

# 基于多域信息增强的轻量级图像超分辨率网络

寇旗旗<sup>1</sup>, 刘规<sup>2</sup>, 江鹤<sup>2</sup>, 陈亮亮<sup>2</sup>, 程德强<sup>2</sup>

(1. 中国矿业大学计算机科学与技术学院, 江苏 徐州 221116; 2. 中国矿业大学信息与控制工程学院, 江苏 徐州 221116)

**摘要:** 针对单幅图像超分辨率重建任务中单域特征的重建能力受限以及深度卷积神经网络参数众多、计算量大导致的难以部署到移动端的问题, 提出了一种基于多域信息增强的轻量级图像超分辨率网络。从空域、频域和转换域 3 个维度入手, 设计了空域多路大核特征提取技术、局部信息增强注意力、频域分频特征增强技术以及转换域基于高频特征模拟技术。通过不同特征域的信息处理, 针对全局与局部的低频和高频特征进行精准优化, 从而提升模型在细节恢复与图像重建中的表现。与现有先进算法在公认基准数据集上进行充分的实验对比和分析, 结果表明所提网络模型能够实现优异的重建效果, 且在性能与效率之间也实现了出色的平衡。

**关键词:** 计算机视觉; 超分辨率; 多域信息增强; 注意力; 轻量级

中图分类号: TP391.41

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2025059

## Lightweight image super-resolution network based on multi-domain information enhancement

KOU Qiqi<sup>1</sup>, LIU Gui<sup>2</sup>, JIANG He<sup>2</sup>, CHEN Liangliang<sup>2</sup>, CHENG Deqiang<sup>2</sup>

1. School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China

2. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China

**Abstract:** Aiming to solve the problems that the reconstruction capability of single-domain features was limited and deep convolutional neural networks used in existing single-image super-resolution reconstruction tasks were difficult to deploy on mobile terminals due to the large number of parameters and high computational requirements, a lightweight image super-resolution network based on multi-domain information enhancement was proposed. Initiating from three dimensions, a set of innovative techniques had been developed, including multi-path large kernel feature extraction in the spatial domain, local information enhancement attention, frequency-domain feature enhancement through frequency splitting, and transformation-domain prior-guided high-frequency feature simulation. By processing information across different feature domains, both global and local low-frequency and high-frequency features were optimized, significantly improving the model's performance in detail recovery and image reconstruction. Extensive experimental comparisons and analyses with the existing advanced algorithms on the recognized benchmark datasets demonstrate that the proposed network model can achieve remarkable reconstruction results while enjoying a high trade-off between performance and efficiency.

**Keywords:** computer vision, super resolution, multi-domain information enhancement, attention, lightweight

收稿日期: 2024-09-25; 修回日期: 2024-12-11

通信作者: 程德强, chengdq@cumt.edu.cn

基金项目: 国家自然科学基金资助项目 (No.52204177, No.52304182); 济宁市重点研发基金资助项目 (No.2021KJHZ013, No.2023KJHZ007)

**Foundation Items:** The National Natural Science Foundation of China (No.52204177, No.52304182), The Key Research and Development Program of Jining City (No.2021KJHZ013, No.2023KJHZ007)

## 0 引言

单幅图像超分辨 (SISR, single image super-resolution) 重建技术旨在从低分辨率 (LR, low-resolution) 图像重建出其相应的细节丰富、视觉质量更好的高分辨率 (HR, high-resolution) 图像<sup>[1]</sup>。借助 SISR 技术,省去了使用更高端设备拍摄制作的环节,节约了大量的成本。

SISR 的主流方法主要分为三类,分别为基于插值的方法<sup>[2-3]</sup>、基于重建<sup>[4-5]</sup>的方法和基于学习<sup>[6-7]</sup>的方法。基于插值的方法通过在原有的像素点之间插入新的像素点来扩大图片的尺寸。该类方法运行速度非常快,但结果常常存在阶梯噪声和模糊效应。基于重建的方法通常采用高维的先验知识来约束解空间,涉及复杂的优化问题和耗时的重建过程,当尺度因子变大时,性能会迅速下降。基于学习的方法依赖于 LR 和 HR 图像的几何相似性,很大程度上依赖于人工提取特征,缺乏自动学习数据特征的能力,对复杂场景的处理效果相对较差。

近年来,卷积神经网络 (CNN, convolutional neural network) 在图像超分辨率 (SR, super-resolution) 领域中的应用日益增多。基于卷积神经网络的方法旨在自动提取与任务相关的深层特征,实现图像质量的提升。例如, SRCNN (super-resolution convolutional neural network)<sup>[8]</sup> 仅由 3 个卷积层组成,却展现出了显著的性能提升,并取得了优异的重建效果。SRCNN 的提出不仅标志着超分辨率模型向深度学习的转变,同时还确立了一种三阶段的典范架构,包括浅层特征提取、非线性特征映射和图像重建阶段。为解决 SRCNN 使用双三次插值扩展原始图像作为输入而带来的高计算负担问题, ESPCN (efficient sub-pixel convolutional network)<sup>[9]</sup> 引入像素洗牌策略,通过在网络末端建立由像素洗牌层和卷积层构成的亚像素卷积层,对输出进行上采样,显著提升了网络的计算效率。随后,为了进一步提高网络性能并获取更好的重建效果,研究人员设计了更宽和更深的超分辨率网络,如 EDSR (enhanced deep super-resolution)<sup>[10]</sup>、MDFN (multi-scale dense feature fusion network)<sup>[11]</sup> 等模型,它们通过增加网络深度和宽度来探索更丰富的特征表示,从而达到优化超分辨率重建的目的。

尽管现有研究表明网络深度与性能之间存在正相关关系,但深层网络的训练成本较高。为解决此

问题,注意力机制被引入,以实现重要信息的选择性关注。例如, RCAN (residual channel attention network)<sup>[12]</sup> 结合了通道注意 (CA, channel attention)。HAN (holistic attention network)<sup>[13]</sup> 通过合并通道注意力和空间注意力以提升网络性能。然而,这些方法虽然增强了网络能力,但是也提高了对计算资源的需求。因此,在追求性能的同时,轻量级网络设计成为研究重点。TCSR (transformer convolutions image super-resolution)<sup>[14]</sup> 提出了领域注意模块,通过滑动窗口提取远程依赖关系,保持轻量级架构下的网络性能水平。VAN (visual attention network)<sup>[15]</sup> 提出大核注意力 (LKA, large kernel attention),通过特定卷积组合提升视觉任务性能,并在多个视觉任务中展现出了优于 Transformer 的性能。然而,随着卷积核尺寸的增大,深度卷积层可能会导致计算和内存消耗显著增加。为此,大核可分离注意力 (LKSA, large kernel separable attention)<sup>[16]</sup> 被提出,它将 LKA 中的深度卷积层的二维卷积核分解为级联的水平和垂直一维核,既降低了模型的计算复杂度,又保持了模型性能。与标准的 LKA 设计相比, LKSA 增大了卷积核尺寸,促使模型关注物体的形状信息而非纹理信息,从而在保持模型性能的同时降低模型计算复杂度。

以上网络通过增强全局信息提取提升网络性能,然而却忽略了除空域外的其他领域信息。多项研究<sup>[17-19]</sup> 证明了与空间域互补的频域以及其他域信息有利于提高图像重建网络性能,对超分任务具有重要意义。HFFN (high-frequency focused network)<sup>[20]</sup> 通过调整频域的高频低频资源分配,选择性增强高频信息并减少低频冗余,合理利用多层互补信息减少重建伪影子。CRAFT (cross-refinement adaptive feature transformer)<sup>[21]</sup> 从频域入手,通过高频增强模块、窗口注意力以及混合模块协同捕获高频信息,提高高频特征表征能力。ESRT (efficient super-resolution transformer)<sup>[22]</sup> 采用 CNN 与 Transformer<sup>[23]</sup> 融合的策略,并设计了高频滤波模块以提高模型效率。此外, LBNNet (lightweight bimodal network)<sup>[24]</sup> 结合对称的 CNN 以及 CNN 与 Transformer 的集成方式,从空域角度优化网络性能。HRN (heterogeneous residual network)<sup>[25]</sup> 通过引入频域特征设计了跨域异构残差模块,实现了频域与空域的信息互补,提高了模型鲁棒性。DTSR

(dual translations super-resolution) [26]设计了域距离感知与域距离监督机制,打破合成低分辨率数据漂移对模型性能的限制。HDNet (high-resolution dual-domain network) [27]在空间域、谱域和频域进行建模,通过互补的多域学习机制,避免了维度坍塌,获得了高内部分辨率,从而改善了重建质量。MSSR (multiscale spectral super-resolution) [28]将多域空间信息与多尺度光谱信息融合,解决空间信息与光谱信息提取不平衡的问题。MDFLSR (multi domain feature learning super-resolution) [29]基于2D、3D单元设计多域特征学习,交替探索空域特征和频域特征,并进行多域特征融合,进一步实现多域信息互补。DSFormer (dual-domain self-supervised transformer) [30]提出了基于双域自监督模型、基于图像域和k空间域实现自监督学习策略进行MRI图像重建,以减轻获取全采样训练数据的成本。Learn++ [31]模型能够在图像域和投影域上同时执行图像恢复和正弦图修复操作,从而充分挖掘投影数据与重建图像之间的潜在关系。MIST-Net (multi-domain integrative swin transformer network) [32]通过提取数据域和图像域的深度特征,从稀疏视图投影中重建高质量的CT图像。上述网络虽然均为通过多域进行特征提取以提高网络性能,但多数网络仅仅通过双域进行信息增强,如MDFLSR、DSFormer、Learn++、SFHN (spatial-frequency domain hybrid network) [33]等,只关注2个特定的特征域,无法充分捕获数据的多样性和复杂性,且应用领域多为高光谱图像超分和医学图像重建。MDFLSR和HDNet虽然关注多个域信息,但应用领域仅限于高光谱图像超分。HFFN、CRAFT、ESRT和LBNNet应用于传统RGB图像超分,但仅仅关注频域或空域信息。因此,对于传统RGB图像超分的多域信息提取仍然有很大的研究空间。

