

基于双向协同训练的 PolSAR 机场跑道半监督检测方法

韩 萍, 张致峥, 周杰龙

(中国民航大学智能信号与图像处理天津市重点实验室, 天津 300300)

摘 要: 针对极化合成孔径雷达 (Polarimetric synthetic aperture radar, PolSAR) 图像跑道检测中标注数据稀缺引发的模型表征能力退化问题, 提出一种双向协同训练的半监督师生模型, 特别是设计了一个助教模块, 通过构建蒸馏损失和反馈损失进行模型联合训练, 突破传统单向蒸馏的层级限制。助教模块通过对比模型间的推理结果反馈尚未完全挖掘的特征信息, 并利用同级特征图生成方向性特征向量, 构建方向性损失辅助学生模型进行高效训练。在美国 UAVSAR 数据集上的实验结果表明, 在标注数据有限的条件下, 本文方法的跑道区域检测精度达到 83.11%, 相比于 Unet、D-Unet 和 Unet++ 系列模型分别提高了 15.63%, 6.46% 和 17.25%。

关键词: 半监督学习; 协同训练; 极化 SAR 图像; 跑道区域检测; 语义分割

中图分类号: TP391 **文献标志码:** A

A Semi-supervised Detection Method for Airport Runways in PolSAR Images Based on Bidirectional Co-training

HAN Ping, ZHANG Zhizheng, ZHOU Jielong

(Tianjin Key Lab for Advanced Signal Processing, Civil Aviation University of China, Tianjin 300300, China)

Abstract: A bidirectional co-training teacher-student framework is developed to mitigate the performance degradation caused by the scarcity of labeled polarimetric synthetic aperture radar (PolSAR) runway detection data. Within this framework, a teaching assistant module is constructed to integrate distillation loss and feedback loss. Underutilized feature representations are identified through a systematic comparison of model inferences and the generation of directional feature vectors. Experimental results demonstrate that a detection accuracy of 83.11% is achieved by the proposed method on the UAVSAR dataset, with improvements of 15.63%, 6.46%, and 17.25% being observed over the Unet, D-Unet, and Unet++ baselines, respectively.

Key words: semi-supervised learning; co-training; polarimetric synthetic aperture radar (PolSAR); runway detection; semantic segmentation

引 言

极化合成孔径雷达 (Polarimetric synthetic aperture radar, PolSAR) 能在低能见度的天气条件下获得高分辨率的地物图像, 并且包含丰富的极化信息, 广泛应用于灾害监测、资源勘探等领域^[1-3], 因此解

译 PolSAR 图像具有重要的研究价值。

近年来,深度学习在图像分类、语义分割等领域都取得了巨大的成功,其中自医学图像处理领域提出的 Unet 语义分割网络^[4]因其强大的泛用性在 PolSAR 语义分割领域中得到较好的应用^[5]。但是由于 PolSAR 图像中常常包含很多与跑道具具有相似散射特征和几何特征的河流、桥梁等干扰地物目标,以及 PolSAR 图像中固有的散斑噪声,导致检测结果存在一定程度的虚警以及在跑道边缘区域会产生漏警的情况。针对以上问题,研究人员从提升模型性能方面开展了很多相关工作。文献[6]在 Unet 网络结构基础上加入了底层空洞卷积模块进行多尺度特征融合,扩大了感受野,增强了上下文信息。文献[7]采用 ResNet101 编码器结构并设计了多重注意力机制解码器,有利于捕捉跑道区域的局部细节和全局信息。以上方法均能减少虚警和漏警的影响,提升跑道的检测精度。除此之外,标注样本不足、标签制作成本过高也成为尚待解决的难点。语义分割网络优异的性能往往依赖于大量的标注数据,尤其是在 SAR 图像等语义分割任务中,像素级的人工标注费时费力,标注的准确性也难以保障。为了应对这些挑战,很多基于样本生成的半监督学习方法开始应用于标注样本不足的语义分割任务,比如利用生成对抗网络(Generative adversarial network, GAN)通过图像合成的方法增加训练样本^[8],利用混合未标记样本生成增强样本^[9]等。上述方法虽能通过利用未标签数据中的丰富信息提升模型的泛化能力,但这些数据增强手段需要依赖能够对未标注数据做出较为理想预测的网络模型作为支撑,而在有限标注数据训练下的网络往往存在特征提取的单一性,难以从海量无标注数据中捕捉深层次语义信息。

针对以上问题,本文提出一种基于双向协同训练的半监督机场跑道检测模型,主要贡献可以概括如下。(1)设计了一种基于双向协同训练的半监督机场跑道检测模型,特别是设计了一个助教模块挖掘学生模型中的有用信息,这些特征包含着丰富的数据表示和潜在模式。通过将这些有用信息反馈给教师模型,为教师模型提供了额外的知识补充,拓宽教师模型的思路,改变了传统教师模型仅单方面输出知识的局限,从而实现了一种更加动态和交互性的知识流动机制,有助于提升整个模型的性能和对跑道区域检测的准确性。(2)基于 UniMatch 半监督模型的思想设计蒸馏损失,引入方向性损失引导学生模型更高效地学习教师模型中蕴含的知识,挖掘跑道区域的深层特征。(3)通过训练策略和无标签样本优化等方式提高检测网络的整体效果,并通过泛化性对比实验验证模型在复杂场景下的跑道检测能力。

