

中图法分类号: TP751.1 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-15

论文引用格式: Polarimetric SAR Image Classification via Semantic Partitioning and Differentiated Contrastive Learning[J/OL]. Journal of Image and Graphics, XXXX: 1-15. DOI: 10.11834/jig.260220. (语义划分与差异化对比学习的极化SAR图像分类方法[J/OL]. 中国图象图形学报, XXXX: 1-15. DOI: 10.11834/jig.260220.) [DOI: 10.11834/jig.260220]

语义划分与差异化对比学习的极化SAR图像分类方法

摘要: 目的 针对高分辨率极化合成孔径雷达(Polarimetric Synthetic Aperture Radar, PolSAR)图像由于空间非平稳性和散射异质性导致的边界混淆及区域语义不一致问题,提出了一种语义划分与差异化对比学习的PolSAR图像分类方法,旨在提升模型对复杂地物的判别能力和边界保持能力。方法 首先,以层次语义模型(Hierarchical Semantic Model, HSM)为理论指导,结合超像素分割技术对图像进行语义划分,精准划分边界敏感区域与非边界稳定区域并构建掩码引导机制;其次,设计双分支异构特征提取网络,利用卷积神经网络(Convolutional Neural Network, CNN)捕获边界区域的细粒度散射特征,同时借助基于超像素的图卷积网络(Superpixel-based Graph Convolutional Network, SGCN)建模非边界区域的长程空间依赖关系;最后,引入自适应边界权重的差异化对比学习策略,通过对边界难分样本的强化监督和跨区域特征对齐,实现特征空间的协同优化。结果 在三组公开极化SAR数据集上进行了实验验证。在西安数据集中,本文算法的总体准确度(Overall Accuracy, OA)达到94.58%,相较于其他主流对比方法,OA提升了0.57%至5.48%;在Flevoland数据集中,OA达到99.63%,且在绝大多数地物类别上取得了最优分类精度;在Oberpfaffenhofen数据集中,OA达到98.65%,显著缓解了郊区类别的错分问题。参数分析表明,5%的训练样本比例可有效平衡精度与人工标注成本。结论 本文提出的分类方法通过“语义分区-异构建模-对比增强”的深度融合,有效解决了传统方法中区域信息与特征学习脱节的瓶颈,在显著提升分类精度的同时,表现出优良的边界保持能力与空间一致性。

关键词: 极化合成孔径雷达;图像分类;层次语义模型;区域划分;对比学习;图卷积网络;异构特征融合

Polarimetric SAR Image Classification via Semantic Partitioning and Differentiated Contrastive Learning

Abstract: **Objective** High-resolution Polarimetric Synthetic Aperture Radar (PolSAR) image classification is a critical task in remote sensing, providing essential data for land-use monitoring, disaster assessment, and environmental resource management. PolSAR imagery characterizes the physical scattering mechanisms and spatial geometric structures of ground objects through multi-channel coherent observations. However, high-resolution PolSAR images often exhibit significant spatial non-stationarity and scattering heterogeneity due to the inherent imaging mechanism and the interference of speckle

收稿日期: 2026-04-16; 修回日期: 2026-07-17

* **通信作者:** 金海燕, 女, 博士, 教授, 博士生导师, CCF 杰出会员, 主要研究方向为图像处理、计算机视觉、智能信息处理。Email: jinhaiyan@xaut.edu.cn; 李军怀, 男, 博士, 教授, 博士生导师, CCF 杰出会员, 研究方向为复杂网络计算、物联网技术、智能信息处理。Email: lijunhuai@xaut.edu.cn

基金项目: 国家自然科学基金面上项目(62471387: 基于黎曼流形测度学习的高分辨极化SAR图像分类方法研究); 陕西省教育厅青年创新团队科研计划项目(23JPI11: 黎曼空间下散射矩阵测度学习的极化SAR图像分类);

Supported by: National Natural Science Foundation General Program (62471387: Research on High-Resolution Polarimetric SAR Image Classification Method Based on Riemannian Manifold Measure Learning); Shaanxi Provincial Department of Education Young Innovative Team Scientific Research Program Project (23JPI11: Polarimetric SAR Image Classification Based on Scattering Matrix Measure Learning in Riemannian Space);

©中国图象图形学报版权所有

noise. Traditional deep learning classification paradigms predominantly utilize isotropic feature processing, applying unified convolutional kernels or network structures across the entire spatial domain. Such models fail to account for the essential differences between boundary and non-boundary pixels in terms of statistical distribution and spatial correlation. Standard convolutional operators are susceptible to interference from heterogeneous neighboring pixels when processing boundary areas, leading to edge positioning deviations and classification "overflow". Conversely, for large-scale homogeneous non-boundary areas, local feature extraction operators struggle to capture long-range spatial dependencies, limiting the ability to maintain semantic consistency. To address these limitations, this study proposes a novel PolSAR image classification method based on region partitioning and differentiated contrastive learning, aimed at optimizing feature representation for complex heterogeneous scenes.

Method The proposed framework follows a systematic "prior layer-backbone layer-supervision layer" architecture. First, as a topological prior, a Hierarchical Semantic Model (HSM) is employed for spatial deconstruction. Based on Marr's visual computation theory, the HSM interprets structural clues to generate a polarization sketch and subsequently partitions the image into boundary-sensitive regions and non-boundary stable regions (comprising homogeneous and structured heterogeneous areas). These results are used to generate binary region masks $M(x, y)$ that provide spatial routing for subsequent feature extraction. Second, a dual-branch heterogeneous backbone network is constructed to adapt to the distinct scattering characteristics of these regions. A Convolutional Neural Network (CNN) branch utilizes small receptive fields to capture fine-grained scattering mutations and local details in boundary areas. Simultaneously, a Superpixel-based Graph Convolutional Network (SGCN) is designed for non-boundary areas. By treating superpixels generated by the Pol_AS LIC algorithm as graph nodes and utilizing symmetric revised Wishart distance for edge weighting, the SGCN effectively models long-range spatial dependencies and maintains global semantic consistency. A feature enhancement module precedes the dual branches to suppress speckle noise. The features from both branches are then integrated through a channel concatenation and a Convolutional Block Attention Module (CBAM) to suppress redundant information and enhance discriminative power. Third, an adaptive boundary-importance differentiated contrastive learning mechanism is introduced as the supervision layer. High-dimensional fused features are mapped to a compact contrastive space. The system learns pixel-level adaptive weights, assigning higher importance to boundary pixels to strengthen supervision on difficult-to-classify samples. The total loss function incorporates a cross-entropy classification loss and a differentiated contrastive loss, which applies distinct constraints to boundary, non-boundary, and cross-region samples to optimize the structure of the feature space.

Result The effectiveness of the proposed method was evaluated on three widely used public PolSAR datasets: Xi'an (China), Flevoland (Netherlands), and Oberpfaffenhofen (Germany). We compared our model against five state-of-the-art methods—Super_RF (Superpixel-based Random Forest), DF GCN (Deep Fuzzy Graph Convolutional Networks), AMS-M2ESL (Adaptive Mask Sampling and Manifold-Euclidean Subspace Learning), 3DCNN (Three-Dimensional Convolutional Neural Network), and PolMPCNN (Polarimetric Multi-Path Convolutional Neural Network)—as well as an ablation model, RPHNet (Region Partition Based Hybrid Deep Network). In the Xi'an dataset experiment, our method achieved an Overall Accuracy (OA) of 94.58% and a Kappa coefficient of 0.9106, outperforming the second-best method, PolMPCNN, by 0.57% in OA. Quantitative analysis showed significant improvements in grass and building categories. For the Flevoland dataset, which features 15 complex crop types, the proposed method achieved an OA of 99.63% and an Average Accuracy (AA) of 99.57%. It successfully suppressed the salt-and-pepper noise and boundary confusion observed in comparative methods, particularly in the Wheat 2 and Rape categories. In the Oberpfaffenhofen dataset, the model reached an OA of 98.65%, effectively mitigating the misclassification of suburban areas while maintaining clear edges for forests and open areas. Parameter analysis indicated that a 5% training sample ratio provides an optimal balance between classification accuracy and manual labeling costs. Furthermore, runtime analysis demonstrated that while the training time (505.30 s) is slightly higher due to the contrastive learning module, the inference speed (15.80 s) remains highly efficient for large-scale tasks.

Conclusion This study presents a comprehensive end-to-end framework for refined PolSAR image classification by integrating region partitioning, heterogeneous feature learning, and differentiated contrastive enhancement. By utilizing the Hierarchical Semantic Model as a spatial guide, the network successfully overcomes the limitations of isotropic feature modeling. The dual-branch CNN-SGCN architecture ensures that fine-grained boundary details and global non-boundary structures are captured simultaneously. Moreover, the adaptive contrastive

tive learning mechanism provides targeted supervision for difficult boundary pixels, significantly enhancing the discriminative power of the feature space. Experimental results across multiple bands (C and L) and varied scenes confirm that the proposed method achieves superior classification performance, boundary maintenance, and spatial consistency compared to existing state-of-the-art models. This research provides a robust and effective solution for the fine-grained analysis of complex heterogeneous PolSAR imagery.

