

文章编号 1004-924X(2024)06-0843-14

基于多尺度自适应注意力的图像超分辨率网络

周 颖^{1,2*}, 裴盛虎¹, 陈海永^{1,2}, 许士博¹

(1. 河北工业大学 人工智能与数据科学学院, 天津 300130;

2. 河北省控制工程技术研究中心, 天津 300130)

摘要:针对大多数图像超分辨率重建方法利用单尺度卷积进行特征提取,导致特征提取不充分的问题,提出基于多尺度自适应注意力的图像超分辨率网络。为充分利用各个层次特征中的上下文信息,设计了多尺度特征融合块,其基本单元由自适应双尺度块、多路径渐进式交互块和自适应双维度注意力依次串联组成。首先,自适应双尺度块自主融合两个尺度的特征,获得了更丰富的上下文特征;其次,多路径渐进式交互块以渐进的方式交互自适应双尺度块的输出特征,提高了上下文特征之间的关联性;最后,自适应双维度注意力自主选择不同维度注意力细化输出特征后,提高了输出特征的鉴别力。实验结果表明,在 Set5, Set14, BSD100 和 Urban100 测试集上,本文方法在 PSNR 和 SSIM 定量指标上相比于其他主流方法相均有提升,尤其对于纹理细节难以恢复的 Urban100 测试集,本文方法在比例因子为 $\times 4$ 时,相较于现有最优方法 SwinIR, PSNR 和 SSIM 指标分别提升了 0.05 dB 和 0.004 5;在视觉效果方面,本文的重建图像具有更多的纹理细节。

关键词:超分辨率;多尺度特征;注意力机制;自适应权重;渐进式信息交互

中图分类号: TP391.4 **文献标识码:** A **doi:** 10.37188/OPE.20243206.0843

Image super-resolution network based on multi-scale adaptive attention

ZHOU Ying^{1,2*}, PEI Shenghu¹, CHEN Haiyong^{1,2}, XU Shibo¹

(1. School of Artificial Intelligence, Hebei University of Technology, Tianjin 300130, China;

2. China Hebei Control Engineering Research Center, Tianjin 300130, China)

* Corresponding author, E-mail: zhouying2007@163.com

Abstract: Aiming at the problem that most image super-resolution methods cannot fully extract features by using single-scale convolution, an image super-resolution network based on multi-scale adaptive attention is proposed. To fully use the contextual information in each hierarchical feature, a multi-scale feature fusion block was designed, whose basic unit consists of an adaptive dual-scale block, a multi-path progressive interactive block, and an adaptive dual-dimensional attention sequentially in series. Firstly, the adaptive dual-scale block autonomously fused the features of two scales to obtain richer contextual features; secondly, the multi-path progressive interactive block interacted the output of the adaptive dual-scale block in a progressive way to improve the correlation between the contextual features; lastly, the adaptive dual-dimensional attention autonomously selected different dimensions of the attention to refine the output fea-

收稿日期: 2023-07-23; 修订日期: 2023-09-19.

基金项目: 国家自然科学基金资助项目(No.U21A20482, No.62073117); 河北省自然科学基金资助项目(No.F202202064)

tures, which makes the output features more discriminative. The experimental results show that on Set5, Set14, BSD100 and Urban100 test sets, the method of this paper improves the PSNR and SSIM quantitative metrics compared to other mainstream methods, especially for the Urban100 test set, where texture details are difficult to be recovered, the method of this paper improves PSNR and SSIM metrics by 0.05 dB and 0.004 5 respectively compared to the existing optimal method, SwinIR, with the scaling factor of $\times 4$; in terms of visual effect, the reconstructed images in this paper have more texture details.

Key words: super-resolution; multi-scale feature; attention mechanism; adaptive weights; progressive information interaction

1 引言

单幅图像超分辨率(Single Image Super-Resolution, SISR)是图像处理中的关键技术之一,其目的是通过低分辨率(Low-Resolution, LR)图像重建高分辨率(High-Resolution, HR)图像,在视频监控^[1]、卫星遥感^[2]和医学成像^[3]等不同领域有强烈的需求。传统 SISR 方法主要有基于插值的方法、基于重建的方法和基于学习的方法,但这些方法重建的超分辨率图像质量较低。随着深度学习的到来,卷积神经网络(Convolutional Neural Networks, CNNs)因其显著的特征表达能力受到了越来越多的关注。卷积超分辨率网络(Super-Resolution Convolutional Neural Network, SRCNN)^[4]首次将深度学习应用于 SISR 任务,促进了大量基于 CNNs 的 SISR 研究的涌现,但因其输入为 LR 图像上采样后的图像导致网络仍具有较高的计算复杂度。为实现更快的运行速度,Shi 等^[5]提出了一种快速卷积超分辨率网络,通过在 LR 图像空间中提取特征来减少计算量,并使用反卷积作为网络最后的上采样操作。然而,浅层的网络由于存在严重的梯度消失问题,早期基于 CNNs 的 SISR 方法的性能无法得到进一步提升。

深度卷积神经网络因其高级特征中包含更多的信息量进一步提高了 SISR 任务的重建性能。Kim 等^[6]提出深度卷积超分辨率网络(Super-Resolution Using Very Deep Convolutional Network, VDSR),通过在网络中引入残差学习,解决了梯度消失问题,相较于早期的浅层网络其性能得到了显著提升;Tai 等^[7]提出深度递归残差网络(Deep Recursive Residual Network, DRRN),利用全局跳跃连接和局部残差学习减

少了参数量,并表现出良好的性能;Lim 等^[8]提出增强型深度超分辨率网络(Enhanced Deep Super-Resolution Network, EDSR),通过去除网络中的批量归一化(Batch Normalization, BN)层,进一步提高了模型的性能。但上述方法均使用单尺度卷积进行特征提取,难免会造成重要信息的丢失,程德强等^[9]将多尺度特征融合并采用密集连接以恢复图像的高频细节;Cai 等^[10]将多感受野分支的特征进行通道拼接以学习局部区域特征;Liu 等^[11]采用递归多尺度网络将不同尺度分支的输出作为其他尺度分支的输入;许娇等^[12]将多尺度特征与残差网络结合以实现图像的重建过程。

随着网络深度与宽度的增加,特征图的维数也在增加,注意力机制的引入提高了网络处理更多特征信息的能力。Niu 等^[13]在残差块中同时加入了层注意力与通道空间注意力,大大提高了 SISR 性能;王杰等^[14]将空间注意力与空间金字塔网络相结合以增强重建图像过程中的特征表达能力;Su 等^[15]提出全局可学习注意力以修复重建图像中受损的非局部纹理细节;Behjati 等^[16]引入方向方差注意力来提高 LR 特征在长距离空间中的依赖性以提高特征的鉴别性。然而,上述注意力忽略了通道与空间维度信息的相关性和跨维度注意力之间的交互作用而难以鉴别重要特征。

本文在上述研究的基础上提出了一种基于多尺度自适应注意力的图像超分辨率重建网络(Imageresolution network based on multi-scale adaptive attention, MAAN)。MAAN 使用多尺度特征融合块(Multi-scale Feature Fusion Block, MFFB)作为特征提取的基本层,交叉学习特征提取中的多尺度信息,并将每个 MFFB 的

输出进行密集连接以充分融合浅层特征与深层特征。MFFB的基本单元(Basic Unit, BU)由自适应双尺度块(Adaptive Dual-scale Block, ADB)、多路径渐进式交互块(Multi-path Progressive Interactive Block, MPIB)和自适应双维度注意力(Adaptive Dual-dimensional Attention, ADA)串联组成。其中,ADB获得了丰富的上下文特征;MPIB提高了ADB的输出中上下文特征之间的关联性;ADA将一维注意力与二维注意力相结合,并自主选择跨维度之间的信息进行交互,进一步提高了特征的鉴别力。

