

基于注意力生成对抗网络的图像超分辨率重建方法^①



丁明航, 邓然然, 邵 恒

(长安大学 信息工程学院, 西安 710064)

通讯作者: 丁明航, E-mail: 447396777@qq.com

摘 要: 针对现有基于深度学习的图像超分辨率重建方法, 其对细节纹理恢复过程中容易产生伪纹理, 并且没有充分利用原始低分辨率图像丰富的局部特征层信息的问题, 提出一种基于注意力生成对抗网络的超分辨率重建方法. 该方法中生成器部分是通过注意力递归网络构成, 其网络中还引入了密集残差块结构. 首先, 生成器利用自编码结构提取图像局部特征层信息, 并提升分辨率; 然后, 通过判别器进行图像修正, 最终将图像重建为高分辨率图像. 实验结果表明, 在多种面向峰值信噪比超分辨率评价方法的网络中, 所设计的网络表现出了稳定的训练性能, 改善了图像的视觉质量, 同时具有较强的鲁棒性.

关键词: 超分辨率重建; 生成对抗网络; 注意力网络; 残差网络; 特征提取

引用格式: 丁明航, 邓然然, 邵恒. 基于注意力生成对抗网络的图像超分辨率重建方法. 计算机系统应用, 2020, 29(2): 205-211. <http://www.c-s-a.org.cn/1003-3254/7283.html>

Image Super-Resolution Reconstruction Method Based on Attentive Generative Adversarial Network

DING Ming-Hang, DENG Ran-Ran, SHAO Heng

(School of Information Engineering, Chang'an University, Xi'an 710064, China)

Abstract: The existing image super-resolution reconstruction method based on deep learning is easy to generate pseudo texture, and the rich local feature layer information in the original low-resolution image is not fully utilized. In order to improve image quality, a super-resolution reconstruction method based on attentive generative adversarial is proposed. The generator part of the method is constructed by attention recursive network, and a dense residual block structure is also introduced in the network. First, the generator extracts the local feature layer information of the image by using the self-encoding structure to improve the resolution. Then, the image is corrected by the discriminator. Finally, the image is reconstructed into a high-resolution image. In a variety of networks for peak signal-to-noise ratio super-resolution evaluation methods, the experimental results show that the designed network exhibits stable training performance, improves the visual quality of the image, and has strong robustness.

Key words: super-resolution reconstruction; generative adversarial network; attention network; residual network; feature extraction

图像超分辨率是一种典型的不适定^[1]问题, 其作为一种基础的视觉问题, 在学界和人工智能公司中引起

了越来越多的关注, 并被广泛地引用于安全和监控成像, 医学成像和图像生成等领域. 图像超分辨率目的在

① 收稿时间: 2019-07-13; 修改时间: 2019-08-20; 采用时间: 2019-08-27; csa 在线出版时间: 2020-01-16

于通过模糊的低分辨率 (Low Resolution, LR) 图像生成视觉质量较好的高分辨率 (High Resolution, HR) 图像. 为了解决这个不适应问题, 已经提出了众多的图像超分辨率算法, 包括基于插值的^[2], 基于重构的^[3]和基于学习的方法^[4-10].

其中基于插值的方法^[11]能较好地保护图像的边缘, 但该方法不考虑图像的内容, 只是简单地进行像素点之间的计算, 容易丢失大量细节信息, 从而导致模糊. 另一种基于重构的超分辨率方法^[12]结合了图像的降质模型, 解决了基于插值方法无法引入先验信息的问题. 该方法虽然对复杂度低的图像效果较好, 但对纹理结构丰富的图像则效果一般. 随后研究人员提出了基于学习的方法, 如 Dong C 等^[13]提出的 SRCNN. 该方法结构简单, 结果相较于其他方法更精确, 但是会损失许多图像的细节信息, 同时数据来源并不多, 所恢复出的图像纹理在视觉上还是无法令人满意.

为了进一步提高视觉质量, 生成对抗网络^[14] (Generative Adversarial Network, GAN) 被引入超分辨率领域, 以改善图像超分辨率后失真的视觉效果. 然而, SRGAN^[15]由于网络结构复杂, 不能稳定训练更深层的网络, 因此其结果仍与真实图像存在差距. 随后, Wang XT 等^[16]研究了 SRGAN 网络体系结构的两个关键组成部分: 对抗损失和感知损失, 并对每个部分进行了改进, 得到了一个增强的 SRGAN (ESRGAN). 该方法可以训练出更深层的网络, 提取更多纹理细节特征, 但在

图像恢复时容易产生更多与真实图像不一致的伪纹理. 同时该方法无法充分利用低分辨率局部特征层信息, 因此, 在无法生成更加自然图像的同时浪费了深层网络的优良设计.

