

控制与决策

Control and Decision

基于类金字塔图残差网络的图像超分辨率重建

赵小强, 王泽, 宋昭漾, 蒋红梅

引用本文:

赵小强, 王泽, 宋昭漾, 蒋红梅. 基于类金字塔图残差网络的图像超分辨率重建[J]. *控制与决策*, 2024, 39(3): 786–794.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2022.1255>

您可能感兴趣的其他文章

Articles you may be interested in

[基于自注意力生成对抗网络的图像超分辨率重建](#)

Image super-resolution reconstruction based on self-attention GAN

控制与决策. 2021, 36(6): 1324–1332 <https://doi.org/10.13195/j.kzyjc.2019.1290>

[基于双分支特征融合的场景文本检测方法](#)

A scene text detection based on dual-path feature fusion

控制与决策. 2021, 36(9): 2179–2186 <https://doi.org/10.13195/j.kzyjc.2020.0002>

[基于条件对抗生成孪生网络的目标跟踪](#)

Conditional generative adversarial siamese networks for object tracking

控制与决策. 2021, 36(5): 1110–1118 <https://doi.org/10.13195/j.kzyjc.2019.1215>

[基于图卷积网络的行为识别方法综述](#)

A survey of action recognition methods based on graph convolutional network

控制与决策. 2021, 36(7): 1537–1546 <https://doi.org/10.13195/j.kzyjc.2020.0514>

[基于改进DenseNet网络的人体姿态估计](#)

Improved DenseNet network for human pose estimation

控制与决策. 2021, 36(5): 1206–1212 <https://doi.org/10.13195/j.kzyjc.2019.1218>

基于类金字塔图残差网络的图像超分辨率重建

赵小强^{1,2,3†}, 王泽¹, 宋昭漾¹, 蒋红梅^{1,2,3}

(1. 兰州理工大学 电气工程与信息工程学院, 兰州 730050; 2. 甘肃省工业过程先进控制重点实验室, 兰州 730050; 3. 兰州理工大学 国家级电气与控制工程实验教学中心, 兰州 730050)

摘要: 针对基于深度学习的图像超分辨率重建算法大多侧重于从大量外部训练数据中学习, 而忽视图像本身的内部知识以及过于关注局部特征的问题, 提出一种基于类金字塔图残差网络的图像超分辨率重建算法. 首先, 该算法构建的残差图卷积结构利用一种预生成图结构的方式将提取的特征图转换为预生成图结构的顶点来构成图结构数据, 从而通过图卷积来学习特征自身内部的拓扑结构, 同时使用残差学习适度地加深图卷积网络以提高重建性能; 其次, 该算法构建的类金字塔多空洞卷积结构, 通过充分利用不同大小的感受野, 避免了不能完全覆盖所有像素点的缺陷, 更好地融合不同尺度的特征信息; 最后, 经过大量实验验证, 所提出的算法显著优于主流超分辨率方法, 有着更好的客观和主观度量结果.

关键词: 图像复原; 超分辨率重建; 图卷积网络; 空洞卷积; 残差学习; 金字塔结构

中图分类号: TP391

文献标志码: A

DOI: 10.13195/j.kzyjc.2022.1255

引用格式: 赵小强, 王泽, 宋昭漾, 等. 基于类金字塔图残差网络的图像超分辨率重建[J]. 控制与决策, 2024, 39(3): 786-794.

Pyramid-like graph residual network for image super-resolution reconstruction

ZHAO Xiao-qiang^{1,2,3†}, WANG Ze¹, SONG Zhao-yang¹, JIANG Hong-mei^{1,2,3}

(1. College of Electrical Engineering and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China; 2. Key Laboratory of Gansu Advanced Control for Industrial Processes, Lanzhou 730050, China; 3. National Experimental Teaching Center of Electrical and Control Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: To address the problems that the most of deep learning-based super-resolution reconstruction algorithms focus on learning from a large amount of external training data while ignoring the internal knowledge of an image itself and focusing too much on local features, a pyramid-like graph residual network is proposed for image super-resolution reconstruction. Firstly, the algorithm builds a residual graph convolution structure, which converts the extracted feature maps into vertices of pre-generated graph structures to constitute graph structure data by using a kind of pre-generated graph structure, so as to learn the internal topology of the features themselves by graph convolution, and the residuals are used to learn a moderate deepening graph convolution network to improve the reconstruction performance. Then, the algorithm builds a pyramid-like multi-dilated convolution structure, which avoids the defect of not completely covering all pixel points by making full use of different sizes of perceptual fields and better fuses feature information at different scales. Finally, experimental results show that the proposed algorithm significantly outperforms the mainstream super-resolution algorithms with better objective and subjective metric results.

Keywords: image restoration; super-resolution reconstruction; graph convolution network; dilated convolution; residual learning; pyramid structure

0 引言

单幅图像超分辨率重建^[1](single image super-resolution, SISR)是计算机视觉领域中一项重要的

图像处理技术,在卫星遥感、医学成像和人脸识别等诸多领域有着广泛应用,其旨在从相应的低分辨率(low resolution, LR)图像中重建高分辨率(high

收稿日期: 2022-07-13; 录用日期: 2022-11-10.

基金项目: 国家自然科学基金项目(62263021); 国家重点研发计划项目(2020YFB1713600); 甘肃省科技计划项目(21YF5GA072, 21JR7RA206).

责任编辑: 张国山.

