

引用格式: LI Haijun, KONG Fancheng, MU Junjie, et al. Infrared Ship Image Generation Algorithm Based on ISE-StyleGAN [J]. Acta Photonica Sinica, 2022, 51(12):1210004

李海军,孔繁程,牟俊杰,等.基于 ISE-StyleGAN 的红外舰船图像生成算法[J].光子学报,2022,51(12):1210004

基于 ISE-StyleGAN 的红外舰船图像生成算法

李海军¹,孔繁程¹,牟俊杰¹,刘霄²,杜贞斌²,林云³

(1 海军航空大学 岸防兵学院, 山东 烟台 264001)

(2 烟台大学 计算机学院, 山东 烟台 264005)

(3 烟台大学 教务处, 山东 烟台 264005)

摘要:针对军用背景下实地采集红外舰船数据难度大且图像质量差的问题,提出红外舰船增强 StyleGAN 的图像生成算法。该算法在 StyleGAN 基础上引入自注意力机制,加强生成器对图像纹理细节以及长距离像素关联信息的关注,对生成器最后模块的分辨率进行调整,并缩减噪声模块的输入量,利用小波判别器减少图像伪影的产生,采用双时间尺度更新规则和 Adam 优化训练过程,引入 WGAN-gp 损失函数提升收敛效率。通过目视判读以及客观评价指标验证得出,该算法生成的红外舰船图像质量较好,同真实红外图像相似度较高;经过四种经典目标检测算法测试,mAP 同原始数据集相比分别提升 15.4%、17.1%、11.9%、9.0%,同另外三种生成式对抗网络算法相比,取得更有效的扩充效果。

关键词:红外图像;生成式对抗网络;图像生成;自注意力机制;小波判别器

中图分类号:TP391.4

文献标识码:A

doi:10.3788/gzxb20225112.1210004

0 引言

红外成像制导相对于传统红外制导方式抗干扰能力强,动态范围更明显,是当前精确制导武器主要制导手段之一,精确制导的主要目标是精准发现,精确打击,因此红外制导的关键因素在于目标的检测。传统红外目标检测一般先对图像区域进行选择,利用尺度不变性特征变换(Scale Invariant Feature Transform, SIFT)和方向梯度直方图(Histogram of Oriented Gradient, HOG)等方法进行特征提取,通过一系列的逻辑运算,然后再使用分类器进行区分,这样的方法需手工设计,时间复杂度高,并且鲁棒性差。近年来,随着计算机视觉领域的不断发展,深度学习逐渐应用在目标检测领域中,根据实际领域需求,不断发展成单阶段和两阶段的目标检测算法。当前,将深度学习应用至红外目标检测领域越来越受到关注,红外目标检测大多通过可见光领域的算法实现,模型的训练对于数据的数量和质量的要求很高,而复杂环境下的完整红外舰船图像难以获取,所以对基于深度学习的红外舰船检测精度有所影响。近年来出现基于人工智能的深度生成模型,包括变分编码器(Variational Auto-Encoders, VAE)^[1]、深度信念网络(Deep Belief Network, DBN)^[2]、生成式对抗网络(Generative Adversarial Networks, GAN)^[3]等,其中 GAN 可以通过无监督学习实现对目标特征的提取,在生成模型中其具有较强的泛化能力,已发展出多种变体^[4-7],基于 GAN 的图像生成成为研究热点。针对使用 GAN 对小样本进行图像生成训练时易出现过拟合的问题,UTKARSH O 等^[8]提出基于锚点的策略,增强潜在空间对于图像中不同区域拥有不同的真实感,能够保持源图像和生成图像的差异性和相对相似性。LIU R 等^[9]通过引入同框架兼容增强潜在的对比损失,使得相邻潜在空间生成的图像相似度高,而不同潜在空间生成的图像差异性高,生成图像能够保证视觉质量且更加多样化。SAR 图像中存在相干斑噪声,导致其条件生成对抗网络模型(Conditional Generative Adversarial Nets, CGAN)出现

基金项目:军队研究生资助课题(No.JY2020B121),山东省自然科学基金(Nos.ZR2021MF133,ZR2017MF00.19)

第一作者:李海军(1965—),男,教授,博士,主要研究方向为航空导弹作战与运用。Email:19804473@qq.com

通讯作者:孔繁程(1998—),男,硕士研究生,主要研究方向为红外图像处理与目标检测。Email:879445564@qq.com

收稿日期:2022-05-18;录用日期:2022-07-06

<http://www.photon.ac.cn>

崩溃, GUO J Y等^[10]通过归一化相干斑噪声解决了模型崩溃的问题。

目前,利用GAN进行数据生成的研究对象仍是可见光图像为主^[11],红外图像生成领域的研究方法集中于可见光图像转红外图像,通过红外图像进行生成处理的研究较少,对红外背景下舰船数据生成的研究更是少之又少。由于实地采集红外舰船图像难度较大、费效比高、数据量较少,为能更好地提升基于深度学习的红外目标检测精度和效果,本文提出基于样式的生成对抗网络(Style Generative Adversarial Nets, StyleGAN)的红外舰船图像生成算法,通过训练生成式对抗网络模型,得到高质量红外舰船图像,能够提供红外舰船数据。根据红外舰船图像特点,采用StyleGAN作为基础架构,在生成器引入自注意力机制,增强舰船图像细节,精简噪声模块,以免图像风格混乱;同时,调整模块分辨率大小,使模型更适合本文图像;利用小波判别器通过小波分解提取图像特征,减少图像伪影的产生,使用双时间尺度更新规则和Adam优化器优化。训练过程,同时再引入WGAN-gp损失函数提升收敛效率。

1 算法原理与网络结构设计

1.1 生成式对抗网络

生成式对抗网络GAN本质是对抗双方生成器和判别器通过零和博弈不断学习、不断改进,最终生成能够以假乱真的本不存在的数据。生成器主要进行样本伪造,训练过程中要不断提高伪造水平,尽可能与真实样本相同;判别器用来判断生成器生成样本的真实性,通过计算生成样本与真实样本的差值去引导生成器不断减小差异。GAN基本结构如图1所示。

