

基于DRGAN和支持向量机的合成孔径雷达图像目标识别

徐英, 谷雨, 彭冬亮, 刘俊

引用本文:

徐英, 谷雨, 彭冬亮, 等. 基于DRGAN和支持向量机的合成孔径雷达图像目标识别[J]. *光学精密工程*, 2020, 28(3): 727–735.

XU Ying, GU Yu, PENG Dong-liang, et al. SAR ATR based on disentangled representation learning generative adversarial networks and support vector machine[J]. *Optics and Precision Engineering*, 2020, 28(3): 727–735.

在线阅读 View online: <https://doi.org/10.3788/OPE.20202803.0727>

您可能感兴趣的其他文章

Articles you may be interested in

采用改进尺度不变特征变换在多变背景下实现快速目标识别

Fast object recognition under multiple varying background using improved SIFT method
光学精密工程. 2015, 23(8): 2349–2356 <https://doi.org/10.3788/OPE.20152308.2349>

基于吉林一号遥感图像的星载目标快速识别系统

Target fast matching recognition of on-board system based on Jilin-1 satellite image
光学精密工程. 2017, 25(1): 255–262 <https://doi.org/10.3788/OPE.20172501.0255>

基于改进SIFT特征提取的车标识别

A vehicle logo recognition algorithm based on the improved SIFT feature
光学精密工程. 2018, 26(5): 1267–1274 <https://doi.org/10.3788/OPE.20182605.1267>

基于小波包和BP神经网络的周界入侵防御系统目标识别

Target recognition of perimeter intrusion defense system based on wavelet packet and BP neural network
光学精密工程. 2015, 23(10z): 806–813 <https://doi.org/10.3788/OPE.20152313.0807>

中高轨道目标的地基光电监视

Ground-based photoelectric surveillance for mid-high orbit target
光学精密工程. 2017, 25(10): 2584–2590 <https://doi.org/10.3788/OPE.20172510.2584>

文章编号 1004-924X(2020)03-0727-09

基于 DRGAN 和支持向量机的 合成孔径雷达图像目标识别

徐 英, 谷 雨*, 彭冬亮, 刘 俊

(杭州电子科技大学 自动化学院, 浙江 杭州 310018)

摘要:为解决 SAR 图像目标识别中样本缺乏和方位角敏感问题,提出了一种基于 DRGAN 和 SVM 的 SAR 图像目标识别算法。首先,采用多尺度分形特征对 SAR 图像进行增强,经过分割得到目标二值图像,基于 Hu 二阶矩估计目标的方位角。然后对估计得到的目标方位角进行量化编码,结合原始图像作为输入,对设计的 DRGAN 模型参数进行训练与优化。由于 DRGAN 中的深度生成模型将目标姿态与外观表示进行解耦设计,故可利用该模型将 SAR 图像样本变换到同一方位角区间。基于变换后的训练样本分别提取归一化灰度特征,利用 SVM 训练分类器。采用 MSTAR 数据集在多个不同操作条件下对提出的算法进行测试,实验结果表明,在带变体的标准操作条件下,能够达到 97.97% 的分类精度,优于部分基于 CNN 模型分类精度,在 4 种扩展操作条件下的分类精度分别为 97.83%, 91.77%, 97.11% 和 97.04%, 均优于传统方法的分类精度。在 SAR 图像目标方位角估计存在一定误差的情况下,训练得到的 GAN 模型作为 SAR 图像目标旋转估计器,能够使得在不进行复杂样本预处理的前提下,仍然取得较高的 SAR 图像目标识别精度。

关键词:合成孔径雷达图像;目标识别;生成对抗网络;方位角估计;支持向量机

中图分类号:TP753 文献标识码:A doi:10.3788/OPE.20202803.0727

SAR ATR based on disentangled representation learning generative adversarial networks and support vector machine

XU Ying, GU Yu*, PENG Dong-liang, LIU Jun

(School of Automation, Hangzhou Dianzi University, Hangzhou Zhejiang 310018, China)

* Corresponding author, E-mail: guyu@hdu.edu.cn

Abstract: To solve the problem of insufficient training samples and sensitivity of the target's aspect angle for Synthetic Aperture Radar (SAR) Automatic Target Recognition (ATR), a target recognition algorithm for SAR images based on a DRGAN and Support Vector Machine (SVM) was proposed in this paper. First, a multiscale fractal feature was used to enhance the input SAR image, and the target binary image was obtained through threshold segmentation. Hu second moments were used to estimate the aspect angle of the target. Second, the estimated angle was quantized into a vector, and the parameters of the designed DRGAN model were trained and optimized using these vectors and original images. Because the deep generative model in DRGAN was designed by disentangling the target's pose from its representation, the aspect angles of SAR image samples could be transformed into the same interval through this model. Normalized gray features were extracted

收稿日期:2019-07-19;修订日期:2019-09-04.

