



计算机科学

COMPUTER SCIENCE

基于频率驱动的多尺度图像超分辨率方法

杨红菊, 张子扬, 李尧

引用本文

杨红菊, 张子扬, 李尧. 基于频率驱动的多尺度图像超分辨率方法[J]. 计算机科学, 2026, 53(5): 218-227.

YANG Hongju, ZHANG Ziyang, LI Yao. Frequency Driven Multi-scale Image Super-resolution Method [J]. Computer Science, 2026, 53(5): 218-227.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[面向视觉Transformer的对抗样本攻击综述](#)

Survey of Adversarial Sample Attacks for Vision Transformer

计算机科学, 2026, 53(5): 404-418. <https://doi.org/10.11896/jsjcx.250600065>

[联合小波分析与频域注意力的高精度人体姿态估计](#)

High-accuracy Human Pose Estimation Combining Wavelet Analysis and Frequency-Domain Attention

计算机科学, 2026, 53(5): 228-236. <https://doi.org/10.11896/jsjcx.250800025>

[融合AEMD与趋势交叉注意力的多尺度Transformer油价预测框架](#)

Multi-scale Transformer Oil Price Prediction Framework with AEMD and Trend Cross-attention

计算机科学, 2026, 53(5): 157-163. <https://doi.org/10.11896/jsjcx.250900086>

[融合图信息瓶颈与Transformer的时序知识图谱推理方法](#)

Enhancing Temporal Knowledge Graph Reasoning Method with Graph Information Bottleneck and Transformer

计算机科学, 2026, 53(4): 393-405. <https://doi.org/10.11896/jsjcx.250400050>

[基于Transformer架构的RNA二级结构预测方法](#)

Prediction Method of RNA Secondary Structure Based on Transformer Architecture

计算机科学, 2026, 53(3): 375-382. <https://doi.org/10.11896/jsjcx.250100005>

基于频率驱动的多尺度图像超分辨率方法

杨红菊 张子扬 李尧

山西大学计算机与信息技术学院 太原 030006

摘要 单图像超分辨率是一种旨在从单张低分辨率图像重建高分辨率图像的关键技术,被广泛应用于医学影像增强、卫星遥感、安防监控及数字媒体内容优化等领域。其核心在于恢复图像丢失的高频细节,以提升视觉质量。然而,该任务因其病态逆问题的本质而面临诸多挑战:传统插值方法难以重建复杂细节;基于卷积神经网络的方法虽能提取局部特征,但在全局建模上不足;而 Transformer 虽擅长长程依赖捕捉,却对高频信息细化能力有限,导致重建图像边缘模糊,纹理失真。为此,提出了一种创新的模型,通过引入小波变换实现多尺度图像分解,优化高频信息提取,并设计三大核心模块:小波细化模块,增强高频细节处理;移位矩形特征增强模块,捕获全局上下文;多尺度小波融合模块,整合高频先验与全局特征。该方法显著提升了纹理和边缘清晰度,兼顾了局部细节与整体一致性。实验结果表明,该模型在 Set5, Set14 和 BSD100 等基准数据集上超越了现有技术,平均峰值信噪比提升约 0.3 dB,主观视觉质量亦更优。该研究不仅有效解决了高频信息恢复难题,还为单图像超分辨率领域提供了新思路,具有重要的学术与应用价值。

关键词: 单图像超分辨率;Transformer;卷积神经网络;小波变换;高频信息优化;多尺度处理

中图分类号 TP183

Frequency Driven Multi-scale Image Super-resolution Method

YANG Hongju, ZHANG Ziyang and LI Yao

School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China

Abstract Single-image super-resolution is a critical technique aimed at reconstructing high-resolution images from single low-resolution inputs. It plays a pivotal role in various fields such as medical image enhancement, satellite remote sensing, security surveillance, and digital media content optimization. The core of SISR lies in restoring the lost high-frequency details to improve visual quality. However, this task faces numerous challenges due to its ill-posed inverse problem nature: traditional interpolation methods struggle to reconstruct complex details; convolutional neural network-based approaches can extract local features but fall short in global modeling; while Transformers excel at capturing long-range dependencies, they have limited capabilities for refining high-frequency information, leading to blurred edges and distorted textures in reconstructed images. To address these issues, this study proposes an innovative model that leverages wavelet transformation for multi-scale image decomposition, optimizing the extraction of high-frequency information. Three key modules are designed: the wavelet refinement module enhances the processing of high-frequency details; the shifted rectangular feature enhancement module captures global context; and the multi-scale wavelet fusion module integrates high-frequency priors with global features. This method significantly improves texture clarity and edge sharpness, balancing both local details and overall consistency. Experimental results demonstrate that the proposed model outperforms existing techniques on benchmark datasets such as Set5, Set14, and BSD100, achieving an average peak signal-to-noise ratio improvement of approximately 0.3 dB, along with superior subjective visual quality. This research not only effectively tackles the challenge of high-frequency information recovery but also provides new insights into the field of single-image super-resolution, holding significant academic and practical value.

Keywords Single-image super-resolution, Transformer, Convolutional neural networks, Wavelet transform, High-frequency information optimization, Multi-scale processing

1 引言

是计算机视觉和图像处理领域的一个重要研究方向,旨在从低分辨率(Low-Resolution, LR)图像恢复出高分辨率(High-Resolution, HR)图像,同时保持图像的细节和视觉质量。

单图像超分辨率(Single-Image Super-Resolution, SISR)

到稿日期:2025-07-08 返修日期:2025-09-26

基金项目:国家自然科学基金(62376145);山西省自然科学基金(202303021211024);国家自然科学基金重点项目(U24A20323)

This work was supported by the National Natural Science Foundation of China(62376145), Natural Science Foundation of Shanxi Province, China(202303021211024) and Key Project of the National Natural Science Foundation of China(U24A20323).

通信作者:杨红菊(yhju@sxu.edu.cn)

SISR 在很多领域中都有比较广阔的应用前景,如视频流媒体^[1]、视频监控^[2],以及医学成像^[3]和卫星遥感^[4]领域。尽管它拥有比较高的应用价值,但是该领域还存在着很多有待解决的问题:传统方法难以有效捕捉高频细节,导致重建图像的边缘和纹理模糊,影响整体清晰度;部分算法无法高效整合多尺度信息,在一定程度上限制了重建效果。

在早期,随着数字图像处理的发展,基于插值的方法成为单图像超分辨率的主流技术。这些方法通过数学公式在低分辨率图像中插入额外的像素来提高分辨率。一些代表性的方法包括最近邻插值^[5]、双线性插值^[6]和双三次插值^[7]。

近年来,深度学习的快速发展推动了超分辨率技术的突破。卷积神经网络(Convolutional Neural Network, CNN)成为一些研究人员喜欢使用的方法。例如, VDSR^[8]通过使用非常深的网络结构来提升超分辨率的性能; SRResNet^[9]创造性地引入了残差学习机制,解决了在训练过程中所面临的梯度消失等棘手问题。但是,其仍然存在着一些局限性,尤其在处理大尺度的图像内容时,由于 CNN 的感受野是有限的,它难以捕捉图像中的长距离依赖关系,导致重建的图像可能在全局结构上存在不一致性,尤其是在处理重复纹理或长距离边缘时。

