

中图法分类号: TP751; TP18 文献标识码: A 文章编号: 1006-8961(2026)04-1044-17

论文引用格式: Liu Z Y, Yang Y, Huang S Y and Wang S Z. 2026. Recursive super-resolution network with adaptive feature fusion. Journal of Image and Graphics, 31(4):1044-1060(刘紫阳, 杨勇, 黄淑英, 王书昭. 2026. 自适应特征融合的超分辨率重建循环网络. 中国图象图形学报, 31(4):1044-1060)[DOI:10.11834/jig.250332]

自适应特征融合的超分辨率重建循环网络

刘紫阳¹, 杨勇², 黄淑英^{3,4*}, 王书昭⁵

1. 天津工业大学控制科学与工程学院, 天津 300387; 2. 天津工业大学计算机科学与技术学院, 天津 300387;
3. 天津工业大学软件学院, 天津 300387; 4. 沧州市天津工业大学研究院, 沧州 061000;
5. 天津工业大学电子与信息工程学院, 天津 300387

摘要: 目的 深度卷积神经网络在单幅图像超分辨率(single image super-resolution, SISR)领域取得显著进展。但过度依赖加深网络结构易造成特征冗余与利用不足, 限制细节重建。为此, 提出一种基于自适应特征融合的循环网络(recursive network based on adaptive feature fusion, AFFRN), 有效提取并融合不同深度的图像特征, 实现更丰富的纹理重建。**方法** 网络中, 通过构建一个自适应特征融合模块(adaptive feature fusion module, AFFM)实现不同深度特征的有效融合。在AFFM中, 设计了3个不同结构的分支: 1) 细节注意力分支(detail attention branch, DAB), 用于捕捉显著细节特征; 2) 细节探索分支(detail exploration branch, DEB), 负责挖掘深层特征信息; 3) 权重分配分支(weight assignment branch, WAB), 自适应地为DAB和DEB提取的特征分配动态权重, 从而实现更加精细的特征融合。同时, 引入循环机制, 将AFFM在每次迭代中提取的特征反馈用于下一次的迭代中, 逐步提升特征表达能力。**结果** 在5个公开测试数据集上的评估结果表明, 所提出的方法在不同放大倍数下取得了不同程度的提升。当放大倍数分别为2、3和4时, 在Urban100数据集上相比先进的自适应特征选择调制网络(adaptive feature selection modulation network, AFSMNet), 峰值信噪比(peak signal-to-noise ratio, PSNR)分别提升0.75 dB、0.53 dB和0.54 dB, 重建图像在视觉上具有明显优势。**结论** 所提出的AFFRN网络通过结合循环机制与基于动态权重分配的多分支特征融合, 实现了更丰富的细节重建, 获得更高质量的超分辨率重建结果。

关键词: 图像超分辨率重建; 循环网络; 自适应特征融合; 细节注意力; 权重分配

Recursive super-resolution network with adaptive feature fusion

Liu Ziyang¹, Yang Yong², Huang Shuying^{3,4*}, Wang Shuzhao⁵

1. School of Control Science and Engineering, Tiangong University, Tianjin 300387, China; 2. School of Computer Science and Technology, Tiangong University, Tianjin 300387, China; 3. School of Software, Tiangong University, Tianjin 300387, China;
4. Cangzhou Institute of Tiangong University, Cangzhou 061000, China; 5. School of Electronics and Information Engineering, Tiangong University, Tianjin 300387, China

Abstract: Objective Single-image super-resolution (SISR) focuses on generating a high-resolution (HR) image from a

收稿日期: 2025-07-25; 修回日期: 2025-10-31; 预印本日期: 2025-11-07

* 通信作者: 黄淑英 huangshuying@tiangong.edu.cn

基金项目: 国家自然科学基金项目(62072218); 天津市自然科学基金项目(24JCZDJC00130, 25JCZDJC00540); 河北省自然科学基金项目(F2025110006, F2025110010); 沧州市天津工业大学研究院项目(TGCYY-Z-0303)

Supported by: National Natural Science Foundation of China (62072218); Tianjin Municipal Natural Science Foundation (24JCZDJC00130, 25JCZDJC00540); Natural Science Foundation of Hebei Province, China (F2025110006, F2025110010); Project of Cangzhou Institute of Tiangong University (TGCYY-Z-0303)

single low-resolution (LR) input by recovering the texture and structural details that are often lost during image degradation. In recent years, SISR has been widely applied in various domains, including computer vision, image processing, and public surveillance. Current SISR methods include traditional approaches—such as interpolation, reconstruction, and example-based techniques—that rely on prior knowledge but often struggle to recover fine details and textures. By contrast, deep learning-based methods, particularly convolutional neural networks, have made significant progress. However, many of these methods improve performance primarily by increasing network depth, which introduces feature redundancy and inefficient utilization of multilayer features, thereby resulting in artifacts such as texture distortion and edge inconsistency. This work addresses these challenges by proposing an adaptive feature fusion recursive network (AFFRN), which incorporates an adaptive feature fusion module (AFFM) consisting of three specialized branches. AFFRN facilitates efficient information propagation and progressive feature enhancement by employing a recursive refinement mechanism; thus, it can significantly improve texture reconstruction quality. **Method** The proposed AFFRN consists of three main stages. First, the input LR image is processed by several convolutional layers to extract initial low-level representations, which primarily capture basic edge and texture information. These feature maps are passed through AFFM, which performs dynamic and depth-aware fusion of multilevel features. Within AFFM, three specialized branches work in parallel to extract and integrate complementary feature representations: 1) The detail attention branch (DAB), which focuses on enhancing shallow-layer features using attention mechanisms to preserve fine textures and edges; 2) The detail exploration branch (DEB), which captures deep contextual information through densely connected blocks combined with downsampling and upsampling to enable modeling of long-range dependencies and global structures; 3) The weight assignment branch (WAB), which, as the core innovation, adaptively assigns weights to features from DAB and DEB based on their relative importance to facilitate selective fusion that suppresses redundancy and emphasizes discriminative information. Moreover, AFFM is integrated into a recursive framework, where the fused features from each iteration are fed back into the module for subsequent refinement. This recursive refinement, which is critical for accurate super-resolution reconstruction, facilitates coarse-to-fine structural recovery and gradual enhancement of hierarchical features. Finally, the aggregated and enhanced feature representations are passed through convolutional and transposed convolutional layers to restore spatial resolution and reconstruct the output HR image. This design allows AFFRN to balance depth, efficiency, and representational power effectively, thereby resulting in superior performance on texture-rich and structurally complex image regions. **Result** AFFRN is trained on 900 images selected from the DIV2K dataset, with data augmentation being performed through random horizontal and vertical flips to enhance data diversity. Training is performed using the Adam optimizer, with L1 loss serving as the loss function. Performance is evaluated on five public benchmark datasets: Set5, Set14, BSD100, Urban100, and Manga109. HR images are downsampled using bicubic interpolation with scale factors of $\times 2$, $\times 3$, and $\times 4$ to simulate LR inputs. The Y channel of the YCbCr color space is used for evaluation, with peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) serving as the assessment metrics. AFFRN is compared with 19 representative SISR methods. Experimental results demonstrate that AFFRN achieves the highest average PSNR and SSIM values across all scaling factors. Moreover, it recovers highly accurate textures and fine structural details. On the Urban100 dataset, AFFRN achieves PSNR gains of 0.75 dB, 0.53 dB, and 0.54 dB over AFSMNet at scaling factors of $\times 2$, $\times 3$, and $\times 4$, respectively. On the Manga109 dataset, the proposed method also outperforms AFSMNet by 0.11 dB, 0.17 dB, and 0.38 dB at the same scales. Ablation studies further validate the effectiveness and necessity of each module within the proposed network. Removing the core WAB from AFFM results in a PSNR decrease of 0.24 dB, thereby emphasizing the critical role of the proposed WAB in performance improvement. **Conclusion** This study proposes AFFRN, a network specifically designed for the SISR task. The architecture integrates a recursive learning strategy with an AFFM, which plays a critical role in progressively enhancing feature representations across multiple depths. Within the AFFM, the DAB is responsible for refining shallow features to preserve high-frequency textures, whereas the DEB captures deep semantic information by modeling global contextual dependencies. WAB adaptively allocates importance weights to merge these complementary features effectively, thereby promoting informative feature integration while suppressing redundancy. The recursive framework further strengthens feature propagation by enabling iterative refinement, thereby allowing the model to improve detail recovery progressively over multiple stages. Comprehensive evaluations on five benchmark datasets demon-

strate that, compared with existing state-of-the-art methods, AFFRN achieves competitive performance in quantitative metrics and visual quality. Ablation studies further validate the necessity and effectiveness of each component, especially highlighting the critical role of the WAB in enhancing reconstruction accuracy. These results confirm that the proposed adaptive multibranch fusion combined with recursive refinement can effectively alleviate feature redundancy and inefficiency issues, thereby leading to rich texture recovery and highly accurate super-resolution results.

Key words: image super-resolution reconstruction; recursive network; adaptive feature fusion; detail attention; weight assignment

0 引言

单幅图像超分辨率(single image super-resolution, SISR)任务旨在通过一幅低分辨率(low resolution, LR)图像重建出细节丰富的高分辨率(high resolution, HR)图像。该技术在计算机视觉领域得到了广泛应用,如公共安全监控(Wang等,2024;Li等,2025;邱德粉等,2023)、遥感观测(张涛等,2024;Yan等,2025)、医学成像(Lyu等,2024;Huang等,2024b)和目标识别(Hsu和Wu,2025)等。SISR重建方法通常分为两类:传统方法和基于深度学习的方法。传统方法主要包括基于插值的方法(Keys,1981)、基于重建的方法(Park等,2003)和基于样例学习的方法(Yang等,2012),该类方法大都存在耗时和纹理重建缺失的问题,很难满足实际应用中实时处理的需求及对重建HR图像的质量需求。

深度卷积神经网络(convolutional neural network, CNN)的不断进步及其优异的特征提取能力,使得基于深度学习的方法吸引了众多研究者的关注。基于深度学习的SISR重建是利用大量的数据学习LR到HR图像间的映射关系。Dong等人(2014)通过堆叠3个卷积层构建了一个SRCNN(super-resolution convolutional neural network)模型来学习LR到HR图像的映射,其性能指标远超过传统方法结果。随后,研究人员尝试通过增加网络的深度以提升模型性能。然而,随着网络层数的增加,训练时可能会出现梯度消失或爆炸的问题,导致模型难以收敛,训练不稳定,参数优化变得困难,进而限制了性能的提升。为了缓解深度神经网络中普遍存在的梯度消失现象,He等人(2016)提出了残差网络(residual network, ResNet)通过引入跳跃连接使得信息和梯度能够在网络中更有效地传播,从而实现了更深层次网络的稳定训练。受此启发, Kim等人

