

# 基于自适应双分支的图像超分辨率重建算法

张艳, 孙明磊, 孙叶美\*, 徐富杰

天津城建大学计算机与信息工程学院, 天津 300384

**摘要** 近年来,注意力机制广泛应用于图像超分辨率重建,极大地提高了重建网络的性能。为了更有效地利用注意力机制,提出一种基于自适应双分支的图像超分辨率重建算法。该算法设计的自适应双分支模块包括注意力分支和无注意力分支,通过自适应权重层动态平衡双分支的权重,舍弃冗余特征以使两个分支达到自适应平衡;其次,设计通道重组坐标注意力模块,通过通道重组的方式实现跨组特征交互,关注不同网络层特征之间的相关性;最后,设计双层残差聚合模块,构建两层嵌套的残差结构,提取残差块中的深层特征,更有效地提高网络的特征提取能力,提高重建图像的质量。在标准数据集上的大量实验验证了所提方法具有更好的重建效果。

**关键词** 图像处理; 超分辨率重建; 自适应权重; 通道重组; 残差聚合

中图分类号 TP391.41

文献标志码 A

DOI: 10.3788/LOP232007

## Image Super-Resolution Reconstruction Algorithm Based on Adaptive Two-Branch Block

Zhang Yan, Sun Minglei, Sun Yemei\*, Xu Fujie

School of Computer and Information Engineering, Tianjin Chengjian University, Tianjin 300384, China

**Abstract** Recently, attention mechanisms have been widely applied for image super-resolution reconstruction, substantially improving the reconstruction network's performance. To maximize the effectiveness of the attention mechanisms, this paper proposes an image super-resolution reconstruction algorithm based on an adaptive two-branch block. This adaptive two-branch block designed using the proposed algorithm includes attention and nonattention branches. An adaptive weight layer would dynamically balance the weights of these two branches while eliminating redundant attributes, thereby ensuring an adaptive balance between them. Subsequently, a channel shuffle coordinate attention block was designed to achieve a cross-group feature interaction to focus on the correlation between features across different network layers. Furthermore, a double-layer residual aggregation block was designed to enhance the feature extraction performance of the network and quality of the reconstructed image. Additionally, a double-layer nested residual structure was constructed for extracting deep features within the residual block. Extensive experiments on standard datasets show that the proposed method has a better reconstruction effect.

**Key words** image processing; super-resolution reconstruction; adaptive weight; channel shuffle; residual aggregation

## 1 引言

图像超分辨率(SR)旨在从退化的低分辨率(LR)图像中重建出高分辨率(HR)图像。目前其已成为图像处理领域的一个重要研究分支,在遥感图像<sup>[1]</sup>、光学图像<sup>[2-3]</sup>、目标检测<sup>[4]</sup>等领域有广泛应用。

近年来,基于深度学习的图像超分辨率重建方法取得了显著的成功。Dong等<sup>[5]</sup>首次将卷积神经网络

(CNN)应用于单幅图像超分辨率重建,提出SRCNN。但由于SRCNN结构简单,重建性能受到一定的限制, Kim等<sup>[6]</sup>提出深层超分辨率重建网络(VDSR),通过残差学习的方式增加网络的层数,解决了深度增加导致梯度消失现象的问题。随后,为了降低模型复杂度, Ahn等<sup>[7]</sup>引入级联残差网络(CARN),通过使用级联操作在残差网络上实现更高效的超分辨率重建。

注意力机制根据输入特征计算特征向量的不同权

收稿日期: 2023-08-29; 修回日期: 2023-10-04; 录用日期: 2023-10-30; 网络首发日期: 2023-11-17

基金项目: 天津市科技特派员项目(20YDTPJC01310)

通信作者: \*sunyemei1216@163.com

重,以表示特征的不同重要程度,已广泛应用于图像分割、目标检测等计算机视觉任务。Hu 等<sup>[8]</sup>提出了压缩和激发(SE)模块来学习通道之间的关系,根据通道所包含的特征信息为通道分配不同的权重。随后卷积块注意力模块(CBAM)<sup>[9]</sup>对通道注意力和空间注意力进行结合,实现自适应特征细化。Guo 等<sup>[10]</sup>提出大核卷积注意力(LKA)机制,利用大尺寸卷积核<sup>[11]</sup>更好地表示图像的全局结构特征并增大感受野,同时在图像超分辨率重建这类低级视觉任务中注意力机制也被广泛使用。Zhang 等<sup>[12]</sup>提出残差通道注意力模块,将通道注意力集成到残差模块中,充分利用特征之间的依赖性,取得了显著的性能提升。Zhao 等<sup>[13]</sup>提出了一种像素级注意力网络(PAN)来增强特征之间相关性的学习。在此基础上,Chen 等<sup>[14]</sup>提出了 A2N 网络,进一步研究将注意力机制用于图像超分辨率重建的方法,实验证明对于 SR 网络,使用注意力机制并不一定能够提升重建效果,盲目地使用注意力机制甚至会导致重建性能下降。而且目前基于注意力机制的方法主要考虑通道和空间位置的相关性,忽略了不同网络层特征之间的相互依赖性。

针对上述问题,本文提出了一种基于自适应双分支的图像超分辨率重建算法(ATSR),其更有效地利用特征信息,提高图像重建性能。提出自适应双分支模块(ATB),该模块包括注意力分支和无注意力分支。通过自适应权重层根据输入特征动态地调整两个分支的权重,有效增强有用特征和抑制冗余特征,从而更有效地利用注意力机制。其中注意力分支采用大核卷积注意力,结合自注意力和大尺寸卷积获得较大的感受野,并通过对大核卷积进行有效分解,以较小的计

算量实现图像全局特征的表示;而无注意力分支旨在学习被注意力分支忽略的特征信息,在保证降低模型参数量的同时,提高超分辨率重建的性能。设计通道重组坐标注意力模块(CSCAB),通过坐标注意力模块关注通道-空间两种维度的特征信息,提高特征的表达能力,并通过通道重组操作实现沿通道维度的跨组特征交互,充分关注各自适应双分支模块所产生的不同网络层特征之间的相关性,解决超分辨率重建图像的纹理趋于平滑的问题。设计双层残差聚合模块(DRAB),通过嵌套的残差结构,充分聚合残差分支上的特征信息,减小特征信息传播过程中的衰减,从而增加残差模块中深层特征的利用率,提升网络的特征提取性能,并引入 CSCAB 捕获跨通道信息和位置敏感信息,促进更具表现力特征的生成。

## 2 所提方法内容

所提基于自适应双分支的图像超分辨率重建网络包括自适应双分支模块、通道重组坐标注意力模块与双层残差聚合模块,整体网络结构如图 1 所示。首先,采用一个  $3 \times 3$  卷积层进行浅层特征提取,并将浅层特征信息送入自适应双分支模块,得到中间层特征  $F_i$ ;然后,利用通道重组坐标注意力模块对提取的所有中间层特征  $F_i$  进行通道重组,生成更具表现力的特征信息;同时,双层残差聚合模块通过双层嵌套的残差结构充分利用深层特征信息,进一步提高特征的利用率;最后,对通道重组坐标注意力模块和双层残差聚合模块的输出特征进行元素相加,再进行上采样,并通过跳跃连接与输入图像进行元素融合,最终重建出 SR 图像。