基于此,本文提出了一种融合多域信息增强的轻量级图像超分辨率网络。该网络综合空域、频域及转换域信息,优化了网络架构,不仅能够提取长距离依赖关系,增强细节提取能力,而且具有线性计算复杂度,显著提高了重建效果和计算效率。空域为图像的原始空间或物理空间,也就是图像数据在其原始表示形式中的空间排列,空域信息直接反映了图像的空间布局 and 低频特征。本文对于空域的信息提取主要通过多路大核注意力 (MLKA, multi-

channel large kernel attention) 模块来实现。该方法将LKA与多尺度学习结合,以较小的计算成本实现多尺度关联,提高了模型对图像的整体布局及局部内容的关注度。此外,结合像素注意力与深度可分离卷积,提出局部信息增强注意力 (LIEA, local information enhancement attention),在降低参数数量的同时,关注并强调图像细节。转换域定义为在空域维度上,采用特殊构造的卷积操作来模拟高频特征提取的特征映射空间。频域转换(如傅里叶变换或小波变换)通常需要较大的计算量,且对图像中的噪声非常敏感,尤其是在高频部分。因此,本文引入了转换域这一概念,将高频信息映射进入转换域,模拟频域中的高频特征提取。本文对于转换域的信息提取通过多路大核可分离注意力 (MLKSA, multi-channel large kernel separable attention) 模块来实现。该方法将LKSA与多尺度学习相结合,提出MLKSA模块,使用多尺度“条状”卷积,对图像横纵方向上的成分进行选择性和处理,实现了对局部高频细节和长距离依赖关系的有效抓取,对图像进行方向性约束并关注图像中物体的局部轮廓,从而提升重建质量。本文中的频域被定义为高频信息增强块 (HFEB, high frequency information enhancement block) 的映射空间。传统的小波变换、快速傅里叶变换 (FFT, fast Fourier transform) 等频域信息提取技术只能对全局频域进行分析,无法细化局部高频特征,因此在处理细节丰富的区域时存在一定的局限性。本文提出了HFEB,进行频域分频信息处理,高频信息被分为全局高频与局部高频,使用最大池化层和像素卷积捕获全局高频特征,使用深度可分离卷积层和激活函数来提取局部高频特征,实现了不同区域高频特征的精确捕捉和优化。实验评估显示,本文提出的算法在客观指标及主观感受上均优于目前主流算法,并且在参数量方面相较于其他轻量级算法更加高效。

综上所述,本文方法在模型尺寸、计算效率与性能间取得了卓越平衡,不仅能够提取长距离依赖关系,而且具有线性计算复杂度,显著提高了计算效率,为轻量级超分辨率领域贡献了一种有效解决方案,与Transformer框架和传统卷积框架相比,该方法在泛化性上更具优势。总体而言,本文的主要贡献如下。

1) 从空域出发设计了MLKA与LIEA机制,有

效解决了空间域局部信息无法精确提取的难题。

2) 从转换域出发设计了MLKSA, 模拟高频特征提取, 避免了传统高频特征提取过程中的噪声干扰问题。

3) 从频域出发, 设计HFEB对高频信息进行分频处理, 实现了不同区域高频特征的精准捕捉和优化, 提升了图像的重建质量。

4) 探索了一种基于多域融合的轻量级超分辨率网络, 针对不同特征域设计了不同的特征提取思路, 该网络在性能与效率之间实现了多维度的平衡。

## 1 网络模型

### 1.1 网络概述

如图1所示, 本文提出的多域信息增强的轻量级超分辨率重建网络为一个端到端网络, 由三部分组成: 由 $3 \times 3$ 卷积层构成的浅层特征提取模块、由6个多域信息增强特征提取块MDIEB (multi domain information enhanced block) 级联构成的深度特征提取模块以及2个 $3 \times 3$ 卷积层与像素洗牌层构成的重建模块。

首先,  $1 \times 1$ 卷积不能捕获特征的空间相关性, 可能导致低频信息丢失, 而 $5 \times 5$ 卷积被使用在超分辨率任务中的浅层特征提取阶段会引入较大的感受野, 造成无关信息引入, 因此本文重新设计了浅层特征提取模块, 仅由一个 $3 \times 3$ 卷积层构成, 低分辨率图像 $I_{LR}$ 经 $3 \times 3$ 卷积层后得到浅层特征 $X_s$ , 该过程可表示为

$$X_s = \phi(I_{LR}; \theta) \quad (1)$$

其中,  $\phi(\cdot)$ 为卷积操作,  $\theta$ 为可学习参数。

随后, 将获得的浅层特征 $X_s$ 输入深层特征提取模块, 该模块由6个级联的MDIEB组成, 以实现深层特征( $X_d$ )提取, 该过程可表示为

$$X_d = \varphi(X_s; \delta) \quad (2)$$

其中,  $X_d$ 为深层特征,  $\varphi(\cdot)$ 为深层特征提取模块的深层特征提取操作,  $\delta$ 为可训练的学习参数。

最后, 通过跳跃连接机制将浅层特征 $X_s$ 与深层特征 $X_d$ 融合, 随后经重建模块得到重建后的超分辨率图像 $I_{SR}$ 。该过程可表示为

$$I_{SR} = \psi((X_s \oplus X_d) \uparrow; \eta) \quad (3)$$

其中,  $\psi(\cdot \uparrow)$ 为 $3 \times 3$ 卷积层、像素洗牌层和 $3 \times 3$

卷积层级联构成的重建模块,  $\eta$ 为重建模块的可学习参数,  $\oplus$ 为元素相加。

以上3个操作可简化为

$$I_{SR} = f_{\text{sisr}}(I_{LR}; (\theta, \delta, \eta)) \quad (4)$$

其中,  $f_{\text{sisr}}(\cdot)$ 为整个多域信息增强超分辨率网络的函数。

在损失函数方面, 本文采用 $\mathcal{L}_1$ 损失作为整个模型的损失函数,  $\mathcal{L}_1$ 损失函数对数据敏感, 能够促使模型快速收敛, 具体计算为

$$\mathcal{L}_1(\beta) = \frac{1}{N} \sum_{i=1}^N |I_{HR}^i - I_{SR}^i| \quad (5)$$

其中,  $\beta$ 表示神经网络的参数集,  $N$ 表示训练图像对个数,  $I_{HR}^i$ 表示第 $i$ 个真实的高分辨率图像,  $I_{SR}^i$ 表示经本文算法重建后第 $i$ 个超分辨率图像。

### 1.2 多域信息增强模块

轻量级超分辨率重建网络的设计宗旨旨在减少计算资源占用和压缩模型参数规模的前提下, 提升图像重建的效率与质量。因此, 本文从空域、频域及转换域的角度出发: 从空域角度出发提出了MLKA; 从频域角度出发提出了高频信息增强块(HFEB); 从转换域角度出发提出了MLKSA。并将其综合集成为多域信息增强特征提取块(MDIEB), 使网络能够关注图像的长距离依赖性、局部细节、高频纹理以及纵横向的频率成分。显著增强了网络对关键图像特征的捕捉能力, 包括图像的整体结构、细节信息、纹理特征及物体的轮廓4个方面, 极大地提升了网络重建图像的质量, 实现了更高效率的超分辨率重建。在模型尺寸、计算效率与重建性能之间达到了很好的平衡。

如图1所示, 多域信息增强网络框架由5个部分构成, 分别是MDIEB、MLKSA、MLKA、高频特征增强块(HFEB)与LIEA。MLKA与MLKSA均引入了膨胀卷积, 可能导致网络阻塞, 因此在每一层MDIEB融入了层归一化(LN, layer normalization)机制, 有效缓解了大核膨胀卷积导致的训练困难问题, 稳定了训练梯度波动并加速了训练进程。给定输入特征 $X \in \mathbb{R}^{C \times H \times W}$ , 输入特征首先经过2个并行的特征提取分支进行处理, 第1个分支是由LN层、MLKSA和MLKA组成得到特征 $N \in \mathbb{R}^{C \times H \times W}$ , 第2个分支包含HFEB得到特征 $Y \in \mathbb{R}^{C \times H \times W}$ 。此外, 为了充分利用多尺度信息并维持训练过程的稳定性, 还引入了跳跃连接机制。

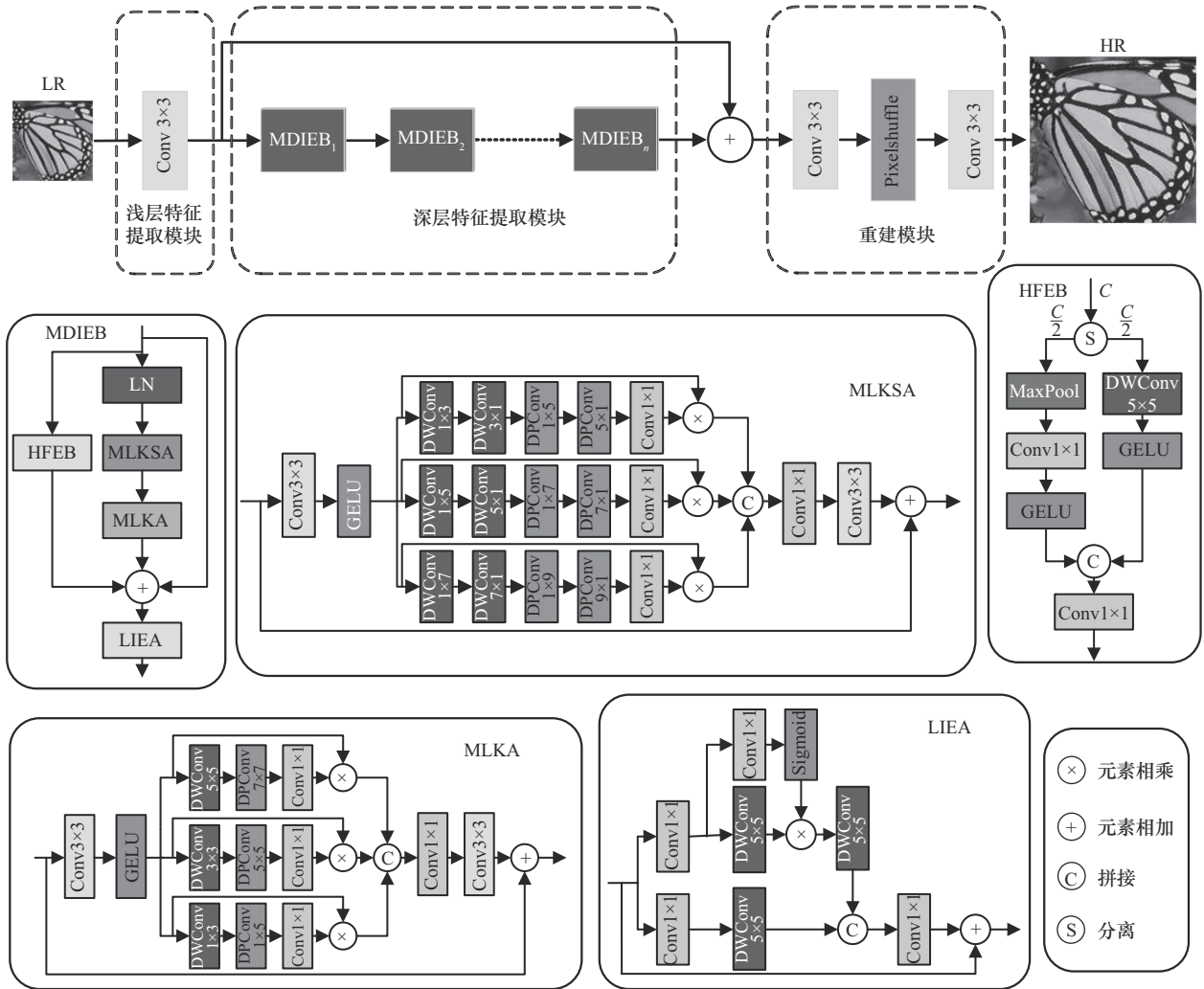


图1 多域信息增强网络框架

最终，将来自不同分支的特征进行逐元素加和，并将结果传递至LIEA模块以产生最终的输出特征  $X_{out} \in \mathbb{R}^{C \times H \times W}$ ，整个过程可表示为

