

## 高效全局注意网络的图像超分辨率重建

王庆庆<sup>1,2</sup>, 辛月兰<sup>1,2\*</sup>, 赵佳<sup>2</sup>, 郭江<sup>1,2</sup>, 王浩臣<sup>2</sup><sup>1</sup>青海师范大学计算机学院, 青海 西宁 810001;<sup>2</sup>省部共建藏语智能信息处理及应用国家重点实验室, 青海 西宁 810001

**摘要** 针对现有高效超分辨率重建算法大多集中于减少参数量, 缺乏对层次特征的关注, 存在图像高维特征不能得到充分利用的问题, 提出一种高效全局注意网络。该网络的主要思想是设计交叉自适应特征块对图像进行不同层次的深度特征提取, 以改善图像高频细节信息缺失的问题。另外, 构造了近邻像素重构块, 将空间关联性和像素分析相结合, 进一步促进边缘细节信息的重建。此外, 还提出一种多阶段动态余弦热重启训练策略, 通过动态调整学习率以避免模型过拟合, 提高训练过程的稳定性并优化网络性能。大量实验结果表明, 所提方法在 Set5 等 5 个基准数据集上相比于其他先进网络, 峰值信噪比 (PSNR) 和结构相似性 (SSIM) 性能指标平均提高了 0.51 dB 和 0.0078, 参数量和浮点运算量 (FLOPs) 平均降低了  $332 \times 10^3$  和  $70 \times 10^9$ 。综上, 所提方法在拥有较低复杂度的同时, 获得了更好的性能指标和视觉效果, 实现了网络高效化。

**关键词** 图像超分辨率重建; 高效全局注意; 层次特征; 像素重构; 训练策略

中图分类号 TP391.41

文献标志码 A

DOI: 10.3788/LOP232053

## Efficient Global Attention Networks for Image Super-Resolution Reconstruction

Wang Qingqing<sup>1,2</sup>, Xin Yuelan<sup>1,2\*</sup>, Zhao Jia<sup>2</sup>, Guo Jiang<sup>1,2</sup>, Wang Haochen<sup>2</sup><sup>1</sup>The College of Computer, Qinghai Normal University, Xining 810001, Qinghai, China;<sup>2</sup>The State Key Laboratory of Tibetan Intelligent Information Processing and Application, Xining 810001, Qinghai, China

**Abstract** To address the prevalent focus on reducing the parameter counts in current efficient super-resolution reconstruction algorithms, this study introduces an innovative efficient global attention network to solve the issues regarding neglecting hierarchical features and the underutilization of high-dimensional image features. The core concept of the network involves implementing cross-adaptive feature blocks for deep feature extraction at varying image levels to remove the insufficiency in high-frequency detail information of images. To enhance the reconstruction of edge detail information, a nearest-neighbor pixel reconstruction block was constructed by merging spatial correlation with pixel analysis to further promote the reconstruction of edge detail information. Moreover, a multistage dynamic cosine thermal restart training strategy was introduced. This strategy bolsters the stability of the training process and refines network performance through dynamic learning rate adjustments, mitigating model overfitting. Exhaustive experiments demonstrate that when the proposed method is tested against five benchmark datasets, including Set 5, it increases the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) performance metrics by an average of 0.51 dB and 0.0078, respectively, and trims the number of parameters and floating-point operations (FLOPs) by an average of  $332 \times 10^3$  and  $70 \times 10^9$  compared with leading networks. In conclusion, the proposed method not only reduces complexity but also excels in performance metrics and visualization, thereby attaining remarkable network efficiency.

**Key words** image super-resolution reconstruction; efficient global attention; hierarchical feature; pixel reconstruction; training strategy

收稿日期: 2023-09-05; 修回日期: 2023-10-06; 录用日期: 2023-10-20; 网络首发日期: 2023-11-07

基金项目: 国家自然科学基金项目 (61662062)、青海省自然科学基金面上项目 (2022-ZJ-929)

通信作者: \*xinyue001112@163.com

# 1 引言

图像超分辨率(SR)重建旨在从给定的低分辨率(LR)图像中重建出含有清晰细节特征的高分辨率(HR)图像。由于这个经典问题伴随着现代电视、手机等清晰度设备的不断发展,因此,提出一种轻量且有效的网络以重建高质量图像,并在这些设备上更好地显示重建图像是具有广泛意义的。

近年来提出的高效超分辨率网络大多试图减少参数量或减小计算复杂度。为了达到轻量化效果,通常采用具有共享权重的递归网络,但由于其复杂模块的拓扑结构并没有简化,从而不会减少运算时间和操作次数。为了找到轻量化和网络性能之间的平衡点,众多研究者持续探索基于深度卷积神经网络(CNN)和轻量级网络的方法,以更好地满足实际和商业需求。深度卷积神经网络因强大的建模能力吸引了诸多研究者进行深入探索。然而,这些网络依赖更深或更复杂的堆叠方式分层提取特征,导致整体网络的参数量升高,难以应用于实际场景。轻量级网络因计算资源受限,往往会注重更少的参数量和更快的运算速度。近年来,主流轻量级网络趋向于采用嵌入单一注意力机制的方式探索特征之间的相关性。这种方法尽管能增强局部模块的特征表达能力,但大多采用级联方式或简单将浅层和深度提取后的特征直接聚合以恢复图像,缺少对层级间特征的充分利用,导致重建效果不够显著。

为此,本文综合考虑当前先进高效网络模型,试图在轻量化和性能之间实现最优平衡,由此提出一种高效全局注意网络(EGAN)。首先考虑到局部与全局上下文信息间的相关性,从整体图像开始,对纵横区域的相关关系进行建模,并按照各个区域的关联线索进行自适应分配权重,提出了交叉自适应特征块(CAFB)。为实现更高质量的图像重建效果并减少模型参数量,提出了近邻像素重构块(NPRB),将图像空间关联性和像素分析相结合,有效提取出边缘和细节,重建出更高质量图像。为了进一步加速模型收敛并提高最终图像重建的准确性,提出了一种新的多阶段动态余弦热重启训练策略,在每个阶段,网络模型都可以享受来自多阶段动态调整学习率的优点,从而不断提高网络性能和收敛速度。结合CAFB、NPRB及所提热重启训练策略,EGAN实现了网络高效化。

本文的主要贡献包括4个方面:提出一种高效全局注意网络(EGAN)以解决图像超分辨率重建问题,与现有先进方法相比,EGAN有效提高了网络性能,并在轻量化和性能之间取得了更好的平衡;提出了交叉自适应特征块(CAFB),从全局多层次信息中提取并聚合图像特征以提高特征表达能力;提出了近邻像素重构块(NPRB),以促进边缘细节信息的重建,恢复更高质量图像;提出了一种多阶段动态余弦热重启训

练策略,利用阶段性动态调整学习率的方式提高模型的收敛速度,进一步提升了重建效果。

# 2 相关工作

图像SR主要分为基于插值的方法<sup>[1]</sup>、基于重建的方法<sup>[2]</sup>和基于学习的方法<sup>[3]</sup>。由于深度学习的引入,基于学习的方法取得了突破性进展。主要从高效图像超分辨率、多尺度特征提取和SR训练策略的角度讨论相关工作。

