doi:10.6041/j.issn.1000-1298.2026.17.003

基于高分二号影像和改进 DeepLabV3 + 的冬小麦种植区提取模型研究

宋荣杰　汪　强　史汶龙　屠晓雨　张宏鸣
（西北农林科技大学信息工程学院，陕西杨凌 712100）

摘要：准确获取冬小麦的空间分布信息是长势监测和产量预估的基础。针对复杂种植环境下冬小麦遥感提取面临的光谱混淆严重、频域信息利用不足以及田块边缘细节丢失等问题，提出了一种基于国产高分二号影像和改进 DeepLabV3 + 的冬小麦种植区提取模型 RSMANet（Residual SCConv MFMSA network with adaptive feature fusion）。首先，通过引入空间和通道重构卷积（Spatial and channel reconstruction convolution，SCConv）设计 SC - ResNet18 替换原有主干网络，提高模型对空间细节和语义特征的提取能力；其次，在编码阶段设计多尺度多频注意力（SCConv with multi-scale and multi-frequency attention module，SCC - MFMSA）模块，该模块主要包括尺度分解、多频通道注意力（Multi-frequency channel attention，MFCA）和多尺度空间注意力（Multi-scale spatial attention，MSSA）3 部分，通过双注意力协同提取影像频域和空域特征，有效区分冬小麦冠层纹理与背景高频噪声，并提高模型的多尺度感知能力；最后，在特征融合阶段引入多尺度注意力融合（Multiscale attention fusion，MAF）模块，实现局部细节特征和全局语义信息的动态融合，通过增强通道、空间和多尺度层次的特征表达和语义关联，改善田块边界模糊和光谱混淆问题。试验结果表明，RSMANet 模型能有效提升冬小麦种植区提取精度，IoU 和 F1 值分别达到 89.09% 和 94.52%，较原始 DeepLabV3 + 模型分别提升 3.19、1.39 个百分点，与其他主流模型（如 CMTFNet、SACANet、TransUNet 等）相比分别提升 1.08 ~ 3.81 个百分点和 0.23 ~ 1.74 个百分点。

关键词：冬小麦；种植区提取；高分二号影像；DeepLabV3 + ；多尺度多频注意力模块
中图分类号：TP751；S127　　**文献标识码**：A　　**文章编号**：1000-1298(2026)17-0031-11　　**OSID**：

Extraction Model Study of Winter Wheat Planting Areas Based on GF - 2 Imagery and Improved DeepLabV3 + Model

Song Rongjie　Wang Qiang　Shi Wenlong　Tu Xiaoyu　Zhang Hongming
(*College of Information Engineering*, *Northwest A&F University*, *Yangling*, *Shaanxi* 712100, *China*)

Abstract: Accurate acquisition of winter wheat spatial distribution information is fundamental to growth monitoring and yield prediction. To address the problems of severe spectral confusion, insufficient utilization of frequency-domain information, and loss of field edge details in remote sensing extraction of winter wheat under complex planting environments, an improved DeepLabV3 + model named RSMANet was proposed for extracting winter wheat planting areas by using Chinese GF - 2 imagery. Firstly, SC - ResNet18 was designed to replace the original backbone network by introducing the spatial and channel reconstruction convolution (SCConv), which improved the model's ability to extract spatial details and semantic features. Secondly, a multi-scale and multi-frequency attention (SCC - MFMSA) module was constructed in the encoding stage, consisting of three components: scale decomposition, multi-frequency channel attention (MFCA), and multi-scale spatial attention (MSSA). By synergistically extracting frequency-domain and spatial-domain features of the imagery through a dual-attention mechanism, this module effectively distinguished winter wheat canopy texture from background high-frequency noise, and improved the model's multi-scale perception ability. Finally, a multi-scale attention fusion (MAF) module was introduced in the feature fusion stage to achieve dynamic fusion of local detailed features and global semantic information. By enhancing feature representation and semantic correlation at channel,

收稿日期：2026 - 04 - 28　修回日期：2026 - 05 - 14
基金项目：农业农村部科技项目和国家重点研发计划项目(2020YFD1100601)
作者简介：宋荣杰(1981—)，女，副教授，博士，主要从事机器学习与农业遥感研究，E-mail：srj@nwsuaf.edu.cn

spatial, and multi-scale levels, the problems of blurred plot boundaries and severe spectral confusion were improved. Experimental results demonstrated that the RSMANet model effectively improved the extraction accuracy of winter wheat planting areas; the IoU and F1 score reached 89.09% and 94.52%, respectively, which were 3.19 and 1.39 percentage points higher than those of the original DeepLabV3 + model, and 1.08 ~ 3.81 and 0.23 ~ 1.74 percentage points higher than those of other mainstream models (such as CMTFNet, SACANet, and TransUNet). The research results can provide a reference for winter wheat planting area extraction based on high-resolution imagery.

Key words: winter wheat; planting area extraction; GF - 2 imagery; DeepLabV3 +; multi-scale and multi-frequency attention

0 引言

冬小麦是中国主要粮食作物之一,准确获取其空间分布与种植面积对于长势监测、产量预估及农业政策制定具有重要意义[1]。传统的农业调查方法依赖人工勘测,耗时费力且在大尺度区域调查中难以保证时效性[2]。随着对地观测技术的发展,卫星遥感已成为获取农作物种植信息的有效手段[3]。现有研究大多采用 MODIS、Landsat 和 Sentinel - 2 等中低空间分辨率遥感影像,利用深度学习技术实现大尺度作物种植区提取[4-7]。我国西北地区耕地环境复杂,尤其是冬小麦种植区往往呈现地块破碎化、边界不规则的分布特征,中低分辨率影像混合像元现象严重[8],难以满足现代农业地块级精细化制图需求。

高分辨率遥感影像具有丰富的几何结构与纹理细节,为细碎田块的精准提取提供了数据基础[9],但同时也面临新的挑战:一是光谱变异性[10]增强与类间特征混淆加剧。高分辨率带来的丰富细节往往伴随着高频噪声[11],特别是在冬小麦拔节期,作物冠层郁闭度高[12],其光谱特征与同期生长的油菜、杂草极为相似,导致同谱异物现象严重[13];二是多尺度特征提取与边界保持困难。受小农户分散种植模式影响[14],冬小麦田块尺度差异显著,且田埂、道路穿插其中。传统卷积神经网络限于局部感受野[15],在提取特征时难以兼顾大尺度田块的内部一致性与小尺度田块的边缘完整性,导致提取结果中存在椒盐噪声[16]或边界粘连。

为解决高分影像中冬小麦种植区光谱混淆严重及田块边缘模糊问题,已有研究尝试对语义分割后的结果进行修正,如宋德娟等[17]通过 RefineNet 耦合最大后验概率模型,利用概率统计规律平滑粗糙的冬小麦田块边缘。但此类后处理方法高度依赖初分割质量,难以从根本上解决特征提取阶段的误判问题。为此,李旭青等[18]在 HRNet 架构中嵌入 OCR 和坐标注意力模块,通过增强像素级的上下文语义关联,提升了对零散田块的感知能力。然而,单一的空间注意力机制因聚焦局部空间特征强化、缺乏全局语义关联建模,在应对大范围背景干扰时,往往难以兼顾全局语义信息,导致在复杂背景下全局语义表征能力不足。蔡玉林等[19]进一步引入自注意力机制,利用其长距离依赖建模能力,增强目标作物与复杂背景的语义区分度。然而,现有局部与全局特征结合应用中,大多采用简单拼接或逐像素相加方法,忽略了不同层级特征在语义和分辨率上的差异,难以兼顾局部细节与全局上下文特征的有效融合,导致模型对多样化农田场景适应性不足。