基金项目:国家自然科学基金面上项目资助(No. 61771177);国防基础科研项目资助(No. JCKY2018415C004)

from these transformed training samples, and an SVM classifier was trained using these features. MSTAR database was used to test the performance of the proposed algorithm under different operating conditions. The experimental results demonstrate that the classification accuracy reached 97.97% under standard operating conditions with variants, which is superior to some methods based on a convolutional neural network. The proposed algorithm can achieve classification accuracies of 97.83%, 91.77%, 97.11% and 97.04% under four extended operating conditions, respectively, which are better than traditional methods. Despite some errors during the estimation of the aspect angle of the object in the SAR image, the trained GAN model acting as rotation estimator of SAR objects still achieves better SAR object recognition performance under the condition in which no complex data preprocessing methods are used.

Key words: Synthetic Aperture Radar (SAR) image; target recognition; generative adversarial networks; aspect angle estimation; Support Vector Machine(SVM)

1 引言

合成孔径雷达 (Synthetic Aperture Radar, SAR) 图像自动目标识别 (Automatic Target Recognition, ATR) 是 SAR 图像解译的一个重要研究方向,通过对 SAR 图像进行处理,实现对感兴趣目标的自动检测、鉴别与分类,是导弹末制导、海洋监视与矿藏探测等应用的重要环节之一^[1]。

SAR 传感器通常搭载在航空器和航天器等运动平台,利用合成孔径原理实现高分辨率微波成像,故 SAR 图像数据获取成本高。SAR 图像具有目标方位角敏感性,在不同操作条件(成像参数)下 SAR 图像中目标外观差异大。SAR 图像具有较强的相干斑噪声,严重影响目标的分割检测质量。虽然研究学者已经提出了很多 SAR ATR 算法,但这些算法通常在特定的操作条件下才能够取得理想的分类结果,距离成熟应用还有一段距离^[2]。对于 SAR ATR 问题,在目标方位角不完备及成像参数变化较大情况下导致训练样本有限时,如何提高 SAR 图像的目标识别性能,是值得深入研究的一个重要问题^[3]。

特征提取与选择、分类器优化设计是影响识别精度的两个关键要素^[4]。目前用于 SAR 图像目标识别的典型分类器包括模板匹配、支持向量机 (Support Vector Machine, SVM)^[5]、稀疏表示^[6]和集成学习^[7]等。随着深度学习理论和技术的飞速发展,卷积神经网络 (Convolution Neural Networks, CNN) 作为一类典型的监督深度学习模型,在目标检测与识别方面取得了优于传统方法的成果,已经开始应用于 SAR 图像目标识别。

在模型结构设计方面,文献[8]参照 VGGNet 结构,设计了一种用于 SAR ATR 的 CNN 结构,但由于采用较小的卷积核 (3×3),所以对 MSTAR (Moving and Stationary Target Acquisition and Recognition) 数据集中 10 类含变体目标的分类精度仅为 93.16%。文献[9]提出的模型 (A-ConvNets) 使用了较大的卷积核 ($5 \times 5, 6 \times 6$),且采用全局平均池化代替全连接层,由于减小了模型参数,故在数据增强的情况下对 10 类不含变体目标的分类精度能够达到 99.13%。文献[10]通过实验分析了加性和连接特征融合结构对基于 CNN 的 SAR 图像识别性能的影响,并提出了一种基于多特征决策级融合的 CNN 模型结构 (Fusion-CNN),对 10 类不含变体目标的分类精度达到了 99.42%。文献[11]研究了 CNN 模型中第一层,即空间特征提取层,对识别精度的影响,设计了一种基于多尺度空间特征提取的深度 CNN 结构 (MultiScale-CNN),对 10 类含变体目标的分类精度达到了 97.69%。

将 CNN 用于 SAR ATR 时,为解决数据集中训练样本不足的问题,通常采用数据增强手段获得足够多的训练样本来提高模型的泛化能力。文献[3]提出了一种基于目标散射中心的 SAR 图像增强方法。文献[5]研究了仿射变换和弹性形变两种数据增强手段对 SAR 图像识别精度的影响。文献[8]采用平移、添加噪声和角度合成等操作获取增强后的训练样本集,实验结果表明这 3 种数据增强手段能极大提高模型的分类精度。文献[9]通过随机裁剪的方法,从分辨率为 128×128 的图像中提取分辨率为 88×88 的子图像。除数据增强手段外,采用深度生成模型生成样本也是一种有效增加

样本数量的方法。为解决目标域 SAR 图像方位角缺失的问题,文献[12]首先采用光线追踪原理生成多种方位角下的 SAR 仿真图像。为提高仿真图像的质量,基于 CycleGAN 算法^[13]将图像由仿真域转换到目标域,实现了目标域图像的数据增强。文献[14]提出了可旋转隐空间的思想,通过输入具有不同方位角的两幅 SAR 图像,采用编码器-解码器结构作为输入图像的隐空间特征表示,利用该特征学习得到一个旋转变换矩阵,然后利用学习得到的矩阵生成多种方位角下的 SAR 图像。文献[15]设计了基于深度生成神经网络的目标特征空间零样本学习模型,设计的模型包括构造器、生成器和解析器 3 个模块,其中构造器将输入的样本标签转换到目标特征空间,然后结合目标的方位角信息通过生成器重构目标图像。文献[16]通过改进 Wasserstein 自编码器模型结构和重构误差提高了 SAR 图像样本生成的质量。