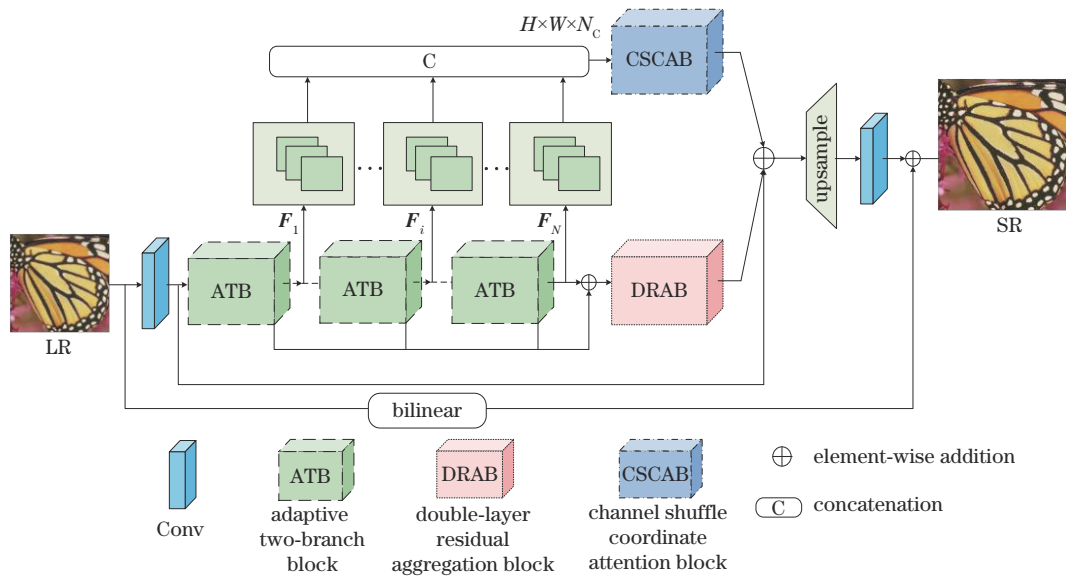


图 1 基于自适应双分支的图像超分辨率重建网络

Fig. 1 Network of image super-resolution reconstruction based on adaptive two-branch

### 2.1 自适应双分支模块

自适应双分支模块采用自适应权重层<sup>[15]</sup>动态分配

注意力分支和无注意力分支的权重,自适应地选择重要的特征信息,平衡注意力分支和无注意力分支,从而

更有效地利用注意力机制。自适应双分支模块结构如图 2 所示。设  $\mathbf{x}_n$  和  $\mathbf{x}_{n+1}$  分别为第  $n$  个自适应双分支模块的输入和输出,对于给定的输入特征  $\mathbf{x}_n$ ,通过  $1 \times 1$  卷积层得到双分支的输入特征:

$$\mathbf{x}'_n = \text{Conv1}(\mathbf{x}_n), \quad (1)$$

式中:Conv1 代表  $1 \times 1$  卷积; $\mathbf{x}'_n$  表示注意力分支和无注

意力分支的输入特征。

为更好地表示全局特征,注意力分支采用 LKA,充分利用自注意力和大尺寸卷积的优点对图像进行全局特征提取,更好地捕获上下文信息。为解决大尺寸卷积计算量大的问题,对大核卷积进行分解,分解过程如图 3 所示。

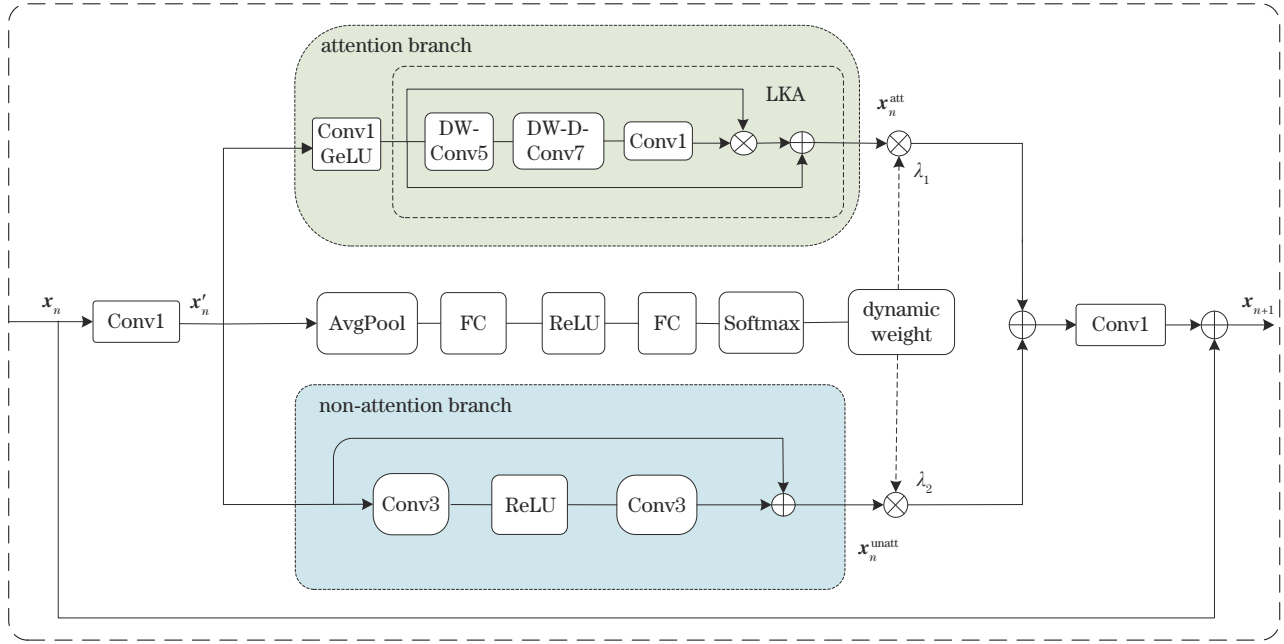


图 2 自适应双分支模块

Fig. 2 Adaptive two-branch block

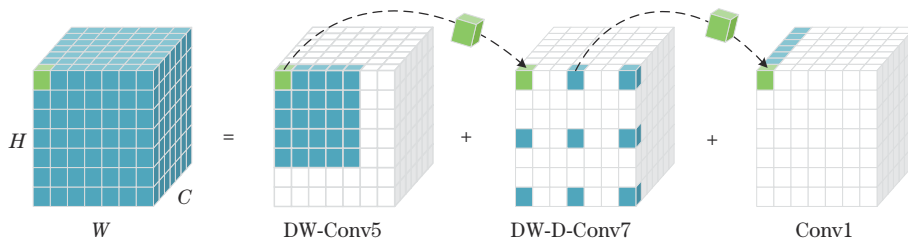


图 3 大核卷积分解过程

Fig. 3 Decomposition process of the large kernel convolution

LKA 将  $19 \times 19$  的大核卷积分解为 3 个部分,包括核大小为  $5 \times 5$  的空间局部卷积(DW-Conv5)、核大小为  $7 \times 7$ 、膨胀率  $d$  为 3 的空间全局卷积(DW-D-Conv7)和核大小为  $1 \times 1$  的通道卷积(Conv1),其中  $7 \times 7$  的空间全局卷积的感受野大小为 19。对于注意力分支,首

