

doi: 10.7690/bgzdh.2022.07.011

## 基于多尺度 GAN 网络的 SAR 舰船目标扩充

黄琼男<sup>1</sup>, 朱卫纲<sup>1</sup>, 刘渊<sup>2</sup>, 李佳芯<sup>1</sup>, 杨莹<sup>1</sup>

(1. 航天工程大学电子与光学工程系, 北京 101416; 2. 中国人民解放军 61646 部队, 北京 100192)

**摘要:** 针对构建合成孔径雷达 (synthetic aperture radar, SAR) 图像舰船目标数据集的过程中, 某些舰船类型样本不足的问题, 提出一种基于多尺度生成对抗网络 (IC-ConsinGAN) 的 SAR 舰船目标扩充方法。通过将注意力机制引入并行多阶段多尺度 GAN 网络中, 提取 SAR 舰船目标的关键特征, 抑制背景特征, 使得生成的 SAR 图像舰船目标不仅具有精细化结构, 而且弥补了单幅图像生成过程中多样性不足的问题。实验结果表明: SIFID 指标比原始 ConsinGAN 网络模型下降了 0.02, 将扩充数据加入到 SAR 舰船目标识别任务中, 10 类舰船目标平均识别率提升了 8.4%, 证实了 IC-ConsinGAN 模型的有效性, 具有一定的工程应用价值。

**关键词:** 生成对抗网络; 合成孔径雷达; 注意力机制; 多尺度; 舰船目标识别

**中图分类号:** TN971 **文献标志码:** A

## SAR Ship Target Expansion Based on Multiscale GAN Network

Huang Qiongnan<sup>1</sup>, Zhu Weigang<sup>1</sup>, Liu Yuan<sup>2</sup>, Li Jiaying<sup>1</sup>, Yang Ying<sup>1</sup>

(1. Department of Electronic and Optical Engineering, Space Engineering University, Beijing 101416, China;

2. No. 61646 Unit of PLA, Beijing 100192, China)

**Abstract:** In order to solve the problem of insufficient samples of some ship types in the process of constructing synthetic aperture radar (SAR) image ship target data set, a SAR ship target expansion method based on multiscale generated countermeasure network (IC-ConsinGan) is proposed. By introducing the attention mechanism into the parallel multi-stage multiscale GAN network, the key features of SAR ship targets are extracted and the background features are suppressed, so that the generated SAR image ship targets not only have a refined structure, but also make up for the lack of diversity in the process of generating a single image. The experimental results show that the SIFID index is 0.02 lower than that of the original ConsinGan network model, and the average recognition rate of 10 types of ship targets is improved by 8.4% when the extended data is added to the SAR ship target recognition task, which confirms the effectiveness of the IC-ConsinGan model and has certain engineering application value.

**Keywords:** generative countermeasure network; synthetic aperture radar; attention mechanism; multiscale; ship target recognition

### 0 引言

合成孔径雷达 (SAR) 具有全天时工作、全天候成像、作用距离远和穿透能力强等特点, 在能见度极低的气象条件下仍能获得高分辨率 2 维图像<sup>[1]</sup>, 在军事、民用等领域具有重大意义。现阶段, 影响检测性能的关键因素主要有算法和数据 2 方面:

1) 学者专注于检测算法的改进和开拓; 2) 学者致力于数据的研究。数据作为推动深度学习技术发展的“原油”, SSDD、HRSID、LS-SSDD 等数据集<sup>[2-6]</sup>的出现为 SAR 舰船目标积累了大量的数据; 然而, 由于各国对非合作目标采用相应的反侦察技术, 导致采集此类目标样本有限, 需要对特定的样本数据进行扩充。

目前, 数据扩充的手段多种多样, 旋转、平移、

添加噪声等数据扩充方法并未从根本上改变图像内部的信息, 只是在图像的几何层面进行了形状的改变; 计算机建模仿真的方法精确度较低, 对于很多实际场景下的非合作目标, 建立精准的几何模型异常困难; 因此, 合理地扩充 SAR 数据势在必行。

