

基于多尺度生成对抗网络的 SAR 图像样本增广

李诗怡^{1,2,*}, 付光远¹, 崔忠马², 杨小婷², 汪洪桥¹, 陈雨魁²

¹火箭军工程大学作战保障学院, 陕西 西安 710025;

²北京遥感设备研究所, 北京 100854

摘要 针对军用无人飞行器对海上重要舰船合成孔径雷达图像获取困难的问题,提出了一种从单一图像学习图像内部分布的无条件图像生成网络。该网络采用金字塔式多尺度生成对抗网络(GAN)思想,在每一层金字塔中都有一个 GAN 负责该尺度下图像块的生成和判别,且每个 GAN 具有相似的结构。生成器前端采用不同大小卷积核连接的 Inception 模块获取不同尺度下的图像特征,为了充分利用这些特征,加入了残差密集模块;判别器采用马尔科夫判别器的思想,捕捉不同尺度下的图像分布。将所有生成的图像制成数据集用于训练不同的目标检测算法,结果表明,训练后模型的平均精度得到了一定的提升,验证了该网络模型的有效性。

关键词 数字图像处理; 金字塔结构; 残差密集; 多尺度; 生成对抗网络

中图分类号 TP751

文献标志码 A

doi: 10.3788/LOP57.201018

Data Augmentation in SAR Images Based on Multi-Scale Generative Adversarial Networks

Li Shiyi^{1,2,*}, Fu Guangyuan¹, Cui Zhongma², Yang Xiaoting², Wang Hongqiao¹, Chen Yukui²

¹College of Operational Support, Rocket Force University of Engineering, Xi'an, Shannxi 710025, China;

²Beijing Institute of Remote Sensing Equipment, Beijing 100854, China

Abstract To solve the problem that it is difficult for military unmanned aerial vehicles to acquire synthetic aperture radar images of important ships at sea, this paper introduces an unconditional image generation network which can learn the internal distribution of images from a single image. The network adopts the idea of a pyramid of multi-scale generative adversarial networks (GAN). In each layer of pyramid, there is a GAN responsible for the generation and discrimination of image blocks at this scale, and each GAN has a similar structure. The head of generator contains Inception modules connected with different sizes of convolution kernels to obtain image features at different scales. In order to make full use of these features, a residual dense block is added. The discriminator uses the idea of Markov discriminator to capture images distribution at different scales. All the generated images are made into data sets for training different target detection algorithms, the results show that the average accuracy of the model is improved to a certain extent, which verifies the effectiveness of the network model.

Key words digital image processing; pyramid structure; residual dense; multi-scale; generative adversarial networks

OCIS codes 100.2000; 280.6730

1 引言

深度学习是近年来备受关注的话题,但其需要大量的样本数据才能发挥优势^[1],而合成孔径雷达(SAR)图像由于设备限制导致数据来源不够丰富。

随着我国海洋事业发展的需求和海上威胁的不断增加,增强海洋监测技术在军用和民用方面都具有重大意义。考虑到军用无人飞行器拍摄的海上舰船 SAR 图像较少的问题,从已有 SAR 图像进行样本增广,以获得大量有效的 SAR 图像,对提高海上舰

收稿日期: 2020-02-13; 修回日期: 2020-03-05; 录用日期: 2020-03-09

基金项目: 国家自然科学基金青年科学基金(61403397)、陕西省自然科学基金基础研究计划(2015JM6313)

* E-mail: www.ryqlm@qq.com

船检测、识别能力十分必要^[2]。

基于计算机绘图的 SAR 仿真系统,如 Balz 等^[3]采用一种在图形处理器(GPU)上栅格化的方法进行快速仿真;Hammer 等^[4]提出了一种基于射线跟踪方法的模拟器,可模拟高分辨率的小场景,但这两种模拟器的精度都容易受到几何精度和电磁计算简化的影响。在目标分类领域,常见的数据增广方法,如剪裁、旋转、平移,都无法取得期望的效果,一方面是因为目标分类对图像平移、旋转不敏感^[5],另一方面是目标分类方法通常将图像整体进行变换,忽略了图像内的场景和目标变动。

生成对抗网络(GAN)是近年来深度学习的热门话题之一,其在自然语言处理、计算机视觉等方面均有应用,且不断往其他领域扩展。受博弈论中的零和博弈启发,GAN 将生成问题当作生成器(generator)和判别器(discriminator)两个网络之间的对抗问题^[6],即给定生成器噪声,将生成器生成的数据作为输出,并用判别器判别生成数据的真假。在博弈过程中,不断改进生成器,以生成更逼真的数据;判别器不断提高自身分辨数据真假的能力。但 GAN 存在训练困难、梯度消失、模型崩溃等问题^[7]。可通过改变 GAN 的结构提升网络的性能,如条件生成对抗网络(CGAN)^[8]可生成指定标签的数据;InfoGAN^[9]在 CGAN 的基础上将生成器输入的噪声进行分解,使数据的生成过程可控,并对生成结果有一定的解释性;Pix2Pix^[10]也是 CGAN 的一种变体结构,将图像映射到另一张图像上,成功将 GAN 应用到图像翻译领域;而循环一致性生成对抗网络(CycleGAN)^[11]、StarGAN^[12]解决了图像翻译的配

对问题,将图像到图像的一对一映射转换为多领域之间的转换。此外,GAN 在超分辨率、音频生成等领域也取得了巨大的进展。其次,可通过改变目标函数提升网络的性能。如 Wasserstein 生成对抗网络(WGAN)^[13]利用 Wasserstein-1 距离代替原始 GAN 中的 JS(Jensen-Shannon)散度,从而改进生成器和判别器的目标函数;带梯度惩罚的 WGAN(WGAN-GP)^[14]使用正则化形式对判别器进行约束,解决了模型崩溃、梯度消失的问题,使 GAN 能更稳定的进行学习。

