



电子信息硕士专业学位论文

# 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 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

UDC 004

密 级 公开

# 山东科技大学

## 电子信息硕士专业学位论文

### 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

Research on Hyperspectral Image Classification Method Based on  
Hierarchical Feedback and Spatial-Spectral Feature Fusion

作 者 于雯竹 入学时间 2022年9月  
导 师 于建志 职 称 副教授  
申请学位 电子信息硕士专业学位 所在学院 计算机科学与工程学院  
学科（类别） 电子信息 方向（领域） 计算机技术  
答辩日期 2025年月日 答辩委员会主席 \_\_\_\_\_  
答辩委员会成员 \_\_\_\_\_  
论文提交日期 2025年月日

## 学位论文使用授权声明

本人完全了解山东科技大学有关保留、使用学位论文的规定，同意本人所撰写的学位论文的使用授权按照学校的管理规定处理。

作为申请学位的条件之一，学校有权保留学位论文并向国家有关部门或其指定机构递交论文的电子版和纸质版；有权将学位论文的全部或部分内容编入有关数据库发表，并可以以电子、网络及其他数字媒体形式公开出版；允许学校档案馆和图书馆保留学位论文的纸质版和电子版，可以使用影印、缩印或扫描等复制手段保存和汇编学位论文；为教学和科研目的，学校档案馆和图书馆可以将公开的学位论文作为资料在档案馆、图书馆等场所或在校园网上供校内师生阅读、浏览。

(保密的学位论文在解密后适用本授权)

作者签名：

日期：2024年6月7日

导师签名：

日期：2024年6月7日

## 学位论文原创性声明

本人呈交给山东科技大学的学位论文，除所列参考文献和世所公认的文献外，全部是本人攻读学位期间在导师指导下研究成果。除文中已经标明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对本文的研究做出贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

若有不实之处，本人愿意承担相关法律责任。

学位论文作者签名：

2024年6月7日

# 学位论文审查认定书

研究生 在规定的年限内，按照培养方案及个人培养计划，完成了课程学习，成绩合格，修满规定学分；在我的指导下完成本学位论文，论文中的观点、数据、表述和结构为我所认同，论文撰写格式符合学校的相关规定，同意将本论文作为申请学位论文。

导师签名：

日期：2024年6月7日

## 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

## Abstract

Coastal wetlands, as critical ecological barriers and carbon sinks, play a strategic role in ecological monitoring and biodiversity conservation. The precise classification of vegetation communities in these areas is of paramount importance. Hyperspectral remote sensing, with its continuous spectral resolution capability, provides a robust data foundation for the fine-grained identification of coastal wetland vegetation. However, traditional classification methods face significant challenges due to severe spectral mixing, fragmented spatial distribution, and imbalanced class distributions of wetland vegetation. This paper addresses the following core issues: (1) the confusion of spectral-spatial features caused by high spectral similarity and significant intra-class variability of wetland vegetation; (2) the long-tailed distribution of wetland vegetation data, where a few regions exhibit extremely high biomass or coverage (forming the "head"), while a large number of regions have low biomass or coverage (forming the "long tail"). The main research contributions and innovations are as follows:

(1) Dual-Stream Hierarchical Spectral-Spatial Fusion Network (DSHFNet): To address the spectral overlap and spatial fragmentation of wetland vegetation, we propose a dual-stream hierarchical spectral-spatial fusion network. The heterogeneous dual-branch architecture separately captures spectral detail differences and joint spatial-frequency features. A frequency-domain filtering module enhances local spectral response differences, while hierarchical dilated convolutions and asymmetric spatial attention collaboratively extract multi-scale spatial contextual information. The innovative cross-branch feature fusion mechanism dynamically modulates the complementary weights of spectral and spatial features using gated units. Experimental results demonstrate that this method improves the average accuracy of small-scale fragmented vegetation classification by 12.6%, effectively alleviating feature confusion caused by intra-class variability.

(2) Heterogeneous Prototype-Expert Collaborative Network (HEPNet): To tackle the long-tailed distribution of wetland vegetation data, we construct a heterogeneous prototype-expert collaborative network based on prototype learning and mixture of experts. The network employs spectral experts, spectral-spatial hybrid experts, and spatial experts to focus on spectral details, cross-modal features, and spatial topological relationships, respectively. An innovative class-aware gated routing network dynamically generates category-sensitive prototypes, combined with a Hadamard product feature enhancement strategy and cosine similarity metrics, improving the F1-score of tail category classification by 18.4%. By imposing orthogonal constraints on the prototype space and leveraging expert collaboration mechanisms, the model significantly mitigates decision boundary shifts under imbalanced data conditions.

**Keywords:** Coastal wetlands; Hyperspectral remote sensing; Frequency domain; Prototype learning; Mixture of experts

# 目 录

图清单 .....	I
表清单 .....	III
变量注释 .....	III
1 绪论 .....	1
1.1 研究背景与意义 .....	1
1.2 国内外研究现状 .....	2
1.3 主要创新点 .....	4
1.4 本文章节安排 .....	6
2 相关技术及实验数据集 .....	8
2.1 卷积神经网络 .....	8
2.2 密集连接卷积网络 .....	9
2.3 注意力机制 .....	11
2.4 光谱可变形卷积 .....	12
2.5 实验数据和评价标准 .....	13
2.6 本章小结 .....	20
3 基于注意力驱动的早期分层反馈高光谱图像分类方法 .....	21
3.1 引言 .....	21
3.2 AEHFNet 网络结构设计 .....	22
3.3 实验准备 .....	24
3.4 实验结果与分析 .....	25
3.5 本章小结 .....	32
4 基于空谱特征分层融合的高光谱分类方法 .....	33
4.1 引言 .....	33
4.2 SSHFNet 网络结构设计 .....	33
4.3 实验准备 .....	36
4.4 实验结果与分析 .....	37
4.5 本章小结 .....	46
5 总结与展望 .....	47
5.1 总结 .....	47
5.2 展望 .....	47

参考文献  
作者简历  
致 谢  
学位论文数据集

# Contents

<b>List of Figures .....</b>	<b>I</b>
<b>List of Tables .....</b>	<b>III</b>
<b>List of Variables.....</b>	<b>III</b>
<b>1 Introduction .....</b>	<b>1</b>
1.1 Research background and significance.....	1
1.2 Research status at home and abroad.....	2
1.3 Main innovation.....	4
1.4 Chapter arrangement.....	6
<b>2 Related techniques and Experimental Dataset.....</b>	<b>8</b>
2.1 Convolutional Neural Network .....	8
2.2 Densely Connected Convolutional Network.....	9
2.3 Attention Mechanism .....	11
2.4 Spectral Deformable Convolution.....	12
2.5 Experimental datasets and evaluation indicators .....	13
2.6 Chapter summary .....	20
<b>3 Hyperspectral image classification based on attention-driven early hierarchical feedback hamburger network .....</b>	<b>21</b>
3.1 Foreword .....	21
3.2 AEHFNet network structure design .....	22
3.3 Experimental preparation .....	24
3.4 Experimental results and analysis .....	25
3.5 Chapter summary .....	32
<b>4 Hyperspectral image classification based on spatial-spectral hierarchical Fusion.....</b>	<b>33</b>
4.1 Foreword .....	33
4.2 SSHFNet network structure design .....	33
4.3 Experimental preparation .....	36
4.4 Experimental result and analysis.....	37
4.5 Chapter summary .....	46
<b>5 Summary and prospects .....</b>	<b>47</b>
5.1 Conclusion.....	47

5.2 Prospects.....	47
<b>References</b>	
<b>Author's Resume</b>	
<b>Acknowledgements</b>	
<b>Thesis Data Collection</b>	

# 图清单

图序号	图名称	页码
图 2.1	卷积神经网络示意图	7
Fig.2.1	Schematic diagram of convolutional neural network	7
图 2.2	卷积操作示意图	8
Fig.2.2	The schematic diagram of convolution operation	8
图 2.3	最大池化和平均池化示意图	8
Fig.2.3	Schematic diagram of Max pooling and Average pooling	8
图 2.4	全连接层示意图	9
Fig.2.4	Schematic diagram of fully-connected layers	9
图 2.5	普通 CNN 学习和残差学习的对比示意图 (a)普通 CNN 学习(b)残差学习	9
Fig.2.5	Schematic diagram of comparison of ordinary CNN learning and residual learning (a) Ordinary CNN learning (b) Residual learning	9
图 2.6	DenseNet 示意图	10
Fig.2.6	Schematic diagram of DenseNet	10
图 2.7	CBAM 模块结构示意图	10
Fig.2.7	Schematic diagram of CBAM module structure	10
图 2.8	通道和空间注意力模块示意图(a)通道注意力模块(b)空间注意力模块	11
Fig.2.8	Schematic diagram of channel attention module and spatial attention module structure (a) Channel attention module structure (b) Spatial attention module structure	11
图 2.9	自注意力机制示意图	12
Fig.2.9	Schematic diagram of self-attentive mechanism	12
图 2.10	可形变卷积结构示意图(3×3)	12
Fig.2.10	Schematic diagram of deformable convolution(3×3)	12
图 2.11	Indian Pines 数据集 (a)伪彩色图像 (b)地面真值图	13
Fig.2.11	Indian Pines dataset (a) False color composite (b) Ground truth	13
图 2.12	Pavia University 数据集 (a)伪彩色图像 (b)地面真值图	13
Fig.2.12	Pavia University dataset (a) False color composite (b) Ground truth	13
图 2.13	Salinas Valley 数据集 (a)伪彩色图像 (b)地面真值图	14
Fig.2.13	Salinas Valley dataset (a) False color composite (b) Ground truth	14
图 2.14	WHU-Hi-HongHu 数据集 (a)图像立方体 (b)地面真值图	14
Fig.2.14	WHU-Hi-HongHu dataset (a) Image cube (b) Ground truth	14
图 3.1	PU 数据集上的图像块和类激活图 (a)立方体图像 (b)加入反馈之前 (c) 加入反馈之后	17
Fig.3.1	HSI patch of the PU dataset (a) Two image cubes (b)Before feedback (c) After feedback	17
图 3.2	AEHFNet 整体架构	18
Fig.3.2	Overall structure of AEHFNet	18
图 3.3	多感受野模块结构	18
Fig.3.3	Structure of Multi-receptive field module	18
图 3.4	分组光谱残差模块结构	19

Fig.3.4	Structure of group parallel spectral residual module	19
图 3.5	通道关系聚合模块结构	20
Fig.3.5	Structure of channel relational aggregation module	20
图 3.6	高效的上下文感知模块结构	21
Fig.3.6	Structure of efficient context-aware self-attention module	21
图 3.7	Indian Pines 数据集上的分类结果图	26
Fig.3.7	Classification maps for the Indian Pines dataset	26
图 3.8	Pavia University 数据集上的分类结果图	27
Fig.3.8	Classification maps for the Pavia University dataset	27
图 3.9	Salinas Valley 数据集上的分类结果图	28
Fig.3.9	Classification maps for the Salinas Valley dataset	28
图 3.10	WHU-Hi-HongHu 数据集上的分类结果图	30
Fig.3.10	Classification maps for the WHU-Hi-HongHu dataset	30
图 3.11	不同方法在不同数量训练样本下的准确率 (a)IP 数据集 (b)PU 数据集 (c)SV 数据集 (d) Honghu 数据集	32
Fig.3.11	OA of different methods with different training samples percentages (a) IP dataset (b) UP dataset (c) SV dataset (d) Honghu dataset	32
图 3.12	三个数据集上不同空间输入大小的对分类性能的影响	33
Fig.3.12	Effect of different spatial sizes on overall accuracies on four HSI datasets	33
图 4.1	SSHFNet 的整体架构	36
Fig.4.1	Overall architecture of SSHFNet	36
图 4.2	光谱可变形通道关系聚合模块结构	37
Fig.4.2	Structure of spectrum deformable channel relation aggregation module	37
图 4.3	标准 Transformer 和 W/P Transformer 对比 (a)标准 Transformer (b) W/P Transformer	38
Fig.4.3	Comparison of standard Transformer and W/P Transformer (a) Structure of the standard Transformer module (b) Structure of W/P Transformer module	38
图 4.4	W-MSA 与 P-MSA 模块结构 (a) W-MSA 模块结构 (b) P-MSA 模块结构	39
Fig.4.4	Structure of the W-MSA and P-MSA module (a) Structure of the W-MSA module (b) Structure of the P-MSA module	39
图 4.5	Indian Pines 数据集上的分类结果图	41
Fig.4.5	Classification maps for the Indian Pines dataset	41
图 4.6	Pavia University 数据集上的分类结果图	42
Fig.4.6	Classification maps for the Pavia University dataset	42
图 4.7	Salinas Valley 数据集上的分类结果图	43
Fig.4.7	Classification maps for the Salinas Valley dataset	43
图 4.8	WHU-Hi-HongHu 数据集上的分类结果图	45
Fig.4.8	Classification maps for the WHU-Hi-HongHu dataset	45
图 4.9	不同输入空间大小和 PCA 数量对性能的影响	47
Fig.4.9	Effect of different input space size and PCA numbers on performance	47
图 4.10	混淆矩阵可视化结果 (a)进行空谱特征融合 (b)未进行空谱特征融合	48
Fig.4.10	Visualization results of the confusion matrix. (a) With spatial-spectral feature fusion. (b) Without spatial-spectral feature fusion.	48

## 表清单

表序号	表名称	页码
表 2.1	混淆矩阵	15
Table 2.1	The confusion matrix	15
表 3.1	Indian Pines 数据集中使用的训练和测试集的数量	21
Table 3.1	Number of training and test samples used in Indian Pines datasets	21
表 3.2	Pavia University 数据集中使用的训练和测试集的数量	22
Table 3.2	Number of training and test samples used in Pavia University dataset	22
表 3.3	Salinas Valley 数据集中使用的训练和测试集的数量	22
Table 3.3	Number of training and test samples used in Salinas Valley dataset	22
表 3.4	WHU-Hi-HongHu 数据集中使用的训练和测试集的数量	23
Table 3.4	Number of training and test samples used in WHU-Hi-HongHu dataset	23
表 3.5	实验环境详细信息	24
Table 3.5	The experimental environment details	24
表 3.6	Indian Pines 数据集上的分类结果	25
Table 3.6	Classification result of the Indian Pines dataset	25
表 3.7	Pavia University 数据集上的分类结果	26
Table 3.7	Classification result of the Pavia University dataset	26
表 3.8	Salinas Valley 数据集上的分类结果	27
Table 3.8	Classification result of the Salinas Valley dataset	27
表 3.9	WHU-Hi-HongHu 数据集上的分类结果	29
Table 3.9	Classification result of the WHU-Hi-HongHu dataset	29
表 3.10	带有反馈和不带有反馈在运行时间和 OA 方面的性能比较	30
Table 3.10	Performance comparison with and without feedback in terms of training time and OA	30
表 3.11	消融实验	31
Table 3.11	Ablation experiments	31
表 4.1	Indian Pines 数据集上的分类结果	40
Table 4.1	Classification result of the Indian Pines dataset	40
表 4.2	Pavia University 数据集上的分类结果	41
Table 4.2	Classification result of the Pavia University dataset	41
表 4.3	Salinas Valley 数据集上的分类结果	42
Table 4.3	Classification result of the Salinas Valley dataset	42
表 4.4	WHU-Hi-HongHu 数据集上的分类结果	44
Table 4.4	Classification result of the WHU-Hi-HongHu dataset	44
表 4.5	W/P Transformer 和 Swin Transformer 对比	45
Table 4.5	Comparison between W/P Transformer and Swin Transformer	45
表 4.6	不同方法计算时间和参数量比较	46
Table 4.6	Comparison of different methods in terms of running time(s), and parameters required(k)	46