Key words: Polarimetric Synthetic Aperture Radar (PolSAR); Image Classification; Hierarchical Semantic Model (HSM); Region Partitioning; Contrastive Learning; Graph Convolutional Network (GCN); Heterogeneous Feature Fusion

论文引用格式: DOI: 10.11834/jig.260220

1

2 李雨珂^{1,2†}, 徐林婧^{1,2†}, 石俊飞^{1,2*}, 金海燕^{1,2}, 李军怀^{1,2}

1. 西安理工大学计算机科学与工程学院 陕西
西安 710048

2. 陕西省网络计算与安全技术重点实验室 陕西
西安 710048

3 Yuke Li^{1,2}, Linjing Xu^{1,2}, Junfei Shi^{1,2}, Haiyan Jing^{1,2}, Junhuai Li^{1,2}

1. School of Computer Science and Engineering,
Xi'an University of Technology, Xi'an, Shaanxi,
710048, China;

2. Shaanxi Key Laboratory of Network Computing
and Security Technology, Xi'an, Shaanxi, 710048,
China

0 引言

极化合成孔径雷达 (Polarimetric Synthetic Aperture Radar, PolSAR) 通过多通道相干观测, 能够精准刻画地物的物理散射机制与空间几何结构。在高分辨率遥感制图与环境监测任务中, 极化 SAR 图像分类是实现精准制图、灾害评估及资源管理的关键技术 (Li 等, 2024)。然而, 受成像机制及相干斑噪声的影响, 高分辨率极化 SAR 图像在空间维度上呈现出显著的非平稳性与散射异质性。这种异质性导致不同地物在边界地带表现出高度非线性的特征重叠 (Bi 等, 2024), 给分类任务带来了巨大挑战。因此, 研究一种面向空间异质性、具备强边界保持能力且

分类精度高的极化 SAR 图像分类方法具有重要的科学价值。

为了应对上述挑战, 学术界已涌现出大量研究工作。从特征提取的角度来看, 早期的分类研究主要建立在物理散射机制建模基础上。(Lee 等, 1999) 建立了基于 Wishart 分布的统计分类模型。(Cloude 和 Pottier, 1997) 提出了基于特征矢量的分解方法, 通过熵和平均散射角等参数描述地物的散射随机性。(Freeman 和 Durden, 1998) 提出的三分量分解模型则将地物散射解构为表面散射、体散射和偶极子散射, 为地物区分提供了核心物理特征。此外, (An 等, 2010) 进一步完善了极化分解理论以应对取向角偏移问题。

从网络模型的角度来看, 卷积神经网络 (Convolutional Neural Network, CNN) 凭借强大的非线性表达能力在极化 SAR 领域得到了广泛应用。(Li 等, 2024) 通过复数卷积神经网络实现了极化相干矩阵原始相位的提取。(Tian 等, 2025) 利用空-谱联合特征显著提升了农作物的分类精度。为了挖掘非规则空间关联, 一些工作将研究重心转向图神经网络 (Wu 等, 2026)。DFGCN 方法利用深度模糊图卷积网络挖掘样本间的空间及物理信息, 提升了异质区域的分类精度 (Liu 等, 2021)。PolMPCNN 方法则融合多路径结构与极化旋转核, 自适应地捕捉地物的旋转角度特征 (Cui 等, 2021)。也有研究通过结合自注意力机制与图卷积网络, 进一步优化了极化 SAR 地物分类的空间一致性 (Wang 等, 2023)。(Shi 等, 2025) 以及 (Jiang 等, 2025) 的研究也证明了深度卷积架构在处理极化空间上下文中的优越性。

近年来, 流形学习与对比学习成为了提升特征判别性的前沿方向。由于极化协方差矩阵具有对称正定属性, AMS-M2ESL 方法结合自适应掩模采样与流形-欧氏子空间学习框架来提取强区分性特征 (Li 等, 2023)。在少样本场景下, 对比学习展现出巨大

潜力,(Guan等,2024)利用多尺度对比增强机制提升了模型的泛化能力。(Ren等,2022)则通过自监督对比预测任务增强了极化特征的鲁棒性。针对极化SAR图像分类中标记样本稀缺的难题,研究者已从迁移学习、元学习等角度进行了系统梳理(Wang等,2024)。此外,(Jiang and Zhang,2025)提出了多维图对比学习方法,通过跨网络对比、跨视图对比与邻居对比相结合的多维度对比损失,有效提升了节点表示质量。

然而,已有方法在处理复杂场景时仍然存在边界混淆及区域语义不一致等瓶颈。一方面,现有深度学习方法多采用各向同性的特征处理范式,未充分考虑边界像素与非边界像素在统计分布上的本质差异,导致边缘定位偏差;另一方面,标准卷积算子在处理大面积匀质区域时,难以捕捉长程空间依赖关系,语义一致性保持能力有限(Ding等,2024)。

受到层次化视觉理解思路的启发(Zhang等,2024),本文将层次语义模型(Hierarchical Semantic Model, HSM)引入极化SAR分类任务。本文的主要贡献如下:

1)提出语义划分与差异化对比学习的分类网络框架:依托HSM对图像进行空间解构,构建集空间分区、异构特征表征与跨尺度对比增强于一体的端到端架构。

2)设计区域掩码引导的异构特征学习:提出区域掩码引导机制,利用CNN捕捉边界细粒度特征,并借助基于超像素的图卷积网络(SGCN)建模非边界区域的长程空间关联。构建自适应边界权重的差异化对比学习机制:针对边界难分样本设计差异化对比约束,通过权重学习强化对重点区域的监督,提升了特征空间的类内紧凑性与类间可分性。实验证明,本文方法在保持地物边界清晰度的同时,显著提升了复杂场景下的分类精度与鲁棒性能。

1 本文方法

1.1 网络模型框架

为了实现极化SAR图像边界与非边界区域的精细化特征建模,并将区域划分先验与特征学习过程进行深度融合,本文提出了一种语义划分与差异化对比学习的极化SAR图像分类方法。该方法严格遵循“语义划分→特征适配提取→差异化对比监

督→损失联合优化”的算法主线,构建了“先验层-主干层-监督层”三级递进的端到端架构,具体执行过程如下:

首先,先验层通过区域掩码将图像解构为边界与非边界区域,为异构特征学习提供空间先验。

其次,主干层构建CNN-SGCN双分支网络:卷积神经网络(Convolutional Neural Network, CNN)提取边界区域的细粒度散射特征,超像素的图卷积网络(Superpixel-based Graph Convolutional Network, SGCN)建模非边界区域的长程空间依赖。随后经注意力机制加权融合,并辅以特征增强与权重正则化,输出高判别性深层特征。

最后,监督层引入自适应边界差异化对比学习,对边界难分样本赋予更高权重,通过拉近类内距离、推开类间距离,增强异质区域判别能力。

通过上述三大模块的深度协同,该方法形成了从原始极化特征输入到像素级分类结果输出的端到端链路,不仅破解了传统方法中区域先验与特征学习脱节的瓶颈,更有效弥补了通用对比学习在复杂边界场景下的监督盲区。该方法的网络模型框架图如图1所示,各模块的具体设计与实现细节如下。

1.2 语义划分引导的掩码机制

为了有效获取复杂地物的不同语义信息,本文使用层次语义模型(Hierarchical Semantic Model, HSM)(Liu等,2016),将一张复杂场景的极化SAR图像划分为边界、异质和匀质区域三类语义区域,该模型包括极化SAR素描图提取阶段和语义划分阶段。极化SAR素描图有效刻画了地物的几何结构,基于素描线的拓扑结构和图像特征,进一步提取区域图,将复杂场景划分为边界、异质和匀质区域。

基于上述HSM,本文将极化SAR图像像素划分为边界区域、异质区域和匀质区域三类。其中,匀质区域为散射特性一致的地物区域,如水体、农田和裸土等;异质区域为散射特性存在局部变化但语义相对统一的地物区域,如密集植被等;边界区域为两类不同地物的交界界面,如建筑—森林边缘等,其散射特性复杂、类别易混淆,是分类任务中的难点。考虑到匀质区域与异质区域均属于完整地物的语义范围,分类核心目标为保持区域内部语义一致性,因此本文将二者统一归为非边界区域,最终形成“边界区域—非边界区域”的二元特征建模导向,为后续差异化特征提取和对比学习提供区域依据。由HSM生

