doi:10.6041/j.issn.1000-1298.2025.06.041

基于优选多源遥感特征和双分支卷积神经网络的茶园提取方法

林欣怡[1,2]　汪小钦[1]　李蒙蒙[1]　金时来[3]　龙　江[1]　冯晓敏[1]
吴瑞姣[4]　林敬兰[3]　李　琳[3]

(1. 福州大学数字中国研究院(福建), 福州 350108; 2. 中国电建集团福建省电力勘测设计院有限公司, 福州 350003;
3. 福建省水土保持试验站, 福州 350001; 4. 福建省地质测绘院, 福州 350011)

摘要: 准确的茶园分布信息可以为土地利用规划、种植布局优化提供科学的决策支撑,有助于推动茶产业可持续发展。本文基于 GF-2 PMS 影像的 RGB 波段,Sentinel-2 光学影像计算的 NDVI,Sentinel-1 时序 SAR 数据构建的物候特征(包括茶树生长幅度(Growth amplitude,GA)和生长期长度(Growth length,GL)),以及 GF-7 立体像对影像计算的坡向、坡度、曲率,构建了茶园多模态遥感特征,并通过随机森林特征优选出最佳组合。利用双分支网络联合学习策略,以 AMLNet(Attentional multiscale lightweight encoder-decoder network)为第 1 分支,Vanilla AMLNet 为第 2 分支,构建耦合多模态信息的双分支网络模型 MIPBNet(Multi-modal information parallel branch network);利用特征融合模块(Dual-branch feature fusion block,DBFF)在解码器末端进行特征级融合;利用复合损失函数进行优化训练。研究结果表明:NDVI + GA + 坡向 + 坡度组合最能提高茶园分类精度。基于 RGB 数据依次加入 NDVI、GA、坡向、坡度的组合方案,实验结果表明,融合多模态特征后,茶园提取结果漏提和误提现象明显减少,总体精度提升 3.11 个百分点。与典型的语义分割模型 UNet、UNeXt、Segformer 相比,MIPBNet 的单分支 AMLNet 获得了更优的茶园提取结果。

关键词: 茶园提取; 多源遥感; 深度学习; 语义分割; 多模态; 双分支卷积神经网络

中图分类号: TP751; S127　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)06-0446-11　　**OSID**:

Tea Plantation Recognition Method Based on Preferred Multi-source Remote Sensing Features and Two-branch Convolutional Neural Network

LIN Xinyi[1,2]　WANG Xiaoqin[1]　LI Mengmeng[1]　JIN Shilai[3]　LONG Jiang[1]
FENG Xiaomin[1]　WU Ruijiao[4]　LIN Jinglan[3]　LI Lin[3]

(1. *The Academy of Digital China(Fujian)*, *Fuzhou University*, *Fuzhou* 350108, *China*
2. *Power China Fujian Electric Power Engineering Co., Ltd.*, *Fuzhou* 350003, *China*
3. *Soil Conservation Experimental Station of Fujian Province*, *Fuzhou* 350001, *China*
4. *Fujian Geologic Surveying and Mapping Institute*, *Fuzhou* 350011, *China*)

Abstract: Accurate information on tea plantation distribution provides scientific support for land use planning and optimization of planting layouts, contributing to the sustainable development of the tea industry. Multimodal remote sensing features of tea plantation were constructed based on RGB bands from GF-2 PMS imagery, NDVI calculated from Sentinel-2 optical imagery, phenological characteristics derived from Sentinel-1 time-series SAR data, including growth amplitude, GA, and growth length, GL), and slope aspect, slope gradient, and curvature calculated from GF-7 stereo imagery. The optimal feature combination was selected through a random forest feature selection algorithm. A dual-branch network model, multi-modal information parallel branch network (MIPBNet), was built by using a multi-network joint learning strategy, with attentional multiscale lightweight encoder-decoder network (AMLNet) as the first branch and Vanilla AMLNet as the second branch. A feature fusion module (dual-

收稿日期: 2024-11-05　修回日期: 2025-02-05

基金项目: 福建省水利科技项目(MSK202214、MSK202431)和福建省科技创新基金项目(2022C0024)

作者简介: 林欣怡(1998—),女,硕士生,中国电建集团福建省电力勘测设计院有限公司助理工程师,主要从事遥感技术与应用研究,E-mail: 343416403@qq.com

通信作者: 汪小钦(1972—),女,研究员,博士,主要从事遥感信息模型与方法、资源与环境遥感研究,E-mail: wangxq@fzu.edu.cn

branch feature fusion block, DBFF) was utilized for feature-level fusion at the end of the decoder, and a composite loss function was employed for optimization training. The research findings were as follows: the combination of NDVI, GA, slope aspect, and slope gradient best improved classification accuracy and was identified as the optimal multi-modal feature set. When RGB data was sequentially augmented with NDVI, GA, slope aspect, and slope gradient, experiments showed a significant reduction in both omitted and falsely extracted tea plantation areas, with an improvement in overall accuracy (OA) of 3.11%. Compared with typical semantic segmentation models such as UNet, UNeXt, and Segformer, the single-branch AMLNet within MIPBNet achieved superior tea plantation extraction results.

Key words: tea plantations extraction; multi-source remote sensing; deep learning; semantic segmentation; multi-modal; two-branch convolutional neural network

0　引言

茶产业是乡村振兴的特色产业之一,对当地经济和就业具有重要意义。据国家统计局数据显示,2020 年中国实有茶园面积 3.22×10^6 hm^2,2022 年达 3.39×10^6 hm^2。中国茶园面积占全球约 60%,茶叶产量长期位居全球第一[1]。获取准确的茶园分布信息可以帮助农业管理部门更好地规划土地利用、种植布局和资源配置,为茶产业健康发展提供重要支撑。

传统的人工调查和手工绘制茶园方式数据获取成本高、耗时长,利用遥感技术获取茶园分布信息成为高效的监测手段[2]。中等分辨率遥感影像易于获取,被广泛应用于较大尺度范围[3-4]。但是与大田作物相比,茶园多零散分布于地理结构复杂的山区,与自然或人工生态系统(如森林、果园、农田边缘)交错分布,形成了丰富的生态边缘效应。中等分辨率影像在识别茶园方面存在局限[5-6],高分辨率影像(如 ZY-3、GF-1/2 等)能够提供更精细的空间细节,从而有效解决混合像元问题[7-9]。高空间分辨率也带来了同一特征内部异质性的增强,这对分类算法的性能构成了新的挑战[10-11]。

深度学习语义分割网络具备对代表性特征进行层次学习能力,以端到端的方式学习深浅层语义信息,最终生成与输入影像精确对应的分割图[12-15]。在已有的研究中,ZHANG 等[16]提出了一种将卷积神经网络与多层感知机(Multilayer perceptron,MLP)融合的高分辨率影像混合深度学习地表分类方法,构建具有不同上下文输入尺寸的 CNN 模型,以学习多尺度特征,同时保留原始的高分辨率信息;VALANARASU 等[17]集成带轴向偏移的 Tokenized MLP(Tokenized multilayer perceptron)模块,结合 MLP 的 CNN 模型可以有效表示空间特征和光谱特征,取得了优异的分类性能。因此,这种方法不仅提高了分类的准确性,还增强了模型对不同尺度特征的敏感度。然而,作物特征复杂多样,涵盖生长状况、物候规律及对环境条件的需求和响应等多个维度。在面向自然背景复杂的茶园场景,单模态数据信息不够全面[18-21]。因此,综合利用多模态遥感特征以实现高精度茶园分布的智能解译,展现出了广阔的应用前景[4,22]。张文凯等[23]提出的多源特征自适应融合网络,能够充分利用多模态数据的互补优势,加强了深度语义分割网络对视觉相似性的区分能力。通过整合可见光、雷达和地形等多模态遥感数据,可以获取到更全面、多层次的信息。但不同来源的数据集之间存在差异,使特征提取融合变得困难[24-26]。为使分散、孤立的数据能够在统一的数据空间内进行关联和整合,与仅使用单个网络相比,双分支网络联合学习可以灵活地互补不同网络之间的优势,弥补每个网络中存在的推理缺陷。

