

双路径反馈网络的图像超分辨重建算法^①

陶 状¹, 廖晓东^{1,2,3}, 沈江红¹

¹(福建师范大学 光电与信息工程学院, 福州 350007)

²(福建师范大学 医学光电科学与技术教育部重点实验室, 福建省光子技术重点实验室, 福州 350007)

³(福建师范大学 福建省先进光电传感与智能信息应用工程技术研究中心, 福州 350007)

通讯作者: 陶 状, E-mail: 2459262167@qq.com



摘 要: 图像超分辨率重建在安防系统, 小目标检测以及医学图像等有着广泛的应用. 本文提出一种双路径反馈网络来提高图像超分辨重建的性能. 在双路径网络中, 一条路径采用深度残差稠密网络学习重建图像的高频信息, 另一条路径直接在输入图像上通过亚像素卷积层上采样到所需分辨率来给重建图像提供低频信息, 然后将两条路径得到的特征图进行融合来自适应的选取所需要的信息, 接着通过一个反馈型卷积层进行局部循环训练来获得大的感受野. 通过在数据集 DIV2K 上训练, 实验结果表明所提方法的有效性和优越性.

关键词: 超分辨率重建; 双路径反馈网络; 残差稠密网络; 特征融合; 卷积层

引用格式: 陶状, 廖晓东, 沈江红. 双路径反馈网络的图像超分辨重建算法. 计算机系统应用, 2020, 29(4): 181-186. <http://www.c-s-a.org.cn/1003-3254/7344.html>

Dual Stream Feedback Network for Image Super-Resolution Reconstruction

TAO Zhuang¹, LIAO Xiao-Dong^{1,2,3}, SHEN Jiang-Hong¹

¹(College of Photonic and Electronic Engineering, Fujian Normal University, Fuzhou 350007, China)

²(Key Laboratory for Optoelectronic Science and Technology for Medicine (Ministry of Education) Cum. Fujian Provincial Key Laboratory for Photonics Technology, Fujian Normal University, Fuzhou 350007, China)

³(Fujian Provincial Engineering Research Center for Optoelectronic Sensors and Intelligent Information, Fuzhou 350007, China)

Abstract: Image super-resolution reconstruction has a wide range of applications, such as security systems, small object detection, and medical imaging. This study proposes a dual stream feedback network to improve the performance of image super-resolution reconstruction. In the dual-stream network, one path adapts a deep residual dense network to learn the high-frequency information of the reconstructed image, and the other path directly samples the input image to the desired resolution through a sub-pixel convolution layer. Then, the feature maps obtained from the two paths are fused to adaptively selecting the required information. Finally, using a feedback convolutional layer for locally loop training to obtain a large receptive field. By training on the dataset DIV2K, the experimental results show the effectiveness and superiority of the proposed method.

Key words: super-resolution reconstruction; dual-stream feedback network; residual dense network; feature fusion; convolution layer

1 前言

图像超分辨率重建是指从低分辨率图像中获得高

分辨率图像. 在硬件限制的条件下, 往往无法获得高分辨率图像, 这时通过软件的方式来获得就显得尤为重

① 基金项目: 中央引导地方科技发展专项 (2017L3009)

Foundation item: Special Fund of Central Government for Local Science and Technology Development (2017L3009)

收稿时间: 2019-09-01; 修改时间: 2019-09-23; 采用时间: 2019-10-11; csa 在线出版时间: 2020-04-05

要。例如在医学图像中,由于成像设备的制约,产生的图像得不到令人满意的结果,而更换硬件的成本昂贵,通过软件算法生成高清图片帮助医生判断是一种很好的选择。在安防系统中拍摄的视频需要存储,由于高分辨率图像需要更多的存储空间,所以通常都是存储低分辨率图像,但是在需要时重新查看这些视频图像往往需要高清的图像来进行观察分析。在目标检测中,小目标检测更具挑战性,文献[1]验证了通过对小目标区域进行超分辨率重建可以提高检测的性能。总体来说,高分辨率图像具有更好的视觉感受,具有更丰富的细节信息,有助于图像的理解、处理和分析。超分辨率重建逐渐成为图像处理领域研究的热点问题。