1 相关工作

1.1 半监督学习

半监督学习是介于全监督学习和无监督学习之间的机器学习方法^[10]。它利用少量标注数据和大量未标注数据共同完成某些学习任务。利用半监督学习的方法减少了标签标注的成本,引入无标签数据减少了模型对有标签数据的依赖,可提高模型的泛化能力和预测准确性。基于伪标签^[11]和基于一致性正则化^[12]是两种最常见的半监督学习方法,其中 UniMatch 是一种结合了伪标签和一致性正则化的半监督学习方法^[13],训练框架如图 1 所示。训练过程中总损失由有监督损失和无监督损失两部分构成,有标签数据进行全监督训练与标签计算有监督损失,无标签数据分别经过弱数据增强和强数据增强处理后经过模型正向传播得到不同的预测结果,若经过弱数据增强后的分类结果超过预设阈值则将其作为强数据增强样本的伪标签,计算与其余预测结果的交叉熵损失作为无监督损失。UniMatch 使用 Wide ResNet 作为基础网络架构,包含编码器与解码器两部分,框架中弱数据增强的样本经过编码器进行特征提取后进行双分支处理,分支一直接经过解码器得到伪标签,分支二先经过特征扰动器后再经过解码器得到预测结果,特征扰动器的特征扰动策略为 Dropout。UniMatch 的优势在于其训练框架简单,训练过程便于调节,但在训练起始阶段加入无标签数据与有标签数据共同训练对生成伪标签的质

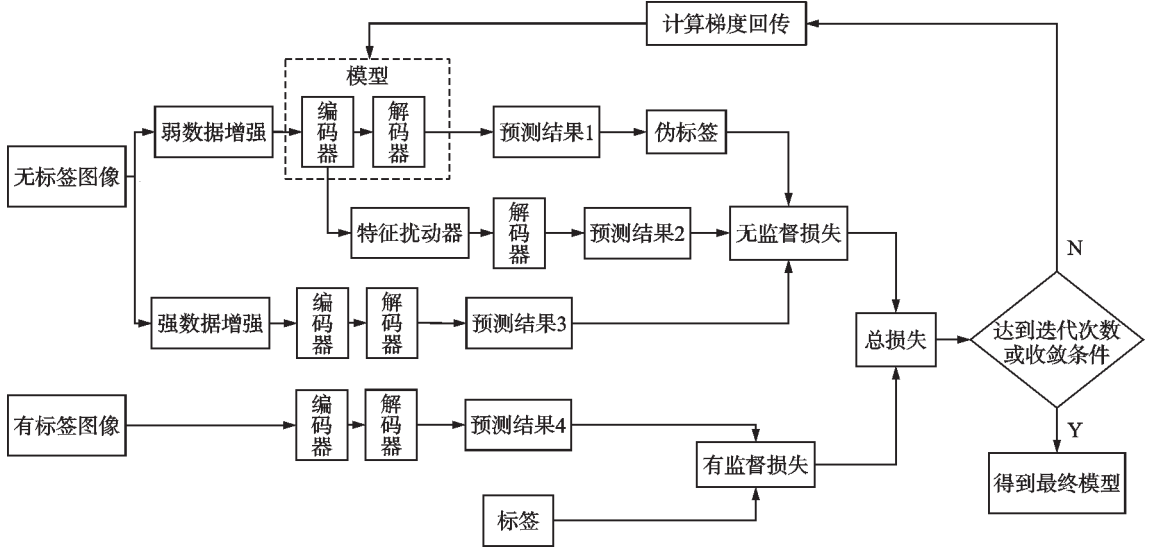


图1 UniMatch 训练框架

Fig.1 UniMatch training framework

量要求较高,模型训练早期阶段往往并不具备良好的伪标签生成能力,不准确的伪标签会影响后续模型收敛速度和最终的检测效果。

1.2 知识蒸馏

经典的知识蒸馏^[14]采用教师-学生模型,如图2所示,教师模型通常是经过预处理的相对复杂的网络,而学生模型结构相对简单。对于训练数据而言,教师模型的特征学习能力更强,将其所学到的知识通过蒸馏的方式传递给学生模型,引导学生模型的参数学习,以此提高学生模型的泛化性能。

知识蒸馏过程中的损失由标签损失和软标签损失两部分组成,标签损失是学生模型与真实标签计算交叉熵损失的结果,交叉熵损失如式(1)所示,软标签损失如式(2)所示,总损失如式(3)所示。

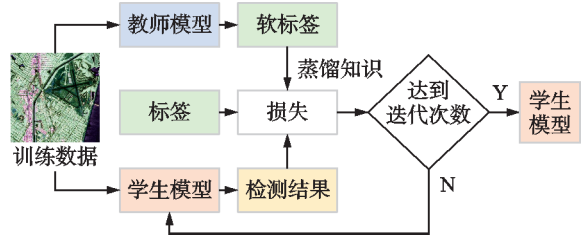


图2 教师-学生模型

Fig.2 Teacher-student model

$$L_{\text{hard}} = - \sum_{j=1}^{N_1} c_j \ln q_j, q_j = \frac{\exp(z_i)}{\sum_k^N \exp(z_k)} \quad (1)$$

$$L_{\text{soft}} = - \sum_{j=1}^{N_1} p_j^T \ln q_j^T, p_j^T = \frac{\exp(v_i/T)}{\sum_k^N \exp(v_k/T)}, q_j^T = \frac{\exp(z_k/T)}{\sum_k^N \exp(z_k/T)} \quad (2)$$

$$L = aL_{\text{soft}} + (1-a)L_{\text{hard}} \quad (3)$$

式(1~3)中: N_1 为总标签数量, c_j 表示第 j 类上的真实值,取值为0或1, q_j 为学生模型预测的属于第 j 类的概率,取值范围在 $[0, 1]$ 之间, z_i 代表学生模型对第 i 个类别的logit值; p_j^T 表示教师模型在温度为 T 时softmax输出在第 i 类上的值, q_j^T 表示学生模型在温度为 T 时softmax输出在第 i 类上的值, z_k 代表学生模型对第 k 个类别的logit值, v_i 代表教师模型对第 i 个类别的logit值; a 为软标签损失权重。

1.3 PolSAR 图像数据表述

在 PolSAR 图像中,地物目标的极化散射特性常常使用极化散射矩阵 S 来描述,表达式为

$$S = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \quad (4)$$

式中: S_{HH} 和 S_{VV} 为共极化分量,分别代表发射与接收极化方式均为水平和发射与接收极化方式均为垂直的复散射系数; S_{HV} 和 S_{VH} 为交叉极化分量, S_{HV} 代表垂直发射水平接收的极化方式的复散射系数, S_{VH} 代表水平方向发射垂直方向接收的极化方式的复散射系数。

