

# 基于视觉基础模型与知识融合的地块级水稻分布遥感提取方法

彭子萌<sup>1</sup>, 段玉林<sup>1</sup>, 余强毅<sup>1</sup>, 吴文斌<sup>1</sup>, 张帅<sup>2</sup>, 赵春雷<sup>1</sup>, 李伯亮<sup>3</sup>,  
张新<sup>3</sup>

1. 中国农业科学院农业资源与农业区划研究所/北方干旱半干旱耕地高效利用全国重点实验室/农业农村部农业遥感重点实验室, 北京 100081;
2. 农业农村部农田建设管理司, 北京 100125;
3. 铁塔智联技术有限公司, 北京 100089

**摘要:** 尽管深度学习方法在地块级作物分布遥感精细提取方面表现出较好效果, 但其高度依赖大量高质量的标注样本, 存在成本高昂、时效性差及跨区域泛化能力不足等瓶颈。为此, 本文提出一种结合视觉基础模型与知识融合的地块分割提示信息优化方法, 将水稻物候特征和光谱指数通过自适应迭代学习转化为视觉基础模型所需的动态提示信息, 有效解决了传统深度学习方法对样本高度依赖的问题。具体的, 首先利用植被指数先验知识剔除非植被区域, 以减少计算量。同时, 针对SAM初步分割结果进行逐地块统计, 更新先验知识阈值范围, 并作为地块分割提示信息二次输入SAM模型, 通过上述迭代学习的方式, 在更新先验知识的同时, 得到最佳的地块级分类识别结果。最后, 本方法创新性提出了采用IoU指标作为迭代分割终止的条件, 构建了兼具精度与效率的自适应学习闭环, 在实现对分割结果稳定性量化监控的同时, 确保模型在达到最优解时自动终止。实验结果表明, 水稻收割期为水稻提取最佳窗口期; 基于该最优时相, 本方法在天津宁河区、黑龙江富锦市、日本新潟市等不同实验区域的制图Kappa系数分别为0.89、0.91和0.86。此外, 与U-Net、DeepLabV3+等典型监督学习模型相比, 本方法在无需针对研究区进行特定模型训练的前提下, 取得了可与监督学习模型媲美甚至更高的提取精度, 充分证实了本文方法的有效性和区域泛化能力, 为实现大范围、低成本、高精度的自动化作物制图提供了新的解决方案。

**关键词:** 遥感精细提取, SAM, 先验知识融合, 水稻制图, 自适应统计学习, 地块级分割, 迭代优化, 提示信息优化

**中图分类号:** P2

**引用格式:** 彭子萌, 段玉林, 余强毅, 吴文斌, 张帅, 赵春雷, 李伯亮, 张新. XXXX. 基于视觉基础模型与知识融合的地块级水稻分布遥感提取方法. 遥感学报, XX(XX): 1-19

PENG Zimeng, DUAN Yuling, YU Qiangyi, Wu Wenbin, Zhang Shuai, Zhao Chunlei, Li Boliang, Zhang Xin. XXXX. A remote sensing method for plot-level rice distribution extraction based on visual foundation models and knowledge fusion. National Remote Sensing Bulletin, DOI:10.11834/jrs.20265396]

## 1 引言

卫星遥感为土地覆盖、作物分布信息提取提供了有力的技术支持。通过利用光谱、物候等特征, 可有效区分像素之间的差异, 由此发展形成了基于像元的遥感信息提取方法 (Tucker, 1979),

目前已经广泛应用于区域、国家乃至全球尺度土地覆盖、作物分布遥感信息提取与制图, WorldCereal (Van Trient 等, 2023) 和 Cropland Data Layer (Bryan 等, 2011) 等数据产品为粮食安全、可持续发展等方面的研究提供了重要的基础数据支持。

收稿日期: 2025-09-25; 预印本: XXXX-XX-XX

基金项目: 国家自然科学基金创新研究群体项目“农业遥感机理与方法”(41921001); 北京市智慧农业创新团队(BAIC10-2025)

第一作者简介: 彭子萌, 研究方向为农业遥感监测。E-mail: 821012530348@caas.cn

通信作者简介: 段玉林, 研究方向为智慧农业。E-mail: duanyulin@caas.cn

地块级作物分布精细提取是开展智慧化田间管理、长势监测与产量评估的重要基础。然而,传统基于像元的方法容易产生“椒盐效应”(Blaschke, 2010),从而影响作物分布遥感制图精度。面向对象的方法通过影像分割等过程,较好地保持了对象内部像素信息的一致性,并突出了对象之间像素信息的差异性,为地块级作物分布精细提取提供了较好技术手段(Hossain等, 2019)。近年来,U-Net、DeepLabV3+等深度语义分割网络的应用,进一步提升了作物遥感分类精度(Chen等, 2018)。然而,现有方法大多依赖大量高质量样本作为训练支撑,其数量与分布直接影响模型的精度与稳定性,限制了方法的区域泛化能力(Waldner等, 2015)。因此,如何在保证精度的同时减少对样本的依赖,提升模型的泛化能力,已成为当前作物遥感精细制图面临的核心挑战。

在主要粮食作物中,水稻遥感精细制图是检验方法稳健性与泛化能力的典型场景。一方面,水稻光谱特征与湿地水体高度相似(Bazzi等, 2019);另一方面,亚洲地区水稻主产区普遍存在地块破碎、经营分散等特征(Fritz等, 2015),增加了样本采集难度与成本。为实现地块级水稻精准提取,现有研究多采用监督学习模型(Zhao等, 2021),结合多源时序遥感影像(Onojeghuo等, 2018)及面向对象的多尺度分割方法(Ding等, 2020; Zhang等, 2023)进行识别。尽管上述方法在局部区域内表现出较高精度,但模型在不同区域易受样本差异影响而产生精度波动(You等, 2021)。为减轻对训练样本的依赖,研究者进一步探索了迁移学习(Hu等, 2022)、弱监督学习(许晴等, 2022)及基于光谱或物候特征的指数判别方法(Chen等, 2023; Zhao等, 2024),但这些方法仍普遍存在稳定性不足和区域泛化能力较弱等问题。

通用视觉基础模型(Segment Anything Model, SAM)为图像分割提供了高效手段(Kirillov等, 2023),在遥感影像分割中受到广泛关注(Gui等, 2024; Janowski等, 2024)。其在无需针对具体任务进行模型训练的情况下,能够生成质量较高的分割结果,为降低遥感应用中的人工标注样本的依赖提供了新的可能。现有研究尝试探索SAM在农业场景中的适应性:郝雪丽等

(2023)、Huang等(2024)将多尺度分割与SAM融合,用于耕地边界提取,在地块边界表达方面取得一定改进,但仍依赖固定规则或单次提示,缺乏依据分割反馈的自适应更新。Chen等(2024)提出RSPrompter方法,通过提示学习改进分割精度,在实例分割任务中表现出较好效果,但其对包含时序信息的水稻遥感高精度提取的适用性未曾验证。Liu等(2025)提出的提示点优化方法主要面向自然图像,其在地块形态复杂、时序变化显著的水稻场景中仍缺乏系统检验。

针对水稻地块级制图的需求,现有基于SAM的研究仍存在三方面局限:一是多依赖静态阈值与单次提示,缺乏基于分割反馈进行动态调整的闭环机制;二是对水稻时序物候特征的利用不足,尚未系统评估模型在不同生育期的分割潜力;三是提示点优化策略在水稻遥感场景中的有效性仍有待验证。因此,如何将SAM的通用分割能力与水稻提取有效结合,构建具备自适应与迭代优化能力的地块级分割框架,是实现高精度、自动化水稻遥感精细化制图亟需解决的关键问题。

为此,本文提出一种融合视觉基础模型与农业先验知识的地块级水稻提取框架。该框架以SAM为核心,构建“分割—筛选—知识更新—再提示”的自适应迭代学习闭环,通过分析高置信度样本的统计特征反向优化阈值,实现从“静态规则”到“动态自适应”的转变,并融合时序物候先验知识约束与更新关键判别阈值。提示环节中,在高置信度分割掩膜的基础上自动生成提示点,通过分割反馈动态调整提示点位置与分布,以增强模型对复杂地块边界识别的稳定性。同时,结合水稻的时序光谱特征,系统对比不同物候期的提取效果,科学选取最优提取窗口。本文提出阈值由静态到自适应、提示由单次到迭代、时相由经验到优选等优化策略,探索了大范围、地块级水稻遥感制图新的技术路径。

## 2 研究区概况与研究数据

### 2.1 研究区概况

本研究选择三个具有代表性水稻种植区,以评估方法的稳健性与泛化能力。天津市宁河区作为核心试验区,是华北平原的典型单季稻集中产区。该区水稻种植制度规律,但地块普遍破碎,

且水稻光谱易与湿地、旱作物等地物混淆，为验证方法的基础性能与算法调试提供了典型场景。为进一步测试方法的泛化能力，选取黑龙江省富锦市和日本新潟市作为验证区。富锦市位于中国三江平原，代表大规模集约化种植模式，区域内地块规模大、形态规整且连片分布。新潟市则代

表东亚精细化农耕模式，其地块级水稻普遍小且形状不规则，与建成区、水网等非农地物交错，形成高度破碎化的农业景观。选取在这三种不同类型的农业景观中进行验证，旨在系统评估本方法的在不同区域的适用性。

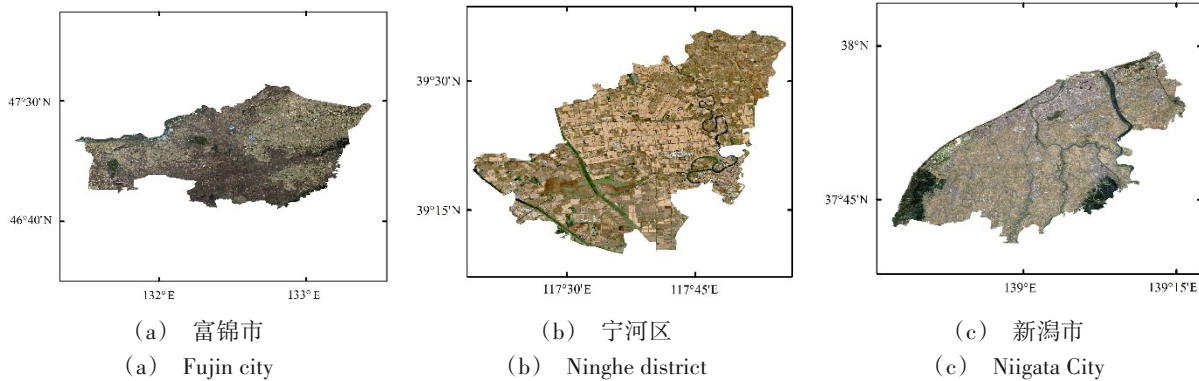


图1 研究区区位图

Fig.1 Location map of the study area

## 2.2 数据预处理

使用 Harmonized Sentinel-2 MSI; Multispectral Instrument, Level-2A 遥感数据产品，该数据已经完成了大气校正、几何校正等预处理。在影像筛选阶段，我们以单景影像的云量小于 20% 作为阈值，并利用 QA60 波段中第 10 位和第 11 位的云与卷云标识进行掩膜去除，同时按尺度因子进行反射率归一化。用  $B_4$  (红)、 $B_3$  (绿)、 $B_2$  (蓝) 波段合成 RGB 影像，并对 RGB 影像做中值滤波处理。采用波段  $B4$  (红)、 $B8$  (近红外) 根据归一化植被指数 (Normalized Difference Vegetation Index, NDVI) 计算公式计算 NDVI 指数，并将地表水体指数 (Land Surface Water Index, LSWI) 中  $B11$  (短波红外) 波段采用双线性插值重采样至 10m，通过波段  $B8$  (近红外)、 $B11$  (短波红外) 计算 LSWI 指数。对于宁河区，选取水稻关键物候