为了解决上述问题,一些基于 Transformer^[10]架构的模型逐渐出现在人们的视野中,并且在很多实验中被证明具有捕捉图像长距离依赖关系的有效性。尽管取得了这些进步,但这些工作主要集中于设计 Transformer 模块以获得全局信息,而忽略了图像频率方面的把控; Transformer 通常在单一尺度上处理图像,难以有效捕捉多尺度信息,这可能导致在恢复图像的全局结构和局部细节时出现不足。为了解决这些问题,引入小波变换^[11]方法。小波变换的多尺度特性使得模型能够同时处理图像的全局和局部特征,这种多尺度特征表示可以更好地模拟人类视觉系统对图像的感知方式,从而提高超分辨率的视觉质量。针对单图像超分辨率任务中细节、纹理和边缘信息的优化需求,本文提出了一种将小波变换融入 Transformer 结构的新模型,为图像处理领域提供了一种创新且有效的解决方案。具体来说,该模型主要包括 3 个部分,即小波细化模块(WRB)、移位矩形特征增强模块(SRFEB)和多尺度小波融合模块(MWFB),这 3 个模块协同工作,共同实现了图像的多尺度^[12]处理,能更好地捕获图像的高频信息,从而获得效果较好的结果。本文的主要贡献包括以下几点。

1)提出了一种全新的单图像超分辨率重建网络架构。该架构在性能上显著优于传统模型,能够更有效地恢复图像细节并提升整体重建质量。

2)提出了一个小波细化模块。通过提取多尺度图像特征,该模块能够有效处理高频细节和多层次特征,从而生成细节更丰富的重建图像。

3)提出了一个多尺度小波融合模块。该模块将小波细化模块的输出与移位矩形特征增强模块的输出进行深度融合。

2 相关工作

2.1 图像超分辨率

近年来,基于 CNN 的图像超分辨率方法相较于传统方

法展现出显著优势,能够更有效地提升算法性能。Dong 等提出的 SRCNN^[13]开创性地将 3 个卷积层引入图像超分辨率任务中,为后续研究奠定了基础。此后,众多研究通过深化网络架构进一步提升性能。Kim 等提出的 VDSR 引入了残差学习机制,成功构建了包含 20 层的深度网络。之后, Lim 等通过简化残差块设计,使得网络深度更深,进一步提升了性能。Zhang 等提出的 RCAN^[14]引入了残差中的残差结构,实现了超过 400 层的深度网络训练。此外,众多研究还探索了空间和通道注意力机制,以改善重建质量。例如,2019 年 Hui 等的工作^[15]、2020 年 Liu 等^[16]的研究,以及 2020 年 Zhou 等^[17]的成果,均在此方面做出了贡献。2020 年, Niu 等提出的 HAN^[18]引入了层注意力模块和通道空间注意力机制,进一步提升了超分辨率效果。

尽管基于 CNN 的方法在图像超分辨率领域取得了显著成果,但它们在全局信息建模方面仍存在局限性,难以有效捕捉图像的长距离依赖关系。

2.2 基于 Transformer 的单图像超分辨率

Transformer^[19]架构在自然语言处理领域取得了巨大成功后,开始被引入到图像超分辨率任务中。Liang 等提出的用于图像恢复的强基线模型 SwinIR^[20]首次将 Swin Transformer 与 CNN 结合,显著提升了超分辨率的性能。由微软和中国科学技术大学研究团队提出的 CSwin Transformer^[21]是一种通用视觉 Transformer,通过在水平和垂直方向上并行计算自注意力,形成十字形窗口,从而在较小的计算量下实现了更大的感受野。这种设计,使得模型能够更高效地捕捉全局信息。Chen 等提出的 HAT^[22]网络结合了通道注意力^[23]和基于窗口的自注意力机制,利用两者的互补优势激活了更多输入像素,提高了重建质量。CRAFT^[24]模型考虑到结合高频先验的重要性,通过结合 CNN 的高频信息提取能力和 Transformer 的全局信息捕捉能力,有效弥补了 Transformer 在高频信息处理上的不足,并通过混合融合机制优化了特征表示。

2.3 结合小波变换的单图像超分辨率

小波变换在图像超分辨率领域的应用背景主要源于其多尺度表示和细节恢复的能力。图 1 展示了小波变换分解图像示例^[25]。将小波变换与深度学习技术相结合,可以更好地解决图像超分辨率中的高频信息缺失和多尺度信息融合问题,同时降低计算复杂度。这种结合为图像超分辨率任务提供了新的思路和方法,具有重要的研究和应用价值。

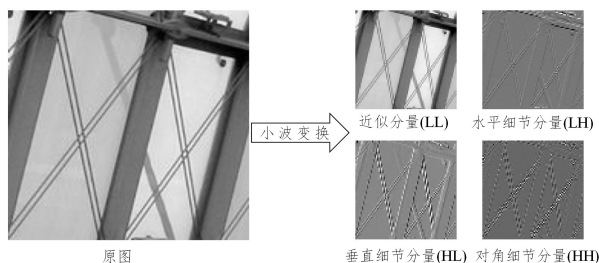


图 1 小波变换

Fig. 1 Wavelet transform

Duan 等提出了一种基于小波域的深度残差网络^[26],不同于其他传统的卷积神经网络直接推导高分辨率图像,该方

法采用多阶段学习策略,首先推理出高分辨率图像对应的小波系数,然后重建超分辨率图像。Huang 等提出了一种基于小波变换的多尺度神经网络 WaveletSRNet^[27],首先将输入图像分解为不同尺度的小波系数,并分别对每个尺度进行超分辨率处理,最后通过小波重构生成高分辨率图像。这种方法有效地利用了小波变换的多尺度特性,提升了超分辨率的性能。Ma 等提出了一种结合小波变换和 SRFeat 网络的单图像超分辨率方法^[28],首先将低分辨率图像输入 SRFeat 网络进行初步超分辨率重建,获得粗略的高分辨率图像,然后利用小波变换对低分辨率图像和粗略高分辨率图像进行特征融合,最终通过逆小波变换得到最终的高分辨率图像。Dharejo 等提出了一种结合小波变换和迁移生成对抗网络的时空单图像超分辨率方法^[29]。具体而言,该方法通过小波变换将低分辨率图像分解为多个频率带,利用迁移 GAN 在不同频率带

中预测高频成分,最后通过逆小波变换重构高分辨率图像。这种方法解决了传统超分辨率方法中高频细节缺失和纹理特征不足的问题,同时通过迁移学习减少了计算成本,并提高了纹理信息的恢复能力。

3 本文方法

该网络主要包括 3 个关键的组成部分,即浅层特征提取模块、小波融合残差组(WFRG)和重建模块,如图 2 所示。浅层特征提取模块包括一个卷积层。小波融合残差组由若干个小波融合模块(WFB)组成,每个 WFB 包括 3 种类型的模块:小波细化块(WRB)、移位矩形特征增强块(SRFEB)和多尺度小波融合模块(MWFB)。首先描述整体的结构;然后分别描述 3 个关键的设计,包括 WRB,SRFEB 和 MWFB。

