

## 基于卷积神经网络的 SAR 图像目标检测算法

杜 兰\* 刘 彬 王 燕 刘宏伟 代 慧  
(西安电子科技大学雷达信号处理国家重点实验室 西安 710071)  
(西安电子科技大学信息感知技术协同创新中心 西安 710071)

**摘 要:** 该文研究了训练样本不足的情况下利用卷积神经网络(Convolutional Neural Network, CNN)对合成孔径雷达(SAR)图像实现目标检测的问题。利用已有的完备数据集来辅助场景复杂且训练样本不足的数据集进行检测。首先用已有的完备数据集训练得到 CNN 分类模型,用于对候选区域提取网络和目标检测网络做参数初始化;然后利用完备数据集对训练数据集做扩充;最后通过“四步训练法”得到候选区域提取模型和目标检测模型。实测数据的实验结果证明,所提方法在 SAR 图像目标检测中可以获得较好的检测效果。

**关键词:** 合成孔径雷达; 目标检测; 卷积神经网络; 训练数据扩充

中图分类号: TN957.51

文献标识码: A

文章编号: 1009-5896(2016)12-3018-08

DOI: 10.11999/JEIT161032

## Target Detection Method Based on Convolutional Neural Network for SAR Image

DU Lan LIU Bin WANG Yan LIU Hongwei DAI Hui

(National Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China)

(Collaborative Innovation Center of Information Sensing and Understanding at Xidian University, Xi'an 710071, China)

**Abstract:** This paper studies the issue of SAR target detection with CNN when the training samples are insufficient. The existing complete dataset is employed to assist accomplishing target detection task, where the training samples are not enough and the scene is complicated. Firstly, the existing complete dataset with image-level annotations is used to pre-train a CNN classification model, which is utilized to initialize the region proposal network and detection network. Then, the training dataset is enlarged with the existing complete dataset. Finally, the region proposal model and detection model are obtained through the pragmatic “4-step training algorithm” with the augmented training dataset. The experimental results on the measured data demonstrate that the proposed method can improve the detection performance compared with the traditional detection methods.

**Key words:** SAR; Target detection; Convolutional Neural Network (CNN); Training data augmentation

### 1 引言

随着合成孔径雷达(SAR)系统应用越来越广泛,需要处理的 SAR 图像数量越来越多,对 SAR 图像处理技术提出了更高的要求,从 SAR 图像中快速并准确地检测目标,是目前的一个技术研究热点。现有的 SAR 图像目标检测方法中,双参数恒虚警(Constant False Alarm Rate, CFAR)检测算法<sup>[1-4]</sup>

应用最为广泛。该方法要求 SAR 图像中目标与背景杂波具有较高的对比度,并且假设背景杂波的统计分布模型是高斯分布,同时需要根据目标大小设置参考窗尺寸,在场景简单的情况下,检测效果较好,但是在复杂场景下检测性能较差。

卷积神经网络(Convolutional Neural Network, CNN)是人工神经网络的一种,它的权值共享网络结构降低了网络模型的复杂度,减少了权值的数量,该优点在网络的输入是多维图像时表现得尤为明显。在 CNN 中,图像作为层级结构最底层的输入,信息再依次传输到不同的层,通过这种多层的学习可以自动发现数据中的复杂结构,避免了传统算法中复杂的特征提取过程。

CNN 在图像识别<sup>[5-11]</sup>方面以惊人的速度发展的同时, Girshick 等人<sup>[12]</sup>把检测问题转化为分类问题,提出 R-CNN 框架结构,首先使用选择搜索

收稿日期: 2016-10-08; 改回日期: 2016-11-24; 网络出版: 2016-12-14

\*通信作者: 杜兰 dulan@mail.xidian.edu.cn

基金项目: 国家自然科学基金(61271024, 61322103, 61525105), 高等学校博士学科点专项科研基金博导类基金(20130203110013), 陕西省自然科学基金(2015JZ016)

Foundation Items: The National Natural Science Foundation of China (61271024, 61322103, 61525105), The Foundation for Doctoral Supervisor of China (20130203110013), The Natural Science Foundation of Shaanxi Province (2015JZ016)

(Selective Search, SS)技术提取候选区域,然后利用 CNN 模型对候选区域提取特征,将这些区域特征使用支持向量机(Support Vector Machine, SVM)分类器做分类识别,并且对候选区域做边框回归来得到更好的检测结果。针对 R-CNN<sup>[12]</sup>中存在的冗余计算、模型训练需要多步操作的问题, Girshick<sup>[13]</sup>进一步提出了 Fast R-CNN 框架结构,整合了整个检测流程,并且对每个图片只进行一次特征提取,大大减少了冗余计算,从而提高了检测速度和检测性能。Fast R-CNN<sup>[13]</sup>中提取候选区域的速度较慢,是整个检测网络的瓶颈, Ren 等人<sup>[14]</sup>又提出了 Faster R-CNN 框架结构,把得到候选区域的工作也借助 CNN 来完成,并且使得候选区域提取网络和目标检测网络共享特征提取层,取得了更好的检测性能。

