

生成对抗网络的血管内超声图像超分辨率重建

吴洋洋, 杨丰, 黄靖, 刘娅琴

南方医科大学生物医学工程学院广东省医学图像处理重点实验室, 广东 广州 510515

摘要:针对超声图像分辨率低导致视觉效果差的问题,本文以超分辨率重建为基础,结合生成对抗网络的方法,生成相对原图更加清晰的血管内超声图像,用于辅助医生诊断与治疗。本方法应用生成对抗网络,生成器生成图像,判别器判断图像真伪。其过程:低分辨率图像经过亚像素卷积层 r^2 个特征通道,产生尺寸大小相同的 r^2 个特征图,对每个特征图中相对应的同一像素重新排列成一个 $r \times r$ 的子块,其对应高分辨率图像中的某一个子块,经过放大,产生 r^2 倍的高分辨率图像。生成对抗网络经过不断优化,获得更优质清晰的图像。将本方法(SRGAN)得出的结果与双立方插值(Bicubic)、超分辨率卷积网络(SRCNN)和亚像素卷积网络(ESPCN)等方法比较,其峰值信噪比(PSNR)和结构相似性(SSIM)分别提高2.369dB和1.79%。因此,我们得知:结合生成对抗网络的图像超分辨率重建能获得很好的血管内超声图像诊断视觉效果。

关键词:血管内超声;超分辨率重建;生成对抗网络;亚像素卷积层

Super-resolution construction of intravascular ultrasound images using generative adversarial networks

WU Yangyang, YANG Feng, HUANG Jing, LIU Yaqin

Guangdong Provincial Key Laboratory of Medical Image Processing, School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China

Abstract: The low-resolution ultrasound images have poor visual effects. Herein we propose a method for generating clearer intravascular ultrasound images based on super-resolution reconstruction combined with generative adversarial networks. We used the generative adversarial networks to generate the images by a generator and to estimate the authenticity of the images by a discriminator. Specifically, the low-resolution image was passed through the sub-pixel convolution layer r^2 -feature channels to generate r^2 -feature maps in the same size, followed by realignment of the corresponding pixels in each feature map into $r \times r$ sub-blocks, which corresponded to the sub-block in a high-resolution image; after amplification, an image with a r^2 -time resolution was generated. The generative adversarial networks can obtain a clearer image through continuous optimization. We compared the method (SRGAN) with other methods including Bicubic, super-resolution convolutional network (SRCNN) and efficient sub-pixel convolutional network (ESPCN), and the proposed method resulted in obvious improvements in the peak signal-to-noise ratio (PSNR) by 2.369 dB and in structural similarity index by 1.79% to enhance the diagnostic visual effects of intravascular ultrasound images.

Keywords: intravascular ultrasound; super-resolution reconstruction; generative adversarial network; sub-pixel convolution layer

图像超分辨率重建(SR)是由一个或多个低分辨率(LR)图像重建出高分辨率(HR)图像的技术,其增加图像中的高频成分,使得图像更加清晰,内容更加丰富^[1]。图像超分辨率重建是一个固有的病态问题^[2],在低分辨率图像中所呈现的细节通常不足以唯一地重建高分辨率图像。解决图像超分辨率重建问题主要有以下3个方面:(1)无混叠上采样图像;(2)去除获取图像时出现的模糊和干扰;(3)多样本的定位和合成^[3]。超分辨率在航空、军事、医学、交通管理等方面均有广泛应用^[4-9]。

当前,结合深度学习的超分辨率重建技术已在计算机视觉领域中得到深入研究。超分辨率卷积网络(SRCNN)是较早提出的一种端到端的超分辨率重建方法,网络结构简单,只有3个卷积层^[10-11]。该方法先使用双立方插值将低分辨率图像放大到目标大小,再通过三层卷积网络做非线性映射,得到高分辨率图像。网络对超参数的取值敏感,易陷入局部极值点,使网络训练不稳定,产生次优结果。SRCNN是先在低分辨率图像上进行上采样插值,让图像放大到与高分辨率图像尺寸相同,再作为网络的输入。深度递归卷积网络提出增加卷积层,使网络感受野扩大。同时为了避免过多网络参数,该文献提出使用递归神经网络^[12]。与SRCNN类似,该网络分为3个模块,分别用于特征提取、特征的非线性变换和图像重建。深度递归卷积网络重建效果比SRCNN有较大提高。深度递归卷积网络输入的是插值

收稿日期:2018-09-09

基金项目:国家自然科学基金(61771233,61271155)

Supported by National Natural Science Foundation of China (61771233, 61271155).