$$N = f_{MLKA}(f_{MLKSA}(f_{LN}(X))) \quad (6)$$

$$Y = f_{HFEB}(X) \quad (7)$$

$$X_{out} = f_{LIEA}(X \oplus N \oplus Y) \quad (8)$$

其中， $f_{MLKA}$ 、 $f_{MLKSA}$ 、 $f_{HFEB}$  与  $f_{LIEA}$  分别表示MLKA、MLKSA、高频特征增强块（HFEB）与LIEA提取特征的操作， $f_{LN}$  与  $\oplus$  分别表示层归一化与元素相加操作。

### 1.2.1 多路大核可分离注意力模块

从转换域特征提取角度，提出了MLKSA，其在CNN中的具体体现为构造特殊形状卷积模拟高频特征提取过程，从而在CNN中直接提取出图像细化特征。具体来说，本文根据观察，图像的

高频信息通常由各区域的边缘、纹理和细节线条构成，设计了X轴和Y轴2个方向的多尺度“条状”卷积，模拟空域图像中的高频特征，通过在纵横2个方向上分别应用“条状”卷积，可以类似频域对图像进行方向性分析，从而有效地对图像在不同频率范围的方向性特征进行处理和分析，避免了传统高频提取（如小波变换、傅里叶变换）过程中的噪声干扰问题。

大核卷积注意力（LKA）把  $K \times K$  大小的普通卷积分解成  $(2d - 1) \times (2d - 1)$  的深度卷积、 $\frac{K}{d} \times \frac{K}{d}$  的深度膨胀卷积和  $1 \times 1$  的点卷积。如图1所示，大核可分离卷积注意力（LKSA）则将LKA中的深度卷积层的二维卷积核进一步分解为级联的水平和垂直一维卷积核。与标准的LKA设计相比，LKSA模块可以实现与标准LKA模块相当的性能，并且

计算复杂度和内存占用更低。给定输入特征  $\mathbf{X}_{in} \in R^{C \times H \times W}$ , 整个LKSA的实现过程可表示为

$$\mathbf{X}_{out} = \mathbf{X}_{in} \otimes \left( f_p \left( f'_{DP} \left( f'_{DW} \left( \mathbf{X}_{in} \right) \right) \right) \right) \quad (9)$$

其中, DW表示深度可分离卷积 (depthwise separable convolution), DP表示点卷积 (pointwise convolution),  $\mathbf{X}_{out} \in R^{C \times H \times W}$ 表示经LKSA分解后得到的输出特征,  $f'_{DW}(\cdot)$ 表示级联的  $(2d-1) \times 1$  与  $1 \times (2d-1)$  的深度可分离卷积,  $f'_{DP}(\cdot)$ 表示级联的  $\frac{K}{d} \times 1$  与  $1 \times \frac{K}{d}$  的深度膨胀卷积,  $f_p(\cdot)$ 表示  $1 \times 1$  点卷积操作,  $\otimes$ 表示元素相乘操作。从波段域的角度分析LKSA, 在波段域中, 图像被分解为多个波段, 每个波段代表不同的频率成分。LKSA通过在纵横2个方向上分别应用卷积, 可以类似地在波段域中对图像进行方向性分析, 可以有效地对图像在不同波段 (即不同频率范围) 的方向性特征进行处理和分析。自然地对像素施加方向性约束, 重点关注图像中物体的形象, 从而提高了图像重建的质量。

MLKSA的结构如图1所示, 它集成了3个不同尺寸的LKSA, 充分发挥了多尺度学习的效能。这种配置使网络在捕捉长距离信息的同时, 也能够细致地捕获局部信息, 并在一定程度上缓解了膨胀卷积引起的信息损耗问题。给定输入特征  $\mathbf{X}_{in} \in R^{C \times H \times W}$ , 该特征首先经  $3 \times 3$  的卷积层与GELU激活函数得到特征  $\mathbf{M} \in R^{C \times H \times W}$ 。接着从水平和垂直2个方向出发, 通过3组不同尺度的大核可分离注意力学习空间上的远程依赖性, 得到相应的特征图  $\mathbf{M}' \in R^{C \times H \times W}$ 、 $\mathbf{M}'' \in R^{C \times H \times W}$  和  $\mathbf{M}''' \in R^{C \times H \times W}$ 。之后将得到的3组特征图按通道拼接并利用  $1 \times 1$  的点卷积对拼接后的特征图进行降维, 再经  $3 \times 3$  的卷积层, 最后通过跳跃连接机制与原输入特征  $\mathbf{X}_{in} \in R^{C \times H \times W}$  按元素相加得到最终输出特征  $\mathbf{X}_{out} \in R^{C \times H \times W}$ 。

上述MLKSA的工作流程可描述为

$$\mathbf{M} = f_\sigma(f_{3 \times 3}(\mathbf{X})) \quad (10)$$

$$\begin{aligned} \mathbf{M}', \mathbf{M}'', \mathbf{M}''' &= f'_{LKSA}(\mathbf{M}) \\ f''_{LKSA}(\mathbf{M}), f'''_{LKSA}(\mathbf{M}) & \end{aligned} \quad (11)$$

$$\mathbf{X}_{out} = f_{3 \times 3} \left( f_{1 \times 1} \left( \mathbf{M}' \otimes \mathbf{M}'' \otimes \mathbf{M}''' \right) \right) \oplus \mathbf{X}_{in} \quad (12)$$

其中,  $f_{3 \times 3}(\cdot)$ 表示  $3 \times 3$  的卷积操作,  $f_\sigma(\cdot)$ 表示GELU激活函数,  $f'_{LKSA}(\cdot)$ 、 $f''_{LKSA}(\cdot)$ 与  $f'''_{LKSA}(\cdot)$ 分别表

示参数为3-5-1、5-7-1与7-9-1的大核可分离注意力函数,  $\otimes$ 和  $\oplus$ 分别表示特征按通道拼接与元素取和,  $f_{1 \times 1}(\cdot)$ 表示  $1 \times 1$  的点卷积操作。

### 1.2.2 多路大核注意力模块

从空域特征提取角度, 本文提出了MLKA, 其在CNN中的具体体现为使用不同尺度的卷积有效地提取图像的局部特征和全局特征, 尤其是在复杂的图像场景中, 能够识别出不同大小、形态的物体或结构, 让网络理解图像的整体布局和内容。

如图1所示, 大核卷积注意力模块分为3个部分: 空间局部卷积 (深度卷积)、空间远程卷积 (深度膨胀卷积) 和通道卷积 ( $1 \times 1$  卷积)。具体来说, 大核卷积注意力把  $K \times K$  卷积分解成  $(2d-1) \times (2d-1)$  深度卷积、 $\frac{K}{d} \times \frac{K}{d}$  维深度膨胀卷积和  $1 \times 1$  点卷积。分解后, 扩大了神经网络的感受野, 以较小的计算成本和参数捕获了图像的远程关系。在获得远程依赖关系后, 即可以估计出一个点的重要性并由此生成注意图。给定输入特征  $\mathbf{X}_{in} \in R^{C \times H \times W}$ , 整个大核卷积注意力的实现过程可表示为

$$\mathbf{X}_{out} = \mathbf{X}_{in} \otimes \left( f_p \left( f_{DP} \left( f_{DW} \left( \mathbf{X}_{in} \right) \right) \right) \right) \quad (13)$$

其中,  $\mathbf{X}_{out} \in R^{C \times H \times W}$ 表示经大核卷积注意力后得到的输出特征,  $f_{DW}(\cdot)$ 、 $f_{DP}(\cdot)$ 与  $f_p(\cdot)$ 分别表示  $(2d-1) \times (2d-1)$  的深度可分离卷积、 $\frac{K}{d} \times \frac{K}{d}$  的深度膨胀卷积与  $1 \times 1$  点卷积操作,  $\otimes$ 表示元素相乘操作。

在图像超分辨任务中, 任何一个点的信息对于恢复出清晰且细节丰富的图像都是不可或缺的, 因此, 本文使用3个不同尺度的大核注意力 (LKA) 构建MLKA, 以捕获图像中不同尺度的关联信息。MLKA促使网络充分关注图像中多尺度长距离信息, 通过让网络理解图像的整体布局和内容, 实现网络性能的提升。

MLKA的结构如图1所示, 给定输入特征  $\mathbf{X}_{in} \in R^{C \times H \times W}$ , 首先经由  $3 \times 3$  的卷积层与GELU激活函数得到特征  $\mathbf{P} \in R^{C \times H \times W}$ , 之后经3组不同尺度的大核注意力学习空间远程依赖关系得到特征  $\mathbf{P}' \in R^{C \times H \times W}$ 、 $\mathbf{P}'' \in R^{C \times H \times W}$  和  $\mathbf{P}''' \in R^{C \times H \times W}$ 。之后将得到的3组特征图按通道拼接后, 利用  $1 \times 1$  的点卷积对拼接后的特征图进行降维, 再经  $3 \times 3$  的卷积层得到特征, 最后将得到的特征通过跳跃连接机