CNN 在光学图像目标检测中表现出优异的性能<sup>[12-14]</sup>,但是在模型训练中需要大量的训练样本。本文利用 CNN 对 SAR 图像进行目标检测。基于 Faster R-CNN<sup>[14]</sup>基本框架,针对训练样本不足的情况,通过已有的完备数据集来辅助场景复杂且训练数据不足的数据集进行检测。基于实测数据的实验结果证明,本文方法在 SAR 图像目标检测中可以获得较好的检测效果。

## 2 算法介绍

本节对算法做具体介绍,2.1 节介绍预训练 CNN 分类模型,2.2 节介绍数据扩充方法,2.3 节介绍候选区域提取网络和目标检测网络。

### 2.1 预训练 CNN 分类模型

CNN 是一种多层网络结构,包含大量的待学习参数。这些待学习参数通常采用随机的初始化方式,这是一个很低的训练起点,因此需要大量的标记训练样本来使该网络很好地收敛,但是,获得如此多的标记训练样本在实际应用中是非常困难的。而使用训练好的 CNN 模型对新网络中的待学习参数初始化,是一个相对较高的训练起点,可以在训

练样本比较少,使得新网络更快地收敛<sup>[9,12]</sup>。

本文的主要目的是训练得到一个基于 CNN 的目标检测模型,但是只有少量训练样本,所以,可以先利用与目标检测任务无关的完备数据集训练得到分类模型,并用该模型来对目标检测模型做参数初始化,然后完成目标检测模型的训练。

本文设计的分类模型结构如图 1 所示,包括 3 个卷积层,卷积核大小依次为 11, 7, 3, 卷积核滑动步长依次为 2, 1, 1, 卷积核个数依次为 64, 128, 256。在前两个卷积层后紧接着最大池化(max pooling)层,池化窗大小为 3, 池化窗滑动步长为 2。在最后一个卷积层后跟着两个全连接层,第 1 个全连接层的维度为 256,第 2 个全连接层维度为 3,最后是 Softmax 层。所有的中间层都使用修正线性单元(Rectified Linear Unit, ReLU)做为激活函数,并且在 ReLU 层后执行批归一化(Batch Normalization, BN)<sup>[15]</sup>,网络的损失函数为负对数似然损失函数(log-loss)。

网络的输入图像尺寸固定为  $128 \times 128$ ,并且每个批次有 128 个训练样本,网络中所有卷积层和全连接层的参数都从均值为 0 标准差为 0.01 的高斯分布中随机采样得到。在整个训练过程中,所有层的学习率都保持一致。初始学习率为 0.01,调整策略是当验证误差不再降低时,就把当前的学习率除以 5,并在网络训练结束前减小 6 次。

### 2.2 数据扩充

使用 2.1 节中的分类模型对检测模型做参数初始化,可以使得 CNN 网络在少量训练样本的情况下更快地收敛,但是该模型的检测效果相对较差,本文提出一种有效的数据扩充方法,与 2.1 节中的预训练方法同时使用,可以训练得到一个更好的目标检测模型。

通常需要执行目标检测的 SAR 图像场景很复杂,包含多种干扰信息,如树木,草坪,建筑物等

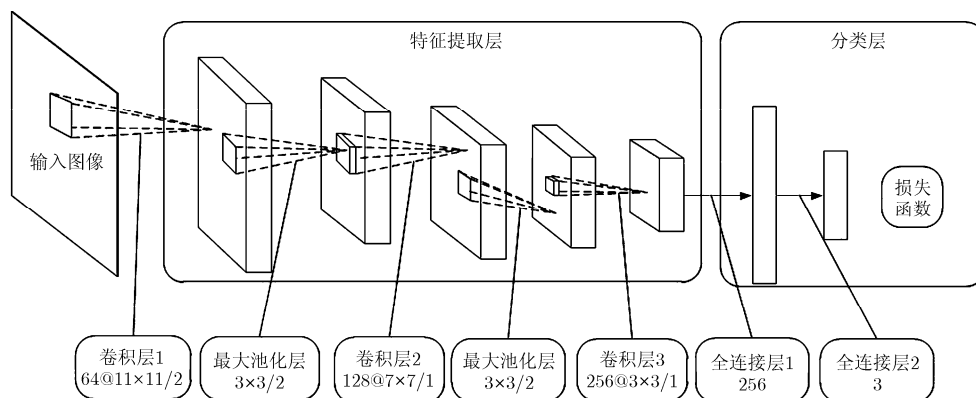


图1 分类模型

等,并且在一幅 SAR 图像中会包含多种姿态的感兴趣目标,如在本文后续实验中我们的感兴趣目标是多种姿态的车辆。想要在这样场景复杂的图像中,利用目标检测模型准确地检测到目标,对于训练数据的质量以及数量要求较高。现有的训练图像很少并且图像质量较差,只使用这些图像训练得到的目标检测模型很难达到检测任务的要求。

