

文章编号:1007-2780(2022)10-1334-11

基于梯度感知的单幅图像超分辨

周 乐¹, 徐 龙², 刘孝艳³, 张鑫泽², 张选德^{1*}

(1. 陕西科技大学 电子信息与人工智能学院, 陕西 西安 710021;

2. 中国科学院 国家天文台, 北京 100101;

3. 西安石油大学 理学院, 陕西 西安 710065)

摘要:随着生成对抗网络在图像超分辨(SR)领域的应用,一些感知驱动的SR方法可以恢复出纹理细节更加丰富的SR图像,有效地缓解了由PSNR主导的SR方法导致重建图像趋于平滑的问题。然而梯度信息作为图像纹理的一种重要表现形式,鲜有SR方法能准确、高效地利用。为此,提出一种基于梯度感知的图像超分辨(GASR)算法,可以更准确地利用梯度信息。一方面,使用梯度域的特征图作为作用在图像域特征图上的卷积核,有效地避免了不同域特征图串联所带来的域冲突问题;另一方面,通过对卷积核尺寸等设计细节的调整使两个分支对应位置所输出的图像域与梯度域特征图的感受野一致。此外,由于实际应用对网络轻量化需求的提高,提出的GASR算法还有效降低了参数量和计算量。与同样利用梯度信息的SPSR相比,GASR最终以约1/6的参数量与1/10的计算量取得了与其相近的性能。在Set14数据集上,LPIPS与PSNR分别提升了0.0022与0.217。实验结果验证了GASR可以在纹理生成与图像平滑度之间取得一个很好的平衡,视觉化效果也验证了GASR不仅可以准确地恢复SR图像,而且在一定程度上缓解了杂乱纹理的生成。

关 键 词:图像超分辨率;梯度;深度学习;生成对抗网络

中图分类号:TP391; TP751.1 **文献标识码:**A **doi:**10.37188/CJLCD.2022-0083

Gradient-aware based single image super-resolution

ZHOU Le¹, XU Long², LIU Xiao-yan³, ZHANG Xin-ze², ZHANG Xuan-de^{1*}

(1. *School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021, China;*

2. *National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China;*

3. *College of Science, Xi'an Shiyou University, Xi'an 710065, China)*

Abstract: With the application of generative adversarial networks in the field of image super-resolution (SR), some perception-driven SR methods can recover SR images with richer texture details, effectively alleviating the over-smoothing problem of the PSNR dominated SR methods. Gradient information is an important representation of image texture. However, few SR methods can make use of this information accurately and efficiently. In this paper, a gradient-aware single image super-resolution (GASR) is proposed, using gradient information better from two aspects. On the one hand, the feature map of gradient

收稿日期:2022-03-11;**修订日期:**2022-03-26.

基金项目:国家自然科学基金(No.61871260)

Supported by National Natural Science Foundation of China(No.61871260)

*通信联系人, E-mail: zhangxuande@sust.edu.cn

domain is used as the convolution kernel imposing on the feature map of image domain, which can effectively avoid the domain conflict caused by the concatenation of feature map of different domains. On the other hand, by elaborating the network details such as convolution kernel size, *etc.*, the image fields output at the corresponding positions of the two branches are consistent with the feel fields of the feature maps in the gradient domain. In addition, the proposed GASR algorithm also effectively reduces the number of parameters and the amount of computation due to the increased demand for network lightweight in practical applications. Compared to SPSR, GASR can achieve the same performance at the cost of about 1/6 of the parameters and 1/10 of the computation of SPSR. On Set14 dataset, LPIPS and PSNR increase by 0.002 2 and 0.217, respectively. The experimental results show that GASR can achieve a good trade-off between texture details and image smoothness. In addition, GASR can not only reconstruct high fidelity SR image, but also alleviate the generation of messy textures.

Key words: image super-resolution; gradient; deep learning; generative adversarial network

1 引 言

图像超分辨(Super Resolution, SR)与图像去噪、图像去雾等问题同为图像恢复领域的研究内容,它们的共同目标是提升图像质量。图像超分辨的特定目标是从低分辨率(Low Resolution, LR)图像恢复得到具有清晰纹理与边缘的高分辨率(High Resolution, HR)图像。根据输入的LR图像个数可以将图像超分辨分为单帧图像超分辨(Single Image Super Resolution, SISR)与多帧图像超分辨(Multiple Image Super Resolution, MISR),本文主要研究SISR,简称SR。在SR中,由于可能会有多幅不同的HR图像退化得到同样的LR图像,因此SR问题是一个难解的病态问题。如今,图像超分辨已经广泛应用于遥感图像^[1]、医学图像等领域,随着智能边缘设备的普及,比如智能手机、VR眼镜等,在这些设备上部署高性能的SR算法已成为实际生产的关键一环。

从SRCNN^[2]首次将CNN引入图像超分辨领域至今,SR性能相较于传统的对自然图像的先验信息进行建模的SR方法有了很大的提升,但已有的高性能SR算法仍然存在两个方面的缺陷:(1)大多数方法以峰值信噪比(PSNR)为主导,仅以像素损失训练网络,会导致得到的SR图像趋于平滑,难以恢复出精细的纹理与尖锐的边缘;(2)模型的计算量与参数量呈现随性能增长的趋势。SR网络模型近些年不断地向更深更宽的方向发展。比如最早将生成对抗网络(GANs)引入图像超分辨的SRGAN^[3],其参数量与每秒浮

点运算次数(Floating-Number Operations, FLOPs)分别为0.73M与7.05G,发展到近期参数量与FLOPs分别为24.80M与265.14G的SPSR^[4],可以看出,参数量和计算量均是成倍增长的。其中FLOPs以 100×100 的LR图像为输入计算得来,参数量与计算量的增长极大地限制了高性能SR算法在智能边缘设备上的应用。