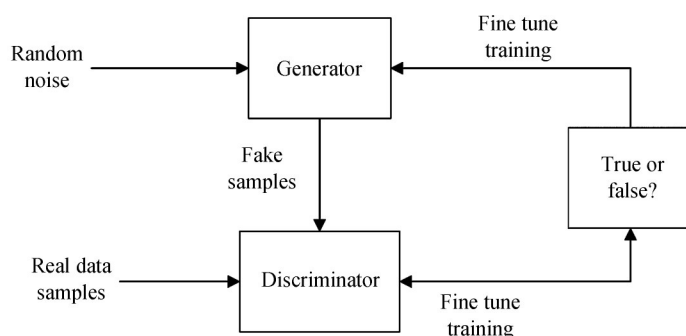


图1 生成式对抗网络结构

Fig. 1 Generative adversarial network structure diagram

原始GAN中目标优化函数为

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

式中, D 表示判别器, G 表示生成器, $x \sim p_{\text{data}}$ 表示 x 来源于真实分布, $z \sim p_z$ 表示随机噪声 z 来源于模拟分布, $E_{x \sim p_{\text{data}}} [\ln D(x)]$ 表示判别器损失, $E_{z \sim p_z} \ln \{1 - D[G(z)]\}$ 表示生成器损失。

训练结束时,对抗双方达到纳什平衡,判断器无法识别生成图像的真伪。原始GAN训练过程中生成器从随机噪声中取样,存在生成样本图像状态不可控的问题,近年来不断改进,其中KARRAS T等提出基于样式的生成式对抗网络(StyleGAN)^[12],StyleGAN移除传统随机噪声,将潜向量(latent code)作为初始输入,采用无监督学习对图像中的语义进行解耦,并采用样式传递方式解决生成图像不可控的问题。

1.2 网络结构设计

StyleGAN模型在高分辨率图像生成以及样式风格变换上有良好的性能,网络结构如图2所示,因此,以StyleGAN为基础结构,针对数据集的特点,对其生成器结构进行改进。StyleGAN能够在可见光自然图像的纹理细节方面有很好的表现,对于红外舰船图像来说,其图案形状、颜色丰富程度以及样式种类远不及可见光自然图像,直接将结构复杂的StyleGAN应用在红外图像生成上难以得到理想的生成图像。因此,在原始StyleGAN基础上对模型进行修改,提出红外舰船增强StyleGAN(Infrared Ship Enhancement StyleGAN, ISE-StyleGAN)模型,以显著提高模型训练的稳定性并获取高质量生成样本。

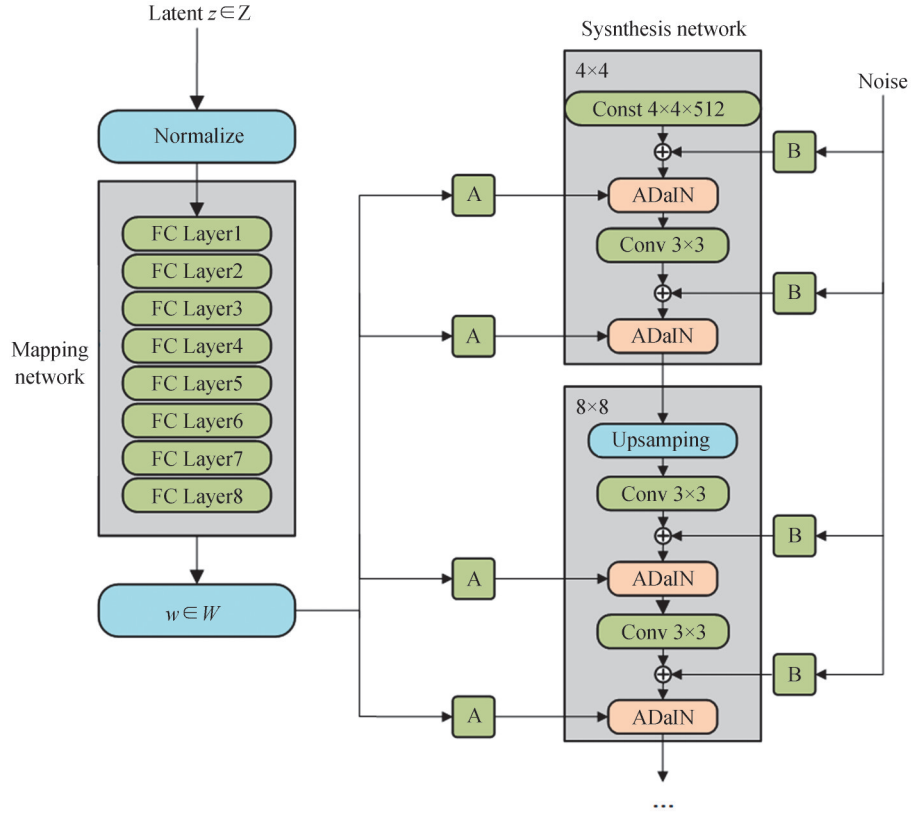


图2 StyleGAN生成器结构
Fig. 2 StyleGAN generator structure

1.2.1 基于自注意力机制的生成器

1)StyleGAN中的感受野大小受限于卷积核,在训练过程中难以从全局角度捕捉红外舰船图像内容,导致图像部分内部信息和长距离像素关联信息丢失,但是如果使用全连接层获取全局信息,会使模型冗余增加,计算量激增。所以,考虑通过引入自注意力机制解决此问题。自注意力生成对抗网络(Self-Attention Generative Adversarial Networks, SAGAN)^[13]首次将自注意力机制引入生成式对抗网络中,生成模型的效果得到极大提升。本文将自注意力引入生成器中,使算法能够在全局领域进行操作,学习到图像中更多的细节和长距离像素关联信息。

如图3所示,自注意力机制将特征映射为三部分,第一部分 $f(x)$ 和第三部分 $h(x)$ 使用 1×1 卷积模块进行 $f(x) = W$ 线性变换,第二部分 $g(x)$ 进行通道压缩,即

$$\begin{cases} f(x) = W_f x \\ g(x) = W_g x \\ h(x) = W_h x \\ v(x) = W_v x \end{cases} \quad (2)$$

式中, W_f, W_g, W_h, W_v 表示四个 1×1 卷积层权重矩阵。

然后将 $f(x), g(x)$ 转换为矩阵形式,进行转置相乘,经过softmax处理得到注意力权重 $\beta_{i,j}$ 为

$$\beta_{j,i} = \frac{\exp(f(x_i)^T g(x_j))}{\sum_{i=1}^N \exp(f(x_i)^T g(x_j))} \quad (3)$$