制与原输入特征  $X_{in} \in R^{C \times H \times W}$  按元素相加得到最终输出特征  $X_{out} \in R^{C \times H \times W}$ 。其中, 3组大核卷积分别是 3-5-1、5-7-1 与 7-9-1, 用于捕获特征之间的不同尺度间的远程依赖信息,  $a - b - 1$  表示级联  $a \times a$  深度卷积 (DWConv)、 $b \times b$  深度扩展卷积 (DPConv) 和点卷积 (PConv)。给定输入特征  $X_{in} \in R^{C \times H \times W}$ , 上述过程可描述为

$$P = f_{\sigma}(f_{3 \times 3}(X)) \quad (14)$$

$$P', P'', P''' = f'_{LKA}(P) f''_{LKA}(P) f'''_{LKA}(P) \quad (15)$$

$$X_{out} = f_{3 \times 3}(f_{1 \times 1}(P' \otimes P'' \otimes P''')) \oplus X \quad (16)$$

其中,  $f_{3 \times 3}(\cdot)$  表示  $3 \times 3$  的卷积操作,  $f_{\sigma}(\cdot)$  表示 GELU 激活函数,  $f'_{LKA}(\cdot)$ 、 $f''_{LKA}(\cdot)$  与  $f'''_{LKA}(\cdot)$  分别表示参数为 3-5-1、5-7-1 与 7-9-1 的大核注意力函数,  $\otimes$  和  $\oplus$  分别表示特征按通道拼接与元素取和操作,  $f_{1 \times 1}(\cdot)$  表示  $1 \times 1$  的点卷积操作。

MLKA 与 1.2.1 节中的 MLKSA 虽然在结构上类似, 但他们的实现功能以及参数量均存在显著差异。MLKSA 代表的转换域操作, 通过构造特殊形状卷积模拟高频特征提取过程, 实现了图像的重要信息 (如物体的轮廓、细节和局部变化) 特征提取, 这些特征均为高频特征。而 MLKA 代表的空域操作, 通过不同尺度的普通卷积核对图像进行处理, 实现了图像局部特征和全局特征的提取, 这些特征多为低频特征。此外, 它们的计算成本和参数量存在显著差异, MLKSA 模块优势在于参数量和内存占用较低, 专注于边缘和细节特征的提取, 而 MLKA 相较于 MLKSA 参数量更大, 且使用标准卷积, 专注于提取全局和局部低频信息。

### 1.2.3 高频信息增强块

从频域特征提取角度, 本文提出了高频信息增强块 (HFEB), 其在 CNN 中的具体体现为使用 HFEB 分支作为高频信息, 并对高频信息进行频域分频处理, 将频率进一步细化为全局高频与局部高频, 实现了不同区域高频特征的精确捕捉和优化。具体来说, HFEB 分支通过 2 个分支, 将特征分割为 2 个通道数相等的特征, 一个支路使用  $5 \times 5$  深度可分离卷积层和 GELU 激活函数提取局部高频特征, 另一个支路利用最大池化层捕获全局高频特征, 并通过一个  $1 \times 1$  的卷积层结合 GELU 激活函数进一步增强高频信息, 从而有效提升细节恢复和特征提取能力。

在图像的频域表征中, 高频分量通常与图像的边缘、纹理以及其他细节特征相对应, 这些特征对于人类视觉系统感知图像质量至关重要。高频信息的丢失会导致图像出现模糊和细节缺失的现象。因此, 在图像重建过程中, 恢复高频信息对于细节和纹理的精确再现是至关重要的。CARFT (cross-refinement adaptive feature modulation transformer) [21] 提出了高频增强剩余块 (HFEB, high-frequency enhancement residual block), 旨在增强高频信息。本文对其进行适当改进, 旨在增强高频信息的同时使其更加轻量化。如图 1 所示, 本文改进的 HFEB 分为 2 个分支, 输入特征  $X_{in} \in R^{C \times H \times W}$  首先被分割为 2 个通道数均等的特征  $X_{in}^1, X_{in}^2 \in R^{\frac{C}{2} \times H \times W}$ , 之后经  $5 \times 5$  深度可分离卷积层和 GELU 激活函数来提取局部高频特征  $X_{out}^1 \in R^{\frac{C}{2} \times H \times W}$ 。接着, 利用最大池化层捕获高频信息, 并通过一个  $1 \times 1$  的卷积层结合 GELU 激活函数来进一步增强这些高频特征, 从而得到增强后的特征  $X_{out}^2 \in R^{\frac{C}{2} \times H \times W}$ 。最后, 将 2 个支路特征拼接融合后经  $1 \times 1$  卷积层得到输出特征  $X_{out} \in R^{C \times H \times W}$ 。整个 HFEB 的工作流程可表示为

$$X_{in}^1, X_{in}^2 = \text{Split}(X_{in}) \quad (17)$$

$$X_{out}^1 = f_{\sigma}(f_{DW}(X_{in}^1)) \quad (18)$$

$$X_{out}^2 = f_{\sigma}(f_{1 \times 1}(f_{\max}(X_{in}^2))) \quad (19)$$

$$X_{out} = f_{1 \times 1}(X_{out}^1 \otimes X_{out}^2) \quad (20)$$

其中,  $\text{Split}(\cdot)$  表示特征按通道分割操作,  $f_{\sigma}(\cdot)$ 、 $f_{DW}(\cdot)$ 、 $f_{1 \times 1}(\cdot)$ 、 $f_{\max}(\cdot)$  和  $\otimes$  分别表示 GELU 激活函数、 $5 \times 5$  深度卷积、 $1 \times 1$  点卷积、最大池化层和特征拼接操作。

### 1.2.4 局部信息增强注意力模块

在图像恢复过程中, 局部信息和长距离信息都是不可或缺的。MLKA 与 MLKSA 通过引入膨胀卷积, 建立特征图的长距离空间关系, 但降低了特征图的局部位置关系。因此, 本节提出了 LIEA 模块, 该模块通过点卷积、深度卷积以及像素注意力机制增强了网络对局部信息的提取能力。这使网络能够聚焦于局部的关键信息, 并在图像中建立局部的关联性, 从而关注到图像的细节部分。如图 1 所示, LIEA 的工作流程如下。给定输入特征

$X_{in} \in \mathbb{R}^{C \times H \times W}$ , 首先经  $1 \times 1$  卷积和  $5 \times 5$  深度卷积构成的局部特征提取支路生成特征  $F_2 \in \mathbb{R}^{C \times H \times W}$ 。此外, 输入特征还经过局部特征增强支路, 首先经  $1 \times 1$  卷积生成特征  $F_1 \in \mathbb{R}^{C \times H \times W}$ , 特征  $F_1 \in \mathbb{R}^{C \times H \times W}$  经  $1 \times 1$  卷积和 Sigmoid 激活函数构成的像素注意力支路生成注意力图, 再与  $5 \times 5$  深度卷积生成的特征相乘, 得到特征  $F'_1 \in \mathbb{R}^{C \times H \times W}$ 。随后经  $5 \times 5$  深度卷积得到局部增强特征  $F''_1 \in \mathbb{R}^{C \times H \times W}$ 。最后, 通过特征拼接操作将 2 个支路特征拼接融合, 经  $1 \times 1$  卷积层进行通道降维操作。此外, 为了保持训练的稳定性, 防止信息丢失, 引入了跳跃连接机制, 最终得到 LIEA 的输出特征  $X_{out} \in \mathbb{R}^{C \times H \times W}$ 。整个过程可表示为

$$F_1 = f_{1 \times 1}(X) \quad (21)$$

$$F'_1 = f_{5 \times 5}(F_1) \otimes f_s(f_{1 \times 1}(F_1)) \quad (22)$$

$$F''_1 = f_{5 \times 5}(F'_1) \quad (23)$$

$$F_2 = f_{5 \times 5}(f_{1 \times 1}(X)) \quad (24)$$

$$X_{out} = f_{1 \times 1}(F''_1 \otimes F_2) \oplus X \quad (25)$$

其中,  $f_{1 \times 1}(\cdot)$  和  $f_{5 \times 5}(\cdot)$  分别表示  $1 \times 1$  点卷积和  $5 \times 5$  深度卷积,  $\otimes$  和  $\oplus$  分别表示特征按通道拼接与元素取和操作。

## 2 实验

### 2.1 数据集与指标

本文采用 DIV2K 作为训练数据集, 其中包含 800 张训练图像, 10 张验证图像。同时使用 4 个基准数据集 Set5<sup>[34]</sup>、Set14<sup>[35]</sup>、B100<sup>[36]</sup> 和 Urban100<sup>[37]</sup> 来评估不同方法的性能, 放大倍数分别为  $\times 2$ 、 $\times 3$  和  $\times 4$ 。采用  $Y$  通道平均峰值信噪比 (PSNR, peak signal to noise ratio)<sup>[38]</sup> 和结构相似度 (SSIM, structural similarity)<sup>[39]</sup> 对图像质量进行评估, 模型复杂度用参数量表示。

### 2.2 执行细节

在模型训练阶段采用双三次下采样从原始的 HR 图像中获取对应的 LR 图像。在训练过程中, 将图像随机裁剪为  $96 \times 96$  大小的图像块, 训练批次大小设置为 16, 特征通道为 48。同时, 执行数据增强操作, 如随机水平翻转或  $90^\circ$  旋转。采用 Adam 优化器训练, 参数  $\beta_1$  与  $\beta_2$  分别为 0.9 与 0.999。初始学习率设置为  $4 \times 10^{-4}$ , 并在迭代  $2 \times 10^5$ 、 $4 \times 10^5$ 、 $6 \times 10^5$

和  $8 \times 10^5$  次时学习率减少一半, 总迭代次数为  $1 \times 10^6$ 。采用  $L_1$  损失函数优化模型。使用 Pytorch1.6 深度学习框架, 编程语言为 python3.7, 在一张 GeForce RTX 3090 GPU 上进行实验。

### 2.3 与先进的方法比较

为验证本文模型的有效性, 与其他先进的轻量级 SISR 方法在 4 个公开数据集上进行比较。这些方法包括 DRCN (deeply-recursive convolutional network)<sup>[40]</sup>、EDSR-baseline<sup>[10]</sup>、CARN (cascading residual network)<sup>[41]</sup>、IMDN (information multi-distillation network)<sup>[42]</sup>、SMSR (sparse mask super-resolution)<sup>[43]</sup>、ESRT (efficient super-resolution transformer)<sup>[22]</sup>、FMEN (fast and memory-efficient network)<sup>[44]</sup>、DRSAN-48m (dynamic residual self-attention network)<sup>[45]</sup>、NGswin (N-gram swin)<sup>[46]</sup>、DiVANet (directional variance attention network)<sup>[47]</sup>、OSFFNet (omni-stage feature fusion network)<sup>[48]</sup> 和 HSRNet (hierarchical image super-resolution network)<sup>[49]</sup>。下面展示定量指标分析和视觉质量结果。

