

# 控制与决策

Control and Decision

## 基于高频信息调制的轻量级单幅图像超分辨率网络

江鹤, 顾豪, 郑州, 乙夫迪, 王子强, 程德强

引用本文:

江鹤, 顾豪, 郑州, 等. 基于高频信息调制的轻量级单幅图像超分辨率网络[J]. *控制与决策*, 2026, 41(1): 201-212.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2024.1471>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### [基于混合模型驱动的红外与可见光图像融合](#)

Infrared and visible image fusion based on hybrid model driving

*控制与决策*. 2021, 36(9): 2143-2151 <https://doi.org/10.13195/j.kzyjc.2019.1749>

#### [基于自注意力生成对抗网络的图像超分辨率重建](#)

Image super-resolution reconstruction based on self-attention GAN

*控制与决策*. 2021, 36(6): 1324-1332 <https://doi.org/10.13195/j.kzyjc.2019.1290>

#### [一种反向递推正弦干扰观测器的设计方法](#)

[Design method of reverse recursive sinusoidal disturbance observer](#)

*控制与决策*. 2021, 36(2): 403-410 <https://doi.org/10.13195/j.kzyjc.2019.0769>

#### [基于FRC框架的红外与可见光图像融合方法](#)

Infrared and visible image fusion based on FRC algorithm

*控制与决策*. 2021, 36(11): 2690-2698 <https://doi.org/10.13195/j.kzyjc.2020.0669>

#### [基于协同聚类和权重注意力稀疏自编码网络的变化检测方法](#)

Change detection approach based on cooperative clustering and weighted attention sparse autoencoder

*控制与决策*. 2021, 36(10): 2442-2450 <https://doi.org/10.13195/j.kzyjc.2019.1633>

# 基于高频信息调制的轻量级单幅图像超分辨率网络

江鹤, 顾豪, 郑州, 乙夫迪, 王子强, 程德强<sup>†</sup>

(中国矿业大学 信息与控制工程学院, 江苏 徐州 221116)

**摘要:** 在图像超分辨率领域, 高频信息的精确度和重建图像的清晰度密切相关。目前主流模型多依赖于稠密连接或简单串行的模块堆叠方法来提取高频信息, 并借助空域注意力机制来优化图像的重建。这种方法虽然有一定的效果, 但不可避免地会造成高频信息的冗余或重建图像精度不足的情况。为此, 引入一种基于高频信息调制的图像超分辨率网络。该网络巧妙地融合了自调制和互调制策略, 以优化图像的重建流程。自调制模块集成了多尺度采样网络与高频引导的空间注意力机制, 有效滤除了冗余信息, 并同步校准了空域与频域的信息。而互调制模块则通过构建自调制信息之间的非线性映射关系, 进一步提升了高频信息的准确性。经过一系列实验验证, 所提出的模型与当前流行的模型相比, 不仅在量化指标上表现出色, 保持了参数的轻量化, 同时具有出色的视觉效果, 充分表明了其在图像超分辨率重建方面的优越性。

**关键词:** 超分辨率; 轻量级; 高频信息调制; 自调制; 互调制; 频域注意力机制

中图分类号: TP391 文献标志码: A

DOI: 10.13195/j.kzyjc.2024.1471

引用格式: 江鹤, 顾豪, 郑州, 等. 基于高频信息调制的轻量级单幅图像超分辨率网络 [J]. 控制与决策, 2026, 41(1): 201-212.

## Lightweight single image super resolution based on high-frequency information modulation

JIANG He, GU Hao, ZHENG Zhou, YI Fu-di, WANG Zi-qiang, CHENG De-qiang<sup>†</sup>

(School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China)

**Abstract:** In the field of single image super resolution, the accuracy of high-frequency information and the clarity of the reconstructed images are closely linked. Current mainstream models rely on dense connections or module stacking to extract high-frequency information, optimizing image reconstruction through mechanisms like spatial domain attention. While effective, these methods often lead to redundancy in high-frequency information and a subsequent degradation in accuracy. To address this issue, this study proposes an image super resolution network that modulates high-frequency information. The network integrates self modulation and mutual modulation strategies to enhance the image reconstruction process. Specifically, the self modulation module combines a multi-scale sampling network with a high-frequency guided spatial attention mechanism, which effectively filters out redundant information while simultaneously calibrating data in both spatial and frequency domains. The mutual modulation module further refines the accuracy of high-frequency information by establishing a nonlinear relationship between the self modulated outputs. Through extensive experimental validation, it is demonstrated that the proposed model has superior performance in terms of quantization metrics and maintains parameter efficiency compared to current popular models. It also delivers outstanding visual results, fully showcasing its advantages in single image super resolution reconstruction.

**Keywords:** super resolution; lightweight; high-frequency information modulation; self modulation; mutual modulation; frequency-domain attention mechanism

## 0 引言

单幅图像超分辨率 (SISR) 是一项经典的计算机

视觉任务, 在医学图像<sup>[1-3]</sup>、遥感监测<sup>[4-6]</sup>、计算摄影<sup>[7-8]</sup>

等领域应用广泛。其核心目标是将低分辨率 (LR) 图

收稿日期: 2024-12-21; 录用日期: 2025-05-26.

基金项目: 国家自然科学基金项目 (52304182, 52204177); 国家重点研发计划项目 (2023YFC2907600, 2021YFC2902701, 2021YFC2902702); 徐州市科技成果转化项目 (KC23369).

责任编辑: 张文安.

<sup>†</sup>通信作者. E-mail: chengdq@cumt.edu.cn.

像重建为高分辨率 (HR) 图像. Dong 等<sup>[9]</sup> 率先将卷积神经网络引入 SISR 任务, 提出了经典的 SRCNN 网络. 经过多年的发展, 大量优秀的图像超分辨率模型相继涌现, 如 EDSR<sup>[10]</sup>、RCAN<sup>[11]</sup> 和 PGRN<sup>[12]</sup> 等. 这些模型在提升图像重建性能的同时, 往往伴随较大的参数量. 轻量级 SISR 模型因参数量低、速度快、可以运行在资源受限的实际场景中, 成为了研究热点. 然而, 轻量级模型面临两个问题: 1) 神经网络中卷积层的核大小相对固定, 易导致神经网络感受野单一; 2) 串联的网络结构缺乏对层间特征依赖关系的关注. 稠密连接可以捕捉层间特征的相关性, 但方式粗糙, 易引发高频信息冗余问题<sup>[13]</sup>. 鉴于此, 本文提出基于高频信息调制的超分辨率重建网络, 主要工作如下:

1) 提出高频信息自调制 (SM) 模块, 将多尺度采样网络与高频引导的空间注意力机制有效结合, 实现了信息的多尺度精确拟合.

2) 构建高频信息互调制 (MM) 模块, 建模高频自调制信息间的非线性映射关系, 提高网络对高频信息的表征能力.

3) 实现基于高频信息调制的超分辨率重建网络 (HIM), 在基准数据集上与流行的轻量级 SISR 模型对比, 其客观指标与视觉效果都取得了最优的水平.

## 1 相关工作

### 1.1 轻量级图像超分辨率模型

在轻量级 SISR 模型的研究中, Kim 等<sup>[14]</sup> 率先提出了基于递归学习的模型 (DRCN), 通过权重共享机制将模型参数控制在 2 M 以内, 取得了较好的效果. 但递归学习常常伴随梯度消失和训练难度大的问题. 为了实现轻量级图像超分辨率模型, Hui 等<sup>[15]</sup> 将信息蒸馏的概念引入 SISR 任务, 提出了信息多蒸馏网络 (IMDN). 该网络通过通道分割的方法减少通道数量, 从而降低网络参数量. 受此启发, 越来越多基于蒸馏技术的轻量化模型被提出, 如 RFDN<sup>[16]</sup> 和 GFEDNet<sup>[17]</sup>, 并取得了优秀的表现. Yu 等<sup>[18]</sup> 则提出了 DIPNet, 通过剪枝技术剔除不必要的连接, 提高了模型的效率. Sun 等<sup>[19]</sup> 结合通道洗牌和大核卷积, 提取不同通道特征之间的相关性; Zhou 等<sup>[20]</sup> 则通过扩大感受野, 引入深度可分离卷积和像素归一化操作, 在减少网络参数的同时提升模型性能; 程德强等<sup>[21]</sup> 提出一种浅层特征调制网络, 探索了浅层特征与深层特征间的映射机制.

简单地堆叠特征提取模块可能导致图像高频信

息无法有效表征, 而使用固定卷积核提取的特征过于单一, 从而影响网络的性能. 此外, 许多模型忽视了特征之间的映射关系, 导致特征间缺乏有效的信息交互. 因此, 如何构建高效的特征映射关系, 依然是当前研究中的一个重要挑战.