先使用  $1 \times 1$  卷积和 GeLU 激活函数校正图像特征,表达式为

$$\mathbf{x}_n^{\text{Conv1}} = \text{GeLU}[\text{Conv1}(\mathbf{x}'_n)], \quad (2)$$

然后,对  $\mathbf{x}_n^{\text{Conv1}}$  通过 LKA 进行全局特征提取,表达式为

$$\mathbf{x}_n^{\text{att}} = \text{Conv}\{\text{DW-D-Conv7}[\text{DW-Conv5}(\mathbf{x}_n^{\text{Conv1}})]\} \times \mathbf{x}_n^{\text{Conv1}} + \mathbf{x}_n^{\text{Conv1}}, \quad (3)$$

式中: $\mathbf{x}_n^{\text{att}}$  为注意力分支的输出特征。LKA 实现了通道-空间两种维度上的自适应,以较少的计算量获得了更丰富的全局特征信息。

无注意力分支采用经典的残差结构,通过残差连接和 2 个卷积层有效增强模型的特征表达能力,其表达式为

$$\mathbf{x}_n^{\text{unatt}} = \text{Conv3}\{\text{ReLU}[\text{Conv3}(\mathbf{x}'_n)]\} + \mathbf{x}'_n, \quad (4)$$

式中: $\mathbf{x}'_n$  和  $\mathbf{x}_n^{\text{unatt}}$  分别表示无注意力分支的输入特征和输出特征;Conv3 代表  $3 \times 3$  卷积;ReLU 表示激活函数。

注意力分支的输出  $\mathbf{x}_n^{\text{att}}$  和无注意力分支的输出  $\mathbf{x}_n^{\text{unatt}}$  分别乘以不同的动态权重  $\lambda_1$  和  $\lambda_2$  后,对应元素相加,最后传递到  $1 \times 1$  卷积,得到模块的输出特征  $\mathbf{x}_{n+1}$ ,

$$\mathbf{x}_{n+1} = \text{Conv1}(\lambda_1 \mathbf{x}_n^{\text{att}} + \lambda_2 \mathbf{x}_n^{\text{unatt}}) + \mathbf{x}_n. \quad (5)$$

## 2.2 通道重组坐标注意力模块

尽管密集连接和跳跃连接可以将浅层信息传递到深层网络中,但是没有充分利用不同网络层特征之间的相互依赖关系。针对此问题,设计了通道重组坐标注意力模块,对不同网络层的聚合特征图进行分组,并分别利用坐标注意力模块<sup>[16]</sup>对特征图进行深层特征提取,最后利用通道重组操作<sup>[17]</sup>实现不同分组特征图之间的特征交互,充分利用了不同网络层特征之间的相关性,模块结构如图 4 所示。

具体来说,对于不同网络层的聚合特征图  $\mathbf{y} \in \mathbf{R}^{N_c \times H \times W}$ ,主要包括维度划分、自适应坐标注意力权重分配和分组特征图聚合 3 个步骤。

### 1) 维度划分

将  $\mathbf{y}$  沿通道维度划分为  $G$  组:

$$[\mathbf{y}_1, \dots, \mathbf{y}_G] = \text{group}(\mathbf{y}), \quad \mathbf{y}_k \in \mathbf{R}^{N_c/G \times H \times W}, \quad k \in [1, G], \quad (6)$$

式中:  $[\mathbf{y}_1, \dots, \mathbf{y}_G]$  表示分组后的特征图;  $\text{group}(\cdot)$  表示分组操作;  $N_c$ 、 $H$  和  $W$  分别表示通道数、空间高度和宽度。

### 2) 自适应坐标注意力权重分配

通过坐标注意力模块学习生成相应的重要性系数  $g_k$ 。对于任意分组特征图  $\mathbf{y}_k$ , 首先对其进行一维的全局平均池化,生成水平和垂直两个方向的特征图,并对

两个输出特征图进行拼接,表达式为

$$\mathbf{f}_k = \delta \left\{ \text{Conv1} \left[ F_{x\text{Avg}}(\mathbf{y}_k), F_{y\text{Avg}}(\mathbf{y}_k) \right] \right\}, \quad (7)$$

式中:  $\mathbf{f}_k$  表示拼接后的特征图;  $F_{x\text{Avg}}(\cdot)$  和  $F_{y\text{Avg}}(\cdot)$  分别表示沿  $x$  和  $y$  方向的平均池化;  $\delta$  代表非线性激活函数。然后将  $\mathbf{f}_k$  沿二维空间维度分解为  $\mathbf{f}_k^h$  和  $\mathbf{f}_k^w$ , 分别编码为两组重要性系数,并通过乘法操作将它们应用于输入特征图,以突出感兴趣的特征,得到坐标注意力模块的输出特征图  $\mathbf{y}'_k$ , 表达式为

$$\mathbf{y}'_k = \mathbf{y}_k \times \sigma \left[ \text{Conv1}(\mathbf{f}_k^h) \right] \times \sigma \left[ \text{Conv1}(\mathbf{f}_k^w) \right], \quad (8)$$

式中:  $\sigma$  代表 Sigmoid 激活函数。

CSCAB 中的坐标注意力模块在增强特征的同时,不仅捕获了跨通道信息,还捕获了位置敏感信息,有助于整个模型更准确地定位感兴趣区域,促进更具表现力特征的生成。此外,优化删除坐标注意力中批归一化操作,提高了重建结果的稳定性和一致性。

### 3) 分组特征图聚合

对坐标注意力模块输出的分组特征图  $\mathbf{y}'_k$  进行聚合,表达式为

$$\mathbf{y}' = \text{shuffle} \left[ \text{Agg}(\mathbf{y}'_k) \right], \quad k \in [1, G], \quad (9)$$

式中:  $\mathbf{y}'$  表示 CSCAB 输出特征图;  $\text{Agg}(\cdot)$  表示聚合操作;  $\text{shuffle}(\cdot)$  表示通道重组操作。通过通道重组操作实现了跨组特征交互,充分关注不同网络层特征  $F_i$  之间的相关性。

