

电子科技大学

UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA

专业学位硕士学位论文

MASTER THESIS FOR PROFESSIONAL DEGREE



论文题目 基于极化信息辅助的 SAR 目标分类方法研究

学科专业 通信工程(含宽带网络、移动通信等)

学 号 202122011004

作者姓名 林小惟

指导老师 杨建宇 教授

学 院 信息与通信工程学院

分类号 _____ 密级 _____

UDC 注 1 _____

学 位 论 文

基于极化信息辅助的 SAR 目标分类方法研究

(题名和副题名)

林小惟

(作者姓名)

指导老师

杨建宇 教授

电子科技大学 成都

(姓名、职称、单位名称)

申请学位级别 专业硕士 学科专业 通信工程(含宽带网络、移动
通信等)

提交论文日期 2024 年 3 月 15 日 论文答辩日期 2019 年 4 月 15 日

学位授予单位和日期 电子科技大学 2024 年 6 月 8 日

答辩委员会主席

评阅人

注 1：注明《国际十进分类法 UDC》的类号。

Research on SAR target classification method based on polarization information assistance

A Master Thesis for Professional Degree Submitted to
University of Electronic Science and Technology of China

Discipline: **Communications engineering**

Author: **Xiaowei Lin**

Student ID: **202122011004**

Supervisor: **Dr. Jianyu Yang**

School: **School of Information and Communication
Engineering**

摘要

合成孔径雷达（Synthetic Aperture Radar, SAR）是一种对地成像雷达，具有全天时、全天候对地物目标进行持续探测的特性。极化 SAR 在 SAR 的基础上，通过发射、接收不同极化方式的电磁波信号，实现对地物目标的极化散射特性探测，具有更丰富的极化信息。极化 SAR 图像解译技术目前已经广泛应用于军事探测、灾害分析、城市规划等多个领域。

日益增多的高分辨、多极化的 SAR 图像数据为复杂场景分析提供了更丰富的地物目标极化信息，为极化 SAR 解译工作带来了积极作用，但带来了新的挑战。直接堆叠所有类型的极化信息存在信息冗余，而单一类型极化信息又无法适用于所有分类目标，这要求特征表示方法能有效利用多类型极化信息。此外，极化 SAR 数据集标注过程中引入的标签噪声要求分类方法能在部分错误标记样本下进行准确的分类。本章针对极化信息提取与标签噪声问题，提出了适用于极化 SAR 的信息提取方法与鲁棒的目标分类方法。本文针对以上问题，开展了相关的研究，主要研究内容如下：

(1) 研究了极化 SAR 相关理论基础。首先，介绍了电磁波极化特性的表征方式。在此基础上，介绍了目标散射特性的集中描述方法以及几种极化矩阵，并引出了极化目标分解方法的相关理论阐述，并总结了不同极化参数的物理含义。

(2) 提出了一种多类型极化信息提取方法。通过构建双通道的极化特征输入网络结构，利用空间通道注意力模块与注意力修正模块，结合跨空间学习方法，实现不同类型极化特征细化与互信息提取，并聚合不同尺寸的极化特征，为分类任务提供更全面、充分的极化信息表示方式，提升分类准确率。

(3) 提出了一种对标签噪声鲁棒的极化 SAR 目标分类方法。通过建立深度网络模型下噪声、干净样本损失分布差异关系，获取样本噪声概率估计，结合边界样本损失增强，实现标签噪声下鲁棒的参数优化过程，同时对边界信息充分利用，提升标签噪声下极化 SAR 目标分类准确率。

使用两组真实极化 SAR 数据集对本文所提出方法进行有效性验证。结果表明，这些方法能够有效应对极化 SAR 图像目标分类中存在的主要问题，实现地物目标的高准确率分类。

关键词：极化合成孔径雷达，极化 SAR 图像，信息提取，目标分类，标签噪声

ABSTRACT

Synthetic Aperture Radar (SAR) is a kind of ground-imaging radar, which has the characteristics of all-weather and all-weather continuous detection of geo-targets. On the basis of SAR, Polarimetric Synthetic Aperture Radar (PolSAR) can detect the polarised scattering characteristics of ground targets by transmitting and receiving electromagnetic wave signals of different polarisation modes, which has richer polarisation information. PolSAR image interpretation technology has been widely used in military detection, disaster analysis, urban planning and other fields.

With the continuous development of PolSAR technology, high-resolution and multi-polarised SAR image data are increasing, which provide richer feature polarisation information for complex scene analysis. The rich polarisation information and increased data volume bring positive effects to polarised SAR interpretation, but also bring new challenges. The rich polarisation information requires feature representations that can efficiently combine different types of polarisation information. In addition, the label noise introduced during the annotation process of polarised SAR datasets requires classification methods to be able to accurately classify in the presence of mislabelling. In this chapter, information extraction methods and robust target classification methods for polarised SAR are proposed to address the polarised information extraction and label noise problems.

The main research work of this paper is as follows:

(1) A polarisation information extractor with superior performance is necessary before conducting research on PolSAR image classification under label noise. Aiming at the problem of multi-type polarisation information characterisation, a polarisation information extraction method based on two-channel attention is proposed. The method extracts polarisation key information independently based on the difference between PolSAR scattering features and target decomposition features using the network structure of dual-channel attention. On this basis, the two types of polarisation information of dual-channel are integrated through information union to mine the complementary information between the two types of features. In addition, the cross-space feature learning method is used to aggregate different scale polarisation features. Jointly, the above methods achieve a better characterisation of the polarisation information of the feature target, thus improving the accuracy of the downstream classification task. Experiments have demonstrated

ABSTRACT

that this method improves the classification task accuracy compared to using a single type of polarisation feature or directly superimposing polarisation features.

(2) Based on the above research, a robust classification method based on mixture model estimation with boundary enhancement is proposed for the PolSAR label noise problem. The method is based on the difference between the loss of noisy samples and clean samples in the deep network model, and fits the loss distribution based on the beta mixture model to achieve the estimation of the sample noise probability. At the same time, the boundary of Pauli pseudo-color image of PolSAR is extracted and inflated based on Sobel operator, and loss enhancement is applied to the samples inside the inflated boundary. Finally, the self-learning loss function is improved to achieve robust learning optimisation by dynamically adjusting the trade-off between prediction and labelling through the sample noise probability.

The methods proposed in this paper are validated with real PolSAR datasets, and the experimental results verify the effectiveness of these methods for target classification under polarised information extraction and label noise, and improve the accuracy of the feature target classification task.

Keywords: Polarimetric Synthetic Aperture Radar (PolSAR), PolSAR image, Polarised information extraction, target classification, label noise

目 录

第一章 基于双通道注意力的极化信息提取方法

1.1 引言

极化 SAR 通过以不同极化方式发射和接收电磁波信号，以实现对地物目标探测，并提供丰富的后向散射信息。使用四种极化方式组合（HH、HV、VH、VV）的形式进行探测可以获得包含目标散射特性的散射矩阵。进一步解析散射矩阵，通过使用具有不同物理含义的散射基进行分解，可获取代表目标不同散射意义的目标分解特征。如何挖掘极化 SAR 数据中的潜在极化信息一直是遥感探测领域的难点与重点。通过采用适合于极化 SAR 数据特性的特征提取技术，使极化特征具有良好的类内相聚、类间远离的可分性，从而提升系统对目标特性的感知能力，成为极化 SAR 图像解译研究中的关键步骤。

根据文献 [?, ?]，没有一个单一特征能在所有分类中均取得完美的表现，每个方法提取的特征具有其长处以及特征见具有互补性。合理设计极化信息提取方法，综合利用不同类型极化特征的有效信息，理论上能大幅度提升目标的特征表征能力，改善分类结果。

目前在极化信息的利用研究中，存在以下问题：首先，绝大多数的极化 SAR 分类方法直接采用相干矩阵作为输入，却忽略了目标分解特征的表示方式 [?, ?] 提供的目标物理散射信息。二，直接堆叠所有的极化特征作为输入会导致信息冗余且计算量大，甚至降低分类准确度 [?].

本章讨论多类型极化特征信息提取方法，提出了一种基于注意力机制的极化信息提取方法。该方法以注意力机制为基础，利用极化散射特征与极化目标分解特征之间的差异性与互补性，构建两种类型特征之间融合表征关系，增强极化信息的表征性能，从而提升下游分类任务的准确率。通过两组真实极化 SAR 数据验证本方法的有效性与优越性。

1.2 注意力机制介绍

注意力机制是机器学习和深度学习中一种关键的技术，其主要目标是在处理信息时实现对输入数据的加权关注，以便网络模型能够更有效地捕捉与任务相关的信息。基于注意力机制的信息提取在自然语言处理、计算机视觉等领域获得了广泛的应用 [?, ?]。注意力机制使用不同的权重来表示输入特征的不同的重要程度，根据关注的角度差异，可以分为通道注意力、空间注意力和混合注意力三种类型。

1.2.1 通道注意力机制

输入深度网络的特征一般使用多维数据表示，通道注意力专注于挖掘不同通道间的关键，通过自适应地调整通道之间的权重，使网络模型能够更加聚焦于对后续任务有益的特征通道，从而提升模型的性能和泛化能力。压缩和激励网络（Squeeze-and-Extraction Networks, SENet）^[?]是最具代表性的通道注意力实现模型。图??为 SENet 的组成结构图，该方法由压缩和激励两个阶段构成。在压缩阶段，通过全局池化操作对输入多维特征进行压缩，将每个通道的信息整合成单一的数值，用于全局感受野的建模。对于维数是 $H \times W \times C$ 的输入特征，压缩操作将其压缩为 $1 \times 1 \times C$ 维，具体如下式所示：

$$z_c = F_{sq}(\mathbf{u}_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u(i,j) \quad (1-1)$$

在激励阶段，利用全连接层和激活函数，学习得到每个通道的权重。得到的权重向量用来对原始输入特征图中的每个通道进行加权，形成加权的特征图。具体如下式所示：

$$\mathbf{s} = F_{ex}(\mathbf{z}, \mathbf{W}) = \sigma(g(\mathbf{z}, \mathbf{W})) = \sigma(\mathbf{W}_2 \sigma * (\mathbf{W}_1 \mathbf{z})) \quad (1-2)$$

其中， σ 表示 ReLu 激活函数^[?]， $\mathbf{W}_1 \in \mathbb{R}^{\frac{C}{r} \times C}$ 且 $\mathbf{W}_2 \in \mathbb{R}^{C \times \frac{C}{r}}$ 。

模型最终通过权重向量 s 来对输入进行重标定得到：

$$\tilde{\mathbf{x}}_c = F_{scale}(\mathbf{u}_c, s_c) = s_c \mathbf{u}_c \quad (1-3)$$

其中， $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_C]$ ， $F_{scale}(\mathbf{u}_c, s_c)$ 表示 s_c 与特征图 $\mathbf{u}_c \in \mathbb{R}^{H \times W}$ 的通道级乘积。

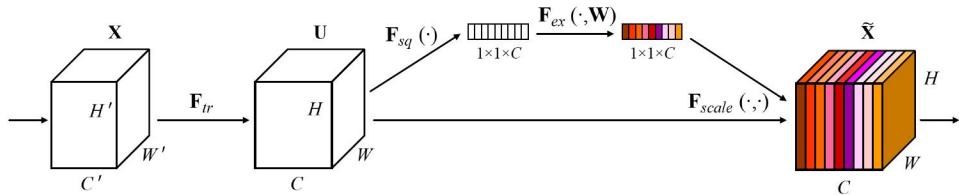


图 1-1 SENet 结构图^[?]

