



电子信息硕士专业学位论文

基于多层次特征交互与空谱特征协同的滨海湿地遥感 精细制图方法研究

Research on the Fine Mapping Method of Remote Sensing of Coastal
Wetlands based on Multi-level Feature Interaction and Spatial-Spectral
Feature coordination

作 者 白雪梅

导 师 崔宾阁 教授

山东科技大学

二〇二五年五月

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

基于多层次特征交互与空谱特征协同的滨海湿地遥感 精细制图方法研究

Research on the Fine Mapping Method of Remote Sensing of Coastal Wetlands
based on Multi-level Feature Interaction and Spatial-Spectral Feature coordination

作 者 白雪梅 入学时间 2022 年 9 月

导 师 崔宾阁 职 称 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 5 月 21 日 答辩委员会主席 路燕

答辩委员会成员 刘荣杰 刘伟 张鹏 陈欣

论文提交日期 2025 年 5 月 28 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。

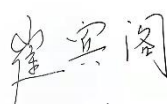
作为申请学位的条件之一，学校有权保留学位论文并向国家有关部门或其指定机构送交论文的电子版和纸质版；有权将学位论文的全部或部分内容编入有关数据库发表，并可以以电子、网络及其他数字媒体形式公开出版；允许学校档案馆和图书馆保留学位论文的纸质版和电子版，可以使用影印、缩印或扫描等复制手段保存和汇编学位论文；为教学和科研目的，学校档案馆和图书馆可以将公开的学位论文作为资料在档案馆、图书馆等场所或在校园网上供校内师生阅读、浏览。

（保密的学位论文在解密后适用本授权）

作者签名：

日期：2025 年 6 月 7 日

导师签名：


日期：2025 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的研究成果。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：

2025 年 6 月 7 日

学位论文审查认定书

研究生 在规定的学习年限内，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：2025 年 6 月 7 日

摘要

黄河口湿地是我国暖温带最广阔、保存最完整的湿地生态系统之一，不仅是自然资源的重要宝库，还在调节气候、环境保护及空气净化方面扮演着不可或缺的角色。此外，它也为众多本土动植物提供了得天独厚的栖息环境。近年来，受人类活动和自然因素的共同影响，黄河口湿地的面积正遭受缩减的严峻挑战。遥感技术凭借其高效性与精确性，在滨海湿地的监测与保护工作中占据了举足轻重的地位。高光谱图像以其卓越的光谱分辨率、丰富的波段信息以及图谱合一的数据形态等特点，因而被广泛用于地物精细分类。本文基于国产“Zhuhai-1”和“高分五号”卫星高光谱遥感图像对滨海湿地植被精细制图方法进行深入研究，主要研究内容和创新成果如下：

(1) 滨海湿地中包含不同尺度的地物类型，单一的卷积操作可能无法充分捕捉和区分不同尺度地物的细微特征。本文提出了一种新的多分支多层次特征交互网络（Multi-branch Multi-level Feature Interaction Network, MMFINet）。首先，该模型利用 ResNet50 作为编码器，有效提取植被图像的低层次细节特征与高层次抽象特征。其次，为了进一步细化编码器的输出特征，本文采用了多尺度注意聚合模块（Multi-Scale Attention Aggregation Module, MSAA），可增强特征提取和融合的能力。然后，模型设计了独特的多分支解码器结构，该结构通过跳层连接将不同分支的特征图进行融合，从而保留更多的细节信息。另外，这种结构不仅负责深入分析并反馈来自不同层级的特征信息，还引入了基于残差单元的全局注意力机制模块（Global Attention Mechanism Module Based on Residual Units, GAM-RU），该注意力机制能够显著提升特征融合过程中的关键信息关注度，进一步增强模型的分类效果。最后，通过增加深度监督优化分类结果。实验结果表明，MMFINet 提升了复杂区域的解译效果，在黄河口滨海湿地两个数据集上的 OA 分别达到了 93.01% 和 92.92%。

(2) 滨海湿地地物空间分布复杂，且不同植物覆被间的光谱特征差异较小，滨海湿地遥感影像解译仍然是一项具有挑战性的任务。本文提出空谱特征协同网络（Spatial Spectral Feature Coordination Network, SSFCNet），它结合了 GCN 和变换器在空间和光谱域特征提取方面的优势，充分利用分类像素的上下文信息，同时在光谱域建立远程依赖关系，有助于实现更精细的分类结果。首先，设计了通道先验卷积注意力 GCN 模块（Channel Prior Convolutional Attention GCN Module, CPCAGCN），在该模块中引入通道感知卷积注意力机制，支持在通道和空间维度上动态分配注意力权重。另外，为了增强局部光谱信息的提取能力，引入了局部光谱特征融合 Transformer 模块（Local Spectral Feature Fusion Transformer Module, LST）。为了减少有价值信息在分层传播过程中的丢失，在变压器中引入了跨层连接机制，增强层间信息传递和自适应特征融合。实验结果表明，SSFCNet 有效提高了滨海湿地光谱相似地物的识别效果，在两个数据集上的 OA 分别为 94.26 % 和 94.07%。

关键词：滨海湿地；精细制图；多分支结构；空谱特征协同

Abstract

The Yellow River Estuary Wetland is one of the most extensive and well-preserved wetland ecosystems in the warm temperate zone of China. It serves not only as a vital repository of natural resources but also plays an indispensable role in climate regulation, environmental protection, and air purification. Additionally, it provides a unique habitat for numerous native flora and fauna. In recent years, the area of the Yellow River Estuary Wetland has been facing severe challenges of reduction due to the combined effects of human activities and natural factors. Remote sensing technology, with its efficiency and precision, has become crucial in the monitoring and conservation of coastal wetlands. Hyperspectral imagery, known for its superior spectral resolution, rich band information, and integrated spectral-spatial data characteristics, is widely used for detailed classification of surface features. This paper conducts an in-depth study on the methods of fine-scale mapping of coastal wetland vegetation based on hyperspectral remote sensing images from the domestic "Zhuhai-1" and "Gaofen-5" satellites. The main research content and innovative achievements are as follows:

(1) Coastal wetlands encompass land cover types at various scales, and a single convolutional operation may not sufficiently capture and differentiate the subtle features of land cover at different scales. This paper proposes a novel Multi-branch Multi-level Feature Interaction Network (MMFINet). Firstly, the model utilizes ResNet50 as an encoder to effectively extract low-level detailed features and high-level abstract features from vegetation images. Secondly, to further refine the output features of the encoder, a Multi-Scale Attention Aggregation Module (MSAA) is employed, which enhances the capability of feature extraction and fusion. Then, the model designs a unique multi-branch decoder structure that fuses feature maps from different branches through skip-layer connections, thereby preserving more detailed information. Additionally, this structure not only analyzes and feeds back feature information from different levels in depth but also introduces a Global Attention Mechanism Module Based on Residual Units (GAM-RU). This attention mechanism significantly improves the focus on key information during the feature fusion process, further enhancing the model's classification performance. Finally, the classification results are optimized by adding deep supervision. Experimental results show that MMFINet improves the interpretation of complex regions, achieving OA values of 93.01% and 92.92% on two datasets of the Yellow River Estuary coastal wetland, respectively.

(2) Coastal wetlands exhibit complex spatial distributions of land cover, and the spectral characteristics among different vegetation covers are often subtle, making the interpretation of remote sensing images of coastal wetlands a challenging task. This paper proposes a Spatial Spectral Feature Coordination Network (SSFCNet), which combines the strengths of GCN and Transformer in extracting features from both spatial and spectral domains. It fully utilizes the contextual information of classification pixels while establishing long-range dependencies in the

spectral domain, contributing to more refined classification results. Firstly, a Channel Prior Convolutional Attention GCN Module (CPCAGCN) is designed, which introduces a channel-aware convolutional attention mechanism to dynamically allocate attention weights across both channel and spatial dimensions. Additionally, to enhance the extraction capability of local spectral information, a Local Spectral Feature Fusion Transformer Module (LST) is incorporated. To mitigate the loss of valuable information during hierarchical propagation, a cross-layer connection mechanism is introduced in the Transformer, enhancing inter-layer information transfer and adaptive feature fusion. Experimental results demonstrate that SSFCNet effectively improves the identification of spectrally similar land covers in coastal wetlands, achieving OA values of 94.26% and 94.07% on two datasets, respectively.

Keywords: coastal wetlands; fine-scale mapping; multi-branch structure; spatial-spectral feature coordination

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 主要研究内容	7
1.4 本文章节安排	7
2 相关技术	9
2.1 卷积神经网络	9
2.2 语义分割	12
2.3 图卷积神经网络	15
2.4 Vision Transformer 网络	错误!未定义书签。
2.5 深度监督	17
2.6 本章小结	18
3 基于多分支多层次特征交互的滨海湿地精细制图方法	19
3.1 引言	19
3.2 MMFINet 网络结构设计	19
3.3 实验准备	24
3.4 实验结果与分析	28
3.5 本章小结	35
4 基于空谱特征协同的滨海湿地精细制图方法	36
4.1 引言	36
4.2 SSFCNet 网络结构设计	36
4.3 实验准备	40
4.4 实验结果与分析	42
4.5 本章小结	48
5 总结与展望	49
5.1 总结	49
5.2 展望	50

参考文献
作者简历
致 谢
学位论文数据集

Contents

List of Figures	I
List of Tables	III
List of Variables	III
1 Introduction	1
1.1 Research background and significance	1
1.2 Research status at home and abroad	2
1.3 Main research contents	7
1.4 Chapter arrangement	7
2 Related techniques	9
2.1 Convolutional Neural Network	9
2.2 Semantic Segmentation	12
2.3 Graph Convolutional Neural Networks	15
2.4 Vision Transformer	错误!未定义书签。
2.5 Deep Supervision	17
2.6 Chapter summary	18
3 A fine mapping method of coastal wetland based on multi-branch multi-level feature interaction	19
3.1 Foreword	19
3.2 MMFINet network structure design	19
3.3 Experimental preparation	24
3.4 Experimental results and analysis	28
3.5 Chapter summary	35
4 A fine mapping method of coastal wetland based on spatial spectral feature coordination	36
4.1 Foreword	36
4.2 SSFCNet network structure design	36
4.3 Experimental preparation	41
4.4 Experimental result and analysis	42
4.5 Chapter summary	48
5 Summary and prospects	49

5.1 Conclusion.....	49
5.2 Prospects.....	50

References

Author's Resume

Acknowledgements

Thesis Data Collection

图清单

图序号	图名称	页码
图 2.1	卷积神经网络示意图	9
Fig.2.1	Schematic diagram of convolutional neural network	9
图 2.2	卷积操作示意图	10
Fig.2.2	Schematic diagram of convolution operation	10
图 2.3	最大池化和平均池化示意图	10
Fig.2.3	Schematic diagram of Max pooling and Average pooling	10
图 2.4	常见的激活函数示意图	11
Fig.2.4	Schematic diagram of popular activation functions	11
图 2.5	全连接层示意图	11
Fig.2.5	Schematic diagram of fully-connected layers	11
图 2.6	语义分割示意图	12
Fig.2.6	Schematic diagram of Semantic Segmentation	12
图 2.7	U-Net 结构示意图	13
Fig.2.7	The structure diagram of U-Net	13
图 2.8	Deeplab3+结构示意图	13
Fig.2.8	The structure diagram of Deeplab3+	13
图 2.9	ResNet 的残差学习模块示意图	14
Fig.2.9	Schematic diagram of the residual learning module of ResNet	14
图 2.10	图卷积结构示意图	14
Fig. 2.10	Schematic diagram of GCN	14
图 2.11	Transformer 中的注意力机制示意图	17
Fig. 2.11	Schematic diagram of attentive mechanism in Transformer	17
图 2.12	深度监督结构示意图	18
Fig. 2.12	Schematic diagram of deep supervision structure	18
图 3.1	MMFNet 网络整体架构	20
Fig.3.1	Overall structure of MMFNet network	20
图 3.2	多尺度聚合模块结构	21
Fig.3.2	Structure of Multi-Scale Attention Aggregation (MSAA) module	21
图 3.3	上采样块结构	22
Fig.3.3	Structure of unsample block	22
图 3.4	基于残差单元的全局注意力机制模块结构	23
Fig.3.4	Structure of global attention mechanism module based on residual units	23
图 3.5	“Zhuhai-1 号”卫星高光谱数据集 (a)假彩色图像(b)地面真值图	24
Fig.3.5	“Zhuhai-1” satellite hyperspectral dataset (a) False color composite (b) Ground truth	24
图 3.6	“高分五号”卫星高光谱数据集 (a)真彩色图像 (b)地面真值图	24
Fig.3.6	“GF-5” satellite hyperspectral dataset (a) True color composite (b) Ground truth	24
图 3.7	“Zhuhai-1 号”卫星高光谱数据集上的分类结果图	29

Fig.3.7	Classification maps for the Zhuhai-1 satellite hyperspectral dataset	29
图 3.8	“高分五号”卫星高光谱数据集上的分类结果图	32
Fig.3.8	Classification maps for the GF-5 satellite hyperspectral dataset	32
图 3.9	不同空间输入大小的对分类性能的影响	35
Fig.3.9	Effect of different spatial sizes on overall accuracies on four HSI datasets	35
图 4.1	SSFCNet 网络整体架构	38
Fig.4.1	Overall structure of SSFCNet network	38
图 4.2	通道先验卷积注意力模块结构	39
Fig.4.2	Structure of channel prior convolutional attention module	39
图 4.3	像素选择块结构	40
Fig.4.3	Structure of pixel selection module	40
图 4.4	标准 2-D 卷积和邻域卷积对比 (a)标准 2-D 卷积 (b)邻域卷积	42
Fig.4.4	Comparison of standard 2-D convolution and neighborhood convolution	42
图 4.5	“Zhuhai-1 号” 卫星高光谱数据集上的分类结果图	44
Fig. 4.5	Classification maps for the Zhuhai-1 satellite hyperspectral dataset	44
图 4.6	“高分五号” 卫星高光谱数据集上的分类结果图	46
Fig.4.6	Classification maps for the GF-5 satellite hyperspectral dataset	46
图 4.7	具有不同数量的 GCN 层数和编码器的 OA 比较	47
Fig.4.7	OA comparison with different numbers of GCN layers and Encoders	47
图 4.8	左侧展示了两个数据集在不同 Patch 大小下的 OA.右侧展示了两个数据集在不同像素选择大小下的 OA	48
Fig.4.8	The left side shows the OA of the two datasets with different patch sizes. The right side shows the OA of the two datasets with different pixel selection sizes.	48

表清单

表序号	表名称	页码
表 3.1	编码器细节	20
Table 3.1	Details of encoder	20
表 3.2	“Zhuhai-1” 卫星高光谱数据集中使用的训练和测试集的数量	25
Table 3.2	Number of training and test samples used in the Zhuhai-1 satellite hyperspectral dataset	25
表 3.3	“高分五号”卫星高光谱数据集中使用的训练和测试集的数量	25
Table 3.3	Number of training and test samples used in the GF-5 satellite hyperspectral dataset	25
表 3.4	混淆矩阵	26
Table 3.4	The confusion matrix	26
表 3.5	实验环境详细信息	27
Table 3.5	The experimental environment details	27
表 3.6	“Zhuhai-1 号” 卫星高光谱数据集上的分类结果	30
Table 3.6	Table 3.5 Classification result of the Zhuhai-1 satellite hyperspectral dataset	30
表 3.7	“高分五号” 卫星高光谱数据集上的分类结果	31
Table 3.7	Classification result of the GF-5 satellite hyperspectral dataset	31
表 3.8	带有交互和不带有交互在 OA 方面的性能比较	33
Table 3.8	Performance comparison with and without interaction in terms of training time and OA	33
表 3.9	表 3.8 不同方法计算时间和复杂性比较	33
Table 3.9	Comparison of different methods in terms of running time and complexity	33
表 3.10	模块实验分析	34
Table 3.10	Module experimental analysis	34
表 3.11	编码器网络的实验分析	35
Table 3.11	Experimental analysis of encoder network	35
表 4.1	“Zhuhai-1 号” 卫星高光谱数据集上的分类结果	44
Table 4.1	Classification result of the Zhuhai-1 satellite hyperspectral dataset	44
表 4.2	“高分五号” 卫星高光谱数据集上的分类结果	45
Table 4.2	Classification result of the GF-5 satellite hyperspectral dataset	45
表 4.3	消融实验	47
Table 4.3	Ablation experiments	47
表 4.4	不同方法计算时间和复杂性比较	49
Table 4.4	Comparison of different methods in terms of running time and complexity	49

变量注释

变量	注释	初现页
A	邻接矩阵	
D	度矩阵	
L_n	对应的对称归一化拉普拉斯矩阵	
x	输入信号	
U	L_n 的特征向量矩阵	
Λ	L_n 的特征值	
θ	参数	
g_θ	特征值 Λ 的函数	
T_K	切比雪夫多项式	
$H^{(l)}$	第 1 层的特征结点矩阵	
σ	非线性激活函数	
$W^{(l)}$	第 1 层的权重矩阵	
\tilde{A}	加了自环的邻接矩阵	
\tilde{D}	\tilde{A} 的度矩阵	
Q	查询向量	
K	键向量	
V	值向量	
d_k	键向量 K 的维度	
Z_{MHSA}	多头自注意力的输出	
W'	投影矩阵	
Z'	MHSA 的输出	
Z	MLP 的输出	
F_{out}	上采样块的输出	
G_s	GAM-RU 模块中空间注意力的输出	
G_c	GAM-RU 模块中通道注意力的输出	
F_{att}	GAM-RU 的最后输出	
f_{in-Bi}	当前分支 Bi 的特征图	
f_{Bi}	分支 Bi 融合之后的特征图	
p'	综合三个分支之后的最后输出	
L_{main}	主损失	
L_{aux}	辅助损失	
λ	超参数	
N	样本总数	10
x_{ii}	样本预测类别和真实类别相同	10
x_{ij}	样本预测类别与真实类别不符	10
x_{i+}	真实图像中第 i 类的样本数量	10
x_{+i}	预测结果中预测为第 i 类的样本数量	10

F_c	CPCA 模块中通道注意力的输出	38
\hat{F}	CPCA 模块的最终输出特征图	39
V_i	特征图中的第 i 个像素节点的特征向量	
V_j	特征图中的第 j 个像素节点的特征向量	11
$Corr(V_i, V_j)$	像素 V_i 和 V_j 之间的相关性权重	11
S_j^i	排序后的相关性权重列表	11
S_j	表示排序后权重列表 S_j^i 中的第 j 个像素特征向量	12
F_m	最终选择的像素特征了向量	15
M	选择的像素数量	15
X_i	第 i 个光谱带的特征	15
X_i^{group}	以第 i 个光谱带为中心的邻域光谱特征组	15
m	邻域的起始光谱带索引	15
W_i	第 i 个光谱带的卷积权重矩阵	20
b_i	第 i 个光谱带的偏置项	20
X_i'	经过邻域卷积操作后的特征	20
X_i''	经过邻域卷积和残差连接后的最终特征	20
F_N	NC 模块的最终输出特征图	20
$\hat{z}^{(l)}$	第 l 层经过 CAF 模块融合后的特征表示	20
z		
w	可学习的网络参数	20

1 绪论

1 Introduction

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

滨海湿地坐落于淡水河流系统与咸水海洋系统相接的过渡带，是陆地与海洋生态系统的重要生态界面^[1]，这类生态系统不仅具有独特的生态结构和显著的经济价值^[2]，更因其对全球气候变化和人类活动的高度敏感性而成为备受关注的生态脆弱区^[3]。

作为生物多样性热点区域，滨海湿地生态系统具有极高的初级生产力，并发挥着多重生态服务功能。其核心功能包括：环境污染物降解、区域气候调节、沉积物淤积与陆地形成、水源涵养、生物多样性保育，以及为人类提供重要的自然资源、土地资源和旅游资源等^[4,5]。然而，近年来，盐沼植被正面临日益严峻的自然与人为压力，导致本地植物群落显著退化，景观破碎化程度不断加深，进而引发生物多样性的急剧下滑，这对于整个生态系统的稳定与可持续性构成了严重威胁^[6]。在此背景下，准确把握滨海湿地植被的物种构成与空间布局，对于监测当地水文气候变动、维护生物多样性、推动滨海湿地的保护与管理、评估滨海湿地生态系统的发展态势等方面具有重要的科学价值和实践意义^[7]。尽管遥感监测技术近年来取得了显著进展，但现有技术手段在滨海湿地监测应用中仍存在明显局限性，难以满足业务部门对高精度、高可靠性和智能化监测的迫切需求。因此，开发更加精准、可靠且智能化的遥感影像解译方法，已成为当前滨海湿地保护与研究中亟待突破的关键技术瓶颈^[8]。

面对黄河三角洲湿地当前面临的诸多复杂挑战，传统野外调查方法因效率低下、成本高昂等局限性，难以满足对湿地状况的实时监测需求，从而在河口湿地的长期跟踪调查中显现出了明显的局限性。相比之下，遥感技术凭借其数据获取迅速、信息涵盖广泛、动态监测能力强以及多平台作业等优势，为湿地监测与研究工作提供了强有力的技术支撑。该技术在湿地植被分类和湿地景观时空动态监测方面发挥着日益重要的作用：一方面，高分辨率遥感影像能够精确识别湿地植被类型，为研究湿地生态系统的物种组成、空间分布格局及其生态功能提供了重要依据；另一方面，其动态监测能力为揭示湿地生态系统的演变规律及其驱动机制提供了宝贵的数据支持^[9,10]。在滨海湿地植被监测这一具体领域，遥感技术同样引起了学术界的广泛关注，并在滨海湿地植被分类^[11]、植被变化动态监测^[7]、湿地退化评估^[12]以及生态恢复成效分析^[13]等多个方面得到了深入且广泛的应用，为滨海湿地的科学管理与保护提供了宝贵的数据支持与决策依据。

高光谱成像技术相较于传统遥感技术具有显著优势。其高光谱分辨率、宽波段覆盖范围以及丰富的光谱信息特征，使其能够精确捕捉地表物体的细微光谱差异。

每个像元所包含的连续光谱曲线，完整记录了地表物体对不同波长电磁波的反射、透射和发射特性。通过深入分析高光谱图像的光谱特征，可以充分利用这些细微差异实现地物的精确识别与分类，从而大幅提升分类精度。将高光谱技术应用于滨海湿地分类研究，不仅能够提高分类准确性，更能深入解析滨海湿地生态系统的组成与分布特征，为湿地资源保护与管理提供科学依据。

作为机器学习领域的重要分支，深度学习近年来在图像分类、目标检测等多个领域取得了突破性进展。在大数据资源与强大计算能力的双重支撑下，深度学习展现出了卓越的特征学习能力，已成为滨海湿地植被遥感监测研究的重要方向^[14,15]。通过将深度学习技术与滨海湿地的实际环境紧密结合，能够实现对湿地生态系统的实时监测与动态评估，及时发现湿地生态系统的变化与异常情况，为湿地资源保护与管理提供科学、及时的决策支持，促进湿地生态系统的可持续发展。

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

黄河口湿地，这一海陆交织的特殊地带，其复杂的植被群落结构和动态变化的环境条件对传统资源调查方法提出了严峻挑战。在这一背景下，遥感技术凭借其高效、大范围的植被信息获取能力，日益成为湿地研究的重要手段。随着高时空分辨率遥感卫星的相继发射，遥感技术在湿地植被监测中的应用深度和广度不断拓展^[16-18]。其中，高光谱遥感技术以其“图谱合一”的独特优势，能够精确获取地表物质连续的光谱特征信息，为实现遥感影像的精细分类与识别提供了可靠的数据基础。这一技术突破使得滨海湿地的高精度解译成为可能，并迅速成为遥感领域的研究热点。国内外学者在该领域取得的一系列创新性成果，有力推动了高光谱遥感技术在滨海湿地监测中的应用向更高水平发展。

1.2.1 高光谱遥感分类研究现状

1983年，美国喷气式推进实验室(JPL)成功研制出首台实验性航空成像光谱仪，这一突破性进展标志着高光谱遥感技术正式进入发展新纪元。随着航空航天技术的飞速发展，一系列先进的光谱仪器相继问世，极大地推动了高光谱遥感技术的进步。高光谱图像凭借其独特的高维光谱信息与空间信息的融合特性，提供了丰富的纹理信息、空间信息及大地几何特征，显著拓展了遥感技术的应用边界，在环境监测、资源勘探、农业评估等领域得到广泛应用。在高光谱图像处理领域，高光谱图像分类无疑是一个研究热点。在高光谱图像处理领域，图像分类始终是研究热点，而如何实现像素级的高效精准分类仍是当前面临的关键挑战。目前，高光谱图像分类主要采用传统机器学习方法和深度学习方法两大类技术路线。下文将分别阐述这两类方法的最新研究进展。

