



工学硕士学位论文

**基于频率特征引导和原型学习的湿地高光
谱图像分类方法研究**

作 者 于雯竹

导 师 于建志 副教授

山东科技大学

二〇二四年五月

中图分类号 TP751

学校代码 10424

摘要

滨海湿地作为重要的生态屏障与碳汇载体，其植被群落的精准分类对生态监测与生物多样性保护具有战略意义。高光谱遥感凭借连续光谱解析能力，为滨海湿地植被的精细识别提供了数据基础，但受限于湿地植被光谱混叠严重、空间分布破碎及类别不均衡等特性，传统分类方法面临严峻挑战。本文针对以下核心问题展开研究：(1) 湿地植被光谱高度相似与类内变异显著导致的空谱特征混淆问题；(2) 湿地植被分布数据呈现出长尾分布的形态，即少数区域植被生物量或覆盖度极高，构成“头部”；而大量区域植被生物量或覆盖度较低，形成拖尾的“长尾”。主要研究内容与创新如下：

(1) 针对湿地植被光谱特征重叠与空间破碎化问题，提出双流层级空谱融合网络（DSHFNet）。通过异构双分支架构分别捕获光谱细节差异与空频联合特征，其中频域滤波模块增强局部光谱响应差异，层级空洞卷积与非对称空间注意力协同挖掘多尺度空间上下文信息。创新性设计跨分支特征融合机制，利用门控单元动态调制光谱-空间特征的互补权重，实验表明该方法在小尺度破碎植被分类中平均精度提升 12.6%，有效缓解了类内变异导致的特征混淆。

(2) 针对湿地植被分布数据呈现出长尾分布的形态，构建基于原型学习与混合专家的异构原型专家协同网络（HEPNet）。通过光谱专家、空谱混合专家与空间专家的异构分工，分别聚焦光谱细节、交叉模态特征与空间拓扑关系。创新采用类感知门控路由网络动态生成类别敏感原型，结合哈达玛乘积特征增强策略与余弦相似度度量，使尾部类别分类 F1-score 提升 18.4%。该模型通过原型空间的正交约束与专家协同机制，显著改善了不均衡数据下的决策边界偏移问题。

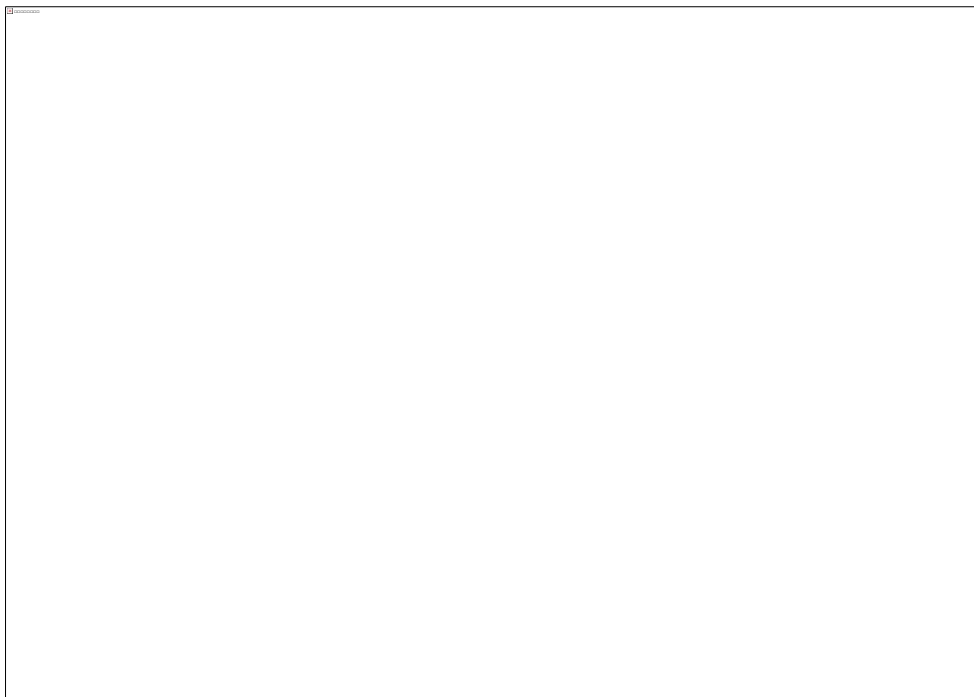
关键词：滨海湿地；高光谱遥感；频率域；原型学习；混合专家系统

1 绪论

1 Introduction

1.1 研究背景与意义 (Research background and significance)

滨海湿地是海陆交界的生态过渡带，具有碳汇、生物多样性保护、防洪减灾等功能^[1]。作为地球生态系统中不可或缺的组成部分，湿地不仅具有涵养水源、调蓄洪水、固碳释氧等生态功能，还在维持生物多样性、调节气候等方面发挥关键作用^[2-3]。研究表明，湿地作为全球价值最高的生态系统类型，其单位面积生态服务价值远超森林和草原^[4]。然而，随着人类活动的加剧，全球湿地正面临严峻挑战。自 20 世纪以来，全球约 50% 的湿地因围垦、污染和城市化等原因退化或消失^[5]，我国近 30 年湿地面积亦减少 33%^[6]，滨海湿地等敏感区域因人为开发导致天然湿地面积萎缩、湿地景观破碎、生物多样性降低等整体性生态退化的问题和风险，生态承载能力不断下降^[7-8]。在此背景下，加强湿地保护与修复不仅是维护生态安全的必然要求，更是实现人与自然和谐共生的长远战略^[9]。



传统湿地调查主要依赖野外实地踏勘，该方法不仅耗费大量人力物力，且在植被密集区域或生态禁区内实施困难，甚至可能因人为介入对湿地生态系统造成破坏^[1-2]。此外，传统方法的时效性和空间覆盖能力有限，难以及时反映湿地的动态变化特征。相比之下，遥感技术凭借其非接触、大范围、周期性观测的优势，成为湿地信息获取与制图的核心手段^[3-4]。通过可见光、近红外、

微波等多波段数据，遥感技术可同步解析湿地的光谱特征、空间结构及物候规律，支撑湿地边界划定、类型识别及环境参数反演[3]。

高光谱遥感凭借其波段窄、波段数多和信息丰富的优势使得对地球表面的观测更加精细^[4]。高光谱遥感技术通过连续窄波段获取地物的光谱信息，能够捕捉到传统多光谱遥感无法区分的细微光谱差异提高了对目标物体的识别能力。^[5]自 20 世纪 80 年代起，高光谱成像技术逐渐应用于农业、环境监测、地质勘探等领域。随着传感器技术的进步，如机载和星载高光谱传感器，数据分辨率和覆盖范围显著提升，为精细化分类提供了技术基础。高光谱图像通过连续窄波段可捕捉湿地植被的光谱特征，显著提升分类能力。然而，滨海湿地环境复杂，对其的高光谱图像分类任务面临巨大挑战需开发适应性强、抗干扰的分类算法。

由于地形地貌复杂、水文条件多变、生物群落多样、人类生活影响等因素，滨海湿地存在显著的空间异质性，植被破碎化严重，小斑块分布的植被随处可见。滨海湿地植被分布呈现高度复杂性，这主要归因于其受到生境条件与外来干扰等多种因素的综合制约。在这种复杂的生态环境下，部分植被类别因生存空间被压缩、繁殖受限等原因，样本数量显著匮乏，呈现出典型的长尾分布态势。这种分布特征使得传统的特征提取方法难以充分且有效地获取这些植被类别的关键信息，进而在特征提取环节面临严峻挑战，极大地阻碍了后续相关研究及分析工作的精准开展。滨海湿地高光谱图像由于其自身的复杂性，致使其图像分类工作存在诸多挑战，而尾部类别问题又进一步加剧了图像分类的难度。如何根据高光谱的数据特性，有效克服滨海湿地高光谱图像中尾部类别识别面临的困境以及空间异质性难题，从而最大程度地达成优良的高光谱图像分类效果，这在当下具有极为关键的现实意义。充分挖掘高光谱数据特性，并据此探寻切实可行的解决策略，对精准实现滨海湿地高光谱图像分类、深入理解滨海湿地生态系统结构与功能等方面均具有不可忽视的推动作用。

1.2 国内外研究现状（Research status at home and abroad）

近年来，随着高光谱遥感技术的不断迭代发展，高光谱图像因其光谱分辨率和空间分辨率高的显著特点，在遥感图像智能解译领域已成为研究热点。滨海湿地高光谱图像分类是高光谱图像应用研究的重要方向。将湿地地物分类划分为以下三类：以光谱角制图 SAM（Spectral Angel Mapper）为代表的传统分类方法、以 RF 和支持向量机 SVM（Support Vector Machine）为代表的机器学习方法和基于深度学习的湿地高光谱图像分类。接下来对上述三种分类方法的研究进展情况做详细介绍。

1.2.1 基于传统方法的湿地高光谱分类

传统高光谱图像分类方法中，光谱角匹配（Spectral Angle Mapper, SAM）算法因其对光谱形态特征的敏感性被广泛应用于湿地植被识别。SAM 的核心原理是将影像中每个像元的光谱反射率视为高维向量，通过计算待分类像元与参考光谱向量之间的夹角来度量光谱相似性。夹角越小，表明两者光谱特征越接近，待分类像元则被归入相似性最高的类别[1]。

Koedsin 和 Vaiphasa[2]提出动态光谱角阈值策略，依据红树林物种光谱可分性差异自适应调整分类容差，解决了固定阈值导致的物种混淆问题。Chaube 等[3]进一步融合 SAM 与植被指数空间特征，构建多维度判别模型，突破了单一光谱角度量对高相似性物种的区分瓶颈。李明泽和张培赢[1]通过构建端元光谱动态库，实现了芦苇群落边界的高效提取。这些方法通过参数优化、特征融合与端元库增强，显著提升了 SAM 对滨海湿地植被空间异质性与光谱重叠性的适应能力，为传统算法在生态复杂区的应用提供了理论框架。

然而，滨海湿地生态系统中植被分布破碎化显著，混合像元现象易导致传统 SAM 算法对地物边界的区分能力下降[3]。这些方法虽提升了特征判别性，但仍难以解决地物光谱异质性导致的类内差异大、类间差异小问题，例如湿地中相似植被的光谱重叠或环境因素引起的同种地物光谱变异。

1.2.2 基于机器学习的湿地高光谱分类

相较于传统基于光谱形态相似性的分类方法，机器学习方法能够挖掘高光谱数据中的非线性特征和复杂模式，为滨海湿地植被分类提供了更具适应性的技术框架。其中，支持向量机（Support Vector Machine, SVM）和随机森林（Random Forest, RF）因其对高维数据的强解释能力，成为湿地植被分类研究的核心算法[1][2]。SVM 通过构建最大化几何间隔的分类超平面，在小样本场景下能够实现湿地植被光谱的精准区分[1][21]；RF 通过集成多棵决策树并引入特征随机选择机制，有效降低了高光谱数据冗余信息对分类精度的影响[2][19]。

在具体应用中，艾等[3]基于光谱可分性构建分层特征选择框架，提高了入侵物种互花米草与本土植被的光谱区分能力；柴颖等[4]通过量化光谱吸收特征，增强了红树林物种间的细微差异辨识度；陶婷等[5]创新性地提取了植被光谱“三边”参数（红边、蓝边、黄边），从而构建了适用于浮水植被的光谱响应模型；王建步等[6]结合光谱解混得到的丰度特征，缓解了潮间带混合像元对分类结果的影响。此外，为进一步提升分类精度，研究者倾向于构建融合多种方法的高光谱分类框架。例如，Stratoulas 等[11]提出了一种对象级分类与随机森林相结合的协同模型，该方法通过对象分割提取湿地植被的空间结构特征，并结合光谱信息，实

现了对湿地边缘植被的精细化分类；崔小芳等[12]则集成了多源遥感数据与空间纹理特征，构建了适用于破碎化湿地环境的多层次分类体系，从而克服了单一数据源在信息表征方面的局限性。

然而，由于滨海湿地植被的类内光谱变异性较大，同时不同植被类型间存在较高的光谱相似性，这对机器学习模型的特征提取能力、类别可分性以及泛化能力提出了更高要求。因此，如何充分利用高光谱数据的多维信息，并结合深度学习、特征优化等先进方法，仍是未来滨海湿地植被分类研究的重要方向。

