

DOI:10.16136/j.joel.2022.06.0693

基于门控卷积神经网络的图像超分辨率重建算法

王文安^{1,2*}, 梁新刚², 刘侍刚^{1,2}

(1. 陕西师范大学 现代教学技术教育部重点实验室, 陕西 西安 710062; 2. 陕西师范大学 计算机科学学院, 陕西 西安 710119)

摘要:近年来,卷积神经网络被广泛应用于图像超分辨率领域。针对基于卷积神经网络的超分辨率算法存在图像特征提取不充分,参数量大和训练难度大等问题,本文提出了一种基于门控卷积神经网络(gated convolutional neural network, GCNN)的轻量级图像超分辨率重建算法。首先,通过卷积操作对原始低分辨率图像进行浅层特征提取。之后,通过门控残差块(gated residual block, GRB)和长短残差连接充分提取图像特征,其高效的结构也能加速网络训练过程。GRB中的门控单元(gated unit, GU)使用区域自注意力机制提取输入特征图中的每个特征点权值,紧接着将门控权值与输入特征逐元素相乘作为GU输出。最后,使用亚像素卷积和卷积模块重建出高分辨率图像。在Set14、BSD100、Urban100和Manga109数据集上进行实验,并和经典方法进行对比,本文算法有更高的峰值信噪比(peak signal-to-noise ratio, PSNR)和结构相似性(structural similarity, SSIM),重建出的图像有更清晰的轮廓边缘和细节信息。

关键词:超分辨率; 门控单元; 门控残差块; 轻量级; 卷积神经网络

中图分类号: TP391 文献标识码: A 文章编号: 1005-0086(2022)06-0637-06

Gated convolutional neural network for image super-resolution reconstruction algorithm

WANG Wen'an^{1,2*}, LIANG Xingang², LIU Shigang^{1,2}

(1. Key Laboratory of Modern Teaching Technology, Ministry of Education, Shaanxi Normal University, Xi'an, Shaanxi 710062, China; 2. School of Computer Science, Shaanxi Normal University, Xi'an, Shaanxi 710119, China)

Abstract: In recent years, convolutional neural networks have been widely used in the field of image super-resolution. The super-resolution algorithm based on convolutional neural network has some problems, such as insufficient feature extraction of image, large number of parameters and difficult training. Therefore, this paper proposes a lightweight image super-resolution reconstruction algorithm based on gated convolutional neural network (GCNN). Firstly, the shallow feature extraction of the original low-resolution image is carried out by convolution operation. Then, the gated residual blocks (GRB) and long and short residual connections fully extract image features, and its high-efficient structure can also accelerate the network training process. The gated unit (GU) in the GRB uses the regional self-attention mechanism to extract the weight of each feature point in the input feature map. And then it multiplies the gate weight by the input feature element by element as the output of the GU. Finally, high-resolution images are reconstructed using sub-pixel convolution and convolution module. Experiments are conducted on Set14, BSD100, Urban100 and Manga109 datasets. Compared with the classical methods, not only does the algorithm in this paper have higher peak signal-to-noise ratio and structural similarity, but also the reconstructed image has clearer contour edges and details.

Key words: super-resolution; gated unit (GU); gated residual block (GRB); lightweight; convolutional neural network

* E-mail: wangwen@snnu.edu.cn

收稿日期: 2021-09-30 修订日期: 2021-11-27

基金项目: 国家自然科学基金(61672333)资助项目

1 引 言

通过单幅低分辨率图像重建出与之对应且与实际情况相符的高分辨率图像的过程叫做单幅图像超分辨率重建^[1]。单幅图像超分辨率重建在医疗、军事、安防等领域有着广泛的应用^[2]。随着深度学习的发展,越来越多基于卷积神经网络的单幅图像超分辨率重建方法相继出现^[3]。