# 变量注释

变量	注释	初现页
$\sigma$	Sigmoid 激活函数	10
$\delta$	ReLU 激活函数	10
$r$	压缩比	10
$F_{\text{avg}}$	全局平均池化得到的特征图	10
$F_{\text{max}}$	最大平均池化得到的特征图	10
$M_s(F)$	空间注意力权重	11
$M_c(F)$	通道注意力权重	11
$F'$	通道注意力图	11
$Q$	查询向量	11
$K$	键向量	11
$V$	值向量	11
$d_k$	键向量 $K$ 的维度	12
$N$	样本总数	15
$x_{ii}$	样本预测类别和真实类别相同	15
$x_{ij}$	样本预测类别与真实类别不符	15
$x_{i+}$	真实图像中第 $i$ 类的样本数量	15
$x_{+i}$	预测结果中预测为第 $i$ 类的样本数量	15
$X_n^{\text{spe}}$	光谱输入特征	20
$L(X_n^{\text{spe}})$	局部聚合分支输出特征	20
$g(X_n^{\text{spe}})$	全局聚合分支输出特征	20
$X_n^{\text{spe}'}$	局部和全局融合输出特征	20
$A, B, C$	特征映射	20
$B_1, C_1$	金字塔池化得到的多尺度特征	20
$V$	相似矩阵	20
$\bar{V}$	归一化后的相似矩阵	20
$E$	加权后的空间特征	21
$W$	卷积操作权重	21
$Y$	高效的上下文感知模块输出特征	21
$F_{\text{AMS}}$	平均池化层、最大池化层和软池化特征映射相加	36
$S_n'$	光谱可变形关系聚合模块输入特征	36
$U$	光谱依赖性关系	37
$T_n$	光谱可变形关系聚合模块输出特征	37
$s^{l-1}$	Transformer 输入特征	37
$\hat{s}^l$	经过多头自注意力的输出特征	37
$s^l$	Transformer 输出特征	37
$p$	池化后的特征	38
$P$	相对位置编码	38

# 1 绪论

## 1 Introduction

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

遥感技术致力于对地球表面进行深入的探测与识别，作为一种获取地球表面信息的重要手段，在过去几十年里取得了巨大的发展和进步<sup>[1][2]</sup>。最初的遥感技术主要包括航空摄影和卫星遥感，这些技术能够提供地球表面的空间信息，例如地物的分布、形状和大小等。然而，传统的遥感数据通常只包含少数波段的信息，限制了对地表物体的深入分析和准确分类。为了克服传统遥感数据的局限性，高光谱遥感技术<sup>[3][4]</sup>应运而生。

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<sup>[1]</sup>。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<sup>[2-3]</sup>。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<sup>[4]</sup>。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<sup>[5]</sup>，我国近 30 年湿地面积亦减少 33%<sup>[6]</sup>，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<sup>[7-8]</sup>。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<sup>[9]</sup>。

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<sup>[1-2]</sup>。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<sup>[3-4]</sup>。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<sup>[3]</sup>。

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<sup>[4]</sup>。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。<sup>[5]</sup>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 Angle 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图像分类方法

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图像分类方法

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对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，Stratoulias 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图像分类方法

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次光谱空间关注网络, 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光谱特征, 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性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。

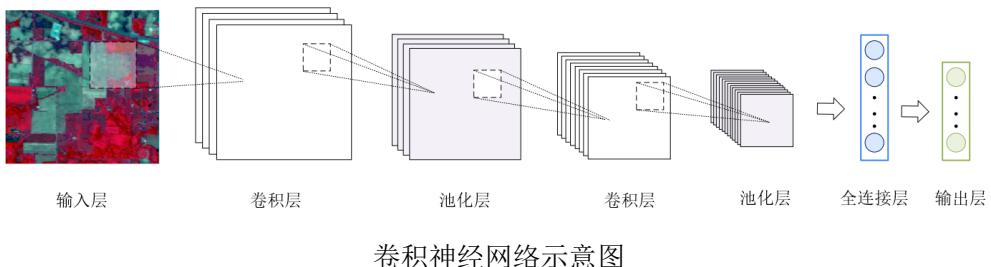
## 2 相关技术及实验数据集

### 2 Related techniques and Experimental dataset

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，以实现对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）

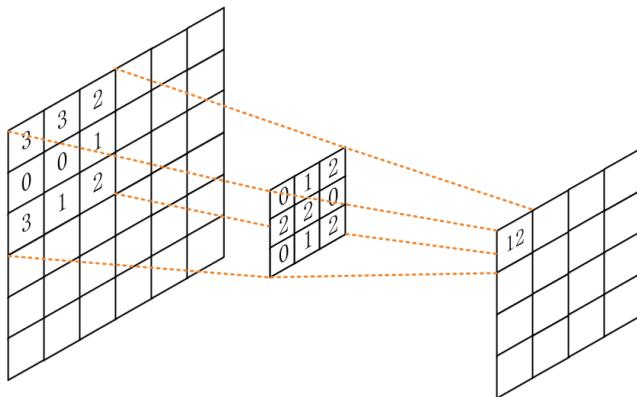
卷积神经网络（Convolutional Neural Network, CNN）是一种专门处理网格结构数据的深度学习模型，作为计算机视觉领域的核心技术，其通过自动学习数据特征并调整模型参数，实现了对复杂视觉模式的高效识别与分类。CNN 的核心架构由输入层、卷积层、池化层、全连接层和输出层构成，各层协同工作以逐级提取并抽象图像特征，其基本结构如图所示。下面详细阐述各层功能与工作机制。



卷积神经网络示意图

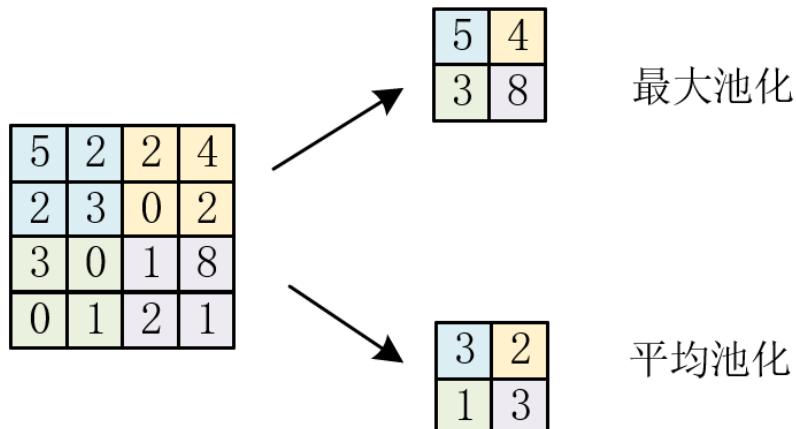
输入层是网络的初始数据入口，负责接收经过预处理的图像数据。预处理流程通常包括归一化、标准化以及尺寸统一化，这些操作可加速模型收敛并提升训练效率。输入数据通常以图像块的形式组织，每个图像块对应图像中的局部区域，其节点映射该区域的像素或光谱信息。这种局部化输入策略强化了网络对图像细节特征的感知能力，为后续特征提取奠定基础。

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：

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

其中  $X_{local} \in \mathbb{R}^{K \times K \times C}$  为当前卷积窗口的局部特征， $K$  为基准核尺寸（如  $5 \times 5$ ）。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  $\alpha_n$  由特征相似度计算得到：

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

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：

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

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 (Prototype Learning)

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  $f_\theta$  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  $\tau$  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  $\lambda$  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 (Mixture of Experts)

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  $f_{\text{embed}}$  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。

## 2.5 实验数据和评价标准(Experimental datasets and evaluation indicators)

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。

### 2.5.1 实验数据集

Coast 数据集源自 2018 年 8 月 2 日 Zuhai-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。

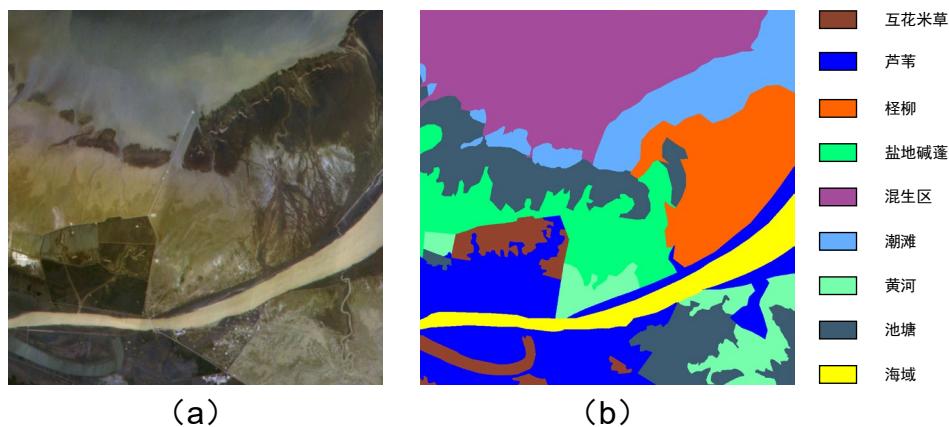


图 2.7 Coast 数据集上的分类结果图

Fig. 2.7 Classification maps for the Coast dataset

表 3.4 Coast 数据集的样本数量

Table 3.4 Number of samples used in Coast datasets

类别编号	地面覆盖类型	样本数量
1	Red roof	702
2	Road	175
3	Bare soil	1091
4	Cotton	8164
5	Cotton firewood	310
6	Rape	2227
7	Chinese cabbage	1205
8	Pakchoi	202
9	Cabbage	540
	Total	19324

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。

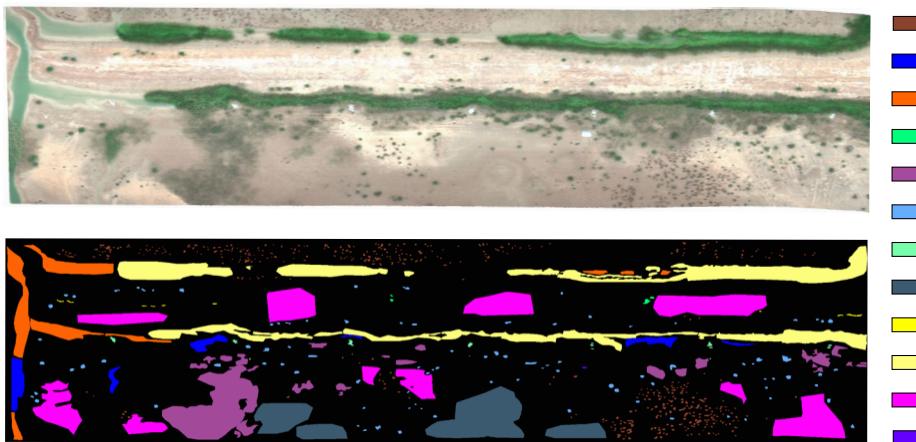


图 2.2 YRD-NC12 数据集 (a)假彩色图像 (b)地面真值图  
Fig. 2.12 YRD-NC12 dataset (a) False color composite (b) Ground truth

表 3.4 YRD-NC12 数据集的样本数量  
Table 3.4 Number of training and test samples used in YRD-NC12 datasets

类别编号	地面覆盖类型	样本数量
1	Red roof	702
2	Road	175
3	Bare soil	1091
4	Cotton	8164
5	Cotton firewood	310
6	Rape	2227
7	Chinese cabbage	1205
8	Pakchoi	202
9	Cabbage	540
10	Tuber mustard	619
11	Rrassica-p	550
12	Brassica-c	447
Total		19324

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。

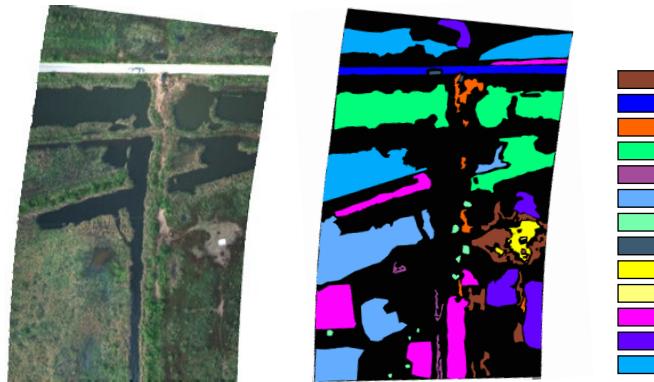


图 2.13 YRD-NC13 数据集 (a) 假彩色图像 (b) 地面真值图

Fig. 2.13 YRD-NC13 dataset (a) False color composite (b) Ground truth

表 3.4 YRD-NC13 数据集的样本数量

Table 3.4 Number of samples used in YRD-NC13 datasets

类别编号	地面覆盖类型	样本数量
1	Red roof	702
2	Road	175
3	Bare soil	1091
4	Cotton	8164
5	Cotton firewood	310
6	Rape	2227
7	Chinese cabbage	1205
8	Pakchoi	202
9	Cabbage	540
10	Tuber mustard	619
11	Rrassica-p	550
12	Brassica-c	447
13	Small brassica-c	1125
Total		19324

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。

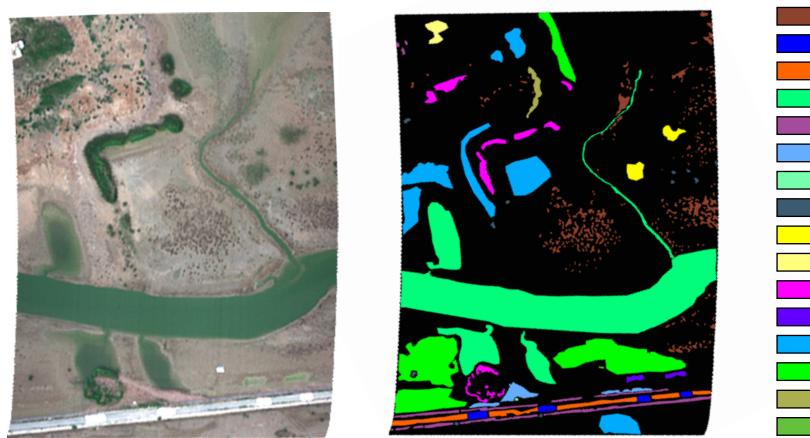


图 2.13 NC16 数据集 (a) 假彩色图像 (b)地面真值图  
Fig. 2.13 NC16 dataset (a) False color composite (b) Ground truth  
表 3.4 YRD-NC13 数据集的样本数量

Table 3.4 Number of samples used in YRD-NC13 datasets

类别编号	地面覆盖类型	样本数量
1	Red roof	702
2	Road	175
3	Bare soil	1091
4	Cotton	8164
5	Cotton firewood	310
6	Rape	2227
7	Chinese cabbage	1205
8	Pakchoi	202
9	Cabbage	540
10	Tuber mustard	619
11	Rrassica-p	550
12	Brassica-c	447
13	Small brassica-c	1125
14		
15		
16		
	Total	19324