已有研究中,GAN 的多种结构变体通常对应特定的任务,且需要一定数量的同类型数据集进行学习。如文献[15]中用 WGAN-GP 生成特定方位角 SAR 图像的新样本,需要多张样本图像作为训练集。本文参考文献[16]中的方法,针对 SAR 图像信息少且纹理简单的特点,提出了一种通过学习单张图像信息,生成有效 SAR 图像的方法。

2 网络结构

2.1 金字塔式多尺度 GAN

采用一种金字塔式多尺度结构 GAN 作为整体框架捕捉图像内部的统计信息,每一层金字塔都有一个简单的 GAN 负责该尺度下图像样本的生成和判别,具体网络结构如图 1 所示。与 GAN 及其变体网络以大量样本数据作为训练集不同,该方法仅需捕捉单一图像在不同尺度下的结构数据作为训练集,在捕获到全局信息的同时注重图像的细节纹理。

图 1 中的 N 为金字塔的尺度,由输入图像的大小决定, $\{x_N, x_{N-1}, \dots, x_0\}$ 为原始图像在不同尺度

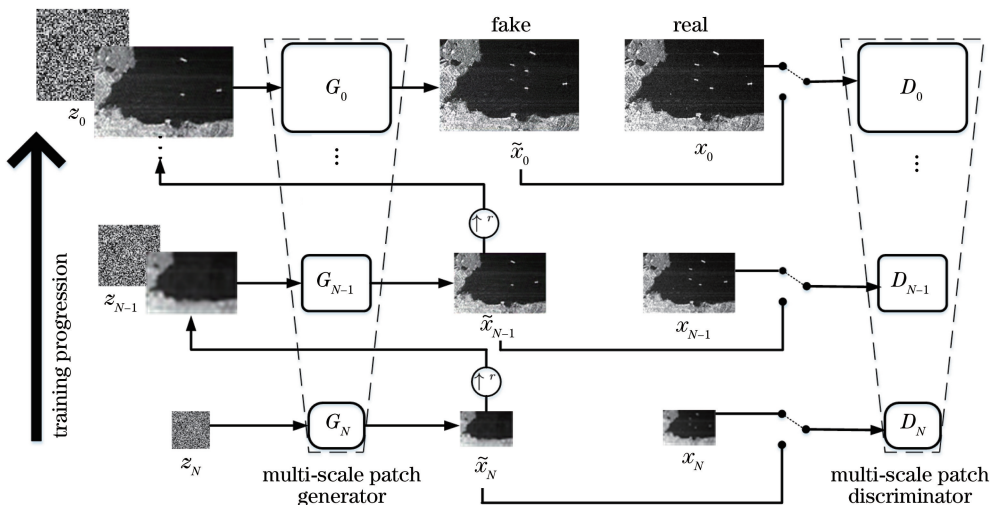


图 1 金字塔式多尺度 GAN 的结构

Fig. 1 Structure of the pyramidal multi-scale GAN

下的下采样结果,采样因子为 $r^n, n \in \{0, 1, 2, \dots, N\}$,且 $r > 1$ 。金字塔模型从最粗的尺度开始训练,第一个生成器 G_N 只需根据输入的高斯空间白噪声 z_N 就能生成图像样本 \tilde{x}_N ,可表示为

$$\tilde{x}_N = G_N(z_N)。 \quad (1)$$

通过噪声 z_N 生成样本 \tilde{x}_N 后,将图像 \tilde{x}_N 上采样的结果与高斯空间白噪声同时输入下一个尺度的生成器 G_{N-1} 中,再由生成器 G_{N-1} 生成新的样本图像 \tilde{x}_{N-1} ,以此类推,每一个尺度生成器生成的图像可表示为

$$\tilde{x}_{n-1} = G_{n-1}[z_{n-1}, (\tilde{x}_n)^{\uparrow r}]，n < N + 1， \quad (2)$$

式中, \tilde{x}_{n-1} 为第 $n-1$ 个尺度下的生成图像, G_{n-1} 为该尺度下的生成器, z_{n-1} 为该尺度下输入的噪声, $(\tilde{x}_n)^{\uparrow r}$ 为第 n 个尺度下生成图像 \tilde{x}_n 的上采样。

将第 n 个尺度下输入的噪声 z_n 和图像 \tilde{x}_{n+1} 的上采样结果同时输入生成器,生成器的主要作用是生成 \tilde{x}_{n+1} 缺少的细节并加入到 \tilde{x}_{n+1} 中,以生成新的样本图像 \tilde{x}_n ,可表示为

$$\tilde{x}_n = (\tilde{x}_{n+1})^{\uparrow r} + f_n[z_n + (\tilde{x}_{n+1})^{\uparrow r}]， \quad (3)$$

式中, f_n 为从生成图像 \tilde{x}_{n+1} 的上采样 $(\tilde{x}_{n+1})^{\uparrow r}$ 和噪声 z_n 到生成细节的映射,具体结构如图 2 所示,其中, BN 为批量归一化, Leaky ReLU 为带泄露修正线性单元,每一尺度的生成器具有相似的结构。在每个尺度上,生成器 G_n 生成的样本图像 \tilde{x}_n 除了输入到下一层的生成器外,还会传递给判别器 D_n ,将样本图像 \tilde{x}_n 与当前尺度下采样得到的图像块进行对比,直到判别器不能判别真假为止。整个训练过程是从下往上进行的,图像块由大到小,从粗糙到精细。