考虑到不同尺度的遥感数据(高分辨率与中等分辨率)在表达地物信息时各有优势,本文综合利用 GF-2 影像、Sentinel-1 SAR 数据、Sentinel-2 光学影像、GF-7 立体像对影像等多源遥感数据,以武夷山为研究区,构建耦合多模态信息的双分支网络模型(Multi-modal information parallel branch network,MIPBNet),深度挖掘并融合茶园的多模态遥感特征,以提高茶园分布提取方法的准确性和鲁棒性。

1　研究区与研究数据

1.1　研究区概况

武夷山市位于福建省北部,是福建省南平市下辖的一个县级市,处于北纬 27°27′~28°04′、东经 117°37′~118°19′之间(图 1)。该市地势多山且较为起伏,山脉交错,河流纵横,茶园、梯田以及山脉之间形成了独特的地貌特征,被誉为中国茶的故乡,这里出产的武夷岩茶是中国最有名的茶叶之一,历史悠久,品质优良。

研究区范围如图 1a 所示,面积为 222.266 km^2。该区域位于武夷山市中部地区,茶树形成了独特的

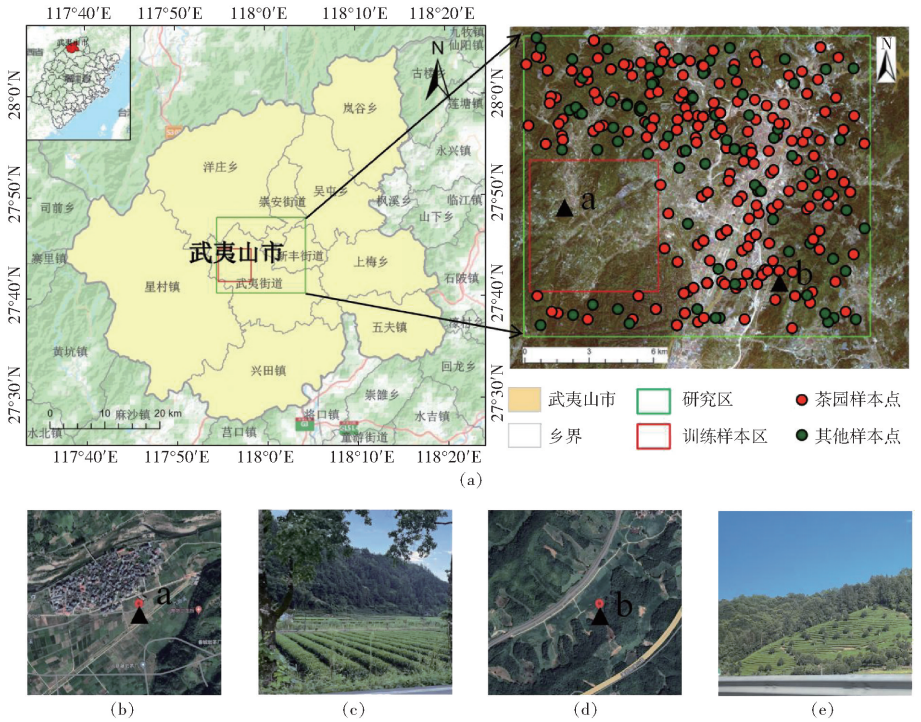

图 1 研究区地理位置

Fig. 1 Geographic location and spatial distribution of study area

灌木丛系统,与通常集中在平坦低地的农田不同,茶树生长在平原、山脉和丘陵等各种地形上,通常位于坡度 30°以下的山坡,由于茶树是一种喜欢光照的植物,阳坡因其光照充足而利于茶叶生长与发育。根据野外实地考察,茶园种植形式多为梯田式种植,如图 1d、1e 所示,平地式较少,如图 1b、1c 所示。

1.2 研究数据

1.2.1 遥感数据及预处理

本文所用数据包括高分二号(GF-2)PMS 影像、Sentinel-1 SAR 数据、Sentinel-2 光学影像、高分七号(GF-7)立体像对影像,具体信息及用途如表 1 所示。在多云多雨的南方广大地区,高分辨率光学影像可能 1~2 a 只覆盖 1 次,通常无法在作物生长最佳时间获取数据。引入中等分辨率时序遥感数据,可以在一定程度上反映作物的生长过程,弥补高分辨率光学影像时间序列的不足。

表 1 数据源信息及用途

Tab. 1 Data source information and use

数据名称	空间分辨率/m	获取时间	主要用途
GF-2 PMS 影像	0.8/3.2	2021-04-17	高分辨率光学影像多尺度特征
Sentinel-1 SAR 影像	10	2021	物候特征构建(通过时序 SAR 数据计算物候特征)
Sentinel-2 光学影像	10	2021	植被指数特征构建(获取年 NDVI 最大值合成图)
GF-7 立体像对影像	1	2022-04-06	地形特征构建(计算坡度、坡向和曲率数据)

(1)GF-2 PMS 影像。预处理包括辐射定标、正射校正、大气校正和 Gram-Scmidt 融合等,获得分辨率 0.8 m 的多光谱影像。

(2)Sentinel-1 时间序列 SAR 数据。采用干涉宽幅(IW)模式下的地面距离探测(Ground range detected,GRD)数据,已经过多视处理和地距转换,空间分辨率 10 m,2021 年 1 月 9 日获得第 1 景,后每隔 12 d 获得 1 景,共获得覆盖研究区影像 30 景 VV 和 VH 极化数据。

(3)Sentinel-2 光学影像。基于 GEE 云平台获取 2021 年的 Sentinel-2 数据的 Level-1C 产品。筛选了云量和冰雪覆盖百分比均低于 10% 的 Sentinel-2 影像,利用 QA60 波段标记去云,对获得的无云影像进行镶嵌合成并裁剪。使用 Sentinel-2 的近红外波段(B8)和红波段(B4)计算得到 NDVI 影像集合,进行最大值合成,获得 2021 年的 NDVI 最大值合成图。

(4)GF-7 立体像对影像。利用 GF-7 前视 PAN 和后视 PAN 的 L1A 影像(经过预处理的辐射校正产品),使用 RSD 软件先进行核线约束 NCC 匹配,然后手动采集关联点,获得空间分辨率为 1 m 的 DEM。

由于影像来自不同的卫星和传感器,以融合后的 GF-2 影像为基准进行地理配准,使相同地物的地理位置对齐,并采用最邻近法,将其他数据空间分辨率统一重采样为 0.8 m。