为了解决网络容易生成伪纹理的问题, 本文首先将注意力生成对抗网络^[17]引入超分辨率方法中, 以此作为超分辨率模型的基础架构. 其次, 为了降低网络结构的复杂性以更多地利用低分辨率的局部特征层信息, 本文还对注意力生成对抗网络中生成器部分进行了改进, 引入了密集残差块结构^[18]. 最后, 在图像超分辨率重建后, 本文还利用峰值信噪比^[19] (Peak Signal to Noise Ratio, PSNR) 和结构相似性^[20] (Structural Similarity Index, SSIM) 这两种普遍的超分辨率评价标准对结果进行定量分析, 对比说明所设计方法的优良性能.

1 相关工作

1.1 生成对抗网络

生成对抗网络 (GAN) 的理论基础源自于博弈论中的二人零和博弈. GAN 强大的图片生成能力, 使其在图像合成、图像修补、超分辨率、草图复原等领域内均有广泛应用. 同样, 生成对抗网络的结构也被应用于本文的基础框架中. GAN 的基本框架包含一个生成器模型 (Generator model, G) 和一个判别器模型 (Discriminator model, D), 如图 1 所示.

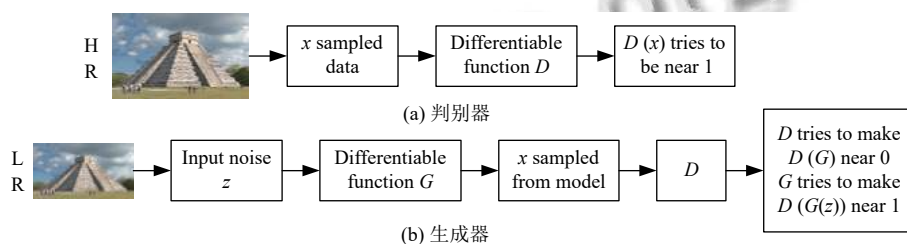


图 1 生成对抗网络基础结构

如图 1 所示的 GAN 的基础框架中, 生成器模型可用函数 $G(x)$ 表示, 判别器模型可用函数 $D(x)$ 表示, 每个函数都有可调参数. 图 1(a) 所示为判别器结构, 其处理流程是从训练高分辨率真实图像中采样作为输入 x , 送入判别器模型 $D(x)$ 中, $D(x)$ 通过自身的训练学习, 尽可能地输出接近 1 的概率值. 图 1(b) 所示为生成器结构, 其处理流程是从训练低分辨率图像中采样作为输入 z , 经

生成器模型生成伪造样本 $G(z)$, 然后将其作为判别器模型的输入. 判别器模型的目的是尽量使 $D(G(z))$ 接近 0, 而生成器模型的目的尽量使它接近 1, 最终在二者的互相博弈中达到平衡.

1.2 残差网络和跳跃连接

残差网络 (Residual Network, ResNet) 和跳跃连接 (skip connection) 结构经研究可以更加容易地训练深层

的网络,因而 GAN 可以通过加入该结构来增加网络层数提升超分辨率图像生成能力. 残差网络和跳跃连接的基本结构如图 2 所示.

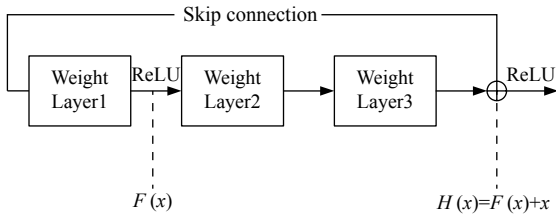


图 2 残差网络和跳跃连接结构

图 2 中, 残差网络在原始的卷积层上增加跳跃连接支路构成基本残差块, 使原始要学习的 $H(x)$ 被表示成 $H(x) = F(x) + x$. 残差网络的残差结构使得对 $H(x)$ 的学习转为对 $F(x)$ 的学习, 而对 $F(x)$ 的学习较 $H(x)$ 容易. 残差网络通过层层累加的残差块结构, 有效缓解了 GAN 深层网络的退化问题, 使得 GAN 更易稳定训练, 更易生成自然真实的图像.

