

机载多维度 SAR 图像地物数据集构建及融合分类方法

郑乃榕¹, 杨子安¹, 施贤正¹, 杨宏^{2,3}, 孙越², 王峰¹

1. 复旦大学 信息科学与工程学院 电磁波信息科学教育部重点实验室, 上海 200433;

2. 中国科学院空天信息创新研究院, 北京 100094;

3. 西北工业大学 电子信息学院, 西安 710129

摘要: 随着合成孔径雷达 SAR (Synthetic Aperture Radar) 成像和深度学习技术的发展, 利用深度学习算法对 SAR 图像进行地物分类得到了广泛关注及应用研究。本文基于对地观测航空遥感系统高分机载数据构建了高分辨率机载多维度 SAR 地物分类数据集 AIR-MDSAR-Map (Airborne Multi-Dimensional Synthetic Aperture Radar Mapping Dataset)。本数据集包含海南省万宁市和江苏省射阳县两个区域的 P、L、S、C 和 Ka 等 5 个波段的多极化 SAR 图像和高分辨率光学图像, 将地物划分为水域、裸地、道路、工业区、林草地、住宅区、种植地、养殖场和其他等 9 类, 并提供精细化的像素级别标签。本文首先利用深度学习中较经典的语义分割方法对 AIR-MDSAR-Map 进行地物分类验证, 同时也检验了不同波段 SAR 图像对各类地物的分类敏感性。然后, 通过不同的融合策略对多维度 SAR 数据进行融合分类, 最终融合方法在多数地物目标上的表现优于单波段, 在综合指标 FWIoU 和 PA 上提升 10%—15%, FWIoU 达 69%, PA 达 81%。该数据集将发布于国家综合地球观测数据共享平台 (<http://www.chinageoss.cn> [2022-06-21]), 能满足不同行业用户和科研用户的需求, 可为多维度 SAR 数据的应用研究提供支撑。

关键词: 遥感, 机载 SAR, 多维度, 地物分类, 深度学习, 语义分割, AIR-MDSAR-Map

中图分类号: TP75/P2

引用格式: 郑乃榕, 杨子安, 施贤正, 杨宏, 孙越, 王峰. 2024. 机载多维度 SAR 图像地物数据集构建及融合分类方法. 遥感学报, 28 (9): 2209–2222

Zheng N R, Yang Z A, Shi X Z, Yang H, Sun Y and Wang F. 2024. Airborne multidimensional SAR land cover dataset and fusion classification method. National Remote Sensing Bulletin, 28 (9): 2209–2222 [DOI: 10.11834/jrs.20242276]

1 引言

遥感是 20 世纪 60 年代兴起的一门综合交叉学科, 遥感数据的来源包括光学、多光谱、高光谱和合成孔径雷达 SAR (Synthetic Aperture Radar) 等。其中, SAR 数据由于其全天候、全天时的优势备受关注 (Moreira 等, 2013)。多维度 SAR 是其基本观测方式在极化、频率、角度和时相维度中的至少两个维度内, 分别获得多个观测量的联合观测技术手段 (吴一戎, 2013)。在 SAR 地物分类的应

用中, 单维度 SAR 数据包含的地物目标信息相对较少, 通常在不同类型地物之间存在一定程度的模糊性 (Shimoni 等, 2009), 而采用多维度 SAR 数据进行地物分类可以获取更丰富的目标信息。因此, 采用多维度 SAR 图像进行地物分类成为 SAR 图像解译的重要研究内容, 也是 SAR 技术在实际应用中的一个重要发展方向 (魏丹 等, 2020)。

SAR 地物分类在植被检测 (Addabbo 等, 2016)、土地利用 (Chen 等, 2013)、城区种植 (Lv 等, 2015) 和海冰监测 (Dierking 和 Busche, 2006) 等

收稿日期: 2022-06-21; 预印本: 2023-01-13

基金项目: 中国高分辨率对地观测专项 (编号: 30-H30C01-9004-19/21)

第一作者简介: 郑乃榕, 研究方向为合成孔径雷达图像的解译与应用。E-mail: nrzheng20@fudan.edu.cn

通信作者简介: 王峰, 研究方向为智能遥感信息获取。E-mail: fengwang@fudan.edu.cn

相关数据链接: <http://www.chinageoss.cn>

领域发挥着不可或缺的作用。传统SAR地物分类提取极化特征可以通过目标分解获得,如Pauli分解(Cloude和Pottier, 1996)、Freeman分解(Freeman和Durden, 1998)等,另一些传统的方法还可以利用边缘检测(Song和Yan, 2017)、极化散射机理(Chen等, 2013)等实现SAR地物分类。也有一些学者采用机器学习的方法,如SVM(Sukawattanavijit等, 2017)、决策树(Mishra和Singh, 2014)等。随着深度学习技术的发展,越来越多的SAR地物分类都通过深度学习方法来实现,如ResNet(He等, 2016)、DenseNet(Huang等, 2017)、VGG16(Simonyan和Zisserman, 2015)等图像分类的方法和FCN(Long等, 2015)、Deeplab(Chen等, 2018b)、注意力(Zhao等, 2022)等语义分割的方法。

相比于传统方法,使用深度学习能处理大量SAR数据并实现高精度的地物分类(吴樊等, 2022; 范吉延等, 2023)。但另一方面,深度学习对数据量要求较高,同时模型的训练需要花费较多时间(O'Mahony等, 2020)。目前,来自不同数据源的地物分类数据集相继被公开,如GID光学数据集(Tong等, 2020),覆盖中国60多个不同的城市,总面积达50000 km²; WHU-Hi数据集(Zhong等, 2020),包含了武汉市多地的高光谱图像; 雄安新区高光谱数据集(岑奕等, 2020),包含了空间分辨率为0.5,光谱范围为400—1000 nm的高光谱影像; AIR-PolSAR-Seg数据集(Wang等, 2022),包含了500张分辨率为8 m的512×512的全极化SAR图像。但是覆盖面积大、分辨率高、地物类别丰富的SAR数据集寥寥无几。可见, SAR数据相比于较容易获得、使用较广泛的光学数据(Tong等, 2020),存在数据量不足的缺点,尤其是机载多维度SAR数据获取更为困难。

本文利用高分专项航空遥感系统获取的多源机载遥感数据,构建了机载多维度SAR数据集AIR-MDSAR-Map(Airborne Multi-Dimensional Synthetic Aperture Radar Mapping Dataset),覆盖中国的海南省万宁市和江苏省射阳县的部分区域,范围约100 km²。数据集包含了P、L、S、C、Ka这5个波段的全极化SAR图像、高分辨率光学图像及其对应的9类地物覆盖类型标签。该数据集提供了同一飞行架次获取的机载多维度SAR数据,满足不同用户的研究与应用需求,可用于不同分辨率、

不同波段、不同极化等对同一地物的几何、散射特性等的分析;同时,多波段SAR图像与高分辨率光学图像互补,可用于分析各波段图像对不同类别地物的分类敏感度;另外,可以根据该数据集多波段多极化的特性研究融合地物分类方法。

在构建数据集的基础上,本文采用深度学习中经典的语义分割方法,如UNet(Ronneberger等, 2015)、SegNet(Badrinarayanan等, 2017)、Deeplab、HRNet等进行评估,并将分类效果较好的HRNet(Wang等, 2021; Sun等, 2019)作为baseline,分别对两个区域的各波段SAR数据集进行进一步实验验证,探究不同波段对不同地物的敏感性,在此基础上,提出由多个波段学习得到的神经网络模型的模型融合方法及多个波段分类结果融合的先验融合方法,融合各波段在不同地物上的分类优势,以期待为多维度SAR地物分类提供思路。

2 多维度SAR地物分类数据集

为得到最终多维度SAR数据对地物分类的最优结果,本文在原始飞行试验获得的数据生成产品图的基础上,首先,对原始数据进行预处理,优化图像质量并生成对应标签;其次,利用神经网络提取地物目标特征,并融合多尺度地物特征对模型进行训练;然后利用训练好的模型权重对测试集进行测试,分析各波段数据对不同地物目标的分类能力;最后,尝试多种融合策略,实现多维度SAR图像地物高精度分类。

