doi:10.6041/j.issn.1000-1298.2026.03.029

# 基于双分支 U－Net 的遥感影像稻田分割方法

王　漫[1]　吴敏琪[1,2]　胡　冬[2]　田明璐[2]　李琳一[2]

(1. 上海第二工业大学计算机与信息工程学院，上海 201209；2. 上海市农业科学院科技信息研究所，上海 201403)

**摘要**：遥感影像中小面积、分布零散的水稻田块难以精准分割。基于线性光谱组合的输入无法挖掘波段间非线性耦合关系，基于堆叠波段影像的输入易引入冗余信息。针对此问题，本文提出一种基于 U－Net 的改进网络，该网络以 RGB 与 NRG 假彩色影像作为双输入，采用双编码器结构提取多模态特征信息，并结合局部金字塔注意力模块与自适应多尺度注意力特征融合模块，显著提升网络对小尺度水稻田块的感知与分割能力。对构建的水稻影像数据集进行实验，表明 DFAU－Net 在分割精度、鲁棒性和效率上表现优异。其 Dice 系数、平均交并比和准确率分别达到 77.54%、86.34% 和 91.48%，较多种主流方法具有明显优势。进一步的消融实验验证了 LPA 模块、AMSADFF 模块和双分支结构的有效性。该方法不仅能提高水稻田块的分割精度，也为复杂背景下的小目标分割提供了有效的解决方案。此外，本研究展示了高分辨率遥感影像在农业监测中的潜力，为精准农业、作物监测及产量估算提供了新的技术路径。综合而言，DFAU－Net 为解决小规模水稻田块分割难题提供了有效的技术支持，具有广泛的实际应用价值。

**关键词**：稻田分割；遥感影像；深度学习；注意力机制；小目标识别

**中图分类号**：TP391　　**文献标识码**：A　　**文章编号**：1000-1298(2026)03-0306-09　　**OSID**：

# Dual-branch U－Net-based Method for Paddy Field Segmentation in Remote Sensing Imagery

WANG Man[1]　WU Minqi[1,2]　HU Dong[2]　TIAN Minglu[2]　LI Linyi[2]

(1. *School of Computer and Information Engineering*, *Shanghai Polytechnic University*, *Shanghai* 201209, *China*
2. *Institute of Agricultural Science and Technology Information*, *Shanghai Academy of Agricultural Sciences*, *Shanghai* 201403, *China*)

**Abstract**: Accurate delineation of small and spatially fragmented paddy fields from remote sensing imagery remains a challenging task. Traditional inputs based on linear spectral combinations are limited in their ability to capture nonlinear couplings across spectral bands, while direct stacking of multi-band imagery often introduces redundant information and increases computational complexity. To address these limitations, an enhanced U－Net-based segmentation framework was proposed. The network employed a dual-input strategy, integrating both RGB and NRG false-color images derived from the Gaofen－2 satellite, and incorporated a dual-encoder architecture to extract complementary multimodal feature representations. To further strengthen the discrimination of fine-scale objects, a local pyramid attention (LPA) module that enabled hierarchical aggregation of local contextual cues was designed, thereby improving the model's sensitivity to small paddy patches with irregular boundaries. In addition, an adaptive multi-scale attention dynamic feature fusion (AMSADFF) module was introduced to dynamically integrate features across multiple scales, mitigating redundancy while preserving the most informative spatial and spectral patterns. By synergizing these mechanisms, the proposed framework—termed DFAU－Net—achieved a robust balance between local detail preservation and global contextual understanding. Experimental evaluations were conducted on a dedicated paddy field dataset constructed from high-resolution Gaofen－2 imagery. Results demonstrated that DFAU－Net consistently outperformed several state-of-the-art segmentation models. Specifically, it achieved a dice coefficient (Dice) of 77.54%, a mean intersection over union (mIoU) of 86.34%, and an overall accuracy (Acc) of 91.48%. These

收稿日期：2025－06－30　修回日期：2025－08－20

**基金项目**：上海市农业科技创新项目(沪农科(I2023005))和上海市农业科学院卓越团队建设计划项目(沪农科卓(2022)015)

**作者简介**：王漫(1967—)，男，教授级高级工程师，主要从事物联网研究，E-mail：wangman@sspu.edu.cn

**通信作者**：李琳一(1968—)，女，研究员，主要从事农业遥感、高通量植物表型采集和分析技术研究，E-mail：lly@saas.sh.cn

results highlighted the superiority of the method in capturing fragmented and small-scale field patterns, where conventional models tended to fail. Overall, DFAU－Net provided a promising solution for accurate agricultural parcel mapping, with potential applications in crop monitoring, yield estimation, and precision agriculture.

**Key words**: paddy field segmentation; remote sensing images; deep learning; attention mechanism; small object recognition

## 0　引言

水稻是我国第一大口粮作物，产量接近粮食总产量的 1/3[1]。准确掌握水稻种植面积，能为粮食产量估测、农业政策制定以及农产品市场调控等提供重要数据基础[2]。遥感技术结合机器学习与深度学习算法，是精确监测水稻种植面积的重要手段，方法的选择取决于稻田空间特征。水稻田块尺寸、形状和分布等特征受地形、水文、耕作方式以及城市化程度等因素影响，有些地区田块面积较大，且集中连片；有些地区田块则“小而散”。长三角都市周边区域受河流、道路等自然或人工设施的分割影响，水稻田块面积小、不规则且分布零散，在卫星遥感影像上呈现细粒度特征，更易受复杂背景影响，分割难度较大。

基于卫星遥感的稻田分割任务首先要选择合适的卫星平台。从卫星影像选择上来看，现有研究[3－6]主要基于 Sentinel 和 Landsat 系列的 10 m 级别中高分辨率卫星平台以及 100 m 级别的 MODIS 低分辨率卫星平台，能够满足分割大面积集中连片的水稻区域需求。但是，在分割具有“小而散”分布特征的水稻田块时会产生很大误差，需要选择更高分辨率的卫星平台，例如高分系列卫星平台[7]。

