

基于注意力机制和边缘感知的田埂提取模型

顾兴健¹ 刘子儒¹ 任守纲^{1,2} 郑恒彪^{2,3} 徐焕良¹

(1. 南京农业大学人工智能学院, 南京 210095; 2. 国家信息农业工程技术中心, 南京 210095;

3. 南京农业大学农学院, 南京 210095)

摘要: 田埂精确提取是数字化农业管理的重要前提。针对由于遮挡、斑秃等因素干扰, 给基于语义分割方法提取田埂带来困难问题, 提出一种基于注意力机制和边缘感知模块的 U-Net 网络实现田埂提取。首先, 将多信息注意力引入 U 型分割网络的下采样中, 增强相邻层之间的上下文信息, 提升对田埂区域语义特征的表示能力。其次, 将边缘感知分割模块应用至 U-Net 解码部分的每一层, 在不同语义特征层提取田埂边缘信息, 提高田埂区域语义分割精度。最后, 联合边缘感知损失与语义分割损失构建联合损失函数, 用于整体网络优化。通过对安徽省淮北市濉溪县小麦基地采集的无人机麦田数据集进行训练和模型验证, 实验结果表明, 本文模型语义分割像素准确率高达 95.57%, 平均交并比达到 77.48%。

关键词: 田埂提取; 语义分割; U-Net; 注意力机制; 边缘感知

中图分类号: TP391.41

文献标识码: A

文章编号: 1000-1298(2023)05-0210-09

OSID:



Ridge Extraction Model Based on Attention Mechanism and Edge Perception

GU Xingjian¹ LIU Ziru¹ REN Shougang^{1,2} ZHENG Hengbiao^{2,3} XU Huanliang¹

(1. College of Artificial Intelligence, Nanjing Agricultural University, Nanjing 210095, China

2. National Engineering and Technology Center for Information Agriculture, Nanjing 210095, China

3. College of Agriculture, Nanjing Agricultural University, Nanjing 210095, China)

Abstract: Accurate extraction of ridges is an important prerequisite for digital agricultural management. However, due to the interference of factors such as occlusion and alopecia areata, it brings challenges for the semantic segmentation method to extract the ridge area. A U-Net segmentation network model was proposed based on a multi-information attention mechanism and an edge-aware module. Firstly, multi-information attention was introduced into the down-sampling of the U-shaped network to enhance the context information between adjacent layers and improve the representation ability of the semantic features of the ridge area. Secondly, the edge-aware segmentation module was applied to each layer of the U-Net decoding part, and the ridge edge information was extracted in different semantic feature layers to improve the semantic segmentation accuracy of the ridge region. Finally, the joint edge-aware loss and semantic segmentation loss were used to construct a joint loss function for overall network optimization. The training and model validation were carried out with the UAV wheat field data set collected by the wheat experimental base in Suixi County, Huaibei City, Anhui Province. The experimental results showed that the pixel accuracy of semantic segmentation of crop plants in different datasets was as high as 95.57%, and the average intersection ratio was 77.48%.

Key words: ridge extraction; semantic segmentation; U-Net; attention mechanism; edge perception

收稿日期: 2022-09-24 修回日期: 2023-01-17

基金项目: 国家重点研发计划项目(2019YFE0125500-04)、国家自然科学基金青年基金项目(32101617)、江苏省农业科技自主创新项目(CX(22)3201)和中国博士后科学基金项目(2022T150327)

作者简介: 顾兴健(1985—),男,讲师,博士,主要从事图像处理 and 深度学习研究, E-mail: guxingjian@njau.edu.cn

通信作者: 郑恒彪(1990—),男,副研究员,博士,主要从事智慧农业和农情遥感监测研究, E-mail: zhenghb@njau.edu.cn

0 引言

近地无人机遥感图像分割实现农田田埂提取是数字化农业管理的重要前提。早期,遥感图像的分割通常是基于阈值、颜色、分水岭等形态学方法^[1-4]。随着深度学习方法的发展,其在遥感领域的应用也越来越广泛。黄亮等^[5]提出一种基于结合 RGB 波段和轻量级 U-Net 实现无人机遥感灯盏花种植信息提取。赵静等^[6]提出将 FCN 用于玉米田垄线区域的提取工作,但面对不同宽度的垄田区域,FCN 的分割精度并不理想,仍需依靠后期的霍夫变换进一步提升精度。杨蜀秦等^[7]提出将 SegNet 深度语义分割网络和霍夫变换结合实现冬小麦种植行的检测。SONG 等^[8]基于 SegNet 网络实现向日葵种植区域的分割,尽管在细节上 SegNet 较 FCN 有了较大保留,并且准确率也提升至 86.2%,但在复杂特征的学习和分割结果的平滑程度上仍存在较为明显的问题。

为使卷积核感受野更好地捕获目标的上下文信息,张永辉^[9]通过 PSPNet 模型对湿地的芦苇草甸区域进行提取,实验结果表明该方法的准确率为 79.2%,尽管 PSPNet 在全局信息的挖掘上确有较大改进,但由于其网络结构复杂,在处理具有多信息的高分辨率图像时其效率有所欠缺。苏健民等^[10]则选择采用 U-Net 网络对城市遥感图像进行道路提取工作,该网络不仅满足了平滑边沿和丰富上下文的需求,同时也具备了轻量的特点,最终使得分割准确率达 90%。在以上语义分割网络的引入下,复杂背景下的目标分割效果均得到了较大提升。其中,针对高分辨率遥感图像大尺度小样本及纹理细节难以分辨等情况,U-Net 因其清晰结构、兼顾深浅层特征以及良好的迁移性,在该领域体现出了较为明显的优势。