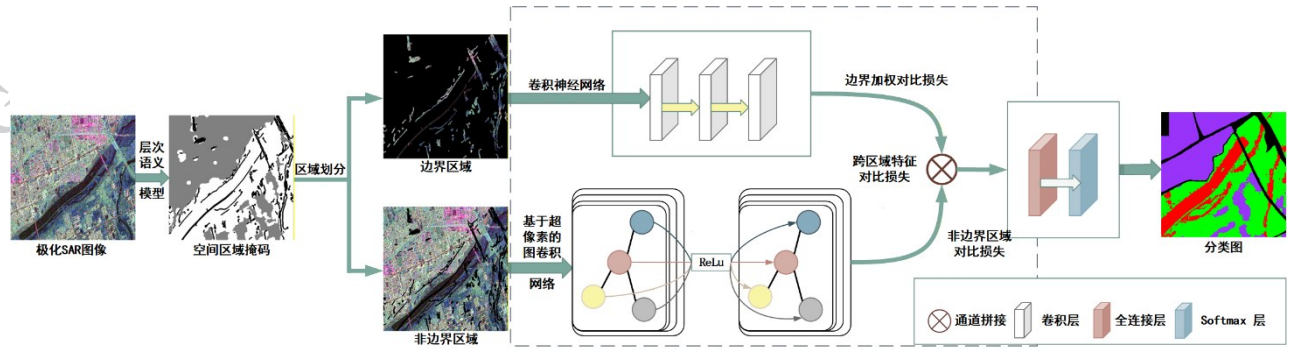


图1 语义划分与差异化对比学习的极化SAR图像分类方法

Fig. 1 Region-Division and Differentiated Contrastive Learning for PolSAR Image Classification

成的区域掩码在网络训练过程中保持固定,不参与反向传播更新,也不随网络参数进行联合优化。

基于上述语义划分,生成与像素级区域标签 $Region(x, y)$ 一一对应的三类区域掩码 $M(x, y)$, 掩码与预处理阶段的区域先验图深度关联,通过二值化处理实现区域像素的精准标记,为特征提取分支提供空间路由指引,其数表达式如公式(1)所示:

$$M(x, y) = [M_b(x, y), M_h(x, y), M_u(x, y)]^T \quad (1)$$

式中, $M(x, y)$ 为像素 (x, y) 处的三维 one-hot 区域掩码向量,第三维分别对应边界区域、异质区域和匀质区域三类区域标签; $M_b(x, y)$ 、 $M_h(x, y)$ 、 $M_u(x, y)$ 分别为边界、异质、匀质区域的掩码向量,均为与输入图像尺寸一致的三维二值向量,满足像素归属于某一区域时掩码值为1,否则为0,且三类掩码向量互斥(单像素仅对应一个掩码值为1)。为适配双分支特征提取需求,将异质与匀质区域统一定义为非边界区域,其掩码为 $M_{non} = M_h + M_u$ 。掩码引导机制通过元素级乘法实现:

$$F_b = F \odot M_b, F_{non} = F \odot M_{non} \quad (2)$$

式中为 F 共享输入特征, F_b 与 F_{non} 分别输入CNN边界分支与SGCN非边界分支, $\odot$ 表示元素级乘法操作。

1.3 掩码驱动的自适应特征提取

针对边界与非边界区域的散射特性差异,结合区域掩码引导策略,本节将构建CNN-SGCN互补特征提取模块。该模块以共享的前端特征增强模块为基础,通过区域掩码实现特征的精准路由,再分别通过SGCN(Kipf和Welling, 2017)建模非边界区域的全局上下文关联、CNN捕捉边界区域的局部散射突变,最终通过混合融合网络与注意力机制实现特征的深度整合。本模块严格遵循“区域划分适配”的核

心设计,形成“特征增强→双分支提取→混合融合→注意力增强”的完整特征学习链路,具体实现如下。

1.3.1 前端特征增强

为统一双分支输入特征的质量并抑制极化SAR图像的相干斑噪声,对原始9维Pauli分解特征进行去噪与高维映射,作为CNN与SGCN子网络的共享输入。设原始极化SAR图像输入特征为 $F_{in} \in \mathbb{R}^{H \times W \times 9}$, 特征增强过程如公式(3)所示:

$$F = \text{ReLU}(\text{Conv}_2(F_{in}, W_2, b_2) + \text{Conv}_1(F_{in}, W_1, b_1)) \quad (3)$$

式中, Conv_1 、 Conv_2 为二维卷积操作, W_1 、 W_2 为卷积权重, b_1 、 b_2 为偏置向量, ReLU 为激活函数,实现非线性映射;输出增强特征 F , 提升特征的判别性。

1.3.2 非边界分支:SGCN全局上下文建模

非边界区域的核心分类目标是实现同类别地物的语义统一,需挖掘超像素尺度的全局上下文信息。本文采用SGCN而非传统GCN,核心优势在于将超像素作为图节点,既降低计算复杂度,又能更好地贴合极化SAR图像的地物散射聚合特性。同时,选用Pol-ASLIC超像素算法(Gao等, 2021),该算法通过极化散射特性优化分割准则,更适配非边界区域的匀质和异质特征分布。

首先,基于Pol-ASLIC的分割结果构建无向图 $G = [V, E, A^c]$, 式中 $V = \{v_1, v_2, \dots, v_N\}$ 为超像素节点集(N 为超像素数量), E 为节点邻接边集, A^c 为加权邻接矩阵。对于超像素数量 N ,是由尺度参数 S 控制,满足 $N = (H \times W)/S$, 式中 H 、 W 分别为图像的高与宽, S 的取值通过参数分析确定(详见2.3节)。进一步,计算第 m 个超像素的平均协方差矩阵 $\bar{C}_m$, 作为节点的初始散射特征,公式如下:

$$\bar{C}_m = \frac{1}{N_m} \sum_{i=1}^{N_m} C_i \quad (4)$$

式中, C_i 为超像素 m 内第 i 个像素的协方差矩阵, N_m 为该超像素内的像素总数。

其次, 采用对称修正 Wishart 距离度量超像素节点间的散射相似性, 结合该距离构建加权邻接矩阵, 实现图模型的权重化, 如公式(5)、(6)所示; 式中局部尺度参数 σ_i 由超像素 i 的 8 近邻(适配极化 SAR 图像超像素分布密度的最优取值)距离均值计算, 如公式(7)所示:

$$d_w(i, j) = \frac{1}{2} \text{tr}(C_i C_j^{-1} + C_j C_i^{-1}) - q \quad (5)$$

$$A_{ij}^c = \begin{cases} \exp(-\frac{d_w^2(C_i, C_j)}{2\sigma_i \sigma_j}), e(i, j) = 1 \\ 0, e(i, j) = 0 \end{cases} \quad (6)$$

$$\sigma_i = \frac{1}{8} \sum_{n=1}^8 d_w(C_i, C_n), n \in N_8(i) \quad (7)$$

上述公式中, $d_w(i, j)$ 表示超像素节点 i 与节点 j 之间的修正 Wishart 距离。 $\text{tr}(\cdot)$ 为矩阵迹运算, $q=3$ 为散射矢量维度, $e(i, j)$ 表示节点邻接性(相邻为 1, 不相邻为 0), $N_8(i)$ 为超像素 i 的 8 近邻节点集。

最后, 采用三层图卷积层对图模型进行特征聚合, 学习非边界区域的全局上下文特征, 第 1 层图卷积操作如公式(8)所示:

$$Z^{(l+1)} = \sigma(\tilde{A}^c Z^{(l)} W^{(l)} + b^{(l)}) \quad (8)$$

式中, $\tilde{A}^c = D^{-1/2} A^c D^{-1/2}$ 为归一化加权邻接矩阵(D 为度矩阵), $Z^{(0)}$ 为由 F_{non} 按超像素求均值得到的初始节点特征, $W^{(l)}$ 、 $b^{(l)}$ 为第 l 层图卷积的权重与偏置。经三层图卷积后, 得到超像素级特征 $Z^{(3)}$, 采用双线性插值法将其映射回像素级, 得到非边界区域特征 $F_{SGCN} \in \mathbb{R}^{H \times W \times 128}$, 实现与边界特征的空间维度对齐。

1.3.3 边界分支: CNN 局部细节捕捉

边界区域存在不同地物之间的散射突变, 需提取精细的局部特征以避免边缘模糊与误分类。本文采用三层浅层 CNN 结构, 其核心依据在于浅层卷积网络具有较小感受野, 能够更加有效地捕捉边界区域的局部细节特征, 同时降低高分辨率极化 SAR 图像训练过程中的过拟合风险。各卷积层均采用 3×3 卷积核, 步长为 1, 填充为 1, 以保证输出特征与输入特征在空间尺寸上保持一致。CNN 子网络以掩码引导后的边界增强特征 F_b 为输入, 通过连续卷积与非线性激活操作提取边界区域的局部散射突变特征, 其计算过程如公式(9)所示:

$$F_{CNN} = 3 \times \text{Conv}(F) \quad (9)$$

式中, $3 \times \text{Conv}(X)$ 表示对输入特征 F 进行 3 次卷积, 经三层卷积提取后, 输出边界区域局部特征 $F_{CNN} \in \mathbb{R}^{H \times W \times 128}$ 。

1.3.4 混合特征融合与注意力增强

为充分融合双分支特征的优势, 本文设计混合融合网络 $Hnet(\cdot)$, 并嵌入 CBAM 通道注意力机制, 实现特征的深度整合与判别性增强, 完成区域划分适配的最终闭环。