[†]通讯作者. E-mail: xqzhao@lut.edu.cn.

resolution, HR) 图像. 目前单幅图像超分辨率重建技术可分为3种类型: 基于插值^[2]的方法、基于重建^[3]的方法和基于学习^[4]的方法. 由于图像超分辨率问题本身是一个多对一的关系, 即同一张HR图像在不同环境下有不同退化方向, 所以早期基于插值的方法因其使用同一个插值核而忽略图像的局部结构, 导致无法分辨图像中存在的局部复杂结构; 基于重建的方法因容易受到输入数据和随机噪声的影响, 导致其重建效果提升不显著, 因此这两种方法都难以改善图像超分辨率重建的效果. 近些年, 深度学习在计算机视觉领域表现出巨大潜力, Dong等^[5]将深度学习应用到超分辨率重建中, 提出了一种卷积神经网络的超分辨率算法 (super-resolution convolutional neural network, SRCNN), 实现了端对端的学习, 但由于只进行了3层卷积操作, 该算法提取的图像信息有限. 针对此问题, Dong等^[6]又提出了基于快速的卷积神经网络超分辨率重建算法 (fast super-resolution convolutional neural network, FSRCNN), 该算法在上采样操作中使用反卷积层代替双三次插值, 而且将网络由3层加深为8层. 随后, 许多研究者在此基础上研究重建效果更好的算法. 网络的加深会带来大量参数的计算, 针对此问题, Lim等^[7]提出了基于增强型深度残差网络的超分辨率算法 (enhanced deep residual network for single image super-resolution, EDSR), 该算法在残差网络的基础上将BN (batch norm) 层移除, 加快网络收敛速度; Li等^[8]提出了基于多尺度残差网络的超分辨率算法 (multi-scale residual network for image super-resolution, MSRN), 该算法利用多尺度特征融合和局部残差学习来充分利用图像的特征; Tai等^[9]提出了基于持续记忆网络的超分辨率算法 (a persistent memory network for image resolution, MemNet), 其利用密集连接结构构建了一个深度神经网络, 使算法具有很强的学习LR图像与HR图像之间映射函数的能力; Ahn等^[10]提出了基于级联残差网络的超分辨率算法 (fast, accurate, and lightweight super-resolution with cascading residual network, CARN), 通过使用局部和全局的级联模块避免了以往使用递归网络造成的信息损失; Wang等^[11]提出了基于自适应加权网络的超分辨率算法 (lightweight image super-resolution with adaptive weighted learning network, AWSRN), 并提出一种自适应加权多尺度重构模块, 在不影响性能的情况下, 通过自适应权值去除一些贡献较低的分支来减少网络参数.

尽管上述基于深度学习的图像超分辨率算法取得了较好的重建效果, 却仍存在一些问题, 这些算法都侧重于从外部训练数据中学习, 而没有充分利用图像本身中的内部知识学习, 同时存在卷积神经网络过于关注局部特征的缺点. 针对以上的问题, 本文提出一种基于类金字塔图残差网络 (pyramid-like graph residual network, PGRN) 的图像超分辨率算法. 首先, 构建一种残差图卷积 (residual graph convolution, RGC) 结构, 将提取的特征图转换为预生成图的顶点来构成图结构数据, 通过图卷积将图拉普拉斯正则化应用于特征图, 使其更加结构化, 进而有助于学习特征图本身的拓扑结构, 同时利用残差学习适当地加深图卷积网络, 以提高图像恢复性能; 其次, 构建一种类金字塔多空洞卷积 (pyramid-like multi-dilated convolution, PMC) 结构, 通过调整扩张率以及在不同支路设置不同金字塔层数的方式, 获得更大的感受野, 从而有助于提取非局部相似特征. 本文算法即保留了卷积神经网络的局部特征提取能力, 又包含了图卷积网络的非局部特征聚合的能力.

1 基础工作

1.1 图卷积网络

图卷积网络 (graph convolutional network, GCN) 的概念最初由Gori等^[12]提出, 并在文献[13-16]中得到进一步阐述. GCN是一种特殊的非局部方法, 通过在非欧几里得空间中建立长距离关系来处理图数据, 因其非局部特征聚合的能力而受到计算机视觉界越来越多的关注. 通过将卷积神经网络从网格数据 (如图像和视频) 拓展到图形结构化数据, 进而对图数据执行图卷积, 其中每个节点可以学习其相邻节点信息的加权平均值, 具体操作如下所示:

$$H^{(l+1)} = \sigma(\tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} H^{(l)} W^{(l)}). \quad (1)$$

其中: $\sigma(\cdot)$ 为激活函数, \tilde{A} 为添加了自连接的邻接矩阵, \tilde{D} 为度矩阵, W 为一个可训练权重的矩阵, $H^{(l)}$ 为第 l 层输出.

最近, Liu等^[17]利用预定义的邻接矩阵来进行图卷积, 用于面部表情的修复; Valsesia等^[18]基于像素构建长距离关系, 并通过为每个图节点分配固定数量邻居的方式, 将图卷积神经网络方法ECC (edge-conditioned convolution) 应用于图像去噪任务; Mou等^[19]通过图注意网络的方式来识别不同相邻节点的重要性, 从而实现图像复原. 尽管上述方法都取得了较好的效果, 但是在训练期间都需要大量计算成本来负担每次迭代中动态生成的图结构. 为了避免该问题, 本文通过一种预生成图结构的方式来避免计算成

本的增加.

1.2 空洞卷积

Yu等^[20]将空洞卷积用于语义分割,通过引入扩张率(dilation rate)这一参数,使得卷积核获得更大的感受野,并保持参数不变.普通卷积操作定义为

$$(I * \varpi)_{(t)} = \sum_{p+q=t} I_{(p)} \varpi_{(q)}. \quad (2)$$

其中: $*$ 表示卷积操作, I 表示输入信号, ϖ 表示卷积核.由此可得空洞卷积操作定义为

$$(I *_l \varpi)_{(t)} = \sum_{p+lq=t} I_{(p)} \varpi_{(q)}. \quad (3)$$

其中: $*_l$ 表示空洞卷积操作, l 表示扩张率.

目前,空洞卷积已成功应用于许多高级视觉任务中.Zhou等^[21]提出了一种级联放大模块,使用不同扩张率的卷积层进行医学图像分割;然后,Brehm等^[22]将空洞卷积引入到图像去模糊任务中,使用多个并行具备不同扩张率空洞卷积来恢复更清晰的图像,并取得了优异的效果.但这些方法所使用的扩张率方式都遵循语义分割规则,因此会存在网络并不能完全覆盖所有像素点的问题,导致图像仍然模糊.为了避免该问题,本文通过设计不同扩张率的方式,以获得几乎完全覆盖感受野的效果.

2 类金字塔图残差网络

2.1 残差图卷积结构

研究^[13-16]表明,图神经网络的主要优点是表达信息之前的传播和数据之间的交互,因此图神经网络是表示输入数据内部关系的强大工具.在卷积神经网络的正向传播过程中,由卷积产生的特征可用于推断其拓扑关系,这是由于在产生的特征中某些特征可能更为重要,而后续卷积层可能更依赖于这些关键特征,这些拓扑关系可以通过图神经网络来表示.因此受文献[23]的启发,本文提出一种RGC结构来学习特征本身的这种拓扑关系.