期的单景 Sentinel-2 遥感影像，未进行跨日多景影像的拼接。而对于富锦市，由于研究范围较广，单景影像无法覆盖整个研究区，因而选取同一成像周期内云量低于 20% 的多景 Sentinel-2 遥感影像，在完成 QA60 云与卷云掩膜后，基于影像重叠区域采用中值合成方式进行拼接。

为应对不同区域水稻物候的显著差异，本研究构建了覆盖其完整生育周期的物候时间窗，涵盖从移栽到收割的各个关键阶段，并以每月上、中、下旬为基本时间单元组织和分析时序影像 (图 2)。依据该策略，针对宁河区，我们在关键物候窗口内筛选出 11 个具有代表性时相影像 (具体时相及精度评估详见 4.2 节及图 9) 进行精细分析；同时，针对富锦市与新潟市，则结合当地事历，在关键窗口期内获取遥感影像。

研究区	4月			5月			6月			7月			8月			9月			10月			11月		
	上旬	中旬	下旬	上旬	中旬	下旬	上旬	中旬	下旬	上旬	中旬	下旬	上旬	中旬	下旬	上旬	中旬	下旬	上旬	中旬	下旬	上旬	中旬	下旬
宁河区																								
富锦市																								
新潟市																								

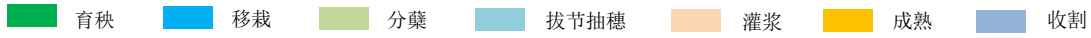


图2 研究区水稻物候历

Fig.2 Rice phenological calendar in the study areas

### 3 研究方法

为减少对高质量人工标注样本的依赖，本文构建了基于视觉基础模型与知识融合的地块级水稻分布遥感精细提取框架（图3）。该框架以自适应迭代优化闭环为核心，通过将视觉通用分割模型（Segment Anything Model, SAM）与先验知识耦合，实现水稻地块级精细制图。

在该框架中，知识融合主要体现在光谱特征与几何约束两个方面。一方面，利用水稻淹水期的光谱特征（LSWI、NDVI）对SAM初始分割结果进行筛选，得到高置信度水稻地块，并将地块的中心点及边缘点作为点提示信息再次输入到SAM图像分割基础模型框架中，将先验知识与视觉基础模型进行了有效结合，使得基于通用分割模型也能实现对特定地物即水稻地块的自动提取。二是通过地块的几何特征约束（面积阈值 $<200\text{m}^2$ ），用于剔除水稻地块分割中的细小地块，进一步提高筛选出的水稻地块提取结果的置信度。

在此基础上，采取迭代优化机制，每一轮迭代首先针对先验知识对SAM的初始分割结果进行筛选，形成高置信度水稻地块；其次，通过两个环节持续优化分割结果，一方面通过高置信度水稻地块生成提示点信息反馈至SAM实现更精细的分割；另一方面通过提取高置信度水稻地块特征，输入到场景自适应学习模块，实现先验知识的更

新，最终实现了“分割—筛选—反馈—优化”的闭环迭代。

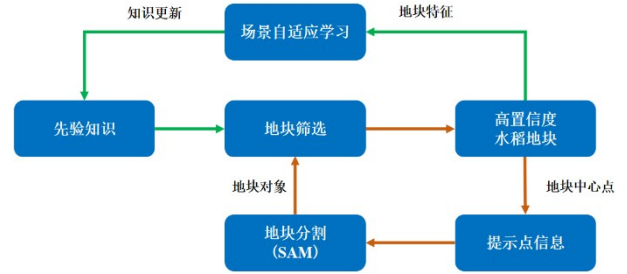


图3 技术流程图

Fig.3 Technical flowchart

#### 3.1 视觉通用分割模型(SAM)

本研究采用通用视觉基础模型（SAM）作为核心的图像分割引擎（Kirillov等，2023）。SAM是一个通用的、可提示的视觉基础模型，它能够基于用户输入，对图像中的任意对象生成高质量的分割掩码，而无需针对特定任务进行额外训练。

SAM的结构主要由三个组件构成（如图4所示）：（1）图像编码器（Image Encoder），采用视觉Transformer（ViT）架构，负责将输入影像提取为高维特征嵌入；（2）提示编码器（Prompt Encoder），将点、边界框等几何提示信息编码为向量；（3）掩码解码器（Mask Decoder），基于高效的Transformer解码器，它融合图像嵌入和提示嵌入，预测出与提示对应的对象掩码。

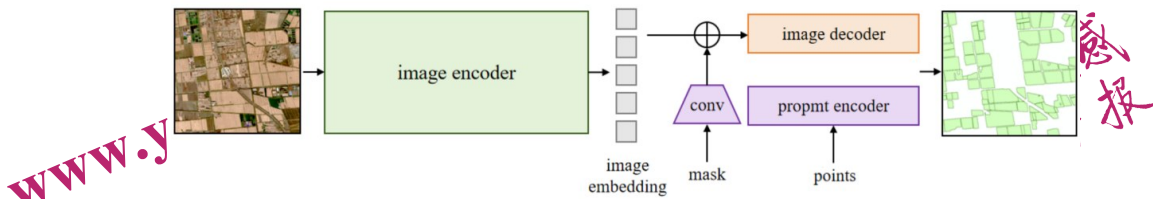


图4 SAM结构图

Fig.4 SAM structure diagram

其工作机制实现了“一次编码，多次解码”的模式，图像仅需通过一次耗时较长的编码器，便可响应多次、快速的提示解码请求。这种对提示信息高效、灵活的响应能力，是本研究构建迭

代优化框架的技术基础。

#### 3.2 植被区域掩膜生成

本研究利用多时相Sentinel-2影像计算归一化植被指数NDVI（Normalized Difference Vegetation

Index), 并通过阈值法生成植被区域掩膜。该方法融合植被生长的时间节律特征 (Han 等, 2022): 某个像元被判定为季节性植被, 需同时满足其在生长高峰期 NDVI 值较高, 而在作物生长初期或土地休耕期 NDVI 值较低。生成掩膜后, 利用植被掩膜对原始多时相影像进行裁剪, 仅保留植被生长区域, 从而有效减少了裸地、建设用地及水体等的干扰。同时, 通过生成植被区域掩膜缩小了后续处理的数据范围, 减少 SAM 模型进行全图分割时的计算负荷, 提升整个流程的运行效率。

### 3.3 基于 SAM 的地块初始分割

基于上述得到的植被掩膜图像, 利用 SAM 模型生成地块对象。在此阶段, 运用 SAM 进行无提示的图像分割, 让模型自动识别并分割出植被掩膜区域内具有独立边界的植被地块。

然而, 在无提示信息输入时, SAM 分割结果中通常存在边界锯齿、局部噪声及区域不连续等问题 (Osco 等, 2023)。若直接用于后续分析将导致水稻提取精度下降。为此, 本研究在初始结果的基础上增加了后处理流程: 首先, 对分割结果进行二值化处理, 并按照连通域面积排序以突出主要地块对象; 其次, 采用形态学运算对对象边界进行修正和平滑处理, 填补内部空洞并去除边缘毛刺; 最后, 通过边缘精修, 进一步提升几何一致性。将优化处理后的分割结果矢量化为地块级多边形数据, 为后续基于水稻时序光谱特征、面积阈值等关键先验知识开展的地块级水稻筛选工作, 奠定了坚实可靠的数据基础。

### 3.4 基于先验知识筛选高置信度水稻地块

#### (1) 水稻淹水期光谱特征

水稻在移栽初期保持浅水层, 是区别大多数旱地作物的典型特征。这一特征已成为遥感监测的关键 (Sakanoto 等, 2007; Teluguntla 等, 2015; Dong 等, 2016)。已有研究表明, 在此阶段, 地表水体指数 (Land Surface Water Index, LSWI) 均值通常高于归一化植被指数 (Normalized Difference Vegetation Index, NDVI) 均值, 同时 LSWI 均值维持在一定的上限范围内 (一般小于 0.45), 可有效区分水稻与旱作物 (Xiao 等, 2006; Chen 等, 2020; Han 等, 2021)。据此, 本研究在宁河区水稻移栽期采用地块均值约束为:  $LSWI_{均值} >$

$NDVI_{均值}$ 、 $LSWI_{均值} < 0.45$ 。初始阈值均来自于被广泛验证的经验范围, 并与上述研究结论一致, 因此即便初始设置存在一定偏差, 亦不致于出现由误筛选样本主导第一轮筛选。其中, LSWI 与 NDVI 的计算公式如下:

$$LSWI = \frac{NIR - SWIR1}{NIR + SWIR1} \quad (1)$$

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (2)$$

为验证上述依据并可视化水稻的阶段性光谱特征, 本研究基于初步筛选得到的高置信度地块级水稻开展“地块一时序”统计: 对每个地块在多时相 Sentinel-2 影像上计算 NDVI 与 LSWI 的像元均值, 形成地块级时序曲线; 随后, 对全部地块的 NDVI 时序曲线、LSWI 时序曲线分别执行逐时相平均操作, 并采用移动平均法平滑, 得到宁河区水稻的 NDVI 与 LSWI 时序变化曲线 (图 5)。

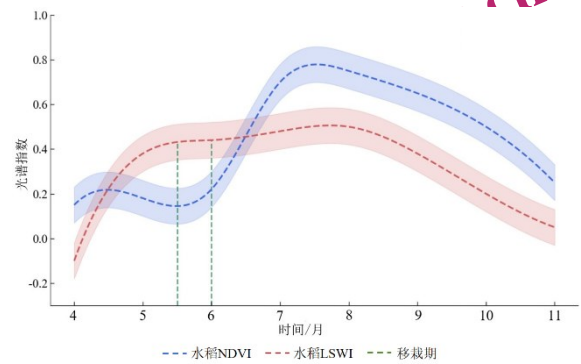


图5 宁河区水稻 NDVI 与 LSWI 时序曲线特征

Fig.5 Temporal characteristics of NDVI and LSWI for rice crops in Ninghe district

#### (2) 水稻旺盛期特征

为排除与水稻在移栽期表现相似的常年水体、湿地及季节性裸地, 本研究引入水稻生长旺盛期的 NDVI 指数特征作为补充条件。已有研究表明, 水稻在该时期的 NDVI 均值通常超过 0.4, 以反映高覆盖度与高叶绿素水平 (Qin 等, 2015; Zhang 等, 2015)。基于上述研究结论并结合宁河区的实测数据, 本研究中地块筛选条件设置为: 地块在旺盛期的  $NDVI_{均值} > 0.4$  且  $NDVI_{最小值} > 0$ , 以确保其具备旺盛的植被生长特征, 并过滤水体和湿地、季节性裸地等非水稻地物。