2.1 成像数据获取及处理

原始数据由高分专项航空遥感系统获取,平台为改装的新舟60遥感飞机,按标准数据生产流程,相对辐射精度为1 dB,绝对辐射定标精度为2.5 dB,天线极化隔离度优于-25 dB。极化定标精度的幅度不平衡优于1 dB,相位不平衡优于15°,极化隔离度优于-30 dB。成像区域位于海南省万宁市和江苏省射阳县,每个区域约50 km²,总面积约100 km²。每个区域均含P、L、S、C、Ka, 5个波段的全极化SAR数据和高分辨率光学数据。其中,多维度SAR的入射角为45°。海南省万宁市的光学数据获取时间是2021年3月28日,多维度SAR数据于2020年12月25日获取;江苏省射阳县的所有数据获取时间均为2021年3月14日。

多维度SAR原始回波数据经数据成像处理、

辐射校正、极化校正和几何校正等预处理后，形成量化的SLC数据，再经过区域网平差和正射校正等几何处理，生成SAR DOM产品。具体流程如图1所示。

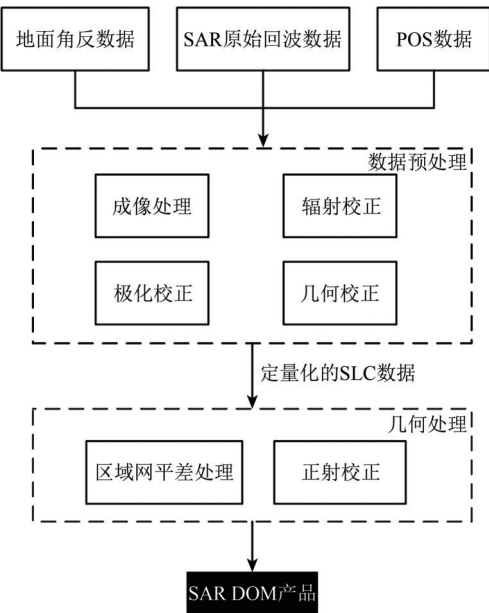


图1 SAR图像生成流程
Fig. 1 The generation process of SAR images

光学载荷飞行获取的数据经图像增强、拼接、粗校正等预处理后，根据空三结果对影像进行密集匹配生成DSM数据，之后进行半自动滤波生产DEM。DEM数据对影像进行正射纠正处理，并在人机交互模式下对影像进行色彩调整、局部修正等操作，最终按图幅输出DOM标准正射影像产品。具体流程如图2所示。

最终的SAR图像是由HH极化、HV极化、VV极化分别作为RGB3个通道而生成的伪彩色图像。原始数据信息见表1。

2.2 地物标签生成及分析

本数据集的标签图像由人工标注生成，利用MATLAB的imageLabeler工具进行标注。参照《土地利用现状分类》国家标准一级层次体系（陈百明和周小萍，2007），结合两个研究区域的地物类型，利用5个波段的全极化SAR图像、高分辨率光学图像和地图对地物目标进行类别判断与标注。将地物类型划分为9类，每类定义及标注完成后的颜色定义如表2所示。为提升标签生成效率，辅助采用半自动标注方法（Zheng等，2023）生成。先由模型进行初步预测，将预测结果作为真值扩充

数据集对模型进行更新，再结合人工校正，如此不断迭代生成所有数据的精细化标签。两个区域最终的SAR图像、光学图像及其对应的标签如图3所示。

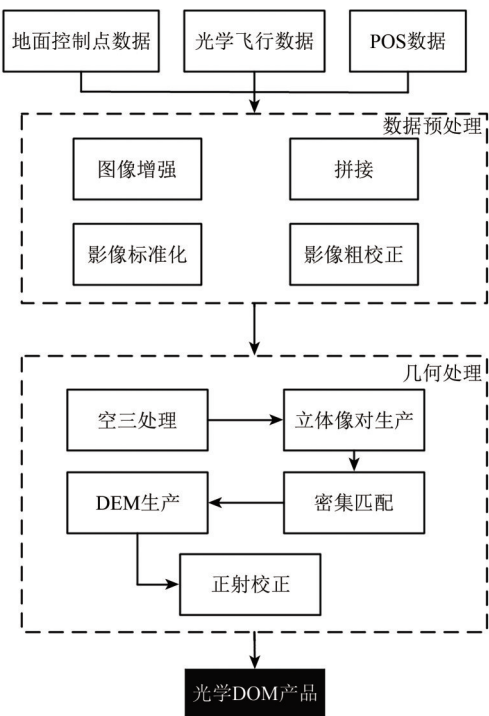


图2 光学图像生成流程
Fig. 2 The generation process of optical images

表1 原始数据信息
Table 1 The information of raw data

数据区域	波段	分辨率	飞行高度/m
海南省万宁市	光学	0.2	2160
	P	1.0	4800
	L	1.0	4800
	S	1.0	4800
	C	0.5	4800
	Ka	0.3	4800
江苏省射阳县	光学	0.2	4000
	P	1.0	4800
	L	1.0	4800
	S	1.0	4800
	C	0.5	4800
	Ka	0.3	4800

从图3可以看出，住宅区与林草地混杂在一起，在训练过程中有利于深度学习分类器的学习，不会出现对某个类别特别偏重的情况。而种植地大部分集中在一块，在深度学习样本切片制作中容易出现整个切片均为同类情况，使分类器在

学习过程中出现类别权重失衡，因此建议可以适当丢弃该类样本。同时，由于地物目标分布的特点，道路占比较少，同时形状狭长，在切片制作过程中可以适当对包含道路的切片进行数据增强。

表 2 类别定义及其对应的 RGB 值
Table 2 Classes definitions and the corresponding RGB values

类别	描述	RGB
水域	河流、水渠、湖泊、水库、坑塘、海面等水体区域	(0, 0, 255)
裸地	泥土地、沙质地、砾石地、岩石地等裸露地表	(139, 0, 0)
道路	铁路、公路、城市道路等典型大马路	(83, 134, 139)
工业区	小、中、大型工厂房屋区域	(255, 0, 0)
林草地	自然林地、人工林地、灌木丛、天然草地、人工草地等	(0, 255, 0)
住宅区	居民楼、办公楼、商场、行政、小平房等非工业区房屋区域	(205, 173, 0)
种植地	果园、茶园、菜园等耕地区域	(139, 105, 20)
其他类	不好区分的类别，如广场等	(178, 34, 34)
养殖场	养殖场	(0, 139, 139)