农田田间自然环境较为复杂,给田块自动提取带来困难。田间因播种不均等原因造成的“斑秃”区域与田埂的纹理特征相似度高,通过单一的基础模型难以得到很好的分割结果。王施云等^[11]提出一种基于空洞卷积的特征融合策略扩大 U-Net 的特征捕获范围,从而提升遥感图像的分割效果,实验证明该方案的分割准确率较原网络提升 3.5%,但该方法在多类且目标差异性大的数据集上体现的效果更为显著,而针对少类且多信息高分辨率的数据时其分割效果却差强人意。同时,由于作物真实生长情况的不可控性,时常会出现田埂被作物遮挡的情况,这势必对后续分割产生显著影响。王斌等^[12]基于 U-Net 对城市道路的提取结果,对断点区域运

用 K-means 聚类算法进行了修复,而该方法并未从网络层面进行改进,而是利用骨架法和邻域检测实现道路修复工作。

传统的将边缘特征或边缘引导的分割方法应用于灰度不均、边缘信息弱的农田遥感图像时,边缘会出现许多非封闭区域的现象。将边缘检测算法和区域分割算法有效地结合,可得到封闭的区域边界。YUAN 等^[13]提出了一种模型无关的后处理方案,即用内部像素的预测代替原来不可靠的边界像素预测,以提高由任何现有分割模型生成的分割结果的边界质量。FREIXENET 等^[14]提出嵌入式结合分割方法和后处理结合分割方法,有效地形成了闭合区域。但该方法对光谱丰富的区域会造成错分现象。GUO 等^[15]提出了一种新的从粗到细的边界细化网络来提取建筑物,保留了层次特征的高级语义和低级空间细节,通过利用 BR 模块中高分辨率特征的空间详细信息,进一步细化了建筑物的预测。但该方法并未在编码部分加强对物体的深层特征的提取。此外,田埂目标区域占比失调的情况在遥感图像中也十分常见。由于田埂自身存在狭窄稀少的特点,在麦田的遥感图像中尤为突出。陈燕生等^[16]基于浅层特征和损失函数对 U-Net 进行了改进,解决了小且零散小宗作物区域分割问题,实验表明该模型的准确率提升 5%,但该方案忽略了对深层信息的运用而仅关注了浅层空间特征的处理效果。因此,为了有效解决田间复杂背景下田埂的提取问题,本文在田间自然环境下,采集麦田的近地高分辨率遥感图像,并提出一种基于多信息注意力机制和边缘感知模块的 U-Net 分割网络,实现大田环境下的田埂的精确提取。模型中通过加入兼顾通道和空间信息的多信息注意力模块(Multi information attention mechanism, MIAM),加强对于麦田斑秃区域和田埂相似特征差异的学习。同时,在上采样过程中的每一阶段增加边缘感知模块,提高田埂区域语义分割精度。通过量化指标评估和可视化结果的方式,对模型及结果进行分析,为后期小麦田块自动划分工作提供理论基础和技术保障。

1 材料与方法

1.1 图像采集

实验时麦田遥感图像采集地点为安徽省淮北市濉溪县小麦基地(面积约 96 667 m²),采集时间分别为 2020 年 5 月 4 日(图 1a)和 2020 年 5 月 27 日(图 1b)。使用大疆(DJI)精灵 Phantom 4 RTK 型无人机设备,其搭载 2 000 万像素的 RGB 相机。设定飞行高度 30 m,飞行路线覆盖整个基地。经过无人

机航拍影像的拼接、正射等处理得到尺寸为 20 788 像素×15 094 像素的小麦实验基地影像。

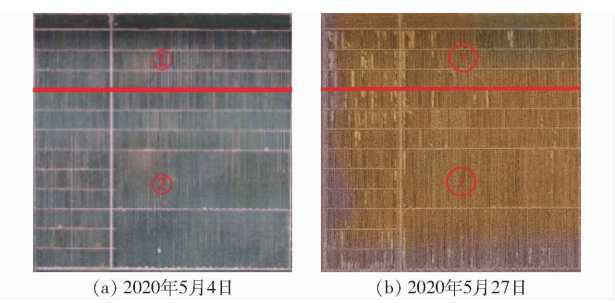


图 1 小麦实验基地遥感图像

Fig. 1 Remote sensing images of wheat experimental base

1.2 训练集和测试集建立

为验证本文所提模型的有效性,实验选取小麦基地遥感图像的区域①(图 1)为训练集数据源,区域②为测试集的数据源。由于模型输入特征维度限制,需要对小麦实验基地图像进行裁剪,生成每幅图像为 512 像素×512 像素×3 通道规格的数据集。利用滑动窗法对两区域进行统一尺寸的图片集获取,其中训练集采用步长为 0.6 倍窗口宽度的重叠滑动方式进行采集,测试集采用无重叠滑动方式进行采集,得到训练数据集 1 456 幅和测试数据集 502 幅。将田埂作为目标提取对象,通过 LabelMe 软件^[17]人工标注田埂区域生成像素级标签文件,如图 2 所示。

为提高模型的泛化能力,本文使用 Pytorch 框架下的数据增强方法。在训练过程中对每个 batch 数据随机选择一种增强方法,无需增加原有数据集数量。主要数据增强手段包括:①旋转:对图像进行水平翻转、垂直翻转 2 种翻转方式,可以防止模型训练过拟合。②随机亮度、对比度增强:改变图像亮度或饱和度,模拟不同时间段的光照差异。