在训练期间,图神经网络每次迭代中生成图结构需要大量计算,并且可能产生不稳定的结果,因此为了提高效率,本文使用Watts-Strogatz模型^[24]来随机生成一种预生成的图结构,以此来减轻动态生成图结构的负担.对于预生成的图结构,将卷积转换为图卷积的操作只需要少量计算成本,且预生成的图结构具有小世界网络特性,其平均最小路径通常很小,并且会生成一些中心节点,以反映特征的重要性.本文根据特征数量来生成相应的中心节点,例如对于一个有128通道的特征图,预生成的图具有128个节点.同时,根据度中心性^[25]可以证明不同的随机图不会导

致图结构的属性存在明显差异,不会导致网络训练结果存在太大偏差.

在RGC结构中,本文将由卷积层产生的特征映射转移至一组独立的节点中,然后通过一个预生成的图结构来连接这些特征映射,此过程如图1所示.此时,由卷积提取的特征将转换为一个结构化的图结构,进而通过图卷积进行处理,将图拉普拉斯正则化应用于特征图,使其更加结构化,进而学习特征图本身的拓扑关系.随后,将数据从节点以相同顺序输回特征映射中,该过程如图2所示.

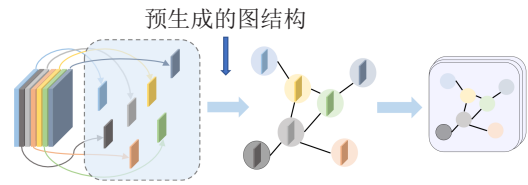


图1 特征图转换为图结构数据

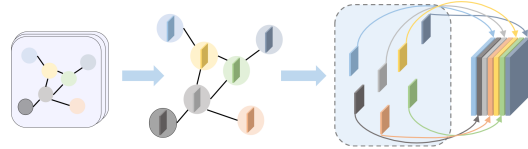


图2 图结构数据转换为特征图

为了避免特征图与图节点之间的转换过程中维度和计算复杂度的增加,本文加入一个图特征维度 F ,设给定输入特征图 $X \in R^{B \times C \times H \times W}$,其中 B 、 C 、 H 、 W 表示输入特征的批量大小、数量、高度和宽度.然后将特征图 $X \in R^{B \times C \times H \times W}$ 转换为 $X^* \in R^{B \times H \times W \times C}$ 形式,并加入 F 作为图特征维度,即 $X^* \in R^{B \times H \times W \times C \times F}$,此时转换后的数据都可以包含在 $C \times F$ 中, X^* 中其他元素都为一个左乘法的特征图.随后将 X^* 通过图卷积,即 $\text{GraphConv}(X^*)$,得到一个新的 $X^* \in R^{B \times H \times W \times C \times F}$,接下来去除图特征维度 F 并将 X^* 转换为 $X^* \in R^{B \times H \times W \times C}$ 形式,最后将 $X^* \in R^{B \times H \times W \times C}$ 转换为 $X \in R^{B \times C \times H \times W}$.

通常图卷积在训练过程中需要通过聚合器从每个节点的相邻节点和更新器收集节点信息来更新节点权重,由于本文使用的是一种无向图的图结构,根据文献[16],本文使用的聚合器如下所示:

$$H = I_N + D^{-1/2} A D^{-1/2} X. \quad (4)$$

其中: A 为无向图的邻接矩阵, I_N 为身份矩阵, N 为节点数量, D 为度矩阵, H 为聚合器.进一步引入一种重新规范化^[16]的技巧来解决图卷积神经网络中的梯度爆炸或消失问题,即根据

$$\begin{cases} \tilde{A} = A + I_N, \\ I_N + D^{-1/2} A D^{-1/2} X \rightarrow \tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} X, \end{cases} \quad (5)$$

可以得到新的聚合器

$$H = \tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} X, \quad (6)$$

其中 \tilde{A} 表示添加了自连接的无向图的邻接矩阵。

此时, 又根据

$$X^{l+1} = H^l \Theta^l, \quad (7)$$

得到RGC结构中所用图卷积的表示, 如下所示:

$$\text{GraphConv}(X) = \tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} X \Theta. \quad (8)$$

其中: $\Theta^l \in R^{C \times F}$ 表示具备 C 个输入通道的第 l 层图卷积的 F 个滤波器的参数矩阵, T^l 表示第 l 层的聚合器, X^{l+1} 表示第 l 层后的卷积矩阵。

为了进一步提高性能, 本文考虑将图神经网络转变为残差图神经网络, 根据文献[26-27]所分析的, 该方法是有用的, 但是本文与文献[27]不同的是将影响超分辨率重建性能的规范化消除, 基于式(8), 得到RGC结构中所用的残差图卷积块(residual graph convolution block, RGCB)的表示, 如下所示:

$$X^{\text{out}} = \text{GraphConv}(\alpha(\text{GraphConv}(X^{\text{in}}))) + X^{\text{in}}, \quad (9)$$

其中 $\alpha(\cdot)$ 表示LReLU激活函数。

2.2 类金字塔多空洞卷积结构

先前的工作^[28]表明, 普通卷积神经网络在进行特征提取的过程中存在过于关注局部特征的缺点, 并在全局信息提取上存在一定的局限性。为此, 本文提出一种PMC结构, 共设置金字塔层数分别为1、2、3的PMC结构, 其中层数为3的PMC结构如图3所示。

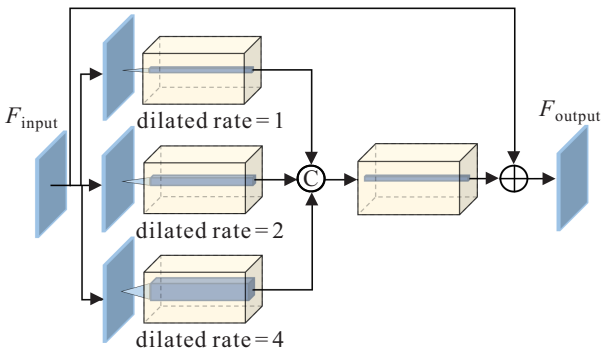


图3 PMC结构

本文所提出的PMC结构使用具有不同扩张率的空洞卷积来获得具有不同感受野的特征, 从而使网络获得更大的感受野, 进而有助于算法提取非局部相似特征并恢复清晰的图像。与以往使用空洞卷积来获取大感受野的超分辨率算法不同的是, 为了避免不能完全覆盖所有像素点的缺陷, 本文通过调整扩张率以及在不同支路设置不同的金字塔层数, 以获得几乎完

