

工程硕士学位论文

**基于深度学习的多尺度马尾藻提取与边缘优化网络方法研究**

Research on Multi-scale Sargassum Extraction and Edge Optimization Networks Based on Deep Learning

**作 者 张艳平**

**导 师 路 燕 副教授**

山东科技大学

二〇二五年六月

中图分类号 TP751 学校代码 10424

UDC 004 密 级 公开

山东科技大学

工程硕士学位论文

**基于深度学习的多尺度马尾藻提取与边缘优化网络方法研究**

Research on Multi-scale Sargassum Extraction and Edge Optimization Networks Based on Deep Learning

作 者 张艳平 入学时间 2022年9月

导 师 路 燕 职 称 副教授

申请学位 工程硕士 所在学院 计算机科学与工程

学科（类别） 计算机技术 方向（领域） 人工智能

答辩日期 2025年5月 日 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： 导师签名：

日 期： 年 月 日 日 期： 年 月 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：

年 月 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：

**摘 要**

马尾藻是一种常见的海洋生态灾害，不仅影响海洋生态系统的健康和稳定，也对人类健康和经济社会发展构成威胁，及时准确获取马尾藻发生和分布范围对马尾藻的预警和防治具有重要意义。由于传统遥感监测方法受限于中低分辨率影像及复杂海洋环境，难以实现多尺度目标精确提取与边缘细节有效保留，而GF-1遥感影像因具有高分辨率多光谱成像能力、大幅宽覆盖、高时间分辨率和稳定的数据产品等特点，使其在遥感监测领域具有广泛的应用潜力，故本文针对马尾藻形态多变、尺度各异的特点及边缘区域水色混淆问题，基于GF-1高分辨率遥感影像，提出两种创新深度学习网络模型，主要研究内容如下：

针对马尾藻形态多样，尺度各异，小尺度马尾藻因其细节信息不足而难以精确提取，提出基于多维重组特征的上下文提取网络（MRF-Net）。该网络通过多维重组特征融合模块（GMFR）实现通道-空间细粒度重组，结合哈达玛乘积与门控机制增强多尺度特征识别能力，引入多尺度上下文提取模块（MOCE）通过不同扩张率的卷积核和不同尺寸的池化操作，来实现对全局与局部特征的有效提取，并通过卷积注意力模块（CBAM）优化特征融合。实验表明，MRF-Net不仅能较准确的识别马尾藻目标，而且对于处理多形态和尺度的马尾藻具有良好鲁棒性，可以有效解决传统方法对小尺度目标漏检及形态适应性差的问题。

针对马尾藻边缘区域生物量密度较低且易于周围水体像素混淆，难以精确提取，提出基于多层注意力和顺序聚合的马尾藻边缘优化网络（SEO-Net）。该网络通过引入多层注意力模块，使用核注意力（KAM）与通道注意力（CAM）的来强化边缘信息，通过引入顺序信息聚合模块（SAB）来融合多层次边界特征，其中SBA模块利用重新矫正注意力单元（RAU）选择性的聚合从低级到高级的特征，从而更精细的提取到目标轮廓。实验表明，SIA-Net能够**实现复杂边界的精细化提取**，生成更加准确的边界信息，有效提升了马尾藻边缘区域的识别精度。

**关键词**：马尾藻探测；多维重组特征；门控机制；注意力机制；语义分割

**Abstract**

Sargassum is a common marine ecological disaster, which not only threatens the health and stability of marine ecosystems, but also poses a risk to human health and economic and social development. Hence, it is of great importance for its early warning and prevention to access timely and accurate information about the occurrence and distribution range of Sargassum. Because the traditional remote sensing monitoring methods are limited by the low and medium resolution images and complex marine environment, it is difficult to realize the precise extraction of multi-scale targets and effective retention of edge details, while the GF-1 remote sensing image has a great range of potential applications in the field of remote sensing monitoring due to its high-resolution multi-spectral imaging capability, large width coverage, high temporal resolution, and stable data products, therefore, this paper centers on the characteristics of variable morphology and different scales and the characteristics of the morphology and scale of Sargassum. Therefore, there are two innovative deep learning network models proposed in this paper based on GF-1 high-resolution remote sensing images to address the characteristics of Sargassum with variable morphology, different scales, and the problem of confusing water color in the edge region, and the main research contents are as follows:

Aiming at the diverse morphology and different scales of Sargassum, the small-scale Sargassum is difficult to be extracted accurately due to its insufficient detail information, a context extraction network based on multidimensional reorganization features (MRF-Net) is proposed. The network realizes channel-space fine-grained reorganization through the multidimensional reorganization feature fusion module (GMFR), combines the Hadamard product with the gating mechanism to enhance the multi-scale feature recognition ability, introduces the multi-scale oriented context extraction module (MOCE) to use the convolution kernel with different expansion rates and the pooling operation of different sizes to capture global and local features, and optimizes the feature fusion through the convolutional attention module (CBAM). Fusion. Experiments show that MRF-Net can not only recognize sargassum targets more accurately, but also has good robustness in dealing with sargassum of multiple morphologies and scales, which can effectively solve the problems of omission detection and poor morphological adaptability of traditional methods for small-scale targets.

Aiming at the edge region of Sargassum where the biomass density is low and easy to be confused with the pixels of the surrounding water body, which is difficult to be extracted accurately, the Sargassum Edge Optimization Network (SEO-Net) based on multi-layer attention and sequential aggregation is proposed. The network enhances edge information by introducing a multilayer attention module using kernel attention (KAM) and channel attention (CAM), and fuses multilevel boundary features by introducing a sequential information aggregation module (SAB), in which the SBA module selectively aggregates features from low-level to high-level by using recalibrated attention units (RAUs) to extract the target contour in a finer way. Experiments show that SIA-Net can realize the fine extraction of complex boundaries, generate more accurate boundary information, and effectively improve the recognition accuracy of the edge region of Sargassum.

**Keywords**: Sargassum detection; Multidimensional reorganized features; Gating mechanism; Attention mechanism; Semantic segmentation

目 录

[**图清单** I](#_Toc192592567)

[**表清单** III](#_Toc192592568)

[**变量注释表** IV](#_Toc192592569)

[**变量注释表** V](#_Toc192592570)

[**变量注释表** VI](#_Toc192592571)

[1 绪 论 1](#_Toc192592572)

[1.1 研究背景与意义（Research Background and Significance） 1](#_Toc192592573)

[1.2 国内外研究现状（The research status at home and abroad） 2](#_Toc192592574)

[1.3 问题的提出（**Statement of problem**） 5](#_Toc192592575)

[1.4 主要研究内容（Main research contents） 6](#_Toc192592576)

[1.5 章节安排（Chapter arrangement） 7](#_Toc192592577)

[2 相关技术 9](#_Toc192592578)

[2.1 植被指数法（**Vegetation Index Method**） 9](#_Toc192592579)

[2.2 语义分割（**Semantic Segmentation**） 10](#_Toc192592580)

[2.3 卷积神经网络（**Convolutional Neural Network**） 10](#_Toc192592581)

[2.4 典型语义分割模型（**Typical Semantic Segmentation Models**） 12](#_Toc192592582)

[2.5 门控机制（**Gating Mechanism**） 16](#_Toc192592583)

[2.6 注意力机制（**Attention Mechanism**） 18](#_Toc192592584)

[2.7 本章小结（Summary） 22](#_Toc192592585)

[3 基于多维重组特征的多尺度马尾藻提取网络 23](#_Toc192592586)

[3.1 引言（Foreword） 23](#_Toc192592587)

[3.2 MRFNet网络结构设计（Design of MRFNet Network Structure） 24](#_Toc192592588)

[3.3 实验准备（Experimental preparation） 28](#_Toc192592589)

[3.4 实验结果与分析（Experimental results and analysis） 32](#_Toc192592590)

[3.5 本章小结（Summary） 38](#_Toc192592591)

[4 基于多层注意力和顺序聚合的马尾藻边缘优化网络 39](#_Toc192592592)

[4.1 引言（Foreword） 39](#_Toc192592593)

[4.2 SEONet网络结构设计（Design of SEONet Network Structure） 40](#_Toc192592595)

[4.3 实验设计（Design of Experiment） 43](#_Toc192592596)

[4.4 实验结果与分析（Experimental results and analysis） 44](#_Toc192592597)

[4.5 本章小结（Summary） 49](#_Toc192592598)

[5 总结与展望 50](#_Toc192592599)

[5.1 总结（Conclusion） 50](#_Toc192592600)

[5.2 展望（Prospects） 50](#_Toc192592601)

[参考文献 52](#_Toc192592602)

[作者简历 60](#_Toc192592603)

[致 谢 61](#_Toc192592604)

[学位论文数据集 62](#_Toc192592605)

Content

[**List of Figures** I](#_Toc192097688)

[**List of Tables** III](#_Toc192097689)

[**List of Variables** IV](#_Toc192097690)

[1 Introduction 1](#_Toc192097692)

[1.3 Statement of problem 5](#_Toc192097693)

[1.4 Main research contents 6](#_Toc192097694)

[1.5 Chapter arrangement 7](#_Toc192097695)

[2 Related technologies 9](#_Toc192097697)

[2.1 Vegetation Index Method 9](#_Toc192097698)

[2.2 Semantic Segmentation 10](#_Toc192097699)

[2.3 Convolutional Neural Network 10](#_Toc192097700)

[2.4 Typical Semantic Segmentation Models 12](#_Toc192097701)

[2.5 Gating Mechanism 16](#_Toc192097702)

[2.6 Attention Mechanism 18](#_Toc192097703)

[2.7 Summary 22](#_Toc192097704)

[3 Multiscale Sargassum Extraction Network Based on Multidimensional Reorganized Features 23](#_Toc192097706)

[3.1 Foreword 23](#_Toc192097707)

[3.2 Design of MRFNet Network Structure 24](#_Toc192097708)

[3.3 Experimental preparation 28](#_Toc192097709)

[3.4 Experimental results and analysis 32](#_Toc192097710)

[3.5 Summary 38](#_Toc192097711)

[4 Sargassum Edge Optimization Network Based on Multi - Layer Attention and Sequential Aggregation 39](#_Toc192097713)

[4.1 Foreword 39](#_Toc192097714)

[4.2 Design of SEONet Network Structure 40](#_Toc192097716)

[4.3 Design of Experiment 43](#_Toc192097717)

[4.4 Experimental results and analysis 44](#_Toc192097718)

[5 Conclusion and prospects 50](#_Toc192097720)

[5.1 Conclusion 50](#_Toc192097721)

[5.2 Prospects 50](#_Toc192097722)

[References](#_Toc164026784)

[Author’s Resume](#_Toc164026785)

[Acknowledgements](#_Toc164026786)

[Thesis Data Collection](#_Toc164026787)

**图清单**

|  |  |  |
| --- | --- | --- |
| 图序号 | 图名称 | 页码 |
| 图2.1 | 卷积神经网络结构图 | 11 |
| Fig. 2.1 | **Figure 2.1** The schematic diagram **of CNN structure** | 11 |
| 图2.2 | 语义分割网络流程图 | 12 |
| Fig. 2.2 | **Figure 2.2** The schematic diagram **of Semantic Segmentation Network** | 12 |
| 图2.3 | FCN网络结构图 | 13 |
| Fig. 2.3 | The schematic diagram of the **FCN** structure | 13 |
| 图2.4 | UNet网络结构图 | 14 |
| Fig. 2.4 | The schematic diagram of the U-Net structure | 14 |
| 表2.5 | 残差单元结构图 | 15 |
| Fig. 2.5 | The schematic diagram of **of the Residual Unit** structure | 15 |
| 图2.6 | LSTM结构图 | 16 |
| Fig. 2.6 | The schematic diagram of the **LSTM** structure | 16 |
| 图2.7 | GRU结构图 | 18 |
| Fig. 2.7 | The schematic diagram of **GRU Structure** | 18 |
| 图2.8 | 点积注意力结构图 | 19 |
| Fig. 2.8 | The structure diagram of  **Dot-Product Attention Mechanism** | 19 |
| 图2.9 | 通道注意力结构图 | 20 |
| Fig. 2.9 | The structure diagram of **Channel Attention Mechanism** | 20 |
| 图2.10 | 空间注意力结构图 | 21 |
| Fig. 2.10 | The structure diagram of Spatial **Attention Mechanism** | 21 |
| 图2.11 | CBAM结构图 | 22 |
| Fig. 2.11 | The schematic diagram of **CBAM Structure** | 22 |
| 图3.1 | 不同尺度的马尾藻遥感影像全局及局部图 | 23 |
| Fig. 3.1 | Global and Local Views of Sargassum at Different Scales in Remote Sensing Images | 23 |
| 图3.2 | MRFNet网络总体结构 | 24 |
| Fig. 3.2 | Overall Structure of the MRFNet Network | 24 |
| 图3.3 | **GMFR**结构图 | 25 |
| Fig. 3.3 | Structural Schematic Diagram of GMFR | 25 |
| 图3.4 | 空间重组单元结构图 | 26 |
| Fig. 3.4 | Structural Diagram of the Spatial Reorganization Unit | 26 |
| 图3.5 | MSDM结构图 | 28 |
| Fig. 3.5 | Structural Schematic Diagram of MSDM | 28 |
| 图3.6 | 基于MRFNet的马尾藻提取流程 | 30 |
| Fig. 3.6 | The Sargassum Extraction Process Based on MRFNet | 30 |
| 图3.7 | 测试图像及真值图 | 32 |
| Fig. 3.7 | Test Image and Ground Truth Map | 32 |
| 图3.8 | 测试图像1马尾藻提取定性结果 | 33 |
| Fig. 3.8 | Qualitative Results of Sargassum Extraction for Test Image 1 | 33 |
| 图3.9 | 测试图像2马尾藻提取定性结果 | 35 |
| Fig. 3.9 | Qualitative Results of Sargassum Extraction for Test Image 2 | 35 |
| 图3.10 | 马尾藻不同阶段网络模型可视化结果 | 37 |
| Fig. 3.10 | Visualization Results of Sargassum at Different Stages of Network Model | 37 |
| 图4.1 | 马尾藻遥感影像和局部放大图 | 39 |
| Fig. 4.1 | Remote Sensing Images of Sargassum and Their Local Magnified Views | 39 |
| 图4.2 | SEONet总体结构图 | 41 |
| Fig. 4.2 | Schematic Diagram of the Overall Structure of SEONet | 41 |
| 图4.3 | 测试图像及真值图 | 44 |
| Fig. 4.3 | Test Image and Ground Truth Map | 44 |
| 图4.4 | 测试图像1马尾藻提取定性结果 | 45 |
| Fig. 4.4 | Qualitative Results of Sargassum Extraction for Test Image 1 | 45 |
| 图4.5 | 测试图像2马尾藻提取定性结果 | 47 |
| Fig. 4.5 | Qualitative Results of Sargassum Extraction for Test Image 2 | 47 |

**表清单**

|  |  |  |
| --- | --- | --- |
| 表序号 | 表名称 | 页码 |
| 表3.1 | 高分一号卫星传感器技术指标 | 29 |
| Table 3.1 | Technical Parameters of GF-1 Satellite Sensors | 29 |
| 表3.2 | 实验环境详细信息 | 31 |
| Table 3.2 | Detailed Information of the Experimental Environment | 31 |
| 表3.3 | **马尾藻探测精度评估的混淆矩阵** | 31 |
| Table 3.3 | Confusion Matrix for Sargassum Detection Accuracy Assessment | 31 |
| 表3.4 | 测试图像1马尾藻提取定量性能对比结果 | 34 |
| Table 3.4 | Comparative Results of Quantitative Performance for Sargassum Extraction on Test Image 1 | 34 |
| 表3.5 | 测试图像2马尾藻提取定量性能对比结果 | 35 |
| Table 3.5 | Comparative Results of Quantitative Performance for Sargassum Extraction on Test Image 2 | 35 |
| 表3.6 | 马尾藻消融实验对比结果 | 36 |
| Table 3.6 | Ablation Experiment Comparison Results for Sargassum | 36 |
| 表4.1 | 测试图像1马尾藻提取定量性能对比结果 | 46 |
| Table 4.1 | Comparative Results of Quantitative Performance for Sargassum Extraction on Test Image 1 | 46 |
| 表4.2 | 测试图像2马尾藻提取定量性能对比结果 | 48 |
| Table 4.2 | Comparative Results of Quantitative Performance for Sargassum Extraction on Test Image 2 | 48 |
| 表4.3 | 马尾藻消融实验对比结果 | 48 |
| Table 4.3 | Ablation Experiment Comparison Results for Sargassum | 48 |

**变量注释表**

|  |  |  |
| --- | --- | --- |
| **变量** | **注释** | **初现页** |
|  | 归一化植被指数 | 9 |
| *NIR* | 近红外波段的反射率 | 9 |
| *RED* | 红光波段的反射率 | 9 |
|  | 差值植被指数 | 9 |
|  | 浮游藻类指数 | 9 |
| *Green* | 图像或数据中绿色通道的数值 | 9 |
| *Blue* | 图像或数据中蓝色通道的数值 | 9 |
|  | 归一化差值水体指数 | 9 |
|  | 权重矩阵 | 17 |
|  | 隐藏状态输入 | 17 |
|  | 当前状态输入 | 17 |
|  | 偏置向量 | 17 |
|  | 权重矩阵 | 17 |
|  | 偏置向量 | 17 |
|  | Sigmoid函数 | 17 |
|  | 激活函数 | 17 |
|  | 权重矩阵 | 17 |
|  | 偏置向量 | 17 |
|  | 遗忘门向量 | 17 |
|  | 前一个时间状态向量 | 17 |
|  | 当前时间输入门向量 | 17 |
|  | 当前时间候选状态向量 | 17 |
|  | 当前时间隐藏状态向量 | 17 |
|  | 权重矩阵 | 17 |
|  | 偏置向量 | 17 |
|  | 当前时间的输出门向量 | 17 |
|  | 重置门的权重矩阵 | 18 |
|  | 更新门的权重矩阵 | 18 |
|  | 当前时间的重置门向量 | 18 |
|  | 当前时间候选隐藏状态向量 | 18 |
|  | 输入特征 | 19 |
|  | 投影矩阵 | 19 |
|  | 投影矩阵 | 19 |
|  | 投影矩阵 | 19 |
|  | 查询矩阵 | 19 |
|  | 键矩阵 | 19 |
|  | 值矩阵 | 19 |

**变量注释表**

|  |  |  |
| --- | --- | --- |
| **变量** | **注释** | **初现页** |
| ϕ( ) | 核函数 | 20 |
|  | 查询向量的第i个元素 | 20 |
|  | 键向量的第i个元素 | 20 |
|  | 值向量的第i个元素 | 20 |
|  | 可学习张量 | 25 |
|  | 特征图 | 25 |
|  | 深度可分离卷积 | 25 |
| *LN* | 层归一化 | 26 |
| *Cat* | 通道维度的拼接 | 26 |
|  | 特征图 | 26 |
| *μ* | 特征图的均值 | 26 |
| *s* | 特征图的标准差 | 26 |
| *γ* | 可学习的缩放参数 | 26 |
| *β* | 可学习的偏移参数 | 26 |
| *ε* | 常数 | 26 |
|  | 每组特征图的缩放参数 | 27 |
| *C* | 通道数 | 27 |
|  | 特征图的重要性权重 | 27 |
|  | Sigmoid函数 | 27 |
|  | 门控机制 | 27 |
| *w1* | 有效特征权重 | 27 |
| *w2* | 冗余特征权重 | 27 |
| *p* | 可学习的参数 | 30 |
|  | 二元标签值 | 30 |
|  | 标签值的概率 | 30 |
|  | 样本的总数量 | 30 |
|  | 真正例 | 31 |
| *FP* | 假正例 | 31 |
|  | 假负例 | 31 |
|  | 精确率 | 31 |
|  | 召回率 | 31 |
|  | 交并比 | 32 |
|  | 注意力权重 | 42 |
|  | 线性映射 | 43 |
|  | 线性映射 | 43 |
|  | 输入特征 | 43 |
|  | 输入特征 | 43 |

