

结合深度学习去噪和超分辨的 SAR 检测识别

王俊¹, 王赛¹, 任俞明¹, 陈德红², 崔闪², 魏少明¹

(1. 北京航空航天大学 电子信息工程学院, 北京 100191;

2. 上海机电工程研究所, 上海 201109)

摘要: 结合深度学习技术的发展, 利用深度卷积神经网络模型, 将特征提取和目标识别集成到一个模型中, 使系统能够自动提取目标特征并给出类别决策。对于低信噪比图像通常需要先去噪, 但是去噪能力的强弱和检测识别精度并不成正比, 本文采用的去噪方法在改善图片质量的同时有效提高了识别准确率, 避免了因细节丢失导致的精度降低。同时为了进一步提高对低信噪比图像的识别效果, 还使用了基于卷积神经网络(convolutional neural networks, CNN)的超分辨网络, 为获得更多特征信息提供条件。另外为解决非完备数据集的问题, 针对MSTAR数据集的稀疏方位角进行了一些研究, 可以在较少训练样本的情况下保持较高的识别率。

关键词: 超分辨; 特征提取; 目标识别; SAR去噪

中图分类号: TN957

文献标志码: A

文章编号: 2096-4641(2020)03-0024-07

Combining Denoising with Super Resolution for Target Detection and Recognition of SAR Image Based on Deep Learning

WANG Jun¹, WANG Sai¹, REN Yuming¹,
CHEN Dehong², CUI Shan², WEI Shaoming¹

(1. School of Electronics and Information Engineering, Beihang University, Beijing 100191, China;

2. Shanghai Electro-Mechanical Engineering Institute, Shanghai 201109, China)

Abstract: Combined with the development of deep learning, using deep convolution network model, this paper integrates feature extraction and target recognition into one model, so that the system can automatically extract target features and give category decisions. Denoising is usually needed for images with low signal-to-noise ratio, but the denoising ability is not proportional to the accuracy of detection and recognition. The denoising method adopted in this paper not only improves the image quality, but also effectively improves the recognition accuracy. It avoids the reduction of accuracy caused by the loss of details. At the same time, in order to further improve the recognition effect of low SNR images, a super-resolution network based on Convolutional Neural Networks (CNN) is used to provide conditions for obtaining more feature information. In addition, in order to solve the problem of incomplete data sets, some research on the sparse azimuth of MSTAR data sets is made, which can maintain a high recognition rate under the condition of fewer training samples.

Keywords: super resolution; feature extraction; target recognition; SAR denoise

0 引言

国内外学者研究的合成孔径雷达(synthetic aper-

ture radar, SAR)图像去噪算法可分为:基于空间滤波的去噪算法、基于变换域滤波的去噪算法和近年来越来越流行的基于深度学习的去噪算法。

收稿日期: 2020-06-18; 修回日期: 2020-08-13

基金项目: 国家自然科学基金(61671035); 重点实验室基金(6142502180103); 武器装备预研基金(6140413030511, 41413020502)

作者简介: 王俊(1972—), 男, 博士, 教授, 主要研究方向为雷达信号处理、FPGA/DSP 嵌入式系统、目标识别与跟踪。

基于空间滤波的去噪方法性能受到滤波窗口大小的影响很大,较小的窗口不能有效地抑制噪声,而较大的窗口在去噪过程中必然会导致图像纹理细节的丢失。此外,在图像的平稳区域,这些滤波算法通常保留了部分散斑噪声或去除了异质区域的弱场景信号^[1]。

近年来,借助深度学习方法去噪的研究越来越多。针对合成孔径雷达图像的斑点噪声,本文在去噪预处理过程中采用了基于深度学习的去噪方法。

传统的合成孔径雷达图像目标识别方法主要有两种,其中最常见的是基于模板匹配的方法。这种方法需要人工提取特征,实现起来较为复杂。第2种方法是基于分类器模型,主要应用有支持向量机、Adaboost 等分类器模型。该方法通过监督学习和训练分类器获得每个目标类别的后验概率,然后进行决策。