此外,现有研究大多聚焦于空间域或通道域的特征增强[17-18],忽略了频域信息在区分冬小麦纹理与高频背景噪声方面的潜力,难以在抑制噪声的同时保留微小边缘细节。同时,在跨尺度特征融合阶段,深层语义特征与浅层细节特征的简单叠加容易引发语义鸿沟,导致不规则田块的边界恢复效果受限。

针对上述问题,本文提出一种基于高分二号影像和改进 DeepLabV3 +[20]的冬小麦种植区提取模型 RSMANet,通过设计 SC - ResNet18 为主干网络,利用空间与通道重构提升对冬小麦种植区特征的捕捉能力,同时降低特征冗余;构建多尺度多频注意力(SCC - MFMSA)模块,融合频域分析与多尺度卷积,在抑制非植被区高频噪声的同时,有效捕捉冬小麦种植区的多尺度上下文特征;通过多尺度注意力融合(MAF)模块动态分配权重,优化深浅层特征的交互,解决冬小麦种植区提取中的边缘模糊问题,实现复杂环境下冬小麦种植区的精细化提取。

1 研究区及数据来源

1.1 研究区概况

研究区位于陕西省扶风县(34°23′30″ ~ 34°43′20″N,107°09′20″ ~ 107°44′40″E),地处关中平原西部,属大陆性湿润季风气候,四季分明,雨量丰沛,年均气温 12.4℃,年均降水量 592 mm,无霜期 209 d。全县粮食作物播种面积 45 755 hm^2,其中冬小麦种植面积约占 3/5。研究区冬小麦一般 10 月

上旬播种，3 月下旬至 4 月上旬拔节，4 月下旬至 5 月上旬抽穗，5 月中旬左右灌浆，6 月上旬成熟收获。

研究区内冬小麦种植分布呈现多样化特点，且地块大小不一、形状不规则。此外，冬小麦地块多与果园、林地、油菜等其他作物地块交错分布，给冬小麦种植区自动化提取造成一定困难。

1.2　遥感数据及预处理

选用扶风县高分二号全色影像和多光谱影像（Gaofen 2 panchromatic and multi-spectral，GF－2 PMS）作为遥感数据源，利用 ENVI 5.6 软件对 GF－2 PMS 影像进行预处理，主要包括辐射定标、大气校正和正射校正等操作；并采用 NNDiffuse pan sharpening 方法对全色（分辨率 1 m）和多光谱影像（分辨率 4 m）进行数据融合，融合后影像分辨率为 1 m，最后裁剪获得研究区影像（图 1），影像覆盖面积 451.31 km^2。

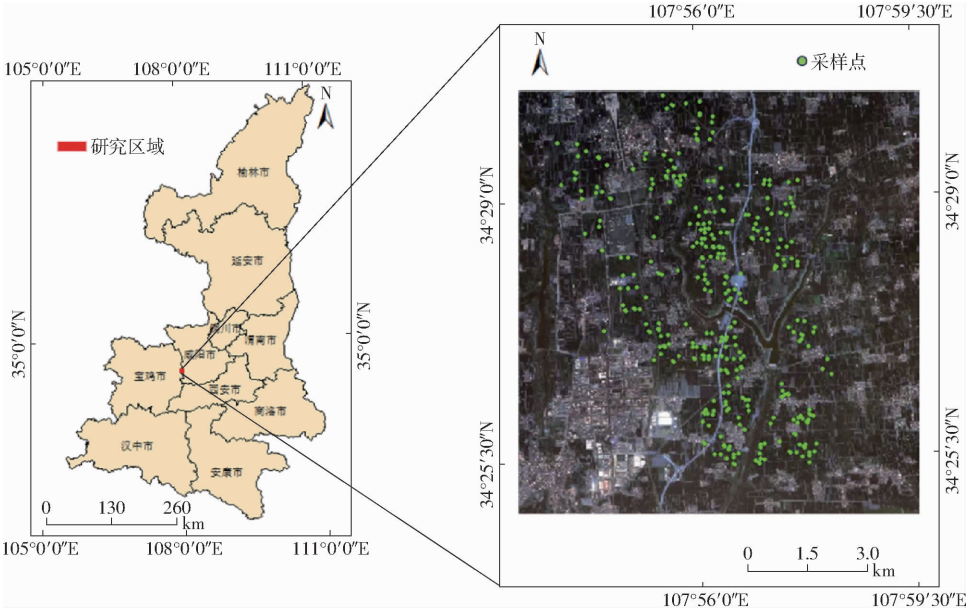

图 1　研究区及遥感影像

Fig. 1　Study area and remote sensing image

影像获取日期为 2024 年 4 月 8 日，此时冬小麦正处于拔节期，植株生长旺盛，冠层郁闭度高，在影像中呈连片的块状分布，地块内部纹理相对平滑。同时期油菜正由现蕾期向开花期转化，其物候特征与冬小麦较为相似。果园、林地呈条带状分布，影像特征表现为粗糙颗粒状纹理结构，与冬小麦种植区光谱差异相对明显。影像中冬小麦与其他作物间仍存在一定程度的光谱混淆，给冬小麦种植区的准确识别带来挑战。

1.3　数据集制作

将研究区地物类型分为冬小麦和非冬小麦 2 类。在影像同期开展实地调研，结合调研结果和目视解译，利用 ArcGIS 软件对影像中的冬小麦种植区进行人工标注，其中冬小麦种植区赋值为 1（图 2 标签图像中红色区域），其余背景赋值为 0（图 2 标签图像中黑色区域）。对影像和标签数据进行滑动窗口裁剪，重叠率 20%，窗口大小 512×512，裁剪后的原始影像、标准假彩色及对应标签如图 2 所示。

为确保数据集类别平衡，将不包括冬小麦种植区的影像及对应标签删除，得到初始影像-标签对共计 1 453 组，并按 6∶2∶2 的比例划分为训练集、验证集和测试集；通过设置固定步长间隔抽取样本，保证三者在空间上互不重叠。然后分别对影像-标签对进行水平翻转、垂直翻转和旋转等数据增强，扩充样本规模，最终得到 5 812 组样本数据，用于模型训练和性能评估。

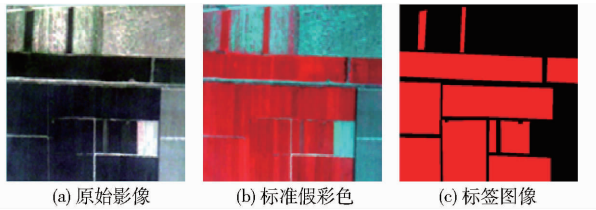

图 2　影像-标签示意图

Fig. 2　Images and labels

2　研究方法

2.1　改进的 DeepLabV3 + 模型结构

本文提出一种改进的冬小麦种植区提取模型 RSMANet，模型结构如图 3 所示。该模型对原始 DeepLabV3 + 网络结构进行以下 3 点改进：引入空间和通道重构卷积（SCConv）对不同尺度上的特征进行重构，减少冗余特征，并将改进后的 SC－ResNet18 网络替换原有主干网络 Xception，提升模

