

基于改进 TransUNet 的黄土高原梯田作业区域提取方法

杨福增^{1,2} 袁敏鑫^{1,2} 许翔虎^{1,2} 王 旺^{1,2} 杨江涛³ 刘志杰^{1,2}

(1. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100; 2. 农业农村部北方农业装备科学观测实验站, 陕西杨凌 712100;

3. 西北农林科技大学水利与建筑工程学院, 陕西杨凌 712100)

摘要: 农田作业区域地图准确构建是实现农机路径规划和导航作业的重要前提。黄土高原梯田田块大小各异、形状复杂多变,并且存在部分凹坑、沟坎和诸多危险作业边界,常用的卫星测点等方法难以准确地提取梯田作业区域,本文以无人机梯田遥感图像为数据基础,提出一种基于多尺度特征提取与融合上采样的改进 TransUNet 模型。在编码器部分,通过引入金字塔压缩注意力模块(Pyramid squeeze attention, PSA),在通道注意力的基础上增强对不同尺度梯田特征提取和融合的能力,并使用残差结构优化 Transformer 层;在解码器部分,引入 Dual up-sample 模块将亚像素卷积层与双线性插值上采样两者融合,提升梯田边界分割精度的同时防止棋盘效应,并在解码器末尾添加通道和空间注意力机制模块(Concurrent spatial and channel squeeze and channel excitation, SCSE),同时对空间和通道维度的信息进行整合增强,有助于图像细节特征逐步恢复。实验结果表明,改进 TransUNet 模型在直长条形、蜿蜒长条形和不规则形 3 类典型梯田测试集上平均像素准确率、F1 值和平均交并比平均分别达 96.0%、96.0% 和 92.3%,3 项指标相较于改进前平均提升 1.8 个百分点,与代表性的 PSPNet、HRNet V2、DeepLab V3+、U-Net 模型相比,3 项指标平均提升 8.3、6.2、5.0、4.2 个百分点。在 3 类单块梯田测试集上,本文模型表现最优,像素交并比平均可达 97.0%。本文方法可为黄土高原梯田环境地图构建和丘陵山地农机导航作业提供参考。

关键词: 梯田; 无人机遥感; 语义分割; TransUNet

中图分类号: S29; TP79

文献标识码: A

文章编号: 1000-1298(2024)12-0278-09

OSID:



Extraction Method of Terrace Operation Area in Loess Plateau Based on Improved TransUNet

YANG Fuzeng^{1,2} YUAN Minxin^{1,2} XU Xianghu^{1,2} WANG Wang^{1,2} YANG Jiangtao³ LIU Zhijie^{1,2}

(1. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China

2. Scientific Observing and Experimental Station of Agricultural Equipment for the Northern China, Ministry of Agriculture and Rural Affairs, Yangling, Shaanxi 712100, China

3. College of Water Resources and Architectural Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China)

Abstract: Accurate map construction of farmland operation area is an important prerequisite for realizing the path planning and navigation operation of farm machinery. The terraced fields on the Loess Plateau have different sizes and complex shapes, and there are some pits, ditches and many dangerous operation boundaries, so it is difficult to accurately extract the terraced operation area by the commonly used satellite point measurement methods. An improved TransUNet model based on multi-scale feature extraction and fusion up-sampling was proposed with the remote sensing images of terraced fields from UAVs as the data base. In the encoder part, the ability of feature extraction and fusion for different scales of terraces was enhanced by introducing the pyramid squeeze attention (PSA) module on top of the channel attention and the Transformer layer was optimized by using the residual structure. In the decoder part, the Dual up-sample module was introduced to integrate the sub-pixel convolutional layer with the bilinear interpolation up-sampling to improve the accuracy of the terraced field boundary segmentation while preventing the checker board effect, and the channel and spatial attention mechanism module

收稿日期: 2024-01-16 修回日期: 2024-03-01

基金项目: 国家重点研发计划项目(2022YFD2001300、2023YFD1000800)和陕西省重点研发计划项目(2022ZDLNY03-04)

作者简介: 杨福增(1966—),男,教授,博士生导师,主要从事农业机器人关键技术和丘陵山地智能农机装备研究,E-mail: yfz0701@163.com

通信作者: 刘志杰(1980—),男,副教授,博士,主要从事智能农机装备和智能控制技术研究,E-mail: liuzhijie@nwfau.edu.cn

(concurrent spatial and channel squeeze and channel excitation (SCSE)) was added at the end of the decoder to integrate and enhance the information of spatial and channel dimensions, which helped to recover the detailed features of the image step by step. The experimental results showed that the mean pixel accuracy, F1 value, and mean intersection over union of the improved TransUNet model can reach up to 96.0%, 96.0%, and 92.3% on average on the test set of three typical terraces, namely, straight and long stripes, meandering stripes, and irregular shapes, respectively, which was an average enhancement of 1.8 percentage points compared with the pre-improvement period, and compared with the representative PSPNet, HRNet V2, DeepLab V3+, and U-Net models, the average improvement of the three indicators was 8.3, 6.2, 5.0, and 4.2 percentage points. On the test set of three types of single terraces, the proposed model performed the best, and intersection over union can reach 97.0% on average. The method can provide a reference for the construction of terraced field environment maps in the Loess Plateau and the navigation operation of agricultural machinery in hilly and mountainous areas.

