

# 基于卷积神经网络梯度和纹理补偿的单幅图像超分辨率重建

黄裕青, 李华锋, 原 铭, 张亚飞

(昆明理工大学信息工程与自动化学院, 昆明 650500)

**摘要:** 现有的单幅图像超分辨率重建算法大都在追求高峰值信噪比(Peak signal-to-noise ratio, PSNR), 在特征提取过程中缺少对图像纹理细节信息的关注, 导致重建图像的人眼主观感知效果不太理想。为了解决这一问题, 本文提出了一种基于卷积神经网络梯度和纹理补偿的单幅图像超分辨率重建算法。具体设计了3条支路分别用于结构特征提取、纹理细节特征提取及梯度补偿, 然后利用所提出的融合模块对结构特征和纹理细节特征进行融合。为防止重建过程中丢失图像的纹理信息, 提出纹理细节特征提取模块补偿图像的纹理细节信息, 增强网络的纹理保持能力; 同时, 利用梯度补偿模块提取的梯度信息对结构信息进行增强; 此外还构建了深层特征提取结构, 结合通道注意力与空间注意力对深层特征中的信息进行筛选及特征增强; 最后利用二阶残差块对结构和纹理特征进行融合, 使重建图像的特征信息更加完善。通过对比实验验证了本文方法的有效性和优越性。

**关键词:** 卷积神经网络; 超分辨率图像重建; 梯度纹理补偿; 注意力

中图分类号: TP751.1

文献标志码: A

## Super-Resolution Reconstruction of Single Image Based on Convolutional Neural Network Gradient and Texture Compensation

HUANG Yuqing, LI Huafeng, YUAN Ming, ZHANG Yafei

(School of Information Engineering and Automation, Kunming University of Technology, Kunming 650500, China)

**Abstract:** The existing super-resolution reconstruction algorithms of single image mostly pursue the peak signal-to-noise ratio (PSNR), and lack the attention to the details of image texture in the process of feature extraction, resulting in poor subjective perception of reconstructed images. In order to solve this problem, this paper proposes a single image super-resolution reconstruction algorithm based on convolutional neural network gradient and texture compensation. Specifically, three branches are designed for structure feature extraction, texture detail feature extraction and gradient compensation, and then the proposed fusion module is used to fuse the structure feature and texture detail feature. To prevent the loss of texture information in the reconstruction process, this paper proposes a texture detail feature extraction module to compensate the texture detail information of the image and enhance the texture retention ability of the network. At the same time, this paper uses the gradient information extracted by the gradient compensation module to enhance the structure information. In addition, this paper also constructs a deep feature

基金项目: 国家自然科学基金(62161015, 61966021)。

收稿日期: 2022-06-29; 修订日期: 2022-10-05

extraction structure, combining channel attention and spatial attention to screen and enhance the information in the deep features. Finally, the second-order residual block is used to fuse the structure and texture features, so that the feature information of the reconstructed image is more perfect. The effectiveness and superiority of the proposed method are verified by comparative experiments.

**Key words:** convolutional neural network; super-resolution image reconstruction; gradient and texture compensation; attention

## 引 言

当图像或视频由传感器产生时,质量通常会受到透镜的清晰度、传输带宽、采样点数和场景细节等因素的限制。单张图像超分辨率(Single image super-resolution, SISR)方法通过从低分辨率(Low resolution, LR)图像中恢复清晰的高分辨率(High resolution, HR)图像来提升图像质量,在自然成像、卫星成像、医学成像、安全监视成像和计算机视觉分析任务<sup>[1-3]</sup>中应用广泛。由于任意低分辨率图像都存在多个解,因此求解超分辨率图像的过程是一个不适定问题。到目前为止,已经出现了多种超分辨率图像重建方法,主要包括基于插值的方法<sup>[4-7]</sup>、基于重建的方法<sup>[8-11]</sup>、基于学习的方法<sup>[12-15]</sup>以及基于深度学习的方法<sup>[16-21]</sup>。

基于插值的方法主要包括双线性插值、最近邻插值与双三次插值<sup>[4]</sup>。这类算法实现简单,然而,它们是通过局部信息来重建高分辨率图像的,当放大倍数较大时,通常效果较差。基于重建的方法是以图像降质模型基础构建HR图像到LR图像的映射,并利用先验知识,如梯度先验<sup>[8]</sup>、边缘先验<sup>[10]</sup>,通过优化损失函数求解得到高分辨率图像。该方法在抑制噪声和边缘保留方面具有较好的效果,但重建效果易受到正则化参数准确性影响,导致纹理细节不能很好的恢复。基于学习的方法是对大量LR图像和其对应HR图像之间的隐含关系进行学习,利用这种隐含关系指导HR图像的重建,例如稀疏表示<sup>[12]</sup>和基于样例的方法<sup>[13]</sup>。虽然基于学习的方法已经取得了显著的改进,但网络的优化过程过于复杂。

近年来,基于深度学习的SISR方法的性能比传统的基于非学习的SISR方法有了显著的提升。基于深度学习的SISR方法将图像超分辨率重建视为一个图像回归任务,通过学习LR和HR图像之间由卷积神经网络(Convolutional neural networks, CNN)表示的端到端图像映射函数来进行超分重建。Dong等<sup>[16]</sup>首次将深度学习应用于图像超分辨率重建,提出了基于卷积神经网络的超分辨率(Super resolution CNN, SRCNN)重建方法,取得了清晰的重建效果。为解决模型参数数量大量增多的问题,研究者们提出了多种解决方案。Dong等<sup>[22]</sup>对SRCNN进行了改进,提出了基于快速卷积神经网络的超分辨率重建(Fast SRCNN, FSRCNN)方法,该方法先在低分辨率阶段进行卷积、提取特征,然后在网络最后进行上采样生成高分辨率图像;Kim等<sup>[21]</sup>构造了超深卷积网络(Very deep convolutional networks, VDSR)。VDSR网络利用残差的思想通过添加一个跳跃连接使网络学习残差参数,而不是全体参数,从而减少了参数量;Zhang等<sup>[23]</sup>提出了基于密集残差密集网络(Residual dense network, RDN)的图像超分辨率重建方法,该方法较好地解决了深度网络中的梯度消失及收敛过慢的问题,以及图像信息在卷积过程中逐级丢失的问题,取得了很好的重建效果。然而,只使用单条支路进行特征提取,可能无法捕获尽可能多的特征信息。为解决此问题,Li等<sup>[24]</sup>提出了一种多尺度残差网络(Multi-scale residual network, MSRN)的超分辨率重建方法,通过利用不同支路不同感受野的卷积核进行特征提取,以获得充足的图像特征信息;Zhao等<sup>[25]</sup>提出像素注意方案构造了一个非常简洁和有效的网络(Pixel attention network, PAN),这种注意方案引入了更少的附加参数,但产生了更好的超分辨率结果;Meng等<sup>[26]</sup>提出了一种梯度信息蒸馏网络(Gradient information distillation network, GIDN),通过信息蒸馏来保持快速、轻量级的

