

# 一种基于融合的单幅图像超分辨率重建

苏进胜,张明军,俞文静

(广州软件学院网络技术系,广东 广州 510990)

**摘要:**图像超分辨率重建是由一幅或者多幅低分辨率图像序列重建与之对应的高分辨率图像。高分辨率图像具有更高像素密度,可以提供更多的图像细节,这些细节往往在一些具体应用场景中起到关键性作用。针对单幅低分辨率重建超分辨率应用问题,提出了一种基于图像融合的方法,该方法选取两种或者两种以上利用生成对抗网络算法进行超分辨率图像的重建算法,然后对它们各自重建的图像进行图像融合。图像融合使用将两幅或多幅图像综合成一幅新的图像。融合能利用两幅(或多幅)图像在时空上的相关性及信息上的互补性,可以使得融合后得到的图像对场景有更全面、清晰的描述,从而更有利于人眼的识别。借鉴了集成学习的思想,该文使用 BasicSR、SRGAN 和 ESRGAN 这三种超分辨率重建算法生成的超分辨率图像进行两两交叉融合进行仿真实验。实验结果表明,这种利用不同生成对抗网络重建的超分辨率图像进行融合简单有效,融合后的超分辨率图像质量在两个指标上 PSNR 和 SSIM 上总体优于融合前的图像质量。

**关键词:**超分辨率重建;图像融合;深度学习;卷积神经网络;生成对抗网络

中图分类号:TP391.41;TP183

文献标识码:A

文章编号:1673-629X(2022)05-0053-05

doi:10.3969/j.issn.1673-629X.2022.05.009

## Single Image Super-resolution Reconstruction Based on Image Fusion

SU Jin-sheng, ZHANG Ming-jun, YU Wen-jing

(Dept. of Network Technology, Software Engineering Institute of Guangzhou, Guangzhou 510990, China)

**Abstract:** Image super-resolution reconstruction is to reconstruct the corresponding high-resolution image from one or more low-resolution image sequences. High-resolution images have higher pixel density and can provide more image details, which often play a key role in some specific application scenarios. Aiming at the application of single-frame low-resolution reconstruction and super-resolution, we propose a method based on image fusion. The respective reconstructed images are image fused. Image fusion uses the integration of two or more images into a new image. Fusion can make use of the temporal and spatial correlation and information complementarity of two (or more) images, which can make the image obtained after fusion have a more comprehensive and clear description of the scene, which is more conducive to human eye recognition. We draw on the idea of ensemble learning and use the super-resolution images generated by the three super-resolution reconstruction algorithms of BasicSR, SRGAN and ESRGAN to carry out two-by-two cross fusion for simulation experiments. The experimental results show that this kind of super-resolution image fusion reconstructed by different generation confrontation networks is simple and effective. The super-resolution image quality after fusion is generally better than the image quality before fusion in terms of PSNR and SSIM.

**Key words:** super-resolution reconstruction; image fusion; deep learning; convolution neural network; generative adversarial networks

## 0 引言

图像超分辨率重建是通过单幅低分辨率图像(low resolution, LR)或图像序列重建超分辨率图像(super resolution, SR)的过程。超分辨率图像有更多的细节信息和画质,这些细节在军事、公共安全、交通安全、气象灾害预警、森林病虫害识别、医学检验成

像和计算机视觉等领域有着重要的应用价值。

单帧图像超分辨率重建(single image super resolution, SISR)<sup>[1-6]</sup>是从观测到的一幅低分辨率图像重建出相应的超分辨率图像,不依赖其他相关图像的超分辨率技术。在监控设备、数字高清、显微成像、视频编码通信、卫星图像遥感、视频复原和医学影像等领

收稿日期:2021-06-02

修回日期:2021-10-08

基金项目:2018年广东省普通高校特色创新项目(2018KTSCX341);2020年广东省省级“创新强校工程”科研项目(2020KTSCX213);2020年信息技术新工科产学研联盟人工智能课程教学试点项目(AEEE2020JG001);2019年广州软件学院科学研究项目(KY201921);2020年广州软件学院院级“质量工程”建设项目(ZYGXK202007)

