

基于注意力引导的轻量级图像超分辨率网络

丁子轩, 张娟*, 李想, 王新宇

上海工程技术大学电子电气工程学院, 上海 201620

摘要 针对图像超分辨率重建网络因结构冗杂、参数量增加导致的计算量过大、训练时间过长等问题, 提出一种注意力引导的轻量级图像超分辨率网络(LAGNet)。LAGNet将随机初始化的自适应权重引入深度残差网络结构, 更大幅度地利用浅层特征信息。其次, 提出注意力引导(AG)模块, 该模块使用高效通道注意力(ECA)模块和空间分组增强(SGE)模块双支路并联结构, 结合通道间关系和空间位置信息特征, 利用注意力引导层动态调整两个分支的权重占比, 准确获取高频特征信息。最后, 使用全局级联连接, 减少网络参数量并加快信息流通速度。使用L1损失函数, 在加快收敛速度的同时防止梯度爆炸。在三个基准数据集上的测试结果表明: 相比其他网络, LAGNet的峰值信噪比平均提高0.39 dB, 模型参数量平均减少24%, 加法操作量和乘法操作量平均减少62%; 在图像视觉效果上整体更为清晰, 细节纹理更自然。

关键词 卷积神经网络; 超分辨率; 注意力机制; 数字图像处理; 计算机视觉

中图分类号 TP391.41

文献标志码 A

DOI: 10.3788/LOP221947

Lightweight Attention-Guided Network for Image Super-Resolution

Ding Zixuan, Zhang Juan*, Li Xiang, Wang Xinyu

College of Electronic and Electrical Engineering, Shanghai University of Engineering Science,
Shanghai 201620, China

Abstract A lightweight attention-guided super-resolution network (LAGNet) is proposed to address issues such as excessive computation and long training time caused by the redundant structure and increased parameters of image super-resolution reconstruction networks. First, the LAGNet introduces randomly initialized adaptive weights into the deep residual network structure to maximize the use of shallow feature information. Second, an attention guidance (AG) module uses the parallel structure of the efficient channel attention (ECA) model and the spatial group-wise enhance (SGE) model, combines the relationship between channels and the spatial location information characteristics, and employs the attention-guide layer to dynamically adjust the weight proportion of the two branches to obtain high-efficiency channel feature information. Finally, the global cascade connection is used to reduce network parameters and speed up information flow. The L1 loss function is used to accelerate convergence speed and prevent gradient explosion. The test results on the three benchmark datasets show that on average the peak signal-to-noise ratio of the LAGNet is increased by 0.39 dB, the model parameters are reduced by 24%, and the addition and multiplication operations are reduced by 62% compared with other networks; the overall visual effect of the image is clear and the detail texture is more natural.

Key words convolutional neural network; super-resolution; attention mechanism; digital image processing; computer vision

1 引言

单幅图像超分辨率(SISR)算法是一种将观测到的低分辨率(LR)图像重建为相应的高分辨率(HR)图像, 使图像具有边缘细节和低频纹理细节的方法。随着SISR技术在医学成像^[1]、卫星成像、安全与监控等

诸多领域显示出巨大潜力和实际应用价值, 越来越多的研究者开始关注与研究这一问题。

传统SISR算法, 例如基于插值的方法, 包括最近邻插值法、双线性插值法、双三次插值法等, 通过填补图像放大后像素的空缺位置恢复图像内容, 达到提高图像分辨率的效果。但是使用这种方法后会得到边缘

收稿日期: 2022-06-29; 修回日期: 2022-08-06; 录用日期: 2022-09-05; 网络首发日期: 2022-09-15

基金项目: 上海市地方院校能力建设项目(21010501500)

通信作者: *zhang-j@foxmail.com

过于平滑的模糊图像,这一点在图像高频信息较多的区域尤为明显。为了解决基于插值的方法的缺点,研究者提出了基于学习的方法,包括基于字典学习的方法^[2]、基于局部线性回归的方法^[3]和基于随机森林的方法^[4]。通过学习从 LR 图像到 HR 图像的映射关系,基于学习的超分辨率方法取得了比插值法更优的结果,获得了丰富的图像高频信息,但是这些方法需要对训练样本施加更多的限制,并且对图像中的噪声极为敏感。

深度学习是当今发展最快的机器学习方法之一^[5]。基于深度学习的卷积神经网络的出现为解决超分辨率问题提供了新思路,许多研究者运用卷积神经网络,得到了优于传统 SISR 算法的结果。Dong 等^[6]提出了基于卷积神经网络的图像超分辨率网络 (SRCNN),此网络是基于深度学习的单幅图像超分辨率网络的开山之作。SRCNN 首先使用双三次插值法将图像放大至目标尺寸,从中提取特征,然后用三层卷积神经网络拟合 LR 图像和 HR 图像之间的非线性映射,最后输出重建后的 HR 图像。尽管 SRCNN 结果优于传统图像超分辨率算法,但过少的网络层数限制其性能的进一步提升。因此,Dong 等^[7]对 SRCNN 进行改进,提出了快速 SRCNN (FSRCNN),得到了优于 SRCNN 的结果。FSRCNN 优势包含 3 点:一是在最后使用反卷积层放大尺寸,因此可以直接将原始的 LR 图像输入到网络中;二是使用更小的卷积核和更多的映射层,获取更丰富的特征信息;三是共享网络中的映射层,提升训练不同放大倍数模型时的效率。Lai 等^[8]提出了拉普拉斯金字塔超分辨率网络 (LapSRN),其结构类似于金字塔,对输入的 LR 图像每级进行一次 2 倍上采样操作,在进行高倍上采样的同时也能得到中间倍和低倍上采样的结果,提升了网络的整体效率,并且 LapSRN 在 8 倍放大倍数时有较优秀的表现。