式中, $\beta_{j,i}$ 表示第 i 个位置对生成第 j 个位置的关系权重, N 表示输入特征图中含有特征的位置数量。

再将 $h(x)$ 与得到的注意力权重 $\beta_{j,i}$ 相乘,再经过 1×1 卷积层处理得到自注意力特征映射 o_j 为

$$o_j = v \left[\sum_{i=1}^N \beta_{j,i} h(x_i) \right] \quad (4)$$

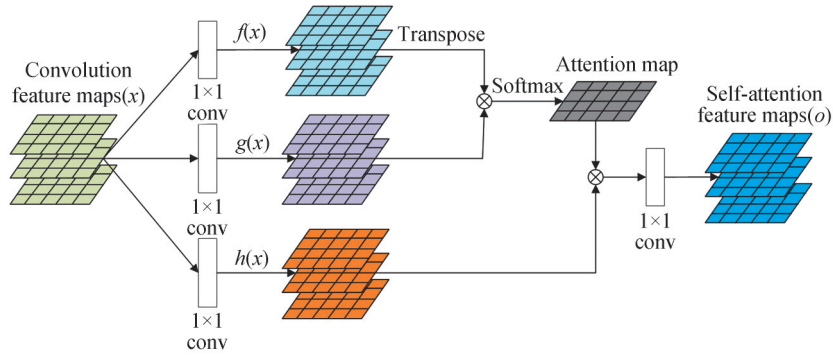


图3 自注意力机制
Fig.3 Self-attention

最后将自注意力特征映射与一比例参数相乘,再同输入特征图加权相加得到最终输出 z_i 为

$$z_i = \gamma o_i + x_i \quad (5)$$

式中, γ 是初值为 0 的预定义系数。

2) 由于红外舰船图像具有色彩单一、样式结构朴素的特点,所以对生成器进行针对性简化,将生成器最后一个模块的分辨率大小设定为 256×256 ,使生成器更适合本文提供的的数据需求,在确保生成图像质量的前提下,能够减少网络所需参数量,提高生成器的运算效率。

3) 减少随机噪声的输入数量。由于红外舰船图像的纹理细节特征没有可见光图像丰富,按照原始 StyleGAN 模型进行图像生成时,噪声过多会引入较多噪点,影响自适应实例归一化^[14]的进行,进而导致生成图像质量下降。因此,生成器每个不同分辨率网络模块的噪声输入均只引入一个噪声模块。最终改进生成器结构如图 4 所示。噪声模块 N 的生成方法表示为

$$N(\omega, x) = \text{randn}(H, W) \times \omega + x, \omega \in R^{\text{channel}}, x \in R^{\text{channel} \times H \times W} \quad (6)$$

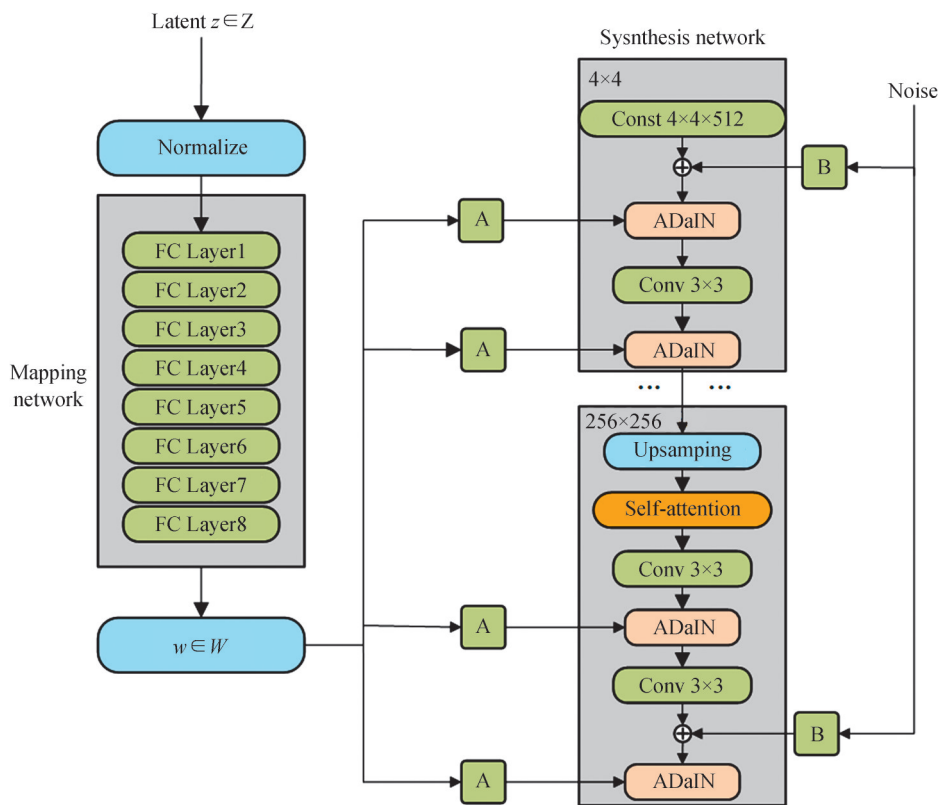


图4 改进的生成器结构
Fig.4 Improved generator architecture

1.2.2 小波判别器

使用小波判别器^[15](Wavelet discriminator)进行伪影抑制。由于实验过程中发现,在 64×64 分辨率的早期训练阶段,可以观察到阻塞伪影,但随着训练的进行,它们逐渐消失。由于判别器判别失败,伪影会使高分辨率合成缺失高频细节,因此需要使用更强的鉴别器来抑制伪影的产生。

小波判别器结构如图5所示,在每一个分辨率模块中,使用一个基于跳跃连接的网络从小波分解中提取特征,并将它们合并到从更高分辨率块派生的特征表示中。判别器对输入图像分层,进行双线性下采样处理,并在每个尺度上检测离散小波分解位置与真实图像的频率差异。这样的小波鉴别器在对抗阻塞伪影方面效果显著。同时,它对分布匹配没有任何负面影响,可有效引导生成器生成丰富的图像细节。

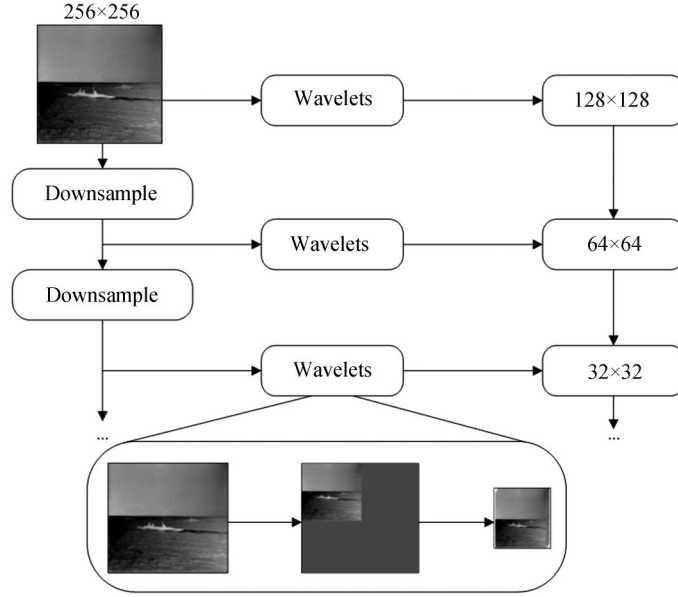


图5 小波判别器结构

Fig. 5 Structure of wavelet discriminator

1.2.3 损失函数

使用WGAN-gp^[16]损失函数提升收敛效率,以此将生成图像 z 同真实样本 x 的Wasserstein-1距离输入至判别器,并基于WGAN的损失标准,用梯度惩罚项替换权重裁剪项,避免梯度爆炸。最终损失函数为

$$L = E_{z \sim P_z} [D(z)] - E_{x \sim P_x} [D(x)] + \lambda E_{\hat{x} \sim P_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2] \quad (7)$$

