

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 基于深度学习的极化 SAR 图像目标分类方法研究

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

基于深度学习的极化 SAR 图像目标分类方法研究

(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 2024 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 Target Classification Method of Polarimetric SAR Images Based on Deep Learning**

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

极化合成孔径雷达（Polarimetric Synthetic Aperture Radar, PolSAR）是一种主动式微波遥感观测传感器，具有全天时、全天候的成像特点。极化 SAR 图像分类方法作为极化 SAR 图像解译中的关键技术，在军事、民用领域具有重大的应用价值。

随着极化 SAR 技术的发展，高分辨、多极化的 SAR 图像数据日益增多，对极化 SAR 图像分类起到了积极作用，但也引入了新的挑战。一方面，单一极化特征无法同时满足所有目标分类需求，但多种极化特征在表示目标特性中存在信息冗余，这要求目标分类方法能对极化 SAR 图像进行有效多极化特征融合。另一方面，极化 SAR 图像数据标注过程中不可避免引入错误标记样本，这要求目标分类方法具备标签噪声鲁棒能力。本文针对以上问题，开展了相关基础研究与方法创新，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 SAR 雷达信号的不同极化特性及常用的表征方式。在此基础上，介绍了针对目标散射特性的极化散射矩阵和二阶统计矩阵两种描述方法。最后，介绍了多种相干与非相干目标分解方法，并总结归纳了不同极化参数的物理含义，为后续研究建立理论基础。

(2) 针对多类型极化特征信息冗余的问题，提出了基于双通道注意力的极化 SAR 图像分类方法。通过构建双通道、多尺度特征融合结构，利用极化 SAR 图像不同极化类型、不同空间位置以及不同特征尺度下的特征相关性，实现了对多类型极化特征的一致信息提取和融合利用，并在此基础上构建深度分类模型，有效提高了极化 SAR 图像的分类准确率。

(3) 针对极化 SAR 图像中标签噪声问题，提出了基于有限混合模型的极化 SAR 图像分类方法。基于有限混合模型理论，建立了噪声标签和准确标签混杂下的损失分布模型，获取对极化 SAR 图像中噪声标签的分布估计。此外，通过结合极化伪装图边缘信息提取方法，使分类模型能够有效利用边缘知识增强识别性能，有效提升了标签噪声下的极化 SAR 图像分类准确率。

利用实测极化 SAR 数据对本文提出的方法进行验证。结果表明，本文所提方法能够有效地应对极化 SAR 图像目标分类中存在的多极化特征信息融合利用和标签噪声下的稳健识别问题，有效提高了极化 SAR 图像的分类准确率。

**关键词：**极化合成孔径雷达，图像分类，多极化特征，有限混合模型



## ABSTRACT

Polarimetric Synthetic Aperture Radar (PolSAR) is an active microwave remote sensing technique known for its all-weather, all-day imaging capabilities. Due to its multi-channel, multi-polarization working characteristics, PolSAR images contain rich scattering information of targets. PolSAR image interpretation techniques have been widely applied in various fields such as military detection, disaster analysis, and urban planning.

In recent years, the increasing availability of high-resolution, multi-polarized SAR image data has provided richer polarization information of ground targets for complex scene analysis, which has positively impacted PolSAR interpretation efforts but also introduced new challenges. Directly stacking all types of polarization information leads to information redundancy, while single-type polarization information is inadequate for all classification targets. This necessitates effective utilization of multi-type polarization information in feature representation methods. Additionally, label noise introduced during PolSAR dataset annotation requires classification methods to accurately classify samples with partial mislabeling. This thesis addresses the aforementioned issues through relevant fundamental research and innovations, with the following main research contents:

(1) Investigation of PolSAR-related theoretical foundations. Firstly, characterization methods of electromagnetic wave polarization properties are introduced. Subsequently, several methods for describing target scattering characteristics and various polarization matrices are discussed. Finally, the theoretical explanation of polarization target decomposition methods is presented, along with a summary of the physical meanings of different polarization parameters.

(2) Addressing the problem of redundancy in multi-type polarization feature information, a PolSAR target classification method based on dual-channel attention is proposed. By constructing a dual-channel PolSAR feature input network structure and employing spatial and channel attention modules, as well as multi-scale learning methods, different types of polarization features are refined and mutual information is extracted. Moreover, aggregation of polarization features at different scales is performed, effectively avoiding the limitations of single features and reducing the redundancy of multi-feature information, thereby improving PolSAR target classification accuracy.

(3) In response to label noise issues, a robust PolSAR target classification method

---

## ABSTRACT

---

under label noise is proposed. By establishing the relationship between the loss distribution differences of noisy and clean samples under a deep neural network model, the noise probability estimation of samples is obtained. Additionally, combined with boundary sample loss enhancement, robust parameter optimization under label noise is achieved. Furthermore, full utilization of boundary information is implemented, resulting in improved PolSAR target classification accuracy under label noise.

The effectiveness and superiority of the proposed methods are validated using real PolSAR data. Results demonstrate that these methods effectively address the problems of redundancy in multi-polarization feature information and label noise in PolSAR image classification, achieving high-accuracy classification of ground targets.

**Keywords:** Polarimetric Synthetic Aperture Radar (PolSAR), PolSAR image, polarization feature, target classification, label noise

## 目 录



# 第一章 基于双通道注意力的极化 SAR 目标分类方法

## 1.1 引言

极化 SAR 通过以不同的极化方式发射和接收电磁波信号，从而实现对地物目标探测，并提供丰富的后向散射信息。通过使用四种极化方式（HH、HV、VH、VV）进行探测，可以获得包含目标散射特性的散射矩阵。使用具有不同物理含义的散射基进行进一步解析，可以获取代表目标不同散射意义的目标分解特征。不同的极化特征从不同的角度描述地物目标的散射特性。挖掘极化 SAR 数据中的有效极化信息一直是遥感探测领域的难点与重点之一。通过采用适合于极化 SAR 数据特性的特征提取技术，可以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之一。

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。而目前在极化 SAR 目标分类研究中，主要存在以下问题：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 [? ]。

本章提出了一种基于双通道注意力的极化 SAR 目标分类方法，讨论了多类型极化特征信息的利用方法。该方法以注意力机制为基础，通过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，构建了两种类型特征之间的融合表征关系，从而增强了极化信息的表征性能，并提高了分类任务的准确率。实验部分通过两组真实极化 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）<sup>[?]</sup>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， $\sigma$  表示 ReLu 激活函数<sup>[?]</sup>， $\mathbf{W}_1 \in \mathbb{R}^{\frac{C}{r} \times C}$  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。

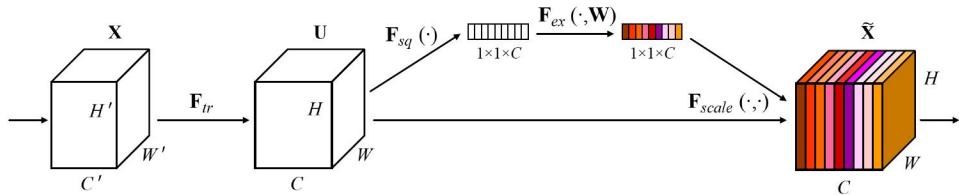


图 1-1 SENet 结构图<sup>[?]</sup>

### 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）<sup>[?]</sup>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  $7 \times 7$  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<sup>[?]</sup>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：

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

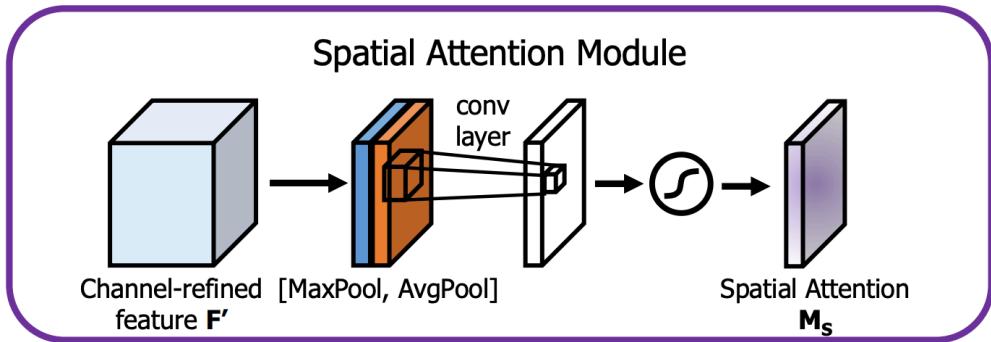


图 1-2 SAM 结构图<sup>[?]</sup>

### 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）<sup>[?]</sup>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,  $\sigma$  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。

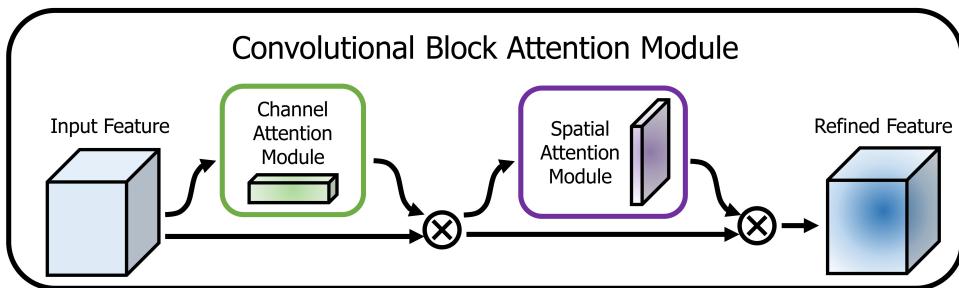


图 1-3 CBAM 结构图 [?]

### 1.3 双通道注意力的极化 SAR 目标分类方法

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。区别于光学图像, 极化 SAR 图像具有独特的数据特征, 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。本章充分考虑了极化 SAR 图像数据特点, 基于后向散射信息与目标分解信息的差异性与互补性, 提出了专适用于极化 SAR 图像的目标分类方法, 即基于双通道注意力的极化 SAR 目标分类方法。该方法充分挖掘两类极化特征的差异信息与互补信息, 以提供更加全面、充分的极化信息表示方式, 从而提升分类准确率。

#### 1.3.1 双通道注意力极化 SAR 目标分类方法框架

如图??所示, 展示了基于双通道注意力的极化 SAR 目标分类算法示意图。该方法构建了端到端的极化特征校准与分类网络结构, 以挖掘多类型极化特征中的隐藏信息为目的, 实现两类型极化特征的信息提取与有效融合, 增强极化特征的可表征性能。

该方法主要分为极化特征校准模块与分类器模块, 分别用于多类型极化特征

细化以及目标分类任务。其主要思路是：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依次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，设计多尺度学习方法，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输入到分类器网络，得到目标分类结果。值得注意的是，该方法中的极化特征校准模块可以作为一个即插即用的插件式模块，应用到其他下游的目标检测、分类识别任务中。

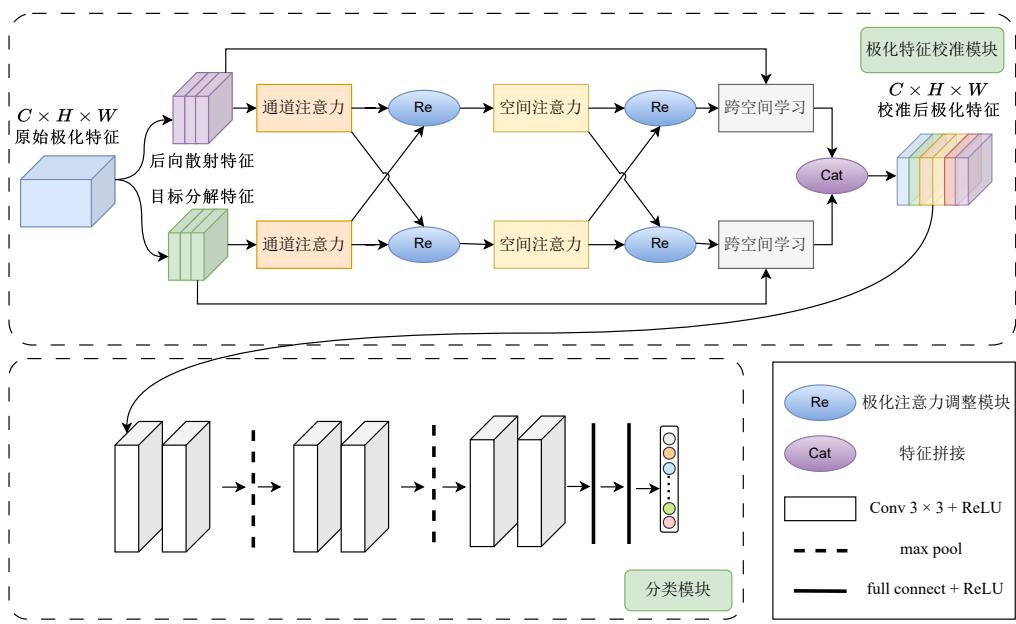


图 1-4 基于双通道注意力的极化 SAR 分类方法示意图

极化特征校准模块输入结构呈现两个独立通道形式，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,  $\text{Re}$  与  $\text{Im}$  分别表示取实部与取虚部运算。 $V$  包含了地物目标完备的后向散射信息, 将  $V$  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	$H, \alpha, A$	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	$P_s, P_d, P_v, P_r$	表面散射、体散射、二次散射、同极化比	4
	$f_s, f_d, f_v$	Freeman 系数	3
Krogager	$ k_s ^2,  k_d ^2,  k_h ^2$	球、双平面、螺旋散射能量	4
Huynen	$A_0, B_0 + B, B_0 - B$	对称、非对称、不规则性信息	3
	$C, D, E$	线性、弯曲、扭转性信息	3
	$F, G, H$	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, 实现由一类极化信息到另一类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, 设计了多尺度学习模块, 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, 实现多尺度的极化特征学习。多尺度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后作为分类器模块输入, 完成目标分类。

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。

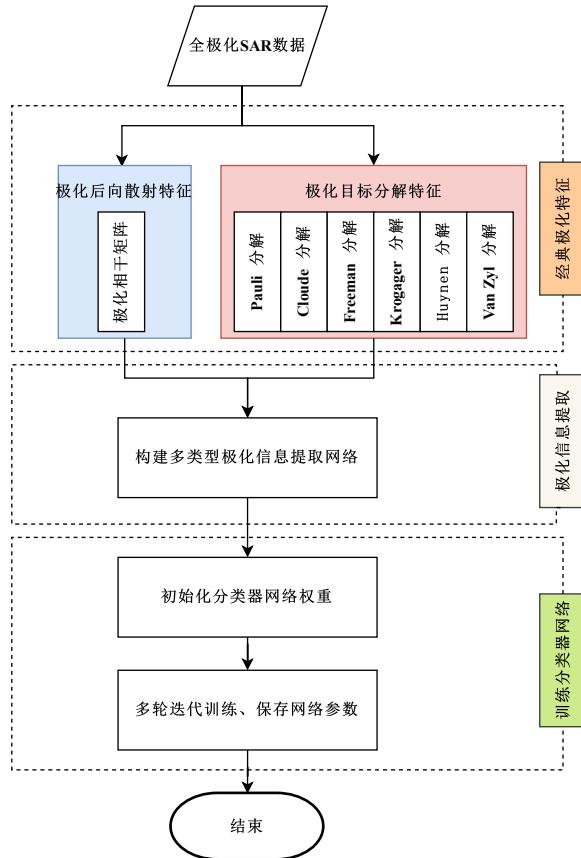


图 1-5 双通道注意力极化 SAR 目标分类算法流程图

### 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：

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

其中， $f^{7 \times 7}$  表示卷积核大小为  $7 \times 7$  的卷积操作； $\sigma$  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  $S_S$ 。