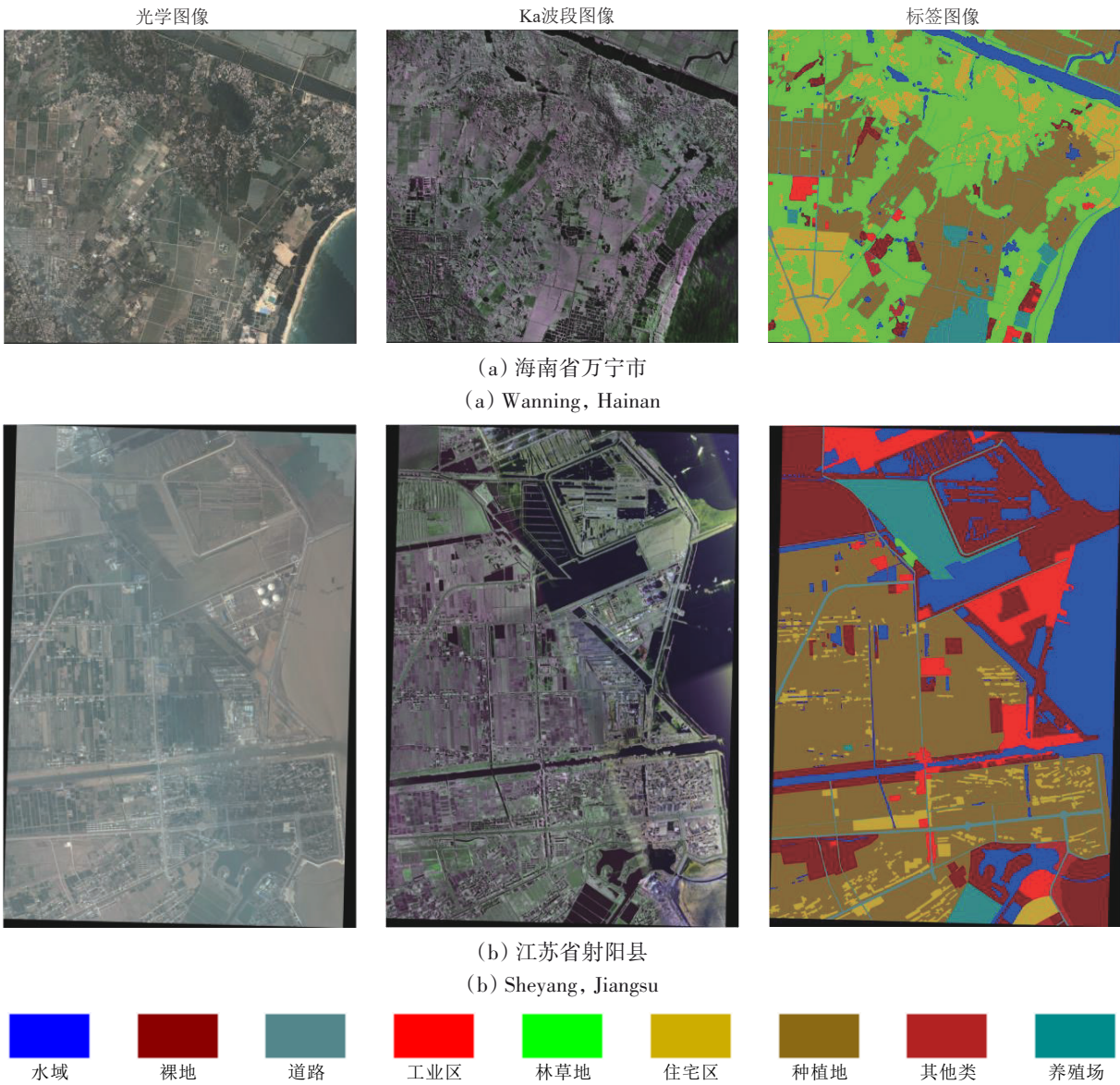


图 3 两个区域经处理后的光学图像、Ka波段图像及其对应标签
Fig. 3 The optical images, Ka band images, and the corresponding labels of the two regions after processing

2.3 数据集切片制作及分析

AIR-MDSAR-Map 数据集为满足不同研究需求，以切片形式发布。由于 AIR-MDSAR-Map 包含的是同一时间、同一地理位置、同一区域范围获取的数据，因此不同分辨率对应的图像大小也不同。为了对齐不同波段的同一切片区域，本数据集以 C 波段为基准，设置切片大小为 1024×1024，不够整除的部分用黑边补齐，其余波段及光学数据的切片大小对应 C 波段进行等比放缩，不够整除部分亦用黑边补齐。切片数量统计如表 3 所示。

表 3 对 AIR-MDSAR-Map 数据集进行了统计展示，随着后续飞行试验的继续进行，本数据集计划同步进行更新扩充。

基于数据分布的角度考虑，本文从每个区域各一景图像中截取测试集样本，Ka 波段选取的测

试集位置信息如表 4 所示。图 4 展示了选取的测试区域在标签图像上的缩略图，以及每个类别在该区域所有数据中及测试集上的所占比例。

表 3 AIR-MDSAR-Map 切片数量统计
Table 3 The slices number statistics of AIR-MDSAR-Map

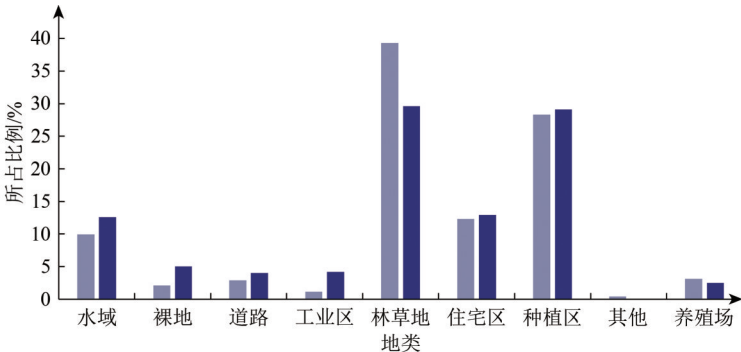
波段	切片尺寸	切片数量
P	512×512	565
L	512×512	565
S	512×512	565
C	1024×1024	565
Ka	1707×1707	565
光学	2560×2560	565
标签	1024×1024	565
总计		3955

表 4 选取的两个测试区域信息
Table 4 The information of two selected regions

区域	图像尺寸(Ka 波段)	选取的测试区域左上角坐标 $[x,y]$	每个区域大小(Ka 波段)	面积所占比例/%
海南万宁 1	15537×17423	$[1398, 6158], [2737, 10477], [7173, 9418], [12947, 12217], [13426, 1279], [4295, 3539]$	2048×2048	9.30
江苏射阳 1	20413×14306	$[1198, 7995], [3620, 449], [5268, 13217], [2996, 16965], [9412, 17864], [3545, 3447], [9886, 10343]$	2048×2048	10.05

由表 4 和图 4 中可见，两个地区分别选取的用来当作测试集类别占比与该区域数据的类别占比基本一致。保证了神经网络的训练集和测试集数据分布的一致性。同时，两个地区的测试集样本数量占比约为 10%，也符合神经网络对训练集和测试集的比例要求。其他所有波段及光学图像

的数据集划分方法均同 Ka 波段，根据分辨率的不同进行等比例放缩。最终，两个区域的 P、L、S、C、Ka 波段图像和光学图像及其所对应的标签图像经划分之后，共同构成本文实验所使用的训练集与测试集。



(a) 海南省万宁市
(a) Wanning, Hainan

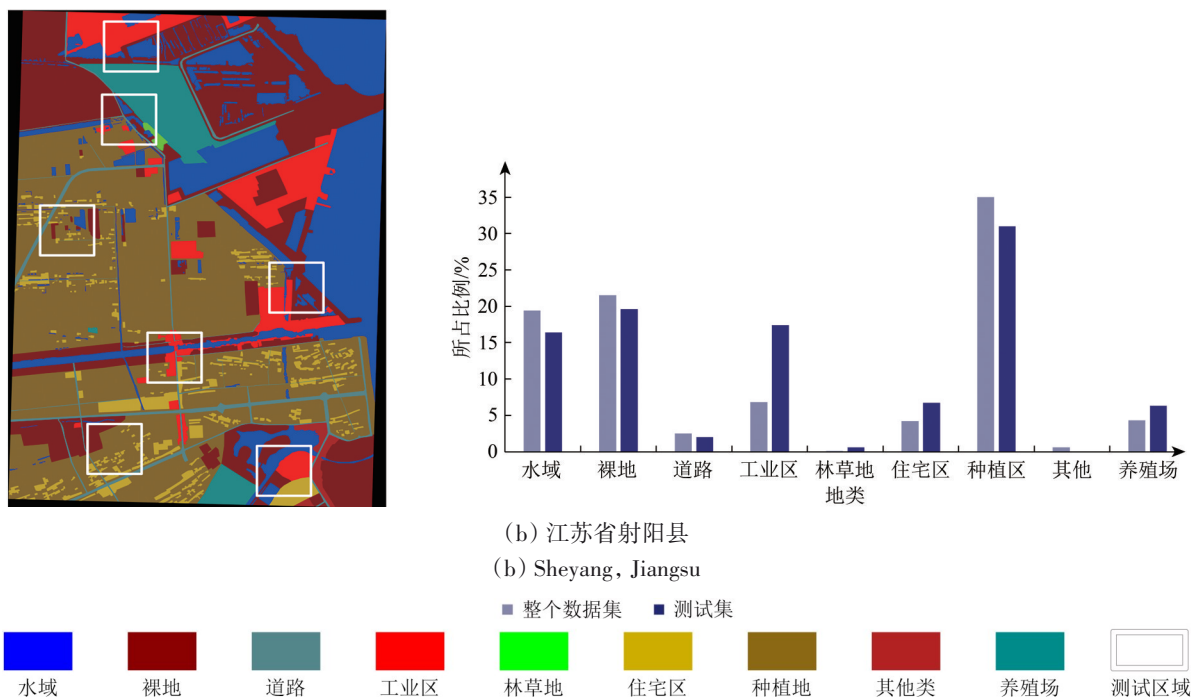


图4 两个区域选取的测试集缩略图及不同地类中所占比例

Fig. 4 The selected test dataset and the proportion of the two regions in different land types