（1）传统机器学习方法

传统的机器学习分类方法主要分为监督分类与无监督分类两大类。监督分类方法通过已知类别的训练样本构建分类器,实现对未知样本的准确识别与分类。这类方法依赖于样本先验知识,运用统计和机器学习算法进行特征提取和模型构建。其优势在于分类精度高,但需要大量训练样本和计算资源支持。常用的监督分类方法包括支持向量机^[19] (Support Vector Machine, SVM)、决策树分类^[20] (Decision Tree Classification, DTC) 等。SVM 是一种基于统计学习理论的二类分类模型。其核心思想是在特征空间中寻找最优超平面,实现两类样本的正确划分并最大化分类间隔,因此也被称为最大间隔分类器。与感知器不同, SVM 不仅追求分类正确性,还强调分类器的泛化能力。潘煜琳等研究者^[21]系统评估了支持向量机、随机森林和 K 近邻三种分类器在湿地信息提取中的应用效果。研究结果表明,通过融合高分五号 (GF-5) 和高分六号 (GF-6) 卫星的多波段特征,结合 SVM 算法可显著提升湿地信息提取精度,其总体分类效果优于其他方法组。DTC 是一种基于树形结构的监督学习算法,通过递归划分特征空间来实现分类任务。该算法构建的层次化树形结构中,内部节点代表属性判断条件,分支对应判断结果,叶节点则代表最终分类结果,具有直观易懂的特点。针对高光谱遥感应用中普遍存在的地物光谱混淆问题,沈聪颖等人^[22]于 2015 年提出了一种创新的水体提取方法。该方法基于 GF-5 卫星影像数据,将单波段阈值法与新型阴影水体指数相结合,构建了改进的决策树模型。与传统单一分类算法相比,该模型能够有效克服山体阴影和裸地等干扰因素的影响,在山区水体提取任务中表现出更高的精度和鲁棒性。

与监督分类方法不同,无监督分类方法无需先验知识,主要依据样本数据的统计特性,通过聚类算法实现样本分类。无其优势在于无需训练样本,能自动发现数据中的隐藏结构和模式,但分类精度可能受样本数据质量和聚类算法选择的影响。常用的无监督方法包括 K-均值^[23] (K-Means) 算法和 ISODATA 算法^[24] 等。K-means 聚类算法是一种基于欧氏距离度量的经典迭代聚类算法,通过最小化类内距离、最大化类间距离实现样本自动分组。作为一种无监督学习方法, K-means 在数据挖掘和机器学习领域具有广泛应用,包括图像分割、文本聚类、社交网络分析等方向。2018 年, K. YASSER 等人^[25]创新性地将 K-means 算法应用于高光谱数据立方体的分析,成功实现了组织光谱差异的精确检测。该研究证实了 K-means 算法在高光谱图像处理中的重要价值,特别是在构建半自动化医学诊断系统方面展现出巨大潜力,能够有效识别和分类正常组织与导管原位癌组织,为医学影像分析提供了新的技术手段。ISODATA 聚类算法是在 K-means 算法基础上改进的先进聚类方法,具有两个显著优势:采用批量样本校正策略提高计算效率;引入动态聚类机制,可根据预设条件自动合并或分裂聚类簇,获得更合理的聚类数目和更优的聚类效果。HONG P 等人^[26]提出了一种创新的改进型 ISODATA 算法,通过融合高光谱图像的频谱特征和空间结

构信息,引入加权矩阵来计算像素间相似度。这种双重信息融合机制能够更精确地判定像素的类别归属,显著提升了高光谱图像分类的准确性。

(2) 深度学习方法

相较于机器学习,深度学习在特征提取方面展现出显著优势。其能够从原始数据中自动挖掘深层次特征^[27-29],无需人工介入进行复杂繁琐的特征工程,在处理复杂数据集时更为高效精确。通过构建多层神经网络架构,深度学习能够敏锐地捕捉到数据中的非线性关联与潜在模式,从而实现对复杂问题的精准建模与预测。这些优势使深度学习算法在遥感图像处理领域展现出广阔应用前景,涵盖城市规划、环境监测、灾害评估等多个重要领域。目前,深度学习算法已成为遥感图像分类领域的核心技术手段,其应用范围随着技术创新和场景拓展不断扩大,为遥感技术的发展注入新的活力。在深度学习的众多算法中,卷积神经网络(CNN)^[30]、循环神经网络(RNN)^[31]以及长短时间记忆网络(LSTM)^[32]等代表性算法尤为突出。

CNN 以其卓越性能在图像特征捕捉方面表现出色,能高效捕捉图像中的局部细节与整体结构特征,实现有效分类。Lee 等人^[33]提出了一种创新的 CNN 模型,巧妙融合空间与光谱特征进行分类,充分利用了高光谱图像的高维度数据优势。为了降低计算成本,Meng 等人^[34]设计了一种轻量级的光谱-空间卷积模块,该模块显著减少了卷积神经网络的参数数量及计算复杂度,实现了高效性与准确性的双重提升。RNN 在高光谱遥感分类中同样展现出了独特优势。它能够处理长序列数据,有效建模光谱数据间的依赖关系,捕捉并利用光谱序列中的时序特征,实现精准的分类任务。Mou 等人^[35]开创性地引入了循环神经网络架构,构建了高性能的高光谱分类模型,为遥感图像分类领域带来了新的突破。

Transformer 模型^[36],凭借其引入的自注意力机制和编码器-解码器架构,有效弥补了 CNN、RNN 及 LSTM 等模型在序列数据处理、全局信息捕捉、并行计算能力和可扩展性方面的局限性。这些独特优势使得 Transformer 能够更精准地把握光谱特征间的长距离依赖关系,充分利用高光谱图像丰富的多维信息,精准提取光谱与空间特征,显著提升分类的准确性。在处理大规模高光谱数据集时,Transformer 展现出更高的效率,能够更迅速地完成任务。He 等人^[37]巧妙地融合 Transformer 与 CNN,构建了一个分类系统,利用 CNN 提取图像空间特征,结合改良的 Transformer 捕捉光谱序列间的复杂关联性。Sun 等人^[38]开发的 SSFTT 方法,旨在深入挖掘并整合空间光谱特征与高级语义特征,进一步提升分类性能。Transformer 模型的兴起激发了计算机视觉领域内众多基于其衍生方法的创新应用。Vision Transformer (ViT)^[39]作为 Transformer 架构在计算机视觉领域的开创性应用,在 ImageNet 图像分类任务中取得了突破性进展。ViT 通过将图像切割成多个小块,并将这些小块映射为向量输入到 Transformer 编码器中,实现了图像分类的新

突破。然而, ViT 的卓越性能依赖于大规模数据集。为了充分发挥 ViT 的潜力, He 等人^[40]提出了跨光谱视觉变换器 (CSiT), 通过双路径视觉变换器框架有效融合粗粒度与细粒度特征, 提升模型特征提取能力。Swin Transformer^[41]则是在 ViT 的基础上进一步拓展, 为计算机视觉任务提供了一个更加通用的基于 Transformer 的主干网络, 能够广泛应用于多种计算机视觉任务。Rahul Ghotekar 等人^[42]提出的结合 GCNN 与 Swin Transformer 的模型, 则充分融合了 GCNN 的空间感知能力与 Swin Transformer 的强大层次化处理能力。该模型不仅提高了分类的准确性, 还成功捕捉到了细节特征和长距离依赖性, 实现了对高光谱图像中土地覆盖类别的更精确识别和映射, 为遥感图像处理领域带来了新的突破。

1.2.2 滨海湿地遥感分类研究现状

近几十年来, 国内外众多学者对遥感技术在湿地信息提取领域的应用进行了广泛且深入的研究。在湿地植被分类的早期探索中, 主要依赖于人工目视解译的方法。这一方法不仅要求解译人员具备极高的专业素养和丰富的实践经验, 而且工作量大、效率低下。随着计算机技术的迅猛发展, 科研工作者们积极寻求更高效、精确的地物识别技术, 以替代传统的目视解译方法。

(1) 基于像素的分类方法

基于像素的分类方法是遥感图像地物分类的早期尝试。该方法将单个像素视为分类的基本单元, 依据每个像素的光谱特征进行分析, 并将提取的特征送入分类器, 或直接根据决策规则进行分类。然而, 滨海湿地地表环境复杂, 植被类型多样, 光谱特征易受潮汐、土壤湿度等因素影响, 导致基于像素的分类方法存在明显局限性。该方法未能充分考虑像素间的空间关联性, 容易受到噪声干扰和地物内部光谱变异的影响, 分类结果常出现“椒盐现象”, 精度难以满足实际需求。常用的算法包括最大似然分类 (Maximum Likelihood Classification, MLC)^[43], 随机森林 (Random Forest, RF)^[44]等。最大似然分类法也称为贝叶斯分类法 (Bayes Classifier) 或极大似然估计法, 是一种基于概率模型的参数估计方法。它通过构建统计模型来计算像素属于某一类别的概率, 从而实现遥感影像的分类。2012 年, 周国琼等人^[45]提出了一种基于模糊 C 均值聚类 (FCM) 改进的最大似然分类方法, 并将其成功应用于遥感影像分类。实验结果表明, 与传统最大似然分类法相比, 改进后的方法显著提升了分类精度, 展现了更强的地物识别能力。这一改进为遥感影像分类提供了一种更为有效的技术手段。RF 是一种基于集成学习的分类方法, 其核心思想是通过构建多个决策树形成“森林”来实现分类预测。该算法采用 Bootstrap 抽样方法, 从原始数据集中有放回地随机选取部分样本, 同时每个节点分裂时随机选择部分特征进行划分, 这种双重随机机制有效增强了模型的泛化能力。李垒等^[46]提出了一种创新的高光谱遥感图像分类方法, 该方法基于随机森林算法, 通过同时利用高光谱图像丰富的光谱特征和空间结

构信息，显著提升了遥感图像的分类精度。

（2）基于对象的分类方法

为了克服基于像素分类方法的局限，人们提出了基于对象的分类方法。该方法将遥感图像分割成一系列对象（或图斑），并对每个对象进行特征提取和分类。基于对象的分类方法能够灵活利用地物的几何信息和结构信息，同时融入人的思维，将丰富的地学知识构建成知识库，从而显著提高了分类的精度。常见的对象方法有最近邻分类（Nearest Neighbor, NN）^[47]，规则基分类（Rule-based Classification, RBC）^[48]。这一方法被众多学者广泛应用于湿地研究中^[49-53]。莫利江等人^[52]采用 Landsat 提供的杭州湾遥感影像数据，通过选取影像中的光谱特征、纹理特征以及空间特征等多个维度，利用 ENVI 软件中的规则样本分割工具，对杭州湾湿地景观进行了信息提取，并取得了卓越的分类准确性。采用面向对象技术的多层次策略提取湿地信息，能够显著降低分类对象的碎片化程度，有效应对因边界模糊和内部异质性较高而产生的“斑驳现象”。该方法结合了地物的空间几何特征和光谱特征，弥补了“同物异谱、同谱异物”现象导致的分类不足，进一步提升了分类结果的精确度^[54,55]。然而，值得注意的是，这种面向对象的方法在特征选择、分类后处理等过程中易受人的主观因素影响，可能导致结果的不确定性。因此，在实际应用中需要谨慎处理这些环节，以确保分类结果的准确性和可靠性。

（2）基于深度学习的分类方法

近年来，深度学习技术的蓬勃发展，为遥感图像地物分类领域带来了革命性的变化。借助深层神经网络模型的构建，深度学习技术能够自动学习并捕捉到遥感图像中错综复杂的特征和模式。该技术充分利用遥感影像中蕴含的光谱信息、空间信息以及上下文信息，实现了对地物的精确分类，显著提升了分类的准确性和效率。Chen 等人^[56]巧妙地运用深度学习技术，对密集的 Sentinel-2 时间序列影像进行了深入剖析。他们成功地从这些影像中提取出了浑浊河口的水线信息，并据此圆满完成了相应的制图任务。这一成果不仅展示了深度学习技术在遥感图像处理中的强大能力，也为河口区域的水环境监测和管理提供了有力支持。Liu 等人^[57]则进一步探索了高光谱与多光谱影像数据的融合应用。他们利用多数据流 CNN 模型，对来自高分五号和 Sentinel 卫星的黄河口湿地与盐城湿地影像进行了深度挖掘和提取。通过融合不同光谱分辨率的影像数据，该模型能够更全面地捕捉到湿地的复杂特征，为湿地资源的保护和管理提供了更为精确和可靠的信息。然而由于滨海湿地的光谱特性受到土壤性质、水文状况、植被类型及其生长环境等多重因素的交织影响，导致“同物异谱”和“同谱异物”的现象广泛存在，现有的遥感监测研究仍面临局限性，湿地的分类仍然是一项具有挑战性的任务。

1.3 主要研究内容（Main research contents）

本研究以山东省黄河口滨海湿地为研究区域，采用 Zhuhai-1 号和高分五号卫星获取的高光谱遥感影像作为数据源。针对滨海湿地遥感解译中地物类型复杂的特点，本研究充分挖掘高光谱遥感影像的多维特征，深入提取图像中的光谱特征信息和空间结构信息。基于此，本文创新性地提出了基于多层次特征交互与空谱特征协同的滨海湿地遥感精细制图方法，主要创新点如下所述：

（1）基于多分支多层次特征交互的滨海湿地精细制图方法

本研究提出了一种面向滨海湿地精细制图的多分支多层次特征交互网络（MMFINet）。首先，设计了多尺度注意力聚合模块（MSAA），该模块通过多尺度特征融合和空间通道注意力机制，增强了特征的语义丰富性和判别力，为解码器提供了更高质量的特征输入。其次，构建了具有特征反馈机制的多分支结构，实现了分支间的交互，其采用特征传播策略，将前序分支的特征图动态反馈至后续分支的各个层级，充分挖掘低层次细节特征与高层次语义特征之间的空间-语义关联，显著提升了模型对滨海湿地复杂地物边界的识别能力。最后，在跳层连接之前，引入了基于残差单元的全局注意力机制模块（GAM-RU），促进跨层特征的有效融合，进一步增强了网络的特征表达能力。

（2）基于空谱特征协同的滨海湿地精细制图方法

本文提出了基于空谱特征协同网络（SSFCNet）用于滨海湿地精细制图。该网络结合了图卷积网络（GCN）和 Transformer 架构，以充分利用高光谱图像（HSI）的空间-光谱特征。首先，设计了一种高效的通道感知卷积注意力机制 GCN 模块（CPCAGCN），通过引入注意力机制，增强 GCN 对光谱特征的提取能力。其次，为了筛选图卷积网络输出的基础特征，设计了一种像素选择模块，从而简化 Transformer 的输入数据。另外，为了聚焦局部光谱信息，设计了局部光谱特征融合 Transformer 模块（LST），通过融合当前光谱带及其邻域的特征，增强了局部光谱信息的提取能力。

1.4 本文章节安排（Chapter arrangement）

本文围绕滨海湿地高光谱遥感图像分类问题展开研究，主要内容分为五个章节，具体安排如下：

第一章：绪论。本章首先阐述了滨海湿地高光谱遥感分类研究的背景及其重要意义。随后，系统概述了当前国内外在高光谱遥感分类领域和滨海湿地遥感分类领域的研究进展。在此基础上，明确了本文的研究目标与核心创新点，并简要介绍了本文的章节安排。

第二章：相关工作介绍。本章详细介绍了本文研究涉及的相关理论与技术基础，

主要包括卷积神经网络、语义分割、图卷积神经网络、Vision Transformer 以及深度监督的基本原理。此外，本章还介绍了实验过程中所使用的数据集以及评价模型性能的指标。本章内容为第三、四章的研究提供了理论基础和技术支持。

第三章：本章提出了一种多分支多层次特征交互网络，用于解决滨海湿地地物多尺度特性带来的分类难题。本章首先介绍基于多分支多层次特征交互网络的整体框架设计，包括多尺度注意力聚合模块，多分支解码器的详细结构的具体实现。随后，阐述了模型的损失函数设计及优化策略。在实验部分，详细描述了实验数据的设置、实验环境的具体配置、对比方法的选取以及实验参数的设定。随后，通过一系列实验，将提出的方法与现有方法进行了定量和定性的对比分析，验证了其优越性。此外，本章还通过消融实验评估了各模块的贡献。

第四章：本章提出了一种空谱特征协同网络，旨在解决滨海湿地地物光谱特征相似性高、空间分布复杂导致的分类挑战。本章首先介绍基于空谱特征协同的整体框架设计，重点阐述了通道先验卷积注意力机制 GCN、像素选择模块以及局部光谱特征融合 Transformer 模块的结构与功能。最后通过一系列实验，验证了本文所提出模型的有效性。包括与现有方法的对比实验、消融实验以及参数敏感性分析。

第五章：本章对全文的研究内容进行了系统总结，归纳了本文在滨海湿地高光谱遥感分类领域的主要贡献与创新点。同时，针对当前研究中存在的不足，为后续研究提供了思路 and 参考。

2 相关技术

2 Related techniques

本章将首先对卷积神经网络的基本构成进行阐述；其次，深入探讨了语义分割技术；为了有效处理具有非欧几里得特性的空间数据，介绍了 GCN；为了能够建立图像 patch 间的长程依赖关系，介绍了 Vision Transformer；为了增强特征的判别能力，接着介绍了深度监督。

2.1 卷积神经网络（Convolutional Neural Network）

卷积神经网络（Convolutional Neural Network, CNN）是一种深度学习架构，这种网络架构具备独特的优势，能够自动捕捉和分析图像中的特征，而无需人为进行特征工程。CNN 可以有效的降低反馈神经网络(传统神经网络)的复杂性，常见的 CNN 结构有 LeNet-5、AlexNet、VGGNet、ResNet、DenseNet 等等；从这些结构来讲 CNN 发展的一个重要方向就是层次的增加，通过这种方式可以利用增加的非线性神经元得出目标函数的近似结构，同时得出更好的特征表达。CNN 的应用主要是在图像分类、目标检测以及图像分割等应用场景应用比较多。图 2.1 展示了卷积网络的示意图，一般由三种网络层组成，分别是输入层、隐藏层和输出层。

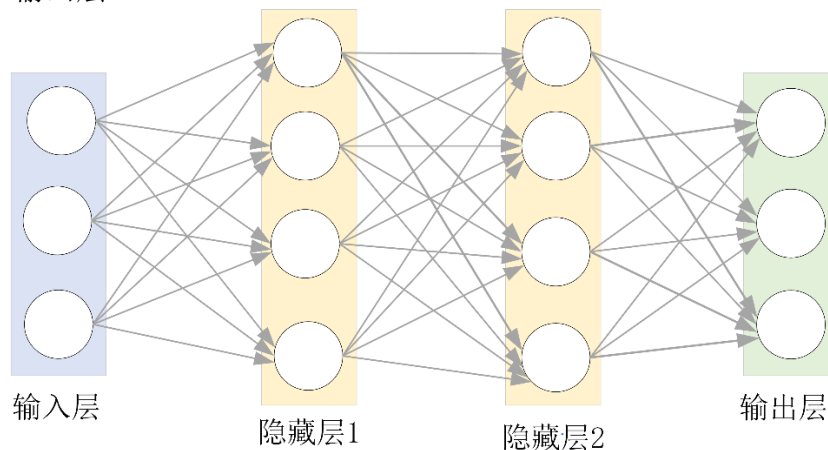


图 2.1 卷积神经网络示意图

Fig. 2.1 Schematic diagram of convolutional neural network

输入层：输入层是 CNN 的第一层，负责接收输入数据。对于图像数据，输入层通常是一个二维或三维数组，其主要任务是将原始数据输入到网络中，同时进行必要的预处理操作，常见的预处理方式主要有三种：去均值、标准化、PCA。CNN 中的隐藏层主要由卷积层、池化层、激活函数、全连接层组成。

卷积层：基于人脑的图片识别过程，可以认为图像的空间联系也是局部的像素联系比较紧密，而较远的像素相关性比较弱，所以每个神经元没有必要对全局图像进行感知，只要对局部进行感知，而在更高层次对局部的信息进行综合操作得出全局信息，即局部

感知。卷积层是 CNN 的核心，负责提取输入数据的特征。它由多个卷积核（或滤波器）组成，每个卷积核负责提取输入数据的一个特定特征。卷积核在输入数据上滑动，计算局部区域的加权和，生成特征图。卷积核的数量决定了输出特征图的数量，卷积核的大小决定了每个卷积核覆盖的输入数据区域大小，步长决定了卷积核在输入数据上滑动的距离，填充则是在输入数据边缘添加的零值，以控制输出特征图的大小。卷积操作的示意图如图 2.2 所示。

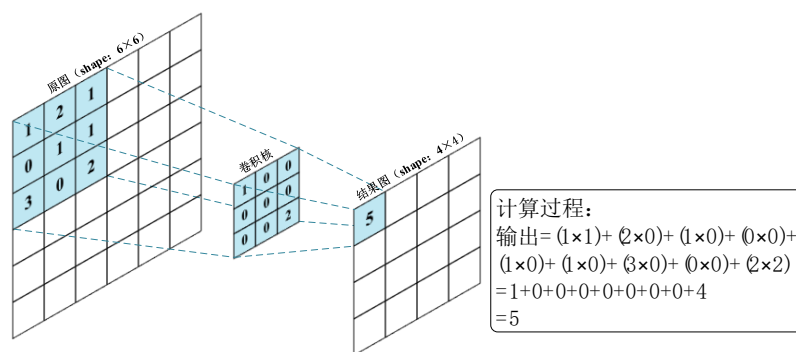


图 2.2 卷积操作示意图

Fig. 2.2 Schematic diagram of convolution operation

池化层：在连续的卷积层中间存在的就是池化层，主要功能是通过逐步减小表征的空间尺寸来减小参数量和网络中的计算，提取重要特征，删除冗余的噪音特征信息；池化层在每个特征图上独立操作。使用池化层可以压缩数据和参数的量，减小过拟合。在池化层中，进行压缩减少特征数量的时候一般采用两种策略：最大池化和平均池化。最大池化和平均池化示意图如图 2.3 所示。

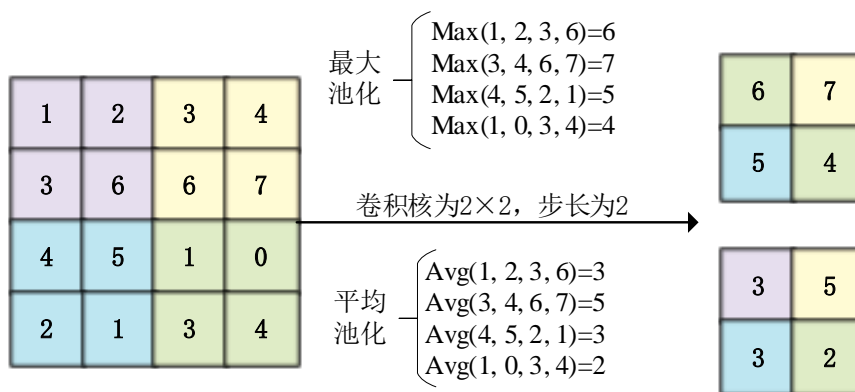


图 2.3 最大池化和平均池化示意图

Fig. 2.3 Schematic diagram of Max pooling and Average pooling

激活层：激活层则通过引入非线性特性，增强网络的表达能力，使模型能够学习复杂的模式和特征。图 2.3 所示为常见的四种激活函数的函数曲线。

全连接层：类似传统神经网络中的结构，FC 层中的神经元连接着之前层次的所有激活输出；换一句话说来讲的话，就是两层之间所有神经元都有权重连接；通常情况下，在 CNN 中，FC 层只会在尾部出现。全连接层示意图如图 2.4 所示。

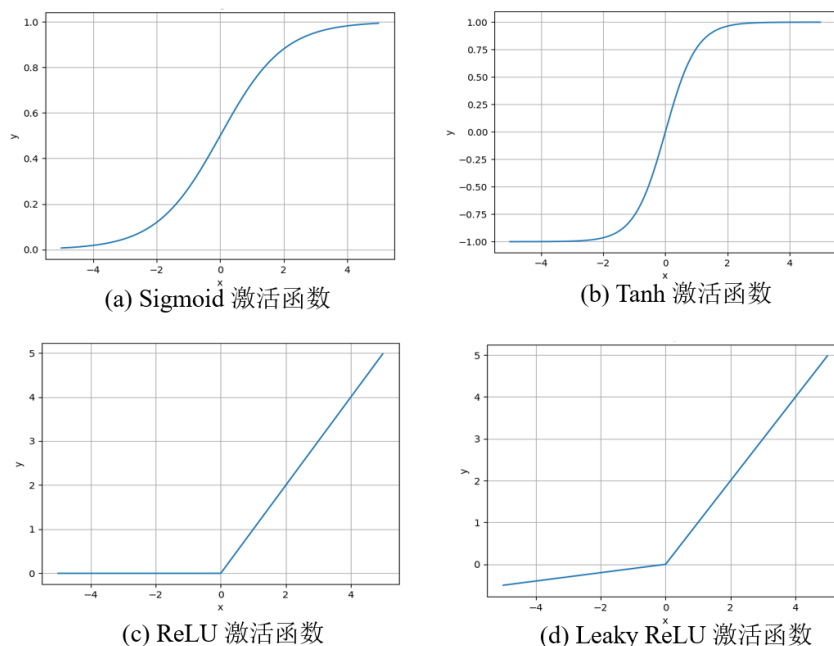
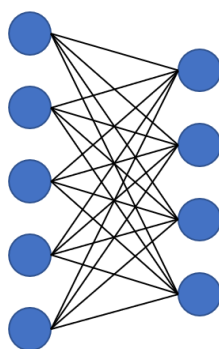


图 2.4 常见的激活函数示意图

Fig. 2.4 The schematic diagram of popular activation functions



网络层1 网络层2

图 2.5 全连接层示意图

Fig. 2.5 Schematic diagram of fully-connected layers

2.2 语义分割 (Semantic Segmentation)

