0.808 2	28.66/0.862 4	33.98/0.947 8
CAMixerSR	753	34.65/0.929 5	30.62/0.847 1	29.26/0.809 3	28.81/0.864 5	34.34/0.949 1
SRFormer-L	861	34.67/0.929 6	30.57/0.846 9	29.26/0.809 9	28.81/0.865 5	34.19/0.948 9
ELAN-light	590	34.61/0.928 8	30.55/0.846 3	29.21/0.808 1	28.69/0.862 4	34.00/0.947 8
MSRA	777	34.65/0.929 1	30.60/0.847 0	29.24/0.809 3	28.86/0.866 4	34.29/0.948 9
CRAFT	744	34.71/0.929 5	30.61/0.846 9	29.24/0.809 3	28.77/0.863 5	34.29/0.949 1
Ours	729	34.79/0.930 4	30.74/0.848 9	29.32/0.810 7	29.05/0.868 6	34.59/0.950 8

注:粗体数字表示最优值

表 3 缩放因子为 4 时不同算法的性能对比

Tab. 3 Quantitative comparison of different methods at scale factor $\times 4$

Models	Params/k	Set5	Set14	BSD100	Urban100	Manga109
		(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)
EDSR	1 518	32.09/0.893 8	28.58/0.781 3	27.57/0.735 7	26.04/0.784 9	30.35/0.906 7
CARN	1 592	32.13/0.893 7	28.60/0.780 6	27.58/0.734 9	26.07/0.783 7	30.47/0.908 4
IMDN	715	32.21/0.894 8	28.58/0.781 1	27.56/0.735 3	26.04/0.783 8	30.45/0.907 5
LatticeNet	777	32.18/0.894 3	28.61/0.781 2	27.57/0.735 5	26.14/0.784 4	—/—
LAPAR-A	659	32.15/0.894 4	28.61/0.781 8	27.61/0.736 6	26.14/0.787 1	30.42/0.907 4
LBNet	742	32.29/0.896 0	28.68/0.783 2	27.62/0.738 2	26.27/0.790 6	30.76/0.911 1
SGGLC-Net-L	868	32.26/0.899 1	28.71/0.784 9	27.69/0.740 2	26.48/0.797 6	30.94/0.914 0
HPUN-L	734	32.31/0.896 2	28.73/0.784 2	27.66/0.738 6	26.27/0.791 8	30.77/0.910 9
SwinIR-light	897	32.44/0.897 6	28.77/0.785 8	27.69/0.740 6	26.47/0.798 0	30.92/0.915 1
CAMixerSR	765	32.51/0.898 8	28.82/0.787 0	27.72/0.741 6	26.63/0.801 2	31.18/0.916 6
SRFormer-L	873	32.51/0.898 8	28.82/0.787 2	27.73/0.742 2	26.67/0.803 2	31.17/0.916 5
ELAN-light	601	32.43/0.897 5	28.78/0.785 8	27.69/0.740 6	26.54/0.798 2	30.92/0.915 0
MSRA	789	32.46/0.898 4	28.86/0.787 6	27.72/0.741 9	26.65/0.803 7	31.08/0.915 7
CRAFT	753	32.52/0.898 9	28.85/0.787 2	27.72/0.741 8	26.56/0.799 5	31.18/0.916 8
Ours	708	32.68/0.901 1	28.96/0.789 7	27.81/0.744 0	26.89/0.808 1	31.53/0.920 5

注:粗体数字表示最优值

根据图 6 可知,在 Image_092 的网格纹理重建中,大多数对比方法都非常模糊,线条形状无法区分。虽然 IMDN 和 SRFormer 在一定程度上恢复了线条的走向,但仍存在明显的错误走向。与其他方法对比,本文所提模型成功地重建出清晰、独立的网格线条,视觉效果在结构上最接近

原图。由图 7 可知,在 Image_011 的平行线条场景中,Bicubic 和 CARN 同样表现出显著的模糊问题。IMDN 和 SRFormer 的重建结果虽然能够区分线条,但在线条边缘处产生了可见的模糊。本文模型则生成了边缘清晰、方向一致的线条。在 Image_076 的夜景建筑重建中,难点在于恢复窗