**变量注释表**

|  |  |  |
| --- | --- | --- |
| **变量** | **注释** | **初现页** |
|  | 3×3卷积 | 43 |
|  | 重新校准注意力 | 43 |

# 1 绪 论

**1 Introduction**

1.1 研究背景与意义（Research Background and Significance）

马尾藻（Sargassum）是一种广泛分布于热带和亚热带海洋的浮游性褐藻，具有复杂的分支结构，以其漂浮特性而闻名，它的分布和生长常受到各种环境因素的多种影响，包括季风、洋流、温度和营养盐等，特别是在海水富营养化的条件下，马尾藻能够迅速繁殖并形成大规模堆积现象，这种现象被称为“黄金潮”[1-2]。漂浮的马尾藻群体，形态多样，通常呈现为零散的小斑块、大面积的带状或片状分布。

近年来，漂浮马尾藻的大规模爆发在全球范围内受到广泛关注[3]，其影响和后果是多方面的，具体如下：第一，破坏环境。马尾藻的过度繁殖和沉积破坏了海洋生态系统的平衡，影响了海洋生物的栖息地和食物链，导致生物多样性下降，马尾藻腐烂分解过程中会不断消耗大量溶解氧，导致水体缺氧，进而影响水质。第二，危害生态。马尾藻的堆积对珊瑚礁生态系统尤其有害，阻碍珊瑚的生长和恢复，影响珊瑚礁的健康，马尾藻覆盖水面会阻碍光线穿透，影响水下植物的光合作用，导致海洋生物死亡。第三，威胁人类健康。马尾藻分解过程中可能释放硫化氢等有害气体，对沿海居民的健康构成威胁，大量马尾藻堆积在海滩上，不仅会影响海滩的美观，还可能携带有害微生物和病原体，影响公共健康。另外，还会威胁着沿海景观、旅游业和水产养殖业，对当地经济产生严重负面影响[4]。

自2011年以来，马尾藻大量涌入加勒比海、墨西哥湾、巴西北部海岸和非洲西海岸，其生物量超过2000万公吨，对当地生态系统和经济造成了严重灾害 [5-6]。2016年12月，大量漂浮的马尾藻漂流至中国江苏沿海的紫菜养殖区，导致高达5亿元人民币的重大经济损失 [7]。2017年5月，浒苔和马尾藻在黄海和东海同时爆发，并在风和海流的作用下向北漂移，使得墨西哥湾和大西洋沿岸成为马尾藻的高发区域[8]。2018年，马尾藻在加勒比海岛屿的海岸线爆发，对当地居民、沿海海洋生态系统和经济部门，尤其是旅游业和渔业，带来了巨大危害和损失[9]。2019年至2025年间，马尾藻的爆发规模不断加剧，在 2025年甚至超过4500万公吨[10]，多个地区因马尾藻堆积导致海滩关闭，旅游收入大幅下降，据估算，仅加勒比地区因马尾藻爆发导致的经济损失累计就已超过数亿美元[12-16]，加之马尾藻腐烂不断释放的有害气体，致使严重危害了沿海居民的生活质量[11]。因此，马尾藻的爆发不仅破坏了海洋生态环境，还对旅游业、渔业和沿海居民的生活质量带来了巨大威胁，这就使及时、准确地监测和预警马尾藻的爆发，精准实施对马尾藻治理的有效措施，成为全球海洋生态保护和经济可持续发展的紧迫任务[3]。

近年来，大量漂浮马尾藻水华在全球许多海域频繁发生，卫星遥感因其在天气和频繁观测方面有着独特的优势，是提供更好的海藻信息的有效工具[17]。目前，粗分辨率卫星图像已被普遍成功地用于研究世界各地的马尾藻事件，并且在全球范围内的马尾藻事件研究中，取得了显著成果。然而，这些中分辨率图像在监测大型藻类爆发的早期阶段存在局限性。例如，在1公里的MODIS图像中，小规模的马尾藻斑块可能仅占据图像像素大小的0.2%，从而导致图像中部分像素 “丢失”[18]。此外，受陆地邻近效应、浅水区域以及高混浊度水域的干扰，对粗分辨率图像的数据质量产生了显著影响，尤其是在沿海水域，进而限制了中分辨率图像在精确监测马尾藻分布和动态变化中的应用[19]。

幸运的是，高分辨率卫星图像能够有效克服上述限制，它能够探测到小规模马尾藻的特征，从而更准确地评估其分布和丰富度。此外，在仅配备蓝、绿、红和近红外（NIR）四个波段的情况下，寻找一种有效的方法从高分辨率影像中提取马尾藻信息显得非常重要[20]。与此同时，在计算机视觉领域这一领域，深度学习技术获得了突破性进展，它不仅能够处理高分辨率影像，还能够提供更具丰富性的地物细节信息、几何结构和纹理特征，这种技术的进步为遥感影像的自动化处理提供了新的可能性。

综上所述，可以将高分辨率遥感影像与深度学习方法相结合，这样不仅可以提升数据处理能力，还能提高遥感影像分析的自动化水平和结果的准确性，这种结合对于多种遥感影像任务中能够展现出显著的优势 [21-22]。

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

卫星遥感是马尾藻检测的重要工具。众多研究者基于卫星遥感展开研究，主要方法有海洋生物光学遥感方法、植被指数法、多传感器遥感检测方法、机器学习方法及深度学习方法。

1. 海洋生物光学遥感方法

为了解海洋生态系统的健康状况和动态变化，通过结合遥感技术和实地采样分析，监测和评估海洋中的生物光学特性，可以更好地了解和管理海洋环境问题。Gernez等人[23]利用潮汐流和卫星测量技术监测叶绿素浓度和悬浮颗粒物，这些数据对于海水养殖至关重要；Zhang等人[24]发现中国近海马尾藻色素含量存在显著差异，并指出特定海域的红藻叶绿素a含量是其他海域的22.6倍；Terasaki等人[25]研究了日本沿海水域马尾藻的岩藻黄素含量的季节性变化；Endo等人[26]分析了裙带菜在不同光照和营养条件下的色素变化，发现在特定环境条件下岩藻黄素和叶绿素a含量的变化趋势。

1. 植被指数法

针对水体中藻类水华的监测和评估，Hu等人[27]提出的浮游藻类指数（FAI）主要用于浮藻藻类提取，能有效区分蓝藻和水生植物，对环境条件改变的敏感性低，结果稳定，并且并且能够对复杂大气环境加以修正，削弱大气、薄云等要素的部分干扰；Wang等人[28]利用改进藻类指数（AFAI）提高了FAI，但它对云不太敏感；Guo等人[29]通过实验证明，差值植被指数（DVI）在减少太阳闪烁和薄云的影响方面表现良好，但会受到边界效应的影响，造成误判；于杰等人[30]根据分析马尾藻的图像和波谱特征，通过实验证明,与DVI法相比，归一化植被指数（NDVI）法不仅能有效减弱大气和海流等噪声的影响，还可以强化藻类水体的边界细节信息，从而提升马尾藻的识别效果；Zhang等人[31]基于绿藻光谱特征分析，利用大气校正后绿带基线高度构建了绿藻指数（GAI），成功实现了黄海和东海浒苔与马尾藻的区分；为进一步优化识别效果，Min等人[32]在此基础上提出了红绿指数（SRG）的斜率特征提取方法，该斜率是根据瑞利校正后的绿色与红色波段反射率斜率计算得出，用以区分浒苔和马尾藻，实验证明GAI和SRG仅适用于黄海近海区域；Mei等人[33]利用多光谱卫星传感器的瑞利校正反射率调谐的区分方法区分不同海水环境条件下浮游浒苔和马尾藻，这种方法比GAI和SRG方法具有更好的识别效果，对海水环境条件不敏感，且不需调整阈值，能够有效提高处理速度。

1. 多传感器遥感监测方法

为监测和评估马尾藻水华，结合不同分辨率和不同传感器的卫星数据。Wang等人 [34]利用中分辨率成像光谱仪图像，根据颜色指数（FUI）评估全球内陆水体的营养状态；zhou等人[35]将FUI应用于Sentinel-2图像来评估武汉水域的营养状态，并发现FUI具有良好的相关性；Kim等人 [36]从2008年到2017年利用地球静止海洋彩色成像仪(GOCI)数据监测了黄海的绿潮和金潮区域；Gower等人[37]利用多边气候影响指标，使用中分辨率成像光谱仪(MERIS和MODIS)提供的卫星图像，对墨西哥湾的马尾藻聚集体进行了探测；Sun等人[38]的研究通过结合MODIS、VIIRS和OLCI的多传感器观测数据，实现了对加勒比海和中大西洋马尾藻分布的连续监测；Wang等人[39]的研究表明，利用陆地卫星OLI和Sentinel-2 MSI传感器的高分辨率数据能够显著增强马尾藻监测的效果；Mei等人[33]从不同的卫星传感器角度分析了浒苔和马尾藻之间的光谱差异，对比寻找合适的卫星传感器来区分浒苔和马尾藻，实验表明GF1-WFV具有更好的区分浒苔和马尾藻的能力。这些研究不仅展示了多传感器遥感监测方法在马尾藻水华监测中的应用，还能够提供更全面和准确的马尾藻水华分布和覆盖范围信息，有助于更好地了解和管理海洋环境问题。

1. 机器学习方法

机器学习在遥感图像处理和大型藻类分类中具有显著的应用潜力。Li等人[40]运用了SHAP，这是一种基于博弈论的解释模型，用于量化每个特征对模型预测的贡献度，同时，他们还利用Bhattacharyya距离来衡量不同类别间的概率分布差异，从而提高了对大型藻类分类的准确性；Chapelle等人[41]开发的支持向量机（SVM）是一种依据统计学习理论的分类技术，它是借助结构风险最小化来实现优化模型；Chen等人[42]利用高分辨率卫星数据，通过分析灰度共生矩阵的四个测度和光谱采样，研究了马尾藻的光谱和纹理特征，结果表明，SVM能够有效利用纹理特征，从而提高马尾藻提取的精度；由Breiman和Cutler提出的随机森林是一种集成学习方法，其通过构建多个决策树来提高分类的准确性和鲁棒性[43-48]，因其高分类精度、快速的处理速度以及对过拟合的强抵抗力，在遥感领域得到了广泛应用[49-53]；Cuevas等人[54]采用随机森林算法来检测尤卡坦半岛东北部地区的马尾藻；Xiao等人[55]也开发了基于随机森林的算法，利用GF-1号卫星图像的光谱差异来区分S.horneri和U. prolifera。

1. 深度学习方法

在海洋遥感领域，深度学习技术在遥感图像处理中具有多样性和有效性，通过结合不同的深度学习架构和算法，可检测和分类如马尾藻等大型藻类。Arellano-Verdejo等人[56]提出了第一个专用于墨西哥加勒比海岸马尾藻检测的深度学习方法ErisNet，即具有卷积和递归层的像素分类器，应用于MODIS光谱带的选择，ErisNet很少考虑相邻像素，因此它不学习由几个像素组成的海藻集合的结构；Wang和Hu [57]使用UNet模型[58]，通过高分辨率卫星数据对小安的列斯群岛和墨西哥湾近海水域的马尾藻进行分割；Yan等人[59]还将UNet应用于中国巢湖的sentinel-2 MSI图像，以识别有害的蓝藻水华，他们将分割与三种基于FAI的自动蓝藻水华提取方法进行了比较:梯度模式、固定阈值和Otsu方法，并显示UNet的准确性更好；Laval等人[60]开发了两种新的深度学习模型，专门用于基于编码器-解码器卷积神经网络(CNN)的海藻检测，一个调谐到Sentinel-2卫星上的多光谱仪器的光谱波段，另一个调谐到Sentinel-3卫星上的海洋和陆地颜色仪器(OLCI)，在使用相同训练从卫星图像检测马尾藻的过程中，这种特定的新方法与ErisNet不同的是CNN考虑了相邻像素，并且比UNet和SegNet具有更少的减少级别，允许丝状物体(如马尾藻聚集)被检测到；Liu等人[61]应用反馈注意力网络（FANet）来掩蔽MODIS影像上的云和云影像素，减少误检并保留更多的有效卫星测量数据。

1.3 问题的提出（**Statement of problem**）

目前，对于马尾藻检测和提取普遍使用中分辨率遥感影像，虽然能够提供一定程度的监测覆盖，但其较低的空间分辨率无法精确捕捉马尾藻的细小结构和边缘特征，导致难以区分马尾藻与其他相似水体成分的光谱信号，同时在监测小规模或局部马尾藻分布时不够精确，影响了对马尾藻水华动态变化的快速响应和详细分析能力。然而，高分辨率遥感影像能够很好的改善上述问题，GF1-WFV遥感影像因具有高分辨率多光谱成像能力、大幅宽覆盖、高时间分辨率和稳定的数据产品等特点，使其在遥感监测领域具有广泛的应用潜力，因此，本文考虑使用GF1-WFV遥感影像。本研究针对马尾藻检测和提取的涉及科学问题如下：

（1）马尾藻因受洋流、季风等自然因素影响，生长形态多变，分布尺度各异。在马尾藻遥感图像中，呈现出个体大小、形状和聚集模式差别较大的特点。以往的方法未能充分利用空间上下文信息，对不同形态或尺度的物体适应性差、泛化能力不强，难以实现对马尾藻精准提取。

（2）由于水体中的悬浮颗粒物、浮游生物以及其他溶解物质会改变水体的光学特性，使马尾藻边缘区域与周围水体的光谱反射特性相似，另外，马尾藻边缘的生物量密度通常是逐渐变化的，形成与周围水体渐变过渡，无清晰界限，导致在遥感影像中马尾藻边缘和背景水体难以区分。加之在海洋洋流和波浪的作用下马尾藻的形态和分布发生变化，使边缘区域变得更加分散和不规则，导致边缘区域较多漏提或错提。

目前为止，还没有形成一套完善的基于深度学习方法的马尾藻信息提取与检测处理流程，但深度学习在处理海量带标签遥感图像方面具有一定优势。

1.4 主要研究内容（Main research contents）

本文针对马尾藻形态多变、尺度各异特点，结合边缘区域生物密度低易于周围水体像素混淆问题，采用GF1-WFV与深度学习相结合的研究方法，主要内容如下：

针对马尾藻形态多样，尺度各异，小尺度马尾藻因其细节信息不足而难以精确提取。提出了一个基于多维重组特征的多尺度上下文提取网络MRF-Net (Context Extraction Network Based on Multidimensional Reorganized Features)，该网络先是引入了多维重组特征融合模块，通过在网络的不同层之间进行通道、高度和宽度的多维特征重组，捕捉和融合来自不同尺度的特征，增强模型对马尾藻不同大小和形状的识别能力，其中利用Hadamard乘积和门控机制进行细粒度的空间重组，突出信息丰富的特征并抑制冗余特征，从而提高对马尾藻的精准识别；随后引入多尺度上下文提取模块，它是利用不同扩张率的卷积核和不同尺寸的池化操作，来获取全局和局部特征，从而捕获不同尺度的上下文信息；最后，通过添加卷积注意力机制以优化多尺度特征的融合。实验结果表明，MRF-Net不仅较为准确度的识别马尾藻目标，而且对于处理多形态和尺度的马尾藻具有良好鲁棒性。

针对马尾藻边缘区域生物量密度较低易于周围水体像素混淆，难以精确提取。提出了一个基于多层注意力和顺序聚合的马尾藻边缘优化网络（Sargassum Edge Optimization Network Based on Multi - Layer Attention and Sequential Aggregation, SEO-Net），SIA-Net的基础网络采用ResNet结构，通过ResNet Block提取特征；然后引入多层注意力模块，通过多个注意力机制组件进一步提取特征，其核心是核注意力和通道注意力，主要用于强化捕捉不同尺度的特征并理解这些特征在图像上下文中的关系；最后引入顺序信息聚合模块，来增强目标边界的细节表现，它是使用选择性边界聚合策略，有选择性地聚合边界信息和语义信息，其中，重新校准注意力单元负责选择性地聚合来自低层（细节丰富）和高层（语义丰富）的特征。实验结果表明，SIA-Net能够**实现复杂边界的精细化提取**，生成更加准确的边界信息，有效提高了马尾藻边缘区域的识别精度。

1.5 章节安排（Chapter arrangement）

本文研究内容设计主要分为五个章节，具体如下：

第一章：绪论。本章主要描述了本研究背景与意义，介绍了马尾藻国内外研究现状，目前马尾藻检测存在的问题，论文主要研究内容和各个章节安排。

第二章：相关技术。本章主要介绍马尾藻检测方法中用到的相关技术，依次包括传统植被指数法、语义分割简介、经典的语义分割网络、门控机制简介、注意力机制简介。本文第三章和第四章的内容是在第二章的基础上展开的。

第三章：基于多维重组特征的多尺度上下文马尾藻提取网络。本章首先介绍了基于多维重组特征的多尺度上下文马尾藻提取网络MRF-Net的主要流程。然后介绍了本章提出的MRF-Net网络模型结构设计。最后介绍了本文实验的研究区域、研究数据、实验配置和评价指标，并通过一系列实验验证了 MRF-Net的有效性：通过定性和定量实验对比马尾藻检测深度学习方法，验证了本文提出的MRF-Net网络在捕获多尺度特征上的优越性，同时通过消融实验验证MRF-Net各模块的有效性，并用可视化结果对网络模型进行有效性分析。

第四章：基于多层注意力和顺序聚合的马尾藻边缘优化网络（SEO-Net），旨在提高对马尾藻边缘区域的识别能力。本章首先介绍了基于多层注意力和顺序聚合的马尾藻边缘优化网络SEO -Net的网络模型整体架构的设计。然后阐述本章相关实验设计。最后对实验结果进行分析，通过定性和定量实验对比马尾藻检测深度学习方法，验证了SEO -Net在边缘提取任务中的优越性，同时通过消融实验验证SEO -Net各模块的有效性。

第五章：总结与展望。本章总结了本文研究工作，并且针对当前研究的不足之处，提出了未来的研究方向。

2 相关技术

2 Related technologies

本章主要讲述此次研究所涉及到的相关技术，主要包括：植被指数法、语义分割、门控机制以及注意力机制等，具体如下。

2.1 植被指数法（**Vegetation Index Method**）

植被指数法是一种在遥感图像分析中广泛应用的技术，它通过计算特定波段的反射率比值或组合来评估植被覆盖情况和生物量，能较好地反映植被覆盖度和生长状况的差异，在遥感监测中一直发挥着重要作用。在海洋生态检测中，尤其是马尾藻，植被指数法可以帮助识别和区分不同类型和浓度的浮游植物。常见的植被指数包括：归一化植被指数（*NDVI*）（见公式2.1）、差值植被指数（*DVI*）（见公式2.2）、浮游藻类指数（*FAI*）（见公式2.3）和归一化差值水体指数（*NDWI*）（见公式2.4）。

（2.1）

*NDVI*可以反映植被的健康状况，其取值范围通常在-1到1之间。其中，正值意味着植被覆盖[62]，*NIR*代表近红外波段的反射率，而*RED*则是红光波段的反射率。

（2.2）

*DVI*是通过计算近红外波段和红光波段的反射率差值来突出植被的存在，对于水体中浮游植物的检测特别有用[63]。

（2.3）

*FAI*专门用于检测水体中的浮游藻类，通过分析绿光和蓝光波段的反射率差异来区分藻类和水生植物[64]。其中，*Green*和*Blue* 分别表示图像或数据中绿色、蓝色通道的数值，通常来源于*RGB*颜色模型（红、绿、蓝三原色）。

（2.4）

*NDWI*用于检测水体中的悬浮物质，如泥沙和浮游生物，对于分析水体浑浊度和藻类水华有重要意义[65]。

在马尾藻检测中，借助植被指数法可以帮助识别马尾藻的分布和浓度。例如，*NDVI*和*DVI*可以用于区分马尾藻和周围水体，而*FAI*则可以用于识别浮游马尾藻，它们通过突出马尾藻的光谱特征，使得能够从遥感图像中自动提取马尾藻 [66-67]。