全覆盖感受野的效果。即当金字塔层数为3时, 采用的扩张率为1、2、4; 当金字塔层数为2时, 采用非重叠扩张率为1、3; 当金字塔层数为1时, 采用的扩张率为1。以金字塔层数为3的PMC结构为例, 其具体过程如下所示:

$$\begin{cases} F_{d_1} = H_{d_1}(F_{\text{input}}), \\ F_{d_2} = H_{d_2}(F_{\text{input}}), \\ F_{d_3} = H_{d_3}(F_{\text{input}}). \end{cases} \quad (10)$$

其中: F_{input} 表示输入特征图, $H_{d_1}(\cdot)$ 、 $H_{d_2}(\cdot)$ 和 $H_{d_3}(\cdot)$ 分别表示扩张率为1、2和4的空洞卷积操作, F_{d_1} 、 F_{d_2} 和 F_{d_3} 表示具有不同扩张率的空洞卷积的输出。紧接着, 使用一个 3×3 卷积来融合不同感受野的特征, 如下所示:

$$\begin{aligned} F_{d_{\text{cat}}} &= \text{Concat}(F_{d_1}, F_{d_2}, F_{d_3}), \\ F_{\text{fusion}} &= H_{\text{fusion}}(F_{d_{\text{cat}}}). \end{aligned} \quad (11)$$

其中: $\text{Concat}(\cdot)$ 表示通道拼接操作, $F_{d_{\text{cat}}}$ 表示通道拼接后的特征, $H_{\text{fusion}}(\cdot)$ 表示用于融合特征的空洞卷积, F_{fusion} 表示融合特征。最后将融合的特征叠加到输入特征上得到输出 F_{output} , 如下所示:

$$F_{\text{output}} = F_{\text{fusion}} + F_{\text{input}}. \quad (12)$$

2.3 网络结构

本文提出一种基于类金字塔图残差网络的图像超分辨率算法。该算法网络框架如图4所示, 由3部分组成, 分别为浅层特征提取单元、图融合单元和图像重建单元。其中浅层特征提取单元由一个 3×3 的卷积层组成, 图融合单元由一个类金字塔多空洞卷积结构、两个残差结构、一个残差图卷积结构及一个 3×3 的卷积层组成, 图像重建单元由上采样模块和重建模块组成。

step 1: 假设 I_{LR} 和 I_{SR} 为输入的低分辨率图像和重建的高分辨率图像, 首先通过初始卷积层从低分辨率图像中提取初始特征 x_0 , 如下所示:

$$x_0 = f_0(I_{\text{LR}}). \quad (13)$$

其中: $f_0(\cdot)$ 为浅层特征提取函数, x_0 为提取的特征图。

step 2: 将浅层特征信息传递到层数不同的类金字塔多空洞卷积结构, 这一过程如下所示:

$$\begin{cases} x'_1 = x''_1 = x'''_1 = x_0, \\ x'_2 = f'(x'_1), \\ x''_2 = f''(x''_1), \\ x'''_2 = f'''(x'''_1). \end{cases} \quad (14)$$

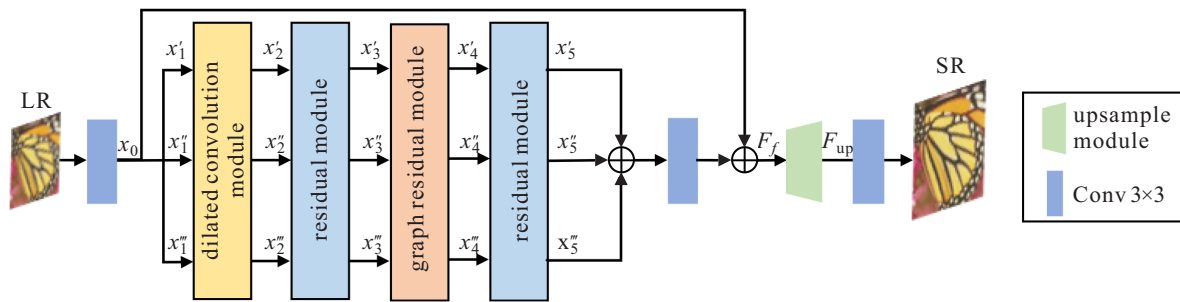


图4 本文算法网络结构

其中: x'_1 、 x''_1 和 x'''_1 分别表示第1、2和3条支路输入的特征图, $f^1(\cdot)$ 、 $f''(\cdot)$ 和 $f'''(\cdot)$ 分别表示层数为1、2和3的类金字塔空洞卷积函数, x'_2 、 x''_2 和 x'''_2 分别表示不同支路所提取的特征图。

step 3: 将提取到的特征图通过由8个残差块组成的残差结构进行进一步的特征提取,如下所示:

$$\begin{cases} x'_3 = (f_R^7(f_R^6(\dots f_R^0(x'_2)\dots))), \\ x''_3 = (f_R^7(f_R^6(\dots f_R^0(x''_2)\dots))), \\ x'''_3 = (f_R^7(f_R^6(\dots f_R^0(x'''_2)\dots))). \end{cases} \quad (15)$$

其中: $f_R^i(\cdot)$ 表示第 i 个残差块, x'_3 、 x''_3 和 x'''_3 表示其提取的特征图。

step 4: 将所提取的特征图转换为图结构数据,并通过由5块残差图卷积块组成的残差图卷积结构提取特征的拓扑关系,使其更加规格化,该过程如下所示:

$$\begin{cases} x'_4 = (f_G^4(f_G^3(\dots f_G^0(x'_3)\dots))), \\ x''_4 = (f_G^4(f_G^3(\dots f_G^0(x''_3)\dots))), \\ x'''_4 = (f_G^4(f_G^3(\dots f_G^0(x'''_3)\dots))). \end{cases} \quad (16)$$