优势,并利用梯度信息提高了超分辨率重建图像的质量;Liu等<sup>[27]</sup>提出了一种新的群跳远连接块(Group skip-connecting network, GSCN)作为图像超分辨率重建的基本构建块,将群卷积与跳跃连接相结合,充分融合丰富的局部特征,从而获得比当前大多数残差块更高的效率;Mehri等<sup>[28]</sup>提出了多路径残差网络(Multi-path residual network, MPRNet),自适应学习最有价值的特征,构建网络以学习高频信息;Tan等<sup>[29]</sup>提出一种简单而有效的自校准特征融合网络(Self-calibrated feature fuse network, SCFFN),该网络通过自校准特征融合,以实现网络性能和适用性之间更好的平衡。然而,上述方法都是将图像的纹理特征信息与结构特征信息视为一个整体,而真实情况下,图像卷积的过程中极易丢失纹理特征,这对重构富含纹理结构信息的高分辨率图像极为不利。

为解决上述问题,近年来研究者们提出了大量基于纹理特征补偿的模型<sup>[30-33]</sup>。这些方法利用多种信息补偿方法对超分辨率重建结果进行补偿,在重构结构信息的同时,防止超分辨率重建过程中丢失纹理细节特征,从而提升了图像的感知性能。具体地,Ledig等<sup>[30]</sup>提出了基于生成对抗网络的超分辨率(Super-resolution generative adversarial network, SRGAN)重建方法,该方法利用生成对抗损失约束生成器与鉴别器,使生成器生成更加逼近真实世界的纹理细节。然而生成对抗网络极易将不属于真实世界的纹理信息引入图像中,为此,Wang等<sup>[31]</sup>提出增强的生成对抗超分辨率网络(Enhanced SRGAN, ES-RGAN)模型,利用密集残差网络进一步提取有效特征,同时为减少参数舍去了批归一化层,从而生成了更逼真、自然的纹理细节信息;Wang等<sup>[32]</sup>提出了空间特征变换生成对抗网络(Spatial feature transform generative adversarial networks, SFTGAN)模型,利用语义先验信息引导LR图像生成HR图像,并利用空间特征变换(Spatial feature transformation, SFT)层来进一步提高纹理恢复的能力;Soh等<sup>[33]</sup>从低分辨率源图像中提取梯度信息,并利用梯度信息对超分辨率过程进行补偿。上述方法都对提升图像的视觉感知效果做出了卓越的贡献。然而,在恢复的SR图像中仍会出现非自然的伪影,且现有的感知驱动方法在性能指标上远低于常用的超分辨率重建方法,这对于评价图像的主观性能极为不利,且影响了方法之间的性能比较。

从以上讨论来看,虽然SISR方法已经取得了很大的进展,但现有的基于深度学习的SISR方法仍然存在一些局限性。这类方法通常利用单支路网络提取图像特征,在损失函数的约束下,使网络提取到由LR图像恢复到HR图像所需的纹理细节特征,从而获得图像高频细节信息,来实现超分辨率重建。然而,这类方法通常需要大规模高分辨率图像参与训练才能保证重建图像的效果,并且通过单支路网络提取特征来同时恢复原始LR图像中包含的所有高频和低频信息存在一定的挑战,存在纹理细节特征丢失的问题,导致超分辨率重建出的图像纹理细节不完整、图像质量不佳等问题。为此,本文提出了一种基于纹理细节嵌入的超分辨率重建网络,通过将图像区分为结构信息与纹理信息分别进行超分辨率重建,防止超分辨率重建过程中丢失图像的纹理信息,并保证性能指标未因产生伪影而下降。网络利用3条支路分别提取图像的结构信息、纹理信息及梯度信息。通过纹理信息与结构信息融合来解决超分辨率重建过程中由于纹理细节丢失而引起的图像质量下降的问题。同时,利用梯度信息对结构信息进行增强,解决由于梯度信息丢失而引起的重建图像边缘模糊的问题。

本文的贡献主要包括:

(1)为解决超分辨率重建过程中纹理细节丢失而引起重建图像质量下降的问题,设计了一种梯度和纹理补偿的超分辨率重建网络。该网络通过不同支路分别对图像的纹理、结构、梯度信息进行提取,并将其有效融入高分辨率重建图像中,有效提升了重建图像的质量。

(2)为使网络捕获感兴趣的结构特征区域与纹理特征区域,引入深层特征提取块。通过通道注意力与空间注意力对结构与纹理特征进行增强,提升关键信息在图像重建中作用。

(3)为更好地融合结构信息与纹理信息,提出了一种二阶残差块结构用于融合模块,其利用不同尺度的卷积核进行特征提取及融合,使融合后的特征能得到更好地表达重建的图像信息。

### 1 网络框架结构

本文提出的基于纹理细节嵌入的超分辨率重建模型如图1所示。该模型主要包括4个模块:纹理细节特征提取模块、结构特征提取模块、梯度补偿模块和融合模块。为防止超分辨率重建过程中纹理信息丢失,本文将纹理特征提取模块提取的纹理特征逐层补偿到结构特征提取模块。纹理特征提取模块与结构特征提取模块包含相似的结构,由1个浅层特征提取块及4个深层特征提取块组成。梯度补偿模块由4个 $3\times 3$ 的卷积和PReLU组成,该模块对原图像进行逐级卷积以提取图像的梯度特征。利用最后得到的梯度信息特征图对结构信息特征图进行增强。然后把增强后的结构信息与纹理信息进行拼接,经过融合模块进行特征增强与特征融合,最后通过反卷积进行重构并利用纹理图进行补偿得到高分辨率图像。

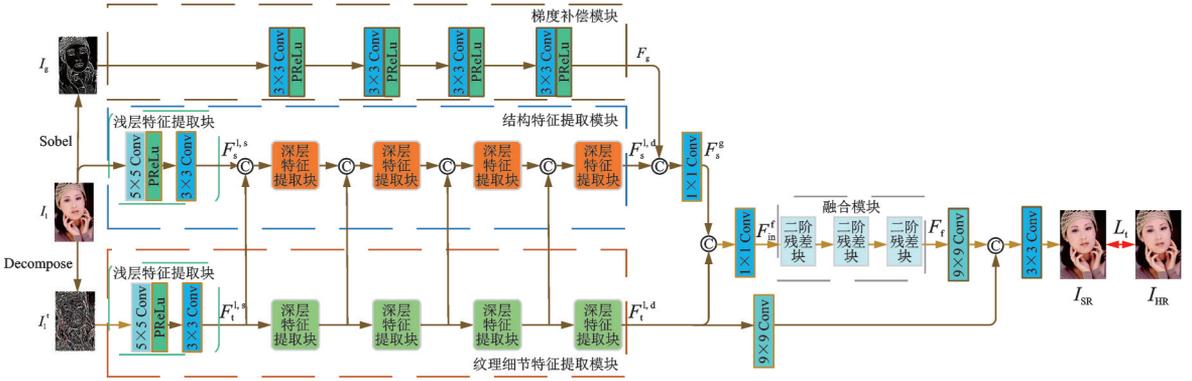


图1 网络总体框架流程图

Fig.1 Flow of network overall framework

具体地,利用纹理特征提取模块与结构特征提取模块从原始的低分辨率图像  $I_t$  及其对应的纹理细节图像  $I_t^1$  中,分别提取结构特征  $F_s^{1,d}$ ,纹理细节特征  $F_t^{1,d}$

$$\begin{cases} F_s^{1,d} = H_s(I_t) \\ F_t^{1,d} = H_t(I_t^1) \end{cases} \quad (1)$$