作者简介:苏进胜(1968-),男,硕士,讲师,研究方向为机器学习、图像处理及计算机视觉。

域都有重要的应用价值。这是一个“病态问题”,因为采用不同的方法重建出的图像不会完全一样。SISR 是一个反向问题,对于一幅低分辨率图像,可以存在不同品质的超分辨率图像与其对应,因此在由低分辨率图像重建超分辨率图像时通常会施加一个先验信息进行规范化约束。在传统的重建方法中,这种先验信息可以通过若干低—超分辨率图像对的实例中习得。而基于深度学习的高分辨率重建方法是通过卷积神经网络直接学习低分辨率图像到超分辨率图像的映射函数。近年来出现了基于生成对抗网络 (generative adversarial networks, GAN)<sup>[7]</sup> 的超分辨率重建的算法。重建方法的关键是学习低分辨率块和超分辨率块之间的映射关系。在重建方法上从早期的稀疏字典方法到现在的深度学习方法,其中使用生成对抗网络重建的效果会更好,但是每一种方法都有其局限性。该文采用了针对不同的生成对抗神经网络重建的超分辨率方法进行图像融合,使用图像融合算法将两幅或多幅图像进行融合重建出相应的超分辨率图像。这种图像融合能利用两幅(或多幅)图像在时空上的相关性及信息上的互补性,使得融合后的图像对一些应用场景有更加全面和清晰的描述。融合两种或多种方法进行重建的图像效果比融合前的单一方法重建的图像效果要好。主要贡献是基于集成学习的思想,提出了一种融合不同超分辨率重建的方法。融合前的方法各有特点。根据超分辨率场景的要求,可以灵活选择不同的算法进行融合,仿真实验证明了该方法的有效性。

## 1 超分辨率重建的常用方法

早期算法常用插值方法,比如最邻近插值、双线性插值、双三次插值。这些滤波算法能产生过于平滑超分辨率图像输出,但是缺乏高频信息。

基于稀疏信号表示的方法<sup>[8]</sup>。图像统计表明图像块可以表示为选择适当超完备字典的稀疏线性组合形式,受此启发,该文寻求每个低分辨率输入图像块的稀疏表示,然后用此稀疏表示的系数来生成超分辨率图像。压缩感知理论<sup>[9]</sup>表明,一幅图像能够在一些条件差的情况下由它的一组稀疏表示系数在超完备字典上进行精确重建<sup>[8]</sup>。

基于学习的单帧超分辨率重建是近年来研究的一个热点问题,又称为图像幻感 (image hallucination)<sup>[10]</sup> 或基于样例 (example-based)<sup>[11]</sup> 的超分辨率重建。它通过使用神经网络方法从训练集样本提取高频信息,从而对未知样本所需信息进行预测,提高图像分辨率的重建质量。

近些年来,随着云计算、大数据和计算机算力的提高,深度学习逐渐兴起,卷积神经网络也被应用到超分

分辨率图像重建应用中。深度卷积神经网络能很好地学习低分辨率图像和超分辨率图像之间的映射关系。卷积神经网络采用单通道卷积核或者多通道卷积核,通过卷积运算逐层提取输入图像的局部特征模式,前向传递给中间的隐藏层,随着隐藏层层数的增加和网络加深而逐步学习到高级的特征,最后经过上采样得以重建高分辨率图像。典型的前馈深度网络包括多层感知机和卷积神经网络 (CNN)。

常用的方法有基于卷积神经网络的方法 (super resolution using convolution neural network, SRCNN) 和基于极深网络的方法 (very deep networks for SR, VDSR)<sup>[12-15]</sup>。深度残差网络<sup>[16]</sup>的引入更是把深度一下子增加至 152 层,更深的网络使用可以明显提高重建图像的质量。SRCNN<sup>[17-19]</sup>首次提出了一种基于深度学习的单幅图像超分辨率重建方法,采用端到端的方式学习超分辨率图像和低分辨率图像之间的映射,可以用一个深度卷积神经网络来表示,通过输入低分辨率的影像输出超分辨率的影像。