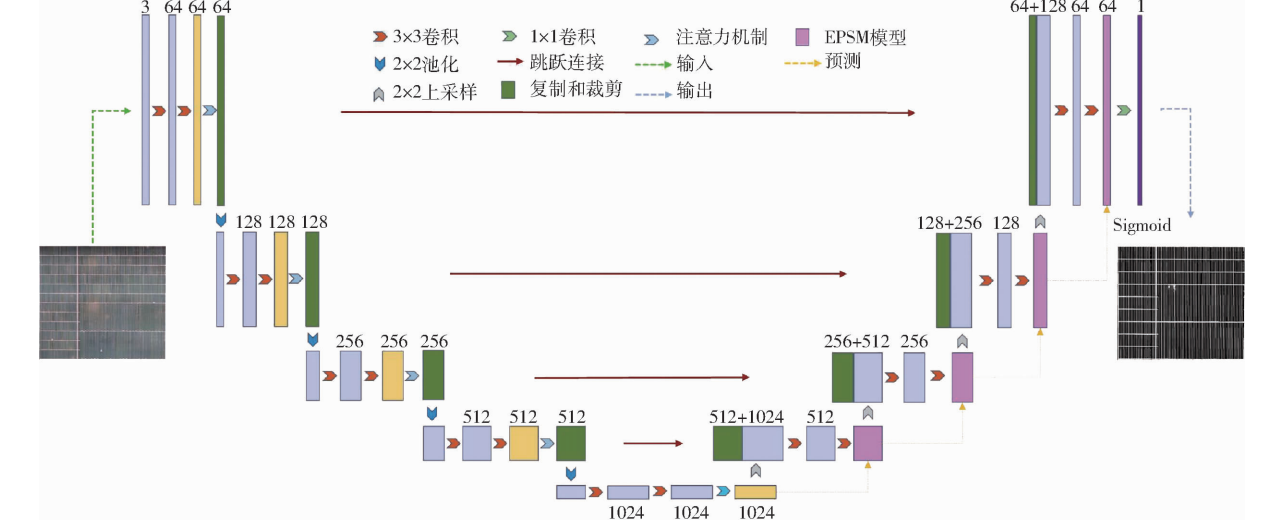


图 3 本文网络结构

Fig. 3 Architecture of network

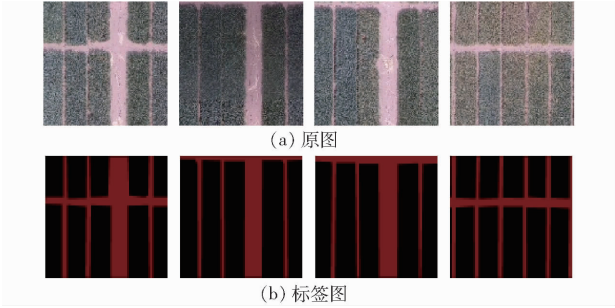


图 2 数据集示意图

Fig. 2 Diagram of dataset

2 麦田图像的语义分割模型

本文提出一种适用于近地高分辨率遥感麦田图像的新型语义分割网络。网络从注意力模块和边缘感知两方面,对 U-Net 进行了改进,主要解决田埂与田块特征差异小及区域占比少的问题,使其能够适用于实际自然环境场景。

2.1 模型结构

设计一种基于 U-Net^[18] 的遥感图像分割模型,如图 3 所示。模型包括 3 部分:①由嵌入多信息注意力模块的卷积块组成的编码器进行特征提取,得到经过加强的新特征 X1'~X5'。②将得到的高维特征通过反卷积进行逐层上采样,使其重塑为原始特征图的分辨率。③通过跳跃连接结构,将编码器提取到的特征信息同上采样的特征图进行拼接,并通过卷积层实现特征融合,同时,在解码器每层的卷积之后都加入边缘感知模块,它可以结合边缘信息和方向信息,对本层的分割结果进行优化。最终通过 1×1 卷积层实现麦田分割。

2.2 融合多信息的注意力模块

注意力机制(Attention)通过合理分配目标对象的特征权重^[19],改善卷积在局部感受野上对特征图

的无差别信息处理,关注有效信息,提升模型鲁棒性。经过 U-Net 分割的麦田图像仍因外界因素干扰导致田埂特征学习不佳,而注意力机制的加入能够改善模型的学习效果,达到良好分割的效果。

本文从通道和空间^[20] 2 个维度分别建立多信息注意力模块 (Multi-information attention module, MIAM),其结构如图 4 所示。该模块主要包含 2 个