1.2.2　样本数据集

模型训练需要与影像数据对应的标签。茶园训练样本数据集制作使用 ArcGIS 软件内置的面矢量构建方法，将人工标注的茶园矢量转换为栅格获得研究区内的初始茶园样本数据集，训练样本位于图 1 研究区红色框区域范围。基于初始训练样本，通过划窗裁剪将影像和对应的标签数据分成 1 200 个 256 像素 ×256 像素的茶园样本数据，用于模型输入。为增加训练样本多样性并改善过拟合问题，经过水平翻转、垂直翻转、顺时针旋转 90° 等数据增强，生成 3 600 个影像和对应的标签。

精度评价数据分为面数据与点数据。面数据来自于在研究区局部区域勾勒的面样本（图 2）。点数据由 ArcGIS Pro 分层随机采样获得（图 1），其中茶园样本点 235 个、其他样本点 119 个。样本点具体获取流程为：①通过提取的茶园分类图分层随机抽样生成 354 个样本点。②基于 Google Earth 地球历史影像数据进行实物判别，修改错误点位属性信息，生成正确的精度评价样本点数据。

2　研究方法

构建耦合多模态信息的双分支网络模型（Multi-modal information parallel branch network，MIPBNet），开展高精度茶园分布提取研究（图 3，图中，GA 为茶树生长幅度；GL 为茶树生长期长度；NDVI 为归一化植被指数）。主要包括茶园特征构建和优选、双分支网络模型构建 2 部分。

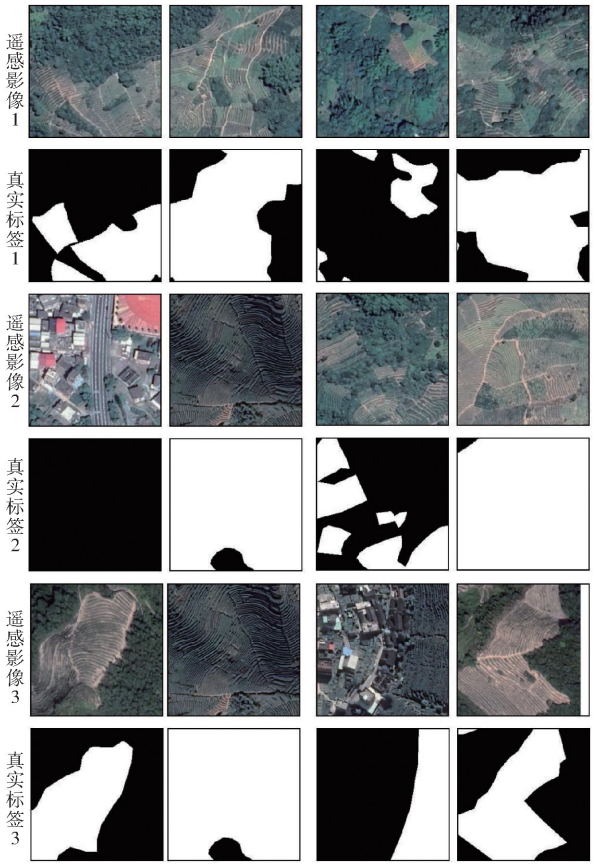

图 2　精度评价面样本

Fig. 2　Sample accuracy evaluation surface

2.1　多模态遥感特征构建与优选

2.1.1　特征构建

（1）基于时序 SAR 数据后向散射的物候特征计算物候指植被长期生长中对环境的适应过程[27]。

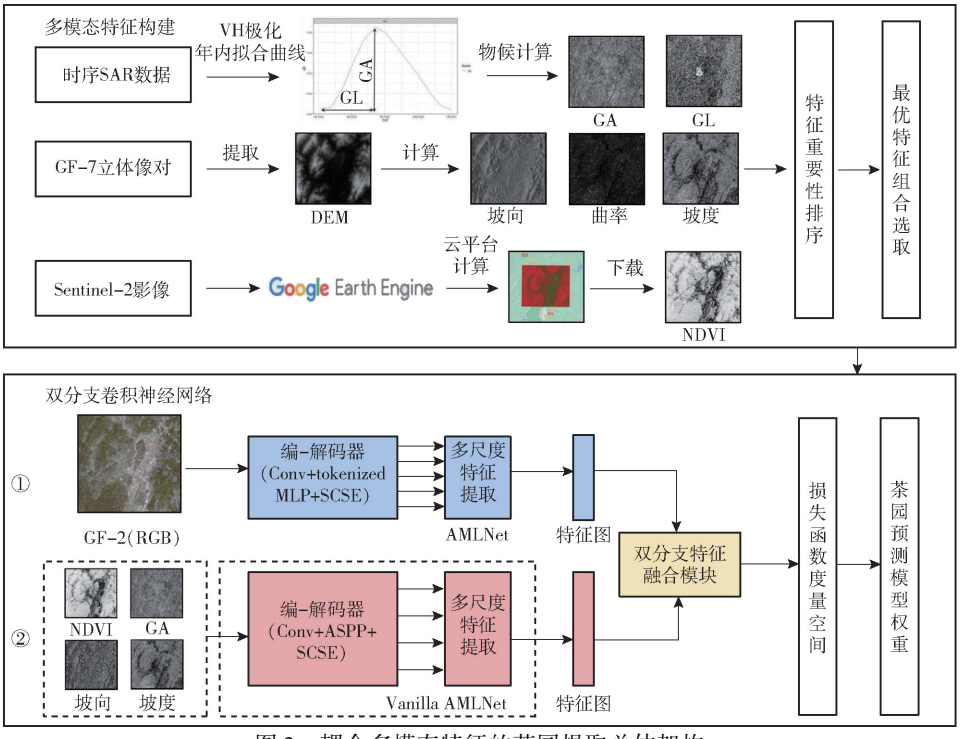

图 3　耦合多模态特征的茶园提取总体架构

Fig. 3　General architecture for tea plantation extraction with coupled multi-modal features

武夷山茶树生长周期分为生长期(3—9 月,嫩芽生长)和霜冻期(9 月至次年 3 月,应对寒冷),这导致其冠层形态变化,从而影响 SAR 后向散射系数。

基于 2021 年时间序列 Sentinel-1A SAR 数据,每类地物随机选择 6 个样本点并取平均,分析了茶园(VH 和 VV 极化)及不同地表类型(不透水面、林地、农田、裸地、水体)的 VH 极化后向散射系数(图 4)。茶园 VV 极化后向散射系数稳定在 -7.6 dB ~ -7 dB,而 VH 极化变化范围更大(-16 dB ~ -13.5 dB),表明 VH 极化对茶园生长周期更敏感。不透水面、林地、农田、裸地、水体 VH 极化后向散射系数分别分布在 -10.49 ~ -10.47 dB、-16 ~ -14.8 dB、-17 ~ -15.90 dB、-17.5 ~ -16 dB、-21.15 ~ -21.11 dB。茶园与其他地类年内特征不同,峰值时间差异显著。