## 2 基于图像融合的超分辨率图像重建方法

### 2.1 基于生成对抗网络的两种超分辨率重建方法分析

大多数超分辨率图像重建方法一般都是放大较小的倍数,例如 2~4 倍,当放大倍数在 4 倍以上时就会出现过度平滑的现象,使得图像出现一些非真实感。近年来,GAN (generative adversarial networks)<sup>[7]</sup> 也被应用到超分辨率重建算法上,取得了非常不错的效果。SRGAN<sup>[1]</sup>首次借助于 GAN 的网络架构生成图像中的细节。SRGAN 的结构如图 1 所示。

生成对抗网络是一种生成模型,是近年来深度学习中有创造性的一种影响深远的方法。生成对抗网络模型主要包括两部分:生成模型和判别模型。生成对抗网络同时训练两个模型:生成模型  $G$  捕捉数据随机分布信息,判别模型  $D$  评估生成样本以假乱真的概率大小。生成模型  $G$  训练过程是使  $D$  最大化的犯错,判别模型最大化地判别出生成样本,等两者分别达到 0.5 的概率,训练结束<sup>[7]</sup>。

对于 GAN,一个简单的理解是可以将其看作零和博弈的过程,该文可以将生成模型和判别模型看作博弈的双方。比如在犯罪分子造假币和警察识别假币的过程中:生成模型  $G$  相当于制造假币的一方,其目的是根据看到的钱币情况和警察的识别技术,去尽量生成更加真实的、警察识别不出的假币。判别模型  $D$  相当于识别假币的一方,其目的是尽可能识别出犯罪分子制造的假币。这样通过造假者和识假者双方的较量和改进,使得最后能达到生成模型能生成尽可能真的钱