应用深度学习技术来对目标进行识别已经取得了一定的成果,本文在此基础上做了一些研究,并用MSTAR数据库中的车辆目标进行了验证。

在对合成孔径雷达图像目标检测的研究中,最常用的是双参数加拿大高等研究院(Canadian institute for advances research, CIFAR)算法,它通过背景杂波的统计特性和滑动窗口的方法,在像素级进行逐一检测。此外,结合背景杂波和目标杂波统计特性的基于广义似然比检验(generalized likelihood ratio test, GLRT)算法,结合图像处理、神经网络和隐马尔可夫模型的目标检测算法也相继问世。

目前,对基于深度学习的合成孔径雷达目标检测研究很少。2016年,Chen 结合滑动窗口和训练卷积神经网络模型进行检测^[2]。通过遍历合成孔径雷达图像的每一个区域,将其发送到人工神经网络进行识别,从而建立了一个完整的识别和检测系统。

除了上述区域建议检测和识别双阶(两级)检测器外^[3-4],还有其他类型的检测器。以 Retinanet 为代表的单级检测器模型包括特征提取网络、检测子网络和识别子网络。与以 faster R-CNN 为代表的传统两级检测器相比,其一级检测器不需要先使用区域建议网络,同时,一级检测器具有计算量小、计算速度快等优点。

本文结合一级检测器和识别网络,构建了端到端 SAR 图像预处理-目标检测识别一体化系统,研究了 SAR 图像预处理对目标检测识别的影响。深度学习模型进一步提高了目标探测和识别的准确性,为雷达目标识别、检测和跟踪提供了新的有效手段。

1 设计论证

SAR 图像的检测识别有着广泛的应用,如何快速有效地进行识别成为了关键,有很多因素都会影响到识别精度。本文提出的方案能够较大幅度地提升性能,主要分为3个部分,分别是超分辨、去噪和检测识别。

在经过超分辨和去噪的预处理之后,检测识别的精度能够进一步提升,图1为整体流程图,后面将会对每个部分进行详细介绍。



图1 整体流程图

Fig. 1 Flow chart

2 去噪分析

本文采用基于卷积神经网络的 SAR 图像去噪方法,利用扩张卷积和残差学习获得较大的感受野,避免网络层过多时效果下降,从而提高网络的性能和效率。同时,采用批归一化方法加快训练过程。总体结构如图2所示。

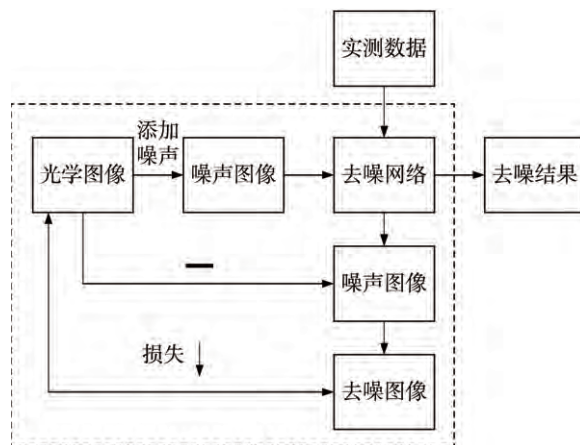


图2 去噪部分结构图

Fig. 2 Structure diagram of denoising

残差网络结构的引入提高了网络的性能,我们不对去噪图像进行训练,而是训练网络学习 SAR 图像的噪声分布特性,然后从带噪图像中减去噪声,得到去噪后的

SAR 图像。使用均方误差作为损失函数,可以表示为

$$\text{loss}(\Theta) = \frac{1}{2N} \|\varphi(y_i, \Theta) - (y_i - x_i)\|_2^2 \quad (1)$$

式中: $\varphi(y_i, \Theta)$ 表示网络学习到的噪声分布特性; $\{(y_i, x_i)\}_{i=1}^N$ 表示 N 个带噪-无噪的图像对。