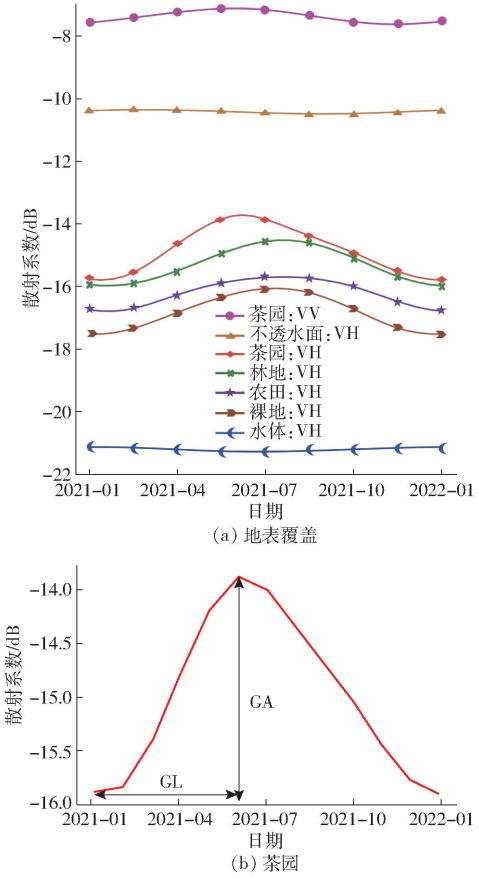

图 4 后向散射系数变化曲线

Fig. 4 Variation curve of backscattering coefficient

如图 4b 所示,构建反映茶树生长幅度(Growth amplitude,GA)和生长期长度(Growth length,GL)的物候特征,计算公式为

$$X = m_{VH} - n_{VH} \tag{1}$$

$$Y = m_{Doy} - n_{Doy} \tag{2}$$

式中 X——生长幅度 Y——生长期长度

m_{VH}——VH 极化局部峰值

n_{VH}——VH 极化局部低值

m_{Doy}——年积日峰值

n_{Doy}——年积日低值

(2)基于 GF-7 立体像对影像的精细地形因子计算

坡度适中、坡向适宜及剖面曲率小之地,利于茶树生长。利用 GF-7 立体像对提取的 DEM 计算得到高空间分辨率的坡度(slope)、坡向(aspect)和曲率(curvature)数据来表征地形特征。

(3)基于 Sentinel-2 光学影像的 NDVI 计算

高质量的植被指数数据可以精准反映地表植被覆盖状况。茶树作为常绿植物,茶园年内 NDVI 最大值可以反映其生长状态。

2.1.2 基于随机森林的特征优选

随机森林(Random forest,RF)[28] 是一种基于 CART 决策树的多分类器集成机器学习方法,能评估每个特征变量对于特定分类任务的重要性。利用基尼指数优选特征,可排除冗余和不相关特征。

2.2 耦合多模态特征双分支网络模型构建

MIPBNet 采用基于编-解码器的双分支网络架构,包括提取 RGB 影像特征的 AMLNet(Attentional multiscale lightweight encoder-decoder network)分支、提取多模态遥感特征的 Vanilla AMLNet 分支,以及双分支特征融合模块(Dual-branch feature fusion,DBFF)。

2.2.1 集成多尺度特征的图像分割网络 AMLNet

AMLNet 模型结构如图 5 所示(图中,H、W 分别表示输入特征图高度、宽度),采用编-解码器体系结构的轻量级深度学习框架。集成带轴向偏移的 Tokenized MLP(Tokenized multilayer perceptron)模块[17] 和通道空间双注意力机制(Spatial and channel squeeze & excitation,SCSE)模块[29],可以减少参数和计算复杂度,并改善连续下采样过程的特征丢失。多尺度特征提取(Multi-scale feature extraction,MSE)模块通过 1×1 卷积调整各层解码器特征通道数为 1,经批归一化处理后,使用双线性插值上采样统一特征尺寸,最后在连接层融合成统一尺度的多通道特征图,由于每个通道特征有不同的感受野,MSE 模块在保留茶园多层级细节信息的同时集成多尺度语义特征。

2.2.2 多模态信息聚合网络 Vanilla AMLNet

构建多模态信息聚合网络 Vanilla AMLNet 用于多模态茶园特征提取。Vanilla AMLNet 模型架构如图 6 所示,通过减少模型层数、引入卷积和空洞空间金字塔池化(Atrous spatial pyramid pooling,ASPP)[30] 模块增强感受野,并优化多尺度特征提取

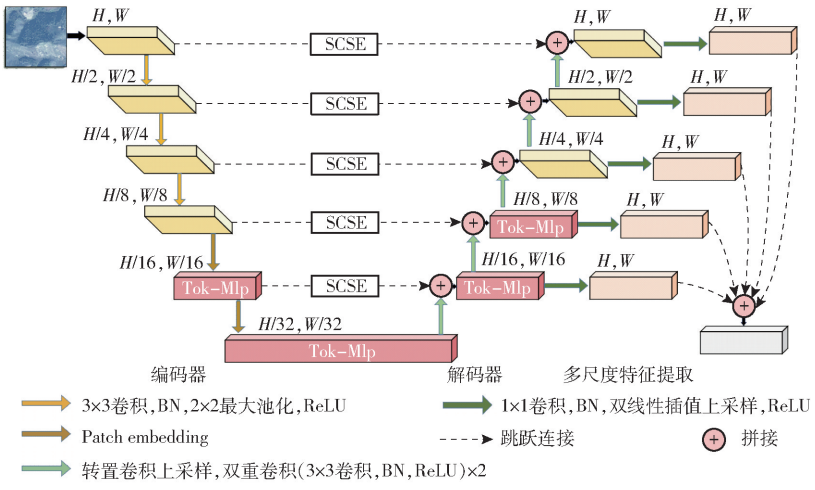

图 5　AMLNet 模型总体结构图

Fig. 5　General structure of AMLNet model

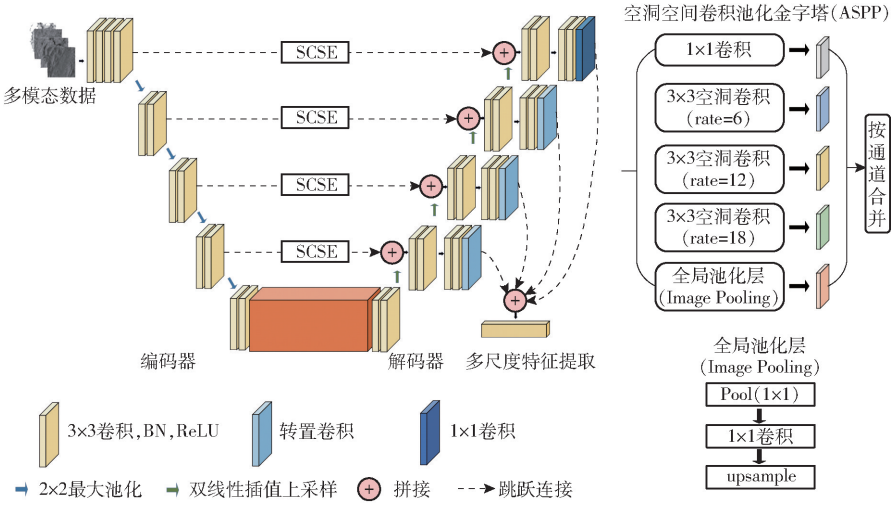

图 6　Vanilla AMLNet 模型结构图

Fig. 6　Vanilla AMLNet model structure diagram

(使用转置卷积替代双线性插值进行上采样),构建了一个高效的主干网络,使得编码器能学习茶园多模态特征的隐性表达,解码器则整合图像高层级特征,最终通过子解码器输出拼接,将不同模态的特征映射到共享的低维空间中。

2.2.3　双分支特征融合模块

不同模态的特征融合是指将来自不同传感器或数据源的特征进行合并,以提高模型的性能和泛化能力[31]。如图 7(图中,C 表示输入影像通道数)所示,采用结合注意力机制的级联特征融合方法,构建了双分支特征融合模块(Dual-branch feature fusion, DBFF),在解码器子网络上聚合 2 个分支提取的特征。DBFF 模块根据输入场景上下文动态地重新校准特征映射,使模型能自适应调整关注点,减少冗余特征和噪声的影响。