生成的对抗网络能够从大量无标签数据中无监督地学习到各种数据特征, 在数据扩充领域有较强的通用性<sup>[7]</sup>。DCGAN<sup>[8]</sup>、WGAN-GP<sup>[9]</sup>等 GAN 变体模型本身在训练时需要大量的训练数据才能取得较好的生成结果, 而特定类型的 SAR 舰船目标的数量远远达不到, 且 GAN 网络对生成图像的尺寸较为敏感, 例如, DCGAN 生成图像的较好尺寸为 128×128 像素, 而笔者所用的数据集高分辨率 SAR 图像舰船尺寸为 800×800 像素。基于以上分析, 笔

收稿日期: 2022-03-20; 修回日期: 2022-04-24

作者简介: 黄琼男 (1995—), 男, 河南人, 硕士, 从事 SAR 舰船目标检测和识别数据集构建、SAR 图像生成研究。E-mail: 878105935@qq.com。

者着眼于对单幅图像进行训练完成扩充任务，与每个阶段都需对所有网络层进行权重优化的 SinGAN<sup>[10]</sup> 相比，采用并行训练机制的 ConsinGAN<sup>[11]</sup> 所需训练时间大大缩短，并且在网络中引入注意力机制模块 (convolutional block attention module, CBAM)<sup>[12]</sup>，重点关注舰船目标所在的区域特征，通过多尺度结构和注意机制的结合，

快速有效地提取 SAR 图像的关键特征。

### 1 ConsinGAN 模型

ConsinGAN 网络采用并行多阶段训练方式，通过对多尺度下的图像结构进行训练，生成所需分辨率大小的图像，同时，注重生成图像的细节纹理等内部信息。ConsinGAN 网络模型结构如图 1 所示。

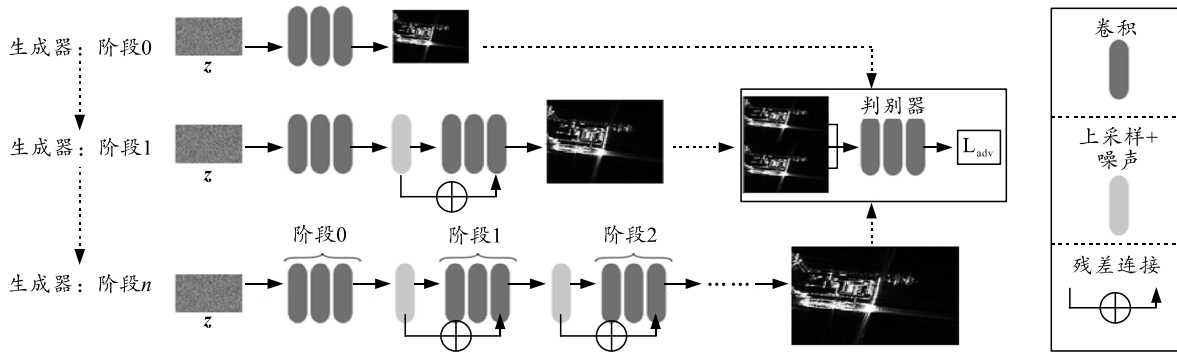


图 1 并行多阶段多尺度 GAN 网络

生成器由 3×3 卷积块和 BN-LeakyReLU 组成，在通过卷积层之前，每个阶段都经过上采样操作并添加额外的噪声，以提高生成图像的多样性。初始通道数设为 128，再利用残差连接重复通过卷积层，最后经过一个 3×3 卷积和 Tanh 激活函数输出生成样本。判别器由 5 个全卷积构成，用下采样的方法获取当前尺度下图像的分布，并将当前阶段生成的图像数据与下采样得到的图像数据输入到判别器网络中计算损失函数。

率，那么生成的图像质量会高些，但是差异性较弱，相反，对较低的生成器采用较小的学习率，那么生成图像的差异性会丰富一些，故采用学习率缩放的方法进行改良。以缩放因子  $\delta$  对学习率  $\eta$  进行缩放，例如阶段  $n$  的生成器  $G_n$  采用的学习率为  $\delta^0\eta$ ， $n-1$  阶段学习率为  $\delta^1\eta$  等， $\delta=0.1$  可以在图像保真度和多样性之间取得较好的平衡。