作者简介:吴洋洋,硕士生,E-mail: wuyang1993@163.com

通信作者:杨丰,教授,博士生导师,E-mail: yangf@smu.edu.cn

后的LR图像。SRCNN与深度递归卷积网络的卷积操作都是在HR图像上进行,计算量较大。亚像素卷积网络(ESPCN)提出亚像素卷积层,核心概念是在LR图像上直接进行卷积运算,在最后一层将图像放大,得到HR图像^[13]。卷积运算由于在LR图像上进行,效率会较高。SR结合车牌自动识别技术将视频的连续帧整合为单幅图像,包含的细节更加丰富^[14]。视频亚像素卷积网络是利用视频相邻帧的相关性,提出使用视频中的时间序列进行高分辨率重建的高效方法^[15]。上述基于学习的超分辨率重建方法中存在图像放大倍数不宜过大(小于4)的问题,若放大倍数在4倍以上(含4倍),易使图像显示结果过于平滑,在图像中缺少细节的真实感。以上方法均使用传统损失函数,将图像超分辨率细节重建问题过于简化。超分辨率生成对抗网络(SRGAN)提出传统的损失函数只能用于单个像素的评价,不足以评价整体图像视觉效果,并在网络中加入计算生成图像与目标图像之间误差的对抗损失,大大提高了图像的真实度,在视觉方面有了较大的进步^[16]。目前多数方法均应用于自然图像^[17-22],然而医学图像与自然图像存在很大的区别,医学图像为灰度图像,图像信息简单、纹理单一、种类少,感兴趣区域更加难以区分。本文针对单幅超声图像,以超分辨率重建为基础,结合生成对抗网络的方法,首次针对血管内超声(IVUS)图像成像较为模糊的特点,将图像放大若干倍,获得优质清晰的超分辨率图像,有效地辅助医生诊断。网络中加入对抗损失函数项,克服了一般模型仅适用于单像素重建的问题。与现有其它超分辨率重建方法比较,改进后模型提升了整体图像视觉效果,细节更加丰富逼真。文中使用均方误差(MSE)、峰值信噪比和结构相似性等三个客观指标与目前常用的三种超分辨率方法比较:Bicubic、SRCNN和ESPCN。从平均意见分(MOS)主观指标进一步验证本算法的重建视觉效果。

1 方法

生成对抗网络是模拟两个模型的对抗过程,其中一个生成器G(Generator),捕获真实数据的分布,生成逼近于真实分布的数据;另一个是判别器D(Discriminator),判断数据是否来自真实分布,输出数据来自于真实分布的概率,即一个0~1之间的数值^[23]。经过互相学习与对抗,G与D之间进行一个最大最小值的博弈。不断更新之后,直到D输出为0.5时,网络达到稳定状态,结束训练。此时,G生成的数据接近于真实分布。G与D的对抗过程如公式(1)所示:

$$\min_G \max_D V(D, G) = E_{x \sim P_{\text{data}}(x)} [\log D(x)] + E_{z \sim P_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

针对生成对抗网络训练过程不稳定的问题,研究者提出了许多改进方法,并取得了较好的效果^[24-29]。

1.1 网络框架

本文算法流程图如图1所示。网络由生成网络和判别网络组成。LR图像经过生成网络放大重建出SR图像。判别网络对SR与HR图像分别进行判断,输出概率,用于真假判断。网络的损失函数加入了D的损失,用来引导G生成高质量图像。

1.1.1 生成网络 生成网络的组成结构如图2G所示。网络有23层,主要由残差块和亚像素卷积层组成,包括卷积、批归一化、ReLU激活层,其中16层为跳跃连接叠加的残差块,每层卷积核大小为3×3,特征通道64个,步长为1。输入图像先经过卷积层的特征提取,再由两层亚像素卷积层进行放大处理,最后经过一层核大小为1×1的卷积,重建出4倍大的SR图像。在生成网络结构中,残差块如图3所示,每一层的结果与上一层的输出使用跳跃连接相叠加。再经过一层卷积,其结果与第一层卷积的结果叠加,可以防止底层特征丢失。单个亚像素卷积层如图4所示,网络的输入是原始低分辨率图像,输出是2倍大的输入图像,其过程为:LR图像经过4个特征通道的卷积操作,得到4张特征图,再对每个相同位置所对应的像素进行有序排列,让每个小块合并成长宽放大为2倍的图像。经过两次上述同样结构的亚像素卷积层,最终获得4倍SR放大图像。