通过选择不同波段的光谱信息，或利用它们构建不同的指数特征，例如归一化植被指数（NDVI）、增强型植被指数（EVI）[8]、归一化差值水体指数（NDWI）[9]、修正的 NDWI（MNDWI）[10]、裸土指数 BSI[11]、归一化差值建筑指数（NDBI）[12]可以增强水稻与背景之间的可分性。然而，基于先验知识构建的线性光谱组合无法挖掘波段之间的非线性关系，也没利用影像的空间结构与纹理特征。单纯增加输入波段数量会产生噪声与冗余信息，降低模型的分割精度与效率。利用与水稻强相关的近红外波段构建出假彩色影像，在不过多产生冗余信息的同时引入影像的空间及纹理特征，结合深度学习模型挖掘非线性特征，能够提高分割精度。

随着卷积神经网络的发展，语义分割模型能够自动提取影像中的非线性高阶特征。在各种深度学习模型[13－18]中，U－Net 模型的 U 形结构能有效地将底层纹理信息与高层语义信息集成在一起，在只需要少量参数的情况下便能对图像实现有效分割，适用于高分辨率遥感影像中的小目标分割任务。但在 U－Net 中，从浅层提取的低级特征和从深层提取的抽象特征没有得到充分交互。其采用逐层下采样的结构，随着图像的分辨率逐渐降低，模型对小面积水稻田块的捕捉能力也逐渐变弱。为解决以上问题，本文充分结合高分二号卫星（GF－2）亚米级高分辨率多光谱影像特性，设计一种基于 U－Net 框架的改进分割网络，以用于分割都市周边区域细粒度水稻田块。

## 1　材料与方法

### 1.1　研究区介绍

本文选取的研究区位于上海市金山区与浙江省嘉兴市嘉善县交界区域，如图 1 所示。该地区水稻田块普遍呈现面积较小、空间分布零散，且与居民建筑、道路等非农业地物高度交织的典型特征，具有较高的分割难度。研究使用的遥感数据来源于中国高分辨率对地观测系统高分二号卫星（GF－2），拥有 3.2 m 多光谱分辨率。

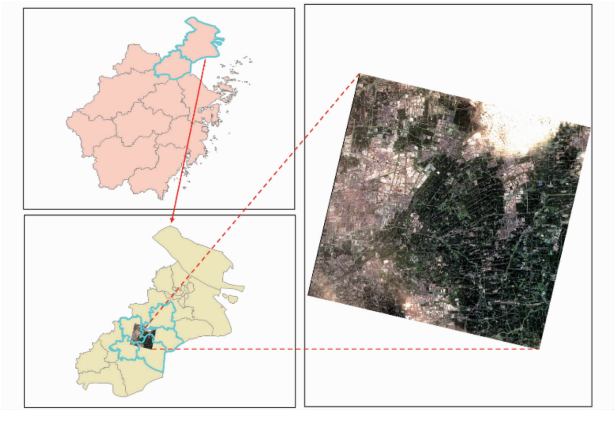


图 1　研究区图

Fig. 1　Map of study area

### 1.2　数据集构建

在数据标注过程中，在研究区域内随机选取若干地块，并结合遥感领域专家经验、Google Earth 高分辨率影像及其他辅助资料对 GF－2 多光谱影像进行综合判读解译，确保标签的准确性与一致性。基于 ArcGIS 软件制作高质量的像素级语义分割标签，随后采用滑动窗口裁剪策略，将大幅影像切分成尺寸为 128 像素×128 像素的深度学习样本。使用

数据增强来扩大样本量,包括沿不同维度的随机翻转和随机移位。最终构建出的数据集共包含 848 幅裁剪影像,将该数据集命名为“GF2RiceSeg”。为保证模型训练和评估的科学性及公平性,将数据集随机划分为训练集、验证集和测试集。其中 70% 样本用于模型训练;20% 样本用于验证模型性能;剩余未参与任何训练过程的样本作为独立测试集,用于评估模型的泛化能力。在训练阶段,采用了 SGD 优化器,训练轮次为 100 轮,权值衰减系数为 0.000 1,初始学习率为 0.001,批次大小为 8。网络架构是在 PyTorch 框架下实现的,所有实验都是在使用 NVIDIA GeForce RTX 4080 GPU、Intel Core i9 - 13900KF CPU 和 32 GB RAM 的标准工作站上进行。训练环境为 Python 3.9,Cuda 12.3,Torch 2.1.0。

### 1.3 U - Net 网络

U - Net 采用了一种 U 形的网络结构,包括收缩路径(编码器)和扩展路径(解码器),两者之间通过跳跃连接相连,架构如图 2 所示。收缩路径通过多次卷积和池化操作提取特征,而扩展路径则通过多次上采样和卷积操作恢复图像的分辨率。跳跃连接将编码器中的特征图与解码器中的对应特征图进行拼接,帮助恢复更多的图像细节信息。