一些用于分类的数据集,其图像场景简单、杂波单一,且每幅图像只有单个目标,可以将这些图像处理后加入原有的训练集中,来共同训练目标检测模型。

扩充数据时需要注意的问题有:(1)两种数据集中目标尺寸存在差别,需要将新加入的图像做尺寸变换,使其目标尺寸与原有训练集中目标尺寸基本一致。新加入的训练样本如果不做尺寸变换,会对原训练集中提供的目标尺寸信息带来一定干扰,使得模型训练出现偏差,影响检测精度;(2)在进行数据扩充时,需要控制扩充数据的数量,使得模型训练可以从扩充数据中获得需要的信息,但是又不会因为扩充数据量太大,导致模型更加偏向于新加入的扩充数据,使得原有训练集中的信息被弱化。一般控制扩充训练样本数为原训练数据的2倍左右。

### 2.3 候选区域提取网络和目标检测网络

基于 CNN 的目标检测算法主要包括两个步骤:(1)从输入图像中提取候选区域;(2)使用目标检测网络对候选区域进行分类和边框回归,得到最终的检测结果。本文中,候选区域提取使用区域建议网络(Region Proposal Network, RPN)<sup>[14]</sup>,目标检测使用 Fast R-CNN 检测器<sup>[13]</sup>。下面,分别介绍 RPN 和 Fast R-CNN 检测器的具体实现。

**2.3.1 RPN** RPN 的核心思想是利用 CNN 从输入图像中直接提取候选区域,并且给每个候选区域一个判断为目标的得分。基于分类模型的基本框架来设计 RPN<sup>[14]</sup>,特征提取层与分类模型的特征提取层

保持一致,在最后一个卷积层输出的卷积特征图上滑动小网络,小网络全连接到卷积特征图  $n \times n$  的区域上,然后将该区域映射为一个低维的特征向量,再将这个特征向量作为两个同级卷积层(边框回归层和分类层)的输入。由于小网络是滑动窗口的形式,且滑动窗口的参数对于卷积特征图的所有位置都是相同的,这种结构可以通过卷积核尺寸为  $n \times n$  的卷积层实现,边框回归层和分类层使用卷积核尺寸为  $1 \times 1$  的卷积层实现,ReLU 应用于  $n \times n$  卷积层的输出,模型结构如图 2 所示。

(1)RPN 模型训练:为了训练 RPN,需要给每个参考区域分配一个标签(是/不是目标),分配正标签给两类参考区域:(a)与某个真值(ground truth)区域有最高的交集并集之比(Intersection over Union, IoU)的参考区域;(b)与任意真值 ground truth 区域有大于 0.7 的 IoU 交叠的参考区域;分配负标签给与所有真值区域的 IoU 交叠都低于 0.3 的参考区域,非正非负的参考区域不参与 RPN 模型训练。

RPN 模型使用多任务损失函数,包括分类损失函数和边框回归损失函数两部分,定义为

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{\text{cls}}} \sum_i L_{\text{cls}}(p_i, y_i) + \lambda \frac{1}{N_{\text{reg}}} \sum_i y_i L_{\text{reg}}(t_i, t_i^*) \quad (1)$$

其中,  $\lambda$  表示平衡系数,用于控制两个损失函数比重;  $N_{\text{cls}}$  表示 mini-batch 中输入样本的个数;  $N_{\text{reg}}$  表示输入的所有参考区域的个数;  $i$  表示一个 mini-batch 中的参考区域的索引;  $L_{\text{cls}}$  表示分类损失函数,选用两类(目标和非目标)的对数似然损失函数,  $p_i$  表示第  $i$  个参考区域预测为目标概率,  $y_i$  表示第  $i$  个参考区域的真实标签;  $L_{\text{reg}}$  表示边框回归损失函数,选用平滑 1-范数作为损失函数,平滑 1-范数对于离群点更加鲁棒,控制梯度的量级使得训练时不