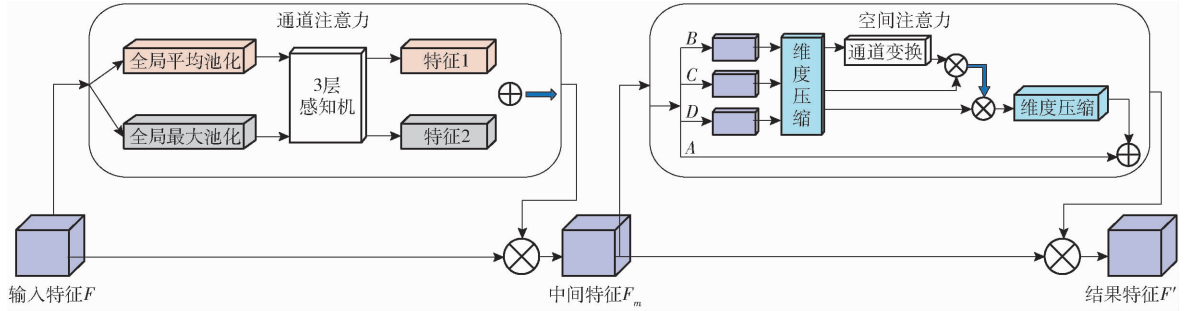


图 4 多信息注意力模块结构示意图

Fig. 4 Structure diagram of multi-information attention module

通道注意力模块由平均池化 (Average pooling)、最大池化 (Max pooling) 和多层感知机 (Multi-layer perception, MLP) 组成。经过 2 种不同的池化方法实现对特征图的空间汇聚,生成 2 类通道注意力向量。由共享的 MLP 网络对所得 2 类注意力向量生成 2 个维度 $C \times 1 \times 1$ 的注意力向量,其中 C 表示通道数。最后通过加权和 Softmax 函数获得通道维度上具有更强指向性的新特征 F_m ^[24]。该模块的计算公式为

$$M_c = \sigma(\text{MLP}(\text{AvgPool}(F))) + \text{MLP}(\text{MaxPool}(F)) \quad (1)$$

式中 M_c ——通道注意力 σ ——Sigmoid 函数
MLP——特征共享网络
AvgPool(F)——平均池化特征
MaxPool(F)——最大池化特征

空间注意力模块则利用潜在信息关系建立^[21-22],在局部特征上捕获更为丰富的上下文内容。该模块由 1×1 卷积、重塑层 (Reshape) 组成。将前一阶段所得特征 F_m 作为本模块输入特征 A ,通过 3 个 1×1 卷积映射成 3 个新特征 $\{B, C, D\} \in \mathbf{R}^{C \times N}$ 。通过点积矩阵分别对特征 B, C ,以及二者的空间注意力特征 $S \in \mathbf{R}^{N \times N}$ 与特征 D 进行关系矩阵计算,得到新空间注意力 $T \in \mathbf{R}^{C \times N}$,最后通过对空间注意力 T 与初始特征图 A 进行点积运算,完成空间维度上的新特征 E 标定。综上,空间注意力模块表达式为

$$E_j = \alpha \sum_{i=1}^N \left(\frac{\exp(B_i C_i)}{\sum_{i=1}^N \exp(B_i C_i)} \right) + A_j \quad (2)$$

分支,一个是基于非局部 (Non-local) 思想的空间注意力模块^[21-22],另一个是基于压缩激发 (Squeeze-and-excitation, SENet) 思想的通道注意力模块^[23]。将编码器输出特征图 F 依次通过通道注意力模块和空间注意力模块,分别获取在通道维度上具有更强指向性和注意力信息的中间特征 F_m ,以及融合空间和通道注意力信息的复合特征 F' 。

式中 i, j ——特征图中像素点位置

N ——像素点总数

α ——尺度系数,且初始值为 0,在训练过程中 α 将被逐步分配到更大权重

2.3 边缘感知的分割模块

参考文献[25]将边缘提取和语义分割 2 个任务同时训练有助于提高分割精度。为充分利用不同层的语义信息,在解码器部分每层上采样特征后都加入边缘感知的分割模块 (Edge perception based segmentation module, EPSM)。图 5 给出了边缘感知模块 EPSM 的整体结构。

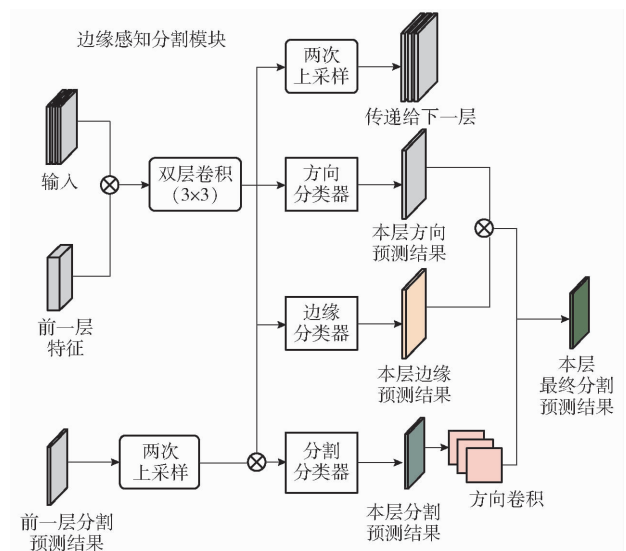


图 5 边缘感知分割模块结构图

Fig. 5 Architecture of edge perception based segmentation module

每个 EPSM 模块输入包括:上一层输出特征、上一层分割结果以及同尺度编码器跳跃连接的特征。

将同尺度编码器特征与上层的输出特征沿通道进行融合,并经过 2 个卷积模块(步长为 1,卷积核为 3×3 卷积层、批归一化层和整流线性单元)得到融合特征。

融合特征经过 2 次上采样操作,作为下一层的特征输入。另外,将融合特征分别送入边缘预测分类器、分割预测分类器和方向预测分类器,得到边缘预测 $\hat{E}$ 、分割预测 $\hat{S}$ 及方向预测 $\hat{D}$ 。每个分类器由 1×1 卷积层、归一化层、整流线性单元 3 部分组成。将融合特征与上层分割结果作为分割分类器的输入,达到同时借助低分率分割结果、同尺度编码特征及深度特征实现高分辨率的分割。