1.2.2 空间注意力机制

空间注意力机制是深度学习处理图像和空间数据中的注意力机制方法。其主要目的是通过对输入数据的不同空间位置引入不同的权重，赋予模型具备灵活关

注对下游任务重要的区域的能力，提升模型对空间结构的感知能力。空间注意力模块（Spatial Attention Module, SAM）^[?]是一个经典的运用空间注意力机制的方法。如图??所示，SAM 的主要思想是首先利用最大池化层和平均池化层获得两个全局的特征图，然后通过拼接操作将两个特征图进行拼合，再利用一个 7×7 的卷积核将拼合的特征图转化成单通道的特征，最后使用 sigmoid 激活函数^[?]得到空间注意力权值，并与原始输入进行相乘得到最终大小与输入相同的输出。空间注意力的计算公式如下式所示：

$$\begin{aligned} \mathbf{M}_s(\mathbf{F}) &= \sigma(f^{7 \times 7}([AvgPool(\mathbf{F}); MaxPool(\mathbf{F})])) \\ &= \sigma(f^{7 \times 7}(\mathbf{F}_{avg}; \mathbf{F}_{max})) \end{aligned} \quad (1-4)$$

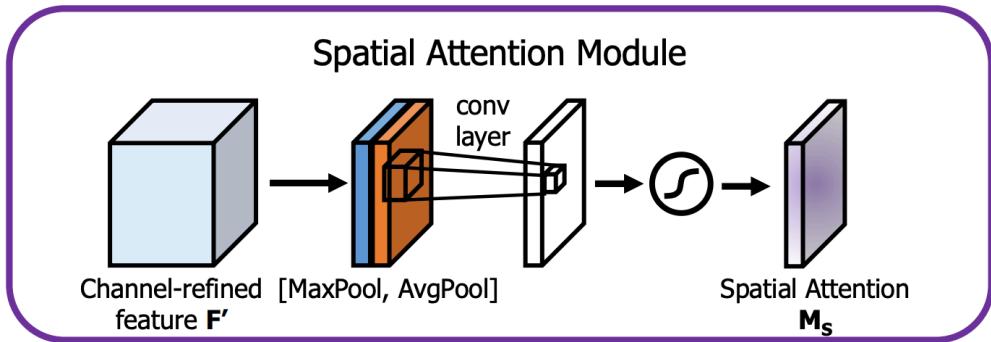


图 1-2 SAM 结构图^[?]

1.2.3 混合注意力机制

混合注意力机制是综合多个注意力模块来处理数据的方法，通过对多个不同类型的注意力机制的融合，来增强深度网络模型对输入数据的建模能力，以更加灵活、全面地捕获输入数据的关键信息。卷积块注意力模块（Convolutional Block Attention Module, CBAM）^[?]是一个综合了通道和空间注意力机制的经典混合注意力方法。如图??所示，CBAM 的主要网络架构有串联的通道注意力模块和空间注意力模块构成。通过依次使用通道和空间注意力模块，分别在通道和空间维度学习数据的关键信息，增强模型对输入数据的感知能力。其中，上一小节已经介绍了空间注意力模块的计算流程，而通道注意力机制的计算公式可以表示如下：

$$\begin{aligned} \mathbf{M}_c(\mathbf{F}) &= \sigma(MLP(AvgPool(\mathbf{F})) + MLP(MaxPool(\mathbf{F}))) \\ &= \sigma(\mathbf{W}_1(\mathbf{W}_0(\mathbf{F}_{avg}^c)) + \mathbf{W}_1(\mathbf{W}_0(\mathbf{F}_{max}^c))) \end{aligned} \quad (1-5)$$

其中， σ 表示 sigmoid 函数， $W_0 \in \mathbb{R}^{C/r \times C}$ 、 $W_1 \in \mathbb{R}^{C \times C/r}$ 均是多层感知机的权重参数。

因此，CBAM 的计算流程可以表示为：

$$\mathbf{F}' = \mathbf{M}_c(\mathbf{F}) \otimes \mathbf{F} \quad (1-6)$$

$$\mathbf{F}'' = \mathbf{M}_s(\mathbf{F}') \otimes \mathbf{F}' \quad (1-7)$$

其中， \otimes 表示逐元素乘法。

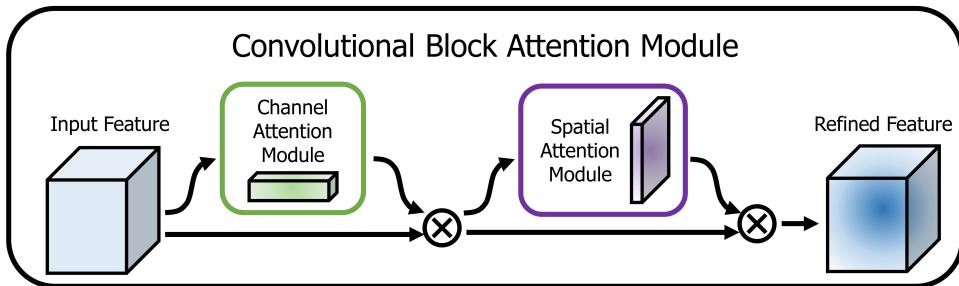


图 1-3 CBAM 结构图 [?]

1.3 双通道注意力的极化信息提取方法

注意力机制方法已广泛应用于图像处理领域，尤其在计算机视觉和深度学习任务中，这些方法通过在网络结构中引入注意力机制，对图像中特定区域进行加权关注，使模型能够更加关注关键的信息，从而提升分类性能。然而，由于极化 SAR 图像具有独特数据特征，包括复杂的多通道信息和多维度的散射机制，直接将通用的注意力机制方法应用于极化 SAR 图像处理任务中无法取得理想的效果。为了更好地适应极化 SAR 图像处理的需求，本章基于后向散射信息与目标分解信息的差异性与互补性，提出了双通道注意力机制的网络结构，充分挖掘两类极化特征的差异信息与互补信息，为下游任务提供更加全面、充分的极化信息表示方式，提升分类任务的准确率。

1.3.1 双通道注意力极化信息提取网络框架

图??是基于双通道注意力的极化信息提取算法示意图。该方法以挖掘多类型极化特征中的隐藏信息为目的，基于极化 SAR 数据中的后向散射特征与目标分解特征的差异性与互补性，实现两类型极化特征的信息提取与有效融合。算法主要思路是：首先，为了充分挖掘两类不同类型极化特征的有效信息，设计双通道网络结构，在两个独立的通道中分别利用注意力方法增强有效信息而抑制无关冗余

信息。其次，利用极化注意力调整模块，基于两类型特征的互补特性，对空间和通道注意力图进行联合动态修正，充分挖掘两类极化特征之间的互补信息。之后，考虑到极化 SAR 图像具有较大的空间尺寸，设计跨空间学习方法，旨在聚合不同尺寸的极化特征。最后，将重新调整的极化特征进行拼接，形成注意力增强的极化特征。值得注意的是，本章提出的极化信息提取方法可以作为一个即插即用的插件式模块，应用到下游的目标检测、分类识别任务中。