为了解决PSNR主导的SR方法恢复的高分辨图像趋于平滑的问题,Justin等人^[5]将内容损失引入到训练SR网络的损失函数中,使用含有语义信息的特征图计算损失,使得超分辨图像的感知质量得到改善,有效减少了人工伪影。在生成对抗网络(GAN)提出之后,Ledig等人^[3]又将对抗损失引入训练SR网络的损失函数中,提出SRGAN,得到了更加逼真的重建图像。近期提出的SPSR利用LR图像的梯度图为重建过程提供丰富的结构信息,得到了具有更好感知质量的SR图像。虽然SPSR的性能优于先前的感知驱动的SR方法,但其存在参数量与计算量过大的问题。

为了缓解SR领域存在的问题,本文利用图像梯度信息,构建较轻量的双分支SR网络,其中包括SR分支与梯度分支。梯度分支为SR分支提供梯度信息,帮助SR分支重建出具有更好感知质量的SR图像。同时,为了更充分、准确地利用图像的梯度信息,首先要避免SPSR中将图像域特征图与梯度域特征图串联所带来的域冲突问题。本文将梯度域特征图作为作用在图像域特征图上的卷积核使用,如此可以有效避免域冲

突所带来的性能损失;其次,为了使得用作卷积核的梯度域特征图更加符合卷积核特性,在SR分支与梯度分支之间添加了两个过渡卷积层;最后,通过调整网络中的卷积核大小,使得两个分支对应位置所输出的图像域和梯度域特征图的感受野一致,达到更准确地利用梯度信息的目的。通过有效的网络结构设计,本文所提出的SR算法在感知质量与PSNR之间取得了一个比较好的平衡。除了性能提升之外,本文利用梯度感知获得了性能优势,通过减少网络通道数与基础构建块数目有效减少了SR网络的计算量与参数量,与SPSR相比,以将近1/6的计算量与1/10的参数量取得了与SPSR相近的性能。

2 相关工作

相比于传统的对先验信息进行建模的SR方法,基于深度学习的SR方法在性能上有了很大的提升。SRCNN^[2]使用3层卷积层首次将CNN引入SR领域,并使用传统的稀疏表示的思想对这3层CNN的原理进行了解释,实现了从传统SR方法到基于深度学习的SR方法的过渡。之后, Kim等人^[6]结合残差网络提出了更深的超分辨率网络VDSR,得到了更好的SR结果。Zhang等人^[7]指出CNN主要是对图像的先验信息进行建模,并论证了单一的模型可以处理图像去块、去噪以及超分辨率问题。上述基于深度学习的图像超分辨率模型的输入都是LR图像经过双三次插值之后的粗糙SR图像,再通过后续的网络细化,得到具有较好纹理细节的SR图像,称之为预上采样框架。然而,这样的预上采样框架不仅会导致网络内部计算量的成倍增加,并且还不利于感受野的有效扩充。

为了解决上述问题, Dong等人^[8]提出了加速的SRCNN,称之为FSRCNN,直接将未插值的LR图像作为网络的输入,在网络末端使用一个反卷积层执行图像的上采样,称之为后上采样框架。后上采样框架不仅可以减少计算量,扩充感受野,还可以充分利用CNN端到端学习精度高的优势,提高SR性能。使用此框架的性能优越的SR方法还有很多。RCAN^[9]使用通道注意力机制构建了超过400层的SR网络, RDN^[10]使用密集连接策略构建SR网络,均获得了较高的PSNR值。

RFANet^[11]利用残差特征融合与增强的空间注意力模块再次提升了PSNR值。然而,以提升PSNR为目标的SR方法会产生趋于平滑的SR图像,难以恢复出复杂的纹理与尖锐的边缘,影响了SR图像的视觉感知质量。为此, Justin等人^[5]在将内容损失引入到训练SR网络的损失函数中,使得超分辨率图像的感知质量得到改善,有效减少了人工伪影。在生成对抗网络(GAN)^[12]提出之后, Ledig等人^[3]又将对抗损失引入训练SR网络的损失函数中,得到了更加逼真的重建图像。之后, Wang等人^[13]使用残差密集块对SRGAN加以改进,提出了ESRGAN,获得了具有更好视觉质量的SR结果。Cheng等人提出的SPSR^[4]利用梯度信息为重建过程提供了丰富的结构信息,使得SR网络可以更好地恢复出图像的细节纹理。然而, SPSR将图像域的特征图与梯度域的特征图相串联,会带来域冲突的问题,从而导致性能的损失。本文提出梯度感知的图像超分辨率重建,将梯度域的特征图作为作用在图像域特征图上的卷积核,以此避免域冲突所带来的性能损失,实验论证了此策略的有效性。

近些年的SR研究除了注重性能的提升之外,为了将高性能的算法部署在智能边缘设备上,也逐渐开始关注模型的轻量化问题。ClassSR^[14]将LR图像分块处理,按照恢复难度将其分为3个类,分别使用具有不同容量的模型处理这3类图像块,有效地减少了已有模型的计算量。AdderSR^[15]利用在图像识别领域非常有效的AdderNet优化SR模型,使用能量消耗较小的加法器替代已有网络中的乘法操作,达到节省计算资源的目的。SMSR^[16]利用稀疏表示理论减少SR模型在图像的平滑区域所消耗的计算量。从近年来SR领域的研究趋势可以看出,模型的轻量也是评估SR算法不可忽视的重要因素之一,因此, GASR不仅关注所重建出的SR图像的感知质量,同时关注模型的轻量化。