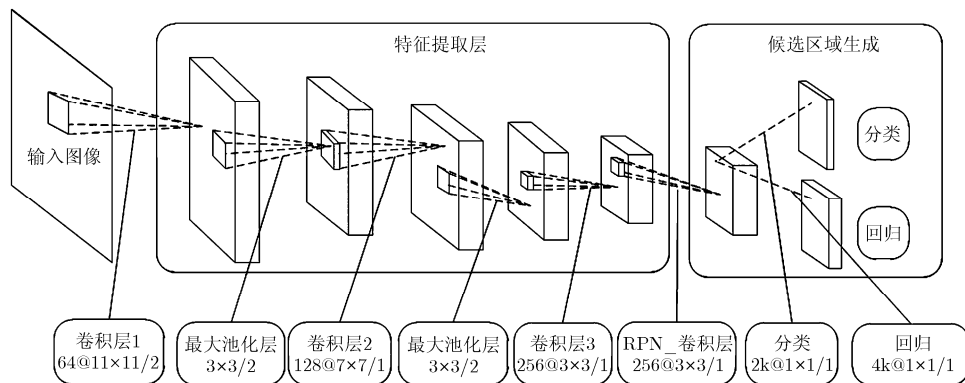


图2 RPN网络

容易跑飞； $t_i$  是一个向量，表示预测的候选区域的 4 个参数化坐标， $t_i^*$  是一个向量，表示真值区域的 4 个参数化坐标。

式(1)中的分类损失函数  $L_{cls}$  具体形式为

$$L_{cls}(p_i, y_i) = -\ln p_i^{y_i} \quad (2)$$

其中， $p_i^{y_i}$  表示第  $i$  个参考区域预测为真实标签  $y_i$  的概率值。

式(2)中的边框回归损失函数  $L_{reg}$  具体形式为

$$L_{reg}(t_i, t_i^*) = \sum_{s \in \{x, y, w, h\}} L_1(t_i^s - t_i^{s*}),$$

$$L_1(x) = \begin{cases} 0.5x^2, & |x| \leq 1 \\ |x| - 0.5, & \text{其它} \end{cases} \quad (3)$$

其中， $L_1(x)$  表示平滑 1-范数，候选区域的参数化坐标  $t_i = (t_i^x, t_i^y, t_i^w, t_i^h)$ ，真值区域的参数化坐标  $t_i^* = (t_i^{x*}, t_i^{y*}, t_i^{w*}, t_i^{h*})$ ，具体形式为

$$\left. \begin{aligned} t_i^x &= (x - x_a)/w_a, & t_i^y &= (y - y_a)/h_a \\ t_i^w &= \log(w/w_a), & t_i^h &= \log(h/h_a) \end{aligned} \right\} \quad (4)$$

$$\left. \begin{aligned} t_i^{x*} &= (x^* - x_a)/w_a, & t_i^{y*} &= (y^* - y_a)/h_a \\ t_i^{w*} &= \log(w^*/w_a), & t_i^{h*} &= \log(h^*/h_a) \end{aligned} \right\} \quad (5)$$

其中， $t_i^x, t_i^y$  表示相对于参考区域尺度不变的平移量， $t_i^w, t_i^h$  表示对数空间中相对于参考区域的宽与高，同理适用于  $t_i^{x*}, t_i^{y*}, t_i^{w*}, t_i^{h*}$ ； $x, y, w, h$  分别表示预测候选区域中心点纵横坐标 ( $x, y$ )，宽  $w$ ，高  $h$ ； $x_a, x^*$  分别表示参考区域、真值区域的横坐标，纵坐标、宽、高采取同样的定义方式。

(2)RPN 模型训练细节： 由图 2 可知，RPN 是一个全卷积网络(Fully Convolutional Network, FCN)<sup>[16]</sup>，可以通过反向传播算法和随机梯度下降法端到端训练。训练阶段，随机地抽取 256 个参考区域，构成一个 mini-batch 作为输入，其中采样的正负参考区域的比例是 1:1，如果正参考区域个数小于

128，就用负参考区域填补这个 mini-batch。

网络中所有新层的参数都从均值为 0 标准差为 0.01 的高斯分布中随机采样得到，特征提取层的参数利用分类模型中特征提取层的参数来初始化。训练模型的初始学习率为 0.001，每 10000 次迭代把当前的学习率除以 10，最大迭代次数为 40000。

通过 RPN 提取到的候选区域之间会有大量的重复，基于候选区域的分类得分，对其采用非极大值抑制(Non-Maximum Suppression, NMS)去除重复的候选区域。我们在执行过 NMS 之后的候选区域集中中选取得分最高的前  $N$  个候选区域作为 Fast R-CNN 检测器的输入。

**2.3.2 Fast R-CNN 检测器：** 基于 2.1 节中的分类模型设计 Fast R-CNN 检测器<sup>[13]</sup>，包括 3 个步骤：

(1)在最后一个卷积层输出的卷积特征图后加入 RoI 池化层，设置参数 ‘H’ 和 ‘W’ 与分类模型中第 1 个全连接层维度兼容；(2)将最后一个全连接层替换为  $n + 1$  类的分类层，并增加一个边框回归层与分类层并联；(3)修改输入数据，输入数据包括两个部分，第 1 部分是  $M$  个图像，第 2 部分是由这  $M$  个图像通过 RPN 生成的候选区域集合。Fast R-CNN 检测器模型如图 3 所示。

把图像以及该图像通过 RPN 得到的候选区域输入到 Fast R-CNN 检测器中，首先得到该图像的在最后一个卷积层的卷积特征图，然后求得每个候选区域在该卷积特征图上的映射区域，并用 RoI 池化层将这些映射区域统一到相同的大小，映射区域经过全连接层得到特征向量，特征向量分别经由两个并列的层(分类层和边框回归层)，得到两个输出向量：第 1 个是分类结果，对每个候选区域输出离散概率分布  $p = (p_0, p_1, \dots, p_n)$ ， $n$  代表总的类别个数；第 2 个是每类目标的参数化坐标，同式(4)中说明。