型对冬小麦种植分布特征的提取能力;在模型编码器中设计多尺度多频注意力模块(SCC - MFMSA),引入频域分析机制区分植被纹理与背景噪声,并通过通道和空间注意力强化不同尺度冬小麦地块的特征表达;受 CMTFNet 启发,在解码器特征恢复阶段,设计多尺度注意力融合(MAF)模块,替换原有的特征简单拼接方式,通过对编码器的浅层特征和解码器的深层特征进行动态加权融合,提升模型在复杂场景下的特征表达能力和语义分割精度。

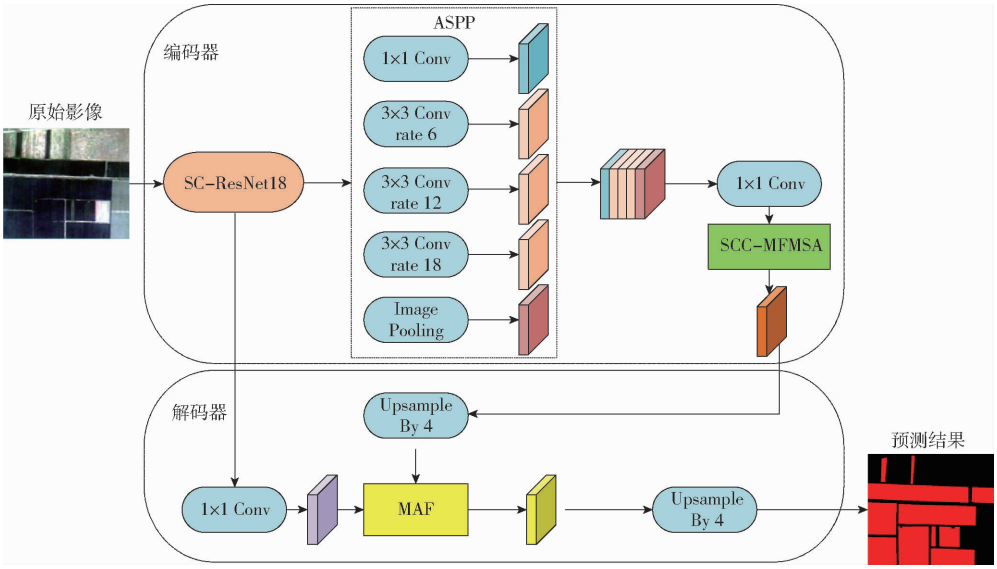

图 3 RSMANet 网络架构

Fig. 3 Framework of RSMANet

对于输入的 4 波段遥感影像,RSMANet 编码器首先利用改进后的 SC - ResNet18 主干网络提取浅层空间特征和深层语义特征。其中,通过 ASPP 模块捕捉深层语义特征中的多尺度上下文信息,随后将其输入至 SCC - MFMSA 模块,引入频域分析机制抑制高频噪声干扰,同时增强模型对影像长距离依赖关系的建模能力。在解码阶段,将经过上述模块增强并上采样后的深层语义特征,与经 1 ×1 卷积降维后的浅层空间特征输入 MAF 模块进行动态融合,以增强空间、通道和多尺度特征的表示,有效恢复目标区域边缘和空间分布信息,提升复杂种植场景下冬小麦种植区提取精度。

2.2 主干网络改进

原始 DeepLabV3 + 的主干网络是基于深度可分离卷积的 Xception[21] 网络,该网络能够有效减少参数计算量,但其深度卷积和逐点卷积会导致部分通道信息丢失,且容易引入冗余特征,限制了网络的特征提取能力,增加了模型后续处理的复杂性。本文采用具有残差结构的 ResNet18 网络作为主干网络,减少细节特征丢失,并在残差模块中融入 SCConv[22] 卷积块以减少冗余特征并增强有效特征表达,进而构建 SC - ResNet18 主干网络(图 4)。与 ResNet50 等深层网络相比,ResNet18 参数量更少、计算开销更低,能在满足冬小麦特征表达能力的同时降低过拟合风险,并提升高分辨率影像分割任务的推理效率。

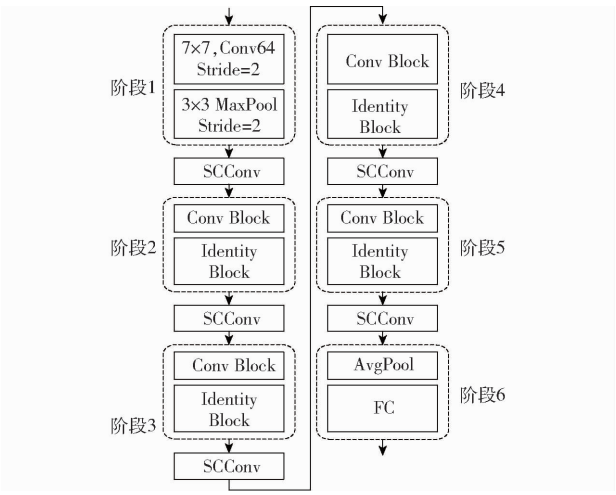

图 4 SC - ResNet18 结构图

Fig. 4 SC - ResNet18 structure

SC - ResNet18 网络结构如图 4 所示,包括对输入特征进行预处理的阶段 1 以及随后的阶段 2 至阶段 6 共 6 个网络层,并且在阶段 1 至阶段 5 每个阶段后融入 SCConv 卷积块。输入特征首先经过阶段 1 的卷积、最大池化处理,随后进入阶段 2 至阶段 5 通过残差连接和双 3 ×3 卷积设计,在保持特征图维度不变的同时,逐层深化特征表示。最后经过阶段 6 的平均池化和全连接层处理,得到最终的输出特征。其中,SCConv 模块结构如图 5 所示,阶段 2 至阶段 5 中的 Conv Block 和 Identity Block 模块如

图 6 所示。

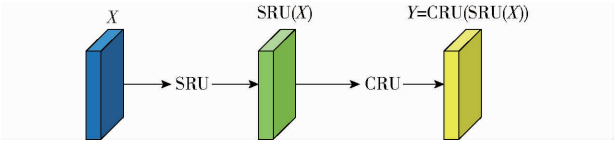

图 5　SCConv 模块结构

Fig. 5　SCConv structure

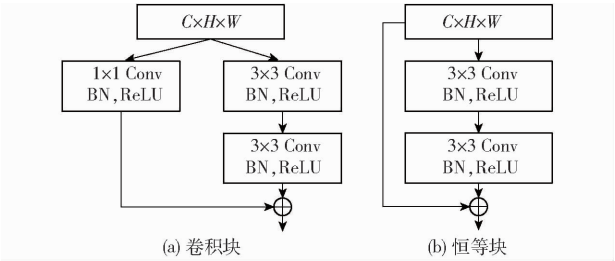

图 6　卷积块和恒等块结构

Fig. 6　Structures of convolutional block and identity block

SCConv 通过空间重构单元(Spatial reconstruction unit,SRU)和通道重构单元(Channel reconstruction unit,CRU)两部分减少冗余特征并降低计算复杂度。SRU 采用分离重构的方法来抑制空间冗余,CRU 采用分裂变换融合的方法来消除通道冗余。对于输入的特征 X,SCConv 首先通过 SRU 操作获得空间细化特征 SRU(X),然后利用 CRU 操作获得通道细化特征 Y。