图像超分辨率重建算法可以分为3类,基于插值的方法、基于重建模型的方法、基于学习的方法。其中基于插值的方法计算速度很快,但是重建的高分辨率图像的质量不好。基于重建模型的方法通常采用复杂的先验知识来限制可能的解空间,这种方法具有生成灵活和锐化细节的优势,但是当比例因子增加时,重建的性能迅速下降,并且速度通常很慢。基于学习的方法例如马尔科夫随机场、随机森林、稀疏编码等直接构建从低分辨率图像到高分辨率图像的映射关系,具有快速计算和杰出性能的优势,近年来基于深度学习的方法逐渐成为当前最具优势的算法。

由于大量可用的数据和高效计算硬件的开发,基于深度学习的图像超分辨率重建得到迅速的发展。第一次将深度学习用在图像超分辨率任务上的是 Dong 等人^[2]提出的 SRCNN,他们首先用插值法将低分辨率图像上采样到所需的高分辨率图,然后使用3层卷积神经网络进行特征提取、转换和重建得到最后的高分辨率图像,使用均方差损失函数训练网络。SRCNN用插值法进行上采样带来的细节平滑效应往往导致输入图像的结构发生错误估计,此外在一开始进行上采样会大量增加后续操作的计算量。基于这些缺点,文献[3]提出一种叫做亚像素卷积层的高效上采样方法,他们首先利用卷积操作来增加特征图的通道数,然后将特征图的像素重新组合,达到上采样的目的,其方法如图1所示。这种方法可以在网络的最后将低分辨率特征图变换到高分辨率图,网络的主体在低分辨率图上进行计算,从而大大减少了计算量加快了运行速度。

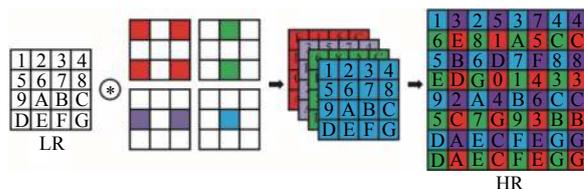


图1 亚像素卷积层示意图

从另一个方面考虑, SRCNN 只用了3层卷积神经网络,然而在计算机视觉的其他任务中,更宽更深的网络拓扑结构通常会带来更优的结果。VDSR^[4]是在超分辨率重建任务中第一个提出的非常深的模型。首先在结构上 VDSR 使用的是20层卷积神经网络,鉴于可以有效增加感受野的最小卷积核大小是3,所以在 VDSR 中所有的卷积核大小都设为3,为了有效训练深的网络, VDSR 使用相对较高的初始学习率来加速收敛并且使用了梯度裁剪技巧来防止梯度爆炸。此外 VDSR 中还提出了一个重要的观点,在低分辨率图像到高分辨率图像重建的过程中,高分辨率图像包含大量的低分辨率图像的信息,因此网络所需要学习的是低分辨率图像的高频部分,这可以通过残差操作来实现即将浅层特征图通过跳转连接和网络后端的抽象特征图相加。在之后深度学习的发展中,性能更好的残差网络和稠密网络先后被运用在图像超分辨率重建的任务中。

最近基于残差稠密网络^[5]的图像超分辨率重建算法被提出。残差稠密模块如图2所示,一方面它利用长跳转残差操作来避免在训练中梯度消失的问题,另一方面模块中所有的卷积层都被传递到后面所有的层中,这种特征重复利用的方式提高了网络的学习效率,最后所有特征图沿着通道方向拼接在一起,然后用 1×1 的卷积层压缩通道,自适应的选取有用特征。残差稠密网络在图像超分辨率中表现优异,本文所提的双路径反馈网络是对残差稠密网络的改进,进一步提高其性能。