3 梯度感知图像超分辨率重建

3.1 GASR框架

本文构建相对轻量的双分支图像超分辨率网络, SR分支用于提取LR图像的深度特征, 梯度分支用于提取LR图像梯度图的深度特征。SR

分支是图像超分辨的主分支,梯度分支作为图像超分辨的辅助分支,为SR分支提供梯度信息,帮

助SR分支重建出更准确的纹理信息。GASR整体结构如图1所示。

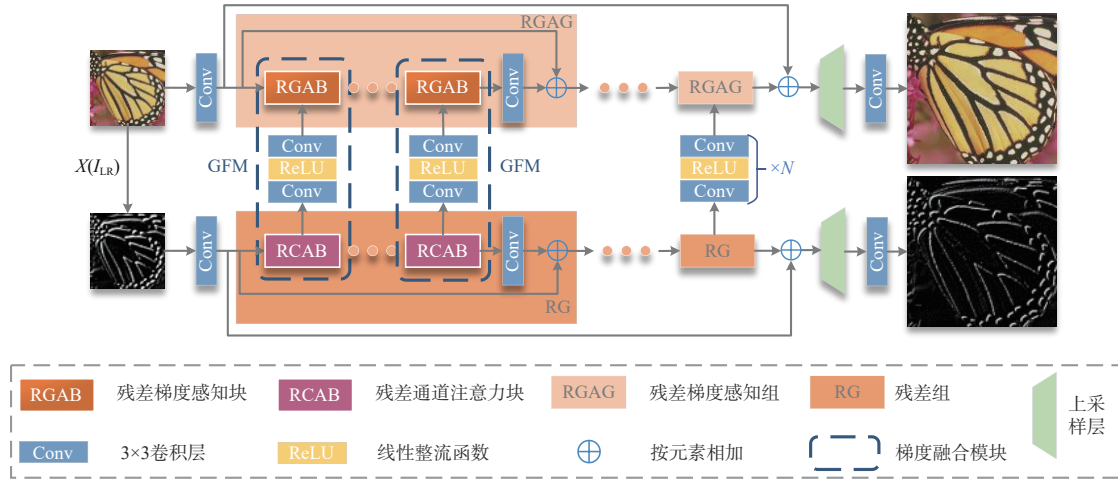


图1 GASR的总体框架

Fig. 1 Overall framework of GASR

首先,提取LR图像及其梯度图的浅层特征:

$$F_0 = H_I(I_{LR}), \quad (1)$$

$$F_{G0} = H_G(X(I_{LR})), \quad (2)$$

其中 $H_I(\cdot)$ 与 $H_G(\cdot)$ 分别表示SR分支与梯度分支的浅层特征提取函数,均为一个 3×3 的卷积层; $X(\cdot)$ 是计算梯度图的函数。将提取的LR图像的浅层特征及其梯度图的浅层特征作为主干网络的两个输入。SR分支与梯度分支的主干网络分别由M个残差梯度感知组(Residual Gradient Aware Group, RGAG)和M个RCAN中的残差组(Residual Group, RG)组成,而RGAG与RG分别由N个RGAB与N个RCAB组成。此外,在主干网络中还存在一条全局跳跃连接。LR图像及其梯度图的深层特征提取可以表示为:

$$F_{Im} = B_{Im}(F_{I(m-1)}, H_{Gm}) = B_{Im}(\dots(B_{I1}(F_0, H_{G1}))\dots, H_{Gm}), \quad (3)$$

$$F_{Gm} = B_{Gm}(F_{G(m-1)}) = B_{Gm}(\dots(B_{G1}(F_{G0}))\dots), \quad (4)$$

其中, $B_{Im}(\cdot, \cdot)$ 与 $B_{Gm}(\cdot, \cdot)$ 为第m个RGAG和RG的函数表示; F_{Im} 是第m个RGAG的输出; H_{Gm} 表示第m个RG中每个RCAB的输出特征图,此特征图是RGAG的输入之一; $F_{G(m-1)}$ 和 F_{Gm} 分别是第m个RG的输入和输出。式(3)表明,GASR的两个分支并不相互独立,即 B_{Im} 需要以第m个RG的中间特征图 H_{Gm} 为另一个输入以融合梯度信

息。最后,将深层特征 F_{Im} 和 F_{Gm} 分别作为SR分支和梯度分支重建层的输入:

$$I_{SR} = R_I(F_{Im} + F_0) = \Delta_I(I_{LR}), \quad (5)$$

$$G_{SR} = R_G(F_{Gm} + F_{G0}) = \Delta_G(X(I_{LR})), \quad (6)$$

其中, $R_I(\cdot)$ 和 $R_G(\cdot)$ 分别表示SR分支和梯度分支的重建函数, $\Delta_I(\cdot)$ 和 $\Delta_G(\cdot)$ 分别表示SR分支与梯度分支。

与SPSR类似,本文采用像素损失、感知损失和对抗损失相结合的方式训练网络,使网络能够重建出具有更好视觉质量的SR图像。其中,像素损失可以有效地提升PSNR,但会导致SR图像趋于平滑,难以恢复出尖锐的边缘以及复杂的纹理。典型的像素损失有L1损失与L2损失,研究表明,L1损失相较于L2损失具有更好的收敛性,并且恢复出的SR图像视觉质量更好,因此选用L1损失作为像素损失:

$$L_{SR}^{Pix_I} = E_{I_{SR}} \| \Delta_I(I_{LR}) - I_{HR} \|_1, \quad (7)$$

$$L_{SR}^{Pix_G} = E_{I_{SR}} \| X(\Delta_I(I_{LR})) - X(I_{HR}) \|_1, \quad (8)$$

$$L_G^{Pix_G} = E_{G_{SR}} \| \Delta_G(X(I_{LR})) - X(I_{HR}) \|_1, \quad (9)$$

其中, $L_{SR}^{Pix_I}$ 与 $L_{SR}^{Pix_G}$ 分别是作用在SR分支输出的SR图像及其梯度图上的损失, $L_G^{Pix_G}$ 是作用在梯度分支输出的SR梯度图上的损失。