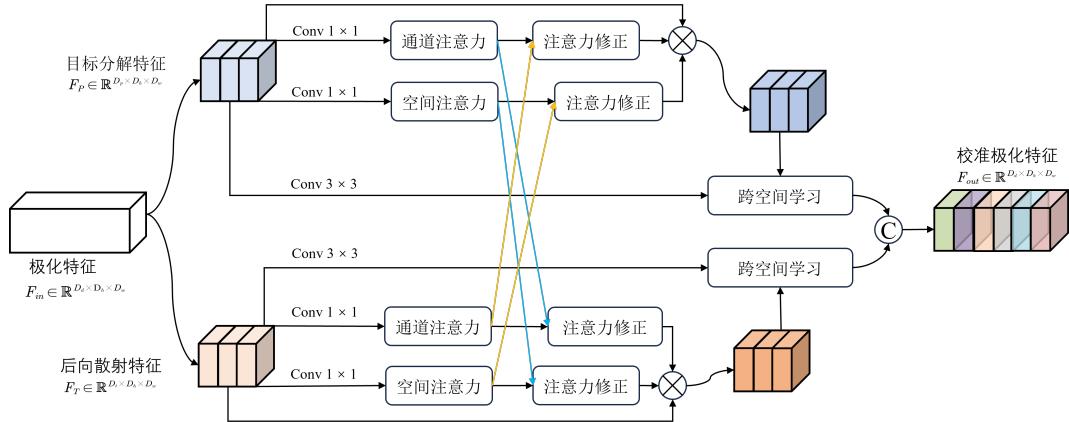


图 1-4 基于双通道注意力的极化信息提取算法示意图

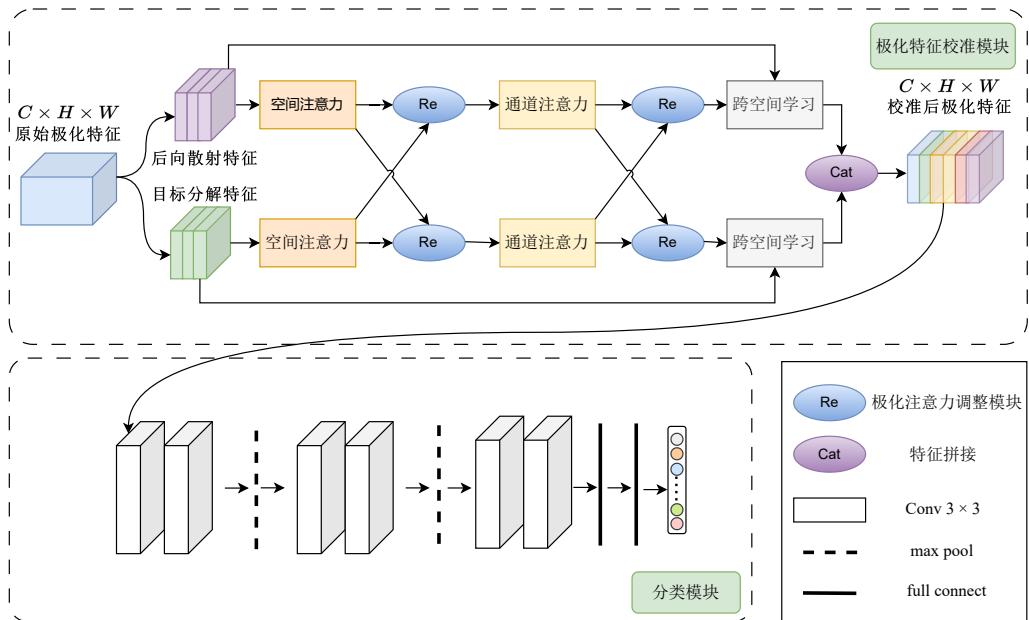


图 1-5 基于双通道注意力的极化 SAR 分类方法示意图

根据图??，网络输入为两个独立通道形式，分别输入具有目标完备后向散射信息的极化相干矩阵和具有不同物理散射特性的极化目标分解特征。两路分支输入特征如下所示：

(1) 通道 1: 后向散射特征

根据公式??, 极化相干矩阵 T 非主对角线元素均是复数, 而主对角线元素均是实数, 并且关于主对角线共轭对称。因此, 可以使用矩阵的上三角 6 个元素代表极化相干矩阵, 考虑到复数性质, 使用一个包含 9 个元素的一维数组表示极化相干矩阵, 具体如下式所示:

$$\mathbf{V} = [T_{11}, T_{22}, \operatorname{Re}(T_{33}), \operatorname{Re}(T_{12}), \operatorname{Re}(T_{13}), \operatorname{Im}(T_{23}), \operatorname{Im}(T_{33}), \operatorname{Im}(T_{12}), \operatorname{Im}(T_{13}), \operatorname{Im}(T_{23})] \quad (1-8)$$

其中, Re 与 Im 分别表示取实部与取虚部运算。 \mathbf{V} 包含了地物目标完备的后向散射信息, 将 \mathbf{V} 作为通道 1 的输入。

(2) 通道 2: 目标分解特征

不同的极化目标分解特征从不同的物理层面反映了目标的散射特性, 通道 2 选择多种经典的极化目标分解特征作为输入, 分别为 Pauli 分解、Cloude 分解、Freeman 分解、Krogager 分解、Huynen 分解共计 32 维目标分解参数。具体特征参数如表??所示。

表 1-1 通道 2 使用的极化目标分解特征参数表

分解方法	特征参数	物理意义	特征维数
Pauli	$ a ^2, b ^2, c ^2$	奇次、偶次、二面角散射	3
Cloude	H, α, A	散射熵、平均散射角、反熵	3
	$\lambda_1, \lambda_2, \lambda_3$	极化相干矩阵特征值	3
Freeman	P_s, P_d, P_v, P_r f_s, f_d, f_v	表面散射、体散射、二次散射、同极化比 Freeman 系数	4
Krogager	$ k_s ^2, k_d ^2, k_h ^2$ $A_0, B_0 + B, B_0 - B$	球、双平面、螺旋散射能量 对称、非对称、不规则性信息	4
Huynen	C, D, E F, G, H	线性、弯曲、扭转性信息 螺旋、沾合、方向信息	3
Van Zyl	f_v, f_d, f_s	体散射、二次散射、表面散射系数	3

为了高效的提取单独通道内的有效极化信息, 使用空间和通道注意力方法对输入的极化特征进行细化, 以激发有效的重要信息而抑制无关的冗余信息, 以加权的方式对输入极化特征进行校准。空间与通道注意力模块的具体网络结构及计算流程将在??小节中介绍。

为了有效融合两类极化特征, 挖掘双通道内不同特征的互补信息, 设计了极化注意力修正模块, 以极化信息一致性为引导, 通过残差连接的形式, 实现由一路极化信息到另一路极化信息的融合修正。极化注意力修正模块的具体网络结构及计算流程在??小节中介绍。

鉴于极化 SAR 目标具有空间尺寸差异大的特点，为了聚合不同尺寸的极化特征，设计了跨空间学习模块，以不同卷积核大小分的卷积层提取不同尺寸的空间信息，结合全局平均池化以及非线性函数映射操作，实现跨空间的极化特征学习。跨空间学习模块的具体网络结构及计算流程在??小节中介绍。

最后，通过特征拼接，将校准后的两类型极化特征进行聚合，作为后续分类模块的输入。

综合以上算法描述，图??对本章算法结合分类器进行目标分类的整体流程进行总结。

1.3.2 空间和通道注意力模块

注意力方法通过强调或抑制输入特征中的不同部分，在深度学习领域，被广泛应用到各种任务中。本文通过引入空间和通道注意力方法，旨在对输入的极化信息动态调整关注度，提升模型对关键信息的感知能力。图??为本方法中的空间注意力模块网络结构。空间注意力方法采用卷积操作来获取全局信息，从而赋予其更多的非线性捕捉能力，更好的拟合不同空间位置之间的复杂相关性，显著减少参数量和计算量。计算公式如下所示：

$$A_s = \sigma(\tilde{f}^{7 \times 7}(AvgPool(S); \Gamma(MaxPool(S)))) \quad (1-9)$$

其中， $\tilde{f}^{7 \times 7}$ 表示卷积核大小为 7×7 的卷积操作； σ 表示 Sigmoid 激活函数。空间注意力通过最大池化和平均池化两个池化操作得到大小为 $2 \times H \times W$ 的特征图，最大池化从图像中提取强的局部信息，平均池化提取全局平均统计特征。对特征图采用卷积运算，转换为 $1 \times H \times W$ 的单一通道特征，再利用 Sigmoid 激活函数将通道特征图转为对应的空间注意力图 A_s ，最后计算注意力图 A_s 与原始输入特征的乘积得到空间注意力信息。计算公式如下：

$$S_S = A_s \otimes S \quad (1-10)$$

其中， \otimes 表示逐元素乘积。通过空间注意力方法，在空间维度进行有效信息激发而抑制无效信息，得到空间注意力图校准的极化特征 S_S 。

相比于空间注意力方法使用卷积层层与激活函数的注意力提取方法，通道注意力方法通过全连接和加和操作，完成对通道维度信息的提取。全局最大池化与平均池化操作对输入特征图的空间依赖性进行拆解，并通过逐个学习每个通道生成反映各个通道重要性的特征图。随后，将得到的两个 $C \times 1 \times 1$ 的特征图进行拼接后使用全连接和加和运算后，再使用 Sigmoid 激活函数，形成成一维的通道注意

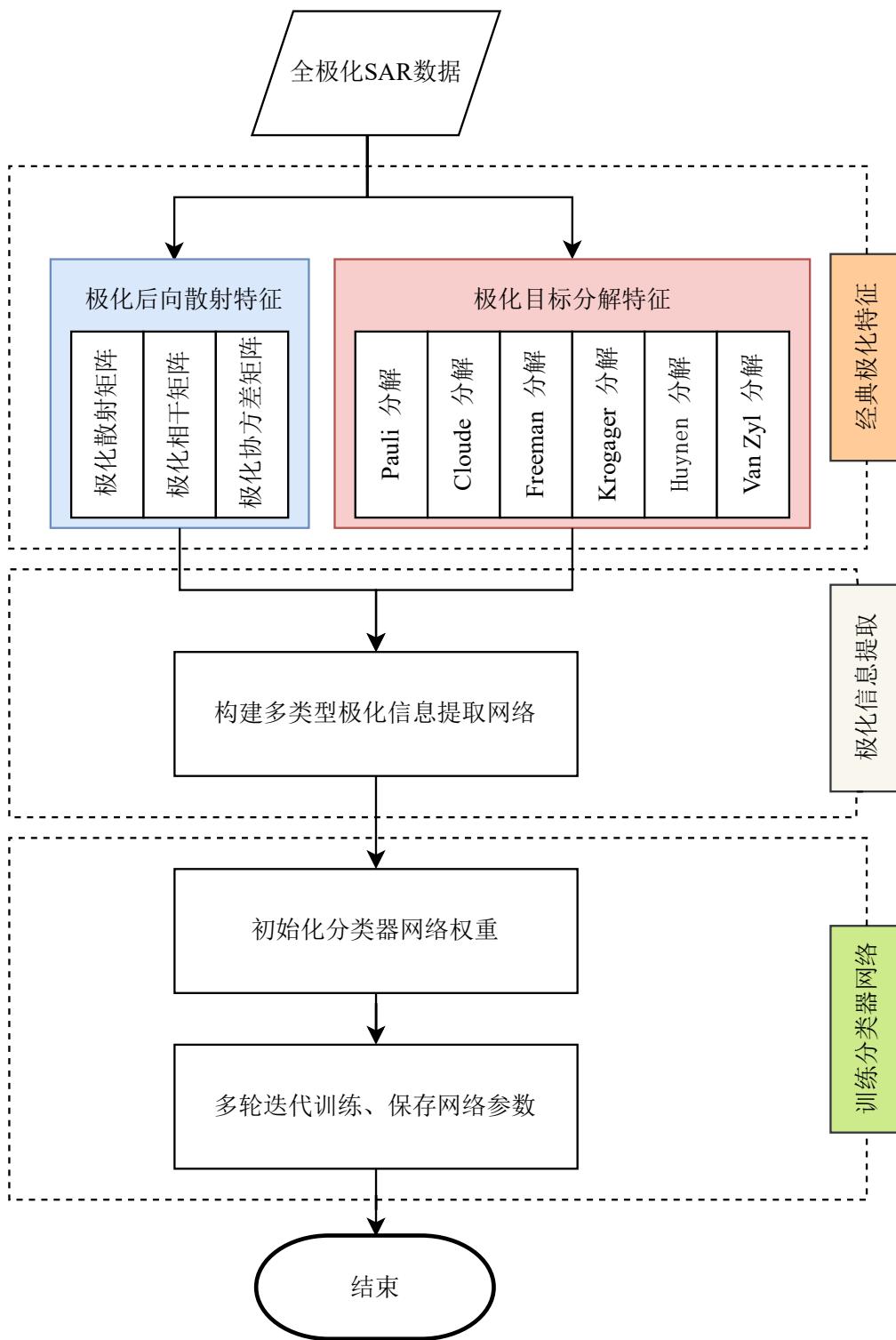


图 1-6 多类型极化信息融合目标分类方法

力图, 用于表示不同通道特征图的重要性。计算如下式所示:

$$A_C = \sigma(\Gamma(\text{AvgPool}(S)) + \Gamma(\text{MaxPool}(S))) \quad (1-11)$$