2 注意力生成对抗网络设计

基于生成对抗网络的基础框架, 本文将注意力递归网络引入到生成器中, 并将常规残差块替换为密集残差结构, 通过此设计来尽可能生成在视觉质量上更为理想的图像. 同时, 改进判别器和感知损失函数, 以此作为用于平衡视觉质量和峰值信噪比 (PSNR) 的网络策略.

2.1 生成器设计

如图 3 所示, 本文所设计的生成器主要由浅层特

征提取网络, 注意力递归网络, 非线性映射网络和最后的上采样网络组成. 图 3 中, I_{LR} 和 I_{SR} 分别表示网络的输入和输出. 具体地, 浅层特征提取网络使用两个卷积层 (Conv) 来提取低分辨率图像的浅层信息. 第一个 Conv 层从 I_{LR} 输入中提取特征 F_{-1} . 然后对 F_{-1} 做进一步的浅层特征提取, 输出为 F_0 , 用公式可以表示为:

$$F_{-1} = H_{SFE1}(I_{LR}) \tag{1}$$

$$F_0 = H_{SFE2}(F_{-1}) \tag{2}$$

其中, $H_{SFE1}(\cdot)$ 和 $H_{SFE2}(\cdot)$ 表示卷积运算.

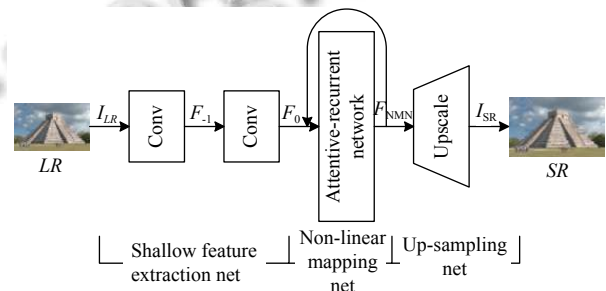


图 3 生成器结构

在对原始输入 I_{LR} 进行浅层特征提取后, 其输出 F_0 作为新的输入被送入注意力递归网络中. 详细的注意力递归网络结构如图 4 所示. 注意力递归网络的主要目的是在输入的图像中提取出需要恢复的纹理细节, 这些纹理细节将会被之后的非线性映射网络所增强并用于高分辨率图像的生成. 因此能否重建出逼真自然的图像, 取决于低分辨率纹理细节提取的好坏.

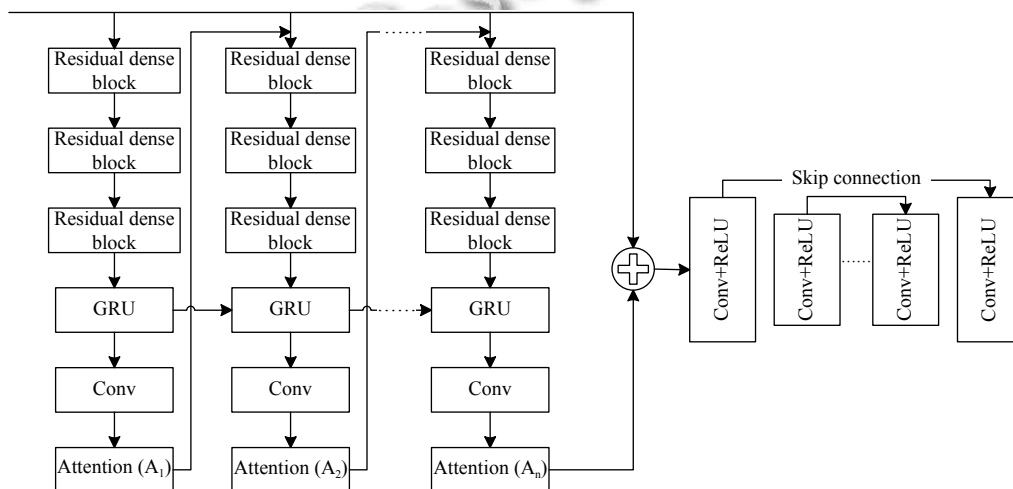


图 4 注意力递归网络结构

如图4中的架构所示,本文采用递归网络来生成视觉注意力.递归网络中的每个时间步长均包含3层密集残差块(Residual Dense Block, RDB),1个循环门控单元(Gated Recurrent Unit, GRU)和1层卷积层,以此用来生成二维注意力图像.每个时间步长所学习到的注意力图都是一个取值介于0到1之间的矩阵,其中每个矩阵元素值越大,表示注意力就越大.在图4中, A_n 为注意力学习过程生成的可视化注意力图,其中 n 代表递归的次数.随着递归次数的增加,训练出的注意力图就越能凸显细节纹理和相关结构.