每个区域的训练集和测试集图像尺寸较大，为适配深度学习网络，进一步通过滑窗生成训练集样本切片。本文采用有重叠滑窗的方法进行取样，以512×512为样本切片的尺寸，256为步长进行滑窗。各波段具有不同分辨率，因此，滑窗生成的样本切片数量也不相同，各波段生成的训练集样本切片数量详见表6。当然，如前所述，本数据集公开发布的样本切片没有重叠，且测试区域没有固定，用户可以根据自身需求进行划分。

2.4 高分辨率语义分割网络

采用语义分割方法对SAR图像进行像素级分类，通常经典的语义分割的网络结构为编解码结构。如UNet (Ronneberger等, 2015) 和SegNet (Badrinarayanan等, 2017) 不断对输入图像进行卷积和下采样操作提取深层语义特征，再经过上采样操作及跳层堆叠，不断恢复分辨率，最终实现像素分类。又如Deeplab (Chen等, 2018a) 系列，采用空洞卷积 (Chen等, 2018b) 和深度可分离卷积 (Howard等, 2017) 在编码器部分扩大卷积核的感受野和特征的局部联系，提高分类器性能。但是这种FCN (Long等, 2015) 式的网络结构，在下采样降低分辨率的过程中会不可避免地产生无法逆转的信息丢失。因此，在语义分割中，不

断提取深层语义特征的同时保持高分辨率显得十分必要。

本文采用上述部分算法对该数据集进行评估，最后再利用其中表现较好HRNet (Sun等, 2019; Wang等, 2021) 进行进一步的实验论证。HRNet具有多个分支，每个分支具有不同的分辨率表示能力。高分辨率分支可以提取低层语义特征，对测试阶段生成清晰详细的高分辨率边界发挥重要作用，低分辨率分支提取高层语义特征，在对图像进行类别识别起到关键作用。在HRNet提取特征的过程中，将多个分支的信息进行融合，使得全程具有高分辨率表示，这也是HRNet在多种任务中均表现良好的重要原因之一。HRNet结构如图5所示。本文使用的损失函数为交叉熵损失函数Loss，具体计算方法如下：

$$\text{Loss} = - \sum_{i=1}^n \left(y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i) \right) \quad (1)$$

式中， n 为类别数， y_i 为第 i 类目标的真实结果， $\hat{y}_i$ 是第 i 类目标的预测结果。

2.5 地物分类评价指标

本文采用的评价指标是语义分割中常用的交并比IoU (Intersection over Union)、频权交并比FWIoU (Frequency Weighted Intersection over Union) 和像素准确率。

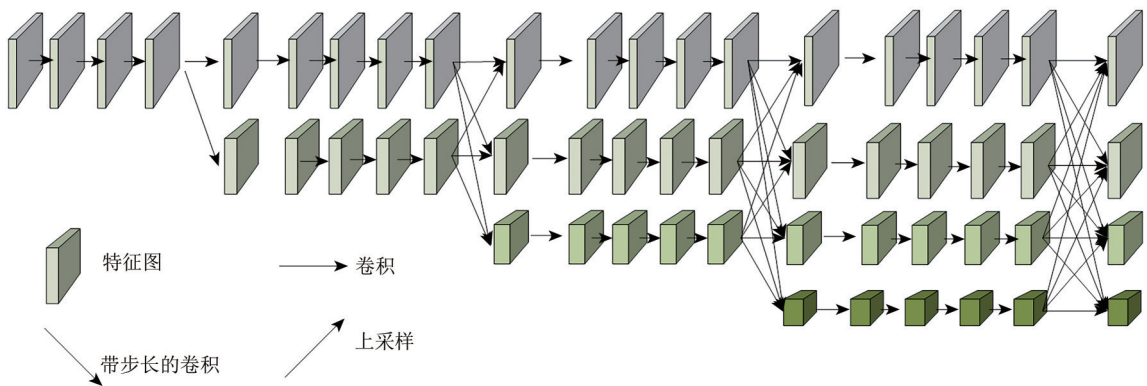


图5 高分辨率网络 HRNet(Wang 等,2021)
Fig. 5 HRNet (Wang et al., 2021)

交并比 IoU：测试结果与真值之间交集和并集的比值。交并比越接近 1，说明测试结果与真实标签越接近，分类效果越好。交并比计算方法如下。

$$\text{IoU} = \frac{\text{target} \cap \text{prediction}}{\text{target} \cup \text{prediction}} \tag{2}$$

频权交并比 FWIoU：根据每个类别出现的频率对各个类的 IoU 进行加权求和。具体计算方法如下：

$$\text{FWIoU} = \frac{1}{\sum_{i=1}^n \sum_{j=1}^n p_{ij} + \sum_{j=1}^n p_{ji} - p_{ii}} \sum_{j=1}^n p_{ij} p_{ii} \tag{3}$$

像素准确率 PA：测试正确的像素数量占总像素数量的比例，具体计算方法如下：

$$\text{PA} = \frac{\sum_{i=1}^n p_{ii}}{\sum_{i=1}^n \sum_{j=1}^n p_{ij}} \tag{4}$$

式（3）和式（4）中， n 为类别数， p_{ij} 表示属于 i 类而被预测为 j 类。

3 融合分类方法与结果分析

前文具体描述了 AIR-MDSAR-Map 的构建、本文采用的深度学习方法及地物分类结果的评价指标。本节将介绍实验参数设置，实验过程和实验结果等，包括经典算法在该数据集上的对比，单波段地物分类和多波段融合地物分类。

3.1 单波段全极化 SAR 地物分类

3.1.1 基准方法实验结果和分析

为评估该数据集，本文对比了主流的语义分割方法。为控制变量，采用分辨率较高且切片数量适中的 C 波段进行实验验证。采用的算法有 UNet、SegNet、Deeplab 和 HRNet。各算法的表现结果如表 5 所示。

表5 基准方法在C波段上地物分类的结果

Table 5 The land cover classification results of baseline algorithms in C band

算法模型	IoU									PA
	水域	裸地	道路	工业区	林草地	住宅区	种植地	其他	养殖场	
UNet	0.577	0.225	0	0.041	0.553	0.328	0.581	0	0.017	0.632
SegNet	0.634	0.258	0.011	0.420	0.192	0.301	0.656	0	0.013	0.611
Deeplab	0.624	0.337	0.048	0.243	0.310	0.298	0.655	0	0.041	0.623
HRNet	0.724	0.563	0.483	0.612	0.612	0.512	0.736	0	0.663	0.784

由表 5 可知，UNet、SegNet、Deeplab 在本数据集上的综合表现不如 HRNet，这是由网络结构决定的。HRNet 能全程保持高分辨率，而其他经典算法都是 FCN 式的结构，在不断的下采样中，难以避免地产生了信息丢失。在地物类别的表现中，由于道路比较狭长，且两边有的树木导致道

路不清晰，这正是分类算法在道路上的表现较差的原因。

3.1.2 各波段分类结果和分析

由上小节可知，HRNet 在本数据集上的分类表现突出，因此选择 HRNet 进行下一步的实验验证

证分析。首先对 AIR-MDSAR-Map 中的每个波段单独进行地物分类以验证不同波段对不同地物的分类敏感性，并测试算法性能以进一步探究多波段融合方法。在训练阶段，以 HRNet18 为训练网络，为了提高模型的拟合能力，迁移了来自 ImageNet 数据集（Deng 等，2009）的预训练模型，