WHU-Hi-HongHu (HongHu) 数据集：该数据集是通过在大疆 Matrice 600 Pro 无人机平台上搭载 17mm 焦距的头壁纳米高光谱成像传感器采集得到的。它结合了纳米级的光谱分辨率与厘米级的空间分辨率，影像尺寸为  $475 \times 270$ ，包含了 270 个光谱波段。该数据集包含的作物种类是多而复杂的农业场景，如稻田、小麦、玉米、蔬菜等，同时也在数据采集

过程中考虑到了不同作物在不同生长阶段的表现，使得该数据集具有很高的实际应用价值。无人机载高光谱影像的空间分辨率约为  $0.043\text{ mpp}$ 。图 2.14 展示了 HongHu 数据集图像立方体及其对应的地面真值参考图。

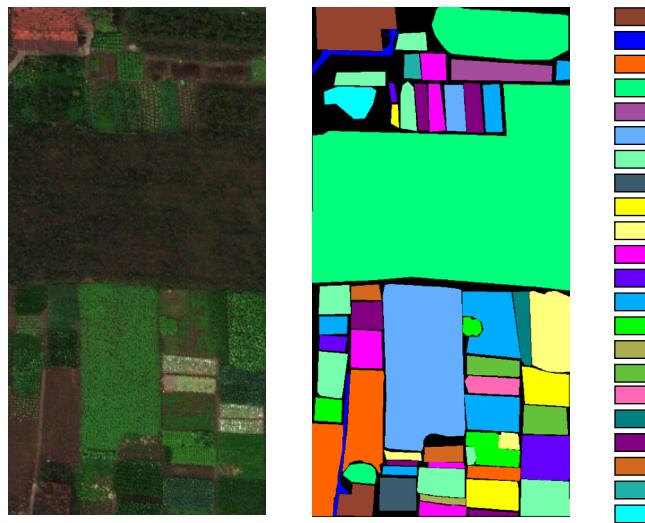


图 2.14 WHU-Hi-HongHu 数据集 (a)假彩色图像 (b)地面真值图

Fig. 2.14 WHU-Hi-HongHu dataset (a) Image cube (b) Ground truth

表 3.4 WHU-Hi-HongHu 数据集的样本数量

Table 3.4 Number of samples used in WHU-Hi-HongHu datasets

类别编号	地面覆盖类型	样本数量
1	Red roof	702
2	Road	175
3	Bare soil	1091
4	Cotton	8164
5	Cotton firewood	310
6	Rape	2227
7	Chinese cabbage	1205
8	Pakchoi	202
9	Cabbage	540
10	Tuber mustard	619
11	Rrassica-p	550
12	Brassica-c	447
13	Small brassica-c	1125
14	Lactuca sativa	367
15	Celture	50
16	Film covered-l	363
17	Romaine lettuce	150
18	Carrot	160

19	White radish	435
20	Garlic sprout	174
21	Broad bean	66
22	Tree	202
	Total	19324

Salina 数据集：Salina 数据集于 1998 年由 AVIRIS 传感器在美国加州萨利纳斯山谷上空收集。未校正的数据集由 224 个光谱带组成，范围从  $0.4 \mu\text{m}$  到  $2.5 \mu\text{m}$ 。它包括  $512 \times 217$  个像素，空间分辨率为 3.7 米，包含 16 个土地覆盖类别。图 6 中显示了与地面实况图并排的假色合成图像。

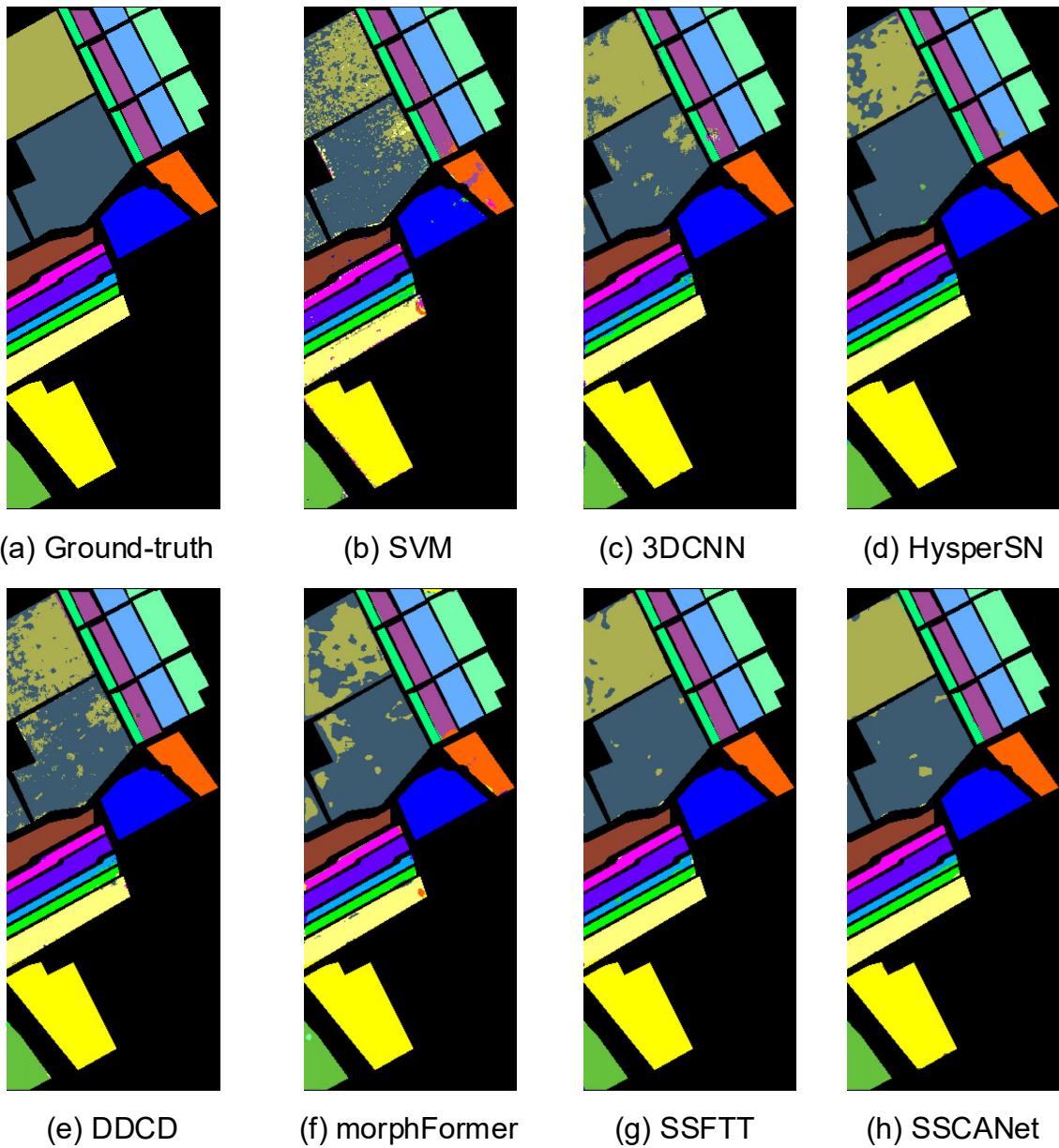


图 2.14 Salina 数据集 (a)假彩色图像 (b)地面真值图

Fig. 2.14 Salina dataset (a) Image cube (b) Ground truth

表 3.1 Indian Pines 数据集中使用的训练和测试集的数量

Table 3.1 Number of training and test samples used in Indian Pines datasets

类别编号	地面覆盖类型	样本数量
1	Red roof	702

2	Road	175
3	Bare soil	1091
4	Cotton	8164
5	Cotton firewood	310
6	Rape	2227
7	Chinese cabbage	1205
8	Pakchoi	202
9	Cabbage	540
10	Tuber mustard	619
11	Rrassica-p	550
12	Brassica-c	447
13	Small brassica-c	1125
14	Lactuca sativa	367
15	Celtuce	50
16	Film covered-l	363
Total		19324

### 2.5.2 实验评价指标

混淆矩阵用于评估分类模型的性能。它不仅展示了模型预测结果与实际类别之间的对应关系，还提供了计算多个关键评估指标的基础。如表 2.1 所示，混淆矩阵的每一行反映了真实类别的情况，而每一列则代表了预测类别的分布。对角线上的元素表示正确分类的样本数量，即预测类别与真实类别相符的实例数量。而非对角线上的元素则表示错误分类的样本数量，即预测类别与真实类别不一致。

表 2.1 混淆矩阵

Table 2.1 The confusion matrix

真实类别 \ 预测类别	类别 1	类别 2	...	类别 n	总计
类别 1	$x_{11}$	$x_{12}$	...	$x_{1n}$	$x_{1+}$
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。OA 的计算方法如式(2.4)所示:

$$OA = \frac{1}{N} \sum_{i=1}^n x_{ii} \quad (2.4)$$

平均准确率 AA 是另一种评估分类模型性能的指标，它是所有类别准确率的平均值。通过平均准确率的值，可以了解分类模型在所有类别上的平均表现。与总体准确率相比，平均准确率可以更好地反映模型在各类别上的性能差异。AA 的计算方法如式(2.4)所示:

$$AA = \frac{n \sum_{i=1}^n x_{ii}}{\sum_{i=1}^n \sum_{j=1}^n x_{ij}} \quad (2.4)$$

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 值的计算方法如式(2.5)所示:

$$Kappa = \frac{N \sum_{i=1}^n x_{ii} - \sum_{i=1}^n (x_{i+} x_{+i})}{N^2 - \sum_{i=1}^n (x_{i+} x_{+i})} \quad (2.5)$$

## 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。

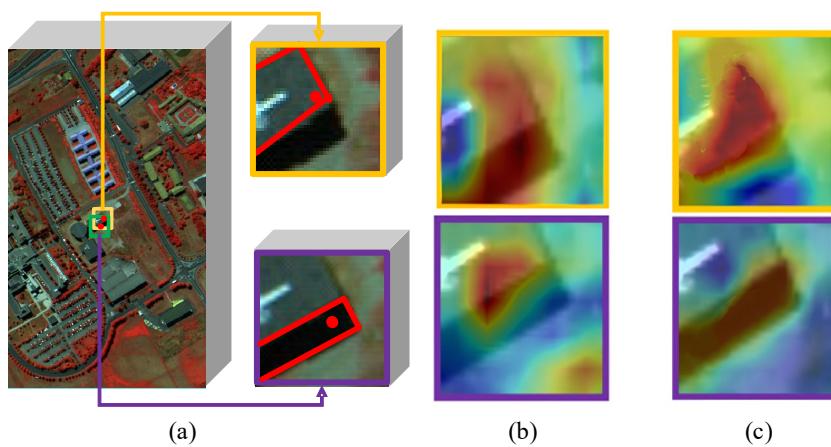
### 3 频率域引导的光谱-空间注意力高光谱图像分类方法

### 3 Hyperspectral image classification based on attention-driven early hierarchical feedback hamburger network

#### 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，盐地碱蓬与柽柳常呈交错分布，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  $1 \times 1$  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)

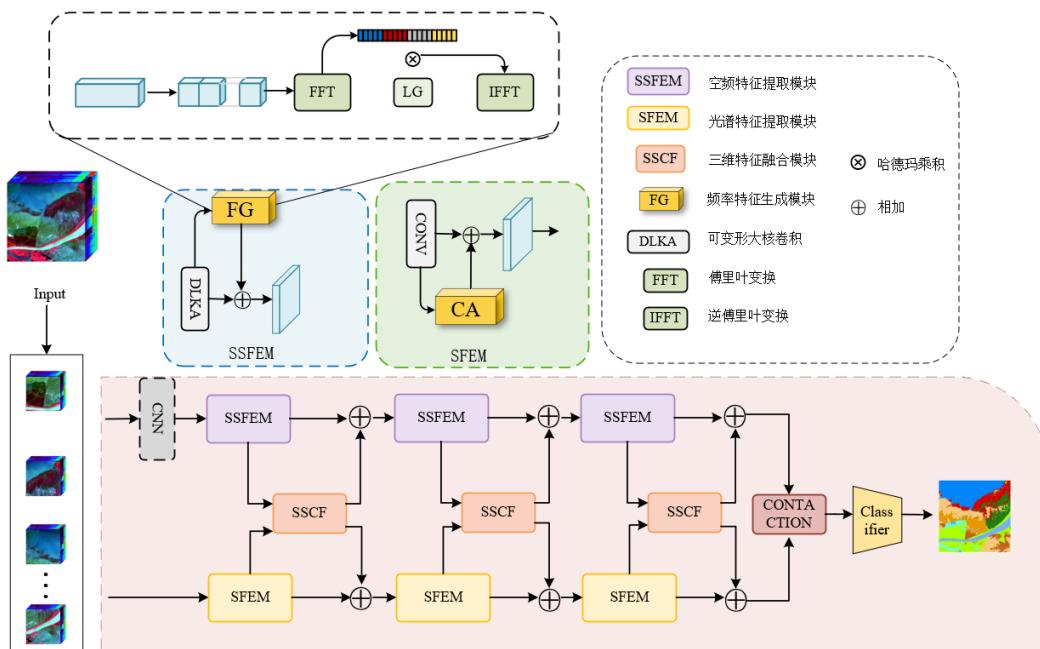
滨海湿地植被分类植被种类高度相似、光谱特征重叠以及环境干扰等挑战，导致部分类别的提取不完整或遗漏。针对这一问题，本章节提出了一种基于空谱协同注意力网络 (SSCANet) 的滨海湿地植被精细分类方法。SSCANet 采用多级特征蒸馏架构，结合频率域引导的多尺度特征提取与空谱协同融合机制，在复杂湿地环境中构建具有鲁棒性的深度特征表示，提高分类精度和环境适应性。

#### 3.2.1 AEHFNet 的整体架构

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 (MRFM)

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  $1 \times 1$  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：

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

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：

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

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。

图 3.3 光谱注意力增强模块结构

Fig.3.3 Structure of

### 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  $\Delta p$  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。

图 3.4 分组光谱残差模块结构

Fig.3.4 Structure of group spectral residual module

### 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 $\lambda$ :

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

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 (Experimental preparation)

#### 3.3.1 实验数据

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。

#### 3.3.2 实验环境

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。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

#### 3.3.3 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<sup>[11]</sup>、3DCNN、HybirdSN<sup>[33]</sup>、DDCD<sup>[71]</sup>、MorphFormer<sup>[48]</sup> 和 SSFTT<sup>[52]</sup> 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) DDCD：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。

### 3.3.4 实验参数设置

在训练阶段，损失函数选用 L2 normalized，同时设定批大小为 64，学习率及权重衰减分别为  $1e-4$  和  $1e-6$ 。为了自适应地调整参数学习率并加速网络权重的更新，采用了 Adam 优化器，初始学习率定为 0.003。训练迭代次数设置为 200 次，将空间大小设置为  $13 \times 13$ ，并使用交叉熵损失函数进行反向传播。为了保证结果的准确性，各方法的分类结果都被执行了 10 次，最终的分类结果是 10 次实验的平均值。

## 3.4 实验结果与分析 (Experimental results and analysis)

### 3.4.1 实验结果

IP 数据集的分类结果：表 3.6 显示了 AEHFNet 方法与其他方法的数值精度比较。该方法的 OA 值比 SVM 高 24.79%，SVM 没有利用空间邻域信息，鲁棒性较差，OA 值较低，仅为 70.69%。AEHFNet 对都是 3 个训练样本的 Alfalfa、Grass-past-mowed 和 Oats 的 OA 值分别达到 98.85%、98.56% 和 97.42%，表明 AEHFNet 能有效处理不平衡样本。可视化结果如图 3.7 所示。通过目视比较，AEHFNet 获得的分类图是最干净、最接近地面真值图的。