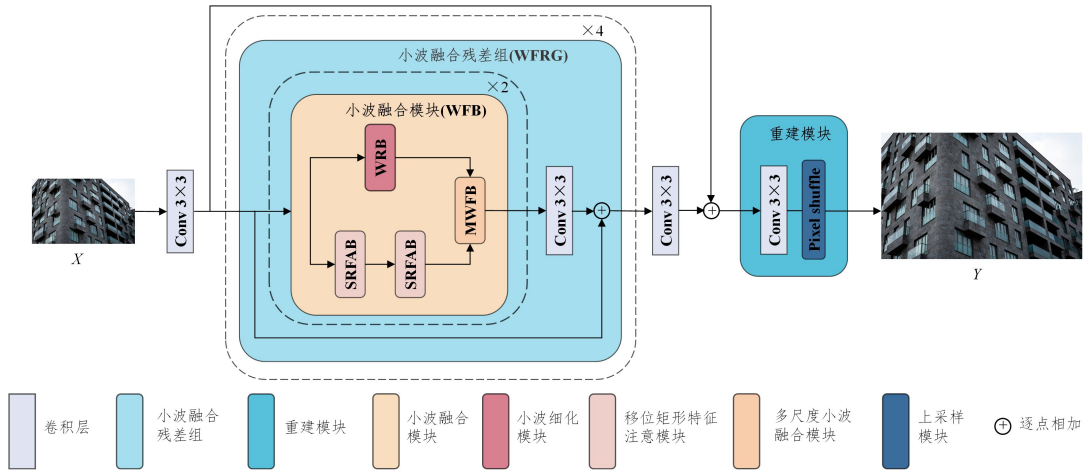


图 2 模型结构

Fig. 2 Model architecture

3.1 模型概述

首先,将 LR 图像输入至一个尺寸为 3×3 的卷积层,利用该卷积层的运算特性提取图像的浅层特征。随后,把所获取的浅层特征馈入一系列精心设计的小波融合残差组架构中,以此深入学习图像的深层特征信息。在最后一个小波融合残差组处理完毕后,引入一个 3×3 的卷积层执行特征聚集操作;同时,在浅层特征提取阶段完成之后,在与最终输出之间构建残差连接,为后续的模型训练流程提供便利,助力模型收敛优化。最终,凭借一个由 3×3 卷积层与混洗层^[30]共同构成的重建模块,生成最终的超分辨率图像。

3.2 小波细化模块

该模块的目的是提取多尺度^[31-33]图像特征,处理高频细节和多层次特征,得到一个细节更丰富的图像。该模块分为两个分支,包括小波变换分支和边缘引导注意力分支,如图 3 所示。具体地,首先将输入的特征分成两部分,分别将其输入到两个分支中进行处理。

$$\mathbf{X}_m^w, \mathbf{X}_m^e = \text{Split}(\mathbf{X}_m) \quad (1)$$

其中, \mathbf{X}_m^w 和 \mathbf{X}_m^e 分别代表小波变换分支和边缘引导注意分支的输入。对于小波变换分支,首先使用一个 3×3 的卷积层对输入进行初步特征提取,之后进入小波变换的过程,通过小波变换将输入图像分解为多层次、多频率的特征表示,将图像

分为 4 个子带。

$$\mathbf{W}\{\mathbf{X}_m^w\} = \{LL, LH, HL, HH\} \quad (2)$$

其中, LL 是低频部分,包含图像的粗略结构; LH 是水平高频部分,包含图像的水平边缘信息; HL 是垂直高频部分,包含图像的垂直边缘信息; HH 是对角线高频部分,包含图像的对角边缘信息。小波变换的输出 \mathbf{W} 被重塑成一个形状为 $(B, C \times 4, H, W)$ 的张量。

$$\mathbf{W}_{\text{rescaled}} = \text{Reshape}(\mathbf{W}, (B, C \times 4, H, W)) \quad (3)$$

这里,将原来的 4 个频带的通道数合并到一个维度中。对每个子带应用小波域卷积^[34]提取特征。

$$\mathbf{W}_{\text{conv}} = \text{Conv2d}(\mathbf{W}_{\text{rescaled}}) \quad (4)$$

卷积后的结果 \mathbf{W}_{conv} 会通过尺度调整模块进行尺度调整。这个模块通过乘以一个缩放因子来调整卷积结果的幅度 $\mathbf{W}_{\text{scaled}} = \text{Scale}(\mathbf{W}_{\text{conv}})$,在尺度调整后,卷积结果被恢复回小波变换的原始形状。

$$\mathbf{W}_{\text{final}} = \text{Reshape}(\mathbf{W}_{\text{scaled}}, (B, C, H, W, 4)) \quad (5)$$

之后进入小波重建阶段,将小波分解后的多尺度特征重新合成完整图像,从最高层开始,将处理后的低频和高频特征组合,使用重建滤波器执行逆小波变换 $\mathbf{X}_{\text{recon}} = \text{IWT}(\mathbf{W}_{\text{final}})$ 。最后,将在像素域直接操作的卷积部分与小波域的重建特征结合。

$$\mathbf{X}_{\text{out}}^W = \mathbf{X}_{\text{recon}} + \text{Conv}2d(\mathbf{X}_{\text{in}}^W) \quad (6)$$

对于边缘引导注意分支,将输入的特征 \mathbf{X}_{in}^W 分别通过通道注意力、边缘增强卷积和空间注意力处理,之后再三特征融合。在通道注意力分支的处理中,输入特征 \mathbf{X} 通过全局平均池化得到通道级别的统计信息,通过一个多层感知机生成通道注意力权重。

$$\mathbf{A}_{\text{channel}} = \sigma(\text{MLP}(\text{GAP}(\mathbf{X}))) \quad (7)$$

其中,GAP 是全局平均池化操作,MLP 是多层感知机, σ 是 Sigmoid 函数。将通道注意力权重应用到输入特征图上,得到输出 $\mathbf{X}_{\text{channel}}$ 。

$$\mathbf{X}_{\text{channel}} = \mathbf{X} \odot \mathbf{A}_{\text{channel}} \quad (8)$$

在边缘增强卷积分支处理中,输入的特征 \mathbf{X} 会经过该卷

积得到增强图像边缘信息后的特征 \mathbf{X}_{edge} 。

$$\mathbf{X}_{\text{edge}} = \text{EdgeConv}(\mathbf{X}) \quad (9)$$

其中,EdgeConv 表示一个边缘增强卷积。在空间注意力分支的处理中,输入特征图 \mathbf{X} 通过一个卷积层生成空间注意力图,空间注意力图通过 Sigmoid 函数进行归一化,得到权重。

$$\mathbf{A}_{\text{spatial}} = \sigma(\text{SACConv}(\mathbf{X})) \quad (10)$$

其中, σ 是 Sigmoid 函数,SACConv 表示空间注意力卷积层。将空间注意力权重应用到输入特征图上,得到输出 $\mathbf{X}_{\text{spatial}}$ 。

$$\mathbf{X}_{\text{spatial}} = \mathbf{X} \odot \mathbf{A}_{\text{spatial}} \quad (11)$$

之后,将三者进行特征融合,得到最终的输出 $\mathbf{X}_{\text{out}}^E$ 。

$$\mathbf{X}_{\text{out}}^E = \mathbf{X}_{\text{edge}} + \mathbf{X}_{\text{channel}} + \mathbf{X}_{\text{spatial}} \quad (12)$$