2.3　多尺度多频注意力模块(SCC－MFMSA)

SCC－MFMSA 模块主要包括尺度分解、多频通道注意力(Multi-frequency channel attention,MFCA)和多尺度空间注意力(Multi-scale spatial attention,MSSA)3 部分(图 7)。

SCC－MFMSA 模块利用 MFCA 提取多频域特征并增强关键频率信息表达,利用 MSSA 提取多尺度空间边界信息,通过为每个尺度分支独立设计空间注意力生成器,动态平衡背景抑制与目标区域特征增强,降低背景信息对冬小麦种植区提取的干扰。该模块通过双注意力协同提取影像频域和空域特征,增强模型对冬小麦种植区域及边界信息的多尺度感知能力。

SCC－MFMSA 模块对现有 MFMSA[23] 模块进行如下改进:针对影像中冬小麦种植区多样化空间分布特征,将多尺度分支的数量增加为 4 个,并通过不同膨胀率卷积与动态下采样策略对输入特征进行处理,获取冬小麦种植区不同尺度的特征和长距离空间关系;在多尺度分支中引入 SCConv 替换原有普

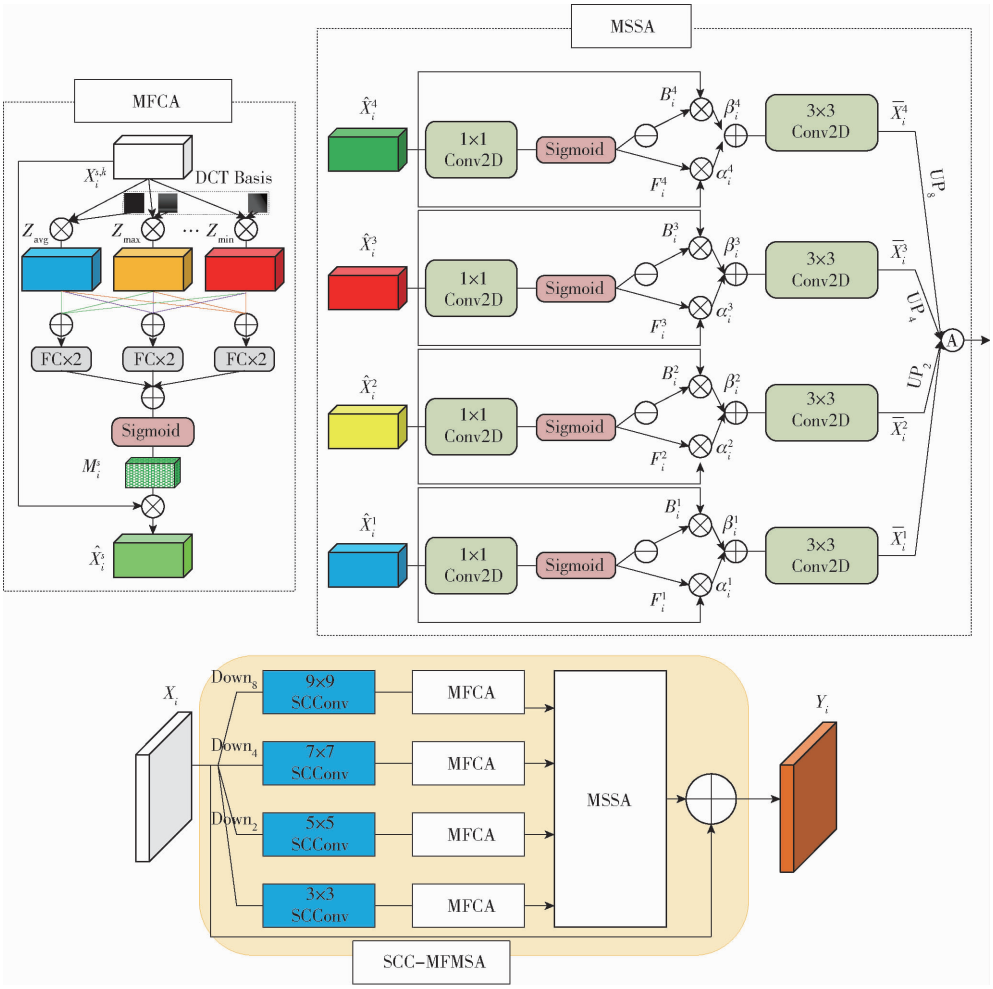

图 7　SCC－MFMSA 结构图

Fig. 7　Structure of SCC－MFMSA

通二维卷积操作,通过空间和通道特征重构聚焦冬小麦种植区关键信息并减少冗余特征响应,提升模型对冬小麦种植区细节纹理和上下文信息的捕捉能力。

2.3.1 尺度分解

为从输入特征 X_i 获得多尺度特征,对于第 s 尺度分支($1\leqslant s\leqslant 4$),利用通道和分辨率降低比 γ($\gamma\in(0,1)$)降低输入特征图 X_i 的分辨率和通道。因此,第 s 尺度分支处的每个输入特征图 X_i^s 表示为

$$X_i^s=\begin{cases}\mathrm{Conv2D}_{2s+1}(\mathrm{Down}_s(X_i))\in\mathbf{R}^{C_s\times H_s\times W_s} & (s=1)\\ \mathrm{SCConv}(\mathrm{Conv2D}_{2s+1}(\mathrm{Down}_s(X_i)))\in\mathbf{R}^{C_s\times H_s\times W_s} & (s=2,3,4)\end{cases} \tag{1}$$

式中 $\mathrm{Down}_s(\ )$——比例因子 2^{s-1} 的下采样操作

C_s——第 s 个尺度分支的通道数

H_s——第 s 个尺度分支的高度

W_s——第 s 个尺度分支的宽度

$\mathrm{Conv2D}_{2s+1}(\ )$——卷积核为 $2s+1$ 的 2D 卷积

SCConv( )——图 5 所示的卷积操作

为兼顾感受野扩展与计算开销,第 s 尺度分支卷积的膨胀率设为 $d_s=s(1\leqslant s\leqslant 4)$。

2.3.2 多频通道注意力(MFCA)

对于每个尺度分支处的特征 $X_i^{s,k}$,通过二维离散余弦变换(Two-dimensional discrete cosine transform,2D DCT)产生的基图像 D 来表征,即

$$\boldsymbol{X}_i^{s,k}=\sum_{h=0}^{H_s-1}\sum_{w=0}^{W_s-1}(X_i^s)_{h,w}D_{h,w}^{u_k,v_k} \tag{2}$$

式中 u_k——$\boldsymbol{X}_i^{s,k}$ 水平方向的频率索引

v_k——$\boldsymbol{X}_i^{s,k}$ 垂直方向的频率索引

此外,在第 s 尺度分支处的 D 定义为

$$D_{h,w}^{u_k,v_k}=\cos\left(\frac{\pi h}{H_s}\left(u_k+\frac{1}{2}\right)\right)\cos\left(\frac{\pi w}{W_s}\left(v_k+\frac{1}{2}\right)\right) \tag{3}$$