## 2.1 高效图像超分辨率

基于CNN的方法在高效图像SR任务中表现出了优异性能。其中,SRCNN<sup>[3]</sup>通过放大目标尺寸和利用三层卷积网络拟合非线性映射,得到高分辨率图像结果,是深度学习应用于图像超分辨率重建上的开山之作。随后,VDSR<sup>[4]</sup>在深层网络结构中首次将全局残差引入SR进行训练,以提升重建效果。然而,这些方法都使用插值处理后的LR图像作为网络输入,从而不可避免地增加了模型的计算复杂度;其次,经预处理的LR图像并未给恢复HR图像提供额外有用的信息。因此,当前的超分辨率技术主要是在网络末层<sup>[5]</sup>进行图像尺寸恢复的,以有效地降低计算成本。其他工作,如SAN<sup>[6]</sup>、HAN<sup>[7]</sup>、NLSA<sup>[8]</sup>、RCCN<sup>[9]</sup>,充分利用了空间或通道维度挖掘图像信息。尽管这些方法产生了显著的效果,但它们均以参数量大和计算复杂度为代价,导致在实际中难以得到运用。IDN<sup>[10]</sup>通过融合长短路径的信息,增强了对局部特征的代表能力。CARN<sup>[11]</sup>提出一种具有组卷积的级联残差网络,使信息更高效地向上传递。IMDN<sup>[12]</sup>提出了一种轻量级的信息多分化网络,通过信息蒸馏块逐步提取高维特征。LAGNet<sup>[13]</sup>提出了一种基于注意力引导的轻量级网络,通过结合双支路并联的通道注意力和空间分组增强模块捕捉图像特征信息。LMDFFN<sup>[14]</sup>结合轻量级特征提取和多种卷积策略,保持了较低的模型参数量和计算复杂度。ShuffleMixer<sup>[15]</sup>引入通道分割和洗牌操作并结合移动设备卷积块,构建了一个端到端的可训练网络以实现高效SR设计。尽管这些方法参数量较少,但与大型SR网络相比,这些方法并未对不同层次的特征进行区分性利用。

为此,所提EGAN利用快速全局注意特征提取的优势,即在多个层次上全面捕获并利用图像的全局和细节信息对特征的长期依赖性进行建模,同时保持了较低的计算复杂度和高效的运行性能,在较短的时间内快速、准确地提取全局注意特征,实现图像的高质量重建。

## 2.2 多尺度特征提取

近年来多尺度特征已广泛应用于各种高级视觉任务中,证明了其对提升模型性能的有效性。SMSR<sup>[16]</sup>综合利用不同尺寸的卷积核进行自适应特征检测和尺度特征融合,有效地捕获了图像中的高频细节信息

以实现高质量的图像重建。DLGSANet<sup>[17]</sup>提出多头动态局部自注意及稀疏全局自注意来动态提取局部特征,并为全局特征探索产生更好的自我关注。然而这些方法大多以参数量大和计算复杂度为代价获得多尺度特征。因此,最近的工作开始探索用更小的成本获取丰富的特征表示。陈伟杰等<sup>[18]</sup>提出一种层次信息自适应聚合网络(HIAAN),通过多尺度信息聚合和注意力重组实现了对特征的自适应校准。

与现有方法不同,本文提出了交叉自适应特征块和近邻像素重构块,利用自适应权重分配对纵横区域的关系进行建模,通过从全局多层次信息中提取聚合特征并结合空间关联性和像素分析,强化模型对图像高维特征的注意力,使其可以更快地捕获图像特征间的依赖关系,进而加强网络模型对细节信息的提取和重建能力。

### 2.3 SR 训练策略

根据机器学习理论,良好的预测结果来自于网络架构、训练数据和优化策略的协同作用。以往SR相关工作<sup>[19]</sup>主要侧重于网络架构的优化,而关于促进性能提升的训练策略的重要性内容却很少被探讨。通常这些SR网络采用Adam优化器最小化L1和L2损失函数进行训练或使用额外数据集延长训练时间;为了进一步提高训练的鲁棒性,还使用了较小学习率和batch size。最近的相关工作表明,先进的训练策略可以显著提高网络性能,通过优化现有网络架构以达到甚至超越新架构的性能水平。RRCAN<sup>[20]</sup>重新改进了RCAN,并证明增加训练迭代可以明显提高网络性能的结论。BSRN<sup>[21]</sup>通过数据增强并采用余弦学习率衰减帮助训练跳出局部最小值。FasterRFDN<sup>[22]</sup>证明对数据集进行预训练和通过最小化L1和L2损失函数对网络进行微调可以有效提高网络泛化能力的结论。

受上述工作启发,本文旨在进一步优化SR网络的训练策略,以充分发挥其潜力。由此,提出一种多阶段动态余弦热重启训练策略,通过阶段性调整学习率重新训练特征权重,避免模型过早进入局部最优解,确保模型能更好地捕捉全局特征信息,从而提高SR网络性能。

## 3 方法内容

首先介绍所提EGAN的整体网络架构,随后详细探讨关键组成部分,即交叉自适应特征块和近邻像素重构块,最后叙述一种新的多阶段动态余弦热重启训练策略。

### 3.1 网络架构

所提EGAN的整体架构如图1所示。整个网络由浅层特征提取、基于CAFB的深层特征提取和基于NPRB的HR图像重建组成。用 $I_{LR}$ 、 $I_{SR}$ 和 $I_{HR}$ 分别表示低分辨率图像、重建图像和高分辨率图像,并将 $I_{LR}$ 和 $I_{SR}$ 作为网络的输入和输出。

EGAN遵循一个简单的拓扑结构。在输入HR图像重建模块之前,从浅层特征提取和深层特征提取中得到的输出具有全局和局部残差连接。具体而言,给定退化图像 $I_{LR} \in \mathbf{R}^{3 \times H \times W}$ ,其中 $H$ 和 $W$ 分别表示图像的高和宽。首先采用由单个 $3 \times 3$ 卷积组成的浅层特征提取模块,用 $H_{SF}(\cdot)$ 表示,以提取局部特征 $F_s \in \mathbf{R}^{C \times H \times W}$ ,其中 $C$ 为层级特征的通道数,表达式为

$$F_s = H_{SF}(I_{LR}). \quad (1)$$

然后,将 $F_s$ 传入深层特征提取模块,深层特征提取模块具体由多个CAFB组成。则第 $n$ 个CAFB的输出可以表示为

$$F_{l,n} = \begin{cases} H_{CAFB}(F_s), & n = 1 \\ H_{CAFB}(F_{l,i-1}), & 1 < i \leq n \end{cases}, \quad (2)$$

式中: $F_{l,n}$ 表示第 $n$ 个CAFB的输出特征; $H_{CAFB}(\cdot)$ 表示CAFB的映射函数; $n$ 表示CAFB的个数,取 $n = 12$ 。此外,还使用一个 $3 \times 3$ 卷积层以平滑逐渐细化的深度特征图。接着将 $F_{l,n}$ 和 $F_s$ 整合作为输入,送入HR图像重建模块,即NPRB。则第 $m$ 个NPRB的输出可以表示为

$$F_{l,m} = \begin{cases} H_{NPRB}(F_{l,n} + F_s), & m = 1 \\ H_{NPRB}(F_{l,1}), & m = 2 \end{cases}, \quad (3)$$

式中: $F_{l,m}$ 表示第 $m$ 个NPRB的输出特征; $H_{NPRB}(\cdot)$ 表示NPRB的映射函数。接着利用一个单独的 $3 \times 3$ 卷积层并结合 $I_{LR}$ 得到最终输出 $I_{SR}$ 。 $I_{SR}$ 可以表示为

$$I_{SR} = H_{smooth}(F_{l,m}) + I_{LR}, \quad (4)$$

式中: $H_{smooth}(\cdot)$ 表示一个 $3 \times 3$ 卷积操作。

SR损失函数旨在最小化 $I_{SR}$ 和 $I_{HR}$ 之间的差异。其中,均方误差(MSE)损失、L1损失和L2损失广泛应用于图像优化。尽管这些方法可以提高客观评价指标,但它们的优化结果往往会有过于平滑的细节纹理,缺乏感官上的图像真实感。因此,为避免图像趋于光滑并减少计算量,选择较为稳定的鲁棒损失函数Charbonnier损失<sup>[23]</sup>,相较于其他损失,Charbonnier损失可以更好地处理异常值。当有接近零点的值,由于常数的存在,Charbonnier损失可以避免梯度消失;当有远离零点的值,由于开方,可以避免梯度爆炸。具体损失函数 $\mathcal{L}_S(I_{HR}, I_{SR}; \theta)$ 可以表示为

$$\mathcal{L}_S(I_{HR}, I_{SR}; \theta) = \frac{1}{N} \sum_{i=1}^N \rho(I_{HR,i} - I_{SR,i}), \quad (5)$$

$$\hat{\theta} = \arg \min_{\theta} \frac{1}{N} \sum_{i=1}^N \mathcal{L}_S[I_{HR,i}, H_{EGAN}(I_{LR,i})], \quad (6)$$