1.1.2 判别网络 判别网络结构图如图2D所示。先将HR图像或者G生成的SR图像进行特征提取,包括卷积、LReLU激活层和批归一化,卷积核个数呈2的指数倍增长,从64增长到512,每个相同数量的卷积核有不同步长(1和2)的两个卷积层,再通过两层全连接层处理和sigmoid函数激活,最终得出判别结果。经过与G的不断对抗,直到D得出的数值逐渐趋于0.5,网络训练稳定。在LReLU之后,网络中使用了批归一化,其作用是避免梯度消失。

1.2 损失函数

SRGAN的损失函数如公式(2)所示:

$$l^{SR} = l_{MSE}^{SR} + l_{VGG}^{SR} + 10^{-3} l_{Gen}^{SR} \quad (2)$$

前两项为内容损失,是传统评价指标,主要针对单个像素,不足以评价整体图像的视觉效果。第三项为对抗损失,加入对抗损失,使整体图像视觉效果大大提高。

第一项为MSE,如公式(3)所示:

$$l_{MSE}^{SR} = \frac{1}{r^2 WH} \sum_{x=1}^r \sum_{y=1}^H \left(I_{x,y}^{HR} - G_{\theta_c} \left(I_{x,y}^{LR} \right) \right)^2 \quad (3)$$

其中W、H分别表示图像的长和宽, I^{LR} 表示低分辨率图像, I^{HR} 表示原始高分辨率图像。根据 I^{HR} 与G生成的图像之间的误差得到MSE,可以提高信噪比,但导致高频细节的缺失。

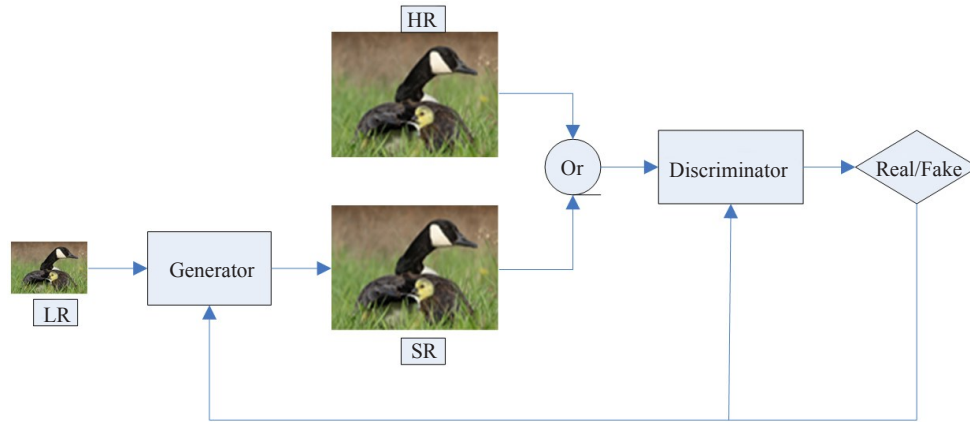


图1 网络流程图
Fig.1 Flow chart of the algorithm.