为了进一步提升图像超分辨率的性能, Kim 等^[9]提出了深度递归神经网络 (DRCN), DRCN 利用跳跃连接加深了结构,通过增加网络感受野大小提升了整体性能。Kim 等^[10]发现低分辨率图像和高分辨率图像具有相关性,通过学习它们之间的残差信息能有效解决超分辨率问题,基于此原理,提出了深层卷积超分辨率网络 (VDSR)。此网络利用深度神经网络和残差学习解决单图像超分辨率问题,并且使用自适应梯度裁剪来加速模型的训练。Tai 等^[11]在 DRCN 的基础上利用参数共享原理设计了深度递归残差网络 (DRRN),增加了由多个残差单元组成的局部残差递归块,通过在残差单元中共享参数,减少经过深层网络后图像细节的丢失,得到了视觉效果更优的超分辨率图像。

上述网络通过不断堆叠网络层数、增加网络参数的方法实现了更好的超分辨率性能,然而过深的网络结构也限制了网络在实际应用中的潜力,存在计算设备要求过高、网络训练时间过长等问题,然而减少网

络参数量和提升网络整体性能很难同时实现。为了解决这些问题,本文提出了一种基于注意力引导的轻量级图像超分辨率算法 (LAGNet),在精简网络结构的同时获得了极具竞争力的超分辨率结果,并且在重建图像质量和网络参数数量之间取得了良好的平衡。LAGNet 的峰值信噪比 (PSNR) 相较于其他优秀算法有显著提升,在参数量 (Parameters) 和乘法操作量和加法操作量 (Multi-Adds) 上也有良好表现。

本文主要工作如下:提出了一种基于注意力引导的轻量级图像超分辨率网络 (LAGNet);设计并提出了注意力引导 (AG) 模块,该模块使用高效通道注意力 (ECA) 模块和空间分组增强 (SGE) 模块双支路并联结构,并利用注意力引导层动态调整两个支路的权重比例,将学习到的通道间关系和空间位置信息特征相结合,快速获得原始图像中的高频特征信息;在 AG 模块间使用全局级联连接,精简网络结构的同时加快信息流通速度,进一步减少网络参数量。

2 相关工作

2.1 轻量级超分辨率网络

自 VDSR^[10]将深度残差网络结构引入超分辨率问题后,堆叠网络层数和增加残差模块数量便成为这一领域的研究趋势。基于此策略提出的网络,如增强的深度残差网络 (EDSR)^[12]将残差模块数量增加到 32 个,深度残差注意力网络 (RCAN)^[13]将网络层数堆叠到 400 层,均获得了较好的超分辨率结果。随着超分辨率网络结构的冗余和参数量的激增,构建轻量级超分辨率网络模型、提升图像超分辨率算法的实际应用价值成为一个亟待解决的问题。

Ahn 等^[14]提出了一种轻量级且高效的级联残差网络 (CARN),通过在卷积组之间加入多个残差连接,减小网络计算的复杂度,但是实验结果表明,直接将残差连接应用到超分辨率算法中的效果还有待提高。Li 等^[15]提出的极小型图像超分辨率网络 (s-LWSR) 通过删除部分激活层压缩模块参数量,实现了极少的参数量,代价是超分辨率图像细节部分失真。Tian 等^[16]提出了一种轻量级图像超分辨网络 (LESRCNN),该网络采用由 1×1 和 3×3 卷积组成的异构结构减少低频信息冗余,通过重建模块将低频特征转换为高频特征,进一步降低网络的参数量。Chu 等^[17]提出了神经网络搜索架构,并基于此架构提出了一种基于神经网络搜索架构的轻量级超分辨率网络 (FALSr),利用神经网络搜索架构代替人工调整参数,实现参数量和模型复杂度之间的平衡。Tian 等^[18]提出了基于 CNN 的非对称图像超分辨率网络 (ACNet),通过由 3×1 、 1×3 和 3×3 卷积组成的非对称卷积结构突显局部关键特征,减少信息冗余的同时加快训练速度。

2.2 注意力机制

注意力机制最开始应用于机器翻译领域,随后扩

展到语义分割、目标检测^[19]、图像分类^[20]等领域。由于注意力机制在减少参数量和提高网络效率方面表现优秀,在计算机视觉领域也有大量的应用。压缩与激励网络(SENet)^[21]中首次提出了通道注意力选择机制,通过选择通道并赋予通道不同权重提取特征图的关键信息。Wang 等^[22]在 SENet 的基础上提出了高效通道注意力网络(ECANet),设计了不降低维度的局部交叉通道交互策略,该策略使用通道维数函数,自适应地确定一维卷积核大小,并确定局域交叉通道相互作用的覆盖范围。Li 等^[23]提出了空间分组增强(SGE)网络,该网络对通道进行分组,为每个通道组生成不同的特征因子来调整通道组的重要性,从而使每个组可以有针对地增强学习能力,并抑制可能出现的噪声。以上研究均表明,注意力机制在减少网络参数量、加快网络运行效率和提升网络准确率方面具有良好的效果。

3 图像重建算法框架

3.1 LAGNet 重建算法结构

LAGNet 结构如图 1 所示,由浅层特征提取模块、

非线性特征映射模块和重建模块组成。 I^{LR} 和 I^{SR} 分别代表网络的输入和输出。为减少网络的参数量,首先通过大小为 3×3 的卷积层从输入 I^{LR} 中提取浅层特征 F_{SF} ,表达式为

$$F_{SF} = C_{3 \times 3}(I^{LR}). \quad (1)$$

随后浅层特征 F_{SF} 经过非线性特征映射模块,完成信息提取,得到图像深层特征 F_{DF} ,表达式为

$$F_{DF} = H_{AG}^n \left\{ H_{AG}^{n-1} \left[\dots H_{AG}^1(F_{SF}) + \mu_x^1 F_{SF} \right] + \mu_x^{n-1} F_{SF} \right\} + \mu_x^n F_{SF}, \quad (2)$$