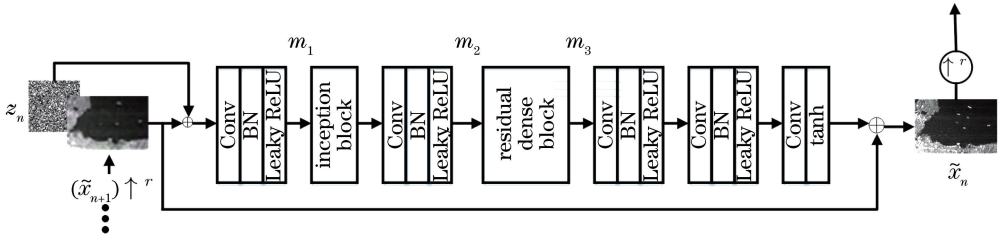


图 2 生成器的结构

Fig. 2 Structure of generator

2.1.1 Inception 模块

生成器将输入的图像用尺寸为 3×3 的卷积核卷积后,其输出通道数变为 192,将图 2 中的输出 m_1 作为 Inception 模块^[17]的输入,如图 3 所示。Inception 模块将输入作为四个分支,用不同尺寸的卷积核进行卷积后,在特征维度上拼接到一起。传统卷积层的输入只和一种尺寸的卷积核卷积,输出的维度是固定的,所有输出特征均匀分布在固定范围内。而本方法采用的 Inception 模块通过尺寸分别为 1×1 、 3×3 (两个)、 5×5 的卷积核提取特征,因此输出的特征分布更紧密,从而减少冗余信息、加快收敛速度,更快提取出目标舰船的独立特征。此外,因为感受野不同,该模块能得到不同尺度下的特征,获取的特征更丰富。

Inception 模块中有三条分支对输入进行卷积运算(卷积核尺寸为 1×1),对于单像素点, 1×1 的卷积运算(Conv)相当于该像素点在所有特征上进行了一次全连接运算,即全连接形式,如图 4 所示。在卷积运算之后,进行 BN,然后通过激活函数,以更好地模拟非线性特征,每个像素点可在相同的感

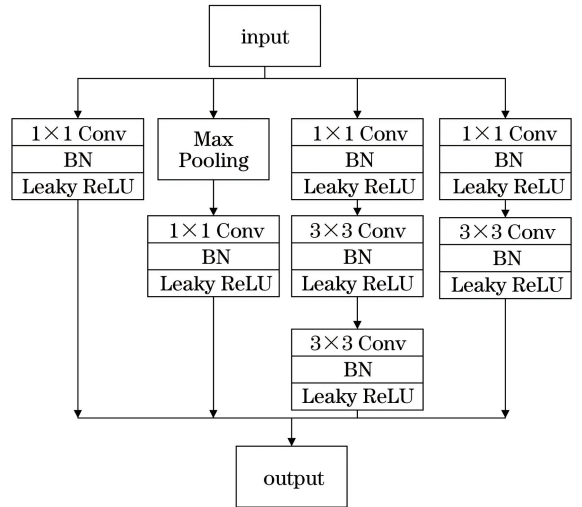


图 3 Inception 模块

Fig. 3 Inception block

受野中叠加更多的卷积,从而提取更多的特征。

为缩短本模型学习单张图像的时间、降低计算成本,使用 1×1 卷积,减少提取的特征数后, Inception 模块还将 5×5 卷积分解为两个 3×3 卷积,以提升模型的运算速度。

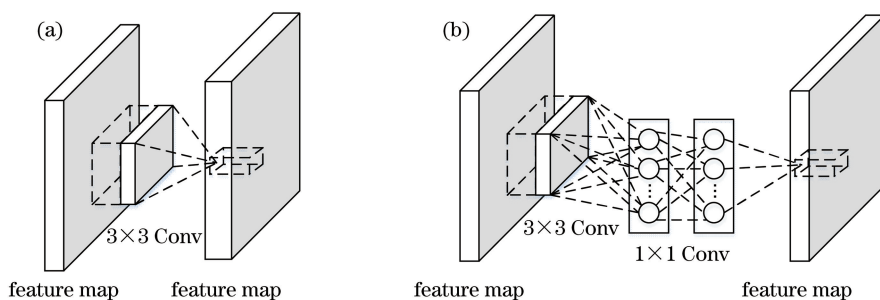


图4 线性卷积层与 1×1 卷积层。(a)线性卷积层;(b) 1×1 卷积层

Fig. 4 Linear convolutional layer and 1×1 convolutional layer. (a) Linear convolution layer; (b) 1×1 convolution layer

2.1.2 残差密集模块

Inception 模块提取不同尺度的特征得到的输出经下采样后,进入残差密集模块。实验采用的残差密集模块^[18]可以充分利用 Inception 模块获取的不同尺度舰船特征,提高生成器生成图像的质量。残差密集模块由五个卷积层密集连接,每个卷积层可以向所有后

续层传递信息。文献[19]已经证明去除 BN 层可以减少计算复杂度并提高网络性能,因此,生成器在残差密集模块中去除所有 BN 层,如图 5 所示,其中, m_2 为图 2 中卷积模块的输出。在计算该结构的输出时进行残差缩放,将残差乘以 0 到 1 之间的常数后,加入到主路径的输出中,可以有效避免网络结构的不稳定^[20-21]。

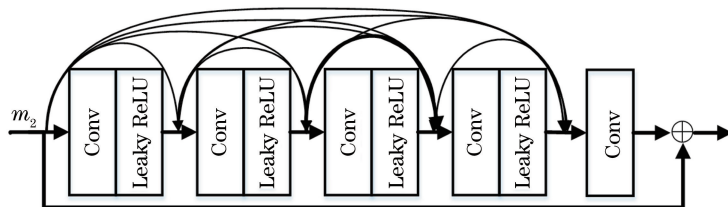


图5 残差密集模块

Fig. 5 Residual dense block

2.2 生成器的结构

为提取 SAR 图像中更多的特征,在原有生成器中的 5 个全卷积网络中加入密集残差模块和 Inception 模块,如图 2 所示。将 Inception 模块置于生成器前侧,有利于提取到不同尺度的特征,将提取到的特征进行下采样后输入残差密集模块,最后将得到的结果进行上采样后输出。