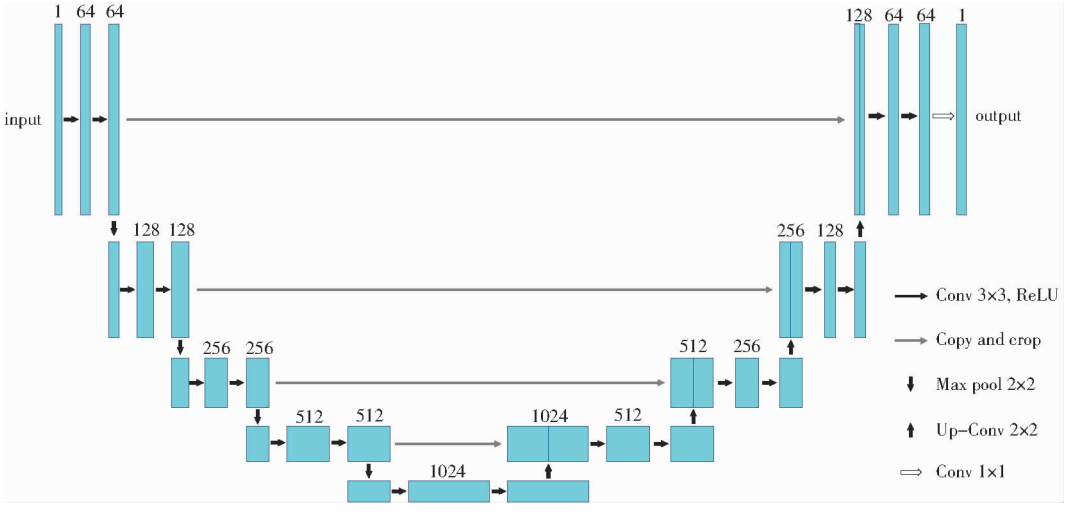



图 2 U - Net 网络架构图

Fig. 2 U - Net network architecture diagram

### 1.4 DFAU - Net 模型

本文提出的模型针对 U - Net 进行一系列改进,具体包括:①引入局部金字塔注意力模块(Local pyramid attention,LPA),引导网络更关注稻田区域。②设计自适应多尺度注意力特征融合模块(Adaptive multi-scale attention dynamic feature fusion,AMSADFF),增强网络对不同尺度小面积稻田的感知能力。③设计双分支编码器架构,将 RGB 影像作为编码器主分支的输入,利用近红外、红光与绿光波段合成 NRG 假彩色影像作为辅助分支的输入。RGB 影像能够提供丰富的空间结构和纹理信息,而假彩色影像通过融合近红外波段增强了水稻与背景(如水体、土壤等)间的光谱差异,将二者进行交互融合能有效提升模型在稻田与背景间的判别能力。基于上述改进,本文将该模型命名为双分支特征注意力 U - Net 网络(Dual-branch feature attentive U - Net,DFAU - Net),用于小尺度水稻田块分割研究。

#### 1.4.1 模型结构改进

本文所设计的 DFAU - Net 结构如图 3 所示。具体而言,以 RGB 影像作为编码器主分支的输入 $x_0$,以 NRG 影像作为编码器辅助分支的输入 $y_0$,解码器的最终输出记为 $z_0$。网络整体设计为 4 层次结构,采用 Layer$i$ 表示第 $i$ 层,其中 $x_i$、$y_i$和 $z_i$分别表示主分支、辅助分支和解码器在第 $i$ 层的特征。

在 Layer$i$ 的编码阶段,$x_i$和 $y_i$分别经过独立的 Down 模块处理,模块包含最大池化、卷积、批归一化和 ReLU 激活操作。为了充分融合 RGB 和 NRG 2 种影像的互补特征信息,将 2 个分支的输出沿通道维度进行拼接。通过这种融合方式,模型既能保留 RGB 影像中的空间及纹理细节信息,又能引入 NRG 影像中对水稻目标更加敏感的光谱特征,从而增强模型的特征判别能力。拼接后的特征进一步输入到 LPA 模块,通过通道降维和注意力加权,引导网络聚焦于与水稻田块最相关的特征通道,同时抑制冗余和噪声信息。LPA 模块的输出作为下一层主分支编码器的输入 $x_{i+1}$,以确保传递的是语义丰富的融合特征。辅助分支 $y_{i+1}$则直接继承自 $y_i$,保持原始近红外光谱特性,有助于在特征融合中维持模态多样性,避免特征过度扭曲。编码阶段 Layer$i$ 层输出特征的形状 $E_i$变化过程为

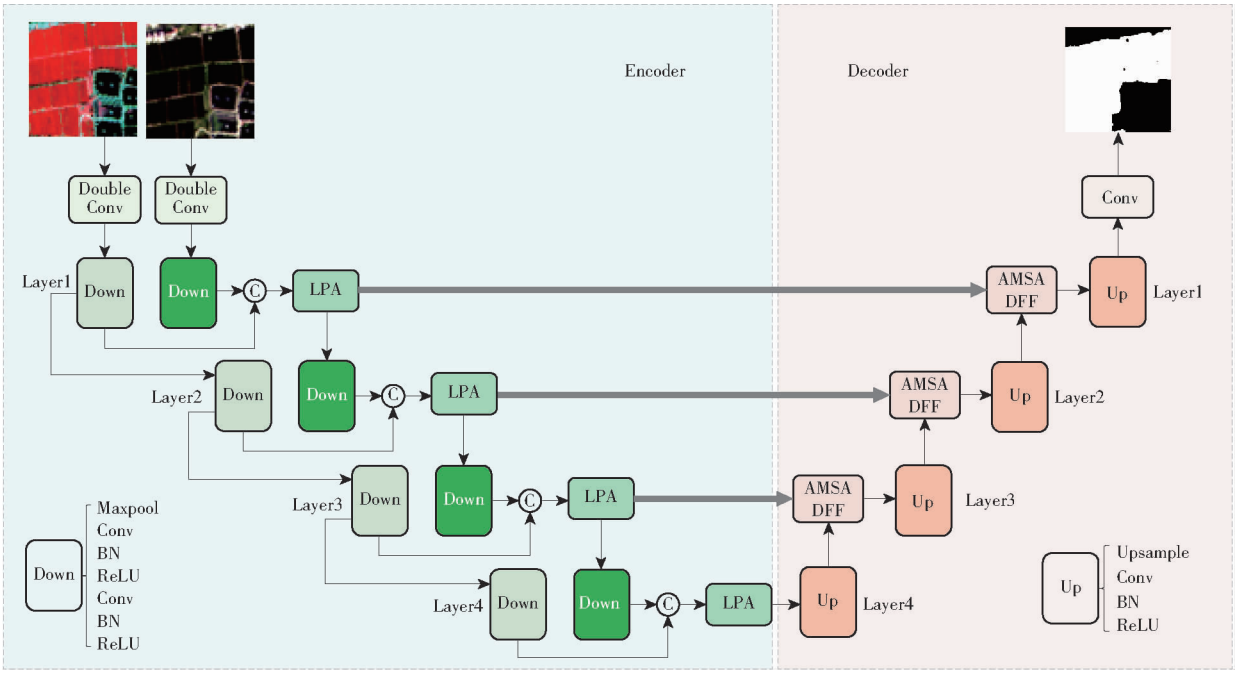



图 3　DFAU－Net 架构图

Fig. 3　DFAU－Net architecture diagram

$$E_i=\begin{cases}(B,128\times2^{i-1},H/2^i,W/2^i) & (1\leqslant i\leqslant3)\\(B,128\times2^{i-2},H/2^i,W/2^i) & (i=4)\end{cases}\tag{1}$$