方向分类器为每个边界像素点计算其到所属对象中心的方向。以 45° 为间隔划分 8 个方向,分别是上、右上、右、右下、下、左下、左、左上 8 个方向。采用包含 8 个 3×3 卷积核的卷积层来自动细化分割预测权重,而这个卷积层中每个卷积核初始化为一个元素为 1,其余元素为 0。例如,给定一个方向预测为左上角的边界像素,目标是将该像素的方向预测构建置信度值替换为其左上角的置信度值。由边缘分类器结果计算二进制边缘预测

$$\hat{E}_p = \text{Sigmoid}(\hat{E}) > 0.5 \tag{3}$$

使用 Softmax 应用于方向预测

$$\hat{D}_p^k = e^k / \sum_{i=1}^8 e^i \tag{4}$$

其中 $\hat{D}_p^k$ 表示像素 k 的方向预测, i 表示上、右上、右、右下、下、左下、左、左上 8 个方向。

利用生成的方向图,通过从边界像素沿该方向移动一定距离来识别相应的内部像素并进行替换生成新的边缘图。将新生成的边缘图用于校正分割预测结果的边缘,生成每层的最终分割预测结果。

2.4 网络损失函数

模型训练需同时最小化田埂像素分类损失、边缘感知损失、边缘像素方向损失,因此本模型损失函数定义为

$$L = \sum_{j=1}^4 \text{CE}(S_j^{\text{final}}, I_j(S)) + \gamma \sum_{j=1}^4 \{ \text{CE}(E_j, I_j(E)) + \text{CE}(D_j, I_j(D)) \} \tag{5}$$

由于解码器部分的每个 EPSM 模块分割图、方向图和边缘图与原始输入图的分辨率都不一致,需要采用双线性插值函数 $I_j(*)$,使得原始图像分辨率与 EPSM 的输出分辨率一致。其中 CE 表示交叉熵损失函数, $I_j(S)$ 、 $I_j(E)$ 、 $I_j(D)$ 分别表示经过下采样的真实的像素类别、边缘像素和边缘方向。 S_j^{final} 、

E_j 、 D_j 分别表示每层 EPSM 模块预测的像素类别、边缘像素和边缘方向。考虑到边缘区域只占图像的小部分,设置 $\gamma = 1$ 平衡非边缘和边缘区域之间的损失。

2.5 模型训练

2.5.1 实验配置

实验基于 Linux Ubuntu 18.04 操作系统使用 Python 3.7 和 Pytorch 深度学习框架来实现,实验配备 16 GB 内存和 A100-SXM4 显卡。

2.5.2 模型训练及实验参数

卷积层选用 Kaiming^[26] 均匀分布初始化方法,同时固定初始权值,便于后期对比实验。模型学习率(Learning rate)设置为 0.001,训练轮数(Epoch)为 45 次,迭代次数(Iteration)共计 111 510 次,批处理量(Batch size)设置为 3,使用 Adam 优化模型。训练时损失值变化曲线如图 6 所示,本文模型训练损失值为 0.004, U-Net 模型训练损失值为 0.006。本文模型在收敛速度以及最小收敛值上都有一定的改进。

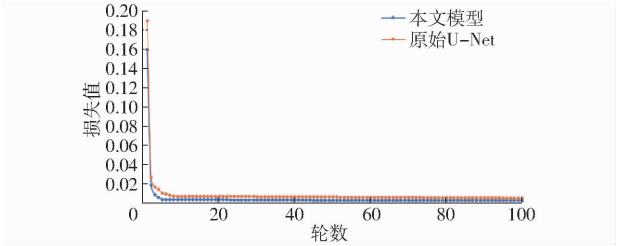


图 6 训练过程中损失值变化曲线

Fig.6 Loss changes during training

3 实验结果与分析

基于无人机采集的近地麦田数据集进行实验,在保持性能的同时降低对数据集的需求量。为验证本文模型的优势,分别将模型中的注意力模块和损失函数模块同现有方法进行比对,并通过指标和局部可视化的分割结果加以展示。此外,针对形态学方法的优势,也将通过与几种分割网络比对的形式加以体现。

3.1 分割评价指标

选取平均交并比(Mean intersection over union, MIoU)、平均像素准确率(Mean pixel accuracy, MPA)、平均 Dice 相似系数(Mean Dice similariy coefficient, MDSC)、平均召回率(Mean recall, MR)作为评价指标。平均像素准确率为预测正确像素与总像素比率均值,平均交并比为预测值与真实值的交并集比率均值,平均 Dice 相似系数为预测值与真实值的 Dice 相似度均值,平均召回率为样本正例被预测正确的比率均值。

3.2 注意力模块性能对比

以 1 456 幅带标注图像作为训练集,以 502 幅图像作为测试集,针对注意力模块进行评估,并以测试集均值作为最终结果。表 1 为各注意力算法下的分割精度对比结果,由表 1 可知,原始 U - Net 模型加入 MIAM 模块后,取得了 93.85% 的平均像素准确率以及 74.12% 的平均交并比,相较原始的 U - Net 以及附加 DA、CBAM 和 CA 注意力模块后的分割网络均有明显提高,平均像素准确率分别提升 2.84、5.36、2.23、0.88 个百分点,而平均交并比分别提升 9.26、14.8、7.17、3.99 个百分点。同时,与其它需每层加入注意力模块的网络比较而言,本文所提出的注意力模块能够实现仅在编码器的第 3、4 层添加,就可得到高于原始精度的结果,这也正是本文方法的优势所在。