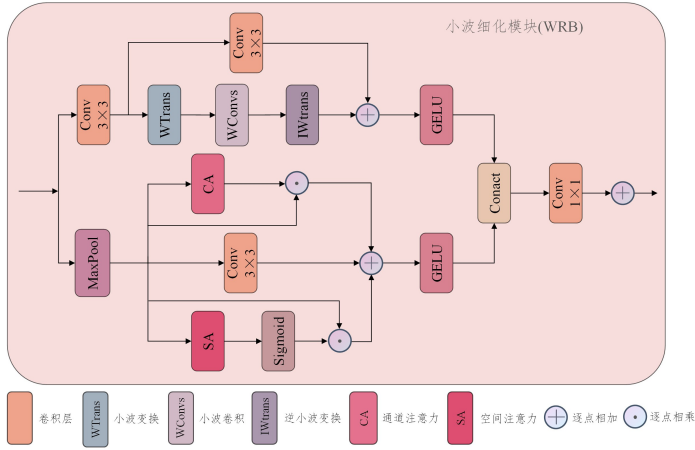


图3 小波细化模块

Fig. 3 Wavelet refine block

3.3 移位矩形特征增强模块

该模块通过引入移位矩形窗口自注意力机制,有效降低了计算复杂度,同时增强了模型对局部特征和上下文信息的捕捉能力。该模块结合了深度可分离卷积和倒置模块,提升了特征表示的丰富性和准确性。如图4所示,给定一个输入 $\mathbf{X}_{\text{in}} \in R^{H \times W \times C}$,将它划分为 $(H \times W) / (rh \times rw)$ 矩形窗口,其中 rh 和 rw 表示矩形窗口的高度和宽度。

对于第 i 个矩形窗口特征 $\mathbf{X}_i \in R^{(rh \times rw) \times C}$,计算查询、键和值。

$$\mathbf{Q}_i = \mathbf{X}_i \mathbf{W}_i^Q, \mathbf{K}_i = \mathbf{X}_i \mathbf{W}_i^K, \mathbf{V}_i = \mathbf{X}_i \mathbf{W}_i^V \quad (13)$$

其中, $\mathbf{W}_i^Q, \mathbf{W}_i^K$ 和 \mathbf{W}_i^V 分别代表投影矩阵。自注意力可以表示为:

$$\text{Attention}(\mathbf{Q}_i, \mathbf{K}_i, \mathbf{V}_i) = \text{Softmax}\left(\frac{\mathbf{Q}_i \mathbf{K}_i^T}{\sqrt{d}} + \mathbf{B}\right) \mathbf{V}_i \quad (14)$$

\mathbf{B} 是动态相对位置编码^[35]。此外,对值进行了深度可分离卷

积和倒置模块的操作,以增强局部提取能力。为了从不同的轴捕获信息,使用水其中,平和垂直窗口两种类型的矩形窗口。将注意力中心分成两个相等的组,并分别计算每组中的自我注意力,然后连接两组的输出以获得最终输出。该过程可以表示为:

$$\text{Rwin-SA}(\mathbf{X}) = \text{Concat}(\mathbf{V-Rwin}, \mathbf{H-Rwin}) \mathbf{W}^P \quad (15)$$

其中, $\mathbf{V-Rwin}$ 和 $\mathbf{H-Rwin}$ 分别表示垂直和水平矩形窗口注意力。在这个过程中,还会把 \mathbf{V} 送入一个深度可分离卷积和倒置模块进行处理,得到补充的局部特征 $\mathbf{L}_{\text{enhanced}}$ 。将经过注意力增强的特征和局部补充模块输出的特征结合,再送入一个多层感知机中进行进一步的特征转换。整个过程可以描述为:

$$\mathbf{X} = \text{Rwin-SA}(\text{LN}(\mathbf{X}_{\text{in}})) + \mathbf{L}_{\text{enhanced}} \quad (16)$$

$$\mathbf{X}_{\text{out}} = \text{MLP}(\text{LN}(\mathbf{X})) + \mathbf{X} \quad (17)$$

其中,LN 代表 LayerNorm 层。

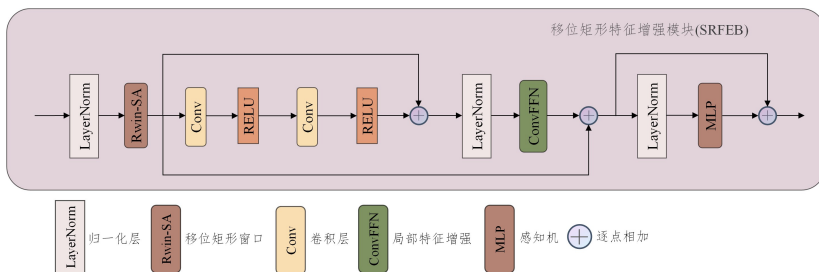


图4 移位矩形特征增强模块

Fig. 4 Shifted rectangle feature enhance block

3.4 多尺度小波融合模块

为了更好地融合小波细化模块的输出和移位矩形特征增强模块的输出全局特征,设计了一个混合块,这个模块更好

地结合了 CNN 和 Transformer 各自的优点。具体来说,如图 5 所示,将小波细化块的输出表示为高频先验查询,移位矩形特征增强模块的输出表示为关键字、值。

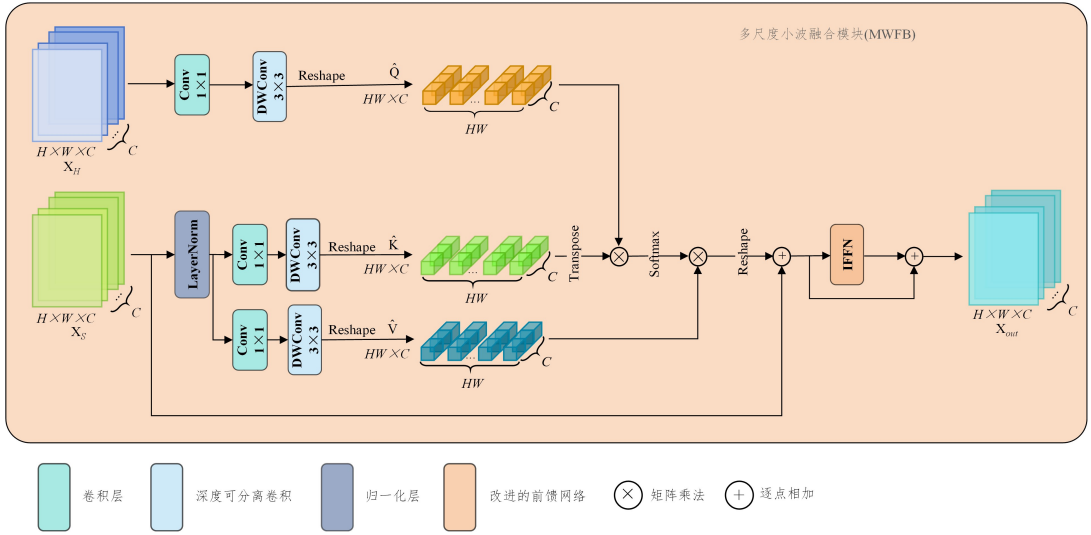


图 5 多尺度小波融合模块

Fig. 5 Multiscale wavelet fusion block

首先,小波细化模块输出的高频先验被一个 1×1 卷积层投影为一个全局特征表示,随后通过 LayerNorm 层对其进行归一化处理,然后对其进行与生成查询 Q 一样的操作,生成关键值 $K \in R^{H \times W \times C}$ 和值 $V \in R^{H \times W \times C}$ 。之后,对三者分别进行重塑操作,得到 $\hat{Q} \in R^{C \times (HW)}$ 、 $\hat{K} \in R^{C \times (HW)}$ 和 $\hat{V} \in R^{C \times (HW)}$ 。计算它们的相互注意力:

$$Attention(\hat{Q}, \hat{K}, \hat{V}) = \text{Softmax}\left(\frac{\hat{Q}\hat{K}^T}{\alpha}\right)\hat{V} \quad (18)$$