2.2.4　复合损失函数构建

损失函数是评估模型预测与真实标签之间差异的指标,而复合损失函数结合不同的损失计算方式,可以平衡不同类别样本对模型训练的影响,提升整体预测性能。

自定义复合损失函数结合了二进制交叉熵损失值 l_{bce} 和 Dice 损失值 l_{dice}。其中,交叉熵(Binary cross-entropy, Bce)损失函数在梯度下降计算中能有效避免梯度弥散问题[32];Dice 系数(Dice coefficient, Dice)损失函数可以衡量预测结果与真实标签之间重叠程度,在计算交并比时忽略大量背景像素[33]。茶园样本中背景占比高,可以结合 l_{bce} 和 l_{dice} 处理训练不稳定和正负样本不平衡的现象[34-36]。

给定一个带有 N 个像素 $x_M^i(i=1,2,\cdots,N)$ 的真实掩膜影像 I_M。设 y_M^i 是像素 $x_M^i \in \{0,1\}$ 的标签值,$\hat{y}_M^i$ 是在 Sigmoid 函数激活后的预测的掩膜标签,因此,二进制交叉熵损失值 l_{bce} 定义为[37]

$$l_{\text{bce}} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_M^i \ln \hat{y}_M^i + (1-y_M^i)\ln(1-\hat{y}_M^i)\right] \tag{3}$$

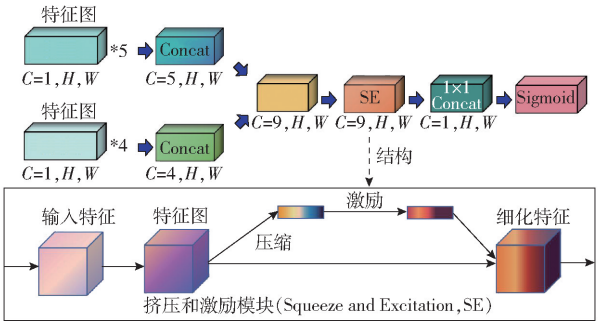

图 7 双分支特征融合模块

Fig. 7 Dual-branch feature fusion module

设 $\boldsymbol{P}$ 和 $\boldsymbol{T}$ 表示预测的掩膜影像 $\hat{I}_M$ 和真实掩膜影像 I_M 的标签向量，$|\boldsymbol{P}|$ 和 $|\boldsymbol{T}|$ 表示 $\boldsymbol{P}$ 和 $\boldsymbol{T}$ 范数，$|\boldsymbol{P}\cap\boldsymbol{T}|$ 为交集范数，因此，Dice 损失值 l_{dice} 定义为[38]

$$l_{\text{dice}}=1-\frac{2|\boldsymbol{P}\cap\boldsymbol{T}|+\varepsilon}{|\boldsymbol{P}|+|\boldsymbol{T}|+\varepsilon} \tag{4}$$

式中，取 $\varepsilon=10^{-5}$ 从而避免分母出现零值。

则复合损失函数 l_M 为

$$l_M=0.5l_{\text{bce}}+0.5l_{\text{dice}} \tag{5}$$

2.3 精度评价方法

利用混淆矩阵计算精确率(Precision)、召回率(Recall)、总体精度(Overall accuracy)、F1 值(F1-Score)、误提率和漏提率对茶园提取精度进行评价。

3 结果分析

3.1 多模态特征优选结果

深度学习模型的地物提取效果受输入数据有效性影响大[39]。基于 RF 对构建的茶园遥感特征进行筛选。通过 1 046 个样本点(茶园 414 个，其他 632 个)评估特征重要性(以特征重要性指标度量，图 8)，发现 NDVI 对茶园识别最重要，GA 次之，坡度、坡向影响较小，GL、曲率贡献度最低。

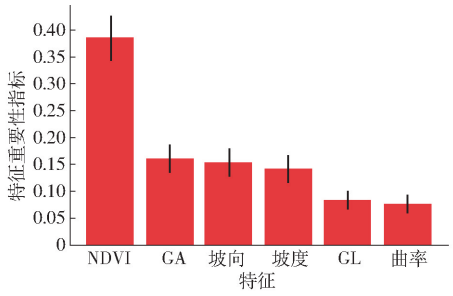

图 8 特征重要性柱状图

Fig. 8 Feature importance chart

基于 RF 分析不同特征组合下总体分类精度的变化，结果如表 2 所示。当仅使用 NDVI 作为特征时，总体精度为 67.52%；随着加入 GA、坡向、坡度，总体精度逐渐提升至 74.84%。进一步添加 GL 和曲率后，总体精度反而下降。因此，NDVI + GA + 坡向 + 坡度组合最能提高分类精度，被选为深度学习网络辅助输入数据。

表 2 不同特征组合的随机森林结果分类精度

Tab. 2 Classification accuracy for random forest results with different combinations of features

特征	总体精度/%
NDVI	67.52
NDVI + GA	62.74
NDVI + GA + 坡向	70.70
NDVI + GA + 坡向 + 坡度	74.84
NDVI + GA + 坡向 + 坡度 + GL	73.56
NDVI + GA + 坡向 + 坡度 + GL + 曲率	74.52

3.2 训练策略及参数设置

模型输入数据按照比例 9∶1 划分为训练集和验证集。使用初始学习率 10^{-4} 的 Adam 优化器，学习率衰减方式为余弦退火，训练批次大小设置为 8，迭代次数为 200。

基于复合损失函数衡量模型预测性能，每个批次将输出误差反向传播到网络参数以拟合样本的输出，并使用交并比(Intersection over union，IoU)在每个批次中衡量模型的性能，当验证集 IoU 最高时保存模型最优权重文件。

3.3 茶园提取实验

基于 MIPBNet 提取茶园时，Vanilla AMLNet 分支依次在 RGB 基础上融合 NDVI、GA、坡向及坡度数据，结果如图 9 所示。仅依赖 RGB 数据会导致误提与漏提现象；增入 NDVI 与 GA 后，漏提情况有所缓解；进一步纳入坡向与坡度信息，道路、居民区等误提大大减少(图 9 红圈处)，茶园图斑的完整性更高。这表明，逐步整合多模态数据能高效利用信息互补优势，提高茶园提取准确性。

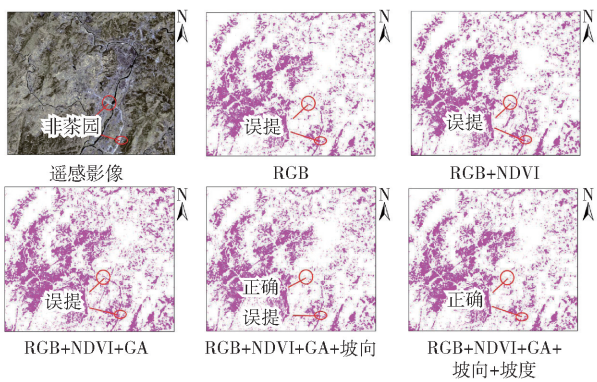

图 9 不同输入参数组合实验提取结果

Fig. 9 Extracted results of ablation experiments are plotted

茶园精度评价结果如表 3 所示，随着多模态遥感特征的逐步引入，总体精度持续上升。相较于仅使用 RGB 影像，加入 NDVI 和 GA 后总体精度和 F1

表 3　精度评价结果

Tab. 3　Accuracy evaluation results　%

数据	精确率	召回率	总体精度	F1 值	误提率	漏提率
RGB	90. 17	93. 77	89. 55	91. 94	9. 83	6. 22
RGB + NDVI	93. 01	91. 03	89. 55	92. 01	6. 99	8. 97
RGB + NDVI + GA	93. 22	94. 02	91. 52	93. 62	6. 78	5. 98
RGB + NDVI + GA + 坡向	95. 79	91. 52	91. 74	93. 61	4. 21	8. 48
RGB + NDVI + GA + 坡向 + 坡度	96. 44	92. 34	92. 66	94. 35	3. 56	7. 66