首先, 对边界局部特征 F_{CNN} 与非边界全局特征 F_{SGCN} 进行通道拼接, 得到初始融合特征, 如公式(10)所示:

$$F_{cat} = \text{Concat}(F_{CNN}, F_{SGCN}) \quad (10)$$

式中, $\text{Concat}(\cdot)$ 为通道拼接操作, $F_{cat} \in \mathbb{R}^{H \times W \times 256}$ 。

其次, 通过 $Hnet(\cdot)$ 的两层卷积层实现双分支特征的深度融合, 消除特征域差异:

$$F_{fusion} = \text{ReLU}(\text{Conv}_5(\text{ReLU}(\text{Conv}_4(F_{cat}, W_4^f, b_4^f)), W_5^f, b_5^f)) \quad (11)$$

式中, Conv_4 、 Conv_5 为融合卷积权重, b_4^f 、 b_5^f 为对应偏置, 输出融合特征 $F_{fusion} \in \mathbb{R}^{H \times W \times 64}$ 。

最后, 引入 CBAM 通道注意力机制, 对融合特征的关键通道进行加权增强, 抑制冗余特征干扰, 如公式(12)、(13)所示:

$$M_c(F_{fusion}) = \sigma(\text{MLP}(\text{AvgPool}(F_{fusion})) + \text{MLP}(\text{MaxPool}(F_{fusion}))) \quad (12)$$

$$F_{final} = F_{fusion} \otimes M_c(F_{fusion}) \quad (13)$$

式中, $\text{AvgPool}(\cdot)$ 、 $\text{MaxPool}(\cdot)$ 为全局平均池化与全局最大池化操作, $\text{MLP}(\cdot)$ 为两层全连接网络, $\sigma(\cdot)$ 为 Sigmoid 激活函数, $M_c(F_{fusion})$ 为通道注意力权重, $F_{final} \in \mathbb{R}^{H \times W \times 64}$ 为最终判别性特征。综上, 本模块的特征学习过程可统一表示为:

$$F_{final} = Hnet(\text{Concat}(\text{CNN}(F_b), \text{SGCN}(F_{non}))) \otimes M_c \quad (14)$$

该过程实现了区域划分、双分支特征提取与注意力增强的深度协同, 既精准捕捉了边界区域的散射突变细节, 又充分挖掘了非边界区域的空间关联特性, 为后续分类决策与差异化对比学习提供了高质量的特征支撑。

1.4 边界感知的差异化对比学习增强机制

针对极化 SAR 图像边界区域散射特性复杂、类别易混淆, 以及传统对比学习监督机制未充分适配区域空间差异的问题, 本节构建边界感知的区域差

异化对比学习特征增强机制,实现对边界难分区域的强化监督与全域特征分布的对齐优化。本机制分为对比特征映射、自适应边界权重学习、差异化对比损失构建三步,具体实现如下。

1.4.1 对比特征映射

将高维深度融合特征 $\mathbf{F}_{final} \in \mathbf{R}^{H \times W \times 64}$ 通过两层全连接层映射至低维对比特征空间,在保留高阶判别信息的同时降低对比学习的计算复杂度,得到紧凑化的对比特征 $\mathbf{Q} \in \mathbf{R}^{H \times W \times 128}$,映射公式为:

$$\mathbf{Q} = \text{ReLU}(\text{FC}_2(\text{ReLU}(\text{FC}_1(\mathbf{F}_{final}, \mathbf{W}_1^{fc}, b_1^{fc})), \mathbf{W}_2^{fc}, b_2^{fc})) \quad (15)$$

式中, FC_1 、 FC_2 为全连接层, $\mathbf{W}_1^{fc}$ 、 $\mathbf{W}_2^{fc}$ 为全连接权重, b_1^{fc} 、 b_2^{fc} 为对应偏置。

1.4.2 自适应边界重要性权重学习

为了精准分类边界像素,本文提出了自适应边界权重学习策略。首先,基于边界区域掩码 $\mathbf{M}_b(x, y)$ 实现难分像素定位:掩码值为1的边界像素为分类难分像素,掩码值为0的非边界像素为易分像素。其次,通过多层全连接网络对特征 $\mathbf{Q}$ 进行像素级权重学习,为难分像素分配更高的权重,实现强化监督,初始自适应权重的计算如下:

$$\tilde{\omega}_i = \sigma(\mathbf{W}_3^o \delta(\mathbf{W}_2^o \delta(\mathbf{W}_1^o \mathbf{Q}(i) + b_1^o) + b_2^o) + b_3^o) \quad (16a)$$

式中,该网络以归一化后的对比特征 $\mathbf{Q}(i)$ 为输入,通过多层全连接映射生成像素级原始权重 $\tilde{\omega}_i$ 。式中, $\mathbf{W}_1^o$ 、 $\mathbf{W}_2^o$ 、 $\mathbf{W}_3^o$ 和 b_1^o 、 b_2^o 、 b_3^o 分别表示权重学习网络中各层的权重和偏置, $\delta(\cdot)$ 表示 ReLU 激活函数, $\sigma(\cdot)$ 表示 Sigmoid 激活函数。

为了保证边界难分像素在对比学习中获得更高的监督强度,而非边界像素不会被过度加权,本文对学习到的权重 $\tilde{\omega}_i$ 进行修正,得到 $\hat{\omega}_i$,即

$$\hat{\omega}_i = 1 + \mathbf{M}_b(i) \cdot \tilde{\omega}_i \quad (16b)$$

式中, $\mathbf{M}_b(i)$ 表示第 i 个像素的边界掩码取值。当像素属于边界区域时 $\mathbf{M}_b(i) = 1$,其权重被增强为 $1 + \tilde{\omega}_i$,当像素属于非边界区域时 $\mathbf{M}_b(i) = 0$,其权重保持为 $\tilde{\omega}_i$ 。

由于自适应权重在训练过程中可能出现过大或过小的极端取值,直接用于损失函数可能导致训练不稳定。因此,本文利用当前训练批次中全部有效像素的平均权重对 $\hat{\omega}_i$ 进行归一化,得到最终权重 ω_i ,定义如下:

$$\omega_i = \frac{\hat{\omega}_i}{\frac{1}{|\Omega|} \sum_{j \in \Omega} \hat{\omega}_j + \varepsilon} \quad (16c)$$

式中, Ω 表示当前训练批次中的有效像素集合, $|\Omega|$ 为有效像素数量, ε 为防止分母为0的极小常数。最终得到的自适应边界重要性权重 ω_i 将用于后续差异化对比损失计算:一方面在边界加权对比损失中作为像素级权重,强化边界难分样本的监督;另一方面在权重正则化损失中用于约束权重分布,避免极端权重导致训练不稳定。

1.4.3 差异化对比损失构建

以改进的有监督对比损失 (InfoNCE 变体) (Guan 等, 2024) 为基础,本文构建面向极化 SAR 图像区域特性的差异化对比学习损失。与传统对比学习对所有像素采用统一损失约束不同,本文针对边界、非边界区域分别定义对比损失,并提出了跨区域的对比损失,提升特征空间的类内紧凑性和类间可分性,具体定义如下:

边界区域损失: 边界区域通常位于不同地物类别的交界处,散射特性复杂且类别混淆程度较高。为增强边界特征的判别能力,本文在边界像素集合 B 内构造加权 InfoNCE 损失,并引入自适应边界重要性权重 ω_i ,以强化边界难分样本的监督,其定义为:

$$L_b = \frac{1}{|B|} \sum_{i \in B} \omega_i \left[- \frac{1}{|P_B(i)|} \sum_{p \in P_B(i)} \log \frac{\exp(q_i^T q_p / \tau)}{\sum_{a \in B, a \neq i} \exp(q_i^T q_a / \tau)} \right] \quad (17a)$$

式中, L_b 为边界加权对比损失。对于任一边界锚点像素 $i \in B$, $P_B(i)$ 表示边界区域内与锚点像素类别相同的正样本集合,即 $P_B(i) = \{p | p \in B, p \neq i, y_p = y_i\}$ 。分母中的 $a \in B, a \neq i$ 表示边界区域内除锚点像素 i 以外的全部候选样本, q_i, q_p, q_a 分别为锚点像素 i 、正样本像素 p 、以及候选像素 a 的低维对比特征向量。该损失通过自适应权重 ω_i 对边界难分样本进行加权,从而增强模型对边界区域细粒度散射差异的判别能力。

非边界区域损失: 非边界区域内部散射特性相对稳定,主要目标是保持同类地物区域内的语义一致性。因此,本文在非边界像素集合 N 内构造基础

InfoNCE 损失,使同类非边界样本在特征空间中更加紧凑,其定义为:

$$L_{nb} = \frac{1}{|N|} \sum_{i \in N} \left[-\frac{1}{|P_N(i)|} \sum_{p \in P_N(i)} \log \frac{\exp(q_i^T q_p / \tau)}{\sum_{a \in N, a \neq i} \exp(q_i^T q_a / \tau)} \right] \quad (17b)$$