通常使用 Pauli 分解将散射矩阵 S 分解为各 Pauli 矩阵的复数形式的加权和,用伪彩色图像描述地物信息,表达式为

$$\begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} = \frac{a}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{b}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \frac{c}{\sqrt{2}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \frac{d}{\sqrt{2}} \begin{bmatrix} 0 & -j \\ j & 0 \end{bmatrix} \quad (5)$$

$$a = \frac{S_{HH} + S_{VV}}{\sqrt{2}}, b = \frac{S_{HH} - S_{VV}}{\sqrt{2}}, c = \frac{S_{HV} + S_{VH}}{\sqrt{2}}, d = \frac{S_{HV} - S_{VH}}{\sqrt{2}} \quad (6)$$

在满足互易定理时, $S_{HV} = S_{VH}$, $d = 0$ 。此时得到伪彩色合成图,红色通道 $= |b|^2$,绿色通道 $= |c|^2$,蓝色通道 $= |a|^2$ 。在 PauliRGB 的 3 个通道中,蓝色通道通常表示各向同性奇次散射成分,它捕捉了目标在水平和垂直极化下的共同散射特性;红色通道通常表示各向同性偶次散射成分,它捕捉了目标在水平极化和垂直极化下的差异性散射特性;绿色通道则表示与水平 45° 倾角的偶次散射成分。

2 本文方法

2.1 半监督训练模型

知识蒸馏的本质是对模型能力的迁移,本文设计了一个双向协同训练模型,包含全监督训练和半监督训练两个阶段。全监督训练阶段如图 3 所示。教师模型和学生模型经过独立的全监督训练,损失函数为二元交叉熵损失,如式(7)所示。

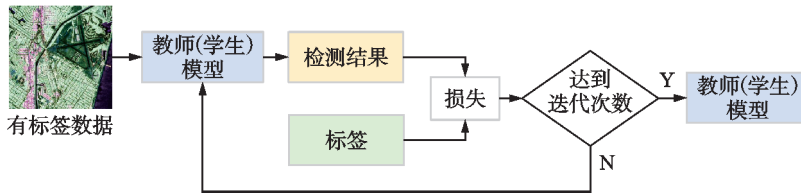


图 3 全监督训练阶段

Fig.3 Supervised training phase

$$\text{loss} = -\frac{1}{N_2} \sum_{i=1}^{N_2} [y_i \ln p_i + (1 - y_i) \ln (1 - p_i)] \quad (7)$$

式中: N_2 为模型预测的总数据量, y_i 为二元标签,取值为 0 或 1, p_i 为模型预测的属于标签 1 的概率,取值范围在 $[0, 1]$ 之间。

全监督训练阶段教师模型和学生模型从不同角度学习数据,充分利用有限的标注数据信息。经过训练后的教师模型和学生模型送入半监督学习阶段,半监督训练阶段如图 4 所示,经过全监督训练后的教师模型和学生模型都具备一定的跑道检测能力,而教师模型因其更为强大的特征提取能力能够对无标签数据中的机场跑道区域做出更准确的分割效果。与传统的知识蒸馏框架不同,由于网络结构不同,经过第一阶段的全监督训练后的学生模型也拥有一些独特的特征理解,因此本文设计了一个助教

和学生模型的下采样尺度保持一致,便于进行特征图的对比。

2.3 助教模块

助教模块的功能是整合教师模型和学生模型的信息并输出模型参数更新所需的反馈损失和蒸馏损失。在反馈损失支路中,助教模块会提取前向传播过程中教师模型和学生模型头部经过归一化处理后的概率图进行对比,通过本文设计的反馈函数计算反馈损失,反馈损失函数如式(8)所示。半监督训练阶段筛选学生模型对教师模型的正反馈有用信息需要考虑两方面的影响:(1)对于两个模型产生的检测结果分歧较大的像素点;(2)需要考虑已经能做出稳定预测的像素点对反馈损失的贡献小,预测不确定性越大的像素点对反馈损失的贡献大。

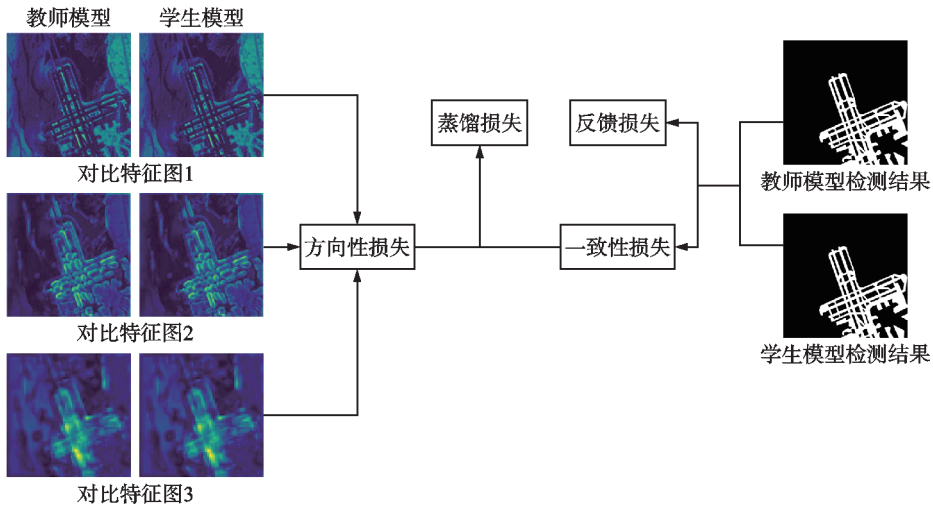


图6 助教模块

Fig.6 Assistant module

$$L_T = \frac{1}{N_3} \sum_{i=1}^{N_3} \lambda_i (p_t^i - p_s^i)^2, \quad \lambda_i = \text{Min} \left(3, \frac{1}{|p_t^i - 0.5| + |p_s^i - 0.5| + \tau} \right) \quad (8)$$