图 7 展示了麦田图像在 U - Net、U - Net + CBAM、U - Net + CA、U - Net + DA 以及本文模型这

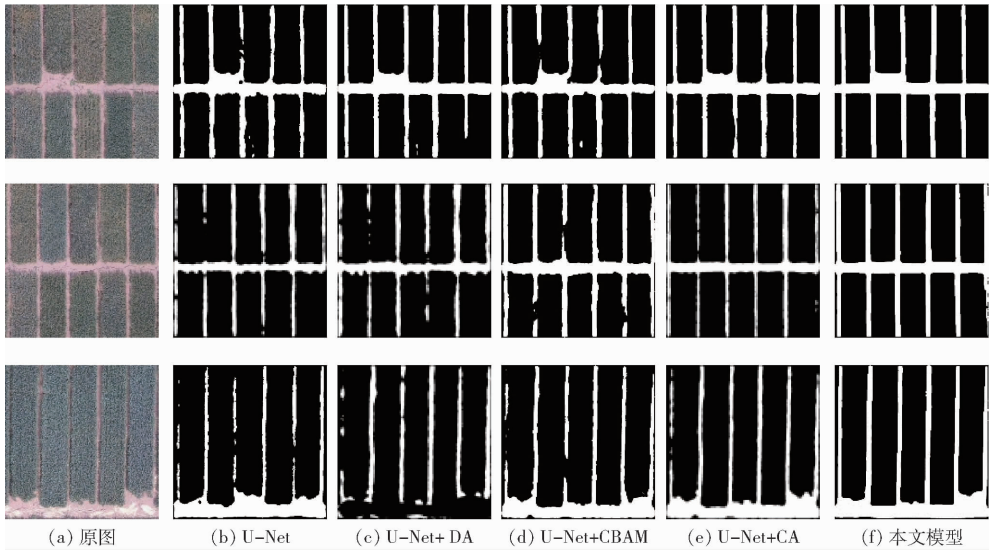


图 7 不同 Attention 模块局部分割图

Fig.7 Local segmentation diagrams of different Attention module

3.3 边缘感知模块性能对比

选取 1 456 幅带标注图像作为训练集,502 幅图像作为测试集。针对 EPSM 模块进行评估,表 2 给出加入 EPSM 模块前后的分割精度对比。由表 2 可知,本文的分割网络加入了 EPSM 模块后,取得了 95.57% 的平均像素准确率以及 77.48% 的平均交并比,相比于之前,平均交并比提升 3.36 个百分点,平均像素准确率提升 1.72 个百分点,平均 Dice 相似系数提升 4.74 个百分点,平均召回率提升 0.69 个百分点。图 8 为局部分割结果,加入 EPSM 模块后,网络改善了边缘模糊与田埂断裂问题。

为更好地评估本文模型,在数据集和网络参数相同的条件下,本文选取了 PSPNet、SegNet、U - Net 3 个经典分割模型同本文模型进行比较和评估。实

表 1 附加不同注意力模块的网络分割结果

Tab.1 Network segmentation results for different attention modules				%
模型	平均召回率	平均 Dice 相似系数	平均像素准确率	平均交并比
U - Net	75.92	78.27	91.01	64.86
U - Net + DA	56.36	62.17	88.49	50.32
U - Net + CBAM	80.91	79.82	91.62	66.93
U - Net + CA	81.24	82.01	92.97	70.13
本文模型	89.48	82.79	93.85	74.12

5 个模型上的局部分割结果,可以看出,只使用原始 U - Net 所得的分割效果会出现较多的田埂缺失和断裂问题,而在加入 DA、CBAM、CA 和 MIAM 这些注意力模块后,在一定程度上解决了上述问题。其中,相较其余几种注意力,在加入本文的多信息注意力模块 (Multi-information attention module, MIAM) 后,能够获得最为理想的分割效果。

表 2 加入 EPSM 模块前后的网络分割结果

Tab.2 Network segmentation results before and after adding EPSM module				%
模型	平均召回率	平均 Dice 相似系数	平均像素准确率	平均交并比
U - Net	75.92	78.27	91.01	64.86
U - Net + MIAM	89.48	82.79	93.85	74.12
本文模型	90.17	87.53	95.57	77.48

验证明,本文所提出的模型平均像素准确率达 95.57%,较 PSPNet、SegNet 和 U - Net 均有较大程度提升,详细的对比结果如表 3 所示。选取表中前 3 个模型进行分析,尽管 PSPNet 拥有强大的捕获上下文信息能力,但是在处理本文这种高分辨率的遥感图像上,其效果却不如另外两者;虽然 SegNet 和