式中　$B$——批次大小

$H$、$W$——特征图的高、宽

在解码器阶段，传统 U－Net 网络直接通过跳跃连接将编码器特征与解码器特征拼接，这种连接方式连接存在特征尺度单一的问题，在识别不同尺寸的小面积水稻田块时不准确。为此采用 AMSADFF 模块替代传统跳跃连接。AMSADFF 模块分别接收来自 Layer$i$＋1 解码器上采样后的特征和当前层编码器经过 LPA 模块处理后的特征，利用多尺度动态特征融合机制，有效增强模型对不同尺寸水稻田块的感知能力。融合后的特征经过 Up 模块（上采样、卷积、批归一化、ReLU）处理得到 $z_i$。为降低过拟合风险并提升推理效率，对 Layer4 进行简化设计。该层 Down 模块仅执行下采样压缩而不增加特征图数量，在 Up 模块之前不设置 AMSADFF 模块，从而简化高层特征的处理流程。整个解码过程 Layer$i$ 层输出特征的形状 $D_i$ 变化过程为

$$D_i=(B,128\times2^{i-2},H/2^{i-1},W/2^{i-1})\tag{2}$$

### 1.4.2　LPA 模块

在遥感影像分割任务中，理想的分割网络只会关注水稻田块。当影像中存在大量背景（如道路、建筑、林地等）时，会产生许多干扰信息，从而影响模型对水稻田块的分割精度。另外，在影像上并非所有区域对分割任务的贡献都是同等重要的。因此，本研究引入局部金字塔注意力（LPA）模块[19]来引导网络更精确地聚焦在水稻田块上，同时抑制无用的信息，结构如图 4 所示。

LPA 模块包括 2 个块：空间注意力（SA）块和通道注意力（CA）块。利用 SA 模块[20]作为 SA 块，通

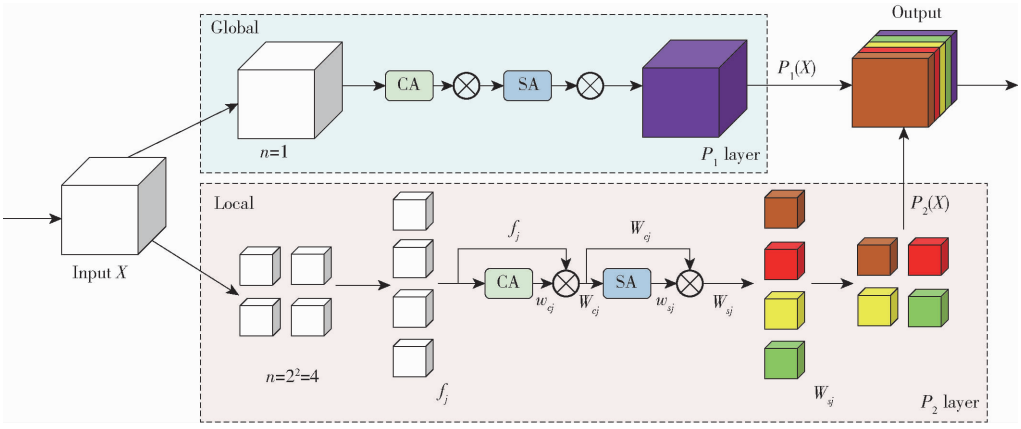



图 4　LPA 模块结构图

Fig. 4　Structure diagram of LPA module

过空间信息寻找特征中与任务相关的部分。采用“挤激”模块(SE 模块)[21]作为 CA 块,为通道分配不同的权重来衡量特征的重要性。LPA 模块整体采用金字塔结构 $P$ 来学习输入特征 $X$ 的信息,其中 $P_i$ 表示第 $i$ 层金字塔。首先通过金字塔第 1 层得到全局注意力图。随后金字塔剩余层将特征图分割成更小的维度,并通过局部注意力机制得到更准确的水稻田块边界。然后将局部信息合并到当前层的注意力图中。最后将不同层注意力图 $P_i(X)$ 相互融合得到最终的注意力图。

### 1.4.3 AMSADFF 模块

在遥感影像分割任务中,理想的分割网络应能精准地定位出各种尺寸的水稻田块。然而,由于本研究区的水稻田块呈现出面积小且零散分布的特征,导致传统卷积网络难以准确定位目标区域。对于小面积的水稻田块而言,随着编码器不断进行下采样特征提取,特征数量不断增加。与此同时,特征图变得越来越粗糙,小面积水稻田块的纹理信息大量丢失。一种简单的方法是同时减少下采样次数与网络层数,尽可能地保留小水稻田块的纹理信息,但这会丢失大量深层特征。另一种方式是只减少下采样次数,不改变网络层数,在保留深层特征的同时获得更大的特征图尺寸,但是会产生巨大的计算代价。考虑到浅层特征中含有丰富的纹理信息,可以将浅层特征与深层特征进行融合,让模型获得更加全面的特征感知能力。YANG 等[22]提出了一种多尺度特征融合方法(DFF),该方法通过自适应平均池化捕获全局信息,并利用注意力机制动态地选择和融合多尺度特征。该方法中单一尺度的池化操作无法处理不同尺寸的水稻田块,且在特征融合时未能充分考虑网络信息的流动性和梯度消失问题。为此,提出一种自适应多尺度注意力特征融合(AMSADFF)模块,用于增强网络对不同尺寸小面积水稻田块的感知能力,如图 5 所示。