DONG 等首先将卷积神经网络应用到单幅图像超分辨率重建上,提出了 3 层卷积神经网络(super-resolution convolutional neural network, SRCNN)^[4]。为了提高图像重建精度,KIM 等通过增加网络深度并引入可以防止网络梯度消失或梯度爆炸的残差学习机制,提出了 20 层的深度卷积神经网络算法(very deep convolutional network for super-resolution, VDSR)和深度递归卷积神经网络算法(deeply-recursive convolutional network, DRCN)^[5]。但是 SRCNN 和 VDSR 将输入图像首先进行了上采样,使输入图像变大再进行卷积等操作,这导致算法的运算量增加和训练难度大。因此,为克服 SRCNN 和 VDSR 运算量大的缺点,LEDIG 等^[6]在网络的末端对图像进行上采样还将残差网络结构(residual network, ResNet)^[7]应用到图像超分辨率重建中,提出了超分辨率残差网络(super-resolution residual network, SRResNet),大大降低了网络的运算量。但 SRResNet 在残差块中使用批量归一化层(batch normalization, BN)导致了网络重建精度没有明显提高。随后,LIM 等在网络结构中删除 BN 层,设计了更加高效的残差块结构,并提出了宽超分辨率网络(enhanced deep residual networks for super-resolution, ED-SR)^[8]和深超分辨率网络(multi-scale deep residual networks for super-resolution, MDSR)^[8],因此这两个算法的重建精度有了较大的提高。近年来,涌现出越来越多的图像超分辨率重建算法,ZHANG 等提出了基于稠密连接结构的超分辨率网络(residual dense network, RDN)^[9]和基于通道注意力机制(channel attention, CA)的超分辨率网络 RCAN^[10]。

现有的基于卷积神经网络的图像超分辨率重建方法都平等地对待同一通道不同区域的特征,这严重限制了网络特征提取的能力^[11]。同时仅仅使用残差机制虽然能够改善梯度消失和梯度爆炸的问题,但随着深度的增加,网络将难以收敛,梯度消失和梯度爆炸仍然可能会出现。为了解决这个问题,本文提出了一种基于门控卷积神经网络(gated convolutional neural network, GCNN)的

图像超分辨重建算法,主要贡献如下:1) 将 DIV2K 和 Flickr2K 两个数据集合并作为训练集,采用随机水平反转和垂直翻转等策略对训练集增强,提高网络的泛化能力;并且将大尺寸图像剪切为小尺寸图像作为输入,提高网络计算速度;2) 设计了一种新颖的门控单元(gated unit, GU),门控残差块(gated residual block, GRB)和门控残差组(gated residual group, GRG)。GU 能为不同区域的特征生成不同的门控权值,将 GU 嵌入残差模块允许网络对不同区域使用不同权值进行计算,增强了网络的特征提取能力;3) 使用了长短残差连接,有效促进了网络的收敛,降低了梯度消失和梯度爆炸的风险。4) 使用亚像素卷积和残差学习机制转换低频特征为高频特征;同时对于 $\times 3$ 和 $\times 4$ 放大倍率网络使用迁移学习策略,以 $\times 2$ 放大倍率网络权重开始训练,节省了网络训练时间。

2 本文方法

2.1 GU

CA 虽然可以用不同的权值控制高低频通道信息的通路,但是对于同一个通道不同特征区域,其仍然使用了平等对待的策略,这限制了网络的性能。本文提出的 GU 对特征图不同区域的特征采用不同的权值控制信息,通过与卷积模块相结合极大地增强了网络的特征提取能力和学习能力。GU 的具体结构如图 1 所示。

对于一幅图像,高频信息与低频信息的分界线通常在几个像素范围内,大的感受野将干扰 GU 分析不同区域的能力。所以,本文设置 GU 中两个卷积层的卷积核大小均为 1×1 。提取门控权值的式为:

$$M_{GU} = S(R(X \otimes W_{GU,1}) \otimes W_{GU,2}), \quad (1)$$

式中, X 为 GU 的输入特征, $W_{GU,1}$ 和 $W_{GU,2}$ 分别表示第一层和第二层卷积权重, $\otimes$ 表示卷积运算, $R(\cdot)$ 和 $S(\cdot)$ 分别表示 ReLU 和 Sigmoid 函数, M_{GU} 为最终的门控权值矩阵。