为了训练 Fast R-CNN 检测器，需要给候选区

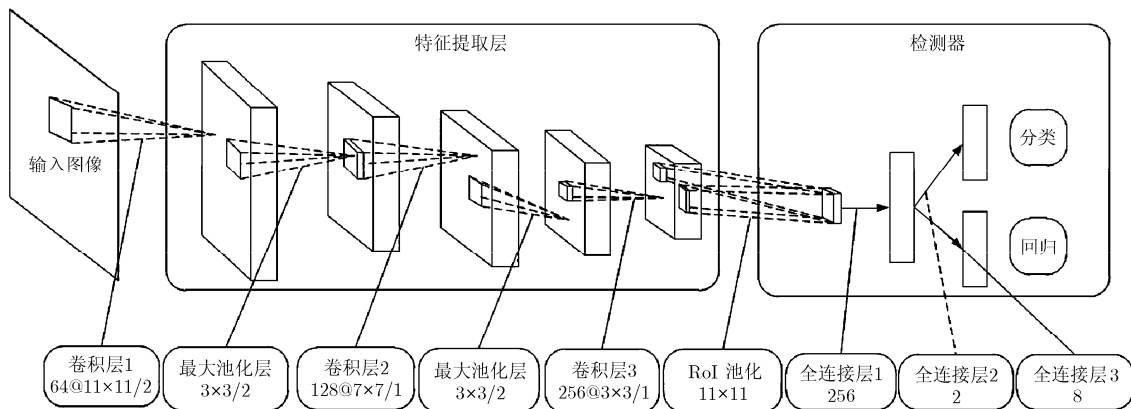


图 3 Fast R-CNN 检测器

域添加类别标签,指定与真值交叠不少于 0.5 的候选区域为该类的正样本,剩下的作为背景,模型损失函数也为多任务损失函数,与式(1)一致。

网络中特征提取层的参数利用分类模型中特征提取层的参数来初始化,训练模型的初始学习率为 0.001,每 10000 次迭代把当前的学习率除以 10,最大迭代次数为 60000。

### 2.3.3 RPN 与 Fast R-CNN 检测器共享特征提取层

到目前为止,RPN 和 Fast R-CNN 检测器都是独立训练的,根据各自任务来调整特征提取层参数,可以通过交替训练来使两个网络共享特征提取层,具体步骤如下:

第 1 步 使用预训练的分类模型对 RPN 做参数初始化,微调网络参数,得到候选区域提取模型。

第 2 步 利用第 1 步中的训练得到 RPN 对训练数据集提取候选区域,作为 Fast R-CNN 检测器的输入,使用分类模型对 Fast R-CNN 检测器做参数初始化,微调网络参数,得到 Fast R-CNN 检测器。

第 3 步 用第 2 步得到的 Fast R-CNN 检测器对 RPN 做参数初始化,并且固定特征提取层参数不变,只微调小网络中卷积层参数,得到一个更优的候选区域提取模型。

第 4 步 利用第 3 步中训练得到 RPN 对训练数据集提取候选区域,作为 Fast R-CNN 检测器的输入,并且固定特征提取层参数固定,微调 Fast R-CNN 检测器中全连接层的参数。这样,两个网络共享特征提取层,构成一个统一的检测网络。

## 3 实验结果与分析

### 3.1 数据及实验设置介绍

为了验证本文方法的有效性,在实验中使用了 MSTAR 数据集和 MiniSAR 数据集。

MSTAR 数据集中每幅图像的尺寸为  $128 \times 128$ ,包含 3 大类 7 种型号的目标,数据集对每一种型号的目标都采集了大量不同方位的图像,方位角为  $0^\circ \sim 360^\circ$ ,方位角间隔约为  $1^\circ \sim 2^\circ$ 。实验中,选择  $17^\circ$  俯仰角的目标图像作为分类模型的训练样本, $15^\circ$  俯仰角的目标图像作为测试样本。

MiniSAR 数据集中图像的尺寸为  $2510 \times 1638$ ,图像中包含多种姿态的目标,并且包含多种干扰信息,如树木,草坪,建筑物等。实验中,从每幅图像中截取图像子块作为训练图像,图像子块的短边像素尺寸约为 300,长边像素尺寸约为 500,图 4 给出了两个训练图像的示例。

由于 MiniSAR 训练集与 MSTAR 数据集中的图像尺寸和目标尺寸都不同。若直接将 MSTAR 图像

拉伸至 MiniSAR 图像的尺寸,两组图像的目标尺寸仍然存在差异,不符合 2.2 节中对扩充数据的要求。因此,我们首先在 MSTAR 图像周围填充像素,改变图像的尺寸;然后再将其拉伸,使得扩充数据集中目标尺寸与 MiniSAR 训练集中目标尺寸基本一致,图 5 给出了两个扩充图像的示例。