式中: $S$ 表示目标上采样尺度因子; $\theta$ 表示整个网络中的可更新参数集; $N$ 表示每批处理中训练样本的数量; $\rho(x) = \sqrt{x^2 + \epsilon^2}$ 表示Charbonnier惩罚函数,即L1范数的可微变体,常数 $\epsilon^2$ 根据经验设置为 $1 \times 10^{-6}$ ; $H_{EGAN}(\cdot)$ 表示EGAN的映射函数。

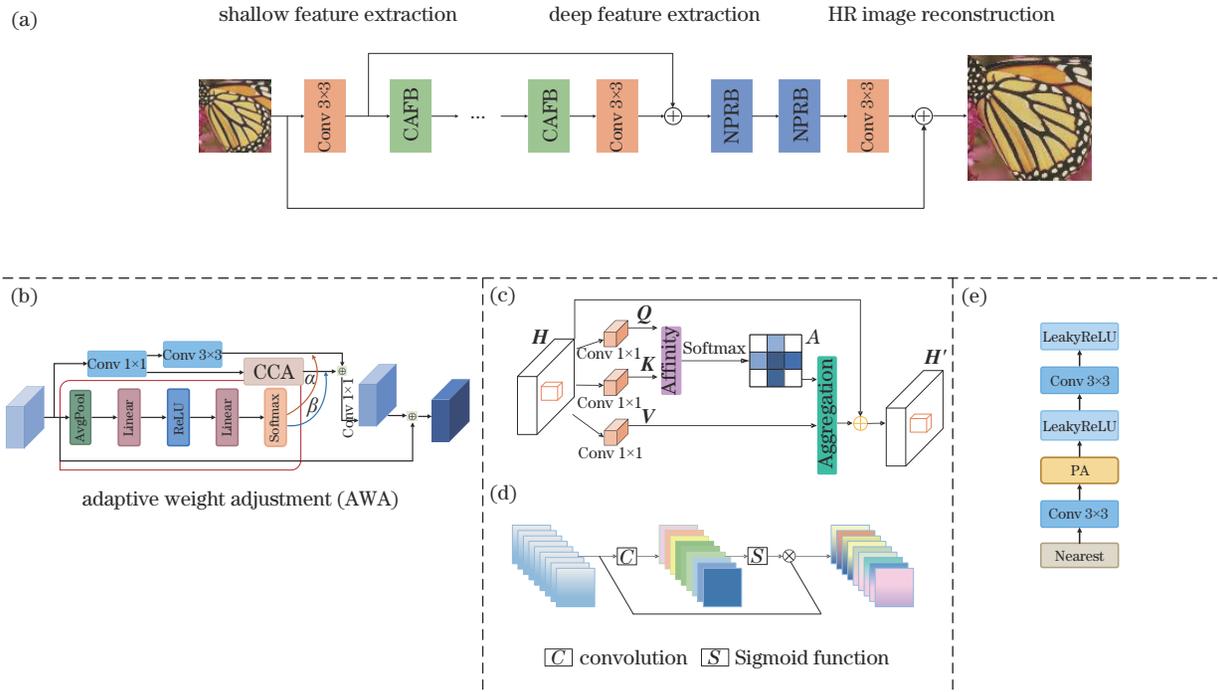


图1 网络架构。(a) 高效全局注意力网络；(b) 交叉自适应特征块；(c) 十字交叉注意力；(d) 像素注意力；(e) 近邻像素重构块  
Fig. 1 Network architecture. (a) EGAN; (b) CAFB; (c) CCA; (d) PA; (e) NPRB

### 3.2 交叉自适应特征块

为了探索局部与非局部上下文信息,以获得更好的特征表示,提出了交叉自适应特征块(CAFB)。如图1(b)所示,该模块包含两个部分:十字交叉注意力(CCA)块<sup>[24]</sup>和自适应权重调节(AWA)块。CAFB通过对层次特征进行自适应校准,捕获纵横方向的上下文特征,并利用自适应权重调节块选择性聚集不同层次特征,最后进行特征融合,借此实现层次特征的区分性利用。

十字交叉注意力块考虑到特征之间的长距离依赖性可以提供更加密集的上下文信息,以获得更具表征性的局部特征。由此,本文在交叉自适应特征块中引入十字交叉注意力块以提取图像周围像素信息,具体表现为通过计算目标特征像素点与特征图中其他点之间的相互关系,对目标像素点的特征进行加权,得到更加有效的目标特征。CCA具体结构如图1(c)所示。

首先,对输入特征 $H$ (尺寸为 $C \times W \times H$ )分别采用3个 $1 \times 1$ 卷积生成特征 $Q, K, V$ ,其中 $Q$ 和 $K$ 尺寸均为 $C' \times W \times H$ ,为了对通道进行降维,减少计算量,使得 $C'$ 小于 $C$ , $V$ 的尺寸保持不变。得到特征的具体过程可以表示为

$$Q, K, V = H_{\text{Conv}1 \times 1}(H), \quad (7)$$

式中: $H_{\text{Conv}1 \times 1}(\cdot)$ 表示一个 $1 \times 1$ 卷积操作。其次,对得到的特征 $Q$ 和 $K$ 进行Affinity操作,目的是获取特征图中每一个像素点与其同行同列像素点的关系,具体操作如图2所示。则其过程可以表示为

$$d_{i,\mu} = Q_{i,\mu} \Omega_{i,\mu}, \quad (8)$$

式中: $Q_{i,\mu}$ 表示特征图 $Q$ 中心位置的特征向量; $\Omega_{i,\mu}$ 表示特征图 $K$ 中位置 $i,\mu$ 处的特征向量; $d_{i,\mu}$ 表示对 $Q_{i,\mu}$ 与 $\Omega_{i,\mu}$ 分别进行向量相乘,组成一个尺寸为 $[(H+W-1) \times W \times H]$ 的特征图。

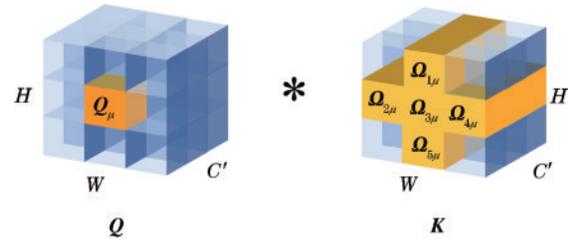


图2 Affinity操作

Fig. 2 Affinity operation

接着,对得到的特征图 $d_{i,\mu}$ 进行Softmax操作,目的是对每个像素点与纵横像素之间的位置关系进行归一化处理,以获得新的特征图,具体过程可以表示为

$$A = f_{\text{Softmax}}(d_{i,\mu}), \quad (9)$$

式中: $A$ 表示注意力特征图; $f_{\text{Softmax}}(\cdot)$ 表示Softmax函数。

最后,对注意力特征图 $A$ 与特征图 $V$ 进行Aggregation操作,通过对各个位置进行向量相乘相加,将上下文信息添加到局部特征 $H$ 中,目的是增强局部特征的像素级表示,并根据空间注意力图选择性地聚合上下文特征,以获得更大的上下文感受野,从而生成新的特征图。这些特征表示不仅实现了相互增益,而且对图像超分辨率重建任务更具鲁棒性。