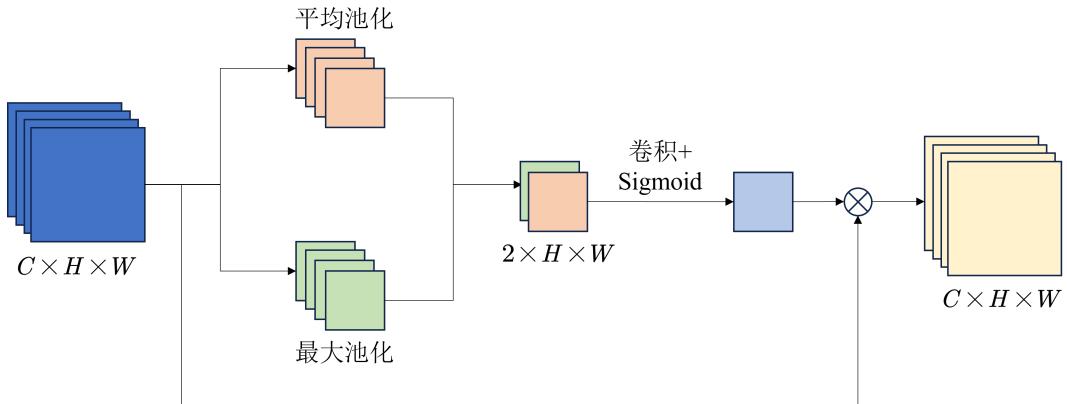


图 1-7 空间注意力模块结构

其中， σ 表示 Sigmoid 激活函数； Γ 表示两层神经网络。将通道注意力图与输入极化特征相乘，得到通道注意力图校准的极化特征。计算公式如下：

$$S_C = A_C \otimes S \quad (1-12)$$

其中， \otimes 表示逐元素乘法。

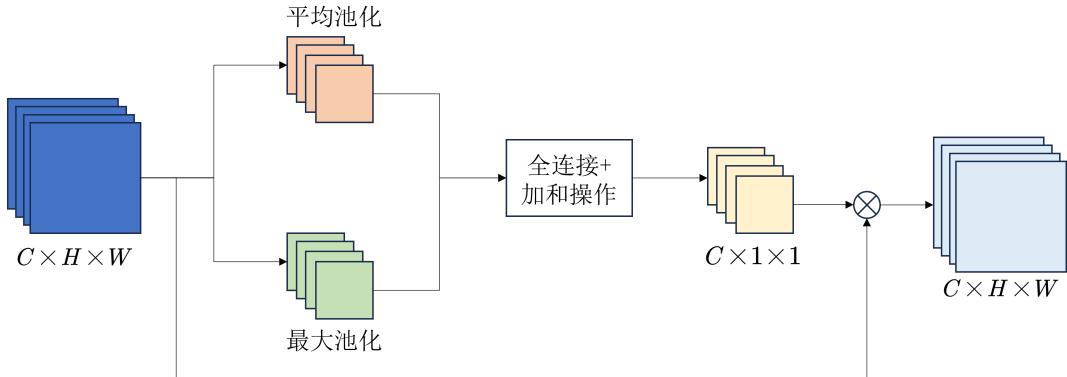


图 1-8 通道注意力模块结构

1.3.3 极化注意力修正模块

基于双通道注意力架构对极化 SAR 输入特征提取空间、通道注意力特征图，对于同一类型目标，校准后的两种类型极化信息存在一致性与互补性。为了保持特征一致性和挖掘极化互补信息，图??展示的极化注意力修正模块，以极化信息一致性为引导，对输入的极化特征进行动态调整。极化注意力修正模块利用输入的一路极化特征作为主要引导，对另外一路特征进行修正。其中主要包含卷积与

残差运算，计算公式如下所示：

$$S'_1 = f^{1 \times 1} (f^{1 \times 1} (S_1)) \times f^{1 \times 1} (f^{1 \times 1} (S_1)) \quad (1-13)$$

$$S'_2 = f^{1 \times 1} (S_2) \quad (1-14)$$

$$S_{out} = S_1 + (S'_1 \times S'_2) \quad (1-15)$$

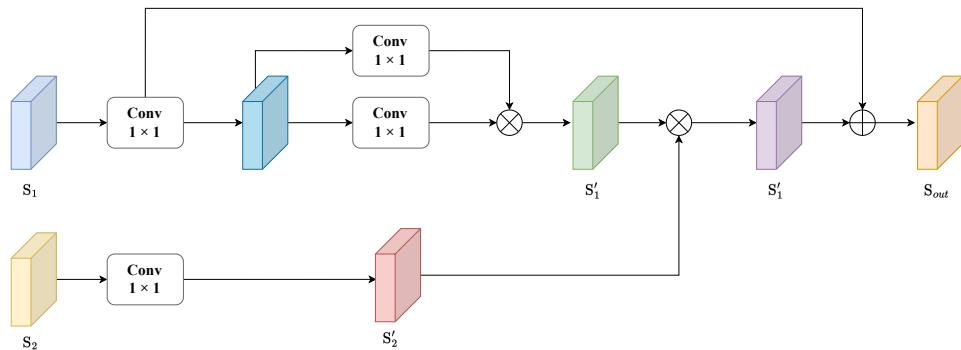


图 1-9 极化注意力修正模块网络结构示意图

1.3.4 跨空间学习模块

跨空间学习模块的网络结构如图??所示。跨空间学习模块提供了一种不同空间维度方向的极化信息聚合方法，来实现多尺度下的极化特征聚合。引入两个分支的张量，分别是 1×1 分支的输出和 3×3 分支的输出。随后利用二维全局平均池化对 1×1 分支的输出中的全局极化空间信息进行编码，用于编码全局信息和建模远程依赖关系。二维全局池化操作可以表示为：

$$z_c = \frac{1}{H \times W} \sum_j^H \sum_i^W x_c(i,j) \quad (1-16)$$

在以上二维全局平均池化的输出处采用二维高斯映射的自然非线性函数 Softmax 来拟合线性变换，进而提升计算效率。通过将上述并行处理的输出与矩阵点积运算相乘，得出了第一个空间注意力图。同样利用二维全局平均池化在 3×3 分支编码全局空间信息，将每组内的输出特征映射计算为生成的两个空间权重值的集合，然后使用 Sigmoid 函数映射成空间位置对应的权重关系。通过捕获像素

级的成对关系，突出显示所有像素的全局上下文信息。计算公式如下：

$$S_1 = f^{1 \times 1}(S) \times \text{AvgPool}(f^{3 \times 3}(S)) \quad (1-17)$$

$$S_2 = f^{3 \times 3}(S) \times \text{AvgPool}(f^{1 \times 1}(S)) \quad (1-18)$$

$$S' = \sigma(S_1 + S_2) \quad (1-19)$$

其中， σ 表示 sigmoid 函数。

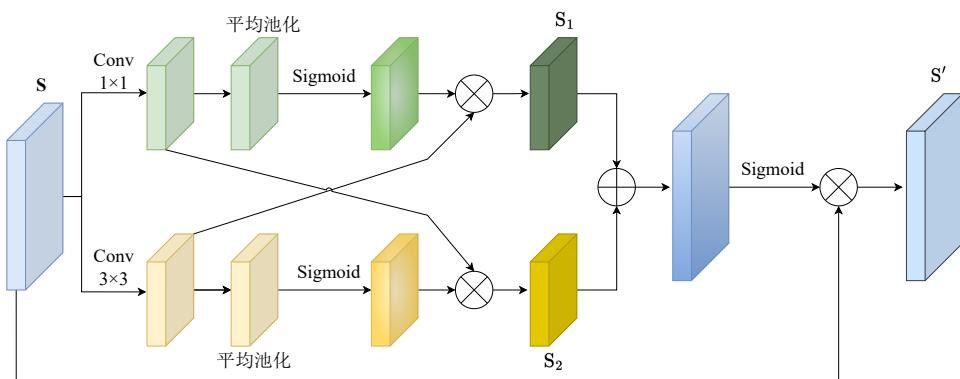


图 1-10 跨空间学习模块网络结构示意图

1.4 实验结果与分析

1.4.1 实验模型介绍

本章实验所使用的计算环境为一台 CPU 为 Intel Core i7-8700K 和配备了 NVIDIA GPU (GeForce RTX 3090, 24G) 的计算机设备。操作系统采用 Ubuntu 20.04 LTS。深度学习框架选择 PyTorch，版本为 1.9.0，同时依赖 CUDA 深度神经网络库 (cuDNN) 版本 8.0.5。在科学计算方面，实验使用 NumPy 库，版本为 1.19.5。

