光电技术及应用

DOI：10.16818/j.issn1001－5868.2023060201

基于改进生成对抗网络的红外图像增强算法

吴国瑞[1*]，王　峰[2]，周平华[1]，马　晨[3]，赵　伟[3]，康智强[3]

（1. 太原理工大学 电子信息与光学工程学院，山西 晋中 030600；2. 太原理工大学 电气与动力工程学院，太原 030000；
3. 中国电子科技集团公司第三十三研究所，太原 030000）

摘　要： 针对红外图像细节分辨率不高、目标边缘模糊等，提出一种基于改进生成对抗网络的红外图像增强算法。首先，基于编码解码网络 U-Net 构建生成器，优化 U-Net 跳跃连接方式，融合全局上下文模块，实现全局和局部特征的上下文建模；然后，基于胶囊网络构建鉴别器，结合 Res2Net 改进胶囊网络结构，并对胶囊网络全连接层进行反卷积重构，实现多尺度图像特征提取，减少模型参数冗余。实验表明，与当前主流算法相比，该算法能有效突出细节信息、抑制噪声，提高图像分辨率和视觉效果。

关键词： 深度学习；红外图像增强；生成对抗网络；胶囊网络；U-Net

中图分类号： TP391　**文章编号：** 1001－5868(2023)05－0782－06

Infrared Image Enhancement Algorithm Based on Improved Generative Adversarial Network

WU Guorui[1], WANG Feng[2], ZHOU Pinghua[1], MA Chen[3], ZHAO Wei[3], KANG Zhiqiang[3]

(1. College of Electronic Information and Optical Engineering, Taiyuan University of Technology, Jinzhong 030600, CHN;
2. College of Electrical and Power Engineering, Taiyuan University of Technology, Taiyuan 030000, CHN;
3. 33rd Institute of China Electronics Technology Group Corporation, Taiyuan 030000, CHN)

Abstract: In order to address the problems of poor detail resolution and blurred target edges in infrared images, an image enhancement method based on improved Generative Adversarial Network is proposed. Firstly, the generator was constructed based on the codec network U-Net, optimizing the U-Net skip connection method and fusing the global context module to achieve contextual modelling of global and local features. Secondly, the discriminator is constructed based on the Capsule Networks, the capsule network structure is improved by combining with Res2Net structure and the fully connected layer of the Capsule Networks was deconvolutionally reconfigured to achieve multi-scale image feature extraction and reduce model parameters redundancy. The experimental results show that, compared with the current mainstream algorithms, the algorithm in this paper can effectively highlight the detail information, suppress the noise, and improve the image resolution and visual effect.

Key words: deep learning; infrared image enhancement; generative adversarial network; capsule networks; U-Net

0　引言

近年来，随着红外技术快速发展，红外成像已广泛应用于军事监测、航空工业和医疗卫生[1]等领域。但在实际应用中，红外图像受采集设备限制和恶劣环境影响，存在细节模糊、对比度低等问题，对后续

收稿日期：2023－06－02.

基金项目：山西省留学人员科技活动项目(20230063)；山西省重点研发计划项目(202102150101008).

* **通信作者：**吴国瑞　**E-mail：** 2654910062@qq.com

红外图像分析和目标检测造成不利影响。因此提高红外图像质量，增强图像细节特征变得尤为重要。

红外图像增强算法主要通过抑制背景噪声，消除无用场景信息，突出图像整体或细节特征来达到改善图像质量的目的。红外图像增强算法主要可分为传统红外图像增强算法和基于深度学习的红外图像增强算法。其中传统图像增强算法大多以直方图均衡为基础，使红外图像像素均匀分布在灰度级范围内，以此来提高图像对比度。在此基础上，许多学者针对直方图均衡算法进行改进，如基于非均匀直方图的图像增强算法[2]通过设置相对边缘强度指标，抑制相对均匀区域中的噪声，锐化图像边缘；Zhang 等[3]通过改进 Sigmoid 函数建立局部熵映射直方图，使局部熵映射直方图自适应地分割为多个子直方图进行均衡，实现红外图像增强。除上述增强算法外，还包括基于 Retinex 的增强算法[4-5]、基于小波变换的增强算法[6]和反锐化掩膜算法[7]等。