### 3.2 扩充数据对检测性能的影响

本节通过多组实验来验证扩充数据对检测性能的影响,为了比较的公平性,所有的检测模型训练都选用 2.1 节中的分类模型作为训练起点。检测阶段,固定 NMS 值为 0.3,将得分值大于 0.7 的切片判定为目标,采用滑窗的方式对测试图像逐块进行检测。设置 5 组实验来验证本文提出的数据扩充方法对检测性能的影响,具体如下:

**实验 1** 只使用 MiniSAR 训练集作为训练样本,训练样本个数为 40;

**实验 2** 使用 MSTAR 数据集做扩充,MiniSAR 与 MSTAR 的数据比例为 1:1;

**实验 3** 使用 MSTAR 数据集做扩充,MiniSAR 与 MSTAR 的数据比例为 1:2.5;

**实验 4** 使用 MSTAR 数据集做扩充,MiniSAR 与 MSTAR 的数据比例为 1:4;

**实验 5** 使用 MSTAR 数据集做扩充,MiniSAR 与 MSTAR 的数据比例为 1:6;

其中,测试图像如图 6 所示,图中的矩形框是人工标注的真值。

从表 1 中 5 组实验的对比结果可以看出,在未加入扩充数据时,训练样本中某些姿态的目标有缺失,因此会有较多的目标漏检。加入扩充数据后,模型对于目标的识别性能会增强,随着扩充数据的增多,检测到的目标个数显著增加,但是,由于扩充数据引入的杂波信息,会带来一定数量的虚警,在 MiniSAR 与 MSTAR 的数据比例为 1:2.5 时,检测性能到达最优。当扩充数据继续增加,虚警数量快速增加,使得模型的检测性能开始下降。

所以,添加适当比例的扩充数据可以显著提高检测性能。

针对表 1 中的检测指标做如下说明:

$$F1\text{-Score} = \frac{2 \times \text{精度} \times \text{召回率}}{\text{精度} + \text{召回率}} \quad (6)$$

$$\text{精度} = \frac{TP}{TP + FP}, \quad \text{召回率} = \frac{TP}{NP} \quad (7)$$

由于篇幅限制,只给出实验 3 的检测结果。如图 7 所示。

### 3.3 与双参数 CFAR 的检测性能比较

本节通过与双参数 CFAR 做比较来验证算法的

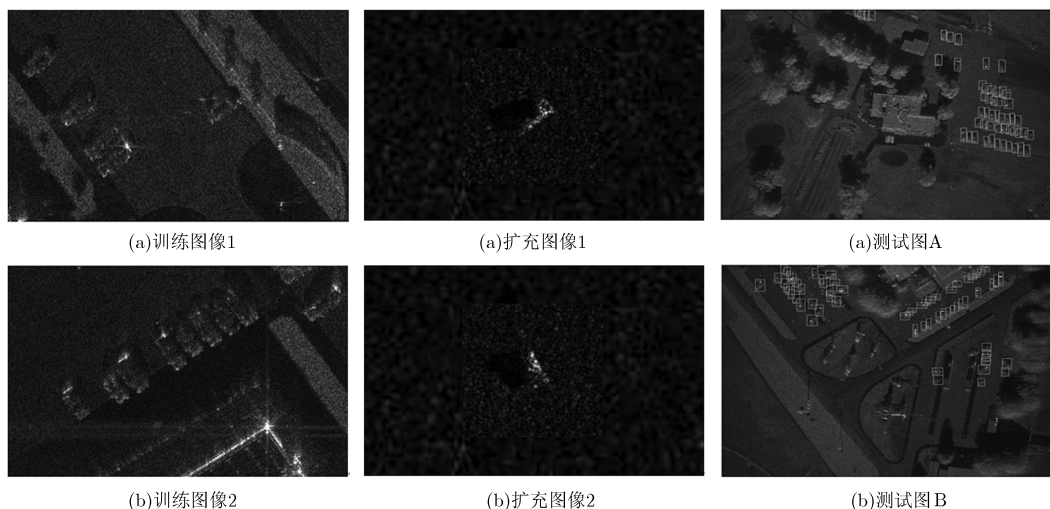


图 4 MiniSAR 训练集

图 5 MSTAR 扩充训练集

图 6 测试图像

表 1 扩充数据对检测性能的影响

测试图像及指标	图 A	图 B	精度	召回率	F1-Score
目标个数(NP)	52	64	/	/	/
实验 1					
检测到的目标个数(TP)	43	45	<b>0.8073</b>	<b>0.7586</b>	<b>0.7822</b>
虚警目标个数(FP)	10	11			
实验 2					
检测到的目标个数(TP)	44	52	<b>0.8136</b>	<b>0.8276</b>	<b>0.8157</b>
虚警目标个数(FP)	8	14			
实验 3					
检测到的目标个数(TP)	46	55	<b>0.8016</b>	<b>0.8707</b>	<b>0.8348</b>
虚警目标个数(FP)	8	17			
实验 4					
检测到的目标个数(TP)	48	56	<b>0.7429</b>	<b>0.8966</b>	<b>0.8126</b>
虚警目标个数(FP)	16	20			
实验 5					
检测到的目标个数(TP)	45	56	<b>0.7113</b>	<b>0.8707</b>	<b>0.7830</b>
虚警目标个数(FP)	18	23			