然后,将 M_{GU} 与输入特征 X 逐元素相乘,即:

$$F_{GU} = E_{GU}(X) = M_{GU} \odot X, \quad (2)$$

式中, $\odot$ 表示逐点相乘, F_{GU} 为 GU 最终的输出特征, $E_{GU}(\cdot)$ 为 GU 函数。

图 2 对比了 Set5^[4] 数据集“Baby”图像在本文网络 GRG 1—4 GU 的 4 个门控权值图,关于本文网络结构信息将在 2.3 节详细介绍。图中门控权值取值范围在 $[0, 1]$,图像全亮点权值为 1,表示信息通路全开,图像全黑的点权值为 0,表示信息通路关闭。可以看出,随着 GU 在网络层的加深,网络关注的区域信息越来越少,这说明深层的 GU 更关注图像的细

节信息。

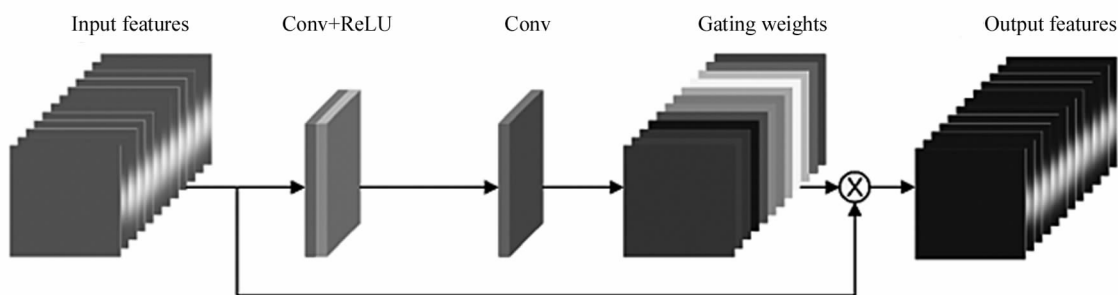


图1 GU网络结构示意图

Fig. 1 The network architecture of GU



图2 不同网络层 GU 的门控权值图

Fig. 2 The GU weight map of different network layers

2.2 GRB

GRB 通过将 GU 内嵌到残差块中,增强了网络特征提取能力和加速了网络的训练过程,具体网络结构如图 3 所示。首先将输入特征 X_n 输入到 GU,然后将 GU 的输出和 X_{n-1} 进行连接操作,减少细节信息损失:

$$X_{n-1,0} = E_{\text{cat}}(E_{\text{GU}}(X_{n-1}), X_{n-1}), \quad (3)$$

式中, $E_{\text{cat}}(\cdot)$ 表示连接函数, $X_{n-1,0}$ 为中间特征图。

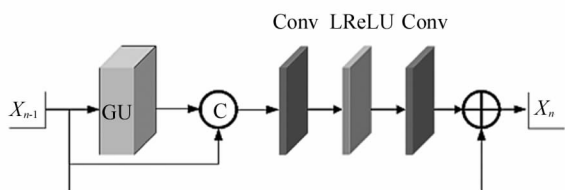


图3 GRB 结构

Fig. 3 GRB structure

然后,对 $X_{n-1,0}$ 依次进行第一层卷积操作、LReLU 激活操作和第二层卷积操作,进一步增强特征,公式为:

$$X_{n-1,1} = LR(X_{n-1,0} \otimes W_{\text{GCRB},1}) \otimes W_{\text{GDRB},2}, \quad (4)$$