初始化学学习率为 0.01，损失函数为交叉熵损失，训练样本从输入的 512×512 像素的切片中随机截取 256×256 大小。训练参数设置详见表 6。部分 Ka 波段的训练样本见图 6。本文实验均于 Linux 系统上，采用两块 NVIDIA GeForce RTX 3090 Ti GPU 实现。

表 6 地物分类的训练参数设置
Table 6 The training parameter settings of land cover classification

实验编号	数据集	样本数量	样本尺寸	训练尺寸	网络结构	预训练	Epoch	Batch size	初始学习率	Flip
1	光学	18377	512	256	HRNet18	ImageNet	200	16	0.01	True
2	P	582								
3	L	582								
4	S	582								
5	C	2687								
6	Ka	7909								

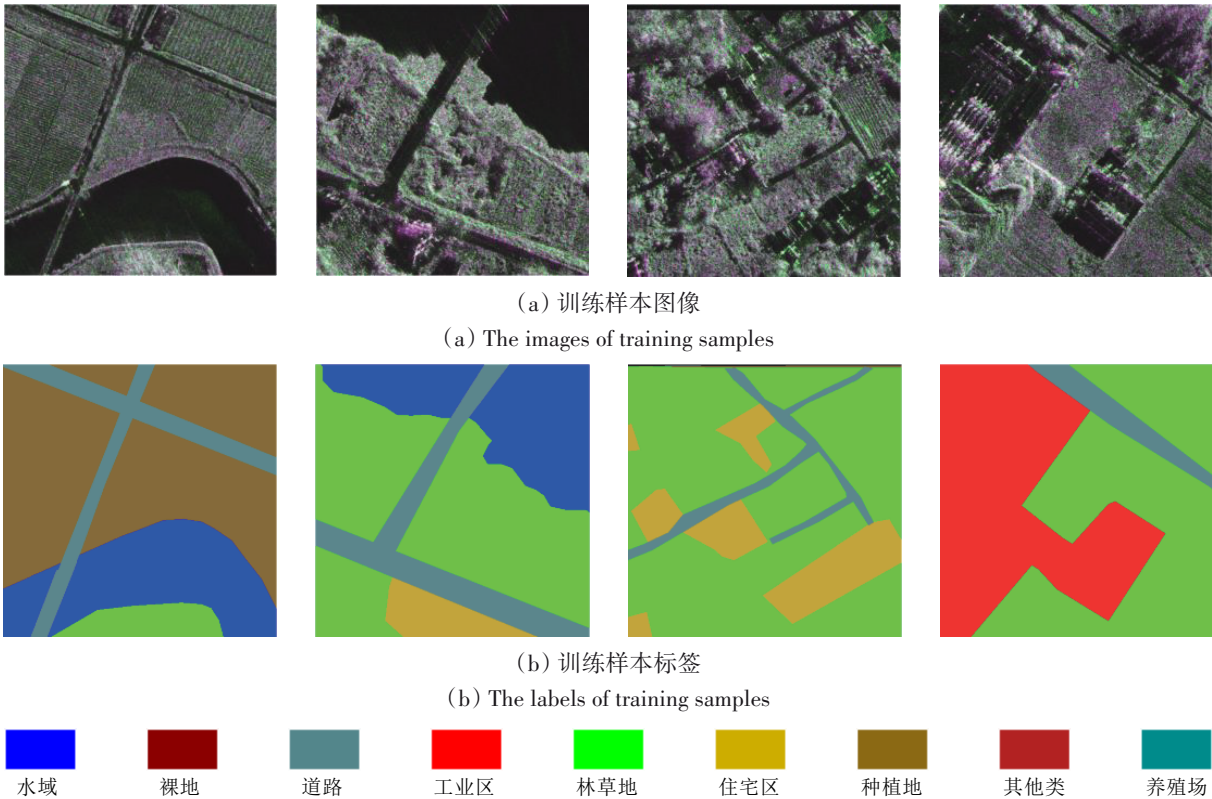


图 6 Ka 波段训练样本示例
Fig. 6 The training samples of Ka band

各波段在测试集上的结果见表 7。分析可得：（1）在测试集上，C 波段数据训练得到的模型对工业区拟合最好，S 波段的模型对林草地和养殖场拟合最好，光学数据对住宅区拟合最好，其余类别在 Ka 波段上能得到最优分类结果。在综合指标上，Ka 波段表现最好，C 波段的结果接近于 Ka 波段。值得一提的是，光学数据上训练得到的模型的拟

合能力不如 C、Ka、S 等波段的原因是：光学数据上不同种类的目标特征较为相似，如部分水域的颜色特征为绿色，与林草地、种植地等较为相似，有一定概率被分类器错分。（2）P、L、S 3 个波段分辨率及数据量等参数均相等，但结果显示，S 波段得到的模型在地物分类的效果上明显优于其他两个波段。

表 7 HRNet18在单波段测试集上地物分类的结果
Table 7 The land cover classification results of HRNet18 in single-band testing dataset

波段	训练耗时/h	IoU									FWIoU	PA
		水域	裸地	道路	工业区	林草地	住宅区	种植地	其他	养殖场		
光学	54	0.708	0.494	0.495	0.513	0.598	0.582	0.719	0.042	0.550	0.6202	0.7644
P	2	0.6151	0.4433	0.3056	0.4472	0.5444	0.4515	0.6485	0	0.5575	0.5461	0.7035
L	2	0.5854	0.4703	0.2755	0.4512	0.5180	0.4317	0.6626	0	0.5565	0.5434	0.7031
S	2	0.7303	0.5082	0.4038	0.4904	0.6260	0.4790	0.7405	0	0.7494	0.6298	0.7691
C	8	0.7241	0.5634	0.4831	0.612	0.6124	0.512	0.7362	0	0.663	0.6482	0.7842
Ka	17	0.7647	0.5773	0.5220	0.5810	0.6130	0.5474	0.7518	0	0.5024	0.6540	0.7878

综上所述，若只考虑分类结果，则选择 Ka 波段进行地物分类最优；若结合数据量、数据分辨率，选择 C 波段和 S 波段进行地物分类最优；若考虑训练时间，S 波段明显最优。

3.2 多波段全极化 SAR 融合地物分类

对于高分辨率机载多波段 SAR 图像实现地物分类，每个波段的 SAR 数据对不同的地物特征有不同的敏感性。因此，每个波段对不同地物目标的分类效果也有优劣。

3.2.1 模型融合

由表 7 分析可知，不同波段数据训练得到的模型对不同地物目标的分类能力不同。本文提出的模型融合方法由多个单波段 SAR 训练得到的模型进行融合，具体见图 7 所示，用来融合的可进行组合。

模型融合的像素分类计算方法如下所示：

$$\text{class} = \max \left(\sum_b f_r(M_b, D_b) \right)_{\text{axis} = \text{classes}} \quad (5)$$

式中， M 表示模型， D 表示数据， b 代表波段，可

取单波段中的任意个波段。 $f_r(M_b, D_b)$ 表示测试之后对概率矩阵进行 resize 操作。将多个波段的概率矩阵 resize 到相同尺寸，再进行加权分类。采用不同波段融合的分类结果见表 8。

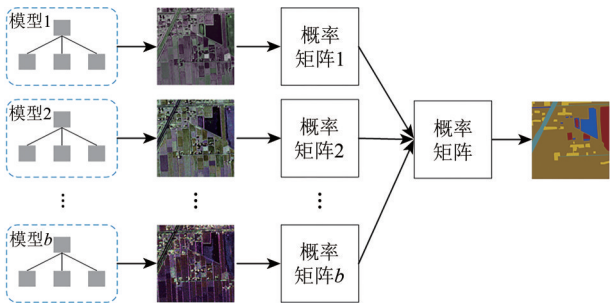


图 7 模型融合策略
Fig. 7 The strategy of model fusion

对比表 7 和表 8 可见：利用各个波段训练的模型进行模型融合，得到最终地物分类结果在水域、裸地、道路、林草地、种植地等 5 个地物类别上达到结果最优；在综合指标上，FWIoU 与 PA 分别达到 0.6866 和 0.8130，比单波段的实验结果高出 10%—15%。