学习率作为深度学习模型训练的关键参数之一，其值的选择对于模型的收敛速度至关重要。在训练过程中，如果学习率设置的过大或者过小，均会可能给模型的分类准确度产生负面影响。通常情况下，在模型的训练初始阶段，采用较大的学习率能够使模型快速收敛到最优点附近。随着训练的执行，逐渐减小小学习率，以更加精确地接近最优点。在本章所采用的模型中，初始学习率为 3×10^{-4} ，在训练至第 30 至 60 个 epoch 期间，学习率经过衰减变为原值的 0.1 倍。这里的 epoch 表示训练集中素有样本完成一次正向传递和反向传播的过程，本章中模型训练的 epoch 设置为 100。Batch size 表示每次训练中选择的样本数量，其值的大小会影响网络的优化速度和执行效率。在本章模型中，batch size 被设置为 64。

本章基于双通道注意力的极化信息提取器 (Dual-attention Polarization informa-

tion extractor, 简记为 DP), 可以作为插件式组件应用到下游的分类任务中。为了验证本章方法的有效性与优越性, 基于 DP 构建端到端的极化 SAR 图像特征提取与分类方法 (记为 DP-CNN), 网络结构如图??所示, 分为极化信息提取和分类器两个部分。从输入的高维原始极化特征出发, 利用 DP 模块可以获得全局信息并且嵌入到分类器中。原始极化特征中有价值的信息被激发, 而没有价值的信息被抑制。当具备了重新校正的极化特征之后, 基于深度神经网络实现极化 SAR 图像的分类。在本节的实验方法中, 使用一个类似 vgg 的卷积架构来实现特征提取与分类流程。具体而言, 该卷积架构包括三个带有 ReLU 激活函数的 3×3 卷积层、三层最大池化函数、两层全连接网络以及一层 SoftMax 激活层构成。

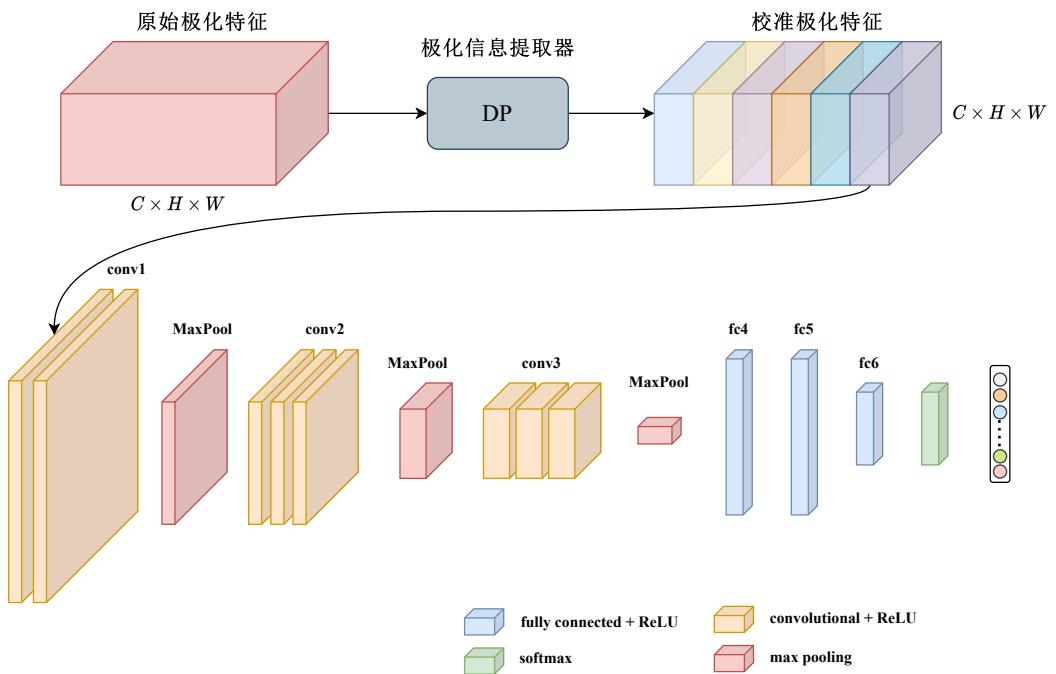


图 1-11 DP-CNN 网络结构示意图

以交叉熵损失函数^[?]为目标, 通过反向传播算法训练模型参数。交叉熵损失时分类问题中最常用的损失函数之一, 衡量了分类模型的预测值与真实标签之间的差异性, 是一种用于优化分类模型的目标函数。

综上所述, 本章的 DP-CNN 方法将输入的原始极化特征 x 映射为预测概率 $p \in \mathbb{R}^C$, 其中, C 表示类别的个数。 x 对应的中心像素预测标签可以通过选择概率最高的类别, 即向量 p 的最大值索引来预测。其计算公式如下:

$$H(p, q) = \sum_{i=1}^n p(x_i) \log(q(x_i)) \quad (1-20)$$

其中, $p(x)$ 表示真实值分布概率, $q(x)$ 表示模型预测分布概率。交叉熵值的变化与

模型的训练效果密切相关，优越的训练效果会让预测概率分布逐渐趋近于真实值概率分布，相应的交叉熵值会逐渐减少。Sigmoid 和 Softmax 损失函数是两个被广泛应用的交叉熵损失函数。Sigmoid 损失函数主要应用于多标签分类任务，其中分类目标可以同时拥有多个标签。这一损失函数模拟了模型对多个独立事件的概率预测，每个事件的概率值落在 $[0, 1]$ 区间内。Softmax 损失函数被广泛应用在多类别分类任务中，其中每个样本仅能关联一个类别。Softmax 函数将预测模型的原始输出转化为表示类别概率的分布，确保所有类别的概率之和为 1，并且模型的输出是互斥的。本章的极化 SAR 图像目标分类任务属于多分类语义分割任务，每个像素都有唯一的正确类别。因此选择多分类交叉熵损失函数，即 Softmax 损失函数作为模型的损失函数，为模型训练提供有力的优化目标。

为了对本章提出的极化信息提取方法进行全面地评估和对比，选择了多种替代方案进行比较，主要涉及两个方面的变化：一方面是对特征输入的处理改变，另一方面是对极化信息提取模块的替代策略。首先在特征输入方面，验证了不同极化特征表示对目标分类任务的影响。利用本章提出的基于双通道注意力的极化信息提取方法结合 CNN 分类模块进行分类、仅使用极化相干矩阵中的元素结合 CNN 分类方法（记为 CNN-T）、仅使用极化目标分解特征结合 CNN 分类方法（记为 CNN-P）、基于散射特征和分解特征简单叠加结合 CNN 的分类方法（记为 CNN-F）作为不同的对比方法，验证本章方法在极化特征表示方面的有效性。其次是在极化信息提取模块层面进行替换，利用基于压缩和激励网络结合 CNN 分类方法（记为 CNN-SE）和基于空间通道注意力结合 CNN 分类方法（记为 CNN-CBAM）作为不同的对比方法，验证本章方法在极化特征表示方面的优越性。在每组实验中，从每个类别选择 1% 的带标签像素，以这些带标签像素为中心，在其周围利用 15×15 的窗口截取图像，形成训练集的样本表示。

1.4.2 精度评价方法

精度评价是对实际数据和模型分类结果进行比较的重要步骤，旨在确定分类模型的准确性，是衡量分类结果可靠性的关键指标。混淆矩阵（Confusion Matrix）通常作为遥感图像分类准确性能的评判指标，并且可以通过混淆矩阵计算得到多种常用的评价参数指标，例如总体分类准确率（Overall Accuracy, OA）、各个类别分类准确率、各个类别平均分类准确率（Average Accuracy, AA）、Kappa 系数等。

混淆矩阵是一个 $n \times n$ 的矩阵，其中 n 表示数据集的类别数量。混淆矩阵的行表示实际类别，列表示预测类别。其中，每个元素 (i, j) 表示实际属于类别 i 的样本被预测为类别 j 的数量。混淆矩阵主对角线元素表示被正确分类的样本，非主对

角线表示分类错误的样本。

在极化 SAR 图像分类结果精度评价中，可以基于混淆矩阵定义以下指标：

1. 总体分类准确率 (OA):

$$OA = \frac{\text{主对角线元素之和}}{\text{混淆矩阵所有元素之和}} \quad (1-21)$$

2. 生产者精度:

$$\text{生产者精度} = \frac{\text{类别对应的主对角线元素}}{\text{类别所在列总和}} \quad (1-22)$$

3. 使用者精度。

$$\text{使用者精度} = \frac{\text{类别对应的主对角线元素}}{\text{类别所在行总和}} \quad (1-23)$$

4. 错分误差。错分误差是指被分类模型错误地划分为用户感兴趣的类别，实际上属于另一类别的样本数量，反映了模型在预测时产生的误报情况。

$$\text{错分误差} = 1 - \text{使用者精度} \quad (1-24)$$

5. 漏分误差。漏分误差是指本应该属于地表真实分类的样本，但是由于模型未能正确分类而被判为其他类别的数量，反映了模型在预测时产生的漏报情况。

$$\text{漏分误差} = 1 - \text{生产者精度} \quad (1-25)$$

6.Kappa 系数 Kappa 系数是一种通过多元统计方法来评价分类精度的指标，旨在量化分类模型的性能相对于完全随机分类的优越性。该系数通过考察混淆矩阵的对角线元素以及总体分布情况，提供了对分类结果误差的全局度量。具体计算公式如下：

$$K = \frac{p_0 - p_e}{1 - p_e} \quad (1-26)$$

其中， p_0 表示总体分类精度，由主对角线元素之和除以所有样本数量计算得到； p_e 表示某一个类别地表真实样本总数与该类中被分类样本总数之积对所有类别求和除以总样本数的平方。将混淆矩阵中的具体元素带入上式，可以得到：

$$K = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{ii} + x_{+i})}{N^2 - \sum_{i=1}^r (x_{ii} + x_{+i})} \quad (1-27)$$

