

〈太赫兹技术〉

基于生成式对抗网络的太赫兹图像增强

张鹏程^{1,2}, 何明霞^{1,2}, 陈 硕^{1,2}, 张洪桢^{1,2}, 张欣欣^{1,2}

(1. 天津大学 测试计量技术及仪器国家重点实验室, 天津 300072;

2. 天津大学 天津大学精密仪器与光电子工程学院, 天津 300072)

摘要: 太赫兹扫描成像中, 由于激光器功率波动和仪器振动等原因, 导致图像对比度较低, 成像质量有待提高, 且目前针对太赫兹图像的处理还停留在传统算法阶段。本文结合深度学习思想, 提出了一种基于生成式对抗网络的图像增强方法。通过对训练集图像引入模糊和噪声, 学习低质量图像和高质量图像之间的映射关系, 并将其应用在真实太赫兹图像中。实验结果表明, 与双边滤波、非局部均值滤波等传统算法相比, 本文方法可在改善图像细节的基础上显著提高图像对比度, 且视觉体验良好, 这为太赫兹图像增强提供了新思路。

关键词: 太赫兹图像; 神经网络; 图像增强; 图像对比度

中图分类号: TP751

文献标识码: A

文章编号: 1001-8891(2021)04-0391-06

Terahertz Image Enhancement Based on Generative Adversarial Network

ZHANG Pengcheng^{1,2}, HE Mingxia^{1,2}, CHEN Shuo^{1,2}, ZHANG Hongzhen^{1,2}, ZHANG Xinxin^{1,2}

(1. State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China;

2. School of Precision Instrument and Optoelectronics Engineering, Tianjin University, Tianjin 300072, China)

Abstract: In terahertz scanning imaging, the image contrast is low due to laser power fluctuation and instrument vibration, and the imaging quality needs to be improved. At present, the processing of terahertz image is still in the traditional algorithm stage. In this paper, an image enhancement method based on Generative Adversarial Network is proposed, which includes the idea of deep learning. By introducing blur and noise into the training set image, the mapping relationship between low-quality images and high-quality images is learned and applied to real terahertz images. The experimental results show that, compared with traditional algorithms such as bilateral filtering and non-local mean filtering, this method can significantly improve the image contrast on the basis of improving image details, and has a good visual sense, which provides a new idea for terahertz image enhancement.

Key words: terahertz image, neural network, image enhancement, image contrast

0 引言

太赫兹成像是一种利用太赫兹波作为信号源的成像技术^[1-2], 与 X 射线相比, 太赫兹波对生物体没有电离辐射, 所以非常适用于安检、无损检测等非接触式成像应用。由于成像过程中激光器功率的波动以及其他因素的影响, 如今的太赫兹成像系统得到的图像分辨率、对比度都比较低。通过改进硬件设备来提高成像质量成本比较高昂, 而利用图像处理技术来提高成像质量已经成为一种高效实用的

方法^[3]。

目前, 太赫兹图像的降噪增强算法还停留在传统算法阶段, 如经常使用的小波变换、非局部均值滤波^[4]、双边滤波^[5]等算法。这些传统算法虽有一定的效果, 但缺点也比较明显。如经小波变换处理后的图像灰度级会偏离原始图像的灰度级, 会造成边界模糊和噪声放大; 双边滤波仅考虑了图像的局部信息, 处理后的图像整体效果欠佳; 非局部均值滤波虽考虑了图像的整体信息, 但不能根据具体情况自适应调节滤波参数。

收稿日期: 2019-09-29; 修订日期: 2019-11-04.

作者简介: 张鹏程 (1995-), 男, 河南周口人, 硕士生, 研究方向为太赫兹成像及图像处理, E-mail: 1298216729@qq.com.