生成器的具体配置如表 1 所示,其中,Conv_block 由卷积、BN 和 Leaky ReLU 激活函数组成,在最粗的尺度上,输入和输出的通道数 C 为 32,每减少四个尺度,将 C 增加 2 倍。

2.3 判别器的结构

判别器采用文献[16]中的结构,采用马尔科夫判别器(patch-GAN)思想,由 5 个全卷积层构成,用下采样方法捕捉当前尺度下的图像分布,如图 6 所示。将当前尺度生成的样本图像和下采样得到的图像进行对比,并计算损失函数。判别器在最粗尺度上进行卷积操作输出的通道数为 32,每减少四个尺度增加 2 倍,卷积核的尺寸为 3×3 。

3 损失函数

模型训练从最粗的尺度开始,按图 1 所示的多尺度结构训练,每训练一个尺度,该尺度的 GAN 就会被固定。第 n 个 GAN 的训练损失包括对抗损失 $S_{\text{loss_adv}}$ 和重建损失 $S_{\text{loss_rec}}$,可表示为

$$S_{\text{loss_train}} = \min_{G_n} \max_{D_n} S_{\text{loss_adv}}(G_n, D_n) + \alpha S_{\text{loss_rec}}(G_n), \quad (4)$$

式中, α 为重建损失在训练损失中所占的权重,实验中取 10。

对抗损失:在每个尺度下生成器都有一个马尔科夫判别器,以判别该尺度下生成图像的真假^[10]。实验采用 WGAN-GP 损失^[14],用 Wasserstein-1 距离代替原始 GAN 的 JS 散度,并将判别器的 Lipschitz-1 约束正则化作为目标函数的惩罚项,解决了损失函数梯度消失问题,提高了模型的稳定性。马尔科夫判别器判别图谱的平均值可表示为

表 1 生成器的参数

Table 1 Parameters of the generator

Block	Operation	Convolution kernel	Input channel	Output channel	
Head	Conv_block	3×3	3	192	
Inception block	Block 1	Conv_block	1×1	192	64
	Block 2	Conv_block	1×1	192	96
		Conv_block	3×3	92	128
	Block 3	Conv_block	1×1	192	16
		Conv_block	3×3	16	32
		Conv_block	3×3	32	32
Block 4	Conv_block	1×1	192	32	
Connection part	Conv_block	3×3	C		
	Conv	3×3	C	C	
Residual dense block			Leaky ReLU		
		Conv	3×3	256	C
				Leaky ReLU	
		Conv	3×3	384	C
				Leaky ReLU	
		Conv	3×3	512	C
tail				Leaky ReLU	
		Conv_block	3×3	C	C
		Conv_block	3×3	C	C
		Conv	3×3	C	3

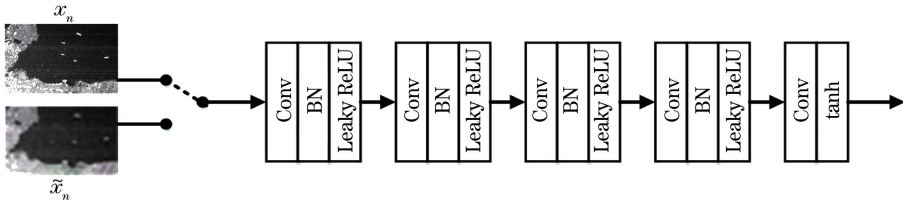


图 6 判别器的结构

Fig. 6 Structure of the discriminator

$$S_{\text{loss_adv}} = \text{adv}\{E_{x_n \sim p_{\text{data}}}[D(x_n)] + E_{\tilde{x}_n \sim p_g}[D(\tilde{x}_n)]\} + \lambda E_{\hat{x} \sim p_{\hat{x}}}[\|\nabla D(\hat{x})\| - 1]^2, \quad (5)$$

式中, adv 为取平均值函数, x_n 为原始图像的下采样结果, \tilde{x}_n 为由噪声生成的样本, p_{data} 为真实图像的分布, p_g 为生成图像的分布, $x_n \sim p_{\text{data}}$ 表示 x_n 服从 p_{data} 分布, $\tilde{x}_n \sim p_g$ 表示 \tilde{x}_n 服从 p_g 分布, x_{data} 为真实样本的集中区域, x_g 为生成样本的集中区域, \hat{x} 在 x_{data} 和 x_g 中随机插值, $\hat{x} = \epsilon x_{\text{data}} + (1 - \epsilon)x_g$, $\epsilon \in [0, 1]$, $\hat{x} \sim p_{\hat{x}}$ 表示 \hat{x} 服从 $p_{\hat{x}}$ 分布。

$D(x_n)$ 为判别器判别真实图像时输出的结果, $D(\tilde{x}_n)$ 为判别器判别生成图像时输出的结果。 E 为期望, ∇ 为梯度算符, $\|\cdot\|$ 为 L1 范数, λ 为梯度损失函数的权重, 实验中取 0.1。

重建损失: 为了生成原始图像 x 的特定噪声图谱, 令 $\{z_N^{\text{rec}}, z_{N-1}^{\text{rec}}, \dots, z_0^{\text{rec}}\} = \{z^*, 0, \dots, 0\}$, 其中, z_n^{rec} 为第 n 个尺度重建的噪声, $n \in \{0, 1, 2, \dots, N\}$, z^* 为训练中固定的噪声图谱, \tilde{x}_n^{rec} 为用噪声图谱在第 n 个尺度生成的图像。当 $n < N$ 时, 重建损失可表示为

$$S_{\text{loss_rec}} = \|G_n(0, [\tilde{x}_n^{\text{rec}}] \uparrow^r) - x_n\|^2. \quad (6)$$