(2016a)首次将残差学习的概念引入超分辨率(super-resolution, SR)重建任务中,提出了一个基于非常深的卷积神经网络的SR重建模型VDSR(very deep super-resolution network)。为了控制模型参数的数量,研究人员将递归机制引入网络结构设计中。例如, Kim等人(2016b)提出了深度递归卷积网络(deeply-recursive convolutional network, DRCN),该网络通过重复使用相同的卷积层实现多次递归,从而在不增加额外参数的情况下加深了网络的深度。网络深度的增加使得网络难以实现特征的长期依赖, Tai等人(2017)提出一种非常深的持久内存网络MemNet(very deep persistent memory network),将网络深度增加至80层。该网络通过引入一个包含递归单元和门控单元的内存块来挖掘持久性内存,从而有效地捕捉长期依赖关系。除了增加网络层数外,为了更好地利用提取的不同深度特征, Huang等人(2017)提出密集连接网络DenseNet(dense convolutional network),在不同层之间采用跳跃连接,将不同深度的特征信息进行有效传递和融合,克服了梯度消失的问题,获得了较好的性能。Lim等人(2017)受DenseNet启发,设计了一个增强的深度残差SR网络EDSR(enhanced deep super-resolution network),通过去除残差模块中的批归一化层,同时增加网络的深度和宽度,使重建出的HR图像无论是视觉效果还是客观指标都达到同时期的最优。

为了提高网络的特征学习效率,使其专注于某些重要特征的学习,注意力机制被引入到图像SR重建领域。Zhang等人(2018)提出残差注意力网络RCAN(very deep residual channel attention network),利用特征图通道之间的相互依赖性,为每个通道的特征生成不同的注意力权重。在此基础上, Dai等人(2019)提出二阶注意力图像SR网络SAN(second-order attention network),该网络在一阶注意力的基础上引入了二阶注意力机制,相比于RCAN,性能再次

得到提升。Transformer(Vaswani等,2017)作为一种基于自注意力机制的网络结构,凭借其在建模长距离依赖关系方面的优越性能,逐渐引入到计算机视觉领域。Yang等人(2020)提出了一种基于纹理Transformer的SISR方法TTSR(texture Transformer network for image super-resolution),该模型依靠Transformer中的注意力机制,有效地实现了高质量纹理信息的搜索与迁移,最终获得性能优越的纹理表征。

传统的Transformer结构,通常存在局部特征建模能力不足、计算复杂度较高等问题,尤其在处理高分辨率图像时计算和内存开销较大,对硬件资源要求较高。然而,已有部分研究针对这些问题进行了改进。例如,ESRT(efficient super-resolution transformer)(Lu等,2022)通过设计高效的轻量化Transformer结构,在保持特征表达能力的同时显著降低了计算成本;SPAN(swift parameter-free attention network)(Wan等,2024)则引入无参数注意机制,有效减少了计算开销。这些方法在一定程度上缓解了传统Transformer的局限性。

目前,尽管大部分基于深度学习的方法通过增加网络深度或复杂度提升重建性能,但这种方式往往带来严重的特征冗余和信息传递瓶颈,导致多层次特征融合效果不佳,进而引发重建图像的细节缺失、伪影和纹理扭曲等问题。针对这一挑战,本文提出了一种自适应特征融合循环网络(adaptive feature fusion recursive network, AFFRN),通过设计多分支差异化建模与通道级动态权重分配机制,实现对不同深度特征的高效融合与精细提取。具体而言, AFFRN设计一个自适应特征融合模块(adaptive feature fusion module, AFFM),该模块由3个功能互补且并行的分支组成:细节注意力分支(detail attention branch, DAB)专注于浅层特征的精细提取与增强,确保图像的细节和边缘信息得到有效保留;细节探索分支(detail exploration branch, DEB)通过结合上下采样操作与密集连接结构,深入挖掘深层语义信息及全局结构特征;权重分配分支(weight assignment branch, WAB)基于通道级动态权重分配机制,自适应调节并融合来自DAB与DEB的特征响应,有效抑制冗余信息并强化两类特征的互补性,从而提升整体特征表达能力。同时,引入循环机制,将AFFM嵌入循环结构中,使得网络能够在多次迭代中逐层整合浅层与深层特征,实现信息的高效传递

与递进式细节恢复。

本文的主要贡献如下:1)提出一种AFFRN用于SISR重建任务,该网络通过引入循环机制联合多分支自适应融合策略,显著提升了多层特征的协同表达与传递效率,从而增强了SISR的重建质量。大量实验结果表明, AFFRN在性能方面优于当前先进的SR重建模型。2)在每个迭代层引入AFFM,设计功能互补的DAB和DEB,分别针对浅层细节和深层语义特征进行差异化建模,提升多尺度特征的捕获与表达能力。DAB通过加权浅层特征,强化图像细节信息;DEB结合上下采样与密集连接,有效挖掘深层特征中的丰富信息。3)在AFFM中设计WAB分支,通过通道级动态权重自适应分配,实现浅层与深层特征的最优融合,有效减少冗余信息,增强细节重建能力,进而提升整体网络性能。

1 提出的方法

1.1 整体结构

针对现有SISR重建方法存在的图像细节重建不足及边缘模糊的问题,本文提出一种AFFRN来实现SISR重建,整体架构如图1(a)所示。该网络主要包含初始特征提取层、AFFM和图像重建层。首先, LR图像通过初始特征提取层,即两个 3×3 卷积层和 1×1 卷积层来实现初始特征的提取,得到初始特征图,该操作表示为

$$\mathbf{F}_0 = C_1(C_3(C_3(I_{LR}))) \quad (1)$$

式中, C_1 和 C_3 分别表示卷积核大小为 1×1 和 3×3 的卷积层。

然后,所提取的初始特征图 \mathbf{F}_0 被送入AFFM。为了增加网络的特征提取能力,网络对AFFM采用一种循环结构,将每次执行AFFM学习到的特征图向前回传作为下一个AFFM输入的一部分,对浅层特征进行特征补充,并精细化所学习到的特征,渐进式地实现特征由粗到细的特征重建和校正,从而获得增强的细节特征。该模块的首次输入只接收初始特征图 \mathbf{F}_0 (即 \mathbf{F}_0^n),之后每次循环, AFFM的输入都包含初始特征图 \mathbf{F}_0 和上一次AFFM执行输出的特征图 $\mathbf{F}_{\text{affm}}^{n-1}$,这里 n 表示当前循环的次数, N 表示设定的执行循环次数。在AFFM中,输入特征图通过一个 1×1 的卷积层在通道方向上进行特征的整合和降维,

其操作可表示为

$$\mathbf{F}_{\text{affm}}^n = \mathbf{M}_{\text{affm}}([\mathbf{F}_0, \mathbf{F}_{\text{affm}}^{n-1}]) = \mathbf{M}_{\text{affm}}(\mathbf{F}_i^n) \quad (2)$$

式中, $\mathbf{F}_{\text{affm}}^n$ 是 AFFM 第 n 次循环输出的融合特征图, \mathbf{M}_{affm} 表示 AFFM 的操作, $[\]$ 表示合并操作, \mathbf{F}_i^n 表示 AFFM 的输入, 即 \mathbf{F}_0 和 $\mathbf{F}_{\text{affm}}^{n-1}$ 合并后的特征图。

最后, 重建特征图 $\mathbf{F}_{\text{affm}}^n$ 被送入到图像重建层, 即由一个 3×3 卷积层和一个反卷积层构成的重建模块中, 来获得重建后的细节图像。最后, 通过一个跳

跃连接构建网络的残差结构, 将经过上采样的 \mathbf{I}_{LR} 与重建的细节图像相加, 得到重建的 HR 图像 \mathbf{Y}_{SR}^n , 该过程表示为

$$\mathbf{Y}_{\text{SR}}^n = R_d(\mathbf{F}_{\text{affm}}^n) + f_{\text{UP}}(\mathbf{I}_{\text{LR}}) = D(C_3(\mathbf{F}_{\text{affm}}^n)) + f_{\text{UP}}(\mathbf{I}_{\text{LR}}) \quad (3)$$

式中, \mathbf{Y}_{SR}^n 表示 AFFM 循环第 n 次时网络的重建图像, $R_d(\cdot)$ 表示图像重建层的操作, $f_{\text{UP}}(\cdot)$ 表示上采样操作, $D(\cdot)$ 表示反卷积层。

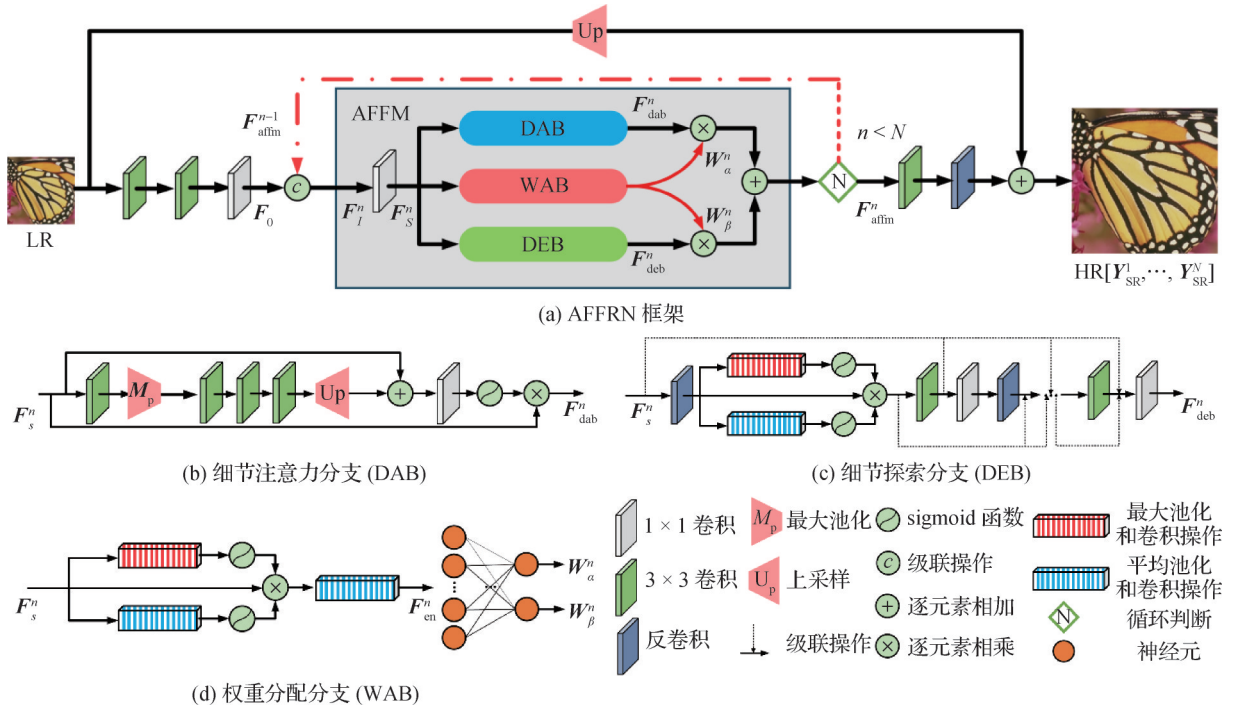


图1 AFFRN网络整体框架

Fig. 1 Overall framework of the AFFRN

((a) network architecture; (b) detail attention branch; (c) detail exploration branch; (d) weight assignment branch)