2 网络模型

MAAN结构如图1所示,在浅层特征提取部分,给定输入LR图像 $I^{LR} \in \mathbb{R}^{3 \times H \times W}$,使用 3×3 卷积提取 I^{LR} 中的浅层特征 $F_0 \in \mathbb{R}^{C \times H \times W}$:

$$F_0 = \varphi(C_{3 \times 3}(I^{LR})), \quad (1)$$

其中: $C_{3 \times 3}(\cdot)$ 为 3×3 卷积操作, $\varphi(\cdot)$ 表示参数整流线性单元(Parametric Rectified Linear Unit,

PReLU)激活函数。

在深层多尺度特征提取与融合部分,将 F_0 输入到多个堆叠的MFFB中,使输出特征更具上下文信息:

$$F_8 = M_8(F_7) = M_8(M_7(\cdots F_0)), \quad (2)$$

其中: $M_i(\cdot)$ 和 $F_i \in \mathbb{R}^{C \times H \times W}$ 分别代表第 $i \in [1, 8]$ 个MFFB和第 i 个MFFB的输出特征。为提高输出特征中上下文信息的关联性,将每个 F_i 在通道维度上进行级联并通过 1×1 卷积处理获得不同层次之间的融合特征:

$$F_{\text{fuse}} = C_{1 \times 1}([F_1, F_2, \dots, F_8]), \quad (3)$$

式中: $C_{1 \times 1}(\cdot)$ 为 1×1 卷积操作, $[\cdot]$ 代表在通道维度上进行级联操作。利用全局残差学习将浅层特征 F_0 与深层特征 F_{fuse} 融合:

$$F = F_0 + F_{\text{fuse}}. \quad (4)$$

在高分辨率图像重建部分,利用上采样因子将 F 重建为高分辨率图像 $I^{SR} \in \mathbb{R}^{3 \times H \times W}$:

$$I^{SR} = C_{3 \times 3}(\text{ps}(C_{3 \times 3}(F))), \quad (5)$$

其中,ps表示pixel-shuffle上采样因子。

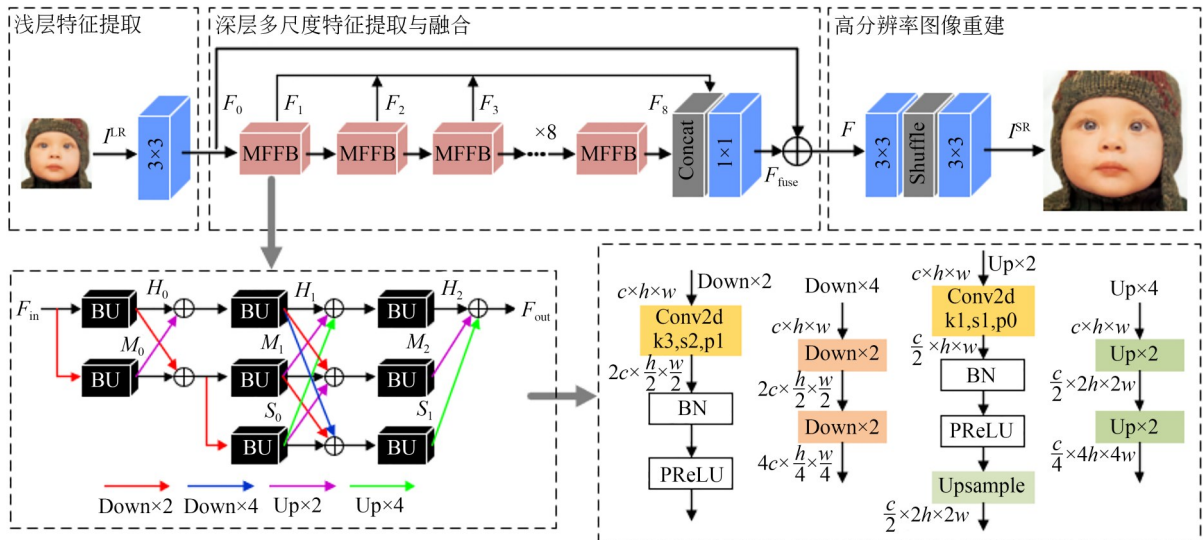


图1 MAAN网络架构

Fig. 1 Network architecture of MAAN

2.1 多尺度特征融合模块(MFFB)

在MAAN中,设计了MFFB作为深层特征提取的基本传递块,MFFB的结构如图1左下部分所示,第一行是LR特征流,第二行是1/2倍LR特征流,第三行是1/4倍LR特征流。3个特征流

通过上采样和下采样操作交叉学习LR空间不同尺度特征的信息。MFFB中2次交叉学习的使用充分提取了每个层次特征包含的上下文信息,将3个特征流分别定义为 $H_0, H_1, H_2, M_0, M_1, M_2$ 和 S_0, S_1 ,传递过程如下:

$$\begin{cases} H_0 = f_{BU}(F_{in}) \\ M_0 = f_{BU}(f_{down \times 2}(F_{in})) \end{cases}, \quad (6)$$

$$\begin{cases} H_1 = f_{BU}(H_0 + f_{up \times 2}(M_0)) \\ M_1 = f_{BU}(f_{down \times 2}(H_0) + M_0) \\ S_0 = f_{BU}(f_{down \times 2}(f_{down \times 2}(H_0) + M_0)) \end{cases}, \quad (7)$$

$$\begin{cases} H_2 = f_{BU}(H_1 + f_{up \times 2}(M_1) + f_{up \times 4}(S_0)) \\ M_2 = f_{BU}(f_{down \times 2}(H_1) + M_1 + f_{up \times 2}(S_0)) \\ S_1 = f_{BU}(f_{down \times 4}(H_1) + f_{down \times 2}(M_1) + S_0) \end{cases}, \quad (8)$$

其中: $f_{BU}(\bullet)$ 为 MFFB 的基本单元, $f_{down}(\bullet)$ 和 $f_{up}(\bullet)$ 分别为下采样和上采样操作, 具体操作和维度变化如图 1 右下部分所示。最后, 将交叉融合后的多尺度特征进行相加得到输出特征:

$$F_{out} = H_2 + f_{up \times 2}(M_2) + f_{up \times 4}(S_1). \quad (9)$$

MFFB 的基本单元 BU 如图 2 所示, BU 由自适应双尺度块 ADB、多路径渐进式交互块 MPIB 和自适应双维度注意力 ADA 串联组成, 利用 ADB 实现不同尺度的自适应融合, 将融合特征通过 MPIB 进一步加强后, 采用 ADA 找出应该强调突出的区域; 为加快模型的收敛速度, 引入了残差学习。

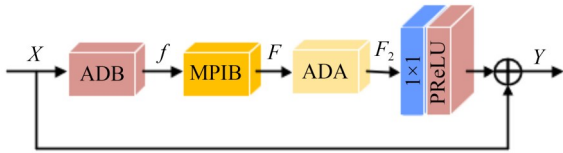


图 2 BU 结构图

Fig. 2 Architecture of BU

2.2 自适应双尺度块(ADB)

2.2.1 ADB 结构

现有的多尺度 SR 方法虽然在网络架构上有所不同, 但只是以线性的方式简单地叠加或拼接多个尺度的特征, 导致特征信息的冗余与模型参数数量的加大。为此, 本文设计了 ADB 以自适应地给两个尺度的特征分配权重, 实现双尺度特征的自适应融合。