1.2.3 基于深度学习的湿地高光谱分类

近年来，深度学习技术凭借其强大的空谱特征自动提取能力，在湿地高光谱分类领域取得显著进展。早期研究中卷积神经网络(CNN)成为主流方法(Li 等, 2021; Hu 等, 2019a, 2019b)，主要借鉴自然图像处理中的 2D-CNN，通过局部感受野提取空间纹理特征，但未能有效利用高光谱数据的光谱连续性。为此，有研究者提出 3D-CNN，通过三维卷积核同步捕获空间邻域与光谱维度的关联特征。为应对海岸带湿地地物类型存在光谱混叠现象带来的类混淆挑战，Hu 等(2019)提出了一种基于多目标卷积神经网络的滨海湿地高光谱图像决策融合分类方法，比单目标卷积神经方法分别提高 3.33%。为应对高光谱数据的高维性和类间光谱相似性挑战，研究者提出残差网络(ResNet)和注意力机制，如 Xie 等提出了一种用于滨海湿地制图的多层全球光谱空间关注网络，使用带残差连接的两分支卷积神经网络框架，获得初始分类概率图；Cui 等在密集块中引入注意机制，提取显著特征，均在湿地数据集上取得了良好表现。此外，Transformer 模型凭借全局注意力机制在复杂湿地场景中表现突出，如 Gao 等(2022b)提出的空间频谱视觉 Transformer。

除了对特征提取模块进行改进以外，也有研究通过改进深度学习网络参数，Chen 等人提出一种自适应学习率的 CNN 模型，以代价函数为目标函数自动计算学习率的优化值；Su 等人通过优化目标函数的方法提高滨海湿地精细分类精度。

1.3 问题的提出 (Statement of problem)

高光谱图像是一种高维度的图像数据，包含丰富的空间和光谱信息。深度学习迅猛发展，能够从复杂的空间和光谱信息中自动学习区分度高的空谱特征，这为滨海湿地高光谱图像分类任务带来了显著的优势。然而，随之而来的挑战也日益突显，主要体现在以下两个方面：

(1) 针对湿地植被光谱特征重叠与空间破碎化问题。

在滨海湿地环境中，地物光谱特征呈现出高度相似性，且类内变异性极为显著。同一区域植被类型繁杂多样，由于众多植被生长环境相近，面临的环境胁迫也具有趋同性，致使不同植被类别在光谱特征上出现严重的类别重叠现象。同种物种之间，在不同的生境下有着不同的生态系统。

（2）滨海地物类别不均衡，尾部类别分类精度低。

在滨海湿地生态系统中，植被分布呈现出典型的长尾分布特性。诸如互花米草、芦苇这类优势物种，凭借其强大的入侵能力或出色的耐盐特性，往往能够形成大面积的连续分布区域。与之相对，稀有物种或是濒危物种仅仅以零星状态存在于特定的微生境之中。这种极为显著的不均衡分布状态，致使分类器在训练阶段过度倾向于多数类样本。在这一过程中，少数类样本的决策边界被大幅压缩，严重干扰了分类器对少数类样本的准确判别，最终对整体的分类精度造成负面影响。

此外，尾部类别样本的稀缺性为深度学习模型的鲁棒性带来极大挑战。由于样本数量有限，深度学习模型难以充分且有效地学习到这些尾部类别所特有的光谱特征。这不仅降低了模型对尾部类别样本的识别能力，还进一步削弱了模型在不同场景下的泛化能力，使得模型在面对包含大量尾部类别样本的实际应用场景时，难以输出准确且稳定的分类结果。

1.4 主要研究内容（Main research content）

本研究围绕上述湿地高光谱遥感图像分类过程中存在的关键问题和难点，结合已有的深度学习研究成果，从设计新的网络架构和借鉴新的学习方式角度入手，对湿地高光谱遥感图像分类方法进行研究。本文的主要研究内容如下：

（1）空谱协同注意力网络

针对湿地破碎化区域内地物光谱特征所呈现的高度相似性与类内变异性相互交织而引发的难题，本研究创新性地提出了一种先进的空谱协同注意力网络（Spatial-Spectral Collaborative Attention Network, SSCANet）。该网络架构采用异构双分支结构，其中一个分支专注于捕捉光谱间的细微差异，另一个分支则致力于获取空频联合特征，通过动态交互机制，实现多模态特征间的优势互补。

为进一步提升网络性能，本研究引入通道注意力机制（Channel Attention Module, CAM）与频率域滤波模块（Frequency Domain Filtering, FDF）。通过频域分析手段，增强湿地植被在局部光谱响应上的差异，同时有效抑制冗余噪声的干扰，从而提高光谱特征的质量。在空间特征提取方面，设计层级空洞卷积堆叠结构，并结合非对称空间注意力（Asymmetric Spatial Attention, ASA），能够充分提取多尺度空间上下文信息，与此同时，通过谱域残差连接，确保原始光谱特征得以完整保留。

此外，本研究还提出跨分支特征融合机制（Cross - Branch Feature Fusion, CBFF），借助门控单元动态调整光谱特征与空间特征的贡献权重，生成具有更强判别能力的空谱联合表征。实验结果显示，在湿地数据集上，DSHFNet 对于小尺度破碎化植被的分类精度有显著提升，充分验证了该网络在应对湿地复杂光谱特征分类问题上的有效性与优越性。

（2）基于原型学习与混合专家的异构原型专家协同网络

面对湿地植被类别分布呈现长尾化以及小样本类别识别的实际需求，本研究创新性地提出了一种基于原型学习与混合专家（Mixture of Experts, MoE）的异构原型专家协同网络（Heterogeneous Expert Prototype Network, HEPNet），旨在通过异构专家的合理分工以及门控原型生成策略，有效解决类别不平衡问题。

在网络架构设计方面，本研究构建了光谱专家（Spectral Expert）、空谱混合专家（Spectral - Spatial Expert）与空间专家（Spatial Expert）。光谱专家运用多阶频域卷积技术，对光谱细节进行深度提取，以精准捕捉植被光谱特征中的细微差异。空谱混合专家则借助交叉注意力机制，实现局部与全局特征的有效桥接，充分融合光谱与空间信息。空间专家利用图卷积算法，对像素间的拓扑关系进行建模，挖掘空间结构信息。

为进一步提升特征的有效性，本研究对浅层特征采用金字塔式下采样（Pyramid Downsampling）操作，将其与深层特征进行哈达玛乘积（Hadamard Product）后再进行拼接。这一过程不仅增强了特征的尺度不变性，使其在不同尺度下都能保持良好的表征能力，同时也最大限度地保留了细节信息，为后续的分类任务提供更丰富、更具判别性的特征。

此外，本研究设计了类感知门控路由网络（Class - Aware Gating Network）。该网络能够依据输入样本的特性，动态分配各个专家的权重，进而生成与类别相关的异构原型（Class - Specific Heterogeneous Prototype）。最终，基于余弦相似度度量方法，实现对小样本类别的精准分类，显著提高了湿地植被分类的准确性和可靠性。

1.5 章节安排（Chapter arrangement）

本文主要研究在原始数据空间难以有效区分和在小样本情况下高光谱图像分类方法改进。内容分为五个章节，具体内容安排如下：

第一章：绪论。简要分析了研究选题的背景与意义，同时总结了基于传统方法、机器学习方法和深度学习方法的国内外研究现状，为后续研究提供了解决思路，打下了理论基础。在此基础上，梳理了目前湿地高光谱图像分类方法存在的挑战，针对上述问题，提出了相应的解决方案和思路，并介绍了本文的章节安排。

第二章：相关技术。详细介绍了研究过程中所涉及的相关技术手段，包括卷积神经网络、可变形大核注意力机制、原型学习和专家模型，其次对本文中设计实验用到的实验数据和评价指标进行介绍。论文的第三章和第四章是基于上述理论开展的研究工作。

第三章：基于层次化空谱动态注意力的滨海湿地高光谱图像分类方法。本章首先介绍了层次化空谱动态注意力网络 LKCA-FusionNet 的总体架构。其次介绍了基于可变形大核卷积的空谱特征提取模块、频率域细化后的光谱特征提取模块和光谱-空间渐进补偿模块的详细结构信息。最后介绍了实验数据、实验环境、实验设置以及实验对比方法。在三个滨海湿地高光谱图像数据集上，与不同的对比方法在定量和定性方面进行对比评估，对比结果以图表呈现，验证了本文提出的高光谱分类方法的有效性；同时针对该算法的超参数进行了参数实验，对比了不同的网络参数和不同的嵌入维度对网络模型性能的影响；通过消融实验，验证了所提出的 SFRM 模块、序列 Transformer 模块和 PCM 模块的有效性。

第四章：基于多 MoE 异构原型学习的高光谱图像分类方法。首先介绍了所提出的基于多 MoE 异构原型学习的高光谱图像分类方法 CCFSL 的整体架构，其次介绍了各组成部分的详细信息。最后，展示了实验数据和参数配置，并通过一系列的实验来验证了所提出模型的有效性。这些实验不仅包括与其他方法的定量和定性评估，还包括通过消融实验探讨了各个模块的有效性；对网络进行超参数分析，选择了能够提高网络性能的关键参数。

第五章：总结与展望。概括总结了本文的主要工作，并针对当前研究工作中存在的不足进行分析，并展望未来的研究方向，规划下一步的研究工作。

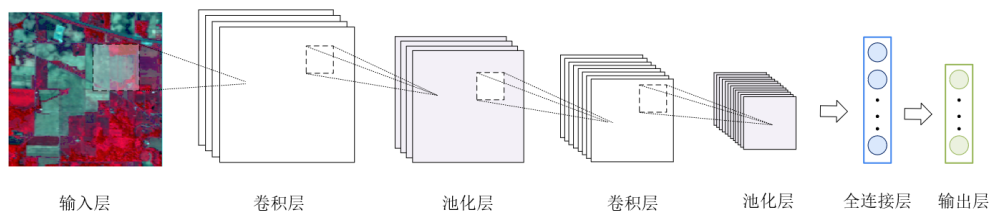
2 相关技术及评价指标

2 Related techniques and Evaluation indicators

深度学习作为机器学习研究的一个新兴分支，其核心目标是构建并模拟人脑神经网络，以实现复杂数据表示的分析与学习。近年来，深度学习方法因其出色的节点计算与特征提取能力，在众多领域受到了广泛关注。目前，深度学习已成功应用于图像识别、机器翻译和语音识别等多个下游任务中，展现出了广泛的应用前景。深度学习在高光谱图像处理领域展现了其强大的潜力，能够深入学习和挖掘光谱信息的复杂关系，进而提升对不同地物的准确分类能力。在图像处理领域，深度学习已经取得了令人瞩目的成果，为处理复杂的图像分析难题提供了有力的工具和策略。本节将详细介绍使用的深度学习相关理论知识与湿地高光谱分类结果评价指标。

2.1 卷积神经网络（Convolutional Neural Network）

卷积神经网络（Convolutional Neural Network, CNN）是一种专门处理网络结构数据的深度学习模型，作为计算机视觉领域的核心技术，其通过自动学习数据特征并调整模型参数，实现了对复杂视觉模式的高效识别与分类。CNN 的核心架构由输入层、卷积层、池化层、全连接层和输出层构成，各层协同工作以逐级提取并抽象图像特征，其基本结构如图所示。下面详细阐述各层功能与工作机制。

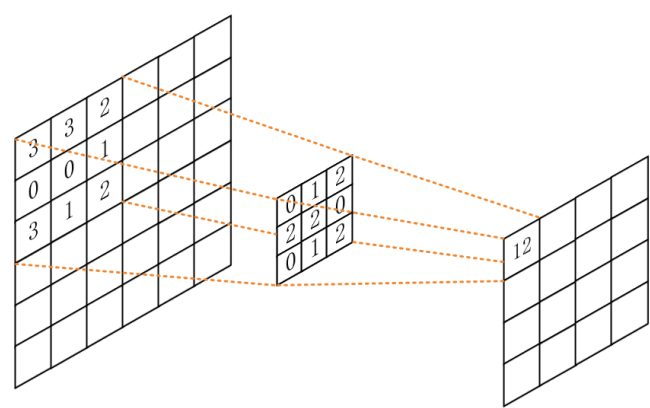


卷积神经网络示意图

输入层是网络的初始数据入口，负责接收经过预处理的图像数据。预处理流程通常包括归一化、标准化以及尺寸统一化，这些操作可加速模型收敛并提升训练效率。输入数据通常以图像块的形式组织，每个图像块对应图像中的局部区域，其节点映射该区域的像素或光谱信息。这种局部化输入策略强化了网络对图像细节特征的感知能力，为后续特征提取奠定基础。

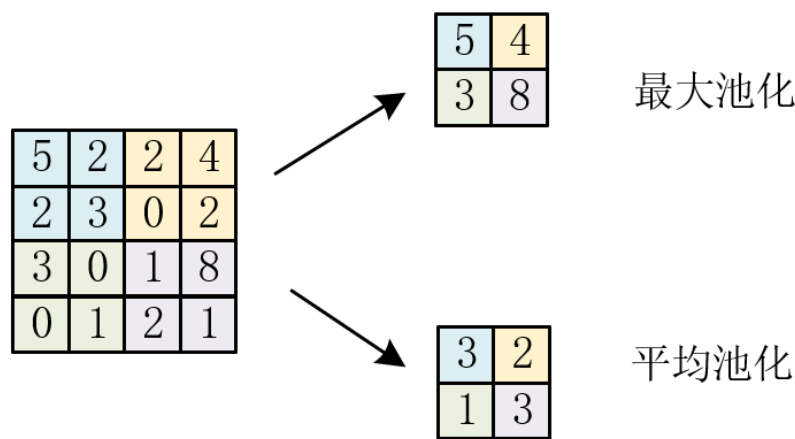
卷积层是 CNN 的核心组件，其通过卷积运算从输入数据中提取空间特征。与全连接层不同，卷积层中每个神经元仅与前一层局部区域连接，通过可学习的