基于深度学习的图像增强算法大致可分为学习前增强和学习后增强[8]。学习前增强[9]指研究人员对原有数据进行变换来增加数据多样性，从而提高算法模型的鲁棒性和网络泛化能力，解决模型过拟合问题。学习后增强[10]也被广泛研究，Dong 等[11]提出一种超分辨率卷积神经网络，通过建立低分辨率和高分辨率图像映射关系，简化网络训练过程，实现图像高质量重建。自 Goodfellow 等[12]在 2014 年提出生成对抗网络(GAN)之后，其在图像数据增强、高质量图像生成领域展现了卓越能力。Chen 等[13]利用彩色图像作为输入样本，改进生成器和鉴别器模型，在 GAN 损失函数中加入正则化参数，缓解生成器网络过拟合，约束模型生成更高质量红外图像。上述算法复杂度低、网络模型训练灵活，但皆采用传统卷积架构，无法兼顾图像的全局特征和局部信息，难以凸显红外图像的细节特征和边缘纹理。

根据上述算法的特点与缺陷，本文提出一种基于改进生成对抗网络的红外图像增强算法。在生成器中将编码解码网络 U-Net[14]与全局上下文模块相结合，有效提取图像的全局与局部特征，增强图像特征表达能力；在鉴别器中将胶囊网络[15]与 Res2Net[16]结合，对胶囊网络全连接层进行反卷积网络重构。最后将本文算法与当前主流算法进行对比实验，结果表明所提算法在客观指标和主观评价中都具有明显优势。

1　本文算法

1.1　生成对抗网络

生成对抗网络是一种无监督深度学习模型，产生于博弈论。其结构主要可分为生成器(G)和鉴别器(D)两部分，如图 1 所示。其中生成器主要功能是捕捉随机噪声 z 的分布，产生具有相同分布、可以欺骗鉴别器的数据样本。鉴别器是一个二元分类器，用于评估生成数据是否属于真实数据。两者之间不断进行博弈对抗，最终生成对抗网络达到纳什平衡状态，此时鉴别器无法区分网络生成样本和真实样本。其博弈过程具体表达式为：

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

式中，x 表示输入到鉴别器的真实样本，$D(x)$表示真实样本 x 的概率，z 表示随机噪声，$G(z)$表示生成器所产生的虚假图像，$D(G(z))$表示生成器产生虚假图像概率。

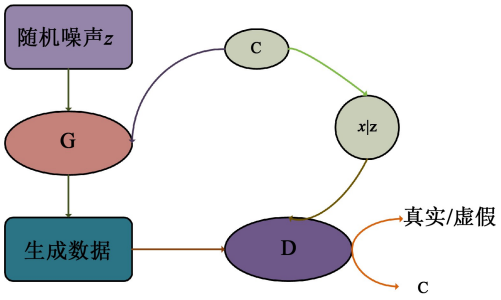

图 1　生成对抗网络结构图

1.2　生成器模型

U-Net[14]是 2015 年由 Ronneberge 提出的 U 型网络结构，是一种基于编码器和解码器的全卷积神经网络。U-Net 由上采样和下采样模块构成，不包含全连接层。其中编码器由多个卷积模块组成，通过下采样中最大池化层压缩特征图尺寸，将特征图传递到解码器中。在解码器中，每个卷积模块不仅输入上采样模块输出的特征图，还接收编码器对应模块输出的特征图，保证输出图像和原始图像的相关性。

针对 U-Net 中卷积层仅关注区域特征，作用于局部感受野，本文在 U-Net 跳跃连接中引入全局上下文模块[17]（GCBlock），如图 2 所示。利用 GCBlock 中自注意力机制建立远距离像素之间的依赖关系，通过计算像素与像素之间的函数值获取像素之间的相关性。全局上下文模块执行过程如下：

1)采用卷积和激活函数获取自注意力权重,然后进行注意力池化获取全局特征;2)通过 1×1 卷积特征变换获得通道依赖性;3)将全局上下文特征与所有位置特征融合。全局上下文运算公式如式(2)所示:

$$g(x)=X_i+\boldsymbol{W}_{v2}R\left(\mathrm{LN}\left(\boldsymbol{W}_{v1}\sum_{j=1}^{N_p}\frac{\mathrm{e}^{\boldsymbol{W}_kX_j}}{\sum_{m=1}^{N_p}\mathrm{e}^{\boldsymbol{W}_kX_m}}X_j\right)\right) \tag{2}$$