#### (3) 地块级水稻面积阈值

在水稻时序光谱特征的基础上, 为进一步过

滤噪声小斑块与异常合并图斑，本研究引入面积阈值：

$$\text{Valid}(r) = \begin{cases} 1, & S_{\min} \leq \text{Area}(r) \leq S_{\max} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

式中， $r$ 代表分割后的地块； $\text{Area}(r)$ 为该地块的面积函数； $S_{\min}$ 和 $S_{\max}$ 分别为地块面积的下限与上限阈值。 $\text{Valid}(r) = 1$ 表示保留该地块，0表示剔除。结合Sentinel-2遥感影像的10m分辨率与宁河区地块规模特征，本研究将 $S_{\min}$ 下限阈值设为200m<sup>2</sup>，以去除微小斑块，同时设置合理的上限值以识别由边界模糊引起的异常合并图斑。通过时序光谱特征与地块阈值特征，初步获得高置信度的地块级水稻集合（Boschetti等，2014；Qiu等，2015）。

### 3.5 先验知识阈值动态更新

在水稻时序光谱特征及面积阈值等先验知识初步筛选的基础上，将高置信度地块作为核心样本集，统计其关键特征分布，以95%置信区间自适应生成阈值，并在“提示—分割”的循环中动态更新，具体更新规则如下。

淹水期差值动态阈值：

$$\Delta = \text{LSWI} - \text{NDVI} \quad (4)$$

$$\tau_{\Delta,t} = \max(0, \mu_{\Delta,t} - 1.96\sigma_{\Delta,t}) \quad (5)$$

其中 $\mu_{\Delta,t}$ 和 $\sigma_{\Delta,t}$ 分别为第 $t$ 轮迭代核心样本集特征分布的均值与标准差。通过取动态阈值下界与0的最大值（即 $\max(0, \mu_{\Delta,t} - 1.96\sigma_{\Delta,t})$ ），在保证统计显著性的同时，满足 $\text{LSWI}_{\text{均值}} > \text{NDVI}_{\text{均值}}$ 的水稻淹水期光谱响应物理规律。

旺盛期NDVI动态阈值：

$$\tau_{\text{NDVI},t} = \mu_{\text{NDVI},t} \pm 1.96\sigma_{\text{NDVI},t} \quad (6)$$

地块面积动态阈值：

$$\tau_{\text{Area},t} = \mu_{\text{Area},t} \pm 1.96\sigma_{\text{Area},t} \quad (7)$$

式中， $\tau_{\text{NDVI},t}$ 与 $\tau_{\text{Area},t}$ 分别表示第 $t$ 轮迭代计算得到的旺盛期NDVI阈值区间与地块面积阈值区间； $\mu_{\text{NDVI},t}$ 与 $\sigma_{\text{NDVI},t}$ 为该轮核心样本集在旺盛期NDVI的均值与标准差； $\mu_{\text{Area},t}$ 与 $\sigma_{\text{Area},t}$ 为核心样本集地块面积的均值与标准差。系数1.96对应95%的置信区间。最终将筛选结果转化为提示信息反馈至SAM中，驱动其生成新的分割掩膜。本小节阐述了动态阈值更新机制的构建思路与实现流程。

### 3.6 高置信度水稻地块提示点生成

将筛选后的高置信度水稻掩膜高效转化为精

准的提示信息，是驱动SAM模型迭代优化的核心环节，并在后续的迭代框架中被循环调用。SAM支持点、框和掩膜三类提示，但已有研究表明，点提示在多轮迭代中具有更高的灵活性，能有效避免掩膜提示可能带来的过拟合问题（Li等，2023）。因此，本研究选择将每一轮迭代生成的高置信度掩膜转化为点集，作为SAM的主要输入。

正向提示点集由地块几何中心点 $P_c$ 与边界点集合 $P_b$ 共同构成。其中，中心点用于确保提示位于目标地块内部，提供水稻的核心信息；边界点则能传递地块形态和空间分布特征，以引导SAM在边缘区域完成更精细的轮廓描绘。

#### (1) 中心点提取

对于每个筛选后的高置信度掩膜区域，首先计算其几何中心点作为正向提示：

$$P_c = \left( \frac{1}{N} \sum_{i=1}^N x_i, \frac{1}{N} \sum_{i=1}^N y_i \right) \quad (8)$$

其中 $(x_i, y_i)$ 为掩膜像元坐标， $N$ 为掩膜内像元总数。该点能够代表整个地块的空间位置。

(2) 边界点提取为了增强提示信息的空间约束性，从掩膜的边界像元集合中均匀选取若干特征点：

$$P_b = \{ p_j \in \partial M \mid j = 1, 2, \dots, k \} \quad (9)$$

其中 $\partial M$ 表示掩膜的边界像元集合， $p_j$ 为选取的第 $j$ 个边界特征点矢量， $k$ 为边界点总数。该策略在保持整体形态的同时，能有效传递边界几何特征，从而提升分割结果的完整性与边界精度。

此外，非植被区域在前期分割中已通过阈值掩膜处理并过滤。为保证提示信息的一致性，本研究仅将其作为少量负向提示点输入SAM，用于抑制非目标对象。

### 3.7 水稻地块分割迭代优化

为修正筛选结果中存在的边界误差、内部空洞及地块粘连等问题，并补充因被遗漏的真实水稻地块，本研究构建了“知识更新—掩膜筛选—迭代优化”的分割框架。该框架通过一个动态反馈闭环，使分割结果的完整性与准确性在迭代中逐步提升，直至收敛。

在第 $t$ 轮迭代中，SAM利用输入影像 $I$ 和提示集合 $P_t$ 进行分割，并结合先验知识参数集 $\Theta_t$ 的约束，生成新的筛选掩膜集合：

$$M_{t+1} = \text{SAM}(I, P_t) \cap \text{Filter}(\Theta_t) \quad (10)$$

其中,  $\text{Filter}(\Theta_t)$ 表示基于第 $t$ 轮先验知识参数的筛选函数。该过程的核心在于,随着迭代的进行, $\Theta_t$ 不断被优化,使得 $\text{Filter}$ 函数的判别愈发精准,从而实现掩膜集合 $M$ 的数量扩充与质量提升。

随后,新生成的掩膜集合 $M_{t+1}$ 将驱动下一轮迭代的“先验知识”与“分割提示”进行更新:

(1) 先验知识自适应更新:系统基于 $M_{t+1}$ 所覆盖的全部像元,重新计算其水稻光谱特征、面积阈值的统计分布,并据此更新知识参数集,得到 $\Theta_{t+1}$ :

$$\Theta_{t+1} = \text{Update}(\Theta_t, \text{Statistics}(M_{t+1})) \quad (11)$$

其中,  $\text{Statistics}(M_{t+1})$ 代表从新掩膜集合中计算的统计特征(均值 $\mu_{t+1}$ 和标准差 $\sigma_{t+1}$ ),  $\text{Update}$ 函数则根据这些统计量生成更紧凑的阈值区间。

(2) 提示信息动态生成:基于更新后的掩膜集合 $M_{t+1}$ 与非植被区域 $R_{neg}$ ,按照3.5节所述方法生成新一轮的点提示集合 $P_{t+1}$ ,用于启动下一次分割:

$$P_{t+1} = \text{GeneratePrompts}(M_{t+1}, R_{neg}) \quad (12)$$

由公式(10)至(12)定义的反馈闭环,实现了先验知识在迭代中的动态更新。其过程如下:在初始阶段( $t=0$ ),掩膜集合 $M_0$ 及提示 $P_0$ 由经验阈值生成,分割结果较为粗糙。进入第一轮迭代( $t=1$ )后,以 $M_0$ 作为高质量样本集,计算其关键特征的统计分布,并据此生成更为精确的动态阈值,进而筛选出更优的掩膜集合 $M_1$ 并构建出更新后的提示 $P_1$ 。随着迭代进行( $t=2, 3, \dots, n$ ),非水稻样本逐步被剔除,提示点质量逐步提高,阈值趋于收敛并接近真实的地块特征范围。该迭代过程通过数据反馈不断修正阈值,直至分割结果稳定。

先验知识动态更新机制贯穿于引入的特征维度。例如,用以区分水稻与水体的关键指标 $\Delta$ ( $\text{LSWI}-\text{NDVI}$ )的判别阈值,水稻生长旺盛期 $\text{NDVI}$ 以及地块面积约束,均根据每一轮迭代后高置信度样本的统计特征进行自适应调整。该机制确保了初始阶段的“粗筛选”能在迭代中通过统计驱动实现“精收敛”,显著提升本文方法对不同区域农田景观特征的普适性。该反馈闭环机制的核心在于,通过迭代过程实现提示信息与分割结果的相互驱动和逐步精化,从而使地块边界趋于平滑,遗漏区域得到有效补全。

为确保迭代过程的科学收敛并防止过拟合,本研究设计了双重终止准则。基于分割稳定性的收敛判据:当连续两轮迭代结果的交并比达到预设阈值时,即 $\text{IoU}(M_{t+1}, M_t) \geq \theta_{\text{IoU}} = 0.95$ ,表明分割边界已趋于稳定;基于面积变化率的收敛判据:当相邻迭代间掩膜面积变化率低于阈值时,即 $\frac{|\text{Area}(M_{t+1}) - \text{Area}(M_t)|}{\text{Area}(M_t)} \leq \theta_{\text{area}} = 0.01$ ,表明分割已收敛。同时设置最大迭代次数 $T_{\text{max}} = 4$ 作为硬约束,避免计算资源的无效消耗。同时,参考 $\text{MeanShift}$ 聚类的收敛思想,若迭代过程中掩膜的几何中心基本不再移动或边界几乎完全重合,也可视为收敛,与 $\text{IoU}$ 的判据相一致。结合实验结果,在宁河区与富锦市,3轮迭代即可收敛,而在新濤市碎片化地块区则为4轮。整个基于SAM与知识融合的自适应迭代优化流程归纳在算法1中。

### 3.8 实验设计与精度验证

#### 3.8.1 实验设计

为确定水稻提取的最优时相,本研究以天津市宁河区为实验区构建等距渔网,采取随机抽样方法选取100个实验区(1024m×1024m,面积约1.05km<sup>2</sup>),用于开展覆盖水稻生育期的多时相影像对比实验,以确定水稻提取的最优时相。图6展示了实验样方的空间分布情况。

#### 3.8.2 真值标注方法

为检验方法的跨区泛化能力,并确保验证样本的一致性与可比性。本研究在宁河区、富锦市和新濤市三个区域分别制作了100个512×512像素大小的真值样本,用于方法精度评价。真值样本制作基于水稻收割期的Sentinel-2伪彩色影像与谷歌高分辨率影像,通过人工目视解译结合已有水稻分布产品及 $\text{NDVI}/\text{LSWI}$ 时序曲线识别典型水稻地块,并利用Labelme等交互式标注工具勾画地块边界,生成二值化掩膜作为地块真值样本。