### 2 IC-ConsinGAN 模型

网络模型的训练过程：首先，从第 0 阶段粗略的分辨率开始多次迭代训练，输入为随机噪声向量  $z$ ，输出为低分辨率图像上的映射，训练阶段  $n$  达到收敛后，通过添加 3 个卷积层来增大生成器的大小。SinGAN 是将前一个阶段生成器生成的图像作为输入，由于生成器与生成器之间用图像作为传输介质，严重影响了网络训练的速度，且限制了生成器之间的交互；鉴于此，此网络结构设置为每个阶段都将上一阶段的特征图作为输入。然后进行阶段 1 训练，从原始的上采样特征，到新添加卷积层的输出间添加残差连接，重复此过程  $n$  次，直到达到所需的输出图像分辨率。

注意力机制源于人类视觉对不同信息敏锐度的研究，主要目标是从众多信息中筛选出对当前任务最有用的信息。笔者采用 CBAM 模块，依次通过通道注意力模块和空间注意力模块得到特征图，使得模型可以在通道上学到“是什么”，在空间上学到“在哪里”。对于图像生成任务，这种“关注”能够使得网络快速定位到目标区域，提升目标的关键特征并且抑制用处不大的特征，从而提高生成图像的质量。一般情况下，不同通道特征对目标生成的效果并非同等重要，有些通道所提取的仅仅是背景区域的特征。如图 2 所示，经卷积得到的特征图中，不同通道含有的信息不同，通道 52 和通道 59 很好地响应了舰船目标的信息，而通道 64 所提取的均为背景特征。ConsinGAN 网络对特征提取时，给各个通道分配了相同的权重；因此，笔者在原网络的基础上增添通道注意力机制，对目标响应较高的通道分配大的权重，使得网络更加关注那些对目标特征响应较高的通道。

利用特征图作为输入会导致训练过拟合，即训练出的模型不会生成任何“新”图像，而只是生成原图像。为解决过拟合问题，一方面，在训练的每个阶段都添加额外的噪声，以提高生成图像的多样性；另一方面，对于不同的生成器采用不同的学习率进行训练，如果对较低的生成器采用较高的学习

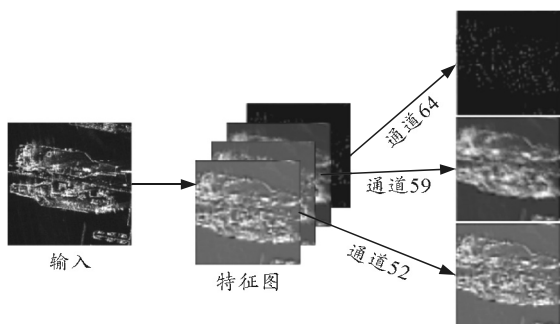


图 2 不同通道特征图对目标特征的响应

对于通道注意力机制的原理，输入为经过卷积

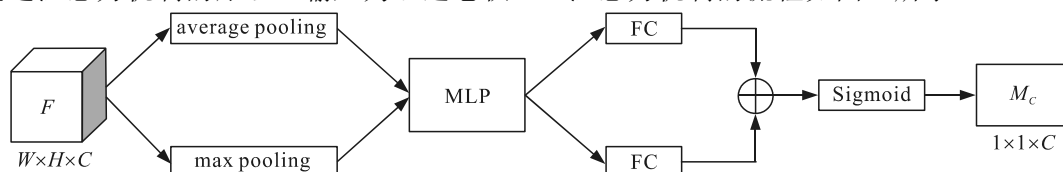


图 3 通道注意力机制

从空间分布的角度分析，传递给生成器一种信息：仅“关注”图像切片中舰船目标特征，而不用过多地关注黑色场景区域特征。利用卷积神经网络进行特征提取时，会将空间区域特征同等对待。当生成器将生成的数据送入到判别器中，可能会因为其与实际图像之间存在较大的差距，而当生成器将生成的场景数据送入到判别器中，其与实际数据无差别，极易发生模式坍塌现象，最终生成一张张黑色场景图像数据；因此，直接采用未经“选择”的卷积网络来生成 SAR 舰船目标，一般不能直接获得较好的结果。笔者希望通过某种方法来实现对空间分布的区别对待，使网络增加对目标区域的关注，减少对场景区域的关注。空间注意力机制在减少场景区域的干扰方面有着很好的效果。通过给舰船目