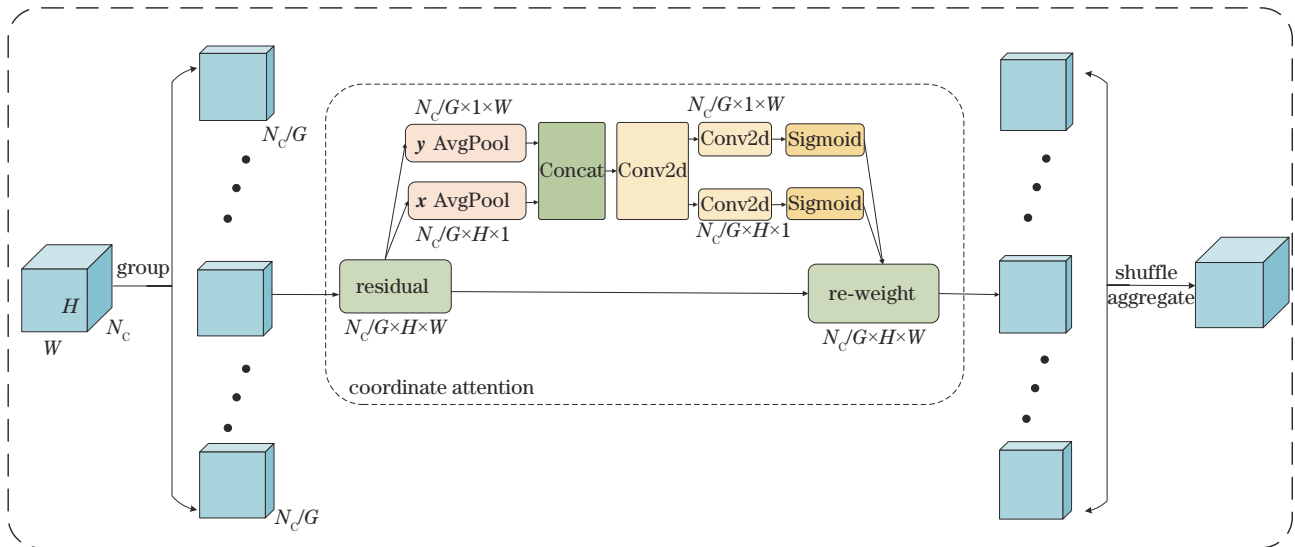


图 4 通道重组坐标注意力模块

Fig. 4 Channel shuffle coordinate attention block

## 2.3 双层残差聚合模块

简单的堆叠特征提取模块过深时,有时会在得不到反向传播结果的情况,导致前面模块的权重参数无法得到有效更新,使得模型的特征表现能力退化。本文设计残差聚合模块(RAB),每个RAB由4个残差模块构成,其中每个残差模块(RB)包括ReLU激活函数和2层卷积层。对于第  $i$  个残差模块的输入特征

$\mathbf{x}_{\text{Resi}}$ , 通过残差模块获得残差聚合特征  $\mathbf{f}_{\text{Resi}}$ , 表达式为

$$\mathbf{f}_{\text{Resi}} = \text{Conv3} \left\{ \text{ReLU} \left[ \text{Conv3}(\mathbf{x}_{\text{Resi}}) \right] \right\} + \mathbf{x}_{\text{Resi}}, \quad (10)$$

$$i = 1, 2, 3, 4.$$

残差模块外也同样采用残差连接,并通过跳跃连接串联4个残差模块的输出特征,最后通过  $1 \times 1$  卷积充分融合所有特征,得到RAB的输出特征,表达式为



$$f_{\text{RAB}i} = \text{Conv1}(\text{Concat}[f_{\text{Res}1}, f_{\text{Res}2}, f_{\text{Res}3}, f_{\text{Res}4}]) + x_{\text{RAB}i}, \quad i=1, 2, 3, 4, \quad (11)$$

式中:  $x_{\text{RAB}i}$  和  $f_{\text{RAB}i}$  表示第  $i$  个 RAB 的输入和输出特征; Concat 表示合并操作。

RAB 的外部采用与其内部相似的连接结构,并在最后增加 CSCAB,提高残差模块中深层特征的利用率,进而构建双层残差聚合模块(DRAB),其结构如图 5 所示。

$$f_{\text{DRAB}i} = \text{Conv1}\{\text{CSCAB}(\text{Concat}[f_{\text{RAB}1}, f_{\text{RAB}2}, f_{\text{RAB}3}, f_{\text{RAB}4}])\} + x_{\text{DRAB}i}, \quad i=1, 2, 3, 4, \quad (12)$$

式中:  $x_{\text{DRAB}i}$  和  $f_{\text{DRAB}i}$  表示第  $i$  个 DRAB 的输入和输出特

征。DRAB 不仅充分聚合了每个 RAB 提取的特征,也保证了残差模块中特征的长距离有效传播,充分提取更具表现力的特征,从而生成更好的 SR 图像。

## 2.4 损失函数

ATSR 网络采用 L1 损失函数进行训练。L1 损失函数相比于其他损失函数,能更好地重建出图像细节信息并且可以使模型快速收敛。L1 损失函数可以表示为

$$L(\theta) = \frac{1}{WH} \sum_{x=0}^W \sum_{y=0}^H |f_{\text{SR}}(x, y) - f_{\text{HR}}(x, y)|, \quad (13)$$

式中:  $\theta$  为网络模型的参数集;  $W$  和  $H$  分别为图像的宽和高;  $f_{\text{SR}}(x, y)$  和  $f_{\text{HR}}(x, y)$  分别表示重建后图像和真实分辨率图像。

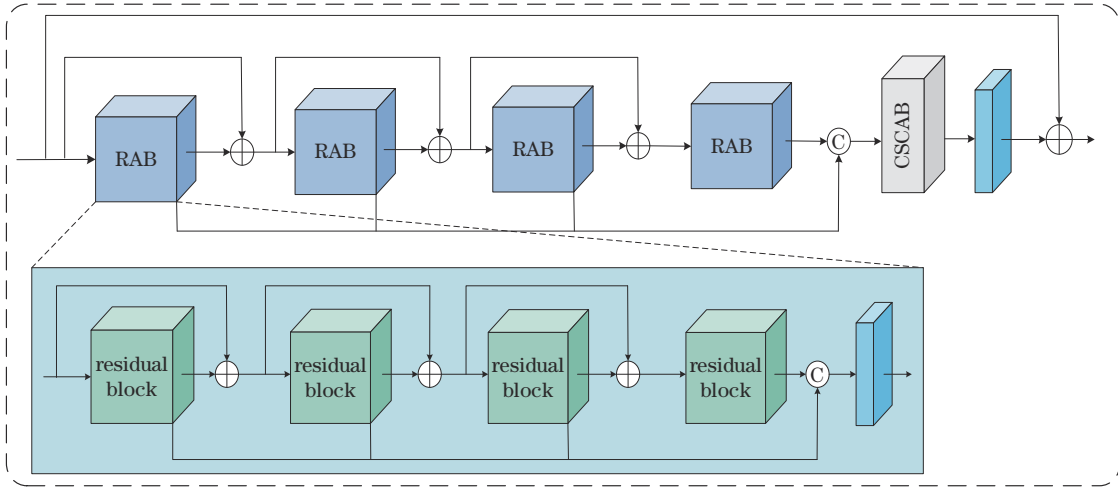


图 5 双层残差聚合模块

Fig. 5 Double-layer residual aggregation block

## 3 实验结果及分析

### 3.1 实验数据集

所提算法使用 DIV2K 数据集<sup>[18]</sup>作为训练数据集进行端到端的模型训练。DIV2K 数据集包括 800 张高质量的训练图像、100 张验证集图像和 100 张测试集图像。在训练过程中,对 800 张训练图像随机旋转  $90^\circ$ 、 $180^\circ$ 、 $270^\circ$  以及水平翻转进行数据扩充。在测试阶段,使用 Set5<sup>[19]</sup>、Set14<sup>[20]</sup>、BSD100<sup>[21]</sup>、Urban100<sup>[22]</sup> 和 Manga109<sup>[23]</sup> 5 个基准数据集,对所提算法 ATSR 与其他经典超分辨率重建算法在测试集上进行图像重建效果对比。