式中, $W_{\text{GCRB},1}$ 和 $W_{\text{GCRB},2}$ 分别表示 GRB 的第一层和第

二层卷积权重, $LR(\cdot)$ 表示 LReLU 函数, $X_{n-1,1}$ 为增强特征。

最后,将 $X_{n-1,1}$ 与输入特征 X_{n-1} 进行残差连接,公式为:

$$X_n = E_{\text{GRB}}(X_{n-1}) = X_{n-1,1} + X_{n-1}, \quad (5)$$

式中, $E_{\text{GRB}}(\cdot)$ 表示门控残差函数, X_n 为 GRB 最终的输出特征。

2.3 网络模型

如图 4 所示,GCNN 主要由图像浅层特征提取、深层特征提取、上采样和图像重建 4 个模块组成。在本文中, I_{LR} 和 I_{SR} 分别表示 GCNN 输入的低分辨率图像和输出的高分辨图像。图像浅层特征提取主要作用是将单通道的灰度图像或者多通道的彩色图像转化为深层特征提取模块需要的多通道特征,其式为:

$$F_0 = E_{\text{SF}}(I_{\text{LR}}), \quad (6)$$

式中, $E_{\text{SF}}(\cdot)$ 表示只有一个卷积层的浅层特征提取函数。

深层特征提取单元由多个串联的 GRG 构成,每个 GRG 中又包括多个串联的 GRB,将 F_0 输入到后面的深层特征提取模块,即:

$$F_{\text{AF}} = E_{\text{DF}}(F_0) = E_{\text{GRC},J}(F_0), \quad (7)$$

式中, $E_{\text{DF}}(\cdot)$ 表示深层特征提取函数, $E_{\text{GRC},J}$ 表示 J

个串联的 GRG 函数。

进一步将增强的深层特征 F_{AF} 输入到上采样模块中,则有:

$$F_{UP} = E_{UP}(F_{AF}), \quad (8)$$

式中, $E_{UP}(\cdot)$ 表示亚像素卷积函数^[12], F_{UP} 为上采样的输出特征。上采样模块放在网络末端减少了网络

中特征信息的传输,提高了网络的运算效率。最后,将上采样特征传入由一个卷积层组成的图像重建模块,即:

$$I_{SR} = E_{GCSR}(I_{LR}) = E_{REC}(F_{UP}), \quad (9)$$

式中, $E_{REC}(\cdot)$ 和 $E_{GCNN}(\cdot)$ 分别表示图像重建和

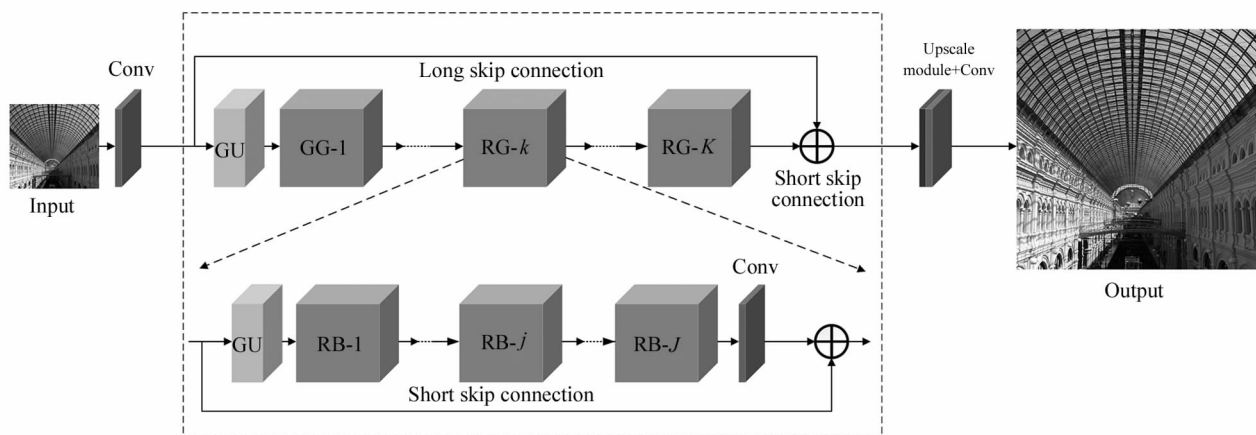


图4 GCNN 网络结构示意图

Fig. 4 The network architecture of GCNN

GCNN 函数。

2.4 网络参数

为了平衡计算资源与网络性能之间的矛盾,本文设置 GCNN 中 GRG 的数量 $J=3$ 。在每个 GRG 中 RGB 的数量 $K=4$ 。GU 中卷积核大小均为 1×1 ,其他卷积核大小为 3×3 。GCNN 深层特征提取模块中卷积通道数为 64。作为图像输出模块,图像重建模块的输出通道数为 3,它表示 RGB 图像的 3 个通道。