式中: $H_s$ 表示结构特征提取模块操作, $H_t$ 表示纹理细节特征提取模块操作。

由于梯度信息丢失会造成重建图像边缘模糊,本文利用梯度信息对结构信息进行增强。利用Sobel算子从原始的低分辨率图像  $I_t$  提取梯度图  $I_g$ ,再经过4个卷积层提取梯度信息得到  $F_g$ ,可以表示为

$$\begin{cases} I_g = f_{Sobel}(I_t) \\ F_g = Conv_g(I_g) \end{cases} \quad (2)$$

式中: $f_{Sobel}$ 表示Sobel算子操作, $Conv_g$ 表示由 $3\times 3$ 卷积与PReLU组成的特征提取操作组合(如图1所示)。结构特征经梯度增强后的输出  $F_s^g$ 可以表示为

$$F_s^g = sconv(\text{cat}(F_g, F_s^{1,d}), k=1) \quad (3)$$

然后,把经梯度信息增强后的结构特征  $F_s^g$ 和纹理特征提取模块的输出  $F_t^{1,d}$ 经过拼接、 $1\times 1$ 卷积,得到的特征  $F_m^f$ 。将  $F_m^f$ 作为融合模块的输入,经过融合模块得到的特征图  $F_f$ ,即

$$\begin{cases} F_m^f = sconv(\text{cat}(F_s^g, F_t^{1,d}), k=1) \\ F_f = H_f(F_m^f) \end{cases} \quad (4)$$

式中  $H_f$ 表示由3个二阶残差块组成的融合模块。

其次,融合模块的输出特征  $F_f$ 通过反卷积进行上采样,并利用纹理特征图  $F_t^{1,d}$ 经反卷积后得到纹理

补偿图  $F_t^{up}$  对高分辨率的融合特征图进行补偿。最后,经  $3 \times 3$  卷积进行重构得到高分辨率图像  $I_{SR}$ , 可以表示为

$$\begin{cases} F_t^{up} = \text{Deconv}(F_t^{1,d}) \\ F_{SR}^h = \text{cat}(\text{Deconv}(F_f), F_t^{up}) \\ I_{SR} = \text{sconv}(F_{SR}^h, k=3) \end{cases} \quad (5)$$

式中 Deconv 代表反卷积操作。

### 1.1 结构特征提取模块

如图 1 所示,结构特征提取模块包含浅层特征提取块和多个深层特征提取块。深层特征提取块通过通道注意力与空间注意力对结构特征进行增强,提升关键信息在图像重建中的作用。同时,通过多尺度卷积提取不同感受野下的特征信息,防止结构信息丢失,使网络能够恢复出包含丰富结构信息的高分辨率图像。

具体地,首先利用浅层特征提取块对源图像进行浅层特征提取,得到浅层结构信息特征  $F_s^{1,s}$ , 即

$$F_s^{1,s} = \text{sconv}(\text{PReLU}(\text{sconv}(I_1, k=5)), k=3) \quad (6)$$

利用浅层纹理特征信息  $F_t^{1,s}$  对  $F_s^{1,s}$  进行补偿作为结构信息提取模块中深层特征提取块的输入。结构特征提取模块中第  $i$  层深度特征提取块的输出  $F_{s,out}^i$  可以表示为

$$F_{s,out}^i = f_{dep}^i(F_{s,in}^i) \quad (7)$$

式中:  $f_{dep}^i$  表示第  $i$  个深层特征提取块,  $F_{s,in}^i$  表示第  $i$  个深层特征提取块的输入,  $F_{s,in}^i = \text{cat}(F_{s,out}^{i-1}, F_t^{i-1})$ ,  $F_{s,in}^1 = \text{cat}(F_s^{1,s}, F_t^{1,s})$ 。结构特征提取模块有 4 个深度特征提取块,其最后的输出为  $F_s^{1,d} = F_{s,out}^4$ 。

深层特征提取块由多尺度特征提取、通道注意力、空间注意力和残差连接组成,其结构如图 2 所示。

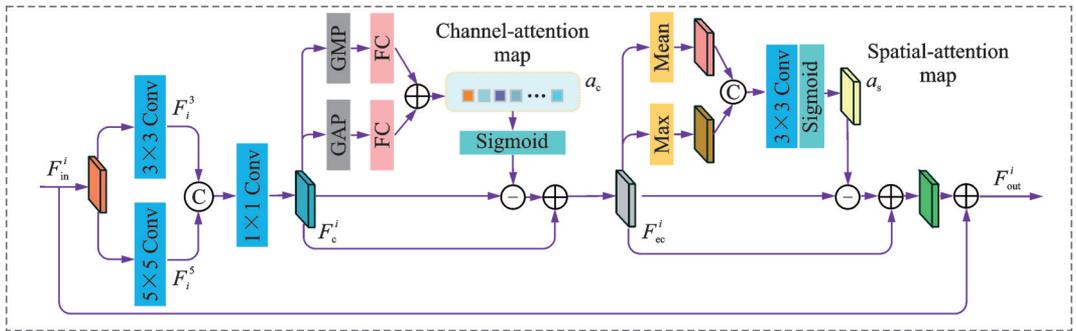


图 2 深层特征提取块结构图

Fig.2 Structure diagram of deep feature extraction block

有研究发现,通过拓宽网络宽度能够更有效地提取特征信息。因此,通过  $3 \times 3$  和  $5 \times 5$  两个不同大小卷积核进行多尺度特征提取,提取得到的特征图分别为

$$\begin{cases} F_i^3 = \text{sconv}(F_{in}^i, k=3) \\ F_i^5 = \text{sconv}(F_{in}^i, k=5) \end{cases} \quad (8)$$

式中:  $F_{in}^i$  表示第  $i$  个深层特征提取块的输入,  $F_i^3$  与  $F_i^5$  分别为第  $i$  个深层特征提取模块中  $3 \times 3$  与  $5 \times 5$  卷积后的结果。

为从低分辨率图像中恢复包含足够高频信息的高分辨率图像,本文结合空间注意力与通道注意力处理经多尺度提取的低分辨率特征。具体地,将多尺度提取得到的特征  $F_i^3$  和  $F_i^5$  拼接后输入通道注意力,从而充分利用多尺度提取到的特征信息,通道注意力的输入  $F_c^i$  可以表示为

$$F_c^i = \text{sconv}(\text{cat}(F_s^i, F_5^i), k=1) \quad (9)$$

在通道注意力操作中,利用全局平均池化与全局最大池化分别对 $F_c^i$ 进行处理,然后经过全连接层和Sigmoid激活函数得到通道注意力 $a_c$

$$a_c = \text{Sigmoid}(\text{FC}(\text{GAP}(F_c^i)) + \text{FC}(\text{GMP}(F_c^i))) \quad (10)$$