Key words: terrace; UAV remote sensing; semantic segmentation; TransUNet

0 引言

农机导航技术是支撑未来智慧农业发展的重要基础^[1-4],该技术能有效减轻劳动强度,提高作业效率,因此被广泛应用于平原地区的各类农业生产活动,其中,准确的田块作业区域地图构建是实现农机路径规划和导航作业的重要前提^[5-8]。农田作业区域是指在一块农田中,为了特定的农业目的(如种植、施肥、灌溉、收获等)而特别划分出来的区域,可以在这个区域内进行一系列农业活动。但黄土高原地区农田多以梯田形式存在,形状和大小复杂多变,田块内有部分凹坑、沟坎,并且存在田埂、坎壁等诸多危险作业边界,而常用的卫星测点等方法难以准确提取梯田作业区域,无法开展农机作业路径规划,所以针对丘陵山区农机导航作业的相关研究鲜有报道。因此,有必要开展针对黄土高原梯田作业地图构建的研究,为实现该地区的农机路径规划和导航作业提供技术支撑。

部分学者利用卫星遥感对梯田面积进行估测^[9-11],但卫星遥感在时空分辨率方面精度有限,难以支撑环境的准确提取^[12]。无人机可以获取特定地点的实时、详细农业信息,极大地弥补了卫星遥感在时空分辨率方面的局限性,使得其在精准农业领域引起了极大关注^[13-15]。传统分割方法主要采用基于颜色、纹理和形状等特征的图像分割算法,如基于阈值的分割算法^[16]、分水岭算法^[17]、边缘检测算法^[18]等。但传统分割方法受纹理、地形等环境因素干扰大、耗时长、分割准确率较低、难以有效地处理大规模图像,适用范围有限。随着深度学习在计算机视觉和自然语言处理任务中取得的巨大成功,其也被引入无人机遥感领域,并在无人机遥感图像分析中取得了较好的表现,如城市土地利用分类、湿地测绘、道路提取等^[19-21]。

本文以无人机遥感图像为数据基础,针对黄土

高原梯田田块大小各异、形状复杂多变,且存在诸多危险作业边界等特点,拟提出一种基于多尺度特征提取与融合上采样的改进 TransUNet 模型。基于本文所提模型对直长条形、蜿蜒长条形和不规则形 3 类典型梯田分别进行提取,并与原始 TransUNet 和具有代表性的 PSPNet、HRNet V2、DeepLab V3+、U-Net 模型进行比较,以期智能农机在黄土高原梯田进行路径规划和导航作业提供技术支撑。

1 数据获取与处理

1.1 研究区域与数据采集

研究区域选择黄土高原具有代表性的甘肃省榆中县龙泉乡梯田,如图 1a 所示。龙泉乡地处黄土丘陵沟壑区,地形以山沟梁峁为主,农田多以梯田形式存在,梯田类型丰富、样本数量大,境内平均海拔 2 160 m。由于受其特殊地形地貌影响,该地区农业机械化、智能化程度相对较低。

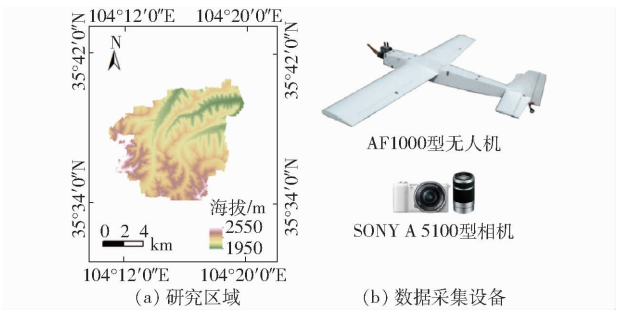


图 1 研究区域及数据采集设备

Fig. 1 Study area and data collection equipment

无人机图像数据采用 AF1000 型固定翼无人机搭载 SONY A 5100 型相机拍摄,如图 1b 所示。机体由可拆卸的模块化机体组成,既方便携带,又可以在短时间内完成组装、起飞,具有续航时间长、飞行高度高等优点。SONY A 5100 型相机有效像素为 2 430 万,单幅图像面积约 340 m × 500 m,拍摄期间天气晴朗,风力 2 ~ 3 级,全程拍摄时长约 24 h。

1.2 数据预处理

无人机航拍图像采用 Agisoft Photoscan 软件导入拼接,利用地面控制点对图像进行几何校正,采用 CGCS2000_3_Gauss_Kruger_CM_105E 投影,选择导出正射影像分辨率为 0.5 m。随机选择 5 块 5 000 像素 × 5 000 像素区域作为本次实验训练集,每个区域实际面积约为 6.25 km²。研究区域的遥感影像复杂,这些区域内除梯田以外,包含大量沟壑、建筑物、各类道路、未开发的斜坡和山脊等干涉元素,部分区域内有积雪覆盖和云层遮挡(图 2a),这

有助于提高模型泛化能力。此外,选择 3 块 1 000 像素 × 1 000 像素具有典型梯田特征的区域作为本次实验测试集,根据梯田形状分别划分为直长条形(地块 1)、蜿蜒长条形(地块 2)和不规则形(地块 3),每个区域实际面积约为 0.25 km²。其中,地块 1 区域内梯田形状相对规整,除梯田外包含较多建筑物和道路等;地块 2 区域内梯田多呈蜿蜒长条形,除梯田外包含部分沟壑及山脊;地块 3 区域内梯田多呈不规则形,并且包含较多的斜坡及山脊。