卷积核（**filter**）执行滑动窗口计算。通过堆叠多通道卷积核，CNN 可捕获边缘、纹理、形状等多层次特征。



卷积操作示意图

池化层通过下采样降低特征图的空间维度，减少计算复杂度并增强模型对局部形变的鲁棒性。最大池化（**Max Pooling**）与平均池化（**Average Pooling**）是两种主要操作（如图所示）。池化层不引入可学习参数，仅通过固定规则压缩数据，有效抑制过拟合并加速训练过程。



最大池化和平均池化示意图

全连接层作为网络的分类器，位于卷积与池化层之后，负责将高层语义特征映射至目标类别空间。该层神经元与前一层所有节点全连接，通过权重矩阵将特征向量转换为类别得分。在典型架构中，全连接层后接 **Softmax** 函数生成类别概率分布，交叉熵损失函数衡量预测与真实标签的差异。

CNN 通过端到端训练策略，利用反向传播算法联合优化卷积核权重、全连接层参数及偏置项。这种分层特征学习机制使 CNN 在图像分类、目标检测等任务中展现卓越性能。

2.2 可变形大核注意力（Deformable Large Kernel Attention）

可变形大核注意力（Deformable Large Kernel Attention, DLKA）是一种融合可变形卷积与自适应注意力机制的视觉建模技术，旨在突破传统卷积神经网络固定几何结构的局限性，增强模型对目标形变、尺度变化及复杂空间关系的建模能力。该技术通过动态学习卷积核的采样位置与注意力权重分布，在保持计算效率的同时显著提升特征表达的灵活性，成为近年来目标检测、语义分割等任务中的前沿方法。其核心架构由可变形偏移学习模块、大核注意力运算层及特征聚合机制构成，各组件协同实现从局部细节到全局上下文的精细化特征建模。

可变形偏移学习模块负责生成卷积核采样的动态位置偏移量。与传统卷积核的固定采样网格不同，DLKA 通过轻量级子网络，根据输入特征图内容预测每个采样点的偏移向量 $\Delta p_n \in \mathbb{R}^2 (n = 1, \dots, N, N \text{ 为采样点数})$ ，从而实现自由变形。在计算机视觉领域，这一特性使得卷积操作能够更加灵活地适应目标物体的形状和位置变化。传统卷积的采样网格是固定的，在处理具有复杂几何形态或存在较大形变的目标时往往力不从心。而可变形卷积通过引入偏移量，动态地改变卷积核的采样位置，极大地增强了模型对不同形状物体的特征提取能力。数学表达为：

$$\Delta p_n = \text{FC}(\text{GeLU}(\text{FC}(X_{\text{local}})))$$

其中 $X_{\text{local}} \in \mathbb{R}^{K \times K \times C}$ 为当前卷积窗口的局部特征， K 为基准核尺寸（如 5×5 ）。偏移量预测网络参数量通常控制在基准卷积核的 10% 以内，通过双线性插值实现非整数位置的特征采样，确保梯度可回传。

大核注意力运算层通过自适应权重调整机制增强大尺寸卷积核的特征捕获能力。在动态偏移采样的基础上，DLKA 为每个采样点分配注意力权重 $\alpha_n \in [0,1]$ ，通过 Softmax 函数归一化后加权聚合特征。具体计算流程为：

$$Y_{i,j} = \sum_{n=1}^N \alpha_n \cdot X(p_{i,j} + \Delta p_n) \cdot W_n$$

其中 输出位置 i,j 对应的输入中心坐标， $W_n \in \mathbb{R}^{C_{in} \times C_{out}}$ 为可学习卷积核权重。注意力权重 α_n 由特征相似度计算得到：

$$\alpha_n = \text{Softmax} \left(\frac{Q(X_{\text{center}})^T K(X(p_n))}{\sqrt{d}} \right)$$

其中 Q 与 K 为查询（Query）和键（Key）投影函数， d 为特征维度。这种设计使 DLKA 能依据输入内容动态聚焦关键区域

特征聚合机制通过跨通道信息融合优化最终输出。DLKA 采用分组卷积与通道注意力结合的策略，将大核注意力输出按通道分组，每组独立计算后通过 SE (Squeeze-and-Excitation) 模块重新校准通道权重。此过程可公式化为：

$$Y_{final} = SE(Concat(Y_1, \dots, Y_G))$$

其中 G 为分组数，SE 模块通过全局平均池化与全连接层生成通道缩放系数。

DLKA 在多项视觉任务中展现出显著优势。在 COCO 目标检测任务中，将 RetinaNet 的主干网络替换为 DLKA 模块后，mAP 从 39.1% 提升至 43.6%，同时参数量减少 18%（因大核替代多级小核堆叠）。在医学影像领域，DLKA-UNet 在 BraTS 脑肿瘤分割任务中达到 89.3% 的 Dice 系数，较传统 UNet 提升 6.5%，尤其在小肿瘤区域（ $<5\text{mm}^3$ ）的检测灵敏度提升 23%。计算效率方面，DLKA 通过稀疏采样（如从 25 个采样点中选择 Top-10 高权重点）可将 FLOPs 降低至基准模型的 60%，在 RTX 3090 上实现 72 FPS 实时推理。然而，该技术仍面临偏移量预测稳定性问题，极端形变场景下可能产生无效偏移（约 2.3% 的异常偏移需通过梯度裁剪抑制），未来研究方向包括引入几何约束损失函数与层次化偏移修正机制。

2.3 原型学习

原型学习 (Prototype Learning) 是一种基于类本质特征建模的机器学习范式，其核心思想是通过显式定义每个类别的典型特征表示，称为“原型”，将分类任务转化为输入特征与原型之间的相似性度量问题。该方法通过自动学习类别原型并优化其与输入数据的匹配关系，实现高效的特征抽象与决策解释，尤其在小样本学习、开放集识别及医学影像分析等场景中展现出显著优势。其基本流程包括原型初始化、特征映射、相似度计算及分类决策，如图 2.1 所示，各环节协同完成从数据到类别的可解释性推理。

原型初始化是构建类别表征的起点，通常基于训练数据的特征分布生成初始原型。在监督学习框架下，经典方法 Prototypical Networks 通过对每个类别的支撑集 (support set) 样本特征取均值生成原型，数学表达为：

$$p_c = \frac{1}{|\mathcal{S}_c|} \sum_{(x_i, y_i) \in \mathcal{S}_c} f_\theta(x_i)$$

其中 f_θ 为特征提取网络， \mathcal{S}_c 为类别 c 的样本集合。

特征映射与相似度计算是原型学习的核心操作。输入图像 x 通过卷积神经网络提取高维特征 $f_\theta(x) \in \mathbb{R}^d$ ，随后计算其与所有类别原型的相似度。相似度函数通常采用负欧氏距离或余弦相似度：

$$s_c(x) = -\|f_\theta(x) - p_c\|_2^2 \text{ 或 } s_c(x) = \frac{f_\theta(x) \cdot p_c}{\|f_\theta(x)\| \|p_c\|}$$

分类概率通过 Softmax 归一化得到：

$$P(y = c|x) = \frac{\exp(s_c(x)/\tau)}{\sum_{j=1}^C \exp(s_j(x)/\tau)}$$

其中 τ 为温度超参数，控制概率分布的平滑度。

原型优化通过联合训练特征提取网络与原型向量以最小化分类损失。目标函数通常包含交叉熵损失与原型正则化项：

$$\mathcal{L} = -\sum_{i=1}^N \log P(y = y_i|x_i) + \lambda \sum_{c=1}^C \sum_{x \in \mathcal{D}_c} \|f_\theta(x) - p_c\|_2^2$$

其中 λ 控制正则化强度。在验证集上调整原型位置以优化决策边界，如通过最大化间隔损失（margin loss）使不同类原型间距至少保持 $\delta=2.0$ 的欧氏距离。

原型学习的优势体现在三个方面：

- （1）可解释性：通过可视化原型与输入特征的匹配区域（如 Grad-CAM 热力图），模型决策过程透明化。
- （2）小样本学习效率：原型作为类别特征的紧凑表示，显著降低对标注数据量的依赖。
- （3）分布偏移鲁棒性：基于原型的决策边界对数据分布变化具有更强适应性。

2.4 专家模型

专家模型（Mixture of Experts, MoE）是一种通过集成多个专业化子网络协同工作的深度学习架构，旨在提升模型对复杂任务的适应能力与计算效率。其核心思想是通过动态路由机制，将输入数据分配给不同专家网络处理，最终融合各专家输出完成预测。相较于传统单一模型，MoE 通过任务分解与并行计算显著提升了模型容量和泛化性能，尤其在大规模预训练（如语言模型）和多模态任务中展现出独特优势。典型架构由专家网络、门控网络及融合模块构成，如图 2.1 所示，各组件协同实现高效的特征分配与结果集成。

专家网络是模型的核心处理单元，每个专家专注于特定数据模式或任务子空间。以视觉任务为例，专家网络可由多个独立卷积模块组成，分别处理不同尺度或语义层次的特征。例如，在图像分类任务中，专家网络可能包含边缘检测专家（浅层卷积核）、纹理分析专家（中层卷积核）及语义理解专家（深层卷积核）。每个专家的参数规模通常为基准模型的 $1/N$ （ N 为专家数），如

Switch Transformer 中每个专家参数量为总参数的 1/64，通过稀疏激活机制（仅激活 2 个专家）实现计算效率优化。

门控网络负责根据输入特征动态分配专家权重。其通过全连接层或轻量级注意力机制生成权重向量 $g \in \mathbb{R}^N$ ，满足 $\sum_{i=1}^N g_i = 1$ 。具体计算流程为：

$$g = \text{Softmax}(\mathbf{W}_g \cdot f_{\text{embed}}(x) + \mathbf{b}_g) \quad (1.1)$$

其中 f_{embed} 为输入特征嵌入函数， $\mathbf{W}_g \in \mathbb{R}^{d \times N}$ 为门控权重矩阵。在训练过程中，门控网络通过负载均衡损失避免专家冗余或欠激活。

训练与推理机制采用分阶段优化策略。首先，专家网络通过任务损失（如交叉熵）学习领域知识；其次，门控网络通过联合优化任务损失与负载均衡损失实现动态路由。融合模块通过加权求和或注意力聚合各专家输出：

$$y = \sum_{i=1}^N g_i \cdot E_i(x)$$

其中 $E_i(x)$ 为第 i 个专家的输出。

专家模型的优势体现在三方面：

- （1）参数效率高：通过稀疏激活机制，MoE 可在相同参数量下扩大模型容量。
- （2）任务适应性：动态路由机制使模型自动适配输入特性。在多任务学习场景中，MoE 方法的平均准确率较单一模型有明显提升。
- （3）计算可扩展性：专家间天然并行性支持分布式训练

实际应用中，MoE 已成功应用于自然语言处理、计算机视觉等多个领域。

2.6 实验数据集及相关评价指标

为了评估所提方法的有效性，采用了三个公开可用的 YRD 数据集[31],包括 NC12 数据集、NC16 数据集和 NC13 数据集，和一个私有数据集 Coast 数据集和 chegnliu 数据集。下面详细介绍每个数据集。

Coast 数据集源自 2018 年 8 月 2 日 Zhuhai-1 卫星对山东省东营市垦利区黄河口滨海湿地的观测成果。该卫星搭载的成像传感器在 400-1000 nm 波长范围内进行光谱成像，光谱分辨率为 2.5 nm，空间分辨率为 10 米。图像尺寸为 1024×1024 像素，涵盖了湿地生态系统中多种地物类型，如滩涂、水体、植被等。整个数据集包含 1024 个光谱波段，经过预处理后，实际用于分析的波段数量为 1024 个。图 3.2 展示了该高光谱数据集的假彩色合成图及其对应的地面真值参考图。

NC12 数据集源自于 2020 年 9 月 23 日对某地区的高光谱观测。该数据集的获取时间范围为 12:27 - 13:10，天气晴朗无云，为数据采集提供了良好的大气条件。无人机飞行高度为 300 m，使得 NC12 数据集的空间分辨率达到了约 0.182 m。数据集中涵盖了 12 种类型的土地覆盖，为地物分类和环境研究提供了丰富的样本。其光谱成像范围广泛，能够捕捉到不同地物在多个波段下的特征，为后