式中, $E_{z \sim P_z} [D(z)] - E_{x \sim P_x} [D(x)]$ 为WGAN损失, $\lambda E_{\hat{x} \sim P_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2]$ 为梯度惩罚项。

1.2.4 双时间尺度更新规则

训练时采用双时间尺度更新规则(TTUR)^[17]。TTUR提出新的训练参数更新策略,将生成器 G 和判别器 D 自动设置不同的学习率,让 D 的收敛速度加快,并且能够平衡两者的训练速度。训练过程是根据生成器和判别器损失函数 L_G, L_D 的随机梯度 $\tilde{h}(\theta, \omega)$ 和 $\tilde{g}(\theta, \omega)$ 来进行的,其中 θ 是生成器中的学习变量参数, ω 是判别器的学习变量参数。由于梯度 $\tilde{h}(\theta, \omega)$ 和 $\tilde{g}(\theta, \omega)$ 是随机的,从图像数据集 $P_{\text{data}}(x)$ 中随机选择 m 个真实样本 $x^{(i)}, 1 \leq i \leq m$,和生成图像数据集 $P_G(z)$ 中 m 个随机选择的生成样本 $z^{(i)}, 1 \leq i \leq m$ 。如果实际梯度为 $h(\theta, \omega) = \nabla_{\theta} L_G$ 和 $g(\theta, \omega) = \nabla_{\omega} L_D$,那么就用随机变量 $M^{(\omega)}$ 和 $M^{(\theta)}$ 定义 $\tilde{h}(\theta, \omega) = h(\theta, \omega) + M^{(\theta)}$ 和 $\tilde{g}(\theta, \omega) = g(\theta, \omega) + M^{(\omega)}$ 。因此梯度 $\tilde{h}(\theta, \omega)$ 和 $\tilde{g}(\theta, \omega)$ 随机近似于实际梯度。TTUR一般假设生成器和判别器更新的学习率为 $a(n)$ 和 $b(n)$,则

$$\begin{cases} \theta_{n+1} = \theta_n + a(n) \cdot (h(\theta_n, \omega_n) + M_n^{(\theta)}) \\ \omega_{n+1} = \omega_n + b(n) \cdot (g(\theta_n, \omega_n) + M_n^{(\omega)}) \end{cases} \quad (8)$$

2 实验结果及分析

2.1 实验数据集

实验数据的获取采用网络爬虫方式,主要是通过网络检索“舰船”、“货轮”等军用民用船只,共4 284张红外舰船图像,其中舰船数量为7 196艘。首先按照8:2的比例将图片划分成训练集和测试集,然后从测试集中随机抽取一半作为验证集,实验前不对红外图像进行任何处理。

2.2 实验配置

本文所有实验均在Ubuntu20.04操作系统中进行,CPU是core i9-12900KF,GPU为显存12G的NVIDIA GeForce RTX 3080ti,python版本为3.7.11,开发工具是pycharm2020.1,框架使用Pytorch1.9.0。实验前首先将自建红外舰船数据集进行批处理,训练参数设置如表1所示。学习率使用TTUR,批训练数设置为32,单次实验训练900个Epochs,使用Adam优化器。

表1 实验训练参数
Table 1 Experimental training parameters

Parameter	Value
Learning rate	TTUR
Optimizer	Adam
Base size	64
Epoch	900

2.3 算法训练过程

ISE-StyleGAN算法实现过程为:

输入:红外舰船数据图像,训练迭代次数 n ,训练步骤数 m

for $t=1$ to n do

for $k=1$ to m do

从数据集分布 $P_{\text{data}}(x)$ 中随机选择 m 个真实样本,记为 $\{x^{(1)}, x^{(2)}, \dots, x^{(m)}\}$

从随机输入 $P_G(z)$ 中随机选择 m 个生成样本,记为 $\{z^{(1)}, z^{(2)}, \dots, z^{(m)}\}$

随机噪声记为: $\hat{x}^{(i)} = \epsilon x^{(i)} + (1 - \epsilon) z^{(i)}$, $\epsilon \in [0, 1]$, $1 \leq i \leq m$

使用梯度上升更新判别器的代价函数

$$\nabla_{\theta_D} \frac{1}{m} \sum_{i=1}^m [D(G(z^{(i)})) - D(x^{(i)}) + \lambda (\|\nabla_x D(\hat{x}^{(i)})\|_2 - 1)^2] \quad (9)$$

end for

再从随机输入 $P_G(z)$ 中采样 m 个生成样本

使用梯度下降更新生成器的代价函数

$$\nabla_{\theta_G} \frac{1}{m} \sum_{i=1}^m -D(G(z^{(i)})) \quad (10)$$

end for

2.4 图像生成效果评价

2.4.1 目视判断评价

为全面展示生成的红外图像,图6中选取了某一图像在各个迭代生成阶段的生成效果。可见在Epoch=100时,图像像素化严重并存在失真现象,无法辨别出舰船图像;当Epoch=300时,图像网格基本消失,但在图像中央已能大致看出舰船轮廓;当Epoch=600时,舰船图像开始显现,轮廓已接近真实清晰,但图像背景略微模糊;当Epoch=900时,整体生成的红外图像非常接近真实图像,舰船目标已能够通过目视判读出。

由图6可以看到,在Epoch从100到900的训练过程中,能够得到不同分辨率的红外舰船图像。因此,可以通过利用低分辨率图像得到舰船的整体轮廓以及部分舰船特征,以此来对舰船目标进行检测识别,并持续跟踪目标;利用高分辨率图像可识别出舰船中的细节特征,因此,可以更具针对性地进行舰船目标的细粒度识别,比如,可对舰船中的驾驶室、雷达、水线等要害部位进行识别定位,利用反舰导弹进行精确打击。