式中, L_{nb} 为非边界基础对比损失。对于任一非边界像素 $i \in N$, $P_N(i)$ 表示非边界区域内与像素 i 类别相同的正样本集合, 即 $P_N(i) = \{p | p \in N, p \neq i, y_p = y_i\}$ 。

跨区域损失:为缓解边界区域与非边界区域之间的特征分布差异,本文以边界像素作为锚点,以非边界像素作为候选样本,构建跨区域对比损失。该损失拉近同类别边界—非边界样本的特征距离,并在归一化对比约束中增大不同类别跨区域样本距离,从而提升跨区域语义一致性和类间可分性,跨区域损失定义为:

$$L_{cross} = \frac{1}{|B|} \sum_{i \in B} \left[-\frac{1}{|P_N(i)|} \sum_{p \in P_N(i)} \log \frac{\exp(q_i^T q_p / \tau)}{\sum_{a \in N} \exp(q_i^T q_a / \tau)} \right] \quad (17c)$$

式中, L_{cross} 为跨区域特征对齐损失。该损失以边界像素 $i \in B$ 作为锚点,以非边界像素集合 N 作为候选样本集合。式中, $P_N(i)$ 表示非边界区域中与边界锚点像素 i 类别相同的正样本集合, 即 $P_N(i) = \{p | p \in N, y_p = y_i\}$ 。

权重正则化损失:为避免自适应边界权重出现过大或过小的极端取值,本文引入权重正则化项,使权重分布保持稳定,从而降低训练不稳定和权重过拟合风险,其定义为:

$$L_{reg} = \frac{1}{|\Omega|} \sum_{i \in \Omega} (\omega_i - 1)^2 \quad (17d)$$

式中, L_{reg} 为权重正则化损失,用于约束自适应权重 ω_i 的数值分布,使其围绕 1 附近波动,避免个别样本权重过大导致模型训练不稳定。

最终,将边界区域损失、非边界区域损失、跨区域损失和权重正则化损失进行加权组合,得到自适应边界差异化对比总损失:

$$L = \alpha L_b + \beta L_{nb} + \gamma L_{cross} + \delta L_{reg} \quad (17)$$

式中, L 为最终的自适应边界差异化对比损失,由上述 4 项损失加权组合得到。 α 、 β 、 γ 和 δ 分别为边界

加权对比损失、非边界基础对比损失、跨区域特征对齐损失和权重正则化损失的平衡系数,通过五折交叉验证确定,取值范围均设置为 $[0, 1]$ 。通过上述差异化约束,模型能够在强化边界难分样本判别能力的同时,保持非边界区域的语义一致性,并实现边界—非边界跨区域特征对齐,从而更有效地适应极化 SAR 图像的空间异质性,提升了特征空间的类内紧凑性与类间可分性。

1.5 总体损失函数

为实现基础分类精度与特征判别性的联合提升,将分类任务监督与对比特征增强监督相融合,构建模型整体优化目标,总体损失函数由主分类损失与区域差异化对比损失加权组成,其数学表达式如公式(18)所示:

$$L_{total} = L_{class} + \lambda L \quad (18)$$

式中, L_{class} 为交叉熵分类损失,以 F_{final} 为输入,负责驱动模型完成像素级地物类别判别,公式为 $L_{class} = -\sum_{c=1}^C y_c \log(p_c)$ (C 为地物类别数, y_c 为真实标签, p_c 为预测概率); L 为自适应边界差异化对比损失,用于提升特征空间的结构化分布与边界区域表征能力;平衡系数 λ 通过五折交叉验

证确定,以三组数据集验证集平均 OA 最高为准,最终选取 $\lambda = 0.10$ 。该取值既能增强边界样本的特征判别性,又避免过大的对比约束破坏分类特征分布。

通过上述总体损失设计,所提框架能够在统一的端到端训练流程中,同步实现区域划分先验利用、异构特征互补学习与自适应对比监督增强,有效解决极化 SAR 图像边界混分、区域异质性问题,最终实现极化 SAR 图像的高鲁棒性分类。

2 实验结果

2.1 实验设置

本文提出了一种区域语义划分与差异化对比学习的极化 SAR 图像分类方法,为了验证该方法的有效性,本文使用了三个公开的被广泛使用的极化 SAR 数据集对其进行验证。同时,本文还使用了五个对比方法和一个消融方法和提出方法进行对比,对比方法分别是 Super_RF (Zhang 等, 2024), DFCCN (Liu 等, 2021), AMS-M2ESL (Li 等, 2023), 3D-CNN

(Li 等, 2023)。, PolMPCNN(Cui 等, 2021)。此外, 为验证差异化对比学习模块的有效性, 本文移除该模块, 仅保留 CNN-SGCN 双分支网络用于极化 SAR 图像分类, 并将其作为对比方法, 记为 RPHnet。

为定量评估模型性能, 我们采用了总体准确度(OA)、平均准确度(AA)与 Kappa 系数进行评估。实验环境设置为 Python3.7(pytorch 1.7.0 GPU), 采用 Win10 系统和 Intel Core i710700 CPU。电脑配置为 64G 内存、NVIDIA GeForce RTX 3070 显卡、8G 显存。在本实验中, 我们将训练比例设置为 5%, 剩余的像素为测试样本。为确保实验公平比较, 所有对比方法均在我们统一的实验设置下(5% 训练样本, 固定随机种子)重新实现。学习速率设置为 $5e-4$, 迭代次数设置为

200。

2.2 实验结果与分析

2.2.1 中国西安数据集实验结果与分析

本实验也使用中国西安数据集。该数据由

SIA-C/X-SAR 系统获取, 为 C 波段全极化影像, 空间分辨率为 $8m \times 8m$, 图像大小为 512×512 像素。图 2(a)与(b)分别展示了该数据的 Pauli RGB 合成图及其对应的真实地物标签。图 2(c) - (g)为五种对比方法在该数据集上的分类结果, 图 2(h)与(i)则分别对应消融实验方法及本文方法的分类结果。本文方法所得混淆矩阵如图 3 所示, 各方法分类性能的定量比较见表 1。

由图 2 可见, 由图 2 可见, Super_RF 方法边缘锯齿明显, 水域分类较差; DFGCN 方法在复杂场景下适应性不足, 建筑与草地混淆严重, 边界噪声明显; AMS-M2ESL 方法噪声较多, 边界轮廓模糊; 3DCNN 方法水域分类较好, 但草地与建筑仍存在误分; PolMPCNN 方法水域精度提升明显, 但草地仍有误分; RPHNet 方法通过语义划分与混合网络实现了稳定分类。本文方法引入自适应边界重要性差异化对比学习机制, 分类精度与空间一致性进一步提升。

表 1 不同方法在西安数据集上的分类准确率(%)

Tab. 1 Classification accuracy of different methods on Xi'an data set (%)

1 类别	Super_RF	DFGCN	AMS-M2ESL	3DCNN	PolMPCNN	RPHNet	本文算法
水域	70.91	84.64	88.99	<u>90.27</u>	95.52	85.05	89.47
草地	94.97	91.79	90.35	93.60	90.95	92.90	<u>94.56</u>
建筑	90.94	87.21	90.27	93.91	<u>97.68</u>	96.64	96.78
OA	89.94	89.10	90.12	93.21	<u>94.01</u>	93.05	94.58
AA	85.61	87.88	89.87	92.60	94.71	91.53	<u>93.60</u>
KAPPA	83.02	81.84	83.68	88.77	90.25	88.51	91.06

注: 加粗字体为每行最优值, 划线字体为每行次优值

不同方法在西安数据集上的定量分析如表 1 所示。传统方法 Super_RF 在水域精度仅为 70.91%, 但在草地类别达到最优的 94.97%。DFGCN、AMS-M2ESL 及 3DCNN 在各向同性特征处理下难以兼顾空间异质性, 整体性能提升有限。PolMPCNN 引入极化旋转核后性能明显改善, 在水域和建筑类别上取得了最优精度。RPHNet 各类别表现均衡稳定。本文方法在保持水域与建筑高精度的同时, 显著提升了草地分类精度, 并取得了最高的 OA(94.58%)与 Kappa 系数(91.06%), 相比六种对比方法, OA 分别提升 4.64%、5.48%、4.46%、1.46%、0.57% 与 1.53%, 验证了本文方法在特征挖掘与判别能力增

强方面的有效性。

2.2.2 荷兰 Flevoland 数据集实验结果及分析

由图 3 可知, 对比方法均存在不同程度的错分与边界模糊问题。Super_RF 在小麦 2 与油菜区域错分严重; DFGCN 因过度侧重全局语义, 边缘判别不足, 导致类别间边界模糊; AMS-M2ESL 在苜蓿与草地混淆明显; 3DCNN 存在大麦与马铃薯的类间误分; PolMPCNN 在小麦 3 区域呈现斑点状噪声, 并将大麦误判为裸土。RPHNet 通过双分支结构实现边界-非边界差异化建模, 各类别边界清晰, 效果较优。本文方法进一步引入自适应边界重要性对比学习, 分类精度与空间一致性再度提升。