ADB 结构如图 3 所示, 给定输入特征图 $X \in R^{C \times H \times W}$, 首先对 X 分别执行两次卷积核为 $3 \times 3, 5 \times 5$ 的卷积处理得到 2 个特征图, 获得 2 个不同尺度的特征 $X_1, X_2 \in R^{2C \times H \times W}$:

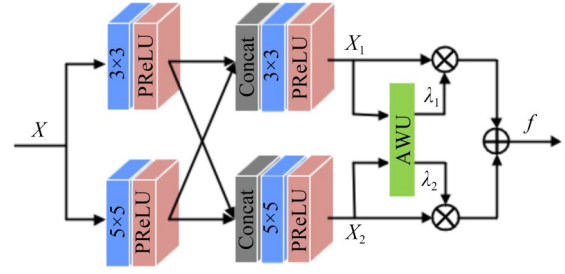


图 3 ADB 结构图

Fig. 3 Architecture of ADB

$$\begin{cases} X_1 = f_1([f_1(X), f_2(X)]) \\ X_2 = f_2([f_1(X), f_2(X)]) \end{cases}, \quad (10)$$

其中: $f_1(\bullet)$ 包含 3×3 卷积和 PReLU 激活函数, $f_2(\bullet)$ 包含 5×5 卷积和 PReLU 激活函数。

然后将 X_1 和 X_2 输入到自适应权重单元 (Adaptive Weights Unit, AWU) 获得各自的权重值 λ_1, λ_2 , 最后利用 λ_1, λ_2 实现双尺度特征的自适应融合:

$$f = \lambda_1 \cdot X_1 + \lambda_2 \cdot X_2, \lambda_1 + \lambda_2 = 1. \quad (11)$$

2.2.2 AWU 结构

AWU 采用加权映射策略, 为降低计算复杂度, 用全局平均池化充当降维算子, 将不同尺度的特征转换为具有全局感受野的统计量, 各统计量通过关键信息筛选与 Softmax 加权映射后获得不同尺度特征的权重。

AWU 结构如图 4 所示, 不同输入 $x_i, i \in [1, n]$, 利用全局平均池化压缩输入特征的空间维度信息后, 得到全局代表特征向量, 将每个特征向量在通道维度上进行拼接并经过 1×1 卷积后, 得到融合特征:

$$x = C_{1 \times 1}([GAP(x_1), GAP(x_2), \dots, GAP(x_n)]), \quad (12)$$

其中, $GAP(\bullet)$ 表示全局平均池化操作。

融合特征经过第 1 个全连接层 (Fully Connected Layer, FC)、整流线性单元 (Rectified Linear Unit, ReLU) 激活函数和第 2 个 FC 后, 得到由 n 个元素组成的初始向量, 该初始向量中元素值的变化范围较大, 因此将该向量经过 Sigmoid 激活函数处理后, 得到各元素值均位于 $[0, 1]$ 之间的简化向量 ϵ, ϵ_i 的大小映射了 Sigmoid 权重参数的大小, ϵ 经过 Softmax 处理后得到各输入特征

的最终权重 ω_i :

$$\epsilon = \sigma(W_2(\delta(W_1(x)))), \quad (13)$$

$$\omega_i = \frac{e^{\epsilon_i}}{e^{\epsilon_1} + e^{\epsilon_2} + \dots + e^{\epsilon_i} + \dots + e^{\epsilon_n}}, \sum_{i=1}^n \omega_i = 1, \quad (14)$$

其中: $\sigma(\cdot)$ 表示 Sigmoid 激活函数, $\delta(\cdot)$ 表示 ReLU 激活函数, $W_1(\cdot)$ 和 $W_2(\cdot)$ 表示两个 FC 层, ϵ_i 表示 ϵ 中的第 i 个元素。

2.3 多路径渐进式交互块(MPIB)

MPIB 将输入特征 $f \in R^{C \times H \times W}$ 在通道维度上均分为 4 组并输入到不同的路径进行处理, 与现有的分组卷积相比, MPIB 不仅减少了因所有通

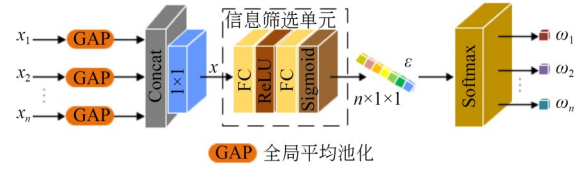


图4 AWU 结构图

Fig. 4 Architecture of AWU

道进行相同处理而造成的特征冗余, 而且增强了处理不同尺度特征的能力。此外, MPIB 将不同路径的输出特征以渐进的方式交互空间信息, 并通过通道混洗(Channel Shuffle, CS)交互通道信息, 具体细节如图 5 所示。

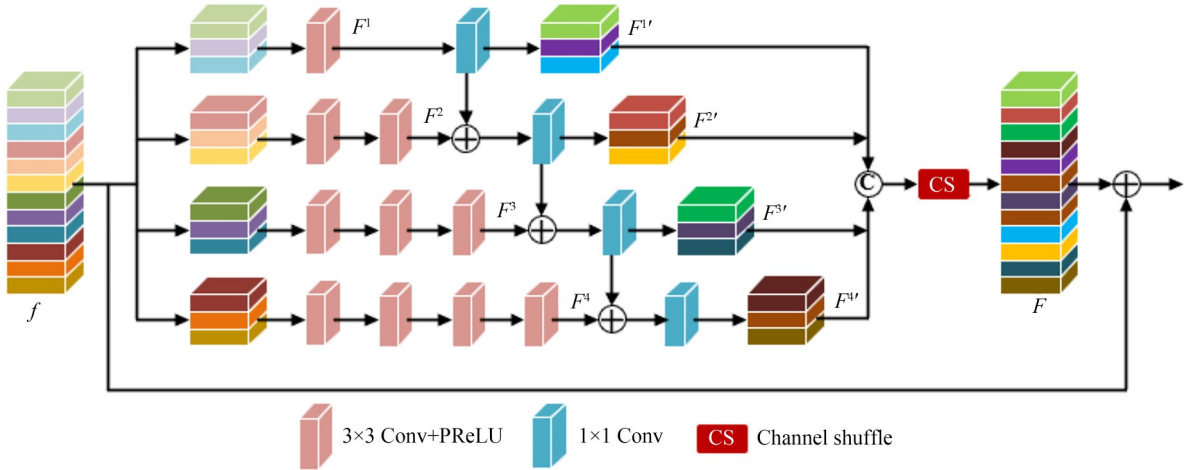


图5 MPIB 结构图

Fig. 5 Architecture of MPIB

MPIB 首先将 4 个子特征分别经过不同数量的 3×3 卷积(卷积数量越多, 感受野越大, 处理大尺度特征的能力越强)处理后得到 $\{F^i\}_{i=1}^4 \in R^{C/4 \times H \times W}$; 然后, F^1 仅进行 1×1 卷积处理, 对于 $i \geq 2$ 的特征 F^i 与相邻路径的特征 $F^{(i-1)'}$ 交互特征在不同通道间的空间信息后采用 1×1 卷积进一步整合特征并调整通道数, 该过程可以描述为:

$$\begin{cases} F^{1'} = C_{1 \times 1}(F^1), i = 1 \\ F^{i'} = C_{1 \times 1}(F^i + F^{(i-1)'}), i \geq 2 \end{cases} \quad (15)$$

最后, 为提高特征的表达能, 将不同路径下的输出特征在通道维度上进行拼接并进行 CS 处理, 交互不同路径的输出特征的通道信息, 进一步增强特征在通道之间的相互作用:

$$F = f + cs([F^{1'}, F^{2'}, F^{3'}, F^{4'}]), \quad (16)$$

其中, $cs(\cdot)$ 表示 Channel Shuffle 操作。

2.4 自适应双维度注意力(ADA)