#### 2.3.1 定量分析

表 1~表 3 给出了本文所提网络模型 MDIESR (multi-domain information enhancement super-resolution) 与其他先进的轻量级 SISR 方法的实验结果。最好和次好的结果分别被粗标显示和加下划线, 其中本文提出的 MDIESR 在除 Urban100 数据集上的  $\times 2$  尺度因子下略逊于 OSFFNet 以外, 其他尺度因子及数据集上均取得了最优的表现。OSFFNet 充分挖掘了浅层网络特征, 通过动态融合增强浅层特征贡献, 然而本文提出的 MDIESR 特征提取层多数集中在深层特征, 无法实现浅层特征与深层特征间的有效互补, 因此浅层特征提取也将成为下一步研究的重要方向。对比经典的 EDSR-baseline 模型, 本文模型在 Urban100 数据集上的  $\times 2$ 、 $\times 3$  和  $\times 4$  尺度因子下分别实现了 0.85 dB、0.57 dB 和 0.55 dB 的显著性能提升, 参数量减少了 30%。此外, 与最近的 DRSAN-48m 模型相比, 本文模型在 Urban100 数据集上的  $\times 2$ 、 $\times 3$  和  $\times 4$  尺度因子下分别提高了 0.19 dB、0.20 dB 和 0.26 dB, 参数量减少了 20%。在近乎同等的参数量的情况下, 对比基于 Transformer 的方法 NGswin, 在 Set5 和 Urban100 数据集的  $\times 3$  尺度因子下, 本文模型分别获得了 0.17 dB 和 0.20 dB 的性

表 1 不同SISR模型在尺度因子为2时的4个基准数据集上的性能比较

模型	参数量	FLOPS	PSNR/SSIM			
			Set5	Set14	B100	Urban100
DRCN <sup>[40]</sup>	1.77×10 <sup>6</sup>	9 788.7×10 <sup>9</sup>	37.63 / 0.958 8	33.04 / 0.911 8	31.85 / 0.894 2	30.75 / 0.913 3
EDSR-baseline <sup>[10]</sup>	1.37×10 <sup>6</sup>	316.3×10 <sup>9</sup>	37.99 / 0.960 4	33.57 / 0.917 5	32.16 / 0.899 4	31.98 / 0.927 2
CARN <sup>[41]</sup>	1.59×10 <sup>6</sup>	222.8×10 <sup>9</sup>	37.76 / 0.959 0	33.52 / 0.916 6	32.09 / 0.897 8	31.92 / 0.925 6
IMDN <sup>[42]</sup>	0.69×10 <sup>6</sup>	158.8×10 <sup>9</sup>	38.00 / 0.960 5	33.63 / 0.917 7	32.19 / 0.899 6	32.17 / 0.928 3
SMSR <sup>[43]</sup>	0.98×10 <sup>6</sup>	—	38.00 / 0.960 1	33.64 / 0.917 9	32.17 / 0.899 0	32.19 / 0.928 4
ESRT <sup>[22]</sup>	0.77×10 <sup>6</sup>	191.4×10 <sup>9</sup>	38.03 / 0.960 0	33.75 / 0.918 4	32.25 / 0.900 1	32.58 / 0.931 8
FMEN <sup>[44]</sup>	0.75×10 <sup>6</sup>	172.0×10 <sup>9</sup>	38.10 / 0.960 9	33.75 / 0.919 2	32.26 / 0.900 7	32.41 / 0.931 1
DRSAN-48m <sup>[45]</sup>	1.19×10 <sup>6</sup>	247.6×10 <sup>9</sup>	38.14 / 0.961 1	33.75 / 0.918 8	32.25 / 0.901 0	32.46 / 0.931 7
NGswin <sup>[46]</sup>	0.99×10 <sup>6</sup>	140.4×10 <sup>9</sup>	38.05 / 0.961 0	33.79 / <u>0.919 9</u>	32.27 / 0.900 8	32.53 / 0.932 4
DiVANet <sup>[47]</sup>	0.90×10 <sup>6</sup>	189.0×10 <sup>9</sup>	<u>38.16 / 0.961 2</u>	<u>33.80 / 0.919 5</u>	<u>32.29 / 0.901 2</u>	32.60 / <u>0.932 5</u>
OSFFNet <sup>[48]</sup>	0.51×10 <sup>6</sup>	83.2×10 <sup>9</sup>	38.11 / 0.961 0	33.72 / 0.919 0	<u>32.29 / 0.901 2</u>	<b>32.67 / 0.933 1</b>
HSRNet <sup>[49]</sup>	1.26×10 <sup>6</sup>	808.2×10 <sup>9</sup>	38.07 / 0.960 7	33.78 / 0.919 7	32.26 / 0.900 6	32.53 / 0.932 0
MDIESR (本文模型)	0.96×10 <sup>6</sup>	<u>111.56×10<sup>9</sup></u>	<b>38.17 / 0.961 3</b>	<b>33.83 / 0.920 0</b>	<b>32.31 / 0.901 3</b>	<u>32.65 / 0.933 1</u>

表 2 不同SISR模型在尺度因子为3时的4个基准数据集上的性能比较

模型	参数量	FLOPS	PSNR/SSIM			
			Set5	Set14	B100	Urban100
DRCN <sup>[40]</sup>	1.77×10 <sup>6</sup>	9 788.7×10 <sup>9</sup>	33.82 / 0.922 6	29.76 / 0.831 1	28.80 / 0.796 3	27.15 / 0.827 6
EDSR-baseline <sup>[10]</sup>	1.55×10 <sup>6</sup>	160.2×10 <sup>9</sup>	34.37 / 0.927 0	30.28 / 0.841 7	29.09 / 0.805 2	28.15 / 0.852 7
CARN <sup>[41]</sup>	1.59×10 <sup>6</sup>	118.9×10 <sup>9</sup>	34.29 / 0.925 5	30.29 / 0.840 7	29.06 / 0.803 4	28.06 / 0.849 3
IMDN <sup>[42]</sup>	0.70×10 <sup>6</sup>	71.5×10 <sup>9</sup>	34.36 / 0.927 0	30.32 / 0.841 7	29.09 / 0.804 6	28.17 / 0.851 9
SMSR <sup>[43]</sup>	0.99×10 <sup>6</sup>	—	34.40 / 0.927 0	30.33 / 0.841 2	29.10 / 0.805 0	28.25 / 0.853 6
ESRT <sup>[22]</sup>	0.77×10 <sup>6</sup>	96.4×10 <sup>9</sup>	34.42 / 0.926 8	30.43 / 0.843 3	29.15 / 0.806 3	28.46 / 0.857 4
FMEN <sup>[44]</sup>	0.76×10 <sup>6</sup>	77.2×10 <sup>9</sup>	34.45 / 0.927 5	30.40 / 0.843 5	29.17 / 0.806 3	28.33 / 0.856 2
DRSAN-48m <sup>[45]</sup>	1.29×10 <sup>6</sup>	133.4×10 <sup>9</sup>	34.59 / 0.928 6	30.42 / 0.844 3	29.18 / 0.807 9	28.52 / 0.859 3
NGswin <sup>[46]</sup>	1.01×10 <sup>6</sup>	66.6×10 <sup>9</sup>	34.52 / 0.928 2	<u>30.53 / 0.845 6</u>	29.19 / 0.807 8	28.52 / <u>0.860 3</u>
DiVANet <sup>[47]</sup>	0.95×10 <sup>6</sup>	89.0×10 <sup>9</sup>	<u>34.60 / 0.928 5</u>	30.47 / 0.844 7	29.19 / 0.807 3	<u>28.58 / 0.860 3</u>
OSFFNet <sup>[48]</sup>	0.52×10 <sup>6</sup>	37.8×10 <sup>9</sup>	34.58 / <u>0.928 7</u>	30.48 / 0.845 0	<u>29.21 / 0.808 0</u>	28.49 / 0.859 5
HSRNet <sup>[49]</sup>	—	—	34.47 / 0.927 8	30.40 / 0.8435	29.15 / 0.806 6	28.42 / 0.857 9
MDIESR (本文模型)	1.07×10 <sup>6</sup>	<u>49.50×10<sup>9</sup></u>	<b>34.69 / 0.929 5</b>	<b>30.58 / 0.846 5</b>	<b>29.25 / 0.808 7</b>	<b>28.72 / 0.863 4</b>

能提升。在计算复杂度方面，每秒浮点运算次数（FLOPS）是衡量计算复杂度的一个重要指标，通常用于描述神经网络或算法执行时的计算量。因此，本文统计了所有算法的FLOPS，并对最优复杂度和次优复杂度的位置进行了加粗和加下划线处理。可以观察到，本文提出的MDIESR网络模型具有次优复杂度。OSFFNet模型的运算量低于所有算法，其原因在于其参数量极少。对于单位参数量的

浮点运算数，MDIESR的表现远超OSFFNet。此外，在同量级参数的算法（如SMSR、NGswin、DiVANet）中，MDIESR的浮点运算数最低，但其指标在不同尺度因子及数据集上均取得了最优的表现。因此，本文提出的MDIESR网络模型在计算复杂度方面表现出色。实验结果充分表明了本文模型无论是在模型指标还是在计算复杂度方面均有着优异的表现。