神经网络得到的特征图  $F=[F_1, F_2, \dots, F_c]$ ，其中： $W$  和  $H$  分别为特征图的宽和高； $C$  为通道数。将该特征图在空间上并行通过一个最大值池化 (max pooling) 和一个平均池化 (average pooling) 处理，分别得到 2 个 1 维向量，然后将 2 个 1 维向量通过一个多层感知机 (multi-layer perceptron, MLP)，并进行逐像素相加的操作，最后将此时学习到的特征通过一个激活函数，使所有通道值和为 1。这样就实现了在通道维度上对原始特征权重的重分配。通道注意力机制的流程如图 3 所示。

标分配更大的权重，给场景区域分配较小的权重，在生成过程中，有效地避免场景区域对舰船目标的干扰。

对于空间注意力机制的原理，输入同样为经过卷积神经网络得到的特征图  $F$ ，特征图上空间点的几何定义为  $L = \{(x, y) | x = 1, \dots, W; y = 1, \dots, H\}$ ，其中  $(x, y)$  为空间点的坐标。然后将该特征图在通道上通过一个最大值池化和一个平均池化处理，将处理后的 2 个特征图按通道维度拼接在一起，之后将含有多个通道的特征图整合到通道数为 1 的一张特征图上，此步骤用卷积操作实现；最后，将得到的特征图通过一个激活函数，使所有通道值和为 1。这样就实现了对空间上每个像素点权重的分配。空间注意力机制的流程如图 4 所示。

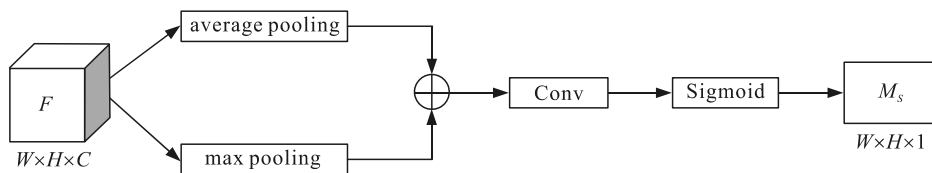


图 4 空间注意力机制

IC-ConsinGAN 网络借鉴 SinGAN 对单幅图像即可进行训练的思想，采用 ConsinGAN 并行多阶段多尺度的基本框架，沿用 WGAN-GP 的损失函数来施加  $L$  约束，最后通过对 SAR 图像与自然图像进行分析，在原网络的基础上加入注意力机制 (CBAM 模块)；此外，为了使网络所提取的特征更加丰富，引入 Inception 模块。原因如下：当训练数据所提供的特征较少时，若直接采用注意力机制模块，可能会导致网络模型虽然提升了舰船目标特征的权重，

但受自身特征的限制仍生成一堆杂乱无章的噪声。为了使网络提取的特征更加丰富，笔者在生成器的浅层网络中引入 Inception V2 模块，使网络模型获得丰富的特征，然后通过注意力机制模块对特征权重重分配，从而实现在不同训练数据上均能取得较好的结果。判别器与 SinGAN 使用的网络结构相同，采用马尔科夫判别器的思想，由于金字塔多尺度架构能够观察图像的全局分布和细节纹理，足以区分生成图像和真实图像，若加入注意力机制只专注于

某个特征，就会传递给生成器一种错觉，认为其余特征不重要，可能会导致只生成单一的场景图像；因此，在判别器中不加入注意力机制。

### 3 实验与分析

在图像生成领域，利用深度学习进行模型训练时，需要强大的硬件和成熟的软件支持。本实验环境为 i7-7700K@4.20 GHz 的 CPU 和 TITAN V 的 GPU，深度学习框架使用 Pytorch 并调用 CUDA 进行加速。