考虑到 SAR 图像具有目标方位角敏感性,本文借鉴人脸识别领域提出的基于解耦表示学习的生成对抗网络模型(Disentangled Representation Learning Generative Adversarial Networks, DRGAN)^[17],首先基于多尺度分形特征估计图像中的目标方位角,对其量化编码后,将训练样本和目标方位角编码向量等作为输入优化 DRGAN 模型参数。然后,将训练得到的 DRGAN 模型视为目标方位角估计器,将训练样本和测试样本归一化到同一方位角量化区间。最后,从变换后的图像中提取归一化灰度特征,采用 SVM 训练分类器。采用 MSTAR 数据集进行模型训练和目标识别实验,验证了提出算法的有效性。

2 基于 DRGAN 和 SVM 的 SAR 图像目标识别流程

本文提出的基于 DRGAN 和 SVM 的 SAR 图像目标识别算法总体框图如图 1 所示。

采用多尺度分形特征对 SAR 图像进行增强,对经过阈值分割后的二值图像基于 Hu 不变矩估计目标方位角。对估计的目标方位角进行量化编码;利用 SAR 训练图像及方位角编码向量对 DRGAN 模型进行参数优化;基于训练得到 DRGAN 模型中的深度生成模型对 SAR 训练图像进行方位角归一化,提取 SAR 图像灰度特征,训练 SVM 分类器;基于训练得到 DRGAN 模型中的深度生成模型对 SAR 测试图像进行方位角

归一化,提取 SAR 图像灰度特征,利用训练得到的 SVM 分类器进行测试。

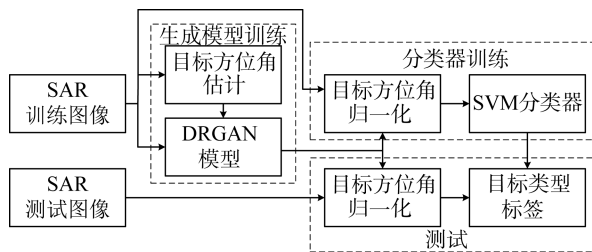


图 1 本文算法流程框图

Fig. 1 Flowchart of proposed algorithm

3 算法实现细节

3.1 基于多尺度分形和 Hu 二阶矩的目标方位角估计

由于 SAR 图像具有较强的相干斑噪声,在计算多尺度特征进行图像增强前,首先采用 5×5 的滤波核对输入 SAR 图像进行均值滤波。

依据分形理论,利用地毯覆盖法在尺度 $\epsilon = 1, 2, \dots, \epsilon_{\max}$ 的情况下计算二维图像表面积 $A(x, y, \epsilon)$ 的公式如式(1)和式(2)所示^[18]:

$$V(x, y, \epsilon) = \sum_{m=x-\epsilon}^{x+\epsilon} \sum_{n=y-\epsilon}^{y+\epsilon} (U(m, n, \epsilon) - B(m, n, \epsilon)), \quad (1)$$

$$A(x, y, \epsilon) = V(x, y, \epsilon) / 2\epsilon, \quad (2)$$

其中, $U(m, n, \epsilon)$ 和 $B(m, n, \epsilon)$ 分别为像素坐标 x, y 处,尺寸为 $(2\epsilon+1) \times (2\epsilon+1)$ 范围内像素的最大值和最小值, $V(x, y, \epsilon)$ 为像素坐标 x, y 处由最大值和最小值曲面围成的三维空间的体积。

定义一个随尺度变化的分形特征 $K(x, y, \epsilon)$, 其计算公式为^[18]:

$$K(x, y, \epsilon) = \exp\left(\frac{\log A(x, y, \epsilon) \log(\epsilon-1) - \log A(x, y, \epsilon-1) \log \epsilon}{\log(\epsilon-1) - \log \epsilon}\right). \quad (3)$$

基于多尺度分形特征 $K(x, y, \epsilon)$ 定义一个度量 $IMFFK(x, y)$ 用于图像增强,如式(4)所示:

$$IMFFK(x, y) = \left(\frac{1}{\epsilon_{\max} - 1} \sum_{\epsilon=2}^{\epsilon_{\max}} K(x, y, \epsilon)^2 - \sum_{\epsilon=2}^{\epsilon_{\max}} \left[K(x, y, \epsilon) - \frac{1}{\epsilon_{\max} - 1} \sum_{\epsilon=2}^{\epsilon_{\max}} K(x, y, \epsilon) \right]^2 \right). \quad (4)$$