#### 3.8.3 精度评价指标

精度评价采用总体精度(Overall Accuracy, OA)、Kappa系数(Kappa)、用户精度(User's Accuracy, UA)、生产者精度(Producer's Accuracy, PA)和 $F_1$ 分数( $F_1$ -score)等指标。设混淆矩阵中真阳性(True Positive, TP)、假阳性

(False Positive, FP)、真阴性 (True Negative, TN)、假阴性 (False Negative, FN) 分别表示像元

分类统计参数, 具体公式如下:

算法 1: 基于SAM与知识融合的自适应迭代地块级水稻提取

```

Input: 时间序列影像  $I$ ; 植被掩膜  $V$ ; 初始阈值  $\sigma_0$ ; 最大迭代次数  $T_{max}$ ; 收敛阈值  $\epsilon$  (IoU),  $\delta$  (Area)
Output: 最终水稻掩膜集合  $M^*$ 

/* 1. 初始化阶段 */
1 用  $V$  裁剪  $I$ ;
2  $M_0 \leftarrow \text{SAM}_{\text{auto}}(I)$ ;
3  $M \leftarrow \text{Filter}(M_0, \sigma_0)$ ;
4  $M_{\text{old}} \leftarrow M, t \leftarrow 0$ ;

/* 2. 迭代优化阶段(双循环) */
5 repeat
6    $t \leftarrow t + 1$ ;
7    $\sigma_t \leftarrow \text{Update}(\text{Statistics}(M, I))$ ; // 循环1: 知识更新
8    $P \leftarrow \text{GeneratePrompts}(M, R_{\text{neg}})$ ; // 循环2: 提示生成
9    $C \leftarrow \text{SAM}(I, P)$ ; // SAM 分割
10   $M_{\text{next}} \leftarrow \text{Filter}(C, \sigma_t)$ ;
11   $\text{IoU}_{\text{score}} \leftarrow \text{IoU}(M_{\text{next}}, M)$ ;
12   $\Delta \text{Area} \leftarrow |\text{Area}(M_{\text{next}}) - \text{Area}(M)| / \text{Area}(M)$ ;
13   $M_{\text{old}} \leftarrow M, M \leftarrow M_{\text{next}}$ ;
14 until ( $\text{IoU}_{\text{score}} \geq 0.95$  且  $\Delta \text{Area} \leq 0.01$ ) 或  $t = T_{max}$ ;

/* 3. 输出 */
15  $M^* \leftarrow M$ ;
16 return  $M^*$ ;

```

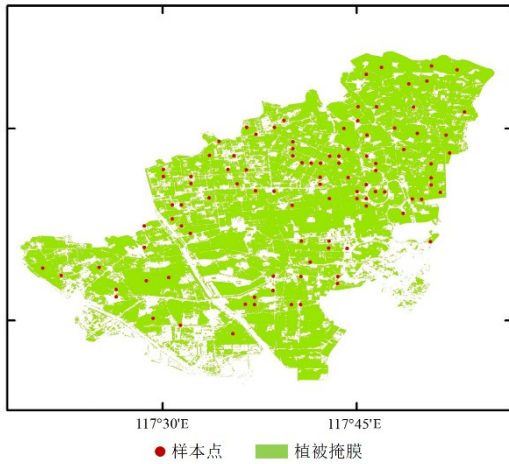


图6 研究区植被掩膜及实验样方空间分布

Fig.6 Spatial distribution of vegetation mask and sample points in the study area

$$O_{\text{A}} = \frac{TP + TN}{TP + FP + TN + FN} \quad (13)$$

$$\text{Kappa} = \frac{p_o - p_e}{1 - p_e} \quad (14)$$

$$\text{UA} = \frac{TP}{TP + FP} \quad (15)$$

$$\text{PA} = \frac{TP}{TP + FN} \quad (16)$$

$$F_1 = \frac{2 \times \text{UA} \times \text{PA}}{\text{UA} + \text{PA}} \quad (17) \text{ 其中, TP 为正确提}$$

取的水稻像元数, TN 为正确分类的非水稻像元数,

FP 为错分为水稻的像元数, FN 为漏分的水稻像元数。 $p_o$  为观测一致率 (即 OA),  $p_e$  为期望一致率。

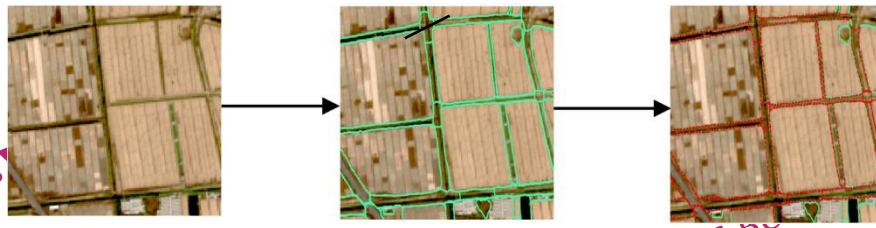
## 4 实验结果

### 4.1 先验知识迭代学习效果

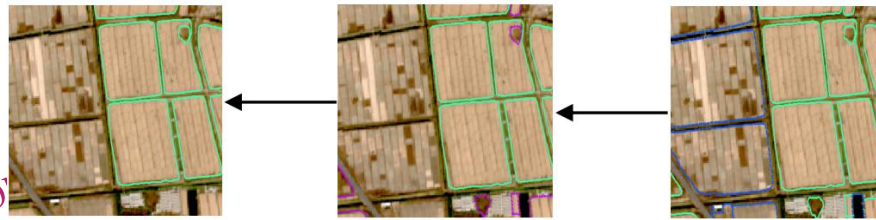
为验证先验知识约束的有效性, 随机抽取 100 个实验区中的典型样区进行可视化展示 (图 7)。处理流程包括: (a) Sentinel-2 原始影像; (b) SAM 全分割; (c) 淹水期光谱特征筛选; (d) 旺盛期 NDVI 筛选; (e) 面积阈值筛选; (f) 高置信度地块级水稻。

结果表明: (1) 淹水期光谱特征有效过滤了与水稻“水面—秧苗”特征不符的对象, 减少了裸地、建设用地和水体等干扰; (2) 旺盛期 NDVI 特征进一步过滤了 NDVI 变化不符合水稻生长规律的旱作物; (3) 通过面积阈值有效去除了噪声小斑块, 并识别异常合并地块。

三类先验知识的综合约束显著提升了地块内像元一致性和边界完整性, 为后续提示分割与迭代优化提供了可靠的输入。实验还表明, 该方法在合理阈值范围内保持稳定效果, 说明其稳健性主要源自多重先验的联合约束, 而非单一参数依赖。



(a) Sentinel-2 原始影像 (b) SAM 全分割 (c) 淹水期光谱特征筛选  
 (a) Raw Sentinel-2 imagery (b) SAM automatic mask generator (c) Spectral feature filtering in flooding



(f) 高置信度地块级水稻 (e) 面积阈值筛选 (d) 旺盛期 NDVI 筛选  
 (f) High-confidence parcel-level rice (e) Area threshold filtering (d) NDVI filtering in peak growth

□ 候选地块边界 □ 光谱特征筛选 □ 旺盛期 NDVI 筛选 □ 面积阈值筛选

图7 典型样区先验知识融合流程图

Fig.7 Flowchart of prior knowledge integration in typical sample area  
 (注: 实线为保留的地块边界, 虚线为各步骤中被剔除的异常地块。)

为进一步揭示本研究提出的统计自适应先验知识框架在迭代优化过程中的动态更新机制, 图8展示了宁河区典型样区内三个关键先验特征(淹水期 LSWI-NDVI、旺盛期 NDVI、地块面积阈值)的概率密度分布及其动态阈值在不同迭代阶段的变化过程。

初始阶段(图8a): 在SAM初次分割后, 所有筛选地块的特征分布呈现较大离散度。例如, 旺盛期 NDVI 的分布覆盖了从0.3到0.8的宽泛区间, 反映了其中混杂了大量非水稻植被。此时, 模型基于整体分布计算出的初始动态阈值范围也相对宽松, 起到了“粗筛选”的作用。

第1—2轮迭代(图8b, 8c): 经过前一轮筛选和SAM的提示分割后, 非地块级水稻被大量过滤。从图中可以清晰地看到, 三个特征的分布均开始向中心区域显著收敛, 标准差( $\sigma$ )减小, 表明筛

选后的样本集内部的一致性增强。

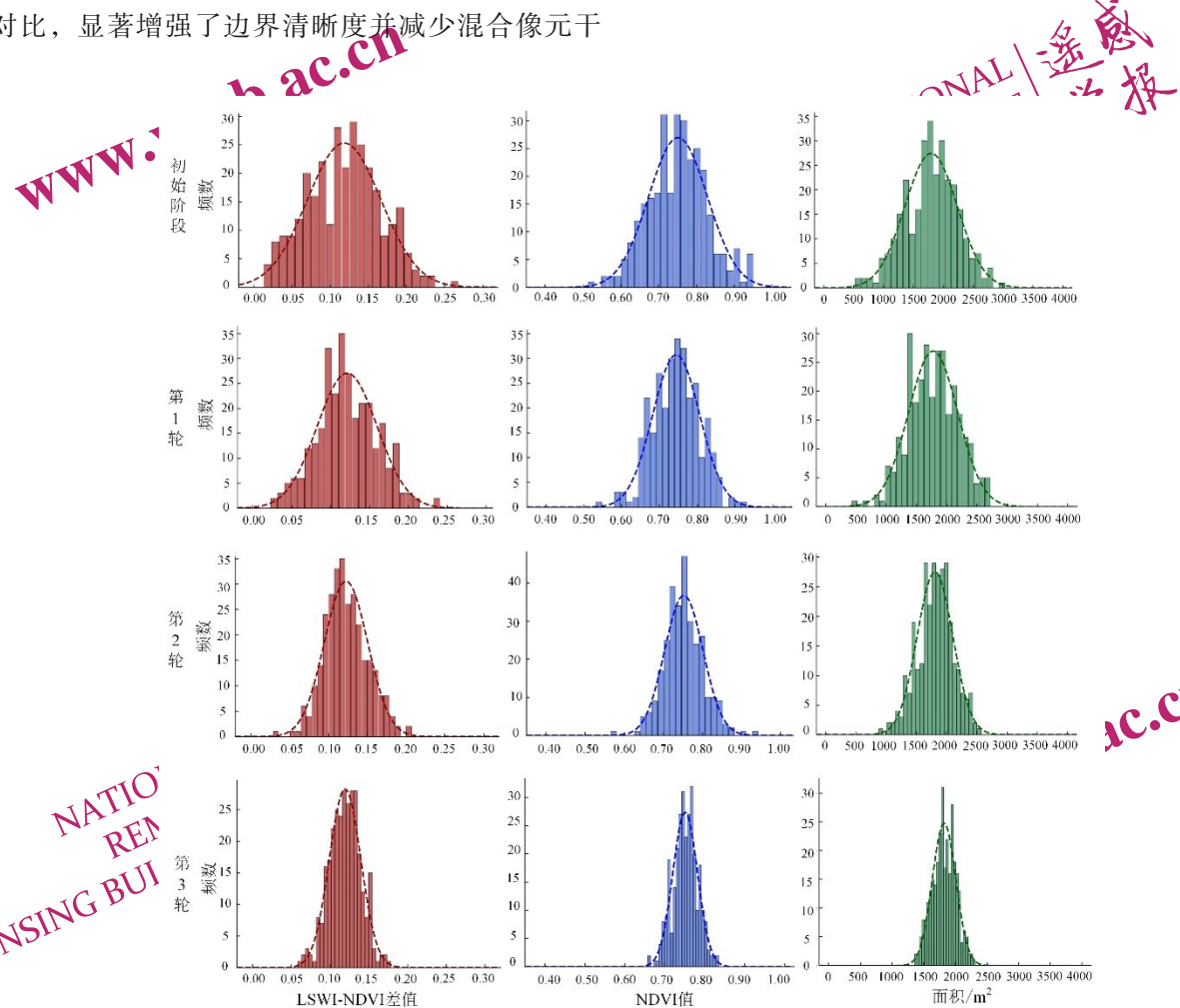
第3轮迭代(图8d): 此时, 特征分布已基本稳定, 与上一轮相比变化极小。动态阈值也收敛至一个稳定的窄区间, 精确地描绘出了该区域地块级水稻的核心特征范围。这标志着迭代过程已达到收敛, 实现了从混杂数据中“学习”并“锁定”目标特征的闭环优化。