随后,对每个 $X_i^{s,k}$ 使用全局平均池化、全局最大池化和全局最小池化策略,并将所有频率的平均、最大和最小池化特征分别融合,生成 Z_d($d\in\{\mathrm{avg},\mathrm{max},\mathrm{min}\}$),即 Z_{avg}、Z_{max} 和 Z_{min},通过使用具有缩减比 r 的两个全连接层 W_1 和 W_2 来产生第 s 尺度分支处的通道注意力图 M_i^s,公式为

$$M_i^s=\sigma\left(\sum_{d\in\{\mathrm{avg},\mathrm{max},\mathrm{min}\}}W_2(\delta(W_1Z_d))\right)\in\mathbf{R}^{C_s} \tag{4}$$

式中 σ——Sigmoid 激活函数

δ——ReLU 激活函数

最后,在第 s 个尺度分支使用 M_i^s 重新校准特征图 X_i^s,即

$$\hat{X}_i^s=X_i^sM_i^s \tag{5}$$

2.3.3 多尺度空间注意力(MSSA)

为每个尺度分支引入两个可学习参数(α_i^s 和 β_i^s),分别控制前景和背景之间的信息流,即

$$\overline{X}_i^s=\mathrm{Conv2D}_3(\alpha_i^s(\hat{X}_i^sF_i^s)+\beta_i^s(\hat{X}_i^sB_i^s)) \tag{6}$$

式中 F_i^s——前景注意力图

B_i^s——背景注意力图

使用上采样(Up_s)匹配分辨率,聚合不同尺度分支的每个细化特征,最后对 X_i 和 MSSA 聚合后的特征应用残差连接,公式为

$$Y_i=X_i+A(\overline{X}_i^1,\mathrm{Up}_2(\overline{X}_i^2),\cdots,\mathrm{Up}_s(\overline{X}_i^s)) \tag{7}$$

式中 $A(\cdot)$——特征聚合函数

Y_i——输出特征

2.4 多尺度注意力融合(MAF)

原始 DeepLabV3 + 仅将浅层特征 S 和深层特征 P 进行简单拼接,缺乏有效的特征融合与交互。因此,本文引入 MAF 模块(图 8)实现局部细节特征和全局语义信息的动态融合,通过增强通道、空间和多尺度层次的特征表达和语义关联,提升模型在冬小麦种植区提取时的适应性和鲁棒性。

首先,对于输入特征 S 和 P,通过引入可学习权重参数动态调整不同特征图的贡献比例,使模型在训练过程中自动优化特征融合方式,充分利用空间细节信息和多尺度语义信息,得到融合后的特征 $\boldsymbol{F}\in\mathbf{R}^{C\times H\times W}$,即

$$\boldsymbol{F}=\mathrm{Conv2D}_5\left(\mathrm{Conv2D}_1(S)\frac{w_0}{w_0+w_1+\varepsilon}+\mathrm{Conv2D}_3(P)\frac{w_1}{w_0+w_1+\varepsilon}\right) \tag{8}$$

式中 w_i——可学习权重参数

ε——随机生成的一个极小值

$\mathrm{Conv2D}_{1,3,5}$——卷积核为(1,3,5)的二维卷积

随后,将 $\boldsymbol{F}\in\mathbf{R}^{C\times H\times W}$ 应用于 3 个分支,分别增强通道、多尺度和空间特征的表示。

对于通道分支,首先对特征进行全局平均池化,生成通道权值分布,并与原始输入逐元素相乘,得到通道注意力特征图;对于多尺度分支,利用 3 组不同膨胀率(6、12、18)的并行卷积提取多尺度特征并融合;对于空间分支,基于混合池化和卷积操作生成空间分布权重。最后通过逐元素求和的方式融合 3 个分支的结果,得到最终的输出特征图。

2.5 试验环境与评价指标

2.5.1 试验环境设置

试验采用的操作系统为 Ubuntu 20.04.3 LTS,GPU 为 NVIDIA - A40,显存容量 48 GB,CPU 为 Intel

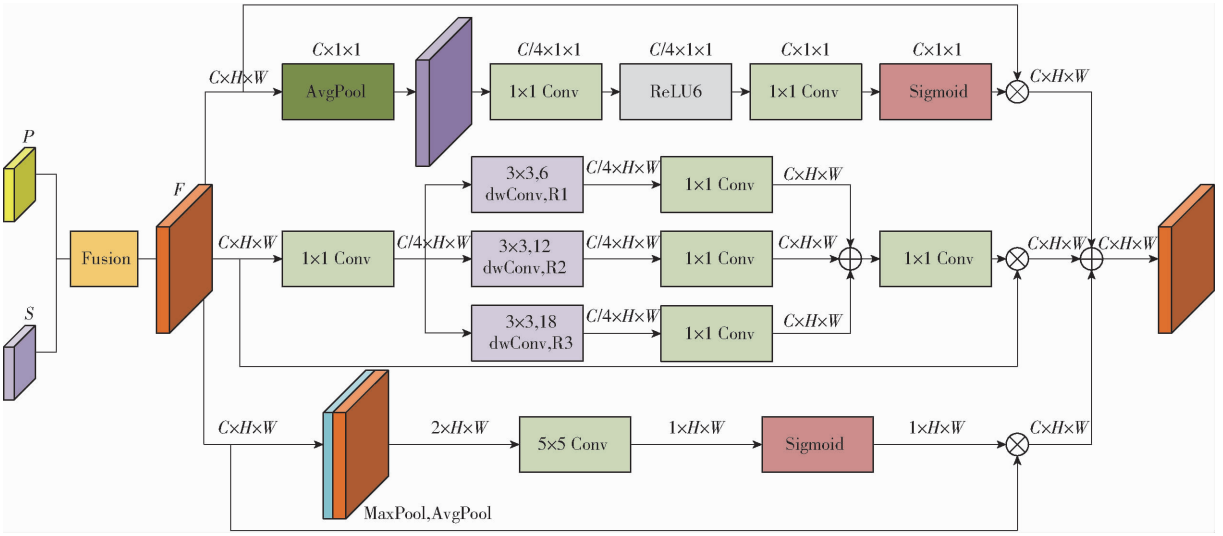

图 8　MAF 模块

Fig. 8　MAF block

Xeon Gold 5318Y CPU 2.10 GHz；Python 版本为 3.9.18，Pytorch 版本为 2.1.1。模型训练参数设置如下：学习率为 0.000 1，批处理大小（Batch size）为 8，epoch 为 100。损失函数采用交叉熵损失和 Dice Loss 构成的混合损失函数，经重复试验设置混合权重为 6∶4。

2.5.2　评价指标

为评估本文冬小麦种植区提取模型性能，采用精确率、召回率、F1 值和交并比（Intersection over union，IoU）[24] 分析不同模型的精度。

3　结果与分析

3.1　不同模型精度对比

为定量分析本文 RSMANet 模型对冬小麦种植区的提取精度，分别选取原始 DeepLabV3 + 模型以及 MCSNet[25]、CMTFNet[26]、SACANet[27] 和 TransUNet[28] 等主流模型进行对比试验，不同模型的提取结果见表 1。其中，MCSNet 通过改进空间卷积提升对狭窄目标的特征提取能力；CMTFNet 综合多尺度机制、通道注意力机制以及多头自注意力机制，增强模型对全局上下文信息的捕获能力；SACANet 采用局部-全局类注意力机制，提高模型的上下文建模能力；TransUNet 利用 Transformer 改进 UNet 模型，增强对目标细节的捕捉能力。