式中: N_3 为总像素点大小, p_t^i 为教师模型的softmax输出在第*i*类上的值, p_s^i 为学生模型的softmax输出在第*i*类上的值, λ_i 为像素影响因子。

用于学生模型参数更新的蒸馏损失由一致性损失和方向性损失两部分组成。在UniMatch半监督框架中使用了噪声抗性更强的弱数据增强检测结果作为强数据增强检测结果的伪标签,本文在半监督阶段中设计的一致性损失 L_a 则由教师模型输出的检测结果作为学生模型检测结果的伪标签,并计算两者的损失,损失函数为二元交叉熵损失。除此之外,本文还引入了一个方向性损失 L_b 辅助蒸馏过程。方向性损失函数如式(9)所示。助教模块分别对教师模型和学生模型各个下采样层级的特征图进行对比,将来自同一层级的特征图在特征空间中计算余弦相似程度,通过反向传播的方式使得学生模型学习的特征逐渐靠近教师模型。

$$L_b = \sum_{i=1}^n [1 - \text{Cos}(x_1, x_2)] \quad (9)$$

式中: n 为对比特征图的数量,Cos为余弦相似度函数, x_1 、 x_2 分别为教师模型和学生模型中的特征向量。

深度神经网络中特征向量存在高维空间中,这些向量的方向蕴含着数据的语义信息。从特征学习的本质来看,对教师模型绝对数值的模仿可能会误导学生模型。尤其是在处理图像分类任务时,图像

的某些特征在特征向量中的表示可能受到多种因素的影响,如光照、颜色对比度等,这些因素可能导致特征向量的绝对数值发生变化,但特征向量的方向所代表的语义关系相对稳定。学生模型仅仅追求绝对数值的模仿可能会学习到与任务无关的表面特征,而方向性的学习则能够聚焦于特征向量之间的深层一致性。

3 实验

3.1 实验数据和网络参数设置

实验数据由美国 UAVSAR(Uninhabited Aerial Vehicle SAR)卫星获取,一共 77 张 PolSAR 图像,包含 93 个机场跑道以及河流、山地、建筑等地物目标,图像大小不一,宽度均为 3 300 像素,长度为 6 000~30 000 像素不等,部分 Pauli 图像及标签图如图 7 所示。其中有标注的数据共有 31 张,利用滑窗得到 832 幅大小为 512 像素 $\times$ 512 像素的图像,按照 8:2 的比例将图像划分为训练集和验证集。无标注的数据由两部分组成,第一部分数据是剩余的 46 张原始 PolSAR 图像,另一部分是人工扩充的无标注数据集,利用滑窗得到 704 幅大小为 512 像素 $\times$ 512 像素的图像,其中人工扩充数据占 192 幅。为了提升模型特征学习的准确性,本文使用 CutMix 的思想扩充无标签数据集^[16]。

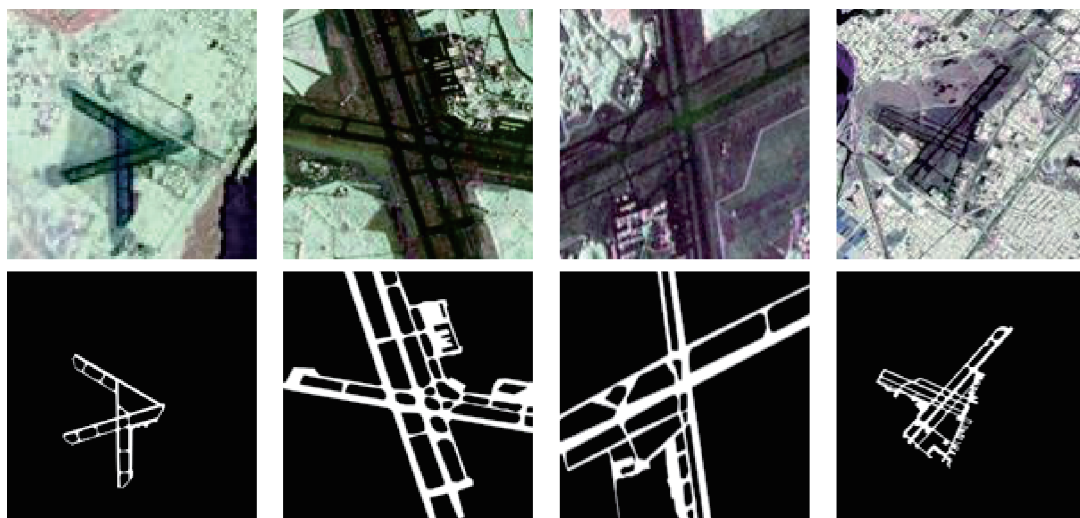


图7 原始图像数据样例(上)及标签图片(下)

Fig.7 Illustration examples of the original image data (top) and its corresponding label image (bottom)

本文采用召回率(Recall)、平均交并比(Mean intersection over union, MIoU)、像素准确率(Pixel accuracy, PA)以及 F_1 -score 共 4 个参数作为评估模型性能的指标。相关实验部署于 Pytorch 1.10 深度学习框架,使用一个 NVIDIA 3090 GPU。训练过程中使用 Adam 梯度下降算法,全监督训练阶段的初始学习率均设置为 0.001,半监督训练阶段教师模型初始学习率设置为 0.000 01,学生模型的初始学习率为 0.000 1。最大训练周期为 250,批量大小为 8。训练截止条件为验证集损失在 6 个训练周期内不再下降或达到最大训练周期。

3.2 对比实验及分析

为了验证本文所提出的双向协同训练方法的有效性,对教师模型及学生模型分别进行训练并得到如表 1 所示的对比结果。从表 1 可以看出,通过半监督训练中双向协同训练过程后教师模型和学生模型的性能都有一定提升,其中教师模型依然比学生模型有更好的检测表现,因此选择教师模型作为最