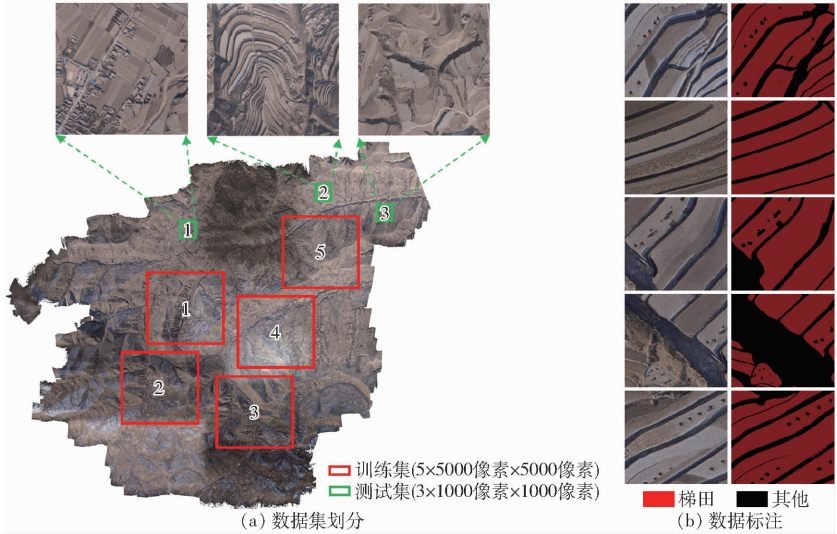


图 2 数据预处理

Fig. 2 Data pretreatment

使用 Labelme 工具对图 2 中所框选的区域进行标注,梯田区域标注为红色 (rgb(255,0,0)),其他区域标注为黑色 (rgb(0,0,0)),如图 2b 所示。标记完成后,使用重叠率 30% 的滑动窗口 (256 像素 × 256 像素) 对原始图像数据和标签数据进行裁剪,并通过旋转和镜像等方式对数据集进行扩充。最终得到 15 000 组样本数据集,按照比例 8:2 划分为训练集和验证集,测试集同样按照上述方式对直长条形、蜿蜒长条形和不规则形 3 种典型梯田区域进行处理,最终得到测试数据 720 组。

此外,由于路径规划需在单个田块内进行,单块梯田区域的预测结果直接影响农机路径规划和导航作业效果,因此额外选择直长条形、蜿蜒长条形和不规则形 3 块不同尺度的单块梯田(图 3)进行预测,得到田块实际面积分别为 2 492.05、8 395.13、2 807.03 m²。

2 研究方法

2.1 TransUNet 模型

TransUNet 是一种新型语义分割模型,通过将 Transformer 融入到 U-Net 网络中,旨在提取全局上

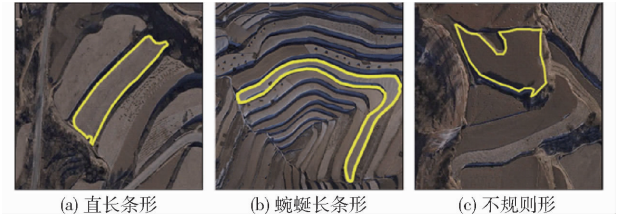


图 3 3 类单块梯田

Fig. 3 Three types of single terrace

下文特征以提高图像分割性能^[22]。Transformer 由多头注意力 (Multi-head attention, MHA) 层组成, MHA 层与卷积神经网络 (Convolutional neural network, CNN) 相比,具有全局感受野,可以从全局角度对上下文关系进行建模。

具体而言,TransUNet 具有编码器-解码器结构。其中,编码器部分由 CNN 模块和 Transformer 层组成,在 CNN 部分,首先借助 ResNet50 网络进行 3 层初步特征提取,并保留每层的输出用于上采样部分的特征融合,Transformer 将 CNN 部分提取的特征图序列化,并借助 Transformer 特有的自注意力机制捕捉图像全局信息。在解码器部分,TransUNet 将上述全局信息逐步上采样之后与底层特征图融合,有效

提高了提取任务的效果。

2.2 TransUNet 提取黄土高原梯田作业区域局限性

由于黄土高原梯田属于典型的不规则田块,田块大小各异、形状复杂多变,并且田块中存在部分凹坑和沟坎,TransUNet 中 CNN 特征提取部分缺乏对多尺度特征信息的有效提取和融合,会导致较小目标被忽略或者较大目标提取效果欠佳。并且,梯田存在诸多危险作业边界,边界提取精度直接影响农机路径规划效果和导航作业安全。然而,在上采样过程中 TransUNet 使用双线性上采样将特征逐步恢复,通常会导致棋盘效应,在相邻图像斑块之间显示不连续或不一致的预测,会降低梯田边界提取效果。

2.3 基于多尺度特征提取与融合上采样的改进 TransUNet 模型

针对黄土高原梯田的田块大小各异、形状复杂

多变,且存在大量性状相似的未开发土地等因素干扰,本文提出基于多尺度特征提取与融合上采样的改进 TransUNet 模型,如图 4 所示。在编码器部分主要由多尺度特征提取模块和残差 Transformer 层 (Residual Transformer, RT) 组成。前者用于不同尺度梯田图像的特征提取和融合,后者用于全局上下文特征的联系。在多尺度特征提取部分,引入金字塔压缩注意力模块 (Pyramid squeeze attention, PSA)^[23] 替换原始 ResNet 中的 3×3 卷积块,该模块在通道注意力的基础上引入了多尺度思想,以增强图像中不同尺度的梯田特征信息 (图 5a)。受残差网络中恒等式块^[24] 的启发,本文优化 TransUNet 堆叠 Transformer 层结构,并引入残差连接 (图 5b),使模型更易于训练,并缓解反向传播过程中梯度消失的问题。