币、识假者判断不出真假的纳什均衡效果(真假币概率都为 0.5)。SRGAN 是使用 GAN 网络中的生成网络

直接生成超分辨率图像。

GAN 网络的结构如图 2 所示。

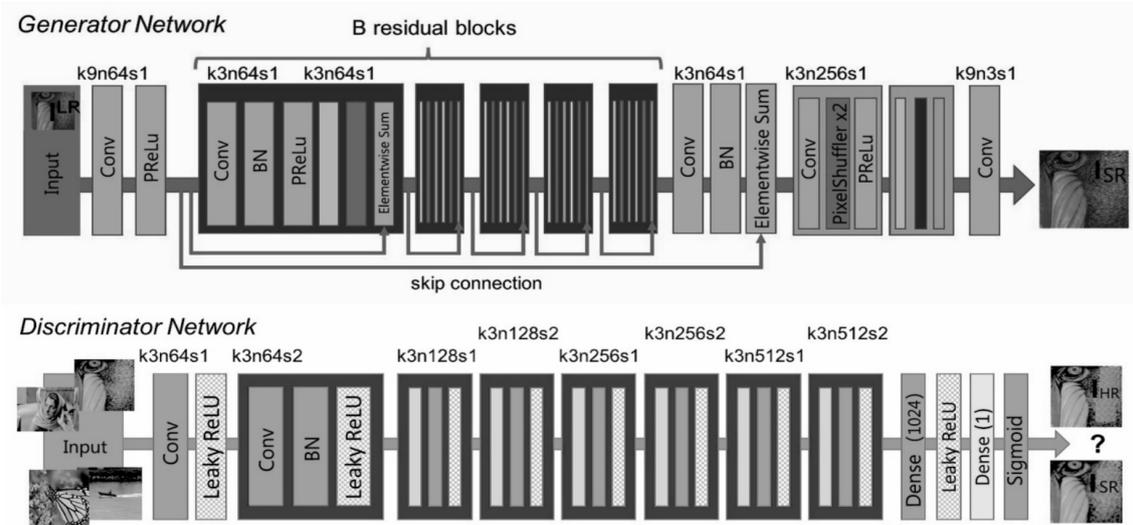


图 1 SRGAN 网络结构

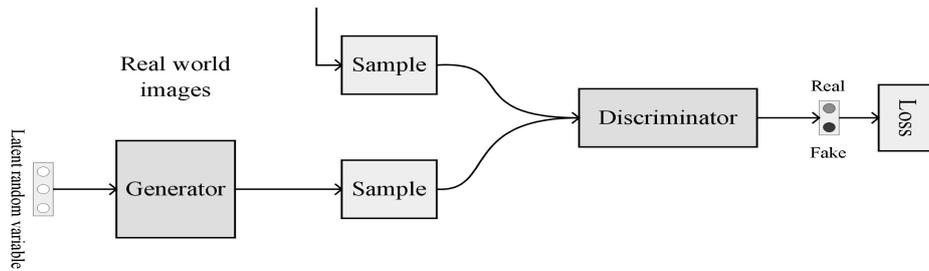


图 2 生成对抗网络结构

GAN 中的生成模型和判别模型一般都是卷积神经网络,网络使用均方误差 (MSE) 指标作为损失函数进行训练,能够获得较高的峰值信噪比 (PSNR),但是这种算法生成的图像往往会丢失一些高频细节信息,视觉体验欠佳。SRGAN 提出提升视觉体验的感知损失 (perceptual loss)<sup>[19]</sup> 和对抗损失 (adversarial loss)<sup>[1]</sup>,从而结合均方误差损失来提升输出图像的质量。

SRGAN 的结果与真实图像之间仍然存在明显的差异,ESRGAN<sup>[20]</sup> 从三个方面去提升这个模型。首先,通过引入密集残差块 (RDDB) 来提升模型的结构,使

之具有更大的容量和更易于训练。去除了批量归一化层,使用了残差缩放 (residual scaling) 等来促进训练一个深层次网络。第二,提出了判别器使用相对平均 GAN (RaGAN),相对平均 GAN 学习判断“是否一个图像相比于另一个更真实”而不是“是否一个图像是真或假”,这样的改进有助于生成器恢复更真实的纹理细节。第三,在 ESRGAN 提出了一种改进的感知损失,通过激活之前使用 VGG 特征而不是在激活后使用。ESRGAN 的网络结构如图 3 所示。

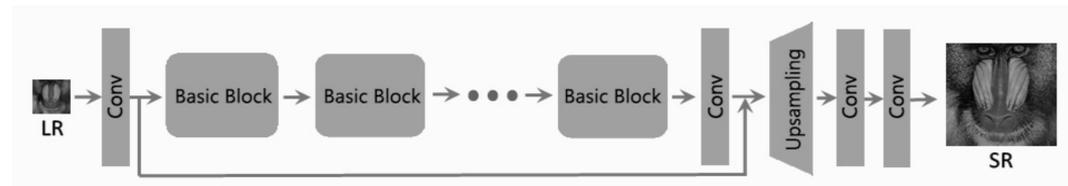


图 3 ESRGAN 网络结构

### 2.2 不同方法重建的超分辨率图像实现融合

图像融合 (image fusion) 是用图像融合算法将两幅或多幅图像合成一幅新的图像。目前应用最多的是像素级融合算法。图像融合能充分利用两幅或多幅图像的相似性及信息上的互补性,使得融合后的图像具备更多的细节信息。实验证明对同一幅图像像素有

细微差别的图像仍然可以进行融合,融合后的图像效果比融合前的效果要好。

子像素卷积方法 (sub-pixel convolution)<sup>[21]</sup> 对提取的图像特征图采用上采样 (upscale) 方式建立超分辨率图像,又叫做像素洗牌 (pixel shuffle),它可以避免带入过多人工因素,重建的图像质量高。

SRGAN 和 ESRGAN 两种方法均使用了生成对抗网络,结合子像素卷积方法进行上采样生成超分辨率图像,生成的图像质量比较高。BasicSR (basic super restoration)<sup>[22]</sup>是一种开源的图像和视频还原工具箱,可以用于超分辨率、去噪等领域。SRGAN 提出感知损失(perceptual loss)和对抗损失(adversarial loss)来提升恢复出的图片的真实感,使得重建的超分辨率图像达

到了照片级效果。正如前文所示,ESRGAN 主要针对 SRGAN 作了三个方面的改进,生成的超分辨率图像比 SRGAN 更逼真、更自然,在纹理上获得了更好的视觉感受。