随着深度学习领域的快速发展,利用神经网络进行图像处理已经成为了研究热点。Dong 等人于2014年提出了一个基于卷积神经网络的图像超分辨率框架 SRCNN (super-resolution convolutional neural network),该框架通过学习低分辨率图像到高分辨率图像之间的映射关系,可有效提高图像的分辨率^[6]。Christian Ledig 等人通过优化损失函数于2017年提出了超分辨率生成式对抗网络框架 SRGAN (super-resolution generative adversarial network),该框架采用了小卷积核和较深的网络结构,使生成图像的分辨率达到更高水平^[7]。

由于神经网络需要大量图像进行训练,对于太赫兹图像来说,由于成像速度和仪器的限制,太赫兹图像数据量相对不足^[8],因此上述神经网络框架更多地应用在可见光图像上,利用神经网络处理太赫兹图像却鲜有报道。本文提出了一种利用 SRGAN 网络对 THz 图像进行增强的算法,根据太赫兹图像对比度低、模糊不清的特点,建立伪 THz 图像库,使其具备真实 THz 图像的特点。通过训练,得到模糊图像到清晰图像的映射关系,并将其应用在真实 THz 图像上,这为太赫兹图像的增强处理提供了新思路。

1 相关理论介绍

1.1 生成对抗网络基本原理

生成对抗网络 GAN (generative adversarial network) 的核心思想源于博弈论中的纳什均衡,其由两部分组成:生成器 G (Generator) 和判别器 D (Discriminator)。生成器的目的是尽量使生成的数据符合真实数据分布,判别器的作用是判断输入数据是来自于真实数据还是生成器的数据。生成器和判别器不断迭代优化自己的生成能力和判别能力,直到二者达到一个纳什平衡。GAN 的计算流程如图 1 所示。

在图 1 中,我们用可微分函数 G 和 D 分别表示

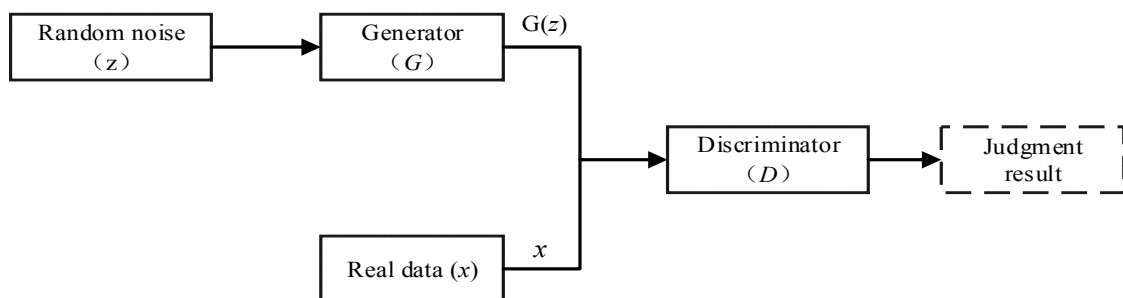


图 1 GAN 流程图
Fig.1 Flow chart of GAN

生成器和判别器。假设真实数据分布为 P_{data} , 随机噪声 z 通过生成器生成尽量符合分布 P_{data} 的样本 $G(z)$ 。对于判别器而言,当输入数据来自于真实数据时,输出为 1; 当输入数据来自于 $G(z)$ 时,输出为 0。判别器 D 的目标是对输入数据做出正确判断,生成器 G 的目标是使自己生成的数据无限趋近于真实数据分布, D 和 G 相互对抗并迭代优化使得二者性能不断提升,最终当 D 的判别能力达到一定程度,且无法准确判断数据来源时,认为生成器 G 已经学到了真实的数据分布。

1.2 GAN 训练机制

GAN 的目标函数描述如下:

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

式中: $E(\cdot)$ 表示期望值的计算; $P_{data}(x)$ 表示真实数据分布; x 为真实样本; $D(x)$ 表示 x 被 D 判断为真实样本的概率; $P_z(z)$ 代表先验分布, z 为采样于该分布的噪声; $G(z)$ 表示噪声 z 通过 G 后生成的样本; $D(G(z))$ 表示生成样本被 D 判断为真实样本的概率。在 GAN 中,生成器的目的是使生成样本尽可能的接近真实样本,即 $D(G(z))$ 越趋近于 1 越好,此时 $V(D, G)$ 会变小;判别器的目的是准确分辨出生成样本和真实样本,即 $D(x)$ 趋近于 1,而 $D(G(z))$ 趋近于 0,此时 $V(D, G)$ 会增大。