为了提升图像的感知质量,Justin等人^[5]使用预训练的VGG网络提取SR图像和HR图像含

有语义信息的中间层特征,计算它们之间的L1损失:

$$L_{SR}^{Per} = E_{I_{SR}} \|\Phi_i(\Delta_I(I_{LR})) - \Phi_i(I_{HR})\|_1, \quad (10)$$

其中, $\Phi_i(\cdot)$ 表示预训练的 VGG 模型第 i 层的输出。

为了在放大因子较大时能够恢复出逼真的纹理细节, Ledig 等人将生成对抗网络(GANs)引入到 SR 任务中, 判别器 D_I 也作为优化的目标之一。受到 SPSR 的启发, 本文不仅使用判别器 D_I 对 SR 分支输出的 SR 图像进行判别, 也使用判别器 D_G 判别其梯度图。优化 D_I 、 D_G 与 Δ_I 的损失函数分别是:

$$L_{SR}^{Dis_I} = -E_{I_{SR}} [\log(1 - D_I(\Delta_I(I_{LR}))) - E_{I_{HR}} [\log(D_I(I_{HR}))], \quad (11)$$

$$L_{SR}^{Adv_I} = -E_{I_{SR}} [\log(D_I(\Delta_I(I_{LR})))], \quad (12)$$

$$L_{SR}^{Dis_G} = -E_{I_{SR}} [\log(1 - D_G(X(\Delta_I(I_{LR}))))] - E_{I_{HR}} [\log(D_G(X(I_{HR})))], \quad (13)$$

$$L_{SR}^{Adv_G} = -E_{I_{SR}} [\log(D_G(X(\Delta_I(I_{LR}))))], \quad (14)$$

其中, $L_{SR}^{Dis_I}$ 与 $L_{SR}^{Adv_I}$ 是作用在 SR 分支输出的 SR 图像上的损失, $L_{SR}^{Dis_G}$ 与 $L_{SR}^{Adv_G}$ 是作用其的梯度图上的损失。GASR 是一个高度对称的双分支网络结构, 即 RGAG 与 RG 在数量上保持一致、结构上保持相似, 如此使得梯度图的深层特征与同样位置处 LR 图像的深度特征在深度上一致, 以便于主分支更好地利用梯度信息。此外, 梯度分支输出梯度图上的 L1 损失也更好地保证了这一点。

3.2 残差梯度感知块(RGAB)

在阐述 RGAB 结构之前, 首先介绍 RCAN 中残差通道注意力块(Residual Channel Attention Block, RCAB)的结构。RCAB 是对基础残差块的改进, 在基础残差块的残差分支尾端添加了通道注意力(Channel Attention, CA), 促使网络可以根据各通道之间的依赖自适应地为不同通道的特征图赋予不同的权重。如图 2 所示, CA 首先通过全局平均池化将输入的大小为 $c \times w \times h$ 的特征图转化为一个 c 维向量, 即将特征在空间维度压缩, 然后通过两个 1×1 的卷积层对其进行通道降维与升维操作, 最后经过 Sigmoid 函数得到每个通道的权重, 将其与输入特征图按通道相乘得到加权的输出特征图。

RGAB 是对 RCAB 的改进, 具体地, 就是在 RCAB 的第一个“卷积层+ReLU”与 CA 之间插入了梯度感知卷积层(Gradient Aware Convolution, GAConv)。如图 2 所示, GAConv 由两个子分支组成。第一个子分支保留原有 RCAB 的卷积层; 第二个子分支为梯度感知分支, 以梯度分支的特征图(图中绿色长方体)为输入, 其尺寸为 $k^2 \times w \times h$ 。输入的梯度特征图每个位置可以组成一个 $k \times k$ 的卷积核(图中橙色), 使用此卷积核在 SR 分支特征图(图中浅灰色)的对应位置进行卷积, 得到此位置的梯度感知结果。最后将两个子分支的输出特征图串联, 通过 1×1 的卷积层恢复通道数后输入到 CA 中, 得到一个 RGAB 的输出。

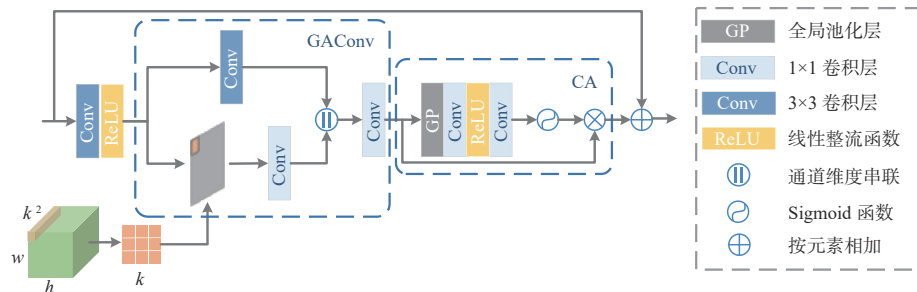


图2 残差梯度感知块结构图

Fig. 2 Residual gradient aware block structure

RGAB 的设计既保留了原有的 RCAB 结构, 同时避免了直接将梯度域特征图与图像域特征图串联而带来的域冲突问题, 也就是将梯度域的特征图作为卷积核作用在图像域的特征图上。实验证明, 将梯度域的特征图作为 SR 分

支卷积核的策略可以有效避免域冲突所带来的性能损失。

3.3 梯度融合模块(GFM)