其中: $f_G^i(\cdot)$ 表示第 i 个残差图卷积块, x'_4 、 x''_4 和 x'''_4 表示其提取的图结构数据特征。

step 5: 将图结构转换为特征图,并通过由8块残差块组成的残差结构,进一步进行深度特征提取,该过程如下所示:

$$\begin{cases} x'_5 = (f_R^7(f_R^6(\dots f_R^0(x'_4)\dots))), \\ x''_5 = (f_R^7(f_R^6(\dots f_R^0(x''_4)\dots))), \\ x'''_5 = (f_R^7(f_R^6(\dots f_R^0(x'''_4)\dots))). \end{cases} \quad (17)$$

其中: $f_R^i(\cdot)$ 表示第 i 个残差块, x'_5 、 x''_5 和 x'''_5 表示其提取的特征图。

step 6: 将提取到的深层特征信息,通过中间卷积层融合,同时通过跳线连接的方式来避免梯度爆炸等情况,如下所示:

$$F_f = f_{3 \times 3}(x'_5 + x''_5 + x'''_5) + x_0, \quad (18)$$

其中 $f_{3 \times 3}(\cdot)$ 表示为中间 3×3 卷积核的卷积层运

算。同时上采样模块通过亚像素卷积对得到特征信息 F_f 进行上采样,如下所示:

$$F_{up} = f_{up}(F_f). \quad (19)$$

其中: $f_{up}(\cdot)$ 表示为上采样运算, F_{up} 表示得到的上采样特征。

step 7: 重建模块 $f_{rec}(\cdot)$ 对上采样特征进行重建,产生高分辨率结果 I_{SR} ,如下所示:

$$I_{SR} = f_{rec}(F_{up}). \quad (20)$$

3 实验结果与分析

3.1 实验配置

本文使用 DIV2K 数据集^[29] 作为训练数据集,其包含 800 张训练图像。训练所用 LR 图像由 HR 图像的双三次下采样获得,同时使用 90° 、 180° 、 270° 旋转以及水平翻转进行数据增强;将 LR 图像裁剪成尺寸为 16×16 的图像块,将 HR 图像裁剪成尺寸为 $16s \times 16s$ 的图像块,其中 s 表示放大因子;使用 Matlab^[30] 生成 $\rho = 0.9$ 的 Watts-Strogatz 图,作为预生成的图结构。对于测试阶段,使用 4 个标准数据集: Set 5^[31]、Set 14^[32]、BSD 100^[33] 和 Urban 100^[34], 包含风景、动物、人像及城市建筑等风格的图像。本文在 RGB 通道上训练网络,在 YCrCb 空间中的 Y 通道上进行测试,使用峰值信噪比 (peak signal to noise ratio, PSNR) 和结构相似度指数 (structural similarity, SSIM) 进行评估。

网络训练所用编程框架为 Pytorch 1.2, 处理器为 Intel(R) Xeon(R) CPU E5-2698v4, 显卡为一张 NVIDIA V100, 32 G 显存, 系统内存为 48 G。本文使用 Adam 优化算法^[35] 进行优化, 其具体参数设置为 $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$ 。使用函数为本文网络的损失函数, 图像批处理大小设置为 16, 整个网络共训练 1 000 个 epoch, 初始学习率为 0.000 1, 每隔 200 个 epoch 学习率降为原来的一半。

3.2 消融分析

为了保证消融实验的公平性, 所有实验训练批次为 400 个 epoch, 并在 Set 5、Set 14 和 BSD 100 数据集上进行放大倍数为 4 时的平均 PSNR 值对比, 各结果

中加粗为最优结果.

3.2.1 不同结构对重建结果的影响

为了验证PGRN各部分的性能, 本文将PGRN分为Baseline、woPMC、woRGC、PGRN四个不同类型. 其中: Baseline表示包含16个残差块的基准网络模型; woPMC表示无PMC结构的模型; woRGC表示无RGC结构的模型; PGRN表示包含PMC结构和RGC结构的模型. 实验结果如表1所示.

表1 不同节点个数对重建性能的影响

method	RGC结构	PMC结构	PSNR		
			Set 5	Set 14	BSD 100
Baseline	×	×	32.01	28.47	27.48
woPMC	✓	×	32.08	28.57	27.55
woRGC	×	✓	32.04	28.50	27.50
PGRN	✓	✓	32.18	28.61	27.57

由表1可知: 与Baseline基准网络模型相比, woRGC模型平均PSNR值提升了0.03 dB、0.01 dB、0.01 dB, woPMC模型平均PSNR值提升了0.07 dB、0.10 dB、0.08 dB, 说明了RGC结构和PMC结构在图像超分辨率重建中的有效性; 此外, PGRN模型与Baseline基准网络模型的实验结果表明, 当通过PMC结构的特征图再通过RGC结构, 会更有利于挖掘图像本身内部的拓扑结构.

3.2.2 图结构中节点个数的影响

图卷积中节点个数的多少能够直接影响其从相邻节点中所学习到的信息, 从而对超分辨率重建性能产生影响, 为了探究最佳实验结果, 本文对预生成图结构中的节点个数进行研究. 对节点个数分别为32、64与128进行比较, 实验结果如表2所示.

表2 不同节点个数对重建性能的影响

method	PSNR		
	Set 5	Set 14	BSD 100
PGRN-32	32.03	28.49	27.51
PGRN-64	32.08	28.56	27.54
PGRN-128	32.18	28.61	27.57

由表2可知, 相较于PGRN-64与PGRN-32, PGRN-128的平均PSNR值分别提升了0.10 dB、0.05 dB、0.03 dB与0.15 dB、0.12 dB、0.06 dB, 说明当预生成图结构中的节点个数增多时, 网络能够更加充分地学习图像内部信息.

3.2.3 残差学习的影响

为了分析将残差学习运用到图神经网络中对图像超分辨率重建性能的影响, 本文将RGC结构中的RGCB分别设置为无残差学习的结构表示GCN, 以及存在残差学习的结构表示ResGCN, 实验结果如表3

所示.

表3 残差学习对重建性能的影响

method	PSNR		
	Set 5	Set 14	BSD 100
GCN	32.09	28.58	27.55
ResGCN	32.18	28.61	27.57

从表3能够得出, 相比于GCN, ResGCN的平均PSNR值提升了0.09 dB、0.03 dB、0.02 dB, 说明将残差学习运用到图神经网络中更有利于网络挖掘图像内部隐含的关联性.