终的推理模型。

表 1 师生模型对比实验结果

Table 1 Comparative experiment of teacher-student model					%
模型	训练阶段	PA	Recall	F_1 -score	MIoU
学生模型	1	64.10	51.71	57.24	67.48
	2	68.89	65.56	66.75	70.55
教师模型	1	79.52	73.33	76.29	82.34
	2	80.35	77.61	80.14	83.11

师生模型的基础是学生模型在训练过程中趋近甚至超越教师模型的性能,对检测结果的采样分析能够清晰地可视化学生模型的学习过程。本文使用数据降维的方式分别对学生模型在全监督训练和半监督训练阶段与教师模型在同一像素位置的特征进行随机采样,结果如图 8 所示,其中 Net 1 代表学生模型,Net 2 代表教师模型。从图 8 可以发现,经过双向协同训练过程后的数据降维结果更加相近,说明教师模型在半监督阶段通过协同训练的方式可以有效指导学生模型进行训练。

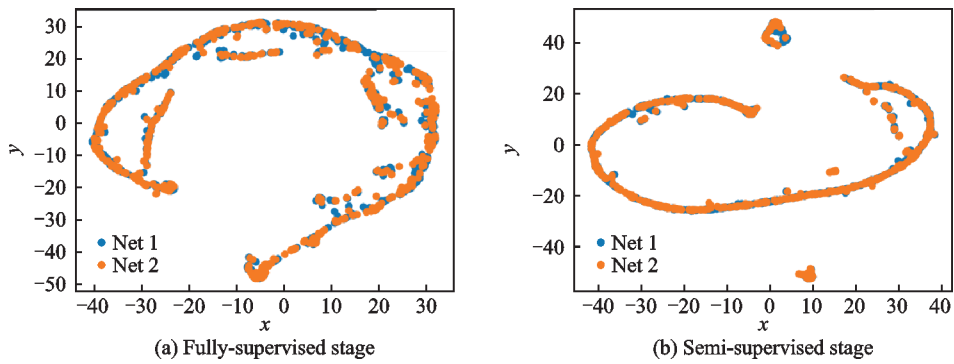


图 8 师生模型降维可视化结果

Fig.8 Visualization of dimensionality reduction results for the teacher-student model

本文通过对师生模型框架进行消融实验验证所设计的一致性损失 Loss 1,方向性损失 Loss 2,反馈损失 Loss 3 的有效性。传统的知识蒸馏框架中教师模型的权重一般不参与更新,因此消融实验以学生模型的指标反映各项损失对最终检测效果的影响,实验结果如表 2 所示。从表 2 可以看出主要指标 MIoU 呈上升趋势。

表 2 主要消融实验

Table 2 Main ablation experiment					%
模型	PA	Recall	F_1 -score	MIoU	
学生模型	64.10	51.71	57.24	67.48	
+ Loss 1	69.15	62.68	70.53	68.79	
+ Loss 2	68.71	68.19	68.09	69.39	
+ Loss 3	68.89	65.56	66.75	70.55	

为了更好地验证本文网络的跑道检测性能,本文选择现有用于跑道检测任务的 Swin-Transforme^[17]、Unet 网络、D-Unet 网络及 Unet++ 网络^[18]进行对比实验,并且为了更好地展示本文方法的检测效果,选取 3 个具有代表性的机场跑道区域进行分析。

实验1为新奥尔良机场,图像尺寸为2 000像素 $\times$ 2 000像素,其光学图、PauliRGB图以及标签如图9所示,检测结果如图10所示,检测结果中跑道区域的放大图如图11所示。如图9(b)所示,场景中跑道区域如红色框所示,此外还包含了河流、山地等容易误检为跑道的地物目标,其中山地区域如橙色框所示,河流如蓝色框所示,公路如黄色框所示。如图10(b)所示,Unet++网络无法将跑道区域与山地、普通道路等地物区分开,绿色箭头处存在大量虚警,而且如图11(b)所示滑行道和主跑道均存在漏警情况。如图10(c)所示,Unet网络相较于Unet++网络表现更好,集中表现在场景中的虚警率有明显降低,但箭头所指处存在少许虚警,总体检测效果一般。如图10(d)所示,D-Unet网络较为完整地检测出跑道区域,但箭头所指处存在少许虚警,而且如图11(d)所示滑行道处存在少量漏警。如图11(e)所示,Swin-Transformer网络可以将机场跑道区域完整地检测出,并且检测区域轮廓清晰,但是从图10(e)中可看出背景中存在较为严重的虚警情况。如图10(f)和图11(f)所示,本文方法能够将机场跑道区域与