Aggregation具体操作如图3所示,其过程可以表示为

$$\mathbf{H}'_{\mu} = \sum_{i=0}^{H+W-1} \mathbf{A}_{i,\mu} \Phi_{i,\mu} + \Phi_{\mu}, \quad (10)$$

式中: $\mathbf{H}'_{\mu}$ 表示输出特征图中的特征向量; $\mathbf{A}_{i,\mu}$ 表示特征图A中通道*i*处位置为 $\mu$ 的特征向量; $\Phi_{i,\mu}$ 表示特征图V中通道*i*处位置 $\mu$ 与其同行同列像素组成的特征向量; $\Phi_{\mu}$ 表示特征图H中位置 $\mu$ 处的特征向量。

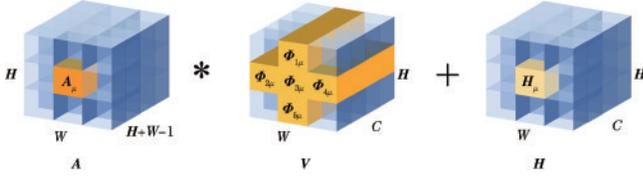


图3 Aggregation操作

Fig. 3 Aggregation operation

为了能更好地识别和强化不同复杂度图像的层次特征,如边缘和纹理信息,构建了一个级联模块即自适应权重调节(AWA)块用于自适应分配权重,与交叉注意力块一起实现特征交互。AWA通过调整不同层次特征的权重以提高图像超分辨率重建效果,具体结构如图1(b)方框标识。自适应权重调节块首先通过平均池化减小输入特征映射的空间尺寸,以降低数据维度和计算复杂性,从而更有效地处理高维特征,同时尽量保留核心特征信息。其次,第一个线性全连接层接收这些经过平均池化处理的特征图,并通过参数优化将特征图投影到新的特征空间。然后,非线性激活函数ReLU用于增强网络的非线性表达能力,从而使模型能逼近复杂的函数关系。经ReLU激活的特征数据传递到第二个线性全连接层,以进一步进行特征转换并为最终输出作准备。最后,Softmax函数将网络输出转换为概率分布,实现对每个特征权重的自适应分配,从而进一步提升网络模型的表达能力,确保了从输入到输出的信息有效传递和转换。

综上,AWA根据CCA计算出的每个特征的重要性进行学习,并自适应调整分配特征权重值 $\alpha$ 和 $\beta$ 于特征图 $\mathbf{X}$ 和 $\mathbf{H}'$ ,从而优化特征的组合和利用。这种灵活性使得模型能够更好地适应不同的图像内容和复杂度,从而提高图像超分辨率重建的质量和效率。AWA的具体过程可以表示为

$$H_{AWA}(\mathbf{F}_s) = \begin{cases} \mathbf{X}(\alpha) \\ \mathbf{H}'(\beta) \end{cases}, \quad (11)$$

式中: $H_{AWA}(\cdot)$ 表示AWA的映射函数; $\mathbf{X}$ 表示经过一个 $1 \times 1$ 和 $3 \times 3$ 卷积操作的特征图; $\mathbf{X}(\alpha)$ 和 $\mathbf{H}'(\beta)$ 分别表示权重值分配给 $\mathbf{X}$ 及 $\mathbf{H}'$ 后的特征图。

最后将得到的特征图聚合并使用 $1 \times 1$ 卷积进行降维,再与输入特征 $\mathbf{F}_s$ 聚合,得到融合特征,过程可以表示为

$$\mathbf{F}_1 = H_{Conv 1 \times 1}[\mathbf{X}(\alpha) + \mathbf{H}'(\beta)] + \mathbf{F}_s, \quad (12)$$

式中: $\mathbf{F}_1$ 表示得到的融合特征。

实验结果证明,CAFB能够有效提取图像周围像素的上下文信息,区分性利用多层次特征,从而提升整体网络的性能。

### 3.3 近邻像素重构块

目前大多SR网络在重构分支中采用了与上采样层相似的结构,即反卷积层或亚像素卷积层,但这种结构存在冗余且效率不高。因此,为进一步提高重建性能,并利用图像中相邻像素的相关性,提出了NPRB,目的是基于低分辨率图像中的像素及其邻近像素来预测高分辨率图像中的像素值,其结构如图1(e)所示。为了增强对特征的代表能力,在卷积层之间引入了像素注意力(PA)机制<sup>[25]</sup>。此外,还采用了最近邻(NN)上采样以进一步减少模型参数量,提高整体网络效率。

不同于通道注意力和空间注意力,像素注意力生成 $H \times W \times C$ 维度的注意力权重,其结构仅仅是 $1 \times 1$ 卷积和Sigmoid激活函数,如图1(d)所示。在像素注意力机制中,通过对输入特征图进行 $1 \times 1$ 卷积,可以将高维特征映射转换为一个低维的权重分布图。这个过程可视为一种全局信息的抽象化表示,它可以捕获输入特征在空间维度上的全局上下文信息。而Sigmoid函数的运用确保了每个像素的注意力权重被约束在0到1之间,这些权重可视为每个像素在特征提取和聚合阶段的相对重要性。因此,在像素注意力机制中, $1 \times 1$ 卷积和Sigmoid激活函数能有效地对每个像素的重要性进行量化,这种结构在增强网络处理图像细节方面起着关键性的作用。

如图1(e)所示,近邻像素重构块首先利用最近邻上采样通过插值生成高分辨率的特征映射,并复制每个像素值至其近邻像素来扩大输入数据尺寸。其次,通过 $3 \times 3$ 卷积操作进行第一次特征提取,旨在捕获上采样后特征映射中的高低级别视觉特性。接着引入像素注意力机制,为每个像素的信息内容分配不同权重,从而使网络能够关注更为重要的区域并削弱其他区域的影响。然后应用LeakyReLU激活函数增加模型的非线性表达能力,并允许一部分负值信息保留,从而避免输出限制在一个特定范围内。最后通过 $3 \times 3$ 卷积进行特征提取并再次应用LeakyReLU激活函数,以获取更丰富的特征表示,进一步提升了网络表达能力。

综上,NPRB能够充分考虑像素间的相关性,并综合利用了多级特征提取、注意力机制和非线性激活,以实现高效且精确的像素级重构。实验结果证明,这种方法能够提高重建图像的质量,尤其在恢复图像纹理和细节方面具有显著优势。

### 3.4 多阶段动态余弦热重启训练策略

图像超分辨率重建任务的复杂性和高度非线性特性导致网络存在收敛速度慢、优化困难及易陷入局部最优等问题<sup>[26]</sup>。本文提出一种新的多阶段动态余弦热重启训练策略,通过在训练的不同阶段使用不同学习

率,提高网络的整体性能并减少训练时间,同时也提供了一种有效的方式来控制和管理训练过程的复杂性。多阶段动态余弦热重启训练策略需要在训练过程中动态调整学习率,通过结合热启动和余弦退火策略的优点,针对图像超分辨率重建任务进行多阶段动态重启,以平衡学习率的探索和利用,在保证稳定性的同时,有效提高网络的训练效率和性能。所提训练策略的可视化结果如图 4 所示,具体过程描述如下。