3.2.4 残差图卷积结构深度的选择

为了探索RGC结构中RGCB个数(表示为 Q)对重建效果的影响, 共设置8组对比实验($Q = 1, 2, \dots, 8$), 结果如表4所示.

表4 不同RGCB个数对重建性能的影响

Q	PSNR		
	Set 5	Set 14	BSD 100
1	32.08	28.57	27.55
2	32.10	28.60	27.56
3	32.13	28.56	27.53
4	32.17	28.60	27.56
5	32.18	28.61	27.57
6	32.17	28.59	27.56
7	32.16	28.58	27.56
8	32.17	28.58	27.56

从表4可以看出, 随着RGCB个数的增加, 网络之间的性能逐渐趋于稳定, 其中 $Q = 5$ 的结果最好, 同时考虑到计算成本, 最终确定 $Q = 5$ 为一个合适的选择. 为了进一步精确RGCB个数的影响, 将第1支路通过5个RGCB后输出的特征图与通过2个RGCB

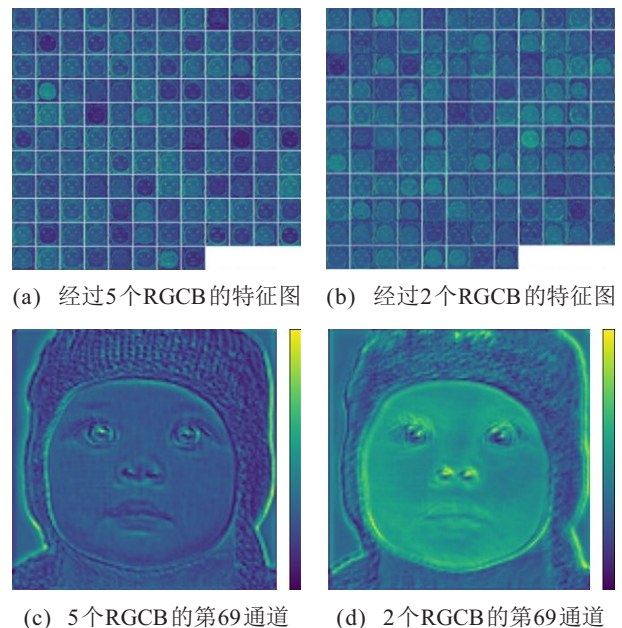


图5 特征图对比

后输出的特征图进行可视化对比,如图5所示.通过图5可以发现,通过5个RGCB后的特征图明显包含更多连续的高频信息,并在很大程度上抑制了低频信息.为了更直观地进行对比,将特征图的第69通道单独拿来对比,可以发现通过5个RGCB后的特征图在脸部边缘与细节和帽子纹理等高频信息上更为突出,说明适当地加深RGC结构,能够使算法更好地学习特征图本身的拓扑关系,提升算法重建性能.

3.3 主观与客观评估

为了验证本文算法的有效性,在Set 5、Set 14、BSD 100和Urban 100标准测试集上,将本文算法与Bicubic、EDSR-baseline^[7]、CARN^[10]、AWSRN-S^[11]、SeaNet-baseline^[36]和Cross-SRN^[37]进行2倍、3倍、4倍的性能对比测试,得出平均PSNR值和平均SSIM值,定量评估结果如表5所示.其中:加粗为最优结果,下划线为次优解.

表5 不同SR算法在放大倍数为2、3、4时的平均PSNR值与SSIM值

method	PSNR/SSIM			
	Set 5	Set 14	BSD 100	Urban 100
Bicubic	33.68/0.9304	30.24/0.8691	29.56/0.8435	26.88/0.8405
EDSR-baseline	37.99/0.9604	33.57/0.9175	32.16/0.8994	31.98/0.9272
AWSRN-S	37.75/0.9596	33.31/0.9151	32.00/0.8974	31.39/0.9207
CARN	37.76/0.9590	33.52/0.9166	32.09/0.8978	31.92/0.9256
SeaNet-baseline	37.99/0.9607	33.60/0.9174	32.18/0.8995	32.08/0.9276
Cross-SRN	38.03/0.9606	<u>33.62/0.9180</u>	32.19/0.8997	32.28/0.9290
PGRN(ours)	<u>38.02/0.9605</u>	33.62/0.9181	<u>32.18/0.8995</u>	<u>32.20/0.9285</u>
Bicubic	30.93/0.8682	27.55/0.7742	27.21/0.7385	24.46/0.7349
EDSR-baseline	34.37/0.9270	30.28/0.8417	29.09/0.8052	28.15/0.8527
AWSRN-S	34.02/0.9240	30.09/0.8376	28.92/0.8009	27.57/0.8391
CARN	34.29/0.9255	30.29/0.8407	29.06/0.8034	28.06/0.8493
SeaNet-baseline	34.36/0.9280	30.34/0.8428	29.09/0.8053	28.17/0.8527
Cross-SRN	34.43/0.9275	30.33/0.8417	<u>29.09/0.8050</u>	28.23/0.8535
PGRN(ours)	<u>34.38/0.9268</u>	<u>30.33/0.8419</u>	29.09/0.8053	<u>28.19/0.8529</u>
Bicubic	28.42/0.8104	26.00/0.7027	26.96/0.6675	23.14/0.6577
EDSR-baseline	32.09/0.8938	28.58/0.7813	27.57/0.7357	26.04/0.7849
AWSRN-S	31.77/0.8893	28.35/0.7761	27.41/0.7304	25.56/0.7678
CARN	32.13/0.8937	28.60/0.7806	27.58/0.7349	26.07/0.7837
SeaNet-baseline	32.18/0.8948	<u>28.61/0.7822</u>	27.57/0.7359	26.05/0.7896
Cross-SRN	<u>32.24/0.8954</u>	28.59/0.7817	<u>27.58/0.7364</u>	<u>26.16/0.7881</u>
PGRN(ours)	32.25/0.8955	28.64/0.7825	27.60/0.7365	26.17/0.7885