前几种 SVM、HybridSN、SSRN 和 FDSSC 方法存在明显的噪声。此外，每张图像中间的蓝色大方块是一个相对难以区分的区域，几乎所有的对比方法都是错误分类的。本文的方法在更大程度上识别了蓝色区域，这也证明了本文的方法的优异性能。主要是由于 AEHFNet 双分支空间光谱分层反馈分支的设计，促使细节信息可以和语义信息进行分层交互，通过两者的交互，模型能够更准确地识别地物的细微差别，从而改善分类效果。

图 3.7 YRD-NC12 数据集上的分类结果图  
Fig. 3.7 Classification maps for the YRD-NC12 dataset

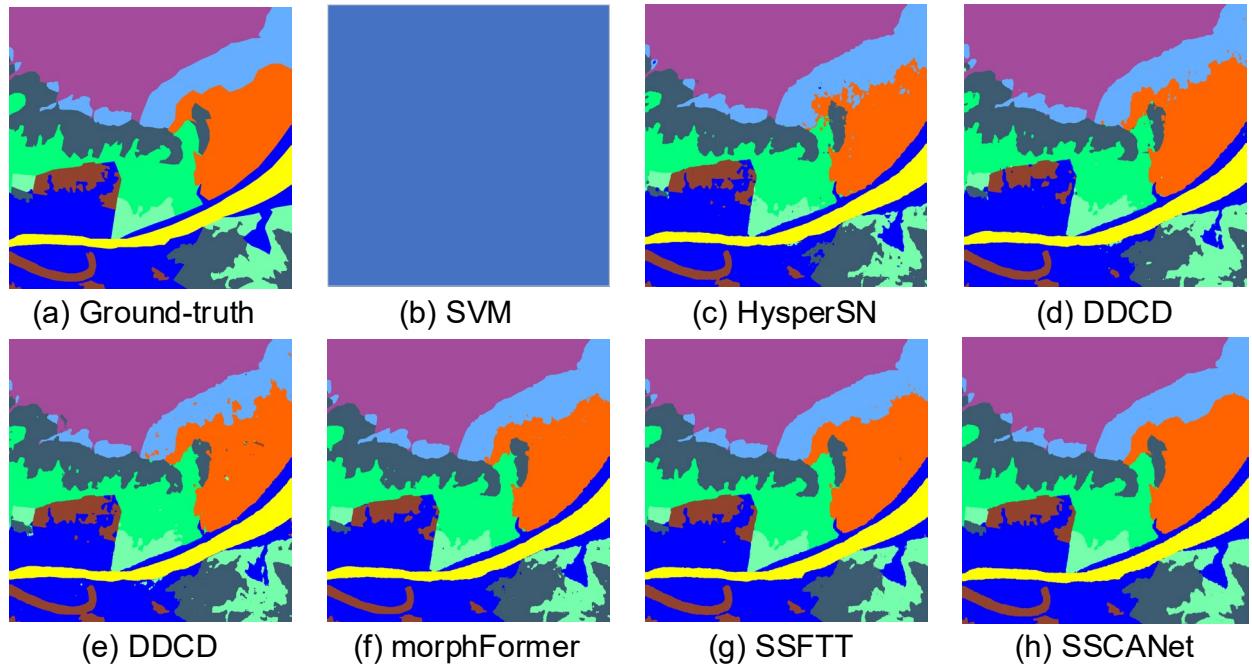


图 4.7 Coast 数据集上的分类结果图  
Fig. 4.7 Classification maps for the Coast dataset

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	DDCD	HybirdSN	MorphFormer	SSFTT	SSCANet
Suaeda glauca	98.34	98.68	98.05	99.68	99.19	99.55	99.52
Cement road	95.92	97.73	97.64	97.46	97.20	94.42	98.39
Asphalt road	96.29	97.07	95.95	96.63	97.45	99.01	98.99
Water	99.9	99.89	99.92	99.98	99.92	99.90	99.88
Stone	94.43	93.42	97.10	95.11	98.27	99.18	97.80
Hay	91.01	93.04	99.01	98.32	98.78	99.60	99.77
Iron	0	71.43	0.00	54.24	100.00	100.00	89.26
Tamarix	50.25	77.73	96.38	93.95	92.39	95.16	97.43
Withered reed	88.32	89.31	92.43	93.46	97.94	98.83	98.88
Reed	86.99	91.72	98.89	95.75	100.00	99.34	98.98
Spartina alterniflora	96.07	98.26	97.99	99.27	99.47	99.11	99.77
Moss	89.97	99.09	99.78	96.43	96.55	96.03	97.20
Dry soil	98.21	97.37	98.37	97.85	99.89	99.87	99.70
Moist soil	99.07	95.98	96.63	98.15	99.60	99.58	99.60
Mudflat	87.45	89.58	96.82	96.24	100.00	99.09	100.00
Standard reflectance cloth	41.43	98.57	100.00	31.12	91.67	98.61	87.98
OA (%)	98.19	97.88	98.50	98.79	99.28	99.52	99.61
AA (%)	82.1	85.81	81.70	85.89	91.74	96.63	97.37
Kappa	97.53	97.10	97.96	98.35	99.02	99.34	99.47

表 3.6 YRD-NC16 数据集上的分类结果

Table 3.6 Classification result of the YRD-NC16 dataset

PU 数据集上的分类结果：表 3.7 展示了 AEHFNet 方法与其他方法在数值精度方面的对比结果。与 IP 数据集不同，PU 的分类结果更好。主要原因是 PU 数据集的类别较少，而每个类别的样本数量较多。从表 3.7 可以清楚地看出，AEHFNet 在 OA、AA 和 Kappa 的值分别为 98.39%，98.05% 和 0.9887，三个指标都优于其他方法。其次是 SSFCN、DBDA 和 FECNet。AEHFNet 的 OA 比上述三种算法分别提升了 1.68%、3.31% 和 1.06%。另外，AEHFNet 在 9 个类别中的 6 个类别上均获得最佳结果，并且在 Meadow、Metal sheets 和 Bricks 等类别中都取得了 100% 的准确率，相比之下，SVM、HybirdSN 和 SSRN 等方法的 OA 较低。

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

第三个实验是在 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，其总体准确率高达 97.73%。与对比方法相比，SSCANet 网络在 AA、OA 和 Kappa 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 3.9 Salina 数据集上的分类结果  
Table 3.9 Classification result of the Salina dataset

类别	SVM	3DCNN	HybirdSN	DDCD	MoF	SSFTT	SSCANet
1	52.38	100	100	100.00	99.90	99.84	99.65
2	72.78	100	99.35	98.14	99.84	98.36	99.97
3	68.27	99.29	99.59	100.00	97.12	100.00	99.85
4	54.67	99.28	98.1	100.00	99.71	98.45	97.16
5	90.57	96.18	99.6	99.65	99.58	97.90	98.8
6	96.19	100	99.97	100.00	97.20	100.00	99.82
7	61.54	100	99.75	100.00	97.27	100.00	99.94
8	98.14	86.27	93.38	96.44	93.73	98.44	99.10
9	61.11	99.39	99.3	99.56	99.64	99.97	99.9
10	66.51	97.65	96.59	96.86	99.03	98.16	99.14
11	86.29	99.39	100	100.00	64.05	99.06	99.81
12	72.66	96.02	96.54	99.90	96.56	99.16	99.90
13	96.22	99.04	97.03	99.76	76.40	99.73	100.00
14	91.48	91.01	92.61	92.13	98.68	86.99	91.89
15	55.75	81.74	89.56	87.98	84.58	87.05	96.84
16	83.33	100.00	99.66	99.71	100.00	100.00	99.89
OA (%)	80.40	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	93.60	96.70	98.85

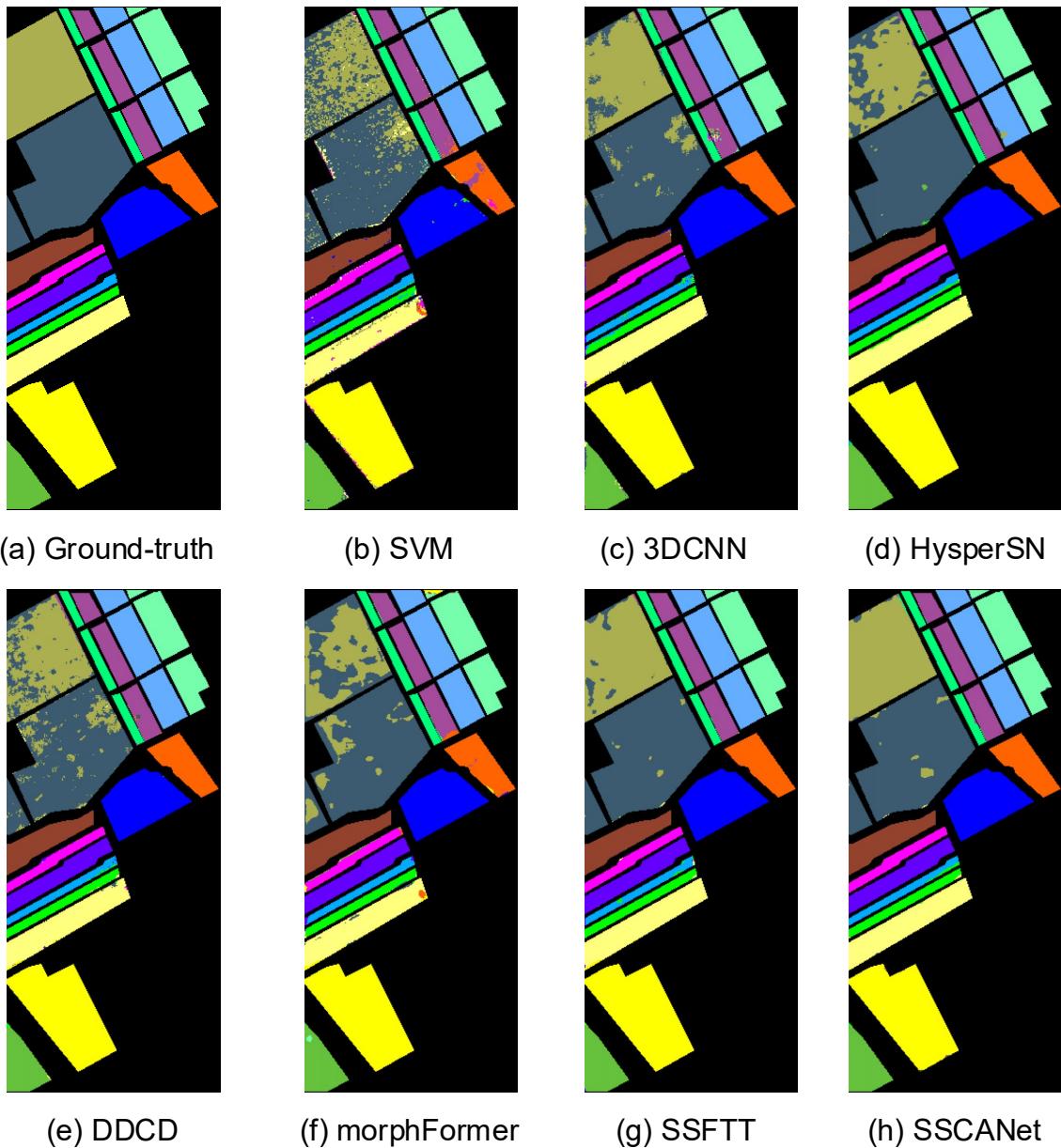


图 3.10 Salina 数据集上的分类结果  
Fig. 3.10 Classification maps for the Salin dataset

### 3.4.2 消融实验

### 3.4.3 训练样本数量对结果的影响

不同方法在不同数量样本训练比例下 OA 值比较：深度学习在一定程度上依赖于训练样本。因此，为了进一步验证所提方法的性能，本文比较了对比方法和本文提出的方法 AEHFNet 在不同训练样本比例下的分类性能。图 3.11 展示了四个数据集上不同训练比例下所有方法的 OA 值比较结果。这里，纵坐标表示 OA，横坐标表示训练样本的比例。AEHFNet 各训练比例的 OA 值都高于其他方法，即使在训练样本非常小的情况下，所提出的方法在不同训练样本百分比下也获得了最好的结果。

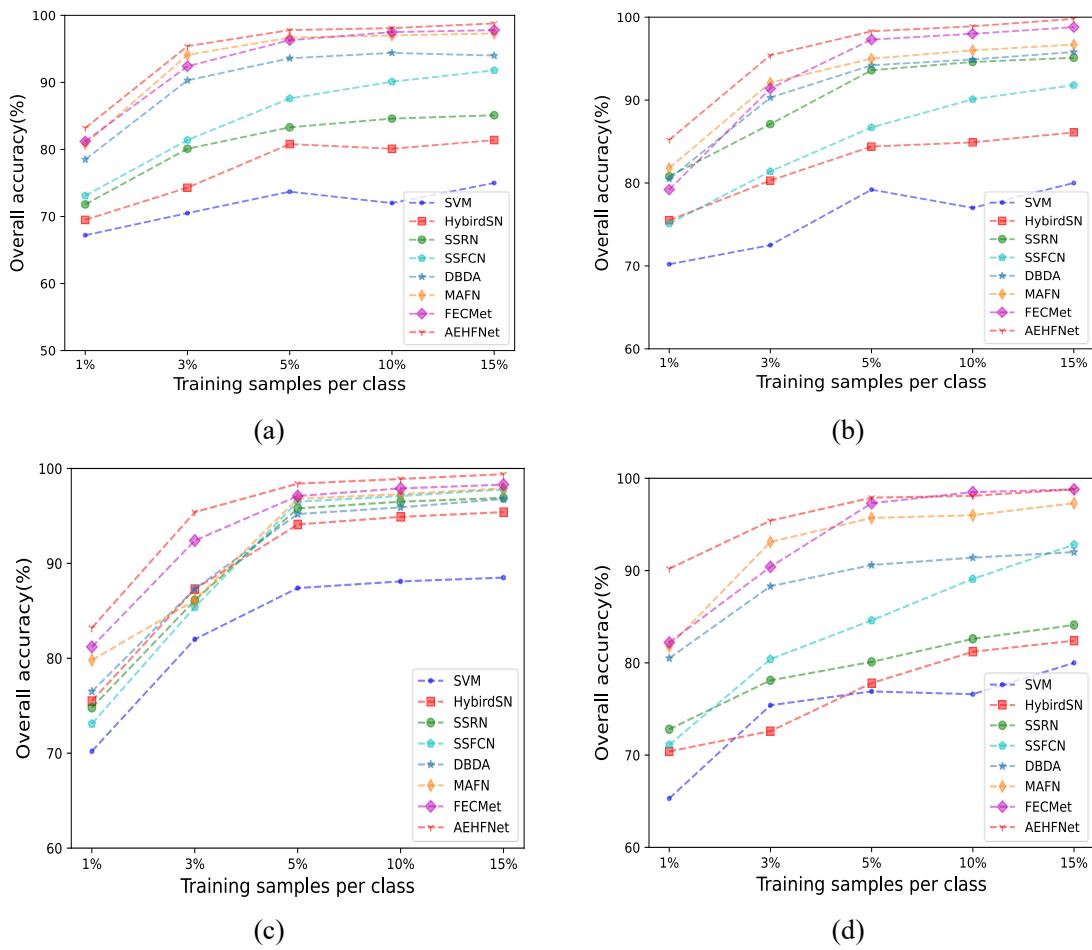


图 3.11 不同方法在不同数量训练样本下的准确率 (a)IP 数据集 (b)PU 数据集 (c)SV 数据集 (d) HongHu 数据集

Fig. 3.11 OA of different methods with different training samples percentages (a) IP dataset (b) UP dataset (c) SV dataset (d) HongHu dataset