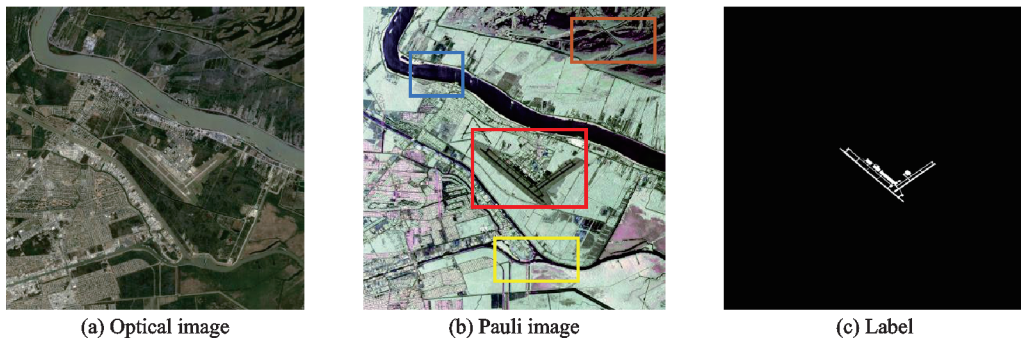


图9 新奥尔良机场跑道区域

Fig.9 Runway regions of the New Orleans International Airport

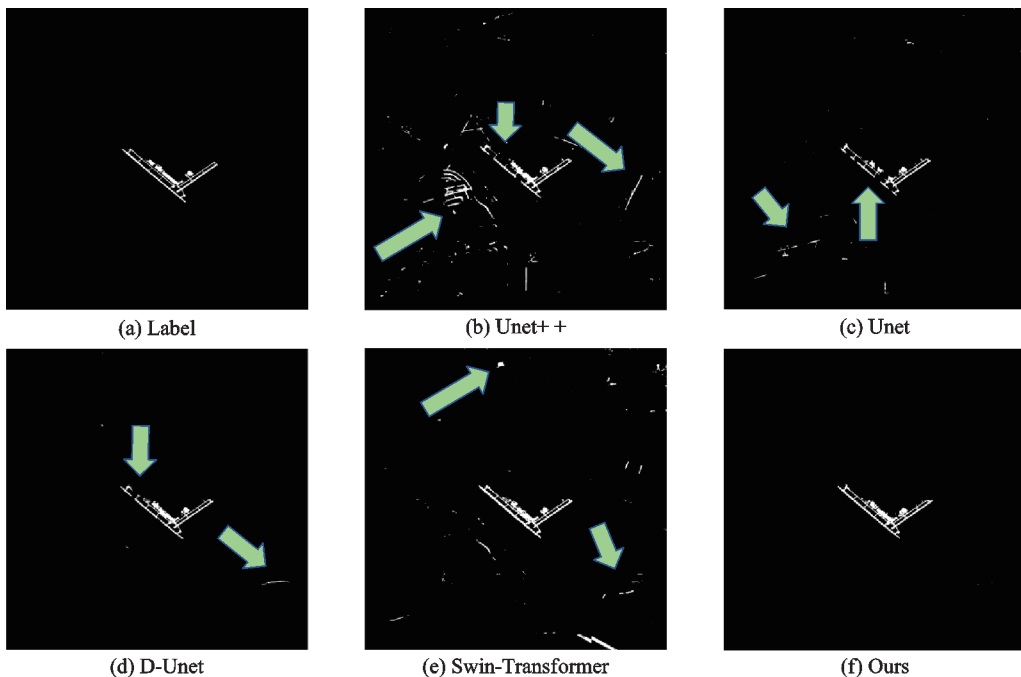


图10 新奥尔良机场场景检测结果

Fig.10 Detection results for the New Orleans airport scene

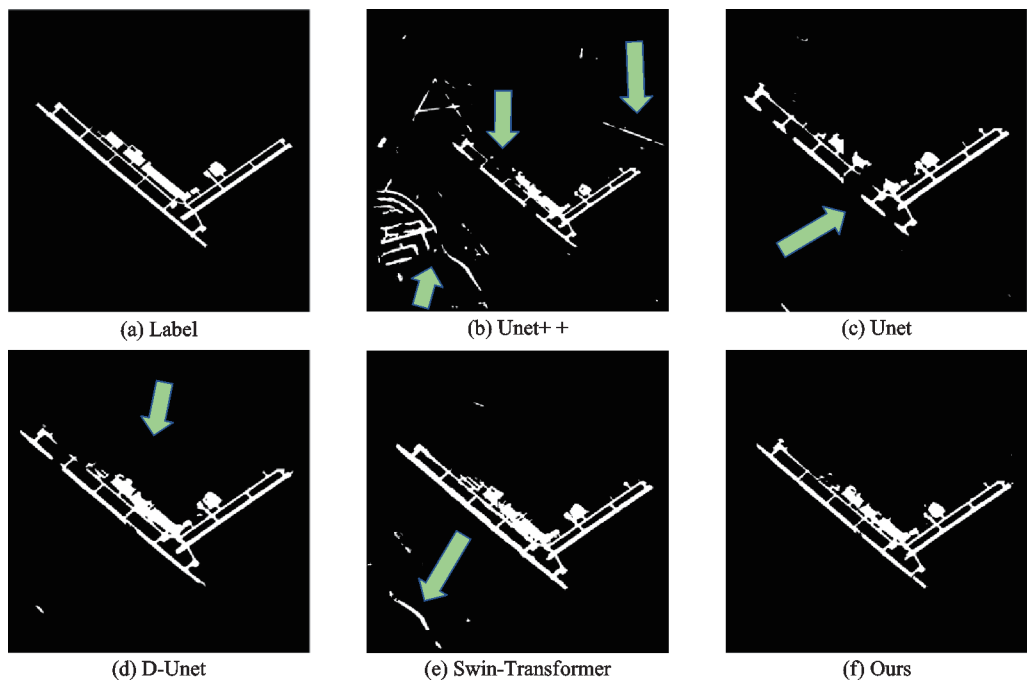


图 11 实验 1 检测结果跑道区域局部放大图

Fig.11 Partial enlarged images of runway area in detection results of Experiment 1

其余地物目标区分开,且轮廓较为完整。表 3 列出了 5 种方法在实验 1 中具体指标的定量比较。可以看出,本文方法的 PA 指标远高于 Unet++ 和 Swin-Transformer 网络,高于 D-Unet 约 3.7%,Recall 指标低于 Swin-Transformer 网络约 6.6%。

表 3 实验 1 评价指标对比

Table 3 Comparison of evaluation indices of Experiment 1				%
方法	PA	Recall	F_1 -score	MIoU
Unet++	31.99	69.68	43.85	63.53
Unet	64.48	41.49	50.50	66.66
D-Unet	75.95	74.96	75.45	80.16
Swin-Transformer	46.61	84.58	60.10	73.47
本文方法-训练阶段 1	73.56	75.65	74.59	80.95
本文方法-训练阶段 2	79.65	78.01	78.82	84.74