由表5可得,分别在放大因子为2、3、4情况下的Set 5、Set 14、BSD 100和Urban 100测试集上,相比其他6种主流算法,本文所提出的PGRN算法都获得了最优或次优的平均PSNR值和平均SSIM值.例如,当放大因子为2时,PGRN算法在Set 14测试集上获得了最高的平均SSIM值,比Cross-SRN算法获得的次优平均SSIM值高0.0001;当放大因子为3时,PGRN算法在BSD 100测试集上获得了最高的平均SSIM值,比Cross-SRN算法获得的次优平均SSIM值高0.0003;特别地,当放大因子为4时,PGRN算法在4种测试集上都获得了最高的平均PSNR值与SSIM值,与次优结果相比分别提升了0.01 dB、0.03 dB、0.02 dB及0.01 dB.实验结果表明,PGRN算法能够更好地适应重建难度更大的放大倍数的图像,同时更善于重建存在不同频段细节特征的图片.本文提出的PGRN算法之所以能够适应难度较大的放大倍数的重建,主要得益于PMC结构及RGC结构,PMC结构可以有效地提取非局部相似特征,RGC结构能够充分学习特

征图内部的拓扑关系,保留更多高频信息.

进一步进行主观效果的评估,图6为本文算法与其他对比算法在Set 14、BSD 100、Urban 100三种数据集上放大4倍的重建视觉效果图.对于Set 14中的图像barbara,其他对比算法的重建图像对于相邻书籍之间的边缘难以分清,具有严重的模糊感,而本文算法重建的图像能够较为清晰地分辨出相邻书籍之间的边缘;对于BSD 100中的图像8023,其他对比算法重建的图像中小鸟羽毛之间的纹理出现不同程度的模糊,而本文算法几乎完美地复原了小鸟羽毛之间的纹理;对于Urban 100中的图像img 005,在大厦顶端的区域,相比于重建质量好的CARN算法的图片,PGRN算法重建的图像不仅避免了几何结构的失真现象,同时构建了更规则的纹理.综上所述,PGRN算法相比主流算法不仅在PSNR值和SSIM值方面实现了性能提升,而且在重建图像的细节方面也具有较好的视觉效果.

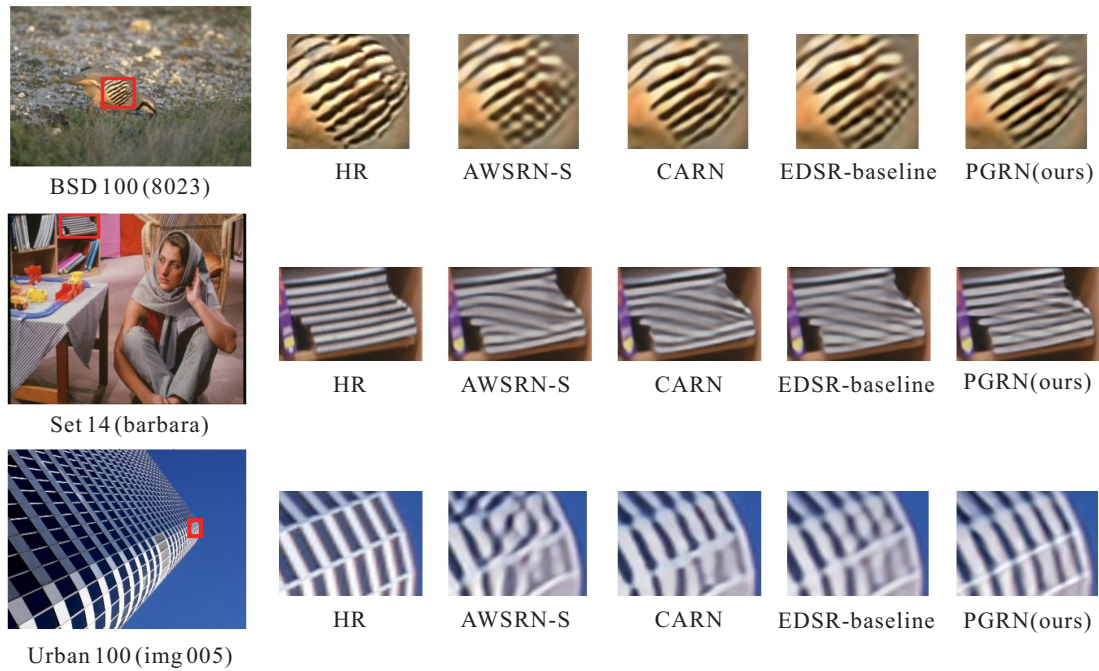


图6 标准测试集下4倍放大视觉效果比较

4 结论

针对基于深度学习的超分辨率重建算法大多侧重于从大量外部训练数据中学习,而忽视图像本身的内部先验知识,以及过于关注局部特征的问题,本文提出了一种基于类金字塔图残差网络(PGRN)的图像超分辨率重建算法。类金字塔多空洞卷积(PMC)结构通过调整扩张率以及在不同支路设置不同金字塔层数的方式,获得几乎完全覆盖感受野的效果,加强了算法捕获非局部相似特征的能力;残差图卷积(RGC)结构使用预生成图的方式,在几乎不增加计算成本的情况下将特征图转换为预生成图结构的顶点来构成图结构数据,进而使用图卷积来学习特征图本身的拓扑结构,同时利用残差学习适当地加深图卷积网络以提高图像恢复的性能。充分的实验结果验证了本文所提出PGRN算法的良好性能。

参考文献(References)

- [1] 王雪松, 晁杰, 程玉虎. 基于自注意力生成对抗网络的图像超分辨率重建[J]. 控制与决策, 2021, 36(6): 1324-1332.
(Wang X S, Chao J, Cheng Y H. Image super-resolution reconstruction based on self-attention GAN[J]. Control and Decision, 2021, 36(6): 1324-1332.)
- [2] 孙旭, 李晓光, 李嘉锋, 等. 基于深度学习的图像超分辨率复原研究进展[J]. 自动化学报, 2017, 43(5): 697-709.
(Sun X, Li X G, Li J F, et al. Review on deep learning based image super-resolution restoration algorithms[J]. Acta Automatica Sinica, 2017, 43(5): 697-709.)
- [3] 杨欣, 费树岷, 周大可. 基于频域、时域相结合的自

适应图像超分辨率重建[J]. 控制与决策, 2012, 27(9): 1343-1347.