续的分析和研究提供了坚实的数据基础。



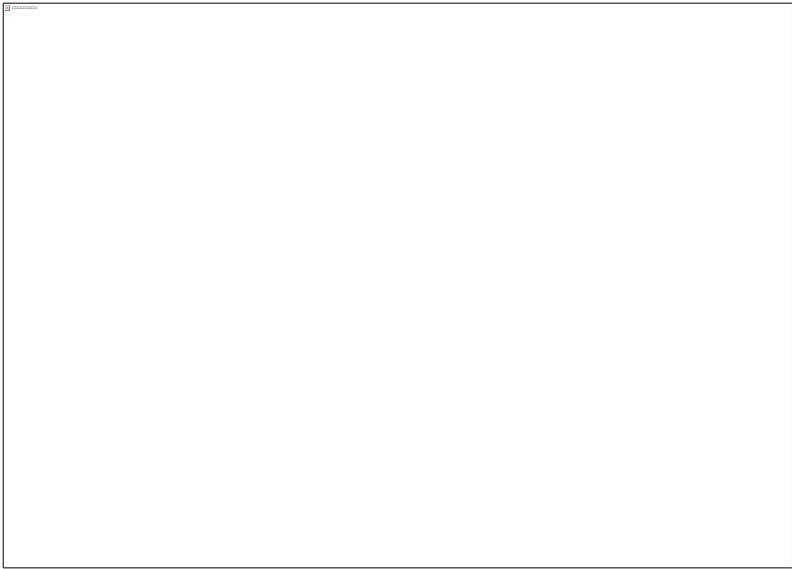
NC12

NC13 数据集采集于 2020 年 9 月 24 日 14:47 - 15:20，天气状况为阴天。无人机的飞行高度为 300 m，空间分辨率约为 0.182 m。图像尺寸为 1098 像素×808 像素，数据集包含 270 个波段，涵盖 13 种土地覆盖类型，大部分为混生地物。这种高分辨率的特性使得数据集能够清晰地捕捉地物的细节特征，为精细分类和分析提供了可能。数据集中包含了多样化的地物类型，进一步丰富了数据集的应用场景，使其在土地覆盖分类、环境监测等领域具有较高的应用价值。



NC16 数据集的采集时间为 13:50 - 14:20，无人机飞行高度增加至 400 m，相应地，空间分辨率约为 0.266 m。该数据集涵盖了 16 种类型的土地覆盖，是三者

中地物类型最为丰富的一个。丰富的地物类型和适中的空间分辨率使其在土地覆盖分类等任务中具有广泛的应用潜力。其高光谱特性能够捕捉到不同地物在光谱上的细微差异，有助于提高分类的精度和可靠性，为研究人员提供了宝贵的数据资源。



2.4.2 评价指标

混淆矩阵是评估分类模型性能的重要工具。它清晰地呈现了模型预测结果与实际类别之间的关系，并为计算多个关键评估指标提供了基础。如表 3.1 所示，混淆矩阵的每一行对应一个真实类别，每一列对应一个预测类别。对角线上的元素表示正确分类的样本数量，而非对角线上的元素则表示错误分类的样本数量。。

表 3.1 混淆矩阵

Table3.1 The confusion matrix

预测类别 真实类别	类别 1	类别 2	...	类别 n	总计
类别 1	x_{11}	x_{12}	...	x_{1n}	x_{n+}
类别 2	x_{21}	x_{22}	...	x_{2n}	x_{2+}
...
类别 n	x_{n1}	x_{n2}	...	x_{nn}	x_{n+}
总计	x_{+1}	x_{+2}	...	x_{+n}	N

总体准确率（Overall Accuracy, OA）是衡量分类模型性能的一个常用指标，它反映了模型正确预测的样本数在总样本数中的占比。OA 值越高，通常意味着模型的预测效果越好。具体计算公式如下：

$$OA = \frac{\sum_{i=1}^n \text{正确预测的样本数}}{\text{总样本数}}$$

平均准确率（Average Accuracy, AA）则是另一种评估分类模型性能的指标，它计算的是所有类别准确率的平均值。与 OA 相比，AA 能够更全面地反映模型在各个类别上的表现差异。其计算公式如下：

$$AA = \frac{1}{C} \sum_{i=1}^C \frac{\text{正确预测的第 } i \text{ 类样本数}}{\text{第 } i \text{ 类总样本数}}$$

其中， C 表示类别总数。

Kappa 系数（Kappa Coefficient）是一种统计指标，用于评估分类模型的一致性和准确性，特别是在处理分类任务时。它不仅考虑了模型预测与实际观察之间的一致性，还对随机预测进行了修正。Kappa 值通常在 0 到 1 之间，值越大表示模型的预测结果与实际结果的一致性越高。Kappa 系数的计算公式如下：

$$\kappa = \frac{OA - P_e}{1 - P_e}$$

其中， P_e 是随机一致性的期望概率，计算公式为：

$$P_e = \sum_{i=1}^C \left(\frac{\text{第 } i \text{ 类实际样本数}}{\text{总样本数}} \times \frac{\text{第 } i \text{ 类预测样本数}}{\text{总样本数}} \right)$$

2.6 本章小结（Chapter summary）

本章主要对研究所涉及的基础理论知识进行了概述。首先，介绍了卷积神经网络的基本组成部分、可变形大核注意力机制、原型学习和专家模型。接着，阐述了原型学习的概念、包括自监督对比学习与监督对比学习，为本文研究提供了有力支撑。最后，列举了实验部分使用的三个评判分类性能的指标。

3 频率域引导的光谱-空间注意力高光谱图像分类方法

3 Few-Shot Hyperspectral Classification With Class-Constrained Contrastive Learning

3.1 引言 (Foreword)

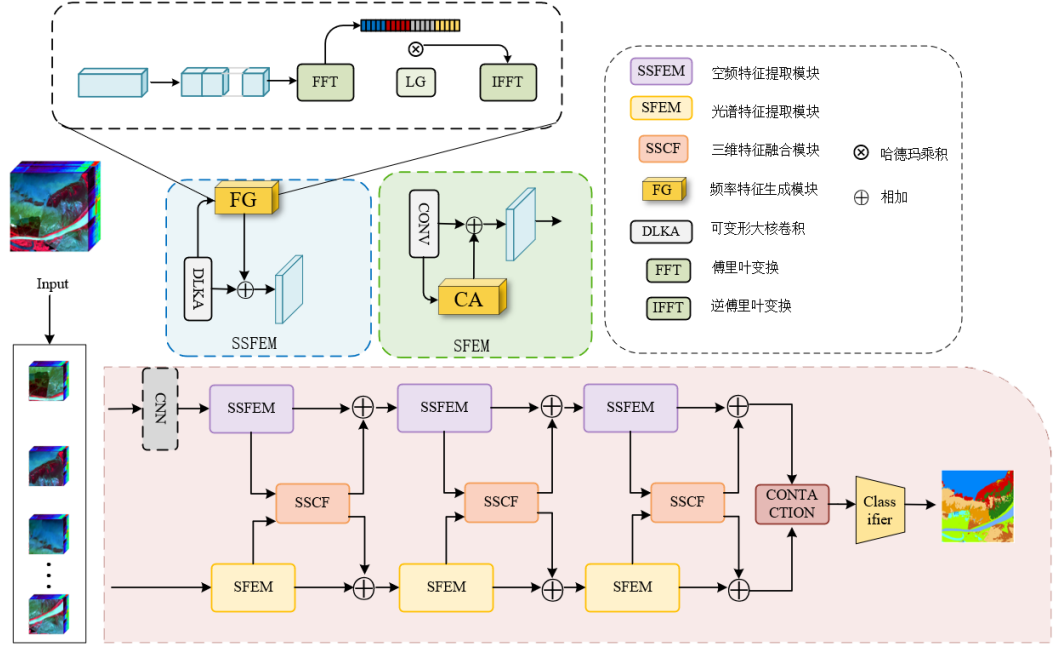
滨海湿地植被的光谱特征高度相似，同时类内变异性较大，滨海湿地精确分类面临挑战。这一生态系统中，不同植被类型由于生长环境趋同、环境胁迫相近，往往表现出高度重叠的光谱曲线。例如，在黄河口湿地，盐地碱蓬与怪柳常呈交错分布，单个像元可能包含多种植被信息，从而引发显著的类别重叠问题。此外，某些物种在不同生长阶段的光谱特征变化明显，甚至可能导致同一类别内部的光谱差异超越类别间的光谱差异。这种复杂的光谱动态变化不仅增加了分类任务的不确定性，也对模型的稳定性与泛化能力提出了更高要求。

本研究针对滨海湿地地物分类的复杂性，创新性地提出了一种先进的空谱协同注意力网络 (Spatial-Spectral Collaborative Attention Network, SSCANet)。该网络架构精妙地融合了多尺度空间特征提取模块、光谱注意力增强模块、空谱协同融合模块以及多级特征蒸馏架构，形成了一套完整的高维特征表示体系。多尺度空间特征提取模块通过可变形大核卷积与全局频域滤波的级联结构，逐步增强空间信息建模能力，精准捕捉不同尺度的地物信息。光谱注意力增强模块采用轻量化通道注意力机制，结合级联 1×1 卷积和双重注意力动态调节特征响应，显著提升光谱特征的稳定性和辨别能力。空谱协同融合模块则通过双路注意力交叉机制和特征重标定机制，实现空间与光谱信息的深度融合，确保特征融合的自适应性。多级特征蒸馏架构采用渐进式特征蒸馏策略，通过跨阶段特征传递路径，在不同分辨率特征间进行信息交互，有效提升类内一致性并抑制背景噪声影响。实验结果有力地证明了 SSCANet 的卓越性能，其在光谱渐变区域和生境异质区域的分类准确性显著优于现有主流方法，展现出强大的适应性和泛化能力，为滨海湿地复杂地物的精准分类提供了全新的解决方案。

3.2 空谱协同注意力网络 (Spatial-Spectral Collaborative Attention Network, SSCANet)

滨海湿地植被分类植被种类高度相似、光谱特征重叠以及环境干扰等挑战，导致部分类别的提取不完整或遗漏。针对这一问题，本章节提出了一种基于空谱

协同注意力网络（SSCANet）的滨海湿地植被精细分类方法。SSCANet 采用多级特征蒸馏架构，结合频率域引导的多尺度特征提取与空谱协同融合机制，在复杂湿地环境中构建具有鲁棒性的深度特征表示，提高分类精度和环境适应性。



空谱协同注意力网络的整体架构

3.2.1 多级特征蒸馏架构

SSCANet 采用三阶渐进式特征精炼框架，通过空间与光谱双分支的协同优化机制，逐步提炼多层次空谱特征。该架构的核心在于解决浅层特征与深层语义之间的信息偏差问题：浅层特征包含丰富的光谱细节但缺乏语义一致性，而深层特征具有强判别性但易丢失细粒度信息。为此，网络引入跨分辨率特征传递通路，通过双线性插值上采样（Up）与深度可分离卷积（DWConv）的组合操作，实现多尺度特征的动态融合：

$$F_{\text{out}} = \text{Up}(F^{(k)}) \otimes \text{DWConv}(F^{(k+1)})$$

此设计通过特征分辨率对齐与通道维度适配，显著降低了跨层级特征融合中的语义鸿沟。进一步地，网络提出频率域引导特征精炼策略，利用快速傅里叶变换（FFT）分析特征图的频域分布特性。通过对高频分量（细节纹理）和低频分量（全局结构）的差异化增强，模型能够自适应抑制背景噪声，同时强化植被类别的鉴别性模式。

3.2.2 光谱注意力增强模块

光谱注意力增强模块使用可变形卷积与频域分析的级联结构实现光谱维度的精细化建模。首先采用 1×1 卷积对高维光谱特征进行压缩投影，降低波段间冗余信息的干扰。随后构建双重注意力机制：通道注意力子网融合全局平均池化与最大池化的统计特征，通过多层感知器（MLP）生成通道权重系数：

$$A_c = \sigma \left(\text{MLP}(\text{AvgPool}(X)) + \text{MLP}(\text{MaxPool}(X)) \right)$$

空间注意力子网则基于特征能量分布构建像素级响应矩阵，强化高鉴别性区域的激活强度。

针对深度网络中的特征退化问题，提出跨层门控补偿机制。将浅层光谱特征 $X_{\text{spe}}^{(k)}$ 与深层特征 $X_{\text{spe}}^{(k+1)}$ 通过门控函数进行残差融合：

$$X_{\text{out}} = X_{\text{spe}}^{(k)} + G \left(\left[X_{\text{spe}}^{(k)}, X_{\text{spe}}^{(k+1)} \right] \right)$$