式中: H_{AG}^n 为第 n 个注意力引导模块,该模块由 AG 模块串联得到; μ_x^n 是随机初始化的自适应权重,可以更大限度地利用浅层特征信息,从而在特征融合阶段提供丰富的信息流。AG 模块间使用全局级联连接,可以在层之间快速传递特征信息,提高网络整体效率。

最后分别将所得图像深层特征 F_{DF} 和低分辨率图像 I^{LR} 输入重建模块,通过 3×3 卷积层后利用亚像素卷积层 f_p 进行上采样操作,加和得到重建后的高分辨率图像 I^{SR} ,表达式为

$$I^{SR} = f_p[C_{3 \times 3}(I^{LR})] + f_p[C_{3 \times 3}(F_{DF})]. \quad (3)$$

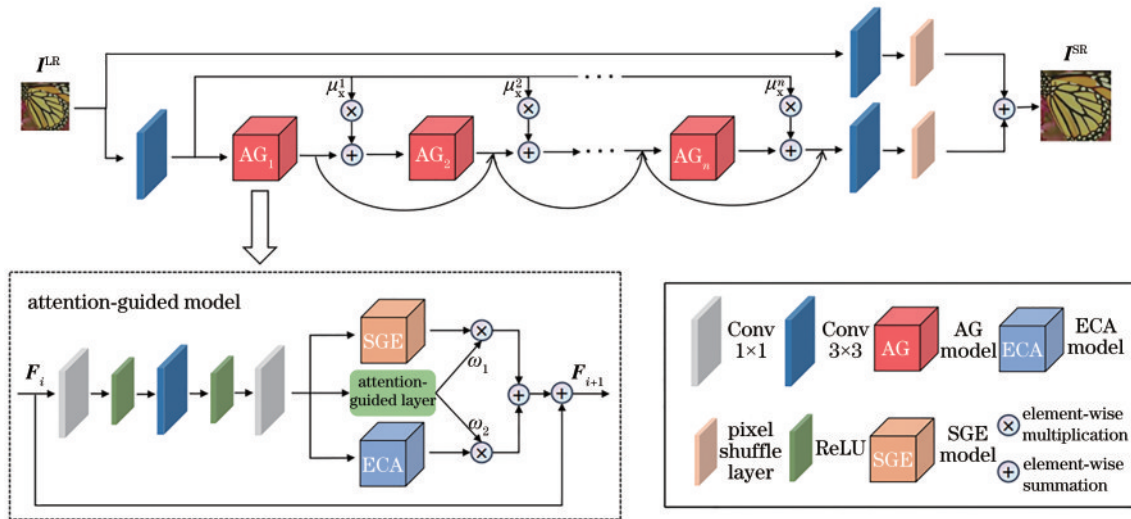


图 1 LAGNet 的结构
Fig. 1 LAGNet structure

3.2 注意力引导模块

AG 模块主要目的是减少网络的参数量和计算时间,具体细节于图 1 左下方虚线框内展示,在经过第 n 个 AG 模块时,输入特征 F_i 首先经过由 2 个 1×1 卷积层、1 个 3×3 卷积层、2 个 ReLU 函数组成的预处理模块,得到预处理结果 F_p ,表达式为

$$F_p = C_{1 \times 1} \left\{ \delta \left[C_{3 \times 3} \left\{ \delta \left[C_{1 \times 1} (F_i) \right] \right\} \right] \right\}, \quad (4)$$

式中: δ 为 ReLU 函数。

并非所有注意力机制都可以提高网络性能,网络中无效且冗余的参数依然会阻碍网络性能的进一步提升,因此提出注意力引导(atg)层^[24]。atg 层能够自动

舍弃部分不重要的注意力特征,动态调整两个模块的权重占比,从而提高网络的表征能力和泛化能力。

atg 层的结构如图 2 所示。预处理结果 F_p 首先通过全局池化操作增加感受野,随后依次通过全连接层、ReLU 函数和全连接层获取图像的特征信息,最后通

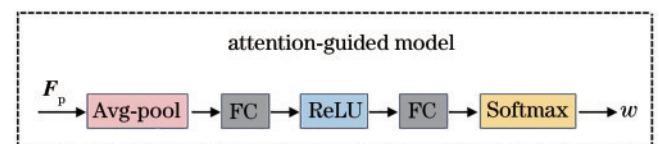


图 2 注意力引导层结构
Fig. 2 Attention-guided layer structure

过 Softmax 函数生成不同模块的动态权重 ω_i , 公式为

$$\omega_i = f_{\text{atg}}(\mathbf{F}_p), \quad i = 1, 2, \quad (5)$$

式中: f_{atg} 为注意力引导层。

将通过 atg 层得到的不同模块权重 ω_1 和 ω_2 分别乘以 ECA 模块和 SGE 模块得到的特征信息, 并进行加和, 得到更全面的有效特征, 最后与输入特征 \mathbf{F}_i 进行加和操作, 得到基于注意力引导的特征图 \mathbf{F}_{i+1} , 公式为

$$\mathbf{F}_{i+1} = \omega_1 \text{ECA}(\mathbf{F}_p) + \omega_2 \text{SGE}(\mathbf{F}_p) + \mathbf{F}_i. \quad (6)$$