算法1 AFFRN 的伪代码过程。

输入: 低分辨率图像 \mathbf{I}_{LR} 。

输出: 重建的高分辨率图像 \mathbf{Y}_{SR}^n 。

1) 初始特征提取。

$\mathbf{F}_0 = C_1(C_3(C_3(\mathbf{I}_{\text{LR}})))$ // 利用卷积层进行特征提取;

$\mathbf{F}_{\text{affm}}^{-1} = \mathbf{F}_0$ // 初始化循环输入。

2) 利用自适应特征融合模块(AFFM)进行循环特征融合与增强。

for $n = 0$ to $N-1$ do;

(1) $\mathbf{F}_i^n = [\mathbf{F}_0, \mathbf{F}_{\text{affm}}^{n-1}]$ // 合并初始特征与上一次循环的输出特征;

(2) $\mathbf{F}_s^n = \text{Conv}_{1 \times 1}(\mathbf{F}_i^n)$ // 特征整合与降维;

(3) 细节注意力分支(DAB);

$\mathbf{F}_{\text{dab}}^n = \text{DAB}(\mathbf{F}_s^n)$ // 提取细节特征;

(4) 细节探索分支(DEB);

$\mathbf{F}_{\text{deb}}^n = \text{DEB}(\mathbf{F}_s^n)$ // 提取深层特征;

(5) 权重分配分支(WAB);

$[\mathbf{W}_\alpha^n, \mathbf{W}_\beta^n] = \text{WAB}(\mathbf{F}_s^n)$ // 计算权重;

(6) 特征融合;

$\mathbf{F}_{\text{affm}}^n = \mathbf{W}_\alpha^n * \mathbf{F}_{\text{dab}}^n + \mathbf{W}_\beta^n * \mathbf{F}_{\text{deb}}^n$ // 加权融合特征;

end for。

3) 图像重建。

$\mathbf{F}_{\text{detail}} = \text{Reconstruct}(\mathbf{F}_{\text{affm}}^n)$ // 重建细节图像;

$$I_{\text{bicubic}} = f_{\text{UP}}(I_{\text{LR}}) // \text{上采样};$$

$Y_{\text{SR}}^n = I_{\text{bicubic}} + F_{\text{detail}}$ // 加权得到最终高分辨率图像。

4) 输出重建图像 Y_{SR}^n 。

为了说明 AFFM 执行循环操作能够学到更丰富的纹理细节, 以 Urban100 (Huang 等, 2015) 数据集中 Img_092 图像进行放大 4 倍的实验结果为例, 如图 2 所示。图 2 展示了 AFFM 的循环次数 n 被依次设置为 1~4 时获得的可视化结果。从图中可以看到, 随着迭代次数的增加, 重建后的 HR 图像包含的纹理信息更丰富、目标边缘更锐化, 且逐步接近真实 HR 图像。

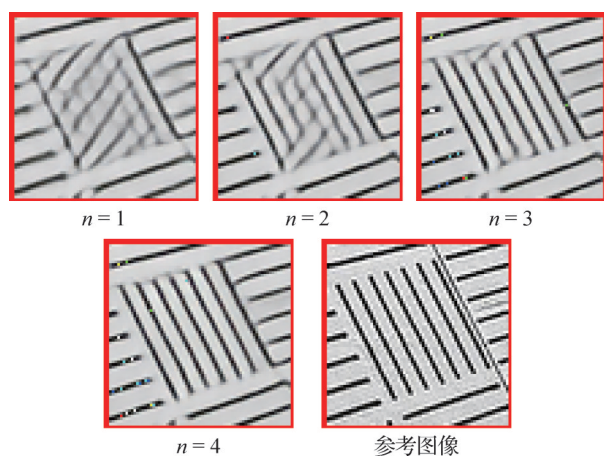


图2 AFFRN 对分辨率×4 重建的中间结果可视化

Fig. 2 AFFRN visualization of intermediate results for ×4 SR

1.2 自适应特征融合模块

在 SISR 任务中, 充分利用不同深度的特征对于提升重建质量至关重要。浅层特征因感受野较小, 更擅长捕捉局部纹理和边缘信息; 而深层特征感受野更大, 更适合建模图像的全局结构与语义内容。现有多数方法通过加深网络并传递浅层特征以实现信息复用, 但缺乏有效的融合机制, 容易导致特征冗余, 限制了浅深层信息的协同作用, 从而影响最终的重建效果。此外, 网络深度的增加也带来了训练难度的提升。因此, 本文提出一种自适应特征融合模块 AFFM, 旨在实现浅层与深层特征的差异化建模和高效融合。该模块主要包含 3 个分支, 即 DAB、DEB 和 WAB。

在 AFFM 中, DAB 聚焦于浅层特征的强化, 借助注意力机制突出细节纹理信息, DEB 则对深层特征进行建模, 通过下采样与上采样操作拓展感受野, 从

而提取更丰富的结构性特征。为了更有效地融合两条分支提取的特征并减少冗余, WAB 在通道维度上对其输出进行自适应加权, 从而兼顾细节与结构信息, 提升最终的图像重建效果。此外, 通过在 AFFM 中引入循环机制, 可逐步增强两条分支的特征提取能力, 并推动其在多次迭代中趋于平衡, 从而实现更稳定、精细的特征融合。该模块具体操作如下

$$F_{\text{dab}}^n = B_{\text{dab}}(C_1(F_i^n)) \quad (4)$$

$$F_{\text{deb}}^n = B_{\text{deb}}(C_1(F_i^n)) \quad (5)$$

$$\{W_\alpha^n, W_\beta^n\} = B_{\text{wab}}(C_1(F_i^n)) \quad (6)$$

$$F_{\text{affm}}^n = M_{\text{affm}}(F_i^n) = W_\alpha^n * F_{\text{dab}}^n + W_\beta^n * F_{\text{deb}}^n \quad (7)$$

式中, B_{dab} 、 B_{deb} 、 B_{wab} 分别表示 DAB、DEB、WAB 这 3 个分支操作。 F_{dab}^n 和 F_{deb}^n 分别表示 DAB 和 DEB 两个分支学习到的特征图, W_α^n 和 W_β^n 是 WAB 学习的权重, 分别分配给 F_{dab}^n 和 F_{deb}^n , * 表示加权操作。

从上述操作中, WAB 通过通道维度自适应地为 DAB 和 DEB 分配权重系数, 动态调整各通道特征在重建过程中的贡献。该机制在抑制冗余特征干扰的同时, 增强了关键信息的表征能力, 实现了浅层与深层特征之间的高效协同融合, 从而提升了图像重建效果。

1.2.1 细节注意力分支

为了更有效地利用图像的浅层特征并保持其完整性, 本文设计了一个细节注意力分支 DAB。该分支通过降低特征图的尺寸, 扩大卷积核的感受野, 以获取更丰富的全局结构信息。同时, DAB 通过融合输入的浅层特征图构建权重图, 对浅层特征进行加权处理, 从而增强细节特征的表达力。

DAB 结构如图 1(b) 所示。具体而言, 输入的特征图 F_i^n 首先经过一个 3×3 的卷积层, 随后经过最大池化层进行下采样, 以扩大后续卷积核的感受野。其中, 池化核尺寸为 7, 步长设为 3。接着, 利用 3 个 3×3 卷积层进一步提取图像的语义特征, 之后通过上采样层恢复特征图尺寸, 并将提取的语义特征与输入特征图相加, 以获得细节更加丰富的特征图。随后, 将该特征图通过 1×1 卷积层实现通道降维, 再经 sigmoid 函数将特征值归一化到 $[0, 1]$ 区间, 生成空间注意力图。最后, 将该注意力图与输入特征图加权融合, 得到细节增强的浅层特征图。DAB 的完整过程可表示为

$$F_{\text{ff}}^n = f_{\text{UP}}\left(C_{3,3}\left(\psi\left(C_3\left(F_s^n\right)\right)\right)\right) + F_s^n \quad (8)$$

$$F_{\text{dab}}^n = \sigma\left(C_1\left(F_{\text{ff}}^n\right)\right) * F_s^n \quad (9)$$

式中, F_s^n 为 DAB、DEB 和 WAB 的输入特征图, 即 $F_s^n = C_1(F_I^n)$, ψ 表示最大池化操作, $C_{3,3}$ 表示 3 个 3×3 卷积层, F_{ff}^n 表示语义特征图与输入的特征图的融合特征图, σ 表示 sigmoid 函数。

1.2.2 细节探索分支

为了深入挖掘图像深层特征, 本文设计了一个细节探索分支 DEB。该分支首先通过注意力机制增强输入特征的表达能力, 随后通过增加网络层数来学习更复杂的深层特征。此外, DEB 引入上采样和下采样操作, 以模拟图像在不同分辨率下的退化与重建过程, 从而捕捉更丰富的特征信息。在该过程中, 引入了多条跳跃连接以实现特征的反向传递与重复利用, 有效缓解了梯度消失现象, 同时提高了模型训练的稳定性与效率。

DEB 的结构如图 1(c) 所示。具体而言, 输入的特征图 F_s^n 首先经过一次反卷积操作以重建特征图, 重建后的特征图通过一个注意力模块实现特征的增强。该注意力模块包含两个互补的注意力分支, 其中一个分支通过最大池化和 1×1 卷积生成加权特征图, 主要强化高频信息; 另一个分支则采用平均池化配合 1×1 卷积生成加权特征图, 侧重于增强低频特征。这两个加权特征图分别与输入特征图进行加权相乘, 获得增强后的特征图。增强后的特征图依次经过 3×3 卷积和 1×1 卷积操作, 进行空间降尺寸与通道降维。接下来, 反卷积操作和 3×3 及 1×1 的卷积操作被多次执行, 采用这种上一下采样操作深入探索更深层次的特征信息。为了实现特征的高效复用, 上一下采样操作层构建了密集连接, 将学习到的不同尺度特征图传递到后续操作层中, 并与相同尺度的特征图进行连接, 从而实现更深层特征的学习。最后, 使用一个 1×1 卷积对特征图进行通道降维, 输出 F_{deb}^n 。DEB 的整个操作可表示为

$$L_0^n = F_s^n \quad (10)$$

$$H_x^n = D\left(C_1\left[L_0^n, L_1^n, L_2^n, \dots, L_{x-1}^n\right]\right) \quad (11)$$

$$H_{1*}^n = F_s^n * \sigma\left(C_1\left(\varphi\left(H_1^n\right)\right)\right) \times \sigma\left(C_1\left(\psi\left(H_1^n\right)\right)\right) \quad (12)$$

$$L_x^n = C_3\left(C_1\left[H_{1*}^n, H_2^n, H_3^n, \dots, H_x^n\right]\right) \quad (13)$$

$$F_{\text{deb}}^n = C_1\left([L_1^n, L_2^n, L_3^n, \dots, L_x^n]\right) \quad (14)$$