值分别提升 1. 97 个百分点和 1. 68 个百分点；进一步引入坡向和坡度后，精确率显著提升，误提率从 9. 83% 降至 3. 56%。最终，RGB + NDVI + GA + 坡向 + 坡度数据组合下总体精度达到 92. 66%，优于仅使用 RGB 影像的 89. 55%，提升 3. 11 个百分点。使用精度评价面数据在 RGB + NDVI + GA + 坡向 + 坡度这一最优数据组合的基础上，MIPBNet 取得总体精度为 92. 66%、F1 值为 94. 35%，实现了高精度茶园自动提取。

值得注意的是，坡度坡向等地理特征并非茶园独有。引入这些特征确实可以增大成功识别茶园的概率，但也可能增加将其他具有相似坡度坡向特征的地物错误识别为茶园的风险。为了平衡这一风险，本文在特征选择中采用 RF 特征优选，组合特征以准确区分茶园与其他地物。

3. 4　对比实验

现有常用的 UNet 模型[40]、UNeXt 模型[17]、Segfomer 模型[41]适用于 RGB 数据，以 MIPBNet 的单分支 AMLNet 和集成多模态数据的 MIPBNet 与其对比。

其中，UNet 作为编码器-解码器基准模型，UNeXt 通过卷积与 MLP 混合结构提升效率，Segformer 利用 Transformer 捕获全局特征。基于 UNet、UNeXt、Segformer 以及提出的 AMLNet 提取研究区茶园。图 10 展示部分区域的茶园提取结果。整体上看，如图 10a 所示，AMLNet、UNet、UNeXt、Segformer 4 种模型均能有效提取茶园。细节上看，如图 10b、10c 所示，AMLNet 提取的茶园边界与真实茶园边界的吻合度更高，在边缘等细节的分割效果更精细。相比之下，UNet 和 Segformer 在提取过程中容易出现图斑粘连的情况，导致茶园区域的误判。而 UNeXt 在提取过程中则出现了较多的孔洞，使得提取结果呈现出一定的破碎化。

由表 4 可看出，MIPBNe、AMLNet、UNet、UNeXt、

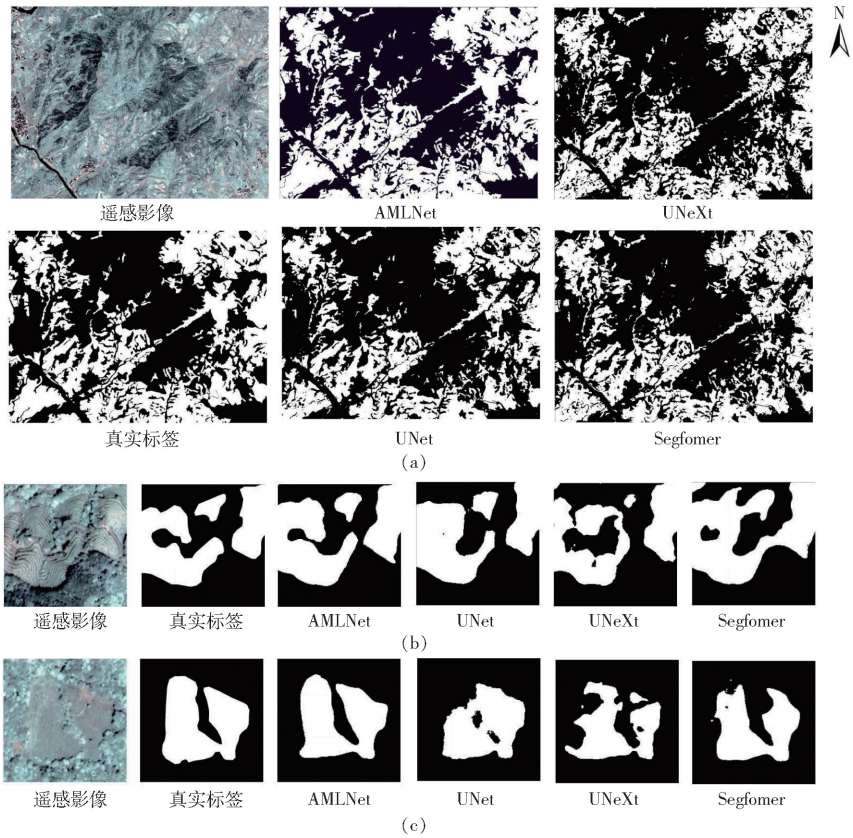

图 10　不同方法的茶园提取结果

Fig. 10　Distributions of tea garden extraction results by different methods

表 4 精度评价结果

Tab. 4 Accuracy evaluation results

方法	参数量	精确率/%	召回率/%	总体精度/%	F1 值/%	误提率/%	漏提率/%
MIPBNet	8.26×10^6	96.44	92.34	92.66	94.35	3.56	7.66
AMLNet	5.54×10^6	90.17	93.77	89.55	91.94	9.83	6.22
UNet	7.65×10^6	93.57	80.40	85.63	86.49	6.43	19.60
UNeXt	1.47×10^6	92.81	77.89	83.91	84.70	7.19	22.11
Segfomer	2.47×10^7	96.63	79.15	86.49	87.02	3.37	20.85

Segformer 的参数量分别为 8.26×10^6、5.54×10^6、7.65×10^6、1.47×10^6、2.47×10^7，MIPBNet、AMLNet、UNet、UNeXt 均采用轻量级结构。精度评价结果表明，在 RGB 数据下，AMLNet 获得了最高的总体精度(89.55%)、F1 值(91.94%)；集成多模态数据的 MIPBNet 的总体精度提升 3.11 个百分点，F1 值提升 2.41 个百分点，具有更优的性能。

4 结论

(1)针对茶园提取任务中面临的复杂性与高精度要求，提出了耦合多模态信息的双分支网络模型 MIPBNet，综合多模态遥感特征，在复杂环境地区实现高精度自动化茶园提取。

(2)针对茶园与其他地物在遥感影像中难以区分、易受环境因素影响等难点，利用多源遥感数据，构建了 NDVI、GA、坡向、坡度、GL、曲率 6 个多模态遥感特征，并基于 RF 筛选出最佳组合，即 NDVI + GA + 坡向 + 坡度。

(3)随着多模态遥感特征的逐步融入，基于 MIPBNet 的茶园提取精度呈上升趋势，茶园提取结果漏提和误提现象明显减少，相比于仅使用 RGB 影像总体精度提升 3.11 个百分点。

(4)与经典的语义分割模型 UNet、UNeXt、Segformer 相比，MIPBNet 的单分支 AMLNet 取得更优的茶园提取结果，总体精度为 89.55%、F1 值为 91.94%，在像素分类以及边缘检测上均展现出了更高的准确性。