语义分割 (Semantic Segmentation) 是很多计算机视觉任务的前提与基础, 旨在为图像中的每个像素分配一个语义类别标签, 从而实现对图像的精细理解和解析。与目标检测和图像分类不同, 语义分割不仅需要识别图像中的物体, 还需要精确地定位物体的边界和形状。语义分割的核心目标是将输入图像划分为多个具有语义意义的区域, 每个区域对应一个特定的类别, 如图 2.6 所示。随着深度学习技术的快速发展, 尤其是卷积神经网络 (CNN) 的出现, 图像语义分割取得了突破性进展。CNN 通过多层卷积和池化操作, 能够自动提取图像的层次化特征, 从而显著提升了语义分割的精度和鲁棒性。传统的图像分割算法主要依赖于手工设计的特征提取方法, 通过分析图像的颜色、纹理、边缘和空间分布等低级语

义信息来指导分割过程。然而，由于这些低级特征的表征能力有限，难以捕捉复杂场景中的上下文信息和高级语义关系，导致分割精度较低。相比之下，基于深度学习的图像分割技术可以自动学习图像的多层次特征表示，能够同时挖掘图像的高级语义信息和低级细节特征，从而实现更精确的分割效果。这种端到端的学习方式不仅显著提升了分割精度，还增强了模型对复杂场景的适应能力，为图像分割任务提供了更为强大的解决方案^[58,59]。



图 2.6 语义分割示意图

Fig. 2.6 Schematic diagram of Semantic Segmentation

2.2.1 经典的深度学习语义分割网络

U-Net: U-Net 是一种经典的卷积神经网络架构，由 Olaf Ronneberger 等人^[60]在 2015 年首次提出，最初主要用于生物医学图像分割任务。凭借其出色的性能表现和灵活的网络结构，U-Net 迅速成为语义分割领域的基准模型，并被广泛应用于医学影像分析和遥感图像解译等多个领域。U-Net 的整体结构呈对称的“U”形，由编码器和解码器两部分构成，通过跳跃连接实现编码器和解码器特征图的多层次融合。这种创新的架构设计通过多层次特征提取与融合机制，能够同时捕获图像的全局上下文信息和局部细节特征，从而实现对高分辨率图像的精细分割。U-Net 的对称结构不仅使其在处理复杂场景时表现出色，还因其高效的端到端训练方式和小样本适应能力，成为许多分割任务的首选模型。自 U-Net 提出以来，研究者们针对其结构和功能进行了多种改进，以适应不同的应用场景和任务需求。例如，U-Net++，3D U-Net，Attention U-Net 等。U-Net 结构示意图如图 2.7 所示。

DeepLabv3+: DeepLabv3+ 是由 Google Research 团队^[61]在 2018 年提出的一种先进的语义分割网络，是 DeepLab 系列的最新版本。它在 DeepLabv3 的基础上引入了编码器-解码器结构，进一步提升了分割精度，尤其是在目标边界细节的处理上表现出色。DeepLabv3+ 结合了空洞卷积和空间金字塔池化等关键技术。DeepLabv3+ 的整体结构由编码器和解码器两部分组成，其核心设计思想是通过多尺度特征提取和精细的特征融合，实现对图像上下文信息和细节信息的充分利用。DeepLabv3+结构示意图如图 2.6 所示。

ResNet: 在 ResNet 提出之前, 深度神经网络的性能随着层数的增加而提升, 但当网络深度达到一定程度时, 性能反而会下降, 这种现象被称为网络退化问题。其主要原因是梯度消失和梯度爆炸问题, 导致深层网络难以有效训练。ResNet 是由微软研究院的 Kaiming He 等人^[62]在 2015 年提出的一种深度卷积神经网络架构。其核心创新在于引入了残差学习机制, 通过跳跃连接解决了深层网络中的梯度消失和退化问题, 使得训练极深的神经网络成为可能。ResNet 通过引入残差学习机制, 允许网络直接学习输入与输出之间的残差, 而不是直接学习复杂的映射关系。根据网络深度的不同, ResNet 衍生出多种变体, 如 ResNet-18、ResNet-34、ResNet-50 等。ResNet 结构示意图如图 2.9 所示。

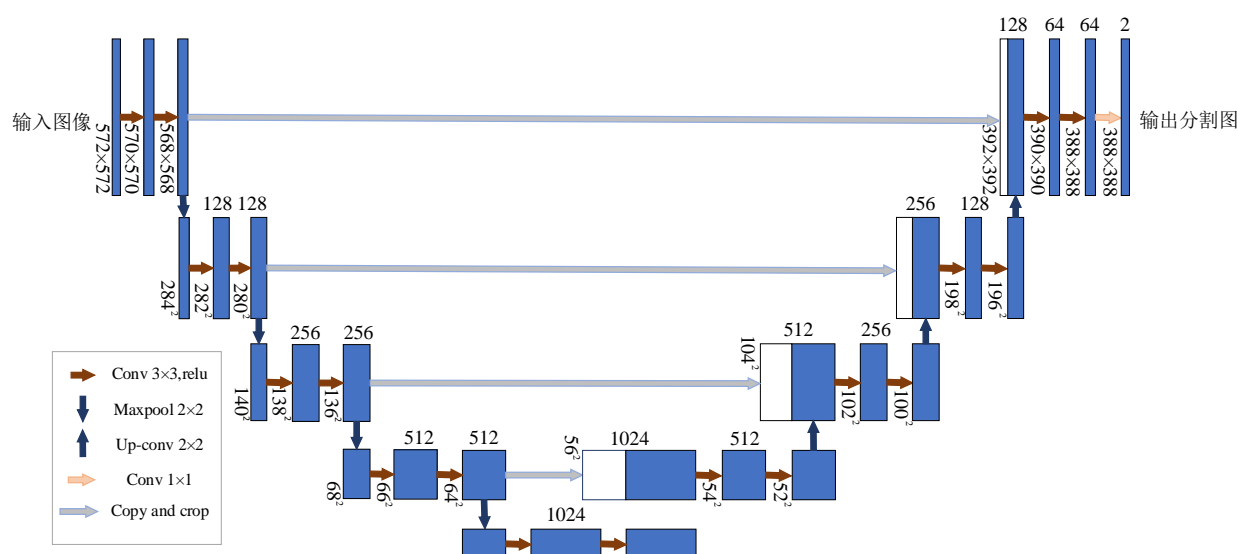


图 2.7 U-Net 结构示意图

Fig. 2.7 The structure diagram of U-Net

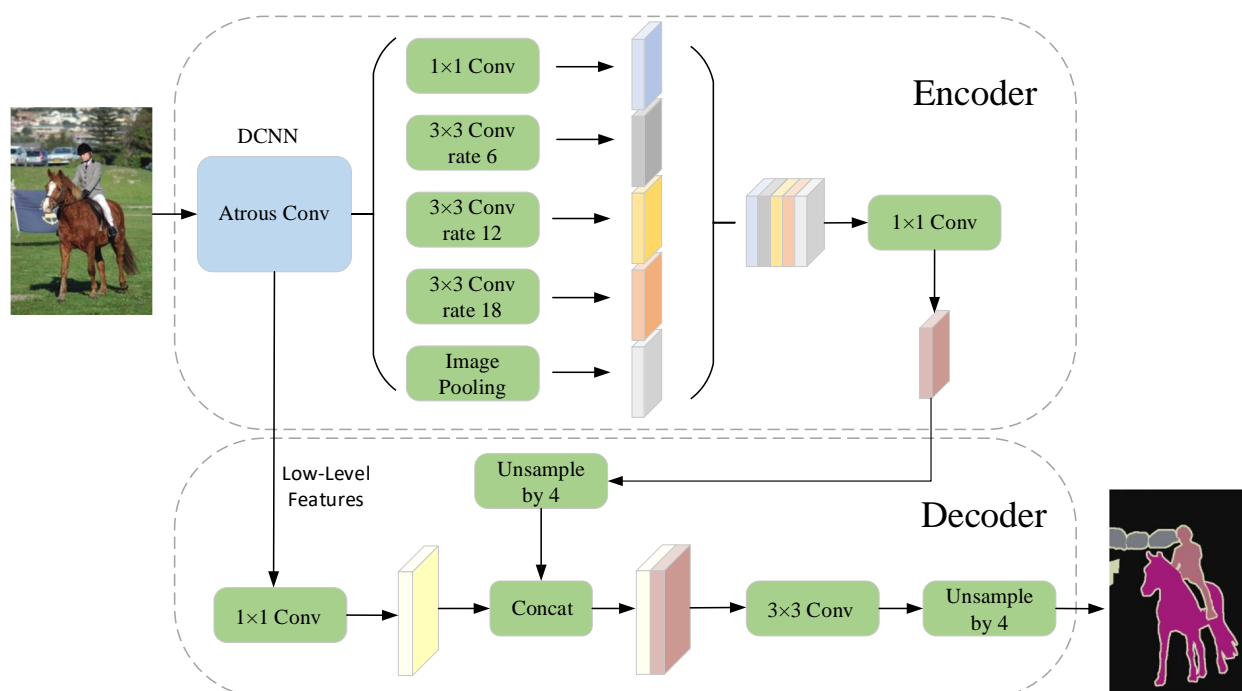


图 2.8 Deeplab3+结构示意图

Fig. 2.8 The structure diagram of Deeplab3+

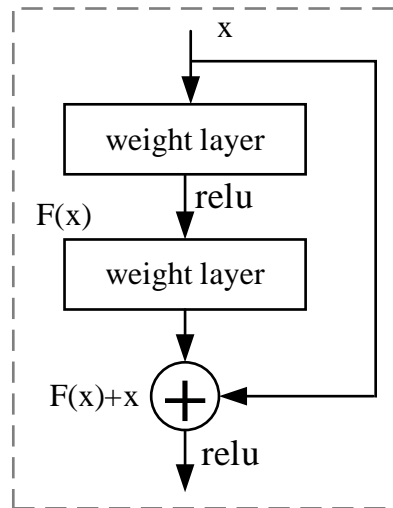


图 2.9 ResNet 的残差学习模块示意图

Fig. 2.9 Schematic diagram of the residual learning module of ResNet

2.3 图卷积神经网络（Graph Convolutional Neural Networks）

GCN 是一种能够直接作用于图并且利用其结构信息的卷积神经网络，其中包括空域图卷积和谱域图卷积两种方法。GCN 的核心思想是将卷积操作从传统数据（如图像或网格）推广到图数据。在图数据中，节点通过边相互连接，形成复杂的网络结构。如图 2.10 所示，图是由节点（vertices）和边（edges）组成的数据结构。GCN 通过聚合每个节点及其邻节点的特征信息，生成节点的有效表示，进而用于各种下游任务。GCN 能够将 HSI 表示为无向图，从而显式地建模不同像素点之间的空间关系。通过将图像中的每个像素视为图中的一个节点，并利用像素之间的空间邻接关系构建边，GCN 能够有效地捕捉像素间的空间依赖性，进而提升像素级分类的性能。具体而言，给定一个图 $G=(V,E,A)$ ，其中 V 表示图中顶点的集合， $|V|=N$ ， E 为图的边集， $A \in R^{N \times N}$ 为图的邻接矩阵，用于描述节点之间的连接关系。通过这种图表示方法，GCN 能够利用像素点之间的空间邻域信息，进一步增强分类模型对空间上下文信息的建模能力，从而显著提高分类精度。

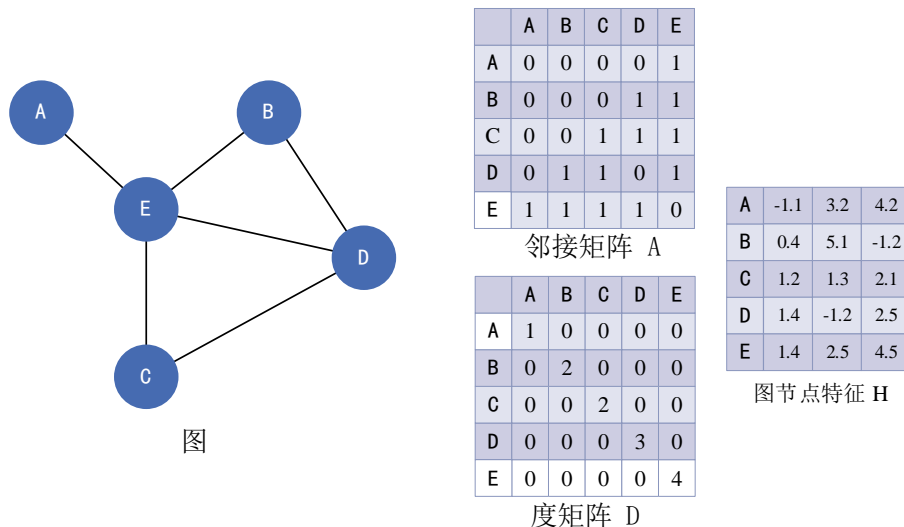


图 2.10 图卷积结构示意图

Fig. 2.10 The structure diagram of GCN

给定邻接矩阵 A 和度矩阵 D ，图的拉普拉斯算子可表示为如公式(2.1)所示：

$$L = D - A \quad (2.1)$$

将拉普拉斯矩阵归一化，可表示为如公式(2.2)所示：

$$L_n = I - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \quad (2.2)$$

其中， I 为单位矩阵， L_n 为对应的对称归一化拉普拉斯矩阵。

为了嵌入节点的表示，引入了卷积核 $g_\theta = \text{dig}(\theta)$ ，其中 $\theta \in R^N$ ；图的卷积公式可表示为如公式(2.3)所示：

$$x * g_\theta = U g_\theta(\Lambda) U^T x \quad (2.3)$$

其中 x 是输入信号， U 是 L_n 的特征向量矩阵， Λ 是 L_n 的特征值， g_θ 可以看作是特征值 Λ 的函数。为了简化计算，采用切比雪夫 Chebyshev polynomial 多项式来进行近似计算， $g_\theta(\Lambda)$ 可表示为如公式(2.4)所示：

$$g_\theta(\Lambda) \approx \sum_{k=0}^K \theta_k T_k(L_n) \quad (2.4)$$

其中 T_k 是切比雪夫多项式。

取一阶切比雪夫多项式进行进一步简化，可表示为如公式(2.5)所示：

$$x * g_\theta = \theta(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}}) x \quad (2.5)$$

其中 $\tilde{A} = A + I$ 和 $\tilde{D} = \sum_j \tilde{A}_{ij}$ 。

基于(2.5)，给出了 GCN 的传播规则，可表示为如公式(2.6)所示：

$$H^{(l+1)} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)}) \quad (2.6)$$

其中 $H^{(l+1)}$ 和 $H^{(l)}$ 分别表示第 $l+1$ 层和第 l 层的值， σ 表示激活函数， $W^{(l)}$ 表示第 l 层的权重矩阵。

2.4 Vision Transformer 网络 (Vision Transformer)

Vision Transformer (ViT) 是一种基于 Transformer 架构的图像分类模型，首次将

Transformer 成功应用于计算机视觉任务。与传统的卷积神经网络（CNN）不同，ViT 完全摒弃了卷积操作，而是通过将图像分割成固定大小的图像块（patches），并将这些图像块线性嵌入为序列化的向量，从而将图像数据转换为适合 Transformer 处理的序列形式。在 HSI 中，光谱序列可以被视为一种典型的序列数据结构。Transformer 模型因其在提取序列特征和构建全局注意力机制方面的卓越能力，从而实现对图像的高效建模。ViT 的主要结构包括以下几个部分：

（1）图像分块与线性嵌入（Patch Embedding）

输入图像 $X \in R^{H \times W \times C}$ （其中 H 和 W 分别表示图像的高度和宽度， C 表示通道数）被分割成 N 个固定大小的图像块，每个图像块的大小为 $P \times P$ 。这些图像块被展平为 N 个向量，其中 $N = \frac{H \times W}{P^2}$ 。然后，每个图像块通过一个线性投影层映射到一个 D 维的嵌入空间，得到图像块的嵌入表示 $z_p \in R^D$ 。

（2）位置编码（Positional Encoding）

由于 Transformer 本身不具备对输入序列顺序的感知能力，ViT 引入了位置编码来保留图像块的空间位置信息。位置编码被添加到图像块嵌入中，以帮助模型理解图像块之间的相对位置关系。

（3）Transformer 编码器（Transformer Encoder）

ViT 使用标准的 Transformer 编码器来处理图像块序列。Transformer 编码器由多头自注意力机制（Multi-Head Self-Attention, MHSA）和前馈神经网络（Feed-Forward Network, FFN）组成。通过多层堆叠，Transformer 编码器能够捕捉图像块之间的全局依赖关系

注意力机制如图 2.11 所示。假设输入序列为 x ，其长度为 m 。首先，通过将位置编码向量与输入序列 x 结合，得到输入向量 p 。随后，序列数据分别通过共享的权重矩阵 W_q 、 W_k 和 W_v 进行线性变换，分别映射为查询向量（Query, Q ）、键向量（Key, K ）和值向量（Value, V ）。基于此，自注意力层（Self-Attention, SA）的计算可以表示为如公式(2.7)所示：

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2.7)$$

其中 d_k 为键向量的维度， $\text{softmax}(\cdot)$ 函数用于计算注意力权重。 Z 是 SA 层的输出，如果有 N 个 SA 层， Z 可以记为 Z_i , $i=1,2, \dots, N$ 。MHSA 的输出可以表示为如公式(2.8)所示：

$$Z_{MHSA} = MHSA(Z_i) = \text{Concat}(Z_1, Z_2, \dots, Z_N)W' \quad (2.8)$$

其中 Z_{MHSA} 是 MHSA 的输出， W' 为投影矩阵。

Transformer 模型通过 MHSA 和多层感知器（MLP）的结合，并辅以层归一化和残差

连接来显著提升模型的性能。因此变压器中的编码模块可以表示为如公式(2.9)(2.10)所示：

$$Z' = MHSA(LN(p)) + p \quad (2.9)$$

$$Z = MLP(LN(Z')) + Z' \quad (2.10)$$

其中 LN 表示层归一化， Z' 和 Z 分别表示 MHSA 和 MLP 的输出。

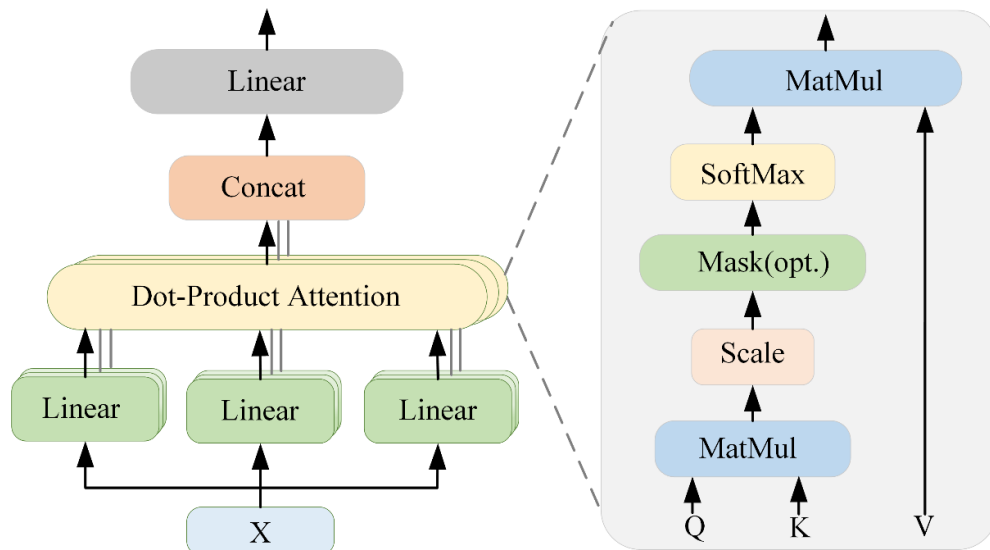


图 2.11 Transformer 中的注意力机制示意图 (左) 多头自注意力 (右) 自注意力
Fig. 2.11 Schematic diagram of attentive mechanism in Transformer. (Left)MHSA (Right)SA

2.5 深度监督（Deep Supervision）

深度监督（Deep Supervision），这一由 Gallagher 等人在 2014 年开创性的深度学习训练技巧，其核心目的在于显著提升模型的性能与训练效率，同时有效应对深度神经网络中普遍存在的梯度消失挑战。在传统卷积神经网络（CNN）的架构中，监督信号往往被局限地施加于网络的末端，即通过一个主分类器来实现对特征的最终判别与分类。然而，深度监督策略则突破了这一局限，它倡导在模型设计的初期阶段就进行深入的规划，精心挑选出 CNN 中的若干个关键中间卷积层，并为这些层分别增设辅助分类器，以此构建一个更为全面且深入的监督体系。这一策略的核心精髓在于，通过在网络内部的多个层次引入监督信号，即辅助损失，来促使模型在训练过程中能够更深入地挖掘并学习到更为丰富且具有高度判别性的特征表示。这些辅助损失在网络的优化过程中发挥着至关重要的作用，它们不仅引导网络向正确的方向前进，还确保了信息在深层网络中的有效传递与特征的精准提取，从而有效规避了梯度消失的问题，显著加速了模型的收敛速度。

深度监督与多任务学习（Multi-task Learning）在形式上虽有一定的相似性，但二者在本质上存在显著差异。在深度监督的框架下，所有的辅助损失与模型最终层的主损失均共享同一个真实标签，而多任务学习则通常涉及多个相关但独立的任务，每个任务可能拥有其独特的标签集和训练目标。最终，网络模型的总训练损失是通过对所有辅助损失与主损

失进行加权求和而得到的，权重的分配可以根据具体的任务需求、模型复杂度以及实验效果等因素进行灵活调整，以达到最佳的模型训练效果。深度监督结构示例如图 2.12 所示，其在网络中间添加了一个分类器进行辅助监督。

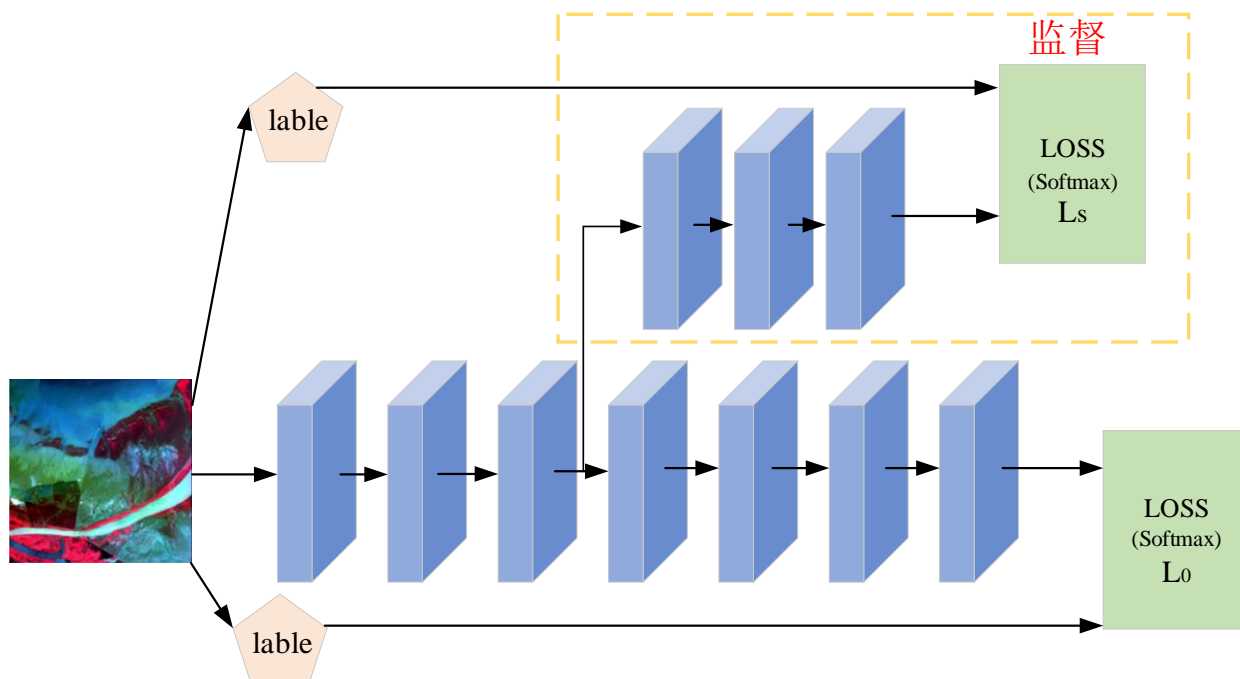


图 2.12 深度监督结构示意图

Fig. 2.12 Schematic diagram of deep supervision structure

2.6 本章小结 (Chapter Summary)

本章系统阐述了本研究工作的理论基础与技术支撑。首先，详细解析了卷积神经网络 (CNN) 的整体架构及其内部组成机制，包括卷积层、池化层和激活层等核心组件的工作原理。其次，深入探讨了语义分割技术、图卷积神经网络 (GCN) 的图结构特征提取机制、Vision Transformer (ViT) 的自注意力机制网络以及深度监督技术的优化策略，为本文的研究提供了坚实的理论支撑和方法论指导。

3 基于多分支多层次特征交互的滨海湿地精细制图方法

3 A fine mapping method of coastal wetland based on multi-branch multi-level feature interaction

3.1 引言 (Foreword)