在先前基于卷积神经网络的超分辨率工作中已经证明了 L_1 损失函数比 L_2 、感知损失、对抗损失等损失函数更有利于网络的收敛^[8],因此本文使用 L_1 损失函数。给定包含了 N 个低分辨率和高分辨率图像对的训练集 $\{L_{LR}^n, I_{HR}^n\}_{n=1}^N$,优化目标函数为:

$$L(\Theta) = \frac{1}{N} \sum_{n=1}^N \| E_{GCNN}(L_{LR}^n) - (L_{HR}^n) \|_1, \quad (10)$$

式中, Θ 表示 GCNN 的参数集。

3 实验

3.1 数据集与实验参数

本文从 DIV2K^[9] 数据集中选取 800 幅图像,从 Flickr2K^[12] 数据集中选取 2500 幅图像,并将这两部分图像合并在一起,作为 GCNN 的训练集。在训练

过程中,使用 Set5 作为验证集。测试集使用的是 Set14^[6]、BSD100^[13]、Urban100^[14] 和 Manga109^[15] 这 4 个基准数据集。本文采用峰值信噪比(peak signal-to-noise ratio, PSNR)和结构相似性(structural similarity, SSIM)作为超分辨率重建图像的定量评价指标。

在训练中,本文将训练集图像裁剪为大小 48×48 的图像块作为网络输入数据,批次大小为 32,在每个批次中使用归一化算法对数据进行预处理,然后将处理后的图像输入到网络模型中。本文使用随机水平翻转、垂直翻转和通道变化策略来增强训练数据,以增强网络的泛化能力。所有放大倍率的网络初始学习率大小设置为 10^{-4} ,并且在学习率衰减 5 次后训练结束。 $\times 2$ 放大倍率的网络每 100 个 epoch 学习率减半, $\times 3$ 和 $\times 4$ 放大倍率的网络使用了 $\times 2$ 模型收敛后的权重作为初始训练权重,学习率每 50 个 epoch 减半。

3.2 实验结果与分析

本文算法 GCNN 与最近几年比较经典的 ED-SR-Baseline, DASR 等^[16] 算法进行了重建精度和可视化结果比较,比较结果如表 1 和图 4 所示。其中 RCAN-Baseline 设置残差组 $J=3$,残差块 $K=6$ 。从表 1 可以看出,随着放大倍率的提高,GCNN 与其他比较方法的平均重建精度差也在增大,这说明 GC-

NN 对于信息严重缺失的图像有较强的重建能力,这在图像质量较高的 Manga109 数据集体现尤为显著。与 EDSR-Baseline,RCAN-Baseline,DASR 几种方法相比,GCNN 在多个测试数据集下均能实现更好的平均重建精度,这证明了 GCNN 有更优异的图像超分辨率性能。

表 1 在测试集下不同超分辨算法的平均 PSNR(dB)/SSIM

Tab. 1 Average results of PSNR (dB),SSIM for different super-resolution methods in the test set