式中,$g(x)$表示全局上下文特征,$\boldsymbol{W}_k$ 和 $\boldsymbol{W}_v$ 表示线性变换矩阵,R 表示 ReLU 激活函数,LN 表示层归一化操作。

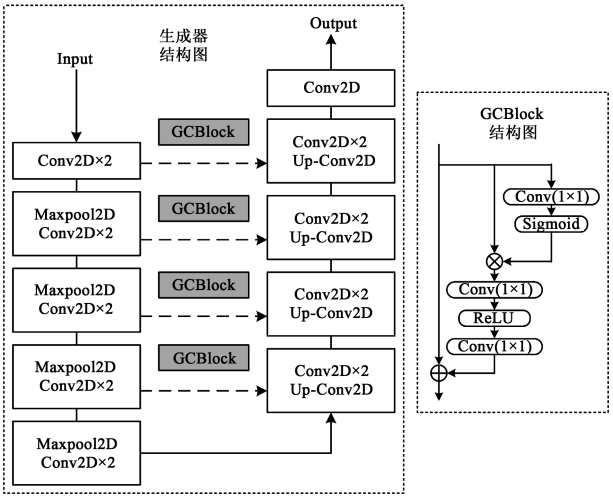

图 2 生成器网络结构图

1.3 鉴别器模型

胶囊网络是 Hinton 于 2017 年提出的一种神经网络模型,通过胶囊来代替传统神经元,从而提高网络表达能力和泛化能力。胶囊网络结构主要由 4 部分组成,即卷积层、主胶囊层、数字胶囊层和全连接层。胶囊网络优势在于通过向量输入输出,能够更好地"理解"图像和语义的关系,具备较强的表达能力,并且通过胶囊建立起的三维图像模型,输出向量更具细节性,更能反应图像状态。

本文在卷积层和第一层胶囊层之间引入 Res2Net 模块,使胶囊网络能对输入图像进行多尺度特征提取,增加感受野范围,从而捕获更有效特征,如图 3 所示。Res2Net 模块是一种新型多尺度骨干网络结构,相比 ResNet 有更好的鲁棒性和泛化能力,如图 3 左侧。输入图像通过 1×1 卷积处理之后的特征图分为 S 个子模块 X_i,其中 $i\in\{1,2,\cdots,S\}$,S 的取值根据图像实际大小所定。每个子模块具有相同尺度大小,除 X_1 之外,其他子模块 X_i 都有对应的 3×3 卷积核,X_i 经卷积计算后的结果作为输入数据添加到下一组残差连接中,具体运算式见式(3):

$$Y_i=\begin{cases}X_i & i=1\\ K_i(X_i) & i=2\\ K_i(X_i+Y_{i-1}) & 2<i\leqslant S\end{cases} \tag{3}$$

式中,K_i 表示卷积操作,Y_i 表示输出数据。

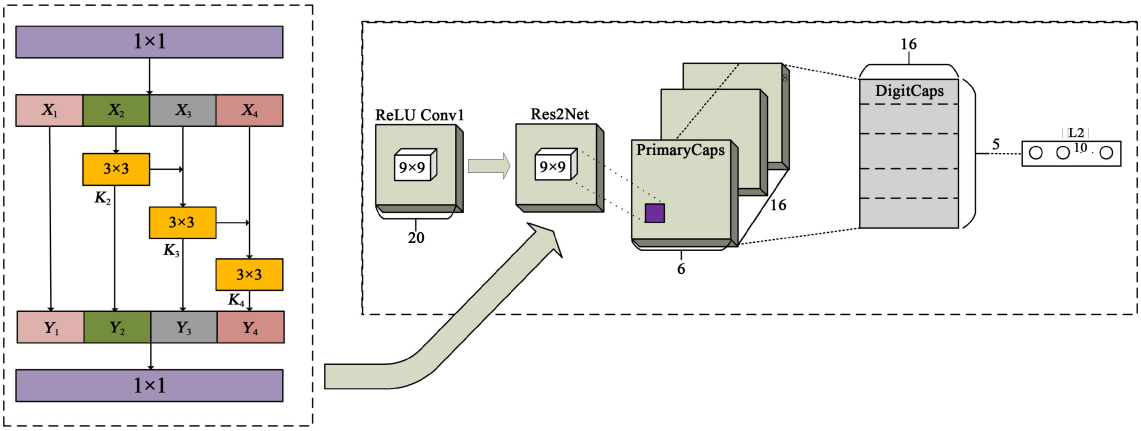

图 3 鉴别器结构图

每个 3×3 卷积操作都可以潜在接受所有前置模块输出的特征信息,每个输出都能增大感受野,使 Res2Net 模块能获取不同数量和不同感受野大小的特征组合,更有利于提取全局特征信息。