如图 1 所示, 梯度融合模块(Gradient Fusion Module, GFM)由两个分支对应位置的 RGAB 与

RCAB以及两个过渡卷积层组成。GFM的输入输出均有两个,分别是梯度域的特征图与感知梯度的图像域特征图。为了尽可能地确保GFM的两个输出深度一致, RGAB和RCAB是结构相似且数量相同的。此外,为了促使RGAB更准确地利用梯度的深度特征,在不考虑CA层的前提下, GFM两个输出特征图的感受野也是一致的。假设输入特征图的感受野为 x ,除了特定的 1×1 卷积层之外, RGAB和RCAB中其余卷积层的核大小均为 $k \times k$,则RCAB输出的梯度域特征图的感受野为 $x + 2(k - 1)$,即图2中绿色长方体所代表的特征图的感受野为 $x + 2(k - 1)$ 。由于图2中浅灰色所表示的特征图感受野为 $x + (k - 1)$,因此, GAConv的输出特征图感受野则为 $x + 2(k - 1)$,与梯度域的输出特征图感受野一致。

此外, GASR并不直接将梯度分支的特征图作为卷积核参数,而是在梯度分支的RCAB与SR分支的RGAB之间添加了两个过渡卷积层。经过大量的有监督训练, GFM中的两个过渡卷积层可以使得最终作为卷积核的梯度域特征图更加满足卷积核特性,获得性能上的提升。实验证明,直接地将梯度分支的特征图作为SR分支的卷积核参数会使性能有一定的损失。

4 实验设置及结果分析

4.1 数据集与细节实现

实验中使用的训练数据集为DIV2K^[17],其中训练集包含800幅2K的高清图像。在图像超分辨常用的测试数据集上测试模型性能,如Set5^[6]、Set14^[18]、BSD100^[19]、General100^[8]等。参考SPSR,选用感知指数(PI)^[20]、学习的感知图像块相似度(LPIPS)^[21]、峰值信噪比(PSNR)和结构相似性(SSIM)^[22]为主要评价指标,其中LPIPS与PI越低表示SR图像的感知质量越好。

在训练过程中,首先在DIV2K的800个LR-HR图像对中随机裁剪16个对应的LR图像块与HR图像块,其中LR图像块大小为 48×48 ,对其进行随机旋转与水平翻转,以变换后的16个图像对为一批对网络中参数进行更新。具体训练步骤有3步:(1)训练放大因子为2的SR模型,即裁剪的HR图像块大小为 96×96 ,损失函数仅保留SR图像上的L1损失;(2)以损失收敛后的2倍的

模型为预训练模型训练放大因子为4的SR模型,即裁剪的HR图像块大小为 192×192 ,同样地仅保留L1损失;(3)结合内容损失与对抗损失对4倍的模型进行微调得到可以生成具有更好感知质量的SR模型。其中第(1)、(2)步是为了保证SR图像中图像内容的正确性,第(3)步可以达到提升SR图像感知质量的目的。损失函数如式(7)~(14)所示,最后一步训练所采用的损失函数为:

$$L = a_1 L_{SR}^{Per} + a_2 L_{SR}^{Pix_g} + a_3 L_{SR}^{Adv_g} + b_1 L_{SR}^{Pix_r} + b_2 L_{SR}^{Adv_r} + c L_G^{Pix_g}, \quad (15)$$

其中: $a_1 = 1, a_2 = 0.01, a_3 = 0.005, b_1 = 0.01, b_2 = 0.005, c = 0.5$ 。优化器使用ADAM优化器^[23],参数设置为 $\beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 10^{-8}$ 。由于2倍的SR模型是随机初始化的,因此初始学习率设置为 10^{-4} ,4倍的SR模型与最终的SR模型均设置初始学习率为 6^{-5} ,每 2×10^5 次迭代学习率衰减为原来的1/2。此外,每1000次迭代在Set14数据集上测试一次,可以直观地观察当前的训练效果。经过反复实验,最终确定RGAG与RG的数目 $M=7$,每个RGAG中的RGAB与RG中的RCAB的数目 $N=14$,未进行特殊说明的卷积核尺寸均为 3×3 ,网络通道数为32。所有实验均使用PyTorch实现,在GeForce RTX 2080 Ti GPUs上运行。

4.2 消融实验

本文对GASR所提出的改进点做消融实验,以支撑本文论点的正确性,消融实验的结果如表1所示,均为Set14上测试结果的平均值。

(1)梯度分支的有效性。本文构建了双分支网络模型GASR,其中梯度分支作为SR分支的辅助分支,为LR图像的超分辨过程提供梯度信息。为了证明梯度分支的有效性,在对比实验中,我们构建了减少通道数与基础构建块数目的RCAN模型,相当于去除GASR中的梯度分支,记为RCAN_sub。对比表1中RCAN_sub与GASR的实验结果可以发现,GASR的LPIPS优于RCAN_sub,并且RCAN_sub在PSNR上也逊色于有梯度分支的GASR。这有效地验证了梯度分支的存在对于更准确地重建高分辨图像的边缘结构起着积极的作用,有助于提升SR图像的视觉质量。

(2)域冲突的影响。为了证明本文所提出的

表 1 消融实验各评价指标结果(Set14 ×4)

Tab. 1 Evaluation index of ablation experiment (Set14 ×4)

Model name	Grad-branch	Non-concat	Interim-conv	Result
RCAN_sub	×	×	×	LPIPS:0.133 2 PSNR:26.815
Model_1	✓	×	×	LPIPS:0.132 0 PSNR:26.822
Model_2	✓	✓	×	LPIPS:0.133 4 PSNR:26.844
GASR	✓	✓	✓	LPIPS:0.129 6 (↓) PSNR:26.857 (↑)