Data set	Magnification	Bicubic	VDSR	DASR	EDSR-Baseline	RCAN-Baseline	GCNN
Set5	×2	33.66/0.93	37.53/0.96	37.87/0.96	37.98/0.96	38.02/0.96	38.11/0.96
	×3	30.39/0.87	33.66/0.92	34.11/0.92	34.36/0.93	34.43/0.93	34.57/0.92
	×4	28.42/0.81	31.35/0.88	31.99/0.88	32.09/0.89	32.27/0.89	32.39/0.89
Set14	×2	30.24/0.87	33.03/0.91	34.34/0.92	33.56/0.92	33.65/0.91	33.80/0.92
	×3	27.55/0.78	29.77/0.83	30.13/0.84	30.28/0.84	30.33/0.84	30.46/0.84
	×4	26.00/0.70	28.01/0.76	28.50/0.78	28.57/0.78	28.67/0.78	28.76/0.78
B100	×2	29.56/0.84	31.90/0.89	32.03/0.90	32.15/0.90	32.21/0.90	32.27/0.90
	×3	27.21/0.74	28.82/0.79	28.96/0.80	29.08/0.81	29.13/0.90	29.20/0.80
	×4	25.96/0.67	27.29/0.72	27.50/0.73	27.56/0.74	27.62/0.73	27.67/0.74
Urban100	×2	26.88/0.84	30.76/0.91	31.49/0.92	31.97/0.93	32.37/0.93	32.44/0.93
	×3	24.46/0.74	27.14/0.82	27.65/0.84	28.14/0.85	28.28/0.85	28.48/0.85
	×4	23.14/0.65	25.18/0.75	25.82/0.78	26.03/0.789	26.29/0.79	26.39/0.79
Manga109	×2	30.80/0.94	37.22/0.97	38.34/0.98	38.54/0.97	38.97/0.97	38.97/0.97
	×3	26.95/0.86	32.01/0.93	33.10/0.93	33.44/0.94	33.78/0.94	33.98/0.94
	×4	24.89/0.79	28.83/0.88	30.12/0.90	30.35/0.90	30.75/0.91	30.90/0.91

在图 5 中,GCNN 相较于 Bicubic、SRCNN、VDSR、SRResNet 4 种方法能够重建更清晰的轮廓细节信息,并且在重建精度上达到最优结果。对于 EDSR-Baseline、RCAN-Baseline、GCNN 3 种方法,重建图像的视觉效果接近,但是 GCNN 相较前两种方法的 PSNR 能够平均超出 0.6 dB,这说明 GCNN 重建图像细节在位置上更贴近真实高分辨率图像。

由于本文的 GRB 中增加了 GU,因此和 EDSR-Baseline 等网络相比,GCNN 单个残差块在参数量上有少量的增加。为了保持网络轻量,GCNN 每一个 GRG 中仅使用了 4 个 GRB。表 2 列出了 GCNN 与对比方法的参数量和网络结构信息,可以看出 GCNN 在保持轻量的基础上具有更好的图像重建结果。

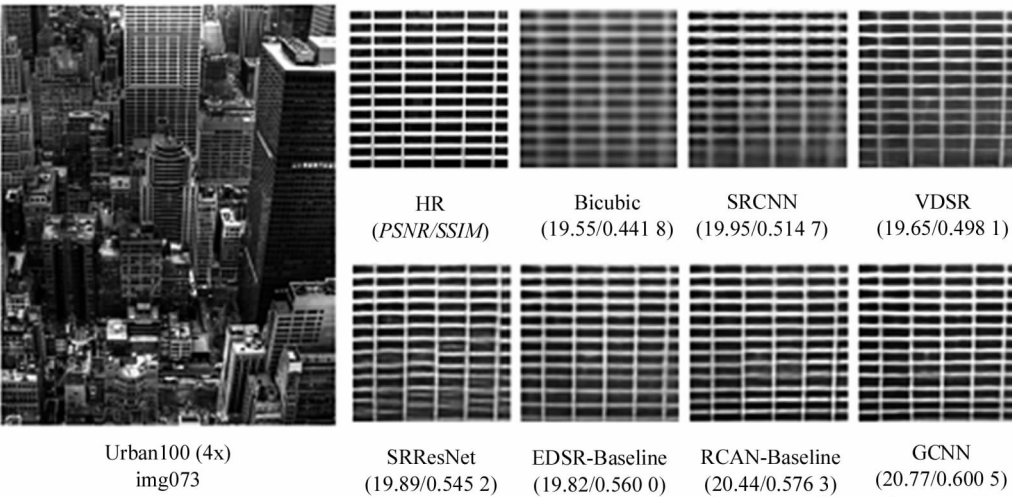


图 5 比例因子为 4 时不同方法的重建可视化结果

Fig. 5 Reconstruction visualization results from different methods with scale $\times 4$