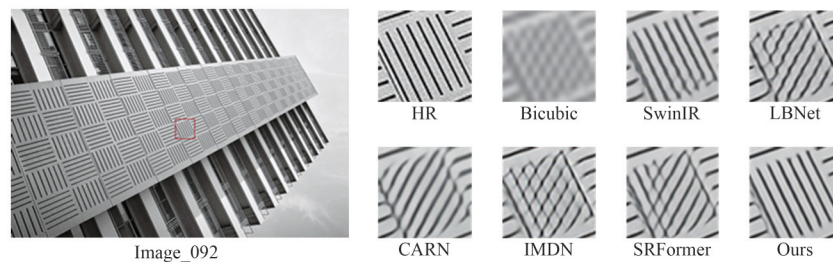


图6 Urban100中 Image_092的重建效果展示图

Fig. 6 Reconstruction results for Image_092 from the Urban100

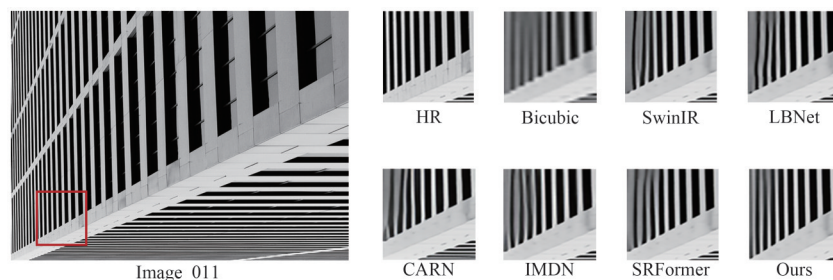


图7 Urban100中 Image_011的重建效果展示图

Fig. 7 Reconstruction results for Image_011 from the Urban100

户框架的精细结构。如图8所示,所有对比方法都难以清晰地重建窗户的网格状结构,恢复的窗户过于平滑,甚至出现纹理错误的现象。本文模型在细节表现上更加完整,具有更清晰的窗户结构和建筑纹理细节,准确地恢复了窗框的几何形状和方向。在图9的老虎皮毛重建中,IMDN的重建

结果显得过于平滑,丢失了大量细节。SwinIR和SRFormer虽然恢复了一部分纹理,但条纹边缘仍然模糊。本文模型重建出了最清晰的虎皮条纹,并且保留了皮毛的精细质感,其结果在视觉上最接近原图。在图10中,包括SwinIR和SRFormer的大多数方法都未能准确恢复这些线条,产生了



图8 Urban100中 Image_076的重建效果展示图

Fig. 8 Reconstruction results for Image_076 from the Urban100



图9 BSD100中 Image_007的重建效果展示图

Fig. 9 Reconstruction results for Image_007 from the BSD100 dataset

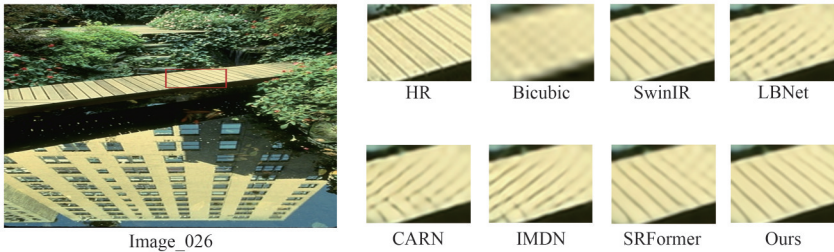


图 10 BSD100中 Image_026 的重建效果展示图
Fig. 10 Reconstruction results for Image_026 from the BSD100

模糊或断裂的现象。相比之下,本文模型成功地抑制了模糊,重建出了连续且清晰的栏杆线条。

综上所述,本文提出的 HFANet 在处理各种具有挑战性的图像内容时,具有更好的视觉效果。这说明通过本文设计的模块能有效增强高频信息,精确地重建复杂细节。

3.3.3 运行时间及 LAM 可视化分析

网络推理时间是衡量轻量化超分辨率算法的一个重要指标。为进一步验证所提出的网络是一个轻量化模型,本文选择定量分析中 IMDN、SwinIR-light、ELAN、CRAFT 模块,在 Urban100 数据集上针对缩放因子为 4 的重建速度进行了对比实验。测试图像尺寸为 64 像素×64 像素,如表 4 所示,平均运行时间是在运行测试 10 次后获得的。