ADA 通过串联一维注意力和二维注意力对不同维度的相互依赖性进行建模, 突出上下文特征中应该强调的区域。两个维度的注意力均由 3 条分支组成, 分别用于捕捉在 (H, W) , (C, W) 和 (H, C) 维度之间的跨维度信息; 并利用 AWU 实现跨维度信息的自主融合。ADA 通过利用通道维度和空间维度之间的相互依赖性, 可以有效地关注上下文特征。具体细节如图 6 所示。

在一维注意力中, 首先利用全局平均池化将输入处理后得到 A_c^1 , 同样由 (C, W) 和 (H, C) 维度分别得到 A_h^1 和 A_w^1 :

$$A_c^1, A_h^1, A_w^1 = \sigma(GAP(F)). \quad (17)$$

然后将 F 分别与 A_c^1, A_h^1 和 A_w^1 相乘得到细化

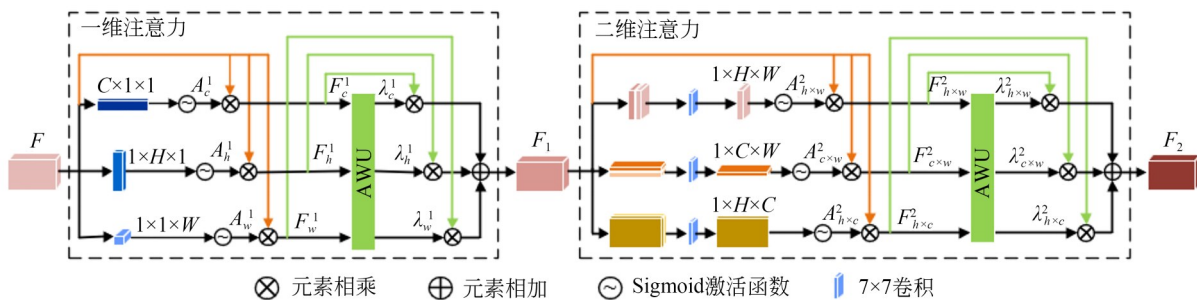


图6 ADA结构图

Fig. 6 Architecture of ADA

特征 F_c^1, F_h^1 和 F_w^1 :

$$F_c^1 = A_c^1 \times F, F_h^1 = A_h^1 \times F, F_w^1 = A_w^1 \times F. \quad (18)$$

最后将每个细化特征输入到 AWU, 获得相应的特征权重值 λ_c^1, λ_h^1 和 λ_w^1 , 利用权重值得到不同细化特征之间的自适应融合特征 F_1 :

$$F_1 = \lambda_c^1 \times F_c^1 + \lambda_h^1 \times F_h^1 + \lambda_w^1 \times F_w^1. \quad (19)$$

在二维注意力中, 同样以 (H, W) 维度为例, 首先对 F_1 在通道轴上分别使用平均值和最大值得到 2 个特征, 将其进行通道拼接并经过 7×7 卷积和 Sigmoid 后得到 $A_{h \times w}^2$, 以类似的方式可以获得 $A_{c \times w}^2$ 和 $A_{h \times c}^2$:

$$A_{h \times w}^2, A_{c \times w}^2, A_{h \times c}^2 = \sigma(C_{7 \times 7}([A(F_1), M(F_1)])), \quad (20)$$

式中: $A(\cdot)$ 和 $M(\cdot)$ 分别为平均值和最大值, $C_{7 \times 7}(\cdot)$ 为核为 7×7 卷积操作。

然后 F_1 分别与所有二维注意力相乘得到细化特征 $F_{h \times w}^2, F_{c \times w}^2$ 和 $F_{h \times c}^2$:

$$F_{h \times w}^2 = A_{h \times w}^2 \times F_1, F_{c \times w}^2 = A_{c \times w}^2 \times F_1, F_{h \times c}^2 = A_{h \times c}^2 \times F_1. \quad (21)$$

最后利用 AWU 获得二维注意力细化特征图的权重值以获得自主融合特征 F_2 :

$$F_2 = \lambda_{h \times w}^2 \times F_{h \times w}^2 + \lambda_{c \times w}^2 \times F_{c \times w}^2 + \lambda_{h \times c}^2 \times F_{h \times c}^2. \quad (22)$$

3 实验结果与分析

3.1 数据集与评价指标

为方便与最先进的方法进行公平对比, 本文采用 DIV2K^[17] 数据集集中的 800 张高质量的 HR 图像作为 MAAN 的训练数据集, LR 图像是对相应的 HR 图像进行双三次线性下采样得到的。采用 4 个常见的基准数据集: Set5^[18], Set14^[19],

BSD100^[20] 和 Urban100^[21] 作为测试集, 在比例因子分别为 $\times 2, \times 3, \times 4$ 上使用峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR)^[22] 和结构相似性指数 (Structure Similarity Index Measure, SSIM)^[23] 评价指标进行测试。

3.2 实施细节

每个 MFFB 中包含 8 个 BU, 模型中所有中间特征图的通道数均为 $C=64$ 。训练前, 通过随机角度旋转或翻转增加训练数据的多样性, 并随机裁剪 48×48 的补丁作为网络的输入 LR 图像; 训练时, 批尺寸大小设置为 16, 初始学习率为 $l_r = 10^{-4}$ 并在每 200 个迭代后减半, 采用自适应力矩估计 (Adaptive Moment Estimation, Adam)^[24] 优化器对网络进行 1000 个迭代训练, 其中 $\beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 10^{-8}$ 。本文网络在 PyTorch 框架上实现, 并在计算机系统配置: Windows10 64-bit, 12th Gen Intel Core i5-12490F CPU, NVIDIA RTX 3090 GPU 上利用均方误差 (Mean Squared Error, MSE) 损失进行训练。

3.3 消融实验

3.3.1 MFFB 数量消融实验

为实现模型参数量 (M: 百万) 与重建性能的均衡, 本文首先探索了 MFFB 的数量对整个网络的影响, 表 1 为 Set5 数据集在比例因子为 $\times 2$ 时网络中含有不同数量 MFFB 时 PSNR 的测试结果与模型参数数量的变化。可以看出, MFFB=7

表 1 MFFB 数量消融实验

Tab. 1 Ablation experiments of the numbers of MFFB

数量	6	7	8	9
参数量/M	7.36	8.52	9.61	10.83
PSNR	38.01	38.20	38.37	38.40

比 MFFB=6 的 PSNR 值增加了 0.19 dB,继续增加 MFFB 的数量, MFFB=8 比 MFFB=7 的 PSNR 值增加了 0.17 dB,参数量增加了 1.09 M,而 MFFB=9 比 MFFB=8 的 PSNR 值仅增加了 0.03 dB,但参数量却增加了 1.22 M。为在重建性能与模型复杂度之间取得平衡,选用 8 个 MFFB 作为最终网络构成。

3.3.2 BU 结构消融实验

本节通过 BU 结构的消融实验证明 BU 中 ADB,MPIB 和 ADA 三个模块的有效性,分别删除 BU 中的 ADB(结构 1)、MPIB(结构 2)和 ADA(结构 3),与完整的 BU 结构进行实验对比,由表 2 上半部分的测试结果可以看出,结构 1、结构 2 和结构 3 的 PSNR/SSIM 评价指标均低于完整的