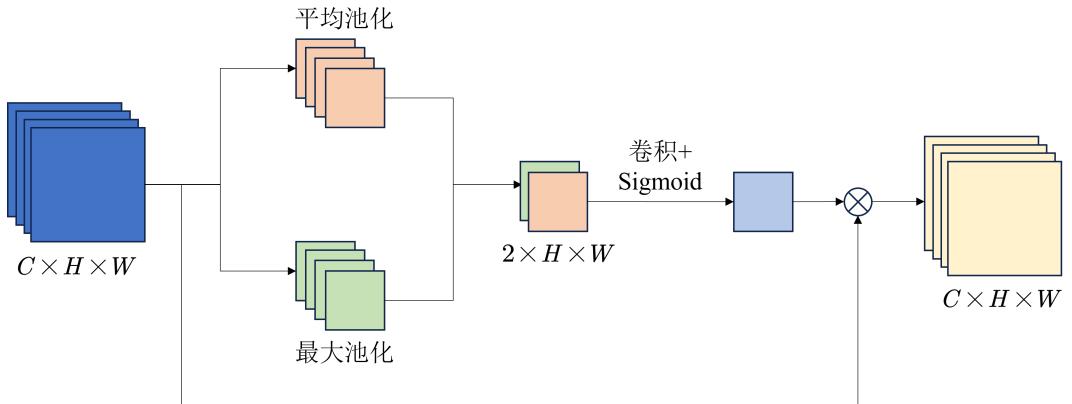


图 1-6 空间注意力模块结构

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：

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

其中， $\sigma$  表示 Sigmoid 激活函数； $\Gamma$  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。

### 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

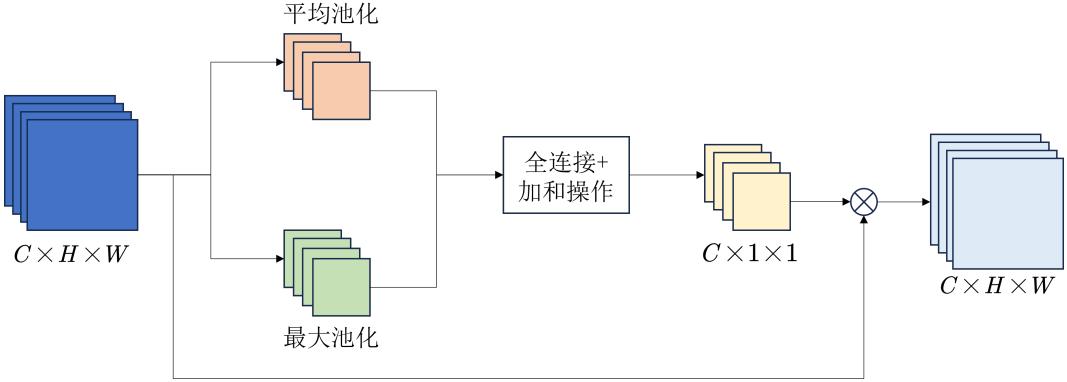


图 1-7 通道注意力模块结构

一致性为引导，对输入的极化特征进行动态调整，增强特征一致性。极化注意力修正模块利用极化特征估计特征映射的位置相关关系，自适应地控制两类注意力图的空间位置信息。以双通道的注意力图  $S_1 \in R^{H \times W}$  与  $S_2 \in R^{H \times W}$  为输入，该模块的计算流程可以表示为：

$$S'_2 = f^{1 \times 1}(S_2) \quad (1-13)$$

$$S'_1 = \text{Ref}(S'_2, \phi(S_1)) \quad (1-14)$$

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

其中， $S_{out}$  表示模块最终输出，运算  $\phi(\cdot)$  表示在特征  $S_1$  引导下估计  $S_2$  的位置相关矩阵，描述了两路特征的空间对应关系， $\text{Ref}(\cdot)$  表示根据位置相关矩阵调整  $S_2$ ，运算  $\phi(\cdot)$  采用类似 softmax 运算操作计算归一化相关矩阵值，具体可以表示为：

$$\Phi_i(j) = \frac{\exp(\theta(S_i)^T \cdot \varphi(S_j))}{\sum_j \exp(\theta(S_i)^T \cdot \varphi(S_j))} \quad (1-16)$$

其中， $i$  是指相关矩阵中某个位置的索引， $j$  是枚举所有位置的索引。 $\Phi_i(j)$  描述了空间位置  $i$  与位置  $j$  之间的对应关系。 $\theta(\cdot)$  与  $\varphi(\cdot)$  均使用卷积核大小为  $1 \times 1$  的卷积层实现。

对于输入注意力特征图位置索引  $i$  处的特征单元，以对应位置的相关元素  $\Phi_i$  为引导，修正过程可以表示为：

$$S'_{1i} = \sum_j \Phi_i(j) \cdot S'_2(j) \quad (1-17)$$

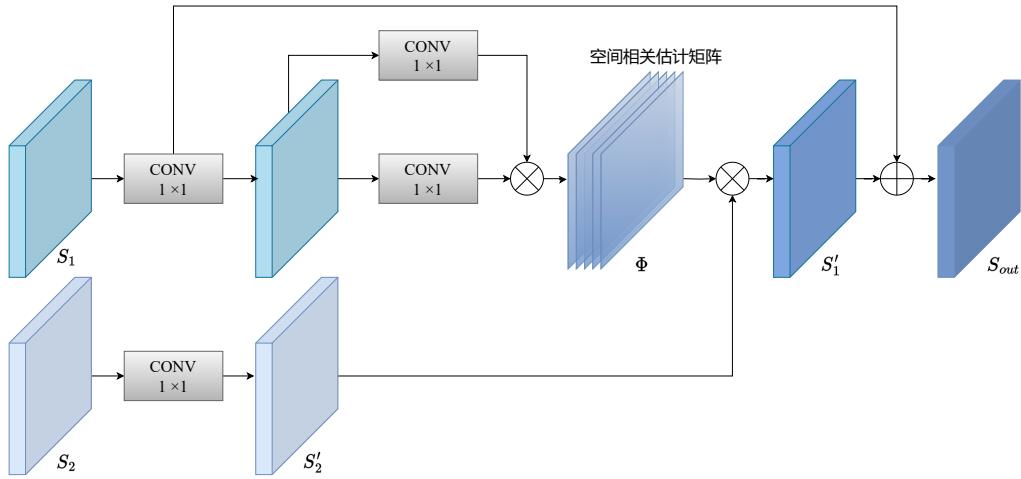


图 1-8 极化注意力修正模块网络结构示意图

### 1.3.4 多尺度学习模块

多尺度学习模块的网络结构如图??所示，提供了一种不同空间维度方向的极化信息聚合方法，以实现多尺度下的极化特征聚合。引入两个分支的张量，分别代表了 $1 \times 1$ 分支的输出和 $3 \times 3$ 分支的输出。通过二维全局平均池化对 $1 \times 1$ 分支的输出进行操作，以编码其中的全局极化空间信息，用于对全局信息进行编码和建模远程依赖关系。二维全局池化操作可以表达为：

$$z_c = \frac{1}{H \times W} \sum_j^H \sum_i^W x_c(i,j) \quad (1-18)$$

$1 \times 1$ 分支特征图在进行二维全局平均池化后，采用自然非线性函数 Softmax 进行二维高斯映射，以拟合线性变换并提升计算效率。通过将上述并行处理的输出进行矩阵点积运算，并与结果相乘，得到了第一个空间注意力图。类似地，利用二维全局平均池化对 $3 \times 3$ 分支进行全局空间信息的编码。这会生成每组内输出特征的两个空间权重值的集合，随后使用 Sigmoid 函数将其映射为相应空间位置的权重关系，突出显示所有像素的全局上下文信息。计算公式如下：

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 $3 \times 3$ 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：

$$S_1 = f^{1 \times 1}(S) \times \text{AvgPool}(f^{3 \times 3}(S)) \quad (1-19)$$

$$S_2 = f^{3 \times 3}(S) \times \text{AvgPool}(f^{1 \times 1}(S)) \quad (1-20)$$

$$S' = \sigma(S_1 + S_2) \quad (1-21)$$

其中， $\sigma$  表示 sigmoid 函数。

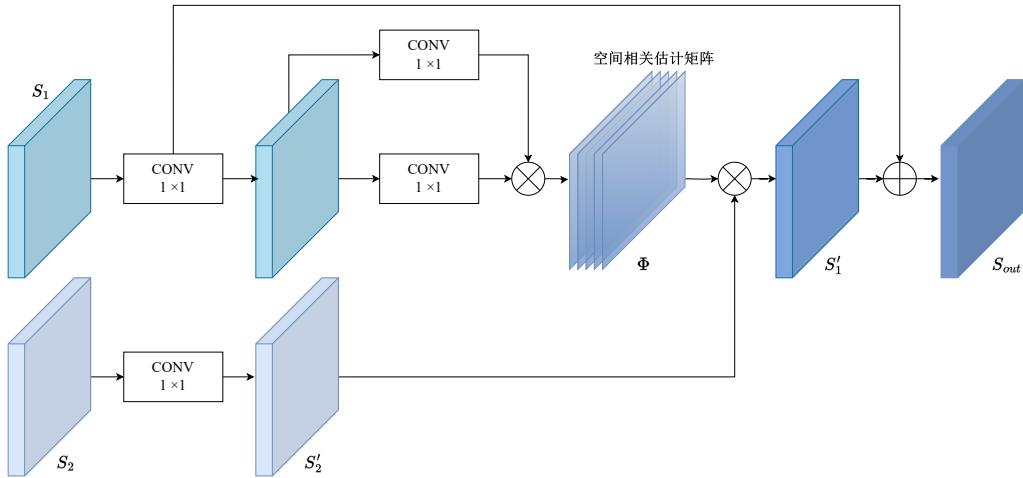


图 1-9 多尺度学习模块网络结构示意图

### 1.3.5 目标分类器模块

在双通道注意力极化信息提取器（Dual Attention Polarization Information Extractor, DP）的基础上，构建端到端的极化 SAR 信息提取与目标分类网络（DP Convolutional Neural Network, DP-CNN）。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，输入到目标分类器模块，通过卷积网络结构，实现对输入特征的分类。校准后的极化特征输入到图??展示的分类模块中。分类器采用类似 VGG 结构的卷积神经网络架构，其中包括三个带有 ReLU 激活函数的  $3 \times 3$  卷积层、两层最大池化函数、两层全连接网络以及一层 SoftMax 激活层。

以交叉熵损失函数<sup>[?]</sup>为目标，通过反向传播算法训练模型参数。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：

$$L = \sum_{i=1}^n p(x_i) \log(q(x_i)) \quad (1-22)$$

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。

整个模型的损失计算流程可以表示为：

- (1) 神经网络分类器最后一层输出每个样本类别的属性值向量；
- (2) 利用 Softmax 函数将属性值向量映射为样本属于每一类的预测概率向量；
- (3) 地面真值标签的 one hot 编码与样本预测概率向量进行交叉熵损失函数运算，并依据损失值进行反向传播优化。

## 1.4 实验结果与分析

### 1.4.1 实验参数设置

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  $3 \times 10^{-4}$ ，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。实验中，使用大小为  $15 \times 15$ ，步长为 1 的滑动窗口对输入图像进行随机截取，选取出 1% 的中心带标签样本作为训练集，而其余的样本作为测试集。

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，从两个对比维度选择了多种替代方案进行比较：一是对特征输入的改变，二是对极化特征校准模块的替换。首先，验证了不同极化特征表示对目标分类任务的影响。对比方法包括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）、以及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）。这旨在验证本章方法在极化特征表示方面的有效性。其次，对极化特征校准模块进行替换。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。这旨在验证本章方法在极化特征表示方面的优越性。

### 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，包括生产者精度（Producer Accuracy, PA）、使用者精度（User Accuracy, UA）、总体分类准确率（Overall Accuracy, OA）、平均分类准确率（Average Accuracy, AA）、Kappa 系数等。

如表??所示， $c$  类样本混淆矩阵是一个  $c \times c$  的矩阵，其中  $c$  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。

表 1-2  $c$  类样本混淆矩阵示意表

		地面真实类别				
		第 1 类	第 2 类	...	第 $c$ 类	预测样本数
预测类别	第 1 类	$n_{11}$	$n_{12}$	...	$n_{1c}$	$n_{1+}$
	第 2 类	$n_{21}$	$n_{22}$	...	$n_{2c}$	$n_{2+}$
	:	:	:	..	:	:
	第 $c$ 类	$n_{c1}$	$n_{c2}$	...	$n_{cc}$	$n_{c+}$
地面真值样本数		$n_{+1}$	$n_{+2}$	...	$n_{+c}$	$N$

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 表示模型正确分类的样本数量占总样本数的比例，衡量分类模型的总体性能水平，计算公式如下：

$$OA = \sum_{i=1}^c \frac{n_{ii}}{N} \quad (1-23)$$

#### 2. 生产者精度（PA）：

PA 是指针对某个具体类别，模型正确预测该类别的样本数占该类别预测样本数的比例，衡量模型对特定类别的分类准确度，计算公式如下：

$$PA = \frac{n_{ii}}{n_{+i}} \quad (1-24)$$

### 3. 使用者精度 (UA):

UA 是指针对某个具体类别，模型正确预测该类别的样本数占该类别实际样本数的比例，衡量模型对特定类别的分类准确度，计算公式如下：

$$UA = \frac{n_{ii}}{n_{i+}} \quad (1-25)$$

### 4.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。计算公式如下：

$$\text{Kappa} = \frac{N \sum_{i=1}^c n_{ii} - \sum_{i=1}^c n_{i+} n_{+i}}{N^2 - \sum_{i=1}^n n_{i+} n_{+i}} \quad (1-26)$$

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 1-3 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  $1024 \times 750$  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、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。

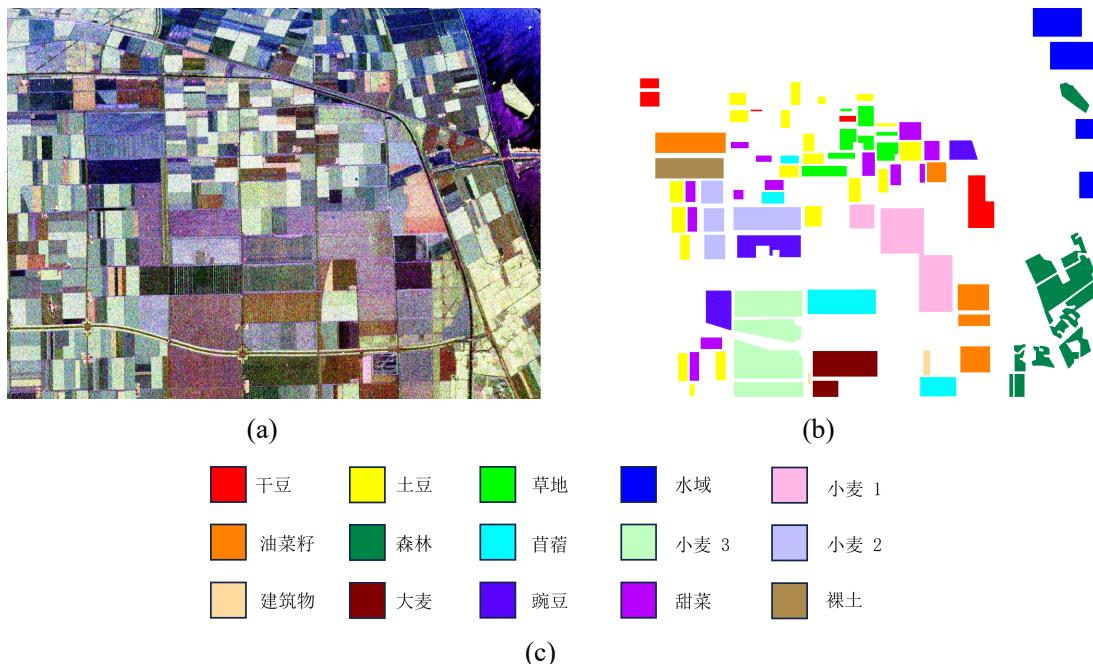


图 1-10 Flevoland 地区实验数据集。(a)Pauli 分解伪彩色图像; (b) 实验数据地面真值; (c) 颜色与类别对应关系

表 1-4 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，其中每个分类结果图下方依次为 A、B、C 区域的局部放大图。根据图??，仅使用散射特征的分类方法，在草地（区域 A）、油菜籽（区域 B）和土豆（区域 C）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。