在这个过程中, α 是一个可以调整的参数,用于学习。把经过细化的特征整合到 X_S 中,得到一个融合后的结果 X_F 。然后,将 X_F 输入一个经过改进的前馈网络^[33]中,这个网络能够更有效地聚合特征。在这个结构中,引入了一个门控机制,它能够更充分地提取图像的空间和通道信息,从而提高性能。整个过程可以表示为:

$$X_F = IAtt(LN(X_S), X_H) + X_S \quad (19)$$

$$X_{MWFB} = IFFN(X_F) + X_F \quad (20)$$

其中, LN 表示 LayerNorm 层, $IAtt$ 表示计算相互注意力, $IFFN$ 表示一个改进的前馈神经网络。

4 实验及结果分析

4.1 实现细节

在模型架构中,进行如下结构配置:采用 4 个递归上下文小波融合残差组(WFRG)和 2 个小波融合模块(WFB)。其中,每个 WFB 模块通过小波细化块(WRB)与移位矩形特征增强块(SRFEB)的组合实现,具体比例为 1:2。网络参数方面,特征通道数配置为 48,注意力头数设定为 6,标准 MLP 层的扩展比例采用 2 倍缩放,而改进型 MLP(IMLP)的扩展比例则提升至 2.66 倍。针对图像处理需求,特别设计了两种非对称窗口尺寸,分别采用 $[4 \times 16]$ 和 $[16 \times 4]$ 的矩形窗口结构来实现多尺度特征提取。

4.2 数据和评估

本研究选用 DIV2K 和 Flickr2K^[37] 数据集作为训练

基础。在验证环节中,采用 5 个标准评估基准: Set5, Set14, BSD100, Urban100 和 Manga109;并针对 3 种不同缩放比例 ($\times 2, \times 3, \times 4$) 进行多尺度性能测试。图像重建质量通过峰值信噪比(PSNR)和结构相似性指数(SSIM)两项指标进行量化评估,同时采用网络参数总量作为模型复杂度的表征指标。

4.3 参数设置

本研究遵循标准超分辨率训练范式,数据预处理阶段通过双三次插值下采样从原始高分辨率(HR)图像生成对应低分辨率(LR)样本。训练配置方面,从训练图像中随机截取 64×64 像素的局部区域作为输入,整个训练周期设定为 50 万次迭代。为提高数据多样性,采用包含水平翻转和多角度旋转($90^\circ, 180^\circ, 270^\circ$)的复合增强策略。优化过程采用 Adam 算法($\beta_1 = 0.9, \beta_2 = 0.999$)最小化 L1 范数损失,设置 8 的批处理规模并初始化学率为 2×10^{-4} 。学习率动态调整方面,在 $2.5 \times 10^5, 4.0 \times 10^5, 4.5 \times 10^5, 4.75 \times 10^5$ 迭代次数节点实施学习率折半策略。实验环境基于 PyTorch 深度学习框架,在 NVIDIA 4080 GPU 上开展训练。

4.4 对比实验

本研究构建了涵盖多代技术路线的综合对比实验体系,在 5 个基准数据集上,通过定量指标(PSNR/SSIM)与定性评估双重验证模型创新性。实验选取超分辨率领域的 11 项里程碑工作:经典 CNN 架构,包含残差学习代表 EDSR^[37]、轻量化标杆 CARN^[41] 和 IMDN^[40],以及动态网络先驱 LatticeNet^[39];Transformer 革新架构,涵盖局部窗口注意力典范 SwinIR^[20] 和高效混合设计 ESRT^[43];混合架构前沿工作,包括频域增强方法 HPUN^[42]、注意力优化方案 LAPAR^[38],以及最新联合框架 CRAFT^[24]、CAFT^[44] 和 SSIU^[45]。所有对比实验均在 NVIDIA 4080 GPU 环境下严格复现原始训练协议,确保跨模型比较的公平性与可复现性。如表 1—表 3 所列,在单图像超分辨率任务的系统性评估中,所提模型在五大基准数据集上展现出较为突出的性能,将最好性能和次好

性能分别用粗体和下划线标出。具体而言:所提模型相较于传统 CNN 架构 EDSR,在 $\times 2$ 和 $\times 4$ 中实现了 0.83 dB 和 0.85 dB 的 PSNR 增益;针对通道注意力优化方法 CARN,在 $\times 2$ 缩放任务中取得了 0.93 dB 的提升;面对 Transformer 架构的竞争,以 0.29~0.34 dB 优势超越了 SwinIR 和 ESRT 等

前沿方法。值得注意的是,Urbn100 因其独特的场景复杂度,已成为衡量超分辨率模型结构保持能力和高频恢复能力的标准,本文模型在该数据集上表现优异,相较于 CRAFT 提高了 0.23 dB 的 PSNR,说明所提模型在高频细节重建方面和伪影抑制方面有一定的竞争力。

表 1 $\times 2$ 放大因子下的性能比较Table 1 Performance comparison at $\times 2$ magnification factor

放大因子	模型	Set5	Set14	BSD100	Urban100	Manga109
		(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)
$\times 2$	EDSR ^[37]	37.99/0.9604	33.57/0.9175	32.16/0.8994	31.98/0.9272	38.54/0.9769
	LAPAR-A ^[38]	38.01/0.9605	33.62/0.9183	32.19/0.8999	32.10/0.9283	38.67/0.9772
	LatticeNet ^[39]	38.06/0.9607	33.70/0.9187	32.20/0.8999	32.25/0.9288	38.71/0.9773
	IMDN ^[40]	38.00/0.9605	33.63/0.9177	32.19/0.8996	32.17/0.9283	38.88/0.9774
	CARN ^[41]	37.76/0.9590	33.52/0.9166	32.09/0.8978	31.92/0.9256	38.36/0.9765
	HPUN-L ^[42]	38.09/0.9608	33.79/0.9198	32.25/0.9006	32.37/0.9307	39.07/0.9779
	SwinIR-light ^[20]	38.14/0.9611	<u>33.86/0.9201</u>	<u>32.31/0.9012</u>	32.60/0.9340	39.12/0.9783
	ESRT ^[43]	38.03/0.9600	33.75/0.9184	32.25/0.9001	32.58/0.9318	39.12/0.9774
	CRAFT ^[24]	38.17/0.9610	33.82/0.9197	32.27/0.9009	32.61/0.9341	<u>39.21/0.9783</u>
	CFAT ^[44]	38.14/0.9608	33.79/0.9192	32.29/0.9010	<u>32.68/0.9343</u>	<u>39.24/0.9782</u>
	SSIU ^[45]	<u>38.20/0.9611</u>	33.83/0.9203	<u>32.32/0.9011</u>	<u>32.72/0.9343</u>	39.20/0.9781
	Ours	38.22/0.9614	33.87/0.9198	32.34/0.9014	32.84/0.9346	39.29/0.9786

表 2 $\times 3$ 放大因子下的性能比较Table 2 Performance comparison at $\times 3$ magnification factor