递归注意力网络中用到的循环门控单元(GRU)包括一个重置门 r_t 和一个更新门 z_t .GRU的计算公式被定义为如下公式:

$$z_t = \sigma(W_z[h_{t-1}, x_t]) \quad (3)$$

$$r_t = \sigma(W_r[h_{t-1}, x_t]) \quad (4)$$

$$\hat{h}_t = \tanh(W[h_t * h_{t-1}, x_t]) \quad (5)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \hat{h}_t \quad (6)$$

其中, h_{t-1} 和 h_t 分别表示上一时刻和当前时刻的隐藏层状态,而 \hat{h}_t 表示当前更新状态. W_z , W_r 分别表示更新门和重置门的权重,*代表卷积运算.将GRU的输出特征送入其后的卷积层中,即可生成二维的注意力图像.在训练过程中,每个时间步长内输入图像均与生成的注意力图像相连,同时它们会作为新的输入被送到递归网络的下一层中.

假设通过 n 层递归注意力网络后的输出为 F_n ,并用如下公式表示:

$$F_n = H_{ATT,n}(F_{n-1}) = H_{ATT,n}(H_{ATT,n-1}(\dots(H_{ATT,1}(F_0))\dots)) \quad (7)$$

其中, H_{ATT} 表示注意力网络的函数, n 代表经过几层递归操作.

在经递归注意力网络进行纹理细节特征提取后,要对这些特征进行映射.所设计的非线性映射网络具有8层conv+ReLU结构,并添加了跳跃连接提升网络训练的稳定性.非线性映射网络可表示为如下公式:

$$F_{NMN} = H_{NMN}(F_n) \quad (8)$$

其中, F_{NMN} 是利用非线性网络函数 H_{NMN} 输出的结果.

最后,在生成高分辨率图像前加入了一个上采样网络,其是由一个卷积层构成的,以此来获得4倍放大的超分辨率图像输出.可表示为如下公式:

$$I_{SR} = H_{UPN}(F_{NMN}) \quad (9)$$

其中, H_{UPN} 为上采样网络函数, I_{SR} 为超分辨率图像输出.

2.2 密集残差块设计

为了进一步提高GAN的图像恢复质量,本文设计了密集残差块结构,并将其应用到注意力递归网络中,其结构如图5所示.如图5(a),首先在原有残差块基础上移除BN层.当训练和测试数据集的统计数据差异很大时,BN层往往会在生成图像中引入伪纹理并限制泛化能力.因此,为了网络的稳定训练,需要移除BN层.同时,移除BN层有助于降低计算复杂性和减小内存开销.如图5(b),在移除BN层后将密集块与跳跃连接进行结合,前面的RDB和后续每一层都的输出都有直接连接,这不仅保留了前馈性质,还充分提取了局部特征层信息,提升了网络的容量.