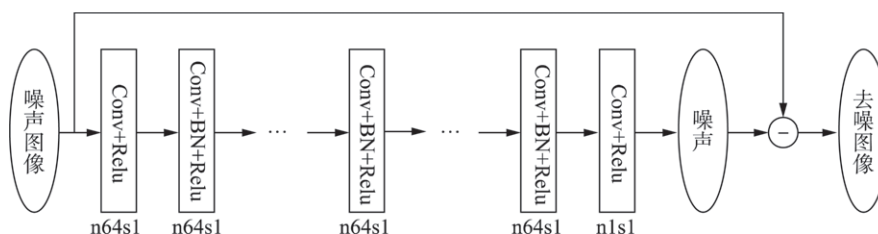


图3 DnCNN 结构

Fig. 3 Structure diagram of DnCNN

在 DnCNN 中,虽然使用了 BN 算法来归一化中间层特征,提高了网络的去噪性能,但网络感受野很小,因此采用扩张卷积的方法来解决这个问题。

扩张卷积是指在原卷积核的相邻元素之间插入一定数量的零元素,达到扩展卷积核的目的,从而扩张网络的感受野。

采用扩张卷积提取特征是为了在网络层数较少的情况下,仍可获得较大的感受野,降低计算负担以及提高运行效率。卷积感受野与层深 i 有线性关系,即

$$F_{\text{depth}-i} = (2i + 1) \times (2i + 1) \quad (2)$$

但是对于扩张卷积,其卷积感受野与层深指数相关,如式(3)所示。

$$F_{\text{depth}-i} = (2^{i+1} - 1) \times (2^{i+1} - 1) \quad (3)$$

层深相同的情况下,扩张卷积可以获得更大的感受野,从而获得更多的特征。一般来说,网络层越深,获得的特征就越多,但也会产生梯度消失问题,从而导致性能下降。为了解决这个问题,使用了跳跃连接将浅层的特征信息直接传输到更高层。这种结构可以有效地减少梯度消失的问题。同时,它能够确保细节不容易丢失。假设在网络中有 k 个跳跃连接,对于

利用残差学习的去噪卷积神经网络(denoising convolutional neural networks, DnCNN)^[5]的具体结构如图3所示。其中,Conv、BN、Relu 分别代表卷积层(convolutional layer)、批量归一化层(batch normalization layer)、非线性激活层(rectified linear unit)。

网络的第 i 层,设 $L^{(i)}$ 为输入图像,则网络的第 $(i+k)$ 层的输出为

$$L^{(i+k)} = F(L^{(i)}, \{W, b\}) + L^{(i)} \quad (4)$$

式中: W 为神经网络中的权重; b 为神经网络中的偏置。

输入和输出的关系大部分时候都是非线性的,因此,要采用非线性函数 ReLU 作为激活函数,如式(5)所示。

$$F(L^{(i)}) = \max(0, L^{(i)}) \quad (5)$$

2 图像超分辨

图像超分辨率重建可以提高检测和识别的准确性^[6]。本文在监督学习的基础上,对原始图像进行下采样,以获得成对的低分辨率和高分辨率图像,然后构建超分辨率模型,并利用配对数据集对模型进行训练,以实现 SAR 图像的超分辨重建。

基于 CNN 或生成对抗网络(generative adversarial networks, GAN)都可以建立超分辨重建模型。CNN 模型为端到端训练,损失函数一般为 $L1$ 或 $L2$ 距离。超分辨必须经过上采样。典型的 CNN 超分辨结构如图4所示。