BU 结构,证明了 BU 中包含的 ADB,MPIB 和 ADA 三个模块的有效性。

为证明 AWU 在 ADB 和 ADA 中均发挥了作用,进行了 AWU 的消融实验,由表 2 下半部分的测试结果可以看出,同时删除 ADA 和 ADB 中的 AWU(结构 4),PSNR/SSIM 评价指标在 4 个基准数据集上均为最低值,而仅删除 ADA 中的 AWU(结构 5)或仅删除 ADB 中的 AWU(结构 6)与结构 4 相比,PSNR/SSIM 评价指标均有所上升,证明 AWU 分别提升了 ADA 和 ADB 的性能;而 ADB 和 ADA 中均不删除 AWU(完整的 BU)与结构 5 或结构 6 相比,PSNR/SSIM 评价指标为最优值,说明在 ADB 和 ADA 中同时加入 AWU 进一步提高了模型的性能。

表 2 BU 结构消融实验结果
Tab. 2 Results of ablation experiments of the structure of BU

结构	Set5		Set14		BSD100		Urban100	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
结构 1	32.35	0.896 3	28.66	0.782 1	27.45	0.727 4	26.58	0.799 8
结构 2	32.39	0.895 9	28.78	0.787 6	27.42	0.729 4	26.64	0.803 3
结构 3	32.43	0.899 1	28.85	0.798 2	27.54	0.731 9	26.59	0.802 8
结构 4	32.69	0.900 2	29.03	0.810 5	27.74	0.745 1	26.92	0.816 7
结构 5	32.76	0.901 8	29.11	0.811 8	27.83	0.746 3	27.03	0.819 5
结构 6	32.77	0.902 0	29.12	0.812 0	27.85	0.746 9	27.08	0.819 4
完整的 BU	32.79	0.902 3	29.15	0.812 8	27.87	0.747 3	27.12	0.820 9

3.3.3 ADB 结构消融实验

本文通过消融实验证明 ADB 中各尺度分支的有效性,其中 3×3,5×5 和 7×7 分别表示 ADB 中只含有 3×3,5×5 或 7×7 卷积 1 条分支,3×3+5×5 表示含有 3×3 卷积和 5×5 卷积(空洞率为 2 的 3×3 卷积代替)2 条分支,3×3+5×5+7×7 表示含有 3 条分支(5×5 卷积由空洞率为 2 的 3×3 卷积代替,7×7 卷积由空洞率为 3 的 3×3 卷积代替)。

表 3 为 Set5 数据集在比例因子为 ×3 时

PSNR 值的测试结果,可以看出,与单尺度分支中性能最好的 7×7 相比,双尺度分支 3×3+5×5 在参数量仅增加了 0.97 M 的情况下,PSNR 值却增加了 0.38 dB;然而,三尺度分支 3×3+5×5+7×7 比 3×3+5×5 的参数量增加了 10.49 M,但 PSNR 值却只增加了 0.14 dB。综合考虑模型大小与实验性能,本文 ADB 选择使用双尺度分支 3×3+5×5,此外,在 3×3+5×5 的基础上加入 AWU,即 3×3+5×5+AWU,与 3×3+5×5 相比,参数量在略微增加 0.19 M 的

表 3 ADB 结构消融实验结果
Tab. 3 Results of ablation experiments of the structure of ADB

分支	3×3	5×5	7×7	3×3+5×5	3×3+5×5+7×7	3×3+5×5+AWU
参数量/M	3.21	5.84	8.53	9.50	19.99	9.69
PSNR	34.26	34.37	34.43	34.81	34.95	34.91

情况下,重建性能却可以达到与 $3 \times 3 + 5 \times 5 + 7 \times 7$ 相近的效果。

3.3.4 MPIB 结构消融实验

本文通过消融实验证明将 MPIB 的输入特征分为 4 组是平衡模型复杂度与重建性能的最佳选择。考虑到中间特征的通道数均为 $C=64$, 为确保分组后各子特征的通道数一致, 将分组数量分别设置为 2, 4, 8, 表 4 为 Set5 数据集在比例因子为 $\times 2$ 时, 对于 MPIB 包含不同分组数量的 PSNR 的测试结果与模型参数数量的变化。可以看出, 分组数为 4 相较于分组数为 2, 在参数量仅增加 1.82 M 的情况下, PSNR 值增加了 0.55 dB; 而分组数为 8 相较于分组数为 4, 参数量增加了 4.94 M, 但 PSNR 值仅增加了 0.16 dB。综合考虑模型的复杂度与重建性能, 将 MPIB 的输入特征分为 4 组作为最佳选择。

表 4 MPIB 分组数量消融实验

Tab. 4 Ablation experiments of the groups of MPIB

组数	2	4	8
参数量/M	7.79	9.61	14.55
PSNR	37.82	38.37	38.53

3.4 对比实验

3.4.1 不同注意力对比实验

为证明本文提出的 ADA 的有效性, 用高效

通道注意力 (Efficient Channel Attention, ECA)^[25], 卷积块注意力模块 (Convolutional Block Attention Module, CBAM)^[26] 和平衡注意力机制 (Balanced Attention Mechanism, BAM)^[27] 替换 MAAN 中的 ADA 进行对比。

表 5 为 4 个基准数据集在比例因子为 $\times 3$ 时的实验结果, 可以看出, 在没有注意力的基本网络中加入通道注意力 (+ECA), 参数量增加了 0.02 M, 4 个基准数据集的 PSNR/SSIM 评价指标增加的均值为 0.10 dB/0.000 8; +CBAM 和 +BAM 与基础网络相比, 虽参数量分别增加了 0.08 M 和 0.13 M, 但 PSNR/SSIM 评价指标的均值分别增加了 0.16 dB/0.001 3 和 0.17 dB/0.001 5, 说明 CBAM 和 BAM 中通道注意力与空间注意力结合的有效性; +ADA 与基础网络相比, 参数量仅增加了 0.04 M, 但 PSNR/SSIM 评价指标的均值却增加了 0.32 dB/0.003 3, 证明了 ADA 通过结合一维注意力与二维注意力和自适应地交互跨维度信息, 不仅比通道-空间注意力具有更少的参数量, 而且表现出了更好的性能, 使网络关注更多的盲区信息。图 7 展示了网络中加入不同注意力时的视觉效果比较, 其中参考图像 (Ground Truth, GT) 为真实高分辨率图像中红色框内区域的放大表示, 可以看出, 本文提出的 ADA 获得了最好的视觉效果, 重建的高分辨率图像具有更多的纹理细节。

表 5 不同注意力的评价指标对比

Tab. 5 Comparison of the evaluation metrics of different attention

方法	参数量/M	Set5		Set14		BSD100		Urban100	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
基础网络	9.65	34.66	0.928 8	30.52	0.847 4	29.12	0.809 7	28.91	0.872 3
+ECA	9.67	34.72	0.929 8	30.67	0.848 3	29.21	0.811 0	29.01	0.872 4
+CBAM	9.73	34.78	0.930 1	30.70	0.848 7	29.26	0.811 4	29.09	0.873 1
+BAM	9.78	34.76	0.929 9	30.72	0.849 3	29.22	0.811 5	29.18	0.873 6
+ADA	9.69	34.93	0.931 7	30.80	0.850 9	29.40	0.813 6	29.37	0.875 3