实验 2 为北京大兴国际机场区域,数据来自于中国高分 3 号数据集,已验证网络的泛化性能。图像尺寸为 2 560 像素×2 560 像素,图像中包含一大一小两个跑道区域,检测结果如图 12 所示。通过对图 12(c~f)的比较可以看到,Unet、D-Unet 网络在面对新的数据样本时检测性能明显下降,主要体现在漏警率上升,而且两个模型均未能检测出位于右侧较小的跑道区域,左侧较大跑道区域的检测结果也不完整。Swin-Transformer 方法能够有效检测出两个跑道区域,但依然面临着虚警率较高的问题,这与实验 1 所反映的模型特性一致。本文方法经过协同训练后检测漏警率有所下降,整体检测效果最佳。表 4 列出了 5 种方法在实验 2 中具体指标的定量比较。可以看出,本文方法的 MIoU 指标远高于 Unet 和 D-Unet,且高于 Swin-Transformer 约 3.4%;Recall 指标低于 Swin-Transformer 约 2.3%; F_1 值和 MIoU 指

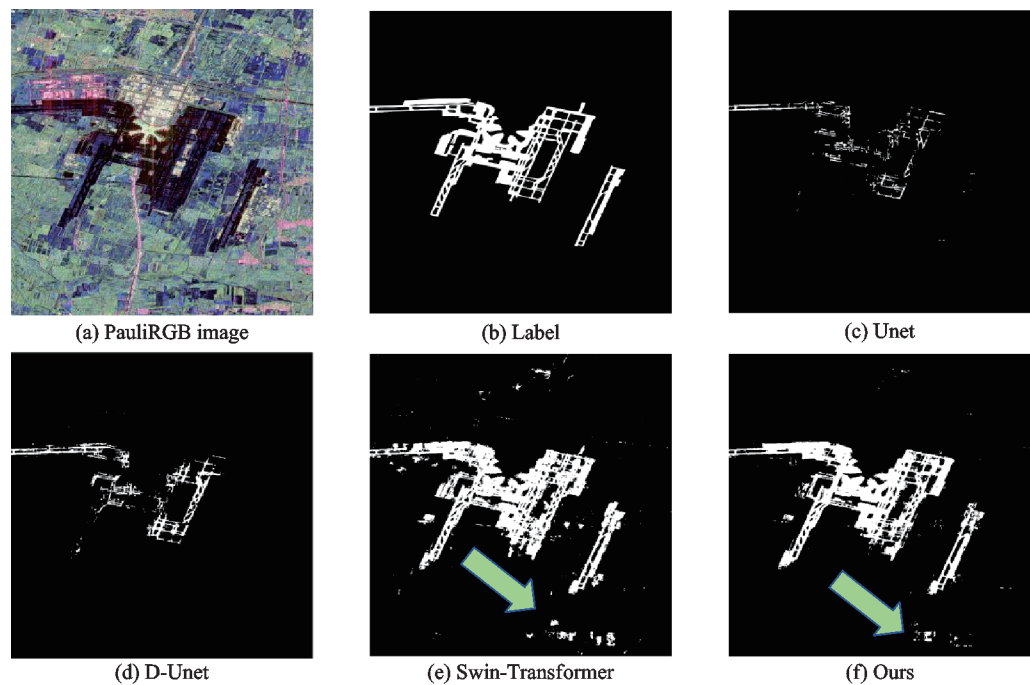


图 12 北京大兴国际机场区域检测结果

Fig.12 Detection results in the Beijing Daxing International Airport area

表 4 实验 2 评价指标对比

Table 4 Comparison of the evaluation indices of Experiment 2					%
方法	PA	Recall	F_1 -score	MIoU	
Unet	93.98	14.33	24.87	53.72	
D-Unet	89.09	24.81	38.81	58.96	
Swin-Transformer	66.67	91.45	77.12	79.12	
本文方法	74.78	89.17	81.35	82.58	

标也均为 5 种方法中最高。

本文使用海沃德机场区域验证双向知识蒸馏过程中学生模型的反馈有效性,如图 13 所示。场景中跑道区域如红色框所示,存在一大一小两个跑道区域,左侧小机场区域检测难度较大,右侧大机场区域结构相对较为复杂,对跑道检测的完整性存在一定难度。此外场景中存在山地和海洋等干扰跑道检测的地物目标,其中山地区域如橙色框所示,海洋及河流区域如蓝色框所示。网络检测结果如图 14 所示,其中 Ours-1 表示第一阶段全监督学习后的教师模型,Ours-2 表示第二阶段学生模型反馈后的教师模型。可以看出橙色山地区域对网络的影响较大,Unet 与 D-Unet 网络均在此区域造成虚警情况。此外 D-Unet 网络对右侧跑道的检测完整性较差,本文教师模型整体检测效果较好,经过半监督训练后的教师模型的虚警率更低。