为解决因全连接层过多造成的参数冗余,对胶囊网络全连接层采用反卷积网络重构,如图 4 所示。通过对数字胶囊层的输出向量进行筛选,只保留正确胶囊层对应向量作为输出,通过一层全连接层,采用 BatchNormalization 优化和 ReLU 激活函数得到特征图,最后通过固定尺寸卷积核进行反卷积操作,得到重构后的生成图像。BatchNormalization 优化方法将神经网络中间层的输入保持均值为 0,方差

为 1 的标准正态分布。其计算公式如下：

$$\mu_{\mathrm{B}}=\frac{1}{m}\sum_{i=1}^{m}x_i \tag{4}$$

$$\delta_{\mathrm{B}}^2=\frac{1}{m}\sum_{i=1}^{m}(x_i-\mu_{\mathrm{B}})^2 \tag{5}$$

$$\hat{x}_i=\frac{x_i-\mu_{\mathrm{B}}}{\sqrt{\delta_{\mathrm{B}}^2+\varepsilon}} \tag{6}$$

$$y_i=\gamma\hat{x}_i+\beta\equiv BN_{\gamma,\beta}(x_i) \tag{7}$$

式中，μ_{B} 为批训练均值，x_i 为第 i 个输入数据，m 为批训练样本数量，式(5)表示方差，式(7)表示对输入数据进行归一化处理，其中参数 γ 和参数 β 为可训练参数，保持输入数据原始分布。该标准化操作可使函数对非线性输入保持敏感，同时避免梯度消失问题，加快收敛速度。

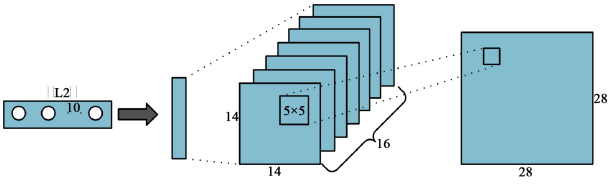

图 4　胶囊网络全连接层重构图

2　实验结果与分析

2.1　环境配置

由于可供选择的红外图像的数据集较少，本文通过将不同复杂背景的红外图像数据集进行整合作为训练数据集，其中包括地/空背景下红外图像弱小飞机目标检测跟踪数据集[18]、红外弱小目标数据集[19]共 7 200 张图片。本文算法模型训练基于 64 位 Windows10 操作系统，使用 Pytorch 深度学习框架，该平台使用 NVIDIA1060 显卡，32 G 运行内存。

2.2　评价指标

评价红外图像增强效果可从两部分看：

(1)人类主观评价，比如红外图像弱小目标是否可清晰呈现，目标区域对比度是否较原图像有所提升，边缘区域是否得到锐化处理；

(2)客观数据指标评价，本文采用峰值信噪比(Peak Signal-to-Noise Ratio，PSNR)，结构相似性(Structural Similarity，SSIM)两个客观指标评价图像质量。PSNR 反映图像像素点误差的大小，用符号 n 表示，其公式如式(8)所示：

$$n=10\times\lg\left(\frac{L^2}{\frac{1}{N}\sum_{i=1}^{N}(I(i)-J(i))^2}\right) \tag{8}$$

式中，N 表示像素数，I 表示原始图像，J 表示重建图像。SSIM 用于衡量两幅图像的相似度，也用来反映图像失真程度，用符号 s 表示，其表达式如式(9)所示：

$$s(I,J)=\frac{(2\mu_I\mu_J+c_1)(2\sigma_{IJ}+c_2)}{(\mu_I^2+\mu_J^2+c_1)(\sigma_I^2+\sigma_J^2+c_2)} \tag{9}$$