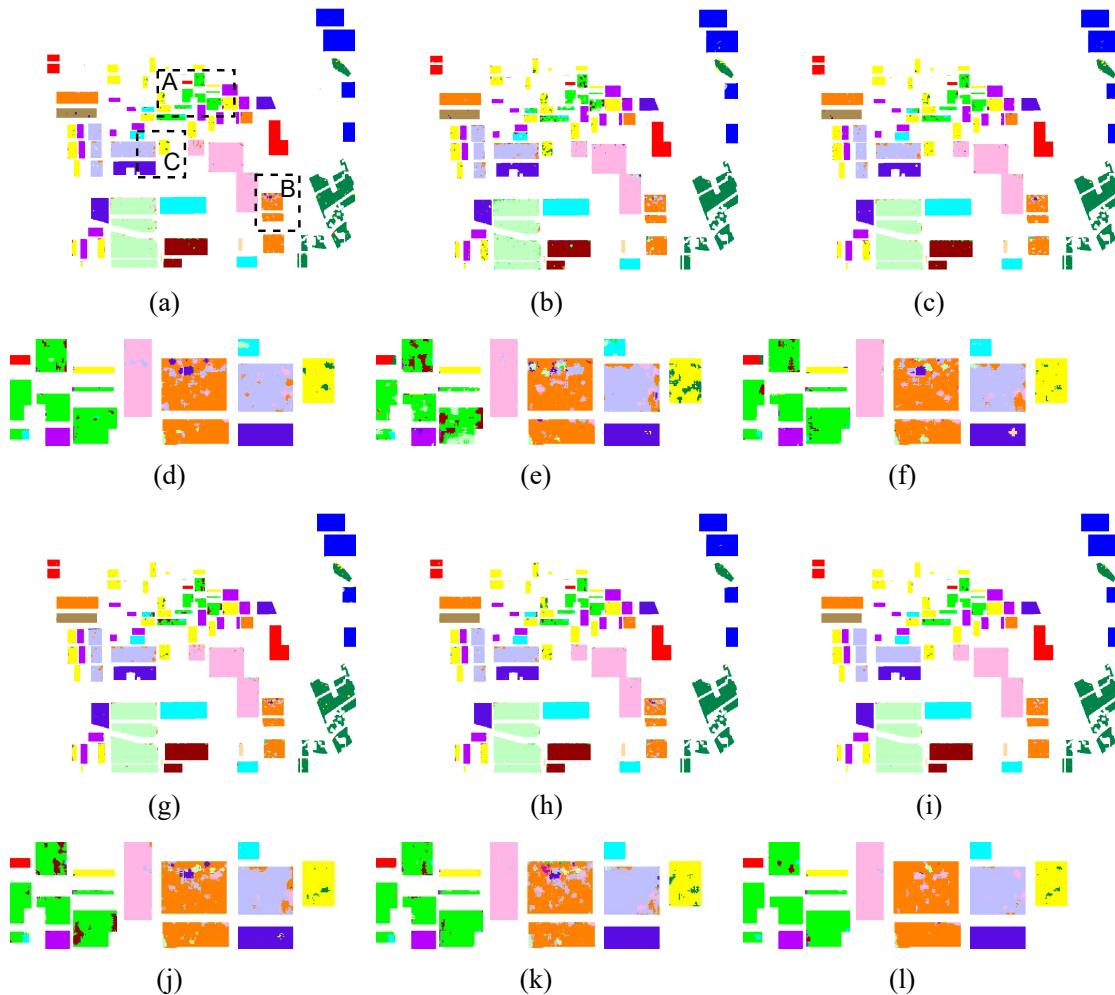


图 1-11 AIRSAR Flevoland 地区数据分类可视化结果图。(a)CNN-T; (b)CNN-P; (c)CNN-F; (d)CNN-T 局部误差图; (e)CNN-P 局部误差图; (f)CNN-F 局部误差图; (g)CNN-SE; (h)CNN-CBAM; (i) 本章方法; (j)CNN-SE 局部误差图; (k)CNN-CBAM 局部误差图; (l) 本章方法局部误差图

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)<sup>[?]</sup> 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，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。

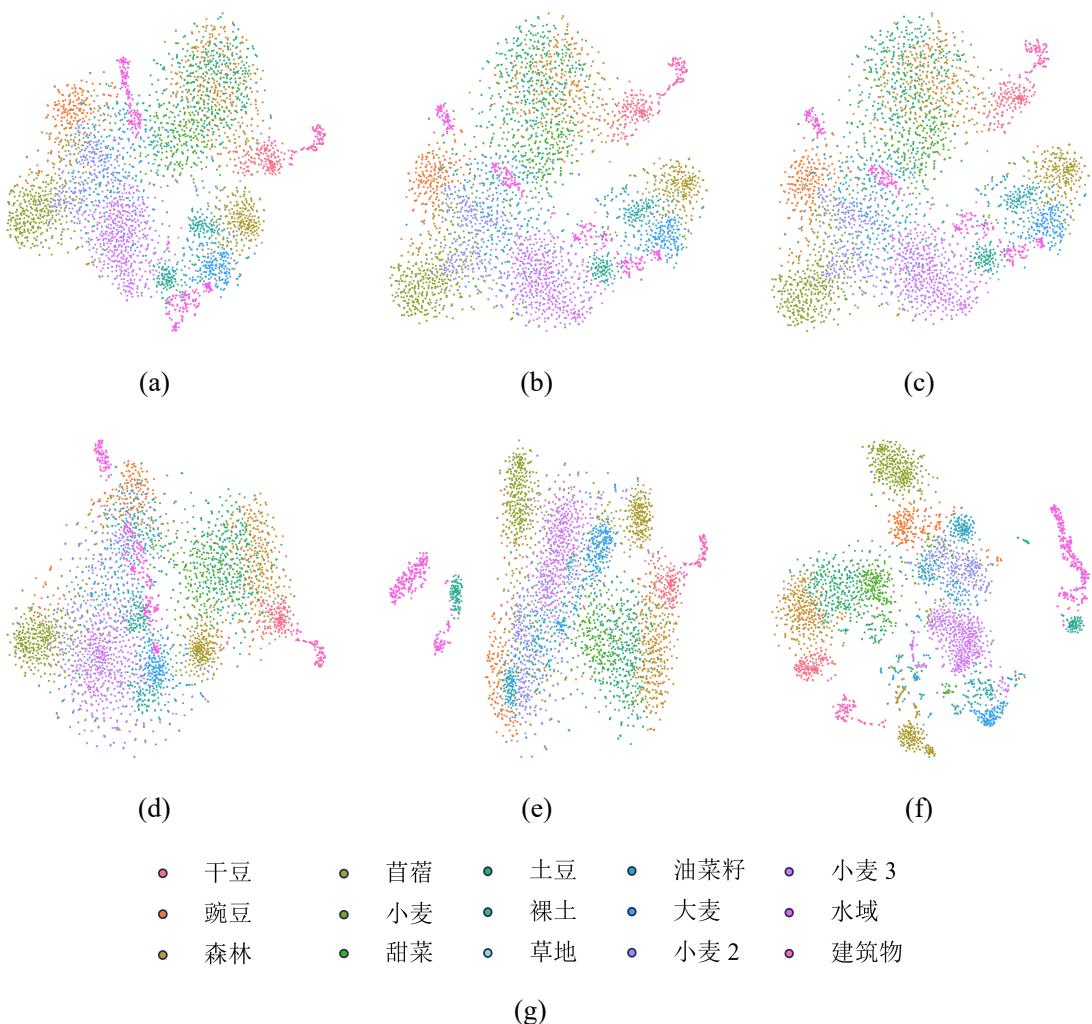


图 1-12 AIRSAR Flevoland 地区数据不同方法特征分布。(a)CNN-T; (b)CNN-P; (c)CNN-F; (d)CNN-SE; (e)CNN-CBAM; (f) 本章方法; (g) 颜色与类别对应关系

#### 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  $1300 \times 1200$ ，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??展示了该数据集的人工标记的地面上真值参考图。地面上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。

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

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

表 1-6 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

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 A、区域 B、区域 C）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

	干豆	豌豆	森林	苜蓿	小麦	甜菜	土豆	裸土	草地	油菜籽	大麦	小麦 2	小麦 3	水域	建筑物
干豆	5889	0	23	0	0	0	1	0	1	0	0	0	0	0	0
豌豆	0	8827	0	0	0	59	1	0	0	7	0	0	0	0	0
森林	1	1	14183	0	0	0	120	0	1	2	3	0	2	2	0
苜蓿	0	0	0	9152	0	0	0	0	171	4	6	0	19	0	0
小麦	0	8	0	0	16560	5	0	0	0	199	0	132	7	0	0
甜菜	3	2	4	4	4	9651	43	0	25	11	5	0	4	0	0
土豆	16	0	108	0	0	23	4666	0	0	0	0	0	1	0	0
裸土	0	0	0	0	0	0	0	2986	0	1	0	0	0	9	0
草地	10	0	0	37	0	5	1	0	5748	0	52	0	24	0	0
油菜籽	0	0	0	0	65	4	0	0	0	1780	0	104	32	0	0
大麦	0	0	6	0	0	0	0	0	131	2	6874	0	6	17	0
小麦 2	0	0	0	0	126	2	0	0	0	215	0	10009	17	0	0
小麦 3	0	0	0	0	10	0	0	0	4	89	1	29	20548	0	0
水域	0	0	0	0	0	0	1	0	0	0	0	0	0	1798	0
建筑物	1	0	0	0	0	0	1	0	0	0	1	0	1	0	337

图 1-13 本章方法在 Flevoland 图像中分类结果混淆矩阵图

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 OA 达到 94.39%，相比于其他验证方法提升了 4.12%。本章方法的散射特征时，在建筑物区域的分类准确率仅为 76.68%，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，直接堆叠极化特征在平均准确率指标反而下降 0.45%，这可能是由于该方法仅仅是将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入无法有效利用两种不同类型特征间的主次关系。而经典注意力方法 CNN-SE

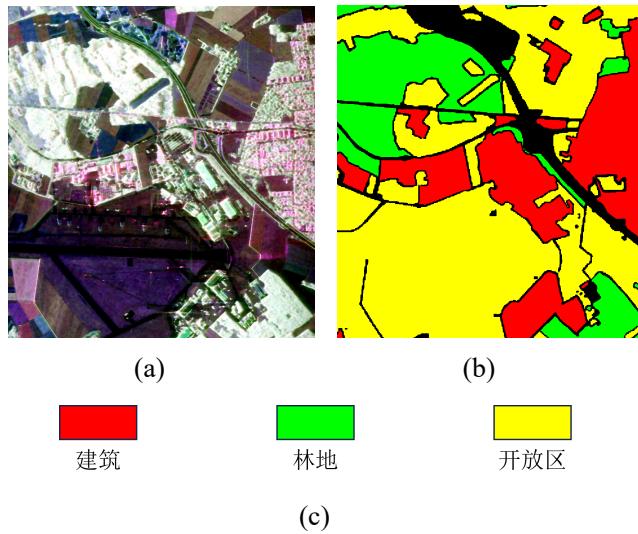


图 1-14 E-SAR Oberpfaffenhofen 区域数据。(a)Pauli 分解伪彩色图像; (b) 实验数据地面真值; (c) 颜色与类别对应关系

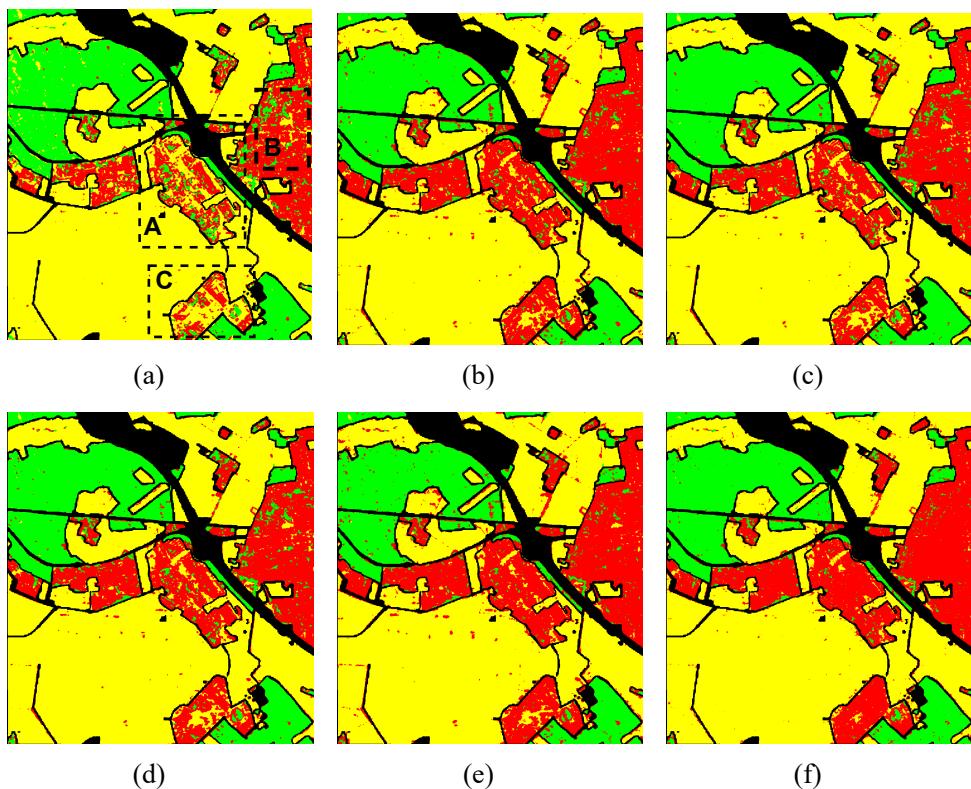


图 1-15 E-SAR Oberpfaffenhofen 地区数据分类可视化结果图。(a)CNN-T; (b)CNN-P; (c)CNN-F; (d)CNN-SE; (e)CNN-CBAM; (f) 本章方法

与 CNN-CBAM 的 OA 分别为 89.13% 和 90.27%，分类准确率提升不足 1%，微弱的准确率提升与计算量的增加表明对极化 SAR 特征优化性能有限。本章方法在