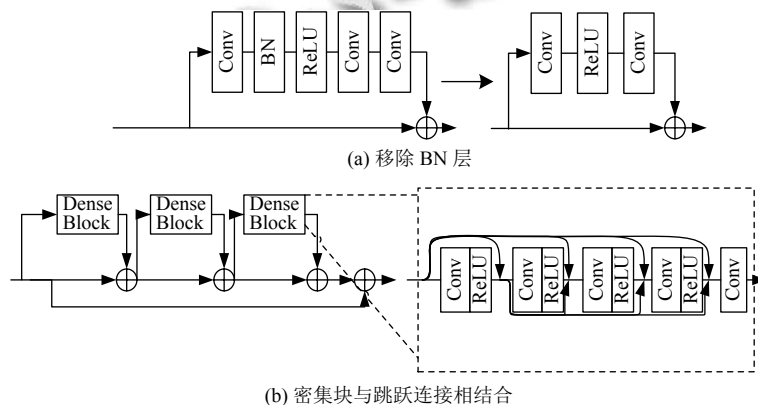


图5 密集残差块结构

2.3 判别器设计

除了改进的生成器结构,本文还改进了判别器.如

图6所示,改进的判别器的网络结构和参数设置与原始SRGAN的判别器不同.具体来说,输入待判别的高

分辨率图片,先经过7层的卷积层抽取图片特征,再将图像数据的维度压平(flatten),最后经全连接层(Fully Connected layer, FC)和 Sigmoid 函数输出判别的结果.本文生成对抗网络经判别器对真实图像与生成图像的判别,反馈到生成器中,使得权重参数变化更加灵敏,以此改善了网络退化的问题.原始 GAN 生成器 G 在初始训练时其性能远不如判别器 D ,这样便会产生模型崩溃(model collapse)问题.为此,本文改进了原始 GAN 的损失函数.该损失函数是基于 L1 范数设计的,同时包含了判别器的重构误差.判别器使用该损失函数在提升图像质量的同时还提升了生成器 G 的学习能力.

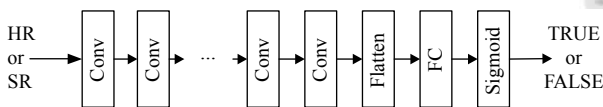


图6 判别器结构

首先图像像素级误差可以认为是服从高斯分布的,基于 L1 范数的损失函数被定义为:

$$L(I) = |I_{HR} - G(I_{LR})| \quad (10)$$

式中,高分辨率图像用 I_{HR} 表示,低分辨率图像用 I_{LR} 表示.需要迭代的生成器和判别器损失函数其优化规则由式(11)表示:

$$\begin{cases} L_D = L_{D_r} - k_t L_{D_f}, \text{ for } \theta_D \\ L_G = L(G(z) - x), \text{ for } \theta_G \\ k_{t+1} = k_t + \lambda_k (\gamma L_{D_r} - L_G) \end{cases} \quad (11)$$

其中,

$$L_{D_r} = L(D(x; \theta_D) - x) \quad (12)$$

$$L_{D_f} = L(D(y; \theta_D) - y) = L((D(G(z; \theta_G)) - G(z; \theta_G)); \theta_D) \quad (13)$$

$$y = G(x; \theta_D) \quad (14)$$

$$\gamma = \frac{E[L(G(z))]}{E[L(x)]} \quad (15)$$

其中,高分辨率图像用 x 表示,低分辨率图像用 z 表示,超分辨率图像用 y 表示,判别器关于高分辨率图像的损失用 L_{D_r} 表示,低分辨率图像的损失用 L_{D_f} 表示. λ_k 表示 k 的增量, k_t 表示 k 的第 t 次迭代结果, k 的值改变可以用来提升生成器的学习能力. γ 是超分辨率图像误差的期望与高分辨率图像的期望值之比,该参数的值可以提升生成图像的质量.

3 实验分析

3.1 评价指标

本文使用峰值信噪比 (PSNR) 和结构相似性 (SSIM) 这两种基础方法作为评价指标,以此来说明图像超分辨率方法性能的优劣.在图像超分辨率评价中, PSNR 的值越高就说明图像分辨率越高,或图像越接近真实高分辨率图像,其是利用均方误差 (Mean Square Error, MSE) 进行计算的:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (16)$$

式中, $m \times n$ 尺寸的图像均用 I 与 K 表示.

PSNR 被定义为:

$$PSNR = 20 \cdot \log_{10} \left(\frac{\text{MAX}_I}{\sqrt{MSE}} \right) \quad (17)$$

一般情况下, PSNR 的单位被设置为 dB, 数值越大图像视觉质量越高.