尽管植被指数法在马尾藻检测中具有十分重要的价值，但也存在一定的局限性。例如，它们可能受到大气条件、水体特性和传感器特性的影响，导致检测结果存在不确定性。此外，单一指数可能无法完全区分马尾藻和其他水体成分，需要结合多种指数方法来提高检测的准确性[68-69]。因此，在本文研究中，我们提出了基于深度学习的新型网络结构，以提高马尾藻检测的准确性和鲁棒性。

2.2 语义分割（**Semantic Segmentation**）

语义分割属于计算机视觉领域的一种图像分割技术，其目标是通过将图像中的每个像素分配到其对应的语义类别中，以识别出图像中的不同对象和区域，它在许多领域发挥着极其重要作用[70]。图像语义分割技术通常分为传统方法与深度学习方法两类。传统方法是通过物体的颜色、纹理、形状及空间位置关系等特征进行提取，再运用聚类、分类和阈值算法等手段实现图像分割[71-72]。然而，这些方法对人工设计特征的依赖性较高，并逐渐暴露出一些局限性。随后，基于深度学习的方法被视为应对图像语义分割难题的极具潜力的途径[73-75]，在自动驾驶、目标检测、医学图像分析、遥感图像分析等多个计算机视觉任务中，该技术展现出广泛的应用价值。

语义分割架构通常可以被视为是由一个编码器-解码器构成，它首先是利用预训练的分类网络作为编码器来捕获输入图像的关键特征，接着解码器通过逐步的特征重构过程，将编码器提取的特征进行放大和细化，最后生成高分辨率的分割结果。这种结构使得语义分割模型不仅能够在像素级别上具备判别能力，还能够将编码器学习到的可判别性特征有效投影到像素空间，从而得到更准确的分割结果。

2.3 卷积神经网络（**Convolutional Neural Network**）

卷积神经网络（convolutional neural network, CNN）是深度前馈神经网络的一种，其核心组成部分是卷积计算，它的网络结构具有多层特性，从而在深度学习领域发挥着极其重要的作用。其主要结构如图2.1所示。



图2.1 卷积神经网络结构图

**Figure 2.1** The schematic diagram **of** CNN structure

1）输入层

卷积神经网络的输入层是接收原始图像数据，这些数据被组织成一个三维数组的形式，其中两个维度分别对应图像的垂直和水平像素，第三个维度则代表颜色通道，例如RGB。

（2）卷积层

卷积层（Convolutions Layer）由多个卷积核（也称为滤波器）构成，主要用于处理来自输入层的三维矩阵数据[76-77]。多个卷积核在输入图像上滑动并计算点积，生成多个特征图，每个特征图捕捉输入数据中的特定特征，最终为网络提供丰富的局部特征信息。

（3）池化层

池化层在卷积神经网络中的关键作用在于能够降低特征图空间维度，通过减少参数数量和计算量，进而简化模型，其主要使用最大池化和平均池化两种方式实现，最大池化是选取覆盖区域内的最大值来保留最显著的特征，而平均池化则是依赖计算区域内的平均值来捕捉整体趋势。

（4）激活函数层

激活函数是神经网络中的非线性映射层，通过添加非线性映射到网络的输出，提升模型对不同数据模式的识别和适应能力。常用的激活函数包括 Sigmoid、ReLU和 Leaky ReLU[78]。

（5）批归一化层

批量归一化层（Batch Normalization，简称BN）是深度学习中常用的一种技术，它主要用于加速深层神经网络的训练过程，并提高模型的稳定性。BN层通过对每一批数据进行归一化处理，使得网络中间层的输入数据分布更加稳定，从而有助于网络更快地学习到有效特征。

2.4 典型语义分割模型（**Typical Semantic Segmentation Models**）

常用的卷积神经网络（CNN）语义分割网络方法主要分为模型训练和模型测试两个阶段。其流程大致如图2.2所示。

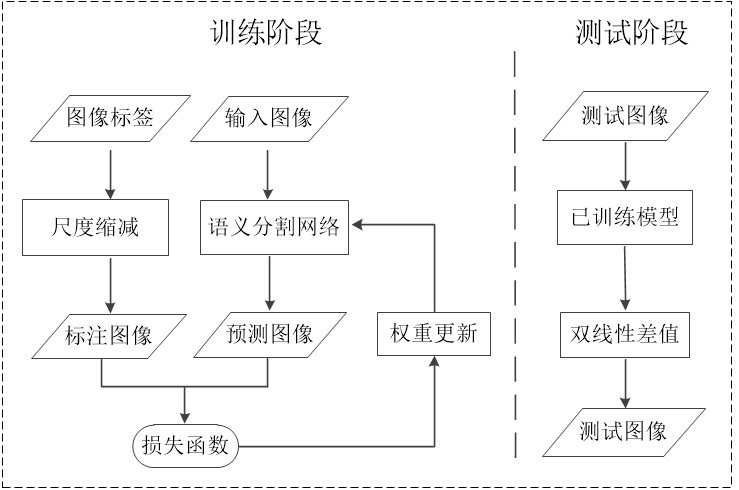


图2.2 语义分割网络流程图

**Figure 2.2** The schematic diagram **of Semantic Segmentation Network**

2.4.1 FCN

全卷积网络（Fully Convolutional Networks，FCN）是用于图像语义分割的一种框架，由Jonathan Long等人在2015年提出的，被视为深度学习在语义分割领域的开创性工作[79-80]。FCN是把传统CNN后面的全连接层替换成卷积层，这样就使得网络的输出是热力图而不是类别。同时，为解决卷积和池化导致图像尺寸变小问题，使用上采样方式对图像尺寸进行恢复。

FCN是首个针对像素级预测端对端的全卷积网络，主要由全卷积和反卷积这两部分组成。全卷积部分通常使用经典的CNN架构，如VGG或ResNet，来提取图像的丰富特征。反卷积部分则是通过上采样得到原尺寸的语义分割图像。上采样部分是使用双线性插值或反卷积（转置卷积）等技术。为了提高分割精度，FCN还采用了跳级结构（skip connections），这种结构通过不同倍数的上采样来融合编码器中不同层次的特征图，比如FCN-32s、FCN-16s和FCN-8s，这样既保留了全局信息又保留了局部细节，从而实现精确的像素级分类。具体结构如图2.3所示。

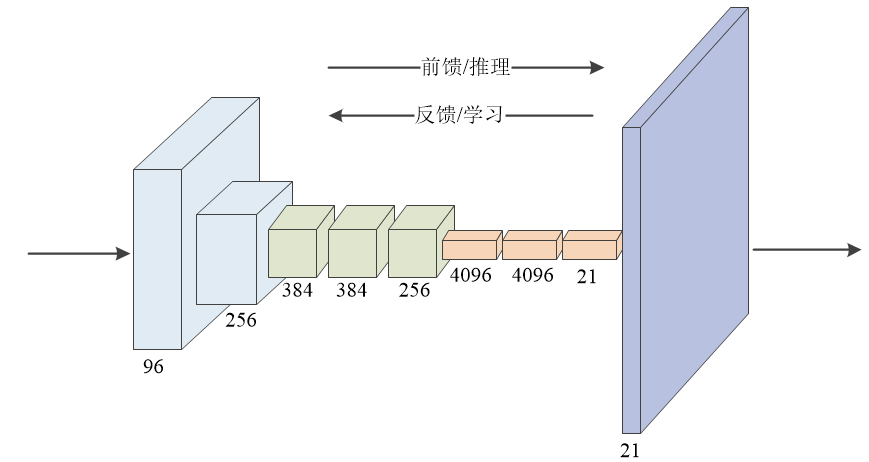


图2.3 FCN网络结构图

**Figure 2.3** The schematic diagram of the **FCN** structure

2.4.2 UNet

UNet是一种用于图像分割的[卷积神经网络](https://zhida.zhihu.com/search?content_id=234974583&content_type=Article&match_order=1&q=%E5%8D%B7%E7%A7%AF%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C&zhida_source=entity)（Convolutional Neural Network，CNN）架构，特别适用于医学图像分割任务。UNet的结构主要由编码器（Encoder）和解码器（Decoder）两部分构成，因其呈现出一条对称的形似英文字母 U的路径而得名，其左边是收缩路径，右边是扩张路径[81]，通过中间的跳跃连接相连，具体结构如图2.4所示。UNet的关键步骤是融合了底层信息的位置信息与深层特征的语义信息，使其在图像分割领域能够表现出卓越的性能。同时，这种设计不仅赋予了网络对输入图像尺寸变化的适应能力，还增强了模型的鲁棒性。



图2.4 UNet网络结构图

**Figure 2.4** The schematic diagram of the U-Net structure

（1）编码器（Encoder）：

编码器部分类似于传统的卷积神经网络，主要用于提取图像的特征。它通过多层卷积操作来逐步减少图像的尺寸，保留更深层次的特征信息。每一层都会进行卷积、激活（ReLU）以及最大池化（Max Pooling），从而提取重要的特征。

（2）跳跃连接（Skip Connections）：

跳跃连接作为编码器与解码器之间的纽带，它负责确保解码器在进行上采样时能够利用编码器中的特征，避免细节信息丢失，从而提升语义分割的精确度和细节表现。

（3）解码器（Decoder）：

解码器部分由多个反卷积块构成，其作用是将编码器提取的特征逐步恢复为与输入图像相同大小的分割结果。这个过程使用了上采样（Upsampling）技术，并通过跳跃连接将编码器中相应层的特征拼接到解码器中，保留了更多的细节信息。

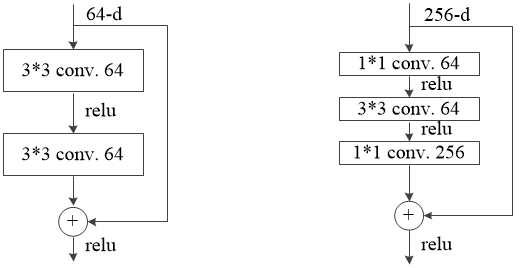
（4）输出层：

输出层是利用一个1x1卷积层，将解码器的输出映射为最终的分割结果，其输出通道数与分割任务中的类别数量相匹配，目的是确保每个像素都能被准确地分类。

2.4.3 ResNet

[ResNet](https://so.csdn.net/so/search?q=ResNet&spm=1001.2101.3001.7020)（Residual Neural Network）由微软研究院的Kaiming He等四名华人提出，通过使用ResNet Unit成功训练出了152层的神经网络，并在ImageNet大规模图像识别竞赛中取得冠军 [82]。为了增强CNN的表征能力，需要不断增加网络的层数，但随着网络深度的增加，出现两个问题：一是梯度消失或梯度爆炸。二是退化问题。由此，ResNet网络增加跳跃连接方式连接一系列残差块，使网络轻松增加了深度且易收敛，打破了以往盲目增加网络深度的操作，具有划时代的意义。常见的残差单元有两种，ResNet中的残差单元主要有两种形式，具体结构如图2.5所示。左边的残差结构是针对层数较少网络，例如ResNet18层和ResNet34层网络，右边是针对网络层数较多的网络，例如ResNet101，ResNet152等。

[ResNet](https://so.csdn.net/so/search?q=ResNet&spm=1001.2101.3001.7020)的相关模型通常都是由各种类型的残差单元组成，其主要优势在于能够缓解梯度消失或者梯度爆炸问题。残差单元通过引入跳跃连接，在进行反向传播时，梯度通过使用恒等函数让其从后继层直接流向先前层，从而缓解了梯度消失或爆炸的问题。



（a）浅层残差示意 图 （b）深层残差示意图

图2.5 残差单元结构图

**Figure2.5** The schematic diagram of **of the Residual Unit** structure

2.5 门控机制（**Gating Mechanism**）

门控机制的灵感来源于[生物神经科学](https://zhida.zhihu.com/search?content_id=241378298&content_type=Article&match_order=1&q=%E7%94%9F%E7%89%A9%E7%A5%9E%E7%BB%8F%E7%A7%91%E5%AD%A6&zhida_source=entity)中对大脑信息处理机制的研究，特别是对记忆和学习过程的研究。门控机制模型主要通过遗忘门、输入门和输出门，灵活控制信息在神经网络中的流动，确保模型能够有效记住重要信息并过滤掉无关信息，从而在处理长序列数据时表现更加稳定和高效。

2.5.1 长短期记忆网络LSTM（Long Short-Term Memory）

传统的[循环神经网络](https://zhida.zhihu.com/search?content_id=241378298&content_type=Article&match_order=1&q=%E5%BE%AA%E7%8E%AF%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C&zhida_source=entity)（RNN）在处理长序列时存在梯度消失或梯度爆炸的问题，导致它们难以学习长期依赖信息。为了解决这个问题，研究者们开始探索新的网络结构。1997年，Hochreiter和Schmidhuber等人[83]提出了第一个成功处理长期依赖问题的神经网络架构长短期记忆LSTM（Long Short-Term Memory），它主要有三个门：遗忘门（Forget Gate）、输入门（Input Gate）和输出门（Output Gate）。具体如图**2.6**所示，它能有效地控制信息的流动，从而能够学习长期依赖关系。

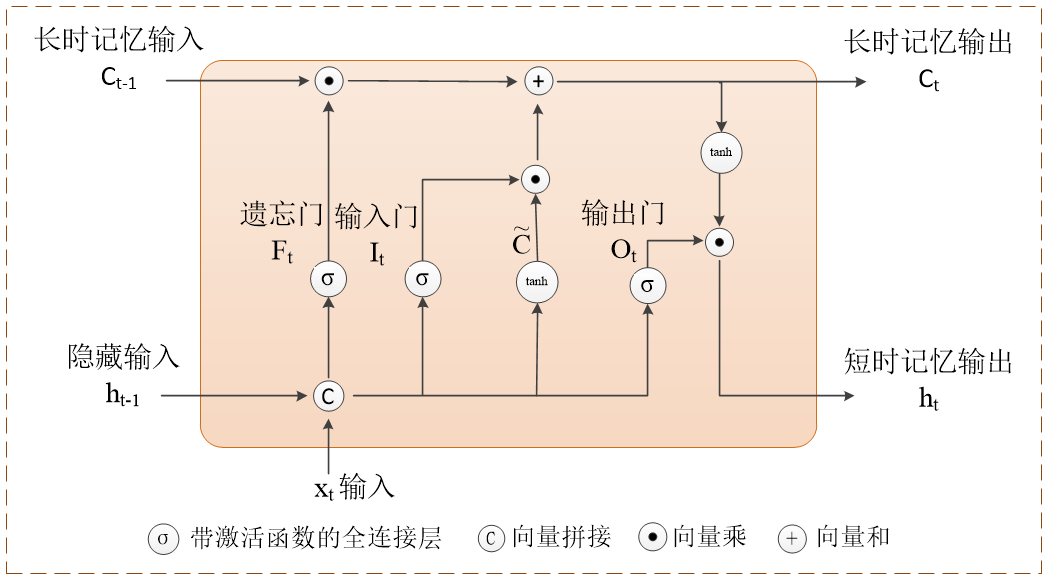


图2.6 LSTM结构图

**Figure 2.6** The schematic diagram of the **LSTM** structure

遗忘门的作用是决定哪些信息应该从单元状态中丢弃或保留。它接收当前输入和上一时刻的隐藏状态作为输入，通过一个Sigmoid激活函数将其映射到0到1之间的值。其中，接近0的值表示对应的信息将被遗忘，接近1的值表示信息将被保留。遗忘门的计算公式如（2.5）所示，其中， 和 分别是遗忘门的权重矩阵和偏置向量。

（2.5）

输入门的作用是决定哪些新的信息将被存储在单元状态中。它同样接收和作为输入，通过Sigmoid函数计算出一个更新比例，同时通过一个*tanh*激活函数对当前输入进行变换，然后将两者相乘得到需要更新的信息。输入门的计算公式如（2.6），当前状态变换公式如（2.7）所示，其中，和是偏置向量。

（2.6）

（2.7）

根据遗忘门和输入门的结果，对信息进行更新。具体公式如下（2.8）所示，其中，表示逐元素相乘。

（2.8）

输出门的作用是决定下一个隐藏状态是什么。它同样接收和作为输入，通过Sigmoid函数计算出一个输出比例，然后将其与经过*tanh*激活函数处理后的信息相乘，得到当前时刻的隐藏状态。输出门的计算公式如（2.9）和（2.10）所示，其中，和分别是输出门的权重矩阵和偏置向量。

（2.9）

（2.10）

2.5.2 门控循环单元GRU（Gated Recurrent Unit）

为简化LSTM的结构并提高计算效率，Cho等人[84]于2014年提出了门控循环单元GRU（Gated Recurrent Unit），它是将LSTM中的三个门简化为两个门，分别是重置门和更新门，重置门决定多大程度上将旧状态信息“重置”，有助于捕获序列中的短期依赖关系，更新门决定多大程度上将旧状态信息传递到新状态，有助于捕获序列中的长期依赖关系。具体如图**2.7**所示。

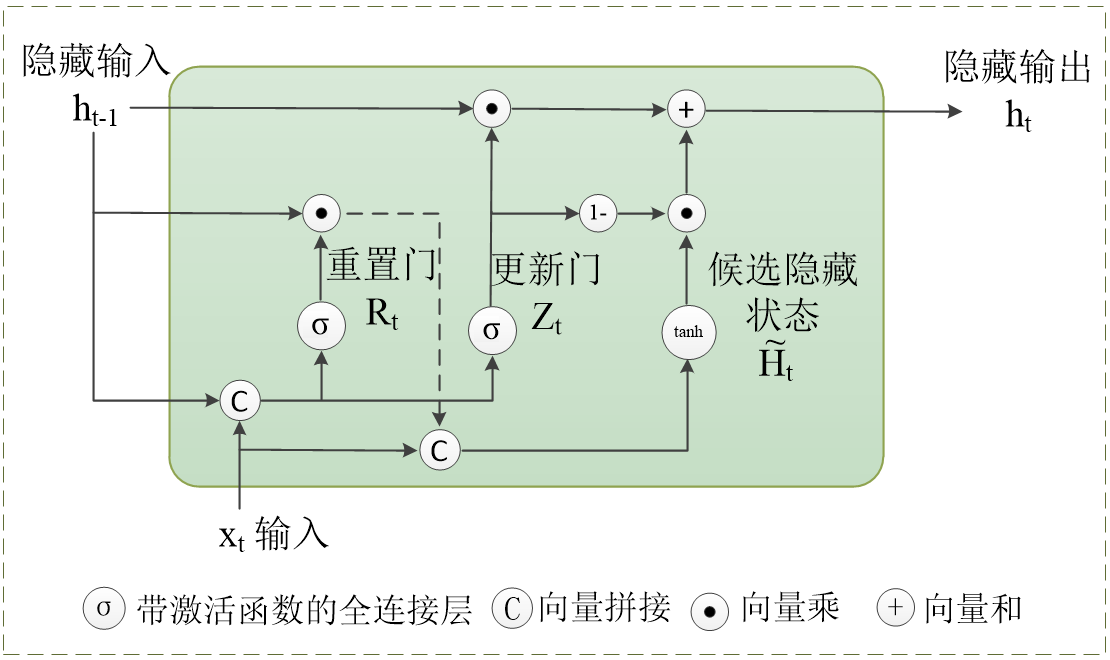


图2.7 GRU结构图

**Figure 2.7** The schematic diagram of **GRU Structure**

重置门的作用是决定前一个时间步的隐藏状态对当前时间步的影响程度。其计算公式如（2.11）所示，其中，是重置门的权重矩阵。当接近0时，过去的隐藏状态将被更多地遗忘，从而允许模型更容易地捕捉到新的输入信息。

（2.11）

更新门的作用是决定有多少过去的信息需要保留到当前时刻，以及有多少当前的输入信息需要被整合到新的隐藏状态中。其计算公式如（2.12）所示，其中，是更新门的权重矩阵，是sigmoid函数，其输出值在0到1之间。当 接近1时，表示更多地保留过去的隐藏状态；当  接近0时，表示更多地使用当前的输入来更新隐藏状态。