#### 3.4.4 消融实验分析

为了验证所提出的 MRFM 模块、CRAM 模块和 ECSV 模块的有效性，在 AEHFNet 上通过移除特定组件进行消融实验以进行比较。表 3.11 显示了在 IP 数据集上的消融实验，当 AEHFNet 移除 MRFM 模块，OA 为 95.44%，性能下降了 2.04%。造成这一现象的原因是 MRFM 模块引入了多种尺度的空洞卷积，以捕获丰富的空间光谱信息，更有效地融合了不同尺度的信息。AEHFNet 移除 CRAM 时，OA、AA 和 Kappa 分别为 94.71%、93.82% 和 0.9279。主要原因是 CRAM、ECSV 多尺度注意力驱动的分层反馈能都使深层的语义信息引导细节信息，使提取的特征可以及时进行修正。AEHFNet 方法 OA、AA 和 Kappa 分别为 97.48%、96.33% 和 0.9713，性能达到了最优。

表 3.11 消融实验  
Table 3.11 Ablation experiments

方法	OA (%)	AA (%)	Kappa
AEHFNet w/o MRFM	95.44	94.66	0.9312
AEHFNet w/o CRAM	94.71	93.82	0.9279

AEHFNet w/o ECSV	95.84	95.21	0.9493
AEHFNet	<b>98.48</b>	<b>97.33</b>	<b>0.9713</b>

### 3.4.5 频率域引导空间特征提取模块有效性分析

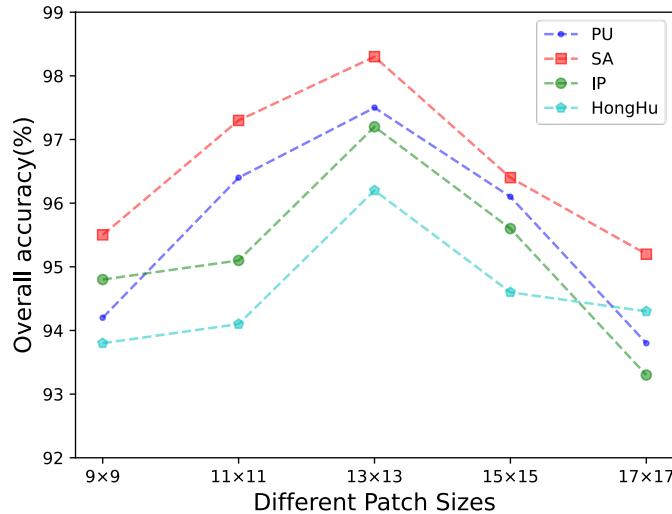


图 3.12 不同空间输入大小的对分类性能的影响

Fig.3.12 Effect of different spatial sizes on overall accuracies on four HSI datasets

## 3.5 本章小结 (Chapter Summary)

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，尤其在盐地碱蓬与柽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滨海湿地高光谱图像分类

### 4 Hyperspectral image classification method based on spatial-spectral features hierarchical fusion

#### 4.1 引言 (Foreword)

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。

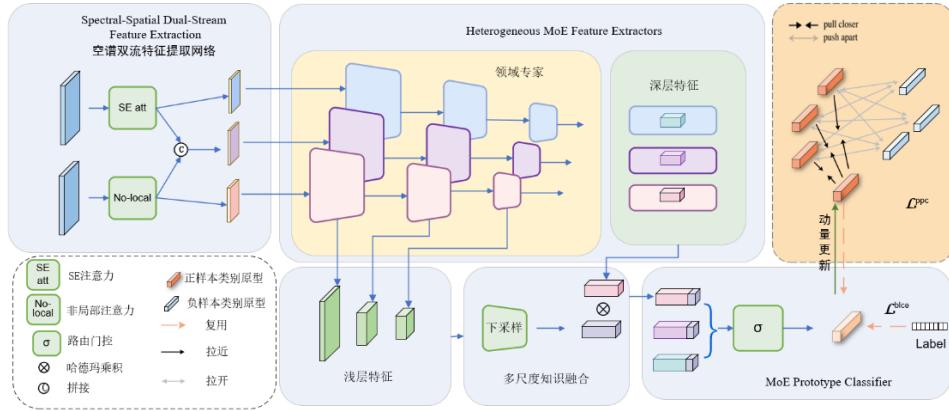


图 4.1 异构专家原型记忆网络整体架构  
Fig. 4.1 Overall architecture of HEPM-Net

#### 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

图 4.2 光谱可变形通道关系聚合模块结构  
Fig. 4.2 Structure of spectrum deformable channel relation aggregation module

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

标准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  $\alpha$  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：

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

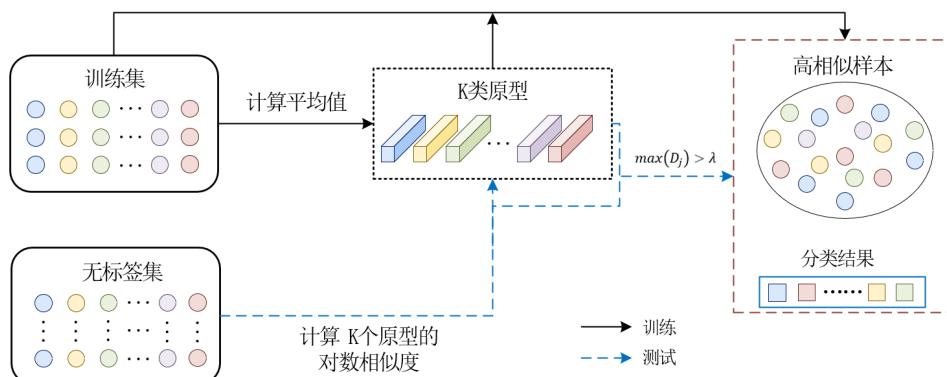


图 原型模块更新结构图

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  $\mathcal{S}(\cdot)$  计算特征与各类型原型的余弦相似度。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 (Experimental preparation)

#### 4.3.1 实验设置

~~在网络的训练过程中，参数更新采用了 Adam 优化器，初始学习率设为 0.0001，并设定权重衰减为  $1e-6$ 。同时，设定批次大小为 32，训练轮数为 200 次，损失函数则选用了多元交叉熵。~~

#### 4.3.2 实验数据

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 3%、1% 和 10% 样本作为训练集。

#### 4.3.3 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

#### 4.3.4 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]、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 (Experimental result and analysis)

### 4.4.1 实验结果

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别判别特征，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。

表 4.1 Indian Pines 数据集上的分类结果

Table 4.1 Classification result of the Indian Pines dataset

类别	SVM	HybirdSN	SSRN	DBDA	MSTNet	SSFTT	T-SSL	SSHNet
Alfalfa	21.34	60.23	64.88	86.36	86.98	86.54	97.85	<b>98.56</b>
Corn-N	47.43	44.38	64.01	91.83	95.25	96.63	96.31	<b>99.89</b>
Corn-M	50.96	80.42	89.65	95.35	96.19	97.32	98.81	<b>100</b>
Corn	35.59	50.21	40.85	86.40	90.77	95.51	99.21	<b>100</b>

Grass-P	64.79	82.88	73.64	99.35	97.25	<b>100</b>	98.99	98.56
Grass-T	96.34	97.47	94.05	97.71	97.98	98.42	95.82	<b>98.56</b>
Grass-P-M	10.36	40.35	50.23	96.29	77.77	96.29	97.16	<b>100</b>
Hay-windrowed	88.23	100	96.52	98.25	97.79	98.55	95.82	<b>100</b>
Oats	19.23	20.25	48.01	70.21	68.94	73.68	94.41	<b>98.77</b>
Soybean-notill	65.66	68.09	67.15	92.06	95.44	95.32	96.56	<b>100</b>
Soybean-mintill	79.24	90.71	93.78	97.49	96.05	95.62	94.91	<b>97.55</b>
Soybean-clean	59.32	73.46	78.19	88.40	95.29	93.06	98.12	<b>98.15</b>
Wheat	68.13	88.66	83.25	98.85	98.56	<b>100</b>	96.51	<b>100</b>
Woods	99.60	94.25	99.92	98.10	97.16	98.16	<b>100</b>	98.56
Buildings-G-T-D	50.26	83.24	91.09	85.44	88.82	97.45	98.45	<b>99.87</b>
Stone-S-T	84.78	27.17	60.89	69.66	79.63	84.68	<b>96.25</b>	95.41
OA (%)	73.69	79.89	83.35	96.79	96.62	97.31	96.48	<b>98.65</b>
AA (%)	55.77	61.93	67.49	91.79	92.90	94.11	95.33	<b>96.95</b>
Kappa	0.6643	0.7685	0.8087	0.9405	0.9614	0.9693	0.9613	<b>0.9846</b>

图 4.5 显示了在 IP 数据集上实验的所有方法的可视化分类结果。可以看出 SSHFNet 的分类图上虽然少数类别有噪声，但是绝大多数类别的分类结果比较好，且优于其他方法。虽然 MSTNet 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。SSHFNet 边缘分类错误问题有所缓解，主要是由于 SSHFNet 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 W/P Transformer 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。

图 4.5 Indian Pines 数据集上的分类结果图

Fig. 4.5 Classification maps for the Indian Pines dataset

表 4.2 表明了 7 种对比方法和 SSHFNet 在 PU 数据集上的实验结果。由表 4.2 可知，本文提出的方法 OA、AA 和 Kappa 值分别为 98.84%，98.27% 和 0.9746，优于比较方法，并且在 9 个类别中的 7 个类别上取得了最好的结果，其中在五个类别中取得了 100% 的分类性能。图 4.6 显示了 PU 数据集上所有方法的分类结果图。MSTNet 在此数据集上的表现比其他数据集差得多，OA 和 Kappa 仅为 91.69% 和 0.9159，说明 MSTNet 模型的泛化能力相对较弱。本文提出的模型以最小的噪声生成了分类结果图，并取得了令人满意的结果。表明本文方法由于提取了图像的多尺度全局和局部特征，因此在保持图像的空间一致性分布的同时，对边缘部分的分类较其他方法也更加准确。

表 4.2 Pavia University 数据集上的分类结果

Table 4.2 Classification result of the Pavia University dataset

类别	SVM	HybridSN	SSRN	DBDA	MSTNet	SSFTT	T-SSL	SSHFNet
Asphalt	91.68	88.62	93.23	95.16	94.23	93.84	97.21	<b>100</b>
Meadows	99.92	99.96	99.96	99.89	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Gravel	10.63	55.36	72.35	88.04	72.35	84.99	96.41	<b>98.91</b>
Tress	71.61	70.43	88.82	88.91	89.82	95.86	95.91	<b>97.01</b>

Metal sheets	99.63	99.89	99.93	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Bare soil	54.27	71.66	93.49	99.41	93.49	99.43	<b>100</b>	<b>100</b>
Bitumen	10.53	32.56	98.32	98.56	99.32	100	<b>100</b>	<b>100</b>
Bricks	35.23	35.79	73.95	94.14	73.95	95.45	97.21	<b>97.89</b>
Shadows	12.32	73.22	97.13	92.75	87.13	94.22	96.91	<b>98.49</b>
OA (%)	76.23	84.45	93.69	95.09	91.69	98.39	97.39	<b>98.84</b>
AA (%)	67.31	77.22	90.92	96.37	90.92	97.74	96.05	<b>98.27</b>
Kappa	0.5251	0.7892	0.9159	0.9537	0.9159	0.9687	0.9687	<b>0.9746</b>

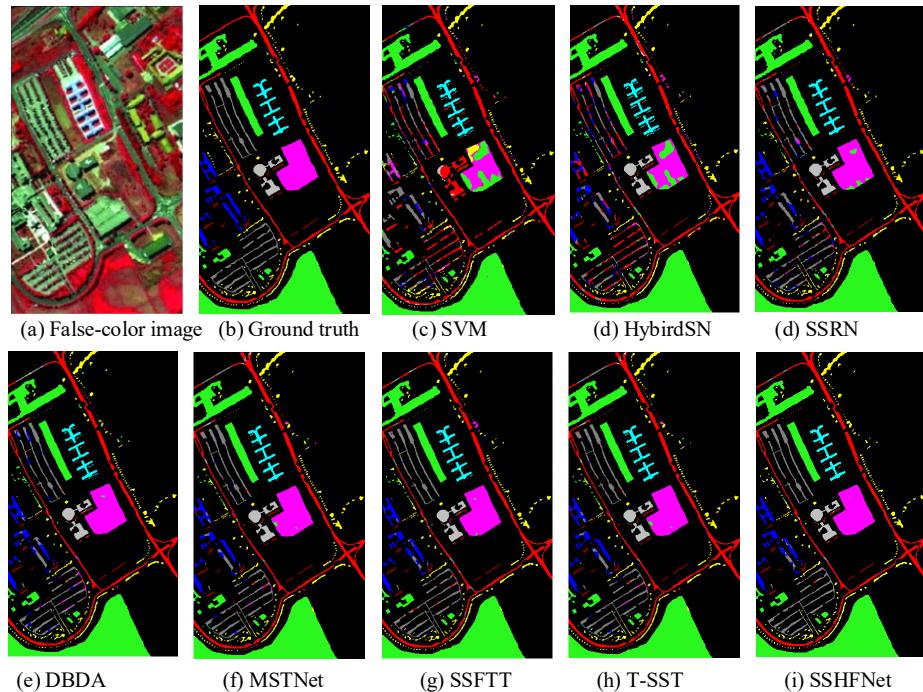


图 4.6 Pavia University 数据集上的分类结果图

Fig. 4.6 Classification maps for the Indian Pines dataset

表 4.3 表明了 7 种对比方法和 SSHFNet 在 SA 数据集上的实验结果。本文可以看到，本文的模型取得了最优的 OA、AA 和 Kappa，分别达到 98.75%、99.65% 和 0.9772。SSHFNet 的双分支空间和空谱特征分层融合的设计确保了信息的充分，从而获得了较高的分类准确率。图 4.7 为实验的可视化结果。本文的方法生成的分类图比比较方法噪声小，更接近地面真实图像，这说明了本文所提出的方法的优越性。其他方法，如 SVM、HybridSN 存在较大的噪声，主要因为 SVM 只利用了光谱特征，忽略了空间信息的重要性，HybridSN 只使用 3D 和 2D 卷积的组合。虽然它有一个很好的局部空间特征提取模型，但它缺乏对捕获长程光谱之间的关系的依赖性。而 MSTNet 方法，它只使用 Transformer 来构建网络，而没有考虑卷积对分类结果的影响，这可能会导致性能不佳。此外，可以发现几种方法错误分类的地物主要集中在 Vinyard-untrained 和 Grapes-untrained，两种地物的光谱曲线几乎相同，对分类造成了极大的难度，还有几种不同种植时间的 Lettuce-romaine 也存在相同的问题。所以，对光谱带相似地物进行分类要尽可能的考虑空间位置信息。SSHFNet 的分类结果在地物边缘处达到了更加平滑的效果。主要是因为 SSHFNet 着重考虑到了空间信息，通