(Yang X, Fei S M, Zhou D K. Self-adapting technology for image super-resolution reconstruction in frequency domain and time domain[J]. Control and Decision, 2012, 27(9): 1343-1347.)

- [4] 吴成东, 卢紫微, 于晓升. 基于加权随机森林的图像超分辨率算法研究[J]. 控制与决策, 2019, 34(10): 2243-2248.
(Wu C D, Lu Z W, Yu X S. Image super resolution reconstruction algorithm based on weighted random forest[J]. Control and Decision, 2019, 34(10): 2243-2248.)
- [5] Dong C, Loy C C, He K, et al. Learning a deep convolutional network for image super-resolution[C]. European Conference on Computer Vision. Zurich: ECCV, 2014: 184-199.
- [6] Dong C, Loy C C, He K M, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295-307.
- [7] 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. Honolulu, 2017: 1132-1140.
- [8] Li J C, Fang F M, Mei K F, et al. Multi-scale residual network for image super-resolution[C]. European Conference on Computer Vision. Cham: Springer, 2018: 527-542.
- [9] Tai Y, Yang J, Liu X M, et al. MemNet: A persistent memory network for image restoration[C]. IEEE International Conference on Computer Vision. Venice, 2017: 4549-4557.
- [10] Ahn N, Kang B, Sohn K A. Fast, accurate, and lightweight

- super-resolution with cascading residual network[C]. Proceedings of the European Conference on Computer Vision. Munich, 2018: 256-272.
- [11] Wang C F, Li Z, Shi J. Lightweight image super-resolution with adaptive weighted learning network[J/OL]. 2019, arXiv: 1904.02358.
- [12] Gori M, Monfardini G, Scarselli F. A new model for learning in graph domains[C]. Proceedings of 2005 IEEE International Joint Conference on Neural Networks. Montreal, 2005: 729-734.
- [13] Scarselli F, Gori M, Tsoi A C, et al. The graph neural network model[J]. IEEE Transactions on Neural Networks, 2009, 20(1): 61-80.
- [14] Gallicchio C, Micheli A. Graph echo state networks[C]. The 2010 International Joint Conference on Neural Networks. Barcelona, 2010: 1-8.
- [15] Li Y J, Tarlow D, Brockschmidt M, et al. Gated graph sequence neural networks[J/OL]. 2015, arXiv: 1511.05493.
- [16] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks[J/OL]. 2016, arXiv: 1609.02907.
- [17] Liu Z L, Li L, Wu Y P, et al. Facial expression restoration based on improved graph convolutional networks[C]. International Conference on Multimedia Modeling. Cham: Springer, 2020: 527-539.
- [18] Valsesia D, Fracastoro G, Magli E. Deep graph-convolutional image denoising[J]. IEEE Transactions on Image Processing, 2020, 29: 8226-8237.
- [19] Mou C, Zhang J, Wu Z Y. Dynamic attentive graph learning for image restoration[C]. 2021 IEEE/CVF International Conference on Computer Vision. Montreal, 2022: 4308-4317.
- [20] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions[J/OL]. 2015, arXiv: 1511.07122.
- [21] Zhou X Y, Zheng J Q, Yang G Z. Atrous convolutional neural network (ACNN) for biomedical semantic segmentation with dimensionally lossless feature maps[J]. 2019, arXiv: 1901.09203.
- [22] Brehm S, Scherer S, Lienhart R. High-resolution dual-stage multi-level feature aggregation for single image and video deblurring[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle, 2020: 1872-1881.
- [23] Xu B Y, Yin H J. Graph convolutional networks in feature space for image deblurring and super-resolution[C]. 2021 International Joint Conference on Neural Networks. Shenzhen, 2021: 1-8.
- [24] Watts D J, Strogatz S H. Collective dynamics of 'small-world' networks[J]. Nature, 1998, 393(6684): 440-442.
- [25] Borgatti S P. Centrality and network flow[J]. Social Networks, 2005, 27(1): 55-71.
- [26] Li G H, Xiong C X, Thabet A, et al. DeeperGCN: All You need to train deeper GCNs[J/OL]. 2020, arXiv: 2006.07739.
- [27] Li G H, Müller M, Thabet A, et al. DeepGCNs: can GCNs go As deep As CNNs?[C]. IEEE/CVF International Conference on Computer Vision. Seoul, 2020: 9266-9275.
- [28] Wang X L, Girshick R, Gupta A, et al. Non-local neural networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 7794-7803.
- [29] Timofte R, Agustsson E, Van Gool L, et al. Ntire 2017 challenge on single image super-resolution: Methods and results[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, 2017: 114-125.
- [30] Higham D J, Higham N J. MATLAB guide[M]. Array Philadelphia: Society for Industrial and Applied Mathematics, 2016.
- [31] 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. Guildford, 2012: 1-10.
- [32] Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations[C]. International Conference on Curves and Surfaces. Berlin, 2012: 711-730.
- [33] 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 8th IEEE International Conference on Computer Vision. Vancouver, 2002: 416-423.
- [34] Huang J B, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars[C]. IEEE Conference on Computer Vision and Pattern Recognition. Boston, 2015: 5197-5206.
- [35] Kingma D P, Ba J. Adam: A method for stochastic optimization[J/OL]. 2014, arXiv: 1412.6980.
- [36] Fang F M, Li J C, Zeng T Y. Soft-edge assisted network for single image super-resolution[J]. IEEE Transactions on Image Processing, 2020, 29: 4656-4668.
- [37] Liu Y Q, Jia Q, Fan X, et al. Cross-SRN: Structure-preserving super-resolution network with cross convolution[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(8): 4927-4939.

作者简介

赵小强 (1969—), 男, 教授, 博士生导师, 从事故障诊断、图像处理、数据挖掘等研究, E-mail: xqzhao@lut.edu.cn;

王泽 (1997—), 男, 硕士生, 从事图像超分辨率重建和图像复原等研究, E-mail: 764741162@qq.com;

宋昭漾 (1995—), 男, 博士生, 从事图像超分辨率重建和图像复原等研究, E-mail: syzlut@126.com;

蒋红梅 (1976—), 女, 副教授, 博士, 从事故障诊断、数据挖掘等研究, E-mail: 25030924@qq.com.