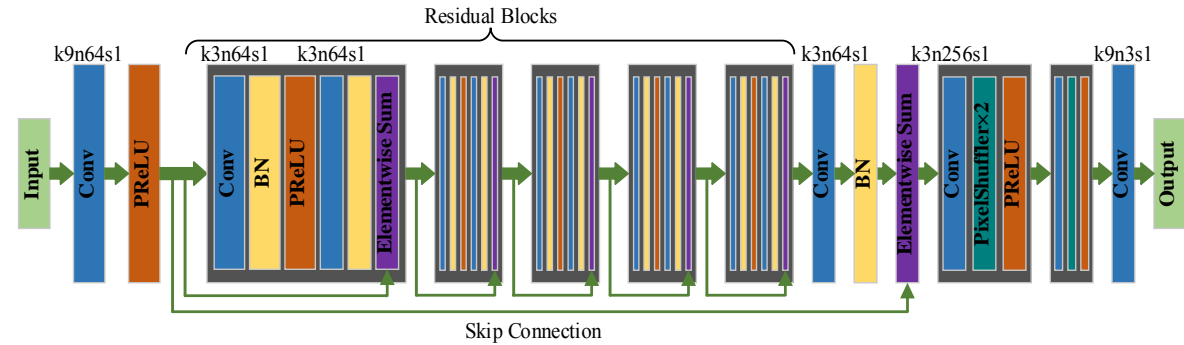
在 GAN 的训练过程中,我们需要训练判别器 D ,使其判断数据来源的准确率达到最大;同时,需要训练生成器 G 使 $\lg(1 - D(G(z)))$ 最小。整个训练过程可以采用交替优化的方法:先固定生成器 G ,训练判别器 D ,使 D 的判别准确率达到最大;然后固定判别器 D ,训练生成器 G ,使 D 的判别准确率达到最小,当且仅当 $P_{data} = P_g$ (由 G 生成的数据分布) 时可得全局最优解。实际训练时,一般对判别器的参数更新 k 次再对生成器的参数更新一次。

2 本文方法

2.1 神经网络结构

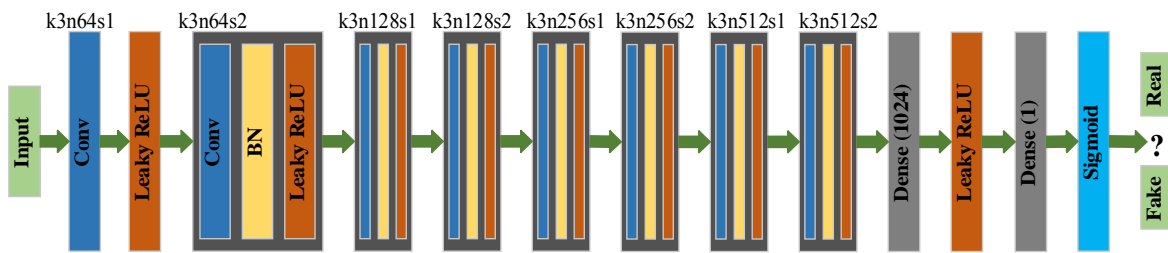
本文使用了 SRGAN 网络结构, 在该网络中采用了较深的网络结构和小卷积核, 使图像的重建效

果达到较高的水平。在以往的研究中, 该框架更多的应用于可见光图像的超分辨率重建, 我们将其应用在太赫兹图像中, 也得到了不错的效果。其网络结构如图 2 所示。



(a) 生成器网络结构

(a) Structure of generator network



(b) 判别器网络结构

(b) Structure of discriminator network

图 2 SRGAN 网络框架结构

Fig.2 Framework of SRGAN

为了解决深层神经网络在训练过程中梯度爆炸等问题, 在生成器网络中引入了残差模块^[9-10]。输入图像经过卷积层和激活函数后, 进入残差模块中。图中每一个残差模块都采用了两层卷积层, 每层有 64 个卷积核, 每个卷积核大小为 3×3 , 卷积层之后利用 BN 层 (Batch Normalization) 和激活函数对输出进行处理。在生成器最后, 采用了两个经过训练的子像素卷积层来提高输入图像的分辨率^[11]。