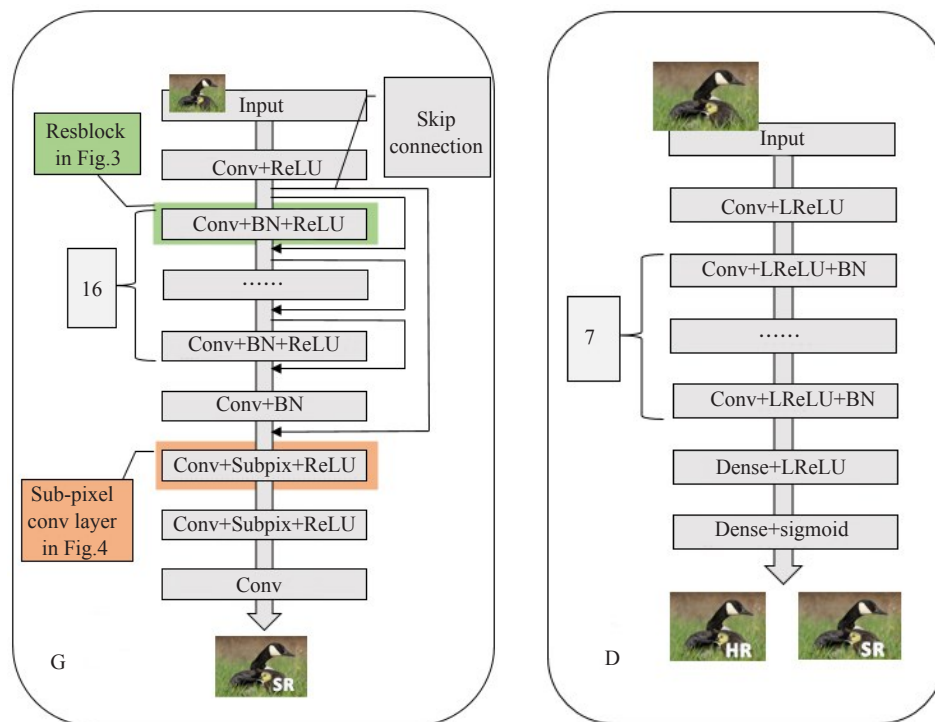


图2 网络结构
Fig.2 Network architecture. The left is the Generator, which involves the Residual Blocks, skip connection and Sub-pixel convolution layers. The right is the Discriminator involving the convolutional layers and full Connection.

第二项为VGG损失,如公式(4)所示:

$$l_{VGG}^{SR} = \frac{1}{W_{i,j} H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta_c}(I^{LR}))_{x,y})^2 \quad (4)$$

$$-\phi_{i,j}(G_{\theta_c}(I^{LR}))_{x,y}^2$$

$\phi_{i,j}$ 表示第*i*个最大池化层前的第*j*个卷积层的特征图。本项将某一层的特征图作为参考,计算其与生成图像之间的欧氏距离^[30],让生成器生成更加逼真的细节。

第三项为对抗损失,如公式(5)所示:

$$l_{Gen}^{SR} = \sum_{n=1}^N -\log D_{\theta_D}(G_{\theta_c}(I^{LR})) \quad (5)$$

*N*表示像素个数。本项计算重建的SR图像被判断为真实HR图像的概率。

2 结果

本文使用SRGAN对IVUS图像进行超分辨率重建,获得更加优质清晰的图像。

2.1 实验环境

实验共有不同帧IVUS图像250张,使用MatlabR2016a做数据的预处理,包括旋转、平移、缩放等方法,将数据扩充10倍得到2500张,其中2000张用

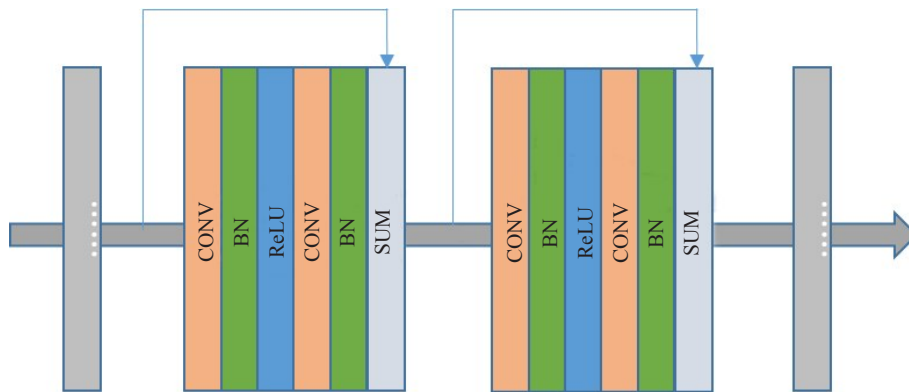


图3 残差块

Fig.3 Residual Blocks. The output of this layer adds its input, and then input to next layer.

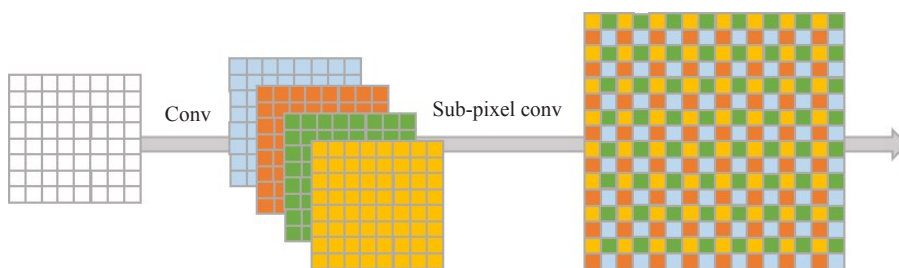


图4 亚像素卷积层

Fig.4 Subpixel convolutional layer.