### 1.2 注意力机制

在图像超分辨率重建任务中, 注意力机制得到了广泛应用, 典型模型包括通道注意力机制的 RCAN<sup>[11]</sup>、增强空间注意力机制的 RFANet<sup>[22]</sup>、像素注意力机制的 PAN<sup>[23]</sup> 等. 为了实现不同模块间提取到特征的有效融合, 程德强等<sup>[24]</sup> 提出了一种基于渐进式感受野的图像重建方法, 实现了网络中特征的高效交互; 韩玉兰等<sup>[25]</sup> 设计了一种树状多分支的残差网络, 每条分支上采用双通道残差策略进行连接, 模块内又融入通道空间注意力机制, 有效地拟合了高频细节特征. 近年来, 基于自注意力机制的模型, 如 SwinIR<sup>[26]</sup>, 也被成功应用于图像超分辨率任务, 通过移动窗口机制模拟远程依赖, 捕捉图像特征与注意力权重之间的交互信息.

尽管上述模型均取得了一定成效, 但仍存在两个主要问题: 1) 这些注意力机制过度关注空域信息, 忽视了频域信息的作用; 2) 由于其自身的局限性, 模型训练时间较长且计算资源消耗较大, 限制了其在实际应用中的效果和可行性. 为此, 本研究提出一种基于高频引导的即插即用空间注意力机制 (HGSA). 实验结果表明, HGSA 在同时关注空域特征和频域特征的基础上, 有效引导模型重建图像的高频细节区域. 与传统的增强空间注意力机制 (ESA)<sup>[22]</sup> 相比, 本研究提出的注意力机制保持低参数量消耗的同时, 还能从多个领域关注对图像重建有益的信息, 大幅度提高了重建图像的质量.

## 2 本研究方法

基于高频信息调制的轻量级图像超分辨率重建网络框架如图 1(a) 所示, 主要包含 3 个部分: 浅层特征提取、深层特征映射以及图像重建.

### 2.1 浅层特征提取

浅层特征提取由一个尺寸为  $3 \times 3$  的卷积核实现, 即

$$\mathbf{F}_S = \mathbf{C}_{3 \times 3}(\mathbf{I}_{LR}). \quad (1)$$

其中:  $\mathbf{I}_{LR}$  为低分辨率图像,  $\mathbf{C}_{3 \times 3}$  为卷积核大小为  $3 \times 3$  的卷积,  $\mathbf{F}_S$  为浅层特征.

### 2.2 深层特征映射

深层特征映射由 6 个高频信息互调制 (MM) 模块以递归方式连接组成, 其中该模块用  $\mathcal{H}_{MM}^6(\cdot)$  表示,

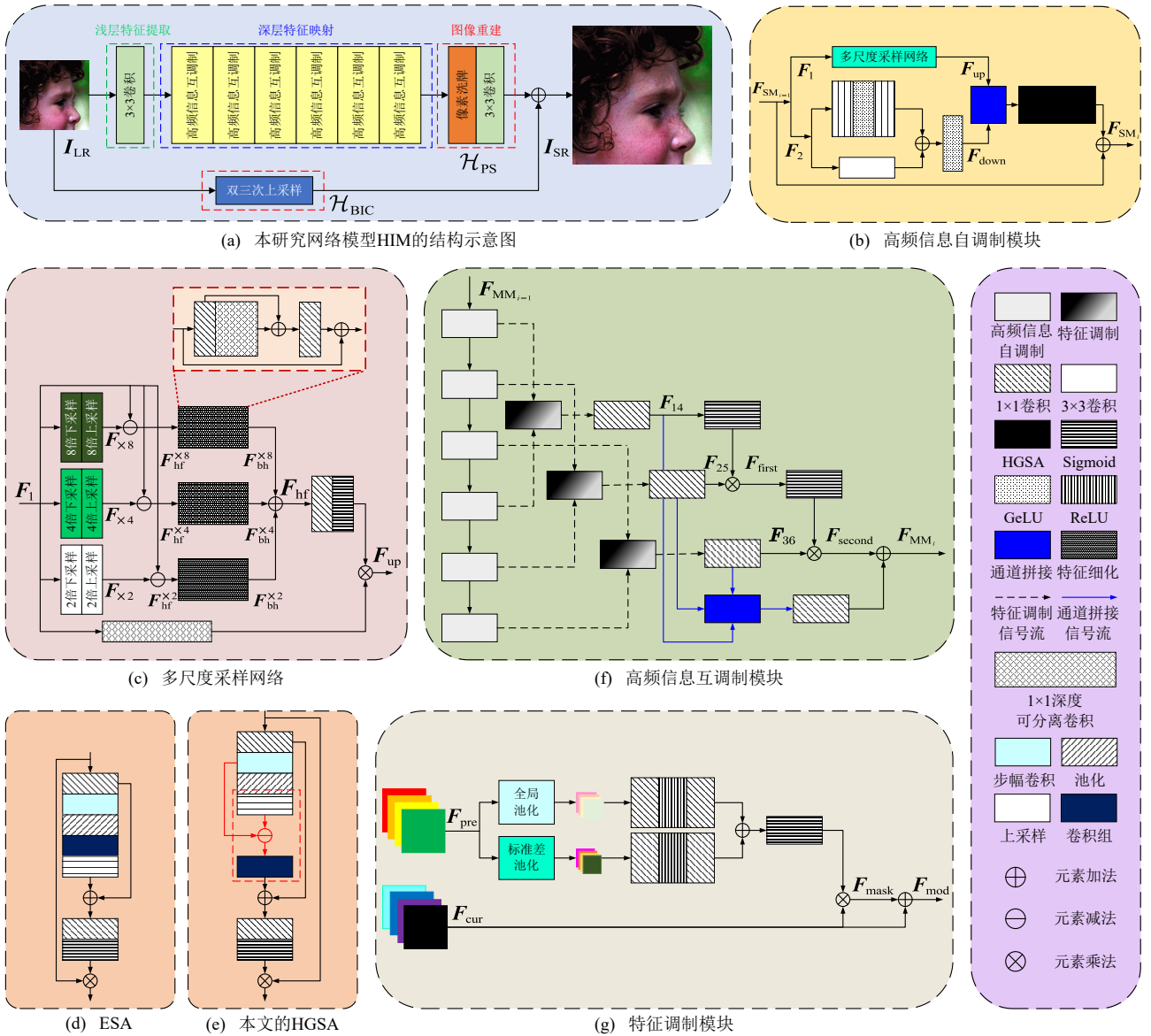


图1 本研究网络模型 HIM 的结构

$F_D$ 是深层特征,有

$$F_D = \mathcal{H}_{MM}^6(F_S). \quad (2)$$

### 2.2.1 高频信息自调制模块

图像退化过程中丢失的大部分都是高频信息,即包含图像边缘、纹理细节等关键信息.为了加强对高频信息的利用,解决轻量级网络高频信息提取不准确的问题,本研究提出一种高频信息自调制(SM)模块,如图1(b)所示.

$$F_1, F_2 = \mathcal{H}_{SPL}(F_{SM_{i-1}}), \quad (3)$$

$$F_{up} = \mathcal{H}_{MSN}(F_1), \quad (4)$$

$$F_{down} = \varphi(C_{3 \times 3}(\varphi(C_{3 \times 3}(F_2)) \oplus C_{3 \times 3}(F_2))), \quad (5)$$

$$F_{SM_i} = \mathcal{H}_{HGSA}(\mathcal{H}_{CAT}(F_{up}, F_{down})) \oplus F_{SM_{i-1}}. \quad (6)$$

式(3)~(6)展示了这一过程.其中: $\mathcal{H}_{SPL}(\cdot)$ 为通道分割; $F_{SM_{i-1}}$ 为该模块的输入, $F_1$ 和 $F_2$ 分别为上

下分支的输入,由 $F_{SM_{i-1}}$ 经过通道分割得到; $\mathcal{H}_{MSN}(\cdot)$ 为多尺度采样网络; $F_{up}$ 和 $F_{down}$ 分别为上、下分支的输出; $\varphi(\cdot)$ 为激活函数 GeLU<sup>[27]</sup>; $\mathcal{H}_{CAT}(\cdot)$ 为特征通道拼接操作; $\mathcal{H}_{HGSA}(\cdot)$ 为高频引导的空间注意力模块; $F_{SM_i}$ 为该模块的输出; $i$ 的取值范围为2~6.