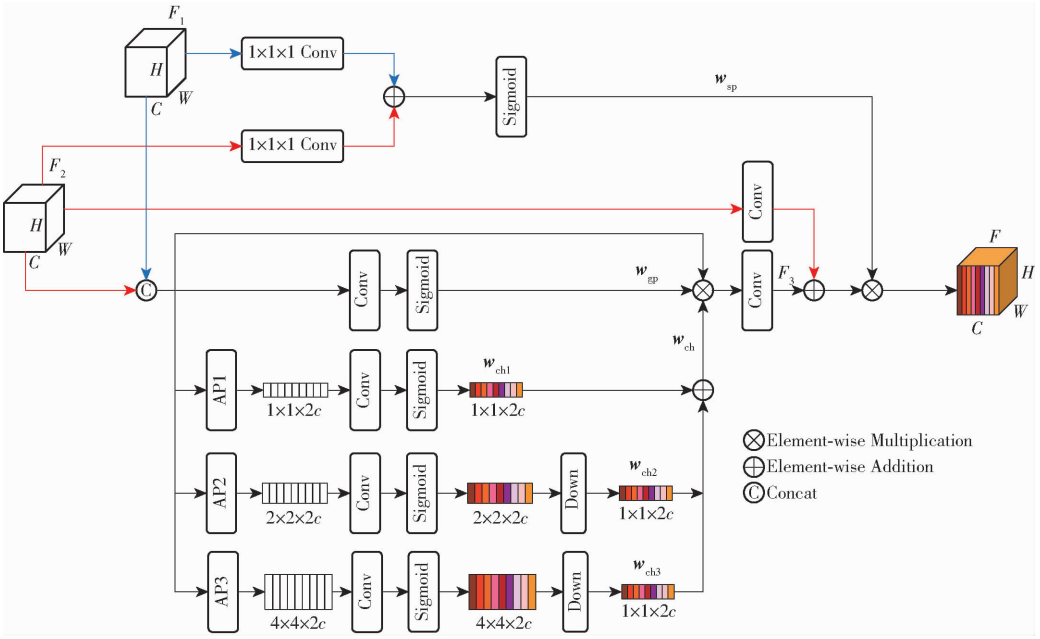



图 5 AMSADFF 模块结构图

Fig. 5 Structure diagram of AMSADFF module

AMSADFF 模块的核心思想在于通过多尺度注意力机制与门控机制协同建模,并结合残差连接机制,在融合浅层纹理特征和高层语义特征的同时,动态选择重要特征区域并抑制冗余特征。

具体来说,浅层特征 $F_1$ 与深层特征 $F_2$ 沿着通道进行拼接。为了提取遥感影像中的不同尺寸水稻田块信息,引入了 3 种不同尺度(1×1、2×2、4×4)的平均池化操作(Avg),并依次通过卷积层(Conv)和 Sigmoid 函数($\sigma$)来得到对应的权重矩阵 $\boldsymbol{w}_{ch1}$、$\boldsymbol{w}_{ch2}$、$\boldsymbol{w}_{ch3}$,即

$$\boldsymbol{w}_{chi}=\sigma(\mathrm{Conv}(\mathrm{Avg}_i([F_1;F_2]))) \tag{3}$$

为了方便后续融合,通过下采样函数($\xi$)将 $\boldsymbol{w}_{ch2}$、$\boldsymbol{w}_{ch3}$ 的尺寸调整至与 $\boldsymbol{w}_{ch1}$ 保持一致,并将三者相加融合为包含不同尺度信息的总权重矩阵 $\boldsymbol{w}_{ch}$ 来增强模型对不同尺寸水稻田块的感知能力,即

$$\boldsymbol{w}_{ch}=\boldsymbol{w}_{ch1}+\xi(\boldsymbol{w}_{ch2})+\xi(\boldsymbol{w}_{ch3}) \tag{4}$$

为了进一步提升特征选择的精度,设计了一种轻量化的门控结构来控制信息的流动。通过 1×1 卷积和 Sigmoid 函数生成门控权重矩阵 $\boldsymbol{w}_{gp}$ 来调整总权重矩阵 $\boldsymbol{w}_{ch}$,将调整后权重矩阵用于拼接后的特征上,即

$$\boldsymbol{w}_{gp}=\sigma(\mathrm{Conv}([F_1;F_2])) \tag{5}$$

随后利用 1×1 卷积对特征通道进行筛选，保留重要的特征，同时丢弃信息量较小的特征，得到筛选后的特征 $F_3$，即

$$F_3 = \boldsymbol{w}_{ch}\boldsymbol{w}_{gp}[F_1; F_2] \tag{6}$$

为了缓解深层网络中可能出现的梯度消失问题，并保留输入特征 $F_2$ 中的结构信息，在通道筛选后引入残差连接。具体而言，使用 1×1 卷积对原始输入特征 $F_2$ 进行映射获得残差信息，与通道筛选后的特征图相加融合，再使用全局空间信息权重 $\boldsymbol{w}_{sp}$ 对融合特征进行调整，生成最终的特征图 $F$，即

$$F = [F_3 + \mathrm{Conv}(F_2)]\boldsymbol{w}_{sp} \tag{7}$$

### 1.5　评价体系

采用分割网络中常用的评价指标来量化提出算法的分割性能，包括 Dice 系数（Dice）、平均交并比（mIoU）和准确率（Acc）。为确保指标能够反映模型在充分收敛状态下的稳定性能，所有指标均基于训练过程后 20 轮的结果取平均值进行统计。此 3 个指标的数值均在区间[0,1]内，数值越大表示模型分割性能越优。

## 2　结果与分析

### 2.1　模型对比

对所提出的 DFAU－Net 模型与其他经典模型（包括 SegNet[23]、DeepLabv3 +[24]、U－Net[25]、U－Net + +[26]、Attention U－Net[27]）在 GF2RiceSeg 数据集上进行可视化对比分析，各模型的评价指标迭代曲线如图 6 所示。与未改进的 U－Net 及其一系列变体相比，DFAU－Net 能够更早收敛并保持稳定，且效果最优。

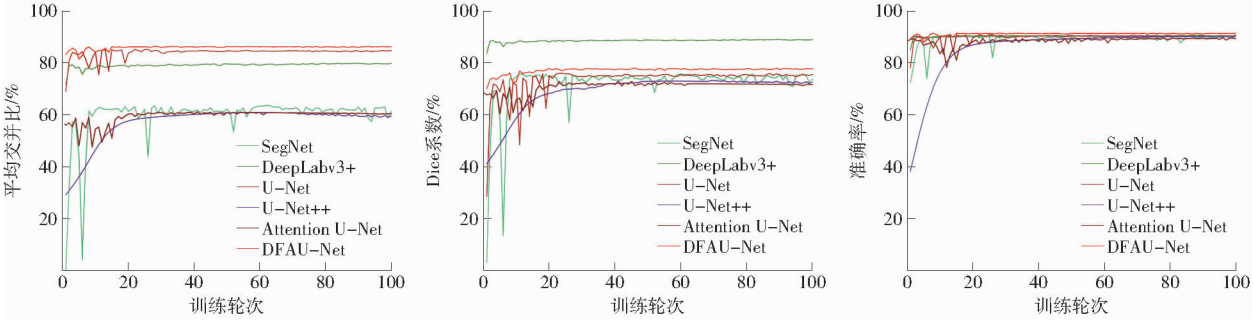



图 6　各模型评价指标迭代曲线

Fig. 6　Iteration curve of evaluation metrics for each model

为确保模型均收敛，取各项评价指标第 81 至 100 轮的值进行统计并计算均值，结果如表 1 所示，所提出的 DFAU－Net 在 GF2RiceSeg 数据集上性能最佳。与经典方法中效果最好的 U－Net 相比，本文方法在测试时 Dice 系数、平均交并比、准确率分别提高 2.03、1.64、0.97 个百分点。