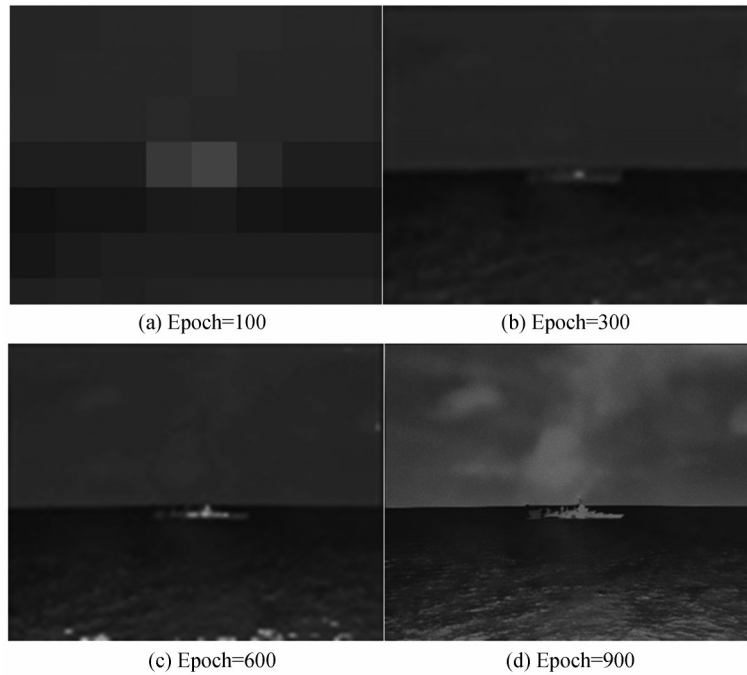
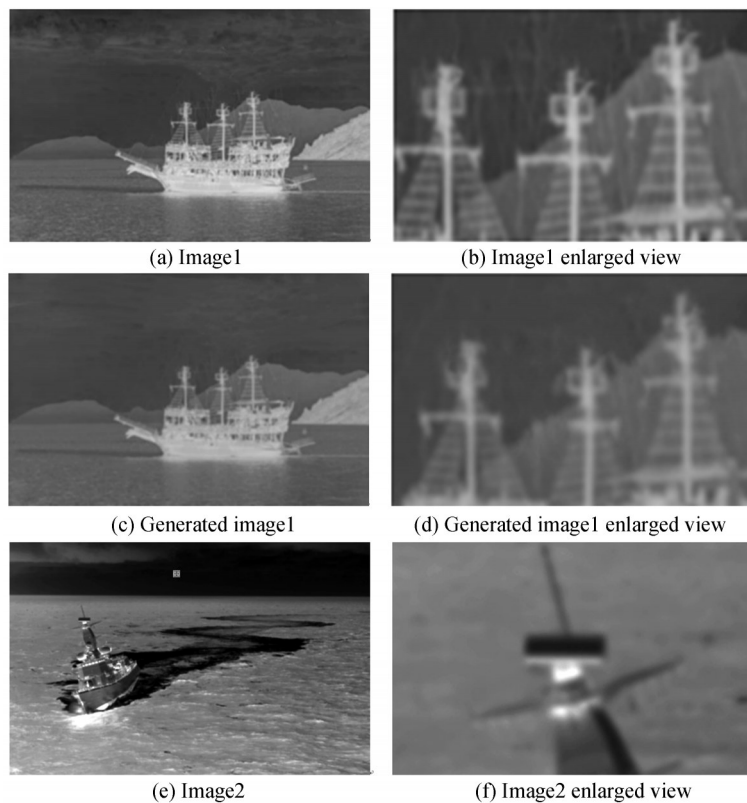


图6 ISE-StyleGAN训练过程
Fig. 6 The training process of ISE-StyleGAN

为进一步验证算法的生成图像质量,对图像中显著位置天线桅杆进行放大对比。图7展示了两幅图的原图、舰船天线桅杆放大图以及ISE-StyleGAN生成图像。从图7(b)、(d)中可以看到虽然生成图像的天线桅杆略有模糊,并且中间的桅杆和右边的桅杆中部横杆一小部分不相同,但是图像整体形状、桅杆位置分布以及其他细节部分基本一致,并且整体的亮度以及对比度比较符合人眼要求。相比图7(f),图7(h)也略微模糊,整体亮度偏灰暗,天线桅杆形状大致相同,能够看出轮廓边缘,较好地地区分出天线桅杆和海面背景。综上,通过将原图和生成图像中显著位置进行对比看出,尽管在整体亮度以及模糊程度上略有差别,但是生



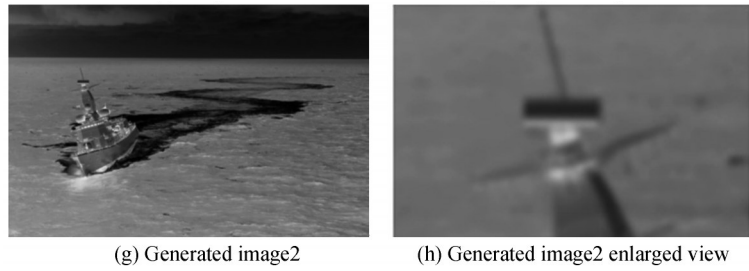


图7 原图与生成图像的对比
Fig. 7 Compare of the original images with the generated images

成图像能够较好地生成图像局部特征,基本还原出原图的效果。

基于相同实验条件、不同场景进行实验,随机选择不同背景的生成红外图像如图8所示,对比原数据、DCGAN、CycleGAN、StyleGAN和ISE-StyleGAN的实验结果。DCGAN生成图像背景部分同原图像相差较多,并且该算法生成的第一张图像中目标尺度大小存在问题,第二张图像中存在空间扭曲,说明仍然存在训练不稳定、模型坍塌的现象;CycleGAN的生成图像对于曝光度较低的背景部分具有较好的生成效果,但是整体的灰度分布、灰度等级存在着较大差异并且有颜色偏移的现象。StyleGAN也能生成表面光滑、质量较好的自然图像,但是部分图中夹杂着一些明显与原始图像特征不符的缺陷特征,存在风格混乱的问题,该算法生成的第一张图中舰船位置颠倒,第四张图中海面出现黑色条纹,图像缺乏真实性。