### 2.2.2 多尺度采样网络

现有的超分辨率网络通过增加网络层数来提取高频信息,但这种方法易导致特征单一且精度较低.基于此,本研究提出一种多尺度采样网络,通过采样的方式来实现高频信息的提取,大大减少了堆叠卷积层所带来的参数量消耗.此外,还引入特征细化模块(FRB),用于进一步细化提取到的高频信息.具体实现细节如图1(c)所示.通过对输入特征进行8倍、4倍和2倍的下采样,以完成多尺度高频信息的提取.这一步骤弥补了主流算法在特征提取操作中,使

用单一卷积核带来的感受野固定、提取特征单一等问题。

$$\begin{cases} \mathbf{F}_{\times 8} = \mathcal{H}_{\text{up}}^{\times 8}(\mathcal{H}_{\text{down}}^{\times 8}(\mathbf{F}_1)), \\ \mathbf{F}_{\times 4} = \mathcal{H}_{\text{up}}^{\times 4}(\mathcal{H}_{\text{down}}^{\times 4}(\mathbf{F}_1)), \\ \mathbf{F}_{\times 2} = \mathcal{H}_{\text{up}}^{\times 2}(\mathcal{H}_{\text{down}}^{\times 2}(\mathbf{F}_1)). \end{cases} \quad (7)$$

其中:  $\mathcal{H}_{\text{down}}^{\times 8}(\cdot)$ 、 $\mathcal{H}_{\text{down}}^{\times 4}(\cdot)$ 、 $\mathcal{H}_{\text{down}}^{\times 2}(\cdot)$ 为8倍、4倍、2倍下采样;  $\mathcal{H}_{\text{up}}^{\times 8}(\cdot)$ 、 $\mathcal{H}_{\text{up}}^{\times 4}(\cdot)$ 、 $\mathcal{H}_{\text{up}}^{\times 2}(\cdot)$ 为8倍、4倍、2倍的双三次上采样,  $\mathbf{F}_{\times 8}$ 、 $\mathbf{F}_{\times 4}$ 、 $\mathbf{F}_{\times 2}$ 为8倍、4倍、2倍采样后的输出特征。

$$\begin{cases} \mathbf{F}_{\text{hf}}^{\times 8} = \mathbf{F}_1 \ominus \mathbf{F}_{\times 8}, \\ \mathbf{F}_{\text{hf}}^{\times 4} = \mathbf{F}_1 \ominus \mathbf{F}_{\times 4}, \\ \mathbf{F}_{\text{hf}}^{\times 2} = \mathbf{F}_1 \ominus \mathbf{F}_{\times 2}; \end{cases} \quad (8)$$

$$\begin{cases} \mathbf{F}_{\text{bh}}^{\times 8} = \alpha(\mathcal{H}_{\text{FRB}}(\mathbf{F}_{\text{hf}}^{\times 8})), \\ \mathbf{F}_{\text{bh}}^{\times 4} = \beta(\mathcal{H}_{\text{FRB}}(\mathbf{F}_{\text{hf}}^{\times 4})), \\ \mathbf{F}_{\text{bh}}^{\times 2} = \gamma(\mathcal{H}_{\text{FRB}}(\mathbf{F}_{\text{hf}}^{\times 2})); \end{cases} \quad (9)$$

$$\mathbf{F}_{\text{hf}} = \mathbf{F}_{\text{bh}}^{\times 8} \oplus \mathbf{F}_{\text{bh}}^{\times 4} \oplus \mathbf{F}_{\text{bh}}^{\times 2}. \quad (10)$$

其中:  $\mathbf{F}_{\text{hf}}^{\times 8}$ 、 $\mathbf{F}_{\text{hf}}^{\times 4}$ 、 $\mathbf{F}_{\text{hf}}^{\times 2}$ 为多尺度采样网络提取到的8倍、4倍、2倍的高频信息;  $\mathcal{H}_{\text{FRB}}(\cdot)$ 为特征细化模块的传递函数. 为了降低参数量,  $\mathcal{H}_{\text{FRB}}(\cdot)$ 由两个卷积核大小为 $1 \times 1$ 的卷积和一个卷积核大小为 $1 \times 1$ 的深度可分离卷积组成;  $\alpha$ 、 $\beta$ 、 $\gamma$ 为各支路的特征细化矩阵;  $\mathbf{F}_{\text{bh}}^{\times 8}$ 、 $\mathbf{F}_{\text{bh}}^{\times 4}$ 、 $\mathbf{F}_{\text{bh}}^{\times 2}$ 依次为8倍、4倍、2倍3条支路的输出特征;  $\mathbf{F}_{\text{hf}}$ 为提取到的高频信息. 最终的高频输出特征为

$$\mathbf{F}_{\text{up}} = \sigma(\mathcal{C}_{1 \times 1}(\mathbf{F}_{\text{hf}})) \otimes \text{DWC}_{1 \times 1}(\mathbf{F}_1). \quad (11)$$

其中:  $\sigma(\cdot)$ 为Sigmoid激活函数,  $\otimes$ 为元素相乘,  $\mathcal{C}_{1 \times 1}(\cdot)$ 为 $1 \times 1$ 卷积,  $\text{DWC}_{1 \times 1}(\cdot)$ 为卷积核大小为 $1 \times 1$ 的深度可分离卷积.

### 2.2.3 高频引导的空间注意力机制

ESA<sup>[22]</sup>通过聚合空间上下文信息,提高了特征的空间表达能力,如图1(d)所示.

ESA只关注空域信息,忽视了对超分辨率任务十分重要的高频信息. 基于此,本研究提出了基于高频引导的空间注意力机制HGSA,如图1(e)所示. 该机制让网络兼顾空域信息的同时,增强其对频域信息的关注度. 首先对输入特征进行池化,然后通过上采样恢复到原始尺寸大小,最后通过元素相减的方式提取出高频信息,并将其作为卷积组的输入,进一步增强高频信息.

### 2.2.4 高频信息互调制模块

当前,端到端网络普遍采用串联结构,这种提取方式存在局限性,在网络训练过程中,反向传播的参数容易梯度消失. 为解决这一问题,EDSR<sup>[10]</sup>引入残

差学习解决了该问题,但它未能有效利用神经网络的前期特征. RDN(residual dense network)<sup>[13]</sup>通过稠密连接,实现了对所有前期特征的复用. 这种无差别的信息处理方式也带来了新的问题——信息冗余. 为了克服这一挑战,提出高频信息互调制模块,旨在通过非线性映射关系的有效表征,实现深层特征映射模块间的精确调控,从而在保持信息丰富性的同时减少不必要的信息冗余.

如图1(f)所示,MM模块的具体实现细节为:将前期3个自调制(SM)模块的输出通过特征调制层生成调制掩码,对后期3个SM模块的输出进行调制,实现对不同阶段高频自调制信息非线性映射关系的建模,具体流程如下:

$$\mathbf{F}_{ij} = \mathcal{C}_{1 \times 1}(\mathcal{H}_{\text{MOD}}(\mathbf{F}_{\text{SM}_i}, \mathbf{F}_{\text{SM}_j})), \quad (12)$$

$$\mathbf{F}_{\text{first}} = \sigma(\mathbf{F}_{14}) \otimes \mathbf{F}_{25}, \quad (13)$$

$$\mathbf{F}_{\text{second}} = \sigma(\mathbf{F}_{\text{first}}) \otimes \mathbf{F}_{36}, \quad (14)$$

$$\mathbf{F}_{\text{MM}_i} = \mathcal{C}_{1 \times 1}(\mathcal{H}_{\text{CAT}}(\mathbf{F}_{14}, \mathbf{F}_{25}, \mathbf{F}_{36})) \oplus \mathbf{F}_{\text{second}}. \quad (15)$$

其中:  $\mathcal{C}_{1 \times 1}$ 是卷积核大小为 $1 \times 1$ 的卷积;  $\mathcal{H}_{\text{MOD}}(\cdot)$ 为特征调制层的传递函数;  $\mathbf{F}_{\text{SM}_i}$ 为当前MM模块中第*i*个SM模块的输出特征;  $\mathbf{F}_{ij}$ 为第*i*个SM模块和第*j*个SM模块的输出经过调制得到的调制掩码, *i*的取值范围是 $1 \sim 3$ ,  $j = i + 3$ ;  $\mathbf{F}_{\text{first}}$ 、 $\mathbf{F}_{\text{second}}$ 、 $\mathbf{F}_{\text{MM}_i}$ 分别为一阶段校准的输出特征、二阶段校准的输出特征以及第*i*个MM模块的输出特征;  $\sigma(\cdot)$ 为Sigmoid激活函数;  $\mathcal{H}_{\text{CAT}}(\cdot)$ 为特征通道拼接操作.

为了更好地对调制后的特征进行融合,受二阶注意力网络SAN<sup>[28]</sup>的启发,本研究对调制后的特征进行多阶段校准融合.

与稠密连接不同,本研究提出的高频信息互调制生成的调制掩码,利用特征调制实现对各层自调制信息映射关系的表征,如图1(g)所示. 这种方式高效利用了前期的特征信息.