#### 4.2 多时相表现与最佳提取窗口

在11个关键物候时相的实验中, 该方法整体精度较高, 平均OA为97.90%, Kappa为0.96, UA、PA和 $F_1$ -score分别为93.80%、94.55%和94.10%。但不同物候阶段精度表现存在差异。育秧—移栽期受水体干扰影响, 精度偏低; 进入拔节孕穗期, 随着冠层快速生长, Kappa提升至0.85-0.90; 灌浆期精度进一步提高; 成熟期保持稳定。收割期表现最佳, Kappa值为0.96。其原因

在于收割期水稻与其他地物形成强烈光谱与纹理对比，显著增强了边界清晰度并减少混合像元干

扰。因此，收割期是水稻提取的最佳窗口。



(a) LSWI-NDVI 差值分布 (b) NDVI 分布 (c) 地块面积分布  
 (a) Distribution of LSWI-NDVI difference (b) Distribution of NDVI (c) Distribution of parcel area

■ 差值直方图 --- 正态分布拟合 ■ NDVI直方图 --- 正态分布拟合 ■ 面积直方图 --- 正态分布拟合

图8 先验知识特征分布与动态阈值的迭代收敛过程

Fig.8 The iterative convergence of a priori knowledge feature distributions and dynamic thresholds

### 4.3 消融实验与组件有效性分析

为系统验证本文所提框架中各核心组件的有效性，本研究设计了逐级递进的消融实验。实验以基线模型开始，依次增加提示点策略、动态阈值更新和迭代机制，评估各模块在逐级引入过程中的边际贡献。

本文消融实验基于固定实验设置下的单次整体实验结果（将随机抽取的100个实验区）视为一个统计整体进行精度评估，而非多次独立重复试验。因此，各模块组合结果不满足传统统计显著性检验（如t-test或p值计算）对样本独立性与重复性的基本假设。基于此，本文采用精度提升增

益分析方法，通过比较逐级消融实验中相邻阶段模型性能的变化幅度，对各模块在引入过程中的边际贡献量化评估。增益百分比计算公式如下：

$$G_k = \frac{V_k - V_{k-1}}{V_{k-1}} \times 100\% \quad (18)$$

其中， $V_{k-1}$ 为第 $k-1$ 阶段（即基准阶段）的精度指标； $V_k$ 为第 $k$ 阶段（引入模块后）的精度指标； $G_k$ 为第 $k$ 阶段的边际增益百分比。

表1展示不同模块组合条件下的消融实验结果。基线模型Kappa系数为0.70，表明静态阈值难以适应复杂遥感场景。引入提示点策略与动态阈值更新机制后，模型性能持续提升，Kappa分别提

升至0.76和0.79，表明基于提示点引导和高置信样本统计反馈的筛选机制有效减少边界误分与类别混淆。引入迭代机制后，模型形成“分割—筛选—知识更新—再提示”的闭环优化框架，整体性能提升最为显著，Kappa系数达0.89。

表2汇总了逐级消融实验的边际增益分析结果。可以看出，各核心模块在OA、Kappa、UA、PA及F<sub>1</sub>-score等指标上均实现稳定正向提升。相比提示点策略与动态阈值更新组合，引入迭代机制带来显著的性能增益，Kappa提升12.6%，PA提升10.52%，表明迭代优化机制在减少漏分、提升分类一致性与结果稳定性方面发挥了关键作用，验证了本文所提出各组件设计的合理性与整体框架的有效性。

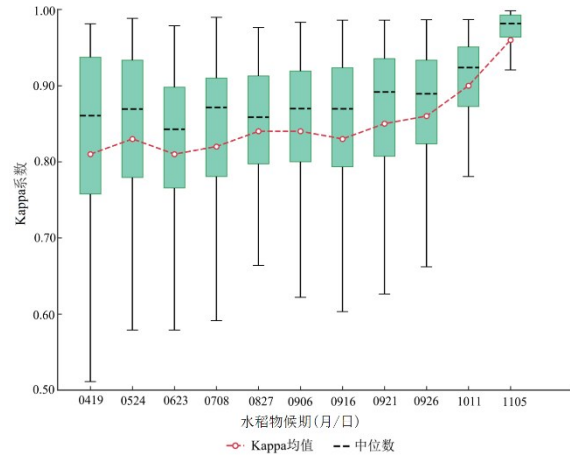


图9 水稻不同物候期Kappa系数的分布特征与变化趋势  
Fig.9 Distribution characteristics and trends of Kappa coefficients across different phenological stages of rice

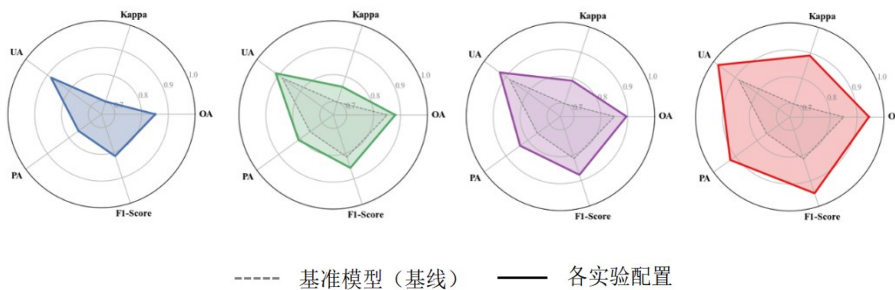
#### 4.4 初始阈值敏感性分析

表1 消融实验结果

Table1 Ablation study results

模块组合			总体指标				
提示点策略	动态阈值	迭代机制	OA/%	Kappa	UA/%	PA/%	F <sub>1</sub> -score/%
×	×	×	85.12	0.70	88.34	75.60	81.47
✓	×	×	88.06	0.76	91.35	80.88	85.17
✓	✓	×	89.54	0.79	92.85	83.52	87.83
✓	✓	✓	<b>94.44</b>	<b>0.89</b>	<b>97.86</b>	<b>92.31</b>	<b>94.90</b>

注：表中“/”代表加入模块；“×”代表去除模块；加粗数值表示同列最佳结果。



(a) 基准模型 (b) 引入提示点策略 (c) 引入动态阈值 (d) 完整模型  
(a) Baseline model (b) Prompt point strategy (c) Dynamic threshold (d) Complete model

图10 消融实验雷达图

Fig.10 Ablation study radar chart

为验证本文方法对初始先验知识阈值设定的依赖性，本研究开展初始阈值的敏感性分析实验。以天津市宁河区为例，在先验知识约束条件下 (LSWI<sub>均值</sub> > NDVI<sub>均值</sub>；LSWI<sub>均值</sub> (淹水期) < 0.45；NDVI<sub>均值</sub> (生长期) > 0.4；NDVI<sub>最小值</sub> > 0) 构建解空间，并在解空间的边界拐点附近随机选取点S1、

S2、S3 (SPT1) 作为初始阈值。

不同初始阈值迭代结果显示，本方法在不同初始阈值设定下均能实现水稻地块的精准提取，结果收敛且精度评价指标Kappa值趋同 (表3)，表明方法对初始阈值的敏感性较弱，具有良好鲁棒性与收敛性。

4.5 模型对比

表2 逐级消融实验边际增益分析结果

Table 2 Marginal gain analysis results of progressive ablation experiments

提示点策略	动态阈值	迭代机制	OA	Kappa	提升百分比/%	PA	F <sub>1</sub> -Score
×	×	×	-	-	-	-	-
√	×	×	+3.45	+8.57	+3.01	+6.98	+4.51
√	√	×	+1.68	+3.95	+1.64	+3.26	+3.12
√	√	√	<b>+5.47</b>	<b>+12.6</b>	<b>+5.40</b>	<b>+10.52</b>	<b>+8.05</b>

注：提升百分比表示逐级消融实验中，相邻阶段相对于前一阶段的边际性能增益，表中加粗数值表示该指标下的最大边际增益。

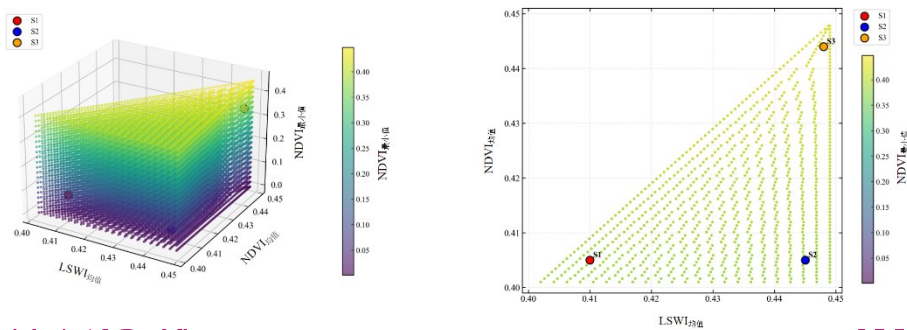


图11 初始阈值解空间及初始阈值点分布

Fig. 11 Initial threshold solution space and initial threshold point distribution

(a)可行解空间三维散点分布 (b)可行解空间XY平面投影区域

(a)3D Scatter Distribution of Feasible Solution Space (b)XY Plane Projection Region of Feasible Solution Space

表3 不同初始阈值条件下的实验结果

Table3 Experimental results under different initial threshold conditions

初始阈值点	LSWI <sub>均值</sub>	NDVI <sub>均值</sub>	NDVI <sub>最小值</sub>	Kappa
S1	0.410	0.405	0.080	0.872
S2	0.445	0.405	0.050	0.878
S3	0.448	0.444	0.350	0.877