式中，μ_I 和 μ_J 分别表示 I 方向上和 J 方向上的均值，σ_I 和 σ_J 为标准差，c_1 和 c_2 为常数。

2.3　实验结果与分析

为证明所提算法的可行性及其优越性，本文对不同场景下的红外图像进行对比实验，图像尺寸为 256×256 和 256×196。本文按照算法步骤训练网络模型，并选取四个场景的红外图像分别在 SSR[20]、MRSCR[21]、ESRGAN[22]、文献[3]和本文算法进行分析比较。四组实验得出的峰值信噪比 PSNR 和 SSIM 数据如表 1 所示，四组图像输入模型得到增强后图像如图 5 所示。

本文对算法增强效果采用主观评估和客观数据分析两种方式进行实验验证，客观数据依据两个数据指标 PSNR，SSIM 进行分析，主观评估采用随机询问不同人对不同图像增强效果的看法。

本文选取不同场景的红外图像增强结果进行主观评估，如图 5 所示。场景一和场景二中含有纹理特征丰富的树木；场景二和场景四中含有复杂多样的道路、车辆和建筑，细节特征较为丰富；场景三的原图中有层次模糊的平原河流，和远处难以分辨的房屋道路，可以作为实验重点观察的对象。在场景一和场景二中 SSR 算法处理后图像相比原图在道路，树木以及边缘细节等效果较佳，但存在背景噪声被过度增强的现象；MSRCR 算法在树木道路等细节处理方面相比本文算法和 SSR 算法效果较差；ESRGAN 算法处理结果相比其他算法更接近原图，并没有进一步增强细节特征；文献[3]算法采用多个子直方图均衡方式增强图像，对比上述算法避免了局部过增强的问题，但与本文算法处理结果相比，在细节处理方面给人直观感受还有待增强；本文算法对树木车辆等复杂目标有较强增强效果，背景中河流平原纹理特征明显。在场景三中，SSR 算法图像中的树木、道路、建筑等有增强效果，但场景四 SSR 算法处理之后的图像背景比较灰暗，出现明显噪声；MSRCR 算法视觉上较场景三和四的原图有些许增强，背景山脉的纹理特征较为清晰；ESRGAN 算法和文献[3]算法较其他算法更加接近原图，仅能看清

楚部分建筑房屋，没有凸显纹理细节；本文算法与其他算法相比对背景噪声抑制较好，前景目标和背景山脉显得更加清晰，树木建筑等细节特征更加明显，而且没有丢失图像本身特征。

客观数据分析方面，图像增强算法利用 PSNR 和 SSIM 两个数据作为指标衡量增强效果，增强后图像的 PSNR 和 SSIM 值越小则增强效果越明显。表 1 中列出了本文算法与其他五种算法在四种不同场景下的具体数据。由表可知，经本文算法增强后图像的 PSNR 和 SSIM 两个指标均小于 MRSCR、SSR、ESRGAN 和文献[6]算法。综上所述，在客观数据分析基础上结合主观评价表明，本文算法较 SSR 算法、MSRCR 算法、ESRGAN 算法和文献[6]算法的 PSNR 和 SSIM 值最小，主观评价最好，取得的红外图像增强效果最好。

表 1 四组实验对比结果

场景	指标	MR-SCR	SSR	ESR-GAN	文献[6]	本文算法
场景一	PSNR	31.12	20.65	32.54	22.68	17.69
	SSIM	0.91	0.79	0.94	0.82	0.56
场景二	PSNR	28.86	22.36	32.42	23.55	14.43
	SSIM	0.90	0.84	0.94	0.86	0.58
场景三	PSNR	21.80	22.23	33.39	26.79	16.21
	SSIM	0.97	0.82	0.94	0.87	0.53
场景四	PSNR	23.32	19.90	32.85	22.68	16.74
	SSIM	0.95	0.82	0.94	0.83	0.69