当 $n=N$ 时,重建损失可表示为

$$S_{\text{loss_rec}} = \|G_N(z^*) - x_N\|^2. \quad (7)$$

可用 \tilde{x}_n^{rec} 确定每个尺度下噪声的标准差 σ_n , 即 \tilde{x}_n^{rec} 和 $\tilde{x}_{n+1}^{\text{rec}}$ 的均方根误差与 σ_n 成正比。

4 实验分析

4.1 实验数据和平台

实验数据:使用 SSDD 公开数据集,数据来源主要有 TerraSAR-X、RadarSat-2 和 Sentinel-1 传感器,有 HH、HV、VV 和 VH 四种极化方式,分辨率为 1~15 m,使用 LabelImg 进行标注,共 1160 张图像,包含 2456 艘舰船^[22]。

实验平台:系统为 Ubuntu18.04,CPU 为 Intel Core i7-9750H,主频为 2.60 GHz,GPU 为 GeForce RTX2060,内存为 16 GB。开发框架为 PyTorch,语言为 Python。

4.2 模型训练

模型训练采用 Adam 自适应算法,设置学习率为 0.0005,一阶动量为 0.5,二阶动量为 0.999,每个尺度的训练次数为 2000 次,当学习次数达到 1600 次时将学习率调整到 0.00005。

4.3 实验结果

图 7 为本方法生成的样本图像,从图 7(a1)~图 7(a4)、图 7(b1)~图 7(b4)可以发现,对于小尺寸舰船,本方法不仅可以学习到舰船的轮廓信息,还可以在生成样本图像时生成不同数量和分布的舰船以及不同的岛屿、海岸线的纹理信息,从而改变海岸地形;从图 7(c1)~图 7(c4)可以发现,对于带噪声的图像,本方法学习噪声分布时不会丢失舰船的特征信息,但生成的噪声缺少颗粒感;从图 7(d1)~图 7(d4)可以发现,本方法在学习大尺寸舰船时可得到舰船的整体轮廓,生成图像中的舰船在图中的数量、位置、大小均不同。

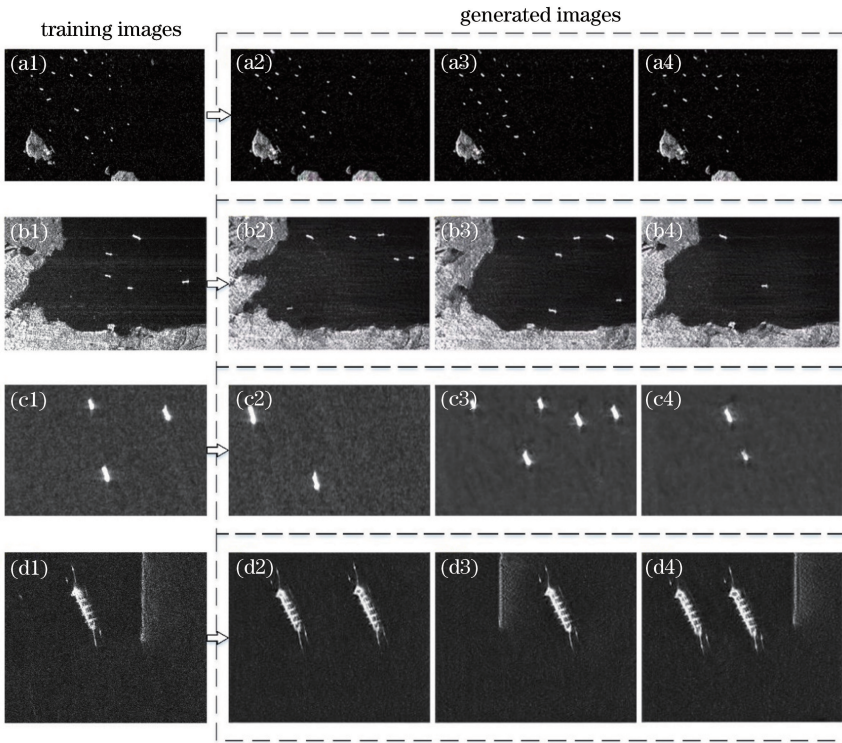


图 7 由单一图像生成的图像。(a)小尺寸舰船的图像 1;(b)小尺寸舰船的图像 2;(c)背景带噪声的图像;(d)大尺寸舰船的图像

Fig. 7 Image generated from single image. (a) Image1 of small ship; (b) image2 of small ship; (c) image with noise in background; (d) image of large ship

本方法改进的重点在于得到更多尺度特征的同时更好地利用图像特征,改进前后生成的图像如图 8 所示。原方法训练生成的样本图像如图 8(b1)、图 8(b2)所示,可以发现,原方法虽然能学习小尺度的舰船特征,但在学习大尺度舰船特征时,

不能将舰船作为一个整体,因此生成的图像出现舰船各部分分离以及与非舰船物体叠加的情况;原方法对舰船轮廓、纹理的学习也有所欠缺,如图 8(b1)和图 8(b2)中,舰船均存在虚假阴影。而本方法在网络首部采用不同尺寸大小的卷积核提取