滨海湿地作为陆地与海洋之间的过渡带，是一个高度复杂且动态变化的生态系统，其地物组成多样，包括多种类型的植被（如互花米草、芦苇等）、水体等。这些地物在空间分布、形态结构和光谱特性上存在显著差异，导致它们在遥感影像中呈现出多尺度的特征。例如，大面积的植被区域或广阔的水域在影像中通常具有较高的光谱响应和明显的纹理特征，容易在视觉上主导整个场景；而小尺度地物，如孤立的小水体或零散的植被斑块，由于其空间范围较小、光谱信息较弱，往往容易被大尺度地物的特征所掩盖，从而在影像解译过程中被忽略或误分类。这种尺度差异不仅增加了滨海湿地地物分类的难度，还可能影响生态监测和环境评估的准确性。因此，如何有效处理遥感影像中的多尺度特征，特别是增强对小尺度地物的识别能力，成为滨海湿地遥感研究中的关键问题。针对这一问题，本章提出了一种多分支多层次特征交互网络（multi-branch multi-level feature interaction, MMFNet），旨在增强模型对滨海湿地遥感影像中多尺度地物特征的辨识与提取能力。具体而言，为了更好地捕捉和融合不同尺度的特征信息，本文引入了一种多尺度注意聚合（Multi-Scale Attention Aggregation, MSAA）模块，同时，帮助解码器通过跳层连接，能够将多尺度特征信息直接传递到解码器，显著提升了模型对复杂地物结构的重建能力和分类精度。设计了多分支解码结构，通过构建层次化特征交互机制，实现了多尺度特征信息的自适应融合与优化，有效提升了网络的特征提取能力和解码精度。该结构将前一个分支生成的特征映射动态反馈到后续分支的各个层次，实现了特征信息的跨层次传递与交互。

3.2 MMFNet 网络结构设计 (MMFNet network structure design)

3.2.1 MMFNet 的整体架构

本文提出了多分支多层次特征交互网络 MMFNet，其总体架构如图 3.1 所示。MMFNet 编码器运用了 ResNet50 架构以捕获图像的多级特征信息。在此基础上，创新性地引入了多尺度注意力聚合模块（MSAA），该模块将来自 CNN 编码器的不同阶段的特征进行拼接，然后通过多尺度卷积对这些特征进行处理，从而提取不同尺度下的细节信息。解码器包含三个并行分支，有效地融合了多尺度特征。具体而言，为了将经过 MSAA 模块精炼的特征图精确映射回原始图像的尺寸，每个分支都精心配备了上采样块（UB）。通过跳层连接来融合不同分支的特征，并将同一层次上的不同分支的特征维度调整为相同。在跳层连接之前，巧妙地嵌入了基于残差单元的全局注意力机制模块（GAM-RU），这一设计不仅促进

了特征之间的深度融合，还进一步提升了网络的整体性能。最终，在 MMFINet 的每个分支的末端，运用了辅助损失函数进行深度监督，充分利用并综合三个分支的各自优势，从而得出更加精准预测结果。

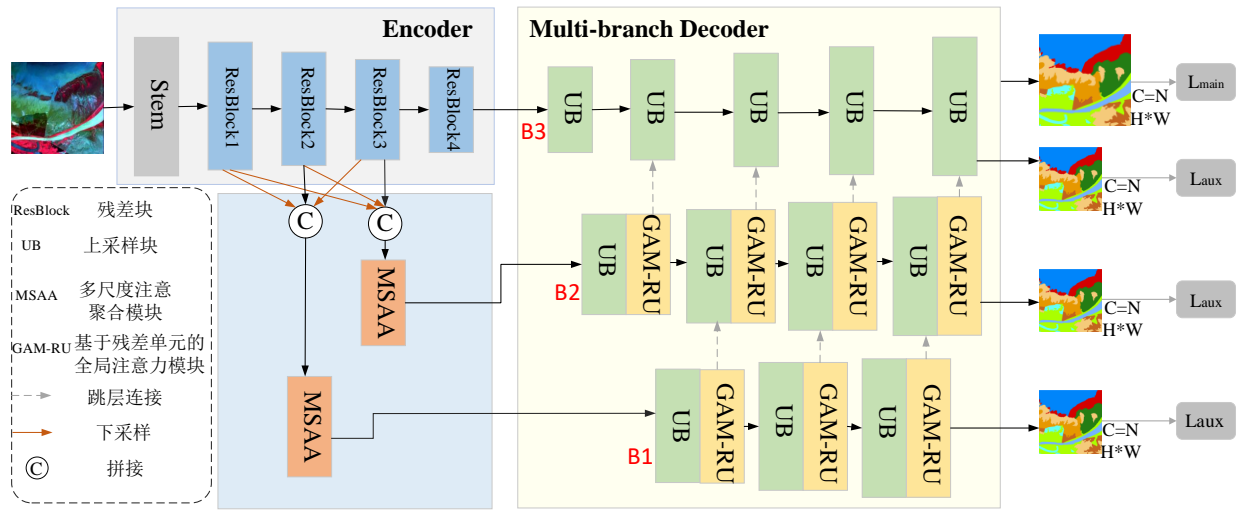


图 3.1 MMFINet 网络整体架构

Fig. 3.1 Overall structure of MMFINet network

3.2.1 编码器网络

本研究以 ResNet50 模型为基础构建编码器，并利用 ImageNet 数据集进行预训练，以充分提取图像的低级特征和高级特征。编码器的设计基于预训练的 ResNet50 模型，通过卷积层、批量归一化（BN）、ReLU 激活函数和池化操作逐步提取特征。编码器的工作过程大概可以表述为：输入图像首先通过编码器的第一个模块，提取最浅层的特征。然后，这些特征通过池化操作降低空间维度，并输入到下一个模块中。每个模块逐步提取更高级的特征，直到最后一个模块提取全局语义信息。此外，在 ImageNet 数据集上进行预训练的模型也可以加快网络的收敛速度，并提高模型的性能。编码器的输出是多尺度的特征图，这些特征图将被输入到多分支解码器中，用于进一步的特征融合。有关编码器的详细信息，如表 3.1 所示。

表 3.1 编码器细节

Table 3.1 Details of encoder

	Stem	ResBlock1	ResBlock2	ResBlock3	ResBlock4
ResNet50	$7 \times 7, 64, \text{stride} = 2$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
Size (H×W)	$\frac{H}{2} \times \frac{W}{2}$	$\frac{H}{4} \times \frac{W}{4}$	$\frac{H}{8} \times \frac{W}{8}$	$\frac{H}{16} \times \frac{W}{16}$	$\frac{H}{32} \times \frac{W}{32}$

3.2.2 多尺度注意力聚合模块 (MSAA)

MSAA 的结构如图 3.2 所示, 如来自 ResNet 编码器级 F_1 、 F_2 和 F_3 的输出被连接为 $\hat{F}_i = \text{Concat}(F_i, F_{i-1}, F_{i+1})$ 。将组合特征 $\hat{F} \in R^{C_1 \times H \times W}$ 输入 MSAA 进行进一步的精细化处理。在 MSAA 中, 是通过空间和通道两条并行路径来实现的。空间细化路径首先从通道投影开始, 利用一个 1×1 的卷积操作将通道数从 C_1 减少到 C_2 , 其中 C_2 是 C_1 的一个比例因子 α 所决定的值。接下来, 多尺度融合策略被采用, 它涵盖了不同内核大小 (如 3×3 、 5×5 、 7×7) 的卷积操作, 以捕捉丰富的空间信息。随后, 通过平均池化和最大池化对空间特征进行进一步的聚合, 并将结果与经过 sigmoid 激活的特征图进行 7×7 卷积和元素级的乘法运算, 从而增强空间特征的表达能力。与此同时, 在通道聚合路径上, 全局平均池化被用来将特征图的维度降低至 $C_1 \times 1 \times 1$ 。紧接着, 一个 1×1 的卷积操作和 ReLU 激活函数被用来生成通道注意力图。这个注意力图随后被扩展以匹配原始输入特征的维度, 并与空间细化路径上得到的特征图进行融合。因此, MSAA 模块通过并行处理空间和通道信息, 显著增强了后续网络层的空间和通道特征。最终, 通过整合 MSAA 模块的输出, 得到了经过优化和增强的特征表示。

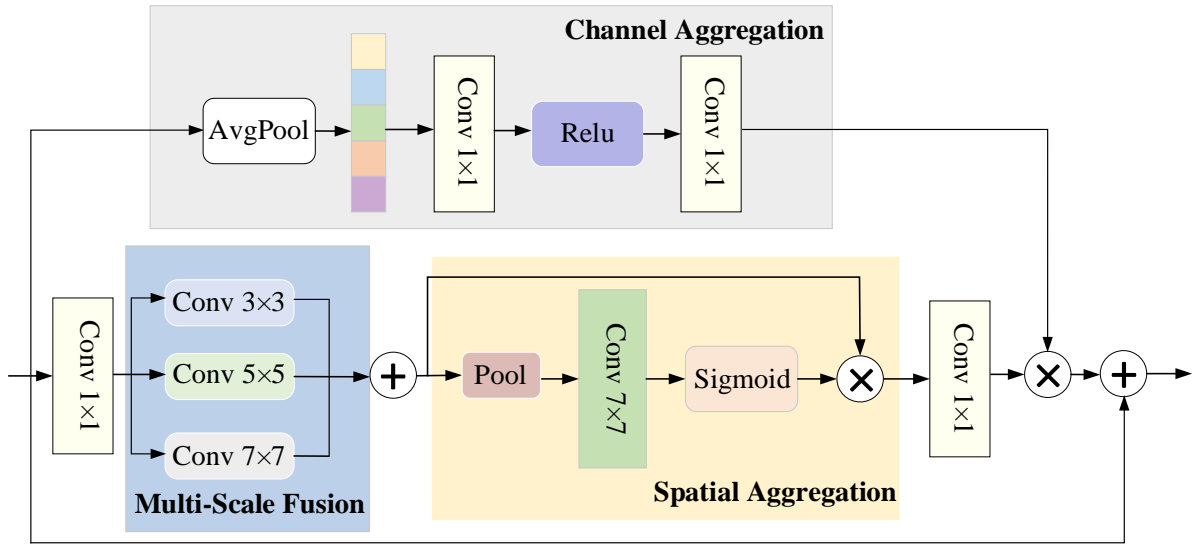


图 3.2 多尺度聚合模块结构

Fig.3.2 Structure of Multi-Scale Attention Aggregation (MSAA) module

3.2.3 多分支解码器

多分支解码器架构巧妙地融合了三个并行且深度各异的分支, 这些分支核心组件包括上采样块 (UB) 与多分支融合模块 (MBFM)。UB 的结构如图 3.3 所示, 采用阶梯式结构设计, 每个阶段处理后图像尺寸逐级翻倍, 同时通过 UB 进一步细化特征图的细节。在上采样块内部, 首先, 采用双线性插值技术提升图像分辨率, 随后, 引入一个由两个 3×3 卷积核组成的细化模块。过程表示如公式(3.1)所示:

$$F_{out} = v(Conv_{3 \times 3}(\delta(Conv_{3 \times 3}(Upsample(F_{in})))))) \quad (4.2)$$

其中, $Upsample$ 为参数为 2 的双线性插值, $Conv_{3 \times 3}$ 为卷积操作 (核大小为 3×3), $\delta(\cdot)$ 为 BN 和 ReLU 激活函数, $v(\cdot)$ 为 BN 和 PReLU 激活函数。

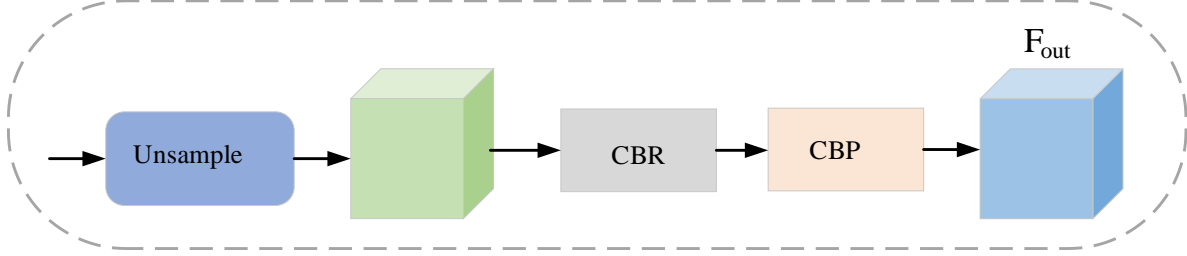


图 3.3 上采样块结构

Fig.3.3 Structure of unsample block

为了充分利用多尺度信息的优势, MBFM 创新性地引入了跳跃连接机制, 这一设计使得不同分支间的特征图得以有效融合和交互, 从而增强特征的细节信息和多尺度信息的利用。具体而言, 解码器的三个分支[B1、B2、B3] (如图 3.1 所示) 通过巧妙的连接策略交互工作: B1 的反馈输出被巧妙地融入 B2 的输入端, 随后, 增强后的特征被传递至同级的 B3 分支。尽管跳跃连接已在一定程度上促进了相邻特征图的融合, 但所得融合特征仍显粗糙。为了进一步增强特征融合的效果, 本研究在跳跃连接之前引入了基于残差单元的全局注意力机制模块 (GAM-RU), GAM-RU 的结构如图 3.4 所示。该模块由空间注意力、通道注意力、特征加权及残差连接三大组件构成。在空间注意力部分, 输入特征 F 首先经过一个 3×3 卷积层进行特征提取, 随后沿通道轴分别执行最大池化和平均池化操作。将所得特征进行拼接, 生成富含信息的特征映射。接着, 通过一个 3×3 卷积层及 sigmoid 激活函数, 精炼出最终的空间权重矩阵。通道注意力的处理流程与空间注意力类似, 但不同之处在于其沿空间维度执行平均池化和最大池化。在最终阶段, 输入特征 F 被同时赋予空间注意力和通道注意力权重, 随后通过跳跃连接将加权后的特征与原始输入特征 F 相结合。整个过程可表示如公式(3.2) (3.3) (3.4)所示:

$$G_s = \sigma(Conv_{3 \times 3}([\max pool(F); avgpool(F)])) = \sigma(Conv_{3 \times 3}([F_{\max}^s; F_{avg}^s])) \quad (4.8)$$

$$G_c = \sigma(Conv_{3 \times 3}([\max pool(F); avgpool(F)])) = \sigma(Conv_{3 \times 3}([F_{\max}^c; F_{avg}^c])) \quad (4.9)$$

$$F_{att} = F \times G_s \times G_c + F \quad (4.10)$$

其中 $\sigma(\cdot)$ 为 sigmoid 激活函数, $Conv_{3 \times 3}$ 为卷积操作 (核大小为 3×3), $\max pool(\cdot)$ 为最大池化操作, $avgpool(\cdot)$ 为平均池化操作。

各分支特征的融合方式计算程序可表示如公式(3.2) (3.3) (3.4)所示:

$$f_{B1} = \delta(Conv_{k \times k, s}[f_{in-B1}]) \quad (4.8)$$

$$f_{B2} = \delta(\text{Conv}_{k \times k, s}[f_{in-B1} \oplus f_{in-B2}]) \quad (4.9)$$

$$f_{B3} = \delta(\text{Conv}_{k \times k, s}[f_{in-B2} \oplus f_{in-B3}]) \quad (4.10)$$

其中 f_{in-Bi} ($i=1、2、3$) 表示为当前分支 Bi 的特征图, f_{Bi} ($i=1、2、3$) 表示的为分支 Bi 融合之后的特征图, \oplus 表示拼接操作, $\text{Conv}_{k \times k, s}$ 表示卷积操作 (核大小为 $k \times k$, 步幅为 s), $\delta(\cdot)$ 为 BN 和 ReLU 激活函数。

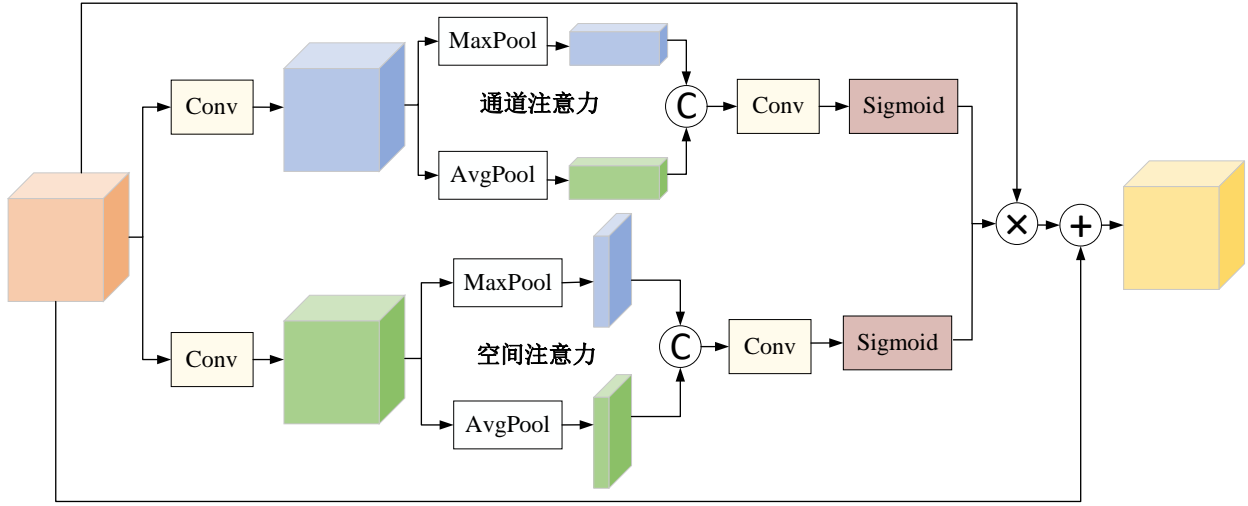


图 3.4 基于残差单元的全局注意力机制模块结构

Fig.3.4 Structure of global attention mechanism module based on residual units

3.2.4 损失函数

采用深度监督从 MMFINet 各分支的最后一层获取输出, 如图 3.1 所示, 综合三个分支, 得到综合结果, 提高分类精度, 对于最后输出可表示为如公式(3.1)所示:

$$p' = \text{Conv}(\sum_{i=1}^3 F^{Bi}) \quad (3.1)$$

其中, F^{Bi} 表示三个分支的特征, Conv 模块通过卷积操作将输入特征映射转换为具有 C 个通道的输出特征图, 其中每个通道对应一个类别的预测结果。总的来说, 最终的分类结果与粗分类结果具有相同的损失函数, 是通过使用标准的交叉熵损失来进行训练的。为了更好地结合主损失 L_{main} 和辅助损失 L_{aux} , 将 L_{aux} 乘以系数 λ 。因此, 总损失可以表示如公式(3.2)所示:

$$L = L_{main} + \lambda(L_{aux1} + L_{aux2} + L_{aux3}) \quad (3.2)$$

其中, λ 是一个超参数, 用于平衡两个损失函数的影响。本章将 λ 设置为 0.5。 L_{main} 和 L_{aux} 均等于 L_{ce} , L_{aux1} 、 L_{aux2} 和 L_{aux3} 是三个分支的粗分割结果。

3.3 实验准备 (Experimental preparation)

3.3.1 实验数据

(1) 遥感图像数据

2018 年 8 月 2 日通过 Zhuhai-1 号卫星在山东省东营垦利区黄河河口湿地采集的 HSI, 光谱波段的波长范围为 400~1000nm。空间分辨率为 10 m, 光谱分辨率为 2.5 nm。图 2.11 展示了该数据集的假彩色图及其对应的地面真值参考图。2013 年至 2021 年间, 研究团队对黄河河口沿海湿地开展了十余次系统性实地考察, 旨在精确监测该区域湿地植被的时空分布特征及生长动态。基于实地调查数据与同期高分辨率遥感影像的协同分析, 研究团队采用人机交互解译方法提取了土地覆盖类型信息, 以确保分类结果的可靠性。根据黄河河口湿地的典型生态系统特征, 研究将土地覆盖类型划分为 9 个类别: 互花米草、混生区、怪柳、芦苇、盐地碱蓬、潮滩、黄河、池塘和海域。图 3.5 展示了该数据集的假彩色图及其对应的地面真值参考图。

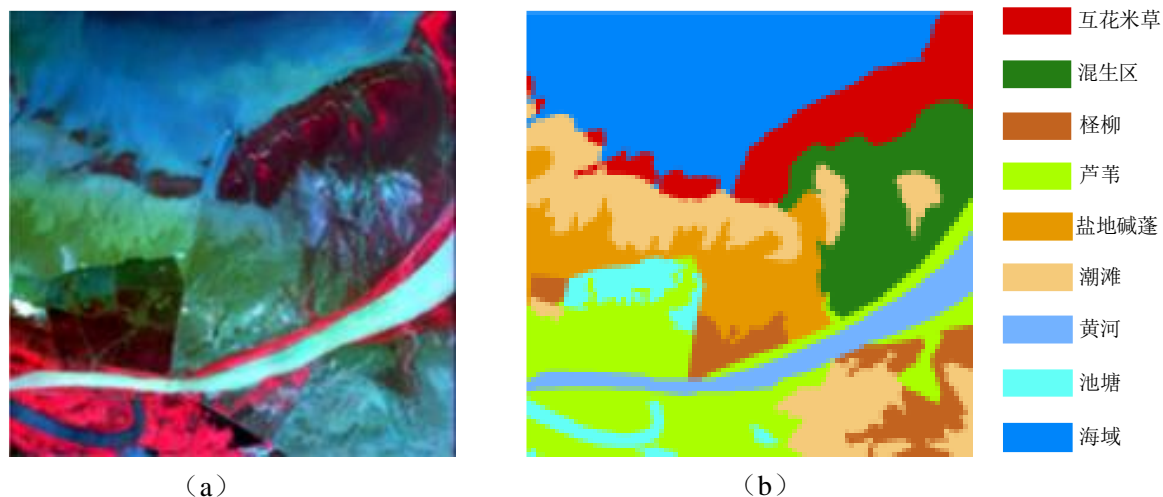


图 3.5 “Zhuhai-1 号” 卫星高光谱数据集 (a) 假彩色图像 (b) 地面真值图

Fig. 3.5 “Zhuhai-1” satellite hyperspectral dataset (a) False color composite (b) Ground truth

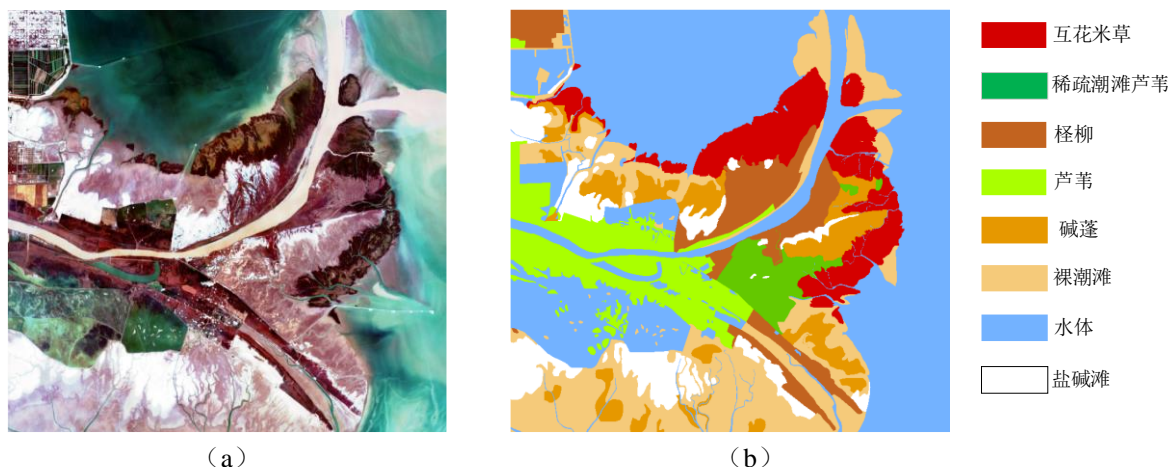


图 3.6 “高分五号” 卫星高光谱数据集 (a) 真彩色图像 (b) 地面真值图

Fig. 3.6 “GF-5” satellite hyperspectral dataset (a) True color composite (b) Ground truth

影像数据由我国"高分五号"卫星于 2018 年 11 月 1 日采集, 拍摄区域为黄河三角洲。"高分五号"卫星是中国高分辨率对地观测系统重大专项的重要组成部分, 是该系列中的第五颗业务卫星, 该卫星于 2018 年 5 月 9 日成功发射升空。光谱波段的波长范围为 390~2513nm, 空间分辨率是 30m, 光谱分辨率为 5nm@390~1029, 10nm@1004~2513nm。根据高分五号遥感影像特点, 结合实地踏勘数据与专家目视解译结果, 本研究地物类型精细划分为 8 个类别, 分别为: 互花米草、稀疏潮滩芦苇、柽柳、芦苇、碱蓬、裸潮滩、水体、盐碱滩。图 3.6 展示了该数据集的真彩色图及其对应的地面真值参考图。

(2) 数据集制作

表 3.2 “Zhuhai-1” 卫星高光谱数据集中使用的训练和测试集的数量

Table 3.2 Number of training and test samples used in the Zhuhai-1 satellite hyperspectral dataset

类别编号	地面覆盖类型	训练样本	测试样本
1	互花米草	11567	84828
2	芦苇	21937	160878
3	柽柳	7651	56108
4	盐地碱蓬	12161	89186
5	混生区	14159	103848
6	潮滩	17499	128330
7	黄河	6919	50745
8	池塘	4818	25340
9	海域	29113	213499
	总共	125824	912762