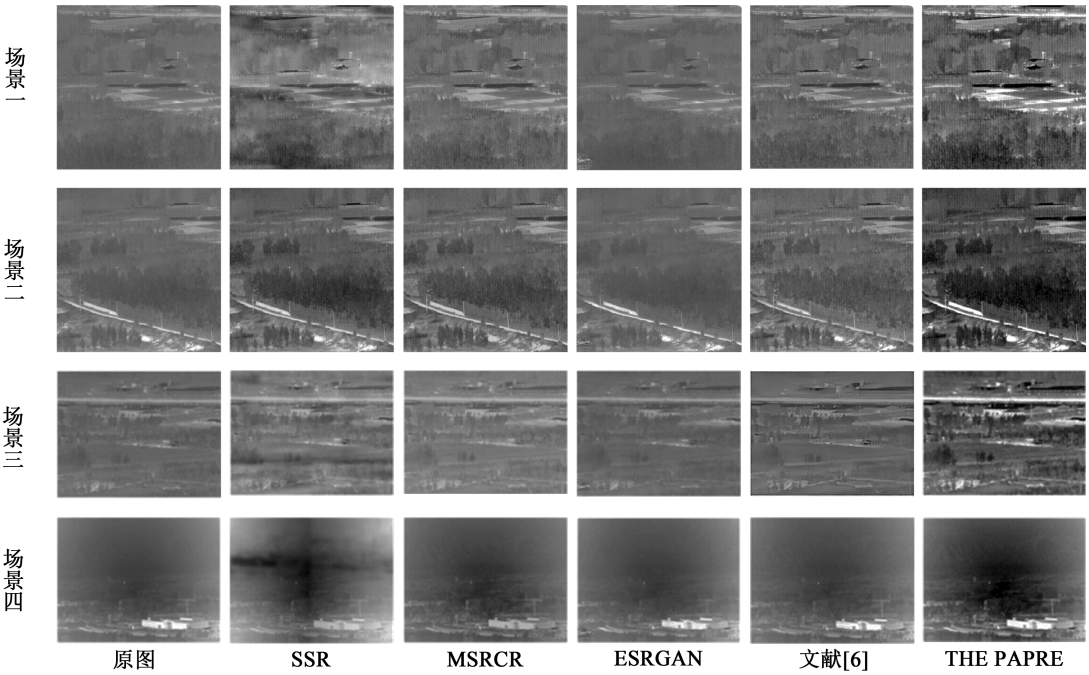

图 5 四组实验中各个算法图像增强效果对比

3 结束语

针对红外图像细节模糊、对比度低等问题，提出了一种基于改进生成对抗网络的红外图像增强算法。该方法基于 U-Net 和胶囊网络构建生成对抗网络；在生成器中将 U-Net 与全局上下文模块相结合，实现局部与全局特征的上下文建模，增强图像细节信息；在鉴别器中将胶囊网络与 Res2Net 结合，使神经网络在更细粒度的层次上提取多尺度图像特征，对胶囊网络的全连接层进行反卷积重构，减少模型参数数量。在不同场景下，将本文算法与主流算法进行对比实验，从视觉主观评价来看，本文算法能有效抑制背景噪声，增强红外图像的细节特征和图像对比度；由峰值信噪比和结构相似性的客观数据可知，本文算法增强后的红外图像质量更好，清晰度更高。

参考文献：

[1] Caroline G B, Martiello M S, Humberto C R, et al. Computer vision system and near-infrared spectroscopy for identification and classification of chicken with wooden breast, and physicochemical and technological characterization[J]. Infrared Phys. Technol., 2019, 96: 303-310.

[2] 吴 婷. 基于非均匀有序直方图和锐化的红外图像增强[J]. 半导体光电, 2020, 41(1): 135-140.
Wu Ting. Infrared image enhancement based on non-uniform ordered histogram and sharpening[J]. 2020, 41(1): 135-140.

[3] Zhang H, Qian W, Wan M, et al. Infrared image enhancement algorithm using local entropy mapping histogram adaptive segmentation[J]. Infrared Phys. Technol., 2022,

120：1-19.

[4] Yang M X，Tang G J，Liu X H，et al. Low-light image enhancement based on Retinex theory and dual-tree complex wavelet transform[J]. Optoelectronics Lett.，2018，14：470-475.

[5] Zhan B C，Wu Y Q. Infrared image enhancement based on wavelet transformation and retinex[C]// 2010 Second Inter. Conf. on Intelligent Human-Machine Systems and Cybernetics，2010：313-316.

[6] 朱庆丰，张　瑞，朱雯青，等. 基于小波变换与像元对目标的短波红外图像增强算法[J]. 半导体光电，2019，40(5)：688-693.