本研究以 U-Net (Ronneberger 等, 2015)、DeepLabV3+和 PSPNet (Zhao 等, 2017) 等主流监督学习模型的分割结果作为参照基准, 衡量本文方法与具备训练样本条件下监督学习模型的表现差距与潜力。选取3种主流监督学习模型进行对比旨在确保参照基准的客观性, 避免受到单个模型的性能影响。

以天津市宁河区的水稻遥感制图为例, 第一步, 基于水稻收割期 Sentinel-2 影像 (B2、B3、B4、B8、B11), 使用 GIS 软件根据宁河区边界遥感影像进行剪裁和分块处理, 裁剪的样本图像尺寸为 512×512 像素。第二步, 随机抽取 5000 张图

片, 结合 Labelme 等样本制作辅助软件进行交互式标注与人工勾画, 生成二值化的掩膜标签图 (即水稻地块与非水稻地块), 并将全部标注样本图片按照 8:2 的比例划分为训练集与测试集。第三步, 对各监督模型采用相同的初始化设置, 基于上述步骤制作的训练样本数据集分别训练模型, 并采用上述步骤制作的测试集对各模型进行精度评价对比, 确保对比测试条件一致。

图 12 展示了不同模型在典型样区的分割结果。从图中可以看出, U-Net 和 DeepLabV3+ 在大田区块内能够较好地识别水稻, 但在道路、沟渠交错及居民点等复杂背景下仍存在一定误分割; PSPNet 在破碎小斑块提取中表现尤为不足, 出现较多漏分与粘连。相比之下, 本文方法能够更准确地刻画田块边界, 并在复杂背景下保持稳定的提取效果。

表 4 列出了不同模型的精度指标。在监督模型经过充分训练的前提下, 本方法在总体精度 (OA=97.90%)、Kappa 系数 (0.96)、用户精度 (UA=93.80%)、生产者精度 (PA=94.55%) 和 F<sub>1</sub>-score

(94.10%) 等指标上均达到了与监督模型相当的水平。其中，PA 指标提升尤为显著，说明本方法有效改善漏分问题。结果显示，本方法在不依赖大

量标注样本的情况下，能够实现与全监督深度学习模型相当甚至略优于其性能的提取精度，具有良好的适应性与鲁棒性。

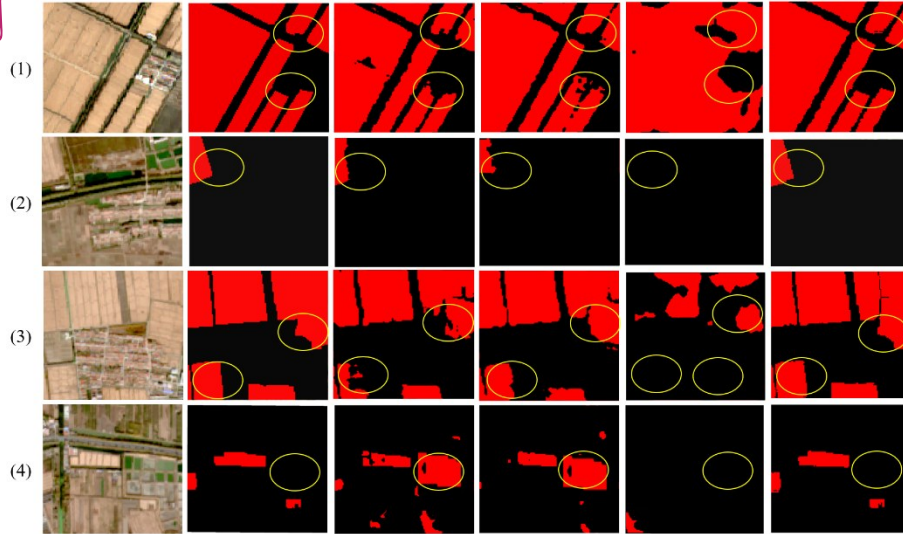


图 12 不同模型对比结果

Fig.12 Comparison results of different models

(a) 卫星影像 (b) 真值 (c) U-Net (d) DeepLab V3+ (e) PSPNet (f) 本文方法

(a) Satellite Image (b) Ground Truth (c) U-Net (d) DeepLab V3+ (e) PSPNet (f) Ours

表 4 不同模型对比结果

Table 4 Comparison results of different models

	OA/%	Kappa	UA/%	PA/%	F <sub>1</sub> -score/%
本文方法	<b>97.90</b>	<b>0.96</b>	<b>93.80</b>	<b>94.55</b>	<b>94.10</b>
U-Net	95.50	0.89	90.00	90.50	90.25
DeepLabV3+	96.20	0.91	92.00	92.60	92.20
PSPNet	92.50	0.73	78.50	75.20	76.80

注：表中加粗数值为同一列最佳结果。

#### 4.6 多区域制图结果与泛化能力验证

基于最优时相遥感影像，完成宁河区 2022 年水稻分布制图 (图 13)。结果显示，水稻主要集中于地势平坦、水系发达的中部乡镇，空间分布格局与当地农业生产活动一致，制图边界清晰、图斑完整。精度评估表明：OA 为 94.44%，Kappa 为 0.89，UA 为 97.86%，PA 为 92.31%，F<sub>1</sub> 分数为 94.9%，体现了方法在区域尺度上的稳健性与适用性。

为检验方法的跨区域泛化能力，在黑龙江富锦市与日本新潟市开展了验证实验，重点评估在统一策略下，场景自适应学习中地块分割提示信息优化的跨区表现。核心试验区 (宁河区) 的多

时相分析确立了一项关键的最优时相选择策略：水稻收割期是水稻精准提取的最佳窗口期。为验证该策略的普适性，本研究将其迁移应用至两个新区域，根据富锦市与新潟市的实际农业生产节律，分别遴选出其水稻收割期的代表性影像 (富锦：2022 年 9 月 26 日；新潟：2023 年 10 月 14 日) 进行无训练样本条件下的地块级水稻精细提取。富锦以大田块为主，动态阈值随当地统计特征自动调整，实现对连片水稻的精准分割，OA 值为 96.80%，Kappa 系数为 0.91；新潟景观破碎、地块偏小，动态面积阈值自动下调以避免漏检，精度仍保持稳定，OA 值为 94.50%，Kappa 系数为 0.86。结果表明，“收割期最优时相”与“场景自适应机制”的结合有效支撑了跨区域的稳健表现。

跨区域验证结果表明，该框架具备出色的泛化能力，主要归因于“最优时相选择策略”与“场景自适应机制”的协同作用。前者保证模型在最佳的数据输入条件下运行，后者支持模型在无需任何人工调参的情况下，自适应不同区域地块尺度与背景特征，从而实现地块级水稻精准提取。三个区域的具体精度对比结果如表 5 所示。

表5 不同研究区水稻提取精度对比

Table 5 Comparison of rice extraction accuracy across different study areas

研究区	OA/%	Kappa	UA/%	PA/%	F1-score/%
宁河区	94.44	0.89	97.86	92.31	94.90
富锦市	96.80	0.91	95.77	96.25	96.01
新潟市	94.50	0.86	92.75	91.43	92.09

#### 4.7 计算效率分析

为验证本方法的计算效率，对关键环节的时间开销进行系统评估。植被掩膜预处理显著提升了大范围影像的处理效率，以宁河区为例，约30%的非植被单元被自动跳过。对比实验结果表明，引入植被掩膜后，宁河区的数据处理时间由“全图SAM”的684秒降至“植被掩膜+SAM”的480秒，计算效率提升约29.8%。

在迭代优化过程中，进一步统计单轮迭代各环节时间。主要计算开销来自SAM推理与分割，耗时557秒，约占单轮总耗时的87%；知识更新环节耗时约13秒，占比约2%；提示生成环节耗时约70秒，占比约11%。表6列出了宁河区迭代的详细时间分布。

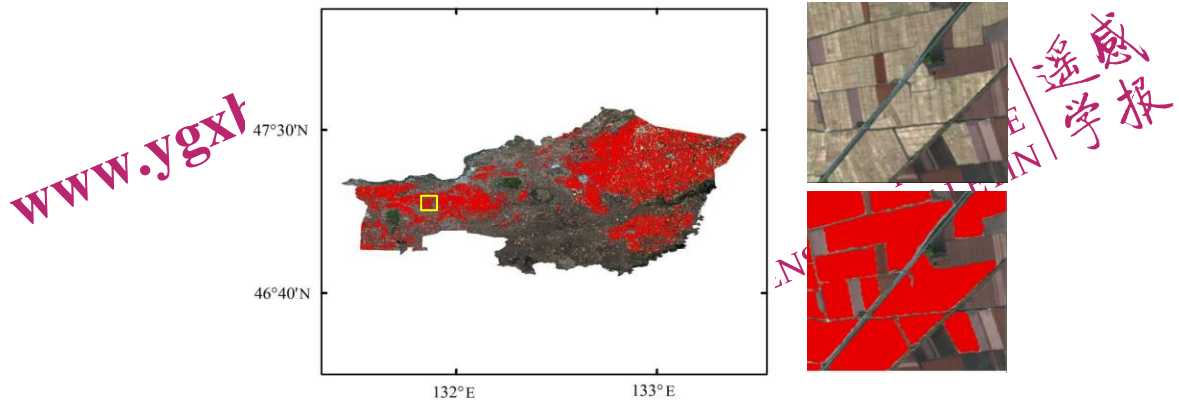
对比不同策略的总耗时，采用迭代和非迭代的方法，总耗时分别为1446秒和480秒。迭代虽会增加计算时间，但带来了稳定的精度提升与边界质量改进。消融结果显示，迭代机制使Kappa由0.79提升至0.89，总体精度（OA）由89.54%提升至94.44%， $F_1$ 分数由87.83%提升至94.90%，使边界更平滑，粘连与小斑块噪声减少，漏分率降低。