表 3.3 “高分五号”卫星高光谱数据集中使用的训练和测试集的数量

Table 3.3 Number of training and test samples used in the GF-5 satellite hyperspectral dataset

类别编号	地面覆盖类型	训练样本	测试样本
1	互花米草	5639	41359
2	芦苇	5762	42262
3	柽柳	4531	33229
4	盐地碱蓬	4190	30733
5	稀疏潮滩芦苇	1888	13852
6	裸潮滩	11318	83000
7	水体	30718	225268
8	盐碱滩	4326	31725
	总共	68372	501428

针对滨海湿地地物种类繁多、分布不均的特点, 为避免全图采样导致的模型过拟合问题, 本研究采用随机采样策略对整幅图像进行处理, 以确保各类地物及其交界区域均能被均衡选取。在构建地面真实样本数据集时, 针对每个地物类别, 随机抽取占该类别总样本量 12% 的样本及其对应的真实标注图作为训练集, 并进一步将图像裁剪为 9×9 像素的

patch 块，而将剩余样本留作测试集。表 3.2-3.3 详细说明了在“Zhuhai-1 号”和“高分五号”卫星高光谱遥感影像两个数据集中的样本划分。

3.3.2 实验评价指标

本研究采用基于混淆矩阵的精度评价体系对滨海湿地精细分类结果进行系统性性能评估，通过构建完整的分类误差矩阵，计算总体分类精度（OA）、平均分类精度（AA）和 Kappa 系数三项指标，实现对分类结果的定量化评价与误差来源分析。混淆矩阵（Confusion Matrix），也称为误差矩阵，是遥感影像分类精度评价中的重要工具，用于直观展示分类果与参考数据之间的对应关系。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 3.4 所示，混淆矩阵是一个 $n \times n$ 的方阵（ n 为分类类别数），其行表示真实类别，列表示预测类别。矩阵中的每个元素表示属于某一真实类别被分类为某一预测类别的样本数量。对角线元素表示正确分类的数量，非对角线元素表示各类别间的混淆情况。在混淆矩阵中， x_{ii} 表示预测结果与真实标签相同，即正确分类的样本数量， $x_{ij} (i \neq j)$ 表示预测结果与真实标签不符，即错误分类的样本数量。 $N = \sum_{i=1}^n \sum_{j=1}^n x_{ij}$ 表示预测样本总数， n 为类别数。 $x_{i+} = \sum_{j=1}^n x_{ij}$ 表示真实图像中第 i 类的样本数量， $x_{+i} = \sum_{j=1}^n x_{ji}$ 表示预测结果中预测为第 i 类的样本数量。通过分析混淆矩阵，研究者可以深入了解分类结果的误差特征，为改进分类算法提供重要依据。

表 3.4 混淆矩阵

Table 3.4 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_{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

总体精度（Overall Accuracy, OA）是评估分类模型性能的核心指标之一，其定义为模型正确分类的样本数量与总样本数量的比值，数值与模型预测性能呈严格正相关关系，即 OA 值越高，表明模型的整体分类性能越优。OA 的计算方法如式(2.11)所示：

$$OA = \frac{1}{N} \sum_{i=1}^n x_{ii} \quad (2.11)$$

平均精度（Average Accuracy, AA）是评估分类模型性能的重要指标，其定义为各类别分类精度的算术平均值，该指标能够有效反映分类模型在各个类别上的平均识别性能，特

别是在类别样本分布不均衡的情况下，AA 相比总体精度（OA）能够更准确地揭示模型在不同类别间的性能差异，为模型优化提供更具针对性的指导。AA 的计算方法如式(2.12)所示：

$$AA = \frac{n \sum_{i=1}^n x_{ii}}{\sum_{i=1}^n \sum_{j=1}^n x_{ij}} \quad (2.12)$$

Kappa 系数（Kappa Coefficient）是一种广泛应用于分类模型性能评估的统计指标，其核心价值在于能够有效衡量分类结果与真实标签之间的一致性程度。该指标通过引入随机一致性概率对分类器的随机预测进行校正。Kappa 值在-1到1之间，一般大于 0，Kappa 值和预测结果成正比。Kappa 值的计算方法如式(2.13)所示：

$$Kappa = \frac{N \sum_{i=1}^n x_{ii} - \sum_{i=1}^n (x_{i+} x_{+i})}{N^2 - \sum_{i=1}^n (x_{i+} x_{+i})} \quad (2.13)$$

3.3.3 实验环境

MMFINet 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

3.3.4 实验对比方法

为了评估所提出模型的性能，本文将 MFFINet 与经典的机器学习方法，如 SVM。和经典的深度学习语义分割方法进行了比较，包括 PSPNet^[63]，SegNet^[64]，UNet^[60]，Deeplabv3+^[61]，MAREsU-Net^[65]，DDRNet^[66]。

(1) SVM: SVM 的核函数能够将数据映射到更高维的空间，从而解决非线性分类问题。

(2) PSPNet: 该网络通过引入金字塔池化模块（PPM），能够有效捕捉图像中的多尺度上下文信息，从而提升复杂场景的解析能力。

(3) SegNet: SegNet 是一种高效且轻量级的语义分割模型，其独特的池化索引技术使其在内存优化和实时性方面具有显著优势。

(4) UNet: U-Net 采用独特的编码器-解码器结构，引入了跳跃连接，将编码器中的特征图与解码器中对应层级的特征图进行拼接。

(5) Deeplabv3+: DeepLabv3+引入了编码器-解码器架构, 结合了 DeepLabv3 的空间金字塔池化模块 (ASPP) 和解码器的上采样路径。

(6) MAResU-Net: MAResU-Net 是一种基于 U-Net 改进的先进图像分割模型, 通过引入多注意力机制和残差连接, 显著提升了特征提取能力和分割精度。

(7) DDRNet: DDRNet 设计了双分辨率分支和特征融合模块, 模型能够同时捕捉细节信息和全局上下文信息在训练过程中。并且引入了深度监督机制, 对中间层的输出进行监督, 以加速模型收敛并提升性能。

3.3.5 实验参数设置

在实验中, 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 0.0005, 权重衰减 0.0004。训练迭代次数设置为 100 次, 批处理大小设置为 32, 并使用交叉熵损失函数进行反向传播, 动量参数设置为 0.9。

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

3.4.1 实验结果

图 3.7 展示了在“Zhuhai-1 号”卫星高光谱数据集上不同方法与 MMFNet 的解译结果。图中的红色方框是各种方法容易出错的地方。图 3.7(a)为真值图, 作为参考基准。图 3.7(b)为机器学习方法 SVM 的解译结果, 通过观察 SVM 的解译结果可以发现, SVM 在桉柳和混生区的分类效果较差, 同时错误地将部分互花米草区域识别为混生区。图 3.7(c)所示为 PSPNet 方法的解译结果, 其在芦苇区域出现了明显的误判, 将部分芦苇错误预测为互花米草和混生区; 此外, 右下角大片面积的芦苇区域也被错误识别为混生区。进一步观察 SegNet、Deeplabv3+、UNet 和 MAResU-Net 的结果 (图 3.7(d)-(g)), 可以发现这些方法在混生区、盐地碱蓬和潮滩之间的分类效果均不理想。同时, 这些方法在芦苇区域均存在明显的噪声。混生区由于包含多种典型地物, 其特征与周围地物具有较高的相似性, 导致这些方法难以明确区分小区域内的地物类别。图 3.7(h)展示了 DDRNet 的结果, 其整体表现优于前述方法, 但在中间红色方框标注的区域, 仍将部分潮滩错误分类为互花米草和混生区。相比之下, MMFNet (图 3.7(i)所示) 通过其多分支多层次交互机制, 能够更好地捕捉多尺度和多层次的语义信息, 结合有效的特征融合和上下文信息利用, 显著改善了分类混淆现象, 整体预测结果更为准确。

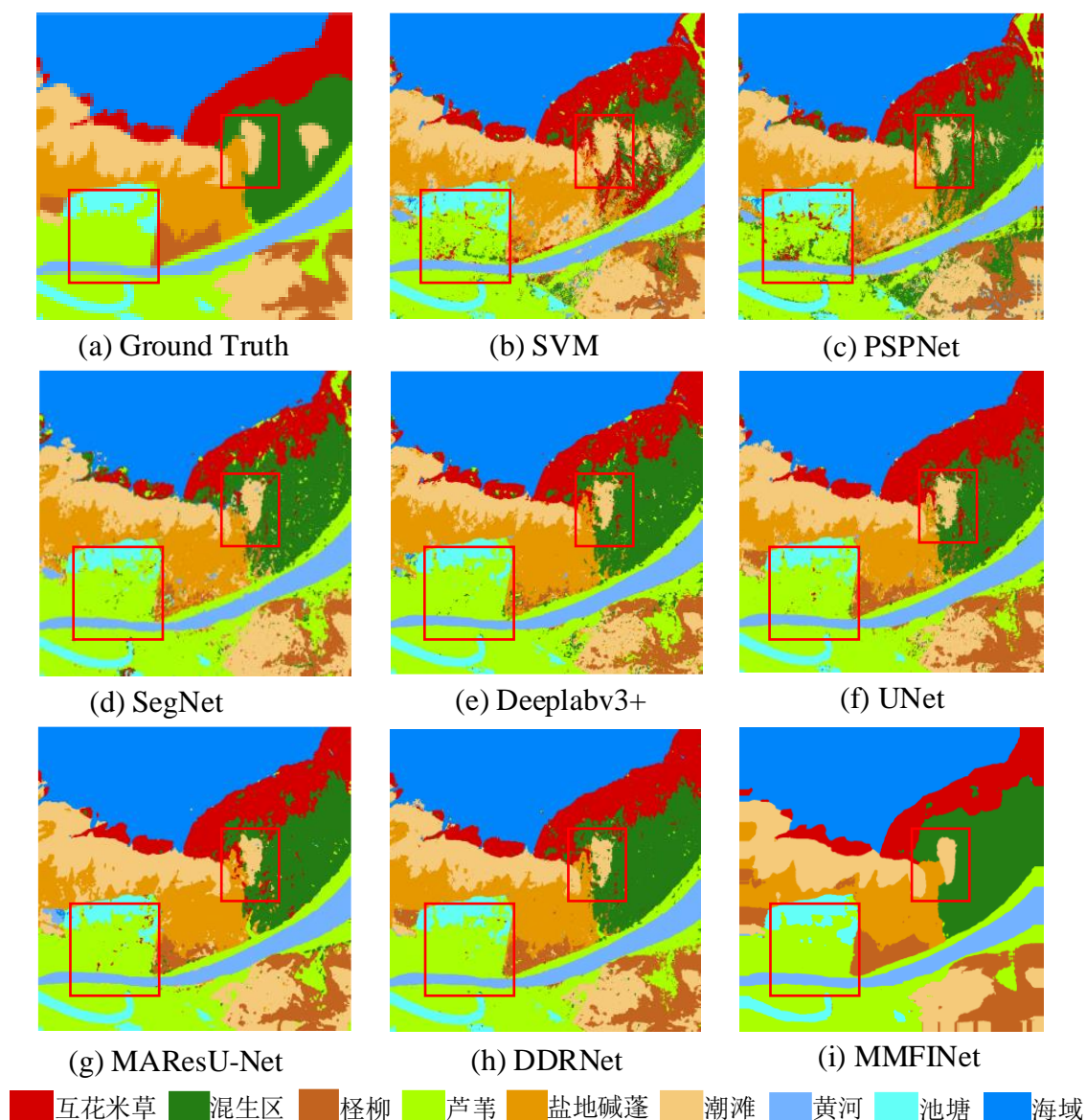


图 3.7 “Zhuhai-1 号”卫星高光谱数据集上的分类结果图

Fig. 3.7 Classification maps for the Zhuhai-1 satellite hyperspectral dataset

表 3.6 展示了 MMFINet 与其他方法在数值精度方面的对比结果，其中加粗数值表示所有方法中的最优值。从表中可以看出，MMFINet 在 OA、AA 和 Kappa 三个指标上均取得了最佳表现，其值分别为 93.01%，92.06%和 0.9186，三个指标都优于其他方法。具体而言，MMFINet 在 OA 值上比 DDRNet、MAREsU-Net 和 UNet 分别提升了 0.4%、1.89%和 2.06%，展现了其在滨海湿地高光谱图像解译任务中的显著优势。此外，MMFINet 在三个类别上的分类精度均达到了最优水平。例如，在混生区类别上，MMFINet 方法的准确率达到 98.47%，远超其他方法。在潮滩类别上，其准确率虽略低于 MAREsU-Net，但差距微乎其微，表明 MMFINet 在该类别上同样具备较强的竞争力。

表 3.6 “Zhuhai-1 号”卫星高光谱数据集上的分类结果

Table 3.6 Classification result of the Zhuhai-1 satellite hyperspectral dataset

类别	SVM	PSPNet	SegNet	Deeplabv3+	UNet	MAREsU-Net	DDRNet	MMFINet
----	-----	--------	--------	------------	------	------------	--------	---------

互花米草	82.09	69.77	70.93	88.93	94.81	90.60	91.41	84.20
芦苇	77.95	78.66	95.12	93.98	92.20	95.23	95.40	91.03
怪柳	31.82	45.61	52.63	54.45	72.55	71.38	83.26	96.54
盐地碱蓬	58.57	75.82	89.01	94.74	93.09	84.72	92.50	86.56
混生区	48.78	83.02	83.96	90.08	82.80	84.17	90.16	98.47
潮滩	85.29	87.02	83.21	70.48	85.08	90.91	85.09	89.30
黄河	95.00	96.08	96.08	92.44	96.52	96.57	97.45	97.92
池塘	89.29	88.89	80.56	82.50	91.25	84.50	87.04	85.00
海域	97.63	99.54	99.08	99.45	90.56	99.39	98.96	99.51
OA (%)	77.37	83.40	87.23	88.22	90.95	91.12	92.61	93.01
AA (%)	74.05	80.49	83.40	85.23	89.65	88.61	91.25	92.06
Kappa	0.7358	0.8064	0.8507	0.8624	0.8946	89.63	0.9138	0.9186

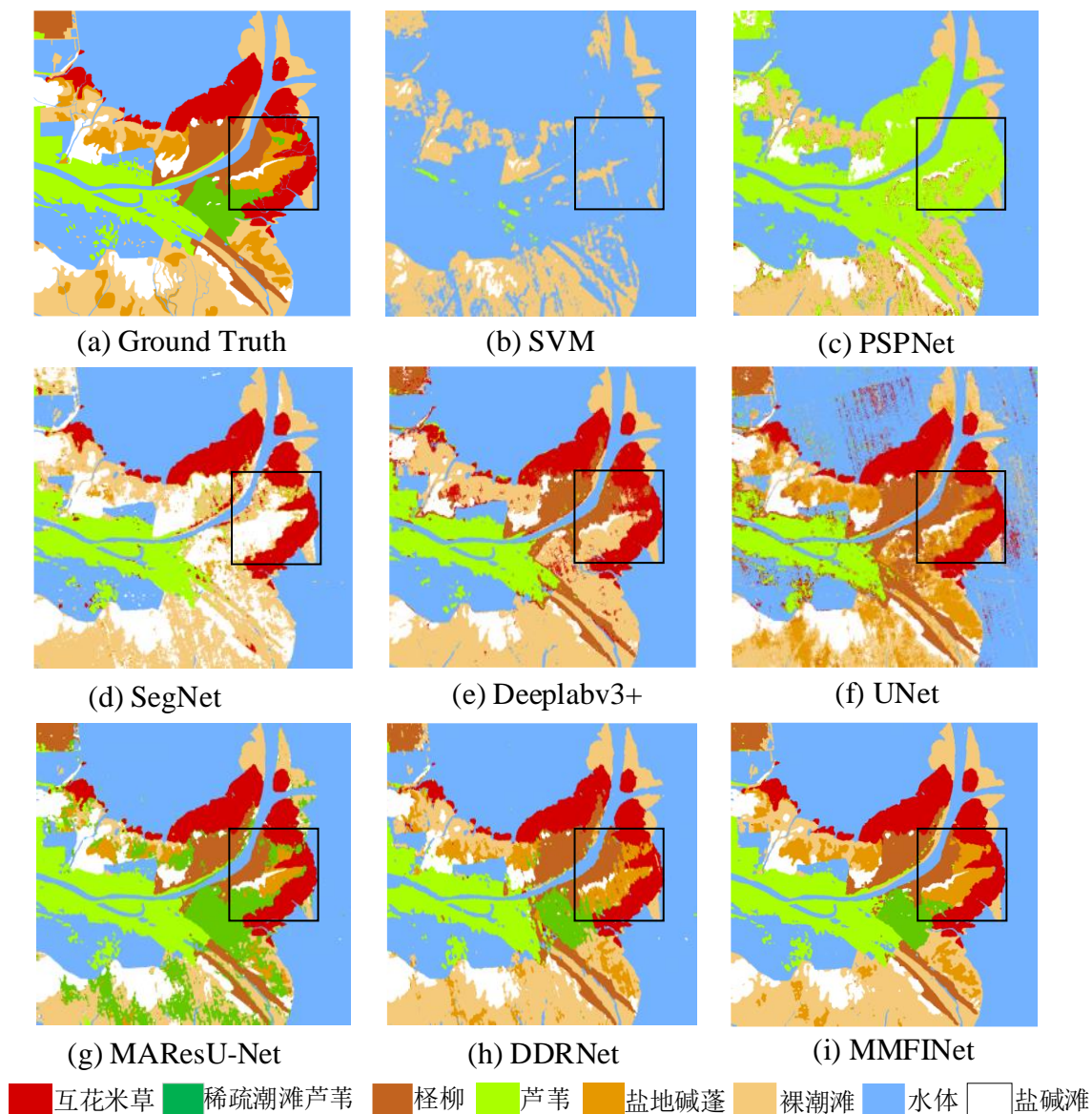


图 3.8 “高分五号”卫星高光谱数据集上的分类结果图

Fig. 3.8 Classification maps for the GF-5 satellite hyperspectral dataset

图 3.8 展示了在“高分五号”卫星高光谱数据集上不同方法与 MMFNet 的解译结果。图中的黑色方框是各种方法容易出错的地方。图 3.8(a)为真值图。图 3.8(b)和图 3.8(c) 分别展示了 SVM 和 PSPNet 的解译结果。从图中可以看出，这两种方法的效果较差。SVM 由于缺乏对高维数据和非线性关系的有效处理能力，仅能识别出裸潮滩和水体等简单类别。PSPNet 虽然具有较强的空间特征提取能力，但对光谱特征的利用不足，仅识别出了芦苇、水体和裸潮滩等类别。图 3.8(d)为 SegNet 的解译结果，其在黑色方框标注的区域中将怪柳类别错误地分成了盐碱滩。图 3.8(e)为 Deeplabv3+的解译结果，其在小块面积区域上出现了明显误判，将部分区域错误分类为互花米草。图 3.8(f)为 UNet 的解译结果，虽然其整体表现优于前几种方法，但结果中仍包含大片面积的噪声。图 3.8(g)为 MAResU-Net 的解译结果，其明显将大片面积的裸潮滩错误分类为芦苇。图 3.8(h)为 DDRNet 的解译结果，其整体效果较好，但在黑色方框标注的区域中，仍将碱蓬错误分类为怪柳。相较于其他方法，MMFNet（图 3.8(i)所示）在包含不同尺度的地物类型显著的复杂环境中，表现出了优越的准确性和鲁棒性，避免了其他方法中常见的误分类现象。

表 3.7 “高分五号”卫星高光谱数据集上的分类结果

Table 3.7 Classification result of the GF-5 satellite hyperspectral dataset

类别	SVM	PSPNet	SegNet	Deeplabv3+	UNet	MAResU-Net	DDRNet	MFFNet
互花米草	0.00	84.12	98.40	88.91	94.30	98.56	95.37	95.74
芦苇	81.77	34.26	91.67	89.58	88.84	86.51	89.93	93.54
怪柳	0.00	0.00	0.00	87.00	90.62	85.23	88.05	92.04
盐地碱蓬	0.00	0.00	9.35	0.00	66.09	60.24	75.12	74.79
稀疏潮滩芦苇	0.00	0.00	0.00	0.00	0.00	91.73	81.91	81.53
裸潮滩	93.63	96.11	77.45	90.67	89.90	93.61	88.67	89.61
水体	98.83	97.35	94.43	97.34	96.65	96.71	96.48	97.46
盐碱滩	83.33	61.42	98.61	72.50	88.12	88.42	87.04	88.33
OA (%)	72.13	73.37	77.93	83.98	89.20	91.83	91.68	92.92
AA (%)	44.70	46.66	58.74	65.75	76.82	87.63	87.82	89.13
Kappa	0.6105	0.6326	0.7044	0.7810	0.8546	89.02	0.8882	0.9047

表 3.7 展示了不同方法在“高分五号”卫星高光谱数据集上的解译结果的具体指标，在综合考虑总体准确率、平均准确率和 Kappa 系数的情况下，本方法展现了较好的性能。从表 3.7 可以清楚地看出，MFFNet 在芦苇和怪柳两个类别上取得了最高的分类准确率，同时其总体精度 OA、平均精度 AA、Kappa 系数三个指标均表现最优，其值分别达到了 92.92%，89.13%和 0.9047。具体来看，SVM 在水体上的精准率达到了最高的 98.83%，但在怪柳、盐地碱蓬等类别上表现较差，分类效果不理想。PSPNet、SegNet、Deeplabv3+和 UNet 方法在某些类别上甚至出现了精准率为 0%的情况，表明这些方法在处理复杂地物分类任务时存在明显不足。MAResU-Net 在互花米草和稀疏潮滩芦苇类别上的精准率较高，

但其整体性能仍不及 MFFINet。总体而言，进一步验证了 MFFINet 在复杂场景解译任务中的优越性。

3.4.2 多分支融合有效性分析

在实验中，本文针对两个数据集深入探讨了分支有交互和无交互情况下 MMFINet 的训练时间与分类性能。实验数据如表 3.8 所示，可以观察到，分别在 “Zhuhai-1 号” 卫星和 “高分 5 号” 卫星高光谱数据集上，引入反馈机制后，模型的 OA 值相较于无反馈情况提升了 2.16% 和 1.85%。此外，尽管引入反馈机制后模型的训练时间略有增加，但这种时间成本的增加完全在可接受范围内，且相较于分类性能的提升，这种代价是值得的。这一结果充分表明，多分支融合模块优化分类性能方面起到了显著作用。多分支融合模块的核心优势在于其能够通过分支间的交互机制，实现多层次、多尺度特征的深度融合。在无交互的情况下，各分支独立提取特征，虽然能够捕捉到不同尺度的信息，但缺乏特征之间的协同优化，导致部分重要信息未能充分利用。而引入反馈机制后，各分支之间能够动态共享和融合特征信息，从而增强了模型对复杂高光谱数据的表征能力。这种交互机制不仅提升了特征的丰富性和判别性，还进一步优化了分类决策过程，使得模型在复杂场景下的分类性能得到显著提升。这一设计不仅验证了多分支结构在分类任务中的有效性，也为后续研究提供了重要的参考价值。未来可以进一步探索更高效的特征交互机制，以在更低的计算成本下实现更高的分类精度。

表 3.8 不带有交互和带有交互在 OA 方面的性能比较

Table 3.8 Performance comparison with and without interaction in terms of training time and OA

	Zhuhai-1		GF-5	
	Training (min)	OA(%)	Training (min)	OA(%)
Without interaction	17.21	90.32	17.73	89.67
With interaction	17.46	93.01	17.82	92.92

3.4.3 计算时间和模型复杂性评估

表 3.9 展示了不同方法运行时间和复杂度比较。从表中可以观察到，Deeplabv3+ 的计算量和参数量相对较高，由于其网络结构设计和使用的模块组件，导致整体复杂度较高，每次迭代都需要较长的计算周期，导致训练和推理时间较长。此外，高参数量的设计还带来了一些潜在问题：一方面，它增加了模型的训练成本，需要更多的计算资源（如 GPU 内存）和更长的训练时间；另一方面，在训练样本有限的情况下，高参数量可能会增加过拟合的风险，从而影响模型的泛化能力。UNet 和 SegNet 网络结构比较简单，采用了经典的编码器-解码器框架，避免了复杂的模块化设计，这种简洁的结构使得它们的 FLOPS 和 Params 显著低于 Deeplabv3+，从而提高了计算效率。MMFINet 的运行时间和复杂度处于

适中状态，这主要是由于引入多分支结构，多分支结构通过引入分支间的交互机制，能够实现多层次、多尺度特征的深度融合，从而提升模型的性能，然而，这种设计也在一定程度上增加了参数量。

表 3.9 不同方法计算时间和复杂性比较

Table 3.9 Comparison of different methods in terms of running time and complexity

数据集	指标	PSPNet	SegNet	UNet	Deeplabv3+	MAResU-Net	DDRNet	MMFNet
Zhuhai-1	Training time (s/epoch)	28.54	18.28	15.78	16.44	13.45	23.18	17.21
	Testing time (s)	6.48	4.78	4.72	5.09	3.79	4.65	3.35
	FLOPS (G)	29.56	28.32	24.28	30.65	26.18	27.21	26.57
	Params (M)	23.38	7.15	8.78	20.37	25.38	20.38	18.56
GF-5	Training time (s/epoch)	29.28	18.59	15.35	16.48	14.13	24.33	17.73
	Testing time (s)	6.37	4.35	4.13	5.42	5.17	4.71	4.65
	FLOPS (G)	26.60	25.32	23.57	29.10	25.98	26.79	25.88
	Params (M)	23.38	7.15	8.78	20.37	25.38	20.38	18.56