过空间光谱双分支设计，空间分支利用 W/P Tranformer 强大的空间全局建模能力，缓解了边缘处分类易出错的问题。

表 4.3 Salinas Valley 数据集上的分类结果

Table 4.3 Classification result of the Salinas Valley dataset

类别	SVM	HybridSN	SSRN	DBDA	MSTNet	SSFTT	T-SSL	SSHNet
Brocoli- g-w-1	96.11	99.50	99.56	97.47	98.56	99.23	<b>100</b>	<b>100</b>
Brocoli- g-w-2	98.56	98.97	97.56	100	99.42	100	98.69	98.23
Fallow	99.21	99.65	99.42	99.64	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>

表 4.3 Salinas Valley 数据集上的分类结果（续）

Table 4.3 Classification result of the Salinas Valley dataset

Class	SVM	HybridSN	SSRN	DBDA	MSTNet	SSFTT	T-SSL	SSHNet
Fallow-r-p	82.84	98.56	99.64	98.84	98.85	97.84	97.20	<b>99.34</b>
Fallow-smooth	95.88	98.13	95.61	98.90	98.17	98.90	96.51	<b>100</b>
Stubble	99.67	99.63	99.72	99.89	100	100	<b>100</b>	<b>100</b>
Celery	99.32	99.91	99.38	99.23	98.91	<b>100</b>	<b>100</b>	<b>100</b>
Grapes-untrained	84.47	87.36	89.97	98.27	97.78	98.27	97.45	<b>98.80</b>
Soil-v-d	97.85	99.90	99.91	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Corn-s-g-w	93.61	93.57	97.46	99.04	99.41	<b>100</b>	<b>100</b>	<b>100</b>
Lettuce-r-4	98.98	99.87	99.34	<b>100</b>	99.81	99.01	<b>100</b>	<b>100</b>
Lettuce-r-5	97.14	95.11	98.69	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Lettuce-r-6	66.12	97.59	98.24	<b>99.88</b>	96.13	98.88	96.74	97.32
Lettuce-r-7	65.01	90.91	85.28	73.49	99.05	93.49	97.81	<b>100</b>
Vinyard-untrained	50.25	83.45	85.44	90.64	94.39	95.64	94.84	<b>96.43</b>
Vinyard-v-t	97.51	98.33	99.73	99.44	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
OA (%)	87.43	94.06	95.82	96.86	95.24	96.86	97.45	<b>98.75</b>
AA (%)	89.29	96.24	96.23	97.37	96.73	97.37	96.23	<b>99.65</b>
Kappa	0.8597	0.9338	0.9432	0.9757	0.9469	0.9757	0.9716	<b>0.9772</b>

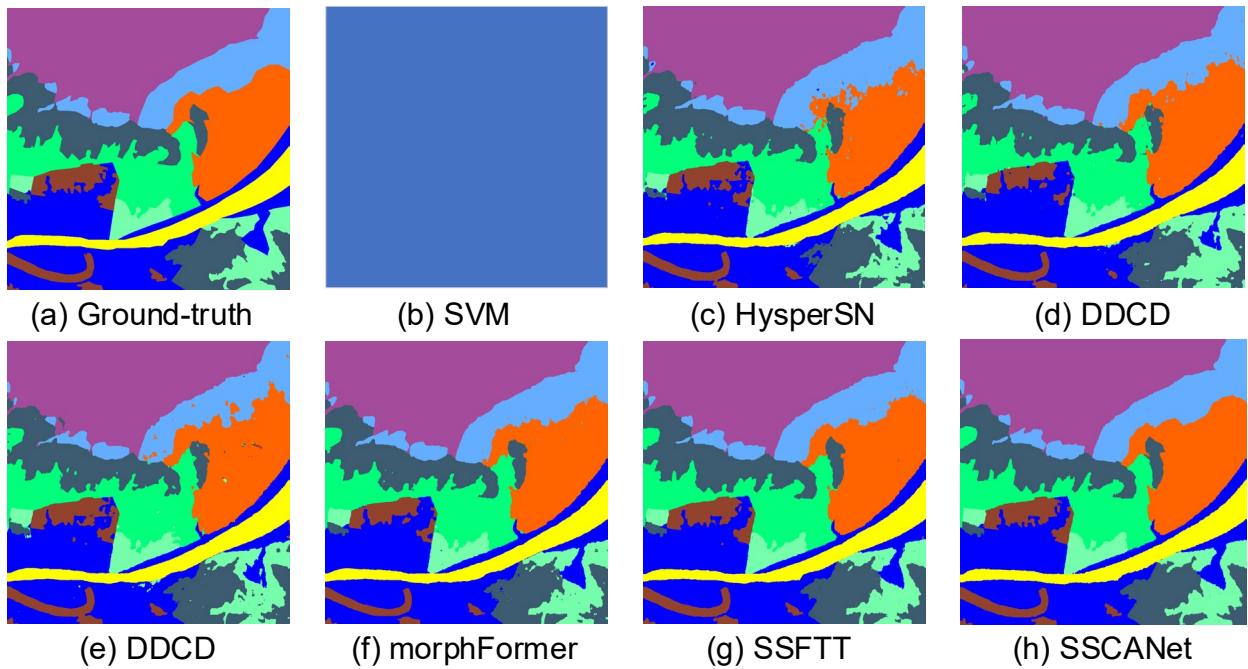


图 4.7 Coast 数据集上的分类结果图  
Fig. 4.7 Classification maps for the Coast dataset

表 4.4 表明了 7 种对比方法和 SSHFNet 在 HongHu 数据集上的实验结果。分类结果的定量结果和可视化结果分别见表 4.4 和图 4.8。从表 4.4 中可以观察到，SSHFNet 的 OA 值达到 98.80%，与 SVM、DBDA、MSTNet 和 AEHFNet 相比，分类准确率有了明显的提高，总体分类精度分别高了 20.32%、6.15%、1.9% 和 2.79%。从图 4.8 可以看出，较多的椒盐噪声呈现在图 4.8(c)中，这是由于 SVM 只考虑了光谱信息，忽略了空间信息。SSRN 和 DBDA 使用光谱和空间特征，这些方法在分类精度方面优于 SVM 和 HybridSN。总之，SSHFNet 得到的分类图最接近真实的地表覆盖。此外，通过观察所有算法的多次实验结果发现提出方法的实验结果相对稳定，这证明了 SSHFNet 具有更强的鲁棒性和泛化能力。

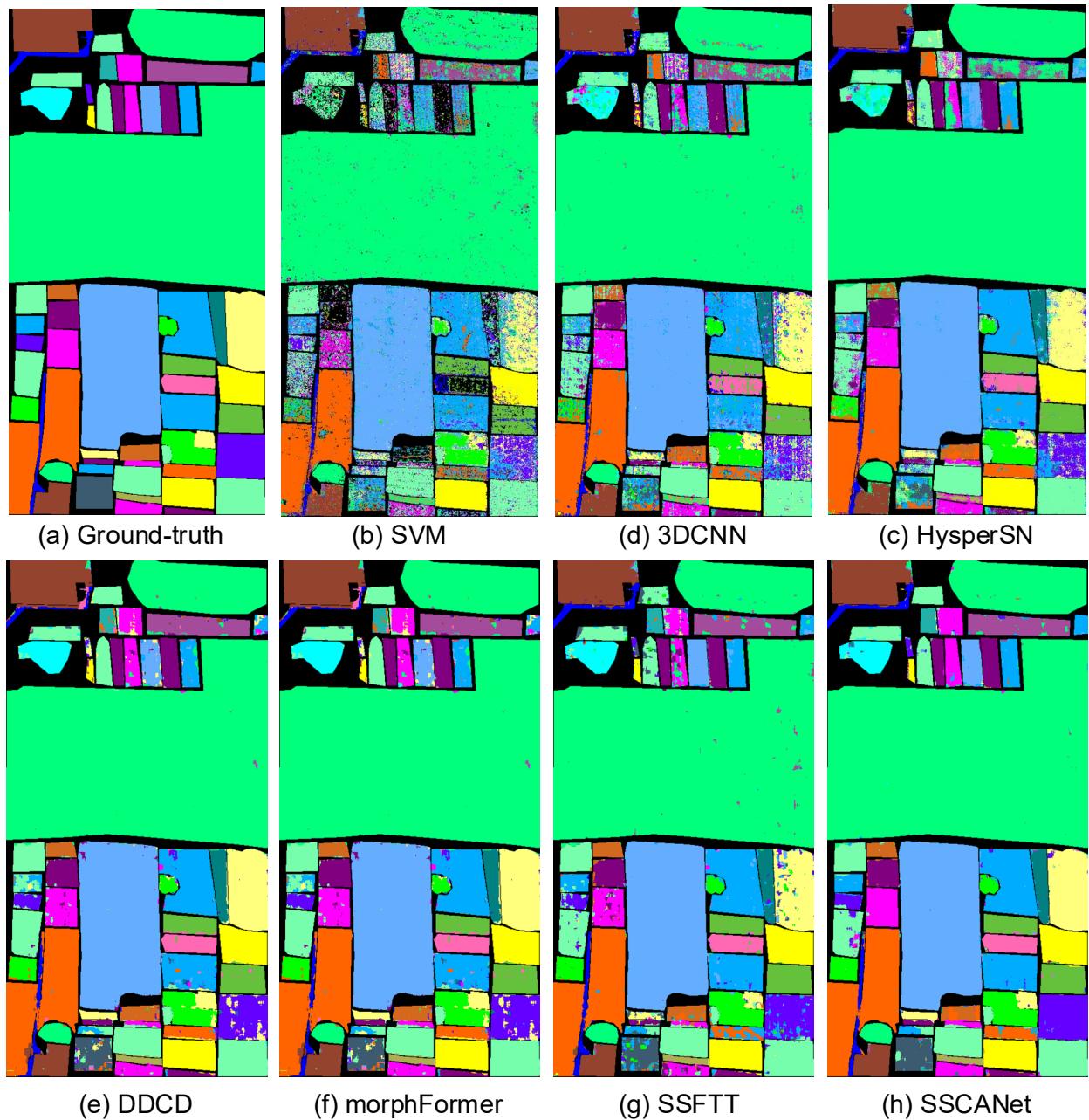


图 4.7 WHU-Hi-HongHu 数据集上的分类结果图

Fig. 4.7 Classification maps for the WHU-Hi-HongHu dataset

表 4.4 WHU-Hi-HongHu 数据集上的分类结果

Table 4.4 Classification result of the WHU-Hi-HongHu dataset

类别	SVM	HybridSN	SSRN	DBDA	MSTNet	SSFTT	T-SSL
1	90.95	94.98	95.12	95.22	95.40	97.64	98.65
2	70.42	73.82	78.81	71.75	90.93	87.28	91.56
3	87.89	83.74	84.37	91.55	93.78	95.63	96.75
4	97.86	96.64	95.86	98.66	99.46	98.99	99.64
5	48.57	68.30	90.06	75.44	96.50	90.47	94.78
6	88.96	89.70	93.66	97.15	99.25	96.93	99.01

7	71.57	78.74	80.09	82.86	88.79	88.86	93.39
8	14.53	48.51	83.57	72.68	93.29	73.23	93.73
9	91.44	98.43	97.92	90.55	97.61	97.43	97.59
10	55.36	79.68	81.36	75.83	89.19	95.96	97.99
11	42.90	76.68	80.12	88.60	95.75	96.36	94.41
12	48.87	59.61	84.26	74.81	89.61	91.93	92.07
13	62.60	66.43	79.33	88.32	96.84	88.24	96.18
14	59.56	79.28	95.42	85.13	97.06	92.10	98.10
15	41.79	91.13	99.50	74.65	91.64	94.86	93.15
16	81.41	91.62	95.49	71.52	97.49	96.63	98.48
17	57.62	84.23	88.95	43.70	85.28	94.09	95.80
18	41.26	78.58	87.81	85.34	88.67	91.76	94.96
19	67.19	84.84	88.44	94.45	93.71	91.28	97.27
20	57.27	72.49	87.10	79.63	89.97	86.06	90.33
21	13.54	29.11	67.12	65.83	92.42	74.94	82.27
22	36.47	80.66	92.96	93.11	97.12	95.13	95.67
OA(%)	82.22	86.67	91.09	91.88	96.13	95.57	97.73
AA(%)	60.36	74.72	74.47	79.49	82.13	90.35	95.02
$\kappa \times 100$	77.32	83.98	88.62	89.72	93.80	94.39	97.14

#### 4.4.2 W/P Transformer 性能分析

本次对比实验探讨了 W/P Transformer 与 Swin Transformer 的训练时间和分类性能方面的表现。通过观察表 4.5，可以发现，尽管两者在分类精度上相近，展现出了相当的性能水平，但在四个数据集上的训练时间上，W/P Transformer 比 Swin Transformer 要缩短 3.42min，展现出了明显的优势。主要原因是 W/P Transformer 采用了基于池化窗口的多头自注意力模块，在保持优越分类性能的同时缩短模型的训练时间。这意味着在相同的条件下，W/P Transformer 在保持较高分类精度的同时，还能更高效地完成模型的训练过程。

表 4.5 Comparison between W/P Transformer and Swin Transformer

Table 4.5 W/P Transformer 和 Swin Transformer 对比

数据集	IP 数据集		PU 数据集		SV 数据集		HongHu 数据集	
	评估指标	Training (min)	OA (%)	Training (min)	OA (%)	Training (min)	OA (%)	Training (min)
W/P Transformer	<b>19.21</b>	98.65	<b>16.51</b>	<b>98.84</b>	<b>13.41</b>	98.75	<b>21.62</b>	<b>98.80</b>
Swin Transformer	22.63	<b>98.97</b>	19.59	98.56	17.32	<b>99.01</b>	24.69	97.81

#### 4.4.3 计算时间和参数量分析

表 4.6 展示了不同方法运行时间和参数量比较。从表中可以观察到，SSRN 的参数量

相对较高，这主要源于其网络结构较深，每次迭代都需要一个较大的计算周期。这种高参数量的设计也带来了一些问题。它不仅增加了模型的训练成本，需要更多的计算资源和时间来训练模型，还可能增加了过拟合的风险，尤其是在训练样本有限的情况下。FECNet 则通过采用基于展开卷积的模型设计，显著减少了参数量。这种轻量级的网络设计使得 FECNet 在保持较好性能的同时，降低了模型的复杂度，提高了计算效率。然而，运行时间作为另一个重要的性能指标，虽然 FECNet 的参数数量适中，但它的运行时间相对较长，这可能影响了它们在实时应用中的性能。SSFTT 结构简单，训练和测试时间都相对较短。AEHFNet 在参数量和运行时间之间取得了良好的平衡。它在保持适度参数量的同时，实现了较短的运行时间。然而，SSHFNNet 的运行时间和参数量稍逊于 AEHFNet，主要是由于 SSHFNNet 设计了轻量级 W/P Transformer，虽然在一定程度上减少了参数量。然而，由于 SSHFNNet 引入了光谱可变形卷积，通过在每个卷积核的参数上添加一个方向向量（或称为偏移量），这在一定程度上增加了参数量。

表 4.6 不同方法计算时间和参数量比较

Table 4.6 Comparison of different methods in terms of running time(s), and parameters required(k)