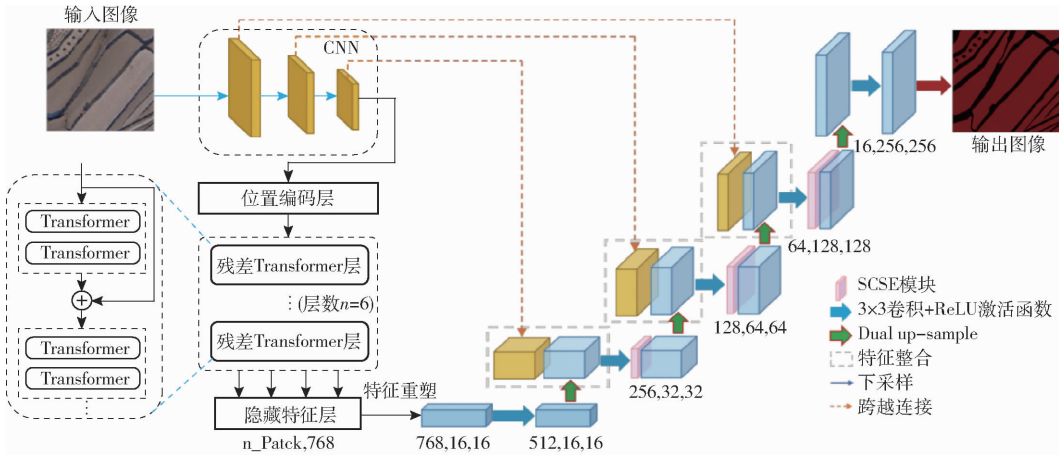


图 4 改进 TransUNet 模型结构图
Fig. 4 Improved TransUNet model structure

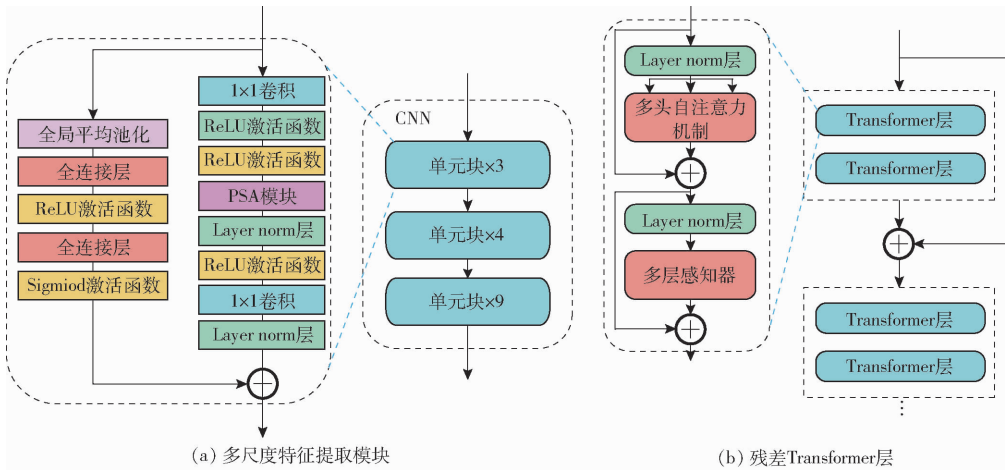


图 5 模型编码器结构图
Fig. 5 Model encoder structure

在金字塔压缩注意力模块中 (图 6), 首先使用 Conv 卷积函数对 S 个初步特征层 (本文 S 设置为 4: F_0, F_1, F_2, F_3) 进行空间多尺度特征提取。该模块使用群卷积的方法, 将特征图通道平均分成 M 份

(M 为群卷积数量), 对每一份特征图分别进行运算后再使用 Cat 函数进行通道合并, 在不增加计算量的情况下以不同的核尺度处理输入张量。计算式为

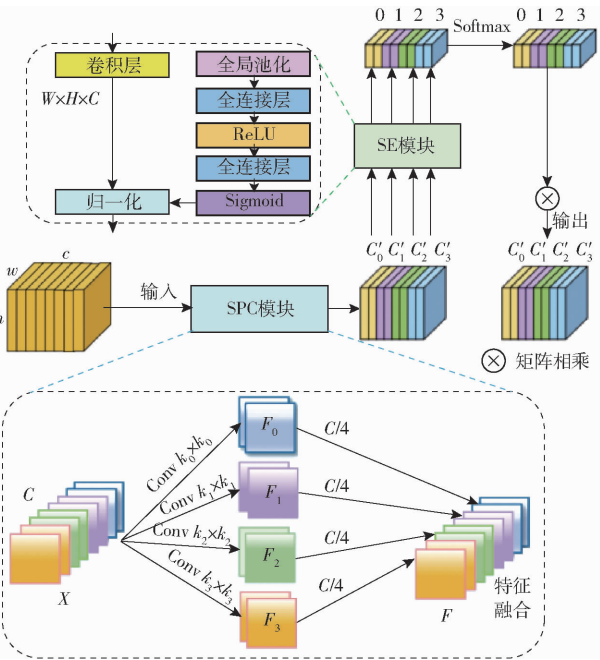


图 6 金字塔压缩注意力模块

Fig.6 Pyramid squeeze attention module

$$C' = \frac{C}{S} \tag{1}$$

$$k_i = 2(i + 1) + 1 \quad (i = 0, 1, \cdots, S - 1) \tag{2}$$

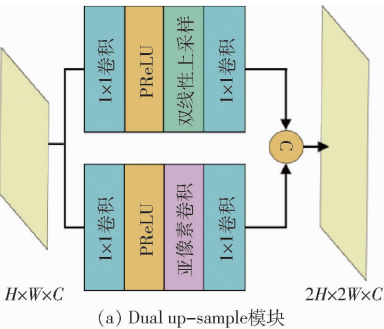
$$G_i = 2^{\frac{k_i - 1}{2}} \quad (i = 0, 1, \cdots, S - 1) \tag{3}$$

$$F_i = \text{Conv}(k_i \times k_i, G_i)(X) \quad (i = 0, 1, \cdots, S - 1) \tag{4}$$

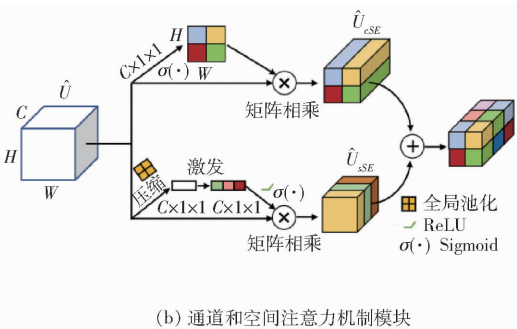
$$F = \text{Cat}([F_0, F_1, \cdots, F_{S-1}]) \tag{5}$$

- 式中 S ——获取不同尺度数量
 C ——输入特征通道数
 C' ——每个尺度获取通道数
 k_i ——第 i 组卷积核大小
 G_i ——第 i 组所对应的群卷积数量
 F_i ——第 i 组空间特征提取结果
 X ——输入卷积特征图
 F ——通道合并后结果