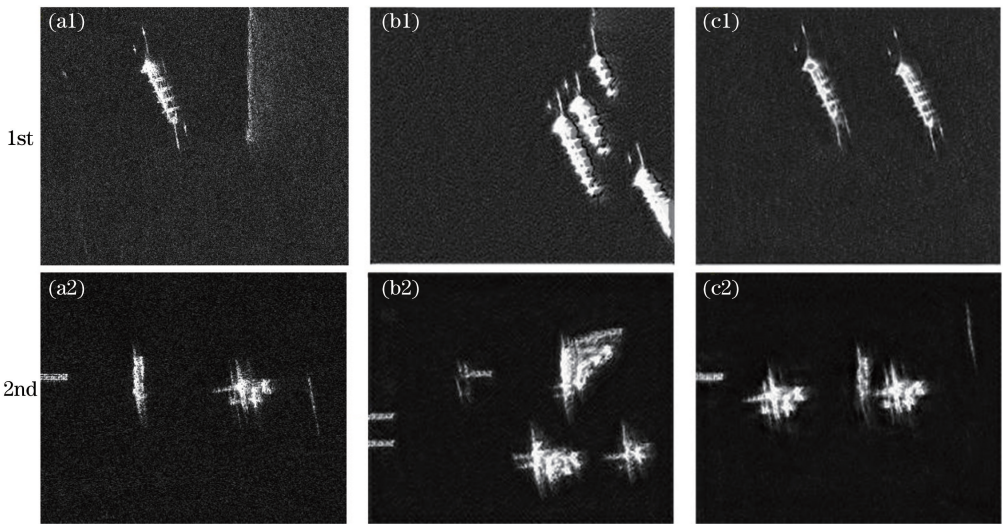


图 8 不同网络生成的图像。(a)用于训练的图像;(b)原网络生成的图像;(c)改进后网络生成的图像

Fig. 8 Images generated by different networks. (a) Images used for training; (b) images generated by the original network; (c) images generated by the improved network

特征,得到图像在不同尺度下的舰船特征,生成图像在保证舰船分布多样性的同时,还可以学习整艘舰船的轮廓。特别是在生成大尺寸舰船时,本方法生成的图像不会出现不完整的舰船以及目标叠加的情况,也没有虚假的阴影。其次,本方法应用了残差密集模块,能更好地利用获取的舰船特征,使生成图像中的舰船纹理更细致。

4.4 结果验证

采用本方法(记为 method1)和传统方法(记为 method2)对样本进行增广,传统增广方法主要对图像进行裁剪、平移、改变亮度、加噪声、旋转、镜像、在图像上添加黑色区域,本方法采用 LabelImg 图像标注软件对生成的样本图像进行人工标注。由于 SSDD 数据集中的数据类别只有舰船一种,因此标注时也只标注舰船,实验共制备了五个数据集。

1) SSDD_method1_20: 从 SSDD 数据集的 1160 张图像中选取 58 张进行训练,主要涉及简单背景下的大、小尺寸舰船和复杂背景下的大、小尺寸舰船。每张图像通过 method1 训练生成 20 张图像,共得到 1160 张图像。

2) SSDD_method1_40: 从 SSDD 数据集的 1160 张图像中选取 29 张图像,每张图像通过 method1 训练生成 40 张图像,共 1160 张图像。

3) SSDD_method2_20: 用相同的 58 张图像,每张图像用 method2 中的任意一种方式进行 20 次增广,共得到 1160 张图像。

4) SSDD_method2_40: 用相同的 29 张图像,每

张图像用 method2 中的任意一种方式进行 40 次增广,共得到 1160 张图像。

5) SSDD_method2_1160: 用 SSDD 数据集中的 1160 张图像,每张图像采用 method2 中的任意一种方式进行 1 次增广,共得到 1160 张图像。

本方法生成的样本主要用于 SAR 图像的目标检测,因此用目标检测算法判定生成的 SAR 图像是否有效。SSD 和 Tiny-YOLO 训练均采用 Adam 自适应算法, batch size 为 16, 初始学习率为 0.00004, epoch 为 10000, 训练集和测试集中的图像数量比为 7:3。由于 SSDD 数据集和标注的生成样本数据集只有一个类别,因此采用平均精度(AP)衡量目标检测算法的训练效果,将 Bounding box 和 Ground truth 的重叠阈值设为 0.5,得到的结果如表 2 所示。

表 2 不同方法生成图像的 AP

Table 2 AP of different methods to generate images

Dataset	AP	
	SSD	Tiny-YOLO
SSDD	0.8097	0.707
SSDD_method1_20	0.7275	0.641
SSDD+ SSDD_method1_20	0.8178	0.726
SSDD+ SSDD_method1_40	0.8101	0.716
SSDD+ SSDD_method2_20	0.8106	0.696
SSDD+ SSDD_method2_40	0.7907	0.672
SSDD+ SSDD_method2_1160	0.8213	0.702

当样本增广的原始图像只有 58 张时,分别用 method1 和 method2 将数据集增广至 1160 张图像,然后与 SSDD 数据集结合得到新的数据集

SSDD + SSDD_method1_20 和 SSDD + SSDD_method2_20, 用这两个新数据集训练目标检测算法 SSD 和 Tiny-YOLO。从表 2 可以看出, 用 method1 和 method2 进行样本增广时, 训练的 SSD 算法 AP 均有上升, 但 method1 上升较多, method2 几乎不变; 用 method2 增广的数据集训练 Tiny-YOLO 时, AP 有所下降; 当原始样本为 29 张图像时, 分别用 method1 和 method2 将数据集增广至 1160 张图像, 用 method2 增广的数据集训练目标检测算法, 两种算法的 AP 相比仅使用 SSDD 数据集有所下降。原因是普通增广方法只能从裁剪、平移、改变亮

度、加噪声、旋转、镜像、在图像上添加黑色区域 7 种方法中选择, 使用单一图像进行 20 次增广时, 存在大量重复样本。而本方法通过单一图像的学习, 由噪声产生样本图像, 不存在大量重复样本的问题。

相比原始 SSDD 数据, 用本方法增广的数据集, 训练得到的模型在目标检测时对虚警、漏检、错检现象都有所改进, 如图 9~图 11 所示。但因每次只训练一张图像, 生成的样本在舰船的多样性上存在不足。对靠近海岸线的大尺度舰船和单张图像中多艘小尺寸舰船的检测效果提升不明显, 如图 12 所示。