实验数据集由 35 景高分三号 SAR 图像制得而成，采用 SAR-AIS 关联匹配以及光学遥感图像和 SAR 图像特征匹配的方法，共制得各类型舰船图像切片 1 175 张，各类型舰船目标所含图像切片数量如图 5 所示，数据集中含有 1 m 和 3 m 分辨率数据，图像切片的尺寸分别为 800×800、256×256 像素。

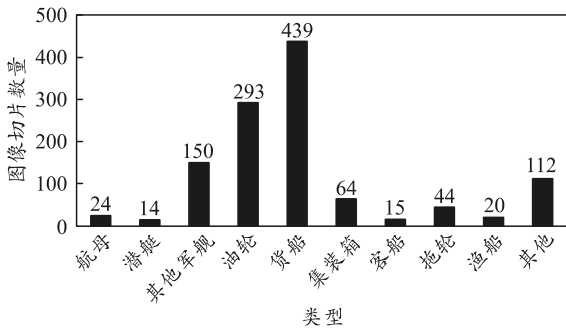


图 5 各类型舰船图像切片数量

分别对航母、油轮、渔船 3 类目标进行生成，生成结果如图 6 所示。

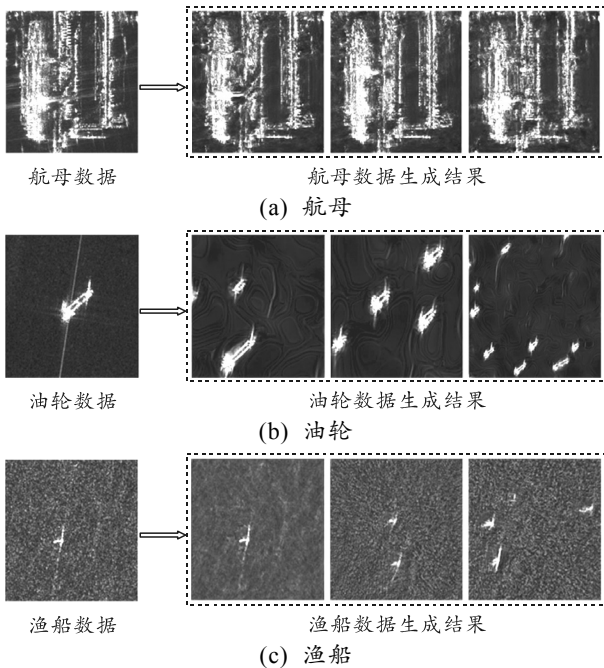


图 6 几种典型特定样本的生成图像

从生成结果可知，笔者提出的针对单幅图像生成的 GAN 网络，不仅可以学习到 SAR 图像的全局信息，而且注重纹理细节信息的学习。舰船目标的轮廓及精细化结构等特征均清晰可见，在生成图像样本中还可以改变舰船目标强散射中心的位置、舰船目标的数量、大小及位置分布等。高分辨率 SAR 图像更精细的细节信息难以生成，对于小尺寸舰船目标的生成效果明显不如大尺寸目标，对于一些特定的样本，网络训练可能会出现过拟合现象。

为了更加直观地展示 IC-ConsinGAN 模型的效果，选用 GAN、WGAN、WGAN-GP、SinGAN、ConsinGAN 和 IC-ConsinGAN 6 种方法分别对航母数据、油轮数据进行对比实验。其中，航母数据 24 张，油轮数据 293 张，后 3 种方法针对单幅图像进行扩充，实验结果如图 7、8 所示。