数据集	指标	SSRN	DBDA	FECNet	MSTNet	SSFTT	AEHFNet	SSHFNNet
IP	Training time (min)	26.36	21.56	28.47	27.45	<b>16.45</b>	<u>18.12</u>	19.21
	Testing time (s)	3.85	6.95	4.12	5.01	<u>2.79</u>	2.85	<b>2.45</b>
	Params (M)	0.45	0.38	0.41	0.35	<b>0.25</b>	<u>0.31</u>	0.38
PU	Training time (min)	24.23	18.59	24.35	26.31	<b>14.13</b>	<u>15.58</u>	16.51
	Testing time (s)	2.31	3.85	6.13	5.32	4.52	<b>4.01</b>	<u>4.31</u>
	Params (M)	0.65	0.52	0.42	0.61	<u>0.28</u>	<b>0.21</b>	0.32
SV	Training time (min)	27.56	20.89	26.28	27.21	15.20	<u>14.46</u>	<b>13.41</b>
	Testing time (s)	8.14	6.51	8.71	9.62	<b>3.11</b>	4.41	<u>3.41</u>
	Params (M)	0.38	0.62	0.39	0.47	0.31	<b>0.23</b>	<u>0.27</u>
HongHu	Training time (min)	25.46	24.18	27.23	26.21	23.34	<b>20.56</b>	<u>21.62</u>
	Testing time (s)	6.04	9.61	6.19	6.13	5.42	<u>4.42</u>	<b>4.13</b>
	Params (M)	0.47	0.54	0.51	0.47	0.53	<b>0.34</b>	<u>0.44</u>

#### 4.4.3 超参数分析

图 4.9 展示了输入不同输入空间大小和 PCA 数量对分类性能的影响。从图中可以清晰地观察到，随着输入空间大小尺寸的逐渐增大，局部空间特征的提取量也相应增加，从而使得初始的 OA 显著上升。然而，当输入空间过大时，会引入过多的干扰像素，导致 OA 逐渐下降。同时，图 4.9 也反映了 PCA 数量对 OA 的影响。正如预期，随着 PCA 数量的增加，OA 总体上呈现出上升趋势，因为前几个主成分通常包含了大量的有用信息。然而，随着 PCA 数量的进一步增加，由于光谱信息的冗余以及学习参数的增多，OA 会略有下降。总的来说，图 4.9 不仅揭示了输入空间大小与 OA 之间的复杂关系，还展示了 PCA 数量对 OA 的影响趋势，为进一步优化特征提取和分类性能提供了有价值的参考。

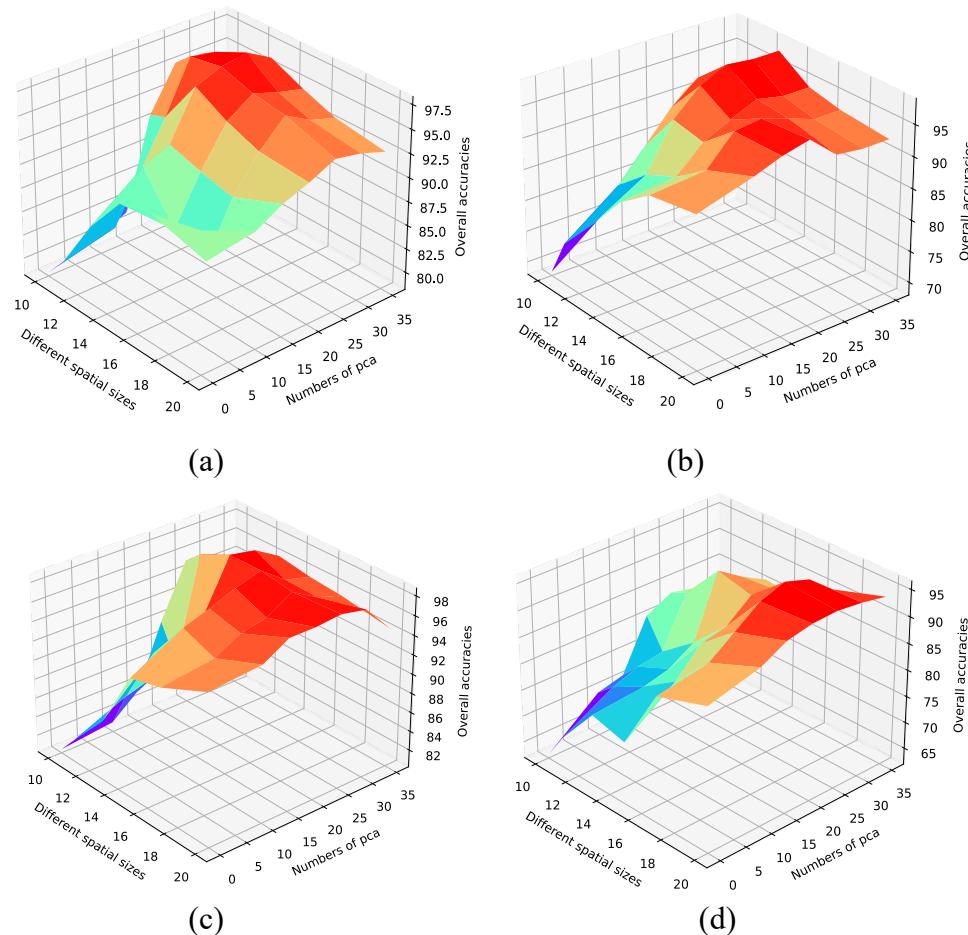


图 4.9 不同输入空间大小和 PCA 数量对性能的影响 (a) IP 数据集 (b) PU 数据集 (c) SV 数据集 (d) HongHu 数据集

Fig. 4.9 Effect of different input space size and PCA numbers on performance (a) IP dataset (b) UP dataset (c) SV dataset (d) HongHu dataset

#### 4.4.3 空谱特征分层融合有效性分析

为了进一步验证光谱补偿机制的有效性，本文在 PU 数据集上对比了进行空谱特征融合和未进行空谱特征融合两种情况下混淆矩阵的可视化结果，分别如图 4.10 (a) 和 4.10 (b) 所示。在这两张图中，纵坐标表示真实的类别标签，横坐标表示预测的类别标签，对角线

表示预测的类别标签与真实的类别标签相同的概率。通过观察和对比，相较于图 4.10 (a) 中空谱特征融合后的准确分类效果，图 4.10 (b) 中未进行空谱特征融合的分类结果存在更多的错误分类情况。这一显著差异直观地展现了空谱特征融合机制在改善分类精度方面的积极作用。

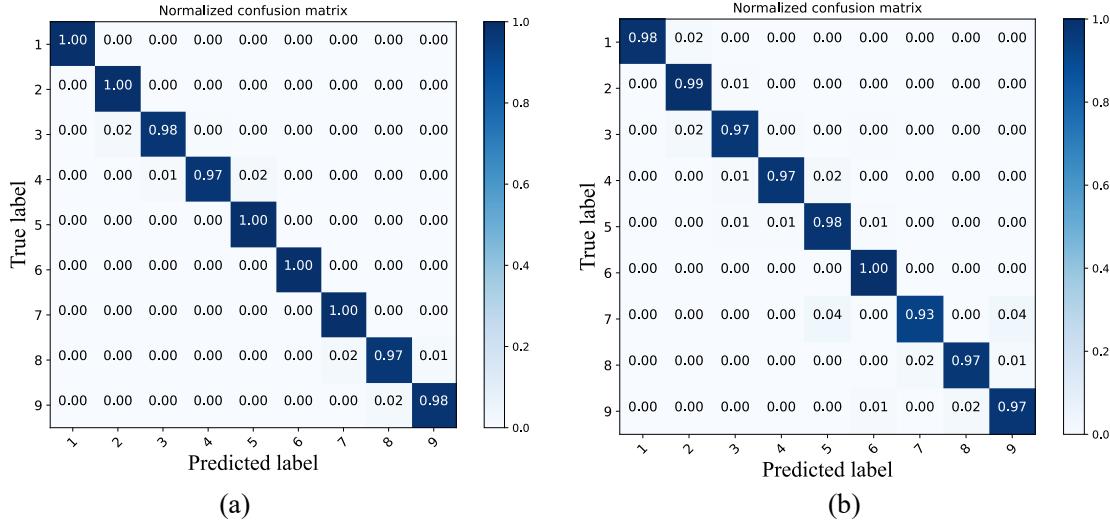


图 4.10 混淆矩阵可视化结果 (a)未进行空谱特征融合 (b)未进行空谱特征融合

Fig.4.10. Visualization results of the confusion matrix (a) With spatial-spectral feature fusion (b) Without spatial-spectral feature fusion

## 4.5 本章小结 (Chapter Summary)

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 Conclusion and prospects

#### 5.1 总结 (Conclusion)

随着航空航天技术的突飞猛进，传感器技术以及高光谱成像技术的日新月异，对地观测系统正经历着前所未有的变革与完善。高光谱图像作为一种能够同时捕捉地物对象光谱信息和空间信息的成像方式，展现出其信息丰富以及图谱高度融合的独特优势。然而，高光谱图像中蕴含的地物信息种类繁多、错综复杂，这也给其后续的数据处理和识别工作带来了极大的挑战，主要表现在：第一，高光谱数据本身存在的高维特性导致的数据冗余，模型训练参数量大以及训练和推理的时间较长也是高光谱图像分类的一大挑战。第二，由于大气散射和周围像元的影响，这些因素导致反射率值发生变化，进而引发高光谱图像的类内变异性。这种类内变异性降低了图像的对比度，使得地物的边缘变得模糊不清，这被称为“邻近像元效应”。模糊的地物边缘给传统的卷积神经网络带来了极大的困扰。

(1) 针对高光谱图像数据高维度导致的高冗余，模块参数量大以及训练和推理时间较长的问题。首先，本文提出了一种基于注意力驱动的分层反馈网络 AEHFNet 用于高光谱图像分类。AEHFNet 鼓励模型获得及时的反馈并进行调整，使模型感受到及时持续的正反馈，从而保持模型学习的动力和积极性。通过不断的反馈和调整，提高学习效率和分类性能。此外，高效的上下文自注意模块融合了金字塔采样策略和非局部块的特性，在充分考虑远程依赖关系的情况下融合了不同上下文信息，从而大大提高了性能，降低了计算成本。最后，AEHFNet 引入了多感受野模块，该模块由不同扩展率的膨胀卷积和全局平均池化组成，以实现不同感受野的叠加效果。

(2) 针对大气散射和周围像元导致的“邻近像元效应”的问题。首先，本文提出了一种基于空谱特征分层融合的网络 SSHFNet 用于高光谱图像分类。SSHFNet 使用轻量级 W/P Transformer 进行空间特征的建模，使得模型能够更好地理解和处理远距离的语义关系，从而缓解边界模糊的问题。其次，光谱分支部分设计了光谱可变形关系聚合模块，该模块融合了软池化和光谱可变形卷积，一方面，通过学习和预测卷积核的偏移量，它可以更好地捕捉光谱数据中的局部变化和细节信息。另一方面，软池化能够减少特征表示中的信息损失和失真。最后，提取的光谱特征在不同层级上和空间特征进行融合，以解决光谱与空间信息融合不足的问题。

#### 5.2 展望 (Prospects)

在当前中国遥感技术持续发展的热潮中，高分系列卫星的广泛应用已经深入人们的日常生活，为众多领域提供了丰富的数据信息。在这样的背景下，高光谱图像分类技术作为遥感信息提取的重要手段，正受到越来越多的关注和重视。因此，下一步的研究计划将从

以下几个方面出发：

- (1) 考虑到高光谱图像往往伴随着其他传感器获取的数据，如激光雷达，未来的研究可以探索将多模态信息融合到网络中，以提升特征提取的效率和分类性能。
- (2) 进一步优化 SSHFNet 的结构和算法，以更好地解决边界模糊问题。考虑引入更加有效的边界信息提取和整合机制，以减少由大气散射和周围像元引起的影响，提高分类结果的准确性和鲁棒性。
- (3) 深入研究光谱信息和空间信息的融合策略，进一步优化光谱分支和空间特征的交互方式，以提高模型对高光谱图像的全局理解能力和分类性能。

## 参考文献

- [1] 李海峰, 郭科. 对地观测技术的发展历史, 现状及应用[J]. 测绘科学, 2010, 35(6): 262-264.
- [2] 张兵. 当代遥感科技发展的现状与未来展望[J]. 中国科学院院刊, 2017, 7: 013.
- [3] T. Wilson and R. Felt. Hyperspectral remote sensing technology (HRST) program[C]. Proceedings of 1998 IEEE Aerospace Conference Proceedings, 1997, 5:193–200.
- [4] 杨哲海, 韩建峰, 宫大鹏, 等. 高光谱遥感技术的发展与应用[J]. 海洋测绘, 2003, 23(6).
- [5] Liang Y, Wu Y, Wang G, et al. Agricultural Hyperspectral Image Classification Based on Deep Separable Convolutional Neural Networks[C]. International Conference on Simulation Tools and Techniques. Springer, Cham, 2022: 403-420.
- [6] Zhou Y, Wetherley E B, Gader P D. Unmixing urban hyperspectral imagery using probability distributions to represent endmember variability[J]. Remote Sensing of Environment, 2020, 246: 111857.
- [7] 王树泽, 张志华, 邓砚学. 基于机载高光谱遥感图像的城市绿地覆盖研究[J]. 激光杂志, 2022, 43(02): 77-81.
- [8] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [9] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.
- [10] He L, Li J, Liu C, et al. Recent advances on spectral–spatial hyperspectral image classification: An overview and new guidelines[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 56(3): 1579-1597.
- [11] Melgani and L. Bruzzone, Classification of hyperspectral remote sensing images with sppuort vector machines, IEEE Trans. geosci. Remote Sens. vol. 42, no. 8, pp. 1778-1790, Aug. 2004.
- [12] A. Samat, P. Du, S. Liu, J. Li, and L. Cheng, E2LMs: Ensemble extreme learning machines for hyperspectral image classification, IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. vol. 7, no. 4, pp. 1060-1069, Apr. 2014.
- [13] J. Li, J. M. Bioucas-Dias, and A. Plaza, Spectral-spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields, IEEE Trans. Geosci. Remote Sens. vol. 50, no. 3, pp. 809-823, Mar. 2012.
- [14] Qiu Y, Song Z, Chen Z. Short-term stock trends prediction based on sentiment analysis and machine learning[J]. Soft Computing, 2022, 26(5): 2209-2224.
- [15] Ham J, Chen Y, Crawford M M, et al. Investigation of the random forest framework for classification of hyperspectral data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2005, 43(3): 492-501.
- [16] G. Licciardi, P. R. Mar, J. Chanussot, and J. A. Benediktsson, Linear versus nonlinear PCA for the classification of hyperspectral data based on the extended morphological profiles, IEEE Geosci. Remote Sens. Lett. vol. 9, no. 3, pp. 447-451, May 2012.