3.2.1 高效通道注意力模块

压缩与激励(SE)模块中的两个全连接层参数冗余, 并且全连接层中的降维操作对通道注意力有负面影响。因此本文在 AG 模块中使用 ECA 模块, ECA 模块在 SE 模块的基础上使用无降维操作的局部跨通道交互策略。该策略通过自适应生成一维卷积核大小, 确定局部跨通道相互作用的覆盖范围, 在减少大量无效参数量和计算量的同时, 得到了更为出色的注意力特征提取结果。ECA 模块结构如图 3(a) 所示, 具体操作如下。

首先对输入的 \mathbf{E}_i 进行全局平均池化操作, 计算自适应的卷积核大小, 公式为

$$\mathbf{Z}_E^C = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W E_i(i, j), \quad (7)$$

式中: C, H 和 W 分别代表输入特征图的通道数、高度和宽度; \mathbf{Z}_E 是通道特征的总体统计信息; \mathbf{Z}_E^C 是通过全局平均池化操作获得的第 C 个特征通道的统计信息。

随后将自适应卷积应用于一维卷积中, 通过 Sigmoid 函数得到每个通道的注意力权重 \mathbf{W}_E :

$$\mathbf{W}_E = \sigma[C_{\text{ID}}(\mathbf{Z}_E^C)], \quad (8)$$

式中: σ 为 Sigmoid 函数。最后生成一个三维矩阵 \mathbf{E} 。(尺寸为 $C \times H \times W$) 作为注意力特征, 输出 \mathbf{E}_o 为输入 \mathbf{E}_i 与通道注意力的乘积, 公式为

$$\mathbf{E}_o = \mathbf{E}_i \cdot \mathbf{W}_E. \quad (9)$$

3.2.2 空间分组增强模块

SGE 模块对通道进行分组, 为每个通道组生成不同的特征因子来调整通道组的重要性, 从而有针对性地学习空间位置信息特征, 并抑制可能出现的噪声。特征因子仅由各组内部的全局和局部特征描述符之间的相似性引导, 因此 SGE 模块几乎没有引入额外的参数量。SGE 模块结构如图 3(b) 所示, 具体操作如下。

首先, 根据通道维度数量, 将特征图划分成 G 组。由于噪声和相似特征的存在, 特征的分布并无明显规律。因此, 利用全局信息进一步加强对关键区域的语义特征学习, 公式为

$$\mathbf{g} = \mathbf{f}_{\text{sp}} = \frac{1}{m} \sum_{h=1}^m \mathbf{x}_h. \quad (10)$$

其次, 利用全局特征 \mathbf{g} , 通过相关操作获得每个特征对应的注意力系数 \mathbf{c}_h :

$$\mathbf{c}_h = \mathbf{g} \cdot \mathbf{x}_h. \quad (11)$$

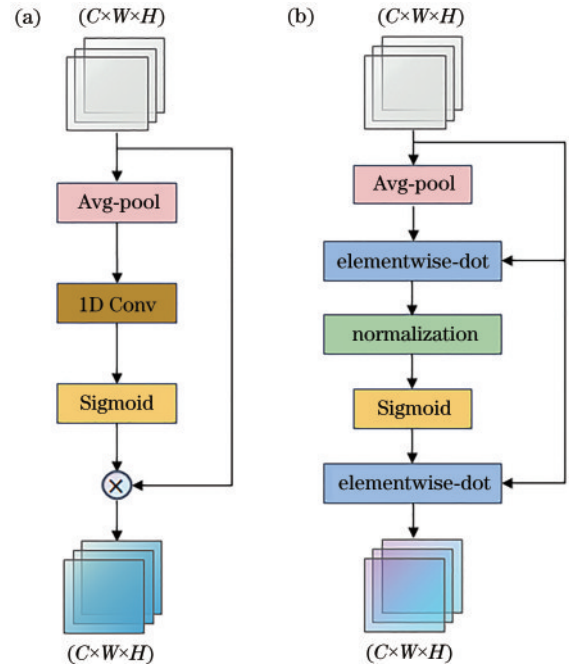


图 3 AG 模块的主要构成部分。(a)ECA 模块结构;(b)SGE 模块结构

Fig. 3 Main structure of AG module. (a) ECA module structure; (b) SGE module structure

为避免不同样本间系数偏置造成的影响, 对注意力系数 \mathbf{c}_h 通过归一化处理层进行归一化操作, 之后通过 Sigmoid 函数得到注意力权重 \mathbf{a}_h , 公式为

$$\mathbf{a}_h = \sigma[f_{\text{BN}}(\mathbf{c}_h)], \quad (12)$$

式中: f_{BN} 为归一化处理层。

最后, 对函数注意力权重 \mathbf{a}_h 与输入特征图 \mathbf{s}_i 进行相关操作, 得到特征图 \mathbf{s}_o :

$$\mathbf{s}_o = \mathbf{a}_h \cdot \mathbf{s}_i. \quad (13)$$

3.3 损失函数

L2 损失函数通常是图像超分辨率领域损失函数的首选, 但 L1 损失函数相较于 L2 损失函数具有更好的收敛性, 能提升训练网络的准确度, 因此实验使用 L1 损失函数训练网络。L1 损失函数计算 LR 图像和 HR 图像每个像素间插值的平均绝对值, 公式为

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \|H_{\text{LAGNet}}(\mathbf{I}_i^{\text{LR}}) - \mathbf{I}_i^{\text{HR}}\|_1. \quad (14)$$

4 实验结果与分析

4.1 实验环境

4.1.1 实验数据集

DIV2K 数据集^[25]由 800 张训练图像、100 张验证图像和 100 张测试图像组成, 实验使用其中的 800 张训练图像进行网络模型训练。为防止网络训练过程中出现过拟合情况, 在图像预处理过程中对训练图像进行数据增强操作, 对图像随机旋转 90°、180°、270°后水平翻转, 得到 3200 张图像的增强数据集。新数据集是基