得到增强后图像采用自适应阈值分割算法得到目标检测结果,即目标二值图像。由于相干斑噪声的影响,分割后的图像中会有多个非连通区域。本文采用面积滤波法,计算目标二值图像中的每个连通区域的包围盒,然后选择面积最大的区域作为候选目标区域,基于 Hu 二阶矩估计目标的方位角^[19]。估计得到目标方位角后,需要对其量化编码才能作为 DRGAN 模型中生成模型 G 的输入。本文将角度范围 $[0^\circ, 360^\circ)$ 均分为 N_a 个区间,每个区间的角度范围 θ 如式(5)所示。此时能够将目标方位角估计值转化为独热(one hot)编码形式 a 。

$$\theta \in \left[i \times \frac{360}{N_a}, (i+1) \times \frac{360}{N_a} \right)_{i=0,1,\dots,N_a-1}. \quad (5)$$

3.2 DRGAN 模型设计与参数训练

GAN(Generative Adversarial Networks)模型包括生成模型 G 和判别模型 D 两部分,通过交替进行训练,优化模型 G 和 D 的参数,使得生成模型 G 能够捕捉到训练样本的分布^[20]。考虑到 SAR 成像过程与目标方位角密切相关,本文选用 DRGAN 模型进行深度生成模型的训练,因为基于 DRGAN 设计的深度生成模型能够实现目标外观表示与目标方位角的解耦。如图 2 所示,生成模型 G 包含编码器 G_{enc} 和解码器 G_{dec} 两部分,

G_{enc} 为一个由卷积构成的 CNN 模型,将输入 SAR 图像 Z 进行编码转换成一个向量,然后结合方位角编码 a 和随机噪声 n 作为编码器 G_{enc} 的输入,如式(6)所示,其中 $\text{concat}()$ 表示向量连接操作。经过由反卷积构成的解码器网络 G_{dec} 生成图像 \hat{Z} :

$$\hat{Z} = G_{\text{dec}}(\text{concat}(n, a, G_{\text{enc}}(Z))). \quad (6)$$

G_{dec} 和 D 具有相同的卷积层结构,但是 D 需要输出两个分支,用于鉴别图像 Z 和 \hat{Z} 。其中一个分支输出 N_l+1 维的目标类别编码, N_l 为样本类别数目;另一分支输出 N_a 维的目标方位角的编码。对于训练图像 Z ,其类别 one hot 编码中前 N_l 维仅目标类别对应的维度为 1,其余均为 0;对于生成图像 \hat{Z} ,其最后一维为 1,其余为 0。对于图像 Z 和 \hat{Z} ,其目标方位角的 one hot 编码均由图像 Z 的目标方位角估计值得到。

考虑到 SAR 图像样本数较少,本文在文献[17]设计模型基础上,将卷积操作的输出通道数减少为原来的一半,设计的判别模型 D 和生成模型 G 如表 1 所示。其中,Conv* 表示 Conv+BN+elu 组合,DeConv* 表示 DeConv+BN+elu 组合,Conv 表示卷积操作,DeConv 表示反卷积操作,BN 表示批归一化操作,elu 表示指数线性单元激活函数,详见文献[17]。

表 1 DRGAN 模型中判别模型 D 和生成模型 G 结构

Tab. 1 Architecture of discriminative model D and generative model G in DRGAN

G_{enc} 和 D 结构			G_{dec} 结构		
层名称	滤波器宽度及步长	输出尺寸	层名称	滤波器宽度及步长	输出尺寸
Input	—	$96 \times 96 \times 1$	FC	—	$6 \times 6 \times 320$
Conv11	$3 \times 3/1$	$96 \times 96 \times 16$	DeConv52	$3 \times 3/1$	$6 \times 6 \times 160$
Conv12	$3 \times 3/1$	$96 \times 96 \times 32$	DeConv51	$3 \times 3/1$	$6 \times 6 \times 128$
Conv21	$3 \times 3/2$	$48 \times 48 \times 32$	DeConv43	$3 \times 3/2$	$12 \times 12 \times 128$
Conv22	$3 \times 3/1$	$48 \times 48 \times 32$	DeConv42	$3 \times 3/1$	$12 \times 12 \times 64$
Conv23	$3 \times 3/1$	$48 \times 48 \times 64$	DeConv41	$3 \times 3/1$	$12 \times 12 \times 96$
Conv31	$3 \times 3/2$	$24 \times 24 \times 64$	DeConv33	$3 \times 3/2$	$24 \times 24 \times 96$
Conv32	$3 \times 3/1$	$24 \times 24 \times 48$	DeConv32	$3 \times 3/1$	$24 \times 24 \times 48$
Conv33	$3 \times 3/1$	$24 \times 24 \times 96$	DeConv31	$3 \times 3/1$	$24 \times 24 \times 64$
Conv41	$3 \times 3/2$	$12 \times 12 \times 96$	DeConv23	$3 \times 3/2$	$48 \times 48 \times 64$
Conv42	$3 \times 3/1$	$12 \times 12 \times 64$	DeConv22	$3 \times 3/1$	$48 \times 48 \times 32$
Conv43	$3 \times 3/1$	$12 \times 12 \times 128$	DeConv21	$3 \times 3/1$	$48 \times 48 \times 32$
Conv51	$3 \times 3/2$	$6 \times 6 \times 128$	DeConv13	$3 \times 3/2$	$96 \times 96 \times 32$
Conv52	$3 \times 3/1$	$6 \times 6 \times 160$	DeConv12	$3 \times 3/1$	$96 \times 96 \times 16$
Conv53	$3 \times 3/1$	$6 \times 6 \times 320$	DeConv11	$3 \times 3/1$	$96 \times 96 \times 1$
AvgPool	$6 \times 6/1$	$1 \times 1 \times 320$			
FC (for D)	—	N_l+1+N_a			