然后,通过挤压和激励(Squeeze and excitation, SE)模块提取具有不同比例特征图的注意力,即通过 SEWeight 函数获得逐通道的注意力向量。通过



(a) Dual up-sample模块



(b) 通道和空间注意力机制模块

图 7 模型解码器结构图

Fig.7 Model decoder structure

Softmax 函数重新校准通道注意力向量来获得多尺度通道权重,Softmax 可以跨通道自适应地选择不同空间尺度,使局部特征与全局特征得到有效融合。最终,将重新校准的权重和相应的特征图进行元素相乘,获得更丰富的多尺度特征信息的精细特征图作为输出。

为提升梯田边界提取精度,本文模型在解码器部分引入来自图像超分辨率领域的亚像素卷积。亚像素卷积可以通过将多通道特征上的单个像素组合成一个特征上的单位,每个特征上的像素就相当于新特征上的亚像素,最终实现从低分辨率图到高分辨图的重构。Dual up-sample 模块^[25]将双线性插值上采样与亚像素卷积层两者融合,能有效改善图像边缘提取精度并缓解棋盘效应,如图 7a 所示。最后,本文模型在每个解码器的末尾添加通道和空间注意力机制模块(Concurrent spatial and channel squeeze and channel excitation, SCSE)^[26],同时对空间维度和通道维度信息进行整合增强,有助于上采样过程中图像边界等细节特征的逐步恢复,如图 7b 所示。

3 试验结果与分析

3.1 模型训练

在图像工作站进行网络模型训练,图像工作站搭配 Intel(R) Xeon(R) Gold 5118 CPU @ 2.30 GHz 2.29 GHz 双处理器、32 GB 运行内存和 NVIDIA GeForce RTX 2080Ti GPU。使用 Anaconda3(64 位)配置虚拟环境,安装 Python 3.7 并选择 Pytorch 作为深度学习框架,将 PyCharm 2017 作为开发平台。将训练集和验证集分为 150 个训练轮次,批量大小为 4。优化器选择 Adam,初始学习率设置为 0.001,随着迭代次数提升,降低学习率至 0.000 1。

基于训练样本数据,对比当前具有代表性的 PSPNet、HRNet V2、DeepLab V3 +、U-Net、TransUNet 与本文模型的训练效果,各模型训练结果如图 8 所示。本文模型在训练损失函数曲线上波动

更稳定、收敛效果更好,并且在平均交并比曲线上能取得更高的分数。

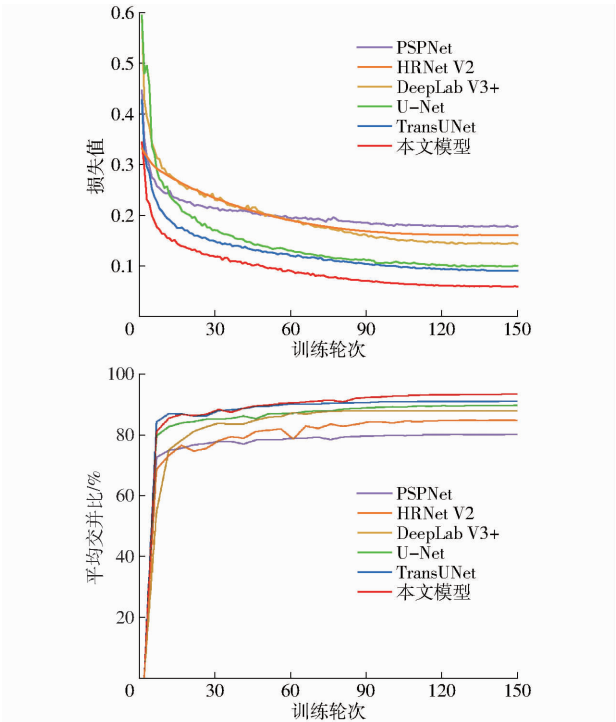


图 8 模型训练曲线
Fig.8 Model training curves

3.2 评价指标

选用 F1 值、平均像素准确率 (Mean pixel accuracy,MPA)、平均交并比 (Mean intersection over union,MIoU)3 个指标来评价网络模型性能,选择像素交并比 (Intersection over union,IoU) 评价模型预测单块梯田区域效果。其中,F1 值为精确率和召回率的调和平均值;平均像素准确率 (MPA) 为测试集上所有类别预测正确数量与总体数量比值的平均值;平均交并比 (MIoU) 为所有类别像素真实值与预测值的交集与并集比值的平均值;像素交并比 (IoU) 为像素真实值、预测值交集与并集的比值。

3.3 消融实验

为验证各模块对网络的贡献,使用原始的 TransUNet 网络为基线模型进行消融实验,如表 1 所示。网络输入尺寸为 256 像素 × 256 像素的梯田图像,使用 MPA、F1 值、MIoU 来衡量网络性能。在消融实验中,基础模块为 PSA、RT、Dual up - sample 和 SCSE。本文采用的策略为,一旦一个模块被证明有效,默认其在随后研究中有效。为了严谨起见,本文选择了几个特殊情况,例如 SETransUNet,以验证不同的模块在添加到原始 TransUNet 时是否仍然有效。

表 1 消融实验							%
Tab.1 Ablation experiment							
模型	PSA	RT	Dual up - sample	SCSE	MPA	F1 值	MIoU
TransUNet	×	×	×	×	94.8	93.9	90.1
TransUNet_PSA	√	×	×	×	95.5	94.8	91.3
TransUNet_RT	×	√	×	×	95.0	94.3	90.4
TransUNet_PSA_RT	√	√	×	×	95.6	95.2	91.6
SETransUNet	×	×	×	√	95.3	94.8	91.1
TransUNet_Du	×	×	√	×	95.1	94.3	90.5
SETransUNet_Du	×	×	√	√	95.6	95.2	91.4
SETransUNet_PSA_RT	√	√	×	√	95.8	95.6	92.0
本文模型	√	√	√	√	96.0	96.0	92.3