式中:FC表示全连接层;Sigmoid表示激活函数,GAP与GMP分别表示全局平均池化和全局最大池化操作。经通道注意力增强后的特征 $F_{ec}^i$ 可表示为

$$F_{ec}^i = (a_c \odot F_c^i) + F_c^i \quad (11)$$

式中 $\odot$ 表示点乘。

在空间注意力中,利用通道注意力增强后的特征图 $F_{ec}^i$ 进行 $C_{\text{mean}}$ 操作(在通道维度对特征图上同一位置的像素值取平均)和 $C_{\text{max}}$ 操作(在通道维度对特征图上同一位置的像素值取最大),并经过 $3 \times 3$ 的卷积和Sigmoid生成空间注意力图,旨在突出局部视图中的高频信息,注意力图 $a_s$ 可以表示为

$$a_s = \text{Sigmoid}(\text{sconv}(\text{cat}(C_{\text{mean}}(F_{ec}^i), C_{\text{max}}(F_{ec}^i)), k=3)) \quad (12)$$

最后,第 $i$ 个深层特征提取模块的输出 $F_{\text{out}}^i$ 可以表示为

$$F_{\text{out}}^i = (F_{ec}^i + F_{ec}^i \odot a_s) + F_{\text{in}}^i \quad (13)$$

## 1.2 纹理细节特征提取模块

由于图像超分辨率重建是一个逆问题,因此很难从LR图像中准确地恢复出其对应的HR图像,并且仅使用一个映射函数来同时恢复原始LR图像中包含的所有高频和低频信息具有挑战性。为解决这一问题,设计了一个纹理细节特征提取模块来学习包括纹理和细节在内的高频信息,以提升最终重建图像的感知质量。一幅图像与其对应的模糊版本之间的差异代表了高频图像,其中包含了大量的高频纹理和细节<sup>[34]</sup>。基于上述发现,将输入的LR图像与其模糊版本求差值,得到其高频分量 $I_1^l$ ,如图3所示。将 $I_1^l$ 作为纹理细节特征提取模块的输入。本文利用经过高斯模糊后的低分辨率图像与原始低分辨率图像 $I_1$ 相减,得到的低分辨率的纹理细节图 $I_1^l$ 可以表示为

$$\begin{cases} I_{\text{blu}} = f_G(I_1) \\ I_1^l = I_1 - I_{\text{blu}} \end{cases} \quad (14)$$

式中 $f_G$ 表示高斯模糊操作。

如图1所示,纹理细节图 $I_1^l$ 分别经过1个浅层特征提取块和4个深层特征提取块进行纹理细节特征提取,深层特征提取块采用的是和结构特征提取块一样的网络结构。浅层特征提取块是由一个 $5 \times 5$ 卷积、一个PReLU激活函数和一个 $3 \times 3$ 卷积组成。 $I_1^l$ 经过浅层特征提取块得到的浅层纹理特征 $F_{\text{t}}^{1,s}$ 可以表示为

$$F_{\text{t}}^{1,s} = \text{sconv}(\text{PReLU}(\text{sconv}(I_1^l, k=5)), k=3) \quad (15)$$

式中:sconv表示等大小卷积操作, $k$ 表示卷积核大小,PReLU表示激活函数。

将提取出的浅层特征 $F_{\text{t}}^{1,s}$ 作为深层特征提取模块的输入。纹理细节特征提取模块中第 $i$ 层深度特征提取块的输出 $F_{\text{t},\text{out}}^i$ 可以表示为

$$F_{\text{t},\text{out}}^i = f_{\text{dep}}^i(F_{\text{t},\text{in}}^i) \quad (16)$$

式中: $f_{\text{dep}}^i$ 表示深层特征提取块, $F_{\text{t},\text{in}}^i$ 表示第 $i$ 个深层特征提取块的输入, $F_{\text{t},\text{in}}^i = F_{\text{t},\text{out}}^{i-1}$ , $F_{\text{t},\text{in}}^1 = F_{\text{t}}^{1,s}$ 。纹理细节特征提取模块有4个深度特征提取块,其最后的输出为 $F_{\text{t}}^{1,d} = F_{\text{t},\text{out}}^4$ 。

## 1.3 融合模块

为了进一步提取具有不同感受野的特征,并增强融合模块的特征提取及信息整合能力,使融合后



图3 图像分解过程的说明  
Fig.3 Explanation of image decomposition process

的特征信息能得到更好的表达,提出了一种新的残差结构用于融合结构和纹理特征,如图4所示。融合模块共包含3个二阶残差块。具体地,所提出的二阶残差块由两个分支组成,分别将两个分支不同尺度的特征拼接起来,并通过残差操作将它们与输入相融合。

首先,把经梯度信息增强后的结构特征  $F_s^g$  和纹理特征提取模块的输出  $F_t^{l,d}$  经过拼接、 $1 \times 1$  卷积,得到的特征  $F_{in}^f$  作为融合模块的输入,经过融合模块得到的特征图  $F_i$  可以表示为

$$\begin{cases} F_{in}^f = \text{sconv}(\text{cat}(F_s^g, F_t^{l,d}), k=1) \\ F_i = H_i(F_{in}^f) \end{cases} \quad (17)$$

式中  $H_i$  表示由3个二阶残差块组成的融合模块。

#### 1.4 损失函数

**重构损失:**通过  $L_1$  损失函数对标签的高分辨率图像与超分辨率重建后的高分辨率图像进行约束,确保两类图像对应的像素点大致相同,可表示为

$$L_{\text{rec}} = \|I_{\text{HR}} - I_{\text{SR}}\|_1 \quad (18)$$

式中  $I_{\text{HR}}$  和  $I_{\text{SR}}$  分别为高分辨率标签图像和超分辨率重建图像。

**梯度损失:**为保证初始的融合结果中具有和标签图像一致的梯度信息,利用如下的梯度损失函数对感知模块作进一步优化,即

$$L_{\text{grad}} = \frac{1}{H \times W} \|\nabla I_{\text{HR}} - \nabla I_{\text{SR}}\|_1 \quad (19)$$

式中  $\nabla$  为梯度算子,  $H$  和  $W$  分别表示图像的高和宽。

总损失  $L_t$  可以表示为

$$L_t = L_{\text{rec}} + L_{\text{grad}} \quad (20)$$

## 2 实验及结果分析

### 2.1 训练数据及预处理

采用 DIV2K 数据集(800 张图像)进行模型训练。本文模型能分别实现超分辨率 2 倍、3 倍和 4 倍的图像重建。首先,对 DIV2K 数据集进行随机裁剪,裁剪出的高清图片构成高分辨率标签集;然后,对高分辨率标签进行下采样构成低分辨率训练集。训练 2 倍上采样因子的网络模型时,为减少内存占用、节省运行时间,高分辨率标签集中图像的大小为  $256 \times 256$ ,低分辨率训练集中图像的大小为  $128 \times 128$ ;训练 3 倍上采样因子的网络模型时,高分辨率标签集中图像的大小为  $192 \times 192$ ,低分辨率训练集中图像的大小为  $64 \times 64$ ;训练 4 倍上采样因子的网络模型时,高分辨率标签集中图像的大小为  $180 \times 180$ ,低分辨率训练集中图像的大小为  $45 \times 45$ 。

### 2.2 网络训练及测试

在训练过程中利用重构损失和梯度损失对生成的高分辨率图像与标签图像进行损失约束,同时将学习率设置为 0.000 1,通过反向传播训练整个网络的参数,直到损失值趋于平缓,说明网络参数收敛,在此经测试将模型训练的 epoch 设置为 2 000。