表 8 不同波段模型融合的地物分类结果
Table 8 The land cover classification results of the fusion of different bands' model

模型融合 采用的波段	IoU									FWIoU	PA
	水域	裸地	道路	工业区	林草地	住宅区	种植地	其他	养殖场		
C、Ka、S	0.7925	0.593	0.5421	0.5988	0.6518	0.5516	0.7752	0	0.6923	0.6842	0.8108
C、Ka、L、P、S、光学	0.7912	0.5892	0.5565	0.5887	0.6487	0.5646	0.7778	0.0153	0.74	0.6866	0.8130
C、Ka、L、P、S	0.7827	0.5825	0.5209	0.5912	0.635	0.5444	0.7745	0	0.7271	0.6781	0.8068

3.2.2 先验融合

以单波段的综合分类结果最优的 Ka 波段为基准，在各个地物类别上单独分类结果最优的波段作为辅助，将不同地物分类结果融合进 Ka 波段的

分类结果中。

在多波段先验融合过程中，当同一个区域出现类别分歧时，需考虑不同地物的优先级。综合 3.1 节分析，先验融合实验的考虑因素见图 8。可

见：(1) 每类地物目标计算得到的IoU均不相同，而IoU越高表明模型对该类别分类准确度越高。因此，每类地物的最大IoU被作为地物优先级的考虑指标之一。如水域与道路发生冲突时，若水域IoU大于道路，则此时水域优先级高于道路。(2) 每类地物目标最优IoU与次优IoU的差值大小，也是地物优先级的考虑指标之一。

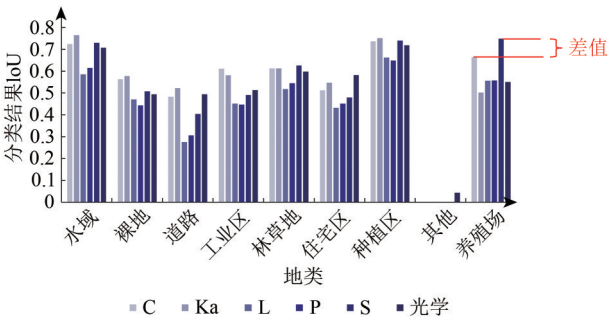


图8 地物类别优先级的定义示意图

Fig. 8 The schematic diagram of land cover priority definition

因此，在先验融合中，不同地物优先级的计算如下所示：

$$\text{priority} = \sum_i \lambda_i \cdot I_i$$

(6)

式中， I_i 为不同的考虑指标所代表的值， λ_i 为权重。本文考虑的类别优先级限定指标有2个，因此式(6)可以简化为

$$\text{priority} = \alpha \cdot I_1 + \beta \cdot I_2$$

(7)

式中， I_1 是不同波段对每个地物目标分类的最优IoU， I_2 是不同波段对每个地物目标分类的最优IoU与次优IoU的差值，二者均为归一化后的值。 α 和 β 分别为其对应的权重。计算得到的值越大说明该类地物在先验融合发生类别冲突时被选择的优先级越高。

由上所述，将在不同地物上分类能力最优的不同波段结果根据地物优先级进行融合。最终的多波段融合结果见表9。

表9 不同 α 与 β 的先验融合的地物分类结果

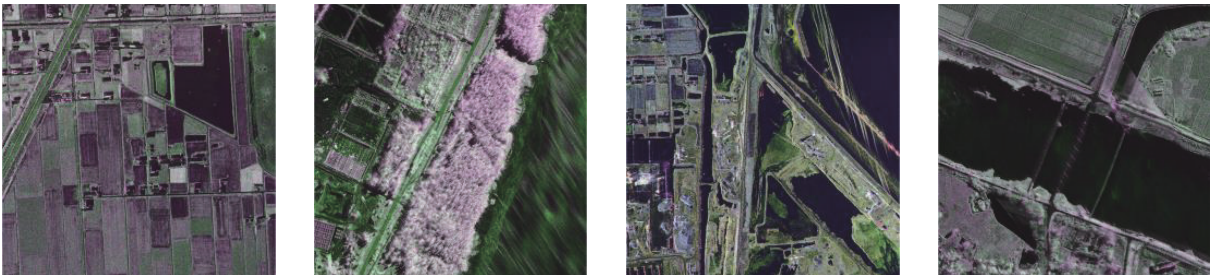
Table 9 The land cover classification results of prior fusion with different α and β

α	β	IoU									FWIoU	PA
		水域	裸地	道路	工业区	林草地	住宅区	种植地	其他	养殖场		
1.0	0	0.7647	0.5864	0.4804	0.6278	0.6085	0.5447	0.7518	0.0015	0.5199	0.659	0.7918
0	1.0	0.8064	0.585	0.5139	0.5994	0.6216	0.5846	0.7599	0.0431	0.679	0.6779	0.8029
0.3	0.7	0.8057	0.5853	0.5141	0.6032	0.6204	0.5846	0.7649	0.0837	0.679	0.6796	0.8053
0.7	0.3	0.8057	0.5864	0.4804	0.6308	0.6195	0.5772	0.7533	0.0015	0.679	0.6773	0.8049
0.5	0.5	0.8057	0.5864	0.5141	0.6462	0.6213	0.6085	0.7649	0.0015	0.679	0.6869	0.8110

将表9与表7的单波段实验结果对比，可见，融合各波段分类器优点的地物分类结果比单波段地物分类结果在综合指标上更优。特别是当 $\alpha=\beta=0.5$ 各波段融合结果在水域、裸地、工业区、住宅区、种植地等5个类别上达到最优，同时，在综合指标上，FWIoU和PA分别达到0.6869和0.8110，均为最优。可见，结合多个波段对不同地物的分类优势，能使最终分类结果达到最优。

图9是不同融合策略的分类结果可视化展示。

其中：图9(a)和图9(b)是Ka波段的图像与真值标签；图9(c)是模型融合的结果；图9(d)是先验融合的结果。对比真值与测试结果可知，模型融合与先验融合的图像在每个地物类型上都表现较好，能正确区分类别得到清晰边界，同时也能将像素正确分类。最重要的是，分类结果能将手动标注遗漏的部分细节正确分类，如水域、住宅区等。