Kappa 系数的大小可以用来表示分类的精度性能，表??描述了 Kappa 系数与模型的分类精度的映射关系。

表 1-2 Kappa 统计值与分类精度映射关系

Kappa 系数	分类精度
<0	较差
0-0.2	差
0.2-0.4	正常
0.4-0.6	好
0.6-0.8	较好
0.8-1	非常好

1.4.3 AIRSAR Flevoland 数据实验

实验数据集选择 NASA/JPL 于 1989 年在 Flevoland 区域采集得到的全极化数据。该数据集是荷兰的一个农业区域遥感数据，作为基准数据集广泛应用于极化 SAR 土地覆盖目标分类研究中。该图像大小为 1024×750 像素，共有 15 种农作物类别，包括茎豆、豌豆、森林、苜蓿、小麦、甜菜、土豆、土豆、裸土、草、油菜籽、大麦、水和少量建筑物。各个农作物目标类别之间的差异较小，相似性较强，因此分类难度较大，容易出现错分漏分的现象。图??和图??分别展示了 AIRSAR Flevoland 数据集的 Pauli 分解伪彩图像以及对应的地面真值标签图像。表??展示了该数据集中每个类别带标签的样本数量。

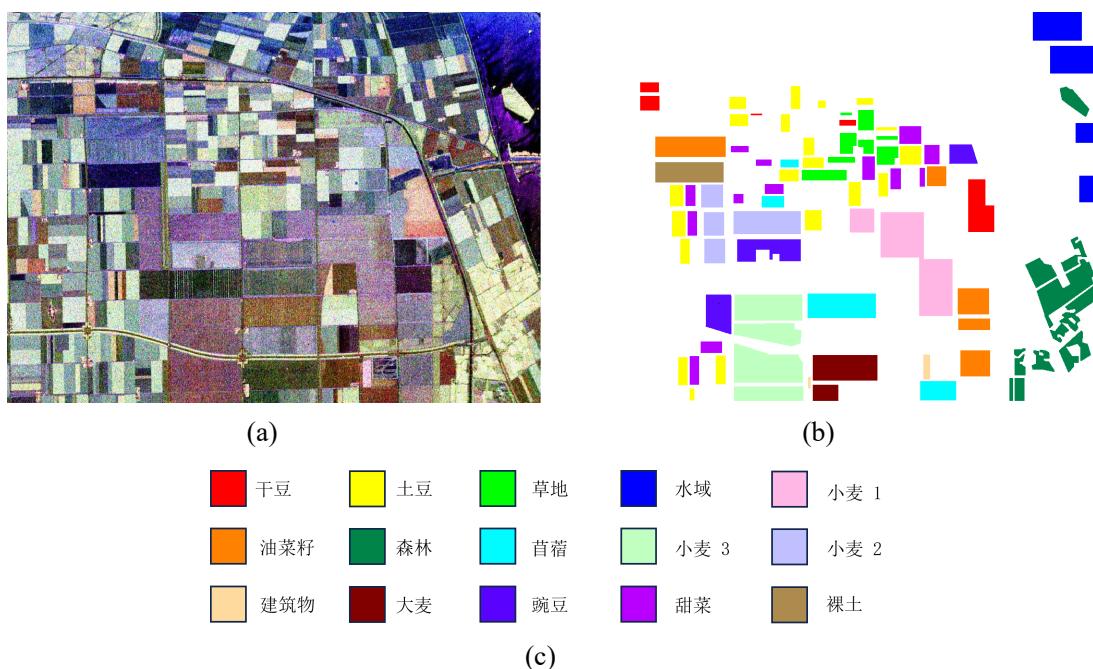


图 1-12 Flevoland 地区实验数据集。(a)Pauli 分解伪彩色图像; (b) 实验数据地面真值; (c) 颜色与类别对应关系

表 1-3 Felvoand 地区实验数据集有标签样本数量

类别	干豆	大麦	裸土	土豆	甜菜	小麦 1	豌豆	苜蓿
数目	6338	7595	5109	16156	10033	11159	9582	10181
类别	草地	小麦 2	油菜籽	小麦 3	建筑物	森林	水域	
数目	7058	16386	13863	22241	735	18044	13232	

图??展示了各个对比方法的可视化分类结果。根据图??，仅使用散射特征的分类方法，在草地、油菜籽区域都有较多的错分样本，这是因为只使用了散射特征而忽略了目标分解特征，没有全面综合利用所有的极化信息导致的。图??展示了仅使用目标分解特征的分类结果，在草地、油菜籽、土豆区域也存在大量的错分样本，这可能是由于没有综合使用极化信息导致的，相干矩阵代表的散射特征是极化 SAR 中最基本、重要的特征。图??展示的简单叠加散射特征与目标分解特征的分类结果，该方法的分类结果几乎与前两种方法相同，仍然在草地、油菜籽区域存在大量错分样本，这反映了直接简单堆叠使用极化特征带来的准确率提升有限。图??与图??展示了使用经典注意力方法的分类结果图，相比于直接叠加特征，并没有带来明显的分类性能提升，少量的错分孤立点与草地中错分的块状区域依然存在，这表明不考虑极化 SAR 数据特征的特征细化方法并不能为极化 SAR 分类任务带来优势。图??展示了基于双通道注意力方法的分类结果，可以看出本方法分类结果更加平滑，错分像素减少，特别是是在小麦、油菜籽区域，类间错分孤立点相对减少，这也验证了本章方法结合散射特征和目标分解特征的有效性，证明本章方法提取的极化信息表示是优越的。

为了进一步探索极化信息提取模块的性能，引入 t-SNE (t-distributed Stochastic Neighbor Embedding) [2] 特征分布图作为衡量极化信息提取模块的可视化评估指标。t-SNE 是一种非线性降维技术，能够有效地将高维数据映射到二维或三维空间，以便更直观地观察样本在特征空间的分布。

图??展示了不同方法特征特征使用 t-SNE 可视化的分布情况。根据图??、图??和图??所展示的仅散射特征、仅分解特征和直接堆叠特征三种不同的特征分布情况，可以看出大多数类别相互之间有交叠的情况，并且类内的样本分布较为散乱，因而容易造成类间错分的情况，导致分类性能优先。而图??和图??展示的使用经典的注意力信息提取方法的特征分布图，可以看出经典的注意力方法由于没有考虑极化 SAR 的数据特性，不同类别之间的重叠情况依然存在，并且类内样本特征分布散乱，证明了在信息提取时不考虑极化 SAR 数据特征对分类模型的性能提升是有限的。根据图??所展示的基于双通道注意力的极化信息提取方法的特征分布图，可以看到相比于其他的特征分布情况，类间的交叠现象有了明显的改善，

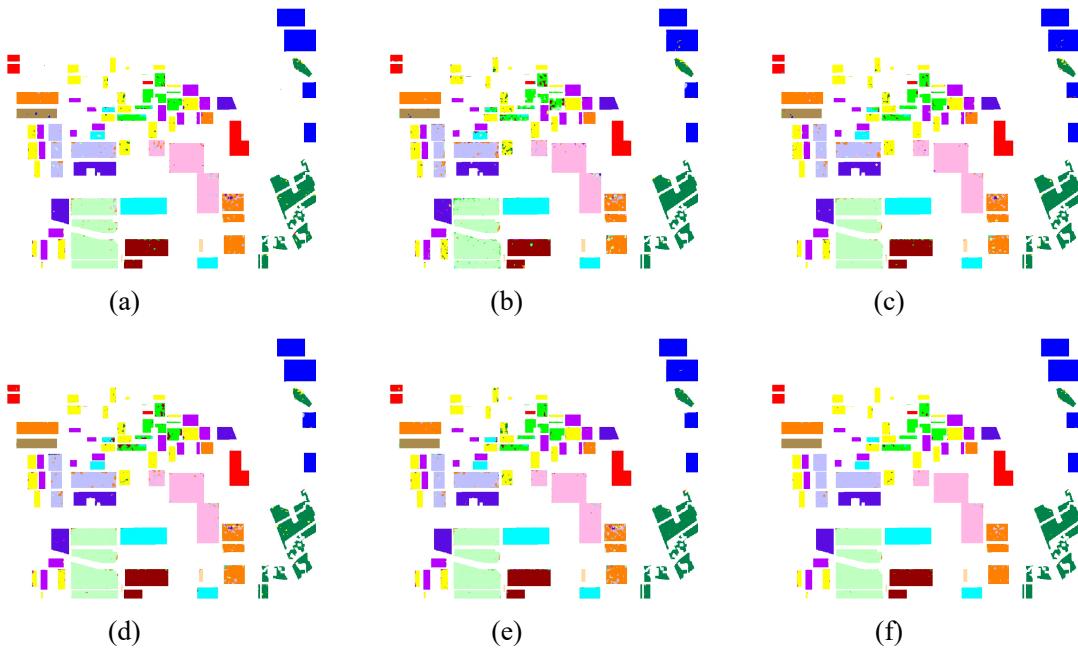


图 1-13 AIRSAR Flevoland 地区数据分类可视化结果图。(a) CNN-T; (b) CNN-P; (c) CNN-F; (d) CNN-SE; (e) CNN-CBAM; (f) 本章方法

同时类内样本特征分布更加紧凑，提升类间可分性，进而带来分类性能上的提升。

表??展示了不同方法的分类数值结果。仅使用散射特征的分类方法，在草地和建筑物区域的分类准确率较低，分别为 89.41% 和 91.17%，这可能是因为仅使用相干矩阵作为极化特征的分类方法，并不能完全反应这两类地物目标的散射特性，导致分类准确率较低。仅使用目标分解特征作为特征的分类方法，在干豆区域达到了 99.66% 的准确率，这是因为目标分解特征中包含有具体物理意义的极化表征，但是在建筑物油菜籽区域的分类准确率较低，仅有 90.2%，反映了目标分解特征与散射特征相互补充，都是必不可少的极化特征表示方法。直接堆叠极化特征的分类方法在分类总体准确率、平均准确率以及大麦等区域准确率反而下降，这表明了直接堆叠极化特征导致特征维数增加，分类器无法有效分清两种类型的主次关系，导致分类准确率有所下降。经典的注意力方法 CNN-SE 与 CNN-CBAM 总体准确率分别为 96.76% 和 96.13%，分类结果准确率指标仅有不足 1% 的提升，表明该方法直接应用于极化 SAR 领域性能提升有限。根据本章方法的分类数值结果，在大多数的类别都具有最高的分类准确率，并且在总体准确率、平均准确率、Kappa 系数分别提升 2.25%、3.12% 和 2.46%，表明了基于双通道注意力的极化信息提取方法对于极化 SAR 目标分类具有优势，验证了本章方法的有效性和优越性。