放大因子	模型	Set5	Set14	BSD100	Urban100	Manga109
		(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)
$\times 3$	EDSR ^[37]	34.37/0.9270	30.28/0.8417	29.09/0.8052	28.15/0.8527	33.45/0.9439
	LAPAR-A ^[38]	34.36/0.9267	30.34/0.8421	29.11/0.8054	28.15/0.8523	33.51/0.9441
	LatticeNet ^[39]	34.53/0.9281	30.39/0.8424	29.15/0.8059	28.33/0.8538	33.53/0.9445
	IMDN ^[40]	34.36/0.9270	30.32/0.8417	29.09/0.8046	28.17/0.8519	33.61/0.9445
	CARN ^[41]	34.29/0.9255	30.29/0.8407	29.06/0.8034	28.06/0.8493	33.50/0.9440
	HPUN-L ^[42]	34.56/0.9281	30.45/0.8445	29.18/0.8072	28.37/0.8572	33.90/0.9463
	SwinIR-light ^[20]	34.62/0.9289	30.54/0.8463	29.20/0.8082	28.66/0.8624	33.98/0.9478
	ESRT ^[43]	34.42/0.9268	30.43/0.8433	29.15/0.8063	28.46/0.8574	33.95/0.9455
	CRAFT ^[24]	<u>34.67/0.9290</u>	<u>30.58/0.8466</u>	29.22/0.8091	28.70/0.8633	<u>34.24/0.9489</u>
	CFAT ^[44]	34.53/0.9276	30.62/0.8471	29.19/0.8085	<u>28.78/0.8634</u>	34.20/0.9488
	SSIU ^[45]	34.65/0.9287	30.56/0.8462	<u>29.24/0.8092</u>	28.74/0.8636	34.22/0.9490
	Ours	34.71/0.9291	30.57/0.8465	29.26/0.8095	28.81/0.8639	34.27/0.9493

表 3 $\times 4$ 放大因子下的性能比较Table 3 Performance comparison at $\times 4$ magnification factor

放大因子	模型	Set5	Set14	BSD100	Urban100	Manga109
		(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)	(PSNR/SSIM)
$\times 4$	EDSR ^[37]	32.09/0.8938	28.58/0.7813	27.57/0.7357	26.04/0.7849	30.35/0.9067
	LAPAR-A ^[38]	32.15/0.8944	28.61/0.7818	27.61/0.7366	26.14/0.7871	30.42/0.9074
	LatticeNet ^[39]	32.18/0.8943	28.61/0.7812	27.57/0.7355	26.14/0.7844	30.41/0.9063
	IMDN ^[40]	32.21/0.8948	28.58/0.7811	27.56/0.7353	26.04/0.7838	30.45/0.9075
	CARN ^[41]	32.13/0.8937	28.60/0.7806	27.58/0.7349	26.07/0.7837	30.47/0.9084
	HPUN-L ^[42]	32.31/0.8962	28.73/0.7842	27.66/0.7386	26.27/0.7918	30.77/0.9109
	SwinIR-light ^[20]	32.44/0.8976	28.77/0.7858	27.69/0.7406	26.47/0.7980	30.92/0.9151
	ESRT ^[43]	32.19/0.8947	28.69/0.7833	27.69/0.7379	26.39/0.7962	30.75/0.9100
	CRAFT ^[24]	32.49/0.8984	28.83/0.7871	27.71/0.7417	26.54/0.7993	31.17/0.9166
	CFAT ^[44]	<u>32.47/0.8981</u>	28.81/0.7869	27.73/0.7414	26.50/0.7988	31.14/0.9164
	SSIU ^[45]	<u>32.51/0.8982</u>	<u>28.84/0.7871</u>	<u>27.75/0.7418</u>	<u>26.58/0.7990</u>	<u>31.19/0.9168</u>
	Ours	32.53/0.8987	28.88/0.7874	27.77/0.7421	26.62/0.7998	31.29/0.9171

图 6 中展示了目视比较的结果($\times 4$)。在一些极具挑战性的场景中,以往的方法往往会在图像恢复过程中出现模糊伪影、失真或纹理恢复不准确等问题。相比之下,本

文方法能够显著减少伪影的产生,同时更好地保留图像的结构和细节。例如,在 img_091 中,大多数现有方法无法准确恢复真实的地砖纹理,有的方法生成了大量不必要的

伪影,而有的则过于平滑,导致图像细节丢失,不够真实;相比之下,本文方法能够更准确地恢复图像结构,并呈现出清晰且自然的纹理。类似的结果也可以在 img_001 和 img_096 中观察到。这主要归功于本文方法在处理图像时,既能够有效处理高频信息,又能够捕获全局上下文信息,这种能力使得恢复后的图像不仅细节丰富,而且整体更加真实自然。

像里每个像素的显著性,以及红框标注的局部区域经超分辨率重建后的图像之间的相关性。通过充分运用更广泛多元的信息,本文模型取得了更为出色的成果。不仅如此,还对扩散指数(DI)进行了研究,该指数能够反映所涉及像素的范围大小,DI值越大,意味着关注的范围越广。相较于 SwinIR 模型,所提模型拥有更高的 DI 值,因而能够捕捉到更多的上下文信息,为精准重建助力。实验结果有力地证实了所提方法的有效性。