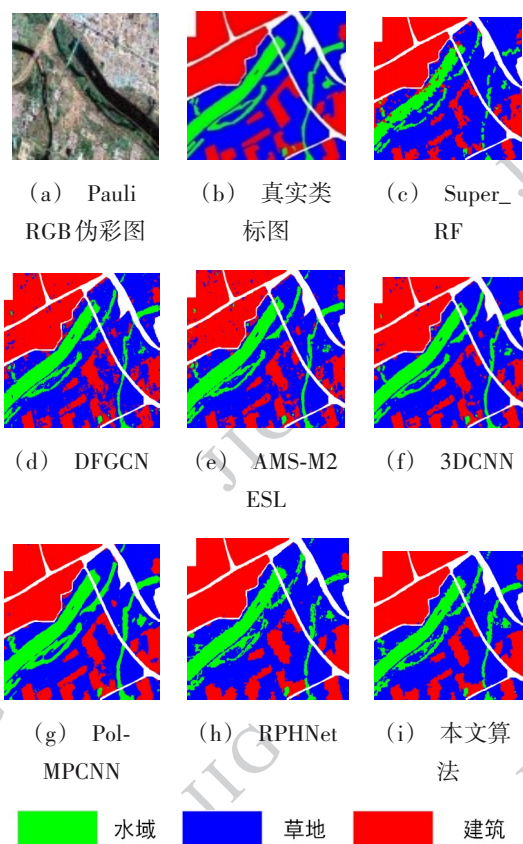


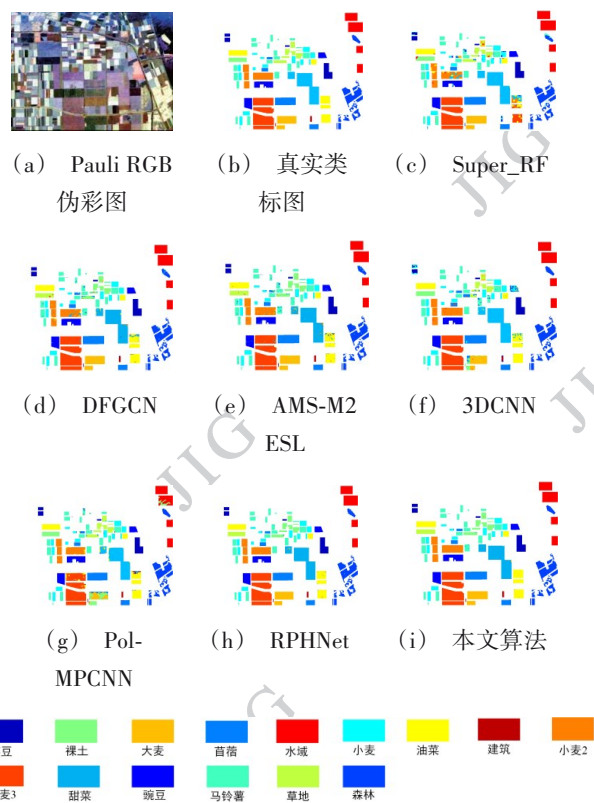
图2 西安数据集分类结果

Fig. 2 Classification results of Xi'an data set

由表2定量结果可知,对比方法普遍存在性能短板或类别偏向性,整体精度受限。其中Super_RF在小麦2和油菜类别上精度极低,DFGCN分类性能不均衡,AMS-M2ESL整体表现平庸,3DCNN和Pol-MPCNN虽在个别类别上达到最优,但受限于其余类别的明显短板,总体精度不高。相比之下,RPHNet性能较为均衡,OA达99.18%。本文方法在绝大多数类别上均取得最优精度,OA、AA及Kappa分别达99.63%、99.57%和99.60%,显著优于所有对比方法,充分验证了该深度融合架构的有效性 with 优越性。

2.2.3 Oberpfaffenhofen数据集实验结果及分析

第三组实验采用的是由德国航空航天中心E-SAR系统获取的L波段全极化SAR数据集,图像空间分辨率为 3×2.2 m,尺寸为 1300×1200 像素,涵盖郊区、林地及开放区域三类地物。数据集总样本数为1560000,其中人工标注样本1385269,未标注样本174731。图4(a)-(b)分别给出该数据集Pauli RGB伪彩色图像与真实地物标签,图4(c)-(i)为五种对比方法、消融模型RPHNet及本文方法的分类



本次实验选用的第二个数据集为荷兰Flevoland地区的AIRSAR C波段全极化数据。该数据集的图像尺寸为 750×1024 像素,空间分辨率达到 $12.1 \text{ m} \times 6.7 \text{ m}$,全图包含157296个标记像素与610704个未标记像素,共涵盖15类地物,包括甜菜、小麦、马铃薯等多种农作物,以及林地、水域、人工建筑等典型地物。图3(a)与图3(b)分别展示了该区域的Pauli RGB伪彩色图像及其对应的标准地物标签图。本次实验选取了五种对比方法及一种消融方法,其定性分类结果如图3(c)至图3(i)所示,各方法的定量评价指标则详细整理于表2中。

图3 Flevoland数据集分类结果

Fig. 3 Classification results of Flevoland dataset

结果可视化,定量评价指标如表3所示。

由图4可知,Super_RF虽区域一致性较好,但边缘锯齿明显;DFGCN在林地区域点状噪声较多,且与郊区混淆严重;AMS-M2ESL对郊区识别误差显著;3DCNN对郊区判别能力较弱;PolMPCNN在林地与开放区域有所提升,但郊区仍有误分。RPHNet郊区误分较少,边界保持相对清晰。本文方法在有效缓解郊区错分的同时,进一步增强了林地与开放区域的边缘精细度,整体视觉效果最优。

表3定量结果表明,对比方法中Super_RF与3DCNN各项指标较为均衡;DFGCN整体指标偏低,林地精度仅89.06%;AMS-M2ESL对郊区识别

表 2 Flevoland 数据集分类精度 (%)

Tab. 2 Classification accuracy of Flevoland dataset (%)

类别	Super_RF	DFGCN	AMS-M2ESL	3DCNN	PolMPCNN	RPHNet	本文算法
茎豆	91.55	<u>99.69</u>	99.18	99.23	99.78	99.58	<u>99.53</u>
豌豆	81.90	99.51	98.32	99.77	99.98	99.71	<u>99.77</u>
森林	78.64	97.65	98.17	98.62	99.96	98.62	<u>99.81</u>
苜蓿	88.01	99.54	97.97	98.78	98.44	<u>99.69</u>	99.93
小麦	79.14	98.06	96.34	97.37	98.01	<u>99.08</u>	99.33
甜菜	82.42	98.22	98.92	99.33	95.51	98.99	<u>99.00</u>
马铃薯	79.77	97.71	96.82	98.81	98.76	99.55	<u>99.52</u>
裸土	74.33	97.53	98.05	99.35	<u>95.70</u>	<u>99.89</u>	100.00
草地	62.00	91.51	95.17	94.25	99.83	98.36	<u>99.01</u>
油菜	53.65	94.24	93.99	95.13	93.55	<u>98.64</u>	99.73
大麦	85.29	99.20	98.45	98.80	95.96	<u>99.70</u>	99.98
小麦 2	55.01	91.45	94.97	95.77	95.96	<u>97.29</u>	99.46
小麦 3	84.36	99.51	98.17	99.03	93.88	<u>99.70</u>	99.82
水域	97.97	99.67	99.88	<u>100.0</u>	96.01	<u>100.00</u>	100.00
建筑	86.34	80.88	91.81	83.29	91.27	<u>95.76</u>	98.62
OA	78.70	97.54	97.38	98.15	96.84	<u>99.18</u>	99.63
AA	78.69	96.29	97.04	97.17	96.80	<u>98.97</u>	99.57
KAPPA	76.73	97.32	97.15	97.98	98.00	<u>99.11</u>	99.60

注:加粗字体为每行最优值,划线字体为每行次优值

能力弱(84.82%),OA 仅 94.41%;PolMPCNN 性能提升明显,OA 达 97.84%;RPHNet 在郊区类别取得最优精度(98.44%),其余指标均为次优。本文方法在除郊区外的各项指标上均为最优,OA、AA、Kappa 分

别达 98.65%、98.49%和 97.70%,较 RPHNet 分别提升 0.23%、0.28%和 0.39%,较 PolMPCNN 分别提升 0.81%、0.83%和 1.38%,综合性能最佳。

表 3 不同方法在 Oberpfaffenhofen 数据集上的分类准确率 (%)

Tab. 3 Classification accuracy of different methods on Oberpfaffenhofen data set (%)