由表 4 可知,本文所提模型的重建速度最快,并且重建质量最高,与 SwinIR 相比参数量下降了 222k,乘加运算次数 (MACs) 下降了 14%,并且平均运行时间 (Average Running Time, ART)

表 4 不同模型在 Urban100 上的平均运行时间

Tab.4 Average running time of different models on Urban100 dataset

Models	Params/k	MACs/G	ART/s	Urban100
				PSNR/SSIM
IMDN	715	40.9	0.022 0	26.04/0.783 8
ELAN	640	54.1	0.020 8	26.54/0.798 2
SwinIR	930	63.6	0.135 2	26.47/0.798 0
CRAFT	753	46.5	0.088 7	26.56/0.799 5
Ours	708	54.9	0.019 5	26.89/0.808 1

注:粗体数字表示最优值

仅为 SwinIR 的 14%。与参数量及 MACs 相近的 ELAN 和 CRAFT 相比,本文模型运行时间分别下降了 6% 和 78%。与 IMDN 相比,虽然 MACs 有所上升,但平均运行时间下降了 11%。

为了进一步验证本文所提模型在细节信息获取方面的先进性,引入 LAM(Local Attribution Map)可视化分析图片结果,具体细节如图 11 所示。LAM 是一种用于超分辨率重建任务中的

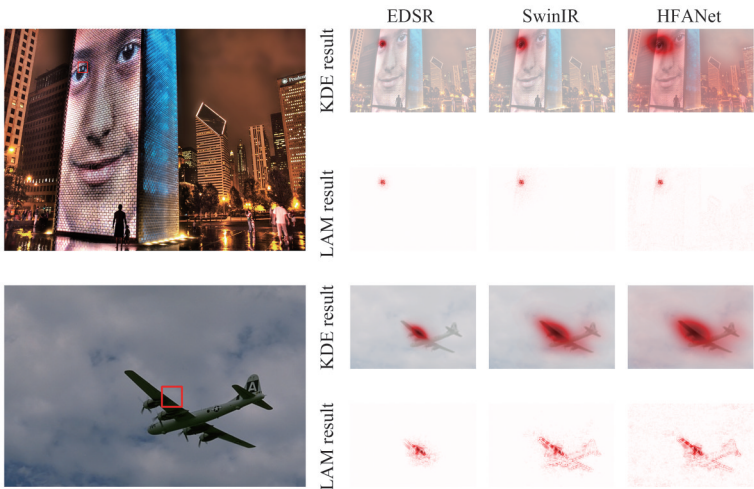


图 11 LAM 可视化效果展示图
Fig. 11 Visual comparison of LAM

可解释性工具,可以定位低分辨率输入图像对重建结果贡献最大的像素。

将本文模型与基于CNN方法的EDSR以及基于注意力机制的方法SwinIR进行对比,由图11可知,EDSR的显著性区域高度集中,表明其感受野受限;SwinIR的区域有所扩大,但受限于窗口注意力机制,呈现出局部化现象。相比之下,本文所提模型在LAM可视化分析中展现出更广泛的红色区域,说明本文模型能够捕获更多的长距离信息,获得了更大的感受野。

综上所述,本文模型在保持轻量化的前提下,实现了重建质量与推理速度的提升。同时,LAM可视化分析说明本文模型所提的HFAB模块能

够使更广泛的像素参与到图像重建过程中,在恢复复杂纹理和结构细节上展现出更优性能。

3.4 消融实验

为了研究不同组件对模型整体性能的影响,对本文所提模型进行消融实验。为确保实验的公平性,所有消融实验均在 $\times 4$ 放大尺度下的Set14和BSD100数据集进行测试。

为了分析每个模块的有效性,本文系统地评估了HFAB模块中各个组件的有效性:双路径增强模块(DPE)、空间细化模块(SRM)和非对称频率感知注意力(AFA)。以完整的HFANet模型作为基线模型(实验4),并依次移除一个组件,实验结果如表5所示。