本研究采用前期特征 $\mathbf{F}_{\text{pre}}$ 的两个常用统计量(即均值和标准差)生成掩码矩阵. 具体而言,首先对 $\mathbf{F}_{\text{pre}}$ 分别进行全局池化和标准差池化,得到前期特征的均值和标准差统计量;然后将两组特征分别经过 $1 \times 1$ 卷积核、非线性激活函数ReLU和Sigmoid激活函数处理;最后将两组特征元素相加,得到调制掩码,即

$$\mathbf{F}_{\text{mask}} = \sigma((\mathcal{C}_{1 \times 1}(\delta(\mathcal{C}_{1 \times 1}(\mathcal{H}_{\text{GAP}}(\mathbf{F}_{\text{pre}})))))) \oplus (\mathcal{C}_{1 \times 1}(\delta(\mathcal{C}_{1 \times 1}(\mathcal{H}_{\text{SDP}}(\mathbf{F}_{\text{pre}}))))), \quad (16)$$

$$\mathbf{F}_{\text{mod}} = \mathcal{H}_{\text{MOD}}(\mathbf{F}_{\text{pre}}, \mathbf{F}_{\text{cur}}) = \mathbf{F}_{\text{mask}} \otimes \mathbf{F}_{\text{cur}} \oplus \mathbf{F}_{\text{cur}}. \quad (17)$$

其中:  $\mathbf{F}_{\text{pre}}$ 、 $\mathbf{F}_{\text{cur}}$ 、 $\mathbf{F}_{\text{mask}}$ 、 $\mathbf{F}_{\text{mod}}$  分别为前期特征、当前特征、调制掩码和调制特征;  $\mathcal{H}_{\text{MOD}}(\cdot)$  为特征调制层的传递函数;  $\mathcal{H}_{\text{GAP}}(\cdot)$ 、 $\mathcal{H}_{\text{SDP}}(\cdot)$  分别为全局平均池化和标准差池化;  $\delta(\cdot)$  为 ReLU 激活函数;  $\sigma(\cdot)$  为 Sigmoid 激活函数。

### 2.3 图像重建

鉴于亚像素卷积网络 (ESPCN)<sup>[29]</sup> 的高效性, 本研究在图像重建时采用了它的像素洗牌层。

$$\mathbf{I}_{\text{SR}} = \mathcal{C}_{3 \times 3}(\mathcal{H}_{\text{PS}}(\mathbf{F}_{\text{D}})) \oplus \mathcal{H}_{\text{BIC}}(\mathbf{I}_{\text{LR}}). \quad (18)$$

其中:  $\mathcal{H}_{\text{PS}}(\cdot)$  为像素洗牌,  $\mathcal{H}_{\text{BIC}}(\cdot)$  为双三次插值,  $\mathbf{I}_{\text{SR}}$  为重建后的超分辨率图像,  $\oplus$  为元素加法。

### 2.4 损失函数

本模型训练中采用对数据波动较敏感的  $L_1$  损失函数

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N \|\mathcal{H}_{\text{HM}}(\mathbf{s}_i) - \mathbf{S}_i\|_1. \quad (19)$$

其中:  $[\mathbf{s}_i, \mathbf{S}_i]_{i=1}^N$  为训练集,  $N$  为样本个数,  $\mathbf{s}_i$  为低分辨率图像块,  $\mathbf{S}_i$  则是与之对应的高分辨率图像块,  $\mathcal{H}_{\text{HM}}(\mathbf{s}_i)$  为重建后的超分辨率图像块。

## 3 实验结果与分析

### 3.1 实验设置

#### 3.1.1 数据集

本研究采用图像超分辨率领域常用的数据集 DIV2K<sup>[30]</sup>, 其中包含 900 张图片. 在训练过程中, 将前 800 张图片作为训练集, 后 100 张图片作为验证集. 为评估本研究网络模型的性能, 在国际通用的 4 个基准数据集上进行测试, 它们分别是 Set5<sup>[31]</sup>、Set14<sup>[32]</sup>、B100<sup>[33]</sup> 和 Urban100<sup>[34]</sup>, 这些数据集中的图像有着复杂的结构和挑战性的纹理。

#### 3.1.2 质量评价指标

本研究用图像  $Y$  通道的峰值信噪比 (PSNR) 和结构相似度 (SSIM) 作为质量评价指标. PSNR 和 SSIM 的值越高, 表明图像的重建质量越好。

$$\text{PSNR} = 10 \times \lg \left( \frac{\text{MAX}^2}{\text{MSE}} \right), \quad (20)$$

$$\text{SSIM} = \frac{(2\mu_1\mu_2 + C_1)(2\sigma_{12} + C_2)}{(\mu_1^2 + \mu_2^2 + C_1)(\sigma_1^2 + \sigma_2^2 + C_2)}. \quad (21)$$

其中:  $\text{MAX} = 255$ ,  $\text{MSE}$  为均方误差;  $\mu_1$  和  $\mu_2$  为真实图像和重建图像的亮度均值;  $\sigma_1$  和  $\sigma_2$  为真实图像和重建图像的方差;  $\sigma_{12}$  为真实图像和重建图像的协方差;  $C_1$  和  $C_2$  为常数。

#### 3.1.3 训练细节

采用随机 90° 旋转和水平翻转进行数据增强. 预

处理后, 训练图像被裁剪成  $48 \times 48$  的小块, 每次随机输入 32 个小块. 实验使用 Adam 优化器<sup>[35]</sup>, 其参数为  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\varepsilon = 10^{-8}$ . 网络的通道数为 48, 初始学习率为  $8 \times 10^{-4}$ . 训练周期为 1 000, 每 200 个迭代周期学习率减半. 实验基于 Pytorch 框架, 并在 NVIDIA 1080Ti GPU 上进行训练和测试。

### 3.2 模型对比

#### 3.2.1 客观量化指标比较

在表 1 中, 将本研究所提出的网络与目前先进的轻量级图像超分辨率重建网络进行了比较. 包括: IMDN<sup>[15]</sup>、FMEN<sup>[36]</sup>、ShuffleMixer<sup>[37]</sup>、LatticeNet<sup>[38]</sup>、DRSAN<sup>[39]</sup>、DLSR<sup>[40]</sup>、DDistill-SR<sup>[41]</sup>、MICU<sup>[42]</sup>、HAFRN<sup>[43]</sup>、LBRN<sup>[44]</sup>. 通过对比指标以及网络参数量分析, 清楚地展示了网络模型在每个数据集上达到的最佳结果。

#### 3.2.2 与 Transformer 模型的比较

选取先进的 Transformer 模型进行比较, 包括 LBNNet<sup>[45]</sup>、NGSwin<sup>[46]</sup>、TCSR-B<sup>[47]</sup>. 由表 2 可知, 本模型仅在缩放因子为 2 时的 SSIM 指标是次优, 在其他情况下, 尤其在恢复难度较高的 3 和 4 缩放因子下, 均是最优. 另外, 本研究使用的注意力是即插即用的, 相比于 Transformer 的自注意力机制, 神经网络的训练成本被大大降低。

从表 1 可以看出, 本研究提出的方法在不同缩放因子下皆取得了最优结果, 该优势在 4 倍的缩放因子下尤为明显. 以细节信息最具挑战性的 Urban100<sup>[34]</sup> 数据集为例, 与 2024 年模型 HAFRN<sup>[43]</sup> 相比, 本研究模型的量化指标 PSNR 和 SSIM 分别提高了 0.43 dB 和 0.010 3. 与参数量接近的 DRSAN<sup>[39]</sup> 相比, PSNR 提高了 0.19 dB, SSIM 提高了 0.003 2. 这是因为随着尺度因子的增大, 图像在退化过程中丢失的高频信息也增多, 导致恢复这些信息的难度增加, 而本研究提出的高频信息调制模块能够有效地恢复这些丢失的高频信息。

#### 3.2.3 主观视觉效果比较

为直观比较不同模型的重建效果, 选取 Urban100 数据集中纹理复杂的图像进行 4 倍放大重建, 并展示视觉效果对比如图 2 所示. 实验结果表明, 本研究模型重建的图像更自然、逼真, 而其他模型的结果与真实图像存在明显差异. 在重复网格区域 (图 2(a)), 本模型能够准确还原图像的排列特征, 其他模型的结果则出现模糊或扭曲变形. 对于建筑物结构 (图 2(b)), 本模型的重建效果细节保留更完整, 特别是在图 2(c) 的网格曲线重建中, 本模型结果与原始图像吻合度