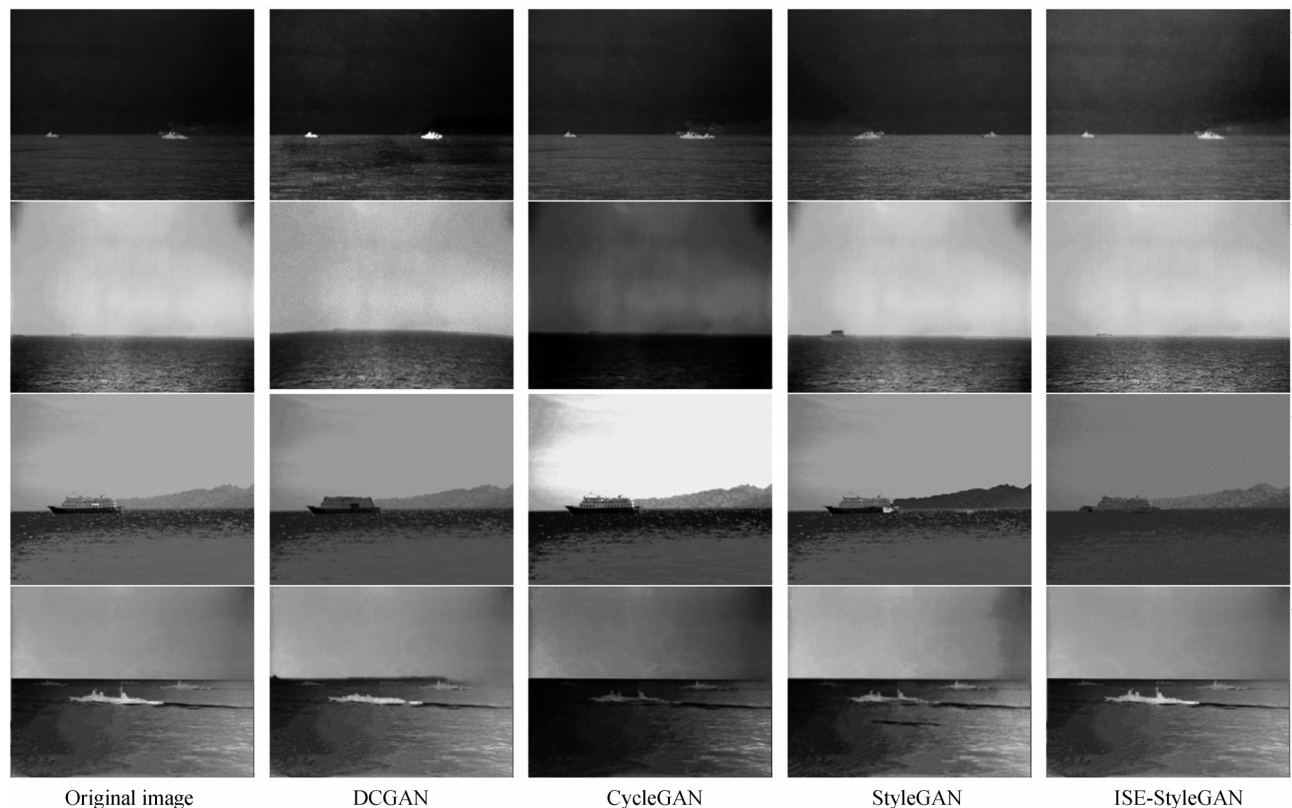


图8 不同生成式对抗网络算法生成图像对比
Fig. 8 Comparison of images generated by different generative adversarial network algorithms

根据红外图像的先验知识以及目视判读得到:1)本文算法在白天、黑夜、云雾等各种气象条件下都能得到较好的舰船和背景的红外成像效果;2)本文算法在舰船目标不同尺度不同数量的条件下也能有较好的红外成像效果;3)图中展示的生成图像满足舰船的大多数特征要求,轮廓清晰可辨,图像纹理细节较明显,整体的灰度分布同真实图像的红外效果非常接近,仅仅在整体亮度以及色差上略有差别。

通过目视图像中特征显著位置对比以及主观评价生成图像可得:使用本文算法所得红外图像能基本展

示出舰船的轮廓以及纹理细节,在灰度分布上与真实图像大致相同,整体上两幅图片的相似度很高,通过目视以及主观判断,生成的红外舰船图像表现较好。

2.4.2 客观指标评价

为进一步验证实验生成图像的有效性,通过客观性能指标峰值信噪比(Peak Signal to Noise Ratio, PSNR)和多层级结构相似性(Multi-Scale Structural Similarity Index, MS-SSIM)^[18]对生成图像进行定量分析。PSNR指标基于原始图像和生成图像中对应像素点的误差大小,主要是衡量生成图像失真程度,分值越高,图像越逼真。由于数据集中图像尺度丰富多样,所以另一个评估指标采用MS-SSIM。同SSIM程序相同,MS-SSIM也是进行原始与生成图像的分辨率、亮度、对比度以及相关结构的比较,得到比较分数然后进行积累,对图像进行下采样,整个过程重复迭代,以此来进行不同图像尺度的多SSIM评价。

表2和表3给出了各个模型对各种背景、尺度的红外舰船图像生成客观指标评价结果,可以看到ISE-StyleGAN的PSNR值以及MS-SSIM值在各种类型的目标下都是最高,说明本文提出的改进算法相比几种经典生成式对抗网络方法生成的红外舰船图像更加逼真,质量更好。同时,ISE-StyleGAN生成的舰船轮廓和细节同原始图像更相似,因此,可以推断出ISE-StyleGAN生成的舰船图像特征同原始图像特征更相似。

表2 生成式对抗网络模型PSNR指标结果

Table 2 PSNR index results of generative adversarial network model

Algorithm	PSNR/dB				
	Single scale	Multiscale	Small target	Cloud and mist interference	Night target
DCGAN	13.254	7.363	6.784	10.436	9.473
CycleGAN	9.792	8.187	6.349	8.342	9.754
StyleGAN	8.835	7.829	4.382	4.993	8.438
ISE-StyleGAN	21.853	19.341	13.740	11.284	15.584

表3 生成式对抗网络模型MS-SSIM指标结果

Table 3 MS-SSIM index results of generative adversarial network model

Algorithm	MS-SSIM/ %				
	Single scale	Multiscale	Small target	Cloud and mist interference	Night target
DCGAN	91.6	85.7	83.4	90.2	93.6
CycleGAN	83.8	88.1	84.7	79.0	89.2
StyleGAN	82.3	87.2	81.9	73.3	86.4
ISE-StyleGAN	97.9	94.5	89.3	92.7	98.5

2.4.3 目标检测算法测试

生成红外舰船图像旨在解决由于红外数据量不足影响目标检测精度的问题。所以,可以通过将生成数据集应用至舰船检测任务来进一步验证生成图像的有效性。验证过程采用不同的数据集,包括原始红外舰船数据集,原始数据集同常规增广的数据组合数据集,DCGAN、CycleGAN、StyleGAN、ISE-StyleGAN生成图像组合数据集分别进行舰船检测训练,数据集具体设计如表4所示。常规增广包括图像平移变换、随机