### 3.2 实验设置及训练策略

实验所用的硬件平台为搭载 Intel Core i9-10900K 处理器并配置 NVIDIA GeForce RTX 3080 Ti 显卡的服务器。软件平台环境为 Ubuntu18.04 操作系统、Python3.6 编译器环境、PyTorch 1.10 深度学习框架。训练过程初始学习率设置为  $10^{-4}$ ,并且每 200 个迭代周期学习率衰减为原来的  $1/2$ 。采用自适应矩估计优化器 (ADAM) 进行训练,使用  $\beta_1=0.9$ 、 $\beta_2=0.999$  和  $\epsilon=10^{-8}$  作为默认参数值。

### 3.3 性能评价指标

采用 YCbCr 空间 Y 通道上的峰值信噪比 (PSNR) 和结构相似指数 (SSIM)<sup>[24]</sup> 评价超分辨率重建结果。PSNR 根据像素点之间的差异来计算和衡量图像的失真程度,PSNR 值越高表示重建图像的质量越好。假设 HR 图像  $f_{\text{HR}}$  和重建图像  $f_{\text{SR}}$  的像素都为  $N$ , PSNR 的计算公式为

$$R_{\text{PSNR}} = 10 \cdot \log_{10} \left\{ \frac{L^2}{\frac{1}{N} \sum_{i=1}^N [f_{\text{HR}}(i) - f_{\text{SR}}(i)]^2} \right\}, \quad (14)$$

式中:  $L$  表示图像的最大像素值,通常取值为 255。

结构相似指数衡量图像之间的结构相似程度,包括亮度、对比度和结构 3 个方面,计算公式为

$$S_{\text{SSIM}} = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma, \quad (15)$$

式中:  $l(x, y)$ 、 $c(x, y)$  和  $s(x, y)$  分别表示两图像之间亮度、对比度和结构的函数表示;  $\alpha$ 、 $\beta$  和  $\gamma$  为三者的权重值,通常将值设定为 1。因此 SSIM 的计算公式为

$$S_{\text{SSIM}} = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (16)$$

式中:  $\sigma_x$  和  $\sigma_y$  表示图像  $x$  和图像  $y$  的方差;  $\mu_x$  和  $\mu_y$  表示均值;  $\sigma_{xy}$  表示两图像的协方差。SSIM 值越接近 1, 则说明重建图像与 HR 图像结构越相似, 图像重建质量越好。

### 3.4 消融实验结果及分析

为验证 ATSR 的有效性, 对 ATB、CSCAB 和 DRAB 结构设计了消融实验。模型 1 为采用单分支结构的 Baseline 模型; 模型 2 在模型 1 的基础上添加自适应双分支模块; 模型 3 在模型 1 的基础上添加通道重组坐标注意力模块; 模型 4 在模型 1 的基础上添加双层残差聚合模块; 模型 5 采用自适应双分支模块, 并添加了通道重组坐标注意力模块; 模型 6 为所提 ATSR 模型。不同模型在 Urban100 测试集上对 2 倍重建图像的平均 PSNR 和 SSIM 值如表 1 所示, 其中符号“√”代表包括该模块, “×”表示不包括该模块, 粗体数字为最高值, 下划线标记数字为次高值。

表 1 消融实验结果  
Table 1 Results of ablation study

Method	ATB	CSCAB	DRAB	PSNR /dB (SSIM)
Model 1	×	×	×	30.87(0.9150)
Model 2	√	×	×	32.08(0.9278)
Model 3	×	√	×	32.03(0.9274)
Model 4	×	×	√	31.98(0.9269)
Model 5	√	√	×	<u>32.28(0.9299)</u>
Model 6(ATSR)	√	√	√	<b>32.44(0.9308)</b>

从表 1 可以看出, 所提出的 3 个模块均有助于提升网络重建性能。模型 2 相较于模型 1, 其 PSNR 和 SSIM 指标分别提高了 1.21 dB 和 0.0128, 表明自适应双分支模块可以通过调整自适应权重来更有效地利用注意力机制, 从而提高超分辨率重建性能。模型 3 的 PSNR 和 SSIM 指标较模型 1 提高了 1.16 dB 和 0.0124, 通道重组坐标注意力模块将所有中间层特征沿通道维度分组, 并通过坐标注意力机制更准确地定位感兴趣区域, 促进更具表现力特征的生成。模型 4 与模型 1 相比, PSNR 和 SSIM 指标提高了 1.11 dB 和 0.0119, 这证明双层残差聚合模块通过嵌套的残差结构, 提升了网络的特征聚合和特征提取能力。模型 5 相比模型 1, PSNR 和 SSIM 指标提高了 1.41 dB 和 0.0149, 进一步证明了在自适应双分支模块的基础上添加通道重组坐标注意力模块的有效性。同时添加了 ATB、CSCAB 和 DRAB 的模型 6 性能最佳。以 Set5 中 butterfly 图为例, 图 6 为 2 倍放大系数下各模型的部分通道输出特征图, 模型 6 中每个通道输出的特征图纹理清晰, 其可以重建出较为清晰的边缘信息。然而相较于模型 6, 其他模型每个通道的输出特征图边缘会出现模糊的现象, 影响重建图像的质量, 进而验证了所提算法的有效性。

为验证 LKA 对算法性能的影响, 在 Set14 和 Urban100 测试集上进行 2 倍超分辨率重建实验, 对大核卷积分解前后的结果进行对比, 如表 2 所示。对大核卷积进行分解后性能更佳, 分解后参数量为分解前的 12%, 在 Set14 和 Urban100 上, PSNR 指标分别提高

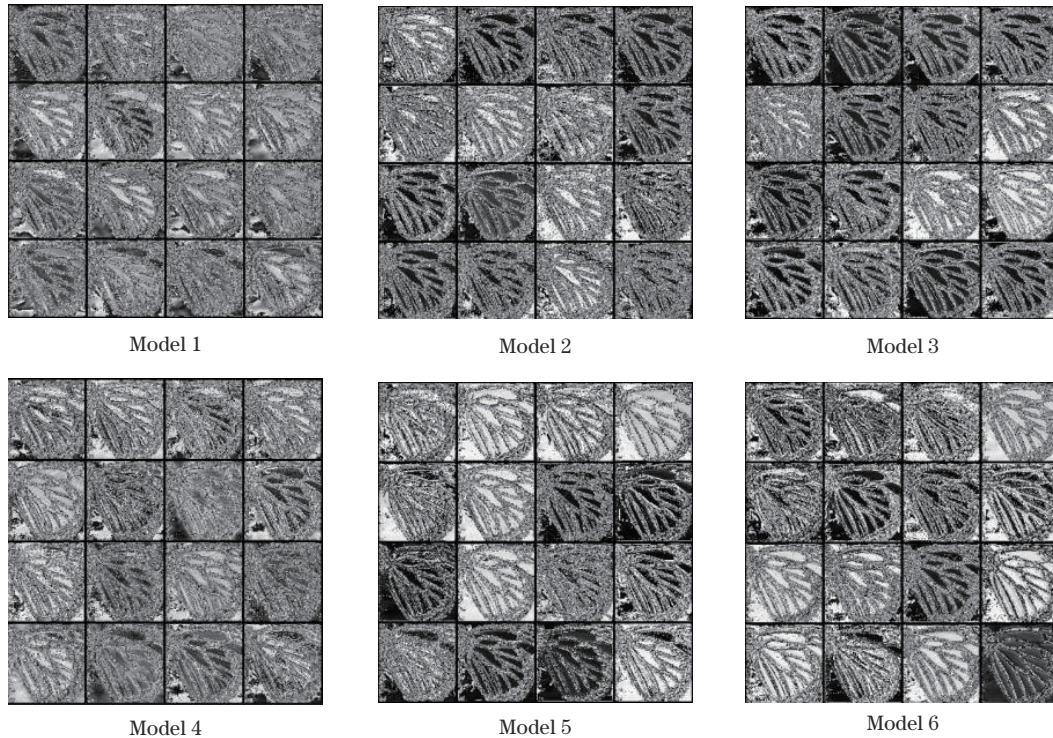


图 6 输出特征图效果对比

Fig. 6 Effect comparison of output feature maps

表 2 大核卷积分解前后所提算法的性能

Table 2 Performance of the proposed algorithm before and after large kernel convolutional decomposition