在测试过程中,分别采用 Set5、Set14、BSD100 和 Urban100 数据集来验证本文方法的有效性。将高分辨率测试图像分别进行 2 倍、3 倍和 4 倍的下采样,获取相应的低分辨率图像,并用于超分辨率重建测试。为验证本文方法的有效性,实验分别选取了 8 种超分辨重建经典方法(Bicubic、SRCNN<sup>[35]</sup>、FSRCNN<sup>[36]</sup>、VDSR<sup>[34]</sup>、LapSRN<sup>[37]</sup>、LESRCNN<sup>[38]</sup>、CDC<sup>[39]</sup> 和 HCN<sup>[40]</sup>)与之进行对比。SRCNN 和 FS-

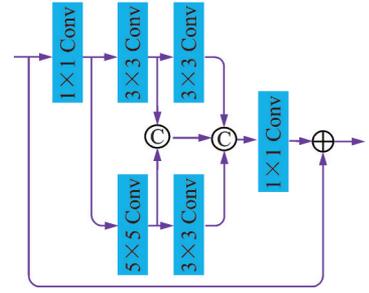


图4 二阶残差块的结构图  
Fig.4 Structure diagram of second-order residual block

RCNN采用91-image数据集进行训练。VDSR、LapSRN、LESRCNN、CDC和HCNN这5种方法均采用DIV2K数据集进行训练,其中HCNN采用Set5、Set14和BSD100来验证,没采用Urban100数据集来验证,其他对比方法都采用Set5、Set14、BSD100和Urban100数据集来验证。

选取峰值信噪比(Peak signal-to-noise, PSNR)和结构相似性(Structural similarity, SSIM)两个客观评价指标来评估超分辨率重建结果质量。PSNR用于衡量重建图像与原图像之间的均方误差,计算单位为dB,PSNR的值越大,表示图像失真越小,即图像重建质量越高;SSIM用于衡量重建图像与原图像的结构相似性,SSIM的值越高,表示重建图像越接近原图像,即图像重建效果更好。最优结果用粗体表示。

### 2.3 不同超分辨率方法对比

表1为分别在Set5、Set14、BSD100和Urban100数据集上采用不同超分辨率重建方法重建结果的定量评价结果。从这些定量评价数据可以看出,本文方法的大部分评价结果达到最优。

表1 2×、3×和4×不同方法重建结果的定量比较

Table 1 Quantitative comparison of 2×, 3× and 4× reconstruction results by different methods

| 方法      | 尺度 | Set5         |                | Set14        |                | BSD100       |                | Urban100     |                |
|---------|----|--------------|----------------|--------------|----------------|--------------|----------------|--------------|----------------|
|         |    | PSNR         | SSIM           | PSNR         | SSIM           | PSNR         | SSIM           | PSNR         | SSIM           |
| Bicubic | 2  | 33.66        | 0.929 9        | 30.24        | 0.868 8        | 29.57        | 0.843 4        | 26.88        | 0.840 3        |
| SRCNN   | 2  | 36.66        | 0.954 2        | 32.45        | 0.906 7        | 31.36        | 0.887 9        | 29.50        | 0.894 6        |
| FSRCNN  | 2  | 37.05        | 0.956 0        | 32.66        | 0.909 0        | 31.53        | 0.892 0        | 29.88        | 0.902 0        |
| VDSR    | 2  | 37.53        | 0.959 0        | 33.05        | 0.913 0        | 31.90        | 0.896 0        | 30.77        | 0.914 0        |
| LapSRN  | 2  | 37.52        | 0.959 1        | 33.08        | 0.913 0        | 31.08        | 0.895 0        | 30.41        | 0.910 1        |
| LESRCNN | 2  | 37.57        | 0.958 2        | 33.30        | 0.914 5        | 31.95        | 0.896 5        | 31.45        | <b>0.920 7</b> |
| CDC     | 2  | 32.35        | 0.876 6        | 29.20        | 0.793 2        | 28.58        | 0.785 6        | 26.06        | 0.776 6        |
| HCNN    | 2  | 37.62        | 0.959 4        | 33.03        | 0.912 7        | 31.91        | 0.896 5        |              |                |
| 本文方法    | 2  | <b>38.10</b> | <b>0.971 6</b> | <b>33.53</b> | <b>0.944 4</b> | <b>32.20</b> | <b>0.903 1</b> | <b>32.02</b> | 0.842 9        |
| Bicubic | 3  | 30.39        | 0.868 2        | 27.55        | 0.774 2        | 27.21        | 0.738 5        | 24.46        | 0.734 9        |
| SRCNN   | 3  | 32.75        | 0.909 0        | 29.30        | 0.821 5        | 28.41        | 0.786 3        | 26.24        | 0.798 9        |
| FSRCNN  | 3  | 33.18        | 0.914 0        | 29.37        | 0.824 0        | 28.53        | 0.791 0        | 26.43        | 0.808 0        |
| VDSR    | 3  | 33.67        | 0.921 0        | 29.78        | 0.832 0        | 28.83        | 0.799 0        | 27.14        | 0.829 0        |
| LapSRN  | 3  | 33.82        | 0.922 7        | 29.87        | 0.832 0        | 28.82        | 0.798 0        | 27.07        | 0.828 0        |
| LESRCNN | 3  | 34.05        | 0.923 8        | 30.16        | 0.838 4        | 28.94        | 0.801 2        | 27.76        | <b>0.842 4</b> |
| CDC     | 3  | 25.29        | 0.296 6        | 23.97        | 0.622 4        | 24.74        | 0.616 5        | 21.96        | 0.608 7        |
| HCNN    | 3  | 33.77        | 0.923 0        | 29.79        | 0.831 8        | 28.84        | 0.798 5        |              |                |
| 本文方法    | 3  | <b>35.07</b> | <b>0.940 6</b> | <b>30.22</b> | <b>0.854 3</b> | <b>29.10</b> | <b>0.791 9</b> | <b>28.06</b> | 0.722 2        |
| Bicubic | 4  | 28.42        | 0.810 4        | 26.00        | 0.702 7        | 25.96        | 0.667 5        | 23.14        | 0.657 7        |
| SRCNN   | 4  | 30.48        | 0.862 8        | 27.50        | 0.751 3        | 26.90        | 0.710 1        | 24.52        | 0.722 1        |
| FSRCNN  | 4  | 30.72        | 0.866 0        | 27.61        | 0.755 0        | 26.98        | 0.715 0        | 24.62        | 0.728 0        |
| VDSR    | 4  | 31.35        | 0.883 0        | 28.02        | 0.768 0        | 27.29        | 0.702 6        | 25.18        | 0.754 0        |
| LapSRN  | 4  | 31.54        | 0.885 0        | 28.19        | 0.772 0        | 27.32        | 0.727 0        | 25.21        | 0.756 0        |
| LESRCNN | 4  | 31.88        | 0.890 7        | 28.43        | <b>0.777 6</b> | 27.47        | 0.732 1        | 25.78        | <b>0.773 9</b> |
| CDC     | 4  | 25.41        | 0.717 7        | 24.01        | 0.616 5        | 24.19        | 0.608 5        | 21.50        | 0.561 0        |
| HCNN    | 4  | 31.39        | 0.884 9        | 28.04        | 0.767 4        | 27.29        | 0.726 0        |              |                |
| 本文方法    | 4  | <b>32.22</b> | <b>0.903 5</b> | <b>28.52</b> | 0.761 1        | <b>27.61</b> | <b>0.707 3</b> | <b>26.02</b> | 0.687 7        |