为了区分真实图像和生成图像, 本文训练了一个判别网络, 体系结构如图 2(b)所示。该网络参考了 Radford 等人总结的神经网络构建建议^[12], 并使用了 LeakyReLU 作为激活函数。该判别网络包含 8 个卷积层, 每个卷积层使用的卷积核尺寸均为 3×3 。和 VGG 网络中一样^[13], 卷积核数量不断增加, 直到从 64 增加到 512 个。每当卷积核数量增加一倍时, 都使用跨步卷积来降低图像分辨率。在得到 512 个特征图之后, 通过两个全连接层和一个最终的 S 型激活函数, 获得样本分类的概率。

2.2 损失函数定义

损失函数的定义对于网络性能至关重要, 该神经网络的损失函数主要由两部分组成: 内容损失和对抗性损失^[14]。其数学表达式如下:

$$L = L_1 + 10^{-3}L_2 \quad (2)$$

式中: L_1 表示内容损失; L_2 表示对抗性损失。

对于内容损失来说, 目前大多数方法通过计算图像的均方误差得到, 这样重建图像具有较高的 PSNR, 但缺点是图像高频信息缺失严重, 整体过于平滑, 视觉体验一般。SRGAN 采用了与视觉体验更吻合的内容损失函数^[15], 表达式如下:

$$L_1 = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\Phi_{i,j}(I^{HR})_{x,y} - \Phi_{i,j}(G_{\theta}I^{LR})_{x,y})^2 \quad (3)$$

式中: $G_{\theta}(I^{LR})$ 表示生成图像; I^{HR} 表示真实图像; $W_{i,j}$ 和 $H_{i,j}$ 代表各个特征图的尺寸。

为了使判别网络适用于风格不同的图像, 除了内容损失以外, 还引入了对抗性损失。对抗性损失

L_2 是基于判别器 D 在所有训练样本上的概率定义的:

$$L_2 = \sum_{n=1}^N -\lg D_{\theta}(G_{\theta}(I^{LR})) \quad (4)$$

式中: $D_{\theta}(G_{\theta}(I^{LR}))$ 表示重建图像 $G_{\theta}(I^{LR})$ 是真实图像的概率^[16]。

3 实验过程及结果分析

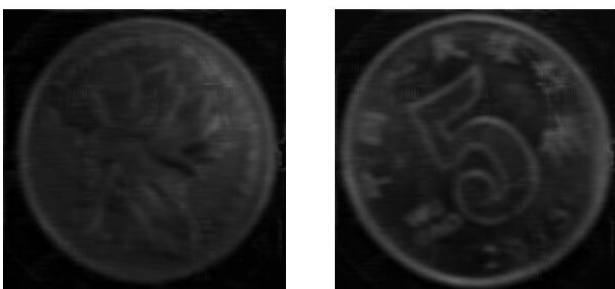
3.1 数据集准备

由于受到太赫兹仪器设备和成像速度的限制,截止到目前并没有足量的数据实现神经网络的训练。本文利用图像处理技术创建了一个图像集,使其具备 THz 图像分辨率低、对比度低的特点。我们利用手机拍摄了 25 张高清样品照片,通过翻转、缩放、裁剪以及不同角度的旋转将数据扩增 40 倍,这 1000 张图片构成了标签图像集。

为了生成对应的训练集,且使训练集具备真实太赫兹图像对比度低等特点,对标签图像进行了如下处理生成训练集图像:

- 1) 对数变换,太赫兹图像一般亮度较暗,对比度较低,通过对标签图像进行对数变换,使其亮度更接近真实太赫兹图像;
- 2) 高通滤波,太赫兹图像一般含有高频噪声,对太赫兹图像进行高通滤波,将提取出来的噪声叠加到标签图像上,使其具备太赫兹图像的噪声特点;
- 3) 模糊处理,由于扫描成像过程中设备的移动和抖动,太赫兹图像会带有模糊,因此需要在标签图像上加上运动模糊和高斯模糊。

经过以上处理,生成的训练图像示例如图 3(a)所示,该图像对比度和分辨率较低,这与 THz 图像的基本特点吻合。但值得注意的是,由于不同的成像系统各有特点,生成的图像也有差别,为了提高本算法的普适性,本文并未对 THz 图像和生成训练图像的相似性进行定量评估。图 3(b)为标签图像示例,本文采用监督学习的方式,学习训练图像到标签图像的映射关系。