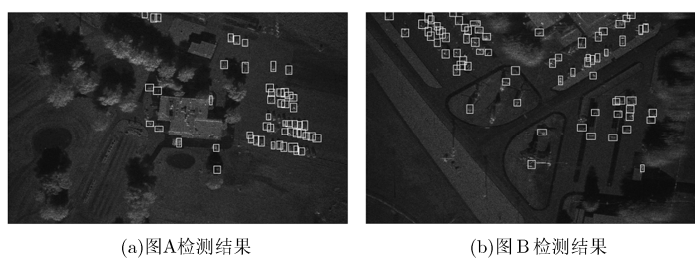


图 7 本文算法检测结果

有效性。其中，双参数 CFAR 参数设置为：保护窗长为 26，背景窗长为 28，虚警概率为  $6.87 \times 10^{-4}$ ；

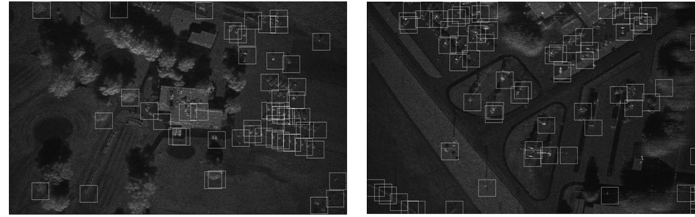
表 2 给出了两种方法的实验结果，双参数 CFAR 检测算法性能相对较差。从图 8 的检测结果可以看出，双参数 CFAR 检测结果中包含虚警数目更多，并且对于相邻目标的区分性较差，不能对每个目标得到准确的目标检测框。主要有两个方面的原因：

第一，双参数 CFAR 检测算法仅仅利用了像素的强度信息，而没有利用目标的结构信息等。单纯的像素强度能提供的信息是比较有限的，而且会对噪声比较敏感，造成了很多的虚警。第二，在进行聚类时，聚类算法将相邻目标上的区域聚为一类，使得对邻近目标的区分表现较差。

本文方法的检测性能优于传统方法，原因主要

表 2 与双参数 CFAR 检测性能对比

测试图像及指标	图 A	图 B	精度	召回率	F1-Score
目标个数 (NP)	64	64	/	/	/
本文方法					
检测到的目标个数 (TP)	46	55	<b>0.8016</b>	<b>0.8707</b>	<b>0.8348</b>
虚警目标个数 (FP)	8	17			
CFAR					
检测到的目标个数 (TP)	45	55	<b>0.6369</b>	<b>0.8621</b>	<b>0.7326</b>
虚警目标个数 (FP)	25	32			



(a)图A检测结果

(b)图B检测结果

图 8 双参数 CFAR 检测结果

在于本方法不仅仅利用了像素强度信息，还利用了目标的结构信息，可以在很大程度上减少虚警目标数。在检测过程中，RPN 可以得到高质量的候选区域，通过边框回归层可以对检测得到的候选区域进行尺寸调整，可以对邻近目标有更精准的定位。

对两种算法的计算量说明如下：

根据文献[17]，双参数 CFAR 的乘法计算量为  $P^h \cdot P^w \cdot (2Q - 1)$ ，加法计算量为  $P^h \cdot P^w \cdot (Q + 3)$ 。其中， $P^h$ ,  $P^w$  分别表示输入图像的长和宽， $Q$  表示保护窗与背景窗之间的杂波像素点个数。

CNN 网络中卷积层的乘法计算量为：

$C_i^2 \cdot K_{i-1} \cdot K_i \cdot \left(\frac{I_i^h - C_i + 1}{S_i}\right) \cdot \left(\frac{I_i^w - C_i + 1}{S_i}\right)$ ，加法计算量为  $(C_i^2 \cdot K_{i-1} - 1) \cdot K_i \cdot \left(\frac{I_i^h - C_i + 1}{S_i}\right) \cdot \left(\frac{I_i^w - C_i + 1}{S_i}\right)$ 。其中， $C_i$  表示第  $i$  层的卷积核大小， $S_i$  表示第  $i$  层卷积核的滑动步长， $I_i^h$ ,  $I_i^w$  表示第  $i$  层输入图的长和宽， $K_i$  表示第  $i$  层卷积核的个数。全连接层可以看成是卷积核大小为 1、滑动步长为 1 的卷积层。