表1 不同缩放因子在4个基准数据集下的指标对比

缩放因子	模型名	参数量/K	期刊	Set5 <sup>[31]</sup>	Set14 <sup>[32]</sup>	B100 <sup>[33]</sup>	Urban100 <sup>[34]</sup>
				PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
×2	IMDN <sup>[15]</sup>	694	ACM MM'19	38.00/0.9605	33.63/0.9177	32.19/0.8996	32.17/0.9283
	FMEN <sup>[36]</sup>	748	CVPR'22	38.10/0.9607	33.78/0.9192	32.23/0.9002	32.46/0.9313
	ShuffleMixer <sup>[37]</sup>	394	NeurIPS'22	38.01/0.9606	33.63/0.9180	32.17/0.8995	31.89/0.9257
	LatticeNet <sup>[38]</sup>	756	TPAMI'23	38.06/0.9607	33.70/0.9187	32.20/0.8999	32.25/0.9302
	DRSAN <sup>[39]</sup>	870	TMM'23	38.13/0.9606	33.72/0.9189	32.24/0.9009	32.41/0.9312
	DLSR <sup>[40]</sup>	322	TCSVT'23	38.04/0.9606	33.67/0.9183	32.21/0.9002	32.26/0.9297
	DDistill-SR <sup>[41]</sup>	657	TMM'23	38.08/0.9608	33.73/0.9195	<b>32.25/0.9007</b>	32.39/0.9301
	MICU <sup>[42]</sup>	–	ESWA'24	37.93/0.9601	33.63/0.9170	32.17/0.8987	32.09/0.9271
	HAFRN <sup>[43]</sup>	496	TCE'24	38.05/0.9606	33.66/0.9187	32.21/0.8999	32.20/0.9289
	LBRN <sup>[44]</sup>	325	ESWA'24	38.08/0.9608	33.57/0.9173	32.23/0.9005	32.35/0.9303
Ours	685	–	<b>38.15/0.9609</b>	<b>33.86/0.9200</b>	<b>32.25/0.9010</b>	<b>32.57/0.9316</b>	
×3	IMDN <sup>[15]</sup>	703	ACM MM'19	34.36/0.9270	30.32/0.8417	29.09/0.8046	28.17/0.8519
	FMEN <sup>[36]</sup>	757	CVPR'22	34.45/0.9275	30.40/0.8435	29.17/0.8063	28.33/0.8562
	ShuffleMixer <sup>[37]</sup>	415	NeurIPS'22	34.40/0.9272	30.37/0.8423	29.12/0.8051	28.08/0.8498
	LatticeNet <sup>[38]</sup>	765	TPAMI'23	34.40/0.9272	30.32/0.8416	29.10/0.8049	28.19/0.8513
	DRSAN <sup>[39]</sup>	768	TMM'23	34.50/0.9278	30.39/0.8437	29.13/0.8065	28.35/0.8566
	DLSR <sup>[40]</sup>	329	TCSVT'23	34.49/0.9279	30.39/0.8428	29.13/0.8061	28.26/0.8548
	DDistill-SR <sup>[41]</sup>	665	TMM'23	34.43/0.9276	30.39/0.8432	29.16/0.8070	28.31/0.8546
	MICU <sup>[42]</sup>	–	ESWA'24	34.38/0.9274	30.35/0.8419	29.10/0.8048	28.14/0.8518
	HAFRN <sup>[43]</sup>	505	TCE'24	34.45/0.9276	30.40/0.8433	29.12/0.8058	28.16/0.8528
	LBRN <sup>[44]</sup>	339	ESWA'24	34.43/0.9276	30.39/0.8429	29.13/0.8059	28.29/0.8545
Ours	780	–	<b>34.58/0.9288</b>	<b>30.52/0.8461</b>	<b>29.19/0.8073</b>	<b>28.54/0.8588</b>	
×4	IMDN <sup>[15]</sup>	715	ACM MM'19	32.21/0.8948	28.58/0.7811	27.56/0.7353	26.04/0.7838
	FMEN <sup>[36]</sup>	769	CVPR'22	32.24/0.8955	28.70/0.7839	27.63/0.7379	26.28/0.7908
	ShuffleMixer <sup>[37]</sup>	411	NeurIPS'22	32.21/0.8953	28.66/0.7827	27.61/0.7366	26.08/0.7835
	LatticeNet <sup>[38]</sup>	777	TPAMI'23	32.18/0.8943	28.61/0.7812	27.57/0.7355	26.14/0.7844
	DRSAN <sup>[39]</sup>	747	TMM'23	32.30/0.8954	28.66/0.7838	27.61/0.7381	26.26/0.7920
	DLSR <sup>[40]</sup>	338	TCSVT'23	32.33/0.8963	28.68/0.7832	27.61/0.7374	26.19/0.7892
	DDistill-SR <sup>[41]</sup>	675	TMM'23	32.29/0.8961	28.69/0.7833	27.65/0.7385	26.25/0.7893
	MICU <sup>[42]</sup>	–	ESWA'24	32.21/0.8945	28.65/0.7820	27.57/0.7359	26.15/0.7872
	HAFRN <sup>[43]</sup>	517	TCE'24	32.24/0.8953	28.60/0.7816	27.58/0.7365	26.02/0.7849
	LBRN <sup>[44]</sup>	345	ESWA'24	32.33/0.8964	28.62/0.7826	27.60/0.7377	26.17/0.7882
Ours	745	–	<b>32.47/0.8983</b>	<b>28.77/0.7854</b>	<b>27.67/0.7395</b>	<b>26.45/0.7952</b>	

注: 粗体为最优, – 表示模型该项参数缺省。

表2 本研究算法与先进 Transformer 算法指标对比

方法名	缩放因子	参数量/K	PSNR/SSIM
LBNNet <sup>[45]</sup>	×2	731	38.05/0.9607
	×3	736	34.47/0.9277
	×4	742	32.29/0.8960
NGSwin <sup>[46]</sup>	×2	998	38.05/0.9610
	×3	1007	34.52/0.9282
	×4	1019	32.33/0.8963
TCSR-B <sup>[47]</sup>	×2	628	38.14/0.9611
	×3	589	34.56/0.9285
	×4	682	32.43/0.8977
Ours	×2	685	<b>38.15/0.9609</b>
	×3	780	<b>34.58/0.9288</b>
	×4	745	<b>32.47/0.8983</b>

注: 粗体为最优。

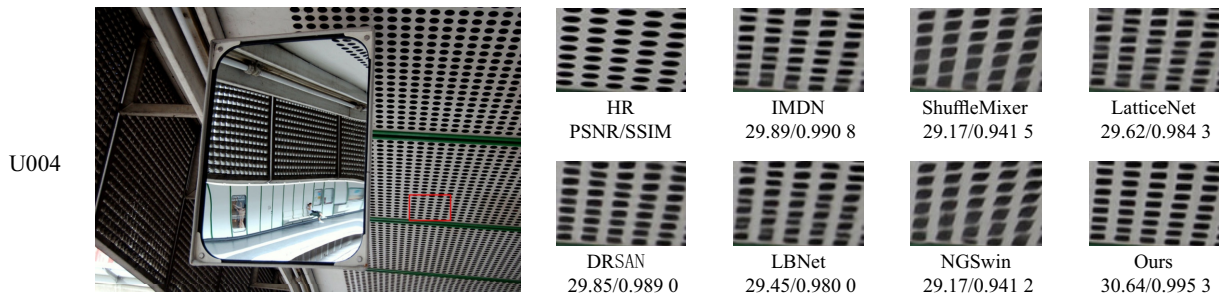
高, 而其他模型存在明显的线条扭曲和模糊问题, 视觉效果较差。

### 3.3 消融实验

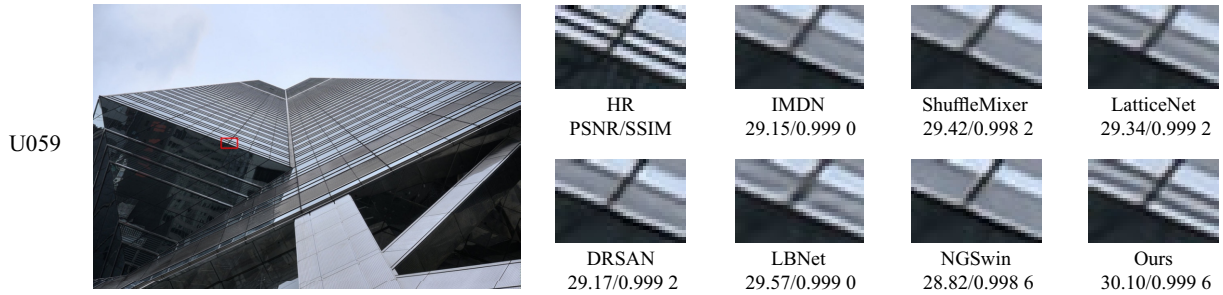
#### 3.3.1 高频信息自调制的必要性分析

本研究提出的高频信息自调制模块, 减少了目前主流算法在重建图像时的冗余信息。为了佐证这一观点, 彰显高频信息自调制模块的必要性, 对基准算法 IMDN<sup>[15]</sup>、Transformer 算法 NGSwin<sup>[46]</sup> 以及本研究算法所重建出的图像进行频谱可视化, 如图 3 所示。

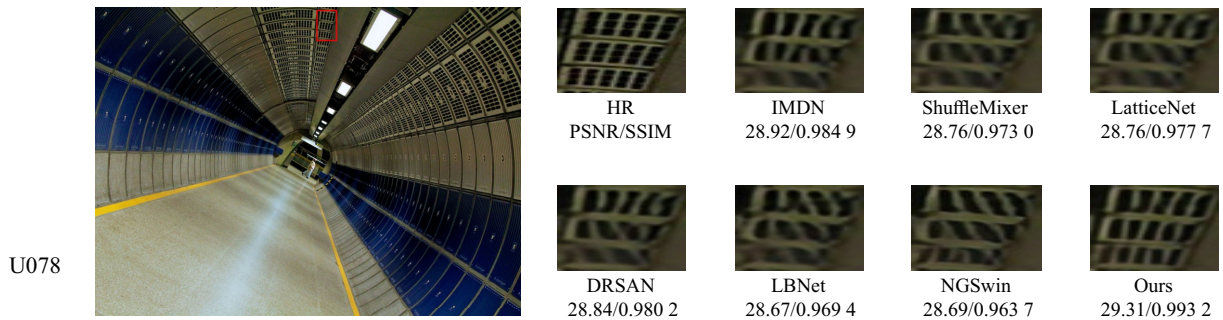
具体步骤是: 以 Urban100 数据集中 img093 图像为样本, 随机采样一个图像小块, 获取图像小块在