参 考 文 献

[1] 殷伟, 张莹, 杨一单, 等. 全球茶叶贸易网络及我国茶叶出口市场格局演变特征分析[J]. 地理与地理信息科学, 2022, 38(2):63 - 70.
YIN Wei, ZHANG Ying, YANG Yidan, et al. Analysis on the characteristics of global tea trade network and pattern evolution of China's tea export market[J]. Geography and Geo-Information Science, 2022, 38(2):63 - 70. (in Chinese)

[2] 冯权泷, 牛博文, 朱德海, 等. 土地利用/覆被深度学习遥感分类研究综述[J]. 农业机械学报, 2022, 53(3):1 - 17.
FENG Quanlong, NIU Bowen, ZHU Dehai, et al. Review for deep learning in land use and land cover remote sensing classification[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3):1 - 17. (in Chinese)

[3] 熊皓丽, 周小成, 汪小钦, 等. 基于 GEE 云平台的福建省 10 m 分辨率茶园专题空间分布制图[J]. 地球信息科学学报, 2021, 23(7):1325 - 1337.
XIONG Haoli, ZHOU Xiaocheng, WANG Xiaoqin, et al. Mapping the spatial distribution of tea plantations with 10 m resolution in Fujian Province using Google Earth Engine[J]. Journal of Geo-information Science, 2021, 23(7):1325 - 1337. (in Chinese)

[4] PENG Y F, QIU B W, TANG Z H, et al. Where is tea grown in the world: a robust mapping framework for agroforestry crop with knowledge graph and sentinels images[J]. Remote Sensing of Environment, 2024, 303:114016.

[5] YAO Z X, ZHU X C, ZENG Y, et al. Extracting tea plantations from multitemporal Sentinel - 2 images based on deep learning networks[J]. Agriculture, 2022, 13(1):10.

[6] XU W H, HUANG S D, WU C, et al. The pruning phenological phase-based method for extracting tea plantations by field hyperspectral data and Landsat time series imagery[J]. Geocarto International, 2020, 37(7):1 - 21.

[7] 柏佳, 孙睿, 张赫林, 等. 基于 GF - 1 和 Sentinel - 2 时序数据的茶园识别[J]. 农业工程学报, 2021, 37(14):179 - 185.
BAI Jia, SUN Rui, ZHANG Helin, et al. Tea plantation identification using GF - 1 and Sentinel - 2 time series data[J]. Transactions of the CSAE, 2021, 37(14): 179 - 185. (in Chinese)

[8] 黄艳红. 基于多源遥感数据的茶园面积提取[D]. 杭州: 浙江大学, 2020.
HUANG Yanhong. Estimating tea plantation area based on multi-source satellite data[D]. Hangzhou: Zhejiang University, 2020. (in Chinese)

[9] TANG Zixia, LI Mengmeng, WANG Xiaoqin. Mapping tea plantations from VHR images using OBIA and convolutional neural networks[J]. Remote Sensing, 2020, 12(18):2935.

[10] TONG Xinyi, XIA Guisong, LU Qikai, et al. Land-cover classification with high-resolution remote sensing images using transferable deep models[J]. Remote Sensing of Environment, 2020, 237:111322.

[11] 廖廓, 聂磊, 杨泽宇, 等. 基于多维卷积神经网络的多源高分辨率卫星影像茶园分类[J]. 自然资源遥感, 2022,34(2):152 – 161.
LIAO Kuo, NIE Lei, YANG Zeyu, et al. Classification of tea garden based on multi-source high-resolution satellite images using multi-dimensional convolutional neural network[J]. Remote Sensing for Natural Resources, 2022,34(2): 152 – 161. (in Chinese)

[12] 张鑫, 姚庆安, 赵健, 等. 全卷积神经网络图像语义分割方法综述[J]. 计算机工程与应用, 2022,58(8):45 – 57.
ZHANG Xin, YAO Qing'an, ZHAO Jian, et al. Image semantic segmentation based on fully convolutional neural network[J]. Computer Engineering and Applications, 2022, 58(8): 45 – 57. (in Chinese)

[13] ZHANG H Y, LIN X Y, LONG J, et al. SegFormer-based cotton planting areas extraction from high-resolution remote sensing images[C]//2023 11th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), 2023.

[14] 梁新宇, 罗晨, 权冀川, 等. 基于深度学习的图像语义分割技术研究进展[J]. 计算机工程与应用, 2020,56(2):18 – 28.
LIANG Xinyu, LUO Chen, QUAN Jichuan, et al. Research on progress of image semantic segmentation based on deep learning[J]. Computer Engineering and Applications, 2020, 56(2): 18 – 28. (in Chinese)

[15] 董秀春, 蒋怡, 杨玉婷, 等. 基于语义分割模型和遥感的柑橘园空间信息提取[J]. 中国农机化学报, 2023,44(1):178 – 184.
DONG Xiuchun, JIANG Yi, YANG Yuting, et al. Spatial information extraction of citrus orchard based on semantic segmentation model and remote sensing[J]. Journal of Chinese Agricultural Mechanization, 2023, 44(1): 178 – 184. (in Chinese)

[16] ZHANG C, PAN X, LI H P, et al. A hybrid MLP – CNN classifier for very fine resolution remotely sensed image classification [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018,140:133 – 144.

[17] VALANARASU J M J, PATEL V M. UNeXt: MLP-based rapid medical image segmentation network[J]. arXiv Preprint, arXiv:2203.04967, 2022.

[18] 欧阳松. 地学知识嵌入的遥感影像深度语义分割方法研究[D]. 武汉: 武汉大学, 2021.
OUYANG Song. The methodology research on geographic knowledge-guided deep semantic segmentation of remote sensing imagery[D]. Wuhan: Wuhan University,2021. (in Chinese)

[19] 刘俊楠, 刘海砚, 陈晓慧, 等. 面向多源地理空间数据的知识图谱构建[J]. 地球信息科学学报, 2020,22(7):1476 – 1486.
LIU Junnan, LIU Haiyan, CHEN Xiaohui, et al. The construction of knowledge graph towards multi-source geospatial data[J]. Journal of Geo-information Science, 2020, 22(7): 1476 – 1486. (in Chinese)

[20] 王志华, 杨晓梅, 周成虎. 面向遥感大数据的地学知识图谱构想[J]. 地球信息科学学报, 2021,23(1):16 – 28.
WANG Zhihua, YANG Xiaomei, ZHOU Chenghu. Geographic knowledge graph for remote sensing big data[J]. Journal of Geo-information Science,2021,23(1):16 – 28. (in Chinese)

[21] LI Yansheng, WU Kang, OUYANG Song, et al. Geographic knowledge graph-guided remote sensing image semantic segmentation[J]. Journal of Remote Sensing,2024,28(2):455 – 469.