图??展示了本章方法分类结果的混淆矩阵可视化图。从图中可以看到，混淆

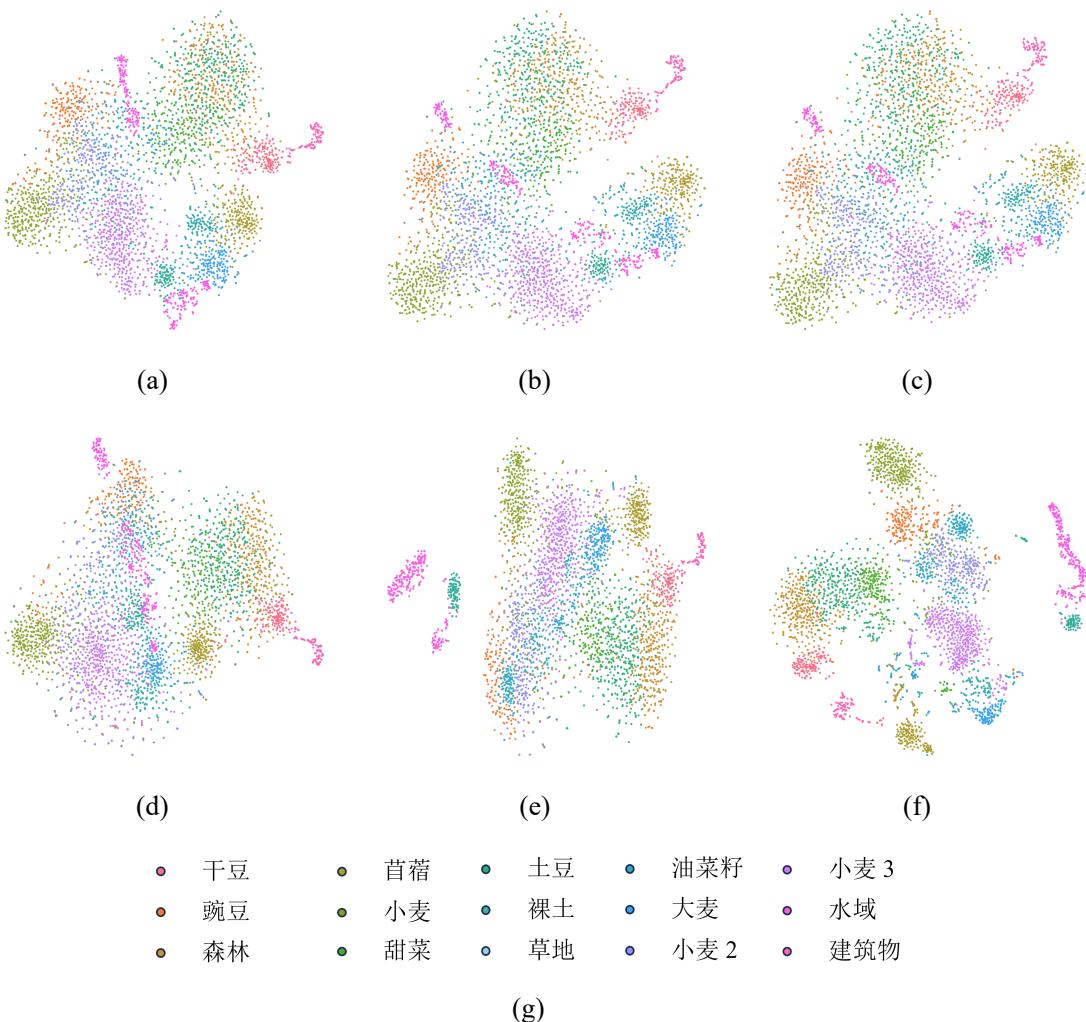


图 1-14 AIRSAR Flevoland 地区数据不同方法特征分布。((a) CNN-T; (b) CNN-P; (c) CNN-F; (d) CNN-SE; (e) CNN-CBAM; (f) 本章方法; (g) 颜色与类别对应关系

矩阵的可视化结果显示本章方法在对角线上具有较高的主对角元素，说明在大多数类别上取得了良好的分类性能。

1.4.4 E-SAR Oberpfaffenhofen 数据实验

该数据集是通过 E-SAR 机载平台在德国 Oberpfaffenhofen 地区拍摄获取到的 L 波段极化 SAR 数据。该数据集经过多视处理，具有高质量的全极化信息，是极化 SAR 目标分类研究的经典数据集之一。数据集在方位向和距离向具有 $3m \times 3m$ 的分辨率，大小为 1300×1200 ，总计包含三种类型的地物目标，分别是：建筑、林地和开放区域。图??展示了该数据集的 Pauli 伪彩色图像，图??展示了该数据

表 1-4 AIRSAR Flevoland 地区数据分类数值结果 (%)

序号	类别	CNN-T	CNN-P	CNN-F	CNN-SE	CNN-CBAM	CNN-DP
1	建筑物	91.17	90.2	94.01	99.79	94.14	98.83
2	油菜籽	94.33	91.78	89.78	94.54	92.32	98.29
3	甜菜	96.57	96.62	96.99	97.92	95.74	98.92
4	干豆	99.66	99.26	98.86	97.08	98.83	99.58
5	豌豆	96.72	97.03	96.54	99.8	95.72	99.25
6	森林	98.13	97.37	96.42	98.85	98.61	99.08
7	苜蓿	96.03	95.96	95.07	96.11	95.48	97.86
8	土豆	96.61	97.03	97.54	97.33	95.59	99
9	裸土	95.02	97.19	98.44	98.51	96.8	99.67
10	草地	89.41	93.11	92.32	96.19	94.25	97.81
11	大麦	93.63	95.45	87.34	95.64	89.34	97.7
12	水域	99.94	99.7	99.66	93.3	99.83	99.99
13	小麦 1	97.53	96.74	97.36	97.27	97.6	97.92
14	小麦 2	93.62	94.48	94.21	97.53	92.25	96.53
15	小麦 3	96.88	97.05	98.06	94.85	98.51	99.36
OA		96.4	96.41	95.9	96.76	96.13	98.65
AA		95.68	95.93	95.51	96.98	95.67	98.65
Kappa		96.07	96.08	95.53	96.48	95.77	98.53

集的人工标记的地面真值参考图。地面真值图中黑色的区域表示未标记的区域。表??展示了各类地物目标样本数量情况。

表 1-5 ESAR Oberpfaffenhofen 数据样本数量

类别	建筑	林地	开放区
数目	269,184	388,503	779,962

图??展示了不同方法的分类结果可视化图像。根据图??展示的仅使用极化相干矩阵为特征的分类结果图，尽管在开放区域的分类效果较好，但是在建筑物区域存在大量被错分的样本，将建筑物区域样本错误分类为开放区域，并且错分的现象较为严重。这可能是因为仅使用极化相干矩阵作为特征时并没有完全利用其他的目标分解特征导致的，说明了综合考虑极化 SAR 图像中的其他极化特性和辅助特征的在分类任务中的重要性。相比之下，图??与图??在建筑物区域中的错分样本相对较少，但是在各类区域中也存在着部分错分的样本。其中将开放区域中的少数像素分类为建筑物的错分情况依然存在，这反映了极化目标分解特征提供的反映目标物理散射特性的信息为分类任务带来了一定的优势，是极化 SAR 图像中的重要的基本特征。图??和图??展示了基于经典注意力方法的特征表示分类结

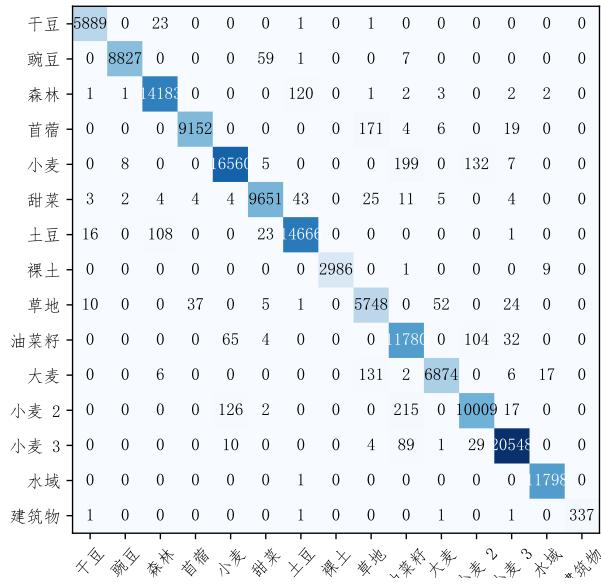


图 1-15 本章方法在 Flevoland 图像中分类结果混淆矩阵图

果，从结果图中可以看出，分类效果仅有微小的变化，错分像素相对减少但是依然存在，这可能是极化 SAR 图像与光学图像存在较大的差异，经典的注意力方法并没有考虑极化 SAR 图像的数据特征特点。图??展示了基于本章提出的双通道注意力极化信息提取方法的分类结果图，从图中可以看出，在建筑物区域的错分样本明显减少，区域内部分类更加平滑，并且每个类中的小块错误区域数量也相对减少，整体分类结果得到明显改善，证明了本章方法的有效性。

图??展示了不同方法数据特征使用 tSNE 可视化的分布情况。根据图??、图??、图??展示的仅散射特征、仅分解特征和直接堆叠的三种特征分布情况，可以看出三种特征分布情况并没有本质性的差异，尽管都可以看出各类之间存在不同的分布情况，但是也能清晰看出，类间存在分布交叠的部分，这也导致了部分样本的错分情况。根据图??和图??展示的使用经典注意力信息提取方法的特征分布图，可以看出虽然类间交叠的情况有细微改善，但是依然有大量的特征存在交叠，这也映照了分类结果改善不大的结果。图??展示了使用本章方法的特征分布情况，可以看出类间交叠的情况得到优化，各类之间分布间距增大，增强了类间可分性，进而带来分类性能的提升。

表??展示了各个方法的分类数值结果。本章提出的基于双通道注意力的极化信息提取方法总体准确率达到 94.6%，相比于其他验证方法提升了 3.73%。本章方法的散射特征时，在建筑物区域的分类准确率仅为 63.16%，这可能是因为建筑物区域目标散射回波差异性大，导致仅使用散射特征时，不能完全表征建筑物目标的特性，进而导致在建筑物区域可分性较低。与仅使用目标分解特征相比，直接