表4 目标检测测试数据集构成

Table 4 Data set composition of object detection test

Number	Algorithm	Source data	Conventionally generated data	Algorithmic data generation	Total
1	None	1 000	0	0	1 000
2	Conventional expansion	1 000	2 000	0	3 000
3	DCGAN	1 000	1 000	1 000	3 000
4	CycleGAN	1 000	1 000	1 000	3 000
5	StyleGAN	1 000	1 000	1 000	3 000
6	ISE-StyleGAN	1 000	1 000	1 000	3 000

缩放、随机裁剪等方法,检测算法采用YOLOv3、SSD、Faster R-CNN、Centernet, epoch=50,使用平均精确率(mAP)作为评价指标。

图9给出了四种算法训练在不同数据集中的mAP变化曲线,对比可知,使用原始数据集进行目标检测,精度结果都不是很理想,并且不同检测算法使用这1000张图片得到的结果相差较大,说明不同检测算法对数据集依赖程度不同。对原始数据集进行扩充后无论是使用常规增广还是通过算法数据生成,目标检测精度均能得到一定程度的提升,同时,不同生成式对抗网络算法得到的扩充数据集在相同目标检测模型中运行,可比较出各个模型的性能差异。使用常规增广方法在不同检测算法上有8%左右的精度提升,其他主流生成算法能将精度提升10%左右,但是,存在使用生成算法的效果不如常规增广的情况。对比表5中序号2和序号5的YOLOv3算法检测精度,常规增广的检测效果提升更明显,对比序号2和序号4使用SSD算法,也是常规增广效果更好。因此,虽然主流生成算法能够丰富红外舰船特征多样性,但是使用该数据集生成的图像质量不高,存在风格混乱以及缺陷特征的问题,会造成海杂波的虚检。从表5可以看出本文算法生成

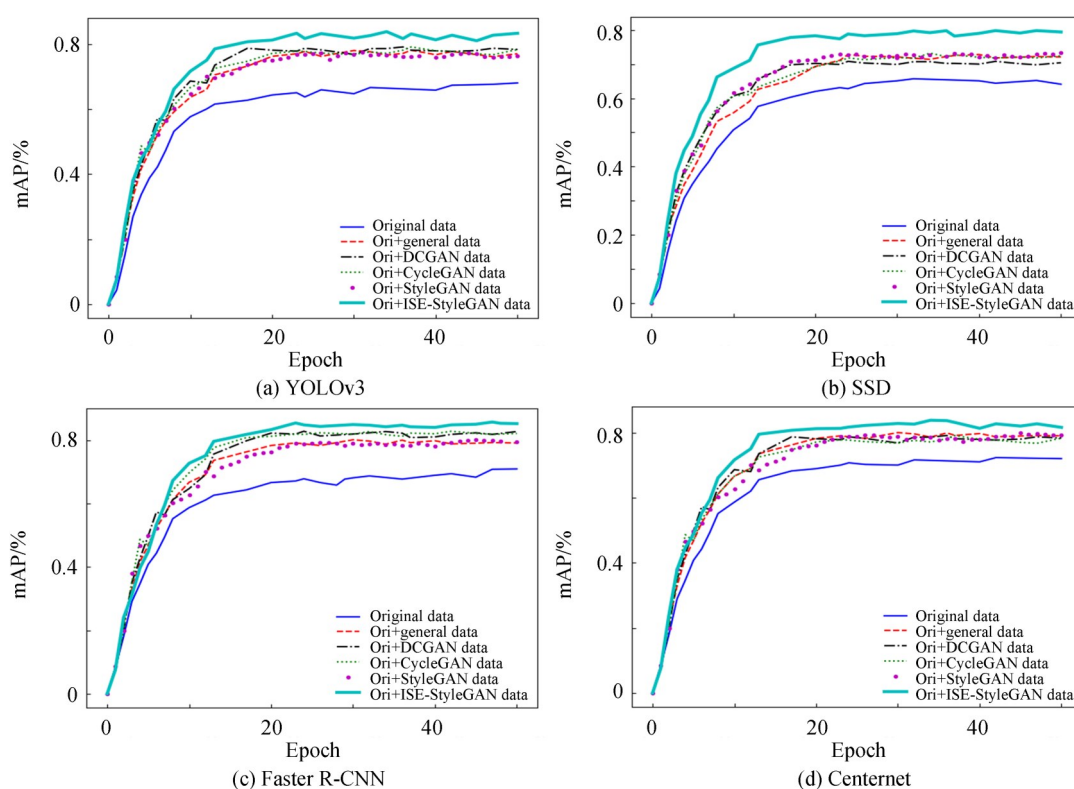


图9 各数据集用不同算法得到的mAP值

Fig. 9 mAP value of each dataset by different algorithms

表5 各数据集目标检测的mAP值

Table 5 mAP value of target detection in each dataset

Number	Algorithm	YOLOv3	SSD	Faster R-CNN	Centernet
1	None	0.679	0.641	0.703	0.727
2	Conventional expansion	0.771	0.733	0.788	0.791
3	DCGAN	0.784	0.712	0.804	0.786
4	CycleGAN	0.788	0.745	0.801	0.787
5	StyleGAN	0.762	0.735	0.797	0.790
6	ISE-StyleGAN	0.833	0.812	0.822	0.817

的数据使用四种目标检测算法测试都能有较好的检测效果,鲁棒性较强。例如在YOLOv3算法中,ISE-StyleGAN扩充后的目标检测网络与原始数据集相比在平均精确率上提升15%左右,验证了基于ISE-

StyleGAN生成红外舰船图像的有效性和可行性。

3 结论

本文针对红外舰船数据难以获取的问题,提出ISE-StyleGAN红外舰船图像生成算法,通过目视判读以及客观评价指标对原始图像以及DCGAN、CycleGAN、StyleGAN、ISE-StyleGAN生成图像进行对比实验。结果表明,本文算法生成的红外舰船图像质量较好,同真实红外图像相似度较高;再经过4种经典目标检测算法测试,本文算法生成图像构建的数据集检测效果均最优。因此,该方法可用于红外图像生成,为红外舰船数据提供来源。