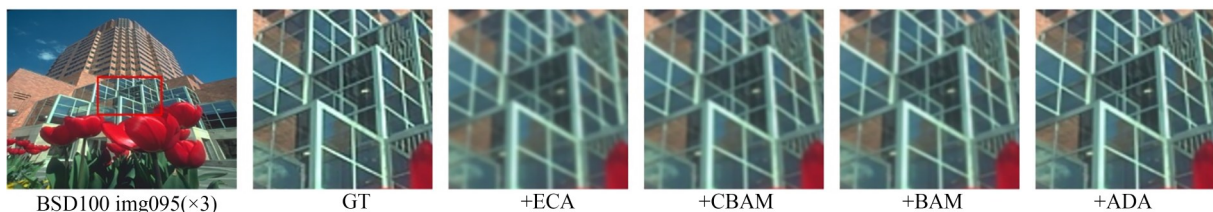


图 7 不同注意力的视觉效果对比

Fig. 7 Comparison of the visual effect of different attention

3.4.2 不同算法对比实验

本文在 Set5, Set14, BSD100 和 Urban100 数据集上将 MAAN 与 10 个代表方法进行比较,包括 Bicubic, SRCNN^[4], VDSR^[6], DRRN^[7]、注意力辅助特征学习网络 (Attentive Auxiliary Feature Learning, A2F-L)^[28], EDSR^[8]、残差密集网络 (Residual Dense Network, RDN)^[29]、高效分组跳跃网络 (Efficient Group Skip Connecting Network, E-GSCN)^[30]、密集残差拉普拉斯网络 (Densely Residual Laplacian Network, DRLN)^[31]、Swin 重建网络 (Image Restoration Using Swin Transformer, SwinIR)^[32] 和联合三边特征滤波网络 (Single Image Super-Resolution Based on Joint Trilateral Feature Filtering, JTF-SISR)^[33]。

表 6 为各方法在 4 个基准数据集上比例因子分别为 $\times 2, \times 3, \times 4$ 的定量评估结果,其中最优值和次最优值分别用粗体和下划线标记。可以看出,本文算法 MAAN 在 3 个不同放大比例任务上均取得了较好的定量数据。在 Set5 数据集的 3 个放大比例任务中,MAAN 相较于次最优算法 SwinIR, PSNR/SSIM 评价指标分别增加了

0.02 dB/0.000 6, 0.04 dB/0.000 5 和 0.04 dB/0.000 2;对于 Set14 数据集,在 $\times 2$ 放大比例任务中,MAAN 比次最优算法 DRLN 的 PSNR 值减少了 0.02 dB,但 SSIM 值却增加了 0.000 7,在 $\times 3$ 和 $\times 4$ 比例任务中,MAAN 相较于次最优算法 SwinIR, PSNR/SSIM 评价指标分别增加了 0.03 dB/0.000 6 和 0.21 dB/0.021 4;对于包含更多复杂纹理细节的 BSD100 和 Urban100 数据集,本文算法 MAAN 在 3 个放大比例任务中的 PSNR 值与 SSIM 值均为最优值,尤其是在 $\times 4$ 放大比例任务中,MAAN 比次最优算法 SwinIR 的 PSNR/SSIM 评价指标分别增加了 0.04 dB/0.001 4 和 0.05 dB/0.004 5,体现了本文方法在恢复高频纹理细节方面的优势。定量数据结果表明,本文算法与不同比例任务中的次最优算法相比,不仅减少了模型的参数量,而且进一步提升了模型的重建性能。

为进一步说明 MAAN 在恢复纹理细节方面的优势,本文将 MAAN 和 Bicubic, SRCNN, VDSR, RDN, DRLN 和 SwinIR 6 个代表性方法在较小数据集 Set14、较大数据集 BSD100 和 Ur-

表 6 比例因子为 $\times 2, \times 3, \times 4$ 时各方法的评价指标对比
Tab. 6 Comparison of evaluation metrics of each method with scale factors of $\times 2, \times 3, \times 4$

比例	方法	参数 量/M	Set5		Set14		BSD100		Urban100	
			PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
$\times 2$	Bicubic	—	33.69	0.928 4	30.34	0.867 5	29.57	0.843 4	26.88	0.843 8
	SRCNN	0.02	36.66	0.954 2	32.42	0.906 3	31.36	0.887 9	29.50	0.894 6
	VDSR	0.67	37.53	0.958 7	33.03	0.912 4	31.90	0.896 0	30.76	0.914 0
	DRRN	0.30	37.74	0.959 1	33.23	0.913 6	32.05	0.897 3	31.23	0.918 8
	A2F-L	1.36	38.09	0.960 7	33.78	0.919 2	32.23	0.900 2	32.46	0.931 3
	JTF-SISR	17.8	38.24	0.961 3	33.83	0.919 6	32.34	0.901 7	32.81	0.934 9
	EDSR	40.7	38.11	0.960 2	33.92	0.919 5	32.32	0.901 3	32.93	0.935 1
	RDN	22.1	38.24	0.961 4	34.01	0.921 2	32.34	0.901 7	32.89	0.935 3
	E-GSCN	15.7	38.21	0.961 2	33.95	0.920 3	32.35	0.901 9	32.94	0.935 7
	DRLN	32.4	38.27	0.961 6	34.28	<u>0.923 1</u>	<u>32.44</u>	0.902 8	33.37	0.939 0
	SwinIR	11.8	<u>38.35</u>	<u>0.962 0</u>	34.14	0.922 7	<u>32.44</u>	<u>0.903 0</u>	<u>33.40</u>	<u>0.939 3</u>
	MAAN	9.61	38.37	0.962 6	<u>34.26</u>	0.923 8	32.49	0.903 2	33.43	0.940 1
$\times 3$	Bicubic	—	30.41	0.865 5	27.64	0.772 2	27.21	0.734 4	24.46	0.741 1
	SRCNN	0.02	32.75	0.909 0	29.28	0.820 9	28.41	0.786 3	26.24	0.798 9
	VDSR	0.67	33.66	0.921 3	29.77	0.831 4	28.82	0.797 6	27.14	0.827 9
	DRRN	0.30	34.03	0.924 4	29.96	0.834 9	28.95	0.800 4	27.53	0.837 8
	A2F-L	1.37	34.54	0.928 3	30.41	0.843 6	29.14	0.806 2	28.40	0.857 4
	JTF-SISR	18.1	34.72	0.929 5	30.57	0.846 5	29.25	0.809 0	28.74	0.864 5

续表 6 比例因子为 $\times 2$, $\times 3$, $\times 4$ 时各方法的评价指标对比Tab. 6 Comparison of evaluation metrics of each method with scale factors of $\times 2$, $\times 3$, $\times 4$

比例	方法	参数 量/M	Set5		Set14		BSD100		Urban100	
			PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
$\times 3$	EDSR	43.7	34.65	0.928 0	30.52	0.846 2	29.25	0.809 3	28.80	0.865 3
	RDN	22.3	34.71	0.929 6	30.57	0.846 8	29.26	0.809 3	28.80	0.865 3
	E-GSCN	15.9	34.69	0.929 4	30.59	0.847 1	29.28	0.809 9	28.87	0.866 9
	DRLN	33.7	34.78	0.930 3	30.73	0.848 8	29.36	0.811 7	29.21	0.872 2
	SwinIR	11.9	<u>34.89</u>	<u>0.931 2</u>	<u>30.77</u>	<u>0.850 3</u>	<u>29.37</u>	<u>0.812 4</u>	<u>29.29</u>	<u>0.874 4</u>
	MAAN	9.69	34.93	0.931 7	30.80	0.850 9	29.40	0.813 6	29.37	0.875 3
$\times 4$	Bicubic	—	28.43	0.802 2	26.10	0.693 6	25.97	0.651 7	23.14	0.659 9
	SRCNN	0.02	30.48	0.862 8	27.49	0.750 3	26.90	0.710 1	24.52	0.722 1
	VDSR	0.67	31.35	0.883 8	28.01	0.767 4	27.29	0.725 1	25.18	0.752 4
	DRRN	0.30	31.68	0.888 8	28.21	0.772 0	27.38	0.728 4	25.44	0.763 8
	A2F-L	1.37	32.32	0.896 4	28.67	0.783 9	27.62	0.737 9	36.32	0.793 1
	JTF-SISR	18.5	32.53	0.898 9	28.80	0.786 7	27.70	0.741 0	26.62	0.802 2
	EDSR	43.1	32.46	0.896 8	28.80	0.787 6	27.71	0.742 0	26.64	0.803 3
	RDN	22.6	32.47	0.899 0	28.81	0.787 1	27.72	0.741 9	26.61	0.802 8
	E-GSCN	15.9	32.53	0.899 2	28.84	0.787 9	27.74	0.742 4	26.71	0.805 0
	DRLN	34.0	32.63	0.900 2	<u>28.94</u>	0.790 0	<u>27.83</u>	0.744 4	26.98	0.811 9
	SwinIR	11.9	<u>32.75</u>	<u>0.902 1</u>	<u>28.94</u>	<u>0.791 4</u>	<u>27.83</u>	<u>0.745 9</u>	<u>27.07</u>	<u>0.816 4</u>
	MAAN	9.72	32.79	0.902 3	29.15	0.812 8	27.87	0.747 3	27.12	0.820 9