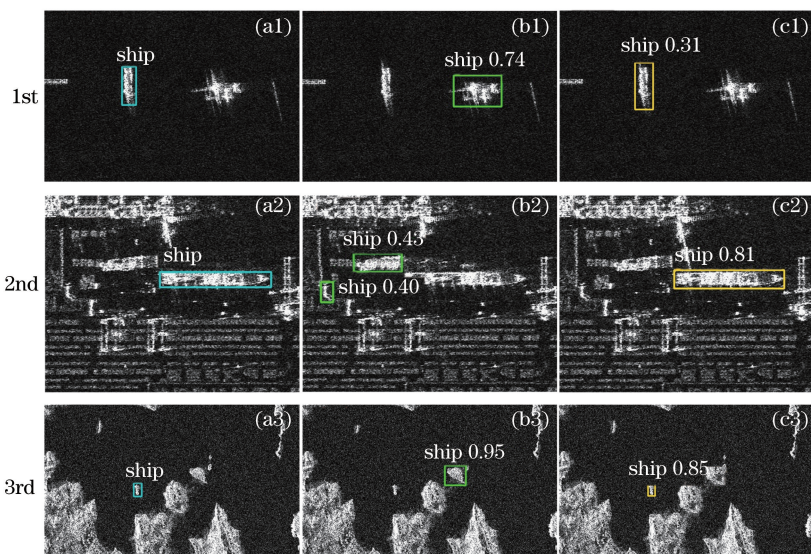


图 9 不同数据集训练的错检情况。(a) 正确检测结果; (b) SSDD 数据集; (c) SSDD 数据集 + 生成的样本数据集

Fig. 9 Error detection of training on different data sets. (a) Correct test results; (b) SSDD data set; (c) SSDD data set + generated sample data set

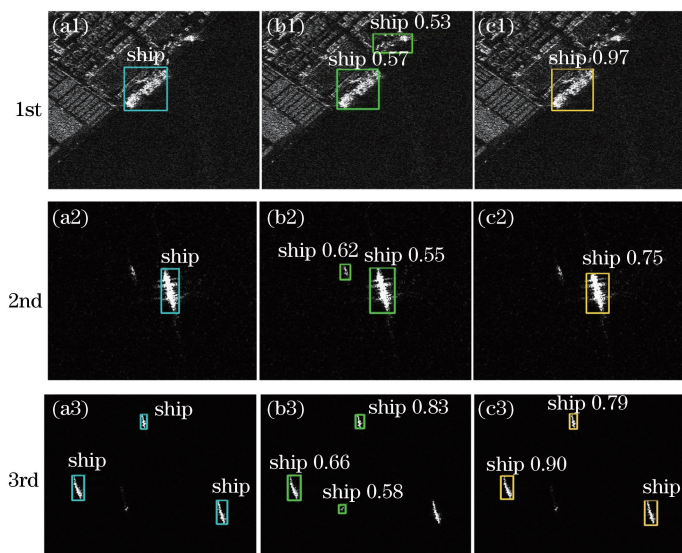


图 10 不同数据集训练的虚警情况。(a) 正确检测结果; (b) SSDD 数据集; (c) SSDD 数据集 + 生成的样本数据集

Fig. 10 False alarms of training on different data sets. (a) Correct test results; (b) SSDD data set;

(c) SSDD data set + generated sample data set

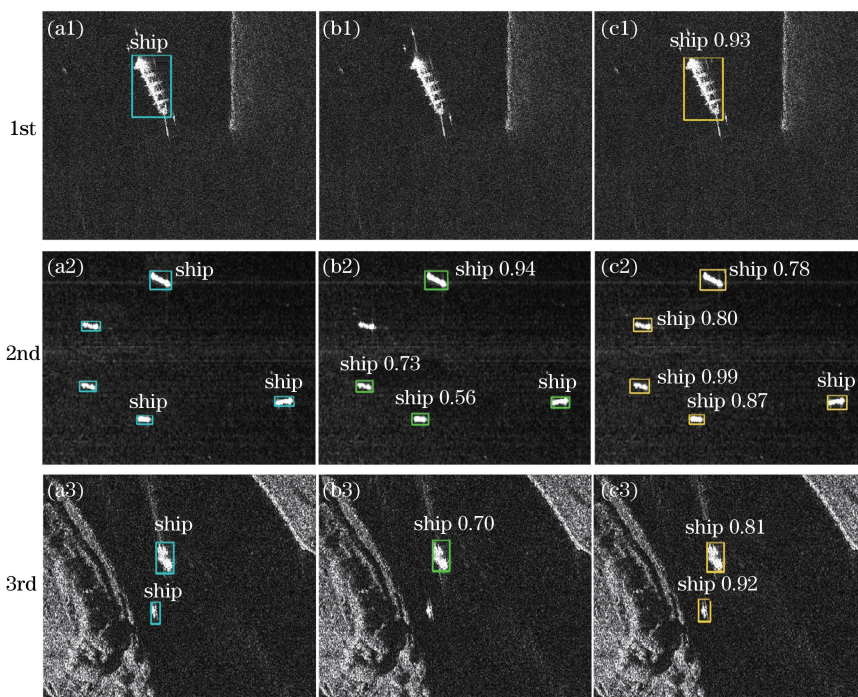


图 11 不同数据集训练的漏检情况。(a)正确检测结果;(b) SSDD 数据集;(c) SSDD 数据集+生成的样本数据集
Fig. 11 Missed detection of training on different data sets. (a) Correct test results; (b) SSDD data set;
(c) SSDD data set + generated sample data set