(a) 不同算法在Img004(来自Urban100)上的视觉效果对比



(b) 不同算法在Img059(来自Urban100)上的视觉效果对比



(c) 不同算法在Img078(来自Urban100)上的视觉效果对比

图2 标准测试集下4倍放大视觉效果比较

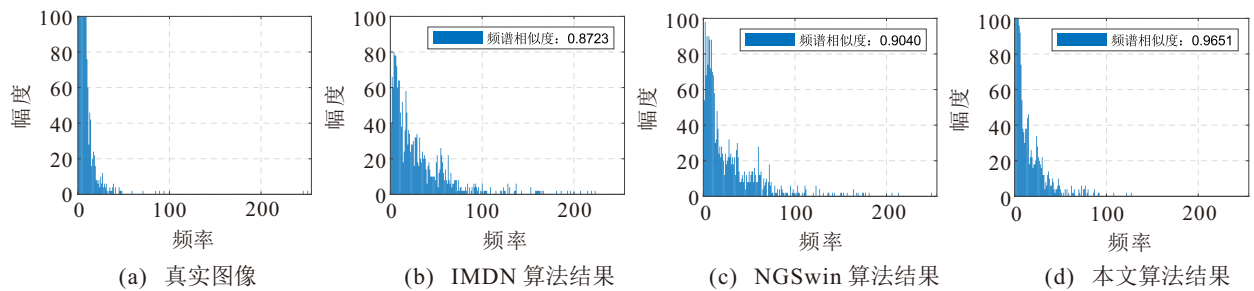


图3 3种算法高频图像的频率直方图

各算法仿真结果中的高频特征图,将高频特征图进行傅里叶变换,得到频域特征图.对频谱移位,将零频率分量移到频谱中心,取绝对值得到幅度谱,绘制出各算法仿真结果的频率直方图.利用余弦相似度计算公式,计算幅度谱间的频谱相似度.

在图3中,当图像频率超过100时,真实图像高频特征包含的信息极其有限.因此,在该频段,正确的模型应当抑制高频信息的生成.然而,IMDN<sup>[15]</sup>和NGSwIn<sup>[46]</sup>算法的结果中在该频段存在大量不恰当的信息,很显然,这些信息是冗余的.本研究提出的

算法引导高频特征以更加准确的方式生成,有效地降低了信息的冗余性,增强了重建的精度.因此,所重建出的图像在纹理方面,给人以最自然的视觉感受,并且在频谱相似度方面高出NGSwIn<sup>[46]</sup>算法约6.76%,表现出最优秀的水平.

### 3.3.2 多尺度采样网络的有效性分析

为了验证所提出多尺度采样网络的有效性,本研究将多尺度采样网络输入的特征作为原始输入,采样因子为2、4、8的分支输出的高频特征图分别表示为2倍、4倍和8倍的高频信息.为了更好地量化

差异,对上述特征图逐像素进行比较.将像素差的绝对值限定在 $0 \sim 5$ 范围内,认为相应像素点相似.结果如图4所示,随着采样因子增加,得到的高频信息与原始高分辨率图像的像素点越相似,对图像重建任务更有帮助.特别是经过自调制的高频输出特征,相似像素点比例为42.52%,明显高于原始输入特征的12.92%.这表明本研究提出的高频自调制模块可显著提升高频信息的准确性,对超分辨率模型的性能有重要贡献.

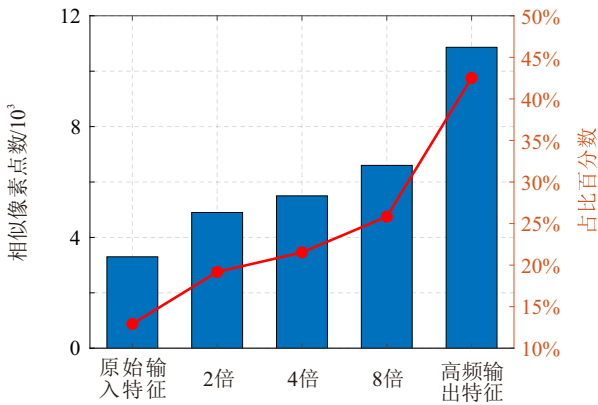


图4 特征图高频信息

### 3.3.3 HGSA 有效性分析

为验证基于高频引导的空间注意力机制的有效性,本节设计3种网络模型进行比较.它们分别是:没有使用注意力的网络None,使用了空间增强注意力ESA的网络,以及使用了高频引导的空间注意力HGSA网络.本研究在缩放因子为4的基准数据集上进行测试.

实验结果如表3所示,空间增强注意力ESA在提升网络性能方面起到了必要的作用.在4个基准数据集Set5、Set14、B100和Urban100上,相较于None,使用了ESA,可以使PSNR分别提高0.04 dB、0.02 dB、0.02 dB和0.05 dB,SSIM值分别提高0.0003、0.0004、

表3 3种网络在4个基准数据集下的指标对比

模型名	Set5 <sup>[21]</sup>	Set14 <sup>[22]</sup>	B100 <sup>[23]</sup>	Urban100 <sup>[24]</sup>
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
None	31.72/0.8885	28.24/0.7726	27.32/0.7274	25.33/0.7590
ESA	31.76/0.8888	28.26/0.7730	27.34/0.7279	25.38/0.7608
HGSA	<b>31.84/0.8898</b>	<b>28.35/0.7746</b>	<b>27.38/0.7290</b>	<b>25.44/0.7633</b>

注:粗体为最优.

0.0005和0.0018.同样的,与ESA相比,HGSA在这4个基准数据集上的量化指标表现更佳,具体表现为相较于ESA,PSNR分别提升了0.08 dB、0.09 dB、0.04 dB和0.06 dB,同时SSIM值也分别增长了0.0010、0.0016、0.0011和0.0025.这一优势源于ESA主要聚焦于空域信息,却忽视了频域信息的重要性.相比之下,本研究提出的HGSA不仅重视空域信息,还深入关注对超分辨率任务至关重要的频域信息,弥补了ESA的不足,实现了性能上的提升.图5展示了经过300轮训练后,3种网络性能曲线的对比情况,可以明显看出,在大约100轮参数更新后,HGSA与ESA之间的性能差距逐渐拉大,且HGSA的性能稳定地保持在ESA之上,进一步验证了HGSA的优越性.

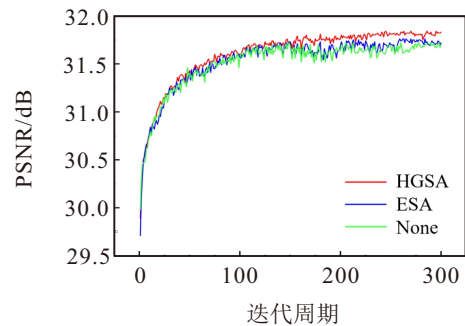


图5 3种网络在Set5数据集上的性能曲线

### 3.3.4 高频信息互调制的有效性分析

为研究MM模块数量对于模型参数量和重建性能的影响,设置3组对比实验,并在缩放因子为4时进行验证.MM模块个数记为 $n$ ,分别为5、6、7,基本通道数都设置为48,每轮输入16个批次,进行300轮训练,测试结果如表4所示.

当 $n$ 为5时,模型参数量最小,同时性能指标较低.当MM模块数量从5增加到6时,在纹理特别复杂的Urban100数据集上,PSNR和SSIM分别增加了0.21 dB和0.0060,重建性能有较大提升.而模型参数量仅增涨了150 K,约为基准模型的16.3%;计算量仅增加了9.54 G,约为基准模型的12.5%.当MM模块数量从6增加到7时,模型参数量增加了150 K,计算量增加了9.54 G,而PSNR却降低了0.40 dB,SSIM降低了0.0117,性能指标明显下降.可以看出,当MM

表4 高频信息互调制模块数量对重建性能的影响

$n$	参数量/K	FLOPs/G	Set5 <sup>[21]</sup>	Set14 <sup>[22]</sup>	B100 <sup>[23]</sup>	Urban100 <sup>[24]</sup>
			PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
5	920	76.30	32.21/0.8948	28.56/0.7798	27.55/0.7344	25.93/0.7794
6	1 070	85.84	<b>32.33/0.8961</b>	<b>28.62/0.7815</b>	<b>27.60/0.7361</b>	<b>26.14/0.7854</b>
7	1 220	95.38	32.05/0.8938	28.44/0.7780	27.47/0.7334	25.74/0.7737

注:粗体为最优.

模块数量增加到6个时,模型拥有了最好的性能,同时拥有相对较小的参数量和计算量.因此,本研究最终选择MM模块数量为6个的模型.

### 3.3.5 高频信息自调制融合方式的有效性分析

本研究提出的高频信息互调制的方法通过调制层对前期高频信息生成掩码,并有效指导后续高频信息的生成.为了验证高频信息互调制方式 Reuse\_mod 的有效性,将其分别与直接连接 Reuse\_none,稠密连接 Reuse\_dense 和跳跃连接 Reuse\_skip 进行对比,这3种网络经过300轮训练,并在Set14数据集上进行了缩放因子为4的测试.