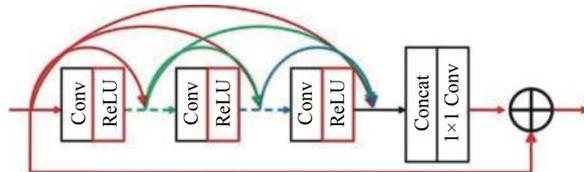


图2 残差稠密模块示意图

由于低分辨率图像与重建的高分辨率图像高度相似,因此低分辨率图像信息对重建的高分辨率图像至关重要。残差稠密网络使用长跳转连接操作将低级特

征传输到网络后端与高级特征进行融合,并没有直接利用输入的低分辨率图像。鉴于这种动机,本文所提的双路径反馈网络一条分支为残差稠密网络,另一条分支使用亚像素卷积算法直接对输入的低分辨率图像进行上采样到所需分辨率,然后两条分支信息融合并通过反馈网络循环训练。这种方法减少了网络学习的难度同时保证了重建图片的质量。

2 方法

本文所提的双路径反馈网络可分解成3部分:首先是基于残差稠密网络的高频信息提取路径,其次是基于亚像素卷积上采样的低频信息提取路径,最后是将两条路径进行信息融合后的反馈网络。

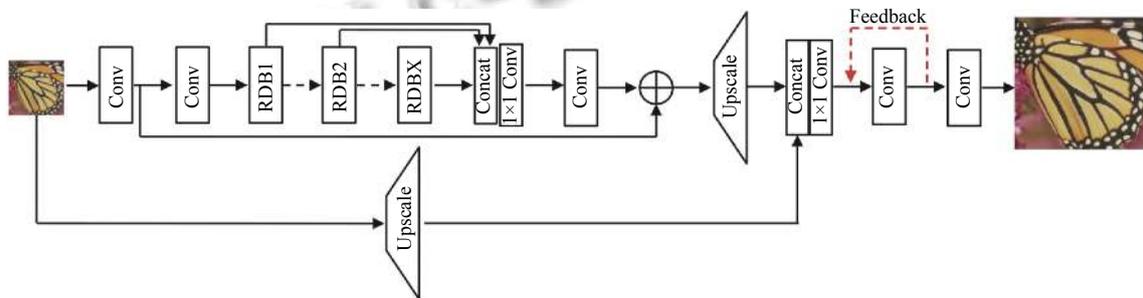


图3 双路径反馈网络的架构示意图

如图3第一条分支所示,浅层特征除了被输入到之后的卷积层,还会通过长跳转连接操作被直接传输到网络的后端,这种残差学习保证了浅层信息的流通,在反向传播时不会由于网络加深而导致的梯度消失影响到浅层参数的更新。

2.2 亚像素卷积上采样路径

一幅图像从频率角度可以看成是高频部分与低频部分的叠加,高频表示的是图像变化剧烈的地方,如轮廓、线条,低频表示的是图像变化缓慢的地方。对于人的视觉系统来说,人对低频部分的变化更敏感,高频部分的变化往往感受不到。因此在图像超分辨率重建中,保证重建图像低频部分的质量是算法需要满足的基本前提,为此本文提出亚像素卷积路径为重建图像提供低频信息,高频部分则由另一条路径学习得到。

如图3第二条分支所示,上采样模块直接作用在原始输入的低分辨率图像上,其过程如图4所示:首先通过 1×1 的卷积扩大输入图像的通道数,然后通过像素值重新安排来扩大分辨率。假设输入的图像通道数为

2.1 残差稠密网络

残差稠密网络^[5]由残差稠密模块堆叠组成,如图3中的上面一条分支所示。输入的是低分辨率图像,首先经过一个卷积层将图像转换到特征域,第一个卷积层提取的是图像的浅层特征,然后将浅层特征输入到第二个卷积层进行特征变换,接着顺序的通过一系列稠密残差模块并且每个残差稠密模块输出的特征图都会被聚集在一起。假设输出的特征图为 $F_x \in \mathbb{R}^{B \times C \times H \times W}$,其中 B 、 C 、 H 、 W 分别表示批量数、通道数、特征图的高和宽,网络中总共堆 K 个稠密残差模块,沿着通道方向将所有特征图拼接在一起得到 $F_a \in \mathbb{R}^{B \times KC \times H \times W}$,具有通道数为 KC 的特征图 F_a 之后通过 1×1 的卷积层通道数被压缩为 C ,保持和浅层特征图的数据维度一致。

C ,上采样因子为 r ,则通过 1×1 卷积需要将通道数扩张到 $C \times r^2$,之后通过像素重排得到的高分辨率图像为 $C \times rH \times rW$ 。