（2.12）

基于更新门和重置门，GRU的隐藏状态更新公式如（2.13）和（2.14）所示，其中，是候选隐藏状态，它是通过将重置门作用于上一个隐藏状态,并与当前输入 一起经过一个*tanh*激活函数得到的。最终的隐藏状态是通过更新门对上一个隐藏状态和候选隐藏状态进行线性组合得到的。

（2.13）

（2.14）

2.6 注意力机制（**Attention Mechanism**）

注意力机制（Attention Mechanism）是深度学习中的一种数据处理方法，广泛应用在自然语言处理、图像识别以及语音识别等各种不同类型的机器学习任务中。其核心思想是让模型能够有选择性地关注输入序列中的不同部分，为输入序列的各个部分分配不同的权重，以此来突出对任务更关键的信息。注意力机制可以视为查询矩阵(Query)、键(key)以及加权平均值构成了多层感知机(Multilayer Perceptron, MLP)，其主要有点积注意力、缩放点积注意力、多头注意力、自注意力和双向注意力等[85]。

2.6.1 点积注意力极其泛化

点积注意力是一种注意力机制，用于计算输入向量之间的相似度，它是通过计算两个向量的点积来度量它们的相似程度，然后对结果进行归一化处理得到注意力权重。

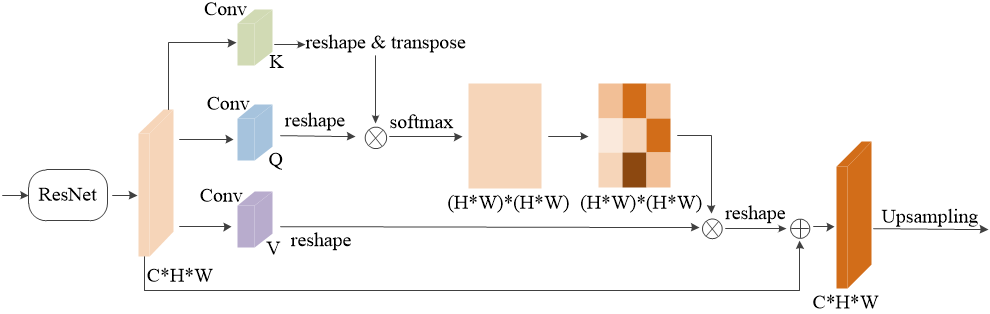


图2.8 点积注意力结构图

**Figure 2.8** The structure diagram of **Dot-Product Attention Mechanism**

如图2.8中所示，它给出了点积注意机制的架构的图示，点积注意力是从由ResNet主干生成的特征映射中捕获长范围上下文信息，并通过跳过连接将细化的特征与原始输入相添加，具体而言，点积注意力机制通过计算查询矩阵（Q）、键矩阵（K）和值矩阵（V）之间的相似性，捕捉特征之间的长距离依赖关系。其核心公式如（2.15）所示，其中，*X* 为输入特征，*Wq、Wk、Wv*为投影矩阵。

（2.15）

点积注意力的输出在通过softmax函数归一化后进行加权求和，如公式（2.16）所示，虽然点积注意力机制在捕捉长距离依赖关系方面表现出色，但是其计算复杂度为 O(N2)，在大规模输入应用中常存在一定的局限性。

（2.16）

为了降低计算复杂度，进一步提出了一种基于核的点积注意力泛化方法。通过引入核函数 ϕ(⋅)，将点积注意力机制泛化为公式（2.17）所示，其中，*ϕ(⋅)*为核函数。

（2.17）

本文选择softplus函数作为核函数，如公式（2.18）所示。通过核函数的引入，点积注意力的计算复杂度降低为 O(N)，显著提升了计算效率。

*softplus(x)=log(1+ex)*  （2.18）

2.6.2 通道注意力

通道注意力（Channel Attention） 是众多注意力机制中的一种，旨在为特征图中的不同通道分配不同的权重，从而增强对任务更有用的通道，抑制不重要的通道。通道注意力的核心思想是通过学习每个通道的重要性，动态调整特征图中各通道的贡献[93]。

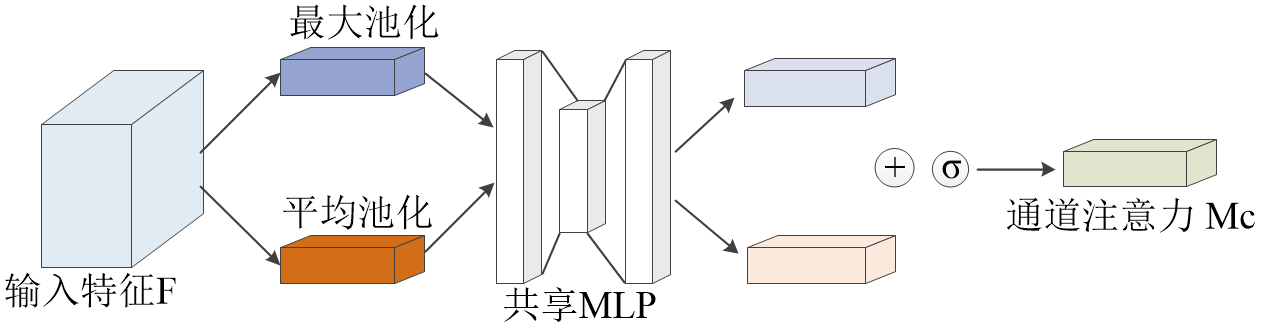


图2.9 通道注意力结构图

**Figure 2.9** The structure diagram of **Channel t Attention Mechanism**

如图2.9所示，通道注意力机制的流程是：首先，对一个尺寸为H×W×C的输入特征图F进行空间维度的全局最大池化和全局平均池化，得到两个1×1×C的特征图；然后，然后，将全局最大池化和全局平均池化的结果，分别送入一个共享的多层感知机（MLP）中学习，得到两个1×1×C的特征图。MLP的第一层神经元个数为 C/r（r为减少率），激活函数为 Relu，第二层神经元个数为C；最后，将MLP输出的结果进行Add操作，接着经过Sigmoid激活函数的映射处理，最终得到通道注意力权重矩阵Mc。

2.6.3 空间注意力

空间注意力机制（Spatial Attention Mechanism）也是众多注意力机制中的一种，它的核心思想是让模型能够聚焦于输入图像或特征图中的关键空间区域，然后为不同区域分配不同的权重，从而增强模型对重要信息的处理能力。

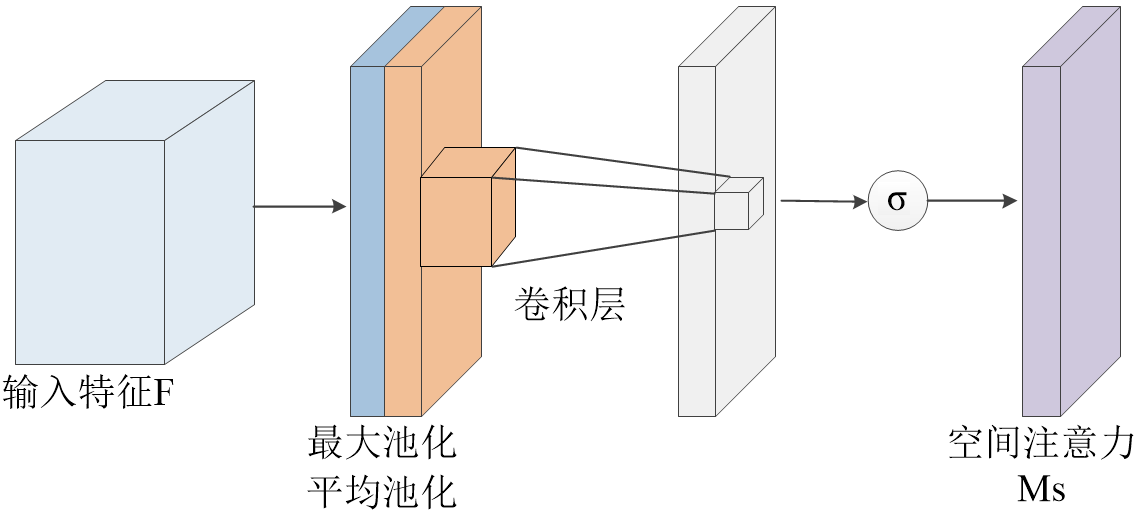


图2.10 空间注意力结构图

Figure 2.10 The structure diagram of Spatial **Attention Mechanism**

如图2.10所示，空间注意力的流程是：首先，对一个尺寸为H×W×C的输入特征图F进行通道维度的全局最大池化和全局平均池化，得到两个H×W×1的特征图；然后，将全局最大池化和全局平均池化的结果，按照通道拼接(concat)，得到特征图尺寸为H×W×2；最后，对拼接的结果进行7×7的卷积操作，得到特征图尺寸为H×W×2，接着通过Sigmoid激活函数 ，得到空间注意力权重矩阵Ms。

2.6.4 CBAM 注意力

CBAM（Convolutional Block Attention Module）是一种轻量级的注意力机制模块，主要作用是通过动态调整特征图的通道和空间权重，增强卷积神经网络（CNN）对关键信息的捕捉能力[94]。

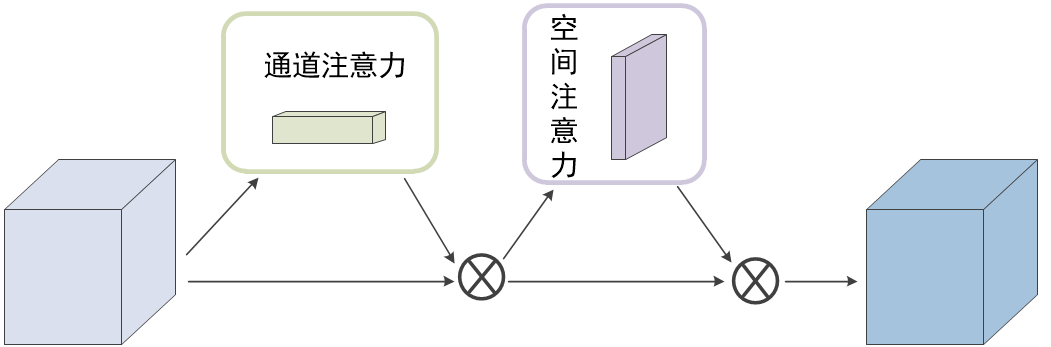


图2.11 CBAM结构图

**Figure 2.11** The schematic diagram of **CBAM Structure**

CBAM的具体结构如图2.11所示，CBAM是由通道注意力模块（Channel Attention Module, CAM）和空间注意力模块（Spatial Attention Module, SAM）串联构成，其中，通道注意力优先处理，它是通过全局平均池化和最大池化聚合空间信息，生成通道权重向量，进而缩放各通道的特征响应，而空间注意力模块则是通过沿通道维度的池化操作生成空间权重图，进一步细化重要区域。这种组合设计不仅轻量级且通用性强，可无缝集成到多种CNN架构中，能够显著提升模型在图像分类、目标检测和语义分割等任务中的性能。

2.7 本章小结（Summary）

本章主要介绍了本文研究内容涉及到的相关技术。先是介绍了几种常见的植被指数法，通过计算特定波段的反射率来评估植被覆盖情况；接着详细阐述了语义分割的概念及其在计算机视觉中的重要性，以及卷积神经网络起到了重要作用；然后，进一步介绍了几种典型的语义分割模型，阐述了不同的网络结构和设计思想。最后，又分别详细介绍了门控机制和注意力机制，门控机制在处理序列数据和捕捉长距离依赖关系方面表现出色，注意力机制通过动态分配权重，能够帮助模型更好地聚焦于图像中的重要信息，其中CBAM（Convolutional Block Attention Module）作为一种轻量级的注意力模块，通过通道注意力和空间注意力的结合，进一步增强了特征融合的效果，从而提升语义分割的性能。总之，通过对这些技术的深入理解，可为后续的语义分割研究和应用提供坚实的基础。

3 基于多维重组特征的多尺度马尾藻提取网络

3 Multiscale Sargassum Extraction Network Based on Multidimensional Reorganized Features

3.1 引言（Foreword）

马尾藻作为一种重要的海洋生物资源，它的分布和动态变化对海洋生态系统具有重要影响。马尾藻的分布、聚集和消散会直接受到水体运动的影响，导致它在形态与尺度上变化多样。因而这种多尺度和形态变化的特性使得传统的遥感图像分析方法在提取马尾藻信息时往往面临精度不足和适应性差的问题，尤其是对小尺度马尾藻的提取，容易造成漏提现象。如图3.1所示，图3.1（a）为大尺度聚集的马尾藻，在遥感图像中光谱信息较强，可以看到呈明显的片状分布，较容易分辨；图3.1（b）为离散分布的小尺度马尾藻，其形状多样，难以精确捕捉。



（b）小尺度马尾藻

（a）大尺度马尾藻

图3.1不同尺度的马尾藻遥感影像全局及局部图

Figure 3.1 Global and Local Views of Sargassum at Different Scales in Remote Sensing Images

由于在门控机制中引入了门控单元，模型可以根据输入数据的特性自适应地调整不同尺度特征的权重，使门控机制不仅能够有效地捕捉大尺度特征的全局信息，还能够保留小尺度特征的细节信息，进而优化了多尺度特征的融合过程，提升了模型对复杂数据的适应性和表达能力。

受门控机制思想的启发，本章提出了一种基于多维重组特征的多尺度马尾藻提取方法，该方法通过多维度特征重组、细粒度特征提取和全局语义信息获取，来提升模型对多尺度目标的适应能力，其过程有效融合了不同网络层的特征，捕捉目标的局部细节和全局总体信息，并且通过抑制冗余特征和高效计算，减少背景干扰并降低计算复杂度。

3.2 MSENet网络结构设计（Design of MSENet Network Structure）

3.2.1 MRFNet网络总体结构

本章提出了一种基于多维重组特征的多尺度马尾藻提取网络，称为 MSENet，专门用于从高分辨率遥感影像中提取马尾藻信息。其总体结构如图3.2所示，MSENet是基于UNet的编码器-解码器架构，为了得到更丰富的语义信息，先是在跳跃连接层引入了多尺度上下文提取模块（MOCE），它可以捕捉多尺度特征的全局上下文信息，接着引入上采样拼接模块（UPCM），它是通过上采样和拼接操作，将低层次的特征和高层次特征进行融合，从而得到不同尺度的马尾藻信息。为了增强复杂情况下特征表达能力，又引入了多维重组注意模块（CBGM），它是由多维特征重组模块（GMFR）和卷积注意力模块（CBAM）两个子模块协调工作，GMFR是通过对多尺度特征进行分组和多维重组，在通道、高度和宽度三个维度上进行精细操作，可以增强网络对有效信息的关注，提高对马尾藻多尺度分布的适应性，而CBAM是从通道维度和空间维度优化特征图，进一步增强对马尾藻重要特征的关注，且抑制背景噪声。

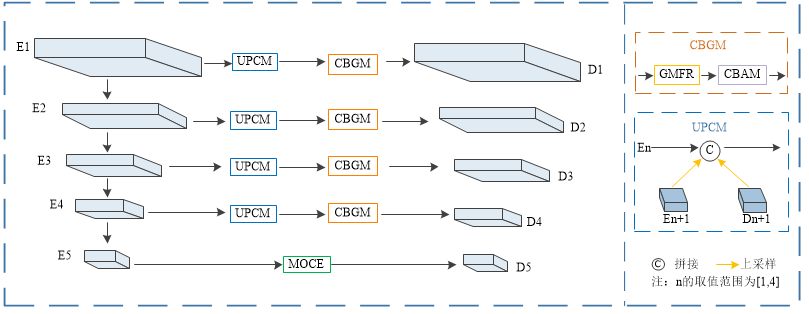


图3.2 MRFNet网络总体结构

Figure 3.2 Overall Structure of the MRFNet Network

3.2.2 多维特征重组模块（GMFR）

考虑到在高分辨率遥感影像中，马尾藻呈现复杂性和多样性分布，且常受到水体、云层等背景信息的干扰，引入了多维特征重构模块（GMFR），它可以通过对特征图在通道、高度和宽度这三个维度的细粒度重组，来增强网络对马尾藻的多尺度特征表达能力，同时抑制冗余信息。G**MFR具体**结构如图3.3所示。

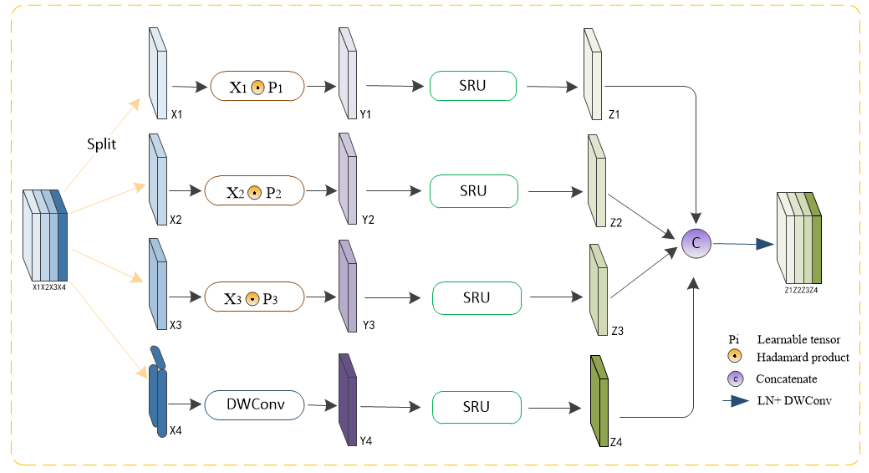


图3.3 **GMFR**结构图

Figure 3.3 Structural Schematic Diagram of GMFR

GMFR模块通过使用哈达玛乘积进行逐元素乘法来增加有效感受域，从而获得全局语义信息，其工作流程主要可分为以下几个步骤：

首先，将输入的多尺度特征图FI和随机初始化的可学习张量 ，用双线性插值调整 的大小以匹配 FI，按通道维度分为四个相等的部分，对于前三组特征图，通过Hadamard积与可学习张量进行逐元素相乘，以增强特征的多维度表达能力，对于第四组特征图，则采用深度可分离卷积（DWConv）进行处理，以降低计算复杂度。具体表达式如（3.1）所示，其中，⊙ 表示哈达玛积，*DW*表示深度可分离卷积。

（3.1）

然后，通过空间重组单元(SRU)重组，并且获得空间细化的特征图Zi。

最后，SRU模块对重组后的特征图进行进一步处理，以生成最终的空间重组特征图 ​。具体来说，SRU将重组后的特征图 沿通道维度拼接，在通过层归一化对拼接后的特征图进行标准化处理，使用深度可分离卷积对标准化后的特征图进行卷积操作，生成最终的空间重组特征图 。这一过程可以表示为公式（3.2）​所示，其中，*DW*是深度可分离卷积，*LN*是层归一化，*Cat*是拼接。