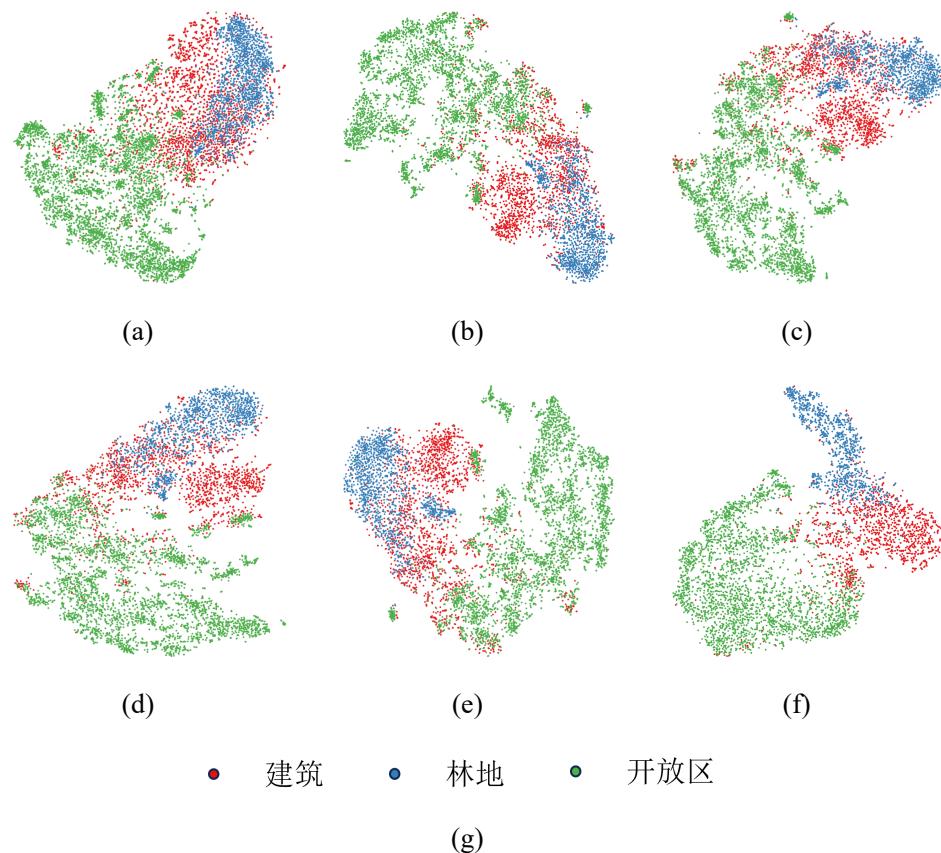


图 1-16 E-SAR Oberpfaffenhofen 地区数据地区数据不同方法特征分布。(a) CNN-T; (b) CNN-P; (c) CNN-F; (d) CNN-SE; (e) CNN-CBAM; (f) 本章方法

OA、AA、Kappa 系数等指标达到了 94.39%、93.28% 和 90.4%，相比于对比方法分别提升了 4.12%、4.3% 和 5.16%，表明本章方法能有效利用多类型极化特征，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 1-7 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	76.68	83.9	80.1	80.42	83.92	<b>90.47</b>
2	林地	90.35	91.99	90.5	91.44	91.61	<b>92.61</b>
3	开放区域	89.83	90.04	91.37	91.63	91.41	<b>96.78</b>
	OA	86.56	88.97	88.52	89.13	90.27	<b>94.39</b>
	AA	85.62	88.64	87.32	87.83	88.98	<b>93.28</b>
	Kappa	81.63	83.51	82.73	83.58	85.24	<b>90.4</b>

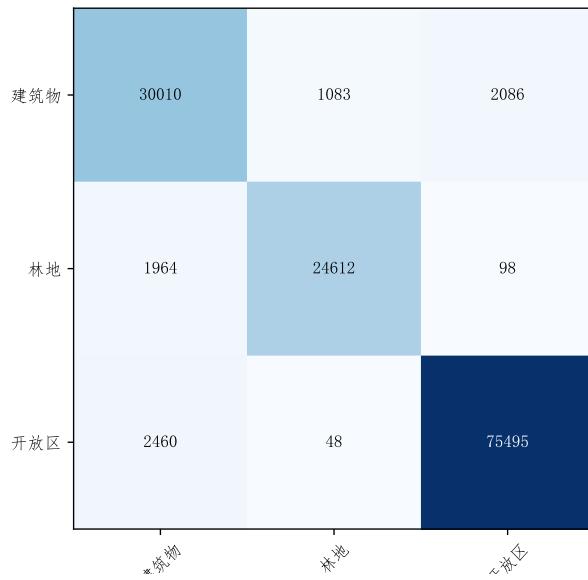


图 1-17 本章方法在 E-SAR Oberpfaffenhofen 图像中分类结果混淆矩阵图

## 1.5 本章小结

本章针对多类型极化特征信息冗余问题，提出了基于双通道注意力的极化 SAR 目标分类方法。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。其次，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，将经过双通道注意力校准后的极化特征输入到分类器，得到最终的分类结果。在两组真实极化 SAR 数据集进行性能测试，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。

## 第二章 基于混合模型的极化 SAR 目标分类方法

### 2.1 引言

深度卷积神经网络（Deep Convolutional Neural Network, DCNN）由于其对大规模数据进行端到端的学习能力以及强大的特征学习能力，在极化 SAR 图像分类领域取得了显著的成果<sup>[?,?]</sup>。不同于光学图像，由于 SAR 的复杂成像机制，极化 SAR 图像不可避免受到相干斑噪声影响，部分图像中目标边缘及内部存在类别模糊和特征信息不可靠问题。极化 SAR 样本标注工作都是通过人工标记完成，很可能因为人工标注错误、自动标注辅助技术误差、地面覆盖类型不准确等因素存在错误标记的情况，引入标签噪声<sup>[?,?]</sup>。DCNN 的优越分类性能依赖于数量充足的高质量标记样本，在存在标签噪声的训练集中训练时，由于其强大的拟合能力甚至会学习到噪声的特征，出现过拟合的现象，导致分类性能下降<sup>[?,?]</sup>。而现有的大多数基于 DCNN 的极化 SAR 图像分类方法都是在理想的训练集中完成模型训练，很少有考虑标签噪声问题。因此研究对标签噪声的极化 SAR 图像分类方法具有重要的价值与意义。

本章着力于含标签噪声的极化 SAR 图像分类方法研究，提出了一种基于混合模型估计的极化 SAR 图像分类方法。首先，根据噪声样本与准确样本的损失函数分布差异特性，使用混合模型对分布差异进行拟合，以估计样本的噪声概率。其次，根据极化 SAR 图像标签噪声更多存在于类别边界区域特性，通过捕获类别边界区域，对边界样本施加额外的惩罚，使分类模型能够有效利用边缘知识增强识别性能。最后，基于改进的自学习优化损失函数，训练过程中通过引入模型的预测来增加感知项，对损失函数进行纠正，完成模型的鲁棒参数优化过程。

### 2.2 有限混合模型

DCNN 模型在含标签噪声样本集上训练时，总是先学习简单的分类规则，之后再花足够的时间来对错误信息进行过拟合<sup>[?]</sup>。相比于干净的标记样本，随机标记的错误样本需要更长的时间来学习，这意味着有噪声样本在训练的前期阶段具有更高的损失，这使得干净样本与噪声样本可以通过损失值的分布情况进行区分。由于对观测值的建模能力，有限混合模型适用于估计干净样本与噪声样本的损失分布估计的场景。

有限混合模型（Finite Mixture Model, FMM）是一种用于描述观测数据生成过程的概率模型<sup>[?]</sup>。该模型假设观测的数据是由有限个概率分布组合而成的混合分

布，而每个成分分布对应一个潜在的类别，混合模型需要考虑这些类别之间的不同权重。

在有限混合模型中，假设观测数据来自于  $K$  个不同的成分，每个成分对应一个概率分布。对于一维数据而言，有限混合模型的概率密度函数可以表示为：

$$f(x) = \sum_{k=1}^K \pi_k \cdot f_k(x) \quad (2-1)$$

其中， $f_k(x)$  表示第  $k$  个成分分布的概率密度函数， $\pi_k$  表示对应成分的权重，并且满足  $\sum_{k=1}^K \pi_k = 1$ 。表示该数据的分布是由多个成分分布的加权和。

## 2.2.1 高斯混合模型

高斯混合模型（Gaussian Mixture Model, GMM）是一种将一个复杂的概率分布建模为多个高斯分布的线性组合的概率模型<sup>[2]</sup>。相比于单一的高斯概率密度函数，GMM 能够更加灵活地适应多峰的数据。GMM 的核心思想是将观察到的数据视为多个高斯分布组成的混合体，每个高斯分布为一个组分量，而所有的分量的加权和形成了整个的混合模型。

对于一个服从高斯分布的连续随机变量  $x$ ，其概率密度函数（PDF）可以表示为：

$$p(x) = \frac{1}{(2\pi)^{1/2} \sigma} \exp \left[ -\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right] = \mathcal{N}(x; \mu, \sigma^2) \quad (2-2)$$

$$(-\infty < x < \infty; \sigma > 0)$$

记作  $x \sim \mathcal{N}(\mu, \sigma^2)$ ，其中， $\mu, \sigma^2$  分别表示  $x$  的均值与方差。

对于由多个高斯变量随机混合形成的部分，记为高斯混合分布。服从高斯混合分布的连续随机变量  $x$  的概率密度函数可以表示为：

$$p(x) = \sum_{m=1}^M \frac{c_m}{(2\pi)^{1/2} \sigma_m} \exp \left[ -\frac{1}{2} \left( \frac{x - \mu_m}{\sigma_m} \right)^2 \right]$$

$$= \sum_{m=1}^M c_m \mathcal{N}(x; \mu_m, \sigma_m^2) \quad (2-3)$$

$$(-\infty < x < \infty; \sigma_m > 0; c_m > 0)$$

其中， $c_m$  表示不同混合类别的权重并且满足  $\sum_{m=1}^M c_m = 1$ 。

在上述的高斯混合分布中，涉及到一组参数，用  $\Theta = \{c_m, \mu_m, \Sigma_m\}$  表示。其中， $c_m$  表示每个分布的权重， $\mu_m$  表示均值， $\Sigma_m$  表示协方差矩阵。在考虑包含多个混合分类的情况下，这一组参数可以用来描述整个混合分布的特征。要获得对混合

分布参数的准确描述，需要根据一组假设来确定这些参数的值，通过最佳的参数，来匹配观测到的数据，进而更加准确的反映真实数据的特征。图??展示了二分量的 GMM 模型概率密度函数示意图：

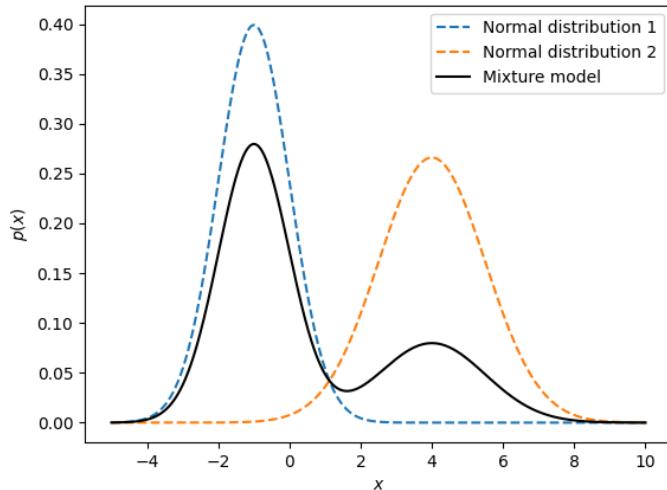


图 2-1 二分量的高斯混合模型概率密度函数示意图

高斯混合模型具有其原理简单、对异常值不敏感等有点，已经成为当前最为普遍使用的混合模型。但是在真实场景下，数据具有非高斯性、非线性等特点，使用高斯混合模型并不能完备地描述这些数据的分布特点，例如倾斜数据、拖尾数据、有界数据等。

### 2.2.2 贝塔混合模型

贝塔混合模型（Beta Mixture Model, BMM）是一种概率模型，将复杂的概率分布建模为多个贝塔分布的线性组合。贝塔分布是一个定义在区间  $[0, 1]$  上的概率分布，通常用于描述在二项分布的贝叶斯估计中<sup>[?]</sup>。

对于一个服从贝塔分布的连续随机变量  $x$ ，其概率密度函数（PDF）可以表示为：

$$p(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} \quad (2-4)$$

其中， $B(\alpha, \beta)$  为贝塔分布的概率密度函数，具体表示为：

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)} \quad (2-5)$$

其中， $\alpha$  和  $\beta$  是贝塔分布的参数， $\Gamma$  表示伽玛函数。对于由多个贝塔分布随机混合形成的部分，记为贝塔混合分布。服从贝塔混合分布的连续随机变量  $x$  的概率密度函数可以表示为：

$$\begin{aligned} p(x) &= \sum_{m=1}^M c_m \frac{x^{\alpha_m-1}(1-x)^{\beta_m-1}}{B(\alpha_m, \beta_m)} \\ &= \sum_{m=1}^M c_m \text{Beta}(x; \alpha_m, \beta_m) \\ (0 < x < 1; c_m > 0; \sum_{m=1}^M c_m = 1; \alpha_m, \beta_m > 0) \end{aligned} \quad (2-6)$$

其中， $c_m$  表示不同混合分类的权重，满足  $\sum_{m=1}^M c_m = 1$ 。

在贝塔混合分布中，参数集为  $\Theta = \{c_m, \alpha_m, \beta_m\}$ 。其中， $c_m$  表示每个贝塔分布的权重， $\alpha_m$  和  $\beta_m$  分别表示每个贝塔分布的形状参数。为了描述整个混合分布的特征，需要通过一组假设，从观测到的数据中确定这些参数的值。这一过程旨在使混合分布的参数最佳匹配观测到的数据，以更准确地反映实际数据的特征。图??展示了二分量的 BMM 模型概率密度函数示意图：