(a) Ka波段图像

(a) The images of Ka band

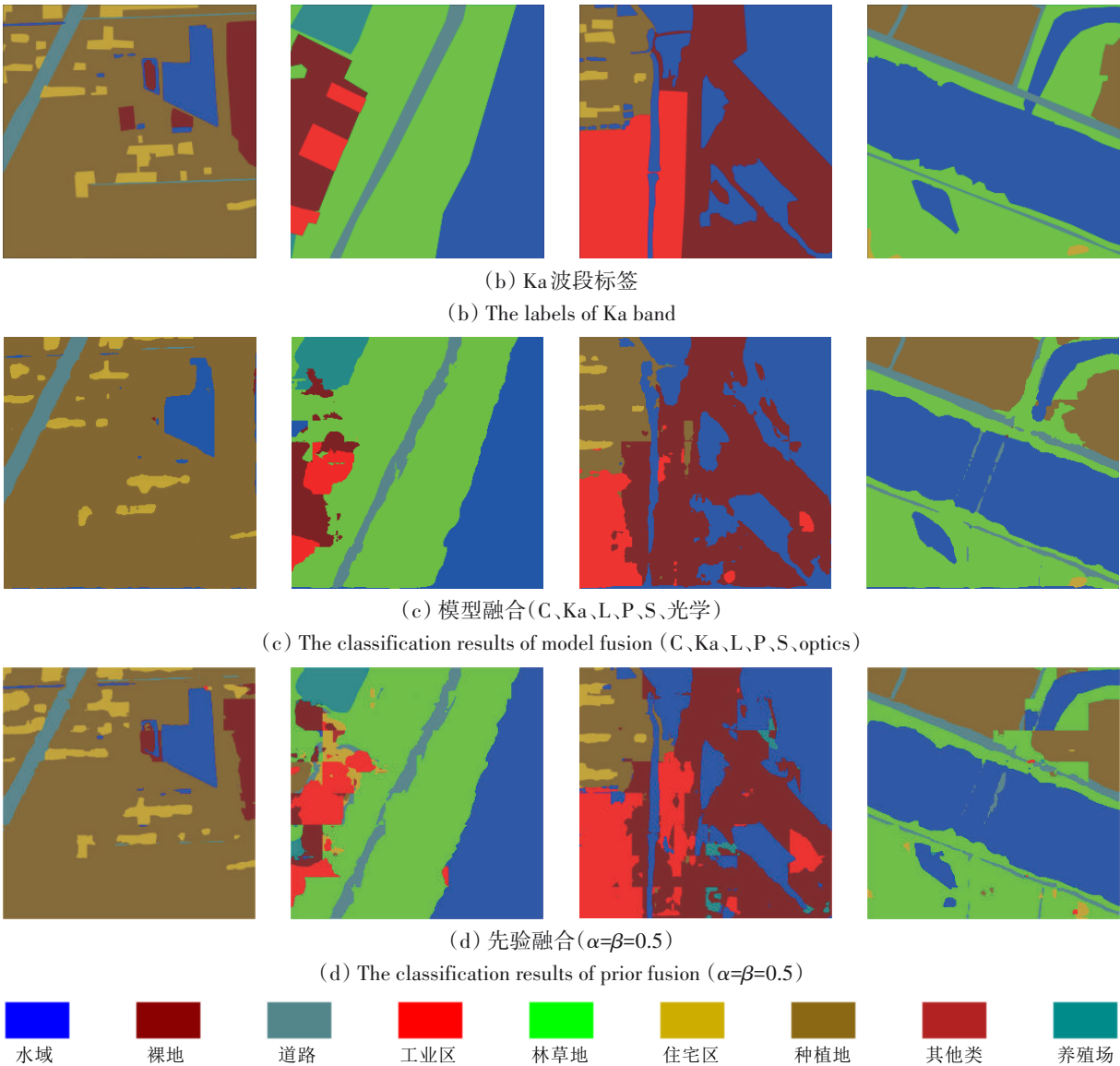


图9 不同融合策略的地物分类结果可视化

Fig. 9 The visualization of land cover classification of different fusion strategies

4 结 论

为探究多维度特征在SAR地物分类上的作用，本文构建了高分辨率机载多维度SAR地物分类数据集 AIR-MDSAR-Map，并在此数据集的基础上应用深度学习方法进行了单波段全极化SAR地物分类试验和多波段全极化SAR融合地物分类试验。主要结论如下：

- (1) 在单波段 SAR 地物分类的中，选择 C 波段和 S 波段进行地物分类最优。若考虑模型训练时间，则利用 S 波段实现地物分类的效果最优；
- (2) 在多波段 SAR 地物分类中，模型融合有选择性地各波段的模型进行融合并分类。先验融合将各波段对不同地物目标的最优分类能力进

行结合，并定义地物类别优先级解决类别冲突。这两种融合策略得到的最终融合结果优于单波段地物分类的结果，在地物分类评价指标上，相比于单波段最高提升 10%—15%。

后续研究中，可以进一步探寻分类过程中进行波段融合或者进行数据分布均衡化来提升分类器的对不同地物目标的分类能力，并在保证分类准确度的情况下降低时间损耗。

参考文献(References)

Addabbo P, Focareta M, Marcuccio S, Votto C and Ullo S L. 2016. Land cover classification and monitoring through multisensor image and data combination//2016 IEEE International Geoscience

- and Remote Sensing Symposium (IGARSS). Beijing: IEEE: 902-905 [DOI: 10.1109/IGARSS.2016.7729228]
- Badrinarayanan V, Kendall A and Cipolla R. 2017. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12): 2481-2495 [DOI: 10.1109/TPAMI.2016.2644615]
- Cen Y, Zhang L F, Zhang X, Wang Y M, Qi W C, Tang S L and Zhang P. 2020. Aerial hyperspectral remote sensing classification dataset of Xiongan New Area (Matiwan Village). *Journal of Remote Sensing* (in Chinese), 24(11): 1299-1306 (岑奕, 张立福, 张霞, 王跃明, 戚文超, 汤森林, 张鹏. 2020. 雄安新区马蹄湾村航空高光谱遥感影像分类数据集. *遥感学报*, 24(11): 1299-1306) [DOI: 10.11834/jrs.20209065]
- Chen B M and Zhou X P. 2007. Explanation of current land use condition classification for national standard of the People's Republic of China. *Journal of Natural Resources*, 22(6): 994-1003 (陈百明, 周小萍. 2007. 《土地利用现状分类》国家标准的解读. *自然资源学报*, 22(6): 994-1003) [DOI: 10.3321/j.issn:1000-3037.2007.06.017]
- Chen L C, Papandreou G, Kokkinos I, Murphy K and Yuille A L. 2018b. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4): 834-848 [DOI: 10.1109/TPAMI.2017.2699184]
- Chen L C, Zhu Y K, Papandreou G, Schroff F and Adam H. 2018a. Encoder-decoder with atrous separable convolution for semantic image segmentation//15th European Conference on Computer Vision. Munich: Springer: 833-851 [DOI: 10.1007/978-3-030-01234-2_49]
- Chen Q, Kuang G Y, Li J, Sui L and Li D G. 2013. Unsupervised land cover/land use classification using PolSAR imagery based on scattering similarity. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3): 1817-1825 [DOI: 10.1109/TGRS.2012.2205389]
- Cloude S R and Pottier E. 1996. A review of target decomposition theorems in radar polarimetry. *IEEE Transactions on Geoscience and Remote Sensing*, 34(2): 498-518 [DOI: 10.1109/36.485127]
- Deng J, Dong W, Socher R, Li L J, Li K and Fei-Fei L. 2009. ImageNet: a large-scale hierarchical image database//2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami: IEEE: 248-255 [DOI: 10.1109/CVPR.2009.5206848]
- Dierking W and Busche T. 2006. Sea ice monitoring by L-band SAR: an assessment based on literature and comparisons of JERS-1 and ERS-1 imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 44(4): 957-970 [DOI: 10.1109/TGRS.2005.861745]
- Fan J Y, Ke C Q, Yao G H and Wang Z F. 2023. Identification of glaciers using fully polarimetric SAR data based on deep-learning. *National Remote Sensing Bulletin*, 27(9): 2098-2113 (范吉延, 柯长青, 姚国慧, 王梓霏. 2023. 基于深度学习的全极化SAR影像冰川边界识别. *遥感学报*, 27(9): 2098-2113) [DOI: 10.11834/jrs.20221541]
- Freeman A and Durden S L. 1998. A three-component scattering model for polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 36(3): 963-973 [DOI: 10.1109/36.673687]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Howard A G, Zhu M L, Chen B, Kalenichenko D, Wang W J, Weyand T, Andreetto M and Adam H. 2017. MobileNets: efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv: 1704.04861* [DOI: 10.48550/arXiv.1704.04861]
- Huang G, Liu Z, Van Der Maaten L and Weinberger K Q. 2017. Densely connected convolutional networks//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE: 2261-2269 [DOI: 10.1109/CVPR.2017.243]
- Long J, Shelhamer E and Darrell T. 2015. Fully convolutional networks for semantic segmentation//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE: 3431-3440 [DOI: 10.1109/CVPR.2015.7298965]
- Lv Q, Dou Y, Niu X, Xu J Q, Xu J B and Xia F. 2015. Urban land use and land cover classification using remotely sensed SAR data through deep belief networks. *Journal of Sensors*, 2015: 538063 [DOI: 10.1155/2015/538063]
- Mishra P and Singh D. 2014. A statistical-measure-based adaptive land cover classification algorithm by efficient utilization of polarimetric SAR observables. *IEEE Transactions on Geoscience and Remote Sensing*, 52(5): 2889-2900 [DOI: 10.1109/TGRS.2013.2267548]
- Moreira A, Prats-Iraola P, Younis M, Krieger G, Hajnsek I and Papathanassiou K P. 2013. A tutorial on synthetic aperture radar. *IEEE Geoscience and Remote Sensing Magazine*, 1(1): 6-43 [DOI: 10.1109/MGRS.2013.2248301]
- O'Mahony N, Campbell S, Carvalho A, Harapanahalli S, Hernandez G V, Krpalkova L, Riordan D and Walsh J. 2020. Deep learning vs. traditional computer vision//Advances in Computer Vision. Cham: Springer: 128-144 [DOI: 10.1007/978-3-030-17795-9_10]
- Ronneberger O, Fischer P and Brox T. 2015. U-net: convolutional networks for biomedical image segmentation//18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Shimoni M, Borghys D, Heremans R, Perneel C and Acheroy M. 2009. Fusion of PolSAR and PolInSAR data for land cover classification. *International Journal of Applied Earth Observation and Geo-information*, 11(3): 169-180 [DOI: 10.1016/j.jag.2009.01.004]
- Simonyan K and Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition//3rd International Conference on Learning Representations. San Diego: [s.n.]
- Song Y H and Hao Y. 2017. Image segmentation algorithms overview.