如图 7 所示,局部归因图(LAM)体现的是低分辨率图

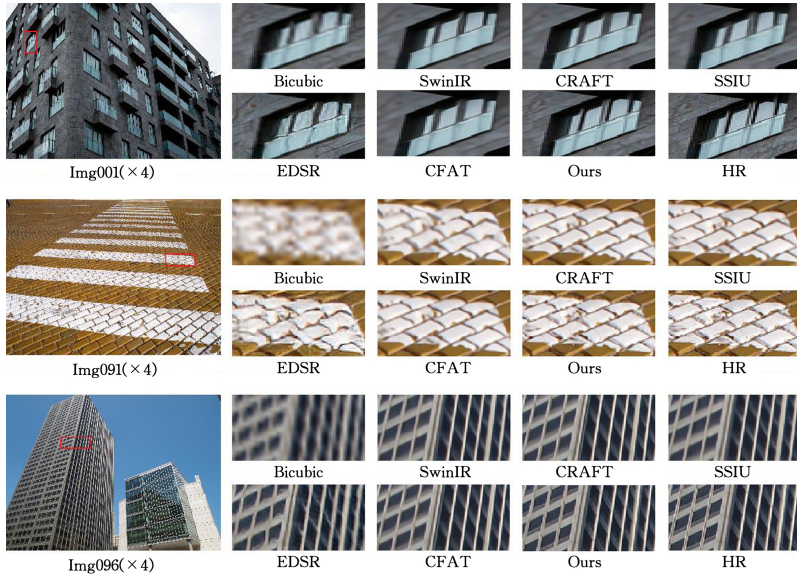


图 6 视觉对比

Fig. 6 Visual comparisons

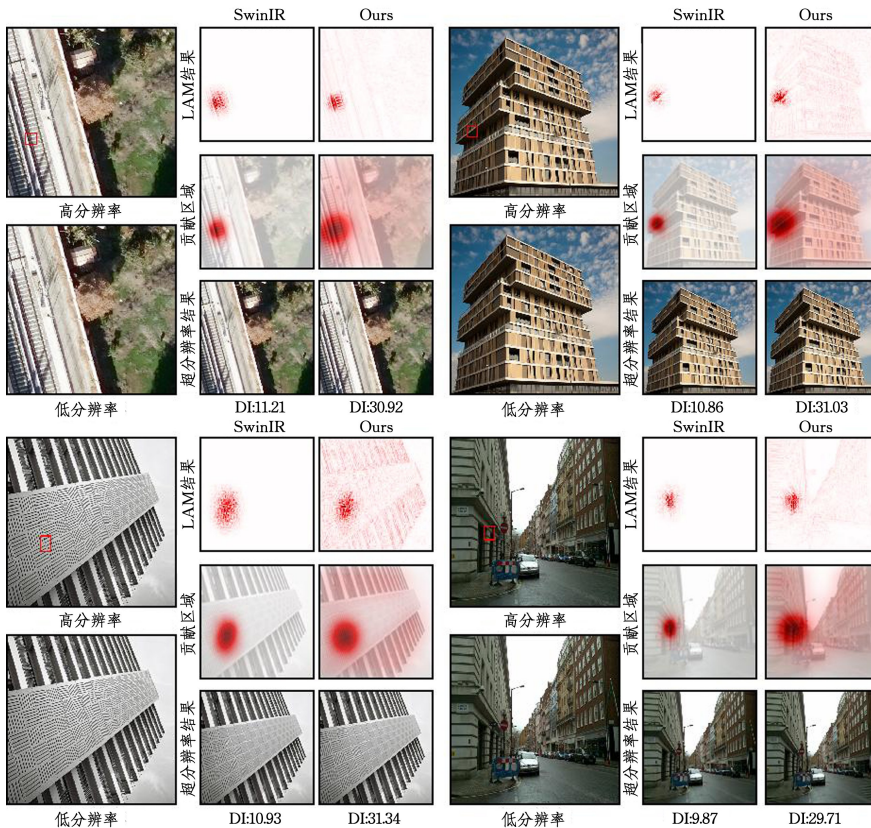


图 7 LAM 结果可视化

Fig. 7 Visualize the results of LAM

此外,图 8 对所提模型和 CRAFT 模型进行上采样之前的最后特征图进行了可视化。可以观察到,所提模型可以获得比 CRAFT 模型更清晰的纹理和边缘,进一步证明了其可以有效增强特征的表达。



图 8 特征图可视化
Fig. 8 Feature map visualization

4.5 消融研究

在数据集 DIV2K 和 Flickr2K 上训练模型,并在 Urban100 上对其进行消融测试。为了公平比较,所有模型都设置相同的实现细节,例如实验参数迭代轮次等。

在 Urban100 数据集($\times 4$)上进行了 4 个消融实验,每个模型的详细性能指标如表 4 所列。通过对比第一行和第四行的结果,可以明显看出移位矩形特征增强模块的有效性。该模块通过增强图像上下文信息的捕捉能力,显著提升了模型性能,具体表现为 PSNR 值提升了 0.21 dB。通过比较第二行和第四行的结果发现,引入小波细化模块后,模型的 PSNR 值得到了 0.11 dB 的提升。此外,该模块还显著改善了图像的细节表现,进一步证明了其在高频信息处理方面的优势。最后,通过对比第三行和第四行的结果,验证了多尺度小波融合模块的有效性。在第三行的模型中,没有采用多尺度小波融合模块来整合小波细化模块和移位矩形特征增强模块的输出,而是通过普通的 3×3 卷积层进行连接。然而,由于这两个模块关注的频率信息不同,直接叠加特征会阻碍网络学习高频和低频分量之间关系的能力。相比之下,多尺度小波融合模块为集成具有不同分布的特征提供了一种高效的解决方案,从而进一步提升了模型的整体性能。

表 4 在数据集 Urban100 上对小波细化模块、移位矩形特征注意模块和多尺度小波融合模块的研究($\times 4$)

Table 4 Research on WRB,SRFAB,and MWFB using the Urban100 dataset($\times 4$)

WRB	SRFEB	MWFB	Concat	PSNR	SSIM
✓		✓		26.41	0.7988
	✓	✓		26.51	0.7992
✓	✓		✓	26.44	0.7990
✓	✓	✓		26.62	0.7998

4.6 复杂性分析

在复杂度的对比研究中,选取了输入大小为 128×128 的样本,来针对所提模型与 SwinIR 和 CRAFT 展开深入分析,复杂性对比如表 5 所列。此次分析全面涵盖了模型参数数量(Params)、浮点运算次数(FLOPs)、GPU 内存占用(GPU Mem)情况以及平均推理时间(Ave Time)这几个关键维度。

表 5 复杂性分析对比

Table 5 Comparative complexity analysis

模型	#Params	#FLOPs	#GPU Mem/ MB	Ave Time/ ms
SwinIR	8.96×10^5	3.17×10^{10}	140.6	71.3
CRAFT	7.74×10^5	2.83×10^{10}	86.1	50.7
Ours	7.62×10^5	2.64×10^{10}	80.6	43.1

表 5 中,GPU 内存的使用量通过官方 PyTorch 函数进行精准测量,而时间成本则是依据 100 次推理运行的结果统计得出。经对比发现,相较于 SwinIR 和 CRAFT,所提模型展现出显著优势;其不仅所需参数更少,浮点运算量更低,而且在 GPU 内存消耗以及推理时间上也实现了大幅缩减。

进一步地,对所提框架的复杂度进行剖析,详细结果如表 6 所列。研究表明,SRFEB 模块对整体复杂度的贡献约为 41%,成为影响复杂度的关键因素之一。与之相比,WRB 模块凭借更少的卷积操作,有效减少了浮点运算次数,进而减轻了计算压力。

表 6 每个模块的复杂性分析

Table 6 Complexity analysis of each block

	w/o WRB	w/o SRFEB	w/o MWFB	总框架
#Params	6.94×10^5	4.46×10^5	5.06×10^5	7.62×10^5
#FLOPs	2.35×10^{10}	1.38×10^{10}	1.96×10^{10}	2.64×10^{10}

表 7 对各模块的耗时进行了计算分析。可以看出,SRFEB 是主要的计算瓶颈,其 48% 的 FLOPs 占比对应 22ms 的耗时,这主要来自窗口注意力计算和深度卷积;MWFB 和 WRB 分别贡献了 26% 和 11% 的 FLOPs,耗时 11ms 和 4ms,主要涉及交叉注意力机制和小波变换操作。

表 7 各模块耗时分析

Table 7 Analysis of module-wise computation time

模块	FLOPs 占比/%	平均推理时间/ms
SRFEB	48	22
MWFB	26	11
WRB	11	4
其他	15	6

结束语 本文提出了一种创新的模型,通过引入小波变换实现多尺度图像分解,并设计了三大核心模块:小波细化模块、移位矩形特征增强模块和多尺度小波融合模块。定量分析显示,在 PSNR 和 SSIM 等多个评价标准中,所提方法均表现出明显优势;定性分析进一步验证了其有效性,重建图像在细节保留、纹理清晰度及边缘锐利度方面均有显著提升;LAM 结果显示,与基线模型相比,所提模型的预测误差更小且稳定,特别是在处理高难度样本时,改进更为显著。

尽管目前方法存在一些不足之处,如处理极端情况下的鲁棒性问题以及计算成本较高等,但未来的研究可以通过改进小波变换算法、增强模型鲁棒性、扩展应用场景等方式进一步完善和发展这一框架。期待在未来能够看到更多基于此框架的新颖研究成果出现,并为图像处理及相关领域带来更多的突破和进步。