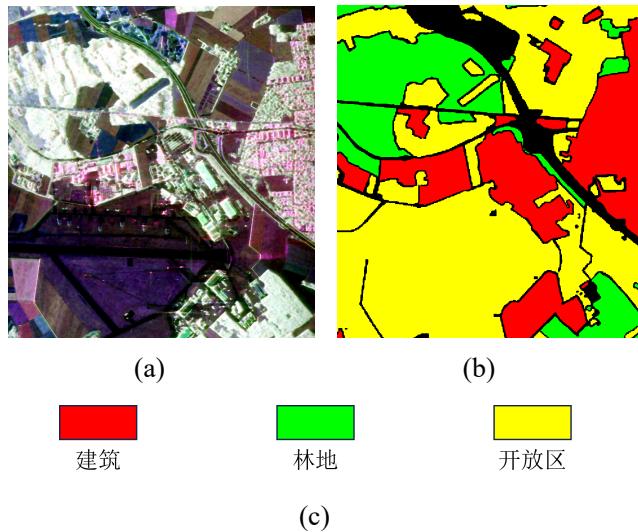


图 1-16 E-SAR Oberpfaffenhofen 区域数据。(a)Pauli 分解伪彩色图像; (b) 实验数据地面真值; (c) 颜色与类别对应关系

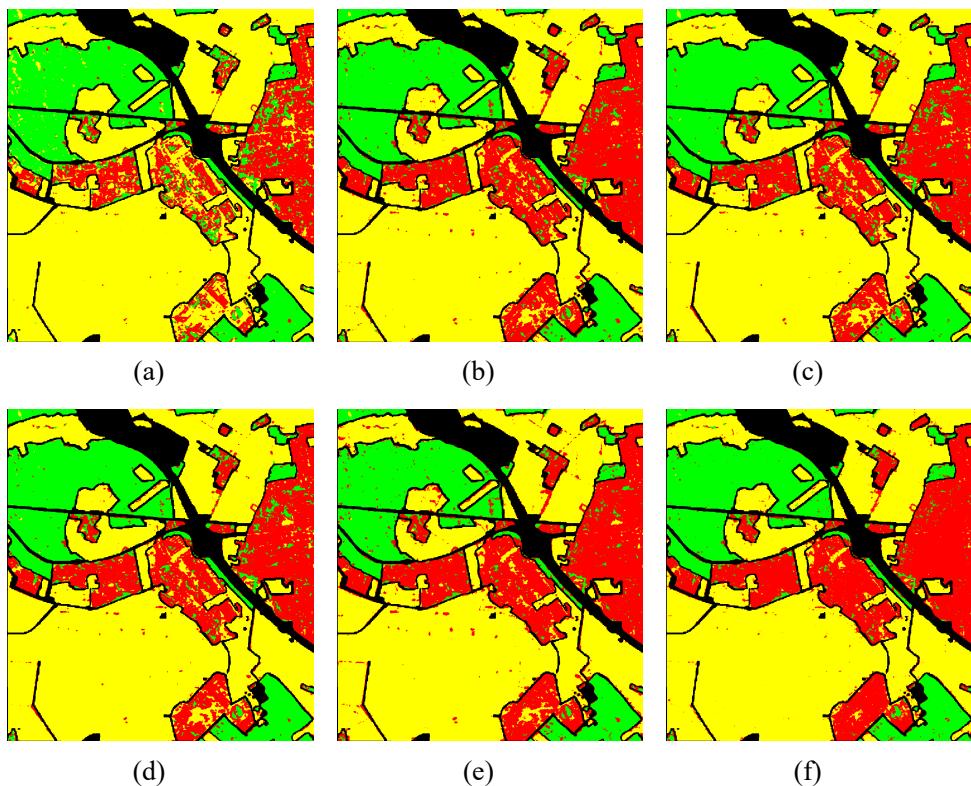


图 1-17 E-SAR Oberpfaffenhofen 地区数据分类可视化结果图。(a) CNN-T; (b) CNN-P; (c) CNN-F; (d) CNN-SE; (e) CNN-CBAM; (f) 本章方法

堆叠极化特征在平均准确率指标反而下降 1.32%，这可能是由于该方法仅仅是对两类的特征进行简单的堆叠，堆叠后的特征维数过高，分类模型对于高维特征输

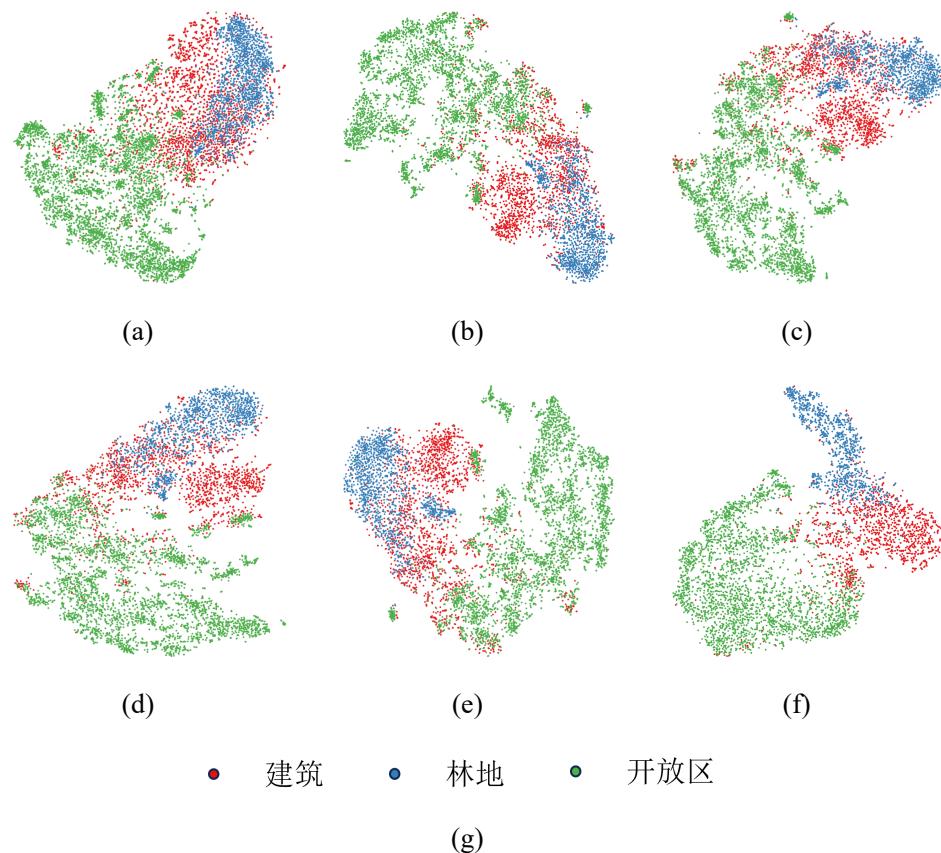


图 1-18 E-SAR Oberpfaffenhofen 地区数据地区数据不同方法特征分布。(a) CNN-T; (b) CNN-P; (c) CNN-F; (d) CNN-SE; (e) CNN-CBAM; (f) 本章方法

入无法有效利用两种不同类型特征间的主次关系。本章方法在总体准确率、平均准确率、Kappa 系数等指标上均有提升，验证了本章方法能有效的提升极化特征的表征方式，提升分类准确度。

图??展示了本章方法分类结果的混淆矩阵可视化图。混淆矩阵的可视化结果显示本章方法在对角线上具有较高的主对角元素，说明在大多数类别上取得了良好的分类性能。

表 1-6 E-SAR Oberpfaffenhofen 地区数据分类数值结果 (%)

序号	类别	CNN-T	CNN-P	CNN-F	CNN-SE	CNN-CBAM	本章方法
1	建筑	96.68	83.9	80.1	80.42	83.92	90.47
2	林地	83.35	91.99	90.5	91.44	91.61	92.61
3	开放区域	89.83	90.04	91.37	91.63	91.41	96.78
	OA	89.56	88.97	88.52	89.13	90.27	94.39
	AA	89.95	88.64	87.32	87.83	88.98	93.28
	Kappa	81.63	83.51	82.73	83.58	85.24	90.4

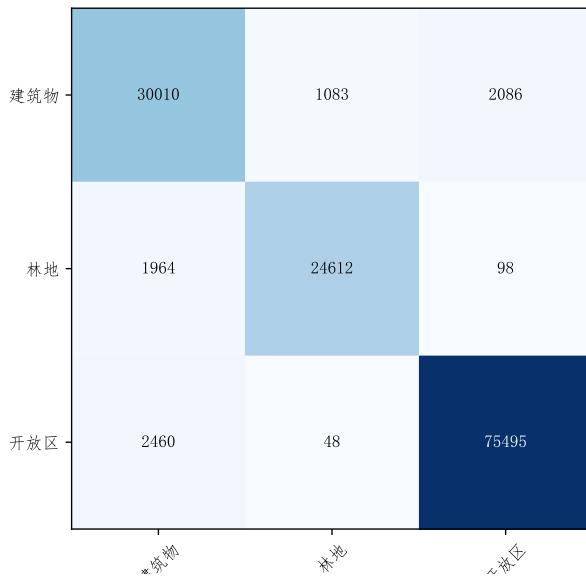


图 1-19 本章方法在 E-SAR Oberpfaffenhofen 图像中分类结果混淆矩阵图

1.5 本章小结

本章针对极化 SAR 有效信息提取问题，提出了一种基于双通道注意力的极化信息提取方法。首先，构建后向散射特征通道和目标分解特征通道的双通道结构，结合空间、通道注意力机制，实现对两种类型极化特征中有效信息的激发而抑制无关信息。然后，为了充分挖掘两类极化特征之间的关联信息，设计了基于散射特征一致性的注意力修正方法，对注意力特征进行动态校正。最后，通过多尺度特征学习模块，聚合不同空间尺寸的极化特征，增强模型对不同空间尺度的特征的感知能力。最后，在实验阶段搭建了端到端的信息提取与分类网络，并从视觉分类效果和数据量化准确率两个层面进行性能评估，结果表明本章提出的信息提取方法能为分类任务提供更加全面、充分的极化信息表示方式，提升极化 SAR 图像的分类精度。

致 谢

感谢党和国家