表5 HFAB模块核心组件的消融实验

Tab. 5 Ablation study on the core components of the HFAB module

Exp.	DPE	SRM	AFA	Params/k	Set14	BSD100
					PSNR/SSIM	PSNR/SSIM
1	×	✓	✓	525	28.87/0.787 8	27.75/0.742 6
2	✓	×	✓	668	28.87/0.787 7	27.75/0.742 7
3	✓	✓	×	673	28.86/0.787 7	27.74/0.742 7
4	✓	✓	✓	708	28.96/0.789 7	27.81/0.744 0

注:粗体数字表示最优值

在实验1中,移除DPE模块中的所有组件。与完整的基线模型(实验4)相比,该变体模型的性能明显下降。在Set14和BSD100上的PSNR和SSIM平均下降了0.07 dB和0.001 7。由可视化结果(图12)可以看出,图像中高频边缘出现了明显的模糊现象,如斑马腿部的条纹边缘。这表明DPE模块通过频率解耦,利用拉普拉斯算子分支对于高频信息的提取,可以有效避免平滑信息对高频细节的干扰,对提升边缘信息的清晰度十分重要。

在实验2中,移除了SRM模块,将DPE的输出直接输入到AFA中。尽管模型参数量有所下降,但其重建性能远低于基线模型。由图12可知,缺失SRM后,模型在处理纹理细节上出现了伪影,这证明了在进行全局信息查询之前,SRM利用可变形卷积对局部特征进行细化处理,能够为注意力机制提供一个信息更丰富的查询(Query),从而显著提升对非刚性纹理的建模能力。缺乏这一步细化后,会导致注意力机制在查询全局信息

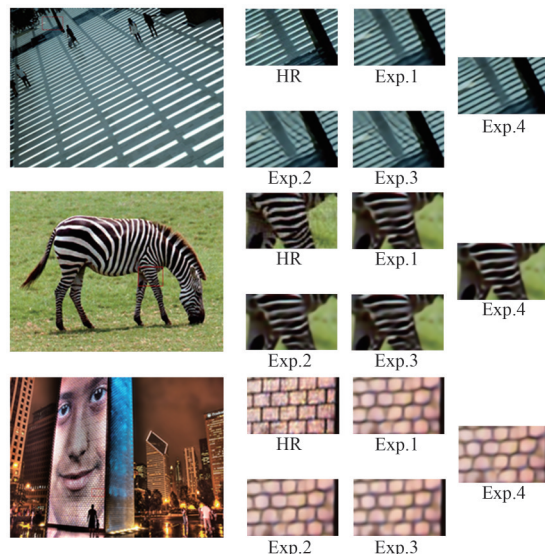


图12 消融实验重建效果展示图

Fig. 12 Visual comparison of the ablation study

时出现偏差,难以匹配到相关的纹理特征。

在实验3中移除了非AFA模块。与基线模型相比,该实验结果在Set14和BSD100上的PSNR

分别下降了 0.10 dB 和 0.07 dB。这是因为 AFA 可通过全局信息检索,利用图像自身的相似性补充局部细节。缺乏这一机制后,模型退化为纯局部网络,无法通过全局信息恢复受损严重的纹理。

为了直观地了解每个核心组件对视觉重建的贡献,本文选择 3 组消融实验可视化结果(图 12),样本分别来自 BSD100、Set14 和 Urban100 数据集。从结果中可以观察到,任何一个核心组件的移除都会导致重建质量的显著下降。而完整模型可以合成最明显、最清晰的线条轮廓,并输出最接近高分辨率参考的视觉效果。

上述可视化实验也为定量实验提供了强有力支持,即 DPE、SRM 和 AFA 在 HFANet 模块中都起着非常重要的作用,它们的合作对于高质量的图像重建是不可或缺的。

为了说明高频滤波器自适应学习的必要性,将高频滤波器的权重固定为初始的拉普拉斯算子,在训练过程中不进行更新。如表 6 中实验 1 所示,相比于实验 3(基线模型,即本文所提模型),在 Set14 和 BSD100 测试集上,其 PSNR 分别下降了 0.09 dB 和 0.06 dB,SSIM 分别下降了 0.001 9 和 0.001 5。这有力地说明,虽然初始化的拉普拉斯算子可以提供一个良好的初始化先验,但是训练过程中禁止网络更新高频滤波器,相当于禁止网络根据数据分布对高频滤波器进行微调,导致性能变差。