式中, H_x^n 表示第 n 次循环的第 x 次反卷积重建尺度

的特征图, H_{1*}^n 表示经过第 n 次循环的第 1 次反卷积, 输入到注意力模块后的增强特征图, L_x^n 表示第 n 次循环的第 x 次卷积的低尺度的特征图, φ 表示平均池化操作。

1.2.3 权重分配分支

在 DAB 中, 浅层特征得到了增强, 而 DEB 则提取了包含丰富浅层和深层信息的特征图。如果直接合并这两种特征图, 则会不可避免地存在较多的特征冗余。因此, 为了有效利用这两种特征, 本文设计了权重分配分支 WAB, 通过通道注意力机制自适应地为 DAB 和 DEB 提取的特征图分配权重系数, 实现了两种特征的整合和补充。

WAB 的具体结构如图 1(d) 所示。输入特征图 F_s^n 首先经过与 DEB 结构相同的注意力模块以增强特征。然后, 通过平均池化在空间维度上对特征图进行压缩, 将二维特征图转换为 1×1 的特征向量, 并调整其长度以匹配输入特征图的通道数。随后, 通过两个全连接层 δ 生成两个权值, 最后经 softmax 函数, 计算出动态权重值 W_α^n 和 W_β^n 。WAB 的整个操作可以表示为

$$F_{\text{em}}^n = F_s^n * \sigma\left(C_1\left(\varphi\left(F_s^n\right)\right)\right) \times \sigma\left(C_1\left(\psi\left(F_s^n\right)\right)\right) \quad (15)$$

$$\{W_\alpha^n, W_\beta^n\} = \text{softmax}\left(\delta\left(\delta\left(\varphi\left(F_{\text{em}}^n\right)\right)\right)\right) \quad (16)$$

式中, F_{em}^n 表示增强后的特征图, δ 表示全连接层, $W_\alpha^n + W_\beta^n = 1$ 。

1.3 损失函数

在 SISR 重建中, 网络训练通常使用 L1 损失函数对重建图像进行约束, 与其他常见的损失函数(如 L2 损失)相比, L1 损失对异常值不敏感, 具有更好的鲁棒性, 并且能够在抑制伪影的同时较好地保留图像细节(Jiang 等, 2023)。在本研究中, 考虑到网络结构采用循环迭代的方式生成重建结果, 在每一次循环输出时, 均利用参考 HR 图像进行 L1 损失计算, 从而在迭代过程中持续引导网络参数更新, 确保最终重建结果的精确性。本文构建网络的 L1 损失函数定义为

$$L = \frac{1}{N} \sum_{n=1}^N \frac{1}{m} \sum_{i=1}^m \left\| (Y_{\text{SR}}^n)^i - I_{\text{HR}}^i \right\|_1 \quad (17)$$

式中, m 是训练图像对的数量, I_{HR} 表示参考 HR 图像, Y_{SR}^n 表示 AFFRN 重建的第 n 个循环的 HR 图像, $\|\cdot\|_1$ 表示 L1 范数。

2 实验与分析

为验证所提方法的有效性,本文与多种先进的SISR重建方法进行了对比,包括: Bicubic, VDSR (Kim等, 2016a)、DRCN (Kim等, 2016b)、LapSRN (Laplacian pyramid super-resolution network) (Lai等, 2017)、CARN (cascading residual network) (Ahn等, 2018)、SRFBN-S (image super-resolution feedback network) (Li等, 2019)、LAPAR-A (linearly-assembled pixel-adaptive regression network) (Li等, 2020)、SMSR (sparse mask super-resolution network) (Wang等, 2021)、ESRT (Lu等, 2022)、ShuffleMixer (Sun等, 2022)、SAFMN (spatially-adaptive feature modulation network) (Sun等, 2023)、LSRN-AED (lightweight super-resolution network based on asymmetric encoder-decoder) (Huang等, 2024a)、SPAN (Wan等, 2024)、AFSMNet (adaptive feature selection modulation network) (Wu等, 2025)、MemNet (Tai等, 2017)、EDSR (Lim等, 2017)、MSLapSRN (deep Laplacian pyramid super-resolution network based on multi-scale training strategy) (Lai等, 2019)、MSRN (multi-scale residual network) (Li等, 2018)和MFSN (multi-scale feature selection network) (Li等, 2024)。为了更好地比较各方法的重建性能,本文展示了多种对比方法的主观视觉效果和客观定量指标,并对所提模型各组成模块进行消融实验,以验证各模块对整体性能贡献。

2.1 数据集与实验设置

在实验中,本文使用广泛应用于SISR任务的DIV2K数据集 (Agustsson和Timofte, 2017)作为训练集,包含1 000幅2 K分辨率的高质量RGB图像数据集,涵盖动植物、自然风景等多样化场景。在本研究中,遵循Liang等人(2021)的研究方法,选取其中900幅HR图像,并采用双三次插值分别以 $\times 2$ 、 $\times 3$ 和 $\times 4$ 的倍数进行下采样,生成对应的LR图像。为了增加网络的泛化性,利用EDSR (Lim等, 2017)中采用的数据集增强方式对图像对进行了数据增强,如随机旋转 90° 、 180° 、 270° 和水平翻转等。为了公平地比较各种方法的性能,各种方法在SISR标准公开数据集上进行测试: Set5 (Bevilacqua等, 2012), Set14 (Zeyde等, 2012), BSD100 (Berkeley segmentation

dataset 100) (Martin等, 2001), Urban100 (Huang等, 2015)和Manga109 (Matsui等, 2017),并采用峰值信噪比 (peak signal-to-noise ratio, PSNR)和结构相似性指数 (structure similarity index measure, SSIM) (Wang等, 2004)对测试结果进行客观定量评价,定量结果均在亮度(Y)通道上进行计算。为了进一步验证本文模型的鲁棒性,在historical (Lai等, 2017)真实数据集上进行了测试。验证集采用Set5数据集,用于模型选择与参数调整。

在模型训练过程中,训练数据被随机裁剪生成大小为 48×48 像素的图像块,批次大小(batch size)设置为16,共训练800个epoch。初始学习率设置为 1×10^{-4} ,并在每200个epoch后减半。adam优化器用于训练网络,其中 $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$ 。本文所提出的模型训练与测试的硬件环境配置为NVIDIA GeForce RTX1080Ti GPU的计算机。软件环境为64位Win10操作系统,采用PyTorch框架。

2.2 消融实验

为了验证所提模型中各模块及参数设置的有效性与合理性,本文设计了两个消融实验,并对实验结果进行分析。

实验1对AFFM中的DAB、DEB和WAB这3个模块的4种不同组合进行了实验,并给出了在测试集Urban100上尺度因子 $\times 4$ 的重建结果的平均PSNR/SSIM值,如表1所示。其中,模型1表示仅包含DAB的模型,模型2为仅包含DEB的模型,模型3为包含DAB和DEB的模型,模型4为本文模型,即包含3个模块的模型。

从表1可知,模型2比模型1获得了更高的指标,主要是因为DAB仅对浅层特征给予关注,DEB则探索了更多的深层特征,这对网络的重建性能更加重要。而模型3结合了DAB和DEB所学习的不同深度特征,实现了更丰富的特征利用,获得了高于模型1和模型2的指标值。模型4中增加了WAB,增益或抑制DAB和DEB学习到的特征,实现高效的特征融合,有效避免了特征冗余所造成的影响,从而获得了更高的客观指标值。

此外,为了更直观地展示各分支在不同类型图像区域中的作用,对DAB和DEB的特征进行了可视化分析,如图3所示。可以看到,DAB的特征图在蝴蝶翅膀等纹理细节丰富的区域表现出显著响应,能够突出高频的边缘与细粒度纹理信息;而DEB的特

表1 在Urban100($\times 4$)数据集, $N=4$ 下的
AFFM的消融实验结果Table 1 The ablation experiment results of AFFM with
 $N=4$ on the Urban100 ($\times 4$) dataset

方法	DAB	DEB	WAB	PSNR/dB	SSIM
模型1	√	×	×	24.86	0.740 9
模型2	×	√	×	26.28	0.791 2
模型3	√	√	×	26.37	0.795 4
模型4(本文)	√	√	√	26.61	0.800 7

注:加粗字体表示各列最优结果。“√”和“×”分别表示采用和未采用。

征图在平滑或低频区域(如翅膀的大面积区域)中响应更强,主要聚焦于捕获全局的语义结构与形状信息。由此可见,DAB有助于增强局部边缘与细节特征,而DEB更倾向于建模整体语义与结构信息。两者在特征关注范围上的互补性,充分验证了在AFFM中联合设计DAB与DEB的合理性与有效性。由此可知,本文所提AFFM结构的有效性,并且其中三支的设计均能为网络的重建性能做出贡献,得到了更好的结果。

实验2是验证AFFM所执行循环次数 N 对模型性能的影响。循环利用AFFM进行特征提取的目的是逐步实现特征的重建和补充。在实验2中,对设定的循环总次数 N 取不同值时的模型进行训练,并在5个标准数据集上对尺度因子 $\times 4$ 的重建结果进行PSNR和SSIM值的对比,如表2所示。从表2可以看到,当 N 从1增加到4时,5个测试集上的PSNR/SSIM值均有所提升,表明适度增加循环深度有助于增强特征表达能力。而当 N 增加到5后,性能指标

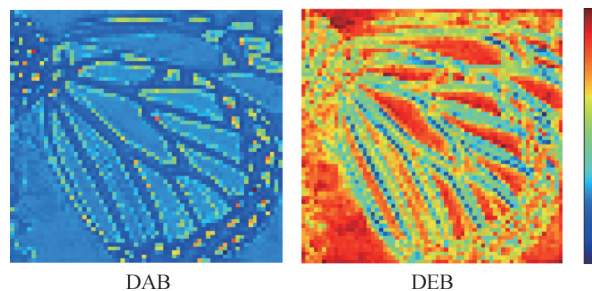


图3 AFFM中DAB与DEB分支输出特征的可视化结果

Fig. 3 Visualization of feature maps from DAB and
DEB in AFFM

反而下降,可能原因是过多的循环引入了冗余特征信息,导致模型过拟合或信息干扰,影响了重建质量。为了在性能与计算效率之间取得平衡,本文最终选择 N 为4作为网络循环深度。此外,该实验同时也验证了所设计的循环机制在逐步补充和增强图像细节方面的有效性,从而使AFFM发挥出最大的性能。