表 1　不同模型对比试验结果

Tab. 1　Comparison of experimental results among different models　%

模型	精确率	召回率	F1 值	IoU
MCSNet	93.21	92.38	92.78	85.28
DeepLabV3 +	93.40	92.86	93.13	85.90
CMTFNet	93.61	93.42	93.53	86.62
SACANet	94.27	92.97	93.62	87.04
TransUNet	94.33	94.25	94.29	88.01
RSMANet	94.87	94.17	94.52	89.09

分析表 1 可知，本文 RSMANet 模型 IoU 和 F1 值分别达到 89.09% 和 94.52%，在冬小麦种植区提取任务中精度高于其他对比模型。与原始 DeepLabV3 + 模型相比，IoU 和 F1 值分别提升 3.19、1.39 个百分点，表明通过设计 SC - ResNet18 主干网络，引入 SCC - MFMSA 模块，并在解码器中改进特征融合方式，本文模型能够有效提升冬小麦种植区的提取精度。

与其他对比模型相比，RSMANet 模型的 IoU 和 F1 值提升幅度分别在 1.08 ~ 3.81 个百分点和 0.23 ~ 1.74 个百分点之间，这主要得益于模型的 SC -ResNet18 网络在特征提取时能够更加聚焦于冬小麦核心特征信息，增强对冬小麦种植区的关注度；同时，模型编码和解码阶段的 SCC - MFMSA 模块和 MAF 模块协同作用，通过在空间、通道和多尺度 3 个维度以及高分影像空间特征和频域特征的联合，强化了网络对冬小麦种植区的特征表达能力，使其能捕捉并恢复关键的全局上下文和边缘轮廓特征信息，从而提升了模型在冬小麦种植区遥感提取任务中的性能。

3.2　消融试验

为验证 RSMANet 模型各模块对网络性能的贡献，本文以原始的 DeepLabV3 + 模型为基线模型，针对 SC - ResNet18、SCC - MFMSA 和 MAF 模块分别设计消融试验，试验结果见表 2。

分析表 2 可知，相较原始 DeepLabV3 + 模型，仅引入 SC - ResNet18 模块时，IoU 提升 2.22 个百分点，F1 值提升 1.02 个百分点，表明 SC - ResNet18 凭借其残差结构增强了主干网络的特征提取能力，

表 2 消融试验结果

Tab. 2 Ablation study result %

SC – ResNet18	SCC – MFMSA	MAF	精确率	召回率	F1 值	IoU
×	×	×	93.40	92.86	93.13	85.90
√	×	×	94.50	93.80	94.15	88.12
×	√	×	94.13	93.49	93.79	87.05
×	×	√	93.86	93.82	93.84	87.28
√	√	×	94.69	94.01	94.34	88.89
√	√	√	94.87	94.17	94.52	89.09

注:“√”表示模块被引入,“×”表示未被引入。

能有效提高冬小麦识别精度;单独引入 SCC – MFMSA 模块时,IoU 和 F1 值分别提升 1.15、0.66 个百分点,表明多尺度多频注意力能有效利用遥感影像在频率域、空间域以及通道的多尺度信息,增强了模型对长距离关系的建模,提升了冬小麦种植区在作物间种区和影像噪声干扰区的可区分性。单独引入 MAF 模块后,IoU 和 F1 值分别提升了 1.38、0.71 个百分点,这是由于多尺度注意力融合模块在特征融合上采样阶段强化了特征表达能力,能够充分利用空间细节信息和上下文语义信息,有效改善冬小麦种植区边界恢复的效果。上述分析表明 3 个模块对模型精度提升均有积极贡献。

同时引入上述 3 个模块(RSMANet)的组合效果最佳,模型 IoU 和 F1 值分别达到 89.09% 和 94.52%,较原始 DeepLabV3 + 模型分别提升了 3.19、1.39 个百分点,说明 3 个模块相互协作,共同提升了模型对冬小麦种植区的识别精度。

3.3 不同模型提取结果分析

为直观对比 RSMANet 模型冬小麦种植区提取效果,分别选取冬小麦与其他作物混合种植区、不同尺度地块种植区、细碎化密集种植区等典型区域,对不同模型提取结果(图 9)进行定性分析。

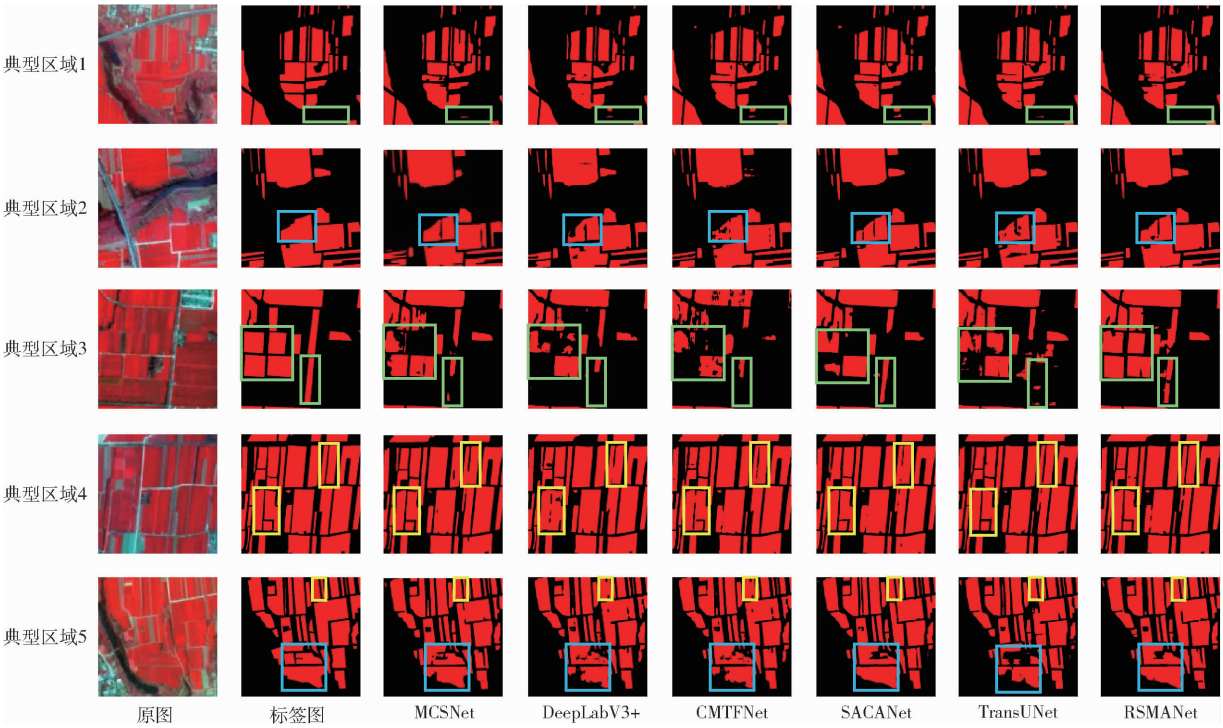

图 9 不同模型冬小麦种植区提取结果

Fig. 9 Extraction results of different models for winter wheat planting areas

由图 9 可知,在冬小麦与其他作物混合种植区(典型区域 1、3),MCSNet、DeepLabV3 + 、CMTFNet、SACANet 等模型均存在不同程度的误分和局部漏检问题,主要表现为将部分油菜种植区(典型区域 1 绿色矩形框区域)误识别为冬小麦种植区,原因在于冬小麦与油菜的光谱特征具有一定相似性。相比之下,本文采用改进的 SC – ResNet18 主干网络强化特征提取能力,提升了冬小麦与光谱特征相近作物的区分能力,有效降低了误分情况。此外,在典型区域 3(绿色矩形框区域),由于不同作物间光谱混淆和噪声干扰,DeepLabV3 + 、CMTFNet 等模型的漏检问题较明显。本文设计的 SCC – MFMSA 模块通过多尺度频率带分解,引入频域特征增强与其他作物的光谱可分性,同时融合目标区域的局部细节与轮廓等空间特征信息,有效降低了噪声干扰导致的漏检,提升了模型在复杂种植环境下的识别能力。