Zhu Qingfeng，Zhang Rui，Zhu Wenqing，et al. SWIR image enhancement based on wavelet transform and pairwise target contrast[J]. Semiconductor Optoelectronics，2019，40(5)：688-693.

[7] Zuo C，Chen Q，Liu N，et al. Display and detail enhancement for high-dynamic-range infrared images [J]. Optical Engineering，2011，50(12)：127401.

[8] 钟梦圆，姜　麟. 超分辨率图像重建算法综述[J]. 计算机科学与探索，2022，16(5)：972-990.

Zhong Mengyuan，Jiang Lin. Review of super-resolution image reconstruction algorithms [J]. J. Front. of Comput. Sci. Technol.，2022，16(5)：972-990.

[9] Hinton G E，Osindero S，Teh Y W. A fast learning algorithm for deep belief nets[J]. MIT Press，2006(7)：1527-1554.

[10] Lin L，Wang R，Wang W，et al. A low-light image enhancement method for both denoising and contrast enlarging[C]// IEEE Inter. Conf. on Image Processing，2015：3730-3734.

[11] Dong C，Loy C C，He K，et al. Image super-resolution using deep convolutional networks[J]. IEEE Trans. Pattern Anal. Mach. Intell.，2016，38(2)：295-307.

[12] Goodfellow I，Pouget-Abadie J，Mehdi M，et al. Generative adversarial networks [J]. Communications of the ACM，2020，63(11)：139-144.

[13] 陈佛计，朱　枫，吴清潇，等. 基于生成对抗网络的红外图像数据增强[J]. 计算机应用，2020，40(7)：2084-2088.

Chen Foji，Zhu Feng，Wu Qingxiao，et al. Infrared image data augmentation based on generative adversarial network[J]. J. Comput. Appl.，2020，40(7)：2084-2088.

[14] Ronneberger O，Fischer P，Brox T. U-Net：Convolutional networks for biomedical image segmentation[C]// Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015，2015：234-241.

[15] Sabour S，Frosst N，Hinton G E. Dynamic routing between capsules[C]// Proc. of the 31st Inter. Conf. on Neural Information Processing Systems，2017：4675-4684.

[16] Gao S H，Cheng M M，Zhao K，et al. Res2Net：A new multi-scale backbone architecture [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence，2019，43(2)：652-662.

[17] Cao Y，Xu J，Lin S，et al. GCNet：Non-local networks meet squeeze-excitation networks and beyond[C]// Inter. Conf. on Computer Vision Workshop (ICCVW)，2019：1971-1980.

[18] 回丙伟，宋志勇，范红旗，等. 地/空背景下红外图像弱小飞机目标检测跟踪数据集[DS/OL]. 中国科学数据(中英文网络版)，2020，5(3)：291-302.

Hui Bingwei，Song Zhiyong，Fan Hongqi，et al. A dataset for infrared image dim-small aircraft target detection and tracking under ground/air background[DS/OL]. 2020，5(3)：291-302.

[19] Wang H，Zhou L，Wang L. Miss Detection vs. False Alarm：Adversarial learning for small object segmentation in infrared images[C]// Inter. Conf. on Computer Vision，2019：8508-8517.

[20] 程铁栋，卢晓亮，易其文，等. 一种结合单尺度 Retinex 与引导滤波的红外图像增强方法[J]. 红外技术，2021，43(11)：1081-1088.

Cheng Tiedong，Lu Xiaoliang，Yi Qiwen，et al. Research on infrared image enhancement method combined with single-scale Retinex and guided image filter[J]. Infrared Technol.，2021，43(11)：1081-1088.

[21] Jobson D J，Rahman Z，Woodell G A. A multiscale Retinex for bridging the gap between color images and the human observation of scenes[J]. IEEE Trans. on Image Processing，2002，6(7)：965-976.

[22] Wang X T，Yu K，Wu S X，et al. ESRGAN：Enhanced super-resolution generative adversarial networks [C]// Computer Vision-ECCV 2018 Workshops，2018：63-79.

作者简介：

吴国瑞(1998—)，男，河南省人，硕士研究生，主要研究方向为图像增强、深度学习；

王　峰(1975—)，男，山西省人，教授，主要研究方向为图像处理、无人机等。