（3.2）

3.2.3 空间重组单元（SRU）

空间重组单元是**GMFR**模块核心组件，用于对特征图进行空间维度的细粒度重组。其结构如图3.4所示。

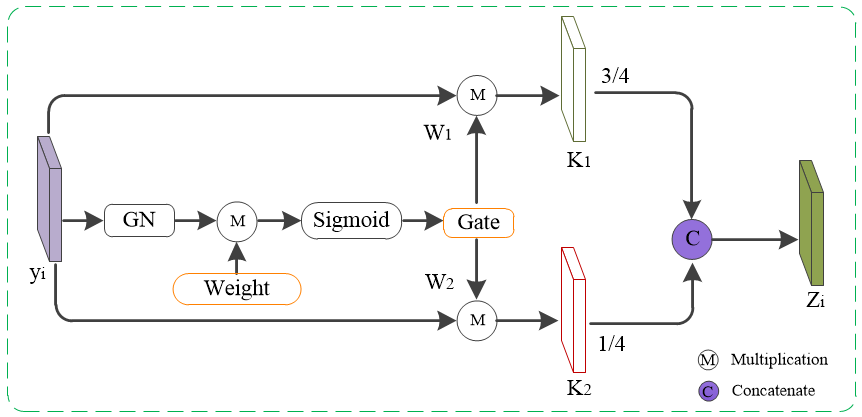


图3.4 空间重组单元结构图

Figure 3.4 Structural Diagram of the Spatial Reorganization Unit

空间重组单元的核心思想是通过对特征图进行分组、标准化和加权，重新组合特征图的空间信息，从而提升网络对马尾藻的特征表达能力。首先，SRU将输入的特征图按通道维度分成若干组，为了消除特征图之间的尺度差异，使网络能够更好地学习到有效的特征，对每组特征图进行标准化处理。标准化公式如（3.3）所示，其中， 是输入的特征图，*μ*和 *σ* 分别是特征图的均值和标准差，*γ* 和 *β* 是可学习的缩放和偏移参数，*ε*是一个小的常数，用于防止除零错误。

（3.3）

其次，SRU通过计算每组特征图的重要性权重来决定哪些特征图包含更多有效信息，哪些特征图是冗余的，权重计算公式如（3.4）所示， 其中， 是每组特征图的缩放参数，C 是通道数。

（3.4）

然后，通过Sigmoid函数和门控机制（Gate），SRU将权重映射到0到1之间，并根据预设阈值（此处使用0.5）将特征图分为有效特征和冗余特征，让SRU能够有效提取特征图中的有用信息，同时抑制冗余信息，从而提升网络的特征表达能力。具体计算过程如（3.5）所示，其中，*S*是Sigmoid函数，*Gate*是门控机制。

（3.5）

最后，SRU模块通过将有效特征和冗余特征分别乘以对应的权重 w1 和 w2，在按通道维度进行重组，以增强特征图的空间表达能力，通常SRU会保留更多的有效特征和较少的冗余特征，从而确保了重组后的特征图能够更好地捕捉有用信息。重组过程如公式（3.6）表示，其中，⊗ 表示逐元素相乘，*Cat*表示通道维度的拼接。通过这种细粒度的重组机制，SRU能够有效提升特征图的质量，同时抑制冗余信息。

（3.6）

3.2.4 卷积块注意力模块（CBAM）

为了进一步增强多尺度特征的融合效果，本章引入了卷积块注意力模块（Convolutional Block Attention Module, CBAM）。首先，G**MFR**通过多维度的特征重组提取了丰富的多尺度特征，为后续处理提供了高质量的特征输入，接着，CBAM通过通道注意力和空间注意力机制，动态调整这些特征的权重，突出对马尾藻特征更重要的通道和空间区域，同时抑制背景信息的干扰。这种结合不仅增强了网络在多尺度特征融合和空间上下文信息捕捉方面的能力，还提高了模型对复杂背景和光照变化的鲁棒性。此外，CBAM的空间注意力机制能够捕捉马尾藻分布的长距离依赖关系，尤其在大范围马尾藻群的提取任务中，帮助网络更好地理解其分布模式。因此，CBAM与G**MFR**的结合使网络更好地适应了不同尺度的马尾藻分布，显著提升了特征表达能力、抗干扰能力和鲁棒性，使网络在各种复杂环境下都能准确提取马尾藻。

3.2.5 多尺度方向上下文提取模块（MOCE）

考虑到马尾藻的形态和分布具有局部密集或大范围稀疏及长距离的空间关联性，传统的方法难以有效捕捉这种全局信息，设计了一种多分支结构MOCE模块来引入丰富的马尾藻信息。具体结构如图3.5所示。

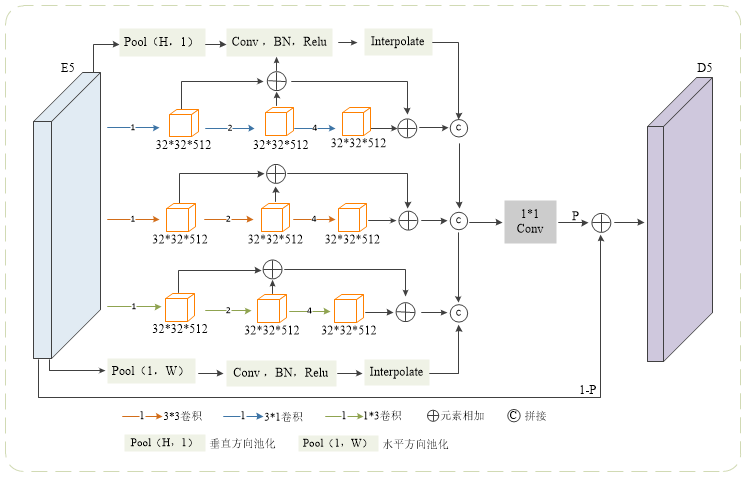


图3.5 MOCE结构图

Figure 3.5 Structural Schematic Diagram of MOCE

MOCE利用具有{1，2，4}的膨胀率的3×3卷积、3×1卷积(水平核)和1×3卷积(垂直核)组成的三个分支，以及由由水平池和垂直池组成的两个分支。卷积分支提取不同尺度的马尾藻特征，汇集分支可以保留马尾藻在垂直和水平方向的全局信息，在利用元素相加融合同一分支中不同接收域的特征映射的操作；然后，连接和卷积操作组合从不同分支输出的特征图；最后，将E5与融合特征图相结合，得到具有丰富马尾藻背景的特征图D5，其中p是一个可学习的参数。

3.3 实验准备（Experimental preparation）

3.3.1 实验数据

本章所使用的实验数据均是由高分一号卫星拍摄的遥感影像。高分一号卫星于2013年12月正式投入使用，该卫星配置了2台分辨率为2m全色以及8m多光谱的高分辨率相机(PMS)和4台分辨率为16m的多光谱中分辨率宽幅相机(WFV)，设置蓝、绿、红、近红外4个波段[95]。GF-1 WFV详细参数如表3.1所示。

表3.1 高分一号卫星传感器技术指标

Table 3.1 Technical Parameters of GF-1 Satellite Sensors

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 传感器 | 波段 | 波长范围 (μm) | 中心波长 (μm) | 空间分辨率 (m) | 幅宽  (km) | 重访周期 (day) |
| GF-1 WFV | 1 | 0.45-0.52 | 0.485 | 16 | 800 | 2 |
| 2 | 0.52-0.60 | 0.560 |
| 3 | 0.63-0.69 | 0.660 |
| 4 | 0.76-0.90 | 0.830 |

3.3.2 研究区域

本章主要选了两个研究区域，均位于中国东部沿海，主要覆盖山东省和江苏省的黄海海域（122°17′ ~ 125°37′E，33°31′ ~ 35°48′N）、江苏省和浙江省的黄海及东海海域（121°57′ ~ 125°14′E，31°51′ ~ 34°08′N），获取时间为2017年5月18日，传感器为GF-WFV1。

3.3.3 实验流程

如图3.6所示，基于MRF的马尾藻提取的总体流程包括三个阶段，分别是数据预处理、模型训练和模型预测。

（1）数据预处理阶段：首先，从遥感影像中提取包含马尾藻的图像区域，并利用ENVI、ArcGIS等遥感图像处理软件裁剪出感兴趣区域（ROI），去除无关背景；其次，通过ArcGIS软件进行专家目视解译生成真值图像并将其转化为二值图像，其中马尾藻区域标记为1，背景标记为0；然后，将裁切后的马尾藻图像进行数据归一化处理，并与对应的真值图像一同裁剪为128×128大小的样本集合；最后，将处理后的图像与对应的标签数据整理成训练样本集合。

（2）模型训练阶段：将预处理后的数据训练马尾藻提取模型，使其能够准确识别马尾藻区域。首先，通过训练数据更新模型的参数，优化模型的性能；然后，计算模型预测结果与真实标签之间的差异（损失值），用于指导参数更新；最后，利用训练样本对模型进行迭代训练，通过反向传播算法不断优化模型参数，从而提升模型对马尾藻区域的提取精度。

（3）模型预测阶段：这一阶段主要是使用训练好的模型对新的遥感图像进行预测，以获取马尾藻的分布信息。具体操作是，它的输入用的是已进行裁剪或归一化等预处理操作，但是没有经过训练的测试图像，接着使用训练好的模型对这些测试图像进行预测，以得到马尾藻的分布概率图，再将预测结果与图像真值标签进行对比，评估模型的精度、召回率及交并比等性能，并输出马尾藻提取结果。

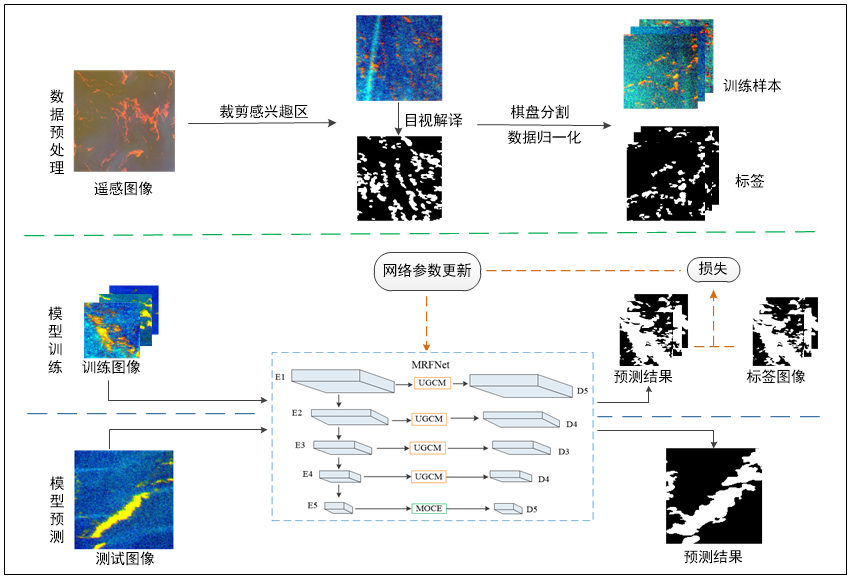


图3.6 基于MRFNet的马尾藻提取流程

Figure 3.6 The Sargassum Extraction Process Based on MRFNet

3.3.4 实验设置

在本章中，提出的 MRFNet 在 Python 3.9.13 和 PyTorch 2.1.2+cu121 深度学习框架下构建，并采用 Adam [96]（Adaptive Moment Estimation）优化器进行模型优化，自适应地调整模型的学习率。所有的实验都在 Ubuntu 18.04 操作系统上进行，使用一个 NVIDIA GeForce RTX 2080Ti 显卡进行网络训练。实验环境详细信息如表3.2所示。实验超参数设置如下：初始化学习率（learning rate）设置为10-4，批处理大小（batch\_size）设置为8，迭代次数（epochs）设置为200。使用二元交叉熵（Binary Cross-Entropy）作为网络损失函数，损失函数表达如式（3.7）所示：

（3.7）

其中，表示二元标签值0或者是属于标签值的概率，*N*为样本的总数量

表3.2 实验环境详细信息

Table 3.2 Detailed Information of the Experimental Environment

|  |  |
| --- | --- |
| 实验环境 | 参数 |
| 系统 | Ubuntu 18.04.5 LTS |
| CPU | Intel(R) Xeon(R) Silver 4214R CPU @ 2.40GHz |
| GPU | NVIDIA GeForce RTX 2080Ti 11GB显存 |
| RAM | 128GB |

3.3.5 评价指标

本文基于混淆矩阵定量评估马尾藻探测结果的精度。马尾藻探测混淆矩阵如表3.3所示。

**表3.3 马尾藻探测精度评估的混淆矩阵**

Table 3.3 Confusion Matrix for Sargassum Detection Accuracy Assessment

|  |  |  |
| --- | --- | --- |
|  |  | 真值图中像元是否为马尾藻水体  True False |
| 预测图中像元 | True | 真正例（TP） 假正例（FP） |
| 是否为马尾藻水体 | False | 假负例（FN） 真负例（TN） |

**其中：真正例（True Positives, TP）**表示马尾藻像元被模型预测为马尾藻水体的数量，**真负例（True Negatives, TN）**表示海水像元被模型预测为海水的数量，**假正例（False Positives, FP）**表示海水像元被模型预测为马尾藻水体的数量，**假负例（False Negatives, FN）**表示马尾藻像元被模型预测为海水的数量。

**基于混淆矩阵可以计算出以下四个评价指标：精确率（Precision, P）、召回率（Recall, R）、F1分数（F1-score, F1）以及交并比（Intersection over Union, IoU）。各指标取值范围为[0,1]。**

**精确率又称查准率，是针对模型预测结果而言，在模型预测为正例的样本中实际为正例的比例。计算方法如式（3.8）所示。**

**（3.8）**

**召回率又称查全率，是针对真值图而言，表示真值图中为正例的样本被模型正确预测为正例的比例。计算方法如式（3.9）所示。**

**（3.9）**

F1分数是精确率和召回率的加权调和平均值，用于综合评估模型的性能。计算方法如式（3.10）所示。

（3.10）

交并比用于衡量模型预测结果与真实结果在空间位置上的重合程度，表示预测结果与真实结果的重叠部分与二者并集部分的比值。计算方法如式（3.11）所示。

（3.11）

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

3.4.1 对比实验

为了验证本章网络提取马尾藻的有效性，分别对MRFNet进行了详细的定性与定量分析。本章主要展现两张测试图像，分别是测试图像1和测试图像2。测试图像1选自山东省和江苏省的黄海海域部分，该图特点是马尾藻呈现密集分布斑块、由于强光照不均，导致暗区边缘模糊及细长弯曲结构变形。测试图像2，选自江苏省和浙江省的黄海及东海海域部分，其图像和真值图如图3.7所示。

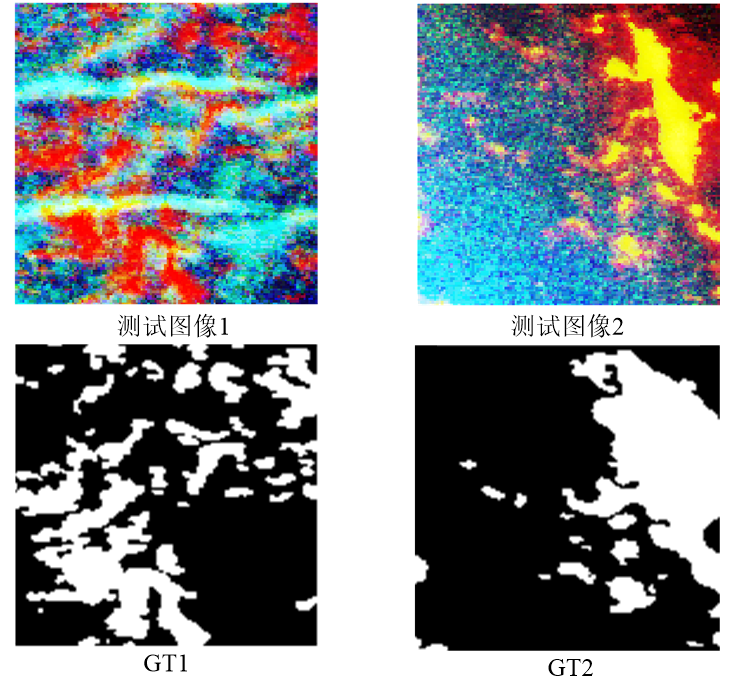


图3.7 测试图像及真值图

Figure 3.7 Test Image and Ground Truth Map

本章实验主要是用与当前流行的深度学习遥感图像语义分割方法UNet、ResNet和SegFormer，以及针对赤潮任务上提出的深度学习模型RTDNet进行对比。实验测试图像可视化结果分别如图3.8和图3.9所示，其中，在图3.8（a）和图3.9（a）中，白色和黑色分别表示马尾藻像素和海水像素，其余图3.8（b-f）和图3.9（b-f）中，绿色和黑色分别表示模型正确提取到的马尾藻像素和海水像素，红色和白色分别表示模型误提和漏提的马尾藻像素。对应的数据结果分别如表3.4和3.5所示，它们是所有测试图像的总体精度评估结果，其中加粗字体为每列精度最优值。具体实验结果如下。

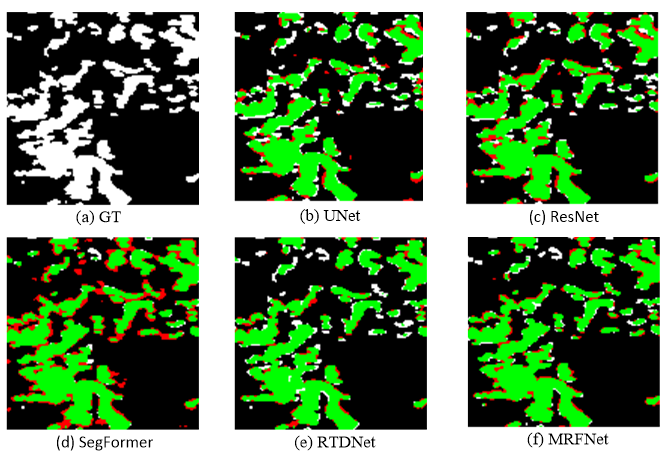


图3.8 测试图像1马尾藻提取定性结果

Figure 3.8 Qualitative Results of Sargassum Extraction for Test Image 1

在图3.8中，从中可以看到，在测试图像1上使用了五种方法对马尾藻提取进行定性对比。UNet和ResNet模型在马尾藻像素的提取上表现相当，但都出现较部分误提和漏提的情况。UNet可能是因为跳跃连接保留了细节，适合小目标场景，但全局感知不足，而ResNet可能是因为其深层特征抽象能力强，但浅层细节丢失，导致低对比度区域误提。SegFormer对马尾藻漏提较少，但是存在大量的误提现象，可能因为其全局建模能力能较好分割密集区域粘连部分，但计算复杂度相对较高。RTDNet在光照不均场景下泛化性差，可能因为它是针对赤潮设计的网络，比较依赖赤潮光谱特征。相比之下， MRFNet模型在马尾藻像素的提取上表现最佳，绿色区域与GT中的白色区域几乎完全重叠，误提和漏提的情况最少。

表3.4 测试图像1马尾藻提取定量性能对比结果

Table 3.4 Comparative Results of Quantitative Performance for Sargassum Extraction on Test Image 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 方法 | 精确率 | 召回率 | F1分数 | IoU |
| UNet | 0.866 | 0.766 | 0.813 | 0.685 |
| ResNet | 0.844 | 0.829 | 0.836 | 0.719 |
| SegFormer | 0.893 | 0.743 | 0.811 | 0.682 |
| RTDNet | **0.908** | 0.848 | 0.877 | 0.781 |
| MRFNet (our) | 0.880 | **0.911** | **0.895** | **0.811** |

在表3.4中，可以看到使用了五种方法在测试图像1上对马尾藻提取进行定量性能对比。MRFNet的召回率（0.911）、F1分数（0.895）和IoU（0.811）指标较其他方法最高，综合表现最优。UNet虽召回率较高（0.766），但其精确率（0.866）和F1分数（0.813）较低，暴露出其原因可能是跳跃连接保留了细节，适合小目标场景，但全局感知不足；ResNet的F1分数（0.836）较好，但对光照不均场景适应性不足（IoU仅0.719）；SegFormer的精确率（0.893）和F1分数（0.811）较高，但召回率（0.743）较低，表明其对细其多样性敏感度不足，RTDNet因依赖特定光谱特征，在复杂光照下泛化性较差，虽然精确率0.908，但其IoU仅0.781。相比之下，MRFNet表现出对马尾藻提取的高精度与高鲁棒性。