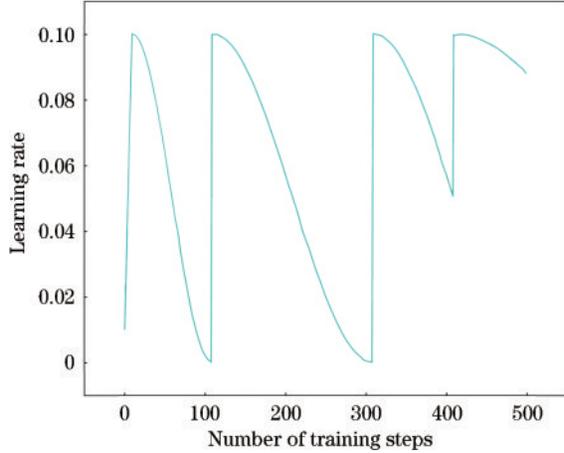


图 4 多阶段动态余弦热重启训练策略的可视化结果

Fig. 4 Visualization results of multistage dynamic cosine thermal restart training strategy

在训练初期,由于模型参数的随机初始化,较大的学习率可能会导致参数更新幅度过大,从而使得模型无法稳定地收敛。为解决此问题,使用较小学习率让模型逐步适应训练数据,避免在优化过程中出现过拟合或振荡问题,然后在 10 个热重启轮数中逐渐增大学习率到预设最大值,以提高网络模型的稳定性和鲁棒性。在模型适应训练数据后,进入多阶段训练周期,即在每个周期内,学习率会从初始值余弦式逐步降低到较小值。这种方式利用初期高学习率帮助网络更快找到优化问题的最优或近似解。随后通过降低学习率进行精细优化,以进一步降低损失函数值,实现对最优解的逼近。最后,在每个训练周期结束时,将重新增大学习率,并进入新的训练周期,目的是让网络模型能够跳出局部最优解,从而进一步优化网络性能。实验结果证明,多阶段动态余弦热重启训练策略能够有效提高网络在图像超分辨率任务中的收敛速度和整体性能,具体过程表示为

$$r_{\text{learn}}(t) = \begin{cases} \frac{t}{T_{\text{warmup}}} r_{\text{initial\_learning}}, & t < T_{\text{warmup}} \\ \eta_{\text{max}} \cdot \frac{1 + \cos \frac{\pi \cdot (t - T_{\text{warmup}})}{T_0}}{2}, & t \geq T_{\text{warmup}} \end{cases}, \quad (13)$$

式中: $t$ 表示周期总数; $T_{\text{warmup}}$ 表示热重启轮数;

$r_{\text{initial\_learning}}$ 表示初始学习率; $\eta_{\text{max}}$ 表示每个循环期间的最大学习率值; $T_0$ 表示初始循环轮数。

## 4 实验结果与分析

### 4.1 数据集及评价指标

采用 DIV2K 数据集作为实验基准数据集。DIV2K<sup>[27]</sup>包含了 1000 张高分辨率图像,其中 800 张图像作为训练集,100 张图像作为验证集,100 张图像作为测试集,其内容覆盖了多种场景和主题,如风景、建筑、人物等<sup>[28]</sup>。为了全面评估所提方法,采用 Set5<sup>[29]</sup>、Set14<sup>[30]</sup>、BSD100<sup>[31]</sup>、Urban100<sup>[32]</sup>、Manga109<sup>[33]</sup>这 5 个公开基准数据集作为测试集,具体内容场景如表 1 所示。

表 1 测试集信息

Table 1 Test set information

Dataset	Number of images	Content scene
Set5 <sup>[29]</sup>	5	Face, bird, butterfly
Set14 <sup>[30]</sup>	14	People, plants and animals, natural scenery, PPT, etc.
BSD100 <sup>[31]</sup>	100	People, plants and animals, and indoor and outdoor environments, natural scenery, etc.
Urban100 <sup>[32]</sup>	100	Cityscape
Manga109 <sup>[33]</sup>	109	Comics, backgrounds, etc.

采用峰值信噪比(PSNR)和结构相似度(SSIM)<sup>[34]</sup>作为评价指标,并在 YCbCr 颜色空间中的 Y 通道<sup>[35]</sup>上作定量比较,此外,还利用参数总量和浮点运算量(FLOPs)评估不同模型的复杂度。

### 4.2 实验环境及细节

实验所使用的计算平台的主要参数:NVIDIA GeForce RTX 3060,显存 12G;实验运行系统环境为 Windows;编程语言为 Python;深度学习框架为 PyTorch。

在 RGB 通道中训练网络模型,同时对 DIV2K 数据集的 800 张高分辨率图像进行双三次下采样处理<sup>[36]</sup>,得到不同尺度的低分辨率图像。此外,还对输入图像进行相应尺度裁剪<sup>[37]</sup>,并通过随机旋转 90°、180°、270°来增加训练数据量<sup>[38]</sup>。具体网络训练参数设置如表 2 所示。网络模型设计方面:CAFB 数量固定为 12 个;NPRB 数量固定为 2 个;网络的基本通道数设置为 64 个。

### 4.3 对比实验

为了验证所提方法的有效性,在相同测试集下,对 EGAN 与其他先进超分算法进行结果对比,包括 Bicubic、VDSR<sup>[4]</sup>、IDN<sup>[10]</sup>、CARN<sup>[11]</sup>、IMDN<sup>[12]</sup>、SMSR<sup>[16]</sup>、LAGNet<sup>[13]</sup>、LMDFFN<sup>[14]</sup>、RCCN<sup>[9]</sup>、ShuffleMixer<sup>[15]</sup>、DLGSANet<sup>[17]</sup>、HIAAN<sup>[18]</sup>。分别从定量指标和定性可

表2 网络训练参数设置

Table 2 Parameter setting for network training

Training program	Parameterization
Optimizer	Adam
$\epsilon^2$	$1 \times 10^{-6}$
$r_{\text{initial\_learning}}$	0.01
$\eta_{\text{max}}$	0.1
$T_0$	100
Batchsize	16
Epoch	500

视化结果两方面进行分析。

表3展示了不同尺度下( $\times 2$ 、 $\times 3$ 、 $\times 4$ )各算法在5个基准测试集上的定量比较结果,其中最优值用加粗表示,次优值用斜体表示。从表3可知:所提EGAN的指标优于SMSR、LAGNet、LMDFFN、RCCN、ShuffleMixer及HIAAN,但与方法DLGSANet取得了相似的结果;当放大因子为2时,所提方法在BSD100测试集上表现最好,PSNR和SSIM比次优的DLGSANet分别高1.62 dB和0.0230;当放大因子为3时,所提方法在5个基准测试集上大多取得优秀表

表3 各算法在5个基准测试集上的定量对比结果

Table 3 Quantitative comparison results of each algorithm on five benchmark test sets