(a) 训练图像示例

(a) Example of training images



(b) 标签图像示例

(b) Example of label images

图3 图像数据集示例

Fig.3 Image dataset example

3.2 训练细节及参数配置

本文神经网络的训练过程基于 64 位 Windows 10 操作系统,使用 Python 语言和 Tensor flow 深度学习框架,在 Google Colab 云计算平台进行训练,该平台配置了 NVIDIA Tesla K80 GPU 加速运算。首先,训练基于 MSE 的 SRResnet 网络,学习率设置为 10^{-3} ,迭代次数为 5×10^4 次。然后在训练 SRGAN 时,将训练好的 SRResnet 网络参数作为 SRGAN 网络的初始化参数,以避免训练过程中出现不必要的局部最优解,学习率设置为 10^{-3} ,迭代次数同样为 5×10^4 次。训练过程中,对生成器和判别器网络进行交替更新,参考 Goodfellow 等人的工作^[16],使用的 k 值为 1。

3.3 实验结果及比较

以上整个训练过程持续了 20 h,并利用 Tensor board 监测了图像 PSNR 值和网络损失函数变化曲线,如图 4 所示。

从图 4 可以看出,随着训练次数的增加,在宏观上,图像的 PSNR 值呈现出逐渐变大的趋势,这说明在训练过程中,生成器生成的图像质量逐渐提高;网络的损失函数随着训练次数的增加呈现出递减的趋势,说明该网络结构及参数配置较好,网络收敛效果显著。

为了更直观地突出本文算法的有效性,我们将训练好的网络应用在真实的太赫兹图像上,并与传统算法处理结果进行比较,效果如图 5 所示。本文的实验对象是一枚经过太赫兹反射成像的硬币,在成像过程中,由于激光器功率波动和仪器设备的振动,原始图像存在一定的噪声和模糊,且对比度较低。经传统算法处理后的图像,可以滤除部分噪声,但图像视觉体验依旧模糊,且对比度提升不明显,图像细节也没有得到有效改善。而经本文算法处理后的图像,滤波效果显著,对比度明显提高,且图像清晰度得到改善,具有更丰富的局部细节。

为了客观评价本文算法的效果,本文以原始图像为参考图像,计算了图 5 中各个图像的峰值信噪比 (peak signal to noise ratio, PSNR), 并利用中心像素与周围四近邻像素的灰度值计算了各个图像的对比度,结果如表 1 所示。对于 PSNR 来说,由于本文算法在图像细节上改变更多,且计算时以原始图像作为参考图像,再加上本文损失函数定义并

未以均方误差为基础,所以 PSNR 会偏低。根据以往经验,PSNR 有时与人眼视觉体验并不相符,所以低 PSNR 并不影响视觉效果。从对比度角度来说,本文算法与传统算法相比,对于提高图像对比度效果非常显著,该结果在图 5 中也有直观体现,这刚好符合本实验室项目的实际需求。