图 13 海沃德机场区域 PauliRGB 图

Fig.13 PauliRGB composite of the Hayward Executive Airport region

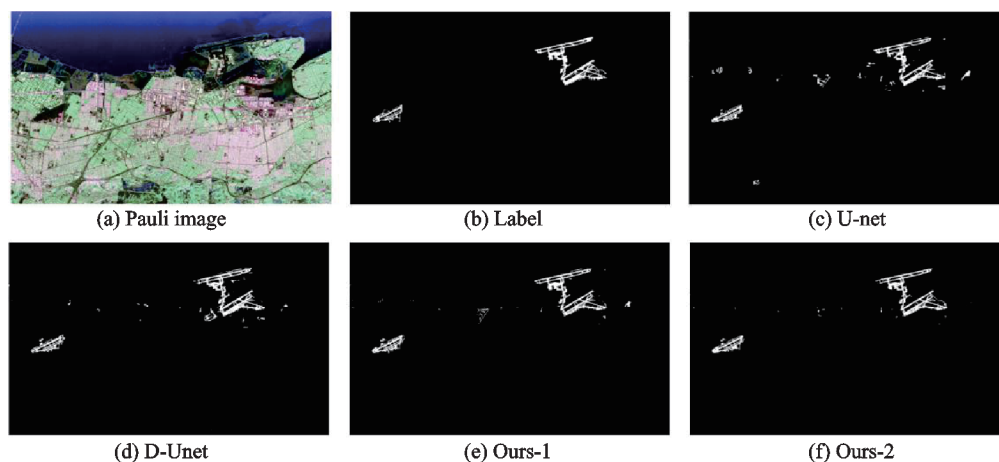


图 14 海沃德机场场景检测结果

Fig.14 Scene detection results of Hayward Airport

4 结束语

本文设计了一种基于双模型协同训练的半监督方法,通过改进模型间的交互机制与联合优化策略,实现对 PolSAR 图像跑道区域的高精度检测。该方法利用无标注数据与双模型并行训练的框架,结合特征互补与知识蒸馏机制,显著提升了模型的特征表达能力。基于美国 UAVSAR 实验数据集的对比结果表明,本文方法在 PolSAR 跑道检测任务中性能优于经典分割网络。同时,通过半监督语义分割策略,本文方法缓解了 PolSAR 图像标注样本稀缺导致的特征提取局限性,为相关领域提供了新的技术思路。当前研究聚焦于跑道区域的二分类检测,未来可进一步扩展至建筑物、植被等多类地物检测任务,验证该方法在复杂场景下的适用性。具体而言,未来工作将探索多分类任务中模型协同训练的优化方式,例如设计类别敏感的权重分配策略以解决样本不平衡问题;结合跨域自适应学习技术,增强模型对不同传感器或地理区域数据的泛化能力;尝试构建多层级特征融合架构,提升复杂场景下多类地物的判别精度。此外,还可以通过引入轻量化模型设计与自动化标注工具推动该方法的工程化应用进程。

参考文献:

- [1] LEE J S, POTTIER E. 极化雷达成像基础与应用[M]. 洪文, 李洋译. 北京: 电子工业出版社, 2013.
LEE J S, POTTIER E. Polarimetric radar imaging from basics to applications[M]. HONG Wen, LI Yang, translated. Beijing: Publishing House of Electronics Industry, 2013.
- [2] 保铮, 邢孟道, 王彤. 雷达成像技术[M]. 北京: 电子工业出版社, 2005.
BAO Zheng, XING Mengdao, WANG Tong. Radar imaging technology[M]. Beijing: Publishing House of Electronics Industry, 2005.
- [3] 张代政. 浅谈通航机场建设的重要性及重大意义[J]. 中国高新区, 2017(15): 233.
ZHANG Daizheng. On the importance and significance of navigation airport construction[J]. Science & Technology Industry Parks, 2017(15): 233.
- [4] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]// Proceedings of Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2015: 234-241.
- [5] YU L, SHAO Q, GUO Y, et al. Complex-valued U-Net with capsule embedded for semantic segmentation of PolSAR image [J]. Remote Sensing, 2023, 15(5): 1371.
- [6] HAN Pin, LIANG Yu. Airport runway area segmentation in PolSAR image based D-Unet network[C]//Proceedings of the

- International Workshop on ATM/CNS. Tokyo, Japan: Electronic Navigation Research Institute, 2022: 127-136.
- [7] CHEN L, TAN S, PAN Z, et al. A new framework for automatic airports extraction from SAR images using multi-level dual attention mechanism[J]. *Remote Sensing*, 2020, 12(3): 560.
- [8] SOULY N, SPAMPINATO C, SHAH M. Semi supervised semantic segmentation using generative adversarial network[C]// *Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV)*. Venice, Italy: IEEE, 2017: 5689-5697.
- [9] OLSSON V, TRANHEDEN W, PINTO J, et al. ClassMix: Segmentation-based data augmentation for semi-supervised learning[C]// *Proceedings of 2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*. Waikoloa, HI, USA: IEEE, 2021: 1368-1377.
- [10] VAN ENGELLEN J E, HOOS H H. A survey on semi-supervised learning[J]. *Machine Learning*, 2020, 109(2): 373-440.
- [11] LEE D H. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks[C]// *Proceedings of Workshop on Challenges in Representation Learning*. [S.l.]: ICML, 2013: 896.
- [12] TARVAINEN A, VALPOLA H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results[C]// *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach, California, USA: ACM, 2017: 1195-1204.
- [13] YANG L, QI L, FENG L, et al. Revisiting weak-to-strong consistency in semi-supervised semantic segmentation[C]// *Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver, BC, Canada: IEEE, 2023: 7236-7246.
- [14] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network[EB/OL]. (2015-03-10). <https://arxiv.org/abs/1503.02531>.
- [15] CAO M, TANG F, JI P, et al. Improved real-time semantic segmentation network model for crop vision navigation line detection[J]. *Front Plant Science*, 2022. DOI: 10.3389/fpls.2022.898131.
- [16] YUN S, HAN D, CHUN S, et al. CutMix: Regularization strategy to train strong classifiers with localizable features[C]// *Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul, Korea: IEEE, 2019: 6022-6031.
- [17] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]// *Proceedings of 2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. Montreal, QC, Canada: IEEE, 2022: 9992-10002.
- [18] ZHOU Z, RAHMAN SIDDIQUEE M M, TAJBAKHS N, et al. Unet++: A nested U-net architecture for medical image segmentation[C]// *Proceedings of International Workshop on Deep Learning in Medical Image Analysis*. Cham: Springer International Publishing, 2018: 3-11.

作者简介:



韩萍(1966-),通信作者,女,教授,研究方向:智能图像处理、SAR目标检测与识别,E-mail: hanping-cauc@163.com。



张致峥(2000-),男,硕士研究生,研究方向:基于深度学习的跑道区域检测,E-mail:15068556818@163.com。



周杰龙(2001-),男,硕士研究生,研究方向:基于深度学习的飞机目标检测,E-mail:13249787968@163.com。

(编辑:张黄群)