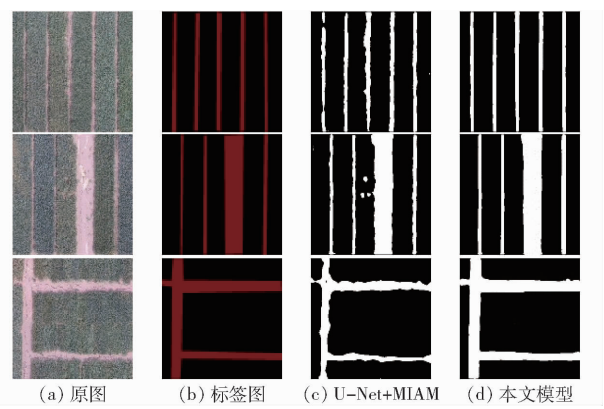


图 8 EPSM 模型效果对比

Fig. 8 EPSM model comparison rendering

U-Net 的分割结果相近,但在图像的细节和边缘信息处理上,U-Net 仍优于 SegNet,这也是本文选择 U-Net 作为基础网络的原因。而本文也通过加

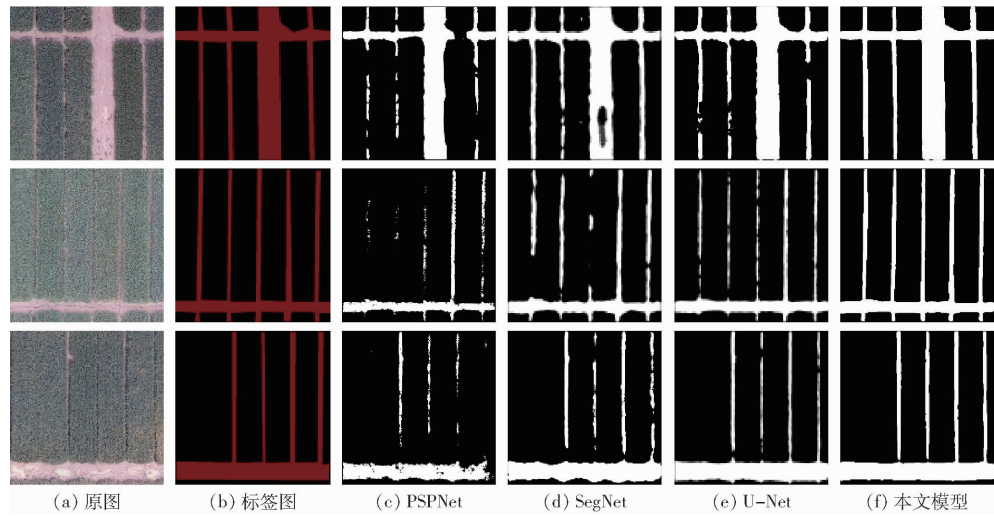


图 9 不同模型分割结果

Fig. 9 Segmentation diagrams of different models

为评估所提模型在不同光照强度下的分割效果,本文选择 PSPNet、SegNet、U-Net 与本文所提模型进行实验比较,在原始测试集和亮度分别降低 40%、50% 和 60% 的数据集上测试,分割准确率如表 4 所示。实验结果表明降低光照强度后,4 种模型的分割效果都有所下降,但本文模型分割准确率仍是最高且下降缓慢。

表 4 不同模型在各测试集上的分割准确率

Tab. 4 Comparison of segmentation performance on different brightness dataset

图像处理方法	PSPNet	SegNet	U-Net	本文模型
原图	83. 67	89. 83	91. 73	95. 57
降低 40% 亮度	62. 03	82. 16	85. 76	90. 16
降低 50% 亮度	53. 46	74. 58	82. 55	86. 93
降低 60% 亮度	47. 74	70. 52	79. 46	82. 15

对测试集内所有数据进行了分割,将所得结果通过裁剪和拼接的方式,还原了初始麦田图像的完

入注意力模块以及边缘感知模块的方式,优化了原始 U-Net 的分割结果,部分结果如图 9 所示。由实验结果可知,本文方案在复杂的大田背景下,对于识别高度相似的田埂特征、解决田埂区域占比少、田埂遮挡等问题均具有较强鲁棒性,同时也体现出模型在训练期间提取到了有用且丰富的特征信息,进而证明本文模型对田埂提取的有效性。

表 3 不同网络模型分割性能对比

Tab. 3 Comparison of classification performance of each network model

模型	平均召回率	平均 Dice 相似系数	平均像素准确率	平均交并比
PSPNet	28. 15	33. 61	83. 16	23. 27
SegNet	80. 15	82. 95	89. 73	69. 84
U-Net	75. 92	78. 27	91. 01	64. 86
本文模型	90. 17	87. 53	95. 57	77. 48

整分割结果图,如图 10 所示。

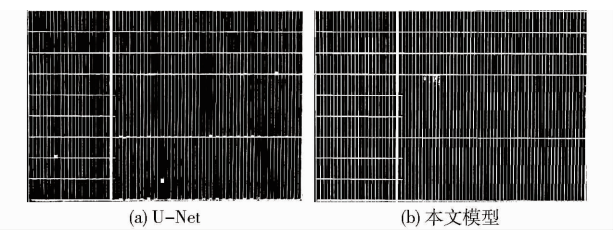


图 10 麦田分割结果

Fig. 10 Segmentation diagrams of wheat field dataset

4 结论

(1)针对无人机高分辨率麦田图像分割易受遮挡物干扰,田埂特征信息少、差异小,以及田埂区域分散且稀少的问题,提出了结合多信息注意力和区域结构损失函数的 U-Net 分割网络模型,并辅以形态学方法,最终实现提升麦田图像分割精度的效果。

(2)采用 U-Net 模型作为原始特征提取网络,

鉴于其关系建模能力好、分割精度高且适合处理大图的特点,为后续模型改进提供基础。

(3)加入注意力模块,分配特征学习过程中的关注点,加强高低阶特征的学习和融合,进一步提高目标特征捕获的准确性。

(4)在模型解码器加入边缘感知分割模块,使高级语义和低级空间细节都得到了保留。同时边缘

感知模块利用方向信息和边界位置引导边缘像素定位至正确位置,提升了边缘分割精度。

(5)实验表明本文模型对无人机高分辨率麦田图像的语义分割平均像素准确率达到 95.57%,平均交并比达到 77.48%,平均召回率达到 90.17%,平均 Dice 相似系数达到 87.53%,满足相关领域的需求。

参 考 文 献

[1] CAO W, QIAO Z, GAO Z, et al. Use of unmanned aerial vehicle imagery and a hybrid algorithm combining a watershed algorithm and adaptive threshold segmentation to extract wheat lodging[J]. Physics and Chemistry of the Earth Parts A/B/C, 2021(6):103016.