2.3 实验结果及分析

为验证所提模型的性能,本文在5个标准数据集上分别对尺度因子 $\times 2$ 、 $\times 3$ 、 $\times 4$ 、 $\times 8$ 下的性能进行了测试,并与一些先进的SISR方法进行客观指标和主观视觉效果上的对比。

2.3.1 客观指标比较

表3—表5显示了各种比较方法在5个标准公开数据集上,尺度因子为 $\times 2$ 、 $\times 3$ 、 $\times 4$ 下的客观指标值。从表中可以看到,本文方法整体性能优于其他对比方法。由于本文模型更加注重对提取特征的增强与利用,因此在纹理复杂的Urban100数据集上表现尤为出色。相比于该数据集中排名第2的方法,所提模型的各项指标提升了0.14~

表2 在基准数据集($\times 4$)上对循环递归深度 N 的消融实验结果Table 2 The ablation experiment results for the recursive depth N on benchmark datasets ($\times 4$)

N	Set5		Set14		BSD100		Urban100		Manga109	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
1	32.18	0.894 5	28.56	0.781 4	27.55	0.735 4	26.04	0.783 9	30.33	0.906 8
2	32.30	0.895 7	28.65	0.782 4	27.58	0.736 4	26.24	0.789 5	30.57	0.909 7
3	32.30	0.896 0	28.65	0.782 7	27.59	0.737 0	26.30	0.792 1	30.68	0.910 9
4(本文)	32.47	0.897 6	28.78	0.786 3	27.68	0.739 7	26.61	0.800 7	31.08	0.915 3
5	32.32	0.896 7	28.71	0.784 3	27.63	0.737 4	26.41	0.794 8	30.79	0.912 1
6	32.27	0.895 6	28.64	0.783 1	27.56	0.736 8	26.34	0.793 0	30.67	0.911 0

注:加粗字体表示各列最优结果。

表3 不同方法在基准数据集上尺度因子为×2下的PSNR和SSIM的平均结果

Table 3 The average PSNR/SSIM results of different methods on benchmark datasets for ×2 SR

方法	年份	Set5		Set14		BSD100		Urban100		Manga109	
		PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
Bicubic	-	33.66	0.929 9	30.24	0.868 8	29.56	0.843 1	26.88	0.840 3	30.80	0.933 9
VDSR	2016	37.53	0.959 0	33.05	0.913 0	31.90	0.896 0	30.77	0.914 0	37.22	0.975 0
DRCN	2016	37.63	0.958 8	33.04	0.911 8	31.85	0.894 2	30.75	0.913 3	37.57	0.973 0
LapSRN	2017	37.78	0.960 0	33.28	0.915 0	32.05	0.898 0	31.15	0.919 0	37.78	0.976 0
CARN	2018	37.76	0.959 0	33.52	0.916 6	32.09	0.897 8	31.92	0.925 6	38.36	0.976 4
SRFBN-S	2019	37.78	0.959 7	33.35	0.915 6	32.00	0.897 0	31.41	0.920 7	38.06	0.975 7
LAPAR-A	2020	38.01	0.960 5	33.62	0.918 3	32.19	0.899 9	32.10	0.928 3	38.67	0.977 2
SMSR	2021	38.00	0.960 1	33.64	0.917 9	32.17	0.899 0	32.19	0.928 4	38.76	0.977 1
ESRT	2022	38.03	0.960 0	<u>33.75</u>	0.918 4	32.25	0.900 1	<u>32.58</u>	<u>0.931 8</u>	39.12	0.977 4
ShuffleMixer	2022	38.01	0.960 6	33.63	0.918 0	32.17	0.899 5	31.89	0.925 7	38.83	0.977 4
SAFMN	2023	38.00	0.960 5	33.54	0.917 7	32.16	0.899 5	31.84	0.925 6	38.71	0.977 1
LSRN-AED	2024	<u>38.06</u>	0.960 7	33.68	<u>0.919 4</u>	32.25	0.900 6	32.33	0.930 8	38.67	0.977 2
SPAN	2024	38.08	0.960 8	33.71	0.918 3	<u>32.22</u>	0.900 2	32.24	0.929 4	38.94	0.977 7
AFSMNet	2025	38.03	0.961 0	33.57	0.917 7	32.18	0.899 8	31.97	0.926 2	38.85	0.977 5
AFFRN (本文)	2025	38.08	<u>0.960 9</u>	33.86	0.920 8	32.25	<u>0.900 3</u>	32.72	0.933 3	<u>38.96</u>	0.977 7

注:加粗、下划线字体表示各列最优、次优结果。

表4 不同方法在基准数据集上尺度因子为×3下的PSNR和SSIM的平均结果

Table 4 The average PSNR/SSIM results of different methods on benchmark datasets for ×3 SR

方法	年份	Set5		Set14		BSD100		Urban100		Manga109	
		PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
Bicubic	-	30.39	0.868 2	27.55	0.774 2	27.21	0.738 5	24.46	0.734 9	26.95	0.855 6
VDSR	2016	33.66	0.921 3	29.77	0.831 4	28.82	0.797 6	27.14	0.827 9	32.01	0.934 0
DRCN	2016	33.82	0.922 6	29.76	0.831 1	28.80	0.796 3	27.15	0.827 6	32.31	0.936 0
LapSRN	2017	34.06	0.924 0	29.97	0.836 0	28.93	0.802 0	27.47	0.837 0	32.68	0.939 0
CARN	2018	34.29	0.925 5	30.29	0.840 7	29.06	0.803 4	28.06	0.849 3	33.49	0.944 0
SRFBN-S	2019	34.20	0.925 5	30.10	0.837 2	28.96	0.801 0	27.66	0.841 5	33.02	0.940 4
LAPAR-A	2020	34.36	0.926 7	30.34	0.842 1	29.11	0.805 4	28.15	0.852 3	33.51	0.944 1
SMSR	2021	34.40	0.927 0	30.33	0.841 2	29.10	0.805 0	28.25	0.853 6	33.68	0.944 5
ESRT	2022	34.42	0.926 8	<u>30.43</u>	0.843 3	<u>29.15</u>	0.806 3	<u>28.46</u>	<u>0.857 4</u>	33.95	<u>0.945 5</u>
ShuffleMixer	2022	34.40	0.927 2	30.37	0.842 3	29.12	0.805 1	28.08	0.849 8	33.69	0.944 8
SAFMN	2023	34.34	0.926 7	30.33	0.841 8	29.08	0.804 8	27.95	0.847 4	33.52	0.943 7
LSRN-AED	2024	<u>34.50</u>	<u>0.928 1</u>	30.45	<u>0.844 0</u>	29.11	<u>0.805 9</u>	28.32	0.854 4	33.36	0.938 5
AFSMNet	2025	34.39	0.927 6	30.36	0.843 0	29.12	0.805 5	28.08	0.849 6	33.74	0.945 1
AFFRN (本文)	2025	34.60	0.928 5	30.45	0.844 6	29.16	0.806 3	28.61	0.860 9	<u>33.91</u>	0.946 5

注:加粗、下划线字体表示各列最优、次优结果。

表5 不同方法在基准数据集上尺度因子为×4下的PSNR/SSIM的平均结果

Table 5 The average PSNR/SSIM results of different methods on benchmark datasets for ×4 SR

方法	年份	Set5		Set14		BSD100		Urban100		Manga109	
		PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
Bicubic	-	28.42	0.810 4	26.00	0.702 7	25.96	0.667 5	23.14	0.657 7	24.89	0.786 6
VDSR	2016	31.35	0.883 8	28.01	0.767 4	27.29	0.725 1	25.18	0.754 0	28.83	0.887 0
DRCN	2016	31.53	0.885 4	28.02	0.767 0	27.23	0.723 3	25.14	0.751 0	28.97	0.886 0
LapSRN	2017	31.74	0.889 0	28.26	0.774 0	27.43	0.731 0	25.51	0.768 0	29.54	0.897 0
CARN	2018	32.13	0.893 7	28.60	0.780 6	27.58	0.734 9	26.07	0.783 7	30.40	0.908 2
SRFBN-S	2019	31.98	0.892 3	28.45	0.777 9	27.44	0.731 3	25.71	0.771 9	29.91	0.900 8
LAPAR-A	2020	32.15	0.894 4	28.61	0.781 8	27.61	0.736 6	26.14	0.787 1	30.42	0.907 4
SMSR	2021	32.12	0.893 2	28.55	0.780 8	27.55	0.735 1	26.11	0.786 8	30.54	0.908 5
ESRT	2022	32.19	0.894 7	28.69	0.783 3	27.69	0.737 9	<u>26.39</u>	<u>0.796 2</u>	<u>30.75</u>	0.910 0
ShuffleMixer	2022	32.21	0.895 3	28.66	0.782 7	27.61	0.736 6	26.08	0.783 5	30.65	0.909 3
SAFMN	2023	32.18	0.894 8	28.60	0.781 3	27.58	0.735 9	25.97	0.780 9	30.43	0.906 3
LSRN-AED	2024	<u>32.31</u>	<u>0.896 6</u>	<u>28.72</u>	<u>0.784 6</u>	27.63	<u>0.738 1</u>	26.25	0.791 1	30.66	<u>0.910 5</u>
SPAN	2024	32.20	0.895 3	28.66	0.783 4	27.62	0.737 4	26.18	0.787 9	30.66	0.910 3
AFSMNet	2025	32.03	0.894 9	28.26	0.783 4	27.62	0.737 5	26.07	0.783 7	30.70	0.909 7
AFFRN (本文)	2025	32.47	0.897 6	28.78	0.786 3	<u>27.68</u>	0.739 7	26.61	0.800 7	31.08	0.915 3

注:加粗、下划线字体表示各列最优、次优结果。

0.22 dB。同时,在其他数据集上同样取得了显著的性能提升,进一步验证了所提方法的有效性与优越性。

在当前的SISR研究中,大多数方法主要聚焦于对LR图像进行×2、×3、×4尺度因子下的重建,而×8尺度因子下的重建效果仍有待提高,本文针对×8尺度因子下的重建也进行了实验对比,其结果如表6所示。从表6可见,所提网络得益于对不同深度特征的有效整合,以及循环机制的逐步增强与补充,其性能得到了显著提升。本文方法在5个标准公开数据集上大多数情况下的PSNR和SSIM指标均优于现有方法,展现出良好的整体性能。但在部分数据集上(如Manga109),PSNR指标的性能略低于ESRT,排名第2。这是因为ESRT采用了高效的Transformer模块,能够更有效地建模图像中相似区域的长程依赖,而本文方法在全局信息依赖关系的建模能力上相对有限。未来将进一步探索模型在处理复杂图像和多样化场景中的适应性与鲁棒性。