3.4.4 消融实验

(1) 不同模块的消融实验分析

为了深入剖析该方法对检测结果的影响机制，本文设计了消融实验，细致探究了网络中核心组件（包括多尺度注意力聚合模块 MSAA、基于残差单元的全局注意力机制 GAM-RU）以及 ResNet50 架构的贡献，如表 3.10 所示。上面提到的所有实验都是“Zhuhai-1 号”卫星高光谱数据集。如表八清晰展现，各模块均对本文的方法产生了不可或缺的正面效应，它们各自在不同层面上发挥了积极作用。只有当所有模块在本文的网络中协同工作时，才能达到最优的性能表现。具体而言，采用 ResNet50 作为预训练（PT）基础的模型，相较于未使用预训练（NPT）的模型，展现出了更为卓越的性能。这一预训练策略使模型在总体精度、平均精度和 Kappa 系数上分别实现了 6%、15%和 6%的显著提升。随后，将 MSAA 和 GAM-RU 融入基于 PT 的模型中，进一步增强了模型的能力。量化评估结果有力地证明了 MSAA 和 GAM-RU 对于推动模型达到高性能的关键作用。最终，当该方法集成了 PT、MSAA 和 GAM-RU 时，其性能达到了巅峰状态，具体表现为总体精度高达 93.01%，平均精度为 92.06%，Kappa 系数为 0.9186。基于这些消融研究的发现，可以确信无疑地得出结论：在本文的网络中部署的所有模块均对滨海湿地精细制图产生了显著的正面影响，共同

推动了性能的飞跃。

表 3.10 不同模块的实验分析

Table 3.10 Module experimental analysis

方法	OA (%)	AA (%)	Kappa
NPT	91.44	89.70	0.8832
PT	92.15	90.82	0.8978
PT+MSAA	92.49	91.35	0.9046
PT+GAM-RU	92.30	91.42	0.9013
MMFINet(Ours)	93.01	92.06	0.9186

(2) 编码器网络结构的消融实验分析

为了评估不同预训练编码器在滨海湿地解译任务中的性能差异，本章基于“Zhuhai-1”和“高分五号”卫星高光谱数据集上开展了消融实验，实验结果如表 3.11 所示。从总体精确度来看，采用 ResNet50 作为模型的编码器具有最高的性能，在 Zhuhai-1 号和 GF-5 号两个数据集上分别达到了 93.01%和 92.92%的精度。相比之下，ResNet18 和 ResNet34 由于网络层较少，特征提取能力有限，难以捕捉复杂的高光谱特征，因此在任务中表现欠佳。而 ResNet101 虽然具有更深的网络结构，能够提取更丰富的特征，但由于其参数量较大，在数据规模有限的情况下容易出现过拟合问题，导致性能下降。而 ResNet50 在模型复杂度和性能之间取得了较好的平衡，其网络深度足以提取高光谱数据中的复杂特征，同时参数量适中，能够有效避免过拟合现象，从而在实验中展现出最佳的性能表现。

表 3.11 编码器网络的实验分析

Table 3.11 Experimental analysis of encoder network

方法	OA(%)	
基线网络	Zhuhai-1	GF-5
ResNet18	89.12	88.03
ResNet34	92.34	92.31
ResNet50	93.01	92.92
ResNet101	91.55	91.11

3.4.5 训练样本数量对结果的影响

为了全面评估 MMFINet 模型的有效性和鲁棒性，本研究采用渐进式训练样本比例策略对 7 种对比算法的总体分类精度（OA）进行了系统性验证。如图 3.9 所示（其中深棕色线条代表 MMFINet 的 OA 值），分别在 Zhuhai-1 和高分五号数据集上设置了 3%、5%、10%、15%和 20%五个梯度的训练样本比例。实验结果表明，随着训练样本数量的增加，各模型的训练过程更加充分，OA 指标呈现明显的上升趋势，这一现象验证了充足训练样本对于模型性能提升的重要性。值得注意的是，在所有对比算法中，MMFINet 展现出显著优越的分类性能（其深红色性能曲线始终位于其他算法之上）。具体而言，当训练样本比例从 3%逐步提升至 20%时，MMFINet 的 OA 值呈现出稳定且持续的增长态势，充分证明

了该模型具有优异的样本利用率。这种卓越的性能表现主要归因于 MMFNet 独特的多尺度信息融合机制；其次，引入深度监督机制进一步优化了特征表示，使得模型在不同训练样本规模下均能保持稳定的分类精度。因此，可以得出结论：MMFNet 不仅具有较高的分类精度，还展现出优异的泛化能力和样本适应性。

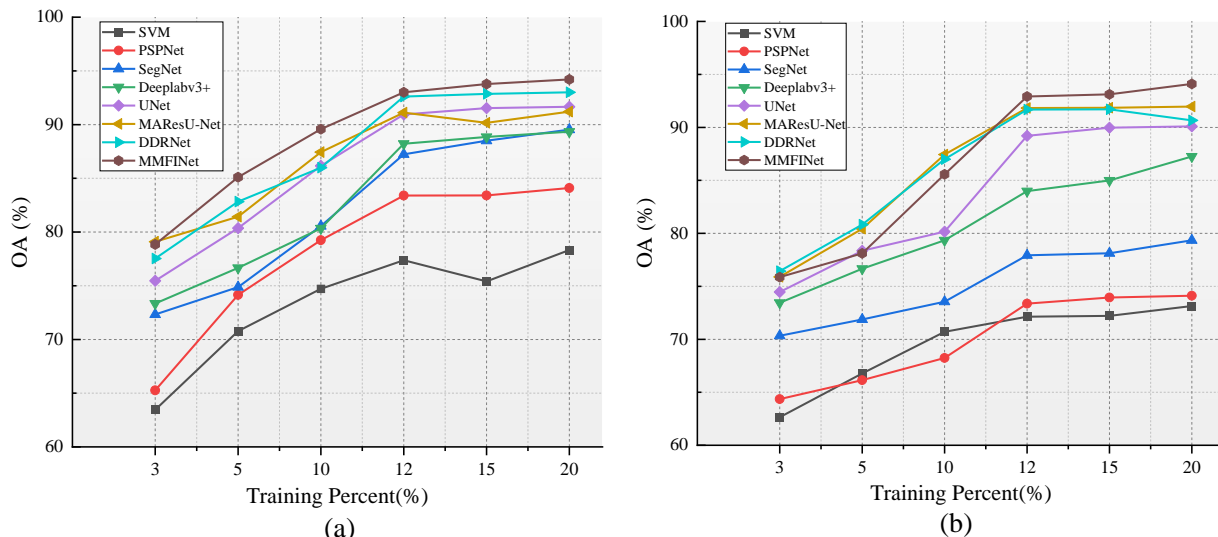


图 3.9 不同空间输入大小的对分类性能的影响

Fig.3.9 Effect of different spatial sizes on overall accuracies on four HSI datasets

3.5 本章小结 (Chapter Summary)

本章针对滨海湿地地物多尺度特性带来的分类挑战，提出了一种创新的多分支多层次特征交互网络 MMFNet，旨在显著提升模型对复杂地物的识别能力。首先，详细阐述了 MMFNet 的整体架构设计，该架构通过层次化的特征提取与交互机制，有效解决了滨海湿地地物尺度差异大的问题。接着，重点介绍了多尺度注意力聚合模块，该模块聚合不同尺度的特征信息，从而增强模型对多尺度地物的表征能力。此外，本章还深入解析了多分支解码器结构，该结构通过并行分支的特征融合策略，充分挖掘细节信息与上下文信息，进一步提升特征表达的丰富性和鲁棒性。为了进一步优化分类结果，引入了深度监督机制，通过在网络的不同层次添加辅助分类器，增强模型的训练稳定性，同时提高解译结果的准确性和可靠性。为验证 MMFNet 的有效性，本章在 Zhuhai-1 号和高分五号卫星高光谱数据集上进行了全面的对比实验和消融实验。

4 基于空谱特征协同的滨海湿地精细制图方法

4 A fine mapping method of coastal wetland based on spatial spectral feature coordination

4.1 引言 (Foreword)

滨海湿地生态系统具有复杂的地物分布特征, 其不同植被类型之间的光谱特征相似度较高, 这种光谱特征的细微差异显著增加了遥感影像解译的难度, 导致地物分类精度降低。由于变压器在捕获长期图像依赖性方面具有优越的能力, 现有的方法大多依赖于 Transformer, 然而, Transformer 模型在应用中也存在一定的局限性。具体而言, 与处理可见光图像相比, 高光谱图像 (HSI) 具有显著的多波段特性。这种丰富的谱域信息虽然提供了更全面的数据特征, 但也可能导致特征提取过程过于分散, 从而难以有效捕捉局部相邻谱带之间的相关性信息, 这在一定程度上削弱了模型对目标分类的贡献度^[67]。相较于传统的卷积神经网络 (CNN), 图卷积神经网络 (GCN) 具有独特的优势。GCN 通过非欧几里得数据空间中对相邻节点特征进行有效聚合, 实现了图结构数据上的卷积运算。这种机制不仅突破了传统 CNN 局限于规则网格结构的限制, 更重要的是建立了空间节点间的长期依赖关系, 从而显著增强了模型对像素点空间上下文信息的表征能力。

针对上述问题, 为充分利用 HSI 的空间-谱域特性, 本研究创新性地整合了 GCN 和 Transformer 模型的优势: 通过 GCN 模型有效捕捉空间域中节点间的局部特征依赖关系, 同时利用 Transformer 的自注意力机制建立谱域序列的长期依赖关系。基于此, 本文提出了一种新型的空谱特征协同网络 SSFCNet。在该方法中, 为了提升 GCN 在空间-光谱域信息处理方面的能力, 本研究引入了通道先验卷积注意模块, 并提出了通道先验卷积注意力机制 GCN (CPCAGCN)。此外, 设计了一个像素选择模块, 用于筛选 GCN 的初始输出特征, 从而优化并简化 Transformer 的输入数据。在 Transformer 特征提取阶段, 提出了邻域卷积方法, 该方法结合了 Transformer 的全局注意力机制, 同时强化了对局部光谱信息的捕捉能力。

4.2 SSFCNet 网络结构设计(SSFCNet network structure design)

4.2.1 SSFCNet 的整体架构

SSFCNet 架构巧妙地融合了通道感知卷积注意力机制 GCN (CPCAGCN) 与局部光谱特征融合变压器 (LST) 两大核心组件, 二者以级联形式集成于网络框架之中。为了深度整合空间与光谱域的信息, 并实现自适应学习能力, 本章引入了高效的通道先验卷积注意力机制 (CPCA)。该机制能够在通道层面与空间维度上灵活地动态分配注意力权重, 从而增强特征表示的有效性。随后, CPCAGCN 模块构建起了像素间的精细空间关系图谱, 并

输出富含空间信息的特征图。通过内置的像素选择块（PSB），精准地筛选出特征图中的分类关键特征，这些特征随即作为输入传递给 LST 模块。在此基础上，进一步设计了一种巧妙的邻域卷积策略（NC），旨在融合 LST 输入特征中的局部光谱域信息，这一步骤对于提升特征表达的全面性和准确性至关重要。最终，为了减少特征在分层传播过程中的信息丢失，还设计了一种跨层连接，以确保多层次特征的有效融合与传递。依托空间光谱域信息的深度挖掘，SSFCNet 得以实现高精度的分类性能。

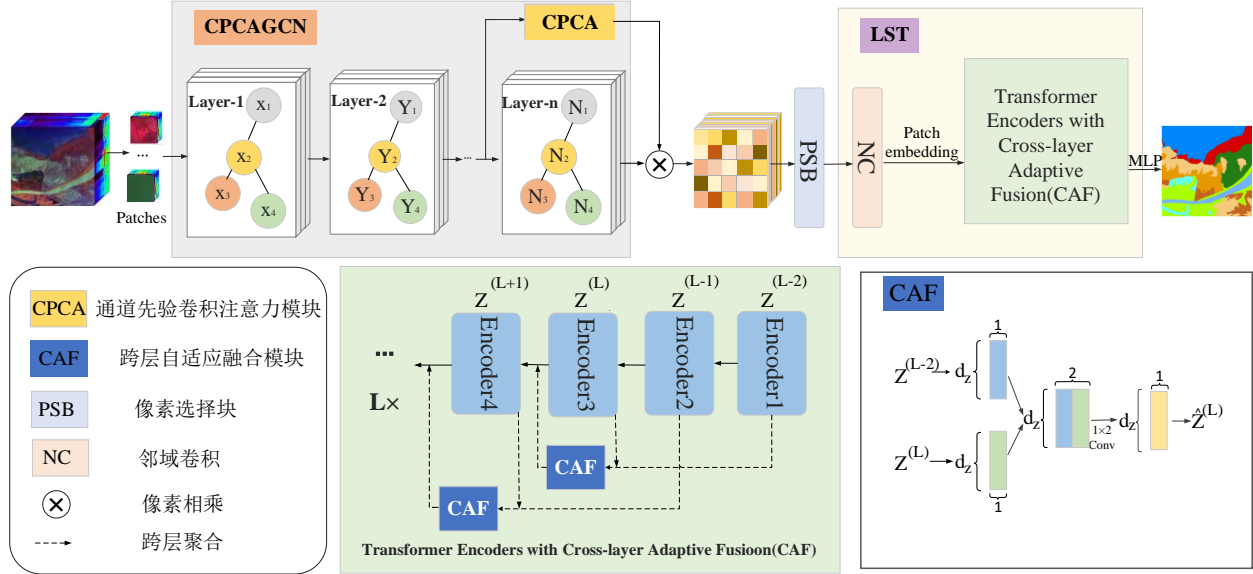


Fig. 4.1 Overall structure of SSFCNet network

4.2.2 通道先验卷积注意力机制 GCN

GCN 主要关注像素点之间的空间关系，然而，忽视了光谱信息的重要性。本文设计了 CPCAGCN 模块，通过增加通道先验卷积注意力模块来提高分类精度，通过在通道和空间维度上动态分配注意力权重，突出显示更有分类价值的光谱信息，抑制冗余信息，实现空间光谱域信息的融合，另外，改模块可以增强局部光谱和空间特征的细节，捕捉每个像素的重要光谱波段和空间区域，以提高遥感图像中高相似性地物的分割精度。

CPCA 如图 4.2 所示，依次执行通道注意力模块和空间注意力模块。给定一个中间特征图 $F \in R^{C \times H \times W}$ 作为输入，通道注意模块 (CA) 首先推断出一个一维通道注意力图 $M_c \in R^{C \times 1 \times 1}$ ，该注意力图用于表征每个通道的重要性。然后将 M_c 与输入特征 F 进行元素相乘，沿空间维度广播信道注意值，得到具有通道注意的细化特征 $F_c \in R^{C \times H \times W}$ 。空间注意模块 (SA) 处理 F_c 生成三维空间注意力图 $M_s \in R^{C \times H \times W}$ 。最终，通过将 M_s 与 F_c 进行逐元素相乘，得到优化后的输出特征 $\hat{F} \in R^{C \times H \times W}$ 。这种双注意力机制的设计不仅增强了模型对通道和空间信息的感知能力，还显著提升了特征的代表效果。总体注意过程可概括为公式(4.1) (4.2):

$$F_c = CA(F) \otimes F \quad (4.1)$$

$$\hat{F} = SA(F_c) \otimes F_c \quad (4.2)$$

通道注意力模块首先通过平均池化和最大池化操作从特征图中聚合空间信息，生成两个不同的空间上下文描述符。这两个描述符随后被送入一个共享的多层感知机（MLP），通过元素级相加得到通道注意力图。为了减少参数开销，共享 MLP 包含一个隐藏层，隐藏激活的大小设置为 $\frac{C}{r \times l \times l}$ ，其中 R 是通道数，r 是降低比率。该过程可以表示为公式(4.3):

$$CA(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \quad (4.3)$$

其中 σ 代表 sigmoid 函数。空间注意力模块处理经过通道注意力处理的特征图，生成 3D 空间注意力图。该模块使用深度卷积模块（Depth-wise Convolution Module）来捕获特征之间的空间关系，同时保留通道之间的关系，降低计算复杂度。深度卷积模块采用多尺度深度可分离条纹卷积核来提取像素之间的空间映射关系，确保在减少计算复杂度的同时有效提取信息。通过 1×1 卷积进行通道混合，生成更精细的注意力图。该过程可以表示为公式(4.4):

$$SA(F) = \text{Conv}(\sum_{i=0}^3 \text{Branch}_i(\text{DwConv}(F))) \quad (4.4)$$

其中，DwConv 表示深度可分离卷积，而 Branch_i ， $i \in \{0, 1, 2, 3\}$ 表示第 i 个分支。分支 0 是身份连接。按照^[68,69]的方法，为了高效捕获多尺度特征信息，采用了两层深度条带卷积来近似替代大核标准的逐通道卷积。通过为每个通道设计不同尺度的卷积核，能够有效提取多尺度的空间特征，从而增强模型对多尺度信息的表征能力。这种设计不仅降低了计算复杂度，还提升了对多尺度特征的适应性。

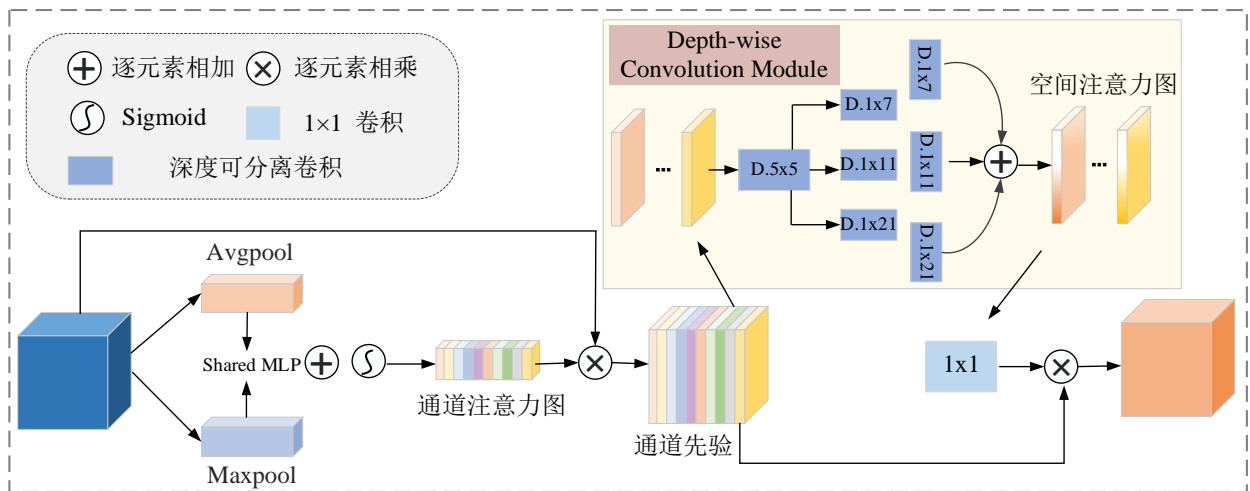


图 4.2 通道先验卷积注意力模块结构

Fig.4.2 Structure of channel prior convolutional attention module

4.2.3 像素选择模块

在视觉 Transformer（ViT）的传统分类任务中，通常采用 patch 嵌入模块将输入图像分

割为不重叠的图像块进行处理。然而，在 HSI 分类任务中，由于需要对每个像素进行独立分类，Transformer 的输入通常被限制为单个像素。为了充分利用 HSI 的空间信息，文献^[63]提出了一种基于 patch 的输入策略，即将待分类像素及其邻域像素共同作为输入。尽管这种方法相较于单一像素输入显著提升了分类性能，但其对邻域像素的无差别引入容易导致边缘像素的误分类问题。针对这一挑战，本文提出了一种基于 GCN 的邻域选择机制，通过 GCN 自适应地选择与中心像素最具相关性的邻域像素构成 Transformer 的输入，如图 4.3 所示。该方法不仅能够有效建立空间长程依赖关系，避免边缘像素的误分类问题，同时显著降低了计算复杂度。

假设有一个分类节点标记为 V_i ，我们首先依据在 GCN 中计算所得的邻接矩阵中各个节点对应的权重值，对这些节点执行排序操作。过程表示如公式(4.5)(4.6)所示：

$$S_j^i = \text{Sort}(\text{Corr}(V_i, V_j)), \quad i, j = 1, 2, \dots, N \quad (4.5)$$

$$V_{ij} = \text{Corr}(V_i, V_j) \quad (4.6)$$

其中， V_{ij} 表示节点 V_i 和 V_j 之间的相关性权重，对所有像素 V_j 与当前像素 V_i 的相关性权重进行排序，从大到小排列。排序后的列表 S_j^i 用于选择与 V_i 相关性最强的像素。N 表示所有节点的个数。取 S_j^i 的前 M 个点构成变压器的输入。该过程如公式(4.7)所示：

$$F_{in}^i = \text{Concat}(S_1, S_2, \dots, S_j, \dots, S_M), \quad j = 1, 2, \dots, M \quad (M \leq N) \quad (4.7)$$

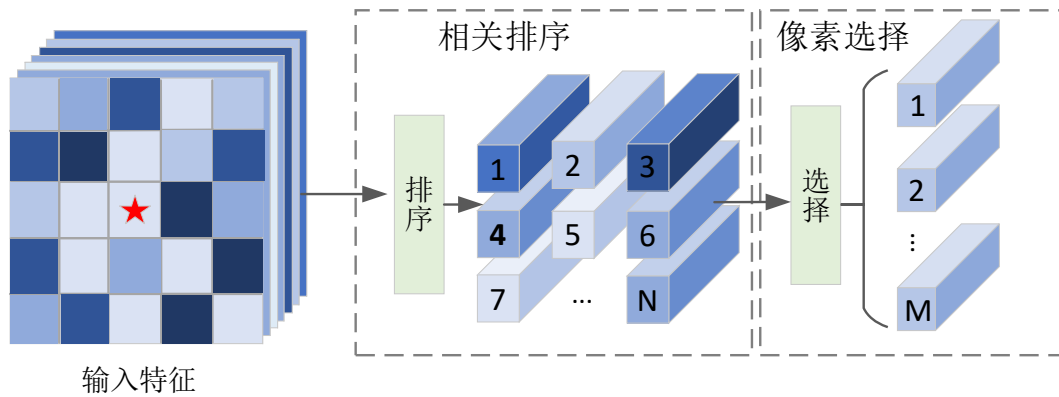


图 4.3 像素选择块结构

Fig.4.3 Structure of pixel selection module

其中 Concat 表示拼接操作，与 V_i 中节点最相关的前 M 个节点形成一个新的输入 F_{in}^i 。相比直接将整个 patch 作为输入，相关性排序方法能够更有利于 GCN 对空间节点进行建模，并从中提取出与当前节点关联性更强的分类特征节点，从而有效地剔除分散的冗余信息，显著降低错误分类的风险。此外，这种方法还大幅减少了计算复杂度，实现了计算资源的优化利用。另一方面，通过引入像素选择模块作为 Transformer 的输入预处理步骤，所获取的特征与 Transformer 的处理机制更为契合。这一改进巧妙地解决了后续处理中可能出现的不同分类节点拥有相同 patch 信息的难题，确保了 Transformer 能够更精准地处理和

分析这些特征。

4.2.4 局部光谱特征融合 Transformer

Transformer 能够有效建模光谱序列中的长距离依赖关系，但在处理高光谱图像时，往往需要对局部光谱信息给予更多关注。由于 HSI 的每个像素都包含了近似连续的光谱信息，相邻光谱波段之间的细微差异往往蕴含着丰富的特征信息。为了在利用变压器捕捉全局依赖关系的同时，更好地融合局部光谱特征，本文创新性地提出了邻域卷积机制。其示意图如图 4.4 所示。具体而言，该方法首先对当前光谱波段及其相邻波段的特征进行融合，通过细致入微的局部特征提取，增强了对光谱细节信息的捕捉能力。随后，采用残差连接的方式，将融合后的特征与原始特征相结合，形成最终的输出。这一设计不仅保留了光谱序列的全局上下文信息，还显著提升了模型对局部光谱特征的敏感度，为精确的光谱分析和特征识别奠定了坚实的基础。该过程如公式(4.8)(4.9)所示：

$$X_{group}^i = Concat(X_m, X_{m+1}, ..., X_i, ..., X_{m+k-1}) \quad (4.8)$$

$$\begin{cases} m = i - \frac{k-1}{2}, & i > \frac{k-1}{2} \text{ and } i < N - \frac{k-1}{2} \\ m = 0, & i \leq \frac{k-1}{2} \\ m = N - k - 1, & i \geq N - \frac{k-1}{2} \end{cases} \quad (4.9)$$

其中， X_{group}^i 为第 i 个谱的相邻谱， k 为相邻谱的个数， N 为谱的总数。 X_{group}^i 邻域的卷积并将残差连接到原始光谱。该过程如公式(4.10)(4.11)(4.12)所示：

$$X_i' = W_i X_{group}^i + b_i, \quad i=1,2,...,N \quad (4.10)$$

$$X_i'' = X_i' + X_i \quad (4.11)$$

$$F_N = Concat(X_1'', X_2'', ..., X_N'') \quad (4.12)$$

其中， W_i 为第 i 个邻域卷积的权重矩阵， b_i 为第 i 个邻域卷积的偏差，得到邻域通道特征 X_i' ；然后在残差拼接后得到输出特征 X'' ，最后对 X'' 进行级联，得到输出特征图 F_N 。