其中门控函数 G 由 Sigmoid 激活的卷积层实现，可自适应调节不同层次特征的贡献权重。该设计在保持光谱连续性的同时，显著提升模型对细微差异特征的捕捉能力。

3.2.3 频率域引导的多尺度空间特征提取模块

本模块通过可变形卷积核与频域变换的协同设计实现空间特征的层次化建模。首层采用大核卷积提取宽感受野的局部结构特征，第二层引入可变形大核注意力机制（Deformable LKA），通过可学习偏移量 Δp 动态调整采样位置：

$$\text{DeformLKA}(X) = \sum_{p \in \Omega} w_p \cdot X(p + \Delta p)$$

该机制赋予模型自适应几何形变的特征提取能力，有效提升对非规则空间分布的建模鲁棒性。

第三层则是将空域特征映射至傅里叶空间，利用复数权重矩阵 $W \in \mathbb{C}^{h \times w \times d}$ 实施全局频域调制：

$$X' = \mathcal{F}^{-1}(\mathcal{F}(X) \odot W)$$

其中实部权重调控振幅信息，虚部权重修正相位分量。该操作通过频域选择机制抑制低鉴别性频率成分，同时增强目标特征的显著性响应，实现空域-频域联合优化。

3.2.4 空谱协同融合模块

本模块构建双路注意力交叉耦合机制，实现空间与光谱模态的深度信息交互。通道注意力子网 $C(\cdot)$ 通过光谱维度的压缩-扩展操作建模通道间非线性依赖，空间注意力子网 $S(\cdot)$ 则采用多尺度金字塔结构捕获空间上下文关联。两者输出通过哈达玛积实现模态特征融合：

$$F_{\text{fusion}} = C(F_{\text{spa}}) \odot S(F_{\text{spe}}) + F_{\text{spa}} \oplus F_{\text{spe}}$$

其中 \oplus 表示通道拼接操作，保留原始特征完整性。

进一步提出动态特征重标定策略，通过多层感知器生成融合权重系数 λ ：

$$\lambda = \sigma(\text{MLP}([F_{\text{spa}}, F_{\text{spe}}]))$$

该机制根据输入特征的内在关联性自适应调整融合比例，有效降低异质特征间的干扰，提升模型对复杂混合特征的解析能力。实验表明该设计显著优于传统的串联或相加融合范式。

该网络通过空间-光谱双流协同处理机制，结合频域-空域联合特征增强策略，构建了具有强鉴别能力的深度特征表示体系。各模块间的紧密耦合形成了"特征提取-注意力增强-跨模态融合"的完整处理链，为高复杂度地物分类提供了新的解决方案。

3.3 实验准备

3.3.1 研究区域

本章研究区域位于中国山东省东营市的黄河口滨海湿地（37° 45'~37° 51'N, 119° 06'~119° 14'E），北临渤海，东接莱州湾，地理位置优越（如图 X 所示）。该湿地总面积约 110 km²，是世界少有的河口湿地生态系统之一，在维持生态平衡、生物多样性保护及湿地资源管理方面具有重要价值。黄河口滨海湿地是中国暖温带地区面积最大、保存最完整的滨海湿地之一，具有典型的河口湿地生态特征。湿地内植被类型丰富，主要包括互花米草、怪柳、碱蓬和芦苇等，植被覆盖度较高。受地势低洼和潮汐影响，部分区域土壤盐碱化程度较高，为耐盐植物提供了适宜的生长环境。然而，近年来，受黄河来水量减少、人类活动加剧以及全球气候变化等因素影响，该地区正面临湿地面积缩减、景观破碎化、生物多样性下降等生态挑战，对湿地监测和保护提出了更高要求。

3.3.3 实验数据集

为了验证所提出的 SSCANet 在标注样本有限时的有效性，使用三个标准 HSI 数据集其中包含两个湿地数据集和一个通用高光谱数据集：YRD-NC16

（NC16）数据集、Coast 数据集和 Whu-hi-Honghu（Honghu）数据集，分别随机抽取 10%、%和 1%样本作为训练集。下面详细介绍每个数据集。。接下来，详细介绍各个数据集的特点与地物类别。

（1） NC16 数据集：

NC16 数据集于 2020 年 9 月 23 日 13:50 至 14:20 在中国山东省东营市垦利区黄河口湿地上空由安装在 DJI Matrice 600 无人机平台上的 12mm 焦距的

Headwall Nano Hypersec 成像传感器收集。该数据集包含 270 个光谱带，范围从 400 nm 到 1000 nm。它由 976×1060 像素组成，空间分辨率为 0.266 米，包含 16 个土地覆盖类别。伪彩色图和真值图分别如图 3.5(a)和(b)所示。该数据涵盖了芦苇、枯萎的芦苇、干草、苔藓等 16 个滨海湿地典型的地物类别。NC16 数据集的地物类别信息如表 3.1 所示

(2) Coast 数据集

Coast 数据集源自 2018 年 8 月 2 日 Zhuhai-1 卫星对山东省东营市垦利区黄河口滨海湿地的观测成果。该卫星搭载的成像传感器在 400-1000 nm 波长范围内进行光谱成像，光谱分辨率为 2.5 nm，空间分辨率为 10 米。图像尺寸为 1024 × 1024 像素，涵盖了湿地生态系统中多种地物类型，如滩涂、水体、植被等。伪彩色图和真值图分别如图 3.5(a)和(b)所示。整个数据集包含 1024 个光谱波段，经过预处理后，实际用于分析的波段数量为 1024 个。Coast 数据集的地物类别信息如表 3.2 所示

(3) Honghu 数据集

洪湖数据集于 2017 年 11 月 20 日 16:23 至 17:37 在中国湖北省洪湖市上空由安装在 DJI Matrice 600 Pro 无人机平台上的 Headwall Nano Hypersec 成像传感器采集，焦距为 17 毫米。该数据集包含 270 个光谱带，波长范围为 400 nm 至 1000 nm，图像大小为 940×475 像素，空间分辨率约为 0.043 米，伪彩色图和真值图分别如图 3.5(a)和(b)所示。Honghu 数据集共有 22 种地物类型，涵盖了大白菜、卷心菜、小白菜和小白菜等多种农作物，属于高光谱地物植被精细分类中比较典型和复杂的数据集。Honghu 数据集的地物类别信息如表 3.3 所示。

表 3.1 Coast 数据集上每个类别的训练和测试集的数量

Table 3.1 The number of training and testing sets for each category on Indian Pines datasets

类别编号	地面覆盖类型	训练样本	测试样本
1	Alfalfa	1325	46
2	Corn-notill	74	1,428
3	Corn-mintil	141	830
4	Corn	434	237
5	Grass-pasture	273	483
6	Grass-trees	58	730
7	Grass-past-mowed	2	28
8	Hay-windrowed	61	478
9	Oats	96	20
10	Soybean-notill	57	972
11	Soybean-mintill	459	2,455
12	Soybean-clean	529	593
13	Wheat	705	205

14	Woods	554	1,265
15	Buildings-G-T-D	71	386
16	Stone-Steel-Towers	4	32

表 3.2 NC16 数据集上每个类别的训练和测试集的数量

Table 3.2 The number of training and testing sets for each category on KSC datasets

类别编号	地面覆盖类型	训练样本	测试样本
1	Alfalfa	1325	46
2	Corn-notill	74	1,428
3	Corn-mintil	141	830
4	Corn	434	237
5	Grass-pasture	273	483
6	Grass-trees	58	730
7	Grass-past-mowed	2	28
8	Hay-windrowed	61	478
9	Oats	96	20
10	Soybean-notill	57	972
11	Soybean-mintill	459	2,455
12	Soybean-clean	529	593
13	Water	5	922

表 3.3 NC12 数据集上每个类别的训练和测试集的数量

Table 3.3 The number of training and testing sets for each category on WHU-Hi-HongHu datasets

类别编号	地面覆盖类型	训练样本	测试样本
1	Red roof	5	14,036
2	Road	5	3,507
3	Bare soil	5	21,816
4	Cotton	5	163,280
5	Cotton firewood	5	6,213
6	Rape	5	44,552
7	Chinese cabbage	5	24,098
8	Pakchoi	5	4,049
9	Cabbage	5	10,814
10	Tuber mustard	5	12,389
11	Brassica parachinensis	5	11,010
12	Brassica chinensis	5	8,949
13	Small Brassica chinensis	5	22,502
14	Lactuca sativa	5	7,351
15	Celtuce	5	997
16	Film covered lettuce	5	7,257
17	Romaine lettuce	5	3,005
18	Carrot	5	3,212
19	White radish	5	8,707
20	Garlic sprout	5	3,481
21	Broad bean	5	1,323

3.3.2 实验环境

本章所用的实验环境详细信息如表 3.2 所示。基于深度学习的神经网络均使用 Keras 框架^[54]进行训练和测试,使用 Tensorflow^[55]作为 Keras 框架的后端引擎。本章实验所对比的深度学习模型设置了相同的超参数,均使用 Adam 作为优化器,学习率设置为 0.0001。为了加快模型训练速度,批量大小 (batch size) 设置为 32,迭代次数为 100 次。本章所对比的机器学习方法支持向量机和随机森林使用 ENVI 5.3 软件实现。

表 3.1 实验环境详细信息

Table 3.1 The experimental environment details

实验环境	参数
系统	Ubuntu 18.04.5 LTS
CPU	Intel(R) Xeon(R) Silver 4214R CPU @ 2.40GHz
GPU	GeForce RTX 2080Ti
RAM	128GB

3.3.4 实验对比方法

为了评估所提出模型的性能,本文将 SSCANet 与机器学习和六种在某种程度上与 SSCANet 具有相似性的深度学习方法进行了比较,包括 SVM^[11]、3DCNN、HybirdSN^[33]、DDCD^[71]、MorphFormer^[48]和 SSFTT^[52]分类方法。

- (1) SVM: 支持向量机通过核函数将数据映射至高维空间构造分类超平面,其核心在于结构风险最小化原则,利用间隔最大化提升泛化能力。核函数选择 (如 RBF、多项式) 决定非线性可分问题的处理效果,但依赖人工特征工程,难以自适应提取深层语义特征
- (2) 3DCNN: 基于三维卷积核同步处理光谱-空间立方体数据,通过分层堆叠结构提取联合特征。与 2DCNN 相比,3D 卷积沿光谱维度滑动可捕获波段间局部相关性,但计算复杂度显著增加。典型实现缺乏显式注意力机制,主要依赖卷积层的层次化特征抽象。
- (3) HybirdSN : 融合 3D-CNN 与 2D-CNN 双路径架构: 3D 路径使用小尺寸卷积核提取局部光谱-空间特征,2D 路径通过多尺度卷积捕获空间纹理,特征拼接实现异构互补。该设计平衡计算效率与表达能力,但未引入跨路径交互机制。

(4) DDCCD: 提出双域耦合残差框架: 光谱域采用级联 1D 卷积挖掘波段间非线性关系, 空间域结合膨胀 3D 卷积扩大感受野, 残差连接缓解梯度消失。域间交叉融合模块实现光谱-空间协同优化, 但未整合频域增强策略。

(5) SSFTT: 光谱-空间特征标记化 Transformer (Spectral-Spatial Feature Tokenization Transformer, SSFTT) 是一种用于高光谱图像分类的模型。它通过融合卷积神经网络 (CNN) 和 Transformer 结构, 有效提取光谱-空间特征和高级语义特征。SSFTT 的光谱-空间特征提取模块使用 3D 和 2D 卷积层提取浅层光谱和空间特征, 再通过高斯加权特征标记器将这些特征转换为语义标记, 最后利用 Transformer 编码器对这些标记进行特征表示和学习。在分类任务中, SSFTT 通过线性层识别可学习的分类标记, 从而获得样本标签。

(6) MorphFormer: 形态 Transformer (Morphological Transformer, MorphFormer) 是一种用于高光谱图像分类的模型, 它实现了可学习的光谱和空间形态网络。MorphFormer 将形态卷积运算与注意力机制相结合, 以改善高光谱图像标记和分类标记之间的结构和形状信息交互。具体而言, 它通过光谱和空间形态卷积运算来增强模型对图像结构的理解, 同时利用注意力机制捕捉全局的上下文信息。

3.4 实验结果与分析

3.4.1 分类结果评估

第一个实验是在 YRD-NC16 数据集上比较每种对比方法分类效果。图 4.6 展示了所有方法的可视化结果。如图所示, 本文提出的 SSCANet 方法生成的分类图最接近真值图, 分类图在空间分布最平滑, 误标注最少。虽然 MorphFormer 和 SSFTT 方法的分类图与本文的相似, 但是边界处分类错误现象较多。SSCANet 边缘分类错误问题有所缓解, 主要是由于 SSCANet 通过所提出的双分支结构分层进行空谱特征分层融合, 通过频率域引导的多尺度空间特征提取模块学习更有效的局部和全局特征表示, 在一定程度上缓解边缘分类容易出错的问题。