ban100上的重建视觉效果进行比较。由于放大倍数越大,重建图像中的纹理细节越难恢复,因此为了进一步说明在重建纹理细节方面的优势,本文重点展示各方法在 $\times 4$ 比例因子上的视觉效果图。

如图 8 所示,对于较小数据集 Set14 的重建结果,可以看出,当 img014 中斑马前肢上有细小的条纹时,Bicubic, SRCNN 和 VDSR 等参数量较少的方法重建后的斑马前肢不仅丢失了条纹细节,而且具有严重的模糊伪影问题;RDN, DRLN 和 SwinIR 等参数量较大的方法,虽然较好地消除了重建图像中的模糊伪影,但仍存在缺失条纹细节的问题;而本文方法 MAAN 与上述方法相比,更好地缓解了伪影问题,重建了更多纹理细节。对于较大数据集 BSD100 和 Urban100 的重建结果,现有方法重建的 img096 中花池地板的模糊感较重,而本文方法重建的图像可以清晰的观察到地板中的纹路信息;现有方法在重建 img044 中天花板时,产生了泛白的背景

信息,而本文方法不仅还原了真实的背景,而且准确重建了天花板中的细小栅线;现有方法重建的 img076 中投影块的结构细节模糊不清,而本文算法准确重建了各个投影块的纹理细节与相邻投影块之间的边缘信息。视觉效果对比表明,本文 MAAN 不仅减轻了模糊伪影问题,而且使重建图像的内容信息更清晰,纹理细节和边缘轮廓更细腻。

3.4.3 模型复杂度对比实验

为全面衡量不同方法的重建性能,将本文方法 MAAN 与非轻量化模型 EDSR, RDN, E-GSCN, DRLN 和 SwinIR 在 BSD100 数据集上比例因子为 $\times 4$ 时的模型参数量、计算量和平均推理时间进行比较。由表 7 可知,与重建性能最优的 SwinIR 相比,MAAN 在保证重建性能的前提下,模型参数量、计算量和平均每张图像的处理时间分别减少了 18.3%, 38.3% 和 24.1%,在模型参数量、计算复杂度与重建性能之间实现了更好的权衡。

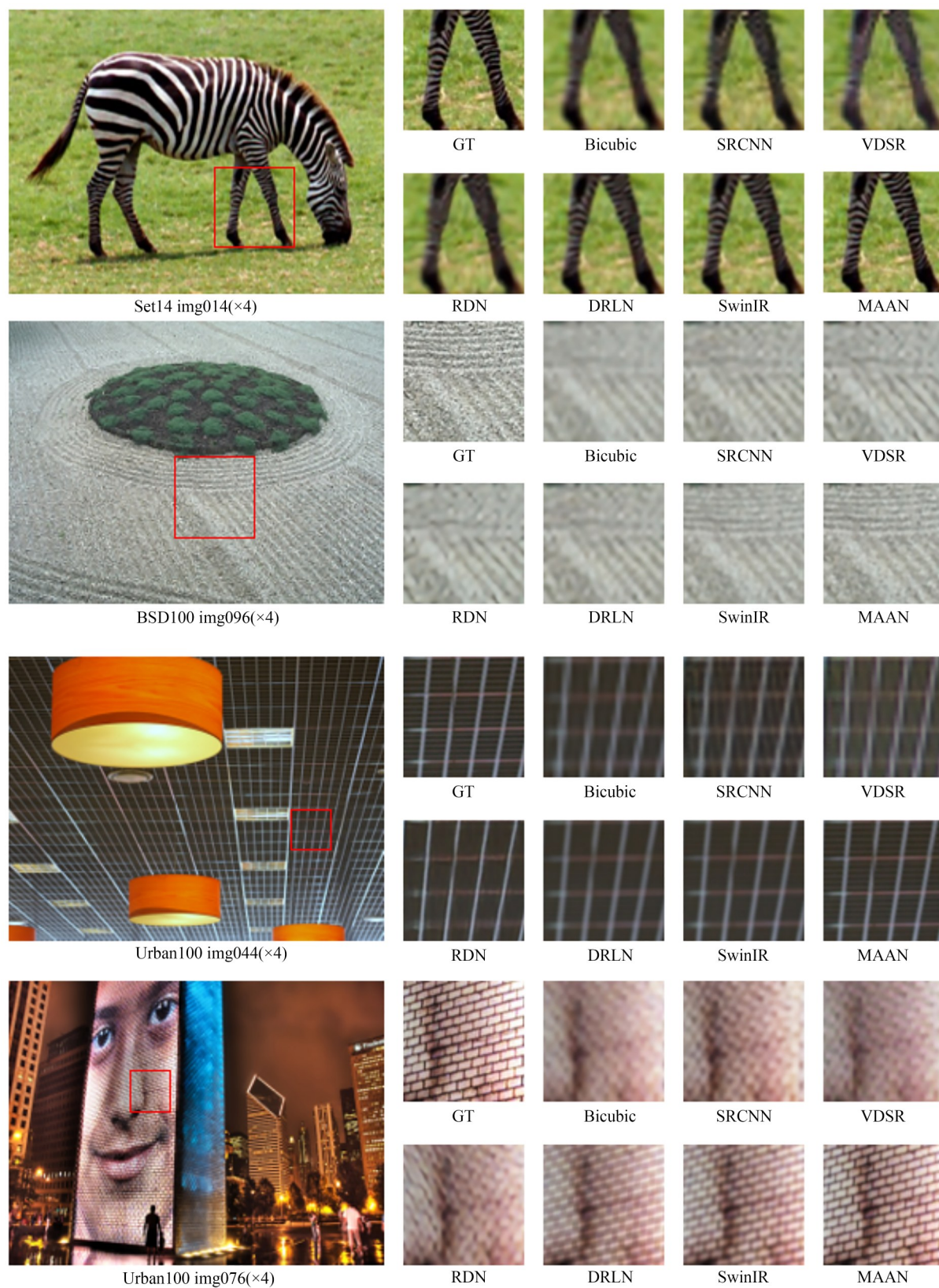
图 8 比例因子为 $\times 4$ 时各方法的视觉效果对比Fig. 8 Comparison of visual effect of each method with a scale factor of $\times 4$

表 7 3 种方法的复杂度对比

Tab. 7 Complexity comparison of the 3 methods

方法	PSNR	SSIM	时间 /s	参数 量/M	计算 量/G
EDSR	27.71	0.742 0	0.338	43.1	516
RDN	27.72	0.741 9	0.224	22.6	209
E-GSCN	27.74	0.742 4	0.136	15.9	144
DRLN	27.83	0.744 4	0.271	34.0	292
SwinIR	27.83	0.745 9	0.112	11.9	128
MAAN	27.87	0.747 3	0.085	9.72	79