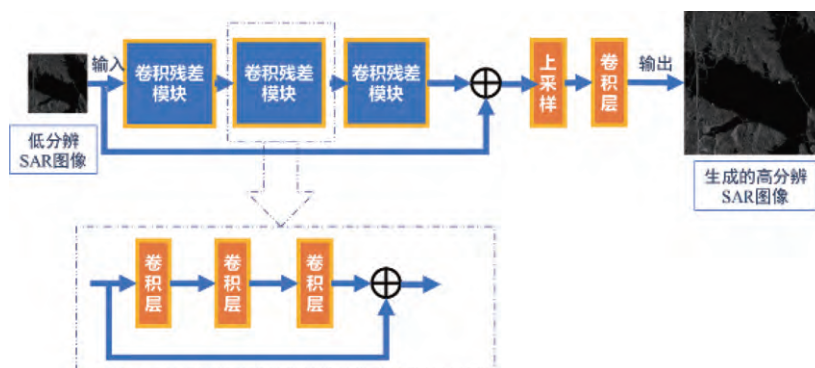


图4 CNN 超分辨结构图

Fig. 4 Super-resolution network based on CNN

基于 GAN 的超分辨率模型结构为:以 CNN 作为生成器,并增加另一个判别器来区分高分辨率图像是真实图像还是生成图像。在训练过程中,对鉴别器和

生成器进行相互训练,如图 5 所示。GAN 模型通过传统距离损失的基础上增加对抗训练来提高生成图像的真实性^[7]。

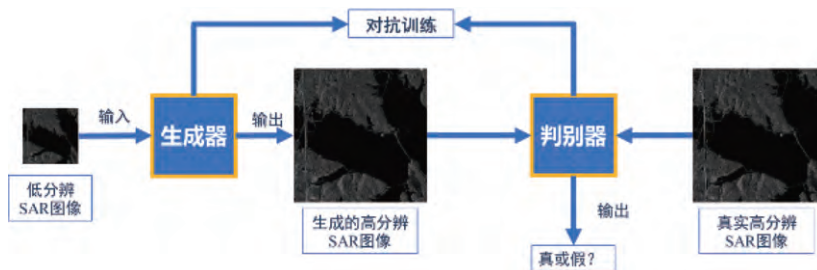


图 5 基于 GAN 的超分辨模型结构图

Fig. 5 Super-resolution network based on GAN

GAN 模型的整体损失函数为

$$l^{\text{SR}} = l_{\text{DIS}}^{\text{SR}} + a \cdot l_{\text{GEN}}^{\text{SR}} \quad (6)$$

式中: $l_{\text{DIS}}^{\text{SR}}$ 为衡量生成器输出与真实图像的差异,采用 L1 距离或 L2 距离,或其他感知损失; $l_{\text{GEN}}^{\text{SR}}$ 为对抗损失函数,分为判别器和生成器两部分; a 为加权系数,取值范围为 0~1。

基于生成器模型与判别器模型的博弈训练,网络可以学习和模仿真实图像的一些难以用数学量化表示的特征,从而提高生成图像的真实性^[8]。

3 检测识别一体化

为了进一步提高 SAR 图像目标检测和识别的精度和效率,本文结合最新的深度学习模型 Retinaet,采用一阶检测器对 SAR 图像目标进行检测识别,并通过精细的特征提取将真实目标和干扰目标区分开来,从而提高模型的抗干扰能力。一级检测器可以实现端到端的训练和预测,如果使用滑动窗口检测时,由于没有区域建议网络的预筛选,需要解决正负样本严重失衡的问题^[9]。通过损失函数可以对正负样本和目标类别进行权衡。对于多尺寸目标检测,可以采用多尺寸特征提取方法,使模型能够适应不同的大小和环境^[10]。

传统的神经网络分类方法面临两个问题:一是由于正负样本的不平衡导致网络模型难以有效地学习目标特征;二是由于“难易”样本的不均衡导致少量难样本特征无法得到有效学习。基于这两个原因,将损失函数公式改为

$$L(p_i) = -\alpha_i (1 - p_i)^\gamma \log(p_i) \quad (7)$$

式中: $p_i = \begin{cases} p, & y = 1 \\ 1 - p, & y = -1 \end{cases}$, $y \in \{1, -1\}$ 代表样本为正样本或负样本; $p \in [0, 1]$ 为神经网络预测认为该样

本属于正样本的概率。

式(7)的核心是对样本进行加权,使得不同样本对损失函数的贡献不同。与传统的交叉熵损失函数相比,损失函数增加了两个加权系数。 α_i 与 p_i 定义类似:正样本取 α ,负样本取 $1 - \alpha$, $\alpha \in [0, 1]$ 。 α_i 可用作网络的超参数,它通常与正样本的出现频率成反比, $(1 - p_i)^\gamma$ 可以根据样本分类的难度对项目进行加权。 γ 作为神经网络模型的超参数,可以调节系数对损失函数的影响。图 6 显示了不同 γ 值下, α_i 取 1 时的损失函数曲线。

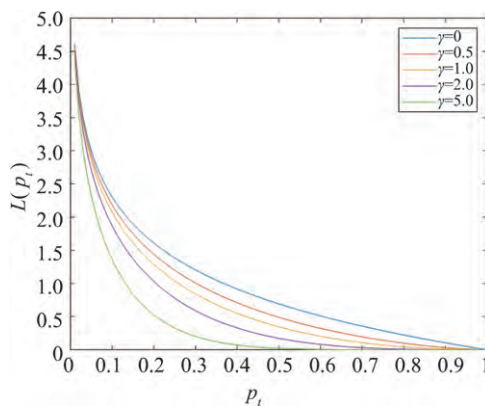


图 6 不同 γ 下损失函数随 p_i 变化值

Fig. 6 The loss function changes with p_i under different γ

在神经网络模型中,较深的隐藏层具有较好的语义,但对于目标检测和识别任务,由于分辨率较低,目标的位置信息会丢失^[11]。而浅层隐藏层包含了足够的位置信息,但难以提取有效的语义信息,不利于分类。利用多尺度特征提取和融合技术可以解决这个问题,其网络的具体结构如图 7 所示。