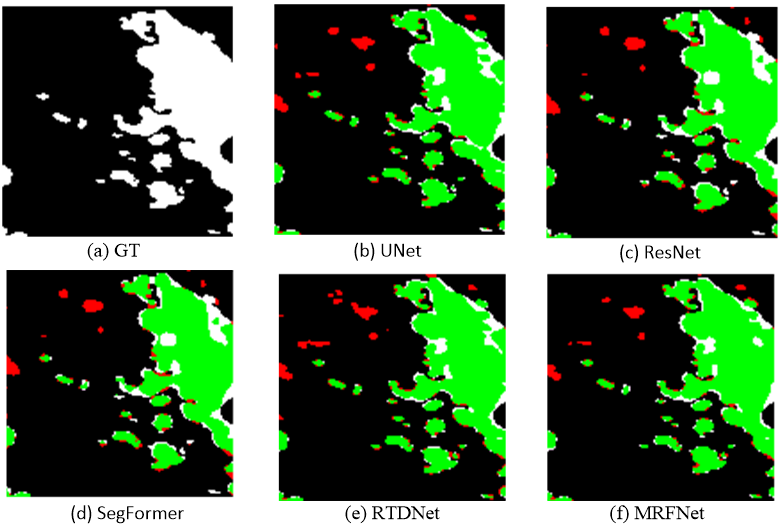


图3.9 测试图像2马尾藻提取定性结果

Figure 3.9 Qualitative Results of Sargassum Extraction for Test Image 2

在图3.9中，可以看到， UNet虽能提取较大连续斑块，但存在边缘模糊和小目标漏提， ResNet通过残差结构保留部分细节，但对暗色水体中的马尾藻易误判为背景， SegFormer凭借Transformer全局建模能力较好分割密集区域，但计算复杂度高且边缘锯齿化， RTDNet依赖特定光谱特征，对形态多变的马尾藻泛化性不足。相比之下，MRFNet绿色区域与GT中的白色区域几乎完全重叠，可以捕捉多尺度特征（如细长弯曲结构和分散斑点），误提和漏提的情况最少，其在马尾藻像素的提取上表现最佳。

表3.5 测试图像2马尾藻提取定量性能对比结果

Table 3.5 Comparative Results of Quantitative Performance for Sargassum Extraction on Test Image 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 方法 | 精确率 | 召回率 | F1分数 | IoU |
| UNet | 0.918 | 0.848 | 0.882 | 0.789 |
| ResNet | 0.885 | 0.788 | 0.834 | 0.715 |
| SegFormer | 0.924 | 0.728 | 0.814 | 0.687 |
| RTDNet | 0.903 | 0.880 | 0.892 | 0.805 |
| MRFNet (our) | **0.926** | **0.889** | **0.907** | **0.830** |

在表3.5中可以看到，在测试图像2上通过五种方法对马尾藻提取定量性能进行对比。MRFNet在精确率（0.926）、召回率（0.889）和F1分数（0.907）上均显著优于其他方法，尤其在复杂背景干扰下（如暗色水体中的马尾藻边缘检测），其IoU较SegFormer提升14.4个百分点，表明其对多尺度特征（细斑点/弯曲带）具有两个的综合处理能力，而SegFormer虽精确率略高（0.924），但召回率骤降17.8个百分点，暴露其对小目标的敏感性不足；RTDNet因依赖特定光谱特征，F1分数（0.892）略低于MRFNet，且IoU差距明显。相比之下，MRFNet表现出最优的特性，凸显了多维重组特征模块（GMFR）、卷积注意力模块（CBAM）和多尺度上下文提取模块（MOCE）的协同作用，其性能优势通过消融实验得到验证。

3.4.2 消融实验

为了验证多维重组特征模块（GMFR）、卷积注意力模块（CBAM）和多尺度上下文提取模块（MSDM）的协同作用有效性，本文对三个模块进行了消融实验。基线网络采用是经典模型UNet。消融实验对应数据指标如表3.6所示，该表中的数据结果为所有测试图像的总体精度评估结果，其中加粗字体为每列精度最优值。

表3.6 马尾藻消融实验对比结果

Table 3.6 Ablation Experiment Comparison Results for Sargassum

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 方法 | 精确率 | 召回率 | F1分数 | IoU |
| Base | 0.918 | 0.848 | 0.882 | 0.789 |
| Base+GMFR | 0.848 | **0.961** | 0.901 | 0.820 |
| Base+MSDA | **0.935** | 0.782 | 0.852 | 0.742 |
| Base+ GMFR+MOCE | 0.910 | 0.878 | 0.894 | 0.809 |
| Base+GMFR+MOCE+CBAM (our) | 0.926 | 0.889 | **0.907** | **0.830** |

在表3.6中，展示了马尾藻消融实验中不同模块组合对模型性能的影响，结果显示，加入GMFR后，模型对小尺度马尾藻的敏感度增强，召回率达0.961上升了13.3个百分点，但精确率略降7.4个百分点，F1分数增至0.901，IoU提高至0.820，MOCE模块虽将精确率提升至0.935，但召回率骤降至0.782，下降19.6个百分点，F1分数和IoU均低于Base模型，表明其全局上下文融合策略可能不利于局部特征的提取，**GMFR与MOCE协同优化**结合后，精确率和召回率趋近于平衡，F1分数达0.894，IoU为0.809，但未超越Base+GMFR，最后在**GMFR和MOCE**基础上加入CBAM后，精确率、F1分数和IoU均为最优，其动态注意力机制有效平衡了精确率与召回率的矛盾，并进一步抑制了背景干扰。实验表明，GMFR是提升小尺度目标检测的核心模块，而CBAM通过自适应特征选择优化了全局性能，两者结合可为高分辨率遥感目标检测提供有效策略。

3.4.3 网络模型可视化结果分析

为了展现网络逐步精细化提取马尾藻的过程，本小节对MRFNet马尾藻探测网络在不同阶段的部分特征图进行了可视化，结果如图3.10所示。

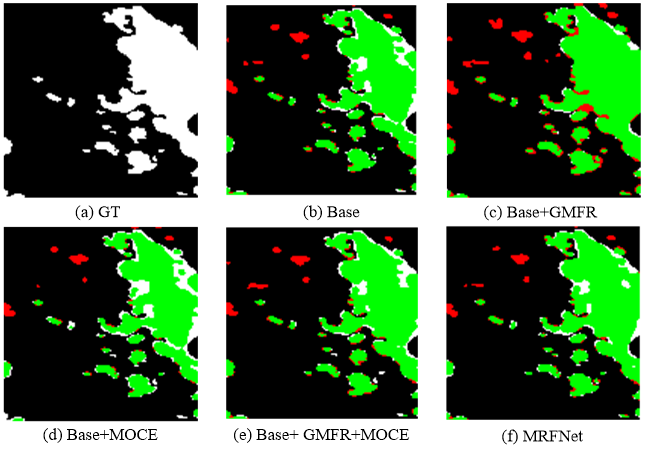


图3.10 马尾藻不同阶段网络模型可视化结果

Figure 3.10 Visualization Results of Sargassum at Different Stages of Network Model

在图3.10中可以看到，图（a）真值标签，呈现是马尾藻和背景，图（b）中可Base模型识别马尾藻出现漏提较多，说明模型缺乏多尺度特征提取能力，难以区分马尾藻与复杂背景（如海水、云层），在引入GMFR后，多尺度特征重组能力增强，绿色区域明显增多且形状更贴近真值分布，说明GMFR通过多维度特征重组（通道/空间维度）和门控机制，有效抑制了背景噪声，其哈达玛乘积和SRU模块增强了局部特征表达，有助于提升检测能力。另外，在Base基础上引入MOCE，可以看到MOCE通过多分支结构捕捉了马尾藻的全局上下文关联，再在Base基础上同时引入GMFR和MOCE能够实现马尾藻局部细节和全局语义信息的平衡，提升模型的完整性和准确性，在此基础上继续引入CBAM，通过动态权重分配，进一步强化了对马尾藻关键特征的关注。实验结果验证，随着模块添加，绿色检测区域显著扩展，即正确提取到的马尾藻增多，即GMFR（多尺度特征提取）和MOCE（全局上下文融合）结合起到了核心作用，在通过CBAM（动态注意力抑制干扰）的辅助作用进一步优化了效果。

3.5 本章小结（Summary）

本章围绕**基于多维重组特征的多尺度上下文马尾藻提取网络（MRFNet）** 展开，系统性地介绍了该网络的设计、实现及其在高分辨率遥感影像马尾藻提取中的应用。首先，在引言部分阐述了马尾藻提取的研究背景及其在海洋生态监测中的重要性，并指出了现有方法的局限性。随后，详细描述了基于MRFNet的马尾藻提取流程，从数据预处理到网络输出的各个环节均得到了清晰的阐述。在网络结构设计部分，重点介绍了MRFNet的核心模块，包括**多维特征重构模块（GMFR）、空间重组单元（SRU）、卷积块注意力模块（CBAM）**和**多尺度方向上下文提取模块（MOCE），**这些模块通过多尺度特征重组、上下文信息融合以及注意力机制，显著提升了网络对马尾藻特征的表达能力，同时有效抑制了复杂背景的干扰。实验部分通过对比分析和消融实验以及网络模型可视化结果分析验证了 MRFNet 的优越性，结果表明该模型在马尾藻提取任务中具有较高的精度和鲁棒性。本章的研究为遥感影像中的马尾藻监测提供了一种高效、可靠的解决方案，并为后续相关研究奠定了坚实的基础。

4 基于多层注意力和顺序聚合的马尾藻边缘优化网络

4 Sargassum Edge Optimization Network Based on Multi - Layer Attention and Sequential Aggregation

4.1 引言（Foreword）

马尾藻的分布通常呈现显著的空间异质性，其生物量并非均匀分布。在集中分布区，马尾藻密度较高，形成明显的漂浮聚集带，光谱特征显著，而在边缘区域，马尾藻密度较低，其光谱特征与周围水体像素高度相似，增加了遥感监测的难度。如图4.1所示。

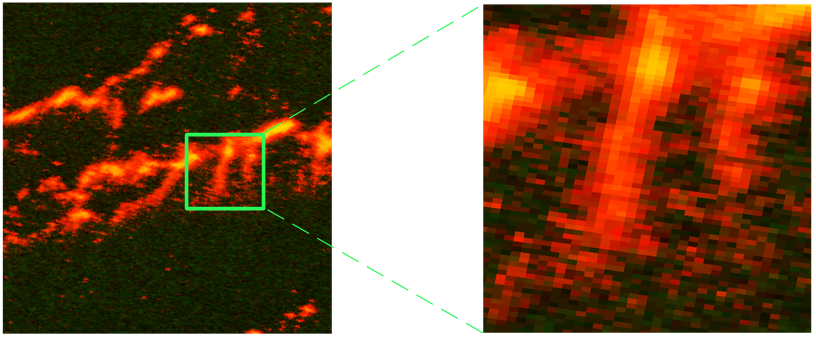


图4.1 马尾藻遥感影像和局部放大图

Figure 4.1 Remote Sensing Images of Sargassum and Their Local Magnified Views

近年来，深度学习技术在遥感图像分割领域取得了显著进展，尤其是基于卷积神经网络（CNN）的方法在复杂场景下的特征提取能力表现出色。然而，现有方法在处理马尾藻边缘区域时仍存在明显局限性，比如传统网络结构对低生物量区域的特征提取能力不足，缺乏对边缘信息的显式利用导致分割边界不精确，同时多层次特征的融合策略不够高效，难以充分表征马尾藻的空间分布特性。

为了解决上述问题，本章提出了一种基于多层注意力和顺序聚合的马尾藻边缘优化网络（Sargassum Edge Optimization Network Based on Multi - Layer Attention and Sequential Aggregation, SEO-Net）。该网络通过以下设计实现对马尾藻边缘区域的精确提取：首先，引入多层注意力机制，增强网络对低生物量区域特征的提取能力；其次，设计边界引导模块，利用边缘信息优化分割结果；最后，采用顺序聚合策略，逐步融合多层次特征，提升网络对马尾藻空间分布的表征能力。

4.2 SEONet网络结构设计（Design of SEONet Network Structure）

4.2.1 SEONet总体结构

本章提出的基于多层注意力和顺序聚合的马尾藻边缘优化网络，称为SEONet，是一个专门设计用于提高马尾藻边界提取精度的深度学习模型。总体结构如图4.2所示，该模型首先利用一个预训练的Backbone Network和一系列ResNet块来提取图像的多级特征。然后，这些特征通过MLA（多层注意力）模块进行处理，MLA由多个AKC（Attentional Kernel and Channel）模块组成，每个AKC模块结合了KAM（Kernel Attention Module）和CAM（Channel Attention Module），以增强特征图的通道和空间信息。最后，在特征经过AKC模块增强后，它们进入SAB（Sequential Aggregation Block），该模块通过1x1卷积层和SBA（Sequential Boundary Aggregation）模块来整合不同ResNet块提取的特征。SBA模块的目的是聚合边界特征和语义信息，以增强目标边界的细节，其中RAU（Re-calibration Attention Unit）负责选择性地聚合低层（细节丰富）和高层（语义丰富）的特征，通过这一过程实现特征的精细调整和融合。

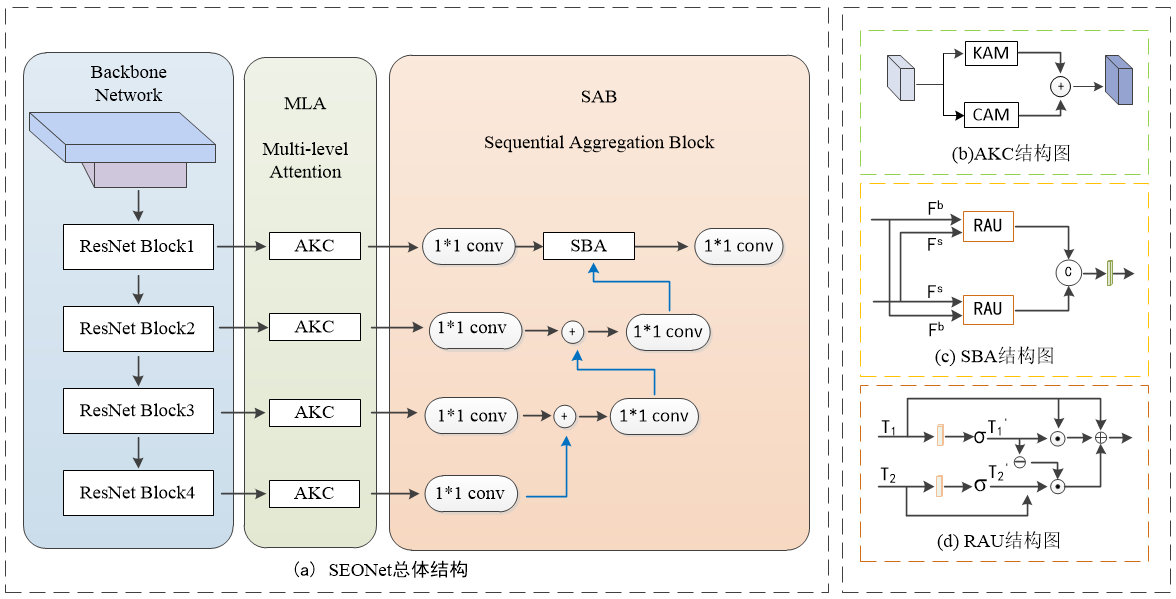


图4.2 SEONet总体结构图

Figure 4.2 Schematic Diagram of the Overall Structure of SEONet

4.2.2 多层注意力模块

考虑到马尾藻边界区域通常包含丰富的细节信息，而全局信息则有助于理解这些细节在整体图像中的位置和意义。为了更全面地捕捉马尾藻的局部细节和全局上下文信息，从而更好地获取马尾藻边界特征，我们引入了多层注意力模块（MLA，Multi-level Attention）。其具体结构如图4.2 (a)所示，该模块通过多个串联的注意力知识组件（AKC）实现多尺度特征提取。每个AKC能够捕捉不同尺度的特征，为模型提供丰富的多尺度信息，这对于识别和提取马尾藻边缘等细节信息至关重要。

如图4.2 (b)所示，每个AKC内部由核注意力（KAM）和通道注意力（CAM）构成，且两者并行工作。KAM专注于识别马尾藻边缘的关键特征，而CAM则理解这些边缘特征在图像上下文中的关系。这种处理策略使得模型能够同时考虑边缘特征及其上下文，从而更准确地提取边缘信息。接着，KAM和CAM的输出通过加法操作进行融合，这不仅整合了关键特征和上下文信息，还增强了模型对马尾藻边缘的识别能力。融合后的特征能够表示更加全面，使模型能够更有效地捕捉马尾藻的细节特征，并在全局上下文中进行精确定位，最终实现更精确的边缘提取。

核注意力机制（KAM）是一种通过核函数计算查询和键之间相似性的注意力机制，它有效避免了直接计算高维点积带来的计算复杂度。KAM具体步骤如下：

首先，使用*softplus*函数作为核函数，*softplus*函数的非零特性保证了即使在输入为负时，梯度也不会消失。表达式如公式（4.1）所示。其次，通过核函数计算查询和键之间的相似性，如公式（4.2）所示，其中，ϕ( ) 为核函数映射。

*softplus(x)=log(1+ex)*  （4.1）

*sim(qi​,kj​)=ϕ(qi​)Tϕ(kj​)*  （4.2）

接着，通过softmax函数对相似性进行归一化，得到注意力权重，如公式（4.3）所示。

（4.3）

然后，使用注意力权重对值矩阵 V 进行加权求和，得到输出特征*KAM*，如公式（4.4）所示。

（4.4）

最后，由于核函数的引入，*KAM*的时间和存储复杂度降低到O(N)，显著减少了计算资源的消耗。

通道注意力机制（CAM）是一种通过全局平均池化和全连接层来捕捉特征通道之间全局依赖关系的注意力机制。因此，让KAM和CAM的输出通过加法操作进行融合，不仅整合了局部边缘特征和全局上下文信息，还增强了模型对边缘特征的识别能力。

4.2.3 顺序聚合模块

为了进一步提升马尾藻边界细节信息的提取精度，考虑充分利用浅层特征与深层特征之间的互补性。浅层特征虽然在语义信息上较为匮乏，但它们在细节丰富度、边界清晰度以及准确性方面具有显著优势。相对地，深层特征蕴含了丰富的语义信息，这对于理解图像内容至关重要。然而，简单地将这些低层与高层特征直接融合可能会引发信息冗余和特征不一致的问题。为了解决这一挑战，引入了顺序聚合块（SAB，Sequential Aggregation Block），该模块通过1x1卷积层和选择性边界聚合（SBA，Selective Boundary Aggregation）模块来有效整合不同ResNet块提取的特征。

SBA模块的核心能力在于其能够有选择性地聚合边界信息和语义信息，从而更精细地刻画目标的轮廓，并重新校准物体的位置。其中，重新校准注意力单元（RAU，Re-calibration Attention Unit）负责选择性地聚合来自低层（细节丰富）和高层（语义丰富）的特征，它首先是在融合前自适应地从两个输入(Fs，Fb)中提取相互表示，如图4.2所示，浅层次和深层次信息通过不同的方式被馈送到两个RAU块中，以弥补高级语义特征缺失的空间边界信息和低级特征缺失的语义信息，接着两个RAU块的输出在3×3卷积后被连接。这种聚合策略实现了不同特征的鲁棒组合，并细化了粗糙特征。RAU块函数过程可以表示为公式（4.5）和（4.6）所示。

（4.5）

（4.6）

其中*T1*、*T2*是输入特征，两个线性映射和sigmoid函数、被应用于输入特征，以将通道维数减少到32并获得特征图T1'和T2'。是点乘。是通过减去特征T1'的逆向运算，将不精确、粗糙的估计提炼为精确、完整的预测图[87]。在把一个核大小为1 × 1的卷积运算作为线性映射过程。因此，SBA的过程可以表述为公式（4.7）所示。

（4.7）

其中，是一个3×3卷积，带有一个批处理归一化层和一个ReLU激活层。在融合来自编码器的第三和第四层之后包含深层次的语义信息，是来自主干的第一层，具有丰富的边界细节。*cat*是沿着通道维度的连接操作。是SBA模块的输出。