表7对比了所提模型与多种主流SR方法在参数量(Params)和浮点运算次数(floating point opera-

tions, FLOPs)方面的表现。表8给出了几种先进算法与所提模型在运行时间(runtime)和显存消耗(GPU memory)上的对比。实验均在放大倍率为×4的条件下,将LR图像重建至1280×720像素进行测试。从表中结果可以看出,虽然AFFRN在参数量和FLOPs上相较部分轻量化模型略有增加,但在相近计算开销下,其重建性能显著优于对比方法。例如,与FLOPs接近的ERST(Transformer架构)相比,AFFRN在Urban100数据集上取得了更高的PSNR和SSIM值,这表明在相似计算复杂度下,基于CNN的AFFRN在高频纹理重建方面同样具备强大的特征表达能力。同时,与同为CNN架构的LAPAR-A相比,AFFRN在精度上也取得了明显提升,进一步验证了所提自适应特征融合与循环机制在CNN框架下的有效性。总体来看,所提模型在保持较低计算负担的同时,兼顾了性能表现与资源效率,体现出良好的精度与效率之间的权衡。

2.3.2 主观视觉效果比较

为了更直观地对不同方法的重建结果进行比较,展示了分辨率×4和×8重建结果的主观效果图。

表6 不同方法在各数据集上尺度因子为×8下的PSNR/SSIM平均结果

Table 6 The average PSNR/SSIM results of different methods on benchmark datasets for ×8 SR

方法	年份	Set5		Set14		BSD100		Urban100		Manga109	
		PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
Bicubic	-	24.40	0.658 0	23.10	0.566 0	23.67	0.548 0	20.74	0.516 0	21.47	0.650 0
VDSR	2016	25.93	0.724 0	24.26	0.614 0	24.49	0.583 0	21.70	0.571 0	23.16	0.725 0
DRCN	2016	25.93	0.674 3	24.25	0.551 0	24.49	0.516 8	21.71	0.528 9	23.20	0.668 6
LapSRN	2017	26.15	0.738 0	24.35	0.620 0	24.54	0.586 0	21.81	0.581 0	23.39	0.735 0
MemNet	2017	26.16	0.741 4	24.38	0.619 9	24.58	0.584 2	21.89	0.582 5	23.56	0.738 7
EDSR	2017	26.96	<u>0.776 2</u>	24.91	<u>0.642 0</u>	<u>24.81</u>	<u>0.598 5</u>	<u>22.51</u>	<u>0.622 1</u>	<u>24.69</u>	<u>0.784 1</u>
MSLapSRN	2018	26.34	0.755 8	24.57	0.627 3	24.65	0.589 5	22.06	0.596 3	23.90	0.756 4
MSRN	2018	26.59	0.725 4	24.88	0.596 1	24.70	0.541 0	22.37	0.597 7	24.28	0.751 7
MFSN	2024	<u>26.97</u>	0.773 8	<u>25.00</u>	0.641 3	24.80	0.596 7	22.45	0.616 1	24.51	0.776 5
AFFRN (本文)	2025	27.27	0.785 5	25.13	0.646 5	24.90	0.600 6	22.75	0.630 0	24.95	0.791 2

注:加粗、下划线字体表示各列最优、次优结果。

表7 不同模型参数量和计算量比较(Urban 100×4)

Table 7 Comparison of parameters and FLOPs of different methods (Urban 100×4)

方法	Params/M	FLOPs/G	PSNR/dB	SSIM
VDSR	0.7	612.6	25.18	0.754 0
DRCN	1.8	9 788.7	25.14	0.751 0
CARN	1.6	90.9	26.07	0.783 7
SRFBN-S	0.5	852.9	25.71	0.771 9
LapSRN	0.8	149.4	25.51	0.768 0
LAPAR-A	0.7	94.0	26.14	0.787 1
MemNet	0.7	623.9	25.50	0.763 0
SMSR	1.0	57.2	26.11	0.786 8
ESRT	0.8	67.7	<u>26.39</u>	<u>0.796 2</u>
ShuffleMixer	0.4	28.0	26.08	0.783 5
SAFMN	0.2	14.0	25.97	0.780 9
LSRN-AED	0.6	33.6	26.25	0.791 1
SPAN	0.5	24.5	26.18	0.787 9
AFSMNet	0.3	13.0	26.07	0.783 7
AFFRN (本文)	1.9	86.4	26.61	0.800 7

注:加粗、下划线字体表示PSNR和SSIM的最优、次优结果。

通过这些对比图,可以明显看出,本文方法在重建细节、减少边缘伪影和避免纹理扭曲方面具有显著优势。

图4展示了所提算法与其他10种比较方法在

表8 不同模型显存消耗和运行时间比较(Urban 100×4)

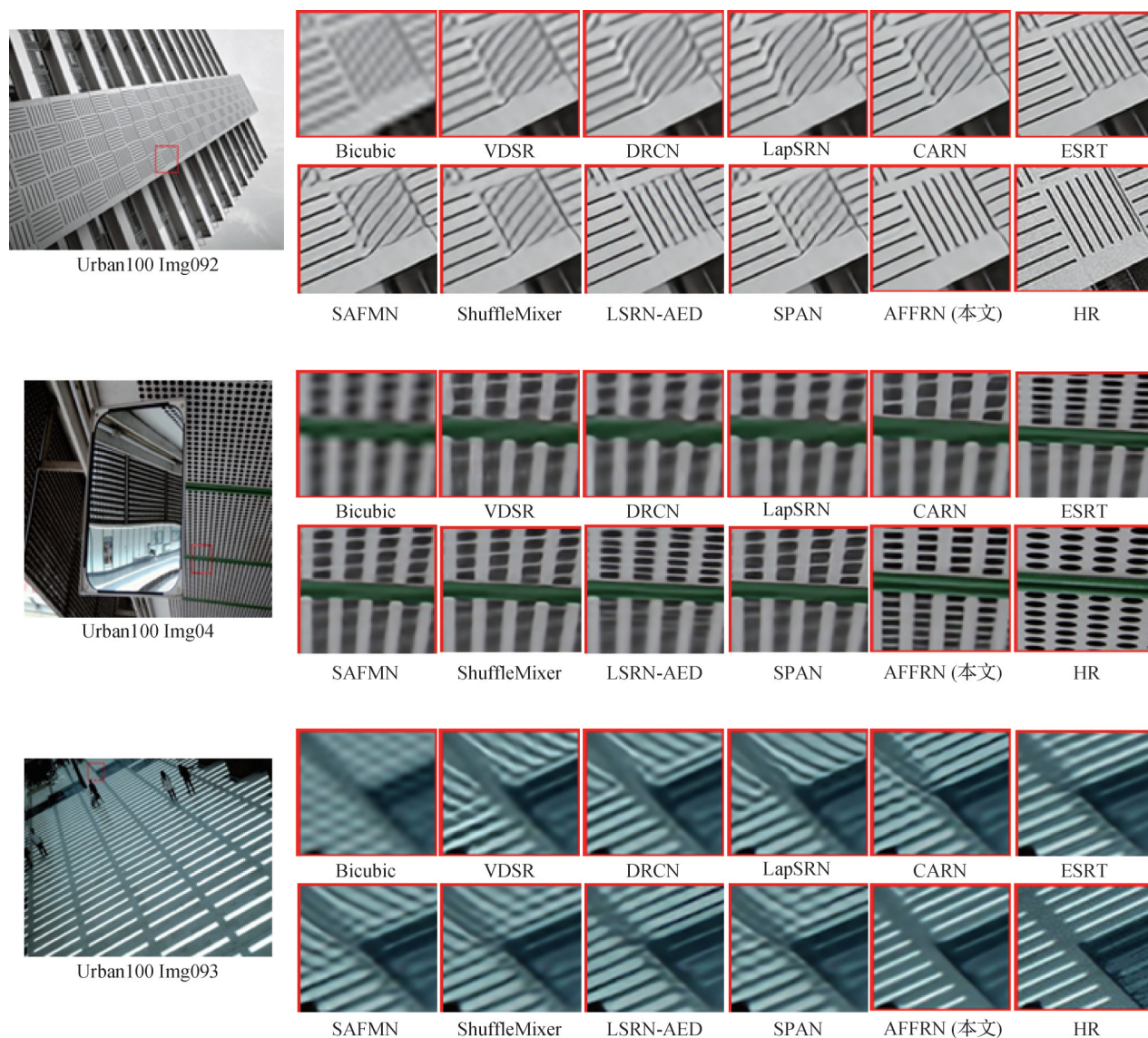
Table 8 Comparison of runtime and GPU memory consumption of different methods (Urban 100×4)

方法	运行时间/ms	显存消耗/G	PSNR/dB	SSIM
ESRT	197.1	14.8	<u>26.39</u>	<u>0.796 2</u>
ShuffleMixer	32.2	1.4	26.08	0.783 5
SAFMN	17.3	1.0	25.97	0.780 9
LSRN-AED	1 447.8	6.0	26.25	0.791 1
SPAN	37.5	1.7	26.18	0.787 9
AFFRN (本文)	370.1	4.0	26.61	0.800 7

注:加粗、下划线字体表示PSNR和SSIM的最优、次优结果。

Urban100数据集集中的3幅图像进行×4超分重建的结果。

为了便于观察,本文对重建结果中的局部区域进行放大展示。Bicubic的结果是采用简单的双三次插值获得的结果,作为基准比较图像。该结果图像是最模糊的,基本没有重建出纹理细节和边缘信息。多数深度学习方法在一定程度上改善了清晰度,但仍存在边缘模糊、纹理方向偏差或形状失真的问题。相比之下,本文方法的重建结果细节更清晰、边缘更锐化,与参考HR图像(ground truth, GT)的纹理细节是最接近的。值得注意的是,与基于Transformer架构的ERST相比,AFFRN在高频纹理区域的恢复更加清晰且结构完整,能够准确还原原始纹理

图4 在Urban100数据集上尺度因子 $\times 4$ 的重建结果比较Fig. 4 Comparison of reconstruction results on the Urban100 dataset for $\times 4$ SR

的方向和细节。这得益于CNN在局部特征建模上的优势,结合自适应特征融合与循环机制,可多次强化高频信息,从而弥补Transformer在细节重建上的平滑化倾向。

图5显示了来自BSD100数据集和Set14数据集中3幅图像的 $\times 4$ 超分重建结果。从图中同样可以观察到本文方法相较于其他方法,重建出了更丰富的纹理信息,并获得了更锐化、更清晰的图像边缘。

图6显示了所提算法与其他比较方法对Manga109数据集中两幅图像进行 $\times 8$ 尺度因子下的重建结果。可以观察到,本文方法在较大的放大倍数下,相较于其他算法,获得了具有清晰边缘和更多纹理的重建结果。例如,在图6 KyokugenCyclone中