超分辨率重建方法很多,该文选用这三种比较典型的方法各自重建超分辨率图像,然后两两进行融合,重建出最后的高分辨率图像。网络结构如图 4 所示。

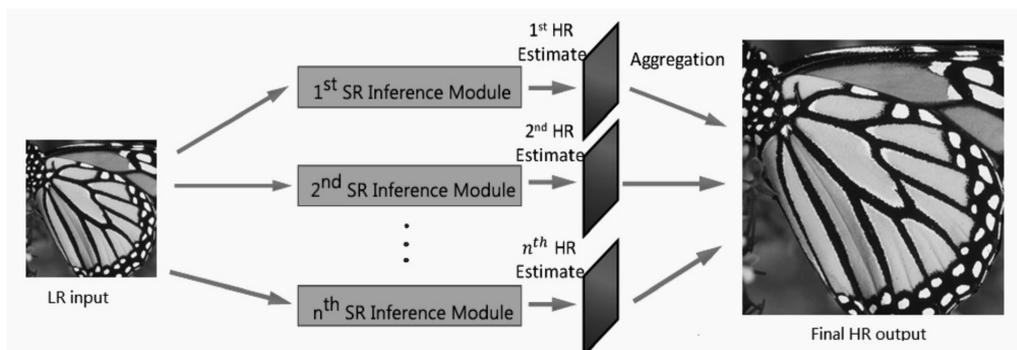


图 4 融合不同超分辨率重建模块的结构

为了验证图像融合后的超分辨率效果,该文采用 BasicSR + SRGAN、BasicSR + ESRGAN 和 SRGAN + ESRGAN 这三种组合,用每一种方法各自生成的超分辨率图像进行融合,从峰值信噪比 (PSNR) 和结构相似性 (SSIM) 两个指标值对比融合后和融合前的超分辨率图像效果。

### 3 图像融合仿真实验

SRGAN 和 ESRGAN 重建的超分辨率图像,不论在像素级还是视觉感官上都到达了非常高的质量。选用数据集 Set5 进行图像融合实验,采用这两种有代表性的算法及 BasicSR 方法在此测试数据集上重建的超分辨率图像两两进行融合,融合权重分别为 0.4 和 0.6。

衡量图像质量标准主要有两个指标:峰值信噪比 (PSNR) 和结构相似性 (SSIM)。PSNR (peak signal to noise ratio) 是应用广泛的一种图像客观评价指标,它是针对像素点间的误差评价图像质量,其值越大图像越接近原始图像质量。SSIM (structural similarity),是另一种图像质量评价指标,它分别从亮度、对比度和结构三方面度量图像相似性<sup>[23]</sup>。

表 1 ~ 表 3 是用三种不同方法对重建的超分辨率图像进行融合仿真实验。在被选用的三种方法上分别进行了 BasicSR 和 SRGAN2 倍上采样超分辨率图像重建、BasicSR 和 SRGAN4 倍上采样超分辨率图像重建及 SRGAN 和 ESRGAN4 倍上采样超分辨率图像重建。主要从衡量图像的两个重要的客观指标上对比了融合前后的图像质量。实验表明用不同方法重建的高分辨率图像进行融合后的效果要比单一方法建立的超分辨率图像的效果好。

表 1 BasicSR 和 SRGAN 融合实验结果 [2 倍上采样]

Set5	SR_x2	SRGAN_x2	Fused_x2
Infered Image			
PSNR	35.89	35.91	36.98
SSIM	0.950 7	0.947 7	0.957 8

表 2 BasicSR 和 SRGAN 融合实验结果 [4 倍上采样]

Set5	SR_x4	SRGAN_x4	Fused_x4
Infered Image			
PSNR	31.41	29.54	31.45
SSIM	0.859 3	0.807 8	0.854 3

表 3 SRGAN 和 ESRGAN 融合实验结果 [4 倍上采样]

Set5	SRGAN_x4	ESRGAN_x4	Fused_x4
Infered Image			
PSNR	29.54	31.47	31.90
SSIM	0.807 8	0.843 6	0.854 9

从表 1 和表 3 可以看出,融合后的 PSNR 和 SSIM 指标都比融合前的要好。从表 2 看融合后的 PSNR 比融合前的好,而 SSIM 指标比融合前的一个好,比另一个要差一些,但是总体上来看融合后的效果要好于融合前的效果。实验结果证明了融合的有效性。