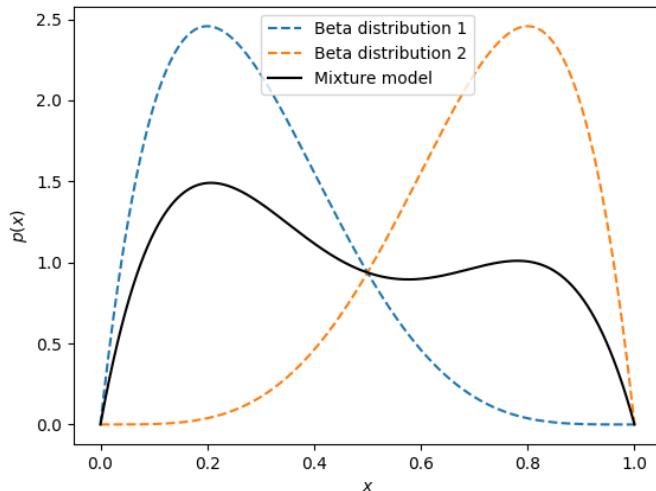


图 2-2 二分量的贝塔混合模型概率密度函数示意图

## 2.3 混合模型与边界增强的极化 SAR 目标分类方法

### 2.3.1 标签噪声问题建模

给定一个包含  $N$  个类别的干净 PolSAR 数据的训练集，表示为：

$$(X, Y^{(c)}) = \{(x_1, y_1^{(c)}), \dots, (x_N, y_N^{(c)})\} \quad (2-7)$$

该数据来源于分布  $D = X \times Y$ 。假设存在一个函数  $\mathcal{F}: Y^{(c)} \rightarrow Y^{(n)}$ , 该函数向标签  $Y^{(c)}$  引入噪声, 将  $\mathcal{F}$  应用于  $(X, Y^{(c)})$  得到一个带有噪声的训练数据集, 表示为:

$$(X, Y^{(n)}) = \{(x_1, y_1^{(n)}), \dots, (x_N, y_N^{(n)})\} \quad (2-8)$$

其中  $Y^{(n)}$  是原始干净标签  $Y^{(c)}$  和被污染的标签  $Y^{(n)} \neq Y^{(c)}$  的组合。

假定标签  $y_i^{(c)}$  被噪声函数污染成  $y_i^{(n)}$ , 存在对称噪声源与非对称噪声源两种类型。对称噪声源模型是将真实标签以相同的概率随机翻转到其他类别, 服从均匀分布, 表达式如下:

$$P(y_i = y_i^{(n)} | y_i^{(c)}) = \frac{1}{N-1} \quad (2-9)$$

非对称噪声源模型是按照根据某种固定的规则进行标签映射操作, 将真实标签以随机概率翻转到其他的某个类别。图??展示了两种类型噪声源的噪声转移矩阵示意图。

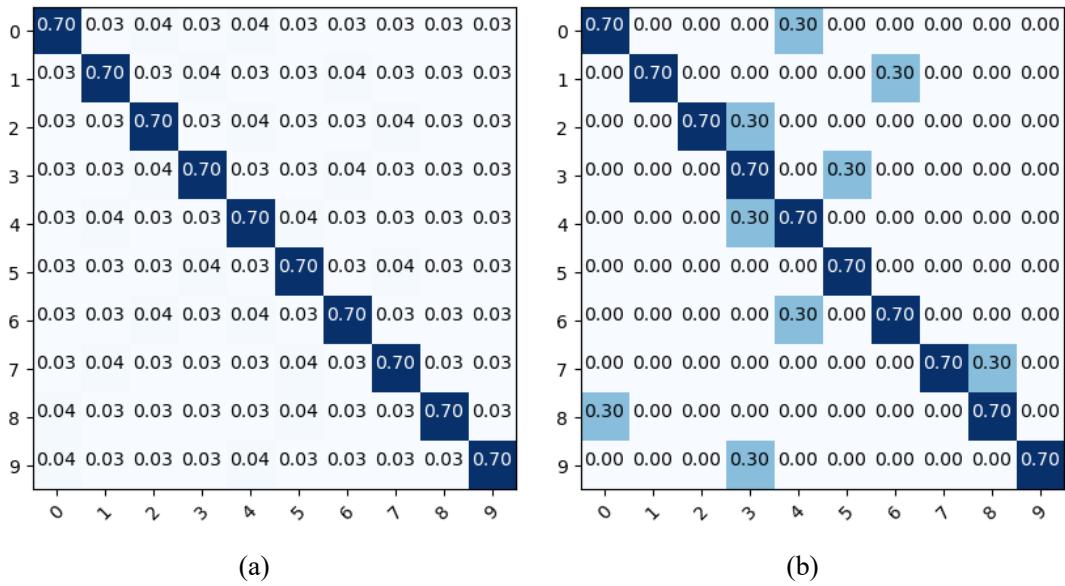


图 2-3 30% 比例噪声转移矩阵示意图。(a) 对称噪声 (b) 非对称噪声

从干净分布  $D$  中随机选择一个测试集  $(X_T, Y_T)$ 。目标是使用带有噪声的数据集  $(X, Y^{(n)})$  训练一个分类器模型  $M: X \rightarrow Y$ , 以确保对干净测试集  $(X_T, Y_T)$  的鲁棒性泛化, 并且在训练过程中无法访问原始的干净标签  $Y_c$ 。

### 2.3.2 混合模型与边界增强的极化 SAR 目标分类算法框架

在极化 SAR 图像目标分类中, 标签噪声的存在通常会导致传统分类方法性能下降, 基于卷积网络的模型往往也会陷入过拟合的问题, 学习到错误的分类规则。

针对标签噪声问题，本章提出了基于混合模型的极化 SAR 图像分类方法，其算法框架如图??所示。区别于传统的交叉熵损失优化方法，本章方法基于混合模型理论对样本损失函数进行参数估计，以区分噪声样本与干净样本。为了降低边界标签噪声样本对模型训练的影响，增强对边界样本的有效信息利用，通过 Sobel 算子对极化 SAR 图像的 Pauli 伪彩图进行边界提取并膨胀，对属于膨胀边界内的样本施加额外的惩罚，增强样本损失。最后，基于改进的自学习优化损失函数，结合模型预测感知项，对噪声标签进行校正，实现模型的鲁棒性训练。

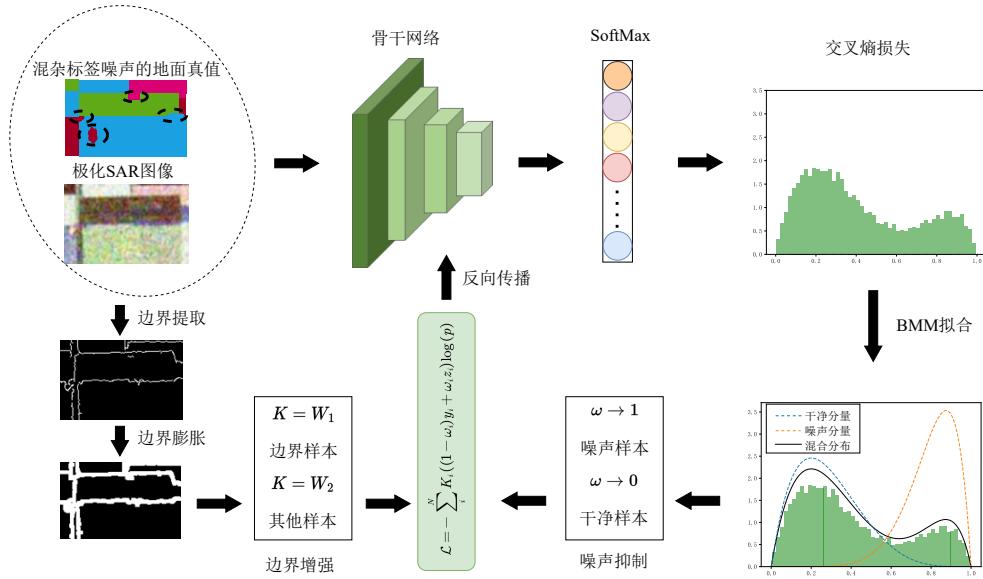


图 2-4 基于混合模型估计的目标分类算法框架图

结合上述算法框架，本章基于混合模型的极化 SAR 图像分类方法可以表示为以下步骤：

(1) 极化特征输入。极化特征具有多种表示形式，选择合适的极化特征对于分类任务具有重要意义。本章研究属于极化 SAR 图像分类问题，将第三章提出的双通道注意力极化信息提取方法作为本算法的前置模块，用于极化特征表示，并输入到后续网络。

(2) 鉴别特征提取与分类。利用卷积神经网络，对输入的极化特征进行特征提取。通过多个堆叠的卷积、池化与激活操作，主干网络学习到极化特征的高级表征。随后，利用全连接层与激活函数，将主干网络的输出映射到不同的类别，得到对样本的预测类别。

(3) 噪声概率估计。计算每个样本预测类别与标签的交叉熵损失，并基于贝塔混合模型拟合样本损失分布。利用损失值来更新迭代混合模型参数，基于更新

后的模型参数计算样本噪声概率。

(4) 边界样本增强。利用 Sobel 算子对极化 SAR 图像的 Pauli 伪彩图进行边界提取并膨胀，对膨胀边界内部的训练样本进行损失函数加权增强，以提高模型对边界信息的学习能力。

(5) 自学习损失函数优化。基于自学习损失函数，利用估计的样本噪声概率，动态调整模型预测与标签真值的权重。利用改进的自学习损失函数计算损失值后，通过反向传播更新模型参数，实现模型的训练优化。

综合上述算法描述，算法流程图如图??所示，下面对各个模块进行详细介绍。

### 2.3.3 噪声概率估计模型

在混合标签噪声数据集训练过程中，深度网络模型通常先学习干净样本特征，然后较长时间才能适应随机嘈杂标签。意味着训练早期，噪声样本相比于干净样本具有更高损失，可以认为训练过程中损失的分布是由干净样本分布与噪声样本分布叠加形成<sup>[?]</sup>。因此，利用二分量混合模型来对样本损失分布进行拟合，完成样本噪声概率估计。

贝塔混合模型具有对二元概率随机变量更加灵活的拟合特性<sup>[?]</sup>，用于拟合混合样本的损失值分布。对于给定的样本  $(x, y)$ ，其中  $x, y, \hat{y}$  分别表示 PolSAR 数据、相应的地面真实标签和  $x$  的分类器预测，交叉熵损失的标准化表示为  $\mathcal{L}_{CE}(\hat{y}, y) = \ell$ 。根据式??与式??，混合标签噪声损失概率密度函数可以表示为：

$$p(\ell) = \lambda_c \cdot p(\ell | \text{clean}) + \lambda_n \cdot p(\ell | \text{noisy}) \quad (2-10)$$

其中， $p(\ell | \text{clean})$  与  $p(\ell | \text{noisy})$  分别表示干净/噪声样本损失分布，具体表示如下：

$$p(\ell | \text{clean}) = \frac{\Gamma(\alpha_c + \beta_c)}{\Gamma(\alpha_c)\Gamma(\beta_c)} \ell^{\alpha_c-1} (1-\ell)^{\beta_c-1} \quad (2-11)$$

$$p(\ell | \text{noisy}) = \frac{\Gamma(\alpha_n + \beta_n)}{\Gamma(\alpha_n)\Gamma(\beta_n)} \ell^{\alpha_n-1} (1-\ell)^{\beta_n-1} \quad (2-12)$$

其中， $\alpha_{c/n}, \beta_{c/n} > 0$  表示干净/噪声样本的贝塔分布参数， $\Gamma(\cdot)$  是伽玛函数， $\lambda_c$  和  $\lambda_n$  分别表示混合系数，表示两个分量的组合系数。

参数集  $\Theta = \{c_{c/n}, \alpha_{c/n}, \beta_{c/n}\}$  描述了整个混合分布的特征。使用最大期望 (Expectation Maximization, EM) 算法，实现混合模型的参数更新迭代。首先，在 E 步中，固定混合系数，损失值  $\ell$  的后验概率  $Q_k(\ell) = p(k | \ell)$  基于贝叶斯规则进行

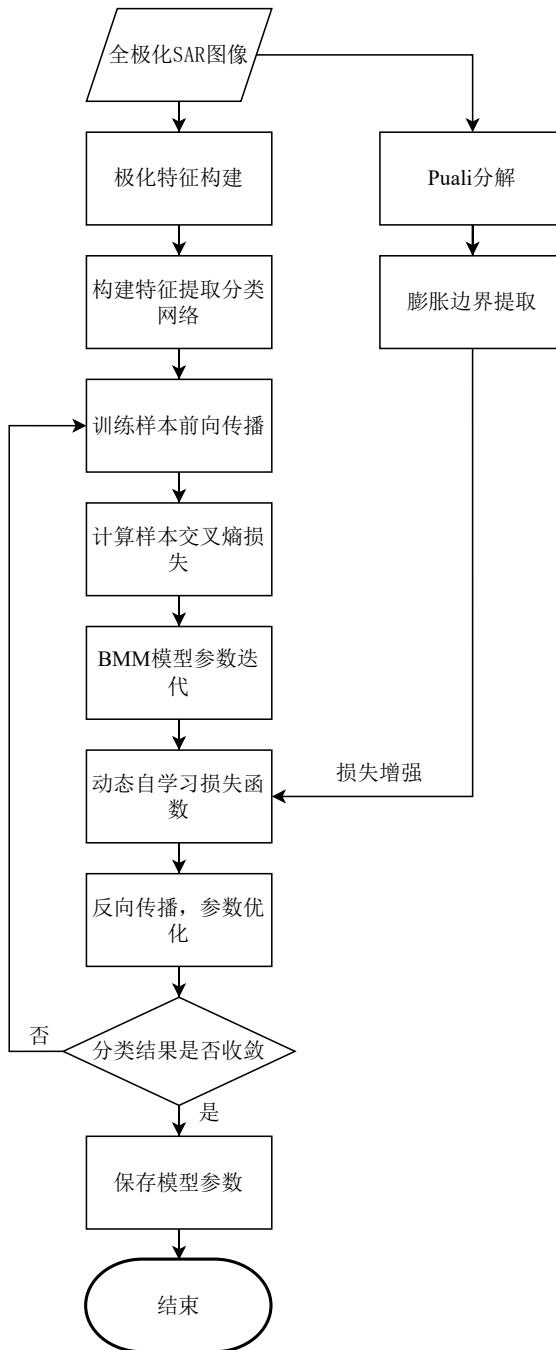


图 2-5 基于混合模型估计的目标分类算法流程图

更新，即

$$Q_k(\ell) = \frac{\lambda_k p(\ell | \alpha_k, \beta_k)}{\sum_{j=0}^1 \lambda_j p(\ell | \alpha_j, \beta_j)} \quad (2-13)$$

其中， $j = 0(1)$  表示干净（噪声）类别。

然后，在 M 步中，基于更新的  $Q_k(\ell)$ ，通过加权版本的矩方法估计  $\alpha_k$  和  $\beta_k$ 。

### 2.3.4 边界增强模块

由于错误标记的噪声样本更多出现在类别交替的边界部分，模型需要对属于边界样本付出更多的关注度。为了加强对边界样本注意力关注和细节增强，对 Pauli 彩图进行 Sobel 算子卷积和展开运算，生成距离边界  $d$  以内的掩膜像素，并将其作为辅助边界损失的目标。对于输入的图像  $I(x, y)$ ，Sobel 算子使用两个  $3 \times 3$  的卷积核与输入图像进行卷积操作，分别计算垂直和水平梯度近似。这两个卷积核如下所示：

$$K_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (2-14)$$

$$K_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (2-15)$$

其中， $K_x$  和  $K_y$  分别表示水平和垂直方向的卷积核。卷积操作的结果分别用  $G_x$  和  $G_y$  表示，其计算过程如下：

$$G_x = I * K_x = \sum_{i=-1}^1 \sum_{j=-1}^1 I(x+i, y+j) \cdot K_x(i, j) \quad (2-16)$$

$$G_y = I * K_y = \sum_{i=-1}^1 \sum_{j=-1}^1 I(x+i, y+j) \cdot K_y(i, j) \quad (2-17)$$

其中，\* 表示卷积操作。通过以上两个卷积操作，得到图像每个像素点的水平和垂直梯度。图像在该相处点出的梯度强度和方向可以通过下式计算：

$$G = \sqrt{G_x^2 + G_y^2} \quad (2-18)$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right) \quad (2-19)$$

其中， $G$  表示梯度强度， $\theta$  表示梯度方向。

通过 Sobel 算子提取到 Pauli 彩图边界后，为了增强对边界区域的细节感知，通过对边界以距离  $d$  进行展开操作，得到膨胀的掩膜边界，用于后续的辅助边界损失。