根据本文设置的实验参数分析可知，本文方法的运算量远大于双参数 CFAR。我们对双参数 CFAR 利用 Matlab 编程，对图 6 中  $2510 \times 1638$  大小的测试图像的平均运算耗时为 13.88 s。为了提高处理速度，对本文方法使用 C++ 编程，在 CPU 模式下执行，根据实验测定，平均运算耗时为 23.63 s。为了进一步提高处理速度，我们使用 CUDA 编程，在 GPU 模式下并行处理，平均运算耗时减少至 4.36 s。可见并行运算大幅提升了运算效率，使其处理速

度可以接受。

实验使用的硬件平台，CPU: Intel Xeon E5-2620 v3 @2.4 GHz, GPU: GTX TITAN X。

## 4 结论

本文研究了训练样本不足的情况下，利用已有的完备数据集辅助检测模型的训练，借助 CNN 模型的强大特征提取能力，充分利用训练数据集提供的多种特征信息，使得基于 CNN 的检测模型可以对目标有更强的检测识别性能。基于实测数据的实验结果证明了本文方法在 SAR 图像目标检测中可以获得较好的检测效果。

## 参考文献

- [1] XING X W, CHEN Z L, ZOU H X, *et al.* A fast algorithm based on two-stage CFAR for detecting ships in SAR images [C]. The 2nd Asian-Pacific Conference on Synthetic Aperture Radar, Xi'an, China, 2009: 506-509.
- [2] LENG X, JI K, YANG K, *et al.* A bilateral CFAR algorithm for ship detection in SAR images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(7): 1536-1540. doi: 10.1109/LGRS.2015.2412174.
- [3] JI Y, ZHANG J, MENG J, *et al.* A new CFAR ship target detection method in SAR imagery[J]. *Acta Oceanologica Sinica*, 2010, 29(1): 12-16.
- [4] ELDHUSET K. An automatic ship and ship wake detection system for spaceborne SAR images in coastal regions[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1996, 34(4): 1010-1019. doi: 10.1109/36.508418.
- [5] KRIZHEVSKY A, SUTSKEVER I, and HINTON G E.

- Imagenet classification with deep convolutional neural networks[C]. Advances in Neural Information Processing Systems, Lake Tahoe, Nevada, USA, 2012: 1097-1105.
- [6] SZEGEDY C, LIU W, JIA Y, *et al.* Going deeper with convolutions[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, Massachusetts, USA, 2015: 1-9.
- [7] ZHANG X, ZOU J, MING X, *et al.* Efficient and accurate approximations of nonlinear convolutional networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, Massachusetts, USA, 2015: 1984-1992.
- [8] HE K, ZHANG X, REN S, *et al.* Deep residual learning for image recognition[C]. Computer Vision and Pattern Recognition, Las Vegas, Nevada, USA, 2016: 770-778.
- [9] ZEILER M D and FERGUS R. Visualizing and understanding convolutional networks[C]. European Conference on Computer Vision. Springer International Publishing, Zurich, Switzerland, 2014: 818-833.
- [10] GOODFELLOW I J, WARDE Farley D, MIRZA M, *et al.* Maxout networks[J]. *International Conference on Machine Learning*, 2013, 28(3): 1319-1327.
- [11] HE K, ZHANG X, REN S, *et al.* Delving deep into rectifiers: Surpassing human-level performance on imagenet classification[C]. International Conference on Computer Vision, Santiago, Chile, 2015: 1026-1034.
- [12] GIRSHICK R, DONAHUE J, DARRELL T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, Ohio, USA, 2014: 580-587.
- [13] GIRSHICK R. Fast R-CNN[C]. Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 1440-1448.
- [14] REN S, HE K, GIRSHICK R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks[C]. Advances in Neural Information Processing Systems, Montréal, Canada, 2015: 91-99.
- [15] IOFFE S and SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]. International Conference on Machine Learning, Lille, France, 2015: 448-456.
- [16] LONG J, SHELHAMER E, and DARRELL T. Fully convolutional networks for semantic segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, Massachusetts, USA, 2015: 3431-3440.
- [17] GAO G, LIU L, ZHAO L, *et al.* An adaptive and fast CFAR algorithm based on automatic censoring for target detection in high-resolution SAR images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2009, 47(6): 1685-1697. doi: 10.1109/TGRS.2008.2006504.
- 杜 兰： 女，1980年生，教授，博士生导师，研究方向为统计信号处理、雷达信号处理、机器学习及其在雷达目标检测与识别方面的应用。
- 刘 彬： 男，1992年生，硕士生，研究方向为雷达目标识别。
- 王 燕： 女，1990年生，博士生，研究方向为 SAR 图像变化检测与目标鉴别。
- 刘宏伟： 男，1971年生，教授，博士生导师，研究方向为雷达信号处理、MIMO 雷达、雷达目标识别、自适应信号处理、认知雷达等。
- 代 慧： 女，1990年生，硕士生，研究方向为雷达目标识别。