表 2 不同方法的参数量和网络模型信息

Tab. 2 Model size and parameter of different methods

Method	Residual group	Residual block	Channel number	Parameter quantity/M
SRCNN	—	—	64	0.11
SRResNet	—	16	64	1.55
DASR	5	5	64	5.8
EDSR-Baseline	—	16	64	1.37
RCAN-Baseline	3	6	64	1.56
GCNN	3	4	64	1.21

4 结 论

现有的图像超分辨率卷积神经网络对于不同区域的特征信息均使用相同的权值进行卷积运算,即使使用了更深的网络,也难以明显提高图像超分辨率精度。本文提出了一种轻量级 GCNN 的超分辨率重建算法,该算法使用了 GU、GRB、GRG 和长短残差连接。GU 通过区域自注意力机制控制特征信息不同区域的权值,将其分别嵌入到门 GRB、GRG 中,显著提高了网络特征提取能力。长短残差连接降低了深度网络的训练难度,提高了网络收敛速度。与 DASR 等近几年经典超分辨率算法相比,本文方法在客观评价指标上实现了更高的精度,在视觉效果上能够有效减少模糊和伪影,重建出更高质量的超分辨率图像。

参考文献:

- [1] YANG W, ZHANG X, TIAN Y, et al. Deep learning for single image super-resolution: a brief review [J]. IEEE Transactions on Multimedia, 2019, 21(12): 3106-3121.
- [2] LAN R, SUN L, LIU Z, et al. Cascading and enhanced residual networks for accurate single-image super-resolution[J]. IEEE Transactions on Cybernetics, 2020, 51(1): 115-125.
- [3] HUI Z, LI J, GAO X, et al. Progressive perception-oriented network for single image super-resolution[J]. Information Sciences, 2021, 546: 769-786.
- [4] DONG C, LOY 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.
- [5] KIM J, LEE J, LEE K. Accurate image super-resolution using very deep convolutional networks[C]//IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 1646-1654.
- [6] LEDIG C, THEIS L, HUSZAR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 4681-4690.
- [7] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.
- [8] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//IEEE Conference on Computer Vision and Pattern Recognition Workshops, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 136-144.
- [9] ZHANG Y, TIAN Y, KONG Y, et al. Residual dense network for image super-resolution[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 2472-2481.
- [10] ZHANG Y, LI K, LI K, et al. Image super-resolution using very deep residual channel attention networks[C]//The 15th European Conference on Computer Vision, September 8-14, 2018, Munich, Germany. Berlin: Springer, 2018: 286-301.
- [11] LAN R, SUN L, LIU Z, et al. Madnet: a fast and lightweight network for single-image super resolution [J]. IEEE Transactions on Cybernetics, 2020, 51(3): 1443-1453.
- [12] SHI W, CABALLERO J, HUSZAR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 1874-1883.
- [13] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-complexity single-image super-resolution based on non-negative neighbor embedding [C]//Proceedings of British Machine Vision Conference, September 3-7, 2012, Surrey, UK. Guildford: BMVC, 2012: 1-10.
- [14] CAO H, MI S. Weighted SRGAN and reconstruction loss analysis for accurate image super resolution[J]. Journal of Physics: Conference Series, 2021, 1903(1): 012050.
- [15] KAR A, BISWAS P. Fast bayesian uncertainty estimation and reduction of batch normalized single image super-resolution network [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 20-25, 2021, Nashville, TN, USA. New York: IEEE, 2021: 4957-4966.
- [16] WANG L, WANG Y, DONG X, et al. Unsupervised degradation representation learning for blind super-resolution [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 20-25, 2021, Nashville, TN, USA. New York: IEEE, 2021: 10581-10590.

作者简介:

王文安 (1995—),男,硕士,主要从事超分辨率重建,图像增强方面的研究。