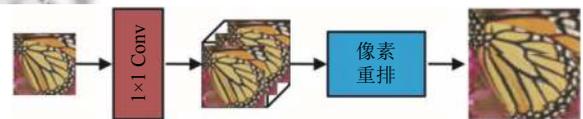


图4 上采样路径示意图

2.3 特征融合与反馈网络

所提方法整体架构的前半部分是双路径网络,一条路径为深度网络来学习高频信息,另一条路径为浅层网络提供低频信息,两条路径分别经过上采样模块后进行信息融合。融合的过程如下:高频路径的特征图为 $F_h \in \mathbb{R}^{B \times G \times H \times W}$,低频路径的特征图为 $F_l \in \mathbb{R}^{B \times C \times H \times W}$,沿着通道方向将 F_h 与 F_l 拼接得到特征图 $F_{hl} \in \mathbb{R}^{B \times (G+C) \times H \times W}$,然后通过卷积核大小为1、通道数为 G 的卷积层自适应的从高、低频信息中选择所需

的特征。

为了进一步提高高分辨率图片重建的质量,将融合后的特征图输入到反馈网络来局部循环训练,反馈网络可以在不增加参数的情况下,扩大特征图的感受野,有利于增强网络的表示能力。反馈网络可表示如下:

$$F_{out}^t = Conv[F_{in}^t] \quad (1)$$

$$F_{in}^t = F_{out}^{t-1} = Conv[F_{in}^{t-1}] \quad (2)$$

其中, $F_{out}^t, F_{in}^t, F_{out}^{t-1}, F_{in}^{t-1}$ 分别表示当前时刻模块的输出、当前时刻模块的输入、上一时刻模块的输出、上一时刻模块的输入。

2.4 损失函数

损失函数用来衡量网络预测的图像与真实图像之间的差异,然后通过梯度下降算法不断迭代优化使损失值越来越小,损失值收敛时网络预测的图像将非常接近真实图像。

本文采用 L_1 损失函数来训练所提的方法,其可以用以下公式表示:

$$L_1(\hat{I}, I) = \frac{1}{HWC} \sum_i \sum_j \sum_k \|\hat{I}_{i,j,k} - I_{i,j,k}\|_1 \quad (3)$$

其中, H, W, C 分别表示图像的高度、宽度和通道数。 \hat{I} 和 I 分别表示网络输出的图像和对应的真实图像。

3 实验

3.1 数据集与训练过程

本文使用 DIV2K 作为训练集。DIV2K 数据集包含 1000 张高质量的图像,其中 800 张为训练集,100 张为验证集,100 张为测试集。本文所提方法使用 800 张

训练图像训练,选取验证集中 10 张图像作为训练过程中的评估,然后在基准数据 Set5, Set14, B100, Urban100 上测试。

残差稠密网络的配置与文献[5]保持一致,反馈网络循环的次数设为 2。所有实验在 Windows 10 系统, Pycharm 开发工具,深度学习开源框架 Pytorch 0.41, NVIDIA 1080ti GPU 上完成,一次完整训练大概需要 1 天时间。

低分辨率图像是在高分辨率图像上通过双三次插值且降采样因子为 4 计算得到的。在训练的过程中,批量大小为 16,高分辨率图像被随机裁剪成大小为 128 的图像块,使用随机水平和垂直翻转来进行数据增强。使用 Adam 优化器更新参数,学习率设为 0.0001,其余参数保持默认值,每个训练周期迭代 1000 次,总共训练 200 个周期。

3.2 性能评估

本文在数值上使用峰值信噪比 (PSNR)^[6] 和结构相似度 (SSIM)^[7] 来评估图像超分辨率重建的性能, PSNR 和 SSIM 的值越大表明图像重建的质量越好。本文所提方法与双三次插值法、SRCNN、RDN 在数值上和视觉上相比较来说明所提方法的性能。

表 1 显示了 PSNR 和 SSIM 在测试集上的结果,插值法和 SRCNN 的重建的原图来自文献[8]的作者提供, RDN 重建的原图无法获得,本文根据网络提供的代码重新训练。在测量 PSNR 和 SSIM 的数值时,本文将原始 RGB 图像转到 YcrCb 色彩空间,然后取 Y 通道进行数值计算, PSNR 和 SSIM 的计算公式采用 scikit-image 提供的函数。