的“高”字,其他方法重建结果很难分辨其中“口”的笔画,本文的结果可以更加清晰地辨认。

为验证模型的泛化能力,图7展示了在真实数据集historical中Img007进行 $\times 4$ SR重建的结果。为了便于观察,本文展示了重建结果中两处纹理丰富区域的放大结果图,如蓝色框和红色框所示。从图中可以看到,本文方法对真实LR图像进行重建,能够获得纹理更丰富、边缘更清晰的重建图像。例如,图中楼房窗户的相关细节,本文的结果可以很好地被辨认,而其他方法的结果比较模糊,许多细节难以辨识。这表明了本文算法在实际应用中依然表现出优异性能,展现了较强的泛化能力和鲁棒性。

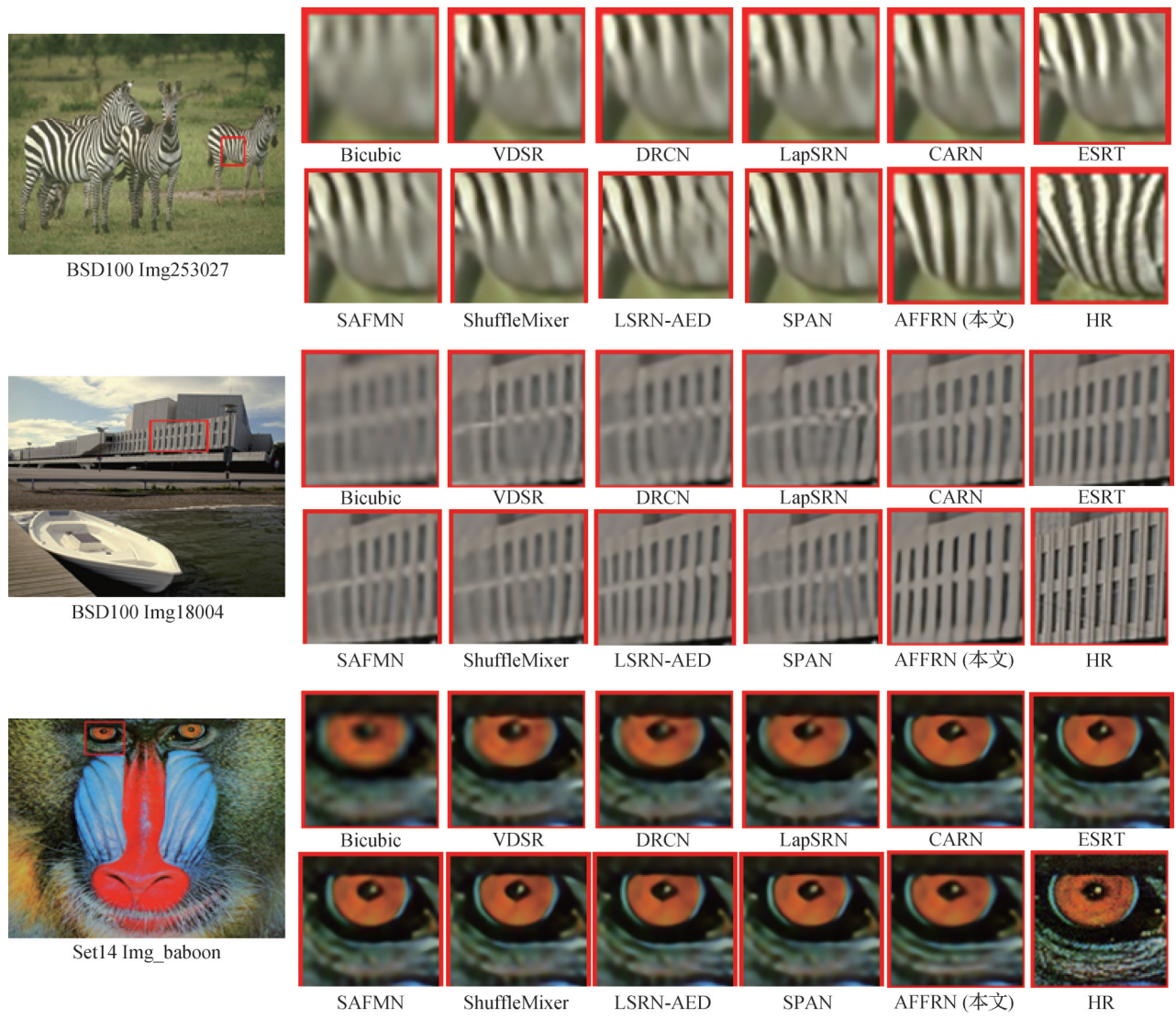


图5 在BSD100数据集和Set14数据集上尺度因子×4的重建结果比较

Fig. 5 Comparison of reconstruction results on the BSD100 and Set14 datasets for ×4 SR

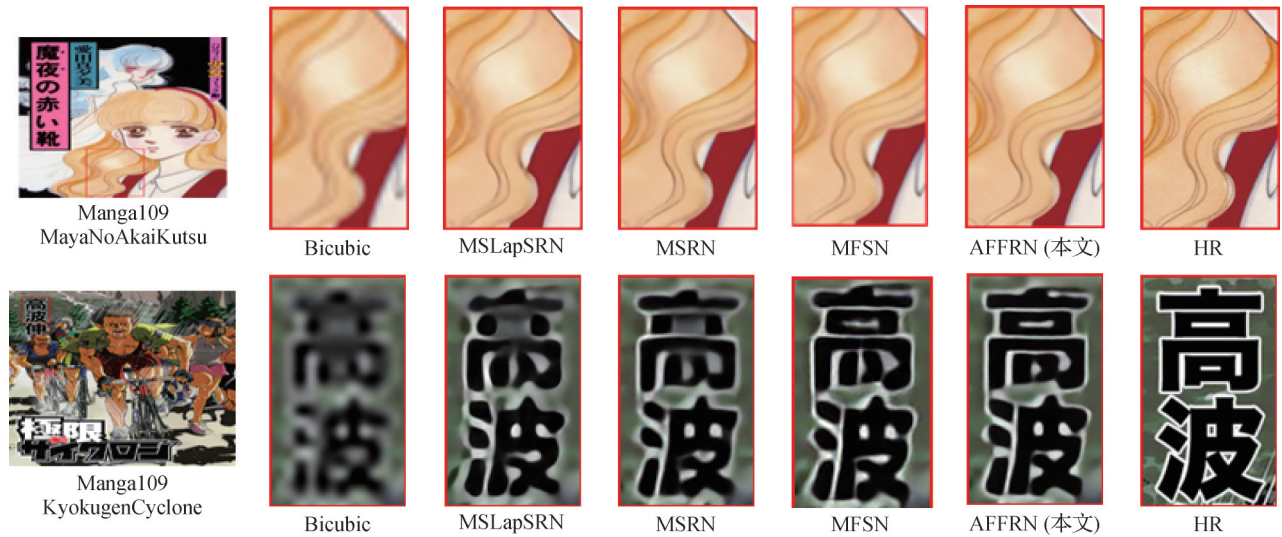


图6 在Manga109数据集上尺度因子为×8的重建结果比较

Fig. 6 Comparison of reconstruction results on the Manga109 dataset for ×8 SR

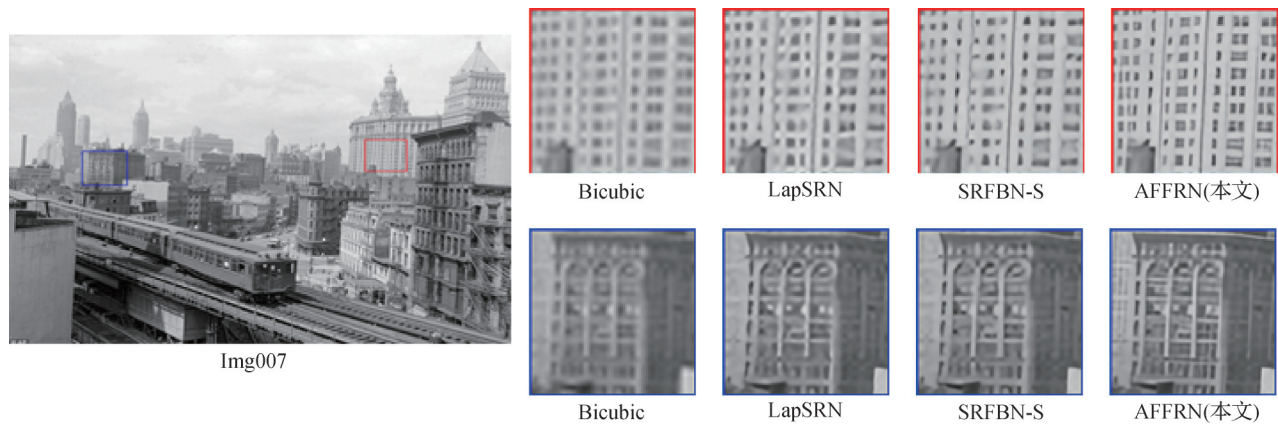


图7 真实图像historical数据集上尺度因子为 $\times 4$ 的重建结果比较
Fig. 7 Comparison of reconstruction results on the real image dataset for $\times 4$ SR

3 结论

本文提出了一种基于自适应特征融合的循环网络的超分辨率重建方法 AFFRN, 通过逐步细化的循环结构实现了图像细节特征的自下而上的提取与增强, 兼顾了重建质量与计算效率。网络采用简洁的卷积层结构进行初始特征提取与图像重建, 保证了整体的轻量化设计。在 AFFRN 中, 核心模块为 AFFM, 其通过嵌入循环结构, 在每一阶段融合并细化不同深度的图像特征。AFFM 包含 DAB 与 DEB, 分别增强浅层细节与挖掘深层结构特征, 并设计基于通道级动态权重机制的 WAB, 自适应分配 DAB 和 DEB 提取特征的权重, 实现了不同深度特征的有效融合, 减少了冗余信息, 显著提升了图像重建质量。在多个公开数据集上与现有先进方法的对比实验结果表明, 本文方法在 PSNR、SSIM 等客观指标以及重建图像的视觉效果方面均取得了显著提升, 验证了所提网络结构的有效性。然而, 尽管本文方法在保持较低计算复杂度的前提下取得了良好的重建性能, 其仍存在一些潜在的局限性。例如, 由于网络设计中采用了固定结构的递归机制, 可能在面对极端退化或跨域图像数据时, 特征的表达能力和泛化性能仍有限; 此外, 尽管 AFFRN 为轻量结构, 但在大尺寸高分辨率图像上的推理效率仍有进一步优化空间。因此, 未来的研究将重点关注提升模型的跨域适应能力与实际部署效率, 同时探索将本文方法扩展至视频超分辨率、实时监控与高层次视觉理解任务中的可行性与应用效果。

参考文献 (References)

- Agustsson E and Timofte R. 2017. NTIRE 2017 challenge on single image super-resolution: dataset and study//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, USA: IEEE: 1122-1131 [DOI: 10.1109/CVPRW.2017.150]
- Ahn N, Kang B and Sohn K A. 2018. Fast, accurate, and lightweight super-resolution with cascading residual network//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 256-272 [DOI: 10.1007/978-3-030-01249-6_16]
- Bevilacqua M, Roumy A, Guillemot C and Morel M L A. 2012. Low-complexity single-image super-resolution based on nonnegative neighbor embedding//Proceedings of the 23rd British Machine Vision Conference. Surrey, UK: BMVA Press: 135.1-135.10 [DOI: 10.5244/C.26.135]
- Dai T, Cai J R, Zhang Y B, Xia S T and Zhang L. 2019. Second-order attention network for single image super-resolution//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA, IEEE: 11057-11066 [DOI: 10.1109/CVPR.2019.01132]
- Dong C, Loy C C, He K M and Tang X O. 2014. Learning a deep convolutional network for image super-resolution//Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland: Springer: 184-199 [DOI: 10.1007/978-3-319-10593-2_13]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Hsu W Y and Wu C H. 2025. Wavelet structure-texture-aware super-resolution for pedestrian detection. Information Sciences, 691: #121612 [DOI: 10.1016/j.ins.2024.121612]
- Huang G, Liu Z, Van Der Maaten L and Weinberger K Q. 2017.