于DIV2K数据集产生的,可以缩小训练集和验证集之间的差距,提取更多有效信息。使用3个标准基准数据集Set5^[26]、Set14^[27]、BSD100^[28]对网络进行测试,3个测试数据集中共有119张图像,包含不同场景下的人物、动植物、自然景观和人造物图像,种类繁多且部分图像细节模糊,重建难度较大。

4.1.2 实验平台与参数配置

硬件部分:处理器型号为8核16线程AMD 3700X,显卡型号为RTX 2070Super,电脑内存为16GB。软件部分:系统为Ubuntu18.04,采用Python 3.6版本编译程序,深度学习框架使用PyTorch 0.4.0版本。

实验在训练过程中的更新轮次设置为2000,训练批次量设置为16,输入和输出通道均设置为64,所有自适应权重的初始值均设置为1。在每个训练批次中,提取大小为 48×48 的LR图像作为输入,同时将相应的HR图像作为测试。训练时采用Adam优化器对网络的权重参数进行更新,更新过程中指数衰减速率设置为 $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$,初始学习速率设置为 10^{-3} ,并且每经过 2×10^5 次反向传播迭代后减半学习速率。

4.2 消融实验

4.2.1 实验评价指标

使用峰值信噪比(PSNR)和结构相似度(SSIM)两个指标对重建图像的亮度通道进行评估。PSNR为信号的最大功率与噪声功率之比,PSNR值越高,说明重建的图像失真度越低,公式为

$$X_{\text{PSNR}} = 10 \times \log_{10} \frac{M_i^2}{X_{\text{MSE}}}, \quad (15)$$

$$X_{\text{MSE}} = \frac{1}{HW} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} [I(i,j) - K(i,j)], \quad (16)$$

式中: X_{MSE} 为原始图像与处理后图像的均方误差; M_i 为图像颜色的最大值,采样的8位图像像素值范围为0到255; $I(i,j)$ 和 $K(i,j)$ 分别为不同图像对应坐标的像素值。SSIM是衡量两幅图像相似度的指标,SSIM值越接近1,说明重建图像和原始图像的相似度越高,公式为

$$X_{\text{SSIM}} = L(X, Y)C(X, Y)S(X, Y), \quad (17)$$

$$L(X, Y) = \frac{2u_x u_y + C_1}{u_x^2 + u_y^2 + C_1}, \quad (18)$$

式中: $L(X, Y)$ 为亮度的估计值; $C(X, Y)$ 为对比度的估计值; $S(X, Y)$ 为结构相似度的估计值; u_x 和 u_y 分别为图像在X维度和Y维度上的均值。

4.2.2 注意力引导模块对网络的影响

为研究AG模块在不同数据集和不同放大倍数条件下的普遍有效性,在其他变量保持不变的前提下,分别在Set5测试集2倍放大倍数、Set14测试集3倍放大倍数、BSD100测试集4倍放大倍数进行实验。对去除AG模块结构的模型和原始LAGNet网络的PSNR最大值进行了对比,如图4所示。

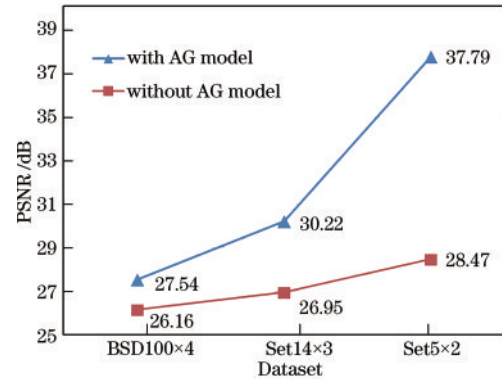


图4 AG模块结构对模型的影响

Fig. 4 Effect of AG module structure on the model

图4三角形折线为使用AG模块模型的PSNR数值变化折线,正方形折线为未使用AG模块模型的PSNR数值变化折线。虽然去除AG模块的模型收敛速度很快,但随着训练的进行出现了过拟合现象,而使用AG模块的模型则没有出现这种情况。实验结果表明,相较于不使用AG模块的网络,使用AG模块的网络在不同测试集和不同放大倍数的情况下的PSNR最大值均有显著提升,网络整体性能也有明显改善。实验过程中也同时记录了三种情况下SSIM的最大值,对比结果表明使用AG模块的网络在SSIM值上平均提升了0.0529。

4.2.3 注意力引导模块数量对网络的影响

为研究AG模块数量对LAGNet的影响,分别设置AG模块数量(N_{AG})为8、12、16、20、24,进行实验,在Set5、Set14、BSD100数据集3倍放大倍数的情况下进行测试。表1显示了AG模块数量对LAGNet网

表1 AG模块数量对网络的影响

Table 1 Influence of the number of AG modules on the network

| Parameter | Dataset | $N_{\text{AG}} = 8$ | $N_{\text{AG}} = 12$ | $N_{\text{AG}} = 16$ | $N_{\text{AG}} = 20$ | $N_{\text{AG}} = 24$ |
|---------------------|---------|---------------------|----------------------|----------------------|----------------------|----------------------|
| PSNR /dB | Set5 | 33.935 | 33.961 | 34.261 | 34.257 | 34.227 |
| | Set14 | 29.312 | 30.117 | 30.224 | 29.941 | 29.725 |
| | BSD100 | 28.452 | 28.625 | 28.933 | 28.561 | 28.027 |
| SSIM | Set5 | 0.9105 | 0.9187 | 0.9253 | 0.9253 | 0.9236 |
| | Set14 | 0.8216 | 0.8369 | 0.8421 | 0.8297 | 0.8262 |
| | BSD100 | 0.7914 | 0.7963 | 0.8024 | 0.7951 | 0.7855 |
| Parameters / 10^3 | | 376 | 422 | 456 | 541.2 | 577.3 |
| Multi-Adds / 10^9 | | 47.7 | 61.6 | 75.4 | 89.3 | 103.4 |