Method	Params / $10^3$	PSNR /dB(SSIM)	
		Set14	Urban100
Before decomposition	11089	33.75	32.34
		(0.9191)	(0.9302)
After decomposition	1308	33.82	32.44
		(0.9193)	(0.9308)

0.07 dB 和 0.10 dB。这说明 LKA 通过对大核卷积进行分解,获得更大的感受野,以较小的计算量实现全局图像特征表示,充分证明了 LKA 对图像超分辨率重建的有效性。

为验证双层残差聚合模块中残差模块的数量对算法性能的影响,使用 2、3、4、5 个残差模块进行实验,其他实验设置和训练策略均保持一致,不同数量残差模块在 Set14 和 Urban100 测试集上进行 2 倍超分辨率重建后的平均 PSNR 和 SSIM 如表 3 所示。

由表 3 结果可知,随着残差模块数量的增加,模型重建性能也会提升。当残差模块数量为 4 时,相比 3 个残差模块,PSNR 指标在 Set14 和 Urban100 上分别提高了 0.16 dB 和 0.27 dB,参数量增加了  $187 \times 10^3$ 。相比残差模块数量为 4,当残差模块数量为 5 时, Urban100 数据集上 PSNR 指标提升不明显,在 Set14

表 3 残差模块数量对所提算法性能的影响

Table 3 Influence of number of RB on the performance of the proposed algorithm

Res module No.	Params / $10^3$	PSNR /dB(SSIM)	
		Set14	Urban100
2	967	33.63(0.9179)	32.16(0.9286)
3	1121	33.66(0.9181)	32.17(0.9287)
4	1308	33.82(0.9193)	32.44(0.9308)
5	1612	33.80(0.9188)	32.48(0.9312)

数据集上甚至出现了指标下降的情况,但是模型参数量增加了  $304 \times 10^3$ 。因此,综合考虑模型性能和参数量,所提算法选用 4 个残差模块。

### 3.5 客观评价指标对比及分析

为客观地评价重建算法的性能,将所提算法与当前主流算法 VDSR<sup>[6]</sup>、CARN<sup>[7]</sup>、IMDN<sup>[25]</sup>、SRMDNF<sup>[26]</sup>、SMSR<sup>[27]</sup>、PAN<sup>[13]</sup>、ARRFN<sup>[28]</sup>、PDAN<sup>[29]</sup>、A2N<sup>[14]</sup>、DRSDN<sup>[30]</sup>和 LESR<sup>[31]</sup>进行了比较。不同算法在 5 种测试集上的平均 PSNR 和 SSIM 值如表 4 所示。从表 4 可以看出,ATSR 方法在各数据集上均表现出良好的重建性能,尤其在 Set5 数据集上,对 3 种不同尺度图片的重建结果均实现最优,在细节信息较多的 Urban100 测试集上也均取得了最高的 PSNR 值。对于其他测试集,ATSR 的平均 PSNR 和 SSIM 值部分小于其他算法的结果,但均接近最优值。

表 4 不同算法在各数据集上的性能

Table 4 Performance of different algorithms on each dataset

Method	Scale	PSNR /dB(SSIM)				
		Set5	Set14	Urban100	BSD100	Manga109
VDSR <sup>[6]</sup>	×2	37.53(0.9590)	33.05(0.9130)	30.77(0.9140)	31.90(0.8960)	37.22(0.9750)
CARN <sup>[7]</sup>		37.76(0.9590)	33.52(0.9166)	31.92(0.9256)	32.09(0.8978)	38.36(0.9765)
IMDN <sup>[25]</sup>		38.00(0.9605)	33.63(0.9177)	32.17(0.9283)	32.19(0.8996)	38.88(0.9774)
SRMDNF <sup>[26]</sup>		37.79(0.9601)	33.32(0.9159)	31.33(0.9204)	32.05(0.8985)	38.07(0.9761)
SMSR <sup>[27]</sup>		38.00(0.9601)	33.64(0.9179)	32.19(0.9284)	32.17(0.8990)	38.76(0.9771)
PAN <sup>[13]</sup>		38.00(0.9605)	33.59(0.9181)	32.01(0.9273)	32.18(0.8997)	38.70(0.9773)
ARRFN <sup>[28]</sup>		38.01(0.9606)	33.66(0.9179)	32.27(0.9295)	32.20(0.8999)	
PDAN <sup>[29]</sup>		38.05(0.9607)	33.65(0.9182)	32.36(0.9300)	32.20(0.8998)	38.71(0.9771)
A2N <sup>[14]</sup>		38.06(0.9608)	33.75(0.9194)	32.43(0.9311)	32.22(0.9002)	38.87(0.9769)
DRSDN <sup>[30]</sup>		38.06(0.9607)	33.65(0.9189)	32.40(0.9308)	32.23(0.9003)	
LESR <sup>[31]</sup>	38.07(0.9606)	33.80(0.9194)	32.42(0.9308)	32.24(0.9000)	38.96(0.9776)	
ATSR		38.11(0.9608)	33.82(0.9193)	32.44(0.9308)	32.24(0.8996)	38.94(0.9775)
VDSR <sup>[6]</sup>	×3	33.67(0.9210)	29.78(0.8320)	27.14(0.8290)	28.83(0.7990)	32.01(0.9340)
CARN <sup>[7]</sup>		34.29(0.9255)	30.29(0.8407)	28.06(0.8493)	29.06(0.8034)	33.50(0.9440)
IMDN <sup>[25]</sup>		34.36(0.9270)	30.32(0.8417)	28.17(0.8519)	29.09(0.8046)	33.61(0.9445)
SRMDNF <sup>[26]</sup>		34.12(0.9254)	30.04(0.8382)	27.57(0.8398)	28.97(0.8025)	33.00(0.9403)
SMSR <sup>[27]</sup>		34.40(0.9270)	30.33(0.8412)	28.25(0.8536)	29.10(0.8050)	33.68(0.9445)
PAN <sup>[13]</sup>		34.40(0.9271)	30.36(0.8423)	28.11(0.8511)	29.11(0.8050)	33.61(0.9448)
ARRFN <sup>[28]</sup>		34.38(0.9272)	30.36(0.8422)	28.22(0.8533)	29.09(0.8050)	
PDAN <sup>[29]</sup>		34.44(0.9276)	30.39(0.8437)	28.34(0.8563)	29.11(0.8063)	33.63(0.9448)



表 4 (续)