表3 不同SISR模型在尺度因子为4时的4个基准数据集上的性能比较

模型	参数量	FLOPS	PSNR/SSIM			
			Set5	Set14	B100	Urban100
DRCN <sup>[40]</sup>	$1.77 \times 10^6$	$9\,788.7 \times 10^9$	31.53 / 0.885 4	28.02 / 0.767 0	27.23 / 0.723 3	25.14 / 0.751 0
EDSR-baseline <sup>[10]</sup>	$1.51 \times 10^6$	$114.0 \times 10^9$	32.09 / 0.893 8	28.58 / 0.781 3	27.57 / 0.735 7	26.04 / 0.784 9
CARN <sup>[41]</sup>	$1.59 \times 10^6$	$90.9 \times 10^9$	32.13 / 0.893 7	28.60 / 0.780 6	27.58 / 0.734 9	26.07 / 0.783 7
IMDN <sup>[42]</sup>	$0.70 \times 10^6$	$40.9 \times 10^9$	32.21 / 0.894 8	28.58 / 0.781 1	27.56 / 0.735 3	26.04 / 0.783 8
SMSR <sup>[43]</sup>	$1.00 \times 10^6$	—	32.12 / 0.893 2	28.55 / 0.780 8	27.55 / 0.735 1	26.11 / 0.786 8
ESRT <sup>[22]</sup>	$0.75 \times 10^6$	$67.7 \times 10^9$	32.19 / 0.894 7	28.69 / 0.783 3	27.69 / 0.737 9	26.39 / 0.796 2
FMEN <sup>[44]</sup>	$0.77 \times 10^6$	$44.2 \times 10^9$	32.24 / 0.895 5	28.70 / 0.783 9	27.63 / 0.737 9	26.28 / 0.790 8
DRSAN-48m <sup>[45]</sup>	$1.27 \times 10^6$	$88.7 \times 10^9$	32.34 / 0.896 0	28.65 / 0.784 1	27.63 / 0.739 0	26.33 / 0.793 6
NGswin <sup>[46]</sup>	$1.01 \times 10^6$	$36.4 \times 10^9$	32.33 / 0.896 3	<u>28.78 / 0.785 9</u>	<u>27.66 / 0.739 6</u>	<u>26.45 / 0.796 3</u>
DiVANet <sup>[47]</sup>	$0.94 \times 10^6$	$57.0 \times 10^9$	<u>32.41</u> / 0.897 3	28.70 / 0.784 4	27.65 / 0.739 1	26.42 / 0.795 8
OSFFNet <sup>[48]</sup>	$0.52 \times 10^6$	$22.0 \times 10^9$	32.39 / <u>0.897 6</u>	28.75 / 0.785 2	<u>27.66</u> / 0.739 3	26.36 / 0.795 0
HSRNet <sup>[49]</sup>	$1.29 \times 10^6$	$203.2 \times 10^9$	32.28 / 0.896 0	28.68 / 0.784 0	27.64 / 0.738 8	26.28 / 0.793 4
MDIESR (本文模型)	$1.04 \times 10^6$	<u>27.89</u> $\times 10^9$	<b>32.49 / 0.898 6</b>	<b>28.84 / 0.786 7</b>	<b>27.73 / 0.739 9</b>	<b>26.59 / 0.800 7</b>

### 2.3.2 视觉质量分析

在具有挑战性的数据集Urban100和B100上的视觉对比如图2~图5所示。从图2~图5中可以看

出,当输入LR图像中存在信息丰富的自相似纹理时,本文提出的MDIESR可以修复严重受损的纹理。通过对比图3~图5中的其他算法重建图像的纹

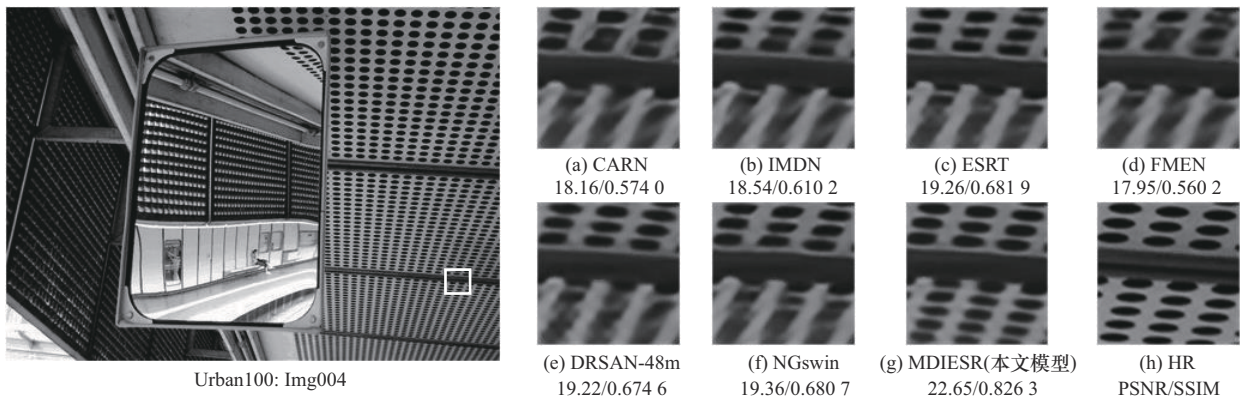


图2 Urban100中Img004重建结果

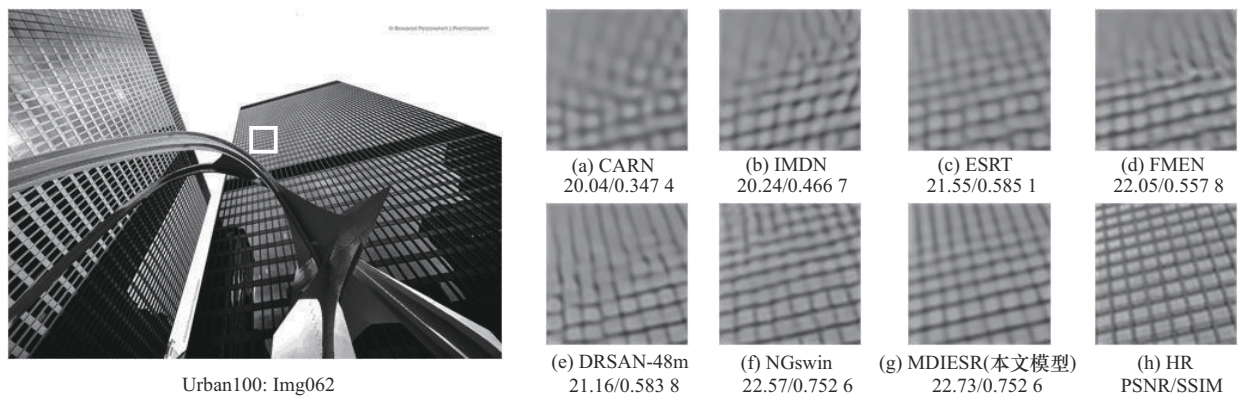


图3 Urban100中Img062重建结果

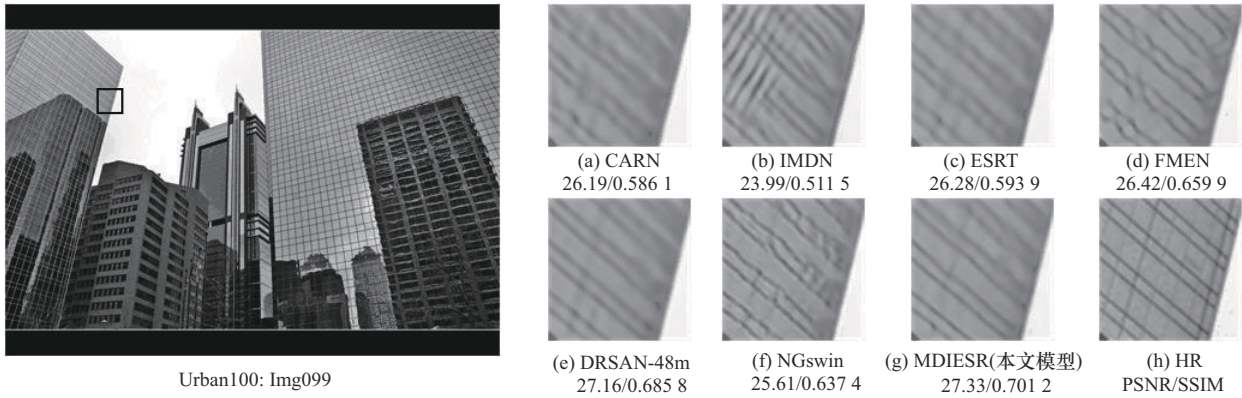


图4 Urban100中Img099重建结果

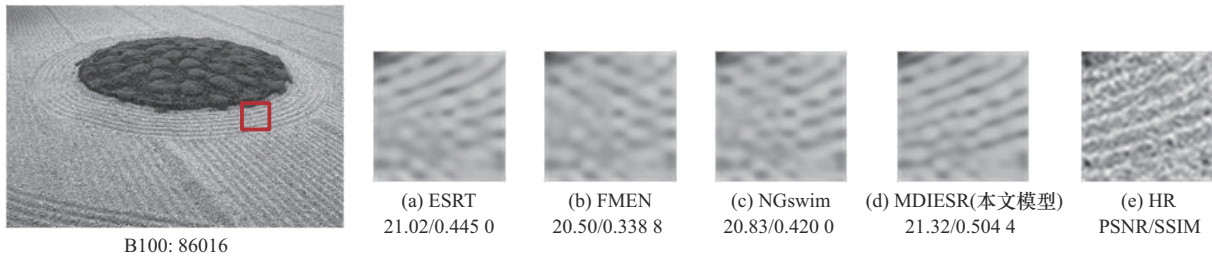


图5 B100中86016重建结果

理，可以看到，MDIESR生成的纹理与HR图像纹理基本相似，但其他算法均难以恢复其轮廓细节。在图2中，其他算法均有严重的失真，而MDIESR从多维度出发，图像恢复质量更高。这些比较表明，本文提出的MDIESR网络模型，在有效地利用空域中局部与全局信息的同时，兼顾了图像的高频信息，同时引入转换域的信息对图像中的点信息重点约束，使网络能够学习到多尺度与多层次的特征。因此能够更好地恢复图像中的纹理信息并重建出视觉效果更优异的图像。

## 2.4 消融实验

本节将对本文中提出的MLKA、MLKSA、HFRB

和LIEA对网络性能的贡献进行详细的分析。在每组实验中，网络结构包含了4个本节提出的特征提取基础块，每个基础块的通道数设定为48，尺度因子设定为 $\times 4$ 。在网络训练过程中，总共迭代 $4 \times 10^4$ 次。在Set5、Set14、B100和Urban100数据集上进行测试，采用PSNR作为评价指标进行性能对比。

### 2.4.1 MDIEB的影响分析

为验证本文提出的MLKSA转换域、MLKA空域、LIEA空域及HFEB频域对特征提取模块MDIEB多域信息增强模块的影响，使用7组消融实验进行验证，实验结果如表4所示。首先验证了

表4 MLKSA、MLKA、LIEA和HFEB对MDIEB的影响

MLKSA 转换域	MLKA 空域	LIEA 空域	HFEB 频域	参数量	PSNR			
					Set5	Set14	B100	Urban100
×	×	×	×	$0.34 \times 10^6$	31.65	28.23	27.34	25.29
√	×	×	×	$0.41 \times 10^6$	31.98	28.44	27.47	25.68
×	√	×	×	$0.44 \times 10^6$	31.99	28.47	27.48	25.71
×	×	√	×	$0.23 \times 10^6$	31.64	28.21	27.31	25.28
√	√	×	×	$0.68 \times 10^6$	32.23	28.63	27.59	26.11
√	√	×	√	$0.69 \times 10^6$	32.29	28.68	27.63	26.26
√	√	√	√	$0.75 \times 10^6$	32.35	28.70	27.64	26.32