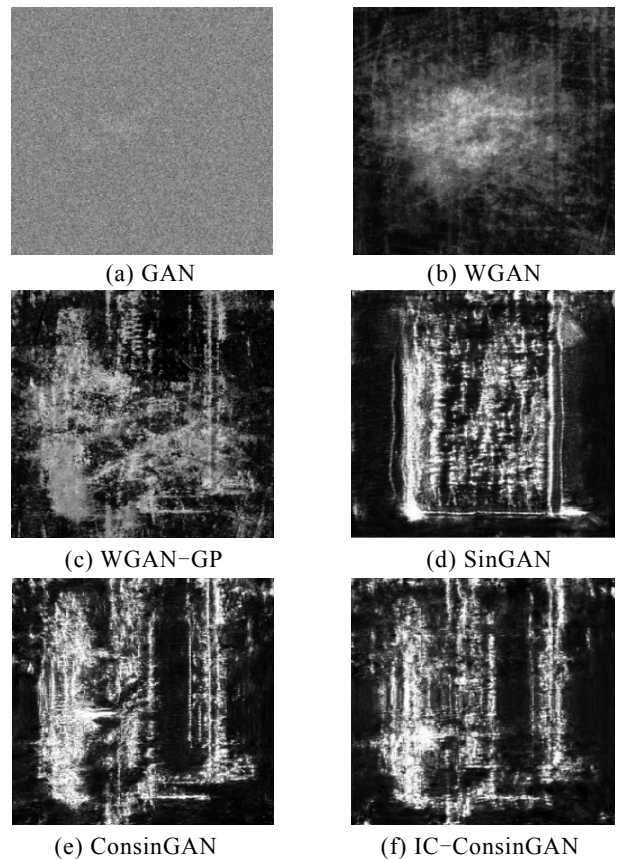
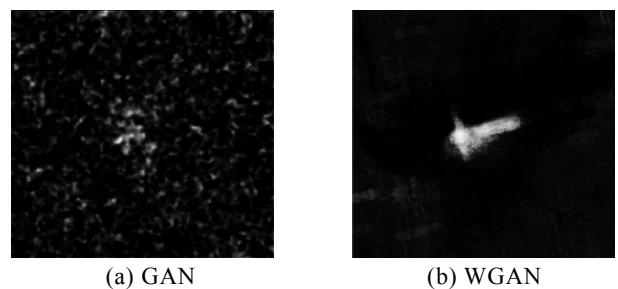


图 7 6 种方法对航母数据生成结果



(a) GAN (b) WGAN

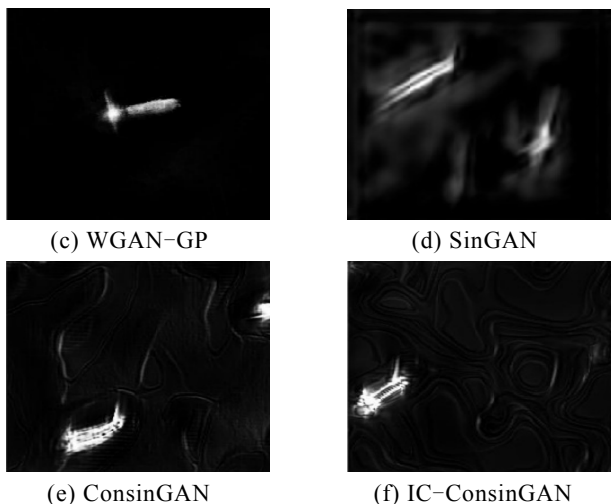


图 8 6 种方法对油轮数据生成结果

从结果中可知，当训练图像较多时，WGAN、

WGAN-GP 的方法才能取得效果，但难以生成大尺寸舰船目标。以单幅图像为训练样本的模型能够取得很好的效果，IC-ConsinGAN 模型对于大、小尺寸舰船目标均能捕获其精细化结构特征。

弗雷歇距离 (fréchet inception distance, FID) 是常用的评价 GAN 图像性能指标之一，由于笔者重点针对单图像模型，所以采用衡量单图像生成的 FID 指标 (single image fréchet inception distance, SIFID)。FID 评价指标是一种距离的度量标准，其值越小，表明生成模型分布与真实分布的差异越小，以此来判断所学习模型的优劣。

6 种方法的 SIFID 指标如表 1 所示，IC-ConsinGAN 模型比原始 GAN 和 ConsinGAN 模型分别降低约 4.27 和 0.02，进一步证实了模型的有效性。

表 1 不同方法的 SIFID 值

方法	GAN	WGAN	WGAN-GP	SinGAN	ConsinGAN	IC-ConsinGAN
SIFID	4.285	2.395	1.155	0.200	0.034	0.014

笔者对 SAR 图像生成是以提高 SAR 图像智能解译为目的，因此，将由不同方法扩充的数据加入到原始训练集中，探究生成图像的优劣及其对识别率的影响。采用 ResNet50 进行分类实验，记原始数据集为 HGDSAR，将各类型的舰船目标随机划分为 80% 训练数据和 20% 测试数据，采用 3 种方法对训练集进行扩充，测试集不变，具体说明如下：