### 2.3.5 损失函数与优化

面对标签噪声问题时，损失函数的精心选择对于引导模型学习训练过程变得极为重要。传统的交叉熵损失完全依赖于分类模型预测类别与给定的训练集标签值进行计算损失，对于训练集中样本标签存在错误时会引导模型参数拟合错误的规则，会鼓励模型学习到错误的信息。为了处理标签噪声问题，本章通过在标准交叉熵损失中引入感知项来辅助调整训练目标。具体而言，损失函数表达式为：

$$\ell_B = \sum_{i=1}^N ((1 - \omega_i)y_i + \omega_i z_i)^T \log(h_i) \quad (2-20)$$

其中， $N$  表示样本数量， $y_i$  与  $z_i$  分别表示样本标签值与分类模型预测值， $\omega_i$  在损失函数中作为权重调整对样本标签值和模型预测值之间的权衡。

如果设置权重  $\omega$  为固定值，并不能有效防止噪声样本的过拟合情况。本章将噪声预测概率与动态损失函数相结合，通过样本属于噪声的概率来动态调整样本标签和模型预测分类的依赖程度。结合 BMM 模型的噪声概率预测，通过将式  $\omega$  设置为由公式??得到的噪声概率  $p(\ell | \text{noisy})$ ，并且在每一个训练轮次结束后使用每个样本的交叉熵损失更新 BMM 模型参数，进而更新对应的噪声概率。因此干净的样本的噪声概率  $1 - \omega_i$  接近 1，更多的依赖于真值标签，而属于噪声样本的损失则由模型分类结果  $z_i$  和预测噪声概率来决定。

## 2.4 实验结果与分析

### 2.4.1 实验参数设置

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、batch-size 等超参数设置与第三章实验参数相同。

实验中，使用两组真实极化 SAR 数据集分别是荷兰 Flevoland 区域和德国 Oberpfaffenhofen 地区数据来验证本章方法的有效性，并利用常规的性能指标总体分类准确率 (OA)、各个类别分类准确率和 Kappa 系数对分类结果进行数值量化。同时对分类的可视化结果进行视觉评估。在给定的极化 SAR 标准数据集中，一部分像素是没有标签的，所以在计算分类准确率的时候只统计数据集中那些有标签的样本被正确分类的百分比，并认为该指标可以表征数据集中整体的分类性能。

实验中，随机从数据集中选择 1% 的样本作为训练集，其余的 99% 样本作为测

试集。为了方便，将本章提出的基于混合模型与边界增强的分类方法简记为 BEL，同时将结合上章极化信息提取方法 DP 作为本章前置模块的方法记为 BEL+DP，[??](#)小节中介绍的特征分类器作为本章方法的主干网络。为了模拟实际场景中出现的对称噪声标签和非对称噪声标签的两种情况，本节实验通过设置噪声转移矩阵的方法，分别在以上两个数据集中合成对应类型的标签噪声。同时，对比了两种噪声源下不同噪声比例对各个方法分类准确率的影响。为了验证本章方法的优越性，选择了多种替代的分类方案进行比较，主要是通过与目前经典的以及优越的分类方法进行对比，包括支持向量机方法（记为 SVM）[\[?\]](#)、基于 Wishart 分布的分类方法（记为 Wishart）[\[?\]](#) 和鲁棒的半监督学习方法（记为 RSL）[\[?\]](#)。

## 2.4.2 AIRSAR Flevoland 数据实验

### 2.4.2.1 对称标签实验

[图??](#)展示了 10% 噪声比例下对称标签噪声的噪声转移矩阵，每个类别正确标记的概率为 90%，10% 的标签噪声均匀地分布在其他类别中。

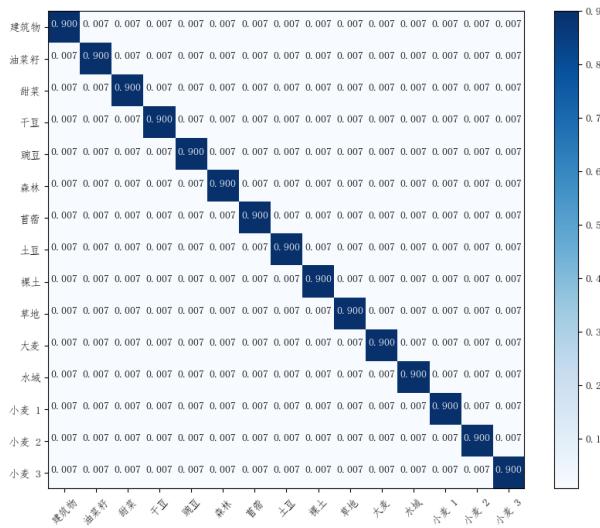


图 2-6 AIRSAR Flevoland 数据 10% 对称噪声转移矩阵

[图??](#)展示了不同实验方法在 10% 对称噪声类型下分类结果图，其中每个分类结果图下方依次为 A、B、C 区域的局部放大图。根据图??、图??展示的 SVM 与 Wishart 分类器的分类结果，在大多数的类别中都存在着严重的错分情况，这可能是由于这两类分类器都无法有效地应对标签噪声的影响，错误的信息对分类规则噪声严重的负面影响，进而导致分类性能下降的现象。而图??展示的在交叉熵损失函数优化的 CNN 分类器的分类结果，可以看到小麦区域和干豆存在大量的错分情况，将小麦错分为水域，将干豆区域错分为土豆，这是因为这两类目标本身散射

特性比较接近，容易造成错分，同时该分类模型并没有对标签噪声进行处理，模型过拟合学习到了错误的类别信息，导致错分的情况。图??展示了使用鲁棒分类方法的分类结果，可以看到大量连续错分的情况得到改善，但是在草地等区域出现了小块的错分样本，将草地错分为大麦，这可能是草地覆盖区域较小，又存在噪声的情况，导致有效的训练样本数量不足，同时 RSL 方法是将判断为噪声的样本直接丢弃也丢失了一部分信息，在这一块区域中的分类性能有限。图??与图??均是本章提出的标签噪声下鲁棒性分类方法的分类结果图，从图中可以直观看出分类效果得到提升，错分的情况大大减少，仅有少数的错分情况，并且使用极化信息提取器后在油菜籽区域的错分情况也得到视觉上的改善，这也验证了本章方法的有效性。

表??展示了不同分类方法在 10% 对称标签噪声比例下的分类数值结果。SVM 与 Wishart 分类方法由于无法有效应对标签噪声的影响，绝大多数类别的分类准确率都低于 60%，存在严重的错分的情况，总体分类准确率较低，分别为 58.98% 和 49.78%。基于交叉熵损失的 CNN 分类器在油菜籽、干豆区域分类准确率较低，但是在其他类别上的分类准确率有了一定的提升，这可能是因为深度网络模型本身对于标签噪声的一定的鲁棒性能，会先学习特征简单的样本，再去学习错误的样本信息，总体分类准确率达到 88.39%。鲁棒性分类学习方法 RSL 由于对判断为噪声的样本进行了丢弃，在大多数类别上准确率取得了不错的成果，但是在建筑物区域的分类准确率较低，这可能是因为建筑物区域本身样本有限，添加噪声之后导致有效的训练样本不足，进而导致了在该类别上的分类准确率较低的情况。BEL 与 BEL+DP 分类准确率相比于其他方法有了较大的提升，两个方法占据了所有类别和所有指标的最大值。BEL+DP 方法在油菜籽、甜菜等区域中分类准确率稍低于 BEL 方法，不足 1% 的差距，但是在建筑物、甜菜等区域中的分类准确率高于 BEL 方法超过 1%，验证了极化信息表征对该场景下性能提升的有效性。通过以上分析证明，本章提出的 BEL 方法对标签噪声的有效性与优越性。

#### 2.4.2.2 非对称标签实验

图??展示了 10% 噪声比例下非对称标签噪声的噪声转移矩阵，每个类别正确标记的概率为 90%，10% 的标签噪声随机映射到另外的一个类别中。

图??展示了 10% 非对称标签噪声下不同对比方法的分类结果，其中每个分类结果图下方依次为 A、B、C 区域的局部放大图。图??与图??依然存在大量的错分情况，并且区域内部存在大量斑点状的错误区域。图??展示了交叉熵损失下的 CNN 分类结果，由于训练集中油菜籽区域部分像素被标记为小麦，导致在分类结果中也存在相同的错分情况，这表明该分类方法在错误标记训练集上过拟合，学

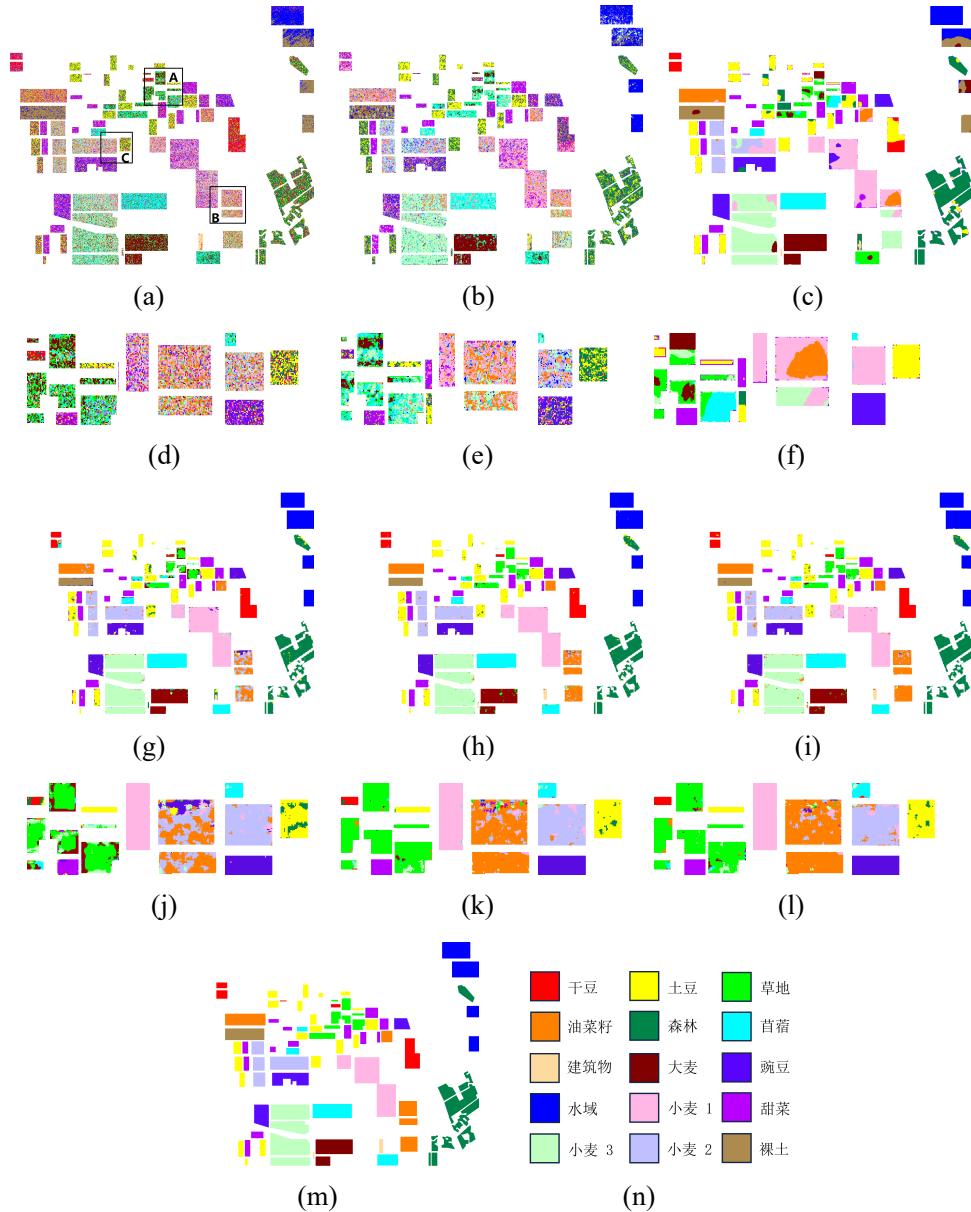


图 2-7 AIRSAR Flevoland 地区数据对称标签噪声下分类可视化结果图。

(a)SVM; (b)Wishart; (c)CNN; (d)SVM 局部误差图; (e)Wishart 局部误差图;  
 (f)CNN 局部误差图; (g)RSL; (h)BEL; (i)BEL + DP; (j)RSL 局部误差图;  
 (k)BEL 局部误差图; (l)BEL + DP 局部误差图; (m)地面真值; (n)颜色与类  
 别对应关系

习了错误的分类规则。图??展示了 RSL 方法的分类结果，可以看出在油菜籽、大麦区域依然存在大块连续的错分区域，将大麦错分为苜蓿，这是因为 RSL 对大量的错误样本过滤性能有限。图??与图??展示了 BEL 与 BEL+DP 方法的分类结果，可以看到由于对称噪声标签的影响，在小麦、林地、油菜籽区域出现了一定的错

序号	类别	SVM	Wishart	CNN	RSL	BEL	BEL + DP
1	建筑物	37.26	32.64	65.07	79.52	86.38	<b>87.75</b>
2	油菜籽	31.35	15.58	33.32	66.64	<b>92.01</b>	90.27
3	甜菜	74.17	9.87	77.49	<b>99.01</b>	95.67	97.28
4	干豆	14.61	14.02	58.6	74.63	<b>96.92</b>	96.62
5	豌豆	50.2	22.9	79.68	96.39	<b>99.11</b>	98.8
6	森林	66.31	41.52	83.57	94.34	96.8	<b>97.7</b>
7	苜蓿	62.92	17.71	59.63	80.53	96.19	<b>98</b>
8	土豆	53.54	26.89	72.74	83.1	95.49	<b>98.42</b>
9	裸土	64.25	11.36	80.46	<b>95.05</b>	91.02	93.12
10	草地	19.64	17.27	55	80.43	87.65	<b>91.66</b>
11	大麦	80.89	10.48	79.28	89.43	96.26	<b>97.49</b>
12	水域	76.9	59.9	49.51	79.32	97.37	<b>99.85</b>
13	小麦 1	59.13	40.18	82.51	91.83	<b>97.71</b>	97.23
14	小麦 2	45.27	46.52	64.23	80.16	<b>94.82</b>	94.22
15	小麦 3	75.74	43.69	73.65	<b>98.85</b>	97.05	98.06
OA		58.98	49.78	69.83	86.13	95.25	<b>96.36</b>
AA		54.15	27.36	67.64	85.94	94.7	<b>95.76</b>
Kappa		56.02	47.16	67.56	85.12	95.47	<b>96.58</b>

表 2-1 AIRSAR Flevoland 地区数据对称标签噪声下分类数值结果 (%)