单独采用 MLKSA、MLKA 和 LIEA 时的网络性能, 对比表 4 中仅采用 4 个残差块 (ResB) 时, 采用 MLKSA 和 MLKA 时网络性能均获得了大幅提升。其中 MLKSA 和 MLKA 在 Set5 数据集上的 PSNR 值分别提升了 0.34 dB 与 0.35 dB。而 LIEA 在参数量降低 32% 的情况下仅付出了 0.3% 的性能代价, 实现了性能与参数量良好平衡。实验结果充分证明了 MLKSA、MLKA 和 LIEA 的有效性。随后验证了同时引入 MLKSA 和 MLKA 的性能, 相比单独引入 MLKSA 和 MLKA 时在 Urban100 数据集上性能分别提升了 0.43 dB 与 0.4 dB。随后在同时引入 MLKSA 与 MLKA 的基础上引入 HFEB, 网络性能在 Urban100 数据集上提升了 0.15 dB, 且仅付出了  $0.01 \times 10^6$  的参数量代价, 实验结果充分表明 HFEB 能够充分关注并利用图像中的高频信息, 极大地提升了网络模型的性能。随后引入了 LIEA 时网络模型性能在 Set5 与 Urban100 数据集上均提升了 0.06 dB。

为了进一步探索 MLKSA 转换域、MKKA 空域以及 HFEB 频域的长距离建模以及高频信息建模的能力, 本文对各个模块的像素归因进行了 LAM 可视化, 如图 6 所示。其中, LAM 表示 LR 中每个像素的重要性与 SR 之间的相似性归因图, BASELINE 为所有卷积核固定为 3 的模型可视化结果, HFEB 为 BASELINE 加入 HFEB 的可视化结果, MLKA 为 BASELINE 加入 MLKA 的可视化结果, MLKSA 为 BASELINE 加入 MLKSA 的可视化结果, MDIESR 为本文的可视化结果。从图 6 中可以明显看出, 与 BASELINE 相比, 加入了 HFEB、MLKA 和 MLKSA 的局部归因范围更大, 这意味着 HFEB、MLKA 和 MLKSA 具有更大的感受野和更强的长距

离建模能力。在 HFEB 中, 归因像素更多地分布在建筑物的窗户边缘, 从而使 LAM 呈现矩形并得出结论, HFEB 更加注重模型的高频信息建模。本文模型 MDIESR 在结合了 HFEB、MLKA 和 MLKSA 后, 归因图进一步扩大, 并且仍然没有忽略高频信息。因此, 本文模型 MDIESR 具备长距离建模和高频信息建模的能力。

实验结果充分证明了单域特征重建具有局限性, 只有将转换域、空域和频域这 3 个域的信息进行融合, 才能实现各个域之间的有效信息互补, 从而提升模型的有效性。

#### 2.4.2 MLKSA 的影响分析

为验证本文提出的 MLKSA 模块的有效性, 设计了 7 组消融实验验证。从不同尺度的 LKSA 配置以及多路策略的应用 2 个维度进行评估, 实验结果如表 5 所示。首先探讨了不同尺度大小的 LKSA 对图像重建质量的影响。实验结果表明, 随着 LKSA 尺度的增加, 模型在 Set5、Set14、B100 和 Urban100 数据集上的 PSNR 呈现出上升趋势。当采用双路策略时, 3-5-1 与 7-9-1 大小的 LKSA 比 5-7-1 与 7-9-1 大小的 LKSA 效果更好。这一现象可以解释为: 在图像超分辨率重建任务中, 长距离的信息关联性与局部纹理信息都是至关重要且不可或缺的, 5-7-1 尺度的 LKSA 在一定程度上未能充分关注局部信息。最后, 当采用三路策略时, 通过结合小、中、大 3 个尺度的 LKSA, 网络能够更为全面地关注并建模图像中的局部细节和长距离依赖关系, 从而提升了网络的泛化能力。因此, 综合考虑计算复杂度、运算成本与性能表现, 最终确定了采用 3-5-1、5-7-1 与 7-9-1 尺度的 LKSA 作为 MLKSA 的配置。

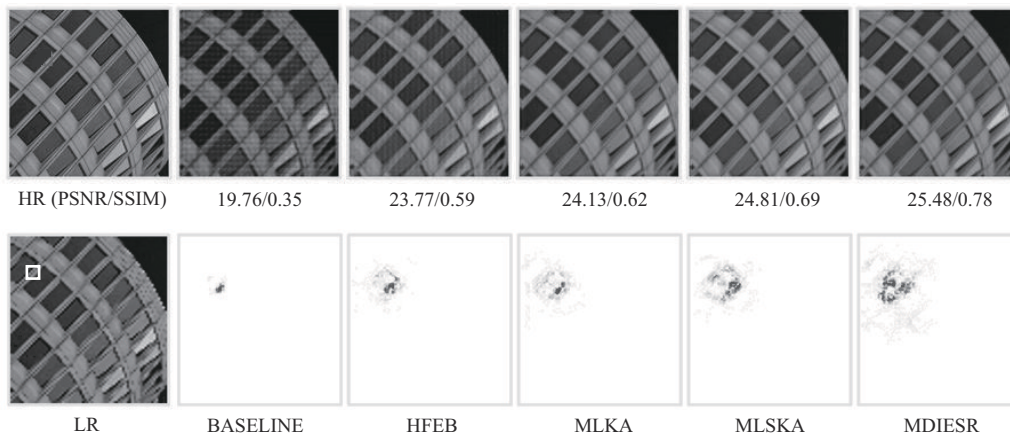


图6 各模块LAM可视化对比结果

### 2.4.3 MLKA 的影响分析

为了验证本文提出的 MLKA 的有效性, 进行了一系列共 7 组的消融实验。这些实验从不同尺度大小的 LKA 以及采用多路策略 2 个维度进行了验证, 实验结果如表 5 和表 6 所示。实验首先考察了不同尺寸的 LKA 对图像重建质量的影响。结果表明, 随着 LKA 尺度的增加, 模型在 4 个基准数据集上的 PSNR 值呈现出一致的上升趋势。在双路配置下, 3-5-1 与 7-9-1 尺度的 LKA 表现出优于 5-7-1 与 7-9-1 尺度的 LKA 的性能, 再次表明了局部信息对于图像超分辨重建的重要性, 同时也再次强调了在图像超分辨率重建任务中, 同时考虑长距离的信息关联性和局部纹理信息的重要性。此外, 采用 3 种

不同尺度的 LKA 构成的 MLKA, 在所有 4 个数据集上均实现了最优的性能。模型的设计考虑了多层次特征的短距离和长程距离建模, 平衡了计算成本、运算开销与模型性能, 并选择了 3-5-1、5-7-1 与 7-9-1 尺度的 LKA 作为 MLKA 模型的配置。

### 2.5 模型复杂度分析

在轻量级超分辨率重建网络模型的设计中, 模型的参数量和计算复杂度是首要考虑的因素。为了充分了解 MDIESR 的模型参数量对性能的影响, 设计了以下 4 组实验, 分别从特征提取块 MDIEB 的输出通道数和个数 2 个方面进行分析实验, 实验结果如表 7 所示。

一方面, 分析了在保持特征提取模块 MDIEB

表 5 LKSA 的尺度与多路机制对 MLKSA 的影响

LKSA3-5-1	LKSA5-7-1	LKSA7-9-1	参数量	PSNR			
				Set5	Set14	B100	Urban100
√	×	×	$0.36 \times 10^6$	31.84	28.34	27.40	25.50
×	√	×	$0.36 \times 10^6$	31.85	28.36	27.42	25.52
×	×	√	$0.36 \times 10^6$	31.86	28.37	27.42	25.54
√	√	×	$0.38 \times 10^6$	31.91	28.40	27.44	25.60
×	√	√	$0.38 \times 10^6$	31.91	28.39	27.45	25.61
√	×	√	$0.38 \times 10^6$	31.93	28.42	27.46	25.67
√	√	√	$0.41 \times 10^6$	31.95	28.45	27.47	25.68

表 6 LKA 的尺度与多路机制对 MLKA 的影响

LKA3-5-1	LKA5-7-1	LKA7-9-1	参数量	PSNR			
				Set5	Set14	B100	Urban100
√	×	×	$0.36 \times 10^6$	31.86	28.38	27.41	25.52
×	√	×	$0.37 \times 10^6$	31.87	28.39	27.42	25.54
×	×	√	$0.38 \times 10^6$	31.90	28.40	27.44	25.61
√	√	×	$0.39 \times 10^6$	31.95	28.44	27.46	25.66
×	√	√	$0.41 \times 10^6$	31.94	28.44	27.45	25.65
√	×	√	$0.40 \times 10^6$	31.96	28.45	27.47	25.67
√	√	√	$0.44 \times 10^6$	31.97	28.46	27.48	25.68

表 7 模型复杂度对模型性能的影响

MDIEB 个数/个	输出通道数	参数量	浮点运算量	PSNR/SSIM			
				Set5	Set14	B100	Urban100
4	48	$0.75 \times 10^6$	$58.35 \times 10^9$	32.34 / 0.896 4	28.69 / 0.785 0	27.63 / 0.738 3	26.30 / 0.793 3
4	64	$1.30 \times 10^6$	$101.22 \times 10^9$	32.46 / 0.898 5	28.76 / 0.785 4	27.68 / 0.740 1	26.47 / 0.796 3
6	48	$1.04 \times 10^6$	$74.95 \times 10^9$	32.46 / 0.898 6	28.81 / 0.786 3	27.71 / 0.740 5	26.53 / 0.799 1
6	64	$1.80 \times 10^6$	$129.75 \times 10^9$	32.52 / 0.899 0	28.82 / 0.786 5	27.72 / 0.740 6	26.62 / 0.801 0

数量不变的情况下, MDIEB的输出通道数对模型性能的影响。实验表明, 当MDIEB个数固定时, 网络模型的性能随着MDIEB输出通道数的增加而提升。具体而言, 在包含6个MDIEB的情况下, 输出通道数为64的网络模型在Set5和Urban100数据集上的PSNR值分别比48输出通道数的模型提升了0.06 dB和0.09 dB。然而, 这也导致了模型参数量和计算量的增加, 两者均增加了73%。