参考文献

[1] MUDUNURI S P, BISWAS S. Low resolution face recognition

- across variations in pose and illumination [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 38(5):1034-1040.
- [2] WAN X F, HAN F, JIANG X Q, et al. Research on super-resolution restoration of video surveillance images [J]. *Laser Journal*, 2014, 35(3):5-8.
- [3] GREENSPAN H. Super-resolution in medical imaging [J]. *The Computer Journal*, 2009, 52(1):43-63.
- [4] MENG F, WU S, LI Y, et al. Single remote sensing image super-resolution via a generative adversarial network with stratified dense sampling and chain training [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 62:1-22.
- [5] JIANG N, WANG L. Quantum image scaling using nearest neighbor interpolation [J]. *Quantum Information Processing*, 2015, 14:1559-1571.
- [6] KIRKLAND E J. *Advanced Computing in Electron Microscopy* [M]. Springer, 2010:261-263.
- [7] GAO S K, GRUEV V. Bilinear and bicubic interpolation methods for division of focal plane polarimeters [J]. *Optics Express*, 2011, 19(27):26161-26173.
- [8] KIM J W, 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*. Piscataway, NJ: IEEE, 2016:1646-1654.
- [9] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2017:4681-4690.
- [10] HAN K, WANG Y H, CHEN H T, et al. A survey on vision transformer [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45(1):87-110.
- [11] ZHANG D S. *Wavelet transform* [M]. Cham: Springer, 2019:35-44.
- [12] SUNKAVALLI K, JOHNSON M K, MATUSIK W, et al. Multi-scale image harmonization [J]. *ACM Transactions on Graphics*, 2010, 29(4):1-10.
- [13] DONG C, LOY C C, HE K M, et al. Learning a deep convolutional network for image super-resolution [C]// *Computer Vision-ECCV 2014: 13th European Conference*. Springer, 2014:184-199.
- [14] ZHANG Y L, LI K P, LI K, et al. Image super-resolution using very deep residual channel attention networks [C]// *Proceedings of the European Conference on Computer Vision (ECCV)*. Cham: Springer, 2018:286-301.
- [15] 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*. Cham: Springer, 2019:2024-2032.
- [16] LIU J, TANG J, WU G S. Residual feature distillation network for lightweight image super-resolution [C]// *Proceedings of Computer Vision-ECCV 2020 Workshops*. Cham: Springer, 2020:41-55.
- [17] ZHOU S C, ZHANG J W, ZUO W M, et al. Cross-scale internal graph neural network for image super-resolution [J]. *Advances in Neural Information Processing Systems*, 2020, 33:3499-3509.
- [18] NIU B, WEN W L, REN W Q, et al. Single image super-resolution via a holistic attention network [C]// *Proceedings of Computer Vision-ECCV 2020: 16th European Conference*. Cham: Springer, 2020:191-207.
- [19] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]// *Advances in Neural Information Processing Systems*. 2017.
- [20] 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*. Piscataway, NJ: IEEE, 2021:1833-1844.
- [21] DONG X Y, BAO J M, CHEN D D, et al. Cswin transformer: a general vision transformer backbone with cross-shaped windows [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2022:12124-12134.
- [22] CHEN X Y, WANG X T, ZHOU J T, et al. Activating more pixels in image super-resolution transformer [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2023:22367-22377.
- [23] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module [C]// *Proceedings of the European Conference on Computer Vision (ECCV)*. Cham: Springer, 2018:3-19.
- [24] 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*. Piscataway, NJ: IEEE, 2023:12514-12524.
- [25] MALLAT S G. Multifrequency channel decompositions of images and wavelet models [J]. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 2002, 37(12):2091-2110.
- [26] DUAN L J, WU C L, EN Q, et al. Image Super-Resolution Algorithm of Deep Residual Network Based on Wavelet Domain [J]. *Journal of Software*, 2019, 30(4):941-953.
- [27] HUANG H B, HE R, SUN Z N, et al. Wavelet-srnet: A wavelet-based cnn for multi-scale face super resolution [C]// *Proceedings of the IEEE International Conference on Computer Vision*. Piscataway, NJ: IEEE, 2017:1689-1697.
- [28] MA C Y, ZHU J W, LI Y J, et al. Single image super resolution via wavelet transform fusion and SRFeat network [J]. *Journal of Ambient Intelligence and Humanized Computing*, 2022, 13(11):1-9.
- [29] DHAREJO F A, DEEBA F, ZHOU Y C, et al. Twist-gan: towards wavelet transform and transferred gan for spatio-temporal single image super resolution [J]. *ACM Transactions on Intelligent Systems and Technology*, 2021, 12(6):1-20.
- [30] SHI W Z, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2016:1874-1883.
- [31] WANG W, SUN Y J, WANG X. Lightweight frequency and spa-

- tial feature fused multi-scale remote sensing scene classification network[J]. Journal of Jilin University(Engineering and Technology Edition), 2025, 55(10): 3361-3371.
- [32] 骆迪,张乾. 基于多尺度注意力的图像超分辨研究[J/OL]. 重庆工商大学学报(自然科学版), 1-14[2025-12-08]. <https://link.cnki.net/urlid/50.1155.N.20240926.0921.002>.
- [33] ZHONG M Y, JIANG L, LI C. Image super-resolution reconstruction algorithm based on multi-scale feature extraction residual network[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2024, 36(1): 76-85.
- [34] FINDER S E, AMOYAL R, TREISTER E, et al. Wavelet convolutions for large receptive fields [C]// Proceedings of the European Conference on Computer Vision. Cham: Springer, 2024: 363-380.
- [35] WANG W X, CHEN W, QIU Q B, et al. Crossformer++: A versatile vision transformer hinging on cross-scale attention [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 46(5): 3123-3136.
- [36] 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. Piscataway, NJ: IEEE, 2022: 5728-5739.
- [37] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Piscataway, NJ: IEEE, 2017: 136-144.
- [38] LI W B, ZHOU K, QI L, et al. Lpar: Linearly-assembled pixel-adaptive regression network for single image super-resolution and beyond [J]. Advances in Neural Information Processing Systems, 2020, 33: 20343-20355.
- [39] LUO X T, XIE Y, ZHANG Y L, et al. Latticenet: Towards lightweight image super-resolution with lattice block [C]// Proceedings of Computer Vision-ECCV 2020: 16th European Conference. Cham: Springer, 2020: 272-289.
- [40] HUI Z, GAO X B, YANG Y C, et al. Lightweight image super-resolution with information multi-distillation network [C]// Proceedings of the 27th ACM International Conference on Multimedia. New York: ACM, 2019: 2024-2032.
- [41] 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 (ECCV). Cham: Springer, 2018: 252-268.
- [42] SUN B, ZHANG Y L, JIANG S Y, et al. Hybrid pixel-unshuffled network for lightweight image super-resolution [C]// Proceedings of the AAAI Conference on Artificial Intelligence. Washington, DC: AAAI, 2023: 2375-2383.
- [43] 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. Piscataway, NJ: IEEE, 2022: 457-466.
- [44] RAY A, KUMAR G, KOLEKAR M H. CFAT: Unleashing Triangular Windows for Image Super-Resolution [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 26120-26129.
- [45] NI Z K, ZHANG Y, YANG W H, et al. Structural Similarity-Inspired Unfolding for Lightweight Image Super-Resolution [J]. IEEE Transactions on Image Processing, 2025, 34: 3861-3872.



YANG Hongju, born in 1975, Ph.D, associate professor, master supervisor, is a member of CCF (No. 10371M). Her main research interests include computer vision and deep learning.

(责任编辑:柯颖)