参考文献

- [1] KINGMA D P, WELLING M. Auto-encoding variational Bayes[J]. arXiv preprint arXiv: 1312.6114.
- [2] HINTON G, SALAKHUTDINOV R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786):504-507.
- [3] GOODFELLOW I, POUGET J, MIRZA M, et al. Generative adversarial nets[C]. *Proceedings of the 27th International Conference on Neural Information Processing Systems*, 2014: 2672-2680.
- [4] DONAHUE J, KRAHENBUHL P, DARRELL T. Adversarial feature learning[C]. *International Conference on Learning Representations(ICLR)*, 2017.
- [5] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. *Computer Science*, 2015.
- [6] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C]. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, IEEE, 2017: 2242-2251.
- [7] CHEN X, DUAN Y, HOUTHOOFT R, et al. InfoGAN: interpretable representation learning by information maximizing generative adversarial nets[C]. *30th International Conference on Neural Information Processing Systems (NIPS)*, IEEE, 2016.
- [8] UTKARSH O, LI Y J, LU J W, et al. Few-shot image generation via cross-domain correspondence[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR)*, 2021: 10743-10752.
- [9] LIU R, GE Y X, CHING L C, et al. DivCo: diverse conditional image synthesis via contrastive generative adversarial network[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021: 16377-16386.
- [10] GUO J Y, LEI B, DING C B, et al. Synthetic aperture radar image synthesis by using generative adversarial nets[J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(7): 1111-1115.
- [11] CUBUK E D, ZOPH B, MANE D, et al. Auto augment: learning augmentation strategies from data[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019: 113-123.
- [12] KARRAS T, LAINE S, AILA T. A style-based generator architecture for generative adversarial networks [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019: 4401-4410.
- [13] ZHANG H, GOODFELLOW I, METAXAS D, et al. Self-attention generative adversarial networks[C]. *Proceedings of the 32st International Conference on Neural Information Processing Systems (NIPS)*, 2018.
- [14] ZHAO C, SHUAI R J, MA L, et al. Dermoscopy image classification based on StyleGAN and DenseNet201[J]. *IEEE Access*, 2021:8659-8679.
- [15] RINON G, DANA C, BAMIT et al. SWAGAN: a style-based WAvelet-driven generative model[J]. arXiv preprint arXiv: 2102.06108.
- [16] ISHAAN G, FARUK A, MARTIN A, et al. Improved training of wasserstein GANs [J]. arXiv preprint arXiv: 1704.00028.
- [17] MARTIN H, HUBERT R, THOMAS U, et al. GANs trained by a two time-scale update rule converge to a local nash equilibrium[C]. *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS)*, 2017:6629-6640.
- [18] WANG Z, BOVIK A C, SHEIKH R H, et al. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.

Infrared Ship Image Generation Algorithm Based on ISE-StyleGAN

LI Haijun¹, KONG Fancheng¹, MU Junjie¹, LIU Xiao², DU Zhenbin², LIN Yun³

(1 Coastal Defense College, Naval Aviation University, Yantai, Shandong 264001, China)

(2 School of Computing, Yantai University, Yantai, Shandong 264005, China)

(3 Office of Academic Affairs, Yantai University, Yantai, Shandong 264005, China)

Abstract: Infrared imaging guidance has a stronger anti-interference ability and a more obvious dynamic range than traditional infrared guidance, which is one of the main guidance means of current precision guidance weapons. In recent years, with the continuous development of computer vision, the application of deep learning to military infrared target detection has attracted more and more attention. Model training has high requirements on the quantity and quality of data. However, it is difficult to obtain complete infrared ship images in the complex environment in the military field, which affects the detection accuracy of infrared ship based on deep learning. Most infrared target detection is achieved through the algorithms of the visible light field. Therefore, the research of GAN-based image generation by GAN data is still mainly visible image. The research on infrared image generation is scarce, and the research on ship data generation under infrared background is even less. In response to these problems, high cost-efficiency ratio, and a small amount of data in field acquisition of Infrared Ship images, this paper proposes ISE-StyleGAN (Infrared Ship Enhancement StyleGAN) algorithm for Infrared Ship image generation. By training the generative adversarial network model, high quality infrared ship image is obtained, which can provide infrared ship data. In this paper, improvements are made based on the StyleGAN model. Firstly, because of the size of receptive field in StyleGAN is limited by the convolution kernel. In this paper, self-attention is introduced into the generator, so that the algorithm can operate in the global domain can learn more details in the image and long-distance pixel association information. Setting the resolution of the last module of the generator to 256×256 can make the generator more suitable for the data requirements provided in this study. On the premise of ensuring the quality of the generated image, the number of parameters required by the network and the amount of random noise input can be reduced, and the computing efficiency of the generator can be improved. As the texture details of infrared ship images are not as rich as those of visible images, too much noise will introduce more noise points during image generation according to the original StyleGAN model, which will affect the normalization of adaptive instances, thus resulting in the degradation of the generated image quality. Therefore, this study only introduces one noise module into the noise input of each network module of different resolutions of the generator. Secondly, a Wavelet discriminator is used to extract image features through Wavelet decomposition and combine them into feature representations derived from higher resolution blocks. In the representation of the characteristics of the image, the discriminator stratifies the input image to perform a bilinear downsampling scale reduction, degree processing, and detection of separation at each scale. Then, by scattering wavelet, the frequency difference between the generated image and the real image is obtained. Such a Wavelet discriminator is very effective against blocking artifacts. Then, TTUR and Adam are used for optimization. In the training process, TTUR can make the generator and discriminator automatically set different learning rates so that the discriminator convergence speed is accelerated and the training speed of the two can be balanced. Finally, WGAN-gp loss function is introduced to improve convergence efficiency. The experimental results of the original data, DCGAN, CycleGAN, StyleGAN and ISE-StyleGAN were compared by visual interpretation in this paper. The infrared images obtained by the algorithm can basically show the ship contour and texture details, and the gray distribution is relatively uniform. Compared to the real image, the overall similarity of the two pictures is high. From the objective indicators, PSNR value and MS-SSIM value are the highest in all types of targets. It shows that the improved algorithm proposed in this paper has better quality and image phase than several classical generative adversarial network methods in generating infrared ship images. At the same time, the outline and details of the ship generated by ISE-StyleGAN are more prominent. Therefore, it can be inferred that the ship image features generated by ISE-StyleGAN are more similar to the original image features. Finally, the validity of the generated image is further verified by applying the generated data set to the ship detection task. Different datasets are used in the verification process, including the original infrared ship dataset, the original dataset and the

conventional augmented data combination dataset. The combined dataset of images generated by DCGAN, CycleGAN, StyleGAN and ISE-StyleGAN were used for ship detection training respectively. Then, the detection algorithm adopted Faster R-CNN, SSD, YOLOv3 and Centernet. Compared with the original dataset, the average accuracy of the expanded ISE-StyleGAN target detection network is about 15% higher than that of the original dataset, which verifies the effectiveness and feasibility of generating infrared ship images based on ISE-StyleGAN.

Key words: Infrared image; Generative adversarial network; Image generation; Self-attention; Wavelet discriminator

OCIS Codes: 100.3008; 010.7295; 110.3000; 110.3008; 150.1135