Method	Scale	PSNR /dB(SSIM)				
		Set5	Set14	Urban100	BSD100	Manga109
A2N <sup>[14]</sup>		34.47(0.9279)	30.44(0.8437)	28.41(0.8570)	29.14(0.8059)	33.78(0.9458)
DRSDN <sup>[30]</sup>		34.48(0.9282)	30.41(0.8445)	28.45(0.8589)	29.17(0.8072)	
LESR <sup>[31]</sup>		34.49(0.9278)	30.42(0.8431)	28.39(0.8567)	29.13(0.8059)	33.76(0.9455)
ATSR		<b>34.53(0.9283)</b>	<b>30.50(0.8441)</b>	<b>28.54(0.8593)</b>	<b>29.18(0.8060)</b>	<b>33.93(0.9463)</b>
VDSR <sup>[6]</sup>		31.35(0.8830)	28.02(0.7680)	25.18(0.7540)	27.29(0.7260)	28.83(0.8870)
CARN <sup>[7]</sup>		32.13(0.8937)	28.60(0.7806)	26.07(0.7837)	27.58(0.7349)	30.47(0.9084)
IMDN <sup>[25]</sup>		32.21(0.8948)	28.58(0.7811)	26.04(0.7838)	27.56(0.7353)	30.45(0.9075)
SRMDNF <sup>[26]</sup>		31.96(0.8925)	28.35(0.7787)	25.68(0.7731)	27.49(0.7337)	30.09(0.9024)
SMSR <sup>[27]</sup>		32.12(0.8932)	28.55(0.7808)	26.11(0.7868)	27.55(0.7351)	30.54(0.9085)
PAN <sup>[13]</sup>	×4	32.13(0.8948)	28.61(0.7822)	26.11(0.7854)	27.59(0.7363)	30.51(0.9095)
ARRFN <sup>[28]</sup>		32.22(0.8952)	28.60(0.7817)	26.09(0.7858)	27.57(0.7355)	
PDAN <sup>[29]</sup>		32.28(0.8957)	28.66(0.7831)	26.27(0.7922)	27.62(0.7378)	30.64(0.9098)
A2N <sup>[14]</sup>		32.30(0.8966)	28.71(0.7842)	26.27(0.7920)	27.61(0.7374)	30.67(0.9110)
DRSDN <sup>[30]</sup>		32.28(0.8962)	28.64(0.7836)	26.30(0.7933)	27.64(0.7388)	
LESR <sup>[31]</sup>		32.28(0.8952)	28.65(0.7827)	26.25(0.7908)	27.61(0.7365)	30.68(0.9100)
ATSR		<b>32.34(0.8973)</b>	<b>28.75(0.7832)</b>	<b>26.37(0.7934)</b>	<b>27.66(0.7370)</b>	<b>30.69(0.9089)</b>

为进一步验证所提算法 ATSR 的有效性,分别与 SRMDNF<sup>[26]</sup>、CARN<sup>[7]</sup>、ARRFN<sup>[28]</sup>、PDAN<sup>[29]</sup>、DRSDN<sup>[30]</sup> 和 LESR<sup>[31]</sup> 6 种超分辨率重建算法在模型参数量和重建性能方面进行对比。在 Set5 数据集上的 4 倍超分辨率重建对比结果如表 5 所示。所提 ATSR 与对比算法参数量均为同一数量级,但 PSNR 和 SSIM 都达到了最优,较次优值,PSNR 和 SSIM 分别提高了 0.06 dB 和 0.0011。所述充分说明了所提算法 ATSR 的有效性。

### 3.6 主观视觉效果对比及分析

由于人的视觉对细节更加敏感,对所提算法的重建图像与其他部分 SISR 算法的重建图像进行了细节对比。图 7 为 Urban100 数据集中 image\_012 图像 ×2 重建结果,CARN、IMDN 和 PAN 算法都无法正常重

表 5 不同算法的不同参数比较

Table 5 Comparison of different parameters of different algorithms

Method	Params /10 <sup>3</sup>	PSNR /dB	SSIM
SRMDNF <sup>[26]</sup>	1552	31.96	0.8925
CARN <sup>[7]</sup>	1592	32.13	0.8937
ARRFN <sup>[28]</sup>	1008	32.22	0.8952
PDAN <sup>[29]</sup>	1587	32.28	0.8957
DRSDN <sup>[30]</sup>	1095	32.28	0.8962
LESR <sup>[31]</sup>	1020	32.28	0.8952
ATSR	1403	<b>32.34</b>	<b>0.8973</b>

建出玻璃建筑的正确纹理,得到的图像均出现了其他方向纹理和纹理模糊的现象,A2N 重建结果则出现纹理扭曲和结构失真的问题,而 ATSR 算法可以准确重建出玻璃建筑的纹理。图 8 为 image\_092 图像 ×3 重



图 7 图像 Urban100\_012 在放大 2 倍下的重建效果对比

Fig. 7 Reconstruction effect comparison of Urban100\_012 at ×2 magnification



建结果,针对细小纹理特征,其他方法均出现了纹理丢失导致边缘模糊的现象,ATSR 仍可以恢复出更多的细节,保留相对更清晰完整的边缘纹理。图 9 为 image\_033 图像×4 重建结果,CARN、IMDN、PAN 和 A2N 处理的图像均出现了严重的结构扭曲,方格结构

均出现了模糊、图像失真的情况,而所提 ATSR 算法的重建图像边缘分明且层次结构清晰。根据 3 种不同尺度重建图像的对比,说明所提 ATSR 相比其他算法能够更有效地提取细节特征,表现出了良好的重建性能。