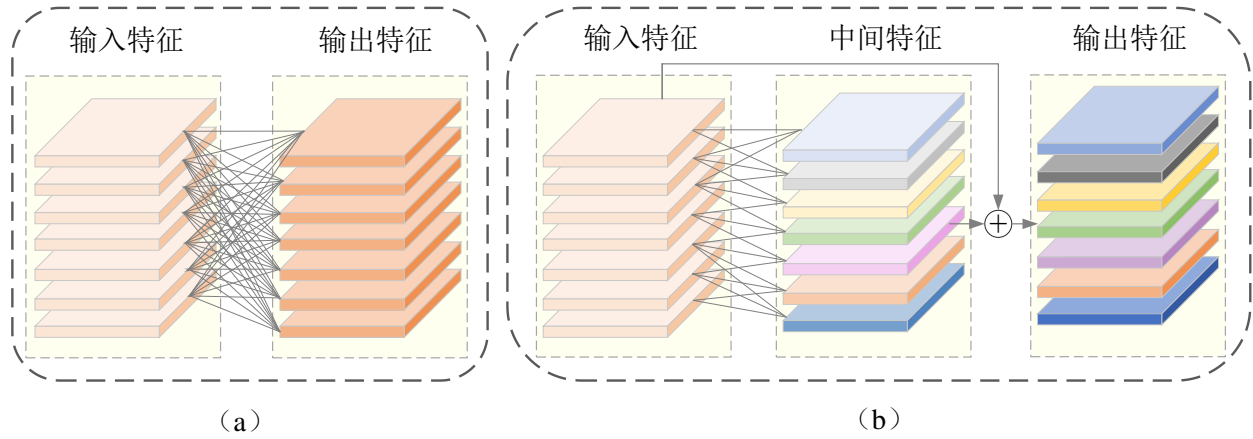


图 4.4 标准 2-D 卷积和邻域卷积对比 (a)标准 2-D 卷积 (b)邻域卷积

Fig.4.4 Comparison of standard 2-D convolution and neighborhood convolution

(a) Standard 2-D convolution (b) Neighborhood convolution

跨层连接能够增强网络层间的信息流动，减少训练过程中的信息损失。在该网络中，还引入了一种跨层连接，自适应学习跨层特征融合，有效平衡了信息的传递与融合。如图 4.1 所示（CAF 模块），设 $z^{(l-2)} \in R^{1 \times d_z}$ 和 $z^{(l)} \in R^{1 \times d_z}$ 分别为第（1-2）和第（1）层的输出。CAF 可以表示为如公式(4.13)所示：

$$\hat{z}^{(l)} \leftarrow \omega \begin{bmatrix} z^{(l)} \\ z^{(l-2)} \end{bmatrix} \quad (4.13)$$

其中 $\hat{z}^{(l)}$ 表示第（1）层与所 CAF 的融合表示， $\omega \in R^{1 \times 2}$ 是自适应融合的可学习网络参数。需要注意的是，CAF 只跳过了一个编码器。这种设计主要是因为浅层网络提取的低级特征与深层网络提取的高级特征之间存在显著的语义差距。如果采用跨度过长的跨层连接（例如跨越两个、三个甚至更多编码器），可能会导致特征融合不充分，甚至造成信息丢失。

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 0.0004，并设定权重衰减为 0.0003。同时，设定批次大小为 32，训练轮数为 100 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 SSFCNet 的有效性，本文选用两个数据集，对其性能进行了全面评估。其中，“Zhuhai-1 号”和“高分五号”卫星高光谱数据集的训练样本数和测试样本数均与第三章

保持一致。

4.3.3 实验对比方法

为了评估所提出模型的性能，本文将 SSFCNet 与经典的机器学习方法 SVM 和语义分割方法 PSPNet，基于 CNN 的分类方法 3D CNN^[70]、RSSAN^[71]、DBDA^[72]，还有基于 Transformer 的分类框架 SpectralFormer^[73]和 SSFTT^[74]进行了比较。

(1) SVM: SVM 的核函数能够将数据映射到更高维的空间，从而解决非线性分类问题。

(2) PSPNet: PSPNet 主要创新点在于金字塔池化模块的设计，通过多尺度的全局上下文信息融合，解决了语义分割任务中的多尺度问题和全局信息缺失问题。

(3) 3D CNN: 3D CNN 使用三维卷积核，在三个维度（宽度、高度、深度）上进行卷积操作，能够同时提取空间和时序特征，适合处理动态数据。

(4) RSSAN: RSSAN 是一种结合了残差学习、光谱-空间注意力机制和多尺度特征融合的先进高光谱图像分类模型。

(5) DBDA: DBDA 通过双分支结构和双重注意力机制，两个分支的特征在后续阶段进行融合，从而同时利用光谱和空间信息。

(6) SpectralFormer: SpectralFormer 是一种基于 Transformer 架构的高光谱图像分类模型，为了减少有价值信息在分层传播过程中丢失的可能性，设计了一种跨层跳过连接。

(7) SSFTT: SSFTT 结合了光谱和空间特征，利用 Transformer 架构进行特征提取和分类，将提取的光谱-空间特征转换为 Token 序列，作为 Transformer 的输入。

4.4 实验结果与分析 (Experimental result and analysis)

4.4.1 实验结果

图 4.5 所示展示了在“Zhuhai-1 号”卫星高光谱数据集上不同方法与 SSFCNet 的解译结果。如图 4.5(b)所示，SVM 由于结构较为简单，无法有效表达复杂的滨海湿地地物特征，导致其在互花米草、芦苇和混生区等类别上出现了严重的混淆分类现象，分类能力明显不足。基于 CNN 的分类方法，如 3D CNN、RSSAN 和 DBDA，通常表现出良好的分类性能，能够在一定程度上捕捉空间和光谱特征。然而，与基于 Transformer 框架的分类网络 SpectralFormer、SSFTT 和 SSFCNet 相比，其在分类精度方面仍存在一定差距。然而，尽管基于 Transformer 的分类器在光谱序列建模方面具有显著优势，能够通过全局注意力机制捕捉长距离依赖关系，但其在空间建模能力和局部光谱特征提取方面的局限性，限制了其性能的进一步提升。本章提出的 SSFCNet（如图 4.5(i)所示）通过结合图卷积网络和 Transformer，充分利用了空间上下文信息和光谱序列特征，显著提升了分类性能。因此，SSFCNet 在滨海湿地地物分类任务中表现最佳，解译结果优于其他对比方法。

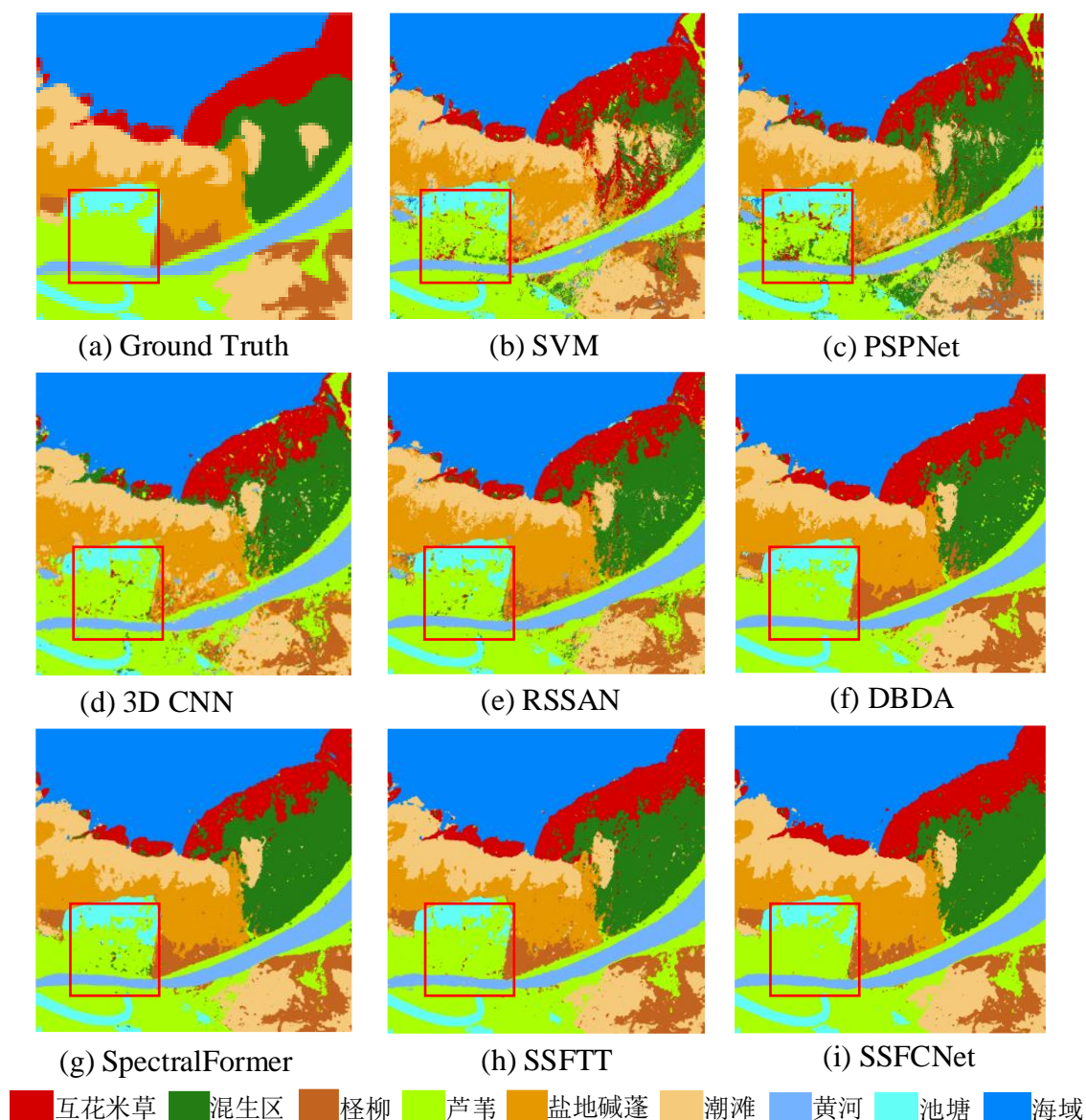


图 4.5 “Zhuhai-1 号”卫星高光谱数据集上的分类结果图

Fig. 4.5 Classification maps for the Zhuhai-1 satellite hyperspectral dataset

表 4.1 表明了 7 种对比方法和 SSFCNet 在“Zhuhai-1 号”高光谱数据集上的实验结果。从表 4.1 中可以明显看出，SSHFCNet 在 OA、AA 和 Kappa 方面优于其他方法，其值分别达到了 94.26%、93.07%、0.9331，展现了其在分类任务中的显著优势。此外，在互花米草、芦苇和混生区达到了最高的精准率。SSFTT 方法整体表现也较为出色，尤其在柽柳、黄河和池塘这三个类别上展现了较好的分类效果。然而，SSFCNet 的总体精度比 SSFTT 高出 0.87%，表明其在全局分类性能上更具优势。

表 4.1 “Zhuhai-1 号”卫星高光谱数据集上的分类结果

Table 4.1 Classification result of the Zhuhai-1 satellite hyperspectral dataset

类别	SVM	PSPNet	3D CNN	RSSAN	DBDA	SpectralFormer	SSFTT	SSFCNet
互花米草	82.09	69.77	71.43	77.92	87.85	76.01	87.79	89.75
芦苇	77.95	78.66	91.78	89.93	91.30	94.47	93.71	95.55
柽柳	31.82	45.61	47.06	69.18	89.17	84.99	87.84	85.88

盐地碱蓬	58.57	75.82	82.72	96.44	94.18	89.14	92.84	94.32
混生区	48.78	83.02	79.79	88.44	84.31	95.76	94.90	96.50
潮滩	85.29	87.02	78.45	73.35	83.40	90.32	88.85	90.05
黄河	95.00	96.08	93.48	95.83	96.35	95.17	98.70	96.75
池塘	89.29	88.89	78.12	92.00	93.02	88.37	93.75	90.03
海域	97.63	99.54	99.48	99.34	99.34	99.27	97.73	98.81
OA (%)	77.37	83.40	83.35	88.31	91.45	92.16	93.39	94.26
AA (%)	74.05	80.49	67.49	86.94	90.99	90.39	92.90	93.07
Kappa	0.7358	0.8064	0.8087	0.8637	0.9006	0.9086	0.9230	0.9331

图 4.6 显示了在“高分五号”卫星高光谱数据集上实验的所有方法的可视化分类结果。从图中可以清晰地看出不同方法在分类性能上的差异。如图 4.6(b)和图 4.6(c)所示，可以看出 SVM 和 PSPNet 方法只能识别出大块面积类别，在光谱和纹理特征相似的类别上表现效果较差。图 4.6(d)显示出了 3D CNN 的解译结果，其在大面积水体区域错误识别为互花米草。图 4.6(e)显示出了 RSSAN 的解译结果，其将柽柳错误识别为芦苇。DBDA、SpectralFormer、SSFTT 方法表现出了较好的分类性能，能够在一定程度上捕捉复杂地物的特征，但解译结果中仍存在一定的噪声。而 SSFCNet（图 4.6 (i)所示）不仅克服了传统 Transformer 在空间建模和局部特征提取方面的不足，还通过深度融合空间-光谱信息，实现了更优的分类效果。

表 4.2 “高分五号”卫星高光谱数据集上的分类结果

Table 4.2 Classification result of the GF-5 satellite hyperspectral dataset

类别	SVM	PSPNet	3D CNN	RSSAN	DBDA	SpectralFormer	SSFTT	SSFCNet
互花米草	0.00	84.12	97.87	76.97	92.14	94.67	97.83	96.38
芦苇	81.77	34.26	94.44	94.58	85.42	92.71	97.25	93.54
柽柳	0.00	0.00	66.67	54.11	93.81	86.09	88.98	92.32
盐地碱蓬	0.00	0.00	60.19	67.05	65.38	78.49	77.21	84.10
稀疏潮滩芦苇	0.00	0.00	0.00	0.00	85.11	80.80	80.18	92.68
裸潮滩	93.63	96.11	86.08	89.05	84.04	87.80	94.13	88.87
水体	98.83	97.35	84.59	96.72	97.26	96.04	96.84	97.44
盐碱滩	83.33	61.42	96.30	77.50	83.33	86.11	82.80	93.48
OA (%)	72.13	73.37	82.49	85.19	90.26	91.50	93.43	94.07
AA (%)	44.70	46.66	73.27	69.55	85.81	87.84	89.40	92.35
Kappa	0.6105	0.6326	0.7716	0.7993	0.8682	0.8861	0.9118	0.9202

表 4.2 表明了 7 种对比方法和 SSFCNet 在“高分五号”卫星高光谱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 系数三个指标均取得了最优效果。其值分别达到 94.07%，92.35%和 0.9202。与 SSFTT 方法相比，SSFCNet 在 OA、AA、Kappa 系数上分别提升了 0.64%、2.95%、0.0084，显著优于其他对比方法。此外，在柽柳、盐地碱蓬和稀疏潮滩芦苇这三个类别上均取得了最高的分类精度，进一步证明了其在复杂地物

分类任务中的优越性。通过深度融合空间-光谱信息，SSFCNet 不仅克服了传统方法的局限性，还在分类性能上实现了显著提升。

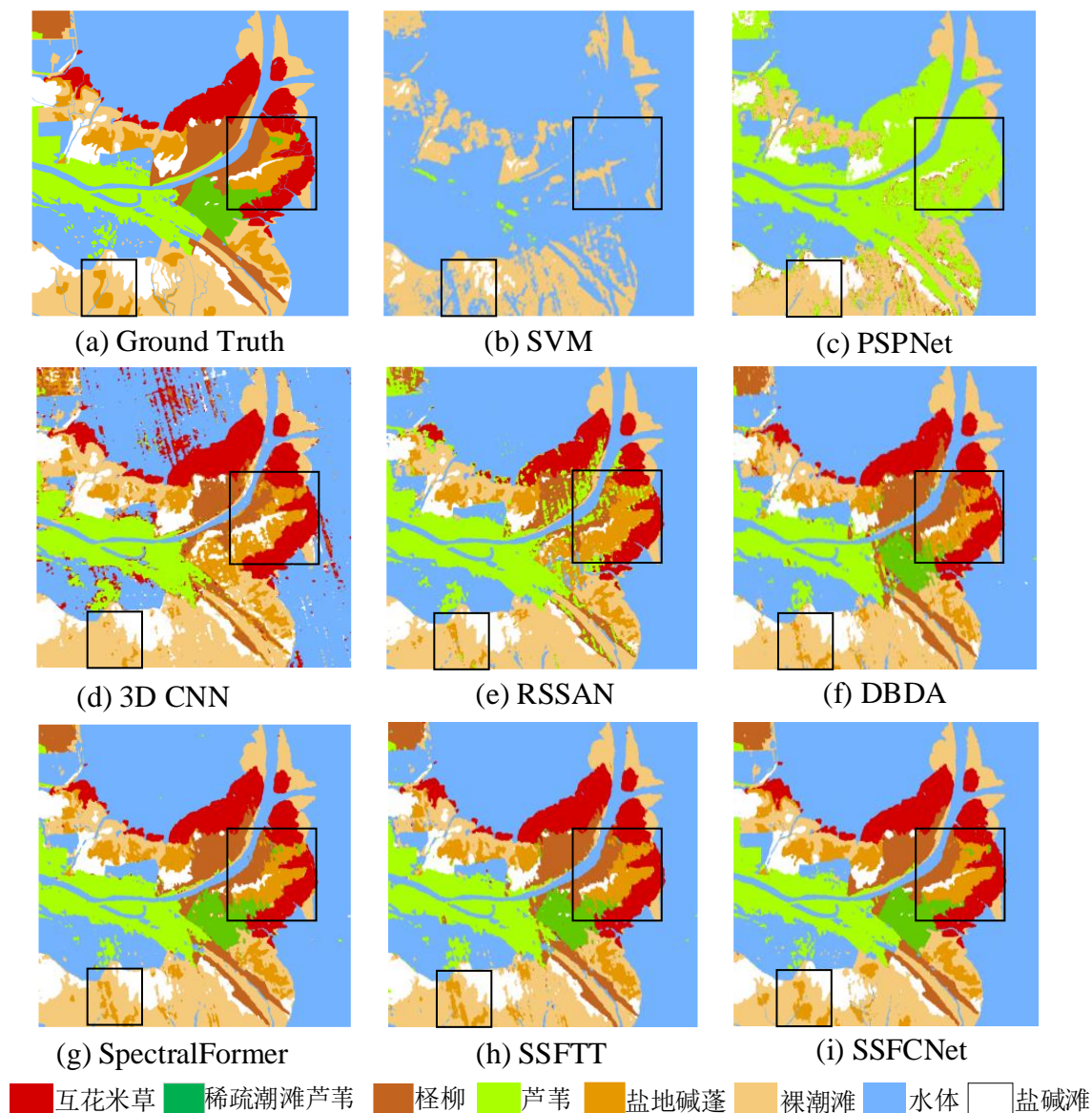


图 4.6 “高分五号”卫星高光谱数据集上的分类结果图

Fig. 4.6 Classification maps for the GF-5 satellite hyperspectral dataset

4.4.2 消融实验

在 SSFCNet 中引入通道先验卷积注意力模块（CPCA），模型能够有效处理复杂的空间结构和丰富的光谱信息。同时，本文提出了邻域卷积模块（NC），以增强对 Transformer 局部光谱域信息的关注能力。为了深入评估这两个模块对 SSFCNet 分类性能的贡献，本文在 Zhuhai-1 号和高分五号数据集上设计了消融实验，实验结果如表 4.3 所示。实验结果表明，即使在未引入邻域卷积模块和光谱注意力机制的情况下，SSFCNet 分类框架仍能取得较好的分类效果，这充分证明了所提出的 GCN 与 Transformer 融合架构在滨海湿地分类任务中的潜力。进一步分析发现，单独引入通道先验卷积注意力模块或邻域卷积模块均能显

著提升分类性能，而两者的结合则实现了更优的分类精度。这一现象表明，通道先验卷积注意力模块能够有效增强特征表示能力，而邻域卷积模块则能够强化 Transformer 对局部光谱信息的捕捉能力，二者协同作用进一步提升了模型的分类性能。

表 4.3 消融实验

Table 4.3 Ablation experiments

			Zhuhai-1			GF-5		
方法	CPCA 模块	NC 模块	OA(%)	AA(%)	Kappa	OA(%)	AA(%)	Kappa
SSFCNet	×	×	92.86	92.15	0.9223	92.88	91.90	0.9171
	√	×	93.07	92.38	0.9241	92.90	92.02	0.9192
	×	√	93.11	92.44	0.9259	93.05	92.18	0.9188
	√	√	94.26	93.07	0.9331	94.07	92.35	0.9202

4.4.3 模型分析

(1) 网络层对实验结果的影响：除了网络结构的设计外，模型中 GCN 和 Transformer 编码器的层数对分类性能具有显著影响。其中，GCN 的最优层数主要取决于所构建的邻接矩阵的稀疏性。过多的 GCN 层数通常会导致过度平滑现象，从而降低分类性能。相反，层数过少则可能无法充分挖掘节点间的高阶关系，导致特征表达能力不足。从图 4.7 的实验结果可以看出，当 GCN 层数设置为 3 层和 4 层时，模型在 Zhuhai-1 和高分五号卫星数据集上的总体分类精度达到最高，这表明适中的 GCN 层数能够有效平衡特征平滑与信息传递之间的关系。对于 Transformer 编码器，当层数分别设置为 5 层时，模型在两个数据集上均取得了最优的分类结果。这一现象表明，Transformer 编码器需要足够的深度来捕获遥感图像中复杂的光谱-空间依赖关系，但过深的层数也可能引入冗余计算和过拟合风险。总而言之，GCN 和 Transformer 编码器的层数优化需要综合考虑数据特性、任务需求以及计算效率，以实现模型性能的最优化。。

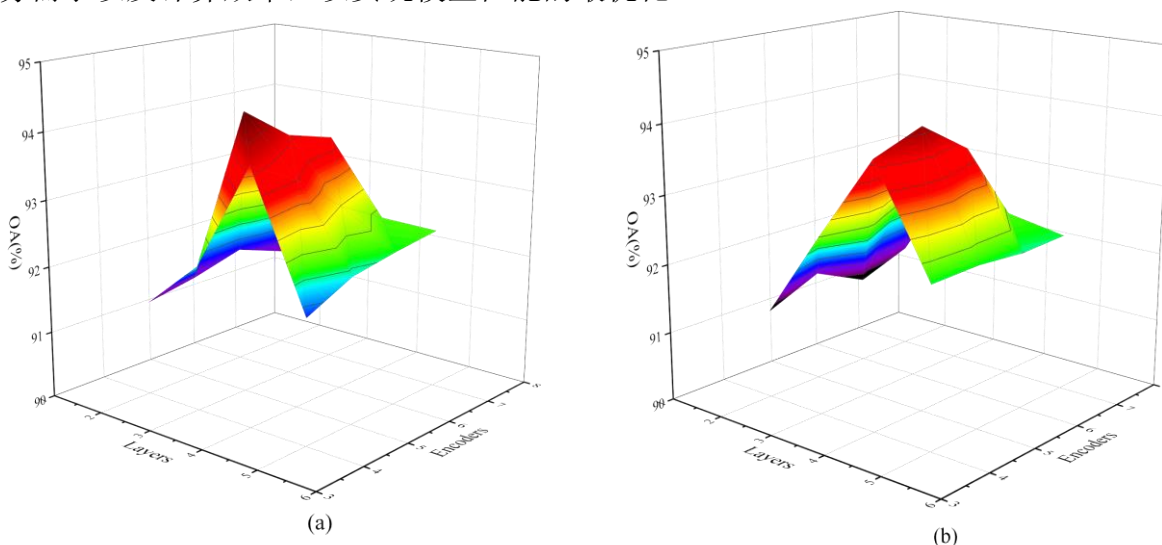


图 4.7 具有不同数量的 GCN 层数和编码器的 OA 比较

Table 4.7 OA comparison with different numbers of GCN layers and Encoders

(2) Patch 尺寸和变压器输入像素对分类精度的影响：本研究通过实验探究了 patch 尺寸和像素选择模块中像素数量对分类性能的影响。图 4.7(a)展示了不同 patch 尺寸对 Zhuhai-1 号和高分五号卫星数据集分类精度的定量影响。理论上，较大的补丁尺寸能够包含更丰富的空间上下文信息，但同时也可能引入更多噪声干扰，导致分类精度下降，并显著增加计算复杂度。实验结果表明，当 patch 尺寸设置为 9×9 时，模型在两个数据集上均取得了最优的分类性能，这证实了 patch 尺寸过大或过小都会对分类效果产生不利影响。此外，通过对像素选择模块中不同像素数量的实验分析表明，在 Zhuhai-1 号和高分五号数据集中，分别选取 19 和 21 个像素时可以获得最佳的分类效果，具体结果如图 4.7(b)所示。

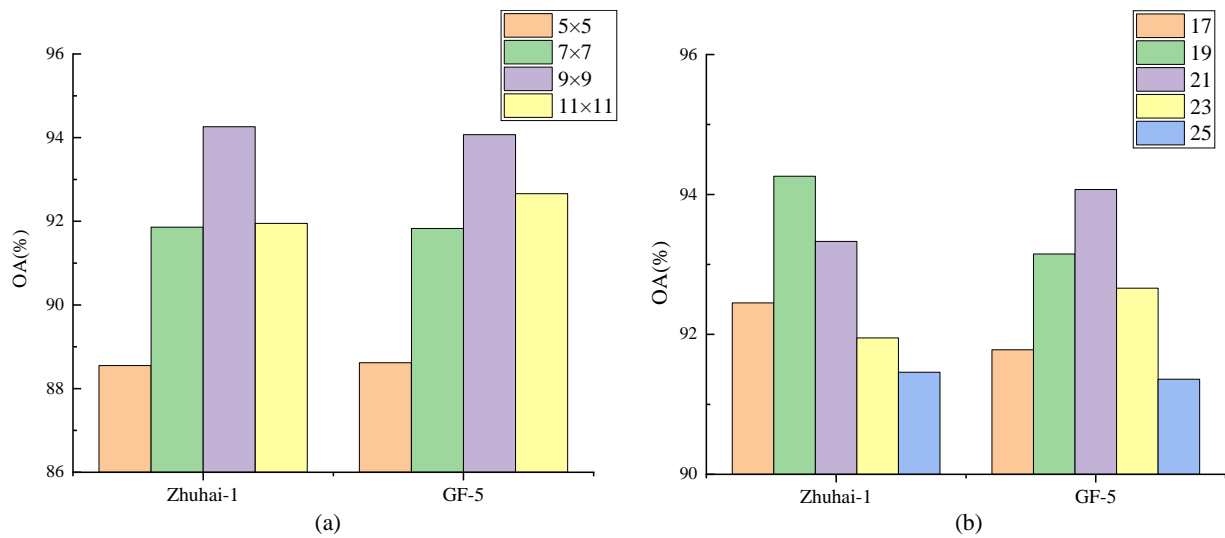


图 4.8 左侧展示了两个数据集在不同 Patch 大小下的 OA.