此外,本文分别选取了Set5、Set14、BSD100、Urban100数据集中的图像进行 $2\times$ 、 $3\times$ 和 $4\times$ 超分重建结果的视觉展示,如图5~7所示。从视觉效果看, SRCNN与FSRCNN方法重建图像的边缘出现了锯齿状纹理;VDSR和CDC方法的重建结果产生了与真实的真实纹理细节不相符的细节特征等。本文提出的方法能更好地拟合真实自然图像,在恢复纹理的同时还增强了重建高分辨率图像的亮度,拥有丰富的纹理细节,并且能去除重建图像的边缘尖锐化现象。实验结果表明,本文方法在视觉效果和定量评价上具有一定优越性。

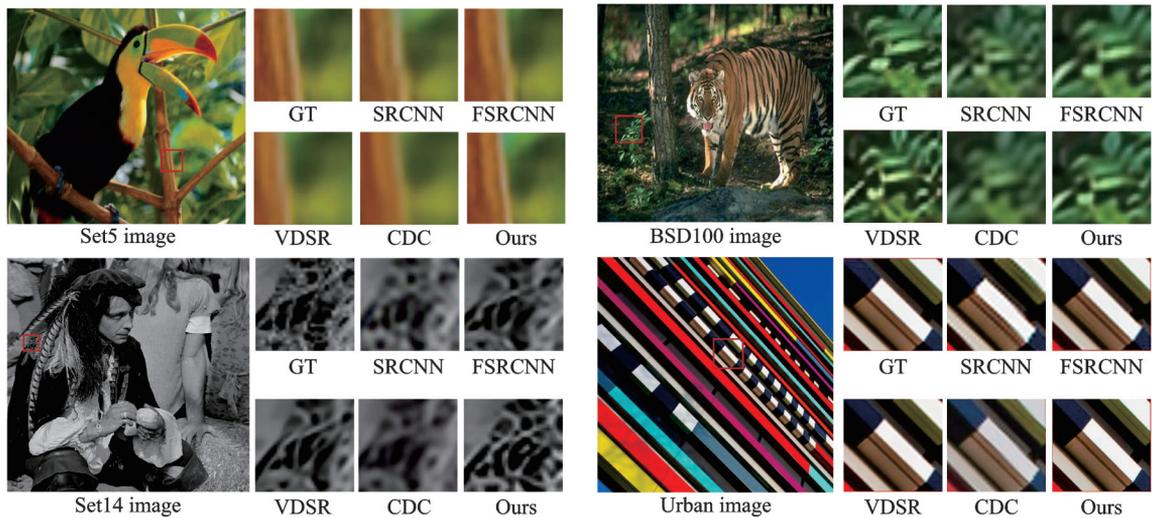


图5 不同对比方法 $2\times$ 超分辨率重建结果比较

Fig.5 Comparison of  $2\times$  super-resolution reconstruction results for different comparison methods

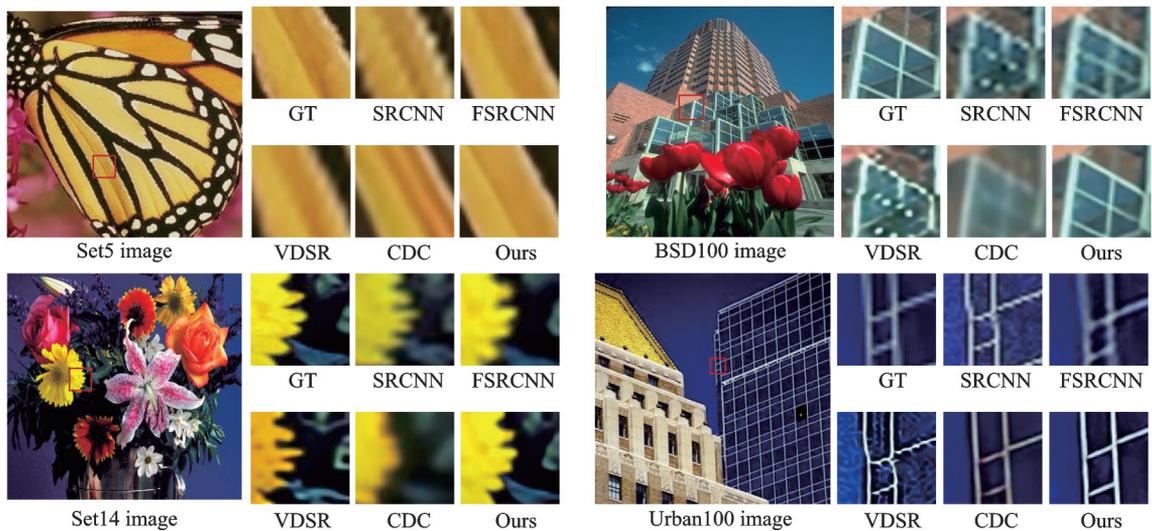


图6 不同对比方法 $3\times$ 超分辨率重建结果比较

Fig.6 Comparison of  $3\times$  super-resolution reconstruction results for different comparison methods

## 2.4 模型有效性分析

为验证本文提出的纹理细节特征提取模块(Texture detail feature extraction module, TDFEM)的有效性,深层特征提取块(Deep feature extraction block, DFEB)、融合模块(Fusion module, FM)、梯度补

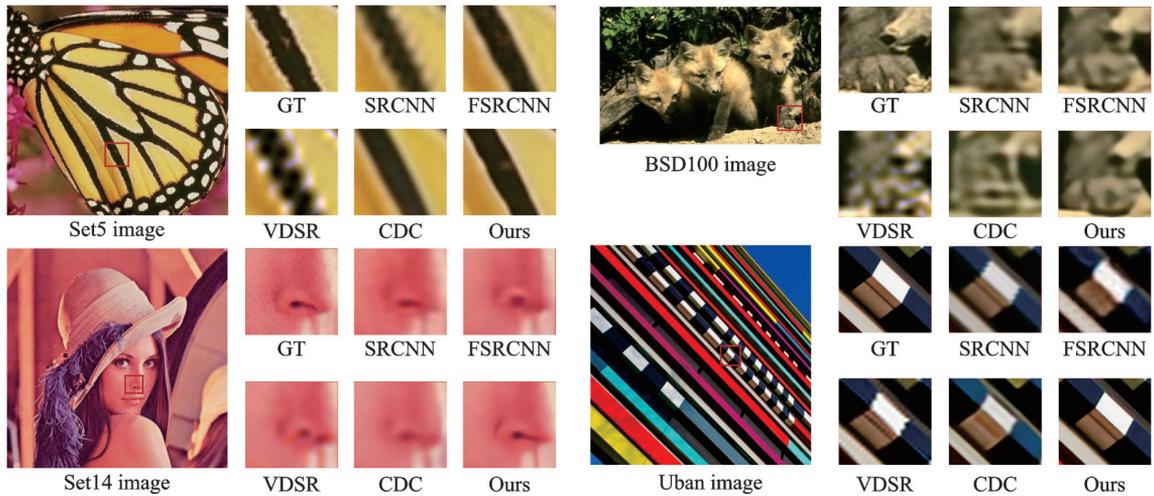


图7 不同对比方法4×超分辨率重建结果比较

Fig.7 Comparison of 4× super-resolution reconstruction results for different comparison methods