来训练,500张用于测试。本文在TensorFlow框架下进行实验,使用2块 Nvidia GTX 1080 GPU,显存为8 GB。实验中生成器残差块层数为16层,初始学习率为0.0001,衰减率为0.1。为了避免参数过大导致训练无法完成,batchsize选择4。

2.2 定量结果

采用Bicubic、SRCNN、ESPCN以及未加入对抗损失的SRGAN-1和加入对抗损失的SRGAN-2等5种方法,分别对IVUS数据进行实验,得到的MSE、PSNR和SSIM计算结果如表1所示。

表1 5种方法定量评价结果比较

Tab.1 Comparison of the quantitative evaluation result of 5 methods

	Bicubic	SRCNN	ESPCN	SRGAN-1	SRGAN-2
MSE	0.0027	0.0033	0.0014	0.0012	0.0015
PSNR	25.4373	24.7513	27.1131	29.4821	28.2343
SSIM	0.9264	0.9236	0.9485	0.9664	0.9511

2.3 定性结果

上述评价指标仅针对单个像素级图像,不足以评价整体重建图像的视觉效果。在此,采用MOS作为图像主观评价指标。MOS由若干测试者对不同超分辨率方法处理后的图像进行评价,并给出相应1到5的分数,能够反映整体图像的视觉效果,1表示图像质量差,5表示图像质量好。由20名测试者对5种方法进行MOS评分,Bicubic、SRCNN、ESPCN、SRGAN-1和SRGAN-2的得分分别为3.2、4.2、3.6、4.4和4.8。图5中,图A~F分别表示Bicubic、SRCNN、ESPCN、未加入对抗损失的SRGAN-1、加入对抗损失的SRGAN-2与真值图,图a-f

从左至右分别为图A~F中红框的放大图。

2.4 结果分析

从表1的定量结果可得,未加入对抗损失的SRGAN-1的各项指标均为最高,加入了对抗损失的SRGAN-2的PSNR和SSIM仅次于前者。从MOS结果可得,加入对抗损失的SRGAN-2分数最高,视觉效果最好,其次为SRGAN-1。由此可知,两种SRGAN网络的PSNR、SSIM和MOS结果均优于Bicubic、SRCNN、ESPCN,两种评价均可证明SRGAN生成效果较好,而Bicubic、SRCNN、ESPCN的结果相对较差。从图5的定性结果可得,图A使用的插值方法结果亮度区分不明