注:“√”表示采用此模块,“×”表示未采用此模块。

金字塔压缩注意力模块在通道注意力的基础上引入多尺度的思想,能有效提高各项指标。如表 1 中的 TransUNet_PSA,相较于原 TransUNet 模型,MPA 上升 0.7 个百分点,F1 值上升 0.9 个百分点,MIoU 上升 1.2 个百分点。残差 Transformer 模块通过优化 Transformer 层结构并加入残差连接,使模型更易于训练,同时各项指标也有一定的提高。Dual up - sample 模块将亚像素卷积层与双线性插值上采样两者融合,提升上采样时特征提取能力与图像边缘提取精度,各项指标有不同程度的提升。通道和空间注意力机制模块能同时对空间维度和通道维度进行信息整合增强,有助于图像细节特征逐

步恢复。如表 1 中 SETransUNet 所示,相较于原 TransUNet 模型,MPA 上升 0.5 个百分点,F1 值上升 0.9 个百分点,MIoU 上升 1.0 个百分点。

3.4 模型性能评价

如表 2 所示,PSPNet 仅增加了金字塔池化模块,总体性能较弱;HRNet V2 改变了传统卷积神经网络自上而下的特征提取工作模式,使得网络在训练时一直保持高分辨率特征图,MIoU 达到 83.4%;DeepLab V3 + 采用空洞卷积,考虑了特征提取和多尺度融合,但缺乏注意力机制,MIoU 达到 85.3%;U - Net使用多层特征融合,MIoU 达到 86.5%;TransUNet 在特征提取部分融合了 CNN 和

Transformer 层,对局部特征和全局特征综合考虑,极大提升网络整体性能,各项指标都有较大程度提高;由于本文模型在 TransUNet 的 CNN 部分增加多尺度特征提取块,并改进 Transformer 层结构,同时优化了传统的上采样方式,并加入混合维度的 SCSE 对空间维度和通道维度进行信息整合增强,因此取得了最佳实验效果,MPA 达到 96.0%,F1 值达到 96.0%,MIoU 达到 92.3%。相较于原 TransUNet 模型,3 项指标平均提升 1.8 个百分点,相较于上述其他模型平均提升 8.3、6.2、5.0、4.2 个百分点。如表 3 所示,采用本文模型预测的 3 块不同尺度单块梯田均能达到较高的像素交并比,平均可达 97.0%,相较于上述 5 种模型平均提升 1.0、2.1、2.8、7.4、11.7 个百分点。

表 2 不同模型提取梯田性能评价结果

Tab.2 Performance evaluation of different models
by terrace segmentation %

模型	MPA	F1 值	MIoU
PSPNet	89.5	89.8	80.1
HRNet V2	91.0	91.2	83.4
DeepLab V3 +	92.0	92.1	85.3
U - Net	92.6	92.5	86.5
TransUNet	94.8	93.9	90.1
本文模型	96.0	96.0	92.3

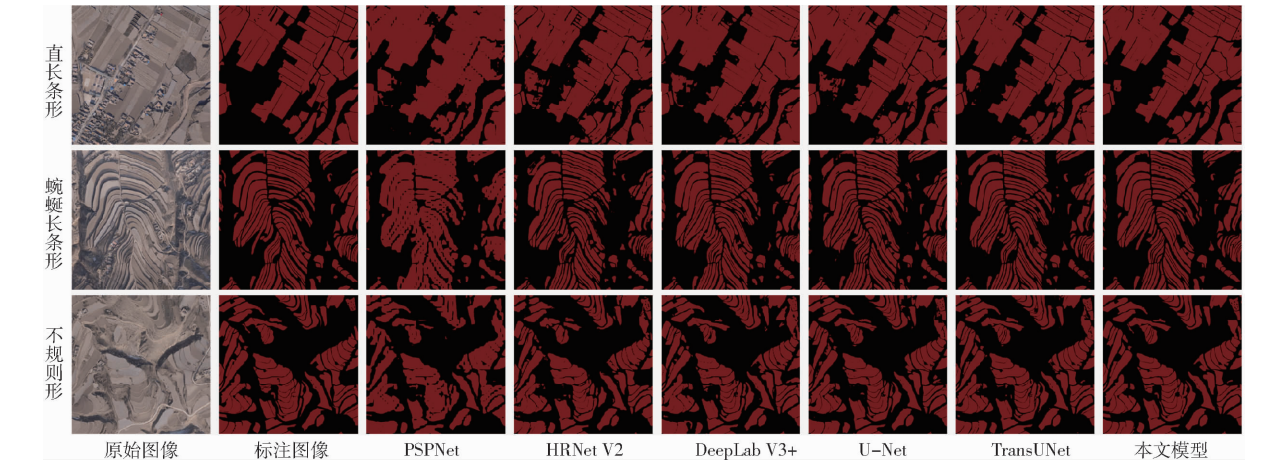


图 9 不同模型梯田测试集预测效果对比

Fig.9 Comparison of terrace test set prediction effects for different models

合的方式代替传统的上采样方案,并加入混合维度的 SE 模块帮助图像细节特征的逐步恢复,总体预测效果最优,梯田间基本无连片区域,错分、漏分现象较少,较小区域被忽略或较大块梯田提取效果欠佳的问题得到有效改善,边界等细节信息更加完整。对预测拼接后的图像掩膜处理得到 3 类单块梯田预测图,如图 10 所示。采用本文模型预测的 3 块不同尺度单块梯田效果均为最佳,梯田区域预测结果更加接近真实作业区域,边界等细节特征预测更