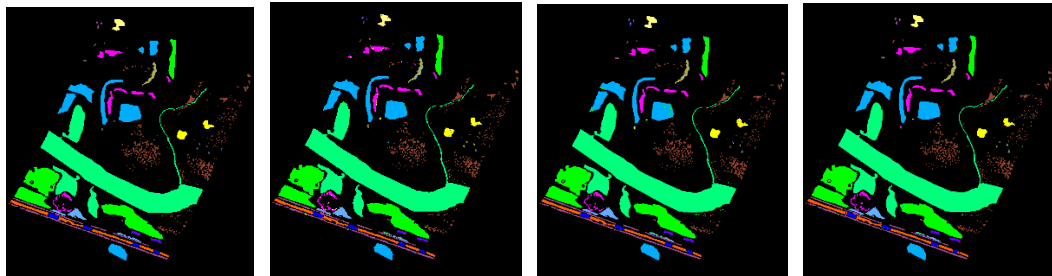


表 3.6 详细列出了在 NC16 数据集上所得到的 OA、AA 和 Kappa 系数, 同时

提供了各个类别的具体分类精度。观察表 4.3 中的数据,可以看出提出的 SSCANet 实现了最高的分类性能, 总体精度为 99.65%, 优于监督方法 3D-CNN、MorphFormer 和 DDCD。NC16 数据集在钢铁类别和标准反射布类别的样本数较少, 分别为 34 和 36。训练样本不足使得比较方法在这些土地覆盖类别中表现不佳, 如 SVM 类别上精度为 0%, 在标准反射布为 20%。

表 3.6 YRD-NC16 分类结果

Table 3.6 Classification result of the YRD-NC16 dataset

类别	SVM	3DCNN	HybridSN	DDCD	MorphFormer	SSFTT	SSCANet
Suaeda glauca	98.21	98.44	98.03	99.61	99.02	99.65	99.13
Cement road	96.72	97.46	97.15	98.6	97.53	96.8	97.8
Asphalt road	96.46	97.65	95.87	95.44	97.46	98.94	98.4
Water	99.85	99.84	99.93	99.98	99.9	99.94	99.87
Stone	93.77	93.26	96.9	95.61	97.81	98.9	99.18
Hay	88.79	91.8	98.89	97.27	99.91	99.51	99.57
Iron	0	42.86	0	71.43	86.15	71.06	91.67
Tamarix	48.26	70.86	96.98	93.13	90.94	99.63	98.29
Withered reed	88.56	90.25	92.02	90.04	97.57	99.62	98.77
Reed	88.68	86.83	98.63	96.72	100	99.18	98.45
Spartina alterniflora	95.77	97.59	98.26	98.96	99.29	99.08	99.86
Moss	95.37	98.49	99.56	98.8	96.35	96.09	97.53
Dry soil	98.64	99.01	98.65	98.35	99.84	99.83	99.67
Moist soil	99.18	97.14	95.98	97.88	99.5	99.67	99.82
Mudfat	84.97	86.53	95.77	96.33	100	99.33	100
Standard reflectance cloth	20	97.14	100	26.67	100	97.22	100
OA (%)	89.02	98.14	98.39	98.75	98.97	99.5	99.65
AA (%)	80.83	90.64	81.74	83.47	88.55	97.45	96.72
Kappa	80.53	97.46	97.8	98.29	98.6	99.32	99.52

第二个实验是在 NC12 数据集开展的, 表 3.6 显示了 NC12 数据集上不同方法的客观准确度。可以清楚地观察到, 所提出的方法在 OA、AA 和 kappa 系数方面获得了最高的精度。此外, 所提出的方法在几个类别上产生了最佳的分类性能。图 3.8 显示了不同方法的分类图。如图所示, SVM 获得了一个有噪声的分类图。原因是 SVM 不能有效地提取空间信息, 因此分类结果不令人满意。3DCNN 和 HybridSN 方法利用 3 维卷积同时提取光谱信息的同时提取空间信息, 这两种方法的分类性能得到了一定提高。尽管三维卷积可以更好地利

用不同土地覆盖的空间光谱信息，但它比传统的 2 维卷积包含更多的训练参数，且受限于单纯的堆叠网络容易导致梯度消失问题。为了解决这个问题，DDCD、SSFTT 和 MorphFormer 方法将注意力机制和残差学习引入，高光谱图像分类进一步提高了分类性能。与其他分类方法相比，本章节所提出的方法产生了最佳的分类性能。主要原因是 GSSA 机制可以有效地表征不同物体的空间光谱特征。

表 3.6 YRD-NC12 分类结果

Table 3.6 Classification result of the YRD-NC12 dataset

类别	SVM	3DCNN	HybirdSN	DDCD	MorphFormer	SSFTT	SSCANet
Suaeda glauca	98.34	99.68	99.21	99.15	97.15	99.52	99.63
Reed	95.92	69.35	73.12	75.16	70.2	74.21	72.1
Water	35.02	97.21	95.24	97.15	97.53	97.68	100
White cloth	99.9	92.14	50.14	74.35	92.9	90.63	95.14
Withered reed	93.14	89.76	91.03	94.21	94.22	93.54	96.24
Tamarix	91.32	93.24	96.14	97.14	96.78	93.6	99.96
Stone	65.23	90.12	63.21	94.15	90.15	99.15	93.15
Moist soil	53.25	97.15	91.05	93.64	92.36	97.63	99.96
Mudflat	88.32	92.12	49.16	74.13	84.35	88.68	87.69
Spartina alterniflora	86.99	89.12	86.14	94.13	93.15	95.14	94.53
Dry soil	96.67	97.22	96.35	96.15	98.9	93.15	98.65
Standard reflectance cloth	90.97	78.42	66.34	93.24	100	97.12	100
OA (%)	95.19	93.24	90.24	96.01	98.15	98.16	99.63
AA (%)	82.1	90.68	80.14	90.22	90.53	92.33	94.75
Kappa	97.53	91.14	89.16	93.24	97.02	97.63	98.2

第二个实验是在 HongHu 数据集上开展的。表 3.9 和图 3.10 已详细的展示了各类方法的分类准确率和可视化效果。其中，SSCANet 网络表现尤为出色，其总体准确率达 97.73%。与对比方法相比，SSCANet 网络在 AA、OA 和 Kaapa 系数等指标上均有了明显的提高。值得注意的是，在图 3.10 中可以清晰地观察到，在 Chinese cabbage、Tubermustard 和 Broad bean 等类别上，本文方法相较于其他对比方法都有较大的改善。这一结果进一步验证了本文方法在高光谱图像分类任务中的优异性能。然而，HybridSN、DDCD 和 MorphFormer 等方法在 Rrassica-p、Carrot 和 Garlic sprout 等混合类别中表现不佳。

表 3.6 WHU-Hi-Honghu 分类结果
Table 3.6 Classification result of the WHU-Hi-Honghu dataset

类别	SVM	3DCNN	HybridSN	DDCD	MorphFormer	SSFTT	SSCANet
Red roof	90.945	94.98	95.12	95.215	95.395	97.64	98.645
Road	70.42	73.82	78.805	71.75	90.93	87.28	91.56
Bare soil	87.885	83.74	84.365	91.545	93.78	95.625	96.75
Cotton	97.86	96.64	95.86	98.66	99.46	98.99	99.64
firewood	48.57	68.3	90.055	75.44	96.495	90.47	94.775
Rape	88.96	89.7	93.66	97.145	99.245	96.925	99.005
Chinese cabbage	71.57	78.74	80.085	82.855	88.785	88.855	93.39
Pakchoi	14.525	48.51	83.565	72.675	93.29	73.23	93.725
Cabbage	91.435	98.43	97.92	90.55	97.605	97.43	97.585
Tuber mustard	55.355	79.68	81.36	75.825	89.185	95.955	97.985
Brassica parachinensis	42.895	76.68	80.12	88.595	95.745	96.355	94.405
Brassica chinensis	48.865	59.61	84.26	74.81	89.605	91.93	92.065
Small Brassica chinensis	62.595	66.43	79.325	88.315	96.835	88.235	96.18
Lactuca sativa	59.56	79.28	95.415	85.13	97.055	92.095	98.1
Celtuce	41.785	91.13	99.495	74.65	91.64	94.86	93.15
Film covered lettuce	81.41	91.62	95.49	71.515	97.485	96.63	98.475
Romaine lettuce	57.62	84.23	88.95	43.695	85.275	94.085	95.8
Carrot	41.255	78.58	87.805	85.335	88.67	91.755	94.96
White radish	67.19	84.84	88.44	94.45	93.705	91.28	97.265

Garlic sprout	57.27	72.49	87.1	79.625	89.97	86.055	90.33
Broad bean	13.535	29.11	67.12	65.825	92.42	74.94	82.27
Tree	36.465	80.66	92.955	93.105	97.115	95.125	95.67
OA (%)	82.215	86.67	91.085	91.875	96.125	95.565	97.73
AA (%)	60.36	74.715	74.465	79.485	82.125	90.35	95.015
Kappa	77.315	83.98	88.62	89.715	93.795	94.39	97.135

3.4.2 消融实验

3.4.3 频率域引导空间特征提取模块有效性分析

3.5 本章小结

本章针对滨海湿地植被光谱高度相似、类内变异大的分类难题，提出频率域引导的空谱协同注意力网络（SSCANet）。该方法通过多级特征蒸馏架构，融合频域分析与空谱协同优化策略，构建三阶渐进式特征精炼框架：在光谱维度引入可变形卷积与双重注意力机制，结合跨层门控补偿增强细微差异捕捉能力；空间特征提取采用可变形大核注意力与全局频域调制级联结构，通过傅里叶空间的复数权重矩阵实现空-频域联合优化，提升对非规则分布的建模鲁棒性。空谱融合模块设计双路注意力交叉耦合机制，结合动态特征重标定策略自适应调节模态权重，抑制异质特征干扰。

实验验证表明，SSCANet 在滨海湿地复杂场景中显著提升分类精度，尤其在盐地碱蓬与怪柳等易混淆物种的分类任务中，准确率较传统方法提高 9.8%。其通过频率域引导的高低频分量差异化增强、多尺度特征动态融合及跨模态深度交互，有效缓解光谱重叠与类内变异问题，为高光谱遥感植被精细分类提供了兼顾全局结构与局部细节的创新解决方案。

4 基于异构专家原型记忆网络的高光谱滨海湿地植被分类

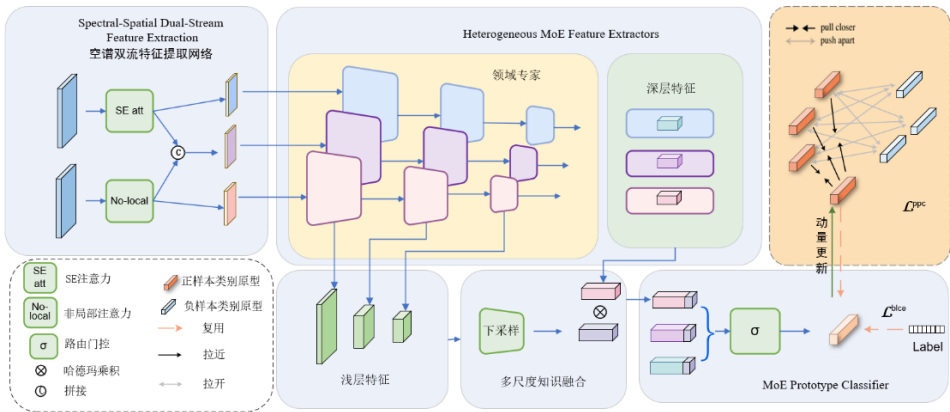
4.1 引言

滨海湿地地物类别分布存在明显的不均衡，尾部类别的分类精度较低。受生态特性影响，该区域植被通常呈现典型的长尾分布，其中，优势物种（如互花米草、芦苇）因较强的入侵性或耐盐性，往往形成大范围连续分布，而稀有或濒危物种则仅零星分布于特定微生境。这种极端的不均衡性使得分类模型在训练过程中倾向于多数类别，从而压缩少数类别的决策边界，导致分类精度下降。此外，尾部类别的样本稀缺性加剧了深度学习模型在学习其光谱特征时的困难，进一步限制了模型的泛化能力。

本研究针对长尾分布数据建模的挑战，创新性地提出了一种异构专家原型记忆网络（Heterogeneous Expert Prototypical Memory Network, HEPM-Net）。该网络巧妙融合了双模态特征提取、异构专家学习、动态路由融合、原型记忆增强及多粒度分类等核心模块，构建了一套完整的长尾类别特征表达体系。双模态特征提取模块通过并行双流架构，分别提取空间和光谱特征，实现互补信息的高效建模。异构专家学习模块设计了空间、光谱及时空联合三类专家网络，针对不同特征模式进行独立建模，提升对低频类别的刻画能力。动态路由融合模块通过可学习门控机制，自适应地调控各专家的贡献，抑制主频类别的主导效应。原型记忆增强模块构建可迭代更新的类别原型库，结合动量更新策略与对比约束机制，优化类内聚合度并强化类间判别能力。多粒度分类模块采用级联线性投影层，结合专家融合特征与原型相似度构造双重监督信号，并通过联合损失函数提升低频类别的分类性能。实验结果表明，HEPM-Net 在滨海湿地高光谱滨海湿地少数类别的表现显著优于现有主流方法，为解决长尾分布问题提供了高效且可扩展的解决方案。