表 1 PSNR 和 SSIM 在测试集上的结果

数据集	插值法 (PSNR/SSIM)	SRCNN (PSNR/SSIM)	RDN (PSNR/SSIM)	本文算法 (PSNR/SSIM)
Set5	28.42/0.8104	30.48/0.8628	32.23/0.8915	32.28/0.8919
Set14	26.00/0.7027	27.50/0.7513	28.68/0.7830	28.73/0.7832
BSD100	25.96/0.6675	26.90/0.7101	27.63/0.7354	27.75/0.7360
Urban100	23.14/0.6577	24.52/0.7221	26.29/0.7921	26.41/0.7925

观察表 1 中的数值可以看出,本文算法在 4 个测试集上 PSNR 和 SSIM 的值都高于其他 3 种方法。双三次插值法由于固有因素重建的图像质量低于基于深度学习的方法。SRCNN 使用浅层的单一结构的卷积神经

网络,在重建图像的质量上面低于深层的网络。本文方法基于 RDN 的改进,在增加了一条直接上采样的路径和局部反馈网络之后,提供了重建图像更多的特征信息,使得的评价指标有小幅上涨,这说明了本文所提算

法的有效性.

图 5 至 图 8 分别展示了在数据集 Set5, Set14, BSD100, Urban100 上的重建效果图. 为了方便观察, 图像的局部区域被放大显示在下半部分. 通过观察放大部分的图像, 在视觉上双三次插值法仅仅通过线性公

式对图像进行 4 倍上采样, 所得图像局部放大后模糊不清. SRCNN 重建的图像效果虽然比双三次插值法好, 但是由于只使用 3 层网络, 重建的结果仍然比较模糊. 本文算法与 RDN 相比, 由于引入了更多特征, 重建的图像显得比较光滑, 看上去更清晰.