Scale	Method	Parameters / FLOPs /		Set5	Set14	BSD100	Urban100	Manga109
		$10^3$	$10^9$	PSNR /dB (SSIM)	PSNR /dB (SSIM)	PSNR /dB (SSIM)	PSNR /dB (SSIM)	PSNR /dB (SSIM)
$\times 2$	Bicubic			33.66(0.9299)	30.24(0.8688)	29.56(0.8431)	26.88(0.8403)	30.80(0.9399)
	VDSR <sup>[4]</sup>	665	613	37.53(0.9587)	33.03(0.9124)	31.90(0.8960)	30.76(0.9140)	37.22(0.9729)
	IDN <sup>[10]</sup>	553	125	37.83(0.9600)	33.30(0.9148)	32.08(0.8985)	31.27(0.9196)	38.01(0.9749)
	CARN <sup>[11]</sup>	1592	223	37.76(0.9590)	33.52(0.9166)	32.09(0.8978)	31.92(0.9256)	38.36(0.9765)
	IMDN <sup>[12]</sup>	694	159	38.00(0.9605)	33.63(0.9177)	32.19(0.8996)	32.17(0.9283)	38.01(0.9749)
	SMSR <sup>[16]</sup>	985	132	38.00(0.9601)	33.64(0.9179)	32.17(0.8990)	32.19(0.9284)	38.76(0.9771)
	LAGNet <sup>[13]</sup>	447	84	37.79(0.9594)	33.40(0.9162)	32.10(0.8991)		
	LMDFFN <sup>[14]</sup>	339		37.92(0.9602)	33.43(0.9163)	32.08(0.8985)	31.85(0.9250)	38.52(0.9766)
	RCCN <sup>[9]</sup>			38.13(0.9610)	33.67(0.9185)	32.22(0.9001)	32.41(0.9302)	
	ShuffleMixer <sup>[15]</sup>	394	91	38.01(0.9606)	33.63(0.9180)	32.17(0.8995)	31.89(0.9257)	38.83(0.9774)
	DLGSANet <sup>[17]</sup>	4730	1097	<b>38.34(0.9617)</b>	<i>34.25(0.9231)</i>	<i>32.38(0.9025)</i>	<b>33.41(0.9393)</b>	<b>39.57(0.9789)</b>
HIAAN <sup>[18]</sup>	631	145	38.04(0.9611)	33.55(0.9186)	32.19(0.9006)	32.24( <i>0.9293</i> )	38.65(0.9773)	
EGAN(Ours)	304	20	<i>38.19(0.9613)</i>	<b>34.40(0.9260)</b>	<b>34.00(0.9255)</b>	<i>32.92(0.9262)</i>	<i>38.70(0.9777)</i>	
$\times 3$	Bicubic			30.39(0.8682)	27.55(0.7742)	27.21(0.7385)	24.46(0.7349)	26.95(0.8556)
	VDSR <sup>[4]</sup>	665	613	33.66(0.9213)	29.77(0.8314)	28.82(0.7976)	27.14(0.8279)	32.01(0.9310)
	IDN <sup>[10]</sup>	553	56	34.11(0.9253)	29.99(0.8354)	28.95(0.8013)	27.42(0.8359)	32.71(0.9381)
	CARN <sup>[11]</sup>	1592	119	34.29(0.9255)	30.29(0.8407)	29.06(0.8034)	28.06(0.8493)	33.50(0.9440)
	IMDN <sup>[12]</sup>	703	72	34.36(0.9270)	30.32(0.8417)	29.09(0.8046)	28.17(0.8519)	33.61(0.9445)
	SMSR <sup>[16]</sup>	993	68	34.40(0.9270)	30.33(0.8412)	29.10(0.8050)	28.25(0.8536)	33.68(0.9445)
	LAGNet <sup>[13]</sup>	456	75	34.26(0.9253)	30.22(0.8421)	28.93(0.8024)		
	LMDFFN <sup>[14]</sup>	341		34.32(0.9264)	30.20(0.8392)	29.03(0.8034)	28.01(0.8483)	33.36(0.9430)
	RCCN <sup>[9]</sup>			34.48(0.9278)	30.37(0.8417)	29.12(0.8054)	28.34(0.8549)	
	ShuffleMixer <sup>[15]</sup>	415	43	34.40(0.9272)	30.37(0.8423)	29.12(0.8051)	28.08(0.8498)	33.69(0.9448)
	DLGSANet <sup>[17]</sup>	4740	486	<i>34.95(0.9310)</i>	<i>30.77(0.8501)</i>	<i>29.38(0.8121)</i>	<b>29.43(0.8761)</b>	<b>34.76(0.9517)</b>
HIAAN <sup>[18]</sup>	640	65	34.44(0.9278)	30.30(0.8421)	29.13(0.8074)	28.27(0.8544)	33.64(0.9453)	
EGAN(Ours)	310	9	<b>34.97(0.9396)</b>	<b>31.01(0.8577)</b>	<b>29.60(0.8190)</b>	<i>29.21(0.8713)</i>	<i>33.81(0.9481)</i>	
$\times 4$	Bicubic			28.42(0.8104)	26.00(0.7027)	25.96(0.6675)	23.14(0.6577)	24.89(0.7866)
	VDSR <sup>[4]</sup>	665	613	31.35(0.8838)	28.01(0.7674)	27.29(0.7251)	25.18(0.7524)	28.83(0.8809)
	IDN <sup>[10]</sup>	553	32	31.82(0.8903)	28.25(0.7730)	27.41(0.7297)	25.41(0.7632)	29.41(0.8942)
	CARN <sup>[11]</sup>	1592	91	32.13(0.8937)	28.60(0.7806)	27.58(0.7349)	26.07(0.7837)	30.47(0.9084)
	IMDN <sup>[12]</sup>	715	41	32.11(0.8934)	28.63(0.7823)	27.58(0.7358)	26.10(0.7846)	30.55(0.9072)
	SMSR <sup>[16]</sup>	1006	42	32.12(0.8932)	28.55(0.7808)	27.55(0.7351)	26.11(0.7868)	30.54(0.9085)
	LAGNet <sup>[13]</sup>	470	68	32.06(0.8912)	28.47(0.7782)	27.54(0.7347)		
	LMDFFN <sup>[14]</sup>	344		32.08(0.8930)	28.46(0.7792)	27.51(0.7341)	25.93(0.7804)	30.25(0.9053)
	RCCN <sup>[9]</sup>			32.24(0.8956)	28.69(0.7833)	27.63(0.7376)	<i>26.31(0.7912)</i>	
	ShuffleMixer <sup>[15]</sup>	411	28	32.21(0.8953)	28.66(0.7827)	27.61(0.7366)	26.08(0.7835)	30.65(0.9093)
	DLGSANet <sup>[17]</sup>	4760	274	<b>32.80(0.9021)</b>	<i>28.95(0.7907)</i>	<i>27.85(0.7464)</i>	<b>27.17(0.8175)</b>	<b>31.68(0.9219)</b>
HIAAN <sup>[18]</sup>	652	37	32.26(0.8964)	28.62(0.7832)	27.59(0.7386)	26.17(0.7886)	30.54(0.9095)	
EGAN(Ours)	312	8	<i>32.35(0.8965)</i>	<b>28.97(0.7910)</b>	<b>28.47(0.7642)</b>	26.04(0.7837)	<i>30.74(0.9145)</i>	

现,特别在 Set14 上相较于 DLGSANet, PSNR 和 SSIM 分别提高了 0.24 dB 和 0.0076;当放大因子为 4 时,所提方法与 DLGSANet 相比,在 BSD100 测试集上 PSNR 和 SSIM 分别提高了 0.62 dB 和 0.0178。同时将所提方法类比实现高效化的层次信息自适应聚合网络(HIAAN),EGAN 在所有放大因子及 5 个基准测试集上分别取差值并求平均,得到 PSNR 和 SSIM 平均提高 0.51 dB 和 0.0078,在参数量和 FLOPs 上平均降低  $332 \times 10^3$  和  $70 \times 10^9$ 。其次,将所提方法与各先进方法进行对比,EGAN 在较少的网络参数量和较低的计算成本下取得了优秀的性能,如图 5 所示。综上所述,EGAN 不仅能在不同尺度下获得较大优势,还具有更少的网络参数量和更低的 FLOPs,从而实现网络轻量化与性能之间的最优平衡,达到网络高效的同时,也充分证明了所提方法的先进性与有效性。

为了更加直观地感受所提改进方法的重建图像质量,从 Urban100 测试集中选取部分图像,与其他先进算法的实验结果作局部放大对比,包括 Bicubic、VDSR、IDN、SMSR、IMDN、ShuffleMixer,如图 6 所示。对于“img092”局部放大图像,IMDN、ShuffleMixer 等在方框中错误预测了建筑物的线条方向,而所提方法恢复的图像块更接近 HR 图像。对于“img093”局部放大图像,其他方法在方框中均无法准确预测地面线条