表5 特征复用有效性、特征互调制高效性分析

模型名	参数量/K	FLOPs/G	特征互调制	特征复用	跳跃连接	稠密连接	PSNR/dB	SSIM
Reuse_none	328	29.12	×	×	×	×	28.24	0.7720
Reuse_dense	370	31.00	×	✓	✓	✓	28.29	0.7736
Reuse_skip	370	31.00	×	✓	✓	×	28.28	0.7734
Reuse_mod	377	31.00	✓	✓	×	×	<b>28.35</b>	<b>0.7746</b>

注:粗体为最优.

### 3.3.6 渐进式特征融合的有效性分析

渐进式特征融合具有层次化信息提取能力高、信息丢失量少等特点.为验证渐进式特征融合的有效性,进行以下消融实验:

1) 使用渐进式特征融合的网络 TS\_Net;

2) 使用单阶段特征融合的网络 OS\_Net;

3) 使用元素相加线性拟合网络 L\_Net.

3种网络结构如图6所示,将3种网络进行300轮训练,并将3种网络在缩放因子为4的标准数据集上进行测试,得到的PSNR值如表6所示.

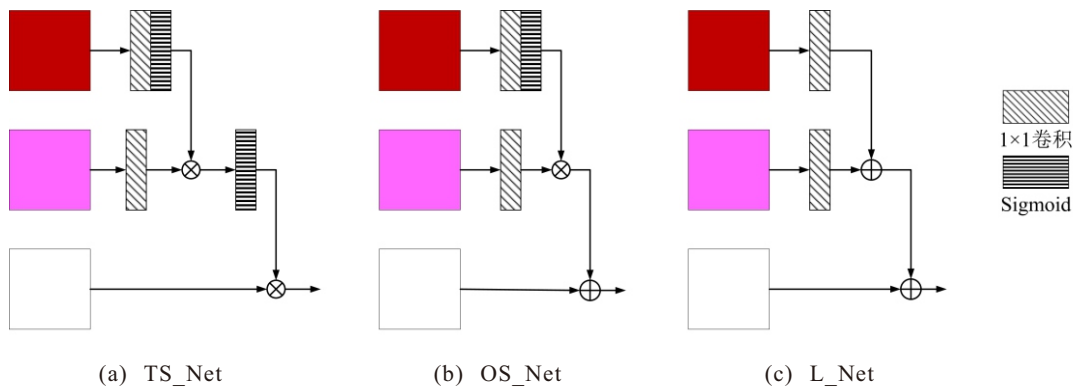


图6 3种特征融合的网络结构

表6 渐进式特征融合有效性分析

方法	参数量/K	Set5 <sup>[21]</sup>	Set14 <sup>[22]</sup>	B100 <sup>[23]</sup>	Urban100 <sup>[24]</sup>
TS_Net	377	<b>31.84</b>	<b>28.35</b>	<b>27.38</b>	<b>25.44</b>
OS_Net	377	31.81(-0.03 dB)	28.34(-0.01 dB)	27.36(-0.02 dB)	25.41(-0.03 dB)
L_Net	377	31.77(-0.07 dB)	28.29(-0.06 dB)	27.29(-0.09 dB)	25.39(-0.05 dB)

从表6可见,3种网络的参数量相同.然而,渐进式特征融合的方法,与单阶段和线性拟合相比,网络的PSNR值均得到了提升.特别值得注意的是,在B100数据集上,相较于线性拟合,渐进式特征融合网络的PSNR值提高了0.09 dB.本研究提出的高频信息互调制模块采用渐进式特征融合的训练方式,

通过逐步整合不同层次的特征信息,使得网络在处理复杂数据时能够更加高效地提取关键特征.

### 3.3.7 调制模型的理论分析

本研究提出的调制连接通过可训练参数矩阵动态调节节点间的信息传递权重.

图7(a)为直接连接,即 $F_{cur} = F_{pre}$ ;图7(b)为稠密

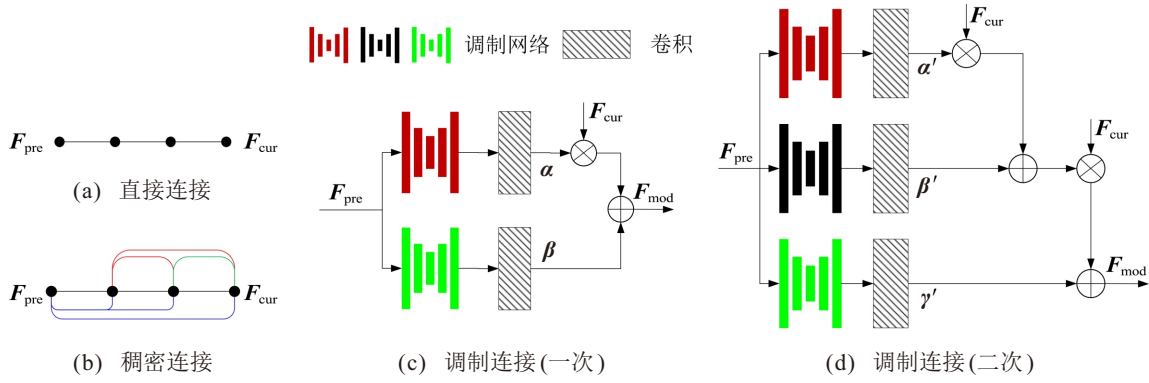


图7 信息节点的连接方式

连接, 即  $F_{cur} = \mu F_{pre}$ , 其中  $\mu$  是参数传递矩阵; 图 7(c) 为本研究提出的调制连接, 即  $F_{mod} = \alpha F_{cur} \oplus \beta$ , 其中  $\alpha$  和  $\beta$  是  $F_{pre}$  生成的掩码矩阵. 特别地, 参数矩阵的可调节性赋予网络自适应筛选有效信息, 校正传递误差的能力, 从而在理论上保证模型能够收敛至更优解. 例如, 取  $\alpha$  为恒等映射矩阵,  $\beta$  为零矩阵, 调制连接退化成直接连接; 取  $\alpha$  为  $\mu$  的广义逆矩阵  $(\mu^T \mu)^{-1} \mu^T$ ,  $\beta$  为零矩阵, 调制连接退化成稠密连接.

此外, 考虑到图像超分辨率重建任务需要对精细的纹理 (高阶特征) 进行建模, 结合模型的计算复杂度, 最终将模型确定为二阶模型,  $F_{mod}$  和  $F_{cur}$  呈二次函数关系, 即  $F_{mod} = \alpha' \otimes F_{cur} \otimes F_{cur} \oplus \beta' \otimes F_{cur} \oplus \gamma'$ , 其中  $\alpha'$ 、 $\beta'$ 、 $\gamma'$  为需要学习的 3 个参数矩阵, 如图 7(d) 所示. 3.3.6 节的消融实验给出了特征融合的最优网络结构, 同时为了防止网络训练时的梯度爆炸, 加入了残差连接和通道拼接, 网络结构由图 7(d) 变成图 1(f) 的网络结构. 其中:  $F_{14} = \alpha'$ ,  $F_{25} = \beta'$ ,  $F_{36} = \gamma'$ . 若取  $\alpha'$  为零矩阵,  $\beta'$  为  $\alpha$  矩阵,  $\gamma'$  为  $\beta$  矩阵, 则模型退化成线性模型. 由此可见, 二次模型是为了拟合神经网络的高阶特征而设计的.

## 4 结论

本研究提出了一种基于高频信息调制的轻量级图像超分辨率重建网络模型. 该模型通过高频信息的自调制模块提高了信息的提取精度, 并通过高频信息的互调制模块建立了自调制信息的非线性映射关系, 从而减少了信息冗余. 本研究设计的模型不仅在网络参数量方面具有一定的优势, 而且在性能方面也表现出色. 在不同缩放因子的基准数据集 Set5、Set14、B100、Urban100 上, 与其他轻量级算法相比, 均实现了 PSNR 和 SSIM 等量化指标的提升.

当然, 特征调制的方式灵活多变, 如何确定一个能够最大化信息利用率的、自适应的网络拓扑结构, 仍是一个值得深入探讨的课题. 未来, 将致力于此方向的研究.