类别	Super_RF	DFGCN	AMS-M2ESL	3DCNN	PolMPCNN	RPHNet	本文算法
郊区	96.79	92.08	84.82	93.43	<u>96.01</u>	98.44	<u>97.32</u>
林地	95.61	89.06	95.70	95.92	<u>98.60</u>	97.46	99.06
开放区域	97.30	98.16	98.20	97.04	98.38	<u>98.73</u>	99.09
OA	96.85	94.91	94.41	95.93	97.84	<u>98.42</u>	98.65
AA	96.57	93.10	92.90	95.46	97.66	<u>98.21</u>	98.49
KAPPA	94.65	91.30	90.43	93.09	96.32	<u>97.31</u>	97.70

注:加粗字体为每行最优值,划线字体为每行次优值

图 4 Oberpfaffenhofen 数据集分类结果 dataset

Fig. 4 Classification results of Oberpfaffenhofen

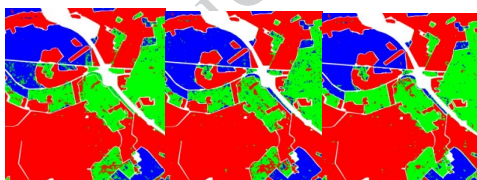
2.3 参数分析

2.3.1 训练样本比例

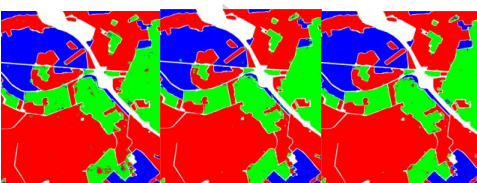


(a) Pauli RGB 伪彩图 (b) 真实类标图 (c)

Super_RF



(d)DFGCN (e)AMS-M2ESL (f)3DCNN



(g)PolMPCNN (h)RPHNet (i)本文算法

郊区 林地 开放区域

由图5可知,三个数据集的OA随训练样本比例增加均呈“初期上升较快、后期趋于平稳”的趋势。5%时西安、Oberpfaffenhofen和Flevoland的OA分别为94.58%、98.63%和99.63%,已接近平台期;继续增至10%后,OA增幅均不足0.2%。综合考虑标注成本与计算开销,5%的样本比例能够在保证精度的同时有效降低人工负担,具备良好的实际应用可行性。

2.3.2 对比学习损失系数

从图6可知,对比损失平衡系数 λ 对各数据集OA影响显著。随着 λ 从0.01增至0.10,三个数据集的OA均持续上升,在 $\lambda=0.10$ 时达到峰值,分别为94.58%、99.63%和98.63%;当 λ 进一步增大至0.25时,三个数据集的OA均出现不同程度下降,分别降至93.27%、98.45%和96.40%。这表明适当的对比损失约束有助于提升特征判别性,但过大的 λ 会导致模型过度关注对比任务而损害分类特征分

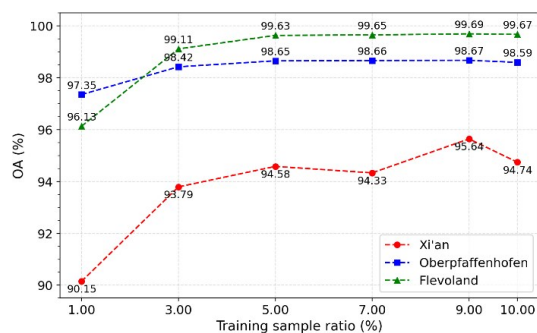


图5 三个实验数据集在不同训练样本比例下的总体分类精度(%)

Fig. 5 Overall classification accuracy of training samples of three data sets at different sampling rates

布。综合五折交叉验证结果,本文最终选取 $\lambda=0.10$,以在分类损失与对比损失之间取得较优平衡。

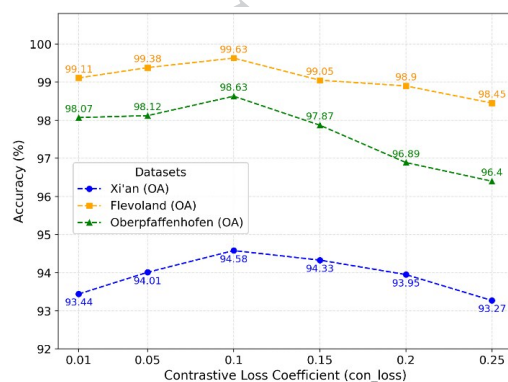


图6 对比损失系数对分类准确率的影响

Fig. 6 The effect of contrastive loss coefficient on classification accuracy

2.3.3 尺度参数S(超像素数量N)

由图7可知,三组数据集的OA随尺度参数S的增大均呈现先升后降的趋势。当S过小时,超像素数量过多,图结构过度细化,不利于区域级语义聚合;当S过大时,超像素粒度过粗,节点易跨越地物边界,削弱了图结构的表达能力。当 $S=300$ 时,三组数据集均达到最高OA,分别为94.58%、99.63%和98.65%,说明该粒度能较好地兼顾全局一致性与边界保持能力。因此,本文最终选取 $S=300$ 作为超像素分割尺度参数。

2.3.4 关键模块有效性验证

为进一步验证本文各关键模块的有效性,本文在西安数据集上进行了消融实验,分别去除HSM区域语义划分模块、SGCN非边界分支、自适应边界权

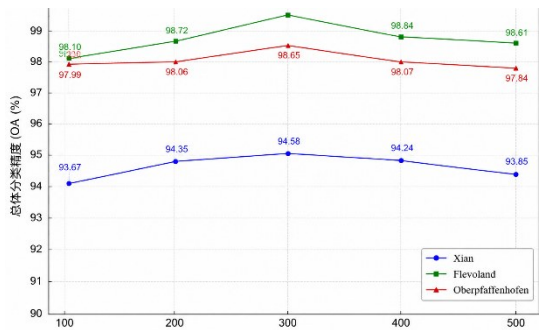


图7 三个实验数据集在不同尺度参数S下的总体分类精度(%)

Fig. 7 Overall classification accuracy (%) of the three experimental datasets under different scalparameters S

重以及跨区域对齐损失,并与完整模型进行对比。

实验结果如表4所示。

表4消融结果可知,完整模型在OA、AA和Kappa上均取得最优结果(94.58%、93.60%、91.06%)。去除HSM模块后,三项指标分别降至90.15%、87.31%和83.63%,表明区域语义先验对差异化建模至关重要;去除SGCN分支后,指标显著下降至82.43%、77.76%和70.50%,验证了SGCN对非边界区域长程依赖建模的关键作用;去除自适应边界权重或跨区域对齐损失后,OA分别降至89.70%和90.28%,说明两者分别通过强化边界难分样本监督和缓解跨区域特征分布差异来提升特征判别性。综上,各模块均对性能提升具有贡献,完整模型通过协同作用取得最优分类效果。

表4 关键模块有效性验证结果
Tab. 4 Effectiveness verification results of key modules

HSM	SGCN	自适应边界权重	跨区域对其损失	OA	AA	Kappa
×	√	√	√	90.15	87.31	83.63
√	×	√	√	82.43	77.76	70.50
√	√	×	√	89.70	87.27	82.97
√	√	√	×	<u>90.28</u>	<u>88.15</u>	<u>83.92</u>
√	√	√	√	94.58	93.60	91.06

注:加粗字体为每行最优值,划线字体为每行次优值

2.4 运行时间分析

表5为各方法在西安数据集上的训练与测试耗时。传统 Super_RF 效率最高(训练 60.33s,测试 1.96s)。深度方法中 DFGCN、3DCNN 和 AMS-M2ESL 训练耗时 121.84s~345.50s,AMS-M2ESL 测试仅 0.12s,推理效率最高;PolMPCNN 因结构复杂,

训练和测试分别高达 21201.26s 和 329.46s,计算成本过高。本文方法在 AMS-M2ESL 基础上引入差异化对比损失,训练耗时由 345.50s 增至 505.30s,但 OA 提升了 4.46%(表1),推理仅需 15.80s,效率较高。总体而言,本文方法在精度、复杂度与推理效率之间取得了合理平衡。

表5 不同方法在西安数据集上的训练和测试时间(s)
Tab. 5 Training and testing time on different methods on Xi'an dataset(s)

时间	Super_RF	DFGCN	AMS-M2ESL	3DCNN	PolMPCNN	RPHNet	本文算法
训练	60.33	165.06	345.50	<u>121.84</u>	21201.26	368.85	505.30
测试	<u>1.96</u>	8.26	0.12	22.80	329.46	15.48	15.80

注:加粗字体为每行最优值,划线字体为每行次优值

2.5 内存开销分析

为评估本文方法的工程部署可行性,统计了推理阶段内存占用,如表6所示。本方法参数量仅 0.619M,对应权重内存约 2.36MB;在 512×512 图像上单次前向推理的激活内存增量约为 71.84MB,总

内存开销约 74MB。结果表明本方法属于轻量级模型,推理内存需求低,具备良好的部署可行性;训练时间为离线开销,不影响部署实时性。

表6 西安数据集上模型推理阶段的内存与时间开销
Tab. 6 Memory and Time Overhead of Model Inference
Stage on Xi'an Dataset