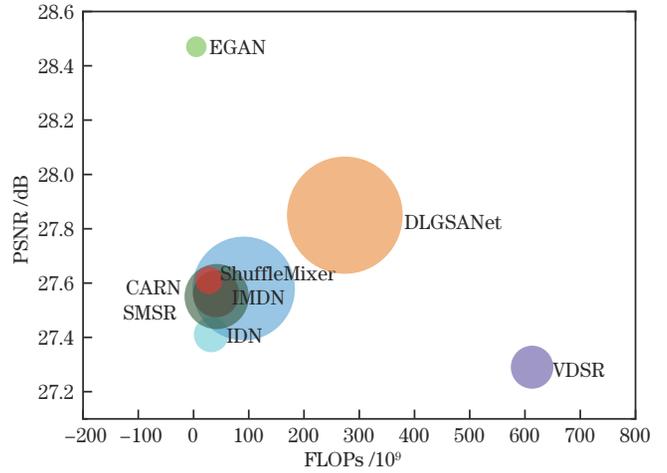


图 5 基于  $\times 4SR$ ,所提 EGAN 与其他方法在 BSD100 数据集上的模型复杂度和性能比较。圆圈大小表示参数量  
Fig. 5 Comparison of model complexity and performance of EGAN proposed in this study with other methods on the BSD100 dataset based on  $\times 4SR$ . The size of the circle indicates the number of parameters

的方向,且出现了一定程度的模糊和失真现象,相比之下,EGAN 恢复的图像更为精准且几乎没有伪影产生,显著优于其他方法,取得了最佳的视觉效果,这充分证明了所提方法的有效性。

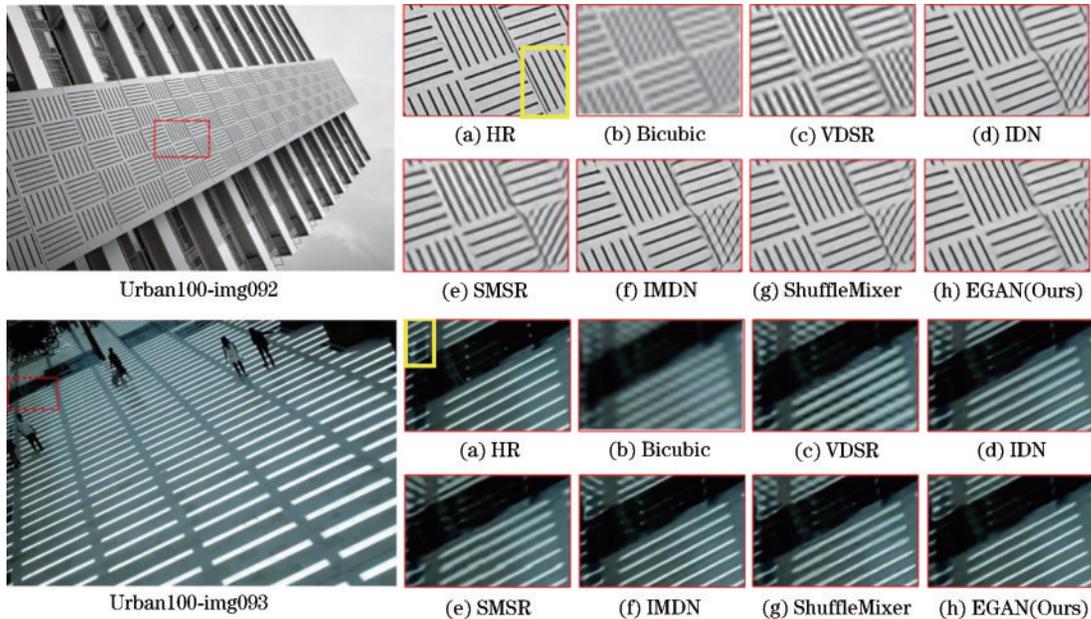


图 6 与先进算法的可视化结果对比

Fig. 6 Comparison of visualization results with advanced algorithms

#### 4.4 消融实验

为进一步研究网络中不同组件对模型性能的影响,对 EGAN 的不同组件分别进行消融实验,以证明组件的有效性,具体包括 CAFB、NPRB 及多阶段动态余弦热重启训练策略。为了进行公平比较,使用包含丰富结构信息的 Urban100 作为测试数据集并在 4 倍

放大因子的图像上进行实验。

1) CAFB。为了证明所提 CAFB 的有效性,对所提方法与两个基线分别进行比较,具体包含 CCA 和 AWA,结果如表 4 所示。可以看到,仅使用 AWA 得到的 PSNR 值为 25.89 dB,引入 CCA 得到的 PSNR 值为 25.96 dB,且这两个基线模型的参数量在增加不多的

表4 CAFB中不同组件对模型性能的影响  
Table 4 Effect of different components in CAFB on model performance

CCA <sup>[24]</sup>	AWA	Number of parameters /10 <sup>3</sup>	PSNR / dB(SSIM)
	✓	<b>289</b>	25.89(0.7831)
✓		305	25.96(0.7834)
✓	✓	312	<b>26.04(0.7837)</b>

情况下,PSNR值均低于CAFB,从而证明了同时使用CCA和AWA的CAFB提高图像重建质量的有效性。此外,还对比了CAFB在整体网络中堆叠数量的消融实验,如表5所示。CAFB数量为12时,模型的参数量在增加不多的同时能获得更优的性能指标。

表5 不同数量的CAFB对模型性能的影响  
Table 5 Effect of different numbers of CAFBs on model performance

Number of CAFBs	Number of parameters /10 <sup>3</sup>	PSNR /dB(SSIM)
4	<b>311</b>	25.43(0.7691)
6	<b>311</b>	25.73(0.7745)
8	312	25.69(0.7762)
10	312	26.02(0.7820)
12	312	<b>26.04(0.7837)</b>
14	312	25.62(0.7723)

表7 热重启策略对模型性能的影响  
Table 7 Effect of thermal restart strategy on model performance

Method	PSNR /dB(SSIM)				
	Set5	Set14	BSD100	Urban100	Manga109
EGAN-Lr	31.91(0.8913)	28.41(0.7778)	28.33(0.7604)	25.73(0.7726)	30.26(0.9088)
EGAN-DCwr	<b>32.35(0.8965)</b>	<b>28.97(0.7910)</b>	<b>28.47(0.7642)</b>	<b>26.04(0.7837)</b>	<b>30.74(0.9145)</b>

表8 热重启策略对IMDN性能的影响  
Table 8 Effect of the thermal restart strategy on IMDN performance

Method	PSNR /dB(SSIM)				
	Set5	Set14	BSD100	Urban100	Manga109
IMDN <sup>[12]</sup>	32.21(0.8948)	28.58(0.7811)	27.56(0.7353)	26.04(0.7838)	30.45(0.9075)
IMDN-DCwr	<b>32.23(0.8948)</b>	<b>28.58(0.7813)</b>	<b>27.56(0.7354)</b>	<b>26.08(0.7859)</b>	<b>30.47(0.9077)</b>

## 5 结 论

提出了一种新颖的高效全局注意网络。通过设计交叉自适应特征块,网络能够有效地融合并利用局部与全局注意信息。同时,所提近邻像素重构块能够减少模型参数量并有效恢复高质量图像信息,进一步增强了模型的重建能力。此外,还提出了一种多阶段动态余弦热重启训练策略,通过在不同阶段适应性调整网络的学习率,从而实现对训练过程的细粒度控制,保证轻量化训练的同时增强网络性能。大量实验结果表

2) NPRB。为了证明所提NPRB的有效性,首先利用SR网络中常用的亚像素卷积(Pixel Shuffle)进行HR图像重建,并使用与该网络相同的设置来训练NPRB以进行公平比较。表6为NPRB与Pixel Shuffle进行上采样的结果比较。所提方法的PSNR指标获得了1.35 dB的增益,由此证明了NPRB对HR图像重建的有效性。