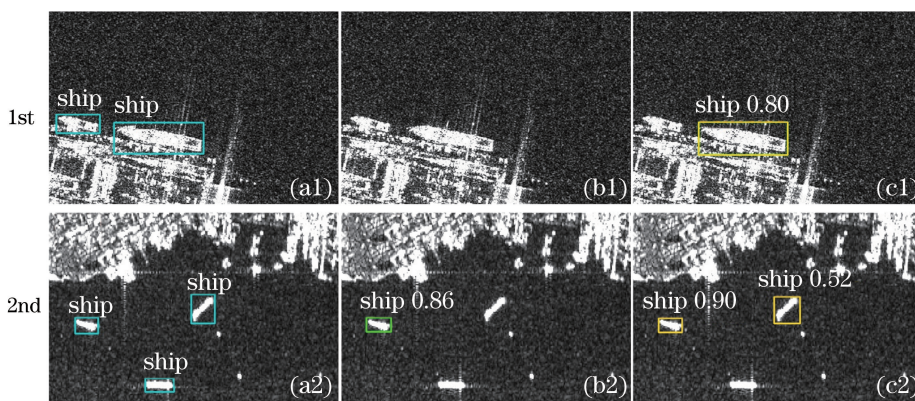


图 12 加入生成数据集后未检测出的结果。(a)正确检测结果;(b) SSDD 数据集;(c) SSDD 数据集+生成的样本数据集
Fig. 12 Undetected result after adding the generated data set. (a) Correct test results; (b) SSDD data set;
(c) SSDD data set + generated sample data set

5 结 论

针对深度学习需要大量样本数据支撑,而 SAR 图像获取较为困难的问题,基于 GAN 提出了一种学习单张图像内部分布的样本增广方法。相比常见的样本增广方法,本方法进行样本数据增广时可以改变图像中目标的数量、相对位置以及背景环境,使生成的样本具有多样性。使用金字塔式的多尺度 GAN 对单张图像进行学习,在生成器前端采用

Inception 模块获取更多尺度的特征,结合残差密集模块充分利用图像特征。将生成的图像标注后作为数据集训练目标检测算法,用训练的模型测试同一测试集,得到的结果较好,证明了本方法生成 SAR 图像的有效性。但本方法每次只能训练一张图像,再次处理另一张图像需要重新训练,增加了训练时间和训练成本。因此,之后的研究中还需考虑能否通过导入具有普适性的预训练模型对具体的图像进行微调,从而达到快速训练的目的。

参 考 文 献

- [1] Wang Z W, She Q, Ward T E. Generative adversarial networks: a survey and taxonomy [EB/OL]. [2020-01-31]. <https://arxiv.org/abs/1701.04862>.
- [2] Yi W, Zeng Y, Yuan Z. Fusion of GF-3 SAR and optical images based on the nonsubsampling contourlet transform [J]. *Acta Optica Sinica*, 2018, 38(11): 1110002.
易维, 曾湧, 原征. 基于 NSCT 变换的高分三号 SAR 与光学图像融合 [J]. *光学学报*, 2018, 38(11): 1110002.
- [3] Balz T, Stilla U. Hybrid GPU-based single-and double-bounce SAR simulation [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2009, 47(10): 3519-3529.
- [4] Hammer H, Schulz K. Coherent simulation of SAR images [J]. *Proceedings of SPIE*, 2009, 7477: 74771G.
- [5] Cubuk E D, Zoph B, Mané D, et al. AutoAugment: learning augmentation strategies from data [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 113-123.
- [6] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [EB/OL]. [2020-01-27]. <https://arxiv.org/abs/1406.2661>.
- [7] Arjovsky M, Bottou L. Towards principled methods for training generative adversarial networks [EB/OL]. [2020-02-05]. <https://arxiv.org/abs/1701.04862>.
- [8] Mirza M, Osindero S. Conditional generative adversarial nets [EB/OL]. [2020-01-30]. <https://arxiv.org/abs/1411.1784>.
- [9] Chen X, Duan Y, Houthoofd R, et al. InfoGAN: interpretable representation learning by information maximizing generative adversarial nets [EB/OL]. [2020-02-03]. <https://arxiv.org/abs/1606.03657>.
- [10] Isola P, Zhu J Y, Zhou T H, et al. Image-to-image translation with conditional adversarial networks [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 5967-5976.
- [11] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 2242-2251.
- [12] Choi Y, Choi M, Kim M, et al. StarGAN: unified generative adversarial networks for multi-domain image-to-image translation [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 8789-8797.
- [13] Ajocsky M, Chintala S, Bottou L. Wasserstein GAN [EB/OL]. [2020-02-04]. <https://arxiv.org/abs/1701.07875>.
- [14] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of Wasserstein GAN [EB/OL]. [2020-01-27]. <https://arxiv.org/abs/1704.00028>.
- [15] Cui Z Y, Zhang M R, Cao Z J, et al. Image data augmentation for SAR sensor via generative adversarial nets [J]. *IEEE Access*, 2019, 7: 42255-42268.
- [16] Shaham T R, Dekel T, Michaeli T. SinGAN: learning a generative model from a single natural image [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE, 2019: 4569-4579.
- [17] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 2818-2826.
- [18] Wang X T, Yu K, Wu S X, et al. ESRGAN: enhanced super-resolution generative adversarial networks [M] // Leal-Taixé L, Roth S, et al. *Computer Vision - ECCV 2018. Lecture Notes in Computer Science*. Cham: Springer, 2018, 11133: 63-79.
- [19] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 2261-2269.
- [20] Lim B, Son S, Kim H, et al. Enhanced deep residual networks for single image super-resolution [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 1132-1140.

- [21] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-ResNet and the impact of residual connections on learning [EB/OL]. [2020-01-30]. <https://arxiv.org/abs/1602.07261>.
- [22] Li J W, Qu C W, Peng S J, et al. Ship detection in SAR images based on generative adversarial network and online hard examples mining [J]. Journal of Electronics & Information Technology, 2019, 41(1): 143-149.
- 李健伟, 曲长文, 彭书娟, 等. 基于生成对抗网络和线上难例挖掘的 SAR 图像舰船目标检测 [J]. 电子与信息学报, 2019, 41(1): 143-149.