**表 1　各类模型性能对比结果**

**Tab. 1　Comparison of performance results for different models**　%

| 模型名称 | Dice 系数 | 平均交并比 | 准确率 |
|---|---|---|---|
| SegNet | 74.06 | 60.76 | 90.40 |
| DeepLabv3 + | 88.82 | 79.89 | 90.68 |
| U－Net | 75.71 | 84.70 | 90.51 |
| U－Net + + | 73.26 | 61.49 | 88.94 |
| Attention U－Net | 69.40 | 57.46 | 88.61 |
| DFAU－Net | 77.54 | 86.34 | 91.48 |

各模型分割效果如图 7 所示，第 1 列为原始遥感图像，第 2 列为对应的真实标签，3～8 列依次为 SegNet、DeepLabv3 + 、U－Net、U－Net + + 、Attention U－Net 以及 DFAU－Net 的预测结果。在图 7a 中，U－Net、Attention U－Net 在水稻田与林地交界处出现误分现象，不能很好地区分边界，DeepLabv3 + 出现了漏分现象，这是导致结果较差的原因。在图 7b 中，由于河流和水稻田块在 RGB 影像上区分不明显，除 DFAU－Net 外的其他模型均将河流错分为水稻田块。在图 7c 中，由于成像受到光线影响，部分温室颜色较暗，除 DFAU－Net 外的其他模型均将温室误分为水稻田块。本文提出的 DFAU－Net 在 3 种情况下，均能够较好地分割出水稻田块，并且不受相似像元的干扰。

### 2.2　消融实验

为验证 DFAU－Net 中各关键模块的有效性，在 GF2RiceSeg 数据集上基于基准模型 U－Net（Base）设计了一系列消融实验，并对典型样本的分割结果进行可视化展示。具体地，在基准模型中逐步引入 LPA 模块（记作 A）、AMSADFF 模块（记作 B′）以及双分支编码器架构（记作 C），以评估其对分割性能的贡献。全部模型均训练 100 轮，结果取最后 20 轮的平均性能指标进行对比，评价指标包括 Dice 系数、平均交并比和准确率，实验结果如表 2 所示。

从结果来看，引入 LPA 模块（模型 Base + A）后平均交并比和准确率分别提高 1.13、0.68 个百分

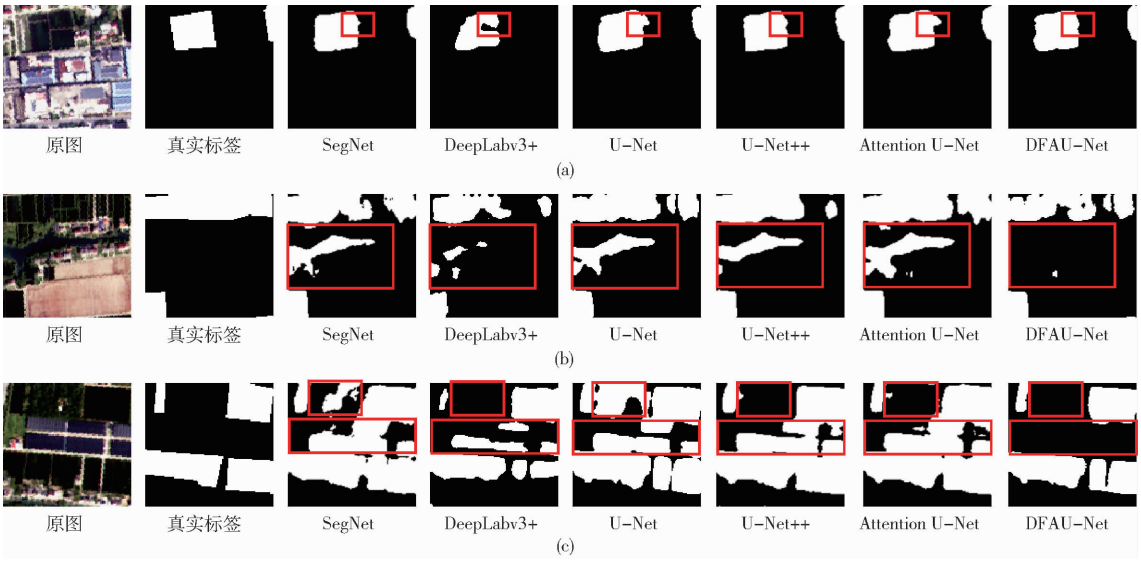



图 7 不同方法结果对比图

Fig. 7 Comparison chart of results from different methods

**表 2 消融实验结果**

**Tab. 2 Ablation experiment results** %

| 模型名称 | Dice 系数 | 平均交并比 | 准确率 |
|---|---|---|---|
| Base | 75.71 | 84.70 | 90.51 |
| Base + A | 74.53 | 85.83 | 91.19 |
| Base + B | 76.20 | 84.77 | 90.43 |
| Base + B′ | 75.45 | 85.62 | 91.08 |
| Base + C | 76.67 | 85.69 | 91.05 |
| Base + A + B′ | 75.43 | 86.04 | 91.29 |
| Base + A + B′ + C | 77.54 | 86.34 | 91.48 |

点,这表明 LPA 模块有助于网络更有效地聚焦于水稻田块区域,对非目标区域的干扰有抑制作用,如图 8 所示。

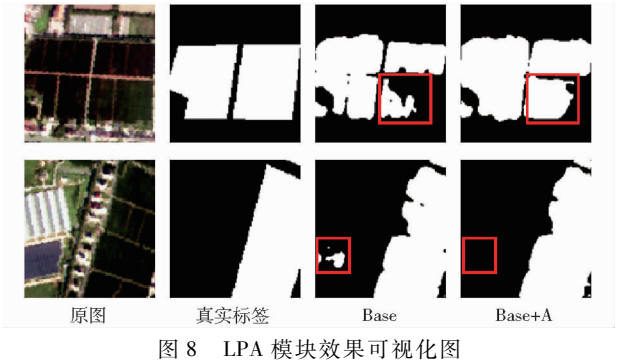



图 8 LPA 模块效果可视化图

Fig. 8 Visualization diagram of LPA module effect

引入 DFF 模块(模型 Base + B)时,模型性能仅略高于基线,说明该模块在处理多尺度目标时仍存在局限。而将 DFF 改进为 AMSADFF(模型 Base + B′)后,模型在平均交并比和准确率方面均有提升,分别升高 0.92、0.57 个百分点。模型能够捕捉到 Base 不能识别的小面积水稻田块,如图 9 所示。说明引入多尺度感知与门控机制可以更好地整合浅层纹理与深层语义信息,增强模型对小目标水稻田块的识别能力。