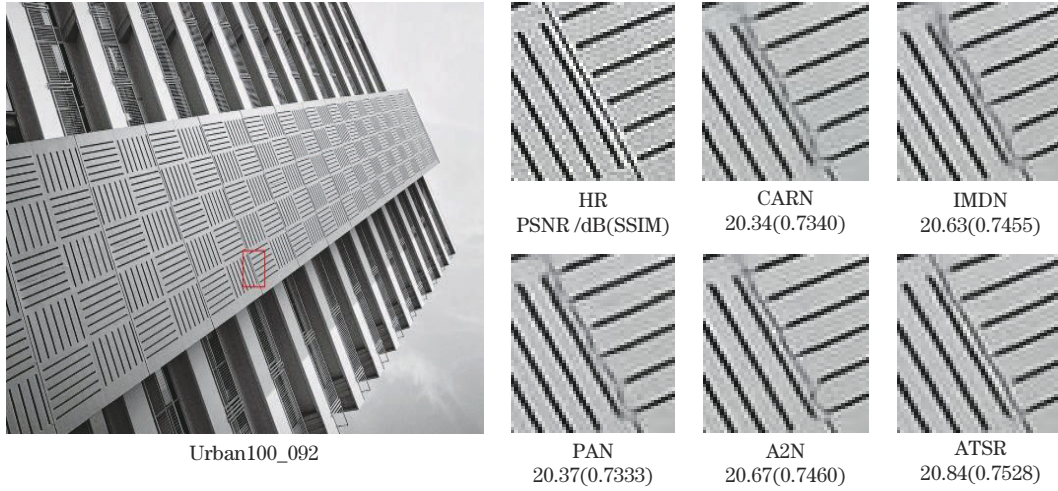


图 8 图像 Urban100\_092 在放大 3 倍下的重建效果对比

Fig. 8 Reconstruction effect comparison of Urban100\_092 at ×3 magnification

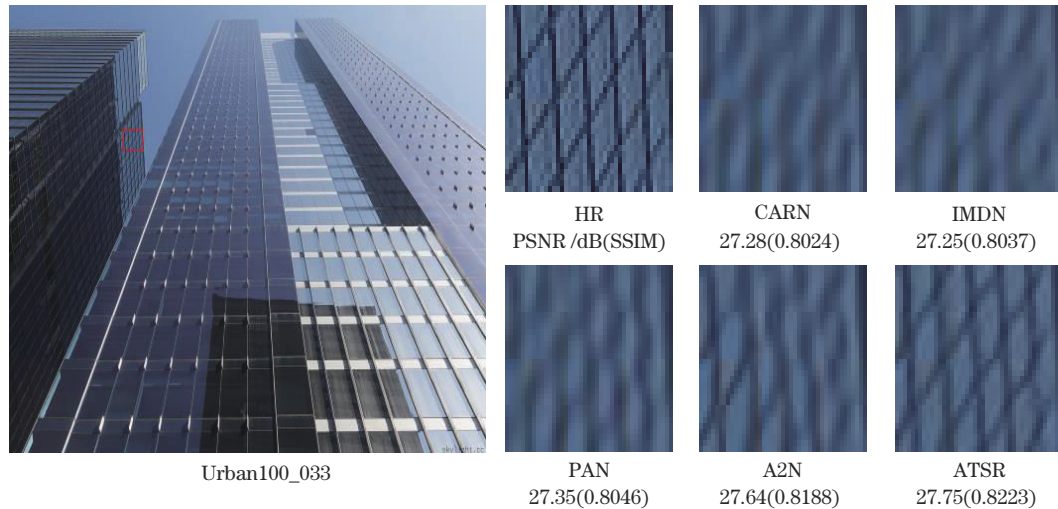


图 9 图像 Urban100\_033 在放大 4 倍下的重建效果对比

Fig. 9 Reconstruction effect comparison of Urban100\_033 at ×4 magnification

## 4 结 论

提出了一种基于自适应双分支的图像超分辨率重建算法。该算法采用自适应双分支模块,动态分配注意力分支和无注意力分支权重,有效增强有用特征和抑制冗余特征。采用通道重组坐标注意力模块促进跨组特征交互,关注不同网络层特征之间的相关性,并通过双层残差聚合模块进行残差特征聚合和特征提取,利用注意力机制提升网络模型的特征表达能力。实验结果表明,所提方法较其他主流超分辨率重建方法获得了更优的重建性能,重建图像表现出更好的视觉效果。在未来的工作中,将进一步研究如何有效利用注意力机制的内

容,以较少的参数量实现更好的重建性能。

## 参 考 文 献

- [1] 李强,汪西原,何佳玮. 基于生成对抗网络的遥感图像超分辨率重建改进算法[J]. 激光与光电子学进展, 2023, 60(10): 1028010.  
Li Q, Wang X Y, He J W. Improved algorithm for super-resolution reconstruction of remote-sensing images based on generative adversarial networks[J]. Laser & Optoelectronics Progress, 2023, 60(10): 1028010.
- [2] Feng R Y, Qian J Y, Peng Y J, et al. Femtosecond infrared optical vortex lasers based on optical parametric amplification[J]. High Power Laser Science and

- Engineering, 2022, 10: e29.
- [3] Reimold M, Assenbaum S, Beyreuther E, et al. OCTOPOD: single-bunch tomography for angular-spectral characterization of laser-driven protons[J]. High Power Laser Science and Engineering, 2023, 11: e68.
- [4] Lin J P, Haberstroh F, Karsch S, et al. Applications of object detection networks in high-power laser systems and experiments[J]. High Power Laser Science and Engineering, 2023, 11: e7.
- [5] Dong C, Loy C C, He K, et al. Learning a deep convolutional network for image super-resolution[M]// Fleet D, Pajdla T, Schiele B, et al. Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer, 2014, 8692: 184-199.
- [6] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1646-1654.
- [7] Ahn N, Kang B, Sohn K A. Fast, accurate, and lightweight super-resolution with cascading residual network[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11214: 256-272.
- [8] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [9] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [10] Guo M H, Lu C Z, Liu Z N, et al. Visual attention network[J]. Computational Visual Media, 2023, 9(4): 733-752.
- [11] Ding X H, Zhang X Y, Han J G, et al. Scaling up your kernels to  $31 \times 31$ : revisiting large kernel design in CNNs [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 11953-11965.
- [12] Zhang Y L, Li K P, Li K, et al. Image super-resolution using very deep residual channel attention networks[M]// Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 294-310.
- [13] Zhao H Y, Kong X T, He J W, et al. Efficient image super-resolution using pixel attention[M]//Karlinsky L, Michaeli T, Nishino K. Computer vision-ECCV 2022 workshops. Lecture notes in computer science. Cham: Springer, 2020, 12537: 56-72.
- [14] Chen H Y, Gu J J, Zhang Z. Attention in attention network for image super-resolution[EB/OL]. (2021-04-19)[2023-03-05]. <https://arxiv.org/abs/2104.09497>.
- [15] Chen Y P, Dai X Y, Liu M C, et al. Dynamic convolution: attention over convolution kernels[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11027-11036.
- [16] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 13708-13717.
- [17] Ma N N, Zhang X Y, Zheng H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design [M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11218: 122-138.
- [18] 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. New York: IEEE Press, 2017: 1122-1131.
- [19] 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, September 3-7, 2012, Surrey. Durham: British Machine Vision Association, 2012: 135.
- [20] Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations[M]//Curves and surfaces. Lecture notes in computer science. Heidelberg: Springer, 2012, 6920: 711-730.
- [21] 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 Eighth IEEE International Conference on Computer Vision, July 7-14, 2001, Vancouver, BC, Canada. New York: IEEE Press, 2002: 416-423.
- [22] 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. New York: IEEE Press, 2015: 5197-5206.
- [23] Matsui Y, Ito K, Aramaki Y, et al. Sketch-based manga retrieval using manga109 dataset[J]. Multimedia Tools and Applications, 2017, 76(20): 21811-21838.
- [24] 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.
- [25] 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, October 21-25, 2019, Nice, France. New York: ACM Press, 2019: 2024-2032.
- [26] Zhang K, Zuo W M, Zhang L. Learning a single convolutional super-resolution network for multiple degradations[C]//2018 IEEE/CVF Conference on

- Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 3262-3271.
- [27] Wang L G, Dong X Y, Wang Y Q, et al. Exploring sparsity in image super-resolution for efficient inference [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 4915-4924.
- [28] Qin J H, Zhang R M. Lightweight single image super-resolution with attentive residual refinement network[J]. Neurocomputing, 2022, 500: 846-855.
- [29] Wu H P, Gui J, Zhang J, et al. Pyramidal dense attention networks for single image super-resolution[J]. IET Image Processing, 2022, 16(12): 3247-3257.
- [30] Cheng G, Matsune A, Du H, et al. Exploring more diverse network architectures for single image super-resolution [J]. Knowledge-Based Systems, 2022, 235: 107648.
- [31] Tang Y G, Zhang X, Zhang X G. An efficient lightweight network for single image super-resolution[J]. Journal of Visual Communication and Image Representation, 2023, 93: 103834.