NATIONAL  
REMOTE  
SENSING BULLETIN | 遥感学报

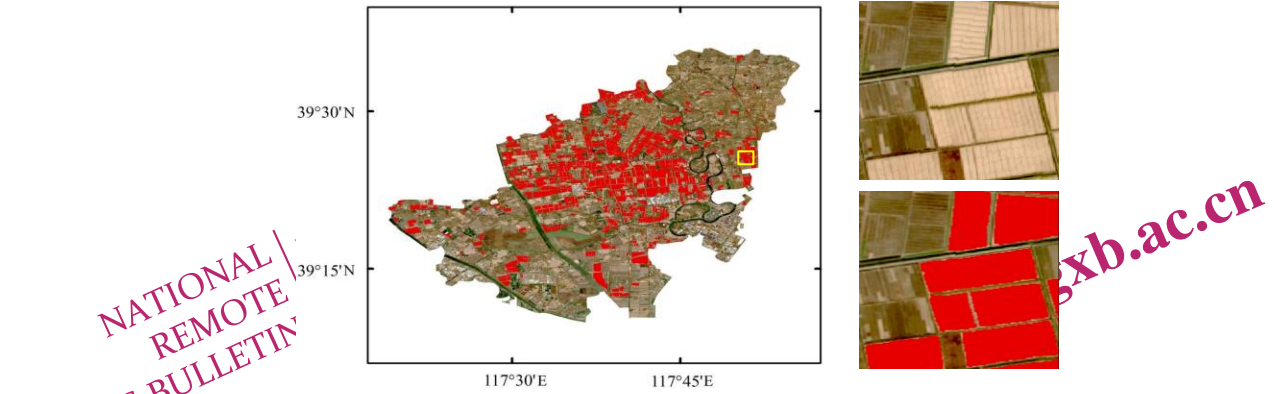
www.ygxb.ac.cn

NATIONAL  
REMOTE  
SENSING BULLETIN | 遥感学报

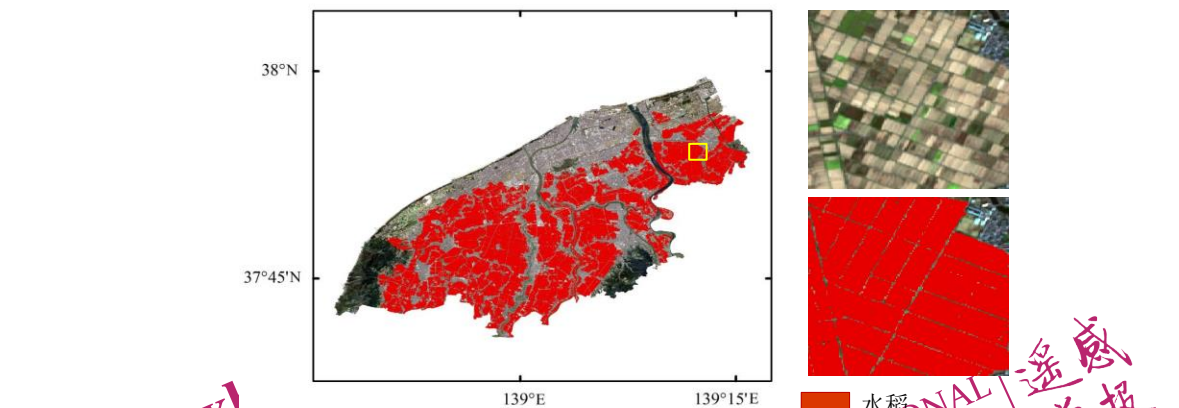
www.ygxb.ac.cn



(a) 富锦市  
(a) Fujin City



(b) 宁河区  
(b) Ninghe District



(c) 新潟市  
(c) Niigata City

图 13 不同研究区水稻空间分布结果及局部细节  
Fig.13 Spatial distribution results and detailed views of rice in different study areas

表6 宁河区迭代各环节时间消耗分布

Table 6 Time consumption distribution across iteration stages in Ninghe District

迭代环节	具体操作	迭代耗时/s	耗时占比/%
SAM分割	基于提示点进行模型推理与分割	557	87
知识更新	统计高置信度地块的光谱/面积特征,更新阈值	13	2
提示生成	将掩膜转换为点/边界点	70	11
单轮迭代总耗时	-	640	100

## 5 结论

针对当前精细化作物分布制图方法高度依赖大量标注样本,面临成本高昂、时效性差及跨区域泛化能力不足等瓶颈,本文提出了融合视觉基础模型(SAM)与农业遥感先验知识的水稻地块提取框架。该框架通过构建“分割—筛选—迭代优化”的自适应学习闭环,利用水稻物候、时序光谱及面积等多维度农业遥感先验知识,动态生成并优化分割提示信息,最终实现水稻的精准提取。

研究表明,本方法在天津宁河、黑龙江富锦及日本新潟等不同农业景观区均取得较高精度,验证了该框架在不同地块规模和种植模式下的有效性、稳健性及区域适用性。消融实验表明,迭代机制对水稻精细化提取具有关键作用,Kappa值由0.79提升至0.89。初始阈值敏感性实验显示,在合理参数范围内,动态阈值和知识更新机制能够在不同初始阈值的设定下保持稳定输出,体现了方法在实际应用中的一定鲁棒性。此外,本研究通过多时相遥感数据分析,量化确定水稻收割期为最佳提取物候窗口,发现此时相下水稻与其他地物光谱、纹理差异呈现最大化,为实际应用中的时相选择提供了依据。

本研究验证了将通用视觉大模型的强分割能力与农业领域先验知识耦合的技术路径,为解决遥感信息提取中的“样本依赖”难题提供可行方案。该框架具备较高的自动化程度与跨区域适应性,可在大范围、低成本条件下实现水稻地块的高精度制图,为粮食安全评估和可持续农业管理等实际应用场景提供支持。

尽管本文方法取得了良好效果,但其仍存在一定的局限性。首先,虽然框架内置阈值迭代自适应更新机制,但初始筛选环节仍依赖经验性先验知识,尚未实现完全的数据驱动初始化。当初

始阈值严重偏离合理范围时,可能导致统计更新过程向错误特征空间漂移,存在难以收敛的风险。其次,本文主要基于时序光谱特征与面积阈值等先验知识,结合SAM图像分割基础模型以实现水稻遥感提取,尚未融合纹理特征、形状指数等判别信息,在光谱特征高度相似的复杂地表条件下仍存在提升空间。此外,当前实验主要针对水稻场景开展,方法在多作物类型和多源数据条件下的适用性仍需进一步验证。

从方法机制、理论完备性及应用拓展的角度来看,本文所提出框架仍具有进一步深化与扩展的空间。在保持分割—筛选—反馈闭环与先验提示协同机制不变的前提下,未来工作将从以下几个方面展开:一是训练面向遥感场景的专用SAM模型,结合高空间分辨率影像,降低初始分割阶段的边界与拓扑误差;二是引入时序图像分割模型,克服单时相分割的固有局限,进一步降低拓扑错误传播风险;三是探索由数据驱动的初始参数自适应生成方法,例如利用无监督聚类或异常检测算法自动确定初始阈值,减少对经验性先验知识的依赖;四是进一步融合纹理特征与形状指数等判别信息,系统评估多物候期、多时相及多波段数据融合在提升分割稳定性与精度方面的潜力;五是该框架从水稻拓展至玉米、小麦等其他主粮作物,并在多熟制农业区开展验证,以评估其作为通用作物提取框架的应用潜力。此外,从理论层面出发,未来还将尝试为“基于高置信样本的迭代提纯”过程构建概率模型,系统分析其收敛条件与错误传播上界,进而更清晰地刻画方法的适用范围与理论极限。

## 参考文献(References)

- Bazzi H, Baghdadi N, El Hajj M, Zribi M, Minh D H T, Ndikumana E, Courault D and Belhouche H. 2019. Mapping paddy rice using Sentinel-1 SAR time series in Camargue, France. Remote Sensing

- ing, 11(7): 887 [DOI: 10.3390/rs11070887]
- Blaschke T. 2010. Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, 65(1): 2 - 16 [DOI: 10.1016/j.isprsjprs.2009.06.004]
- Boryan C, Yang Z, Mueller R and Craig M. 2011. Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, cropland data layer program. Geocarto International, 26(5): 341 - 358 [DOI: 10.1080/10106049.2011.562309]
- Boschetti M, Nutini F, Manfron G, Brivio P A and Nelson A. 2014. Comparative analysis of normalised difference spectral indices derived from MODIS for detecting surface water in flooded rice cropping systems. PLoS ONE, 9(2): e88741 [DOI: 10.1371/journal.pone.0088741]
- Chen H, Li H, Liu Z, Zhang C, Zhang S and Atkinson P M. 2023. A novel greenness and water content composite index (GWCCI) for soybean mapping from single remotely sensed multispectral images. Remote Sensing of Environment, 295: 113679 [DOI: 10.1016/j.rse.2023.113679]
- Chen K, Liu C, Chen H, et al. 2024. RSPrompter: Learning to prompt for remote sensing instance segmentation based on visual foundation model. IEEE Transactions on Geoscience and Remote Sensing, 62: 4701117 [DOI: 10.1109/TGRS.2024.3356074]
- Chen L C, Papandreu O, Kokkinos L, Murphy K and Yuille A L. 2018. Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4): 834 - 848 [DOI: 10.1109/TPAMI.2017.2699184]
- Chen N, Yu L, Zhang X, Shen Y, Zeng L, Hu Q and Niyogi D. 2020. Mapping paddy rice fields by combining multi-temporal vegetation index and synthetic aperture radar remote sensing data using Google Earth Engine machine learning platform. Remote Sensing, 12(18): 2992 [DOI: 10.3390/rs12182992]
- Ding M, Guan Q, Li L, Zhang H, Liu C and Zhang L. 2020. Phenology-based rice paddy mapping using multi-source satellite imagery and a fusion algorithm applied to the Poyang Lake plain, Southern China. Remote Sensing, 12(14): 2259 [DOI: 10.3390/rs12142259]
- Dong J W, Xiao X M, Kou W L, Qin Y W, Zhang G L, Li L, Jin C, Zhou Y, Wang J, Biradar C, Liu J Y and Moore B III. 2016. Mapping paddy rice planting area in Northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. Remote Sensing of Environment, 185: 142 - 154 [DOI: 10.1016/j.rse.2016.02.016]
- Fritz S, See L, McCallum I, You L, Bun A, Moltchanova E, D' Odorico P, Ramankutty N, Roy D, Robinson C S and Obersteiner M. 2015. Mapping global cropland and field size. Global Change Biology, 21: 1980 - 1992 [DOI: 10.1111/gcb.12838]
- Gui B, Bhardwaj A and Sam L. 2024. Evaluating the efficacy of Segment Anything Model for delineating agriculture and urban green spaces in multiresolution aerial and spaceborne remote sensing images. Remote Sensing, 16(2): 414 [DOI: 10.3390/rs16020414]
- Han J, Zhang Z, Cao J and Luo Y. 2022. Mapping rapeseed planting areas using an automatic phenology- and pixel-based algorithm (APPA) in Google Earth Engine. The Crop Journal, 10(5): 1483-1495 [DOI: 10.1016/j.cj.2022.04.013]
- Han J C, Zhang Z, Luo Y C, Cao J, Zhang L L, Cheng B, Zhuang H M, Jing Z and Tao F L. 2021. AsiaRiceMap10m: high-resolution annual paddy rice maps for Southeast and Northeast Asia from 2017 to 2019. Earth System Science Data, 13: 3733 - 3756 [DOI: 10.5194/essd-13-3733-2021]
- Hao X L, Li H B, Duan Y L, Shang G B and Yu Q Y. 2023. Application of SAM-based field parcel extraction method in the effectiveness evaluation of land leveling. China Agricultural Information, 35(5): 1-10 (郝雪丽, 李会宾, 段玉林, 尚国琲, 余强毅. 2023. 基于SAM的田块提取方法在田块平整成效评价中的应用研究. 中国农业信息, 35(5): 1-10 [DOI: 10.12363/issn.1002-1302.20230209])
- Hossain M D and Chen D. 2019. Segmentation for object-based image analysis (OBIA): A review of algorithms and challenges from remote sensing perspective. ISPRS Journal of Photogrammetry and Remote Sensing, 150: 115-134 [DOI: 10.1016/j.isprsjprs.2019.02.009]
- Hu Y R, Zeng H W, Tian F Y, Zhang M, Wu B F, Gilliams S, Li S, Li Y C, Lu Y M and Yang H H. 2022. An interannual transfer learning approach for crop classification in the Hetao Irrigation District, China. Remote Sensing, 14(5): 1208 [DOI: 10.3390/rs14051208]
- Huang Z X, Jing H T, Liu Y M, Yang X M, Wang Z H, Liu X L, Gao K and Luo H F. 2024. Segment Anything model combined with multi-scale segmentation for extracting complex cultivated land parcels in high-resolution remote sensing images. Remote Sensing, 16(18): 3489 [DOI: 10.3390/rs16183489]
- Janowski Ł and Wróblewski R. 2024. Application and evaluation of the AI-powered Segment Anything Model (SAM) in seafloor mapping: A case study from Puck Lagoon, Poland. Remote Sensing, 16(14): 2638 [DOI: 10.3390/rs16142638]
- Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, Xiao T, Whitehead S, Berg A C, Lo W Y, Dollár P and Girshick R. 2023. Segment Anything. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, 4015 - 4026 [DOI: 10.1109/ICCV51070.2023.00371]
- Li Y, Wang D, Yuan C and Hu J. 2023. Enhancing agricultural image segmentation with an agricultural Segment Anything Model adapter. Sensors, 23(18): 7884 [DOI: 10.3390/s23187884]
- Liu N, Xu X, Su Y, Zhang H and Li H-C. 2025. PointSAM: Pointly-supervised Segment Anything Model for remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 63: 5608515 [DOI: 10.1109/TGRS.2025.3529031]
- Onojegun A O, Blackburn G A, Wang Q, Atkinson P M, Kindred D and Miao Y. 2018. Mapping paddy rice fields by applying machine learning algorithms to multi-temporal Sentinel-1A and Landsat data. International Journal of Remote Sensing, 39(4): 1042-1067 [DOI: 10.1080/01431161.2017.1395969]
- Osco L P, de Lemos E L, Wu Q, Bartolo R E, Ji L and Ferreira M P. 2023. The Segment Anything Model (SAM) for remote sensing