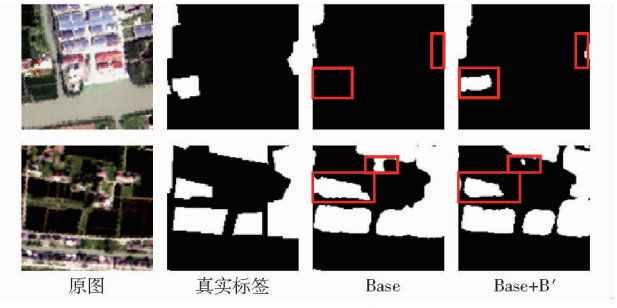



图 9 AMSADFF 模块效果可视化

Fig. 9 Visualization of AMSADFF module effect

引入双分支架构(模型 Base + C)进一步提高了模型性能,其中 Dice 系数提升至 76.67%、平均交并比提升至 85.69%。模型能够较好地识别出 Base 无法识别或者错误识别的与水稻田特征相似的区域,如图 10 所示。这一结果表明,融合 RGB 与 NRG 影像能够有效提高模型对非水稻田块目标的排除能力。

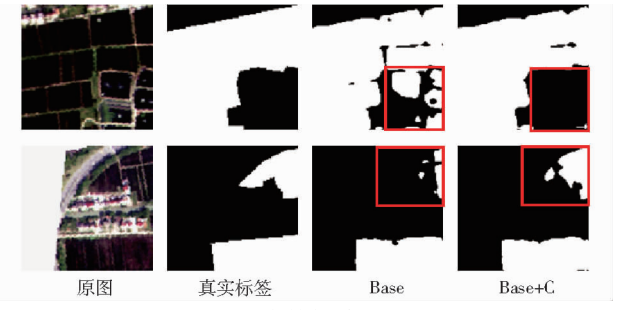



图 10 双分支结构效果可视化图

Fig. 10 Visualization diagram of double-branch structure effect

当联合引入 LPA 与 AMSADFF 模块(模型 Base + A + B′)时,性能进一步提升。在三者同时引入后(模型 Base + A + B′ + C,即 DFAU - Net),模型在各

项指标上均取得最优表现，验证了三者之间具有良好的互补性，协同提升了模型对小面积水稻田块的分割精度。综合定量与可视化结果可见，所提出的 DFAU－Net 在应对复杂背景、小目标识别与模态差异等方面具备优越的鲁棒性。

## 3　讨论

DFAU－Net 的双分支编码器架构分别使用 RGB 影像和 NRG 假彩色影像作为输入，能充分利用不同波段间的互补特性。RGB 影像提供了丰富的空间特征和纹理信息，而 NRG 假彩色影像则融入了近红外波段的光谱特征，能够增强水稻与背景之间的光谱差异。这种设计克服了线性光谱组合无法挖掘波段之间的非线性关系的困难，同时解决了仅依赖多波段堆叠的输入所带来的冗余信息问题，使模型能够更快更准地捕捉到水稻与背景（如道路、建筑物、水体等）之间的特征差异。将双分支编码器提取到的 2 种影像特征进行融合，并利用 LPA 模块的空间与通道注意力机制对特征进行筛选，保留关键特征，使得模型更加关注影像中的水稻田块区域。传统模型采取逐层下采样的结构来增加特征数量，代价是牺牲图像的分辨率。对于大面积田块而言，这种分辨率的损失在一定程度上是可以接受的，因为它有助于提高模型的训练和推理效率。然而，对于小面积的水稻田块而言，影像分辨率的降低会导致重要边缘和纹理信息丧失，从而显著影响分割精度。AMSADFF 模块通过融合低层的纹理特征和高层的语义特征，能够恢复丢失的细节信息。此外，多尺度池化操作的引入使得模型能够在不同尺度上捕捉到下采样过程中不同分辨率特征图上水稻田块的变化，从而提升了模型对小面积水稻田块的感知能力。基于这些原因，DFAU－Net 的分割性能得到了显著提升。

由于水稻遥感影像数据样本稀缺，尽管近期兴起的 Transformer[28]、Mamba[29]、Segment Anything Model（SAM）[30] 以及基于它们的变体 ViT[31]、VMamba[32]、RSPS－SAM[33] 等模型具备强大的特征提取能力，但它们依赖于大规模标注样本，在高分辨率样本稀少的情况下，难以充分发挥其优势。未来工作将进一步探索生成对抗网络（GAN）等深度生成模型在高分辨率遥感影像样本合成与数据增强中的应用，构建多样化的模拟样本以丰富水稻遥感影像数据集。通过引入高质量伪影像与标签，有望缓解样本稀缺对模型训练的制约，为复杂场景下小面积水稻田块的精准识别提供更充分的训练支撑。

## 4　结论

（1）为弥补针对小面积、零散分布的水稻田块分割领域的研究空白，本文充分利用 GF－2 多光谱数据，提出了一种基于 U－Net 的改进分割网络 DFAU－Net。

（2）基于 GF－2 多光谱影像与多源辅助资料，制作了高质量像素级分割标签，构建了包含 848 幅图像样本的 GF2RiceSeg 数据集，为相关研究提供了可复用的实验基础。

（3）验证了近红外波段在提升模型对水稻田块与背景区分能力的有效性。通过构建 RGB 与 NRG 假彩色影像的双分支编码器输入架构，模型在低对比度条件下对水稻田块特征可区分性得到增强。

（4）设计的自适应多尺度注意力特征融合模块通过多尺度池化与门控机制协同作用，能够动态选择并融合不同尺度水稻特征，有效缓解多次连续下采样过程中小面积水稻田块纹理信息丢失的问题，提升了模型对多尺度水稻田块的感知能力。

（5）验证了局部金字塔注意力模块的有效性。通过空间与通道注意力机制协同作用，可以引导网络在特征提取过程中重点关注与水稻田块最相关的区域和通道信息，显著抑制道路、建筑等非目标区域干扰。

（6）对比实验结果表明，DFAU－Net 的 Dice 系数、平均交并比和准确率分别达到 77.54%、86.34% 和 91.48%，显著优于 U－Net、DeepLabv3＋、Attention U－Net 等经典分割网络，充分验证了所提出方法的有效性。

## 参　考　文　献


[1] ZHENG J, WANG W, LIU G, et al. Towards quantification of the national water footprint in rice production of China: a first assessment from the perspectives of single-double rice[J]. Science of the Total Environment, 2020, 739: 140032.