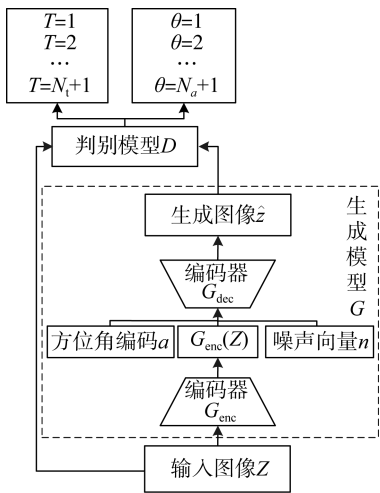


图 2 DRGAN 模型结构框图

Fig. 2 Block diagram of DRGAN model

3.3 基于 SVM 的分类器设计

训练得到 DRGAN 模型后,一方面可以用于生成不同角度的训练样本,另一方面可以将不同角度的图像归一化到同一方位角区间。因为本文估计得到的目标方位角带有一定误差,且对其进行了量化处理,故采用第一种方法生成图像时与输入训练样本图像的分布会有一定差异,故本文采用后一方法对输入图像进行方位角归一化。

基于训练得到的生成模型 G 对输入训练图像 Z 进行变换,此时输入的目标方位角 one hot 编码均设置成一致,例如 $[1, 0, 0, \dots]$,则相当于对输入图像进行方位角归一化。生成变换图像 Z-hat 后,将该图像向量化处理,然后除以 255.0,将图像灰度归一化到 $[0, 1]$ 区间。以此作为输入特征,采用 SVM 进行分类器训练。

4 实验结果分析

为验证本文提出算法的有效性,采用 MSTAR

数据集在多个场景下进行算法测试,并与目前已知的算法分类结果进行对比分析。训练 DRGAN 模型时,输入 SAR 图像的分辨率为 96×96 。采用 Adam 优化器,学习率设定为 0.000 2,批大小设定为 32。目标方位角量化区间设为 $N_a = 30$,随机噪声 n 维度设定为 50。每训练完一次判别模型 D,训练 4 次生成模型 G。训练模型 D 采用的损失函数为交叉熵损失函数。SVM 分类器采用径向基(rbf)核函数,gamma 值设定为 0.000 1,惩罚系数 C 设定为 32,采用“一对多”的分类策略,迭代次数设定为 1 000。

参考文献[8],标准操作条件下 MSTAR 数据中 10 类训练样本和测试样本类别和数目见表 2,将其记做 SOC。本文对测试样本中包含变体和不包含变体情况分别进行了算法性能测试。此外,采用了文献[6]和文献[9]中使用的 4 种扩展操作条件测试提出算法的性能,分别记做 EOC1-EOC4。EOC1 和 EOC2 的配置见文献[9],其中 EOC1 场景对 2S1, BRDM2, ZSU234 和 T72 4 类目标进行分类,用于比较在较大平台俯视角差异情况下的识别精度,EOC2 主要用于测试型号不同情况下的目标识别能力。EOC3 和 EOC4 的配置见文献[6]。在 EOC3 条件下,对 BMP2, T72, BTR60 和 T62 4 类目标进行分类,其中 BMP2 和 T72 存在变体情况,训练时采用 SN9563 和 SN132 型号,测试样本的型号分别为 SN9566, SNC21, SN812, SNS7。EOC4 主要测试分类器对于含变体情况的目标识别能力。与 EOC1 不同,EOC4 仅采用 2S1, BRDM2 和 ZSU234 3 类目标,并没有采用 T72。测试样本分别采用俯视角为 30° 和 45° 下的目标样本图像,其中 BRDM2 和 ZSU234 存在变体情况。本文仅测试俯视角为 30° 情况下的识别精度。

表 2 MSTAR 数据集描述

Tab. 2 Description of MSTAR dataset

目标类型	2S1	BRDM2	D7	T62	ZIL131	ZSU234	BTR60	BMP2			BTR70	T72		
								SN9563	SN9566	SNC21	SNC71d	SN132	SN812	SNS7
训练样本数	299	298	299	299	299	299	256	233			233	232		
测试样本数	274	274	274	273	274	274	195	195	196	196	196	196	195	191