### 4 结束语

提出了一种使用不同生成对抗网络生成的超分辨

率图像进行图像融合重建超分辨率图像的方法,通过对不同方法重建的高分辨率图像进行融合生成最终的超分辨率图像。实验结果表明融合后的图像效果要比融合前采用不同方法单独重建的图像效果好。超分辨率图像重建在医学影像诊断、航天和气象等领域都有广泛的应用。在融合不同的方法上还没有完全采用集成学习的思想实现,在后续研究中将深入研究,实现端到端的神经网络架构生成超分辨率图像。

#### 参考文献:

- [1] LEDIG C, THEIS T, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//2017 IEEE conference on computer vision and pattern recognition (CVPR). Honolulu, Hawaii; IEEE, 2017: 105–114.
- [2] TAI Ying, YANG Jian, LIU Xiaoming. Image super resolution via deep recursive residual network [C]//2017 IEEE conference on computer vision and pattern recognition (CVPR). Honolulu, Hawaii; IEEE, 2017: 2790–2798.
- [3] 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 (CVPR). Honolulu, Hawaii; IEEE, 2017: 1132–1140.
- [4] HUI Zheng, WANG Xiumei, GAO Xinbo. Fast and accurate single image super-resolution via information distillation network [C]//2018 IEEE conference on computer vision and pattern recognition (CVPR). Salt Lake City; IEEE, 2018: 723–731.
- [5] ZHANG Yulun, TIAN Yapeng, KONG Yu, et al. Residual dense network for image super resolution [C]//2018 IEEE conference on computer vision and pattern recognition (CVPR). Salt Lake City; IEEE, 2018: 2472–2481.
- [6] DENG Xin. Enhancing image quality via style transfer for single image super-resolution [J]. IEEE Signal Processing Letters, 2018, 25(4): 571–575.
- [7] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets [C]//Advances in neural information processing systems (NIPS). Montreal, QC, Canada; NIPS, 2014: 2672–2680.
- [8] 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.
- [9] KULKARNI N, NAGESH P, GOWDA R, et al. Understanding compressive sensing and sparse representation-based super resolution [J]. IEEE Trans. Circuits Syst. Video Technol., 2012, 22(5): 778–789.
- [10] FAN W, YEUNG D Y. Image hallucination using neighbor embedding over visual primitive manifolds [C]//2007 IEEE conference on computer vision and pattern recognition. Minneapolis, MN; IEEE, 2007: 1–7.
- [11] FREEMAN W T, JONES T R, PASZTOR E C. Example-based super resolution [J]. IEEE Comput. Graph. Appl, 2002, 22(2): 56–65.
- [12] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//2016 IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas, Nevada; IEEE, 2016: 770–778.
- [13] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C]//2016 IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas, Nevada; IEEE, 2016: 1646–1654.
- [14] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks [C]//2016 international conference on learning representations (ICLR). Caribe Hilton, San Juan, Puerto Rico; [s. n.], 2016.
- [15] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]//2015 international conference on learning representations (ICLR). San Diego, California, US; [s. n.], 2015.
- [16] LIU P, HONG Y, LIU Y. Multi-branch deep residual network for single image super-resolution [J]. Algorithms, 2018, 11(10): 144–159.
- [17] DONG C, LOY C C, HE K, et al. Learning a deep convolutional network for image super-resolution [C]//2014 European conference on computer vision (ECCV). Zurich, Switzerland; Springer, 2014: 184–199.
- [18] 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, 2016, 38(2): 295–307.
- [19] JOHNSON J, ALAHI A, LI Feifei. Perceptual losses for real-time style transfer and super-resolution [C]//2016 European conference on computer vision (ECCV). Zurich, Switzerland; Springer, 2016: 694–711.
- [20] WANG X, YU K, WU S, et al. Esrgan: enhanced super-resolution generative adversarial networks [C]//2018 European conference on computer vision. Munich, Germany; Springer, 2018: 63–79.
- [21] 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]//2016 IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas, Nevada; IEEE, 2016: 1874–1883.
- [22] WANG Xintao, YU KE, CHAN K C K, et al. BasicSR: open source image and video restoration toolbox [EB/OL]. 2020. <https://github.com/xinntao/BasicSR>.
- [23] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Trans. Image Process, 2004, 13(4): 600–612.