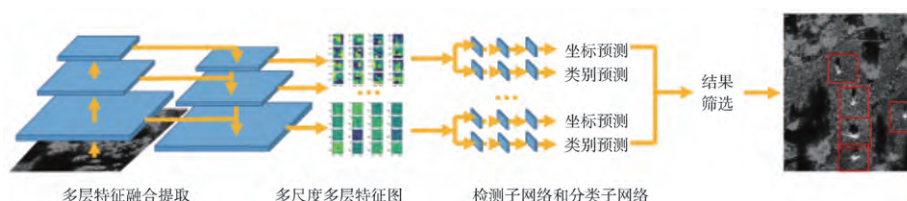


图 7 单阶检测识别一体化网络模型

Fig. 7 Network of single-order detection and recognition

特征提取网络采用特征金字塔网络结构,可提取不同大小的多层特征图像并进行特征融合。基于这些特征图,在每种尺寸的特征图上使用滑动窗口,目标分别由两个子网络检测和分类。分类任务的损失函数采用改进的交叉熵损失函数。检测的损失函数采用传统的 smooth-L1 loss,表达式为

$$L = \sum_{i \in \{x, y, w, h\}} \text{smooth}_{L1}(t_i - v_i) \quad (8)$$

式中: t_i 是网络预测的边框回归结果; v_i 是真实的边框回归结果; $\text{smooth}_{L1}(\cdot)$ 表达式为

$$\text{smooth}_{L1}(x) = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \quad (9)$$

4 实验仿真

4.1 图像去噪

训练阶段,在网络数据集的选取上,采用了 UC Merced LandUse 公开数据集中随机选取的 400 张光学遥感图像作为数据集标签,原始图像大小为 256×256 像素。

在网络训练中,网络的学习率初始化为 0.001,训练和测试过程使用 tensorflow 框架,显卡 Nvidia Tesla V100-DGXS、内存 32GB。一次输入网络的图像样本数量设置为尽可能占满 GPU 显存的数量,在本文中为 128,训练了 50 个 epochs。

为了验证神经网络在真实 SAR 图像相干斑抑制上的性能,图 8 展示了采用 DnCNN 对有噪 SAR 图像进行相干斑抑制后的处理结果以及一些传统去噪方法处理后结果的效果对比图。可以看出,传统去噪算法对图像的纹理细节保持较差,而 DnCNN 在去噪的同时保留了较多的纹理信息^[12]。

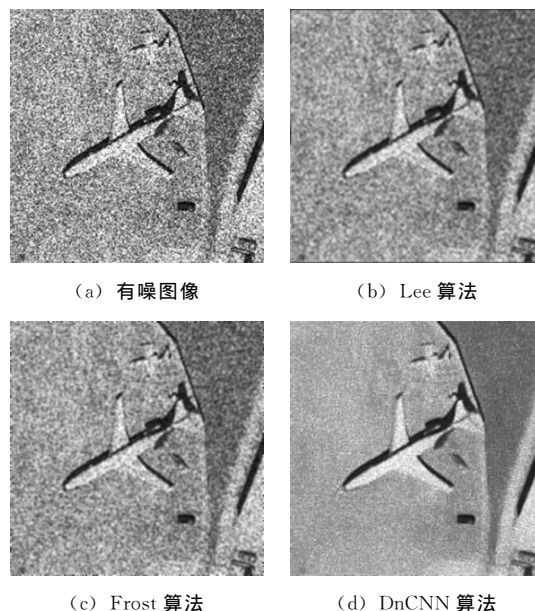


图 8 去噪效果对比图

Fig. 8 Comparison of denoising effects

不同去噪方法的边缘保持指数(edge preservation index, EPI)、峰值信噪比(peak signal to noise ratio, PSNR)指标如表 1 所示。