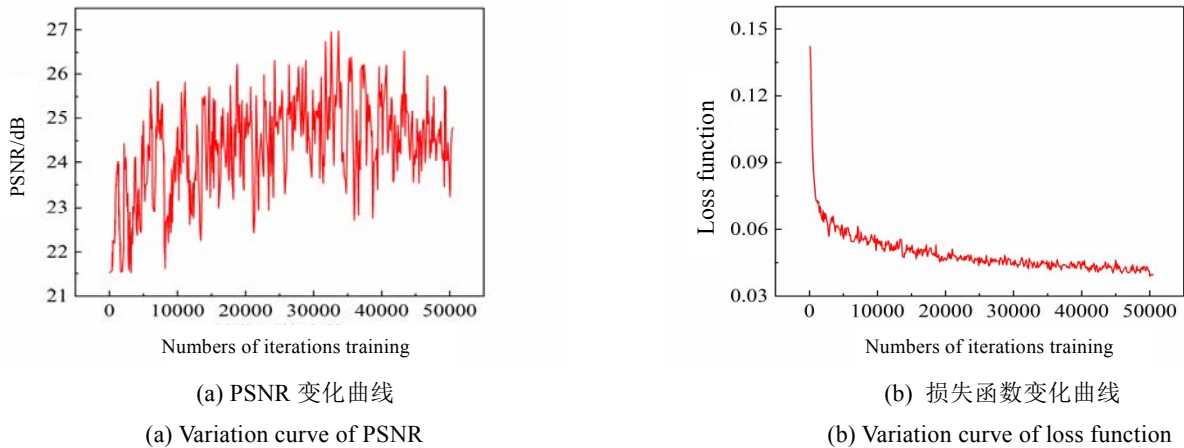


图 4 训练过程中曲线变化

Fig.4 Variation curves change during training

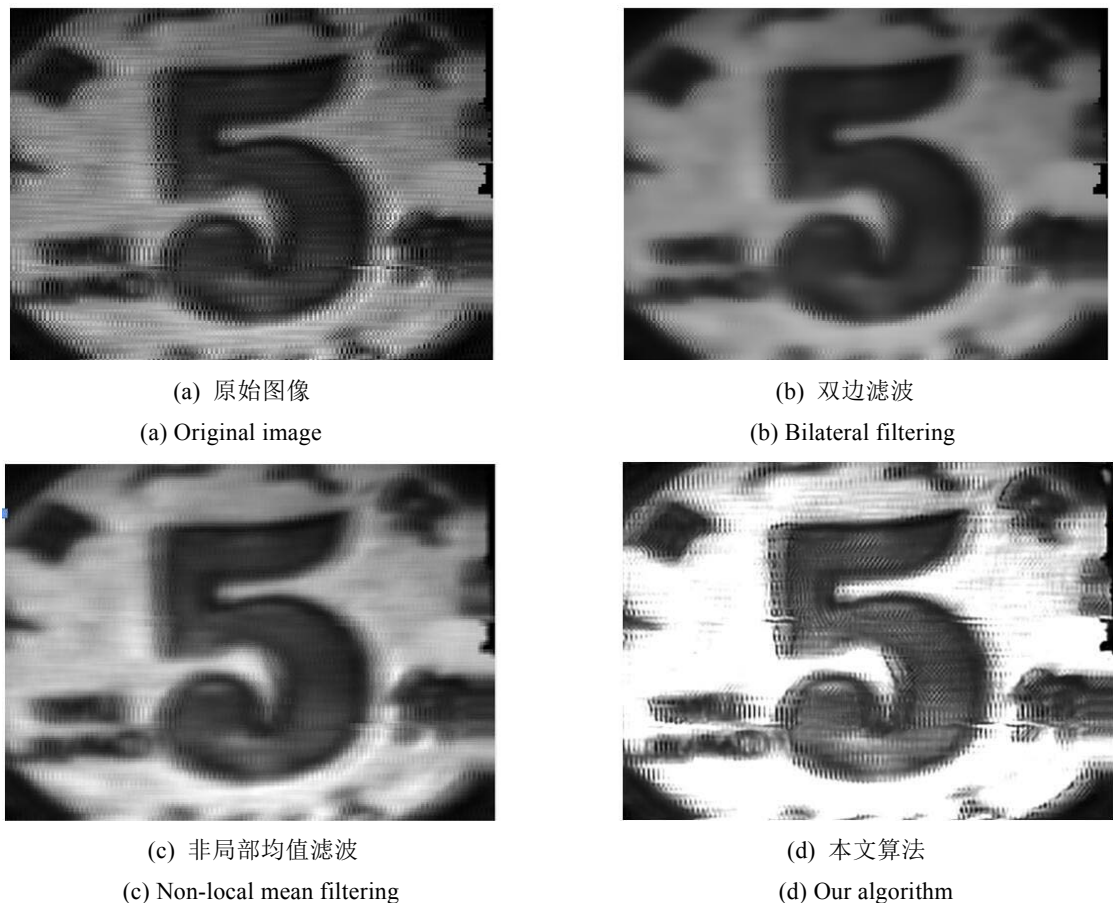


图 5 不同算法实验结果比较

Fig.5 Comparison of experimental results of different algorithms

表1 不同方法 PSNR、对比度计算结果

Table 1 PSNR and contrast calculation results by different methods

	Bilateral filtering	Nonlocal mean filtering	Our algorithm
PSNR	28.18	26.87	25.28
Contrast	178.69	134.48	375.98