络 PSNR、SSIM 和参数数量的影响。可以看出当 AG 模块数量为 16 时, PSNR 与 SSIM 参数均为最高, 且模型参数数量相对较小, 因此实验设置 AG 模块数量为 16。

4.2.4 注意力引导模块结构对网络的影响

为研究 AG 模块结构对 LAGNet 网络的影响, 在

控制变量的前提下, 分别删去 SGE 和 ECA 模块, 替换为常用 SE 模块, 同时设置有无 atg 层的对照实验, 证明 atg 层在减少参数数量、提升网络整体性能方面的作用。在 Set5、Set14、BSD100 数据集放大倍数分别为 2、3、4 的情况下测试 PSNR、SSIM 和参数数量, 得到的结果如表 2 所示。

表 2 AG 模块结构对网络的影响
Table 2 Influence of AG module structure on the network

| AG module structure | Dataset | Scale | PSNR /dB | SSIM | Parameters /10 ³ | Multi-Adds /10 ⁹ |
|---------------------|---------|-------|----------|--------|-----------------------------|-----------------------------|
| ECA+SE | Set5 | 2 | 36.572 | 0.9538 | 429 | 68.5 |
| SGE+SE | Set5 | 2 | 36.937 | 0.9546 | 502 | 85.2 |
| ECA+SGE | Set5 | 2 | 37.584 | 0.9566 | 553 | 89.6 |
| ECA+SGE+atg | Set5 | 2 | 37.792 | 0.9594 | 447 | 83.6 |
| ECA+SE | Set14 | 3 | 29.786 | 0.8317 | 433 | 77.9 |
| SGE+SE | Set14 | 3 | 29.592 | 0.8295 | 505 | 93.3 |
| ECA+SGE | Set14 | 3 | 30.109 | 0.8403 | 526 | 86.1 |
| ECA+SGE+atg | Set14 | 3 | 30.224 | 0.8421 | 456 | 75.4 |
| ECA+SE | BSD100 | 4 | 27.267 | 0.7254 | 439 | 91.6 |
| SGE+SE | BSD100 | 4 | 27.294 | 0.7263 | 509 | 107.6 |
| ECA+SGE | BSD100 | 4 | 27.314 | 0.7262 | 506 | 84.2 |
| ECA+SGE+atg | BSD100 | 4 | 27.541 | 0.7347 | 470 | 67.9 |

4.3 实验结果分析

4.3.1 客观指标分析

对所提 LAGNet 网络模型与目前效果优异的方法进行对比, 其他方法包括 Bicubic、SRCNN^[6]、

FSRCNN^[7]、LapSRN^[8]、DRCN^[9]、VDSR^[10]、DRRN^[11]、CARN^[14]、s-LWSR^[15]、LESRCNN^[16]、FALS^[17]、ACNet^[18]。与上述的所有其他方法相比, LAGNet 网络的 PSNR 和 SSIM 有明显优势, 实验结果如表 3 所示。

表 3 不同方法在 Set5、Set14 和 BSD100 数据集上的测试结果
Table 3 Test results of different methods on Set5, Set14, and BSD100 datasets

| Method | Scale | Parameters /10 ³ | Multi-Adds /10 ⁹ | PSNR /dB | | | SSIM | | |
|----------------------|-------|-----------------------------|-----------------------------|----------|-------|--------|--------|--------|--------|
| | | | | Set5 | Set14 | BSD100 | Set5 | Set14 | BSD100 |
| Bicubic | | | | 33.66 | 30.24 | 29.56 | 0.9299 | 0.8688 | 0.8431 |
| SRCNN | | 57 | 52.7 | 36.34 | 32.45 | 31.36 | 0.9521 | 0.9067 | 0.8879 |
| FSRCNN | | 12 | 6.0 | 36.94 | 32.63 | 31.53 | 0.9558 | 0.9088 | 0.8920 |
| VDSR | | 665 | 612.6 | 37.53 | 33.03 | 31.90 | 0.9587 | 0.9124 | 0.8960 |
| LapSRN | | 813 | 29.9 | 37.52 | 33.08 | 31.80 | 0.9591 | 0.9130 | 0.8950 |
| DRCN | | 1774 | 17974 | 37.63 | 33.04 | 31.85 | 0.9588 | 0.9118 | 0.8942 |
| DRRN | 2 | 297 | 6796.9 | 37.74 | 33.23 | 32.05 | 0.9591 | 0.9136 | 0.8973 |
| CARN-M | | 412 | 91.2 | 37.53 | 33.26 | 31.92 | 0.9583 | 0.9141 | 0.8960 |
| LESRCNN | | 516 | 110 | 37.65 | 33.32 | 31.95 | 0.9586 | 0.9148 | 0.8965 |
| FALS ^R -B | | 326 | 74.7 | 37.61 | 33.29 | 31.97 | 0.9585 | 0.9143 | 0.8967 |
| FALS ^R -C | | 408 | 93.7 | 37.66 | 33.26 | 31.96 | 0.9586 | 0.9140 | 0.8965 |
| ACNet | | 1356 | 501.5 | 37.72 | 33.41 | 32.06 | 0.9588 | 0.9160 | 0.8978 |
| LAGNet | | 447 | 83.6 | 37.79 | 33.40 | 32.10 | 0.9594 | 0.9162 | 0.8991 |
| Bicubic | | | | 30.39 | 27.55 | 27.21 | 0.8682 | 0.7742 | 0.7385 |
| SRCNN | | 57 | 52.7 | 32.39 | 29.30 | 28.41 | 0.9033 | 0.8215 | 0.7863 |
| FSRCNN | | 12 | 5.0 | 33.16 | 29.43 | 28.53 | 0.9140 | 0.8242 | 0.7910 |
| VDSR | 3 | 665 | 612.6 | 33.66 | 29.77 | 28.82 | 0.9213 | 0.8314 | 0.7976 |
| DRCN | | 1774 | 17974 | 33.82 | 29.76 | 28.80 | 0.9226 | 0.8311 | 0.7963 |
| DRRN | | 297 | 6796.9 | 34.03 | 29.96 | 28.95 | 0.9244 | 0.8349 | 0.8004 |