表 3 不同模型单块梯田像素交并比

Tab.3 Intersection over union of single-field
in different models %

地块	PSPNet	HRNet V2	U - Net	DeepLab V3 +	TransUNet	本文模型
直长条形	87.6	95.9	95.6	95.6	97.0	97.1
蜿蜒长条形	83.8	88.0	92.6	94.5	94.7	96.2
不规则形	84.5	84.8	94.2	94.6	96.2	97.6

3.5 可视化分析

如图 9 所示,在直长条形、蜿蜒长条形、不规则形 3 类典型特征梯田测试集上,PSPNet 出现了较多连片区域,在 3 类典型区域测试集上总体提取效果较弱;HRNet V2 出现了较多的错分、漏分区域,整体预测效果相对有所提升;DeepLab V3 + 误提取的情况有所改善,但边界等细节特征比较混乱;U - Net 同样出现了部分误提取现象,并且边界信息相对粗糙;TransUNet 通过融合 CNN 和 Transformer 层,同时对局部特征和全局上下文特征进行提取,相较于只使用 CNN 的模型预测效果更好,但仍存在部分较小区域被忽略或较大块梯田提取效果欠佳,以及梯田边界不完整现象;而本文模型在通道注意力机制基础上增强了对梯田多尺度特征提取和融合的能力,并且采用双线性插值上采样与亚像素卷积层两者融

加精准,尤其是相对复杂的蜿蜒长条形和不规则形两类梯田提升最明显,验证了改进的各模块在解译梯田遥感影像方面的有效性。

4 结论

(1)针对黄土高原梯田形状复杂多变、田块大小各异,且存在诸多危险作业边界等特点,提出了一种基于多尺度特征提取与融合上采样的改进 TransUNet 模型。通过引入 PSA 模块增强了模型对

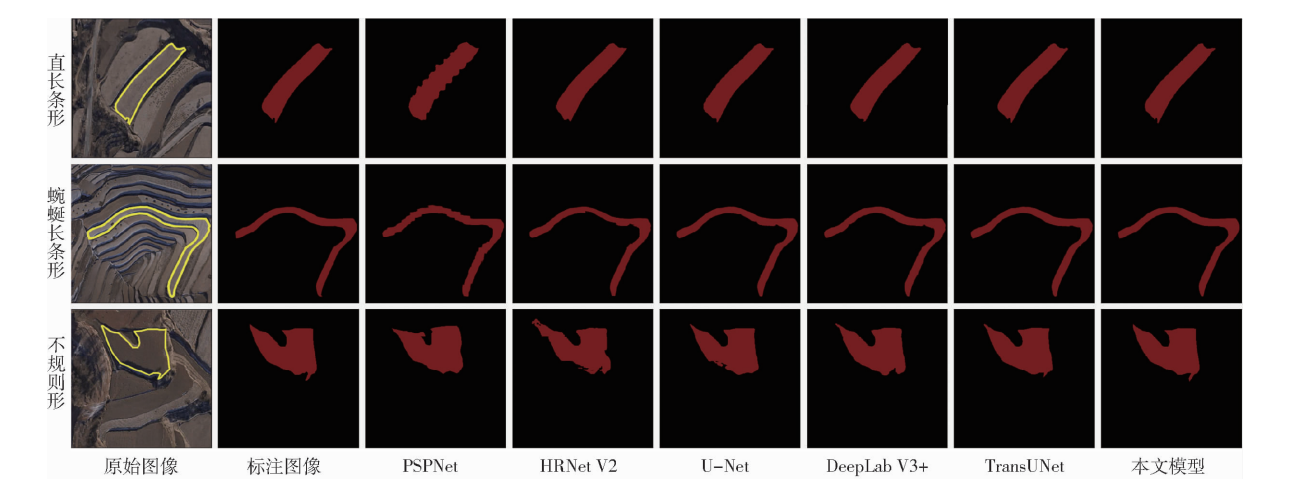


图 10 3 类单块梯田预测效果对比

Fig. 10 Comparison of single-field block predictions for three types of terraces

梯田多尺度特征提取和融合的能力,同时使用残差结构优化了 Transformer 层,并将亚像素卷积层与双线性插值上采样两者融合代替传统的上采样方案,最后在解码器末尾添加 SCSE 模块对空间和通道维度信息进行整合增强,有助于图像细节特征逐步恢复。

(2)在直长条形、蜿蜒长条形和不规则形 3 类典型特征梯田测试集上,本文模型具有更高的预测精度和更好的预测效果,MPA、F1 值和 MIoU 分别可

达 96.0%、96.0% 和 92.3%,相较于改进前平均提升 1.8 个百分点。与代表性的 PSPNet、HRNet V2、DeepLab V3+、U-Net 模型相比,3 项指标平均提高 8.3、6.2、5.0、4.2 个百分点。本文模型在不同尺度单块梯田的预测结果更加精准,边界混乱及错分、漏分现象得到改善,表明该模型能有效应用于黄土高原梯田遥感影像提取中,可为梯田环境地图构建提供参考。

参 考 文 献

[1] 孟志军,王昊,付卫强,等. 农业装备自动驾驶技术研究现状与展望[J]. 农业机械学报,2023,54(10):1-24.
MENG Zhijun, WANG Hao, FU Weiqiang, et al. Research status and prospects of agricultural machinery autonomous driving[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 1-24. (in Chinese)

[2] 周俊,何永强. 农业机械导航路径规划研究进展[J]. 农业机械学报,2021,52(9):1-14.
ZHOU Jun, HE Yongqiang. Research progress on navigation path planning of agricultural machinery[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 1-14. (in Chinese)

[3] 毛文菊,刘恒,王小乐,等. 双导航模式果园运输机器人设计与试验[J]. 农业机械学报,2022,53(3):27-39,49.
MAO Wenju, LIU Heng, WANG Xiaole, et al. Design and experiment of dual navigation mode orchard transport robot[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3): 27-39, 49. (in Chinese)

[4] 刘志杰,王小乐,任志刚,等. 基于虚拟雷达模型的履带拖拉机导航路径跟踪控制算法[J]. 农业机械学报,2021,52(6):376-385.
LIU Zhijie, WANG Xiaole, REN Zhigang, et al. Crawler tractor navigation path tracking control algorithm based on virtual radar model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(6): 376-385. (in Chinese)

[5] 王宁,韩雨晓,王雅萱,等. 农业机器人全覆盖作业规划研究进展[J]. 农业机械学报,2022,53(增刊1):1-19.
WANG Ning, HAN Yuxiao, WANG Yaxuan, et al. Research progress of agricultural robot full coverage operation planning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(Sup. 1): 1-19. (in Chinese)

[6] QIN J, WANG W, MAO W, et al. Research on a map-based cooperative navigation system for spraying-dosing robot group[J]. Agronomy, 2022, 12(12): 3114.