图像结构相似性是一种符合人类主观感受的评价指标.通常情况下, SSIM 的值越靠近 1, 代表超分辨率图像与原始图像结构越相似,也就是失真越小,质量越好. SSIM 被定义为:

$$SSIM = \frac{(2u_f u_{\bar{f}} + C_1)(\sigma_{f\bar{f}} + C_2)}{(u_f^2 + u_{\bar{f}}^2 + C_1)(\sigma_f^2 + \sigma_{\bar{f}}^2 + C_2)} \quad (18)$$

式中, u_f 表示真实高分辨率图像的平均灰度值, σ_f 表示真实高分辨率图像的方差,而 $u_{\bar{f}}$ 表示超分辨率图像的灰度值, $\sigma_{\bar{f}}$ 表示超分辨率图像的方差,最后真实高分辨率图像与超分辨率图像之间的协方差用 $\sigma_{f\bar{f}}$ 来表示, C_1 和 C_2 是常数,这两个常数被设置为, $C_1 = (k_1 * L)^2$, $C_2 = (k_2 * L)^2$, $k_1 = 0.01$, $k_2 = 0.03$, $L = 255$.

3.2 数据集与训练样本

本文利用超分辨率中被广泛使用的 Set5, Set14 和 BSD100 3 个基准数据集进行实验.所有实验均是将低分辨率图像进行 4 倍放大得到超分辨率图像的.训练样本使用来自 ImageNet 数据库的 10 万张图像的随机样本.并在配置有 NVIDIA GTX 1080Ti 的硬件设备上进行网络的训练.低分辨率图像的获取方法是利用 4 倍双三次下采样方法获取.同时,因为网络设计为全卷积网络,所以可以向生成器中送入任意大小的图像.为了优化网络的训练效果,本文使用 Adam 作为网络优化器.并设置初始学习速率为 10^{-4} , 进行 200 次迭代.

- Vision. Singapore. 2014. 111–126.
- 5 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.
 - 6 Tong T, Li G, Liu XJ, *et al.* Image super-resolution using dense skip connections. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy. 2017. 4809–4817.
 - 7 Zhang K, Zuo WM, Zhang L. Learning a single convolutional super-resolution network for multiple degradations. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 3262–3271.
 - 8 胡长胜, 詹曙, 吴从中. 基于深度特征学习的图像超分辨率重建. 自动化学报, 2017, 43(5): 814–821.
 - 9 苏健民, 杨岚心. 基于生成对抗网络的单帧遥感图像超分辨率. 计算机工程与应用, 2019, 55(12): 202–207, 214. [doi: [10.3778/j.issn.1002-8331.1807-0188](https://doi.org/10.3778/j.issn.1002-8331.1807-0188)]
 - 10 段然, 周登文, 赵丽娟, 等. 基于多尺度特征映射网络的图像超分辨率重建. 浙江大学学报(工学版), 2019, 53(7): 1331–1339. [doi: [10.3785/j.issn.1008-973X.2019.07.012](https://doi.org/10.3785/j.issn.1008-973X.2019.07.012)]
 - 11 苏衡, 周杰, 张志浩. 超分辨率图像重建方法综述. 自动化学报, 2013, 39(8): 1202–1213.
 - 12 Yang X, Zhang Y, Zhou DK, *et al.* An improved iterative back projection algorithm based on ringing artifacts suppression. Neurocomputing, 2015, 162: 171–179. [doi: [10.1016/j.neucom.2015.03.055](https://doi.org/10.1016/j.neucom.2015.03.055)]
 - 13 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](https://doi.org/10.1109/TPAMI.2015.2439281)]
 - 14 Goodfellow IJ, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets. Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, QB, Canada. 2014. 2672–2680.
 - 15 Ledig C, Theis L, Huszár F, *et al.* Photo-realistic single image super-resolution using a generative adversarial network. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 105–114.
 - 16 Wang XT, Yu K, Wu SX, *et al.* Esrgan: Enhanced super-resolution generative adversarial networks. In: Leal-Taixé L, Roth S, eds. Computer Vision – ECCV 2018 Workshops. Munich, Germany. 2018. 63–79.
 - 17 Qian R, Tan RT, Yang WH, *et al.* Attentive generative adversarial network for raindrop removal from a single image. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 2482–2491.
 - 18 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.
 - 19 陈晨, 刘明明, 刘兵, 等. 基于残差网络的图像超分辨率重建算法. 计算机工程与应用. <http://kns.cnki.net/kcms/detail/11.2127.TP.20190527.1720.009.html>. [2019-07-05].
 - 20 陈龙彪, 谏雨章, 王晓晨, 等. 基于深度学习的水下图像超分辨率重建方法. 计算机应用, 2019, 39(9): 2738–2743. [doi: [10.11772/j.issn.1001-9081.2019020353](https://doi.org/10.11772/j.issn.1001-9081.2019020353)]