表 3 (续)

| Method | Scale | Parameters /10 ³ | Multi-Adds /10 ⁹ | PSNR /dB | | | SSIM | | |
|----------------------|-------|-----------------------------|-----------------------------|----------|-------|--------|--------|--------|--------|
| | | | | Set5 | Set14 | BSD100 | Set5 | Set14 | BSD100 |
| CARN-M | | 412 | 46.1 | 33.99 | 30.08 | 28.91 | 0.9236 | 0.8367 | 0.8000 |
| LESRCNN | | 516 | 49.1 | 33.93 | 30.12 | 28.91 | 0.9231 | 0.8380 | 0.8005 |
| ACNet | | 1541 | 369 | 34.14 | 30.19 | 28.98 | 0.9247 | 0.8398 | 0.8023 |
| LAGNet | | 456 | 75.4 | 34.26 | 30.22 | 28.93 | 0.9253 | 0.8421 | 0.8024 |
| Bicubic | | | | 28.42 | 26.00 | 25.96 | 0.8104 | 0.7027 | 0.6675 |
| SRCNN | | 57 | 52.7 | 30.09 | 27.50 | 26.90 | 0.8530 | 0.7513 | 0.7101 |
| FSRCNN | | 12 | 4.6 | 30.71 | 27.59 | 26.98 | 0.8657 | 0.7535 | 0.7150 |
| VDSR | | 665 | 612.6 | 31.35 | 28.01 | 27.29 | 0.8838 | 0.7674 | 0.7251 |
| LapSRN | | 813 | 149.4 | 31.54 | 28.19 | 27.32 | 0.8850 | 0.7720 | 0.7270 |
| DRCN | 4 | 1774 | 17974 | 31.53 | 28.02 | 27.23 | 0.8854 | 0.7670 | 0.7233 |
| DRRN | | 297 | 6796.9 | 31.68 | 28.21 | 27.38 | 0.8888 | 0.7720 | 0.7284 |
| s-LWSR ₁₆ | | 144 | 8.3 | 31.63 | 27.92 | 27.35 | 0.8869 | 0.7701 | 0.7287 |
| CARN-M | | 412 | 32.5 | 31.92 | 28.42 | 27.44 | 0.8903 | 0.7762 | 0.7304 |
| LESRCNN | | 516 | 28.6 | 31.88 | 28.44 | 27.45 | 0.8903 | 0.7772 | 0.7313 |
| ACNet | | 1784 | 347.9 | 31.83 | 28.46 | 27.48 | 0.8903 | 0.7788 | 0.7326 |
| LAGNet | | 470 | 67.9 | 32.06 | 28.47 | 27.54 | 0.8912 | 0.7782 | 0.7347 |

对网络参数量和加法乘法操作量进行了整体评估,在 Set5 数据集 4 倍放大倍率的条件下,对所提 LAGNet 与 SRCNN^[6]、FSRCNN^[7]、LapSRN^[8]、DRCN^[9]、VDSR^[10]、DRRN^[11]、CARN-M^[14]、s-LWSR₁₆^[15]、LESRCNN^[16]、ACNet^[18] 共 10 种网络进行对比,证明了 LAGNet 网络在参数量和加法乘法操作量之间取得了平衡,并得到良好的实验结果,如图 5 所示。

4.3.2 主观指标分析

使用 BSD100 测试集图像中赛车背后文字进行视觉效果对比,在放大倍数为 2 的情况下,得到的结果如图 6 所示。可以看出前 4 张图像文字模糊不清,第 5 和第 6 张图细节缺失,后 3 张图细节部分存在过度平滑的情况,所提 LAGNet 对图像整体的恢复更自然,细节方面表现更清晰。

使用 Set5 测试集图像中小孩的编织帽细节部分进行视觉效果对比,在放大倍数为 4 的情况下,得到的结果如图 7 所示。可以看出前 6 张图细节模糊不清,后

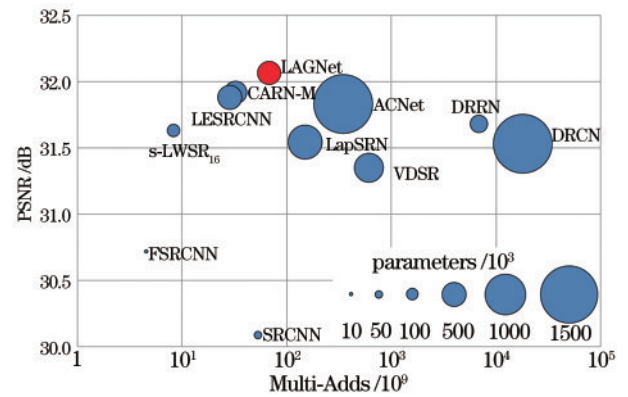


图 5 现有轻量级图像超分辨率算法在 Set5 数据集 4× 上的性能和参数量比较

Fig. 5 Performance and parameters comparison of existing lightweight image super-resolution methods on Set5 dataset with the magnification of 4×