偿模块(Gradient compensation module, GCM)对超分辨率重建图像质量的影响,本节进行了消融实验。以无纹理细节特征提取模块、无深层特征提取块、无融合模块和无梯度补偿模块的网络作为基准(Base);Base加上深层特征提取模块的网络记为Base+DFEB;Base+DFEB加上梯度补偿模块的网络记为Base+DFEB+GCM;Base+DFEB+GCM加上纹理细节特征提取模块的网络记为Base+DFEB+GCM+TDFEM;Base+DFEB+GCM+TDFEM加上融合模块的网络记为Base+DFEB+GCM+TDFEM+FM。利用Urban100数据集中的图像放大3倍进行消融实验。消融实验的定量评价结果如表2所示,其视觉效果如图8所示。

(1) 深层特征提取块的有效性

本文采用深层特征提取块作为结构特征提取块和纹理细节特征提取块的基本框架,为防止超分辨率重建过程中丢失感兴趣的结构特征区域与纹理特征区域从而导致图像质量下降,在Base中引入了深层特征提取块。为了证明其有效性,将Base模型的重建结果与Base+DFEB模型的重建结果进行了比较。如图8所示,Base+DFEB模型由于在重建过程中增强了感兴趣的区域,其重建结果呈现出了更丰富的结构和纹理信息。此外,如表2所示,PSNR值提升了0.13 dB,SSIM值提升了0.007 3。DFEB的引入也使重建结果的客观评价指标取得了一定的改进。上述结果验证了DFEB在本文模型中的有效性。

表2 消融实验的定量评价结果

Table 2 Quantitative evaluation results of ablation experiment

| Method                 | Base | DFEB | GCM | TDFEM | FM | PSNR/SSIM     |
|------------------------|------|------|-----|-------|----|---------------|
| Base                   | ✓    |      |     |       |    | 25.89/0.648 1 |
| Base+DFEB              | ✓    | ✓    |     |       |    | 26.02/0.655 4 |
| Base+DFEB+GCM          | ✓    | ✓    | ✓   |       |    | 26.94/0.694 6 |
| Base+DFEB+GCM+TDFEM    | ✓    | ✓    | ✓   | ✓     |    | 27.26/0.701 6 |
| Base+DFEB+GCM+TDFEM+FM | ✓    | ✓    | ✓   | ✓     | ✓  | 28.06/0.722 2 |

(2) 梯度补偿模块的有效性

在本文模型中,梯度补偿模块被用来增强结构信息。为验证GCM的有效性,将Base+DFEB模型

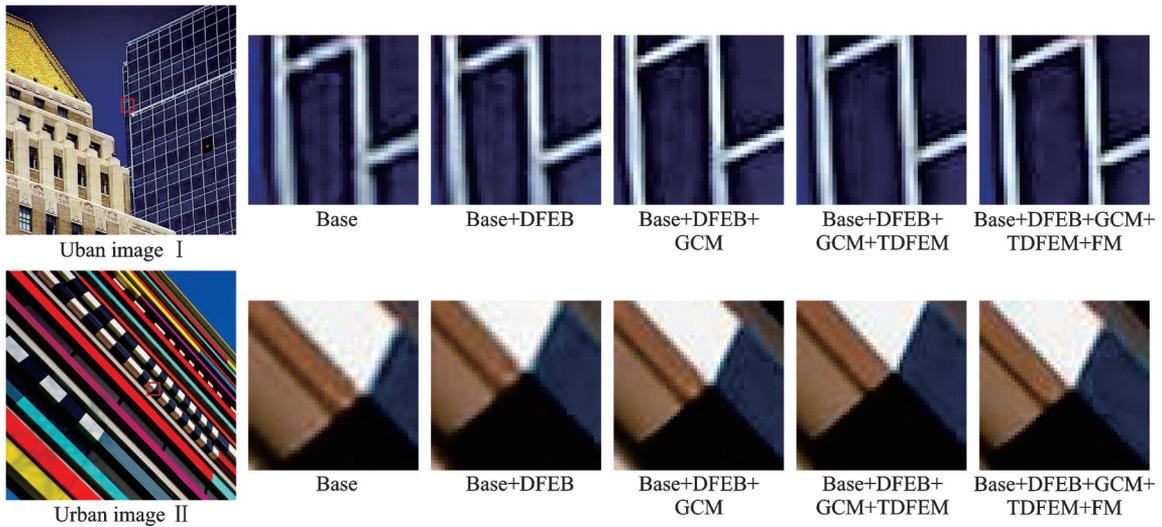


图8 消融实验的视觉效果展示

Fig.8 Visual effect display of ablation experiment

的重建结果与 Base+DFEB+GCM 模型的重建结果进行比较。如图 8 所示, Base+DFEB+GCM 模型在重建过程中增强了结构信息, 使重建结果呈现出更好的结构特征。同时, 由表 2 可知, 引入 GCM 后, 模型的 PSNR 和 SSIM 值分别提升了 0.92 dB 和 0.039 2, 证明了梯度补偿模块的有效性。

### (3) 纹理细节特征提取模块的有效性

本文采用纹理细节特征提取模块对纹理信息进行单独提取, 防止超分辨率重建过程中丢失纹理细节信息导致图像质量下降。在模型中引入了 TDFEM, 为了证明其有效性, 将 Base+DFEB+GC 模型的重建结果与 Base+DFEB+GCM+TDFEM 模型的重建结果进行比较, 详见图 8。Base+DFEB+GCM+TDFEM 模型的重建结果呈现出清晰的纹理细节信息。此外, 如表 2 所示, 引入 TDFEM 后, 模型的 PSNR 和 SSIM 值分别提升了 0.32 dB 和 0.007。上述结果验证了 TDFEM 在本文模型中的有效性, 证明纹理细节特征提取模块能够有效地提取纹理细节信息, 提升了模型的纹理保持能力。

### (4) 融合模块的有效性

为保证经梯度补偿模块增强后的结构信息以及纹理细节特征提取模块提取的纹理信息能被充分利用, 本文构建了融合模块(FM)。为验证 FM 的有效性, 将 Base+DFEB+GCM+TDFEM 与 Base+DFEB+GCM+TDFEM+FM 的重建结果进行比较。由图 8 可以看出, Base+DFEB+GCM+TDFEM+FM 模型在重建过程中充分融合了图像的结构和纹理特征, 其重建结果呈现出更好的结构和纹理细节。同时, 从表 2 可以看出, 其 PSNR 和 SSIM 分别提升了 0.80 dB 和 0.018 6。这证明融合模块能使图像的特征信息更加完善, 从而有效提高了重构图像的质量。

## 3 结束语

本文针对现有单图片超分辨率重建方法中存在单支路网络在恢复原始 LR 图像中包含的所有高频和低频信息时容易丢失纹理细节信息的问题, 利用卷积神经网络设计了一种梯度和纹理补偿的超分辨率重建网络。利用深层特征提取块中的通道注意力与空间注意力, 提升模型对特征图中高频信息的关注能力, 增强图像重建的效果。此外, 利用纹理细节特征提取模块能够有效地提取纹理细节, 改善图像的视觉质量。最后, 使用融合模块进一步融合纹理和结构特征, 能使重建图像的特征信息更加完善, 有效提高了重构图像的质量。实验结果表明, 与目前最先进的高效的 SR 方法相比, 该方法取得了最优的性能。