采用本文设计的 DRGAN 模型中的深度生成模型 G 对训练样本进行角度归一化后的图像如图 3 所示。其中,方位角 a 的 one hot 编码设定为 $[1, 0, \dots, 0]$,随机噪声 n 设定为零向量。从图 3 可以看出,几乎所有的目标方位角均沿水平方向,仅图像的灰度分布有一定差异。方位角归一化后的 SAR 图像位于高维欧氏空间的若干低维流形,这有利于后续分类器的设计。

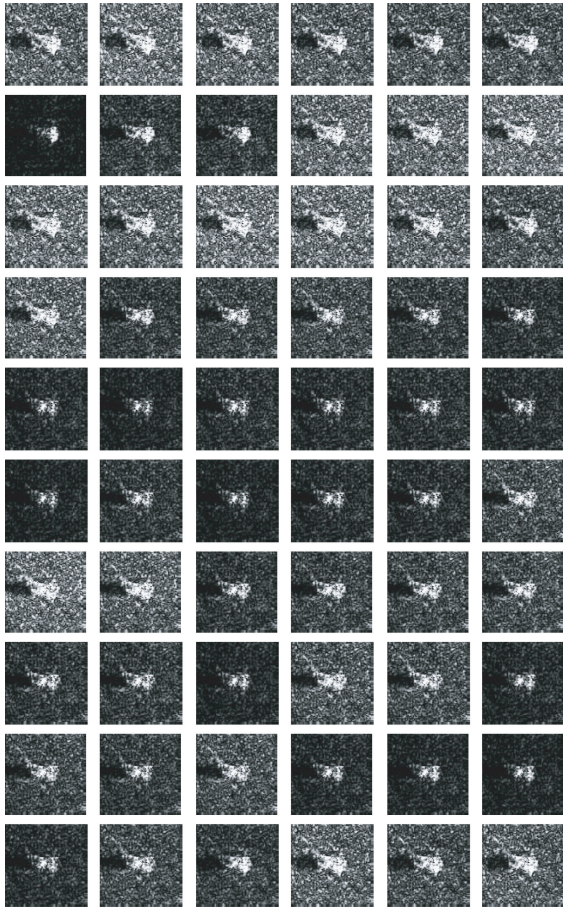


图 3 型号为 2S1 的方位角归一化 SAR 图像(前 60 个训练样本)

Fig. 3 Aspect angle normalized SAR images for 2S1 class (first 60 training samples)

本文算法在 SOC 配置下进行目标识别的混淆矩阵、平均分类精度(Average Accuracy, AA)和总体分类精度(Overall Accuracy, OA)如表 3 所示,其中每一行最右边一列代表该类别的分类精度(Percentage of Correct Classification, PCC)。在测试样本包含变体的情况下总体分类精度达到 97.97%,与目前基于深度 CNN 模型的

方法接近。在 4 个扩展操作条件 EOC1-EOC4 条件下本文算法的识别结果分别如表 4~表 7 所示。从表中可以看出,在扩展操作条件下,本文提出算法仍然取得了优异的分类性能,总体分类精度分别达到了 97.83%,91.77%,97.11%,97.04%,平均分类精度分别达到了 97.83%,91.77%,97.63%,97.19%。这说明了本文提出算法的适用性,在多种操作条件下均能取得较好的识别性能。

表 3 标准操作条件下本文算法的混淆矩阵

Tab. 3 Confusion matrix of our algorithm under standard operating conditions

	1	2	3	4	5	6	7	8	9	10	PCC/%
1	271	0	0	0	0	3	0	0	0	0	98.91
2	0	262	0	0	9	0	3	0	0	0	95.62
3	0	0	274	0	0	0	0	0	0	0	100
4	0	0	0	273	0	0	0	0	0	0	100
5	0	0	0	0	274	0	0	0	0	0	100
6	0	0	0	0	0	274	0	0	0	0	100
7	0	1	0	0	2	0	192	0	0	0	98.46
8	0	0	0	0	0	0	0	553	1	33	94.21
9	0	0	0	0	0	0	0	3	192	1	97.96
10	0	0	0	0	0	0	0	9	0	573	98.45
AA											98.36
OA											97.97

表 4 EOC1 下本文算法的混淆矩阵

Tab. 4 Confusion matrix of our algorithm under EOC1

	2S1	BRDM2	T72(A64)	ZSU234	PCC/%
2S1	285	0	0	3	98.96
BRDM2	0	287	0	0	100
T72(A64)	0	0	288	0	100
ZSU234	22	0	0	266	92.36
AA					97.83
OA					97.83

表 5 EOC2 下本文算法的混淆矩阵

Tab. 5 Confusion matrix of our algorithm under EOC2

	BMP2	BRDM2	T72	BTR70	PCC/%
T72	202	21	2487	0	91.77
AA					91.77
OA					91.77

表 6 EOC3 下本文算法的混淆矩阵

Tab. 6 Confusion matrix of our algorithm under EOC3

	BMP2	T72	BTR60	T62	PCC/%
BMP2	366	26	0	0	93.37
T72	9	377	0	0	97.67
BTR60	1	0	194	0	99.49
T62	0	0	0	273	100
AA					97.63
OA					97.11