4 结论

本文基于前人提出的 GAN 原理, 利用自己建立的图像库, 成功训练了 SRGAN 网络, 并将其应用于太赫兹图像增强处理上。通过进行对照实验, 将本文算法与几种传统算法相比, 实验结果表明, 本文算法在解决太赫兹图像数据量不足的基础上, 可以显著提高太赫兹图像的对比度, 且处理后的图像细节更加丰富, 图像质量及视觉体验比传统算法更加优越。为了提高算法的普适性, 本文对 THz 图像和生成的训练图像之间的相似性只做了定性评价, 这也是今后工作有待改进的地方。

基于深度学习的图像处理算法虽起步较晚, 但与传统算法相比优点突出, 能弥补传统算法的很多不足, 随着人工智能技术的发展, 该领域也将进一步成为研究热点。

参考文献:

[1] 郑显华, 王新柯, 孙文峰, 等. 太赫兹数字全息术的研发与应用[J]. 中国激光, 2014, 41(2): 0209003.
ZHENG Xianhua, WANG Xinke, SUN Wenfeng, et al. Developments and applications of the terahertz digital holography[J]. *Chinese J. Lasers*, 2014, 41(2): 0209003.

[2] 张馨, 赵源萌, 张存林. 被动式太赫兹图像分割算法[J]. 强激光与粒子束, 2013, 25(6): 1597-1600.
ZHANG Xin, ZHAO Yuanmeng, ZHANG Cunlin. Passive terahertz image segmentation algorithm[J]. *High Power Laser and Particle Beams*, 2013, 25(6): 1597-1600.

[3] 徐利民. 高分辨太赫兹图像处理[D]. 西安: 中国科学院研究生院(西安光学精密机械研究所), 2013.
XU Limin. High Resolution Terahertz Image Processing[D]. Xi'an: University of Chinese Academy of Sciences (Xi'an Institute of Optics and Precision Mechanics), 2013.

[4] 李琦, 杨永发, 胡佳琦. 一种用于太赫兹共焦扫描图像复原的复合算法[J]. 红外与激光工程, 2015, 44(1): 321-326.
LI Qi, YANG Yongfa, HU Jiaqi. A composite algorithm used for terahertz confocal scanning image restoration[J]. *Infrared and Laser*

Engineering, 2015, 44(1): 321-326.

[5] 杨永发, 李琦. 双边滤波算法的太赫兹共焦扫描图像去噪应用[J]. 激光与光电子学进展, 2015, 52(12): 121101.
YANG Yongfa, LI Qi. Application of bilateral filtering algorithm on terahertz confocal scanning image denoising[J]. *Laser & Optoelectronics Progress*, 2015, 52(12): 121101.

[6] DONG C, CHEN C L, HE K, et al. *Learning a Deep Convolutional Network for Image Super-Resolution*[M]//Computer Vision – ECCV Springer International Publishing, 2014: 184-199.

[7] Ledig C, Theis L, Huszar 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), IEEE, 2017: 4681-4690.

[8] 卢贺洋, 苏胜君, 袁明辉, 等. 太赫兹图像的超分辨率重建[J]. 红外技术, 2019, 41(1): 59-63.
LU Heyang, SU Shengjun, YUN Minghui, et al. Super-resolution reconstruction of terahertz image[J]. *Infrared Technology*, 2019, 41(1): 59-63.

[9] Johnson J, Alahi A, Li F. Perceptual losses for real-time style transfer and super-resolution[C]//In European Conference on Computer Vision (ECCV), Springer, 2016: 694-711.

[10] Gross S, Wilber M. Training and investigating residual nets, online[EB/OL].[2016-02-04]. <http://torch.ch/blog/>.

[11] 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]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 1874-1883.

[12] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[C]//International Conference on Learning Representations (ICLR), 2016: 1-16.

[13] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[C]//International Conference on Learning Representations (ICLR), 2015: 268-282.

[14] Bruna J, Sprechmann P, Lecun Y. Super-resolution with deep convolutional sufficient statistics[C]//International Conference on Learning Representations (ICLR), 2016: 352-369.

[15] Gatys L A, Ecker A S, Bethge M. Texture synthesis using convolutional neural networks[C]//Advances in Neural Information Processing Systems (NIPS), 2015: 262-270.

[16] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//Advances in Neural Information Processing Systems (NIPS), 2014: 2672-2680.