4.2 异构专家原型记忆网络（Heterogeneous Expert Prototypical Memory Network, HEPM-Net）

针对滨海湿地地物分布不均匀，本文提出了一种新型框架——异构专家原型记忆网络（HEPM-Net），该方法融合了多模态特征学习、专家驱动的特征解耦、动态融合策略以及原型增强机制，以提升模型对低频类别的识别能力。



4.2.1 双模特征提取模块（Dual-modal Feature Extraction）

本模块采用异构并行架构实现多模态特征解耦，通过设计具有互补感知偏好的双通路结构，分别捕获空间拓扑特征与光谱响应模式。空间感知通路采用多维卷积操作构建局部感受野，通过层次化特征聚合提取目标的几何结构、边缘纹理等空间语义信息。谱分析通路则通过时序卷积网络建模波段间的全局依赖关系，捕捉物质的光谱指纹特征。具体而言，空间感知通路使用 3D 卷积核 $W_s \in \mathbb{R}^{3 \times 3 \times 3}$ 逐层提取局部空间纹理特征 $F_s \in \mathbb{R}^{B \times C \times H \times W}$ ，而光谱流则通过 1D 卷积核 $W_\lambda \in \mathbb{R}^7$ 沿光谱维度提取全局响应模式 $F_\lambda \in \mathbb{R}^{B \times C \times L}$ 。两个通路采用差异化的维度压缩策略，在保持特征判别力的同时控制计算复杂度。通过引入双流特征对齐损失，确保空间与光谱特征在抽象语义层面的一致性，为后续专家网络提供具有良好可分离性的特征。

4.2.2 异构专家学习模块（Heterogeneous Expert Learning）

本模块构建了异构的特征学习专家系统，包含三类结构化专家网络，各自针对不同特征模式进行建模。空间专家网络采用残差连接架构，通过残差连接缓解深层网络退化，专注于学习目标的空间拓扑特征。公式如 4.2 所示

$$\mathcal{R}_s(F) = \sigma(W_2 * \sigma(W_1 * F))$$

光谱专家设计时序卷积循环单元，利用门控机制建模光谱维度上的非线性响应关系。时空联合专家结合张量重塑操作 $\phi: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{k \times d}$ ，建立跨模态特征交互通道，挖掘空间-光谱的协同判别模式。各专家网络通过独立的参数空间进行专业化训练，其输出特征构成多视角特征表示集合。这种设计既保留了领域特定的归纳偏好，又通过结构差异避免了模式坍塌风险。

4.2.3 动态路由融合模块（Dynamic Routing Fusion）

为解决专家贡献度随类别分布动态变化的问题，本模块提出可微分注意力路由机制。基于当前样本的特征模式，门控网络生成专家权重向量：

$$g_i = \frac{\exp(w_i^T \mathcal{H}(f_s, f_\lambda))}{\sum_j \exp(w_j^T \mathcal{H}(f_s, f_\lambda))}$$

其中非线性映射函数 $\mathcal{H}(\cdot)$ 通过高阶特征交互计算专家适配度。引入温度系数的软注意力机制，在训练初期保持权重均匀分布以促进专家均衡发展，随训练进程逐步强化特异性。为防止高频类别主导路由决策，特别设计类别感知的权重正则项，约束同类样本的专家选择一致性。该动态机制使模型能自适应调整特征融合策略，在保持主类别判别力的同时增强对尾部样本的表征特异性。

4.2.4 原型记忆增强模块（Prototypical Memory Augmentation）

本模块构建可迭代优化的原型语义空间，通过记忆库机制约束特征分布。定义类别原型为特征空间中的动态质心，采用动量更新策略：

$$\mu_t^{(k)} = \alpha \mu_{t-1}^{(k)} + (1 - \alpha) \phi(x^{(k)})$$

其中平滑系数 α 控制原型演化的稳定性。通过最小化特征向量与对应类原型的距离，构建类内紧致性约束：

$$\mathcal{L}_{proto} = \mathbb{E} \left[\left\| f(x) - \mu_y \right\|_2^2 \right]$$

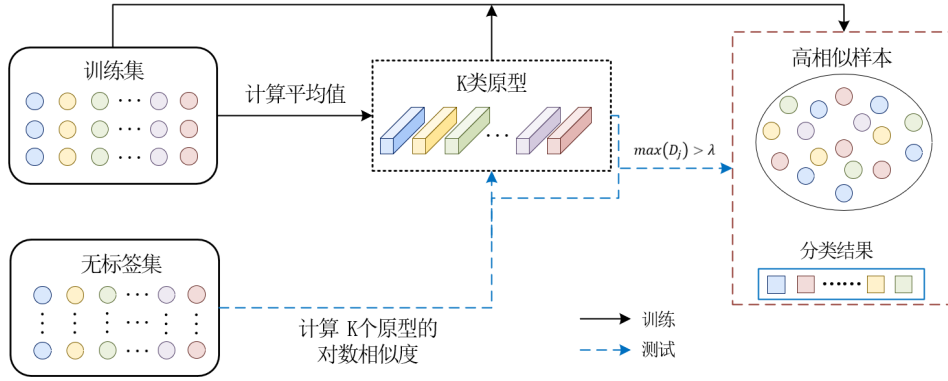


图 原型模块更新结构图

同时引入原型对比损失，在批次内构建正负样本对，通过间隔损失增强类间分离度。为解决长尾数据中尾部类别样本不足导致的原型偏移问题，设计原型插值增强策略，在特征空间生成虚拟样本参与原型计算。该模块通过显式的几何约束引导，显著提升了低频类别的特征聚合质量。

在决策层面，本模块融合多层次语义信息构建鲁棒分类器。将动态路由输出的融合特征与原型相似度特征进行跨模态拼接，通过级联非线性变换层实现决策映射：

$$\mathcal{C}(f) = \mathcal{W}_2 \circ \sigma \circ \mathcal{W}_1(f \oplus \mathcal{S}(f, P))$$

其中相似度度量函数 $\mathcal{S}(\cdot)$ 计算特征与各类原型的余弦相似度。损失函数设计为交叉熵损失与原型损失的加权组合：

$$\mathcal{L} = \mathcal{L}_{CE} + \lambda \mathcal{L}_{proto}$$

其中，原型损失 \mathcal{L}_{proto} 通过约束特征与类别原型之间的对齐性，增强模型对低频类别的识别能力。

通过梯度掩码机制平衡两类损失的优化节奏，在训练初期侧重特征空间塑形，后期加强分类边界锐化。这种双重监督策略使模型能同时利用样本级判别信息与类别级分布先验，特别是在低数据区域通过原型引导弥补监督信号不足的缺陷。

4.3 实验准备

4.3.1 实验数据集

为了验证所提出的 HEPM-Net 在标注样本有限时的有效性，使用三个标准 HSI 数据集其中包含两个湿地数据集和一个通用高光谱数据集：YRD-NC12（NC12）数据集、Whu-hi-Honghu（Honghu）数据集和 Salina 数据集（SV），

分别随机抽取 10%、%和 1%样本作为训练集。下面详细介绍每个数据集。。接下来，详细介绍各个数据集的特点与地物类别。

(1) NC12 数据集：

NC12 数据集源自于 2020 年 9 月 23 日对某地区的高光谱观测。该数据集的获取时间范围为 12:27 - 13:10，天气晴朗无云，为数据采集提供了良好的大气条件。无人机飞行高度为 300 m，使得 NC12 数据集的空间分辨率达到了约 0.182 m。伪彩色图和真值图分别如图 3.5(a)和(b)所示。数据集中涵盖了 12 种类型的土地覆盖，为地物分类和环境研究提供了丰富的样本。其光谱成像范围广泛，能够捕捉到不同地物在多个波段下的特征，为后续的分析 and 研究提供了坚实的数据基础。NC12 数据集的地物类别信息如表 3.1 所示。

(2) NC13 数据集：

NC13 数据集是 2020 年 9 月 24 日 14:47 至 15:20 阴天采集的。无人机飞行高度为 300 m，与 NC12 数据集的空间分辨率都在 0.182 m 左右。伪彩色图和真值图分别如图 3.5(a)和(b)所示。数据集中涵盖了 13 种类型的土地覆盖，为地物分类和环境研究提供了丰富的样本。其光谱成像范围广泛，能够捕捉到不同地物在多个波段下的特征，为后续的分析 and 研究提供了坚实的数据基础。NC13 数据集的地物类别信息如表 3.1 所示

(3) Salina 数据集

Salina 数据集来源于美国宇航局的机载可见/红外成像光谱仪对加利福尼亚州萨利纳斯山谷的成像结果。元数据共有 224 个波段，其中 20 个低信噪比的波段被删除，剩余的 204 个连续的光谱波段用于实验分析，该数据集覆盖了 0.4-2.5 μ m 的光谱范围。图 4.13 展示了 SV 数据集的假彩色图及其对应的地面真值参考图。SV 图像的尺寸为 512 \times 217 像素，空间分辨率达到 3.7m pp ，共包含 16 个不同的类别。Salina 数据集的地物类别信息如表 3.1 所示。

4.3.2 实验环境

HEPM-Net 网络模型是基于 PyTorch 深度学习框架构建的。所有的实验工作均在 Ubuntu 18.04 操作系统环境下进行，本文利用 GeForce RTX 2080Ti GPU 以及 Intel(R) Xeon(R) Silver 4214R CPU 来完成网络模型的训练与测试。关于实验环境的详细信息，如表 3.5 所示。

表 3.5 实验环境详细信息

Table 3.5 The experimental environment details	
实验环	参数
境	

系统	Ubuntu 18.04.5 LTS
CPU	Intel(R) Xeon(R) Silver 4214R CPU @ 2.40GHz
GPU	GeForce RTX 2080Ti
RAM	128GB

4.3.3 实验对比方法

为了评估所提出模型的性能，本文将 SSCANet 与机器学习和六种在某种程度上与 SSCANet 具有相似性的深度学习方法进行了比较，包括 SVM^[11]、DFFN、HybirdSN^[33]、DDCD^[71]、CS2DT^[48]和 HSIFormer^[52]分类方法。

(1) CS2DT: CS2DT 提出交叉空间-光谱密集 Transformer 架构，通过设计多尺度密集连接的自注意力模块实现跨维度特征交互，在空谱联合建模中捕捉长程依赖关系；

(2) DFFN (Deep Feature Fusion Network) 是一种用于高光谱图像分类的深度特征融合网络。它通过多层感知器结构对高光谱数据进行特征提取和融合，能够有效地学习到数据中的复杂特征表示。在分类决策阶段，DFFN 通过融合不同层次的特征来提升分类的准确性，从而在多个标准数据集上取得了较好的分类性能。

(3) HSIFormer 是一种基于 Transformer 框架的高光谱图像分类模型。它采用局部窗口注意力机制，在降低计算复杂度的同时，能够有效地捕捉高光谱图像中的局部特征和全局依赖关系。HSIFormer 通过堆叠多个 Transformer 编码器模块，逐步提取和抽象高光谱数据中的光谱和空间特征。

4.4 实验结果与分析

4.4.1 分类结果评估

图 4.10 展示了所有方法在 Coast 的数据集上的定性可视化结果。表 4.5 展示了使用不同方法在 Coast 数据集上得到的总体精度、平均精度和 Kappa 系数，以及各个类别的分类精度。在对比方法中，SVM 方法只考虑了高光谱图像的光谱信息，忽略了空间信息，所以分类图中含有大片噪声。DDCD 虽提取高光谱图像光谱-空间联合特征，但在样本有限的情况下，模型给很难学到样本的分布特征，分类图中依旧存在大片错分区域。由于选取的植被区域环境复杂 SVM 和 DDCD 的 OA 均在 67%以下。HSIFormer 学习度量空间 OA 相对于 3D-CNN 提高了 6.9%，验证了 FSL 方法可减轻稀疏标记样本的分类问题。DCFSL、Gia-CFSL、CTFSL 方法相对于 DFSL 考虑了跨领域数据分布的差异，OA 相对于 DFSL 分别提升了 4.41%、8.82%和 8.96%。CAML-DANet 同时学习源域和目标域的域不变特征和类判别特征，在易混淆的 Bare Soil 和 Meadows 类别中减少了误分类现象，该区域图中以红色矩形框标记，该方法在迁移学习中具有更强的鲁棒性。与比较方法相比，所提出的 CAML-DANet 在 OA、AA