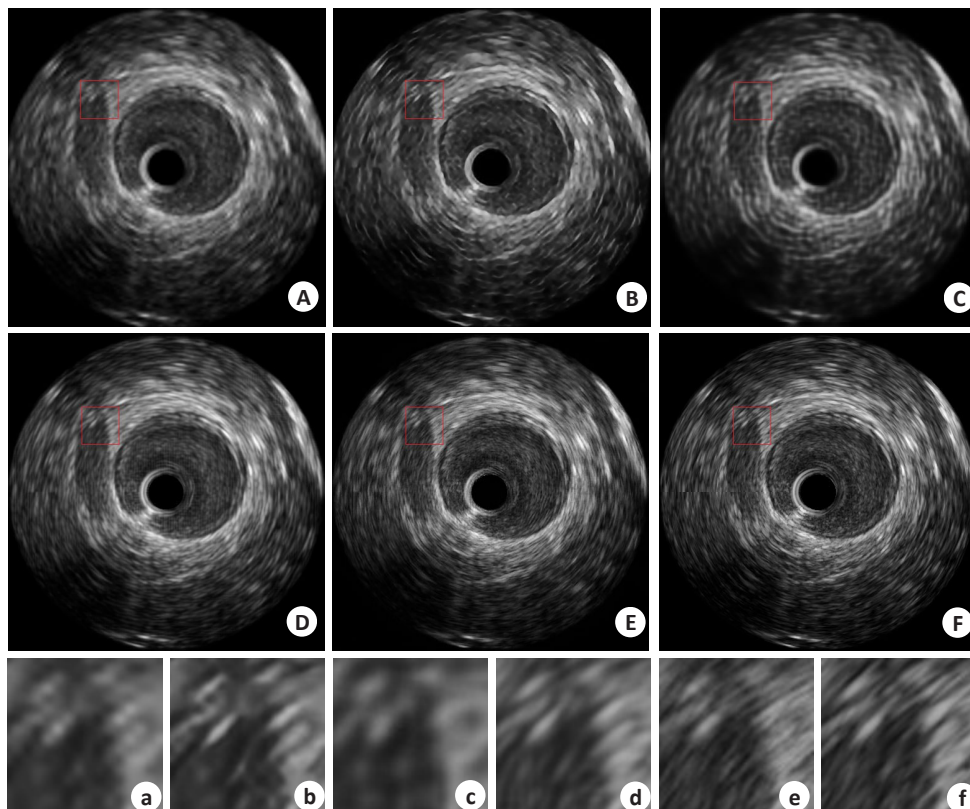


图5 4种方法对比
 Fig.5 Comparison of the 4 methods. A-F: Bicubic, SRCNN, ESPCN, SRGAN-1, SRGAN-2 and ground truth. a-f is the region of interest in the red frame in A-F separately.

显,图像整体较为模糊,分辨率较低;图B使用的SRCNN,相比图A清晰度有所提升,但仍然存在模糊、分辨率低的缺点;图C亮度偏暗,图像整体较为平滑,边界不明显,相比图A清晰度降低;图D亮度接近图F,相比前3种方法清晰度大大提高,但细节不够逼真;图E亮度与图F相似,图像清晰、亮度均匀、边界明显、细节真实,整体视觉效果最接近于真值图。

3 讨论

超声成像在所有医学成像方法中最快速便捷,可达到实时观察动态成像的效果。但超声成像往往较为模糊,图像中掺杂许多斑点噪声,对医生的诊断结果产生干扰。本文针对超声成像模糊的缺点,以超分辨率重建为基础,结合生成对抗网络的方法,对单帧IVUS图像进行超分辨率重建,获得了优质清晰的图像。

本文对SRGAN网络做了详细的介绍。网络由生成网络和判别网络两部分组成。生成网络将图像进行卷积等操作,得到的4张特征图对应位置的像素通过有序排列,生成一张长宽各放大两倍的图像,经过两层亚像素卷积层放大,最终结果为四倍大的图像。再将生成的图像输入判别网络,判别网络对图像的真伪性进行判断,得出一个0~1之间的数值,直到判别网络输出结果趋近于0.5时,训练结束。此时网络生成的图像最接近

于真实图像。

SRGAN方法在均方误差、峰值信噪比和结构相似性的评价结果优于其他方法,这说明生成对抗网络生成图像可以达到并超过在原图基础上插值、计算等方法的效果;在SRGAN中加入对抗损失后,经过平均意见分进行主观评价,生成的图像更加逼近真实图像,其整体图像视觉效果进一步提高,证明本文方法的重建质量优于现有的Bicubic、SRCNN和ESPCN等方法,有利于超分辨率重建。本方法只应用于单幅图像,并未利用多帧连续图像的时间空间相关性,重建难度大。后期工作将针对多帧图像特征的融合与生成,难度降低,并可以从视觉及评价指标等方面进一步改善超分辨率重建效果。

综上所述,SRGAN的超分辨率重建效果优于其他3种方法,加入对抗损失的SRGAN重建的整体图像视觉效果有了更大的提高。我们认为本方法可以产生以假乱真的医学图像,能够代替真实图像。方法也可应用到医学图像分割、病灶检测等方面,为临床诊断起到帮助。

参考文献:

- [1] 李智屹,韩玉兰,刘涛,等. 基于全变差和稀疏表示的图像超分辨率重构[J]. 信息技术, 2018, 41(9): 24-7.
- [2] Baker S, Kanade T. Limits on super-resolution and how to break them [J]. IEEE Trans Pattern Anal Mach Intell, 2002, 24(9): 1167-83.
- [3] Nguyen K, Fookes C, Sridharan S, et al. Super-resolution for