- Densely connected convolutional networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 2261-2269 [DOI: 10.1109/CVPR.2017.243]
- Huang J B, Singh A and Ahuja N. 2015. Single image super-resolution from transformed self-exemplars//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE: 5197-5206 [DOI: 10.1109/CVPR.2015.7299156]
- Huang S Y, Li W, Yang Y, Wan W G and Lai H Z. 2024a. LSRN-AED: lightweight super-resolution network based on asymmetric encoder-decoder. *Soft Computing*, 28(13): 8513-8525 [DOI: 10.1007/s00500-024-09745-5]
- Huang W F, Liao X Y, Chen H, Hu Y, Jia W J and Wang Q. 2024b. Deep local-to-global feature learning for medical image super-resolution. *Computerized Medical Imaging and Graphics*, 115: #102374 [DOI: 10.1016/j.compmedimag.2024.102374]
- Jiang J J, Wang C Y, Liu X M and Ma J Y. 2023. Deep learning-based face super-resolution: a survey. *ACM Computing Surveys*, 55(1): #13 [DOI: 10.1145/3485132]
- Keys R. 1981. Cubic convolution interpolation for digital image processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(6): 1153-1160 [DOI: 10.1109/TASSP.1981.1163711]
- Kim J, Lee J K and Lee K M. 2016a. Accurate image super-resolution using very deep convolutional networks//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 1646-1654 [DOI: 10.1109/CVPR.2016.182]
- Kim J, Lee J K and Lee K M. 2016b. Deeply-recursive convolutional network for image super-resolution//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 1637-1645 [DOI: 10.1109/CVPR.2016.181]
- Lai W S, Huang J B, Ahuja N and Yang M H. 2017. Deep Laplacian pyramid networks for fast and accurate super-resolution//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 5835-5843 [DOI: 10.1109/CVPR.2017.618]
- Lai W S, Huang J B, Ahuja N and Yang M H. 2019. Fast and accurate image super-resolution with deep Laplacian pyramid network. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(11): 2599-2613 [DOI: 10.1109/TPAMI.2018.2865304]
- Li J C, Fang F M, Mei K F and Zhang G X. 2018. Multi-scale residual network for image super-resolution//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 527-542 [DOI: 10.1007/978-3-030-01237-3_32]
- Li L, Zhang Y, Yuan L and Gao X B. 2025. SANet: face super-resolution based on self-similarity prior and attention integration. *Pattern Recognition*, 157: #110854 [DOI: 10.1016/j.patcog.2024.110854]
- Li M H, Zhao Y Q, Zhang F, Luo B, Yang C H, Gui W H, et al. 2024. Multi-scale feature selection network for lightweight image super-resolution. *Neural Networks*, 169: 352-364 [DOI: 10.1016/j.neunet.2023.10.043]
- Li W B, Zhou K, Qi L, Jiang N J, Lu J B and Jia J Y. 2020. LAPAR: linearly-assembled pixel-adaptive regression network for single image super-resolution and beyond//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.: #1708
- Li Z, Yang J L, Liu Z, Yang X M, Jeon G and Wu W. 2019. Feedback network for image super-resolution//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 3862-3871 [DOI: 10.1109/CVPR.2019.00399]
- Liang J Y, Cao J Z, Sun G L, Zhang K, Van Gool L and Timofte R. 2021. SwinIR: image restoration using Swin transformer//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Montreal, Canada: IEEE: 1833-1844 [DOI: 10.1109/ICCVW54120.2021.00210]
- Lim B, Son S, Kim H, Nah S and Lee K M. 2017. Enhanced deep residual networks for single image super-resolution//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, USA: IEEE: 1132-1140 [DOI: 10.1109/CVPRW.2017.151]
- Lu Z S, Li J C, Liu H, Huang C Y, Zhang L L and Zeng T Y. 2022. Transformer for single image super-resolution//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. New Orleans, USA: IEEE: 456-465 [DOI: 10.1109/CVPRW56347.2022.00061]
- Lyu J, Wang S, Tian Y P, Zou J, Dong S J, Wang C Y, et al. 2024. STADNet: spatial-temporal attention-guided dual-path network for cardiac cine MRI super-resolution. *Medical Image Analysis*, 94: #103142 [DOI: 10.1016/j.media.2024.103142]
- Martin D, Fowlkes C, Tal D and Malik J. 2001. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics//Proceedings of the 8th IEEE International Conference on Computer Vision. Vancouver, Canada: IEEE: 416-423 [DOI: 10.1109/ICCV.2001.937655]
- Matsui Y, Ito K, Aramaki Y, Fujimoto A, Ogawa T, Yamasaki T, et al. 2017. Sketch-based manga retrieval using manga109 dataset. *Multimedia Tools and Applications*, 76(20): 21811-21838 [DOI: 10.1007/s11042-016-4020-z]
- Park S C, Park M K and Kang M G. 2003. Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, 20(3): 21-36 [DOI: 10.1109/MSP.2003.1203207]
- Qiu D F, Jiang J J, Hu X Y, Liu X M and Ma J Y. 2023. Guided Transformer for high-resolution visible image guided infrared image super-resolution. *Journal of Image and Graphics*, 28(1): 196-206 (邱德粉, 江俊君, 胡星宇, 刘贤明, 马佳义. 2023. 高分辨率可见光图像引导红外图像超分辨率的Transformer网络. *中国图象*

- 图形学报, 28(1): 196-206 [DOI: 10.11834/jig.220604]
- Sun L, Dong J X, Tang J H and Pan J S. 2023. Spatially-adaptive feature modulation for efficient image super-resolution//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 13144-13153 [DOI: 10.1109/ICCV51070.2023.01213]
- Sun L, Pan J S and Tang J H. 2022. ShuffleMixer: an efficient ConvNet for image super-resolution//Proceedings of the 36th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Inc.: #1259
- Tai Y, Yang J, Liu X M and Xu C Y. 2017. MemNet: a persistent memory network for image restoration//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 4549-4557 [DOI: 10.1109/ICCV.2017.486]
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, et al. 2017. Attention is all you need//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates Inc.: 6000-6010
- Wan C, Yu H Y, Li Z Q, Chen Y H, Zou Y J and Liu Y Q. 2024. Swift parameter-free attention network for efficient super-resolution//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle, USA: IEEE: 6246-6256 [DOI: 10.1109/CVPRW63382.2024.00628]
- Wang C Y, Jiang J J, Jiang K and Liu X M. 2024. Low-light face super-resolution via illumination, structure, and texture associated representation//Proceedings of the 38th AAAI Conference on Artificial Intelligence. Vancouver, Canada: AAAI Press: 5318-5326 [DOI: 10.1609/aaai.v38i6.28339]
- Wang L G, Dong X Y, Wang Y Q, Ying X Y, Lin Z P and An W. 2021. Exploring sparsity in image super-resolution for efficient inference//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 4915-4924 [DOI: 10.1109/CVPR46437.2021.00488]
- Wang Z, Bovik A C, Sheikh H R and Simoncelli E P. 2004. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing, 13 (4): 600-612 [DOI: 10.1109/TIP.2003.819861]
- Wu C, Wang L, Su X and Zheng Z R. 2025. Adaptive feature selection modulation network for efficient image super-resolution. IEEE Signal Processing Letters, 32, 1231-1235 [DOI: 10.1109/LSP.2025.3547669]
- Yan X Y, Chen J C, Xu Q Z and Li W. 2025. Mitigating texture bias: a remote sensing super-resolution method focusing on high-frequency texture reconstruction. IEEE Transactions on Geoscience and Remote Sensing, 63: 1-18 [DOI: 10.1109/TGRS.2025.3547903]
- Yang F Z, Yang H, Fu J L, Lu H T and Guo B N. 2020. Learning texture transformer network for image super-resolution//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 5790-5799 [DOI: 10.1109/CVPR42600.2020.00583]
- Yang S Y, Wang M, Chen Y G and Sun Y X. 2012. Single-image super-resolution reconstruction via learned geometric dictionaries and clustered sparse coding. IEEE Transactions on Image Processing, 21(9): 4016-4028 [DOI: 10.1109/TIP.2012.2201491]
- Zeyde R, Elad M and Protter M. 2012. On single image scale-up using sparse-representations//Proceedings of the 7th International Conference on Curves and Surfaces. Avignon, France: Springer: 711-730 [DOI: 10.1007/978-3-642-27413-8_47]
- Zhang T, Wang B F, Fu Y, Liu S R, Ye J C, Shan P H, et al. 2024. Deep learning-based spectral image super-resolution: a survey. Journal of Image and Graphics, 29(8): 2113-2136 (张涛, 王彬洋, 付莹, 刘松荣, 叶吉超, 单培红, 等. 2024. 基于深度学习的光谱图像超分辨率综述. 中国图象图形学报, 29(8): 2113-2136 [DOI: 10.11834/jig.230747])
- Zhang Y L, Li K P, Li K, Wang L C, Zhong B N and Fu Y. 2018. Image super-resolution using very deep residual channel attention networks//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 294-310 [DOI: 10.1007/978-3-030-01234-2_18]

作者简介

刘紫阳,男,博士研究生,主要研究方向为超分辨率、遥感图像处理与机器学习。E-mail: 2320090166@tiangong.edu.cn

黄淑英,通信作者,女,教授,主要研究方向为遥感图像处理、计算机视觉与机器学习。

E-mail: huangshuying@tiangong.edu.cn

杨勇,男,教授,博士生导师,主要研究方向为遥感图像处理、计算机视觉与机器学习。E-mail: yangyong@tiangong.edu.cn

王书昭,男,博士研究生,主要研究方向为图像处理与机器学习。E-mail: 2320080147@tiangong.edu.cn