1) method1: 从 10 种类型数据中各挑选出 2 张图像进行训练，共挑选出 20 张图像，主要涉及不同分辨率下简单场景大、小尺寸目标和复杂场景大、小尺寸目标。用所挑选数据分别进行旋转 (90°、180°、270°)、添加噪声 (高斯噪声、椒盐噪声、随机噪声等)、翻转 (水平、垂直、对角线等)、调节亮度和对比度等 4 种变换扩充 30 次，共制得 600 张图像。

2) 用同样的 20 张图像经 ConsinGAN 模型分别采样出 30 张图像，共 600 张图像。

3) 用所挑选的 20 张图像经 IC-ConsinGAN 模型分别采样出 30 张图像，共制得 600 张图像。

4 种识别结果用混淆矩阵表示如图 9 所示。

4 种不同的训练集识别率分别为 59.1%、59.9%、65.8%、67.5%，利用 IC-ConsinGAN 模型扩充的方法比原始数据集的识别率提升了 8.4%。用传统的扩充方法并不能使识别率有较大提升，潜艇、拖轮等类型的目标识别率甚至有所下降。在传统的扩充方法中，人为地添加噪声可能会导致网络学习到太多

无关的特征，尽管测试集中含有噪声数据，但人工添加的噪声和 SAR 成像机理所产生的噪声不同，潜艇等目标通常所处场景复杂，添加噪声更增加了识别的难度；改变亮度和对比度的方法可能会改变 SAR 图像中的散射强度，SAR 图像为复数图像，不同 SAR 图像之间的区别更多在于散射强度的不同，而测试集中可能不含有相同或相似散射强度的 SAR 图像数据，导致识别率并未有较大提升。由 HGDSAR\_method2 和 HGDSAR\_method3 识别结果可知，用生成对抗网络的方法能够大大提升 SAR 舰船识别率，其中 IC-ConsinGAN 相比于 ConsinGAN 方法提高 1.7%。可以看出：笔者提出的方法对渔船、拖轮等一些难以识别的小尺寸舰船目标有较大提升，不仅保持了生成大尺寸图像的有效性，而且通过增加注意力机制改善了小尺寸舰船目标的生成效果。该方法对 SAR 图像识别效果有较大的提升。

真实类别 \ 预测类别	航母	军舰	潜艇	油轮	货船	集装箱	客船	渔船	拖轮	其他
航母	4	1	0	0	0	0	0	0	0	0
军舰	0	21	0	1	6	1	0	0	0	1
潜艇	0	0	2	0	1	0	0	0	0	0
油轮	0	2	0	40	11	0	1	0	0	5
货船	0	2	0	15	57	4	0	1	3	6
集装箱	0	0	0	2	6	4	0	0	0	1
客船	0	1	0	0	0	1	1	0	0	0
渔船	0	1	0	1	1	0	0	1	0	0
拖轮	0	2	0	0	1	1	0	0	3	2
其他	0	0	0	3	6	1	0	1	5	7

(a) HGDSAR

预测类别 真实类别	航母	军舰	潜艇	油轮	货船	集装箱	客船	渔船	拖轮	其他
航母	3	1	0	1	0	0	0	0	0	0
军舰	0	18	0	2	7	1	0	0	0	2
潜艇	0	1	0	0	1	0	0	1	0	0
油轮	0	0	0	47	9	0	0	0	1	2
货船	0	2	0	17	64	1	0	1	1	2
集装箱	0	1	0	1	8	3	0	0	0	0
客船	0	1	0	0	1	0	0	0	0	1
渔船	0	1	0	1	0	1	0	1	0	0
拖轮	0	0	0	3	4	0	0	0	0	2
其他	0	1	0	3	11	1	0	1	0	6