- biometrics: A comprehensive survey[J]. *Pattern Recognit*, 2018, 78: 23-42.
- [4] Greenspan H. Super-resolution in medical imaging[M]. New York : oxford university press, 2009: 43-63.
- [5] Hong CQ, Yu J, Tao DC, et al. Image-based three-dimensional human pose recovery by multiview locality-sensitive sparse retrieval [J]. *IEEE Transa Industr Electr*, 2015, 62(6): 3742-51.
- [6] 许丽娜, 何鲁晓. 基于凸集投影的高分四号卫星影像超分辨率重建 [J]. *测绘学报*, 2017, 46(8): 1026-33.
- [7] 曾 凯, 丁世飞. 图像超分辨率重建的研究进展[J]. *计算机工程与应用*, 2017, 53(16): 29-35.
- [8] 柯 健, 徐志京. 基于生成对抗网络的语音增强算法研究[J]. *信息技术与网络安全*, 2018, 37(5): 54-7.
- [9] 曾 超, 陈 雨. 车牌超分辨率重建与识别[J]. *计算机测量与控制*, 2018, 26(3): 244-9.
- [10] Dong C, Loy CC, He KM, et al. Image super-resolution using deep convolutional networks [J]. *IEEE Trans Pattern Anal Mach Intell*, 2016, 38(2): 295-307.
- [11] Umehara K, Ota J, Ishida T. Super-resolution imaging of mammograms based on the super-resolution convolutional neural network[J]. *OJMI*, 2017: 180-95.
- [12] Kim J, Lee JK, Lee KM. Deeply-recursive convolutional network for image super-resolution [C]//2016 Ieee Conference On Computer Vision And Pattern Recognition (Cvpr), 2016: 1637-45.
- [13] Shi WZ, 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), 2016: 1874-83.
- [14] Seibel J, Goldenstein S, Rocha A. Eyes on the target: super-resolution and license-plate recognition in low-quality surveillance videos[J]. *IEEE Access*, 2017, 5: 20020-35.
- [15] Li DY, Liu Y, Wang ZF. Video super-resolution using non-simultaneous fully recurrent convolutional network [J]. *IEEE Transactions Image Process*, 2019, 28(3): 1342-55.
- [16] Ledig C, Theis L, Huszar F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//30th Ieee Conference On Computer Vision And Pattern Recognition (Cvpr 2017), 2017: 105-14.
- [17] Yang J, Wright J, Huang TS, et al. Image super-resolution *via* sparse representation[J]. *IEEE Trans Image Process*, 2010, 19(11): 2861-73.
- [18] Cheong JY, Park IK. Deep CNN-based super-resolution using external and internal examples[J]. *IEEE Signal Process Lett*, 2017, 24(8): 1252-6.
- [19] Feng ZX, Lai JH, Xie XH, et al. Image super-resolution via a densely connected recursive network[J]. *Neurocomputing*, 2018, 316: 270-6.
- [20] Creswell A, White T, Dumoulin V, et al. Generative adversarial networks an overview[J]. *IEEE Signal Process Mag*, 2018, 35(1): 53-65.
- [21] Douzas G, Bacao F. Effective data generation for imbalanced learning using conditional generative adversarial networks [J]. *Expert Syst Appl*, 2018, 91: 464-71.
- [22] Liu YJ, Xu XP, Xu JH, et al. Image super-resolution reconstruction based on disparity map and CNN[J]. *IEEE Access*, 2018, 6: 53489-98.
- [23] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. *NIPS*. MIT Press, 2014: 2672-80.
- [24] Mirza M, Osindero S. Conditional generative adversarial nets [J]. *Computer Sci*, 2014: 2672-80.
- [25] Lin DY, Fu K, Wang Y, et al. MARTA GANs: unsupervised representation learning for remote sensing image classification [J]. *IEEE Geosci Remote S*, 2017, 14(11): 2092-6.
- [26] Wolterink JM, Leiner T, Viergever MA, et al. Generative adversarial networks for noise reduction in Low-Dose CT[J]. *IEEE Trans Med Imaging*, 2017, 36(12): 2536-45.
- [27] Li YC, Xiao NF, Ouyang W. Improved generative adversarial networks with reconstruction loss[J]. *Neurocomputing*, 2019, 323: 363-72.
- [28] Yang JC, Wang ZW, Lin Z, et al. Coupled dictionary training for image super-resolution[J]. *IEEE Trans Image Process*, 2012, 21(8): 3467-78.
- [29] Hong X, Zan Y, Weng F, et al. Enhancing the image quality *via* transferred deep residual learning of coarse PET sonograms [J]. *IEEE Trans Med Imaging*, 2018, 37(10): 2322-32.
- [30] Ren SQ, He KM, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Trans Pattern Anal Mach Intell*, 2017, 39(6): 1137-49.

(编辑:吴锦雅)