表6 NPRB对HR图像重建性能的影响  
Table 6 Effect of NPRB on HR image reconstruction performance

NPRB	Pixel Shuffle	Number of parameters /10 <sup>3</sup>	PSNR /dB (SSIM)
	✓	<b>304</b>	24.69(0.6868)
✓		312	<b>26.04(0.7837)</b>

3) 多阶段动态余弦热重启训练策略。为了证明所提热重启策略的有效性,以EGAN-DCwr为基线,将不同学习策略的变体作为对照组,即EGAN-Lr。总轮数设置为500,学习率设置为0.01,当训练达到300轮时,学习速率减半,实验结果如表7所示。与所提热重启策略相比,EGAN-Lr的使用降低了网络性能。这表明,热重启策略的应用有助于网络在优化过程中跳出局部最小值,提高整体性能。同时,也探究了所提热重启策略的泛化性,即对IMDN采用热重启策略,对比结果如表8所示,实验结果表明,所提热重启策略可以应用于其他现有的SR模型。

明,所提方法,包括整体结构和训练策略,实现了性能和轻量之间的良好平衡,从而实现网络高效化。

## 参 考 文 献

- [1] Zhang L, Wu X L. An edge-guided image interpolation algorithm via directional filtering and data fusion[J]. IEEE Transactions on Image Processing, 2006, 15(8): 2226-2238.
- [2] Rhee S, Kang M G. Discrete cosine transform based regularized high-resolution image reconstruction algorithm [J]. Optical Engineering, 1999, 38(8): 1348-1356.

- [3] Dong C, Loy C C, He K M, 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.
- [4] 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.
- [5] 徐欣宜, 邓慧萍, 向森, 等. 基于特征交互融合与注意力的光场图像超分辨率[J]. 激光与光电子学进展, 2023, 60(14): 1410017.  
Xu X Y, Deng H P, Xiang S, et al. Light field image super-resolution based on feature interaction fusion and attention mechanism[J]. Laser & Optoelectronics Progress, 2023, 60(14): 1410017.
- [6] Dai T, Cai J R, Zhang Y B, et al. Second-order attention network for single image super-resolution[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2020: 11057-11066.
- [7] Niu B, Wen W L, Ren W Q, et al. Single image super-resolution via a holistic attention network[M]// Vedaldi A, Bischof H, Brox T, et al. Computer vision-ECCV 2020. Lecture notes in computer science. Cham: Springer, 2020, 12357: 191-207.
- [8] Mei Y Q, Fan Y C, Zhou Y Q. Image super-resolution with non-local sparse attention[C]// 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 3516-3525.
- [9] Li X. Image super-resolution reconstruction based on residual compensation combined attention network[J]. Journal of Electronics and Information Science, 2023, 8(1): 48-55.
- [10] Hui Z, Wang X M, Gao X B. Fast and accurate single image super-resolution via information distillation network[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: 723-731.
- [11] 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.
- [12] 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.
- [13] 丁子轩, 张娟, 李想, 等. 基于注意力引导的轻量级图像超分辨率网络[J]. 激光与光电子学进展, 2023, 60(14): 1410010.  
Ding Z X, Zhang J, Li X, et al. Lightweight attention-guided network for image super-resolution[J]. Laser & Optoelectronics Progress, 2023, 60(14): 1410010.
- [14] Guo X X, Tu Z C, Li G Y, et al. A novel lightweight multi-dimension feature fusion network for single-image super-resolution reconstruction[J]. The Visual Computer, 2023: 1-12.
- [15] Sun L, Pan J S, Tang J H. ShuffleMixer: an efficient ConvNet for image super-resolution[EB/OL]. (2022-05-30)[2023-05-04]. <https://arxiv.org/abs/2205.15175>.
- [16] 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.
- [17] Li X, Pan J S, Tang J H, et al. DLGSANet: lightweight dynamic local and global self-attention networks for image super-resolution[EB/OL]. (2023-01-05)[2023-05-04]. <https://arxiv.org/abs/2301.02031>.
- [18] 陈伟杰, 黄国恒, 莫非, 等. 层次信息自适应聚合的图像超分辨率重建算法[J]. 计算机工程与应用, 2024, 60(5): 221-231.  
Chen W J, Huang G H, Mo F, et al. Image super-resolution reconstruction algorithm based on hierarchical information adaptive aggregation[J]. Computer Engineering and Applications, 2024, 60(5): 221-231.
- [19] Zhang X D, Zeng H, Zhang L. Edge-oriented convolution block for real-time super resolution on mobile devices[C]// Proceedings of the 29th ACM International Conference on Multimedia, October 20-24, 2021, Virtual Event, China. New York: ACM Press, 2021: 4034-4043.
- [20] Lin Z D, Garg P, Banerjee A, et al. Revisiting RCAN: improved training for image super-resolution[EB/OL]. (2022-01-27)[2023-05-06]. <https://arxiv.org/abs/2201.11279>.
- [21] Li Z Y, Liu Y Q, Chen X Y, et al. Blueprint separable residual network for efficient image super-resolution[C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 19-20, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 832-842.
- [22] Li Y W, Zhang K, Timofte R, et al. NTIRE 2022 challenge on efficient super-resolution: methods and results[C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 19-20, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 1061-1101.
- [23] Lai W S, Huang J B, Ahuja N, et al. Fast and accurate image super-resolution with deep Laplacian pyramid networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(11): 2599-2613.
- [24] Huang Z L, Wang X G, Huang L C, et al. CCNet: criss-cross attention for semantic segmentation[C]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, the Public of Korea. New York: IEEE Press, 2020: 603-612.
- [25] Zhao H Y, Kong X T, He J W, et al. Efficient image

- super-resolution using pixel attention[M]//Bartoli A, Fusiello A. Computer vision-ECCV 2020 workshops. Lecture notes in computer science. Cham: Springer, 2020, 12537: 56-72.
- [26] Kong F Y, Li M X, Liu S W, et al. Residual local feature network for efficient super-resolution[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 19-20, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 765-775.
- [27] 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.
- [28] Zhong M Y, Jiang L. Review of super-resolution image reconstruction algorithms[J]. Journal of Frontiers of Computer Science and Technology, 2022, 16(5): 972-990.
- [29] 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, UK. London: British Machine Vision Association, 2012: 1-10.
- [30] Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations[M]//Boissonnat J D, Chenin P, Cohen A, et al. Curves and surfaces. Lecture notes in computer science. Heidelberg: Springer, 2012, 6920: 711-730.
- [31] 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. ICCV, July 7-14, 2001, Vancouver, BC, Canada. New York: IEEE Press, 2002: 416-423.
- [32] 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.
- [33] 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.
- [34] 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.
- [35] Liu J, Chen C, Tang J, et al. From coarse to fine: hierarchical pixel integration for lightweight image super-resolution[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(2): 1666-1674.
- [36] Zhang X D, Zeng H, Guo S, et al. Efficient long-range attention network for image super-resolution[M]//Avidan S, Brostow G, Cissé M, et al. Computer vision-ECCV 2022. Lecture notes in computer science. Cham: Springer, 2022, 13677: 649-667.
- [37] 彭晏飞, 张曼婷, 张平甲, 等. 聚合残差注意力网络的单图像超分辨率重建[J]. 激光与光电子学进展, 2023, 60(10): 1010017.
- Peng Y F, Zhang M T, Zhang P J, et al. Single-image super-resolution reconstruction aggregating residual attention network[J]. Laser & Optoelectronics Progress, 2023, 60(10): 1010017.
- [38] Chen L Y, Chu X J, Zhang X Y, et al. Simple baselines for image restoration[M]//Avidan S, Brostow G, Cissé M, et al. Computer vision-ECCV 2022. Lecture notes in computer science. Cham: Springer, 2022, 13667: 17-33.