4 结 论

为使重建图像中包含更多的上下文信息,本文提出了一种多尺度自适应渐进式注意力网络

MAAN。提出的多尺度特征融合块MFFB,利用由自适应双尺度块ADB、多路径渐进式交互块MPIB和自适应双维度注意力串联组成的基本单元BU交叉学习多尺度的层次特征。其中,ADB丰富了上下文特征,MPIB增强了上下文特征之间的关联性,ADA提高了特征的鉴别性。定量指标与视觉效果的对比如图所示,本文方法在不同比例的超分辨率重建任务中,与其他主流方法相比效果均有提升,尤其在Set14测试集的 $\times 4$ 比例任务上,本文方法与现有最优方法SwinIR相比,PSNR和SSIM指标分别提升了0.21 dB和0.0214;并且本文方法重建的图像具有更多的纹理细节与边缘信息。然而,MAAN与轻量化模型相比,仍具有较大的参数量,在未来的工作中,在确保重建性能的前提下,将进一步减少模型的复杂度。

参考文献:

- [1] HIJJI M, KHAN A, ALWAKEEL M M, *et al.* Intelligent image super-resolution for vehicle license plate in surveillance applications[J]. *Mathematics*, 2023, 11(4): 892.
- [2] CHEN S L, OGAWA Y, ZHAO C B, *et al.* Large-scale individual building extraction from open-source satellite imagery via super-resolution-based instance segmentation approach[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2023, 195: 129-152.
- [3] WU Q, LI Y W, SUN Y W, *et al.* An arbitrary scale super-resolution approach for 3D MR images via implicit neural representation[J]. *IEEE Journal of Biomedical and Health Informatics*, 2023, 27(2): 1004-1015.
- [4] 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.
- [5] SHI W Z, CABALLERO J, HUSZÁR F, *et al.* Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Washington, USA: IEEE, 2016: 1874-1883.
- [6] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Washington, USA: IEEE, 2016: 1646 - 1654.
- [7] TAI Y, YANG J, LIU X M. Image super-resolution via deep recursive residual network[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017. Honolulu, HI. IEEE, 2017: 2790 - 2798.
- [8] 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 (CVPRW)*. July 21-26, 2017. Honolulu, HI, USA. IEEE, 2017: 1132 - 1140.
- [9] 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.
CHENG D Q, ZHAO J M, KOU Q Q, *et al.* Multi-scale dense feature fusion network for image super-resolution[J]. *Opt. Precision Eng.*, 2022, 30(20): 2489-2500. (in Chinese)
- [10] CAI Q, LI J X, LI H F, *et al.* TDPN: texture and detail-preserving network for single image super-resolution[J]. *IEEE Transactions on Image Processing: a Publication of the IEEE Signal Pro-*

- cessing Society, 2022, 31: 2375-2389.
- [11] LIU F Q, YANG X M, DE BAETS B. A deep recursive multi-scale feature fusion network for image super-resolution [J]. *Journal of Visual Communication and Image Representation*, 2023, 90: 103730.
- [12] 许娇,袁三男. 增强型多尺度残差网络的图像超分辨率重建算法[J]. *激光与光电子学进展*, 2023, 60(4): 3788/LOP212884.
- XU J, YUAN S N. Image super-resolution reconstruction algorithm based on enhanced multi-scale residual network [J]. *Laser & Optoelectronics Progress*, 2023, 60(4): 3788/LOP212884. (in Chinese)
- [13] NIU B, WEN W L, REN W Q, et al. *Single Image Super-Resolution via a Holistic Attention Network*[M]. Computer Vision-ECCV 2020. Cham: Springer International Publishing, 2020: 191-207.
- [14] 王杰,徐国明,马健,等. 轻量级注意力级联网络的偏振计算成像超分辨率重建[J]. *光学精密工程*, 2022, 30(19): 2404-2419.
- WANG J, XU G M, MA J, et al. Polarization computational imaging super-resolution reconstruction with lightweight attention cascading network [J]. *Opt. Precision Eng.*, 2022, 30(19): 2404-2419. (in Chinese)
- [15] SU J N, GAN M, CHEN G Y, et al. Global learnable attention for single image super-resolution [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023: 1-12.
- [16] BEHJATI P, RODRIGUEZ P, FERNÁNDEZ C, et al. Single image super-resolution based on directional variance attention network [J]. *Pattern Recognition*, 2023, 133: 108997.
- [17] AGUSTSSON E, TIMOFTE R. NTIRE 2017 challenge on single image super-resolution: dataset and study[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. July 21-26, 2017. Honolulu, HI, USA. IEEE, 2017: 1122-1131.
- [18] 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. British Machine Vision Association, 2012:135.
- [19] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations [C]. *International Conference on Curves and Surfaces*. Berlin, Heidelberg: Springer, 2012: 711-730.
- [20] ARBELÁEZ P, MAIRE M, FOWLKES C, et al. Contour detection and hierarchical image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(5): 898-916.
- [21] 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 (CVPR)*. June 7-12, 2015. Boston, MA, USA. IEEE, 2015: 5197-5206.
- [22] WANG Y J, LI J H, LU Y, et al. Image quality evaluation based on image weighted separating block peak signal to noise ratio[C]. *International Conference on Neural Networks and Signal Processing*, 2003. *Proceedings of the 2003. December 14-17, 2003. Nanjing*. IEEE, 2003: 994-997.
- [23] 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.
- [24] KINGMA DP, BA J L. Adam: a method for stochastic optimization [J]. *arXiv preprint*, arXiv: 1412.6980, 2014.
- [25] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: efficient channel attention for deep convolutional neural networks[C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 13-19, 2020. Seattle, WA, USA. IEEE, 2020: 11531-11539.
- [26] WOO S H, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]. *Proceedings of the 15th European Conference on Computer Vision (ECCV)*. Berlin, Germany: Springer, 2018: 3-19.
- [27] WANG F Y, HU H T, SHEN C. BAM: A Lightweight and Efficient Balanced Attention Mechanism for Single Image Super Resolution[EB/OL]. 2021: *arXiv*: 2104.07566. <http://arxiv.org/abs/2104.07566>
- [28] WANG X H, WANG Q, ZHAO Y Z, et al. *Lightweight Single-Image Super-Resolution Network with Attentive Auxiliary Feature Learning*

- [M]. Computer Vision-ACCV 2020. Cham: Springer International Publishing, 2021: 268-285.
- [29] ZHANG Y L, TIAN Y P, KONG Y, *et al.* Residual dense network for image super-resolution [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 18-23, 2018. Salt Lake City, UT. IEEE, 2018: 2472 - 2481.
- [30] LIU C, LEI P C. An efficient group skip-connecting network for image super-resolution[J]. *Knowledge-Based Systems*, 2021, 222: 107017.
- [31] ANWAR S, BARNES N. Densely residual Laplacian super-resolution [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(3): 1192-1204.
- [32] 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 (ICCVW)*. October 11-17, 2021. Montreal, BC, Canada. IEEE, 2021: 1833 - 1844.
- [33] ZUO Y F, XIE J C, WANG H, *et al.* Gradient-guided single image super-resolution based on joint trilateral feature filtering [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023, 33(2): 505-520.

作者简介:



周 颖(1970—),女,天津人,博士,副教授,硕士生导师,主要从事图像处理 and 智能控制等方面的研究。E-mail: zhouying2007@163.com



裴盛虎(1998—),男,河北衡水人,硕士研究生,主要从事图像生成、图像超分辨率重建等方面的研究。E-mail: 1650954141@qq.com