- arXiv preprint arXiv: 1707.02051 [DOI: 10.48550/arXiv.1707.02051]
- Sukawattanavijit C, Chen J and Zhang H S. 2017. GA-SVM algorithm for improving land-cover classification using SAR and optical remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 14(3): 284-288 [DOI: 10.1109/LGRS.2016.2628406]
- Sun K, Xiao B, Liu D and Wang J D. 2019. Deep high-resolution representation learning for human pose estimation//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE: 5686-5696 [DOI: 10.1109/CVPR.2019.00584]
- Tong X Y, Xia G S, Lu Q K, Shen H F, Li S Y, You S C and Zhang L P. 2020. Land-cover classification with high-resolution remote sensing images using transferable deep models. *Remote Sensing of Environment*, 237: 111322 [DOI: 10.1016/j.rse.2019.111322]
- Wang J D, Sun K, Cheng T H, Jiang B R, Deng C R, Zhao Y, Liu D, Mu Y D, Tan M K, Wang X G, Liu W Y and Xiao B. 2021. Deep high-resolution representation learning for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10): 3349-3364 [DOI: 10.1109/TPAMI.2020.2983686]
- Wang Z R, Zeng X, Yan Z Y, Kang J and Sun X. 2022. AIR-PolSAR-Seg: a large-scale data set for terrain segmentation in complex-scene PolSAR images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15: 3830-3841 [DOI: 10.1109/JSTARS.2022.3170326]
- Wei D, Li Y and Huang D. 2020. Overview on methods of land classification based on polarimetric SAR images. *Computer Systems and Applications*, 29(11): 29-39 (魏丹, 李渊, 黄丹. 2020. 极化 SAR 图像地物分类方法综述. *计算机系统应用*, 29(11): 29-39) [DOI: 10.15888/j.cnki.csa.007705]
- Wu F, Zhang H, Wang C, Li L, Li J J, Chen W R and Zhang B. 2022. SARBuD1.0: a SAR building dataset based on GF-3 FSII imageries for built-up area extraction with deep learning method. *National Remote Sensing Bulletin*, 26(4): 620-631 (吴樊, 张红, 王超, 李璐, 李娟娟, 陈卫荣, 张波. 2022. SARBuD 1.0: 面向深度学习的 GF-3 精细模式 SAR 建筑数据集. *遥感学报*, 26(4): 620-631) [DOI: 10.11834/jrs.20220296]
- Wu Y R. 2013. Concept on multidimensional space joint-observation SAR. *Journal of Radars*, 2(2): 135-142 (吴一戎. 2013. 多维度合成孔径雷达成像概念. *雷达学报*, 2(2): 135-142) [DOI: 10.3724/SP.J.1300.2013.13047]
- Zhao Q, Liu J H, Li Y W and Zhang H. 2022. Semantic segmentation with attention mechanism for remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5403913 [DOI: 10.1109/TGRS.2021.3085889]
- Zheng N R, Yang Z A, Yang H, Sun Y and Wang F. 2023. Semi-automatic annotation method for semantic segmentation of synthetic aperture radar images//*Proceedings of the 8th China High Resolution Earth Observation Conference (CHREOC 2022)*. Singapore: Springer: 95-101 [DOI: 10.1007/978-981-19-8202-6_9]
- Zhong Y F, Hu X, Luo C, Wang X Y, Zhao J and Zhang L P. 2020. WHU-Hi: UAV-borne hyperspectral with high spatial resolution (H2) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with CRF. *Remote Sensing of Environment*, 250: 112012 [DOI: 10.1016/j.rse.2020.112012]

Airborne multidimensional SAR land cover dataset and fusion classification method

ZHENG Nairong¹, YANG Zi'an¹, SHI Xianzheng¹, YANG Hong^{2,3}, SUN Yue², WANG Feng¹

1. Key Laboratory for Information Science of Electromagnetic Waves (Ministry of Education), School of Information Science and Technology, Fudan University, Shanghai 200433, China;

2. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;

3. School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China

Abstract: With the development of Synthetic Aperture Radar (SAR) imaging and deep learning, the use of deep learning to classify land cover in SAR images has received extensive attention and applied research. In this study, a high-resolution airborne multidimensional SAR land cover classification dataset is constructed on the basis of the high-resolution airborne data of the Chinese Aeronautic Remote Sensing System (CARSS) for Earth observation, namely, AIR-MDSAR-Map (Airborne Multidimensional Synthetic Aperture Radar Mapping Dataset).

The original data are obtained by CARSS, and the platform is a modified Xinzhou 60 remote sensing aircraft. SAR and optical images are generated in accordance with the standard data production process. After imaging processing, radiometric correction, polarization correction, and geometric correction, the original SAR data are preprocessed to form Single Look Complex (SLC) data, and then geometric processing is used to generate SAR DOM products. After image enhancement, splicing, and rough correction, the raw optical data are

preprocessed to generate DSM data, and then semiautomatic filtering is performed to produce DEM. Finally, AIR-MDSAR-Map contains polarization SAR images in bands of C, Ka, L, P, and S and high-resolution optical images in Wanning, Hainan, and Sheyang, Jiangsu, with the spatial resolution ranging from 0.2 m to 1 m depending on the band.

AIR-MDSAR-Map divides the land cover into nine categories and generates fine pixel-level labels through a semiautomatic labeling algorithm. In this study, the classical semantic segmentation methods in deep learning, such as UNet, SegNet, DeepLab, and HRNet, are used to verify the classification of AIR-MDSAR-Map. At the same time, we test the classification sensitivity of different band images to all kinds of land cover objects.

This dataset includes multidimensional SAR images of the same place and time, which can be used for fusion classification research. In this study, multidimensional SAR data are fused and classified through different fusion strategies; model fusion classifies land cover by selectively fusing the models of each band, and the a priori fusion uses the prior information of the classification results in each band to distinguish land cover on defining the priority of objects. These two fusion methods outperform the single band in the performance of some types of land cover and improve the FWIoU and PA by 10%—15%, the FWIoU reaches 69%, and PA is 81%.

AIR-MDSAR-Map can satisfy the research and application requirements of different users and can be used to study the characteristics of the same land cover object with different resolutions, bands, and polarizations. Moreover, it can provide a strong promotion for the development of multidimensional SAR applications. The AIR-MDSAR-Map will be available at the ChinaGEOSS Data Sharing Network (<http://www.chinageoss.cn>).

Key words: remote sensing, airborne SAR, multi-dimensional, land cover classification, deep learning, semantic segmentation, AIR-MDSAR-Map

Supported by China High Resolution Earth Observation Special Funding Project (No. 30-H30C01-9004-19/21)