(b) HGDSAR\_method1

预测类别 真实类别	航母	军舰	潜艇	油轮	货船	集装箱	客船	渔船	拖轮	其他
航母	5	0	0	0	0	0	0	0	0	0
军舰	1	20	0	5	3	1	0	0	0	0
潜艇	0	0	3	0	0	0	0	0	0	0
油轮	0	3	0	46	6	1	2	0	0	1
货船	0	3	0	13	68	1	0	1	1	1
集装箱	0	2	0	2	4	5	0	0	0	0
客船	0	1	0	0	1	0	0	0	0	1
渔船	0	1	0	2	0	0	0	1	0	0
拖轮	0	3	0	1	0	0	0	1	1	3
其他	0	0	0	3	7	2	0	0	3	7

(c) HGDSAR\_method2

预测类别 真实类别	航母	军舰	潜艇	油轮	货船	集装箱	客船	渔船	拖轮	其他
航母	4	1	0	0	0	0	0	0	0	0
军舰	1	20	0	2	4	1	0	0	0	3
潜艇	0	0	3	0	0	0	0	0	0	0
油轮	0	0	0	46	9	2	0	0	1	1
货船	0	4	0	12	62	3	0	0	4	3
集装箱	0	0	0	2	4	7	0	0	0	0
客船	0	0	0	0	1	0	0	0	0	2
渔船	0	0	0	1	0	0	0	3	0	0
拖轮	0	0	0	1	1	0	0	0	5	2
其他	0	0	0	3	6	1	0	0	3	10

(d) HGDSAR\_method3

图 9 4 种训练集 的识别结果

## 4 结论

笔者提出一种针对 SAR 舰船目标生成的 IC-ConsinGAN 模型, 在多阶段多尺度 GAN 网络中引入注意力机制, 有效地提取舰船目标的关键特征, 抑制背景区域特征, 通过与其他 GAN 衍生模型进行对比, 证实了该方法的有效性。将扩充数据加入

到识别网络中, 结果表明: 10 类舰船目标的识别准确率有较大提升, 能够为 SAR 图像智能解译提供高效的数据扩充方法。

## 参考文献:

- [1] 郁文贤, 柳彬, 丁拥科, 等. 高分辨率、多时相 SAR 图像数据集的构建[J]. 测绘通报, 2014(S1): 119-122.
- [2] 李健伟, 曲长文, 邵嘉琦, 等. 基于深度学习的 SAR 图像舰船检测数据集及性能分析[C]//第五届高分辨率对地观测学术年会论文集. 2018: 22.
- [3] WANG Y, WANG C, ZHANG H, et al. A SAR dataset of ship detection for deep learning under complex backgrounds[J]. Remote Sensing, 2019, 11(7): 765.
- [4] 孙显, 王智睿, 孙元睿, 等. AIR-SARShip-1.0: 高分辨率 SAR 舰船检测数据集[J]. 雷达学报, 2019, 8(6): 852-862.
- [5] WEI S, ZENG X, QU Q, et al. HRSID: A high-resolution SAR images dataset for ship detection and instance segmentation[J]. IEEE Access, 2020, 8(1): 120234-120254.
- [6] ZHANG T W, ZHANG X L, KE X, et al. LS-SSDD-v1. 0: A Deep Learning Dataset Dedicated to Small Ship Detection from Large-Scale Sentinel-1 SAR Images[J]. Remote Sensing, 2020, 12(18): 2997.
- [7] 梁俊杰, 韦舰晶, 蒋正锋. 生成对抗网络 GAN 综述[J]. 计算机科学与探索, 2020, 14(1): 1-17.
- [8] RADFORD, ALEC, METZ, et al. Unsupervised representation learning with deep convolutional generative adversarial networks[DB/OL]. <https://arxiv.org/abs/1511.06434>, 2015-11-19.
- [9] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of wasserstein gans[C]//Advances in neural information processing systems. 2017: 5767-5777.
- [10] SHAHAM T R, DEKEL T, MICHAELI T. SinGAN: Learning a generative model from a single natural image[C]//In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV). 2019.
- [11] HINT T, FISHER M, WANG O, et al. Improved Techniques for Training Single-Image GANs[C]// 2021 IEEE Winter Conference on Applications of Computer Vision(WACV). IEEE, 2021.
- [12] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Proceedings of the European conference on computer vision (ECCV). 2018.