表 1 性能指标对比

Tab. 1 Performance index comparison

算法名称	EPI	PSNR
Lee	0.352 9	11.462 9
Frost	0.508 7	12.739 8
DnCNN	0.652 7	17.874 7

4.2 图像超分辨

首先基于 CNN 模型进行训练。数据集采用高分辨率 SAR 舰船检测数据集,因为该数据集 SAR 图像分辨率相对较高。它包含了 31 张 SAR 图像样本,每张大小为 $3\,000 \times 3\,000$ 像素。我们将每张样本分割为 500×500 像素的切片以增加样本数量,并挑选出内容较为

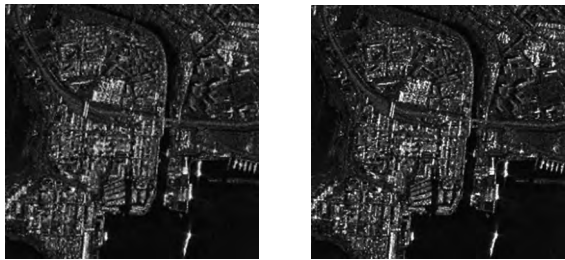
丰富的图像样本进行训练,最终得到 293 张切片。由于数据量较少,因此训练难度较大,在网络结构中采用了大量的残差模块^[13]。对于网络模型中关键的提升分辨率的上采样步骤,采用 sub-pixel convolution 算法,即将 $W \times H \times 4C$ 的特征图张量通过周期筛选重排为 $2W \times 2H \times C$ 的张量。相对于传统的上采样,该算法可以避免信息损失。为构建低分辨-高分辨匹配的数据对,对每张 500×500 像素切片进行双三次插值下采样,从而得到 250×250 像素的低分辨图像,与原始图像组成数据对,从而进行深度学习训练。采用 PSNR 和结构相似性指数(structural similarity index metric, SSIM)指标对结果进行评价,经过初步训练,得到结果如表 2 所示。

表 2 性能指标评价

Tab. 2 Performance index evaluation

算法名称	PSNR	SSIM
双三次插值法	29.77	0.80
CNN 算法	30.26	0.81

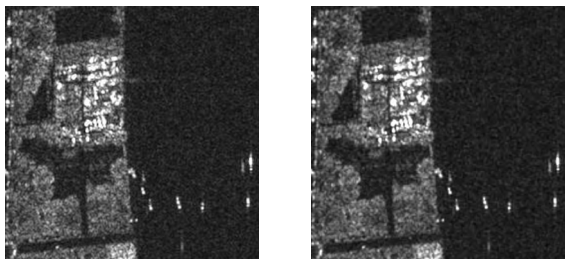
可以看出,基于 CNN 的超分辨算法要优于传统的双三次插值采样算法,图 9 为实际效果。



(a) 双三次插值采样算法 (b) 基于 CNN 的超分辨算法
图 9 超分辨效果对比图

Fig. 9 Comparison of super resolution effects

同时我们也在 SAR 舰船检测数据集(SAR ship detection dataset, SSDD)数据集^[14]上进行了测试,发现超分辨模型可以正常工作,未出现失真,证明虽然训练数据集样本数量极少,但模型仍然有较好的泛化能力,图 10 为测试结果。



(a) 低分辨率图像 (b) 超分辨率图像

图 10 SSDD 数据集测试对比图

Fig. 10 Test comparison based on SSDD data set

4.3 目标识别

使用基于 Retinanet 的多层特征提取-检测识别的单阶网络对 SAR 图像目标进行检测识别。目标检测与识别中,常用的评价指标有查准率(precision)、查全率(recall)、平均精度(average precision, AP)和平均精度均值(mean AP, mAP)^[15], AP 的值越高,分类器的分类效果越好,表 3 展示了直接训练和经过去噪再训练的平均精度。

表 3 去噪识别效果对比

Tab. 3 Comparison of recognition effects after denoising

目标型号	直接训练的 AP	去噪后的 AP
2S1	0.823 3	0.855 0
BMP2	0.835 1	0.918 7
BRDM_2	0.860 5	0.882 7
BTR60	0.844 6	0.862 6
BTR70	0.812 6	0.828 8
D7	0.859 4	0.892 9
T62	0.899 6	0.901 7
T72	0.795 2	0.872 5
ZIL131	0.911 0	0.896 5
ZSU_23_4	0.915 0	0.914 7