[7] GALCERAN E, CARRERAS M. A survey on coverage path planning for robotics[J]. Robotics and Autonomous Systems, 2013, 61(12): 1258-1276.

[8] SUN Q, ZHANG R, CHEN L, et al. Semantic segmentation and path planning for orchards based on UAV images[J]. Computers and Electronics in Agriculture, 2022, 200: 107222.

[9] ZHANG Y, SHI M, ZHAO X, et al. Methods for automatic identification and extraction of terraces from high spatial resolution satellite data (China-GF-1)[J]. International Soil and Water Conservation Research, 2017, 5(1): 17-25.

[10] LUO L, LI F, DAI Z, et al. Terrace extraction based on remote sensing images and digital elevation model in the Loess Plateau, China[J]. Earth Science Informatics, 2020, 13: 433-446.

[11] ZHAO H, FANG X, DING H, et al. Extraction of terraces on the Loess Plateau from high-resolution DEMs and imagery

- utilizing object-based image analysis[J]. ISPRS International Journal of Geo-Information, 2017, 6(6): 157.
- [12] NIU B, FENG Q, CHEN B, et al. HSI-TransUNet: a transformer based semantic segmentation model for crop mapping from UAV hyperspectral imagery[J]. Computers and Electronics in Agriculture, 2022, 201: 107297.
- [13] 张宏鸣,胡勇,杨勤科,等. 基于影像与坡度数据融合的梯田田块分割方法[J]. 农业机械学报, 2018, 49(4): 249-256.
ZHANG Hongming, HU Yong, YANG Qinke, et al. Segmentation method of terraced fields based on image and gradient data [J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(4): 249-256. (in Chinese)
- [14] 杨蜀秦,林丰山,徐鹏辉,等. 基于无人机遥感影像的多生育期冬小麦种植行检测方法[J]. 农业机械学报, 2023, 54(2): 181-188.
YANG Shuqin, LIN Fengshan, XU Penghui, et al. Row detection of multi-growth winter wheat field based on UAV remote sensing image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(2): 181-188. (in Chinese)
- [15] OSCO L P, JUNIOR J M, RAMOS A P M, et al. A review on deep learning in UAV remote sensing[J]. International Journal of Applied Earth Observation and Geoinformation, 2021, 102: 102456.
- [16] 邵志明,王怀彬,董志城,等. 基于近红外相机成像和阈值分割的苹果早期损伤检测[J]. 农业机械学报, 2021, 52(增刊): 134-139.
SHAO Zhiming, WANG Huaibin, DONG Zhicheng, et al. Early bruises detection method of apple surface based on near infrared camera imaging technology and image threshold segmentation method[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(Supp.): 134-139. (in Chinese)
- [17] 丁永军,张晶晶,LEE Won Suk,等. 小波变换与分水岭算法融合的番茄冠层叶片图像分割[J]. 农业机械学报, 2017, 48(9): 32-37.
DING Yongjun, ZHANG Jingjing, LEE Won Suk, et al. Segmentation of tomato leaves from canopy images by combination of wavelet transform and watershed algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(9): 32-37. (in Chinese)
- [18] 杨泽楠,牛海鹏,黄亮,等. 基于 MSR-cut 的高空间分辨率遥感影像边缘检测分割[J]. 农业机械学报, 2021, 52(8): 154-162.
YANG Ze'nan, NIU Haipeng, HUANG Liang, et al. Edge detection segmentation method for high spatial resolution remote sensing image based on MSR-cut[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(8): 154-162. (in Chinese)
- [19] 杨蜀秦,王鹏飞,王帅,等. 基于 MHSA + DeepLab v3+ 的无人机遥感影像小麦倒伏检测[J]. 农业机械学报, 2022, 53(8): 213-219.
YANG Shuqin, WANG Pengfei, WANG Shuai, et al. Detection of wheat lodging in UAV remote sensing images based on multi-head self-attention DeepLab v3+ [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(8): 213-219. (in Chinese)
- [20] YAN C, LI Z, ZHANG Z, et al. High-resolution mapping of paddy rice fields from unmanned airborne vehicle images using enhanced-TransUNet[J]. Computers and Electronics in Agriculture, 2023, 210: 107867.
- [21] YUAN Q, SHEN H, LI T, et al. Deep learning in environmental remote sensing: achievements and challenges[J]. Remote Sensing of Environment, 2020, 241: 111716.
- [22] CHEN J, LU Y, YU Q, et al. TransUNet: transformers make strong encoders for medical image segmentation[J]. arXiv Preprint, arXiv: 2102.04306, 2021.
- [23] ZHANG H, ZU K, LU J, et al. EPSANet: an efficient pyramid squeeze attention block on convolutional neural network[C]// Proceedings of the Asian Conference on Computer Vision, 2022: 1161-1177.
- [24] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [25] FAN C, LIU T, LIU K. SUNet: swin transformer UNet for image denoising[C]// 2022 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 2022: 2333-2337.
- [26] ROY A G, NAVAB N, WACHINGER C. Concurrent spatial and channel 'squeeze & excitation' in fully convolutional networks[C]// Medical Image Computing and Computer Assisted Intervention-MICCAI 2018, 2018: 421-429.