指标	数值
参数量	0.619M
参数内存(FP32)	70.50MB
单次图推理激活内存	71.84MB

注:加粗字体为每行最优值,划线字体为每行次优值

3 结 论

针对复杂极化SAR场景异构区域分类困难问题,本文提出一种语义划分与差异化对比学习的分类模型。首先,基于层次语义模型提取极化SAR复杂场景的语义划分,获得边界、异质和匀质区域掩码,并将异质与匀质区域合并为非边界区域;其次,构建CNN-SGCN双分支网络,分别提取边界区域的局部散射特性和非边界区域的全局特征;最后,设计自适应权重的差异化对比学习机制,通过边界加权对比、非边界语义一致性约束和跨区域特征对齐损失优化分类性能。三组公开数据集实验表明,本文方法在总体精度、边界保持能力和空间一致性方面均取得了较好的分类效果。

同时,本文方法仍存在一定局限性。第一,差异化对比学习会增加训练阶段的计算开销,尤其是在大规模高分辨率影像上需要进一步优化样本采样和批内相似度计算策略;第二,模型性能对对比损失系数为人工选取,无法自适应学习;未来工作将进一步研究跨场景的轻量化学习方法和自适应参数学习,以提升模型在复杂极化SAR场景的泛化能力。

参考文献(References)

An W, Cui Y and Yang J. 2010. Three-component model-based decomposition for polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 48(6): 2732-2739

Bi H X, Kuang Z Z, Li F, Gao J H and Xu C. 2024. Overview of deep learning algorithms for PolSAR image classification. *Chinese Science Bulletin*, 69(35): 5107-5125 (毕海霞, 况祖正, 李凡, 高静怀, 徐晨. 2024. 极化SAR图像分类深度学习算法综述. *科学通报*, 69(35): 5107-5125) [DOI: 10.1360/TB-2024-0087]

Cloude S R and Pottier E. 1997. An entropy based classification scheme

for land applications of polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 35(1): 68-78

Cui Y H, Liu F, Jiao L C, Guo Y W, Liang X F, Li L, et al. 2021. Polarimetric multipath convolutional neural network for PolSAR image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-18 [DOI: 10.1109/TGRS.2021.3071559]

Ding L. 2024. Joint spatio-temporal modeling for semantic change detection in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1-14 [DOI: 10.1109/TGRS. 2024. 3362795]

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

Gao H, Wang C C, Xiang D L, Ye J W and Wang G Y. 2021. TSPol-ASLIC: Adaptive superpixel generation with local iterative clustering for time-series quad- and dual-polarization SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-15 [DOI: 10.1109/TGRS.2021.3126669]

Guan D, Zhang L, Wang Y and Liu J. 2024. Multi-scale contrastive learning for few-shot PolSAR image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 1120-1135

He C, Li Y, Liao M and Yang J. 2020. A novel PolSAR image classification method based on spatio-spectral feature fusion. *IEEE Geoscience and Remote Sensing Letters*, 17(1): 92-96

He M Y, Li Q, Guo Y and Yan H M. 2025. A survey and prospect of intelligent hyperspectral image classification. *Journal of Image and Graphics*, 30(6): 2207-2238 (何明宇, 李谦, 郭阳, 严鸿蒙. 2025. 智能高光谱图像分类研究进展与展望. *中国图象图形学报*, 30(6): 2207-2238) [DOI: 10.11834/jig.240532]

Jiang N N, Zhao W B, Guo J, Zhao Q and Zhu J. 2025. Multi-scale feature extraction with 3D complex-valued network for PolSAR image classification. *Remote Sensing*, 17(15): 2663 [DOI: 10.3390/rs17152663]

Jiang X C and Zhang X W. 2025. Multidimensional graph contrastive learning based on graph enhancement and multi-neural networks. *Journal of Image and Graphics*, 30(9): 3097-3110 (蒋晓灿, 张晓雯. 2025. 图增强与多神经网络结合的多维图对比学习. *中国图象图形学报*, 30(9): 3097-3110) [DOI: 10.11834/jig.240612]

Kipf T N and Welling M. 2017. Semi-supervised classification with graph convolutional networks//*Proceedings of the 5th International Conference on Learning Representations*. Toulon: ICLR

Lee J S, Grunes M R and Kwok R. 1999. Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution. *International Journal of Remote Sensing*, 20(8): 1603-1618

Li M S, Li W, Liu Y K, Huang Y W and Yang G P. 2023. Adaptive mask sampling and manifold to Euclidean subspace learning with distance covariance representation for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 61:

- 1-18 [DOI: 10.1109/TGRS.2023.3265388]
- Li Y, Chen J and Liu L. 2023. 3D-CNN with multi-scale feature fusion for polarimetric SAR image classification. *Remote Sensing*, 15(8): 2145
- Li Y, Yang J F, Zhang H S, Li G and Chen J. 2024. Review of polarimetric synthetic aperture radar remote sensing for land cover classification. *National Remote Sensing Bulletin*, 28(8): 1835-1853 (李煜, 杨静飞, 张鸿生, 李刚, 陈杰. 2024. 极化合成孔径雷达遥感地物分类研究进展. *遥感学报*, 28(8): 1835-1853) [DOI: 10.11834/jrs.20242346]
- Liu F, Shi J F, Jiao L C, Liu H Y, Yang S Y, Wu J, et al. 2016. Hierarchical semantic model and scattering mechanism based PolSAR image classification. *Pattern Recognition*, 59: 325-342 [DOI: 10.1016/j.patcog.2016.02.020]
- Liu H Y, Zhu T W, Shang F H, Liu Y Y, Lv D R and Yang S Y. 2021. Deep fuzzy graph convolutional networks for PolSAR imagery pixel-wise classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 504-514
- Li W M, Xia H, Zhang J D, Wang Y, Jia Y and He Y H. 2024. Complex-valued 2D-3D hybrid convolutional neural network with attention mechanism for PolSAR image classification. *Remote Sensing*, 16(16): 2908 [DOI: 10.3390/rs16162908]
- Ren B, Hou B, Li M and Jiao L. 2022. Contrastive learning-based polarimetric SAR image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-14
- Shi J F, Ji S S, Jin H Y, Li J H, Gong M G and Lin W S. 2025. Content-adaptive multi-region deep network for polarimetric SAR image classification. *IEEE Transactions on Circuits and Systems for Video Technology*, 35 (1): 617-631 [DOI: 10.1109/TCSVT. 2024. 3456480]
- Tian K, Liu X C, Tao D P and Ni J. 2025. One unsupervised feature selection method for the classical linear classifier in land coverage classification with PolSAR imagery. *Computational Intelligence*, 41 (1): e70025 [DOI: 10.1111/coin.70025]
- Wang S, Ji K F and Dong K. 2023. Polarimetric SAR land cover classification combining self-attention mechanism and graph convolutional networks. *Journal of Image and Graphics*, 28(4): 1142-1155 (王爽, 纪凯夫, 董凯. 2023. 自注意力机制与图卷积网络结合的极化 SAR 土地覆盖分类. *中国图象图形学报*, 28(4): 1142-1155) [DOI: 10.11834/jig.220256]
- Wang Z Q, Li Y, Zhang R, Wang J B, Li Y C and Chen Y. 2024. A review of few-shot SAR image classification methods. *Journal of Image and Graphics*, 29(7): 1902-1920 (王智强, 李阳, 张睿, 王建波, 李宇超, 陈勇. 2024. 少样本 SAR 图像分类方法综述. *中国图象图形学报*, 29(7): 1902-1920) [DOI: 10.11834/jig.230563]
- Wu Y, Ji K, Yu W and Su Y. 2022. Deep learning for polarimetric SAR image classification: a review of algorithms and targets. *Remote Sensing*, 14(14): 3326
- Wu Z Z, Wan L, Hong F H, Tang Z D, Sun F, Zou L and Wang X F. 2025. Graph convolutional network integrating skeleton large kernel operators and global context information. *Journal of Image and Graphics*, 30(11): 3604-3616 (吴志泽, 万龙, 洪芳华, 汤正道, 孙斐, 邹乐, 王晓峰. 2025. 融合骨架大核算子和全局上下文信息的图卷积网络. *中国图象图形学报*, 30(11): 3604-3616) [DOI: 10.11834/jig.240353]
- Zhang Q, Xu F and Jin Y Q. 2024. Superpixel-based random forest for high-resolution PolSAR image understanding. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 450-468
- Zhang Z, Guo H, Yang J S, Wang X G and Du Y. 2024. Adversarial network with higher order potential conditional random field for PolSAR image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 1795-1808 [DOI: 10.1109/JSTARS.2023.3322344]
- Zhu X X, Tuia D, Mou L C, Xia G S, Zhang L P, Xu F, et al. 2020. Deep learning in remote sensing: a comprehensive review. *IEEE Geoscience and Remote Sensing Magazine*, 8(4): 8-36

作者简介



李雨珂, 共同一作, 女, 主要研究方向为图像分类, 黎曼分类。邮箱地址: liyuke@stu.xaut.edu.cn



徐林婧, 共同一作, 女, 主要研究方向为对比学习, 图像分类, 邮箱地址: 15529639363@163.com