注:LPIPS 越小越好,PSNR 越大越好

避免不同域特征图串联来缓解域冲突问题策略的有效性,本实验构建了直接将图像域与梯度域特征图在通道维度串联的网络模型,记为 Model_1。Model_1 模型的上下分支均以 RCAB 为基础构建块,梯度分支的每个 RCAB 输出不仅作为下一个 RCAB 的输入,同时要与 SR 分支对应位置的 RCAB 的输出在通道维度串联,再经过一个 1 × 1 卷积层恢复通道数,此输出作为 SR 分支下一个 RCAB 的输入。从表 1 的实验结果可以发现,不同域的特征图直接串联造成了 LPIPS 与 PSNR 不同程度的恶化,表明直接将图像域与梯度域的特征图串联所造成的域冲突问题会导致图像超分辨性能的下降。

(3)过渡卷积层的必要性。本实验将 GASR 中 RGAB 与 RCAB 之间的两个过渡卷积层移至梯度分支的相邻 RCAB 之间,以证明过渡卷积层的有效性。首先在 GASR 中去除两个过渡卷积层,接下来在梯度分支的相邻 RCAB 之间添加两个 1 × 1 卷积层。第一个卷积层将上一个 RCAB 输出的梯度域特征图的通道数转换为 3 × 3 = 9 个通道,此卷积层所输出的特征图不仅用于下一个卷积层的输入,同时也要作为对应位置 RGAB 中卷积层的卷积核参数。第二个卷积层起到恢复通道数的作用,记此网络为 Model_2。由表 1 的实验结果可知,去除了过渡卷积层的 Model_2 在 LPIPS 和 PSNR 指标上均逊色于 GASR,此结果验证了过渡卷积层可以起到使得梯度域特征图更加符合卷积核特性的作用。

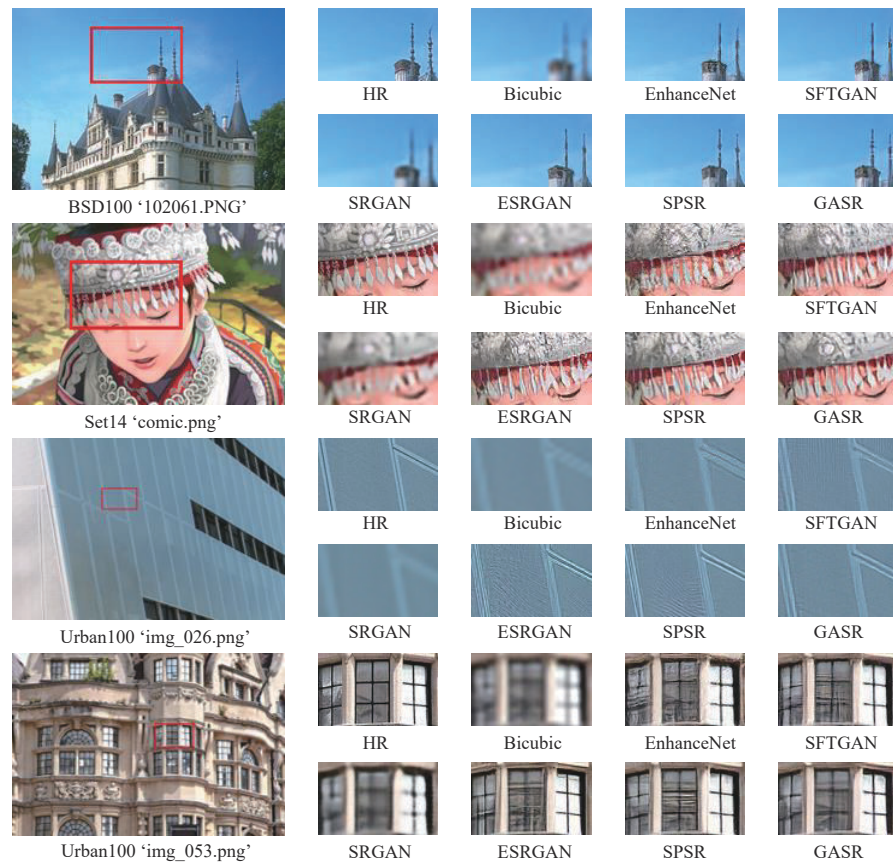
4.3 基准数据集的对比实验

本文将 GASR 的实验结果与之前性能较好的

感知驱动的 SR 算法做定量与定性上的比较,对比算法包括 SFTGAN^[24]、SRGAN^[3]、ESRGAN^[13]、EnhenceNet^[25]与 SPSR^[4]。定量比较结果如表 2 所示,定性比较结果如图 3 所示。

(1)定量对比。从表 2 可以看出,虽然 SPSR 达到了比较好的 LPIPS 与 PI,但其参数量与计算量较大,其中参数量是 GASR 的 5.7 倍,FLOPs 是 GASR 的 10.2 倍,GASR 以绝对的参数量与计算量优势达到了与 SPSR 相近的性能。与 ESRGAN 相比,GASR 也以更少的参数量与计算量取得了更好的性能,其中 ESRGAN 的参数量是 GASR 的 3.9 倍,FLOPs 是 GASR 的 3.4 倍。虽然 SFTGAN 与 SRGAN 参数量与计算量较少,但其性能上逊色于其余的对比算法。此外,与感知驱动的对比 SR 算法相比,GASR 在所有数据集上取得了最好的 PSNR 与 SSIM,表明其在感知质量与平滑度之间取得了更好的平衡。表 2 中的结果表明,本文所提出的 GASR 相较于其余感知驱动的 SR 算法以较少的参数量达到了优越的性能。

(2)定性对比。图 3 展示了 GASR 与其余感知驱动的 SR 算法在视觉效果上的对比。从第 1 组对比图可以看出,SRGAN 所恢复的图像趋于平滑,难以恢复塔尖的细节;EnhenceNet、SFTGAN 与 ESRGAN 虽然能够恢复出塔尖的大体结构,但对于细节的恢复还有待提升;SPSR 虽然能够恢复出部分塔尖的细节,但 GASR 的细节更为准确,与 HR 图像更为贴合。观察第 2、3、4 组对比结果可知,SRGAN 均会产生过度平滑的 SR 结果,EnhenceNet、SFTGAN、ESRGAN 与 SPSR 会产生一些杂乱的纹理,导致 SR 图像严重失真,