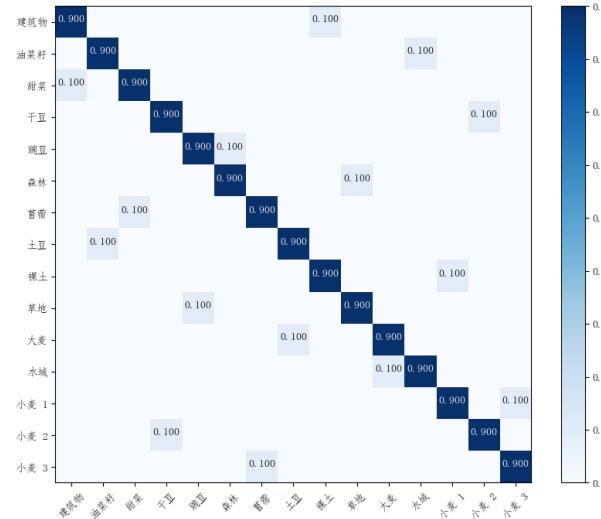


图 2-8 AIRSAR Flevoland 数据 10% 非对称噪声转移矩阵

分情况，但是在其他大部分区域分类结果较平滑，错分样本较少，以上结果证明了本章方法在非对称噪声标签下的有效性与优越性。

表??展示了各对比方法在非对称标签噪声下的分类数值结果。SVM 与 Wishart 作为经典的极化 SAR 分类方法，在非对称标签噪声下分类效果不理想，总体

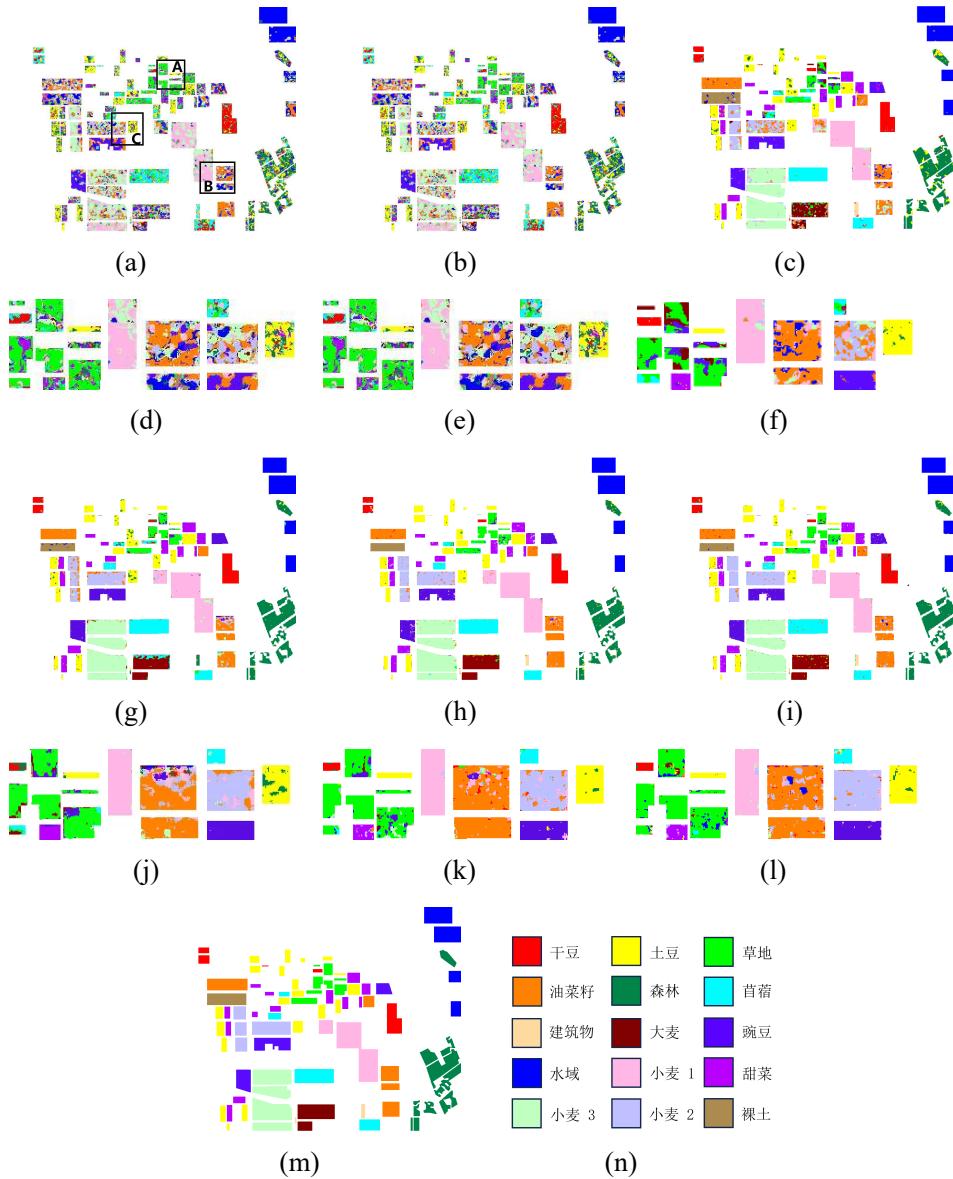


图 2-9 AIRSAR Flevoland 地区数据非对称标签噪声下分类可视化结果图。

(a)SVM; (b)Wishart; (c)CNN; (d)SVM 局部误差图; (e)Wishart 局部误差图;  
 (f)CNN 局部误差图; (g)RSL; (h)BEL; (i)BEL + DP; (j)RSL 局部误差图;  
 (k)BEL 局部误差图; (l)BEL + DP 局部误差图; (m) 地面真值; (n) 颜色与类  
 别对应关系

准确率分别为 56.48% 与 50.32%，这是因为基于 SVM 与 Wishart 的方法分类器模型简单，无法鉴别标签噪声的错误信息。基于交叉熵损失的 CNN 方法虽然整体有了提升，但是在建筑物区域准确率仅为 42.28%，这是因为 CNN 也不具备对标签噪声的鲁棒性能，会学习到错误的分类准则。除了干豆、水域、大麦三个区域略低于 RSL 方法，本章方法在其他指标上均取得了最高值，尤其是建筑物区域获得

了较大的提升，反映了本章提出的方法是有效且优越的。

表 2-2 AIRSAR Flevoland 地区数据非对称标签噪声下分类数值结果 (%)

序号	类别	SVM	Wishart	CNN	RSL	BEL	BEL + DP
1	建筑物	31.73	31.56	42.28	46.24	91.09	<b>96.16</b>
2	油菜籽	33.76	18.23	61.25	82.77	<b>92.34</b>	92.19
3	甜菜	73.98	10.56	74.86	95.23	96.68	<b>97.92</b>
4	干豆	16.23	12.97	85.41	<b>96.8</b>	91.01	94.6
5	豌豆	48.34	24.56	87.73	95.62	97.86	<b>97.39</b>
6	森林	64.28	40.08	82.19	91.36	96.68	<b>97.96</b>
7	苜蓿	60.72	18.23	79.2	79.8	94.43	<b>96.04</b>
8	土豆	52.01	22.14	81.69	<b>95.9</b>	95.68	94.41
9	裸土	64.58	14.2	81.32	98.04	94.01	<b>99.39</b>
10	草地	21.33	19.32	77.84	64.59	92.98	<b>92.4</b>
11	大麦	78.89	9.47	76.23	59.39	89.52	<b>97.24</b>
12	水域	75.42	58.14	68.04	<b>98.2</b>	97.35	97.7
13	小麦 1	57.6	42.18	80.37	<b>94.83</b>	93.97	93.87
14	小麦 2	44.32	44.25	65.17	86.95	83.83	<b>91.55</b>
15	小麦 3	74.36	45.84	70.35	94.2	94.72	<b>98.12</b>
OA		56.48	50.32	75.28	89.88	93.82	<b>96.52</b>
AA		53.17	27.45	74.26	88.91	91.48	<b>95.79</b>
Kappa		54.59	48.03	73.96	88.94	93.26	<b>95.22</b>

图??展示了各个对比方法在两种噪声类型下分类准确率随着噪声比例的变化情况。从图中可以看出，随着标签噪声比例增加，各个方法分类准确率呈现不同程度的下降。在对称标签噪声下，SVM、Wishart、CNN 这三种未考虑噪声标签影响的分类方法随着噪声比例增加下降幅度较快，在方法的分类准确率不高的情况下，从 10% 噪声比例至 50% 分别下降了 22.70%、13.91% 和 18.40%，表明了标签噪声带来的错误信息对分类任务形成了较大的影响。对于 RSL 方法，从 10% 噪声比例至 50% 分类准确率下降了 10.59%，下降幅度相对较小，但是在 50% 噪声比例下分类准确率低于 80%，表明该方法对高比例噪声标签的鲁棒性能有限。BEL 与 BEL+DP 方法从 10% 噪声比例至 50% 分类准确率分别下降了 9.9% 和 8.69%，准确率下降幅度均小于 10%，且在 50% 噪声比例下准确率依然均在 85% 以上。从图中还可以看出，相同噪声比例下，非对称噪声源下的分类准确率要明显低于对称噪声源，SVM、Wishart 分类方法的最差分类准确率已经低于 30%，RSL 方法的最差分类准确率降到了 54.32%，本章分类方法降到了 56.20%。尽管大面积的错误标记样本严重影响了模型分类规则，导致无法根据损失判定噪声概率，但是本章方法的分类准确率指标依然高于对比方法。以上结果证明了本章方法对不同噪声类

型、不同噪声比例均具有良好的鲁棒性能。

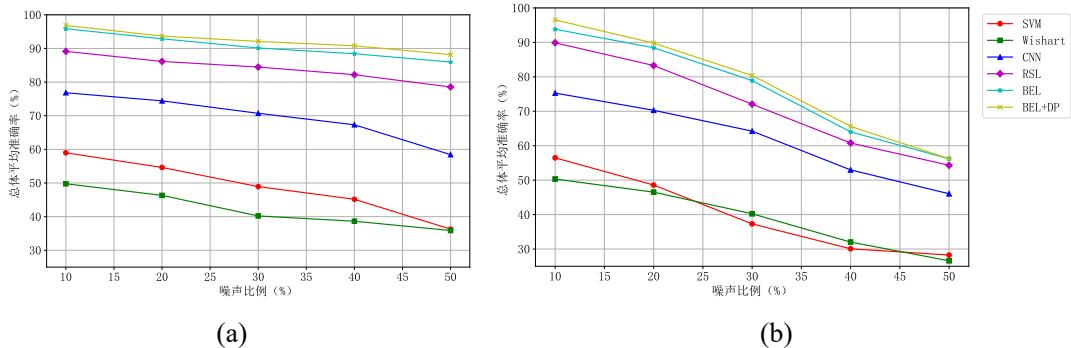


图 2-10 AIRSAR Flevoland 地区数据不同噪声比例下分类准确率结果。(a) 对称噪声源; (b) 非对称噪声源

### 2.4.3 E-SAR Oberpfaffenhofen 数据实验

#### 2.4.3.1 对称标签实验

图??展示了 10% 噪声比例下对称标签噪声的噪声转移矩阵，每个类别正确标记的概率为 90%，10% 的标签噪声均匀地分布在其他类别中。

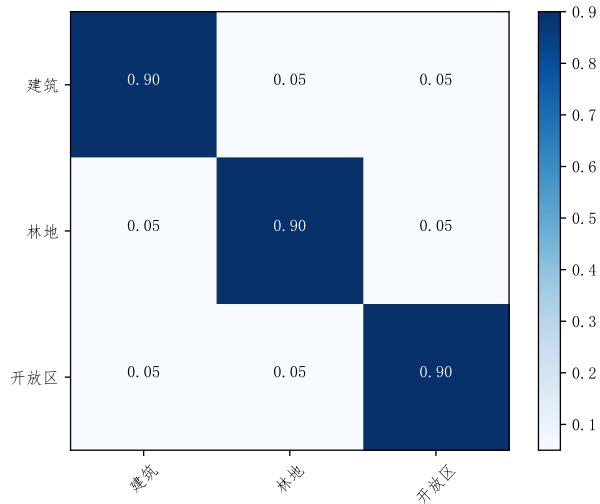


图 2-11 E-SAR Oberpfaffenhofen 数据 10% 对称噪声转移矩阵

图??展示了不同分类方法在 10% 对称标签噪声条件下的分类可视化图。根据图??与??展示的 Wishart 与 SVM 的分类结果，可以看出分类效果较差，尤其是在林地、建筑区域存在大量的错分样本，将建筑物区域错分为开放区，将林地区域错分为建筑区域，一方面是因为 Wishart 与 SVM 分类方法的分类性能有限，对散射

特性复杂的极化 SAR 数据分类结果不理想，另一方面是因为将错误的标签噪声信息归为学习规则之一，学习到错误的分类规则，导致分类性能的进一步下降。根据图??展示的在交叉熵损失函数优化的 CNN 分类器的分类结果，可以看到相比于 SVM 与 Wishart 的分类效果有了一定的提升，但是依然存在着错误分类情况，主要是将建筑物区域分类为开放区，这是因为建筑物区域的散射特性差异大，并且交叉熵损失函数的 CNN 会学习错误的标签信息，造成在训练集上过拟合，进而导致在测试集错分的情况。根据图??展示的 RSL 方法的分类结果，经过对噪声标签概率估计后过滤可能是错误的信息的操作，可以看到分类效果直观好于 CNN 方法，但是依然在建筑区、林地区域存在少量的错误块，这可能是因为 RSL 直接丢弃了错误的样本，丢失了训练样本中的数据分布特征，导致分类性能有限。图??与图??展示的 BEL 与 BEL+DP 方法的分类结果，可以看到错分情况得到改善，在类别区域内部更加平滑，错分块数量减少，证明了本章方法的有效性与优越性。

表??展示了不同实验方法在 10% 比例对称标签噪声下的分类数值结果。其中，SVM 与 Wishart 分类方法虽然在开放区域的分类准确率取得了不错的效果，分别为 95.28% 和 95.21%，但是由于在建筑区域和林地区域分类效果均不理想，导致整体准确率、平均准确率、Kappa 系数三相指标都较低，这是因为噪声标签信息对该两个分类方法的影响，学习到错误的分类规则。RSL 方法在林地区域单向准确率取得了最高的分类准确率为 95.16%，相比于本章方法提高了 3.94%，这可能是因为林地区域散射特性差异大，极化信息提取器在该区域的提升性能不足，同时该区域训练样本充足，直接丢弃所有的可能为噪声的样本并未丢失关键的数据分布特征。RSL+DP 方法在综合指标方面均取得了最高值，分别为 94.55%、93.52% 和 90.66%，相比于其他方法分别提升了 4.76%、5.6% 和 1.23%。综上所述，本章方法在噪声标签条件下的数据集中具有更优越的分类性能。

表 2-3 ESAR Oberpfaffenhofen 地区数据对称标签噪声下分类数值结果 (%)

序号	类别	SVM	Wishart	CNN	RSL	BEL	BEL+DP
1	建筑	75.19	61.35	81.25	75.32	89.98	<b>92.35</b>
2	林地	76.54	74.58	80.24	<b>95.16</b>	91.59	91.82
3	开放区域	95.28	95.21	85.82	93.28	93.39	<b>96.38</b>
	OA	79.07	83.05	83.67	89.79	92.56	<b>94.55</b>
	AA	82.33	77.04	82.44	87.92	91.65	<b>93.52</b>
	Kappa	77.25	70.45	79.14	89.43	89.68	<b>90.66</b>