通过这一精细调整和融合过程，实现了特征表示的优化。通过这种创新的方法，SAB模块不仅显著增强了目标边界的细节表现，还优化了特征融合策略，从而在整体上提高了马尾藻边界提取的精度。

4.3 实验设计（Design of Experiment）

4.3.1 实验数据

本章所使用的实验数据与第三章相同，均是由高分一号卫星拍摄的遥感影像，配置了2台分辨率为2m全色以及8m多光谱的高分辨率相机(PMS)和4台分辨率为16m的多光谱中分辨率宽幅相机(WFV)，设置蓝、绿、红、近红外4个波段。遥感数据集大小仍是128\*128，但在第三章的基础上增加了600张，共1600张。

4.3.2 实验环境与参数设置

本章所用到的实验环境和设备均与第三章实验相同，即所有的实验都在 Ubuntu 18.04 操作系统上进行，**GPU**选用 NVIDIA GeForce RTX 2080Ti 显卡进行网络训练，

**CPU**为Intel Xeon Silver 4214R（主频2.40GHz），支持高并发任务处理，**内存**配置128GB，确保大规模数据加载与多任务运行的流畅性，实验初始化学习率（learning rate）设置为10-4，批处理大小（batch\_size）设置为8，迭代次数（epochs）设置为200。

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

4.4.1 对比实验

为了验证本章网络提取马尾藻的有效性，分别对SIANet进行了详细的定性与定量分析。本章实验主要展现两张测试图像，分别是测试图像1和测试图像2，测试图像1，位于山东省和江苏省的黄海海域，测试图像2，位于江苏省和浙江省的黄海及东海海域，其图像和真值图如图4.3所示。

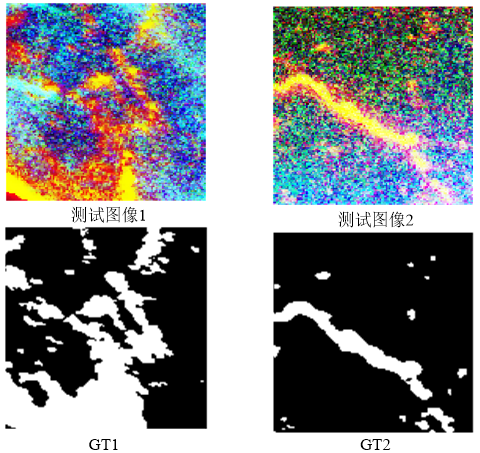


图4.3 测试图像及真值图

Figure 4.3 Test Image and Ground Truth Map

本章实验同第三章主要是用与当前流行的深度学习遥感图像语义分割方法UNet、ResNet和SegFormer，以及针对赤潮任务上提出的深度学习模型RTDNet进行对比。实验测试图像可视化结果分别如图4.4和图4.5所示，其中，在图4.4（a）和图4.5（a）中，白色和黑色分别表示马尾藻像素和海水像素，其余图4.4（b-f）和图4.5（b-f）中，绿色和黑色分别表示模型正确提取到的马尾藻像素和海水像素，红色和白色分别表示模型误提和漏提的马尾藻像素。对应的数据结果分别如表4.1和4.2所示，它们是所有测试图像的总体精度评估结果，其中加粗字体为每列精度最优值。具体实验结果如下。

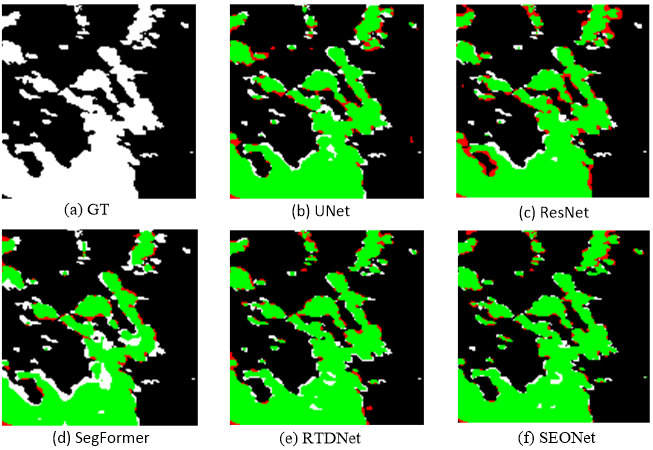


图4.4 测试图像1马尾藻提取定性结果

Figure 4.4 Qualitative Results of Sargassum Extraction for Test Image 1

在图4.4中，是在测试图像1上马尾藻提取定性对比结果。从中可以看到， SEONet在展现出显著的边界细节保留能力，相较于UNet和图ResNet的边界断裂或过度平滑现象（如马尾藻细小分支的缺失），以及SegFormer因全局注意力机制导致的局部模糊和漏检（如边缘区域不完整），而SEONet表现出更少的误提和漏提，可能是由于MLA模块中KAM强化空间细节和CAM抑制背景噪声之间的协同结合起到了作用，此外， RTDNet虽在整体轮廓上与图（a）真值标签接近，但在高曲率或低对比度区域（如藻体末端）仍存在轻微模糊。相比之下， SEONet进一步优化了此类区域的连贯性，最终实现了与GT高度一致的提取效果，可能是因为使用SBA模块的层级特征融合（RAU单元平衡细节与语义）起到了作用。这一结果验证了SEONet在多尺度特征增强与边界优化机制上的有效性。

表4.1 测试图像1马尾藻提取定量性能对比结果

Table 4.1 Comparative Results of Quantitative Performance for Sargassum Extraction on Test Image 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 方法 | 精确率 | 召回率 | F1分数 | IoU |
| UNet | 0.921 | 0.866 | 0.893 | 0.807 |
| ResNet | 0.875 | 0.895 | 0.885 | 0.794 |
| SegFormer | 0.932 | 0.758 | 0.836 | 0.718 |
| RTDNet | 0.930 | 0.914 | 0.922 | 0.856 |
| SEONet (our) | **0.934** | **0.933** | **0.933** | **0.875** |

**在表4.1中，显示的是测试图像1马尾藻提取定量性能对比结果。**实验数据表明，SEONet在精确率（0.934）、召回率（0.933）、F1分数（0.933）和IoU（0.875）四项指标上全面领先，其精确率与召回率高度均衡（仅差0.001），体现了模型在抑制误检与漏检之间的最优平衡。**对比其他模型，UNet因跳跃连接未能充分整合多尺度细节，召回率（0.866）显著偏低（F1=0.893）；ResNet虽召回率较高（0.895），但深层网络过度泛化导致精确率（0.875）不足，误检问题突出；SegFormer依赖全局注意力机制，虽精确率（0.932）接近**SEO**Net，但局部边界响应不足导致召回率（0.758）和IoU（0.718）大幅落后；RTDNet虽表现次优（F1=0.922，IoU=0.856），但其边界特征融合策略对复杂区域的适应性略逊于SIANet的SBA模块，IoU差距达0.019。**SEO**Net的核心优势源于其创新设计，MLA模块通过KAM（空间注意力）聚焦边界关键区域，结合CAM（通道注意力）动态抑制背景噪声，而SBA模块利用RAU单元自适应融合低层细节（如纹理）与高层语义（如形状），并通过1x1卷积压缩冗余特征，最终实现复杂边界的精细化提取。**

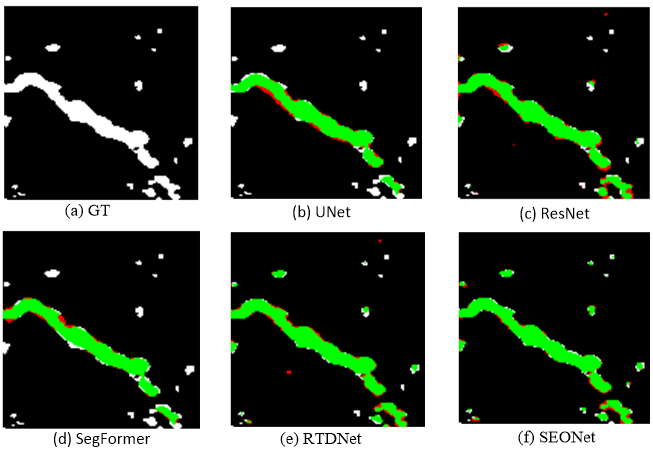


图4.5 测试图像2马尾藻提取定性结果

Figure 4.5 Qualitative Results of Sargassum Extraction for Test Image 2

在图4.5中，展示的是测试图像2马尾藻提取定性结果从中可以看到， SIANet在提取任务中展现出更强的边界适应性与细节还原能力，相较于UNet在高密度藻丛区域出现的轮廓粘连（如相邻藻体边界难以区分），以及ResNet因深层特征泛化导致的局部误检（如将阴影误判为藻体），SEONet通过MLA模块的KAM（空间注意力）精准聚焦藻丝边缘，结合CAM（通道注意力）抑制复杂背景（如水面波纹干扰），显著提升了目标区域的分离度。此外， SegFormer因全局注意力机制对局部纹理的弱响应，导致藻体末端呈现断裂（如箭头所指的碎片化区域），而RTDNet虽在整体轮廓上接近真值标签图，但在低光照区域的模糊边界处（如藻体与背景交界）仍存在轻微漏检。相比之下，SEONet则通过SBA模块的RAU单元动态融合低层纹理特征（如藻丝表面细节）与高层语义信息（如藻体形状），有效修复了此类边界不连续问题，最终在复杂场景下实现了与GT高度一致的分割效果，验证了其多尺度特征增强与边界优化机制的鲁棒性。

表4.2 测试图像2马尾藻提取定量性能对比结果

Table 4.2 Comparative Results of Quantitative Performance for Sargassum Extraction on Test Image 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 方法 | 精确率 | 召回率 | F1分数 | IoU |
| UNet | **0.919** | 0.741 | 0.821 | 0.696 |
| ResNet | 0.859 | 0.874 | 0.867 | 0.765 |
| SegFormer | 0.888 | 0.754 | 0.816 | 0.689 |
| RTDNet | 0.904 | 0.914 | 0.909 | 0.834 |
| SEONet (our) | 0.913 | **0.919** | **0.916** | **0.845** |

在表4.2中，显示的是测试图像2马尾藻提取定量性能对比结果，实验数据表明，UNet虽以最高精确率（0.919）减少误检，但其召回率（0.741）大幅偏低，导致F1分数（0.821）和IoU（0.696）落后，表明跳跃连接在多尺度边界整合上的不足；ResNet召回率较高（0.874），但精确率（0.859）因深层网络对背景干扰（如阴影误检）的敏感性而受限，整体性能（IoU=0.765）未达最优；SegFormer因全局注意力机制弱化局部细节，召回率（0.754）和IoU（0.689）均为最低，尽管精确率（0.888）尚可，藻体末端仍出现明显断裂；RTDNet虽在召回率（0.914）和F1分数（0.909）上接近SEONet，但精确率（0.904）和IoU（0.834）略逊，推测其边界融合策略对高密度藻丛的适应性不足；而SEONet通过MLA模块（KAM强化空间细节、CAM抑制噪声）与SBA模块（RAU动态平衡低层纹理与高层语义）的协同优化，在精确率（0.913）与召回率（0.919）间实现高度均衡（F1=0.916，IoU=0.845），尤其在低对比度区域（如藻体末端）展现出更连贯的分割效果，验证了多尺度特征增强与边界优化机制的有效性。

4.4.2 消融实验

为了验证多层注意力模块（MLA）、顺序信息聚合模块（SAB）协同作用有效性，本章对SEONet进行了消融实验。基线网络采用是ResNet。消融实验结果如表4.3所示，该表中的数据结果为所有测试图像的总体精度评估结果，其中加粗字体为每列精度最优值。

表4.3马尾藻消融实验对比结果

Table 4.3 Ablation Experiment Comparison Results for Sargassum

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 方法 | 精确率 | 召回率 | F1分数 | IoU |
| Base | 0.892 | 0.867 | 0.879 | 0.784 |
| Base+MLA | 0.903 | 0.890 | 0.896 | 0.805 |
| Base+SAB | 0.898 | 0.875 | 0.886 | 0.820 |
| SEONet (our) | **0.913** | **0.919** | **0.916** | **0.845** |

在表4.3中，显示的是马尾藻消融实验对比结果，结果表明，基础模型（Base）在精确率（0.892）、召回率（0.867）、F1分数（0.879）和IoU（0.784）上表现一般；加入**MLA模块**（Base+MLA）后，精确率（0.903）和召回率（0.890）显著提升，主要得益于KAM（空间注意力）抑制背景噪声和CAM（通道注意力）增强目标特征响应，但IoU（0.805）提升有限；加入**SAB模块**（Base+SAB）后，IoU（0.820）因边界融合优化（SBA和RAU单元）提升最明显，召回率（0.875）也小幅增加；而完整SEO**Net**（Base+MLA+SAB）通过模块协同作用，四项指标全面最优（精确率0.913、召回率0.919、F1分数0.916、IoU 0.845），其中MLA模块主导误检与漏检的平衡（精确率+0.021、召回率+0.052），SAB模块强化边界细节融合（IoU+0.061），验证了多级注意力与顺序聚合策略的互补性，为复杂场景下的精细分割提供了有效解决方案。

4.5 本章小结（Summary）

本章针对马尾藻边缘区域生物量密度低、光谱特征与周围水体相似的问题，提出了基于多层注意力和顺序聚合的马尾藻边缘优化网络（SEONet），通过多层注意力模块（MLA）和顺序信息聚合模块（SAB）的协同作用，显著提升了马尾藻边缘区域的识别精度，生成了更加准确的边界信息。实验结果表明，SEONet在精确率、召回率、F1分数和IoU等指标上均优于其他方法，尤其在低对比度区域和复杂背景下的边界提取表现出色。消融实验进一步验证了MLA模块在增强特征提取能力和SAB模块在优化边界细节方面的有效性。本章的研究为马尾藻边缘区域的精确提取提供了新的技术手段，为后续的生物量估算和生态研究提供了可靠支持。

# 5 总结与展望

**5** **Conclusion and prospects**

5.1 总结（Conclusion）

本研究围绕马尾藻的遥感探测问题，基于GF-1高分辨率遥感影像，结合深度学习技术，提出了两种新型网络模型：基于多维重组特征的多尺度上下文提取网络（MRF-Net）和基于多层注意力和顺序聚合的马尾藻边缘优化网络（SEO-Net）。这两种网络模型分别针对马尾藻的多尺度特征提取和边缘区域精确分割问题，取得了显著的效果。

（1）针对马尾藻形态多变、尺度各异的特点，本文提出了MRF-Net网络。该网络通过多维重组特征模块（GMFR）、多尺度上下文提取模块（MOCE）和卷积注意力模块（CBAM）的协同工作，有效捕捉了马尾藻的多尺度特征，增强了对不同形态和尺度马尾藻的识别能力。实验结果表明，MRF-Net在提取马尾藻分布范围时表现出较高的精度和鲁棒性，尤其是在处理复杂背景和多尺度目标时，显著优于传统的遥感图像处理方法和其他深度学习模型。

（2）针对马尾藻边缘区域生物量密度低、光谱特征与周围水体相似的问题，本文提出了SEO-Net网络。该网络通过多层注意力模块（MLA）和顺序信息聚合模块（SAB），结合边界引导策略，优化了马尾藻边缘区域的提取效果。实验结果表明，SEO-Net能够有效提升马尾藻边缘区域的识别精度，生成更加准确的边界信息，为后续的生物量估算和生态研究提供了可靠的技术支持。

综上所述，本文通过结合高分辨率遥感影像和深度学习技术，提出了两种创新的网络模型，成功解决了马尾藻探测中的多尺度特征提取和边缘区域分割难题。研究成果不仅为马尾藻的遥感监测提供了新的技术手段，也为其他类似海洋生态灾害的监测和预警提供了参考。未来的研究可以进一步探索多源遥感数据的融合以及弱监督学习方法的应用，以进一步提升模型的通用性和效率。

5.2 展望（Prospects）

本研究虽在多个测试数据集上取得了良好的结果，但仍存在一些不足之处，需要进一步深入研究。未来的研究方向主要包括以下几个方面：

（1）多源数据融合。目前的研究主要基于单一的 GF-1 遥感影像数据。然而，马尾藻的光谱响应不仅与生物量有关，还与生物的种类、海水温度、盐度以及水质条件等因素密切相关。未来的研究将考虑融合多源卫星遥感数据（如 Sentinel-2、MODIS 等）以及不同马尾藻生物种类的卫星遥感数据，构建通用的马尾藻探测模型，以提高模型对不同环境条件和马尾藻种类的适应性。

（2）弱监督学习和无监督学习方法。本文采用监督学习构建模型，但深度学习在马尾藻提取中对数据和标签的依赖较高，需要大量的人力和物力资源进行标注。为了降低成本，未来研究应考虑引入弱监督学习和无监督学习方法。例如，利用少量标注数据结合大量未标注数据进行半监督学习，或者通过生成对抗网络（GAN）生成伪标签来辅助模型训练，从而提高提取效率并减少对标注数据的依赖。

（3）模型优化与实时性提升。虽然 MRF-Net 和 SEO-Net 在精度上表现出色，但模型的计算复杂度较高，难以满足实时监测的需求。未来的研究将致力于优化网络结构，减少模型参数量和计算量，同时探索轻量化模型在马尾藻检测中的应用，以提高模型的实时性和实用性。