3 者存在细节过度平滑的情况,所提 LAGNet 能恢复更多的纹理细节,整体观感也更符合实际。



图 6 不同方法对 BSD100 数据集中放大系数为 2 的图像的重建效果对比

Fig. 6 Comparison of reconstruction effects of different methods on images with a magnification factor of 2 in the BSD100 dataset



图 7 不同方法对 Set5 数据集中放大系数为 4 的图像的重建效果对比

Fig. 7 Comparison of reconstruction effects of different methods on images with magnification factor of 4 in Set5 dataset

5 结 论

针对图像超分辨率重建算法中网络结构冗余导致计算量过大、网络训练时间过长、计算设备要求过高等问题,提出了一种注意力引导的轻量级图像超分辨率算法。网络模型通过注意力引导机制减少了网络的参数量,通过使用随机初始化的自适应权重和全局级联连接高效利用浅层特征信息,精简网络结构的同时提高了网络整体性能。其中注意力引导模块的双支路并联结构将学习通道间的关系与空间位置信息特征相结合,快速获得原始图像的高频信息。实验结果表明,相较于其他轻量级网络算法,所提 LAGNet 重建算法在客观评价指标上能达到最优,在整体视觉效果上对纹理细节也有更好的表现能力。但在实际应用中往往追求更高的效率,能否在提高运行速度的同时提升模型性能仍需要研究,拟减小网络宽度进行实验,进一步精简网络结构,提升超分辨率重建算法的性能。

参 考 文 献

- [1] 胡芬,林洋,侯梦迪,等. 基于深度学习的细胞骨架图像超分辨重建[J]. 光学学报, 2020, 40(24): 2410001.
Hu F, Lin Y, Hou M D, et al. Super-resolution reconstruction of cytoskeleton image based on deep learning[J]. Acta Optica Sinica, 2020, 40(24): 2410001.
- [2] Yang J C, Wright J, Huang T S, et al. Image super-resolution via sparse representation[J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861-2873.
- [3] Timofte R, de Smet V, van Gool L. A+ : adjusted anchored neighborhood regression for fast super-resolution[M]//Cremers D, Reid I, Saito H, et al. Computer vision-ACCV 2014. Lecture notes in computer science. Cham: Springer, 2015, 9006: 111-126.
- [4] Schulter S, Leistner C, Bischof H. Fast and accurate image upscaling with super-resolution forests[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 3791-3799.
- [5] 王一同,周宏强,闫景道,等. 基于深度学习算法的计算光学研究进展[J]. 中国激光, 2021, 48(19): 1918004.
Wang Y T, Zhou H Q, Yan J X, et al. Advances in computational optics based on deep learning[J]. Chinese Journal of Lasers, 2021, 48(19): 1918004.
- [6] 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.
- [7] Dong C, Loy C C, Tang X O. Accelerating the super-resolution convolutional neural network[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9906: 391-407.
- [8] Lai W S, Huang J B, Ahuja N, et al. Deep Laplacian pyramid networks for fast and accurate super-resolution [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 624-632.
- [9] Kim J, Lee J K, Lee K M. Deeply-recursive convolutional network for image super-resolution[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1637-1645.
- [10] 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, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1646-1654.
- [11] 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, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2790-2798.
- [12] 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, July 21-26, 2017, Honolulu,

- HI, USA. New York: IEEE Press, 2017: 1132-1140.
- [13] 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.
- [14] 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.
- [15] Li B, Wang B, Liu J B, et al. S-LWSR: super lightweight super-resolution network[J]. IEEE Transactions on Image Processing, 2020, 29: 8368-8380.
- [16] Tian C W, Zhuge R, Wu Z H, et al. Lightweight image super-resolution with enhanced CNN[J]. Knowledge-Based Systems, 2020, 205: 106235.
- [17] Chu X X, Zhang B, Ma H L, et al. Fast, accurate and lightweight super-resolution with neural architecture search[C]//2020 25th International Conference on Pattern Recognition (ICPR), January 10-15, 2021, Milan, Italy. New York: IEEE Press, 2019: 59-64.
- [18] Tian C W, Xu Y, Zuo W M, et al. Asymmetric CNN for image superresolution[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 52(6): 3718-3730.
- [19] 王勃凡, 赵海涛. 基于径向基激活函数的高光谱小目标检测[J]. 光学学报, 2021, 41(23): 2311001.
Wang B F, Zhao H T. Small object detection in hyperspectral images based on radial basis activation function[J]. Acta Optica Sinica, 2021, 41(23): 2311001.
- [20] 张祥东, 王腾军, 朱劭俊, 等. 基于扩张卷积注意力神经网络的高光谱图像分类[J]. 光学学报, 2021, 41(3): 0310001.
Zhang X D, Wang T J, Zhu S J, et al. Hyperspectral image classification based on dilated convolutional attention neural network[J]. Acta Optica Sinica, 2021, 41(3): 0310001.
- [21] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [22] 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. New York: IEEE Press, 2020: 11531-11539.
- [23] Li X, Hu X L, Yang J. Spatial group-wise enhance improving semantic feature learning in convolutional networks[EB/OL]. (2019-05-23) [2022-08-09]. <https://arxiv.org/abs/1905.09646>.
- [24] Chen H Y, Gu J J, Zhang Z. Attention in attention network for image super-resolution[EB/OL]. (2021-04-19)[2022-08-09]. <https://arxiv.org/abs/2104.09497>.
- [25] Timofte R, Agustsson E, Gool L V, et al. NTIRE 2017 challenge on single image super-resolution: methods and results[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1110-1121.
- [26] 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: 135.
- [27] 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.
- [28] 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, 2001: 416-423.