表 6 HFAB 模块内部关键设计的消融实验

Tab. 6 Ablation study on key internal designs of the HFAB module

Exp.	Params/k	Set14	BSD100
		PSNR/SSIM	PSNR/SSIM
1	706	28.87/0.787 8	27.75/0.742 5
2	708	28.87/0.788 1	27.74/0.742 9
3	708	28.96/0.789 7	27.81/0.744 0

注:粗体数字表示最优值

为了验证非对称输入设计在 AFA 中的有效性,交换了 Q 和 K、V 的生成源:使用全局上下文生成 Q,并使用精细的局部细节生成 K、V。如表 6 实验 2 可知,相比于实验 3,在 Set14 和 BSD100 测试集上,其 PSNR 分别下降 0.09 dB 和 0.07 dB,SSIM 分别下降 0.001 6 和 0.001 1。该结果充分说明

了本文所设计的利用最具信息量的局部特征查询全局信息库的设计是有效的。

为了证明 SRM 中使用的条状可变形卷积的有效性,设计了 3 个对比实验,实验结果如表 7 所示,用 $\text{conv1} \times 5/5 \times 1$ (实验 1)、 $\text{conv1} \times 7/7 \times 1$ (实验 2)和 $\text{conv1} \times 3/3 \times 1$ (实验 3)代替条状可变形卷积。基线模型(实验 4)在所有配置中取得了最佳性能,与最接近的 $\text{conv1} \times 5/5 \times 1$ (实验 1)相比,本文所提模型在 Set14 和 BSD100 上的 PSNR 和 SSIM 平均获得了 0.08 dB 和 0.001 6 提升。这一提升主要归功于可变形卷积的几何适应能力。标准卷积受限于固定的几何网络,难以捕捉形变特征。而本文采用的方向分离的可变形卷积可以根据特征信息调整采样位置,在保持模型参数量不变的前提下增强了模型对纹理特征的重建能力。

表 7 SRM 中卷积核类型与尺寸的消融实验

Tab. 7 Ablation study on the kernel type and size within the SRM

Exp.	Params/k	Set14	BSD100
		PSNR/SSIM	PSNR/SSIM
1	711	28.87/0.787 7	27.75/0.742 8
2	705	28.88/0.787 9	27.74/0.742 5
3	706	28.87/0.787 9	27.75/0.742 6
4	708	28.96/0.789 7	27.81/0.744 0

注:粗体数字表示最优值

4 结 论

针对现有模型面临计算成本高昂、局部高频纹理的感知有限以及关键信息挖掘不充分的问题,本文提出了一种基于高频融合注意力网络(HFANet)用于超分辨率图像重建。同时,设计了双路径增强模块(DPE),通过引入由拉普拉斯算子初始化的可学习高通滤波器,实现了高、低频特征的解耦与自适应增强,有效提升了网络对边缘纹理信息的提取能力。此外,本文采用有方向分离可变形卷积的空间细化模块(SRM)从 DPE 的输出中捕获非刚性局部纹理。结合非对称频率感知注意力机制(AFA),构建了一种用局部细节驱动的全局信息检索过程,可以有效地利用卷积归纳偏置补充自注意力,大幅度提高模型的性能。高频融合注意力模块(HFAB)和局部窗口自注意力(LWSA)的级联结构还可以实现全局

局部信息的协同处理,从而进一步提高重建细节的稳定性。

实验结果表明,本文方法与其他方法相比获得了更优异的性能。在缩放因子为4时,本文所

提方法在 Manga109 数据集上的 PSNR 和 SSIM 平均提高了约 0.71 dB 和 0.008 1,重建图像在纹理细节和边缘信息上更具优势,并且具有更高效的运行时间。

参考文献:

- [1] SHI W Z, CABALLERO J, LEDIG C, *et al.* Cardiac image super-resolution with global correspondence using multi-atlas PatchMatch [C]//*Proceedings of the 16th International Conference on Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*. Nagoya: Springer, 2013: 9-16.
- [2] 刘郭琦,刘进锋,朱东辉. 基于生成对抗网络的图像超分辨率重建算法[J]. 液晶与显示, 2021, 36(12): 1720-1727.
LIU G Q, LIU J F, ZHU D H. Image super resolution reconstruction algorithm based on generative countermeasure network [J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(12): 1720-1727. (in Chinese)
- [3] WEN W L, REN W Q, SHI Y H, *et al.* Video super-resolution via a spatio-temporal alignment network [J]. *IEEE Transactions on Image Processing*, 2022, 31: 1761-1773.
- [4] 贺兴,王磊,张鹏超,等. 基于多维注意力网络的图像超分辨率重建[J]. 液晶与显示, 2025, 40(7): 1056-1066.
HE X, WANG L, ZHANG P C, *et al.* Image super-resolution reconstruction based on multidimensional attention network [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(7): 1056-1066. (in Chinese)
- [5] JIAN M W, LAM K M. Simultaneous hallucination and recognition of low-resolution faces based on singular value decomposition [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2015, 25(11): 1761-1772.
- [6] JIAN M W, LAM K M, DONG J Y. A novel face-hallucination scheme based on singular value decomposition [J]. *Pattern Recognition*, 2013, 46(11): 3091-3102.
- [7] JIAN M W, LAM K M, DONG J Y. Facial-feature detection and localization based on a hierarchical scheme [J]. *Information Sciences*, 2014, 262: 1-14.
- [8] JIAN M W, LAM K M. Face-image retrieval based on singular values and potential-field representation [J]. *Signal Processing*, 2014, 100: 9-15.
- [9] 陈豪,夏振平,程成,等. 基于Transformer-CNN的轻量级图像超分辨率重建网络[J]. 计算机应用, 2024, 44(1): 292-299.
CHEN H, XIA Z P, CHENG C, *et al.* Lightweight image super-resolution reconstruction network based on Transformer-CNN [J]. *Journal of Computer Applications*, 2024, 44(1): 292-299. (in Chinese)
- [10] DONG C, LOY C C, HE K M, *et al.* Image super-resolution using deep convolutional networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295-307.
- [11] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas: IEEE, 2016: 1646-1654.
- [12] LIM B, SON S, KIM H, *et al.* Enhanced deep residual networks for single image super-resolution [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Honolulu: IEEE, 2017: 1132-1140.
- [13] 陈纯毅,吴欣怡,胡小娟,等. 多尺度注意力融合的图像超分辨率重建[J]. 中国光学(中英文), 2023, 16(5): 1034-1044.
CHEN C Y, WU X Y, HU X J, *et al.* Image super-resolution reconstruction with multi-scale attention fusion [J]. *Chinese Optics*, 2023, 16(5): 1034-1044. (in Chinese)
- [14] 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 (ACMMM)*. Nice: ACM, 2019: 2024-2032.
- [15] 陈清江,陈鹏民. 多维度聚合Transformer的图像超分辨率重建[J]. 光学精密工程, 2025, 33(12): 1955-1970.