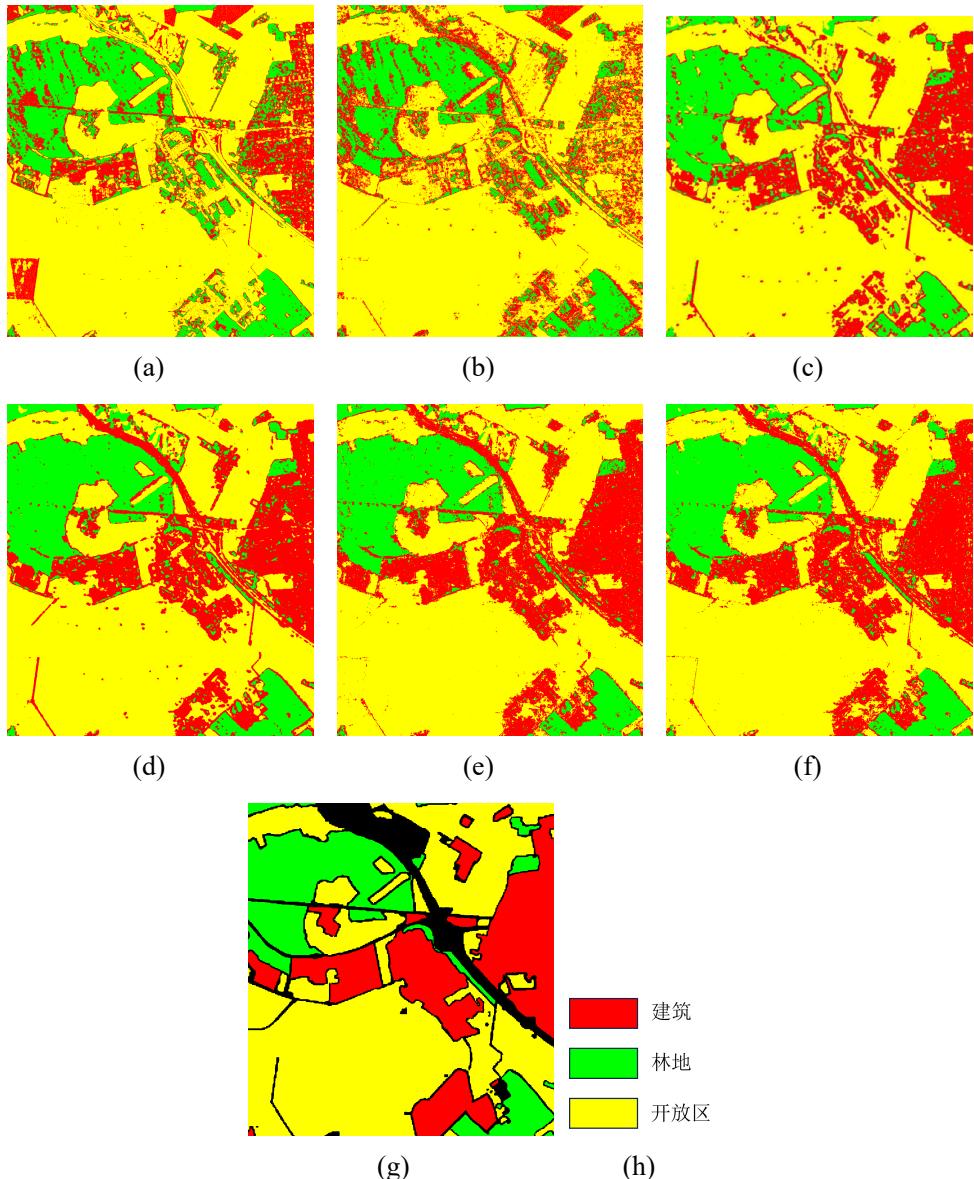


图 2-12 AIRSAR Flevoland 地区数据对称标签噪声下分类可视化结果图。

(a)SVM; (b)Wishart; (c)CNN; (d)RSL; (e)BEL; (f)BEL + DP; (g) 地面真值;

(h) 颜色与类别对应关系

#### 2.4.3.2 非对称标签实验

图??展示了 10% 噪声比例下非对称标签噪声的噪声转移矩阵，每个类别正确标记的概率为 90%，10% 的标签噪声随机映射到另外的一个类别中。

图??展示了 10% 非对称标签噪声下不同分类方法的分类结果。图??与图??展示的 Wishart 与 SVM 的分类结果，可以在林地、建筑区域都存在较多的错误区域，分类效果差。图??为交叉熵损失下的 CNN 分类结果，可以看到受到标签噪声影响，

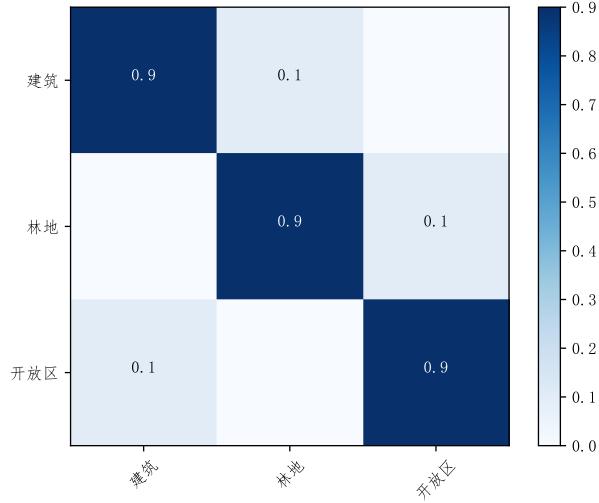


图 2-13 ESAR Oberpfaffenhofen 数据 10% 非对称噪声转移矩阵

开放区域错分为建筑、建筑区域错分为林地的情况存在较多。图??为 RSL 方法的分类结果。。。图??与图??为 BEL 与 BEL+DP 方法的分类结果，可以看出在建筑物区域和林地区域的错分情况减少，分类结果更加平滑。因此，证明了本章方法的有效性与优越性。

表??展示了非对称标签噪声下的分类数值结果。SVM、Wishart 以及 CNN 方法在建筑区域的分类准确率较低，导致总体准确率较低，分别为 76.58%，80.32%，82.49%，这是因为这三个方法均没有筛选错误样本的能力，学习错误的分类规则。本章方法除了林地区域，其他的分类指标均取得最高值，分类准确率均在 90% 以上，且总体分类准确率达到 94.32%，较其他方法提升了 2.74%，这也证明了本章方法具备应对非对称标签噪声的鲁棒性，进一步证明本章方法的有效性。

表 2-4 ESAR Oberpfaffenhofen 地区数据非对称标签噪声下分类数值结果 (%)

序号	类别	SVM	Wishart	CNN	RSL	BEL	BEL+DP
1	建筑	72.38	58.83	78.92	86.73	88.06	<b>90.48</b>
2	林地	74.54	72.43	81.32	<b>94.44</b>	89.44	90.83
3	开放区域	92.28	92.31	83.16	90.53	94.04	<b>97.18</b>
	OA	76.58	80.32	82.49	91.58	91.8	<b>94.32</b>
	AA	79.73	74.52	81.13	90.56	90.51	<b>92.83</b>
	Kappa	74.71	68.93	77.32	86.97	87.88	<b>90.31</b>

图??展示了各个对比方法在两种噪声类型下分类准确率随着噪声比例的变化情况。可以看出，分类准确率在两种类型标签噪声中随着噪声比例的变化趋势与

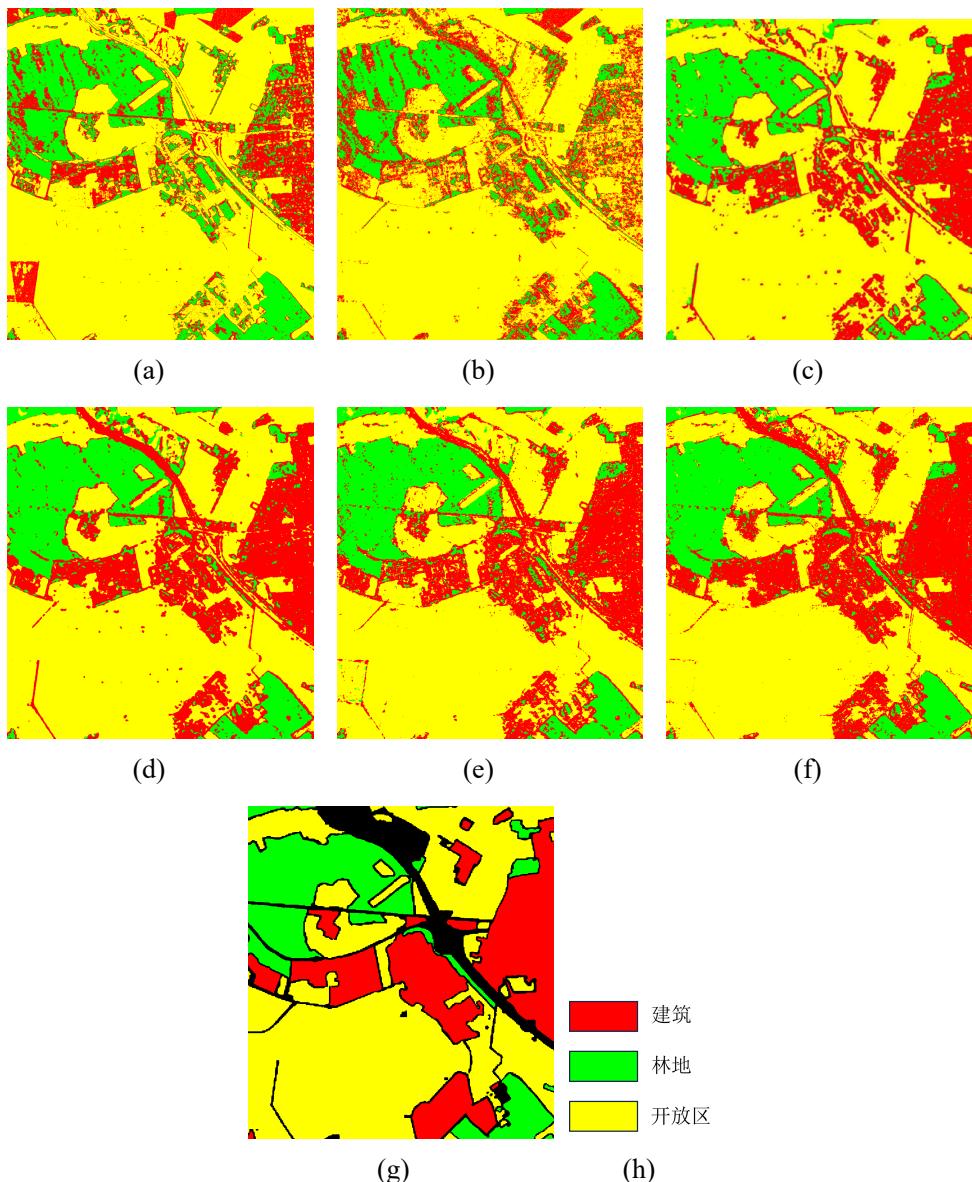


图 2-14 AIRSAR Flevoland 地区数据非对称标签噪声下分类可视化结果图。

(a)SVM; (b)Wishart; (c)CNN; (d)RSL; (e)BEL; (f)BEL + DP; (g) 地面真值;  
 (h) 颜色与类别对应关系

AIRSAR Flevoland 地区实验结果保持相同，均呈现下降趋势，且非对称噪声源下分类准确率普遍低于对称噪声源。在非对称标签噪声下，SVM、Wishart、CNN 这三种方法由于未考虑噪声标签的错误信息，在噪声比例较高时分类规则明显被改变，导致分类准确率大幅度下降，从 10% 的噪声比例至 50% 分别下降了 20.64%、28.82% 和 15.30%。而考虑了标签噪声的 RSL 方法分类准确率下降了 13.29%，表现出对标签噪声具有一定的鲁棒性，但是在 50% 噪声比例下准确率仅为 76.5%，也表现出了该方法的局限性。而 BEL 与 BEL+DP 方法分别下降了 7.62% 和 7.8%，

展现出对标签噪声的优越的鲁棒性能，同时在 50% 噪声比例下依然保持在 85% 左右及以上。在非对称标签噪声下，各个方法准确率出现大幅下降，SVM、Wishart、CNN 的最差准确率均低于 50%，RSL 与本章方法最差分类准确率也几乎相同，都在 65% 左右，进一步表明非对称标签噪声对分类规则影响更大。结合以上数据，表明本章方法对不同类型标签噪声、不同噪声比例下，具有更加优越的噪声标签鲁棒性能。

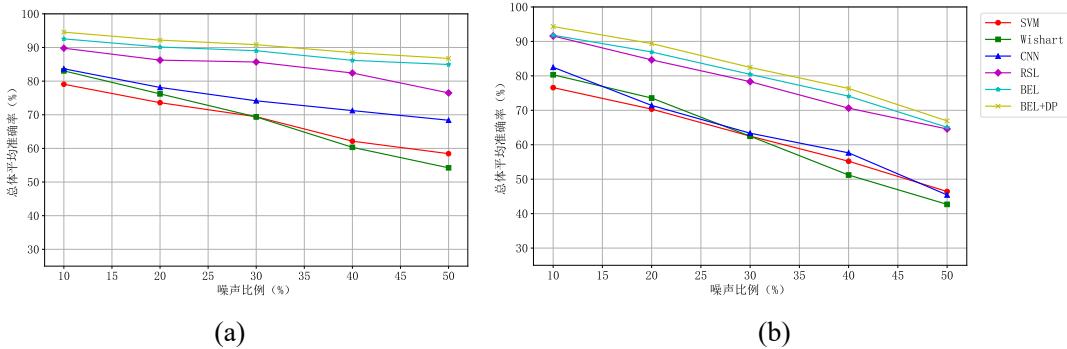


图 2-15 ESAR Oberpfaffenhofen 地区数据不同噪声比例下分类准确率结果。

(a) 对称噪声源 (b) 非对称噪声源

#### 2.4.4 本章小结

针对极化 SAR 图像中存在标签噪声污染问题，本章研究了基于混合模型和边界增强的鲁棒性极化 SAR 目标分类方法。首先介绍了有限混合模型，并分析了两种混合模型的优缺点。在此基础上，通过深度学习在包含标签噪声样本下训练损失分布存在差异，提出了基于混合模型的样本概率估计方法。同时，为了增强模型对边界信息的感知能力，充分利用边界样本信息，对极化 SAR 的 Pauli 伪彩图使用 Sobel 算子提取边界并膨胀，对膨胀边界内样本进行损失增强。最后，提出了动态的自学习损失函数，通过噪声样本概率动态调整模型预测与真值标签的依赖，实现模型的鲁棒性训练过程。最终，本章通过在两个真实极化 SARS 数据集，对两种噪声源进行模拟实验，验证本章方法的有效性与优越性。同时，还对不同噪声比例下各个方法的分类准确率进行了比较，进一步验证本章方法的有效性与优越性。

## 致 谢

行文至此，意味着在成电三年的研究生时光即将画上句号。三年时光，如同一部精彩纷呈的交响乐，旋律中交织着喜悦与忧愁，欢乐与焦虑，每个音符都是我成长路上不可或缺的部分。

感谢杨建宇老师、武俊杰老师、李中余老师以及团队中每一位尊敬的导师。感谢你们的辛苦付出，让雷达探测成像团队形成了良好的科研环境与科研环境，让我们感受到来自团队的强大自信力。

感谢裴季方老师的耐心指导与支持。每当我在学术探索中遇到难题时，总能在百忙之中抽出宝贵时间，以无尽的耐心和细致的关怀，为我排疑解惑。他严谨治学精神和深厚的专业知识，不仅为我指明了科研的方向，更为我的个人成长树立了榜样。同时，感谢吴从文师兄的细心帮助与支持，他的建议和指导对我完成学术论文有着重要的影响。

感谢研究生阶段的所有好伙伴们。感谢你们在这段旅程中的陪伴与支持，我们一起经历风雨，一起分享快乐，你们的存在让这段时间更价丰富多彩。

感谢我的女朋友陈阳一路的鼓励与支持。你是我最坚强的后盾，在我遇到困难和挑战时，你总是给予我最坚定的支持和最温暖的鼓励，感谢你的陪伴与理解。

最后，感谢各位评审专家的辛苦付出。