可以得到,未经过处理的 mAP 为 0.855 6,去噪后的 mAP 为 0.882 6。图 11 为检测识别的结果示意图。

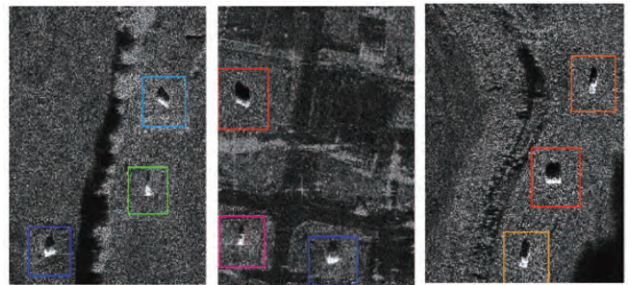


图 11 基于 Retinanet 模型的单阶 SAR 目标检测-识别结果

Fig. 11 Single-order SAR target detection-recognition results based on Retinanet model

5 结束语

本文提出了一种 SAR 图像检测识别方法,首先利用深度学习对数据进行去噪和超分辨,本文使用的去噪网络相比于传统去噪方法有更好的效果,在去噪的同时保留了大多数细节信息,避免了后续检测因去噪导致的精度降低。此外,借助基于 CNN 的超分辨网

络将 SSDD 数据集图像的分辨率提高了一倍,有效改善了图片质量,尤其对于低信噪比图像而言,相比于直接识别,对数据预处理之后可以显著提高识别效果。然后利用 Retinanet 模型,实现了目标的检测识别一体化,在实际的 SAR 图像检测识别场景中能够达到比较好的效果,为低信噪比 SAR 图像的识别提供了方法。

参 考 文 献

- [1] 刘帅奇, 扈琪, 刘彤, 等. 合成孔径雷达图像去噪算法研究综述[J]. 兵器装备工程学报, 2018, 39(12): 106—112.
- [2] CHEN S Z, WANG H P, XU F, et al. Target classification using the deep convolutional networks for SAR images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(8): 4806—4817.
- [3] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137—1149.
- [4] KANG M, LENG X G, LIN Z, et al. A modified faster R-CNN based on CFAR algorithm for SAR ship detection[C]//2017 International Workshop on Remote Sensing With Intelligent Processing (RSIP). Shanghai, China: IEEE, 2017: 1—4.
- [5] ZHANG K, ZUO W M, CHEN Y J, et al. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising[J]. IEEE Transactions on Image Processing, 2017, 26(7): 3142—3155.
- [6] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 1646—1654.
- [7] WANG L G, ZHENG M N, DU W B, et al. Super-resolution SAR image reconstruction via generative adversarial network[C]//2018 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE). Hangzhou, China: IEEE, 2018: 1—4.
- [8] LI Z, YANG J L, LIU Z, et al. Feedback network for image super-resolution[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019: 3862—3871.
- [9] BUCHER M, HERBIN S, JURIE F. Hard negative mining for metric learning based zero-shot classification[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2016: 524—531.
- [10] WEI X F, WANG X Q, CHONG J S. Local region power spectrum-based unfocused ship detection method in synthetic aperture radar images[J]. Journal of Applied Remote Sensing, 2018, 12(1): 1—5.
- [11] HAN P, ZHANG X S, GE P. Crashed airplane detection based on feature fusion in POLSAR image[C]//2012 IEEE 11th International Conference on Signal Processing. Beijing, China: IEEE, 2012: 2003—2006.
- [12] ZHANG Q, YUAN Q Q, LI J, et al. Learning a dilated residual network for SAR image despeckling[J]. Remote Sensing, 2018, 10(2): 196.
- [13] LI J W, QU C W, PENG S J, et al. Ship detection in SAR images based on convolutional neural network[J]. Systems Engineering and Electronics, 2018, 40(9): 1953—1959.
- [14] LIN T Y, GOYAL P, GIRSHICK R B, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318—327.
- [15] 蒋鑫. 基于深度学习的 SAR 图像目标识别分类研究[D]. 成都: 电子科技大学, 2018.