- [17] Li J, Liu M, Ma D, et al. Learning shared subspace regularization with linear discriminant analysis for multi-label action recognition[J]. *The Journal of Supercomputing*, 2020, 76: 2139-2157.
- [18] A. Villa, J. A. Benediktsson, J. Chanussot, and C. Jutten, Hyperspectral image classification with independent component discriminant analysis, *IEEE Trans. Geosci. Remote Sens.* vol. 49, no. 12, pp. 4865-4876, Dec. 2011.
- [19] Y. Qian, M. Ye, and J. Zhou, "Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features," *IEEE Trans. Geosci. Remote Sens.* vol. 51, no. 4, pp. 2276-2291, Apr. 2013.
- [20] L. Shen and S. Jia, Three-dimensional Gabor wavelets for pixel-based hyperspectral imagery classification, *IEEE Trans. Geosci. Remote Sens.* vol. 49, no. 12, pp. 5039-5046, Dec. 2011.
- [21] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.* vol. 43, no. 3, pp. 480-491, Mar. 2005.
- [22] 蒲生亮. 高光谱图像深度学习分类模型研究[J]. *测绘学报*, 2023, 52(1): 172.
- [23] Barbedo J G A. A review on the combination of deep learning techniques with proximal hyperspectral images in agriculture[J]. *Computers and Electronics in Agriculture*, 2023, 210: 107920.
- [24] Sellami A, Tabbone S. Deep neural networks-based relevant latent representation learning for hyperspectral image classification[J]. *Pattern Recognition*, 2022, 121: 108224.
- [25] Khan A, Vibhute A D, Mali S, et al. A systematic review on hyperspectral imaging technology with a machine and deep learning methodology for agricultural applications[J]. *Ecological Informatics*, 2022, 69: 101678.
- [26] Rajendran T, Valsalan P, Amutharaj J, et al. Hyperspectral image classification model using squeeze and excitation network with deep learning[J]. *Computational Intelligence and Neuroscience*, 2022, 2022.
- [27] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* vol. 7, no. 6, pp. 2094-2107, Jun. 2014.
- [28] Y. S. Chen, X. Zhao, and X. Jia, "Spectral-spatial classification of hyperspectral data based on deep belief network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* vol. 8, no. 6, pp. 2381-2392, Jun. 2014 .
- [29] Chang Y L, Tan T H, Lee W H, et al. Consolidated convolutional neural network for hyperspectral image classification[J]. *Remote Sensing*, 2022, 14(7): 1571.
- [30] Lu Z, Liang S, Yang Q, et al. Evolving block-based convolutional neural network for hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-21.
- [31] Paoletti M E, Moreno-Álvarez S, Xue Y, et al. AAtt-CNN: Automatical attention-based convolutional neural networks for hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023.
- [32] Anand R, Khan B, Nassa V K, et al. Hybrid convolutional neural network (CNN) for kennedy space center hyperspectral image[J]. *Aerospace Systems*, 2023, 6(1): 71-78..
- [33] Cao X, Yao J, Xu Z, et al. Hyperspectral image classification with convolutional neural network and active learning[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(7): 4604-4616.

- [34] Bera S, Shrivastava V K. Analysis of various optimizers on deep convolutional neural network model in the application of hyperspectral remote sensing image classification[J]. International Journal of Remote Sensing, 2020, 41(7): 2664-2683.
- [35] Jijón-Palma M E, Kern J, Amisse C, et al. Improving stacked-autoencoders with 1D convolutional-nets for hyperspectral image land-cover classification[J]. Journal of Applied Remote Sensing, 2021, 15(2): 026506-026506.
- [36] Yu C, Han R, Song M, et al. A simplified 2D-3D CNN architecture for hyperspectral image classification based on spatial-spectral fusion[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 2485-2501.
- [37] Li W, Chen H, Liu Q, et al. Attention mechanism and depthwise separable convolution aided 3DCNN for hyperspectral remote sensing image classification[J]. Remote Sensing, 2022, 14(9): 2215.
- [38] 刘金香, 班伟, 陈宇等. 融合多维度 CNN 的高光谱遥感图像分类算法[J]. Chinese Journal of Lasers, 2021, 48(16): 1610003.
- [39] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016 ,770-778.
- [40] 颜铭靖, 苏喜友. 基于三维空洞卷积残差神经网络的高光谱影像分类方法[J]. Acta Optica Sinica, 2020, 40(16): 1628002.
- [41] Zhong, Z.; Li, J.; Luo, Z.; Chapman, M.; Weinberger, Q. Spectral-spatial residual network for hyperspectral image classification: a 3-D deep learning framework. ieee trans. geosci. remote sens. 2017, 56, 847-858.
- [42] Huang G, Liu Z, Laurens V, et al. Densely Connected Convolutional Networks[J]. IEEE Computer Society, 2016, 2261-2269.
- [43] Wang, W.; Dou, S.; Jiang, Z.; Sun, L. A Fast Dense Spectral-Spatial Convolution Network Framework for Hyperspectral Images Classification. remote Sens. 2018, 10, 1068.
- [44] Y. Cui, Z. Yu, J. Han, S. Gao, and L. Wang, Dual-triple attention network for hyperspectral image classification using limited training samples, IEEE Geosci. Remote Sens. Lett., vol. 19, pp. 1–5, 2022.
- [45] Zhai H, Liu Y. Hyperspectral Image Classification based on Atrous Convolution Channel Attention Aided Dense Convolutional Neural Network[J]. IEEE Geoscience and Remote Sensing Letters, 2024.
- [46] Zhao J, Wang J, Ruan C, et al. Dual-Branch Spectral-Spatial Attention Network for Hyperspectral Image Classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024.
- [47] Zhu, M.; Jiao, L.; Liu, F.; Yang, S.; Wang, J. Residual spectral-spatial attention network for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 2020, 59, 449–462.
- [48] R. Li, S. Zheng, C. Duan, Y. Yang, and X. Wang, Classification of hyperspectral image based on double-branch dual-attention mechanism network, Remote Sens. vol. 12, no. 3, p. 582, Feb. 2020.
- [49] Y. Xu, B. Du and L. Zhang, Self-Attention Context Network: Addressing the Threat of Adversarial Attacks for Hyperspectral Image Classification, in IEEE Transactions on Image Processing, vol. 30, pp. 8671-8685, 2021.

- [50] Yu H, Xu Z, Zheng K, et al. MSTNet: A multilevel spectral–spatial Transformer network for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-13.
- [51] H. -C. Li, S. -S. Li, et al. Recurrent Feedback Convolutional Neural Network for Hyperspectral Image Classification, in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 5504405, 2021.3064349.
- [52] C. Shi, D. Liao, T. Zhang and L. Wang, Hyperspectral Image Classification Based on Expansion Convolution Network, in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-16, 2022, Art no. 5528316.
- [53] Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30 (2017).
- [54] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N. Attention is all you need. Adv. Neural Inf. Process. Syst. 2017, 30, 5998–6008.
- [55] He, X.; Chen, Y.; Lin, Z. Spatial-Spectral Transformer for Hyperspectral Image Classification. Remote Sens. 2021, 13, 498.
- [56] Peng Y, Zhang Y, Tu B, et al. Spatial–spectral Transformer with cross-attention for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-15.
- [57] Yu H, Xu Z, Zheng K, et al. MSTNet: A multilevel spectral–spatial Transformer network for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-13.
- [58] Sun L, Zhao G, Zheng Y, et al. Spectral–spatial feature tokenization Transformer for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-14.
- [59] 张艺超, 郑向涛, 卢孝强. 基于层级 Transformer 的高光谱图像分类方法[J]. 测绘学报, 2023, 52(7): 1139.
- [60] He, Xin, Yushi Chen, and Zhouhan Lin. Spatial-spectral Transformer for hyperspectral image classification. Remote Sensing 13.3 (2021): 498.
- [61] Tian, L.; Tu, Z.; Zhang, D.; Liu, J.; Li, B.; Yuan, J. Unsupervised learning of optical flow with cnn-based non-local filtering. IEEE Trans. Image Process. 2020, 29, 8429–8442.
- [62] Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An image is worth  $16 \times 16$  words: Transformers for image recognition at scale.
- [63] Touvron, H.; Cord, M.; Douze, M.; Massa, F.; Sablayrolles, A.; Jégou, H. Training data-efficient image Transformers & distillation through attention. In Proceedings of the International Conference on Machine Learning, Virtual Event, 18–24 July 2021; pp. 10347–10357.
- [64] Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. swin Transformer: Hierarchical vision Transformer using shifted windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Virtual Event, 12 October 2021; pp. 10012–10022.
- [65] Han, K.; Xiao, A.; Wu, E.; Guo, J.; Xu, C.; Wang, Y. Transformer in Transformer. Adv. Neural Inf. Process. Syst. 2021, 34, 1056–1067.
- [66] Wang, W.; Xie, E.; Li, X.; Fan, D.-P.; Song, K.; Liang, D.; Lu, T.; Luo, P.; Shao, L. Pyramid vision Transformer:

- A versatile backbone for dense prediction without convolutions. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Virtual Event, 12 October 2021; pp. 568–578.
- [67] Sun J, Zhang J, Gao X, et al. Fusing spatial attention with spectral-channel attention mechanism for hyperspectral image classification via encoder-decoder networks[J]. *Remote Sensing*, 2022, 14(9): 1968.
- [68] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2016, 770-778.
- [69] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE. 2018: 7132-7141.
- [70] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module[C]//*Proceedings of the European Conference on Computer Vision*. Munich, Germany: Springer. 2018: 3-19.
- [71] Zhong Z, Li J, Luo Z, et al. Spectral-spatial residual network for hyperspectral image classification: A 3-D deep learning framework[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 56(2): 847-858.
- [72] Xu Y, Du B, Zhang L. Beyond the patchwise classification: Spectral-spatial fully convolutional networks for hyperspectral image classification[J]. *IEEE Transactions on Big Data*, 2019, 6(3): 492-506.

# 作者简历

## 一、基本情况

姓名：董雯雯 性别：女 民族：汉 出生年月：1998-11-23 籍贯：山东省济宁市  
1. 2017.09-2021.06 德州学院  
2. 2021.09-2024.06 山东科技大学

## 二、攻读学位期间获奖情况

1. 2021 年 12 月获山东科技大学第八届大学生科技创新大赛三等奖；
2. 2022 年 10 月获山东科技大学二等学业奖学金；
3. 2022 年 8 月获第九届山东科技大学“互联网+”大学生创新创业大赛二等奖；
4. 2022 年 11 月获优秀研究生荣誉称号；
5. 2023 年 4 月获山东科技大学“优秀共青团员”称号；
6. 2023 年 10 月获山东科技大学科技活动先进个人
7. 2023 年 10 月获山东科技大学二等学业奖学金；

## 三、攻读学位期间研究项目情况

1. Cui Binge, **Dong Wenwen**, Yin Bei, et al. Hyperspectral image classification method based on semantic filtering and ensemble learning[J]. Infrared Physics & Technology, 2023, 135: 104949.
2. Liu Rongjie, Cui Binge, **Dong Wenwen**, et al. A refined deep-learning-based algorithm for harmful-algal-bloom remote-sensing recognition using Noctiluca scintillans algal bloom as an example[J]. Journal of Hazardous Materials, 2024: 133721.
3. 国家发明专利：王智慧，贾宪青，**董雯雯**等人发表专利《一种基于机器视觉的工件表面缺陷检测方法和装置.》专利号：CN114581426A
4. 融合地学知识图谱的多生态型滨海湿地植被精细分类方法研究. 国家自然科学基金项目（面上项目）(No.42276185)。参与人员
5. 基于机器视觉的智能制造产品表面缺陷检测系统（软著） 登记号：2021SR1892954

## 致 谢

行文至此，落笔为终，这一年我25岁，学生生涯也就此告一段落，叹时光易逝，惜韶华难追。曾阅过些许感人肺腑的致谢，也曾想过自己要写些什么，但执笔于此，心中思绪万千，百感交集，却不知如何描述。回首过去，每一步都有迹可循，走的每一步都算数。

桃李不言，下自成蹊。我要感谢我的导师崔宾阁教授。崔老师不仅为我提供了严谨的学术指导，还在我遇到困难和挫折时给予了我鼓励和支持。同时也感谢课题组的路燕老师、于建志老师对我的帮助。老师们的敬业精神和深厚的学术造诣，让我深受启发，也激励我在学术道路上不断前进。

春晖寸草，山高海深。感谢父母养育了我，陪我长大，教我明事理，知进退。感谢父母对我多年求学以来一直无条件的付出，给予我无忧无虑的生活，无条件的支持我的所有决定；感谢父母努力生活，辛苦打拼，让我可以有足够的底气和信心去面对未知的生活。愿时光再慢一些，愿生活对父母再温柔一些，愿我可以成为你们的骄傲。

山水一程，有幸遇见。感谢我可爱的舍友们，我们朝夕相伴、相互包容、相互扶持、并肩作战，让我的求学路充满欢声笑语。感谢师兄师姐，师弟师妹们，特别感谢我的同门黄诗涵、张文凤、郭宝涛、刘成龙、孙欢、燕庆杰。正因为有你们，让这三年不这么难熬。这世界有那么多人，恰巧我们成为了朋友，那些不能对父母道出的心事，是你们让它们有了出口，一起为每一份喜悦喝彩，期待彼此成为更好的人。

愿岁并谢，与友长兮。感谢我的朋友们一直以来的陪伴，不厌其烦的听我诉说我的喜怒哀乐。我们一起分享着彼此的青春，相互鼓励、共同奋斗着，祝我们的友谊永远看不到尽头，在未来漫长的日子里，继续相伴而行。

时光清浅，幸之相遇。感谢我的男朋友王家浩，我们是同学，是好朋友，亦是恋人，既见君子，云胡不喜，感谢你带着真诚善良坚定的选择我，让我感受爱与被爱。前路漫漫，祝你，祝我，也祝我们，在未来日子里，携手共进退，一起奔赴更美好的未来。

最后，我要感谢参与我论文评审和答辩的各位老师。他们对我的论文提出了宝贵的意见和建议，使我的研究工作更加完善。他们的严谨治学态度和敬业精神，让我深感敬佩。

## 学位论文数据集

关键词*	密级*	中图分类号*	UDC	论文资助
高光谱图像分类; 早期分层反馈; 注意力机制; 空谱特征融合	公开	TP751	004	
学位授予单位名称*	学位授予单位代码*	学位类别*	学位级别*	
山东科技大学	10424	电子信息硕士专业学位	硕士	
论文题名*		并列题名*		论文语种*
基于分层反馈和空谱特征融合的高光谱图像分类方法研究		Research on Hyperspectral Image Classification Method Based on Hierarchical Feedback and Spatial-Spectral Feature Fusion		中文
作者姓名*	董雯雯	学号*	202183060029	
培养单位名称*	培养单位代码*	培养单位地址	邮编	
山东科技大学	10424	山东省青岛市	266590	
学科专业*	研究方向*	学制*	学位授予年*	
计算机技术	人工智能	三年	2024	
论文提交日期*		2024 年 5 月 28 日		
导师姓名*	崔宾阁	职称*	教授	
评阅人		答辩委员会主席*	答辩委员会成员	
		路燕	刘伟、刘荣杰、张鹏、陈欣	
电子版论文提交格式	文本 <input checked="" type="checkbox"/> 图像 <input type="checkbox"/> 视频 <input type="checkbox"/> 音频 <input type="checkbox"/> 多媒体 <input type="checkbox"/> 其他 <input type="checkbox"/>			
推荐格式: application/msword; application/pdf				
电子版论文出版(发布)者	电子版论文出版(发布)地		权限声明	
论文总页数*	75			
注: 共 33 项, 其中带*为必填数据, 共 22 项。				



电子信息硕士专业学位论文

# 基于分层反馈和空谱特征融合的高光谱 图像分类方法研究

Research on Hyperspectral Image Classification Method based on  
Hierarchical Feedback and Spectral-spatial Feature Fusion

作 者 董雯雯

导 师 崔宾阁 教授

山东科技大学  
二〇二四年五月