图 3 与先进的感知驱动的 SR 方法的视觉比较结果($\times 4$)Fig. 3 Results compared visually with perceptual-driven SR methods ($\times 4$)表 2 与先进的感知驱动的 SR 方法的比较结果($\times 4$)Tab. 2 Results compared with advanced perception-driven SR method ($\times 4$)

Dataset	Metric	Bicubic	SFTGAN	SRGAN	ESRGAN	SPSR	GASR (Ours)
Set5	#Params.		1.8M	0.7M	16.7M	24.8M	4.3M
	FLOPs		0.8G	7.0G	89.7G	265.1G	26.6G
	PI	7.369 9	3.758 7	3.982 0	3.752 2	3.274 3	3.636 4
	LPIPS	0.340 7	0.089 0	0.088 2	0.074 8	0.064 4	0.070 2
	PSNR	28.420	29.932	29.168	30.454	30.400	30.517
Set14	SSIM	0.824 5	0.866 5	0.861 3	0.867 7	0.862 7	0.872 0
	PI	7.026 8	2.906 3	3.085 1	2.926 1	2.903 6	3.027 2
	LPIPS	0.439 3	0.148 1	0.166 3	0.132 9	0.131 8	0.129 6
	PSNR	26.100	26.223	26.171	26.276	26.640	26.857
BSD100	SSIM	0.785 0	0.785 4	0.784 1	0.778 3	0.793 0	0.798 1
	PI	7.002 6	2.377 4	2.545 9	2.479 3	2.351 0	2.427 3
	LPIPS	0.524 9	0.176 9	0.198 0	0.161 4	0.161 1	0.161 6
	PSNR	25.961	25.505	25.459	25.317	25.505	25.698
General100	SSIM	0.667 5	0.654 9	0.648 5	0.650 6	0.657 6	0.664 2
	PI	7.936 5	4.287 8	4.375 7	4.323 4	4.099 1	4.250 7
	LPIPS	0.352 8	0.103 0	0.105 5	0.087 9	0.086 3	0.085 1
	PSNR	28.018	29.026	28.575	29.412	29.414	29.744
General100	SSIM	0.828 2	0.850 8	0.854 1	0.854 6	0.853 7	0.861 4

注:标红、标蓝分别为此数据集上某种性能指标的第1、2名。LPIPS与PI越小越好,PSNR与SSIM越大越好

而 GASR 在平滑度与纹理生成之间取得了更好的平衡,获得了更好的视觉效果。这有效地验证了 GASR 可以更准确地恢复图像的结构信息,产生视觉质量更优的 SR 图像的结论。

4.4 真实图像的对比实验

除了在基准数据集上对 GASR 做定量与定性分析之外,也在真实图像上分析了 GASR 的有效性。本实验选取感知驱动的 SR 算法如 SFTGAN^[24]、SRGAN^[3]、ESRGAN^[13]、USRGAN^[26]、SPSR^[4]与 GASR 做比对。

RealSR^[27]是通过调节相机焦距并利用图像配准算法获得的真实 LR-HR 图像对的数据集。本实验随机选取 RealSR 中的 46 对 LR-HR 图像对作为测试图像,将 GASR 与双三次(Bibubic)插值、ESRGAN 与 USRGAN 进行比较,定量比较结果如表 3 所示。与计算复杂度更高的 ESRGAN 相比,GASR 以复杂度低的优势在各个评价指标上均胜出,而 USRGAN 虽然取得了更高的 PSNR 与 SSIM,但其所得 SR 图像的感知质量不及 GASR。定性比较结果如图 4 所示。SFTGAN 与 ESRGAN 均会产生不同程度的杂乱纹理或失真现象,SRGAN 难以恢复图像的边缘,而 GASR 虽然有少量人工伪影,但其相较于上述 3 种算法所恢复出的图像边缘更加清晰。虽然 SPSR 可

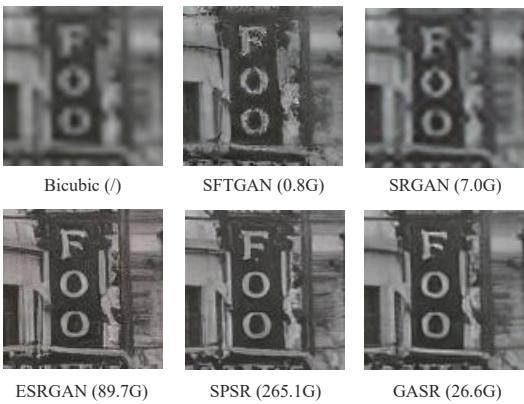


图 4 真实图像上的对比实验结果(×4)

Fig. 4 Comparison of experimental results on real images (×4)

以构建出更清晰的图像边缘,但从图 4 括号中所标注的 FLOPs 看出,SPSR 的计算量是 GASR 的 10 倍左右,这极大地限制了 SPSR 在真实场景中的应用。

5 结 论

本文提出了一种较轻量的梯度感知图像超分辨(GASR)算法,可以更准确地利用图像的梯度域信息,产生具有更好感知质量的 SR 图像。GASR 主要从以下两个方面达到准确利用梯度域信息的目的:首先,将梯度域的特征图作为作用在图像域特征图上的卷积核,如此可以有效地避免不同域特征图串联所带来的域冲突问题;其次,为了保证梯度域特征图和图像域特征图在对应位置的感受野一致,在卷积核尺寸等模型细节上进行了一定设计。最终 GASR 以 4.3M 的参数量与 26.6G FLOPs 的计算量达到了比其他感知驱动的 SR 算法更好的性能,在纹理生成与平滑度之间取得了更好的平衡。