在不同地块尺度的提取结果中(典型区域 2、5),DeepLabV3 + 、CMTFNet 和 TransUNet 等模型均存在地块内部破碎和边缘缺失的问题(蓝色矩形框区域),说明这些模型在处理不同尺度的目标区域

时，难以有效整合多尺度特征信息，对目标区域的识别不够全面。本文通过设计 SCC - MFMSA 模块整合多尺度特征信息，增强模型对局部区域和全局区域的关联推理，提升了模型对目标区域的多尺度感知能力，避免了单一尺度特征导致的识别不完整问题，使得模型能够精准识别从集中连片到分散田块的冬小麦种植区。

在冬小麦密集种植且地块细碎区域（典型区域 4、5），DeepLabV3 + 和 CMTFNet 模型结果中存在地块边界粘连问题（黄色矩形框区域）。相比之下，RSMANet 模型利用 MAF 模块融合并强化了不同通道、尺度和空间维度的特征信息，从而可以更清晰地分辨不同地块间的边界，有效减少了因空间特征模糊导致的边界粘连问题。

3.4　模型复杂度分析

为进一步评估 RSMANet 模型及各改进模块的效率，从参数量、计算量以及推理时间 3 方面分析模型复杂度，试验结果见表 3。

表 3　模型计算复杂度对比

Tab. 3　Computational complexity comparison of models

模型	参数量	浮点运算量	平均延迟/ms
DeepLabV3 +	$4.153\,1\times10^{7}$	$1.526\,04\times10^{11}$	314.62
+ SC - ResNet18	$3.849\,6\times10^{7}$	$1.444\,62\times10^{11}$	286.30
+ SCC - MFMSA	$4.146\,5\times10^{7}$	$1.523\,10\times10^{11}$	313.85
+ MAF	$4.224\,0\times10^{7}$	$1.544\,27\times10^{11}$	316.48
RSMANet	$4.435\,0\times10^{7}$	$1.560\,27\times10^{11}$	320.69

与原始 DeepLabV3 + 相比，单独引入 SC - ResNet18 后，参数量下降 7.31%，浮点运算量下降 5.34%，平均延迟下降 9.01%，表明该结构能够有效削减冗余计算，提升推理效率与实时性；单独引入 SCC - MFMSA 模块时，参数量和浮点运算量基本保持不变，平均延迟有微弱下降，表明该模块在增强多尺度与多频特征建模的同时，未引入额外的计算开销。此外，单独引入 MAF 模块后，各项指标比原始 DeepLabV3 + 有小幅提升，表明动态加权融合操作对整体计算负担影响有限。集成 3 个改进模块后，RSMANet 模型参数量上升 6.79%，浮点运算量上升 2.24%，平均延迟增加 1.93%。结合前文精度对比结果可知，模型在实现精度有效提升的同时，未造成明显的推理效率下降，在精度与计算开销之间取得了较好的平衡。

3.5　模型泛化性试验

为验证 RSMANet 模型对不同生育期冬小麦种植区提取的适用性，选取研究区 2021 年 5 月 9 日的高分二号影像（此时冬小麦处于抽穗期）进行泛化性试验，并与原始 DeepLabV3 +、CMTFNet 和 TransUNet 共 3 个代表模型进行对比，结果如表 4 所示。本试验数据集构建方法同 1.3 节。

表 4　模型泛化性试验结果

Tab. 4　Model generalization results　%

模型	精确率	召回率	F1 值	IoU
DeepLabV3 +	92.78	93.48	93.13	85.24
CMTFNet	93.12	93.60	93.36	86.70
TransUNet	93.43	94.11	93.77	87.29
RSMANet	94.43	94.69	94.56	88.61

分析表 4 可知，RSMANet 模型仍表现出最优性能。相较于原始 DeepLabV3 + 模型，F1 值和 IoU 分别提升 1.43、3.37 个百分点，表明 RSMANet 模型对不同生育期的冬小麦种植提取具有一定的泛化能力。

4　讨论

本研究基于高分二号影像和 RSMANet 模型，实现了复杂种植环境下冬小麦种植区的精细提取。现有冬小麦识别研究大多采用中低分辨率影像、单一空间注意力机制[2,6,9]，难以在高分辨率条件下同时抑制高频噪声并保持细碎地块边界。本文从特征提取、频域建模与多尺度融合 3 方面优化网络结构，能够有效区分冬小麦与其他光谱相近作物，缓解田块内部破碎和地块间边界粘连现象。

RSMANet 模型在高分二号影像上获得较好的分割精度，可满足县域尺度冬小麦地块级制图需要，为精细化农情监测提供参考；研究方法对水稻、玉米等其他粮食作物的高分辨率遥感识别具有借鉴意义。

本研究在拔节期和抽穗期两个不同生育期验证了模型的泛化性，模型在其他冬小麦典型种植区域及其他类型高分影像中的适用性有待进一步验证。此外，研究以单时相影像为主，物候信息及时间序列特征利用不足，后续将开展多源遥感数据融合试验，并可引入轻量化设计，在提升模型精度与鲁棒性的同时满足大范围业务化应用需求。

5　结论

（1）RSMANet 模型的 IoU 和 F1 值分别达到 89.09% 和 94.52%，与 DeepLabV3 + 等主流模型相比分别提升 1.08 ~ 3.81 个百分点和 0.23 ~ 1.74 个百分点，表明本文模型能有效提升冬小麦种植区遥感识别精度。

（2）通过设计 SC - ResNet18 主干网络、多频多尺度注意力（SCC - MFMSA）及多尺度注意力融合（MAF）3 个模块构建改进模型，实现冬小麦种植区

关键特征的有效提取与多尺度动态融合。消融试验结果表明,单独引入各模块均能不同程度提升模型性能,上述 3 个模块协同能够有效增强模型对冬小麦种植区多尺度上下文和边缘细节特征的多维度感知和有效表达,为复杂农田场景下冬小麦精细化提取提供借鉴。

参 考 文 献

[1] 郑希望,王凯,朱野,等. 联合相干性特征的短时序双极化 SAR 冬小麦提取[J]. 遥感信息, 2024, 39(3): 48 - 54.
Zheng Xiwang, Wang Kai, Zhu Ye, et al. Winter wheat extraction based on short time-series SAR data with dual polarizations and coherence characteristics[J]. Remote Sensing Information, 2024, 39(3): 48 - 54. (in Chinese)