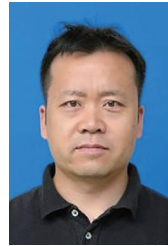
- CHEN Q J, CHEN P M. MDAT: multi-dimensional aggregation Transformer for image super-resolution reconstruction [J]. *Optics and Precision Engineering*, 2025, 33(12): 1955-1970. (in Chinese)
- [16] LIANG J Y, CAO J Z, SUN G L, *et al.* SwinIR: image restoration using swin transformer [C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*. Montreal: IEEE, 2021: 1833-1844.
- [17] LU Z S, LI J C, LIU H, *et al.* Transformer for single image super-resolution [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. New Orleans: IEEE, 2022: 456-465.
- [18] 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 (ECCV)*. Tel Aviv: Springer, 2022: 649-667.
- [19] WANG H, CHEN X H, NI B B, *et al.* Omni aggregation networks for lightweight image super-resolution [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver: IEEE, 2023: 22378-22387.
- [20] ZAMIR S W, ARORA A, KHAN S, *et al.* Restormer: efficient Transformer for high-resolution image restoration [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans: IEEE, 2022: 5718-5729.
- [21] TIMOFTE R, AGUSTSSON E, VAN GOOL L, *et al.* NTIRE 2017 challenge on single image super-resolution: methods and results [C]//*Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Honolulu: IEEE, 2017: 1110-1121.
- [22] BEVILACQUA M, ROUMY A, GUILLEMOT C, *et al.* Low-complexity single-image super-resolution based on nonnegative neighbor embedding [C]//*Proceedings of 2012 British Machine Vision Conference*. Surrey: BMVA Press, 2012: 1-10.
- [23] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations [C]//*Proceedings of the 7th International Conference on Curves and Surfaces*. Avignon: Springer, 2012: 711-730.
- [24] ARBELÁEZ P, MAIRE M, FOWLKES C, *et al.* Contour detection and hierarchical image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(5): 898-916.
- [25] HUANG J B, SINGH A, AHUJA N. Single image super-resolution from transformed self-exemplars [C]//*Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015: 5197-5206.
- [26] MATSUI Y, ITO K, ARAMAKI Y, *et al.* Sketch-based manga retrieval using manga109 dataset [J]. *Multimedia Tools and Applications*, 2017, 76(20): 21811-21838.
- [27] AHN N, KANG B, SOHN K A. Fast, accurate, and lightweight super-resolution with cascading residual network [C]//*Proceedings of the 15th European Conference on Computer Vision (ECCV)*. Munich: Springer, 2018: 256-272.
- [28] LUO X T, XIE Y, ZHANG Y L, *et al.* LatticeNet: towards lightweight image super-resolution with lattice block [C]//*Proceedings of the 16th European Conference on Computer Vision (ECCV)*. Glasgow: Springer, 2020: 272-289.
- [29] LI W B, ZHOU K, QI L, *et al.* LAPAR: linearly-assembled pixel-adaptive regression network for single image super-resolution and beyond [C]//*Proceedings of the 34th Conference and Workshop on Neural Information Processing Systems (NeurIPS)*. Vancouver: Curran Associates Inc., 2020: 1708.
- [30] SUN B, ZHANG Y L, JIANG S Y, *et al.* Hybrid pixel-unshuffled network for lightweight image super-resolution [C]//*Proceedings of the 37th AAAI Conference on Artificial Intelligence*. Washington, D. C. : AAAI, 2023: 2375-2383.
- [31] ZHOU X Q, HUANG H B, HE R, *et al.* MSRA-SR: image super-resolution Transformer with multi-scale shared representation acquisition [C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. Paris: IEEE, 2023: 12665-12676.
- [32] LI A, ZHANG L, LIU Y, *et al.* Feature modulation transformer: cross-refinement of global representation via high-frequency prior for image super-resolution [C]//*Proceedings of the IEEE/CVF International Conference on*

- Computer Vision (ICCV)*. Paris: IEEE, 2023: 12480-12490.
- [33] WANG Y, LIU Y, ZHAO S J, *et al.* CAMixerSR: only details need more “Attention” [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE, 2024: 25837-25846.
- [34] ZHOU Y P, LI Z, GUO C L, *et al.* SRFormer: permuted self-attention for single image super-resolution [C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. Paris: IEEE, 2023: 12734-12745.
- [35] GAO G W, WANG Z X, LI J C, *et al.* Lightweight bimodal network for single-image super-resolution via symmetric CNN and recursive transformer [C]//*Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*. Vienna: IJCAI, 2022: 913-919.
- [36] FAN W S, WANG Y, WANG C, *et al.* Semantic-guided global-local collaborative networks for lightweight image super-resolution [J]. *IEEE Transactions on Instrumentation and Measurement*, 2025, 74: 5026014.

作者简介:



郑龙光,男,硕士研究生,2022年于重庆理工大学获得学士学位,主要从事图像处理方面的研究。E-mail: zhenglongguang23@mails.ucas.ac.cn



朱波,男,博士,正高级工程师,2014年于中国科学院大学获得博士学位,主要从事可见光信息获取、光学遥感等方面的研究。E-mail: irico_zb@qq.com