## 参考文献 (References)

- [1] Wang X Y, Hu Q, Cheng Y S, et al. Hyperspectral image super-resolution meets deep learning: A survey and perspective[J]. *CAA Journal of Automatica Sinica*, 2023, 10(8): 1668-1691.
- [2] Yu Y, She K, Liu J H, et al. A super-resolution network for medical imaging via transformation analysis of wavelet multi-resolution[J]. *Neural Networks*, 2023, 166: 162-173.
- [3] Ren S, Guo K, Zhou X, et al. Medical image super-resolution based on semantic perception transfer learning[J]. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2023, 20(4): 2598-2609.
- [4] Mishra D, Hadar O. Accelerating neural style-transfer using contrastive learning for unsupervised satellite image super-resolution[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 1-14.
- [5] Wen X, Li F, Mao Z X, et al. Super-resolution reconstruction of remote sensing images based on an adaptive fourth-order PDE[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 3434339.
- [6] Liu Y H, Zhong Y F, Shi S N, et al. Scale-aware deep reinforcement learning for high resolution remote sensing imagery classification[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2024, 209: 296-311.
- [7] Jiang H, Asad M, Liu J J, et al. Single image detail enhancement via metropolis theorem[J]. *Multimedia Tools and Applications*, 2024, 83(12): 36329-36353.
- [8] Jiang H, Cai H K, Yang J. Learning in-place residual homogeneity for image detail enhancement[C]. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing. Calgary, 2018: 1428-1432.
- [9] Dong C, Loy C C, He K M, et al. Learning a deep convolutional network for image super-resolution[C]. *Computer Vision — ECCV 2014*. Cham: Springer International Publishing, 2014: 184-199.
- [10] Lim B, Son S, Kim H, et al. Enhanced deep residual networks for single image super-resolution[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, 2017: 1132-1140.
- [11] Zhang Y L, Li K P, Li K, et al. Image super-resolution

- using very deep residual channel attention networks[C]. *Computer Vision — ECCV 2018*. Cham: Springer International Publishing, 2018: 294-310.
- [12] 赵小强, 王泽, 宋昭漾, 等. 基于类金字塔图残差网络的图像超分辨率重建[J]. *控制与决策*, 2024, 39(3): 786-794.  
(Zhao X Q, Wang Z, Song Z Y, et al. Pyramid-like graph residual network for image super-resolution reconstruction[J]. *Control and Decision*, 2024, 39(3): 786-794.)
- [13] 程德强, 陈杰, 寇旗旗, 等. 融合层次特征和注意力机制的轻量化矿井图像超分辨率重建方法[J]. *仪器仪表学报*, 2022, 43(8): 73-84.  
(Cheng D Q, Chen J, Kou Q Q, et al. Lightweight super-resolution reconstruction method based on hierarchical features fusion and attention mechanism for mine image[J]. *Chinese Journal of Scientific Instrument*, 2022, 43(8): 73-84.)
- [14] Kim J, Lee J K, Lee K M. Deeply-recursive convolutional network for image super-resolution[C]. *2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, 2016: 1637-1645.
- [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*. Nice France, 2019: 2024-2032.
- [16] Liu J, Tang J, Wu G S. Residual feature distillation network for lightweight image super-resolution[C]. *Computer Vision — ECCV 2020 Workshops*. Cham: Springer International Publishing, 2020: 41-55.
- [17] Zhang W, Fan Z Q, Song Y, et al. Lightweight image super-resolution with group-convolutional feature enhanced distillation network[J]. *International Journal of Machine Learning and Cybernetics*, 2023, 14(7): 2467-2482.
- [18] Yu L, Li X P, Li Y W, et al. DIPNet: Efficiency distillation and iterative pruning for image super-resolution[C]. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Vancouver, 2023: 1692-1701.
- [19] Sun L, Dong J X, Tang J H, et al. Spatially-adaptive feature modulation for efficient image super-resolution[C]. *2023 IEEE/CVF International Conference on Computer Vision*. Paris, 2023: 13144-13153.
- [20] Zhou L, Cai H M, Gu J J, et al. Efficient image super-resolution using vast-receptive-field attention[C]. *Computer Vision — ECCV 2022 Workshops*. Cham: Springer Nature Switzerland, 2023: 256-272.
- [21] 程德强, 王子强, 张皓翔, 等. 基于浅层特征调制的轻量级单幅图像超分辨率重建[J]. *控制与决策*, 2024, 39(12): 4045-4054.  
(Cheng D Q, Wang Z Q, Zhang H X, et al. Lightweight single image super-resolution based on shallow feature modulation[J]. *Control and Decision*, 2024, 39(12): 4045-4054.)
- [22] Liu J, Zhang W J, Tang Y T, et al. Residual feature aggregation network for image super-resolution[C]. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, 2020: 2356-2365.
- [23] Zhao H Y, Kong X T, He J W, et al. Efficient image super-resolution using pixel attention[C]. *Computer Vision — ECCV 2020 Workshops*. Cham: Springer International Publishing, 2020: 56-72.
- [24] 程德强, 程志威, 马祥, 等. 基于渐进式感受野的轻量级图像超分辨率重建方法[J]. *控制与决策*, 2025, 40(4): 1237-1246.  
(Cheng D Q, Cheng Z W, Ma X, et al. Lightweight image super-resolution reconstruction method based on progressive receptive field[J]. *Control and Decision*, 2025, 40(4): 1237-1246.)
- [25] 韩玉兰, 顾琴, 崔玉杰, 等. 基于树状多分支残差注意力网络的真实场景图像超分辨率重构[J]. *控制与决策*, 2024, 39(10): 3442-3450.  
(Han Y L, Gu Q, Cui Y J, et al. Real-world super-resolution based on residual attention network with tree based multi-branch structure[J]. *Control and Decision*, 2024, 39(10): 3442-3450.)
- [26] Liang J Y, Cao J Z, Sun G L, et al. SwinIR: Image restoration using swin transformer[C]. *2021 IEEE/CVF International Conference on Computer Vision Workshops*. Montreal, 2021: 1833-1844.
- [27] Hendrycks D, Gimpel K. Gaussian error linear units (GELUs)[J/OL]. 2016, arXiv: 1606.08415.
- [28] Dai T, Cai J R, Zhang Y B, et al. Second-order attention network for single image super-resolution[C]. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, 2019: 11057-11066.
- [29] Shi W Z, Caballero J, Huszar F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]. *2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, 2016: 1874-1883.
- [30] Timofte R, Gu S H, Wu J Q, et al. Ntire 2018 challenge on single image super-resolution: Methods and results[C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2017: 1110-1121.
- [31] 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*. Surrey, 2012: 135.1-135.10.
- [32] Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations[C]. *Curves and Surfaces*. Berlin: Springer, 2012: 711-730.
- [33] Martin D, Fowlkes C, Tal D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[C]. *Proceedings of the 8th IEEE International Conference on Computer Vision*. Vancouver, 2001: 416-423.
- [34] Huang J-B, Singh A, Ahuja N. Single image super-

- resolution from transformed self-exemplars[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, 2015: 5197-5206.
- [35] Kingma D P, Ba J. Adam: A method for stochastic optimization[J/OL]. 2015, arXiv: 1412.6980.
- [36] Du Z C, Liu D, Liu J, et al. Fast and memory-efficient network towards efficient image super-resolution[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. New Orleans, 2022: 852-861.
- [37] Sun L, Pan J S, Tang J H. Shufflemixer: An efficient convnet for image super-resolution[J]. Advances in Neural Information Processing Systems, 2022, 35: 17314-17326.
- [38] Luo X, Qu Y, Xie Y, et al. Lattice network for lightweight image restoration[J]. Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(4): 4826-4842.
- [39] 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, 2023, 25: 907-918.
- [40] Huang H, Shen L, He C Y, et al. Differentiable neural architecture search for extremely lightweight image super-resolution[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(6): 2672-2682.
- [41] Wang Y, Su T T, Li Y S, et al. DDistill-SR: Reparameterized dynamic distillation network for lightweight image super-resolution[J]. IEEE Transactions on Multimedia, 2023, 25: 7222-7234.
- [42] Chen Y T, Xia R L, Yang K, et al. MICU: Image super-resolution via multi-level information compensation and U-net[J]. Expert Systems with Applications, 2024, 245: 123111.
- [43] Wang K X, Yang X M, Jeon G. Hybrid attention feature refinement network for lightweight image super-resolution in metaverse immersive display[J]. IEEE Transactions on Consumer Electronics, 2024, 70(1): 3232-3244.
- [44] Hao F W, Wu J S, Liang W Y, et al. Lightweight blueprint residual network for single image super-resolution[J]. Expert Systems with Applications, 2024, 250: 123954.
- [45] 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. Vienna, 2022: 913-919.
- [46] Choi H, Lee J, Yang J. N-gram in swin transformers for efficient lightweight image super-resolution[C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, 2023: 2071-2081.
- [47] Wu G, Jiang J J, Bai Y C, et al. Incorporating transformer designs into convolutions for lightweight image super-resolution[J/OL]. 2023, arXiv: 2303.14324.

#### 作者简介

江鹤 (1990-), 男, 讲师, 硕士生导师, 主要研究方向为图像修复、三维重建, E-mail: [jianghe@cumt.edu.cn](mailto:jianghe@cumt.edu.cn);

顾豪 (2002-), 男, 硕士生, 主要研究方向为图像超分辨率重建, E-mail: [guhao@cumt.edu.cn](mailto:guhao@cumt.edu.cn);

郑州 (2002-), 男, 硕士生, 主要研究方向为图像细节增强, E-mail: [zhengzhou@cumt.edu.cn](mailto:zhengzhou@cumt.edu.cn);

乙夫迪 (2001-), 男, 硕士生, 主要研究方向为图像超分辨率重建, E-mail: [yifudi@cumt.edu.cn](mailto:yifudi@cumt.edu.cn);

王子强 (2000-), 男, 硕士生, 主要研究方向为图像超分辨率重建, E-mail: [wangziqiang@cumt.edu.cn](mailto:wangziqiang@cumt.edu.cn);

程德强 (1979-), 男, 教授, 博士生导师, 主要研究方向为机器视觉、模式识别, E-mail: [chengdq@cumt.edu.cn](mailto:chengdq@cumt.edu.cn).