[2] ZHAO Z, DONG J, YANG J, et al. From rice planting area mapping to rice agricultural system mapping: a holistic remote sensing framework for understanding China's complex rice systems[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2025, 224: 382－397.

[3] BEM P P, CARVALHO O A, CARVALHO O L F, et al. Irrigated rice crop identification in Southern Brazil using convolutional neural networks and Sentinel－1 time series[J]. Remote Sensing Applications: Society and Environment, 2021, 24: 100627.

[4] XUAN F, DONG Y, LI J, et al. Mapping crop type in Northeast China during 2013—2021 using automatic sampling and tile-based image classification[J]. International Journal of Applied Earth Observation and Geoinformation, 2023, 117: 103178.

[5] LUINTEL N, MA W, MA Y, et al. Tracking the dynamics of paddy rice cultivation practice through MODIS time series and

phenorice algorithm[J]. Agricultural and Forest Meteorology, 2021, 307: 108538.
[6] ZHANG J, GAO L, LIU M, et al. Accurate paddy rice mapping based on phenology-based features and object-based classification[J]. Remote Sensing, 2024, 16(23): 4406.
[7] 罗锡文,谷秀艳,胡炼,等. 大田无人化智慧农场农田边界识别技术研究现状与展望[J]. 农业机械学报,2025,56(2):1-18.
LUO Xiwen, GU Xiuyan, HU Lian, et al. Research status and outlook of farmland boundary recognition technology in large-scale unmanned smart farms[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(2):1-18. (in Chinese)
[8] THORP K R, DRAJAT D. Deep machine learning with Sentinel satellite data to map paddy rice production stages across West Java, Indonesia[J]. Remote Sensing of Environment, 2021, 265: 112679.
[9] TU Y, LANG W, YU L, et al. Improved mapping results of 10 m resolution land cover classification in Guangdong, China using multisource remote sensing data with Google Earth Engine[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 5384-5397.
[10] ARJASAKUSUMA S, SWAHYU K S, RAFIF R, et al. Combination of Landsat 8 OLI and Sentinel-1 SAR time-series data for mapping paddy fields in parts of West and Central Java Provinces, Indonesia[J]. ISPRS International Journal of Geo-Information, 2020, 9(11): 663.
[11] HUANG C, YOU S, LIU A, et al. High-resolution national-scale mapping of paddy rice based on sentinel-1/2 data[J]. Remote Sensing, 2023, 15(16): 4055.
[12] KANG J, YANG X, WANG Z, et al. Collaborative extraction of paddy planting areas with multi-source information based on google earth engine: a case study of Cambodia[J]. Remote sensing, 2022, 14(8): 1823.
[13] WEI P, CHAI D, LIN T, et al. Large-scale rice mapping under different years based on time-series Sentinel-1 images using deep semantic segmentation model[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 174: 198-214.
[14] DU M, HUANG J, WEI P, et al. Dynamic mapping of paddy rice using multi-temporal landsat data based on a deep semantic segmentation model[J]. Agronomy, 2022, 12(7): 1583.
[15] JO H W, PARK E, SITOKONSTANTINOU V, et al. Recurrent U-Net based dynamic paddy rice mapping in South Korea with enhanced data compatibility to support agricultural decision making[J]. GIScience & Remote Sensing, 2023, 60(1): 2206539.
[16] ONOJEGHUO A O, MIAO Y, BLACKBURN G A. Deep ResU-Net convolutional neural networks segmentation for smallholder paddy rice mapping using Sentinel 1 SAR and Sentinel 2 optical imagery[J]. Remote Sensing, 2023, 15(6): 1517.
[17] XIA L, ZHAO F, CHEN J, et al. A full resolution deep learning network for paddy rice mapping using Landsat data[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2022, 194: 91-107.
[18] AFAQ Y, MANOCHA A. Multi-resolution-based deep learning approach for rice field monitoring[J]. Canadian Journal of Remote Sensing, 2022, 48(2): 278-298.
[19] DU H, WANG J, LIU M, et al. SwinPA-Net: swin transformer-based multiscale feature pyramid aggregation network for medical image segmentation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 35(4): 5355-5366.
[20] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
[21] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition., 2018: 7132-7141.
[22] YANG J, QIU P, ZHANG Y, et al. D-Net: dynamic large kernel with dynamic feature fusion for volumetric medical image segmentation[J]. arXiv Preprint, arXiv:2403.10674, 2024.
[23] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Segnet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
[24] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 801-818.
[25] RONNEBERGER O, FISCHER P, BROX T, et al. U-Net: convolutional networks for biomedical image segmentation[C]// Medical image Computing and Computer-assisted Intervention-MICCAI 2015: 18th International Conference. Munich, 2015: 234-241.
[26] ZHOU Z, RAHMAN S M, TAJBAKHSH N, et al. Unet++: a nested U-Net architecture for medical image segmentation [C]// Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA, Granada, 2018: 3-11.
[27] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-Net: learning where to look for the pancreas[J]. arXiv Preprint, arXiv:1804.03999, 2018.
[28] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30:5998-6008.
[29] GU A, DAO T. Mamba: linear-time sequence modeling with selective state spaces[J]. arXiv Preprint, arXiv:2312.00752, 2023.
[30] KIRILLOV A, MINTUN E, RAVI N, et al. Segment anything[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023: 4015-4026.
[31] ZHENG S, LU J, ZHAO H, et al. Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 6881-6890.
[32] LIU Y, TIAN Y, ZHAO Y, et al. Vmamba: visual state space model[J]. Advances in Neural Information Processing Systems, 2024, 37: 103031-103063.
[33] LIU Z, LI Z, LIANG Y, et al. RSPS-SAM: a remote sensing image panoptic segmentation method based on sam[J]. Remote Sensing, 2024, 16(21): 4002.