# 参考文献

1. Chen Y, et al. Automatic Extraction Method of Sargassum Based on Spectral-Texture Features of Remote Sensing Images[C]. In: IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019: 3705-3707. DOI: 10.1109/IGARSS.2019.8898131.
2. Pixa C, Manganini K, Myers A, Michel AP. A Low-Cost Entangling Drifter for Sargassum Tracking: Design and Experimental Evaluation[C]. In: OCEANS 2023, Limerick, Ireland, 2023: 1-5. doi: 10.1109/OCEANSLimerick52467.2023.10244472.
3. Zhang H, et al. Detecting Sargassum Bloom Directly From Satellite Top-of-Atmosphere Reflectance With High-Resolution Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-12. Art no. 4206212.
4. Wang M, Hu C. Satellite remote sensing of pelagic Sargassum macroalgae: The power of high resolution and deep learning[J]. Remote Sensing of Environment, 2021, 264: 112631.
5. BBC News. (2018). Sargassum Seaweed Invasion: The Caribbean's Growing Environmental and Economic Crisis. [Online News Article]. Retrieved from [URL].
6. Caribbean Regional Fisheries Mechanism (CRFM). (2019). Economic Impact of Sargassum Seaweed Invasions on Caribbean Fisheries and Tourism. [Regional Report].
7. Florida Department of Environmental Protection. (2020). Economic and Environmental Impact of Sargassum Seaweed on Florida Beaches. [State Report].
8. Jamaica Observer. (2019). Sargassum Seaweed Impact on Tourism and Coastal Communities. [Newspaper Article].
9. Lohrer, A. M., Thrush, S. F., & Gibbs, M. M. (2015). Seaweed Outbreaks: A Global Threat to Coastal Ecosystems and Economies. Annual Review of Marine Science, 7, 313-339.
10. National Marine Environmental Monitoring Center of China. (2018). Impact of Sargassum Bloom on Kelp Farming in Jiangsu Coastal Area. [Internal Report].
11. Pixa, C., Manganini, K., Myers, A., & Michel, A. P. (2023). A Low-Cost Entangling Drifter for Sargassum Tracking: Design and Experimental Evaluation. In OCEANS 2023, Limerick, Ireland, 1-5. DOI: 10.1109/OCEANSLimerick52467.2023.10244472.
12. Sutton, M., Stum, J., Hajduch, G., Dufau, C., Maréchal, J.-P., & Lucas, M. (2019). Monitoring a New Type of Pollution in the Atlantic Ocean: The Sargassum Algae. In OCEANS 2019 - Marseille, Marseille, France, 1-4. DOI: 10.1109/OCEANSE.2019.8867437.
13. [4]Wang, M., & Hu, C. (2021). Satellite Remote Sensing of Pelagic Sargassum Macroalgae: The Power of High Resolution and Deep Learning. Remote Sensing of Environment, 264, 112631.
14. Xiao Y, Liu R, Kim K, Zhang J, Cui T. A Random Forest-Based Algorithm to Distinguish Ulva prolifera and Sargassum From Multispectral Satellite Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-15. Art no. 4201515.
15. [11]Sutton M, Stum J, Hajduch G, Dufau C, Maréchal J-P, Lucas M. Monitoring a new type of pollution in the Atlantic Ocean: the sargassum algae[C]. In: OCEANS 2019 - Marseille, Marseille, France, 2019: 1-4. doi: 10.1109/OCEANSE.2019.8867437.
16. Feng C, Xing Y, Wang M. Colors of macroalgae: distinguishing Ulva prolifera and Sargassum horneri using Sentinel-2 L2A images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. doi: 10.1109/JSTARS.2024.3519615.
17. H. Zhang et al.,“A simple and effective method for monitoring floating green macroalgae Blooms: A case study in the Yellow Sea,” Opt. Exp., vol. 27, no. 4, pp. 4528–4548, 2019
18. C. Hu, L. Feng, R. F. Hardy, and E. J. Hochberg, “Spectraland spatial requirements of remote measurements of pelagic Sargassum macroalgae,” Remote Sens. Environ., vol. 167, pp. 229–246, Sep. 2015.
19. M. Wang and C. Hu,“Automatic extraction of Sargassum features from Sentinel-2 MSI images,” IEEE Trans. Geosci. Remote Sens., vol. 59, no. 3, pp. 2579–2597, Mar. 2021.
20. H. Zhang et al., "Detecting Sargassum Bloom Directly From Satellite Top-of-Atmosphere Reflectance With High-Resolution Images," in IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-12, 2023, Art no. 4206212, doi: 10.1109/TGRS.2023.3295826.
21. Kwon, D., Hong, S.M., Abbas, A., Park, S., Nam, G., Yoo, J., Kim, K., Kim, H.T., Pyo, J., & Cho, K.H. (2023). Deep learning-based super-resolution for harmful algal bloom monitoring of inland water. GIScience & Remote Sensing, 60.
22. Wang, M., & Hu, C. (2021). Satellite remote sensing of pelagic Sargassum macroalgae: The power of high resolution and deep learning. Remote Sensing of Environment, 264, 112631.
23. Gernez, P., D. Doxaran, and L. Barille, Shellfish Aquaculture from Space: Potential of Sentinel2 to Monitor Tide—Driven Changes in Turbidity, Chlorophyll Concentration and Oyster Physiological Response at the Scale of a Oyster Farm. Frontiers in Marine Science, vol.4, Article 137, pp.1 - 16, 2017, doi: 10.3389/fmars.2017.00137.
24. H. Q. Zhang et al.,“Analysis of pigments infloating andfixed Sargassum horneri by high performance liquid chromatography coupled with triple quadrupole tandem mass spectrometry,” (in Chinese), J. Nucl. Agricult. Sci., vol. 33, no. 6, pp. 1173–1180, 2019.
25. M. Terasaki et al.,“Spatial and seasonal variations in the biofunctional lipid substances (fucoxanthin and fucosterol) of the laboratory-grown edible Japanese seaweed (Sargassum horneri Turner) cultured in the open sea,” Saudi J. Biol. Sci., vol. 24, no. 7, pp. 1475–1482, Nov. 2017.
26. H. Endo, Y. Okumura, Y. Sato, and Y. Agatsuma,“Interactive effects of nutrient availability, temperature, and irradiance on photosynthetic pigments and color of the brown alga Undaria pinnatifida,” J. Appl. Phycol., vol. 29, no. 3, pp. 1–11, 2016.
27. Hu C. 2009. A novel ocean color index to detect floating algae in the global oceans. Remote Sensing of Environment 113(10):2118–2129 DOI 10.1016/j.rse.2009.05.012.
28. Wang, M.; Hu, C. Mapping and Quantifying Sargassum Distribution and Coverage in the Central West Atlantic Using MODIS Observations. Remote Sens. Environ. 2016, 183, 350–367.
29. Xing, Q., Guo, R., Wu, L., An, D., Cong, M., Qin, S., & Li, X. (2017). High-Resolution Satellite Observations of a New Hazard of Golden Tides Caused by Floating Sargassum in Winter in the Yellow Sea. IEEE Geoscience and Remote Sensing Letters, 14(10), 1815-1820. doi:10.1109/LGRS.2017.2737079.
30. 于杰,黄洪辉,舒黎明,等.马尾藻遥感信息提取[J].遥感信息,2013,28(02):93-100+105.
31. Z. L. Zhang, “Remote sensing identification of Ulva prolifra and Sargassum and evolution of green tide in the Yellow Sea and the East China Sea,” Ph.D. dissertation, School Geograph. Sci., East China Normal Univ., Shanghai, China, 2014.
32. S. H. Min, J. D. Hwang, H. J. Oh, and Y. B. Son,“Reflectivity characteristics of the green and golden tides from the Yellow Sea and East China Sea,” J. Coastal Res., vol. 90, no. 1, pp. 310–316, 2019.
33. Mei, L., Han, Z., Liu, Y. A remote sensing method for distinguishing between Ulva and Sargassum [J]. Remote Sensing Information, 2016, 31(2): 44-48.
34. Wang, S., Li, J., Zhang, B., Spyrakos, E., Tyler, A. N., Shen, Q., Zhang, F., Kuster, T., Lehmann, M. K., Wu, Y., Peng, D. Trophic state assessment of global inland waters using a MODIS-derived Forel-Ule index [J]. Remote Sensing of Environment, 2018, 217: 444-460
35. Zhou, Y., He, B., Fu, C., Giardino, C., Bresciani, M., Liu, H., … Liang, S. (2021). Assessments of trophic state in lakes and reservoirs of Wuhan using Sentinel-2 satellite data. European Journal of Remote Sensing, 54(1), 461–475. https://doi.org/10.1080/22797254.2021.1960201
36. K. Kim, J. Shin, K. Y. Kim, and J. H. Ryu,“Long-term trend of green and golden tides in the Eastern Yellow Sea,” J. Coastal Res., vol. 90, no. 1, pp. 317–323, 2019.
37. Gower, J.; Hu, C.; Borstad, G.; King, S. Ocean Color Satellites Show Extensive Lines of Floating Sargassum in the Gulf of Mexico. IEEE Trans. Geosci. Remote Sens. 2006, 44, 3619–3625.
38. Sun, Y., Wang, M., Liu, M., Li, Z. B., Chen, Z., Huang, B. Continuous Sargassum Monitoring Across the Caribbean Sea and Central Atlantic Using Multi-Sensor Satellite Observations [J]. Remote Sensing of Environment, 2024, 309: 114223. DOI: 10.1016/j.rse.2024.04.019.
39. Wang, M., Hu, C., Zhang, S., Barnes, B. B., Xie, Y., Cannizzaro, J. P., English, D. C. Mapping and Quantifying Pelagic Sargassum in the Atlantic Ocean Using Multi-Band Medium-Resolution Satellite Data and Deep Learning [J]. Remote Sensing of Environment, 2023, 284: 113337. DOI: 10.1016/j.rse.2023.04.019.
40. Li, D., Gao, Z., Wang, Z., 2022. Analysis of the reasons for the outbreak of Yellow Sea green tide in 2021 based on long-term multi-source data. Mar. Environ. Res. 178, 105649.
41. https://doi.org/10.1016/j.marenvres.2022.105649.
42. Chapelle, P. Haffner, V.N. Vapnik. Support vector machines for histogram– based image classification [J]. IEEE Transaction on Neural Networks, 1999, 10(5), 1055-1064.
43. Y. Chen et al., "Automatic Extraction Method of Sargassum Based on Spectral-Texture Features of Remote Sensing Images," IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019, pp. 3705-3707, doi: 10.1109/IGARSS.2019.8898131.
44. L. Breiman and R. A. Cutler,“Random forests machine learning,” J. Clin. Microbiol., vol. 2, no. 1, pp. 199–228, 2001.
45. D. R. Cutler et al.,“Random forests for classification in ecology,” Ecology, vol. 88, no. 11, pp. 2783–2792, Nov. 2007.
46. Verikas, A. Gelzinis, and M. Bacauskiene,“Mining data with random forests: A survey and results of new tests,” Pattern Recognit., vol. 44, no. 2, pp. 330–349, Feb. 2011.
47. H. Phan, M. Maas, R. Mazur, and A. Mertins,“Random regression forests for acoustic event detection and classification,” IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 23, no. 1, pp. 20–31, Jan. 2015.
48. R. A. Nugrahaeni and K. Mutijarsa,“Comparative analysis of machine learning KNN, SVM, and random forests algorithm for facial expression classification,” in Proc. Int. Seminar Appl. Technol. Inf. Commun. (ISemantic), Semarang, Indonesia, Aug. 2016, pp. 163–168.
49. Paul, D. P. Mukherjee, P. Das, A. Gangopadhyay, A. R. Chintha, and S. Kundu,“Improved random forest for classification,” IEEE Trans. Image Process., vol. 27, no. 8, pp. 4012–4024, Aug. 2018.
50. M. Belgiu and L. Dragut,“Random forest in remote sensing: A review of applications and future directions,” ISPRS J. Photogramm. Remote Sens., vol. 114, pp. 24–31, Apr. 2016.
51. R. Colditz, “An evaluation of different training sample allocation schemes for discrete and continuous land cover classification using decision tree-based algorithms,” Remote Sens., vol. 7, no. 8, pp. 9655–9681, 2015.
52. J. Haas and Y. Ban,“Urban growth and environmental impacts in Jing-Jin-Ji, the Yangtze, river delta and the pearl river delta,” Int. J. Appl.Earth Observ. Geoinf., vol. 30, pp. 42–55, Aug. 2014.
53. J. Stefanski, B. Mack, and O. Waske,“Optimization of object-based image analysis with random forests for land cover mapping,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 6, no. 6, pp. 2492–2504, Dec. 2013.
54. N. Tsutsumida and A. J. Comber,“Measures of spatio-temporal accuracy for time series land cover data,” Int. J. Appl. Earth Observ. Geoinf., vol. 41, pp. 46–55, Sep. 2015.
55. Cuevas E, Uribe-Martínez A, Liceaga-Correa MDLÁ. 2018. A satellite remote-sensing multi-index approach to discriminate pelagic Sargassum in the waters of the Yucatan Peninsula, Mexico. International Journal of Remote Sensing 39(11):3608–3627 DOI 10.1080/01431161.2018.1447162.
56. Xiao, Y., Liu, R., Kim, K., Zhang, J., Cui, T., 2021. A random forest based algorithm to distinguish Ulva prolifera and Sargassum from multispectral satellite images. IEEE Trans. Geosci. Remote Sens. 60, 115. <https://doi.org/10.1109/TGRS.2021.3071154>.
57. Arellano-Verdejo, J.; Lazcano-Hernandez, H.E.; Cabanillas-Terán, N. ERISNet: Deep Neural Network for Sargassum Detection along the Coastline of the Mexican Caribbean. PeerJ 2019, 7, e6842.
58. Wang, M.; Hu, C. Satellite Remote Sensing of Pelagic Sargassum Macroalgae: The Power of High Resolution and Deep Learning. Remote Sens. Environ. 2021, 264, 112631.
59. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, Munich, Germany, 5–9 October 2015; pp. 234–241.
60. Yan, K.; Li, J.; Zhao, H.; Wang, C.; Hong, D.; Du, Y.; Mu, Y.; Tian, B.; Xie, Y.; Yin, Z.; et al. Deep Learning-Based Automatic Extraction of Cyanobacterial Blooms from Sentinel-2 MSI Satellite Data. Remote Sens. 2022, 14, 4763. [CrossRef]
61. Laval, M.; Belmouhcine, A.; Courtrai, L.; Descloitres, J.; Salazar-Garibay, A.; Schamberger, L.; Minghelli, A.; Thibaut, T.; Dorville, R.; Mazoyer, C.; et al. Detection of Sargassum from Sentinel Satellite Sensors Using Deep Learning Approach. Remote Sens. 2023, 15, 1104. <https://doi.org/10.3390/rs15041104>.
62. Liu, M., Wang, M., Sun, Y., & Li, Z. B. (2024). Deep-Learning-Based Cloud Masking on Multispectral Ocean Color Imagery for Floating Macroalgae Monitoring. IEEE Transactions on Geoscience and Remote Sensing, 62, 4200213.
63. Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment, 8(2), 127-150.
64. Hu, C. (2009). A novel ocean color index to detect floating algae in the global oceans. Remote Sensing of Environment, 113(10), 2118-2129.
65. Wang, M., & Hu, C. (2016). Mapping and quantifying Sargassum distribution and coverage in the central west Atlantic using MODIS observations. Remote Sensing of Environment, 183, 350-367.
66. Xu, H., EATwell, C. J., & Gao, B. (2012). A review of remote sensing of suspended sediments in surface waters. ISPRS Journal of Photogrammetry and Remote Sensing, 70, 66-72.
67. H, Y., Liu, R., Kim, K., Zhang, J., & Cui, T. (2022). A Random Forest-Based Algorithm to Distinguish Ulva prolifera and Sargassum from Multispectral Satellite Images. IEEE Transactions on Geoscience and Remote Sensing, 60, 1-15.
68. Feng, C., Xing, Y., & Wang, M. (2024). Colors of macroalgae: distinguishing Ulva prolifera and Sargassum horneri using Sentinel-2 L2A images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.
69. Min, S. H., Hwang, J. D., Oh, H. J., & Son, Y. B. (2019). Reflectivity characteristics of the green and golden tides from the Yellow Sea and East China Sea. Journal of Coastal Research, 90(1), 310-316.
70. Zhang, H., et al. (2023). Detecting Sargassum Bloom Directly From Satellite Top-of-Atmosphere Reflectance With High-Resolution Images. IEEE Transactions on Geoscience and Remote Sensing, 61, 1-12.
71. Haifeng Li, Kaijian Qiu, Li Chen, Xiaoming Mei, Liang Hong, and Chao Tao. "SCAttNet: Semantic Segmentation Network With Spatial and Channel Attention Mechanism for High-Resolution Remote Sensing Images." IEEE Geoscience and Remote Sensing Letters, 2020, pp. 1-5. DOI: 10.1109/LGRS.2020.2988294
72. X. Huang, L. Zhang, and W. Gong,“Information fusion of aerial images and LIDAR data in urban areas: Vector-stacking, re-classification and post-processing approaches,” Int. J. Remote Sens., vol. 32, no. 1, pp. 69–84, Jan. 2011.
73. Y. Yang, S. Hallman, D. Ramanan, and C. C. Fowlkes,“Layered object models for image segmentation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 9, pp. 1731–1743, Sep. 2012.
74. Ronneberger, P. Fischer, and T. Brox,“U-Net: Convolutional networks for biomedical image segmentation,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2015, pp. 234–241.
75. J. Long, E. Shelhamer, and T. Darrell,“Fully convolutional networks for semantic segmentation,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2015, pp. 3431–3440.
76. V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A deep convolutional encoder-decoder architecture for image segmentation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
77. Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. science, 2006, 313(5786): 504-507.
78. Khan S, Rahmani H, Shah S A A, et al. A guide to convolutional neural networks for computer vision[J]. 2018.
79. Xu J, Li Z, Du B, et al. Reluplex made more practical: Leaky ReLU[C]//2020 IEEE Symposium on Computers and communications (ISCC). IEEE, 2020: 1-7.
80. L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam,“Encoder–decoder with Atrous separable convolution for semantic image segmentation,” in Proc. ECCV, 2018, pp. 801–818.
81. E. Shelhamer, J. Long and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 4, pp. 640-651, 1 April 2017, doi: 10.1109/TPAMI.2016.2572683.
82. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. Springer International Publishing, 2015: 234-241.
83. He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[A]. 2016: 770–778.
84. Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735-1780. DOI: 10.1162/neco.1997.9.8.1735.
85. Cho, K., van Merriënboer, B., Bahdanau, D., & Bengio, Y. (2014). On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. arXiv:1409.1259. <https://arxiv.org/abs/1409.1259>
86. Wang, X., Girshick, R., Gupta, A., & He, K. (2018). Non-local neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
87. Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.
88. X. Wang, R. Girshick, A. Gupta, and K. He,“Non-local neural networks,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2018, pp. 7794–7803.
89. H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena,“Self-attention generative adversarial networks,” in Proc. Int. Conf. Mach. Learn., 2019, pp. 7354–7363.
90. Bello, B. Zoph, Q. Le, A. Vaswani, and J. Shlens,“Attention augmented convolutional networks,” in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Oct. 2019, pp. 3286–3295.
91. J. Fu et al.,“Dual attention network for scene segmentation,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 3146–3154.
92. Y. Li et al.,“Attention-guided unified network for panoptic segmentation,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 7026–7035.
93. Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 7132-7141).
94. Liu, M., Wang, M., Sun, Y., & Li, Z. B. (2024). Deep-Learning-Based Cloud Masking on Multispectral Ocean Color Imagery for Floating Macroalgae Monitoring. IEEE Transactions on Geoscience and Remote Sensing, 62, 4200213.
95. S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 3-19.
96. 白照广.高分一号卫星的技术特点[J].中国航天,2013(08):5-9.
97. Kingma D P, Ba J. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.
98. Deng-Ping Fan, Ge-Peng Ji, Tao Zhou, Geng Chen, Huazhu Fu, Jianbing Shen, and Ling Shao. Pranet: Parallel reverse attention network for polyp segmentation. In International conference on medical image computing and computer assisted intervention, pages 263–273. Springer, 2020.

# 作者简历

1. 基本情况

姓名：张艳平 性别： 女 民族：汉

出生年月：1996.08.18 籍贯：山东省临沂市

1. 2017.09-2021.07 德州学院学士
2. 2022.09-2025.07 山东科技大学硕士
3. 攻读学位期间获奖情况
4. 2022年10月获山东科技大学二等学业奖学金
5. 2023年10月获山东科技大学三等学业奖学金
6. 2024年10月获山东科技大学三等学业奖学金
7. 攻读学位期间研究项目情况

# 致 谢

# 行文至此，落笔为终。所有经历皆有学习，二十余载的求学之路渐进尾声，硕士生涯终将拉下帷幕。回首过往，收到硕士录取通知书时满心欢喜，于是带着憧憬出发，来到青岛这个美丽的滨海城市，开始了硕士阶段的学习和生活，曾哭过、笑过、迷茫过，但所有体会和收获皆是幸运，目光所及皆是回忆，感恩生命中的每一次相遇。

# 学位论文数据集

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **关键词\*** | **密级\*** | | | | **中图分类号\*** | | | **UDC** | | | | **论文资助** |
| 马尾藻探测；多维重组特征；门控机制；注意力机制；语义分割 | 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 | | | | | 工程硕士 | | | 硕士 | | |
| **论文题名\*** | | | | | | **并列题名\*** | | | | | **论文语种\*** | |
| 基于深度学习的多尺度马尾藻提取与边缘优化网络方法研究 | | | | | | Research on Multi-scale Sargassum Extraction and Edge Optimization Networks Based on Deep Learning | | | | | 中文 | |
| **作者姓名\*** | | 张艳平 | | | | | **学号\*** | | | 202283060066 | | |
| **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 | | | | | 3年 | | | 2025 | | |
| **论文提交日期\*** | | | | | 2025年5月 | | | | | | | |
| **导师姓名\*** | | 路燕 | | | | | **职称\*** | | | 副教授 | | |
| **评阅人** | | | | **答辩委员会主席\*** | | | | **答辩委员会成员** | | | | |
|  | | | | 崔宾阁 | | | |  | | | | |
| **电子版论文提交格式 文本（√） 图像（ ） 视频（ ） 音频（ ） 多媒体（ ）**  **其他（ ）**  **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\*** | | | | 79 | | | | | | | | |
| **注：共33项，其中带\*为必填数据，共22项。** | | | | | | | | | | | | |