## 参考文献:

- [1] ZHAO Xiaole, ZOU Xueming, ZHANG Tao, et al. Channel splitting network for single MR image super-resolution[J]. IEEE Transactions on Image Processing, 2019, 28(11): 5649-5662.
- [2] MEISHVILI G, JENNI S, FAVARO P. Learning to have an ear for face super-resolution[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2020: 1364-1374.
- [3] ZHANG D, SHAO J, LI X, et al. Remote sensing image super-resolution via mixed high-order attention network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(6): 5183-5196.
- [4] KEYS R. Cubic convolution interpolation for digital image processing[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1981, 29(6): 1153-1160.
- [5] ZHANG Lei, WU Xiaolin. An edge-guided image interpolation algorithm via directional filtering and data fusion[J]. IEEE Transactions on Image Processing, 2006, 15(8): 2226-2238.
- [6] SANCHEZ-BEATO A, PAJARES G. Noniterative interpolation-based super-resolution minimizing aliasing in the reconstructed image[J]. IEEE Transactions on Image Processing, 2008, 17(10): 1817-1826.
- [7] ZHOU Fei, YANG Wenming, LIAO Qingmin. Interpolation-based image super-resolution using multisurface fitting[J]. IEEE Transactions on Image Processing, 2012, 21(7): 3312-3318.
- [8] SUN Jian, XU Zongben, SUN Jian, et al. Image super-resolution using gradient profile prior[C]//Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2008: 1-8.
- [9] MAIRAL J, BACH F, PONCE J, et al. Non-local sparse models for image restoration[C]//Proceedings of 2009 IEEE 12th International Conference on Computer Vision. [S.l.]: IEEE, 2009: 2272-2279.
- [10] TAI Y W, LIU S, BROWN M S, et al. Super resolution using edge prior and single image detail synthesis[C]//Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2010: 2400-2407.
- [11] ZHANG K, GAO X, TAO D, et al. Single image super-resolution with non-local means and steering kernel regression[J]. IEEE Transactions on Image Processing, 2012, 21(11): 4544-4556.
- [12] YANG J, WRIGHT J, HUANG T S, et al. Image super-resolution via sparse representation[J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861-2873.
- [13] TIMOFTE R, DE SMET V, VAN GOOL L. Anchored neighborhood regression for fast example-based super-resolution [C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2013: 1920-1927.
- [14] CHEN C L P, LIU L, CHEN L, et al. Weighted couple sparse representation with classified regularization for impulse noise removal[J]. IEEE Transactions on Image Processing, 2015, 24(11): 4014-4026.
- [15] LIU L, CHEN L, CHEN C L P, et al. Weighted joint sparse representation for removing mixed noise in image[J]. IEEE Transactions on Cybernetics, 2016, 47(3): 600-611.
- [16] DONG C, LOY C C, HE K, et al. Learning a deep convolutional network for image super-resolution[C]//Proceedings of European Conference on Computer Vision. Cham: Springer, 2014: 184-199.
- [17] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. [S.l.]: IEEE, 2017: 136-144.
- [18] ZHANG Y L, LI K, LI K, et al. Image super-resolution using very deep residual channel attention networks[C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich, Germany: Springer, 2018: 286-301.
- [19] LIU J, ZHANG W, TANG Y, et al. Residual feature aggregation network for image super-resolution[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2020: 2359-2368.
- [20] NIU B, WEN W, REN W, et al. Single image super-resolution via a holistic attention network[C]//Proceedings of European Conference on Computer Vision. [S.l.]: Springer, 2020: 191-207.
- [21] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2016: 1646-1654.
- [22] DONG C, CHEN C L, TANG X. Accelerating the super-resolution convolutional neural network[C]//Proceedings of European Conference on Computer Vision. [S.l.]: Springer, 2016: 391-407.
- [23] ZHANG Y, TIAN Y, KONG Y, et al. Residual dense network for image super-resolution[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2018: 2472-2481.
- [24] LI J, FANG F, MEI K, et al. Multi-scale residual network for image super-resolution[C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich, Germany: Springer, 2018: 517-532.

- [25] ZHAO H, KONG X, HE J, et al. Efficient image super-resolution using pixel attention[C]//Proceedings of European Conference on Computer Vision. [S.l.]: Springer, 2020: 56-72.
- [26] MENG B, WANG L, HE Z, et al. Gradient information distillation network for real-time single-image super-resolution[J]. *Journal of Real-Time Image Processing*, 2021, 18(2): 333-344.
- [27] LIU Cong, LEI Pengcheng. An efficient group skip-connecting network for image super-resolution[J]. *Knowledge-Based Systems*, 2021, 222: 107017.
- [28] MEHRI A, ARDAKANI P B, SAPP A D. MPRNet: Multi-path residual network for lightweight image super resolution [C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. [S.l.]: IEEE, 2021: 2704-2713.
- [29] TAN Congming, CHENG Shuli, WANG Liejun. Efficient image super-resolution via self-calibrated feature fuse[J]. *Sensors*, 2022, 22(1): 329.
- [30] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2017: 4681-4690.
- [31] WANG X, YU K, WU S, et al. Esrgan: Enhanced super-resolution generative adversarial networks[C]//Proceedings of the European Conference on Computer Vision (ECCV) Workshops. Munich, Germany: Springer, 2018.
- [32] WANG X, YU K, DONG C, et al. Recovering realistic texture in image super-resolution by deep spatial feature transform [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2018: 606-615.
- [33] SOH J W, PARK G Y, JO J, et al. Natural and realistic single image super-resolution with explicit natural manifold discrimination[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2019: 8122-8131.
- [34] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2016: 1646-1654.
- [35] DONG C, LOY C C, HE K, et al. Image super-resolution using deep convolutional networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 38(2): 295-307.
- [36] CHAO D, CHEN C L, TANG X. Accelerating the super-resolution convolutional neural network[C]//Proceedings of European Conference on Computer Vision. [S.l.]: Springer, 2016.
- [37] LAI W S, HUANG J B, AHUJA N, et al. Deep Laplacian pyramid networks for fast and accurate super-resolution[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2017: 624-632.
- [38] TIAN C, ZHUGE R, WU Z, et al. Lightweight image super-resolution with enhanced CNN[J]. *Knowledge-Based Systems*, 2020, 205: 106235.
- [39] WEI P, XIE Z, LU H, et al. Component divide-and-conquer for real-world image super-resolution[C]//Proceedings of European Conference on Computer Vision. [S.l.]: Springer, 2020: 101-117.
- [40] LIU Bangli, DJAMEL A. Effective image super resolution via hierarchical convolutional neural network[J]. *Neurocomputing*, 2020, 374: 109-116.

#### 作者简介:



黄裕青(1996-),女,硕士研究生,研究方向:图像超分辨率重建,E-mail:2277516094@qq.com。



李华锋(1983-),男,教授,研究方向:计算机视觉与图像处理相关方向。



原铭(1997-),男,硕士研究生,研究方向:图像超分辨率重建及图像融合。



张亚飞(1981-),通信作者,女,副教授,研究方向:模式识别与图像处理,E-mail:zyfeimail@163.com。