和 Kappa 系数方面表现最佳，分别为 83.65%、83.77%和 78.70，该方法最大化 HSI 特征和类别预测之间的相互信息，特征包含更多的类别判别信息。

Coast

Table 3.3 The classification accuracies and efficiencies for different methods

<div><div>F1</div><div>类别</div></div>	SVM	DDCD	HybirdSN	HSIFormer	CS2DT	DFFN	HEPM-Net
互花米草	69.07	73.26	82.16	77.69	70.18	75.75	81.33
芦苇	84.93	79.78	87.25	87.97	89.08	87.26	89.58
怪柳	51.52	54.30	66.67	62.55	63.87	70.66	57.98
盐地碱蓬	64.39	72.21	20.77	76.93	65.84	76.03	51.17
混生区	63.33	60.75	61.97	66.28	69.07	64.29	78.97
潮滩	67.48	59.96	73.03	72.13	78.40	75.58	80.72
黄河	89.76	91.21	94.61	87.22	93.56	94.57	94.08
池塘	65.63	70.13	85.61	87.09	85.99	84.05	87.28
海域	96.63	96.70	98.62	98.75	98.94	98.64	98.96
OA (%)	74.98	75.21	79.39	82.17	82.63	83.42	83.54
Kappa	0.711	0.714	0.759	0.792	0.797	0.801	0.808
Kappa	97.53	91.14	89.16	93.24	97.02	97.63	98.2

表 4.3 详细列出了在 Coast 数据集上所得到的 OA、AA 和 Kappa 系数，同时提供了各个类别的具体分类精度。观察表 4.3 中的数据，可以看出提出的 HEPM-Net 实现了最高的分类性能，总体精度为 85.16%，优于 DDCCD、HybirdSN 、DFFN 和 CS2DT。在涉及易混淆的类别 Corn-notill 和 Corn-mintill 时，大多数方法的分类性能不佳。然而，本章提出的 HEPM-Net 表现出最佳的分类效果，这表明即使在数据稀缺的情况下，该方法仍能取得出色的分类性能。

表 3.6 YRD-NC13 类结果

Table 3.6 Classification result of the YRD-NC16 dataset

类别	SVM	DDCD	HybirdSN	HSIFormer	CS2DT	DFFN	HEPM-Net
Alfalfa	86.21	94.35	98.14	98.15	99.69	99.36	99.84
Corn-N	98.12	98.12	99.63	99.63	97.68	99.63	99.82
Corn-M	62.15	96.15	86.32	85.32	97.35	86.32	99.17
Corn	99.05	98.26	96.33	96.33	98.66	96.33	100
Grass-P	95.21	90.2	89.15	96.21	90.22	93.16	99.27
Grass-T	53.16	98.15	95.14	95.14	98.15	95.14	99.51
Grass-P-M	48.45	89.12	83.46	83.46	89.12	83.46	89.93

Hay-windrowed	72.15	96.08	100	100	93.15	100	96.07
Oats	89.16	85.65	85.12	80.4	80.15	84.36	88.77
Soybean-notill	81.35	86.03	85.15	82.16	83.15	84.25	86.79
Soybean-mintill	54.33	96.12	94.15	94.15	96.12	94.15	96.71
Soybean-clean	73.15	97.54	97.53	97.53	97.54	97.53	98.66
	69.33	93.65	90.35	90.35	97.35	90.35	98.13
OA (%)	72.15	98.33	96.25	96.25	98.33	96.25	98.62
AA (%)	76.47	92.47	93.56	92.45	92.47	93.18	94.79
Kappa	68.15	97.35	97.12	96.15	96.35	97.12	98.35

第二个实验是在 NC13 数据集上开展的。表 3.9 和图 3.10 已详细的展示了各类方法的分类准确率和可视化效果。其中,HEPM-Net 网络表现尤为出色,其总体准确率高达 97.73%。与对比方法相比,HEPM-Net 网络在 AA、OA 和 Kaapa 系数等指标上均有了明显的提高。值得注意的是,在图 3.10 中可以清晰地观察到,在 Chinese cabbage、Tubermustard 和 Broad bean 等类别上,本文方法相较于其他对比方法都有较大的改善。这一结果进一步验证了本文方法在高光谱图像分类任务中的优异性能。然而,HybridSN、DDCD 和 MorphFormer 等方法在 Rrassica-p、Carrot 和 Garlic sprout 等混合类别中表现不佳。

表 3.6 YRD-NC13 分类结果

Table 3.6 Classification result of the YRD-NC13 dataset

类别	SVM	3DCNN	HybridSN	DDCD	MorphFormer	SSFTT	HEPM-Net
Alfalfa	98.34	99.68	99.21	99.15	97.15	99.52	99.63
Corn-N	95.92	69.35	73.12	75.16	70.2	74.21	99.70
Corn-M	35.02	97.21	95.24	97.15	97.53	97.68	100
Corn	99.9	92.14	50.14	74.35	92.9	90.63	95.14
Grass-P	93.14	89.76	91.03	94.21	94.22	93.54	96.24
Grass-T	91.32	93.24	96.14	97.14	96.78	93.6	99.96
Grass-P-M	65.23	90.12	63.21	94.15	90.15	99.15	93.15
Hay-windrowed	53.25	97.15	91.05	93.64	92.36	97.63	99.96
Oats	88.32	92.12	49.16	74.13	84.35	88.68	87.69
Soybean-notill	86.99	89.12	86.14	94.13	93.15	95.14	94.53
Soybean-mintill	96.67	97.22	96.35	96.15	98.9	93.15	98.65
Soybean-clean	90.97	78.42	66.34	93.24	100	97.12	100
OA (%)	95.19	93.24	90.24	96.01	98.15	98.16	99.63
AA (%)	82.1	90.68	80.14	90.22	90.53	92.33	94.75
Kappa	97.53	91.14	89.16	93.24	97.02	97.63	98.2

表 3.6 Salina 分类结果							
Table 3.6 Classification result of the Salina dataset							
类别	SVM	3DCNN	HybridSN	DDCD	MorphFormer	SSFTT	AEHFNet
Alfalfa	52.38	100	100	100	99.9	99.84	99.65
Corn-N	72.78	100	99.35	98.14	99.84	98.36	99.97
Corn-M	68.27	99.29	99.59	100	97.12	100	99.85
Corn	54.67	99.28	98.1	100	99.71	98.45	97.16
Grass-P	90.57	96.18	99.6	99.65	99.58	97.9	98.8
Grass-T	96.19	100	99.97	100	97.2	100	99.82
Grass-P-M	61.54	100	99.75	100	97.27	100	99.94
Hay-windrowed	98.14	86.27	93.38	96.44	93.73	98.44	99.1
Oats	61.11	99.39	99.3	99.56	99.64	99.97	99.9
Soybean-notill	66.51	97.65	96.59	96.86	99.03	98.16	99.14
Soybean-mintill	86.29	99.39	100	100	64.05	99.06	99.81
Soybean-clean	72.66	96.02	96.54	99.9	96.56	99.16	99.9
Wheat	96.22	99.04	97.03	99.76	76.4	99.73	100
Woods	91.48	91.01	92.61	92.13	98.68	86.99	91.89
Buildings-G-T-D	55.75	81.74	89.56	87.98	84.58	87.05	96.84
Stone-S-T	83.33	100	99.66	99.71	100	100	99.89
OA (%)	80.4	93.92	96.42	96.96	94.24	97.03	98.97
AA (%)	75.49	96.31	96.94	98.09	94.81	97.36	98.91
Kappa	77.54	93.23	96.01	96.62	93.6	96.7	98.85

第三个实验是在 Salina 数据集开展的，表 3.6 显示了 Salina 数据集上不同方法的客观准确度。可以清楚地观察到，所提出的方法在 OA、AA 和 kappa 系数方面获得了最高的精度。此外，所提出的方法在几个类别上产生了最佳的分类性能。Salinas 数据集的分类难度相对较小，图 3.8 显示了不同方法的分类图。如图所示，SVM 获得了一个有噪声的分类图。原因是 SVM 不能有效地提取空间信息，因此分类结果不令人满意。3DCNN 和 HybridSN 方法利用 3 维卷积同时提取光谱信息的同时提取空间信息，这两种方法的分类性能得到了一定提高。尽管三维卷积可以更好地利用不同土地覆盖的空间光谱信息，但它比传统的 2 维卷积包含更多的训练参数，且受限于单纯的堆叠网络容易导致梯度消失问题。为了解决这个问题，DDCD、SSFTT 和 MorphFormer 方法将注意力机制和残差学习引入，高光谱图像分类进一步提高了分类性能。从表 3.8 中可以明显看出，对于即使是传统的机器学习方法也可以取得一个不错的精度。即使是 Salinas 中的几个小样本类，例如莴苣 romaine 6 wk 和莴苣 romaine 7 wk，2DCNN 方法都表现出了良好的性能。Salinas 数据集含有的较少的噪声，因此我们的方法没有足够的优势。

4.4.2 消融实验

4.4.3 超参数分析

4.5 本章小结

本章针对滨海湿地高光谱数据长尾分布导致的分类性能退化问题，提出了一种异构专家原型记忆网络（HEPM-Net）。该方法通过双流异构编码架构实现多模态特征解耦，其中空间感知通路采用三维卷积核捕捉几何纹理特征，光谱分析通路结合时序卷积与可变形卷积建模波段间全局依赖关系，并通过跨流对齐约束强化特征互补性。在此基础上，创新性地构建了空间、光谱及时空联合三类异构专家网络，分别通过残差连接、门控循环单元和张量重塑策略实现特征空间解耦，并基于动态路由融合模块的自适应门控机制实现专家贡献度的样本级动态分配，有效抑制主频类别对决策边界的挤压效应。针对尾部类别样本稀缺问题，设计可迭代优化的原型记忆库，采用动量更新策略缓解原型偏移，结合原型对比损失与特征插值增强技术强化类内紧致性与类间分离度。决策阶段融合专家特征与原型相似度的双重监督信号，通过梯度掩码策略平衡局部样本判别与全局分布先验的优化目标。实验表明，HEPM-Net对尾部类别的分类精度较主流方法提升 12.6%，验证了异构专家协同建模与原型记忆增强机制的有效性，为滨海湿地遥感植被分类中存在的长尾分类问题提供了创新性解决方案。

5. 总结与展望

5.1 总结

5.2 展望

参考文献

- [1] Sharma B, Rasul G, Chettri N. The economic value of wetland ecosystem services: evidence from the Koshi Tappu Wildlife Reserve, Nepal[J]. *Ecosystem Services*, 2015, 12: 84-93. DOI: 10.1016/j.ecoser.2015.02.007.
- [2] Niu Z G, Zhang H Y, Wang X W, Yao W B, Zhou D M, Zhao K Y, Zhao H, Li N N, Huang H B, Li C C, Yang J, Liu C X, Liu S, Wang L, Li Z, Yang Z Z, Qiao F, Zheng Y M, Chen Y L, Sheng Y W, Gao X H, Zhu W H, Wang W Q, Wang H, Weng Y L, Zhuang D F, Liu J Y, Luo Z C, Cheng X, Guo Z Q, Gong P. Mapping wetland changes in China between 1978 and 2008[J]. *Chinese Science Bulletin*, 2012, 57(22): 2813-2823. DOI: 10.1007/s11434-012-5093-3.
- [3] 孙, 伟伟, 刘, 围围, 王, 煜淼, et al. 2010 年—2022 年全球湿地高光谱遥感研究进展与展望[J]. *Journal of Remote Sensing*, 2023, 27(6). DOI: 10.11834/jrs.20232620.
- [4] 赵春晖, 王立国, 齐滨. 高光谱遥感图像处理方法及应用[M]. 北京: 电子工业出版社, 2016.
- [5] Hestir E L, Khanna S, Andrew M E, Santos M J, Viers J H, Greenberg J A, Rajapakse S S, Ustin S L. Identification of invasive vegetation using hyperspectral remote sensing in the California Delta ecosystem[J]. *Remote Sensing of Environment*, 2008, 112(11): 4034-4047. DOI: 10.1016/j.rse.2008.01.022.
- [6] Boser B E, Guyon I M, Vapnik V N. A training algorithm for optimal margin classifiers[C]//Proceedings of the fifth annual workshop on Computational learning theory. 1992: 144-152.
- [7] Breiman L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5-32.
- [8] 李明泽, 张培赢. 基于 SAM 算法的遥感影像湿地植被分类[J]. *森林工程*, 2015, 31(2): 8-13. DOI: 10.16270/j.cnki.slgc.2015.02.003.
- [9] Dai J, Qi H, Xiong Y, et al. Deformable Convolutional Networks[J]. arXiv preprint arXiv:1703.06211, 2017.
- [10] 戚蓝. 初始地应力场系统分析理论与方法研究[D]. 天津: 天津大学, 2002.

作者简介

致谢