右侧展示了两个数据集在不同像素选择大小下的 OA

Fig. 4.8 The left side shows the OA of the two datasets with different patch sizes. The right side shows the OA of the two datasets with different pixel selection sizes.

4.4.4 计算时间和模型复杂性分析

为了全面评估不同方法的操作效率和计算消耗，我们在 Zhuhai-1 号和高分五号两个数据集上对表 4.4 列出的方法进行了训练和测试时间的对比实验。通过计算浮点运算次数（FLOPs），我们进一步量化了各模型的复杂度。为了保证实验的公平性，所有方法均在相同的计算平台上进行测试，并采用相同的超参数设置。实验结果表明，本文提出的 GCN 与 Transformer 联合分类架构 SSFCNet 在显著提升分类精度的同时，不可避免地增加了训练和测试的时间成本。然而，与其他先进的深度学习网络相比，SSFCNet 的训练和测试时间仍处于适中水平，完全在可接受的范围内。具体而言，SSFCNet 通过巧妙地融合 GCN 的空间结构建模能力和 Transformer 的光谱特征提取能力，在保证模型性能的同时，有效控制了计算复杂度。其 FLOPs 显著低于许多现有的复杂网络，这主要归功于高效的模块设计和参数优化策略。

表 4.4 不同方法计算时间和复杂性比较

Table 4.4 Comparison of different methods in terms of running time and complexity

数据集	指标	PSPNet	3D CNN	RSSAN	DBDA	SpectralFormer	SSFTT	SSFCNet
-----	----	--------	--------	-------	------	----------------	-------	---------

Zhuhai-1	Training time (s/epoch)	28.54	10.68	12.67	16.84	17.50	13.58	13.92
	Testing time (s)	6.48	3.76	4.08	10.69	9.44	4.73	4.05
	FLOPS (G)	29.56	20.73	20.55	30.55	26.58	30.38	21.87
GF-5	Training time (s/epoch)	29.28	11.56	12.33	14.84	14.13	14.13	11.51
	Testing time (s)	6.37	4.04	4.03	9.96	7.98	5.03	5.54
	FLOPS (G)	26.60	20.89	21.58	31.11	25.78	28.70	20.68

4.5 本章小结（Chapter Summary）

本章节详细介绍了所提出的基于空谱特征协同网络 SSFCNet。该方法创新性地结合了图卷积网络和 Transformer 的优势，分别用于提取高光谱图像中的空间信息和光谱信息，从而实现了滨海湿地复杂地物的高效特征提取与精准分类。首先，本章系统阐述了 SSFCNet 的整体架构设计，该架构引入了两个核心组件：通道感知注意力机制 GCN 和局部光谱特征融合 Transformer。其中，通道感知注意力机制 GCN 模块通过自适应学习通道间的关系，生成包含丰富光谱信息和空间结构信息的特征表示，显著提升了特征的表达能力；而局部光谱特征融合 Transformer 则通过设计的邻域卷积模块，增强了对局部光谱信息的建模能力，实现了对图像光谱域特征的高效提取。这种双模块协同的设计不仅充分挖掘了空谱特征的互补性，还显著提升了模型对复杂地物的分类性能。

5 总结与展望

5 Conclusion and prospects

5.1 总结 (Conclusion)

黄河口湿地位于中国山东省东营市，是黄河入海口形成的独特滨海湿地生态系统，也是我国暖温带地区面积最大、生态完整性最高的湿地之一。作为黄河流域与渤海交汇的生态屏障，黄河口湿地不仅拥有丰富的生物多样性，还因其独特的地理位置和生态环境，成为全球重要的候鸟迁徙中转站和栖息地，被誉为“地球之肾”和“天然碳库”。然而，近年来受人类活动和自然因素的双重影响，黄河口湿地的面积正面临持续缩减的威胁，其生态平衡和生物多样性保护也面临严峻挑战。因此，加强对黄河口湿地的科学监测与生态保护，对于维护区域生态安全、促进可持续发展具有重要意义。遥感技术中的高光谱成像是一种先进的对地观测手段，与传统的多光谱成像相比，高光谱图像具有更高的光谱分辨率，能够提供更加精细的光谱特征，从而实现了对地物的精准识别与分类。滨海湿地地物分类的复杂性和挑战性主要体现在：第一，滨海湿地地物的空间分布具有高度的复杂性和多样性，其地物类型通常呈现多尺度分布特征。第二，滨海湿地中的植被类型由于生理结构和环境干扰等原因，造成光谱特征上往往表现出较高的相似性。

(1) 针对滨海湿地地物多尺度特性导致的分类难题，本文提出了一种多分支多层次特征交互网络 MMFINet。MMFINet 引入多尺度注意力聚合模块，该模块通过空间和通道路径的融合，能够保留更多的细节信息，尤其是在处理小尺度地物时，能够更好地捕捉到细微的结构变化。另外，设计多分支解码器，该结构通过跳跃连接和上采样块，将不同层次的特征图进行融合。这种设计不仅能够有效地将低层次的细节信息与高层次的语义信息有机结合，还在最终输出中保留了丰富的空间细节，从而提高了分类精度。同时，多分支解码器通过设计不同深度的分支，能够自适应地处理不同尺度的特征信息，进一步增强了对多尺度地物的识别能力。

(2) 针对滨海湿地地物空间分布复杂、光谱特征相似性高导致的分类挑战。本文提出了一种空谱特征协同网络 SSFCNet，旨在充分利用高光谱图像的空间-光谱域信息，实现细粒度地物分类。通过引入通道先验卷积注意力机制 GCN 强化光谱-空间特征的提取能力。该模块引入了通道先验卷积注意力，能够动态分配通道权重，聚焦于关键光谱波段，从而更精准地提取具有判别性的光谱特征。并进一步提出了局部光谱特征融合 Transformer 模块，有效增强了 Transformer 对局部特征的捕捉能力，通过融合局部光谱信息，提升了模型对光谱细节的表征能力。实验结果表明，SSFCNet 能够充分挖掘空间和光谱特征的互补性，在滨海湿地地物分类任务中表现出优异的性能，其分类精度显著优于现有主流算法。

5.2 展望 (Prospects)

尽管本文提出的方法在黄河口滨海湿地植被分类中取得了较好的效果，但仍存在一些值得进一步探索的方向。因此，下一步的研究计划将从以下几个方面出发：

- （1）模型泛化能力的提升：当前研究主要基于黄河口滨海湿地的特定数据集，未来可以尝试将模型应用于其他类型的湿地生态系统，验证其泛化能力。
- （2）多源数据融合：本文主要基于高光谱遥感影像进行研究，未来可以尝试融合多源数据（如激光雷达等），以获取更丰富的地物特征信息。
- （3）模型效率的优化：当前模型在计算效率和参数量方面仍有优化空间，未来可以尝试设计更轻量化的网络结构。

参考文献

- [1] Song S , Wu Z , Wang Y ,et al. Mapping the Rapid Decline of the Intertidal Wetlands of China Over the Past Half Century Based on Remote Sensing[J]. *Frontiers in Earth Science*, 2020, 8.
- [2] 贾鹏, 贾静, 李德志. 滨海湿地土壤种子库及其在生态恢复中的潜在价值[J]. *Biology Teaching*, 2024, 49(3).
- [3] Kirwan M L , Megonigal J P . Tidal wetland stability in the face of human impacts and sea-level rise[J]. *Nature*, 2013, 504(7478):53–60.
- [4] 李荣冠, 王建军, 林俊辉. 福建典型滨海湿地: Typical coastal wetlands in Fujian Province[M]. 科学出版社, 2014.
- [5] Delgado L E , Víctor H. Marín. Interannual Changes in the Habitat Area of the Black-Necked Swan, *Cygnus melancoryphus*, in the Carlos Anwandter Sanctuary, Southern Chile: A Remote Sensing Approach[J]. *Wetlands*, 2013, 33:91-99.
- [6] 于淼, 栗云召, 屈凡柱, 等. 黄河三角洲滨海湿地退化过程的时空变化及预测分析[J]. *农业环境与发展*, 2020, 037(004):484-492.
- [7] 李有志, 崔丽娟, 潘旭, 等. 辽河口湿地植物多样性及物种功能型空间分布格局[J]. *生物多样性*, 2015, 23(4):471-478.
- [8] 李彦胜, 张永军. 耦合知识图谱和深度学习的新一代遥感影像解译范式[J]. *Geomatics & Information Science of Wuhan University*, 2022, 47(8):1176-1190.
- [9] Ren K , Sun W , Meng X ,et al. Fusing China GF-5 Hyperspectral Data with GF-1, GF-2 and Sentinel-2A Multispectral Data: Which Methods Should Be Used?[J]. *Remote Sensing*, 2020, 12(5): 882.
- [10] Sun W , Du Q .Hyperspectral Band Selection: A Review[J]. *IEEE Geoscience and Remote Sensing Magazine*, 2019, 7(2):118-139.
- [11] 吴蕴泽, 王健, 燕纬双, 等. 基于 TWDTW 的黄河三角洲湿地植被分类研究[J]. *测绘科学技术*, 2022, 10(3):10.
- [12] 于淼, 栗云召, 屈凡柱, 等. 黄河三角洲滨海湿地退化过程的时空变化及预测分析[J]. *农业环境与发展*, 2020, 037(004):484-492.
- [13] 范青. 基于红树林湿地生态恢复景观规划设计—以珠海大门口水道湿地公园为例[J]. *绿色科技*, 2022, 24(24):135-139.
- [14] Hu Y , Zhang J , Ma Y ,et al. Hyperspectral Coastal Wetland Classification Based on a Multiobject Convolutional Neural Network Model and Decision Fusion[J]. *IEEE Geoscience and Remote Sensing Letters*, 2019:1-5.
- [15] Delancey E R , Simms J F , Mahdianpari M ,et al. Comparing Deep Learning and Shallow Learning for Large-Scale Wetland Classification in Alberta, Canada[J].*Remote Sensing*, 2019, 12(2).
- [16] 曹庆安. 基于时序 NDVI 差值的早稻种植信息识别[J]. *江西科学*, 2023, 41(3):544-549.
- [17] 夏玉松, 周启刚, 李 辉, 等. 基于决策树与 NDVI 时序变化检测的撂荒耕地的地形特征研究—以重

- 庆市巫山县为例[J]. Bulletin of Soil & Water Conservation, 2024, 44(4).
- [18] 尚洁琼. 遥感时序去噪算法的系统评价[J]. 地球科学前沿 (汉斯), 2024, 14(10):1279-1287.
- [19] Hearst M A , Dumais S T , Osman E ,et al. Support vector machines[J].IEEE Intelligent Systems & Their Applications, 1998, 13(4):18-28.
- [20] Kotsiantis, S. B . Decision trees: a recent overview[J]. Artificial Intelligence Review, 2013, 39(4):261-283.
- [21] 潘煜琳, 林辉. 基于国产高分遥感数据的黑龙江三江湿地分类[J]. 中南林业科技大学学报, 2022(009):042.
- [22] 沈聪颖, 甘淑, 李新澳, 等. 基于 GF-5 高光谱特征分析的水体提取方法研究[J]. 全球定位系统, 2021, 46(5):6.
- [23] Wong J A H A .Algorithm AS 136: A K-Means Clustering Algorithm[J].Journal of the Royal Statistical Society, 1979, 28(1):100-108.
- [24] Ball G H , Hall D J .ISODATA, A NOVEL METHOD OF DATA ANALYSIS AND PATTERN CLASSIFICATION[J].stanford research institute, 1965.
- [25] Khouj Y , Dawson J , Coad J ,et al. Hyperspectral Imaging and K-Means Classification for Histologic Evaluation of Ductal Carcinoma In Situ[J].Frontiers in Oncology, 2018, 8.
- [26] Hong P , Ye X F , Yu H ,et al. An improved hyperspectral image classification approach based on ISODATA and SKR method[C]//Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series.International Society for Optics and Photonics, 2016.
- [27] Lecun Y , Bengio Y , Hinton G .Deep learning[J].Nature, 2015, 521(7553):436.
- [28] 周飞燕, 金林鹏, 董军, 等. 卷积神经网络研究综述[J].计算机学报, 2017, 40(6):23.
- [29] Li G , Zhang C , Lei R ,et al. Hyperspectral remote sensing image classification using three-dimensional-squeeze-and-excitation-DenseNet (3D-SE-DenseNet)[J].Remote Sensing Letters, 2020, 11(2):195-203.
- [30] Kim Y .Convolutional Neural Networks for Sentence Classification[C]//2014.
- [31] Elman J L .Finding Structure in Time[J].Cognitive Science, 1990, 14(2):179-211.
- [32] Hochreiter S , Schmidhuber J .Long Short-Term Memory[J].Neural Computation, 1997, 9(8):1735-1780.
- [33] Lee H , Kwon H .Going Deeper With Contextual CNN for Hyperspectral Image Classification[J].IEEE Trans Image Process, 2017, 26(10):4843-4855.
- [34] Meng Z , Jiao L , Liang M ,et al. A Lightweight Spectral-Spatial Convolution Module for Hyperspectral Image Classification[J].IEEE Geoscience and Remote Sensing Letters, 2021.
- [35] Mou L , Ghamisi P , Zhu X X. Deep Recurrent Neural Networks for Hyperspectral Image Classification[J].IEEE Transactions on Geoscience & Remote Sensing, 2017:3639-3655.
- [36] Vaswani A , Shazeer N , Parmar N ,et al. Attention Is All You Need[J].arXiv, 2017.
- [37] He X , Chen Y , Lin Z .Spatial-Spectral Transformer for Hyperspectral Image Classification[J].Remote Sensing, 2021, 13(498).
- [38] 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.
- [39] Dosovitskiy A , Beyer L , Kolesnikov A ,et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale[C]//International Conference on Learning Representations.2021.
- [40] He W , Huang W , Liao S ,et al. CSiT: A Multiscale Vision Transformer for Hyperspectral Image Classification[J].IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15:9266-9277.
- [41] Liu Z , Lin Y , Cao Y ,et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows[J]. 2021.
- [42] R. Ghotekar, K. Shaw, M. Rout,et al. "End-to-End Hyperspectral Image Classification Using Hybrid GCNN with SWIN Transformer," 2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIIIE), Ballari, India, 2023, pp. 1-8,
- [43] Gopinath R A .Maximum Likelihood Modeling With Gaussian Distributions For Classification[C]//Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on IEEE, 1998.
- [44] Breiman L .Random forest[J].Machine Learning, 2001, 45:5-32.
- [45] 周国琼, 段海军, 陈剑鸣, 等. 基于模糊 C 均值聚类改进的最大似然分类法[J]. 科学技术与工程, 2012, 20(19):4.
- [46] 李垒, 任越美. 基于随机森林的高光谱遥感图像分类[J].计算机工程与应用, 2016, 52(24):5.
- [47] Cover T , Hart P .Nearest neighbor pattern classification[J].IEEE Transactions on Information Theory, 2003, 13(1):21-27.
- [48] Pan,Quan,Jiao,et al. Belief rule-based classification system: Extension of FRBCS in belief functions framework[J].Information Sciences: An International Journal, 2015.
- [49] 马宇龙, 林志垒. 基于面向对象和 CART 决策树方法的遥感影像湿地变化检测研究—以龙祥岛地区为例[J]. 福建师范大学学报: 自然科学版, 2017, 33(6):12.
- [50] 庄小燕, 况润元. 基于 Sentinel-2 数据的南矾湿地面向对象决策树分类[J]. 湿地科学, 2024, 22(5):630-640.
- [51] 王娅, 郭继发, 林雨. 基于面向对象方法的天津市滨海新区湿地信息提取[J].山东理工大学学报:自然科学版, 2023, 37(6):25-31.
- [52] 莫利江, 曹宇, 胡远满, 等. 面向对象的湿地景观遥感分类—以杭州湾南岸地区为例[J].湿地科学, 2012, 10(2):8-13.
- [53] 谢锦莹, 丁丽霞, 王志辉, 等. 基于 FCN 与面向对象的滨海湿地植被分类[J]. 林业科学, 2020, 56(8):9-12.
- [54] 杨朝辉, 白俊武, 陈志辉, 等. 利用 Sentinel-2A 影像的面向对象特征湿地决策树分类方法[J]. 测绘科学技术学报, 2019, 36(3):7-9.
- [55] 朱长明, 李均力, 张新, 等. 面向对象的高分辨率遥感影像湿地信息分层提取[J]. 测绘通报,

2014(10):6.

- [56] Chen C , Tian B , Zhou Y Z C W W . Mapping intertidal topographic changes in a highly turbid estuary using dense Sentinel-2 time series with deep learning[J].ISPRS journal of photogrammetry and remote sensing, 2023, 205(Nov.):1-16.
- [57] Liu C , Tao R , Li W ,et al. Joint Classification of Hyperspectral and Multispectral Images for Mapping Coastal Wetlands[J].IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, PP(99):1-1.
- [58] 史特. 基于生成对抗网络的跨域遥感图像语义分割[D].武汉大学,2023.
- [59] 田萱, 王亮, 丁琪. 基于深度学习的图像语义分割方法综述[J]. 软件学报, 2019, 30(2):29.
- [60] Ronneberger O , Fischer P , Brox T .U-Net: Convolutional Networks for Biomedical Image Segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention.Springer International Publishing, 2015.
- [61] Chen L C , Zhu Y , Papandreou G ,et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation[C]//European Conference on Computer Vision.Springer, Cham, 2018.
- [62] He K , Zhang X , Ren S ,et al. Deep Residual Learning for Image Recognition[J].IEEE, 2016.
- [63] Zhao H , Shi J , Qi X ,et al. Pyramid Scene Parsing Network[J].IEEE Computer Society, 2016.
- [64] Badrinarayanan V , Kendall A , Cipolla R . SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation[J].IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017:1-1.
- [65] Li R , Zheng S , Duan C ,et al. Multi-stage Attention ResU-Net for Semantic Segmentation of Fine-Resolution Remote Sensing Images[J]. 2020.
- [66] Yang Z , Bi F , Hou X ,et al. DDRNet: Dual-Domain Refinement Network for Remote Sensing Image Semantic Segmentation[J].IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17[2025-03-03].
- [67] S. Mei, C. Song, M. Ma and F. Xu, "Hyperspectral Image Classification Using Group-Aware Hierarchical Transformer," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-14, 2022, Art no. 5539014,
- [68] Guo M, Lu C, Hou Q, et al. SegNeXt: Rethinking Convolutional Attention Design for Semantic Segmentation. ArXiv, (2022)abs/2209.08575.
- [69] Huang Z , Wang X , Wei Y ,et al. CCNet: Criss-Cross Attention for Semantic Segmentation[J].IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, PP(99):1-1.
- [70] Tran D , Bourdev L , Fergus R ,et al. Learning Spatiotemporal Features with 3D Convolutional Networks[C]//IEEE International Conference on Computer Vision.IEEE, 2015.
- [71] M. Zhu, L. Jiao, F. Liu, S. Yang and J. Wang, "Residual Spectral–Spatial Attention Network for Hyperspectral Image Classification," in IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(1):449-462.

- [72] Li R , Zheng S , Duan C ,et al. Classification of Hyperspectral Image Based on Double-Branch Dual-Attention Mechanism Network[J].Remote. Sens. 2019, 12:582.
- [73] Hong D , Han Z , Yao J ,et al. SpectralFormer: Rethinking Hyperspectral Image Classification with Transformers[J]. 2021.
- [74] 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.

作者简历

一、基本情况

姓名：白雪梅 性别：女 民族：汉 出生年月：1998-04-05 籍贯：山东省菏泽市

- 2020.09-2022.06 济宁学院
- 2022.09-2025.06 山东科技大学

二、攻读学位期间获奖情况

- 2022 年 10 月获山东科技大学二等学业奖学金；
- 2023 年 11 月获优秀研究生荣誉称号；
- 2023 年 10 月获山东科技大学二等学业奖学金；
-
- 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；
- 2023 年 4 月获山东科技大学“优秀共青团员”称号；
- 2023 年 10 月获山东科技大学科技活动先进个人

三、攻读学位期间研究项目情况

- Zhao H, Bai X Y, Zeng Q T, Zhou H, Bai X M. Nested Entity Recognition Method Based on Multidimensional Features and Fuzzy Localization, Neural Processing Letters, 2024, 56(3):1-21
- BiJiaNet-面向海岸带动态检测的智能遥感大模型。2023-2024 年，算法开发人员。
- 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
- 基于国产高分数据和遥感大模型的海岸线自动提取方法研究.山东省自然科学基金项目（面上项目）。参与人员

致 谢

回首求学之路，感慨万千。这一年，我 27 岁，学生生涯即将画上句号。时光荏苒，韶华易逝，曾经的点点滴滴仿佛昨日，而如今已站在人生的新起点上。这段旅程中，有太多的人值得感谢，是他们的支持与陪伴让我走到了今天。

师恩如海，铭记于心。首先，我要衷心感谢我的导师崔宾阁教授。在学术道路上，您严谨的治学态度、渊博的学识和耐心的指导让我受益匪浅。无论是科研中的困惑，还是生活中的迷茫，您总是给予我无私的帮助和鼓励。您的教诲不仅让我在学术上不断进步，更让我学会了如何做人、如何面对挑战。同时，我也要感谢课题组的路燕老师和黄玲老师，感谢你们在实验和论文撰写过程中给予的宝贵建议和支持。

恩深似海，情暖如春。感谢我的父母，你们是我生命中最坚实的后盾。你们的无私付出和无条件的支持，让我能够心无旁骛地追求学业。你们的爱与关怀，是我前行的动力。愿时光温柔以待，愿我能够成为你们的骄傲，用未来的成就回报你们的养育之恩。在这还要感谢我的同卵双胞胎姐姐，是她鼓励我支持我，更是我成长路上的坚实依靠，虽然我们生活中有小吵小闹，但是她永远是最爱的老姐。最后要感谢的是我家大恒子，他不仅是我生命中最特别的存在，更是我未来人生旅途中最亲密的伴侣，我们将携手组建一个充满爱与温暖的小家庭。无论何时何地，他始终以无限的耐心和理解支持着我，他的存在，让我对未来充满期待。

同窗情谊，温暖岁月。感谢我的舍友和师门兄弟姐妹们。我们一起度过了无数个日夜，分享着彼此的喜怒哀乐。你们的陪伴让我的求学之路充满了欢声笑语，也让那些艰难的时刻变得不再孤单。感谢你们在我迷茫时给予的鼓励，在我成功时分享的喜悦。愿师哥师姐们在未来的工作中大展宏图，愿师弟师妹们在科研的道路上勇攀高峰。

未来可期，砥砺前行。学生时代的结束，意味着新的人生篇章即将开启。带着这份感激与期待，我将以更加坚定的信念和饱满的热情，迎接未来的挑战与机遇。愿未来的日子里，我能够不负韶华，勇敢追梦，书写属于自己的精彩人生。

最后，我要感谢参与我论文评审和答辩的各位老师。你们的宝贵意见和建议让我的论文更加完善，也让我对学术研究有了更深的理解。你们的严谨治学态度和敬业精神，将永远激励我在未来的道路上不断追求卓越。

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*	白雪梅	学号*	202283060072	
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	2025	
论文提交日期*	2025 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

基于多层次特征交互与空谱特征协同的滨海湿地遥感 精细制图方法研究

Research on the Fine Mapping Method of Remote Sensing of Coastal
Wetlands based on Multi-level Feature Interaction and Spatial-Spectral
Feature coordination

作 者 白雪梅

导 师 崔宾阁 教授

山东科技大学

二〇二五年五月