另一方面, 当MDIEB的输出通道数保持不变时, 通过增加MDIEB个数也能显著提升网络模型的重建性能。在输出通道数为48的情况下, 拥有6个MDIEB的网络相较于只有4个MDIEB的网络, 在4个数据集上的PSNR值分别提升了0.12 dB、0.12 dB、0.08 dB和0.23 dB, 而参数量仅增加了38.7%, 计算量增加了28.4%。从表7可以观察到, 仅增加MDIEB个数时, 参数量和计算量的增加幅度远小于增加输出通道数时的增幅。因此, 综合考虑模型的参数量和计算量, 本文模型最终采用了包含6个MDIEB, 每个MDIEB输出通道数为48的网络配置作为MDIESR算法的网络模型。这样的设计使MDIESR在模型性能、复杂度和计算开销之间实现了良好的平衡, 以相对较少的参数量和计算量获得了优越的重建性能。

### 3 结束语

本文提出了一种基于多域信息增强的轻量级图像超分辨重建网络, 该网络针对不同特征域设计了多样化的特征提取策略, 旨在通过精细化的信息处理, 优化全局与局部的低频和高频特征, 从而提升模型在细节恢复和图像重建中的表现。从空域角度出发, 将LKA与多尺度学习结合, 以较小的计算成本实现多尺度关联, 提高了模型对图像的整体布局及局部内容的关注度。从频域角度出发, 通过HFEB提供先验, 并进行频域分频处理, 将频率细化为全局高频与局部高频, 实现了不同区域高频特征的精确捕捉和优化。从转换域出发, 设计了MLKSA, 模拟高频特征提取, 避免了传统高频特征提取过程中的噪声干扰问题。此外, 在空域、频域及转换域上权衡了网络的性能和效率, 显著提升了重建图像的质量和效率。与性能相似的其他算法相比, 在参数量相当的情况下, 本文模型也能达到更好的客观标准和视觉效果。在4个基准数据集的

3个尺度因子下, 本文所提MDIESR在峰值信噪比与结构相似度上均超过了DiVANet, 在Urban100数据集上, MDIESR的4倍重建结果在峰值信噪比上提高了0.17 dB, 在结构相似度上也提高了0.0049。在下一步的研究工作中, 将会结合图像退化的机理, 构建更加轻量化的图像超分辨率重建网络。

### 参考文献:

- [1] CHENG D Q, CHEN L L, LV C, et al. Light-guided and cross-fusion U-Net for anti-illumination image super-resolution[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(12): 8436-8449.
- [2] CHANG H, YEUNG D Y, XIONG Y M. Super-resolution through neighbor embedding[C]//Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. Piscataway: IEEE Press, 2004: 1-8.
- [3] MEIJERING E, UNSER M. A note on cubic convolution interpolation[J]. IEEE Transactions on Image Processing, 2003, 12(4): 477-479.
- [4] SUN J, XU Z B, SHUM H Y. Image super-resolution using gradient profile prior[C]//Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2008: 1-8.
- [5] IRANI M, PELEG S. Improving resolution by image registration[J]. CVGIP: Graphical Models and Image Processing, 1991, 53(3): 231-239.
- [6] 寇旗旗, 程志威, 程德强, 等. 基于蓝图分离卷积的轻量化矿井图像超分辨率重建方法[J]. 煤炭学报, 2024, 49(9): 4038-4050.
- [7] KOU Q Q, CHENG Z W, CHENG D Q, et al. Lightweight super resolution method based on blueprint separable convolution for mine image[J]. Journal of China Coal Society, 2024, 49(9): 4038-4050.
- [8] CHEN L L, GUO L, CHENG D Q, et al. Structure-preserving and color-restoring up-sampling for single low-light image[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(4): 1889-1902.
- [9] 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.
- [10] 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 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 1874-1883.
- [11] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE Press, 2017: 1132-1140.
- [12] 程德强, 赵佳敏, 寇旗旗, 等. 多尺度密集特征融合的图像超分辨率重建[J]. 光学精密工程, 2022, 30(20): 2489-2500.
- [13] CHENG D Q, ZHAO J M, KOU Q Q, et al. Multi-scale dense feature fusion network for image super-resolution[J]. Optics and Precision Engineering, 2022, 30(20): 2489-2500.
- [14] ZHANG Y L, LIK P, LIK, et al. Image super-resolution using very

- deep residual channel attention networks[C]//Computer Vision-ECCV 2018. Berlin: Springer, 2018: 294-310.
- [13] NIU B, WEN W L, REN W Q, et al. Single image super-resolution via a holistic attention network[C]//Computer Vision-ECCV 2020. Berlin: Springer, 2020: 191-207.
- [14] WU G, JIANG J J, BAI Y C, et al. Incorporating transformer designs into convolutions for lightweight image super-resolution[J]. arXiv Preprint, arXiv: 2303.14324, 2023.
- [15] GUO M H, LU C Z, LIU Z N, et al. Visual attention network[J]. Computational Visual Media, 2023, 9(4): 733-752.
- [16] LAU K W, PO L M, REHMAN Y A U. Large separable kernel attention: rethinking the large kernel attention design in CNN[J]. Expert Systems with Applications, 2024, 236: 121352.
- [17] QU L Q, ZHANG Y Q, WANG S, et al. Synthesized 7T MRI from 3T MRI via deep learning in spatial and wavelet domains[J]. Medical Image Analysis, 2020, 62: 101663.
- [18] IBRAHIM V, PAUL J S. A frequency domain network for hybrid-domain super-resolution MRI[C]//Proceedings of the International Conference on Signal Processing and Communications (SPCOM). Piscataway: IEEE Press, 2024: 1-5.
- [19] CHEN X Y, WANG X T, ZHOU J T, et al. Activating more pixels in image super-resolution transformer[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2023: 22367-22377.
- [20] WENG X T, CHEN Y, ZHENG Z C, et al. A high-frequency focused network for lightweight single image super-resolution[J]. arXiv Preprint, arXiv: 2303.11701, 2023.
- [21] 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 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2023: 12480-12490.
- [22] LU Z S, LI J C, LIU H, et al. Transformer for single image super-resolution[C]//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE Press, 2022: 456-465.
- [23] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30(5): 256-264.
- [24] 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[J]. arXiv Preprint, arXiv: 2204.13286, 2022.
- [25] JI L, ZHU Q H, ZHANG Y Q, et al. Cross-domain heterogeneous residual network for single image super-resolution[J]. Neural Networks, 2022, 149: 84-94.
- [26] CHEN H G, DONG L, YANG H, et al. Unsupervised real-world image super-resolution via dual synthetic-to-realistic and realistic-to-synthetic translations[J]. IEEE Signal Processing Letters, 2022, 29: 1282-1286.
- [27] HU X W, CAI Y H, LIN J, et al. HDNet: high-resolution dual-domain learning for spectral compressive imaging[C]//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2022: 17521-17530.
- [28] CHEN C, WANG Y C, ZHANG Y X, et al. Remote sensing hyperspectral image super-resolution via multidomain spatial information and multiscale spectral information fusion[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5515216.
- [29] LI Q, YUAN Y, WANG Q. Hyperspectral image super-resolution via multi-domain feature learning[J]. Neurocomputing, 2022, 472: 85-94.
- [30] ZHOU B, DEY N, SCHLEMPER J, et al. DSFormer: a dual-domain self-supervised transformer for accelerated multi-contrast MRI reconstruction[C]//Proceedings of the 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE Press, 2023: 4955-4964.
- [31] ZHANG Y, CHEN H, XIA W J, et al. LEARN++: recurrent dual-domain reconstruction network for compressed sensing CT[J]. IEEE Transactions on Radiation and Plasma Medical Sciences, 2022, 7(2): 132-142.
- [32] PAN J Y, ZHANG H Y, WU W F, et al. Multi-domain integrative Swin transformer network for sparse-view tomographic reconstruction[J]. Patterns, 2022, 3(6): 100498.
- [33] WU Z J, LIU W H, LI J, et al. SFHN: spatial-frequency domain hybrid network for image super-resolution[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(11): 6459-6473.
- [34] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding[C]//Proceedings of the British Machine Vision Conference 2012. Piscataway: IEEE Press, 2012: 1-10.
- [35] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations[C]//Curves and Surfaces. Berlin: Springer, 2012: 711-730.
- [36] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 1646-1654.
- [37] 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 (CVPR). Piscataway: IEEE Press, 2015: 5197-5206.
- [38] HORÉ A, ZIOU D. Image quality metrics: PSNR vs. SSIM[C]//Proceedings of the 2010 20th International Conference on Pattern Recognition. Piscataway: IEEE Press, 2010: 2366-2369.
- [39] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [40] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 1637-1645.
- [41] AHN N, KANG B, SOHN K A. Fast, accurate, and lightweight super-resolution with cascading residual network[C]//Computer Vision-ECCV 2018. Berlin: Springer, 2018: 256-272.
- [42] 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 Press, 2019: 2024-2032.
- [43] WANG L G, DONG X Y, WANG Y Q, et al. Exploring sparsity in im-

age super-resolution for efficient inference[C]//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2021: 4917-4926.

- [44] DU Z C, LIU D, LIU J, et al. Fast and memory-efficient network towards efficient image super-resolution[C]//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE Press, 2022: 852-861.
- [45] PARK K, SOH J W, CHO N I. A dynamic residual self-attention network for lightweight single image super-resolution[J]. IEEE Transactions on Multimedia, 2021, 25: 907-918.
- [46] CHOI H, LEE J, YANG J. N-gram in swin transformers for efficient lightweight image super-resolution[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2023: 2071-2081.
- [47] BEHJATI P, RODRIGUEZ P, FERNÁNDEZ C, et al. Single image super-resolution based on directional variance attention network[J]. Pattern Recognition, 2023, 133: 108997.
- [48] WANG Y, ZHANG T. OSFFNet: omni-stage feature fusion network for lightweight image super-resolution[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2024, 38(6): 5660-5668.
- [49] LIU Y Q, JIA Q, ZHANG J, et al. Hierarchical similarity learning for aliasing suppression image super-resolution[J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(2): 2759-2771.

#### [作者简介]



寇旗旗 (1988-), 男, 河南襄城人, 博士, 中国矿业大学副教授、硕士生导师, 主要研究方向为图像处理/AI视频分析、智能检测与模式识别、图像增强与复原。



刘规 (1999-), 男, 重庆酉阳人, 中国矿业大学硕士生, 主要研究方向为图像处理、图像超分辨率重建。



江鹤 (1990-), 男, 江苏徐州人, 博士, 中国矿业大学讲师, 主要研究方向为图像修复、图像识别和三维重建。



陈亮亮 (1993-), 男, 江苏徐州人, 博士, 中国矿业大学讲师, 主要研究方向为图像处理、图像增强和超分辨。



程德强 (1979-), 男, 河南洛阳人, 博士, 中国矿业大学教授、博士生导师, 主要研究方向为图像智能检测与模式识别、图像处理与视频编码。