- applications: from zero to one shot. *International Journal of Applied Earth Observation and Geoinformation*, 124: 103540 [DOI: 10.1016/j.jag.2023.103540]
- Qin Y W, Xiao X M, Dong J W, Zhang G L, Kou W L, Menarguez M A, Biradar C, Jin C, Zhou Y, Wang J, Moore B and Li X P. 2015. Mapping paddy rice planting area in cold temperate climate region through analysis of time series Landsat 8 (OLI), Landsat 7 (ETM+) and MODIS imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 105: 220-233 [DOI: 10.1016/j.isprsjprs.2015.04.008]
- Qiu B, Li W, Tang Z, Chen C and Qi W. 2015. Mapping paddy rice areas based on vegetation phenology and surface moisture conditions. *Ecological Indicators*, 56: 79 - 86 [DOI: 10.1016/j.ecolind.2015.03.039]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: Convolutional networks for biomedical image segmentation. In: *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 234 - 241 [DOI: 10.1007/978-3-319-24574-4\_28]
- Sakamoto T, Van Nguyen N, Kotera A, Ohno H, Ishitsuka N and Yokozawa M. 2007. Detecting temporal changes in annual flooding in the Mekong Delta from MODIS time-series imagery. *Remote Sensing of Environment*, 109(3): 295 - 313 [DOI: 10.1016/j.rse.2007.01.011]
- Telugunta P, Rye D, George B, Walker J P and Malano H. 2015. Mapping flooded rice paddies using time series of MODIS imagery in the Krishna River Basin, India. *Remote Sensing*, 7(7): 8858 - 8882 [DOI: 10.3390/rs70708858]
- Tucker C J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2): 127 - 150 [DOI: 10.1016/0034-4257(79)90013-0]
- Van Tricht K, Degerickx J, Gilliams S, Lemoine G, McNairn H and Waldner F. 2023. WorldCereal: a dynamic open-source system for global-scale, seasonal, and reproducible crop and irrigation mapping. *Earth System Science Data*, 15(12): 5491 - 5515 [DOI: 10.5194/essd-15-5491-2023]
- Waldner F, Canto G S and Defourny P. 2015. Automated annual cropland mapping using knowledge-based temporal features. *ISPRS Journal of Photogrammetry and Remote Sensing*, 110: 1 - 13 [DOI: 10.1016/j.isprsjprs.2015.10.008]
- Xiao X, Boles S, Froelking S, Li C, Babu J Y, Salas W and Moore B. 2006. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. *Remote Sensing of Environment*, 100(1): 95 - 113 [DOI: 10.1016/j.rse.2005.10.004]
- Xu Q, Zhang J S, Zhang F, Qiu S, Yang Z and Duan Y M. 2022. Applicability of weak samples to deep learning crop classification. *National Remote Sensing Bulletin*, 26(7): 1395-1409 (许晴, 张锦水, 张凤雷, 杨志, 段雅鸣. 2022. 深度学习农作物分类的弱样本适用性. *遥感学报*, 26(7): 1395-1409 [DOI: 10.11834/jrs.20221127])
- You N S, Dong J W, Huang J X, Du G M, Zhang G L, He Y L, Yang T, Di Y Y and Xiao X M. 2021. The 10-m crop type maps in Northeast China during 2017-2019. *Scientific Data*, 8(1): 41 [DOI: 10.1038/s41597-021-00827-9]
- Zhang C, Zhang H and Tian S. 2023. Phenology-assisted supervised paddy rice mapping with the Landsat imagery on Google Earth Engine: experiments in Heilongjiang Province of China from 1990 to 2020. *Computers and Electronics in Agriculture*, 212: 108105 [DOI: 10.1016/j.compag.2023.108105]
- Zhang G L, Xiao X M, Dong J W, Menarguez M A, Qin Y W, Chau D, Biradar C, Zhou Y, Zhang Y, Jin C, Wang J, Doughty R B, Ding M J and Moore B. 2015. Mapping paddy rice planting areas through time-series analysis of MODIS land surface temperature and vegetation index data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 106: 157 - 171 [DOI: 10.1016/j.isprsjprs.2015.05.011]
- Zhao H S, Shi J P, Qi X J, Wang X G and Jia J Y. 2017. Pyramid scene parsing network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2881 - 2890 [DOI: 10.1109/CVPR.2017.660]
- Zhao R, Li Y and Ma M. 2021. Mapping paddy rice with satellite remote sensing: a review. *Sustainability*, 13(2): 503 [DOI: 10.3390/su13020503]
- Zhao Z, Dong J, Zhang G, Yang J, Liu R, Wu B and Xiao X. 2024. Improved phenology-based rice mapping algorithm by integrating optical and radar data. *Remote Sensing of Environment*, 315: 114460 [DOI: 10.1016/j.rse.2024.114460]

## A remote sensing method for plot-level rice distribution extraction based on visual foundation models and knowledge fusion

PENG Zimeng<sup>1</sup>, DUAN Yuling<sup>1</sup>, YU Qiangyi<sup>1</sup>, Wu Wenbin<sup>1</sup>, Zhang Shuai<sup>2</sup>, Zhao Chunlei<sup>1</sup>, Li Boliang<sup>3</sup>, Zhang Xin<sup>3</sup>

1.State Key Laboratory of Efficient Utilization of Arable Land in China, Key Laboratory of Agricultural Remote Sensing (AGRIRS) Ministry of Agriculture and Rural Affairs, Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China;

2.Department of Farmland Construction and Management, Ministry of Agriculture and Rural Affairs of the People's Republic of

China, Beijing 100125, China;

3.Smart Tower Corporation Limited, Beijing 100089, China

**Abstract:** Objective Deep learning methods have demonstrated considerable potential for the fine-scale extraction of plot-level crop distributions from remote sensing imagery. However, their performance is heavily reliant on large volumes of high-quality annotated samples, leading to significant challenges, including high labeling costs, poor timeliness, and limited cross-regional generalization capability. These limitations are particularly pronounced in fragmented agricultural landscapes where plot boundaries are complex. This study aims to address these bottlenecks by proposing a novel framework that achieves accurate plot-level rice mapping without the need for region-specific model training and with minimal dependence on manual annotations. By integrating visual foundation models with agricultural prior knowledge, this research seeks to establish a robust, automated solution for crop mapping in diverse environments. Method We propose a plot segmentation prompt optimization method that couples the Segment Anything Model (SAM) with domain knowledge fusion. The core of our framework is an adaptive, iterative learning closed-loop system comprising "segmentation → screening → knowledge update → re-prompting." First, prior knowledge derived from time-series vegetation indices—specifically the Normalized Difference Vegetation Index (NDVI) and Land Surface Water Index (LSWI)—is applied to mask out non-vegetated areas, substantially reducing the computational load. The SAM is then employed for initial segmentation. Crucially, the framework performs statistical analysis on high-confidence segmented plots to dynamically update the thresholds of prior knowledge (e.g., spectral characteristics during flooding and peak growth stages, and geometric area thresholds). These updated thresholds are automatically translated into refined prompt information, including plot center points and boundary points, which are fed back into the SAM model to guide finer segmentation. Furthermore, we innovatively introduce the Intersection over Union (IoU) metric and area change rate as convergence criteria to automatically terminate the iterations, ensuring both algorithmic stability and computational efficiency. Result The proposed method was systematically validated across three geographically and agronomically diverse rice cultivation regions: Ninghe District (Tianjin, China), Fujin City (Heilongjiang, China), and Niigata City (Japan). The results demonstrated high mapping accuracy and strong generalization capabilities. In Ninghe, the method achieved an Overall Accuracy (OA) of 94.44%, a Kappa coefficient of 0.89, and an F1-score of 94.90%. Similarly, high performance was observed in Fujin (OA: 96.80%, Kappa: 0.91) and Niigata (OA: 94.50%, Kappa: 0.86). Comparative analysis of multi-temporal imagery identified the rice harvest period as the optimal phenological window for extraction due to maximized spectral and textural contrast. Ablation studies revealed that the iterative mechanism significantly improved the Kappa coefficient from 0.79 to 0.89 compared to the baseline. Moreover, in a direct comparison with typical supervised learning models (U-Net and DeepLabV3+) trained on local samples, our method achieved superior extraction accuracy and boundary completeness despite using no training samples whatsoever, proving its robustness against sample scarcity. Conclusion This study presents an effective framework for high-precision, plot-level rice mapping that significantly reduces dependency on annotated samples. By successfully coupling the powerful, generic segmentation capability of SAM with agricultural remote sensing prior knowledge through an adaptive iterative mechanism, the method overcomes the "sample dependence" bottleneck of traditional deep learning. It offers a promising technical pathway for large-scale, cost-effective, and automated crop mapping, providing essential data support for food security assessment and sustainable precision agricultural management.

**Key words:** Remote sensing fine extraction, SAM, prior knowledge fusion, rice mapping, adaptive statistical learning, parcel-level segmentation, iterative optimization, prompt information optimization

**Supported by** Supported by Program for Innovative Research Groups of the National Natural Science Foundation of China (No. 41921001); Beijing Smart Agriculture Innovation Consortium Project (BAIC10-2025)