[2] 张冬韵,吴田军,骆剑承,等. 时空协同的农业种植结构遥感精细制图[J]. 遥感学报, 2024, 28(8):2014 - 2029.
Zhang Dongyun, Wu Tianjun, Luo Jiancheng, et al. Precise cropplanting structure mapping method based on spatial-temporal collaboration of remote sensing[J]. National Remote Sensing Bulletin, 2024, 28(8):2014 - 2029. (in Chinese)

[3] 宁岸,任淼淼,蒋杉,等. 光学遥感技术在农情监测中的应用与展望[J]. 中国农机化学报, 2025, 46(2):199 - 206.
Ning An, Ren Miaomiao, Jiang Shan, et al. Application and prospect of optical remote sensing technology in agricultural monitoring[J]. Journal of Chinese Agricutural Mechanization, 2025, 46(2): 199 - 206. (in Chinese)

[4] 林欣怡,汪小钦,李蒙蒙,等. 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[J]. 农业机械学报, 2025, 56(6):446 - 456.
Lin Xinyi, Wang Xiaoqin, Li Mengmeng, et al. Tea plantation recognition method based on preferred multi-source remote sensing features and two-branch convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(6):446 - 456. (in Chinese)

[5] Wang C, Li K, Yang W, et al. GACDNet: mapping winter wheat by generative adversarial cross-domain networks with transformer integration for zero-sample extraction[J]. Computers and Electronics in Agriculture, 2024, 221: 109012.

[6] Fan L, Xia L, Yang J, et al. A temporal-spatial deep learning network for winter wheat mapping using time-series Sentinel - 2 imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2024, 214: 48 - 64.

[7] Li J, Cai Y, Li Q, et al. A review of remote sensing image segmentation by deep learning methods[J]. International Journal of Digital Earth, 2024, 17(1): 2328827.

[8] 冯权泷,任燕,姚晓闯,等. 基于多源光学雷达数据融合的黄淮海平原冬小麦识别[J]. 农业机械学报, 2023, 54(2): 160 - 168.
Feng Quanlong, Ren Yan, Yao Xiaochuang, et al. Identification of winter wheat in Huang - Huai - Hai plain based on multi-source optical radar data fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(2): 160 - 168. (in Chinese)

[9] Anderegg J, Tschurr F, Kirchgessner N, et al. Pixel to practice: multi-scale image data for calibrating remote-sensing-based winter wheat monitoring methods[J]. Scientific Data, 2024, 11(1): 1033.

[10] 王佳玥,蔡志文,王文静,等. 协同多源国产高分影像和面向对象方法的南方农作物遥感识别[J]. 中国农业科学, 2023, 56(13): 2474 - 2490.
Wang Jiayue, Cai Zhiwen, Wang Wenjing, et al. Integrating multi-source Gaofen images and object-based methods for crop type identification in south China[J]. Scientia Agricultura Sinica, 2023, 56(13): 2474 - 2490. (in Chinese)

[11] Yan X, Chen J, Xu Q, et al. Mitigating texture bias: a remote sensing super-resolution method focusing on high-frequency texture reconstruction[J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 5615218.

[12] 吕吉平,张云帆,侯睿彤,等. 冬小麦生长季蒸散发多时间尺度演变规律及驱动机制解析[J]. 节水灌溉, 2025(12): 105 - 112.
Lü Jiping, Zhang Yunfan, Hou Ruitong, et al. Analysis of multi-time scale evolution and driving mechanism of evapotranspiration in winter wheat growing season[J]. Water Saving Irrigation, 2025(12): 105 - 112. (in Chinese)

[13] Zhai W, Cheng Q, Duan F, et al. Remote sensing-based analysis of yield and water-fertilizer use efficiency in winter wheat management[J]. Agricultural Water Management, 2025, 311: 109390.

[14] 王鹏新,王静怡,郭丰玮,等. 基于遥感多参数和 Stacking 集成学习的冬小麦单产估测[J]. 农业机械学报, 2025, 56(11): 369 - 377.
Wang Pengxin, Wang Jingyi, Guo Fengwei, et al. Yield estimation of winter wheat based on multiple remotely sensed parameters and Stacking ensemble learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(11): 369 - 377. (in Chinese)

[15] Shu X, Wang J, Zhang A, et al. CSCA U - Net: a channel and space compound attention CNN for medical image segmentation [J]. Artificial Intelligence in Medicine, 2024, 150: 102800.

[16] 张依涵,沈占锋,寇雯齐,等. 基于多源遥感数据的林农间作种植结构精细分类[J]. 农业机械学报, 2025, 56(10): 429 - 436.
Zhang Yihan, Shen Zhanfeng, Kou Wenqi, et al. Fine classification under pattern of forest and crops intercropping based on multi-source remote sensing data[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(10): 429 -

436.（in Chinese）

[17] 宋德娟,张承明,杨晓霞,等. 高分二号遥感影像提取冬小麦空间分布[J]. 遥感学报, 2020, 24(5): 596 – 608.
Song Dejuan, Zhang Chengming, Yang Xiaoxia, et al. Extracting winter wheat spatial distribution information from GF – 2 image[J]. Journal of Remote Sensing,2020,24(5):596 – 608. (in Chinese)

[18] 李旭青,吴冬雪,王玉博,等. 基于改进 HRNet 的遥感影像冬小麦语义分割方法[J]. 农业工程学报, 2024, 40(3): 193 – 200.
Li Xuqing, Wu Dongxue, Wang Yubo, et al. Improved HRNet semantic segmentation model for remote sensing images of winter wheat[J]. Transactions of the CSAE, 2024, 40(3): 193 – 200. (in Chinese)

[19] 蔡玉林,刘照磊,孟祥磊,等. 基于 HRNet 和自注意力机制的多源遥感影像水稻提取[J]. 农业工程学报, 2024, 40(4): 186 – 193.
Cai Yulin, Liu Zhaolei, Meng Xianglei, et al. Rice extraction from multi-source remote sensing images based on HRNet and self-attention mechanism[J]. Transactions of the CSAE, 2024, 40(4): 186 – 193. (in Chinese)

[20] 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.

[21] Chollet F. Xception: deep learning with depthwise separable convolutions[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 1251 – 1258.

[22] Li J, Wen Y, He L. SCCONV: spatial and channel reconstruction convolution for feature redundancy[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023: 6153 – 6162.

[23] Nam J H, Syazwany N S, Kim S J, et al. Modality-agnostic domain generalizable medical image segmentation by multi-frequency in multi-scale attention[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024: 11480 – 11491.

[24] Luo Z, Yang W, Yuan Y, et al. Semantic segmentation of agricultural images: a survey[J]. Information Processing in Agriculture, 2024, 11(2): 172 – 186.

[25] Liu Y, Xu X, Dai Z, et al. MCSNet: multi-scene crack segmentation network based on few-shot learning[C]// Proceedings of the 2023 China Automation Congress (CAC), 2023: 3633 – 3638.

[26] Wu H, Huang P, Zhang M, et al. CMTFNet: CNN and multiscale transformer fusion network for remote-sensing image semantic segmentation[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 2004612.

[27] Ma X, Che R, Hong T, et al. SACANet: scene-aware class attention network for semantic segmentation of remote sensing images[C]// Proceedings of the 2023 IEEE International Conference on Multimedia and Expo (ICME), 2023: 828 – 833.

[28] Chen J, Lu Y, Yu Q, et al. TransUNet: Transformers make strong encoders for medical image segmentation[J]. Medical Image Analysis, 2024, 94: 103205.