表 7 EOC4 下本文算法的混淆矩阵

Tab. 7 Confusion matrix of our algorithm under EOC4

	2S1	BRDM2	ZSU234	PCC/%
2S1	285	0	3	98.96
BRDM2	0	419	1	99.76
ZSU234	29	0	377	92.86
AA				97.19
OA				97.04

本文算法与文献中提出算法在多种操作条件下的对比结果如表 8 所示。当采用表 2 中的训练样本进行训练时,本文提出算法的识别精度如表 8 中 OURS 所示。基于本文提出的目标方位角估计算法直接对 SAR 图像进行方位角归一化,然后提取灰度特征,利用 SVM 进行分类,该结果记做 SVM+ROTATED。文献[5]采用经过预训练的 CNN 模型作为特征提取模块,基于 SVM 训练分类器。虽然文献[5]在 SOC 场景下取得了最优的分类精度,但是该算法需要已知目标的精确方位角信息对原始图像进行校正,同时采用了表 2 中的所有训练样本,且采用了更加有效的数据增强手段。

表 8 多种操作条件下各算法的总体分类精度

Tab. 8 Overall accuracies of different algorithms under several operating conditions

算法	分类场景					
	SOC		EOC1	EOC2	EOC3(30°)	EOC4
	含变体	不含变体				
SVM+ROTATED	92.23	95.38	86.19	70.66	90.77	91.29
SVM+CNN ^[5]	99.5					
KSR ^[6]					93.90	93.81
IGT ^[7]	95		85	80		
CHU-Nets ^[3]		99.09				
CNN ^[8]	93.16					
A-ConvNets ^[9]		99.13	96.12	98.93		
Fusion-CNN ^[10]	99.42					
MultiScale-CNN ^[11]	97.69					
OURS	97.97	98.80	97.83	91.77	97.11	97.04
OURS-30	74.68	74.30	92.70	57.93	84.19	90.40

从表 8 中可以看出,本文提出算法在多个扩展操作条件下取得了最高的识别精度,而直接利用估计的方位角对训练样本进行旋转的分类精度不高,故也验证了本文采用的基于 DRGAN 的深度生成模型进行方位角归一化的有效性。为验证 DRGAN 模型的目标方位角解耦学习能力,采用表 2 中每类训练样本的前 30 个样本训练 DRGAN 模型(注:MSTAR 数据集中每 60 个样本使得目标方位角基本能够覆盖 0°~360°),并利用训练模型对这些训练样本和所有测试样本进行方位角归一化处理,目标识别结果如表 8 中 OURS-30 所示。从中可以看出,本文提出算法仅在 EOC2 的情况下识别精度不理想,在其他 4 种操作条件下,均取得了较高的识别精度。这主要是因为 EOC2 场景下测试样本包含了 T72 的 5 个不同子型号(S7, A32, A62, A63, A64),而训练样本的子型号与之不同(SN132),当没有足够训练样本的情况下,深度生成模型对样本分布的学习不是很充分,导致识别精度较低。

与基于稀疏表示和基于深度 CNN 等主流 SAR 图像目标识别方法对比结果可以看出,本文提出算法优于或者接近多数经过数据增强训练得到的 CNN 模型分类精度,要明显高于基于稀疏表示识别算法。

本文采用的 DRGAN 模型是一种深度生成模型,若能准确地估计 SAR 图像中的目标方位角,则可以利用该模型进行训练样本的增广,从而解决 SAR 图像识别中训练样本缺乏的问题,进而通过深度模型的优化设计提高 SAR 图像目标识别的准确率。

5 结 论

本文提出了一种结合 DRGAN 和 SVM 的 SAR 图像目标识别算法,采用 MSTAR 数据集在多种操作条件下验证了提出算法的有效性。由于 SAR 图像具有目标方位角敏感性,故采用

DRGAN 学习能够得到与方位无关的目标外观表示。虽然采用的目标方位角估计算法带有一定的误差,但通过设置较大的量化区间,能够削弱该估计误差对识别精度的影响。从变换后图像中提取特征,采用 SVM 进行分类器设计,能够取得与基于深度 CNN 模型相当的精度,这也验证了采用的深度生成模型的有效性。实验结果表明,本文提出算法在标准操作条件分别取得了 97.97% 和 98.80% 的识别精度,在扩展操作条件下分别取得了 97.83%、91.77%、97.11% 和 97.04% 的识别精度,能够在不进行复杂数据预处理的情况下,取得优异的识别正确率,能够满足 SAR 图像目标识别应用的需求。

参考文献:

- [1] 吴良斌. SAR 图像处理与目标识别[M]. 北京: 航空工业出版社, 2013.
WU L B. *SAR Image Processing and Target Recognition* [M]. Beijing: Aviation Industry Press, 2013. (in Chinese)
- [2] EL-DARYMLI K, GILL E W, MCGUIRE P, *et al.*. Automatic target recognition in synthetic aperture radar imagery: a state-of-the-art review [J]. *IEEE Access*, 2018, 7(7):071598.
- [3] LIN Z, JI K, KANG M, *et al.*. Deep convolutional highway unit network for SAR target classification with limited labeled training data [J]. *IEEE Geoscience & Remote Sensing Letters*, 2017, 14(7): 1091-1095.
- [4] 马世伟, 刘丽娜, 傅琪, 等. 采用 PHOG 融合特征和多类别 Adaboost 分类器的行为识别 [J]. *光学精密工程*, 2018, 26(11): 2827-2837.
MA SH W, LIU L N, FU Q, *et al.*. Using PHOG fusion features and multi-class Adaboost classifier for human behavior recognition [J]. *Opt. Precision Eng.*, 2018, 26(11): 2827-2837. (in Chinese)
- [5] WAGNER S A. SAR ATR by a combination of convolutional neural network and support vector machines [J]. *IEEE Transactions on Aerospace & Electronic Systems*, 2017, 52(6): 2861-2872.
- [6] DONG G G, KUANG G Y. SAR target recognition Via sparse representation of monogenic signal on grassmann manifolds [J]. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2016, 9(3): 1308-1319.
- [7] SRINIVAS U, MONGA V, RAJ R G. SAR automatic target recognition using discriminative graphical models [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2014, 50(1): 591-606.
- [8] DING J, CHEN B, LIU H, *et al.*. Convolutional neural network with data augmentation for SAR target recognition [J]. *IEEE Geoscience & Remote Sensing Letters*, 2016, 13(3): 364-368.
- [9] CHEN S, WANG H, XU F, *et al.*. Target classification using the deep convolutional networks for SAR images [J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2016, 54(8): 4806-4817.
- [10] CHEN S Q, ZHAN R H, HU J M, *et al.*. Feature fusion based on convolutional neural network for SAR ATR [C]. *The 4th Annual International Conference on Information Technology and Applications, ITM Web of Conferences*, 2017: 05001.
- [11] 谷雨, 徐英. 面向 SAR 目标识别的深度卷积神经网络结构设计 [J]. *中国图象图形学报*, 2018, 23(6): 928-936.
GU Y, XU Y. Structural design of deep convolutional neural networks for SAR target recognition [J]. *Journal of Image and Graphics*, 2018, 23(6): 928-936. (in Chinese)
- [12] LIU L, PAN Z, QIU X, *et al.*. SAR target classification with CycleGAN transferred simulated samples [C]. *2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE*, 2018: 4411-4414.
- [13] 杜振龙, 沈海洋, 宋国美, 等. 基于改进 CycleGAN 的图像风格迁移 [J]. *光学精密工程*, 2019, 27(8): 1836-1844.
DU ZH L, SHEN H Y, SONG G M, *et al.*. Im-

- age style transfer based on improved CycleGAN [J]. *Opt. Precision Eng.*, 2019, 27(8): 1836-1844. (in Chinese)
- [14] SAGI K, TOIZUMI T, SENDA Y. Rollable latent space for azimuth invariant sar target recognition [C]. 2018 *IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 2018: 21-24.
- [15] SONG Q, XU F. Zero-shot learning of SAR target feature space with deep generative neural networks [J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(12): 2245-2249.
- [16] WANG K, ZHANG G, LENG Y, *et al.*. Synthetic aperture radar image generation with deep generative models [J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 1-5.
- [17] LUAN T, XI Y, LIU X. Disentangled representation learning GAN for pose-invariant face recognition [C]. *International Conference on Computer Vision & Pattern Recognition*, IEEE, 2017: 1-10.
- [18] LIU J, WEI H, HUANG X Y, *et al.*. A bridge-ship collision avoidance system based on FLIR image sequences [C]. *Advances in Electrical Engineering and Computational Science*, Springer, 2009: 123-133.
- [19] BRADSKI G R. Real time face and object tracking as a component of a perceptual user interface [C]. *Proceedings Fourth IEEE Workshop on Applications of Computer Vision*, IEEE, 1998: 214-219.
- [20] CRESWELL A, WHITE T, DUMOULIN V, *et al.*. Generative adversarial networks an overview [J]. *IEEE Signal Processing Magazine*, 2018, 35(1): 53-65.

作者简介:



徐英(1982—),女,福建浦城人,讲师,2004年于西安电子科技大学获得学士学位,2011年于浙江大学获得博士学位,主要从事遥感图像目标检测与识别方面的研究。E-mail: xuying@hdu.edu.cn



谷雨(1982—),男,吉林双阳人,副教授,2004年,2009年于浙江大学分别获得学士、博士学位,主要从事多源信息融合、遥感图像目标检测与识别方面的研究。E-mail: guyu@hdu.edu.cn