表 3 RealSR 数据集上的比较结果(×4)

Tab. 3 Comparison results on RealSR datasets (×4)

Method	LPIPS	PI	PSNR	SSIM
Bibubic	0.485 3	8.189 0	27.507	0.865 9
ESRGAN	0.428 8	7.597 5	27.799	0.874 4
USRGAN	0.422 6	7.592 8	27.867	0.877 6
GASR(Ours)	0.417 9	7.503 7	27.805	0.876 4

注:标红、标蓝分别为此数据集上某种性能指标的第 1、2 名

参 考 文 献:

[1] 刘涛,钱锋,张葆. 遥感图像的 MAP 超分辨重建[J]. 液晶与显示,2018,33(10):884-892.
LIU T, QIAN F, ZHANG B. MAP super-resolution reconstruction of remote sensing image [J]. *Chinese Journal of Liquid Crystals and Displays*, 2018, 33(10): 884-892. (in Chinese)
[2] DONG C, LOY C C, HE K M, *et al.* Learning a deep convolutional network for image super-resolution [C]//*Proceedings of the 13th European Conference on Computer Vision*. Zurich: Springer, 2014: 184-199.
[3] LEDIG C, THEIS L, HUSZÁR F, *et al.* Photo-realistic single image super-resolution using a generative adversarial network [C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu:

- IEEE, 2017: 105-114.
- [4] MA C, RAO Y M, CHENG Y A, *et al.* Structure-preserving super resolution with gradient guidance [C]//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 7766-7775.
 - [5] JOHNSON J, ALAHI A, LI F F. Perceptual losses for real-time style transfer and super-resolution [C]//*Proceedings of the 14th European Conference on Computer Vision*. Amsterdam: Springer, 2016: 694-711.
 - [6] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 1646-1654.
 - [7] ZHANG K, ZUO W M, CHEN Y J, *et al.* Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising [J]. *IEEE Transactions on Image Processing*, 2017, 26(7): 3142-3155.
 - [8] DONG C, LOY C C, TANG X O. Accelerating the super-resolution convolutional neural network [C]//*Proceedings of the 14th European Conference on Computer Vision*. Amsterdam: Springer, 2016: 391-407.
 - [9] ZHANG Y L, LI K P, LI K, *et al.* Image super-resolution using very deep residual channel attention networks [C]//*Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018: 294-310.
 - [10] ZHANG Y L, TIAN Y P, KONG Y, *et al.* Residual dense network for image super-resolution [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 2472-2481.
 - [11] LIU J, ZHANG W J, TANG Y T, *et al.* Residual feature aggregation network for image super-resolution [C]//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 2356-2365.
 - [12] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, *et al.* Generative adversarial nets [C]//*Proceedings of the 27th International Conference on Neural Information Processing Systems*. Montreal: MIT Press, 2014: 2672-2680.
 - [13] WANG X T, YU K, WU S X, *et al.* ESRGAN: enhanced super-resolution generative adversarial networks [C]//*Proceedings of European Conference on Computer Vision*. Munich: Springer, 2018: 63-79.
 - [14] KONG X T, ZHAO H Y, QIAO Y, *et al.* ClassSR: a general framework to accelerate super-resolution networks by data characteristic [C]//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021: 12011-12020.
 - [15] SONG D H, WANG Y H, CHEN H T, *et al.* Addersr: towards energy efficient image super-resolution [C]//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021: 15643-15652.
 - [16] WANG L G, DONG X Y, WANG Y Q, *et al.* Exploring sparsity in image super-resolution for efficient inference [C]//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021: 4915-4924.
 - [17] AGUSTSSON E, TIMOFTE R. NTIRE 2017 challenge on single image super-resolution: dataset and study [C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Honolulu: IEEE, 2017: 1122-1131.
 - [18] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations [C]//*Proceedings of the 7th International Conference on Curves and Surfaces*. Avignon: Springer, 2010: 711-730.
 - [19] 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 of the Eighth IEEE International Conference on Computer Vision*. Vancouver: IEEE, 2001: 416-423.
 - [20] BLAU Y, MECHREZ R, TIMOFTE R, *et al.* The 2018 PIRM challenge on perceptual image super-resolution [C]//*Proceedings of European Conference on Computer Vision*. Munich: Springer, 2018: 334-355.
 - [21] ZHANG R, ISOLA P, EFROS A A, *et al.* The unreasonable effectiveness of deep features as a perceptual metric [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake

- City: IEEE, 2018: 586-595.
- [22] 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.
- [23] KINGMA D P, BA J. Adam: a method for stochastic optimization [J]. *arXiv*, 2014: 1412.6980.
- [24] WANG X T, YU K, DONG C, *et al.* Recovering realistic texture in image super-resolution by deep spatial feature transform [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 606-615.
- [25] SAJJADI M S M, SCHÖLKOPF B, HIRSCH M. Enhancenet: single image super-resolution through automated texture synthesis [C]//*Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice: IEEE, 2017: 4501-4510.
- [26] ZHANG K, VAN GOOL L, TIMOFTE R. Deep unfolding network for image super-resolution [C]//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 3214-3223.
- [27] CAI J R, ZENG H, YONG H W, *et al.* Toward real-world single image super-resolution: a new benchmark and a new model [C]//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE, 2019: 3086-3095.

作者简介:



周 乐(1999—),女,陕西渭南人,硕士研究生,2020年于陕西科技大学获得学士学位,主要从事图像处理、图像超分辨等方面的研究。E-mail: luoluo3664@sust.edu.cn



张选德(1979—),男,宁夏固原人,博士,教授,2013年于西安电子科技大学获得博士学位,主要从事图像恢复、图像质量评价、视觉追踪、多光谱图像处理等方面的研究。E-mail: zhangxuande@sust.edu.cn