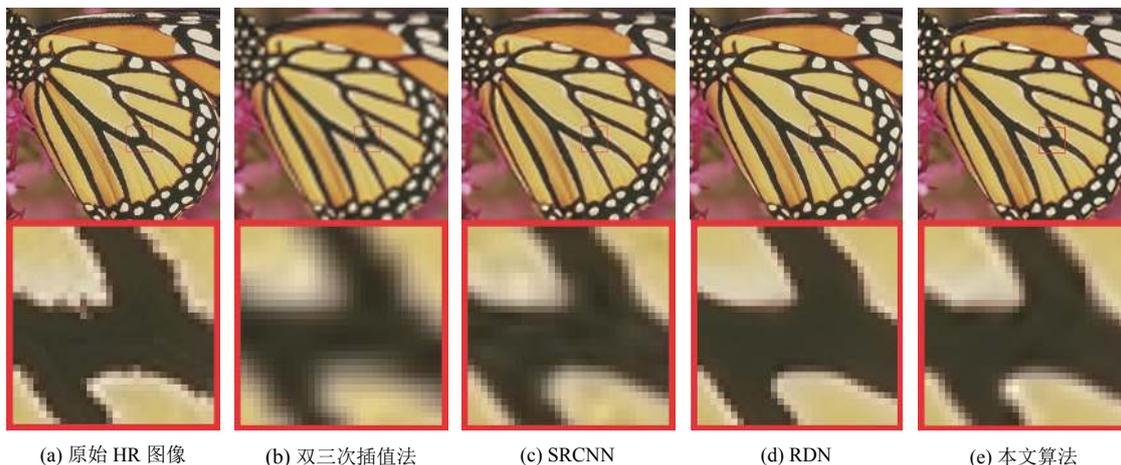


图 5 数据集 Set5 中图像“butterfly”的重建结果展示图

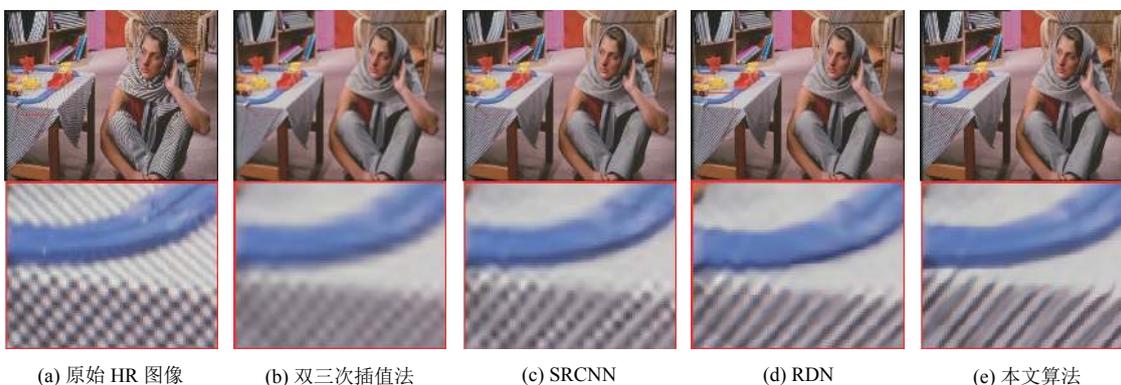


图 6 数据集 Set14 中图像“barbara”的重建结果展示图



图 7 数据集 BSD100 中图像“119082”的重建结果展示图



图8 数据集 Urban100 中图像“img096”的重建结果展示图

3.3 对比实验

本文通过一组对比实验来确定网络的反馈次数。图9显示了实验结果对比图,横坐标为反馈次数,纵坐标为在验证集上的 $PSNR$ 值。随着反馈次数的增加,网络中的感受野增大,网络性能通常会提升。从图中可以看出当反馈次数为2时所提方法在验证集上的 $PSNR$ 值最大,所以本文设定反馈次数为2。当反馈次数大于2时,由于缺少监督可能导致网络性能下降。

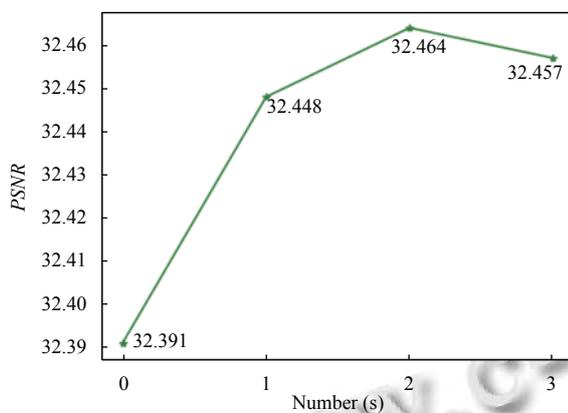


图9 不同反馈次数性能对比图

4 结论

本文提出一种基于双路径反馈网络的图像超分辨率算法,两条路径分别为深度网络和浅层网络,深度网络学习重建图像的高频信息,特别的浅层网络直接给重建图像提供低频信息,有利于训练前期网络稳定的优化迭代。此外模型中所采用的反馈网络利用局部循环训练进一步提高了图像重建的质量。最后本文在4个基准数据集上通过数值和视觉两方面验证了所提方法的有效性。

参考文献

- 1 Bai YC, Zhang YQ, Ding ML, *et al.* Finding tiny faces in the wild with generative adversarial network. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 21–30.
- 2 Dong C, Loy CC, He KM, *et al.* Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295–307. [doi: 10.1109/TPAMI.2015.2439281]
- 3 Shi WZ, Caballero J, Huszár F, *et al.* Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 1874–1883.
- 4 Kim J, Kwon Lee J, Mu Lee K. Accurate image super-resolution using very deep convolutional networks. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 1646–1654.
- 5 Zhang YL, Tian YP, Kong Y, *et al.* Residual dense network for image super-resolution. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 2472–2481.
- 6 Huynh-Thu Q, Ghanbari M. Scope of validity of PSNR in image/video quality assessment. Electronics Letters, 2008, 44(13): 800–801. [doi: 10.1049/el:20080522]
- 7 Wang Z, Bovik AC, Sheikh HR, *et al.* Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 2004, 13(4): 600–612. [doi: 10.1109/TIP.2003.819861]
- 8 Huang JB, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars. Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. 2015. 5197–5206.