[22] 杨劲松. 基于多模态特征融合的遥感图像语义分割方法研究[D]. 北京: 中国矿业大学, 2021.
YANG Jinsong. Research on semantic segmentation of remote sensing images based on multimodal feature fusion[D]. Beijing: China University of Mining and Technology, 2021. (in Chinese)

[23] 张文凯, 刘文杰, 孙显, 等. 多源特征自适应融合网络的高分遥感影像语义分割[J]. 中国图象图形学报, 2022, 27(8):2516 – 2526.
ZHANG Wenkai, LIU Wenjie, SUN Xian, et al. Multi-source features adaptation fusion network for semantic segmentation in high-resolution remote sensing images[J]. Journal of Image and Graphics, 2022,27(8):2516 – 2526. (in Chinese)

[24] 王萌, 王昊奋, 李博涵, 等. 新一代知识图谱关键技术综述[J]. 计算机研究与发展, 2022,59(9):1947 – 1965.
WANG Meng, WANG Haofen, LI Bohan, et al. Survey on key technologies of new generation knowledge graph[J]. Journal of Computer Research and Development, 2022, 59(9): 1947 – 1965. (in Chinese)

[25] 徐有为, 张宏军, 程恺, 等. 知识图谱嵌入研究综述[J]. 计算机工程与应用, 2022,58(9):30 – 50.
XU Youwei, ZHANG Hongjun, CHENG Kai, et al. Comprehensive survey on knowledge graph embedding[J]. Computer Engineering and Applications, 2022, 58(9): 30 – 50. (in Chinese)

[26] 张吉祥, 张祥森, 武长旭, 等. 知识图谱构建技术综述[J]. 计算机工程, 2022,48(3):23 – 37.
ZHANG Jixiang, ZHANG Xiangsen, WU Changxu, et al. Survey of knowledge graph construction techniques[J]. Computer Engineering, 2022, 48(3): 23 – 37. (in Chinese)

[27] 王敏钰, 罗毅, 张正阳, 等. 植被物候参数遥感提取与验证方法研究进展[J]. 遥感学报, 2022,26(3):431 – 455.
WANG Minyu, LUO Yi, ZHANG Zhengyang, et al. Recent advances in remote sensing of vegetation phenology: retrieval algorithm and validation strategy[J]. National Remote Sensing Bulletin, 2022,26(3):431 – 455. (in Chinese)

[28] BREIMAN L. Random forests[J]. Machine Learning, 2001,45(1):5 – 32.

[29] ROY A G, NAVAB N, WACHINGER C. Concurrent spatial and channel 'squeeze & excitation'[C]//Fully Convolutional Networks, Cham, 2018.

[30] CHEN L, PAPANDREOU G, KOKKINOS I, et al. Deeplab: semantic image segmentation with deep convolutional nets,

atrous convolution, and fully connected CRFS[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4):834 - 848.

[31] 段立娟, 孙启超, 乔元华, 等. 基于注意力感知和语义感知的 RGB - D 室内图像语义分割算法[J]. 计算机学报, 2021,44(2):275 - 291.
DUAN Lijuan, SUN Qichao, QIAO Yuanhua, et al. Attention-aware and semantic-aware network for RGB - D indoor semantic segmentation[J]. Chinese Journal of Computers,2021,44(2):275 - 291. (in Chinese)

[32] 刘浩, 骆剑承, 黄波, 等. 基于特征压缩激活 Unet 网络的建筑物提取[J]. 地球信息科学学报, 2019,21(11):1779 - 1789.
LIU Hao, LUO Jiancheng, HUANG Bo, et al. Building extraction based on SE - Unet[J]. Journal of Geo-information Science, 2019, 21(11): 1779 - 1789. (in Chinese)

[33] 常惠, 饶志强, 赵玉林, 等. 基于改进 U - Net 网络的隧道裂缝分割算法研究[J]. 计算机工程与应用, 2021,57(22): 215 - 222.
CHANG Hui, RAO Zhiqiang, ZHAO Yulin,et al. Research on tunnel crack segmentation algorithm based on improved U - Net network[J]. Computer Engineering and Applications, 2021, 57(22): 215 - 222. (in Chinese)

[34] 殷晓航, 王永才, 李德英. 基于 U - Net 结构改进的医学影像分割技术综述[J]. 软件学报, 2021,32(2):519 - 550.
YIN Xiaohang,WANG Yongcai, LI Deying. Suvery of medical image segmentation technology based on U - Net structure improvement[J]. Journal of Software,2021,32(2):519 - 550. (in Chinese)

[35] 于营, 王春平, 付强, 等. 语义分割评价指标和评价方法综述[J]. 计算机工程与应用, 2023,59(6):57 - 69.
YU Ying, WANG Chunping, FU Qiang, et al. Survey of evaluation metrics and methods for semantic segmentation[J]. Computer Engineering and Applications, 2023, 59(6): 57 - 69. (in Chinese)

[36] 张晓宇, 王彬, 安卫超, 等. 基于融合损失函数的 3D U - Net + + 脑胶质瘤分割网络[J]. 计算机科学, 2021,48(9): 187 - 193.
ZHANG Xiaoyu, WANG Bin, AN Weichao,et al. Glioma segmentation network based on 3D U - Net + + with fusion loss function[J]. Computer Science, 2021, 48(9): 187 - 193. (in Chinese)

[37] 吴止锾, 高永明, 李磊, 等. 类别非均衡遥感图像语义分割的全卷积网络方法[J]. 光学学报, 2019,39(4):401 - 412.
WU Zhihuan, GAO Yongming, LI Lei, et al. Fully convolutional network method of semantic segmentation of class imbalance remote sensing images[J]. Acta Optica Sinica,2019,39(4):401 - 412. (in Chinese)

[38] MILLETARI F, NAVAB N, AHMADI S A. V - net: Fully convolutional neural networks for volumetric medical image segmentation[C]//2016 Fourth International Conference on 3D Vision (3DV), 2016.

[39] 张悦琦, 任鸿瑞. 融合特征优选与随机森林算法的 GF - 6 影像东北一季稻遥感提取[J]. 遥感学报, 2023,27(9):2153 - 2164.
ZHANG Yueqi, REN Hongrui. Remote sensing extraction of paddy rice in Northeast China from GF - 6 images by combining feature optimization and random forest[J]. National Remote Sensing Bulletin,2023,27(9):2153 - 2164. (in Chinese)

[40] RONNEBERGER O, FISCHER P, BROX T. U - net: convolutional networks for biomedical image segmentation[C]// International Conference on Medical Image Computing and Computer-assisted Intervention, 2015.

[41] XIE Enze, WANG Wenhai, YU Zhiding, et al. SegFormer: simple and efficient design for semantic segmentation with transformers[J]. Advances in Neural Information Processing Systems, 2021,34:12077 - 12090.

(上接第 421 页)

[25] 廖庆喜,徐阳,袁佳诚,等. 油菜联合收获机切抛组合式纵轴流脱离装置设计与试验[J]. 农业机械学报,2019,50(7): 140 - 150.
LIAO Qingxi, XU Yang, YUAN Jiacheng, et al. Design and experiment on combined cutting and throwing longitudinal axial flow threshing and separating device of rape combine harvester[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019,50(7):140 - 150. (in Chinese)

[26] 连国党,魏鑫鑫,马丽娜,等. 轴流螺旋滚筒式食用向日葵脱粒装置设计与试验[J]. 农业工程学报,2022,38(17):42 - 51.
LIAN Guodang, WEI Xinxin, MA Li'na, et al. Design and experiments of the axial-flow spiral drum threshing device for the edible sunflower[J]. Transactions of the CSAE, 2022, 38(17): 42 - 51. (in Chinese)

[27] PAN F, HU B, LUO X. Design and testing of a shearing and breaking device for mulch film and cotton stalk mixtures[J]. Transactions of the ASABE, 2021(2):64.

[28] 严义其,坎杂,王吉亮,等. 膜杂捆料多级破碎机设计与试验[J]. 农业工程学报,2024,40(2):187 - 195.
YAN Yiqi, KAN Za, WANG Jiliang, et al. Design and experiment of multi-stage crusher for film and miscellaneous bundle materials[J]. Transactions of the CSAE, 2024, 40(2): 187 - 195. (in Chinese)