[2] 戴建国,薛金利,赵庆展,等.利用无人机可见光遥感影像提取棉花苗情信息[J].农业工程学报,2020,36(4):63-71.
DAI Jianguo, XUE Jinli, ZHAO Qingzhan, et al. Extraction of cotton seedling growth information using UAV visible light remote sensing images[J]. Transactions of the CSAE, 2020, 36(4): 63-71. (in Chinese)

[3] 周俊,王明军,邵乔林.农田图像绿色植物自适应分割方法[J].农业工程学报,2013,29(18):163-170.
ZHOU Jun, WANG Mingjun, SHAO Qiaolin. Adaptive segmentation of field image for green plants[J]. Transactions of the CSAE, 2013, 29(18): 163-170. (in Chinese)

[4] 张军国,冯文钊,胡春鹤,等.无人机航拍林业虫害图像分割复合梯度分水岭算法[J].农业工程学报,2017,33(14):93-99.
ZHANG Junguo, FENG Wenzhao, HU Chunhe, et al. Image segmentation method for forestry unmanned aerial vehicle pest monitoring based on composite gradient watershed algorithm[J]. Transactions of the CSAE, 2017, 33(14): 93-99. (in Chinese)

[5] 黄亮,吴春燕,李小祥,等.基于弱监督语义分割的灯盏花无人机遥感种植信息提取[J].农业机械学报,2022,53(4):157-163.
HUANG Liang, WU Chunyan, LI Xiaoxiang, et al. Extraction of *Erigeron breviscapus* planting information by unmanned aerial vehicle remote sensing based on weakly supervised semantic segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(4): 157-163. (in Chinese)

[6] 赵静,曹佃龙,兰玉彬,等.基于 FCN 的无人机玉米遥感图像垄中心线提取[J].农业工程学报,2021,37(9):72-80.
ZHAO Jing, CAO Dianlong, LAN Yubin, et al. Extraction of maize field ridge centerline based on FCN with UAV remote sensing images[J]. Transactions of the CSAE, 2021, 37(9): 72-80. (in Chinese)

[7] 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. Planting 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)

[8] SONG Z, ZHANG Z, YANG S, et al. Identifying sunflower lodging based on image fusion and deep semantic segmentation with UAV remote sensing imaging[J]. Computers and Electronics in Agriculture, 2020, 179:105812.

[9] 张永辉.基于语义分割的滨海湿地遥感图像信息提取[D].青岛:山东科技大学,2020.
ZHANG Yonghui. Information extraction of remote sensing image for coastal wetland based on semantic segmentation[D]. Qingdao: Shandong University of Science and Technology, 2020. (in Chinese)

[10] 苏健民,杨岚心,景维鹏.基于 U-Net 的高分辨率遥感图像语义分割方法[J].计算机工程与应用,2019,55(7):207-213.
SU Jianmin, YANG Lanxin, JING Weipeng. U-Net based semantic segmentation method for high resolution remote sensing image[J]. Computer Engineering and Applications, 2019, 55(7):207-213. (in Chinese)

[11] 王施云,杨帆.基于 U-Net 特征融合优化策略的遥感影像语义分割方法[J].计算机科学,2021,48(8):162-168.
WANG Shiyun, YANG Fan. Remote sensing images semantic segmentation method based on U-Net feature fusion optimization strategy[J]. Computer Science, 2021, 48(8):162-168. (in Chinese)

[12] 王斌,陈占龙,吴亮,等.兼顾连通性的 U-Net 网络高分辨率遥感影像道路提取[J].遥感学报,2020,24(12):1488-1499.
WANG Bin, CHEN Zhanlong, WU Liang, et al. Road extraction of high-resolution satellite remote sensing images in U-Net network with consideration of connectivity[J]. Journal of Remote Sensing, 2020, 24(12):1488-1499. (in Chinese)

[13] YUAN Yuhui, XIE Jingyi, CHEN Xilin, et al. SegFix: model-agnostic boundary refinement for segmentation[C]//European Conference on Computer Vision (ECCV), 2020:489-506.

[14] FREIXENET J, MUNOZ X, RABA D, et al. Yet another survey on image segmentation: region and boundary information integration[C]//European Conference on Computer Vision (ECCV), 2002:408-422.

[15] GUO Haonan, DU Bo, ZHANG Liangpei, et al. A coarse-to-fine boundary refinement network for building footprint extraction from remote sensing imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 183: 240-252.

[16] 陈燕生,赵丽娜,吴亚娟,等.改进 U-Net 的小宗作物遥感图像分割研究[J].科学技术创新,2021(14):11-14.

- [17] RUSSELL B C, TORRALBA A, MUROY K P, et al. LabelMe: a database and web-based tool for image annotation[J]. International Journal of Computer Vision, 2008,77(1-3):157-173.
- [18] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015:234-241.
- [19] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems, 2017:5998-6008.
- [20] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//European Conference on Computer Vision (ECCV), 2018:3-19.
- [21] WANG X, IRSCHICK R, GUPTA A, et al. Non-local neural networks[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2018:7794-7803.
- [22] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020:3141-3149.
- [23] JIE H, LI S, GANG S, et al. Squeeze-and-excitation networks[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018:7132-7141.
- [24] FFARAHNNK-GHAZANI F, BAGHSHAH M S. Multi-label classification with feature-aware implicit encoding and generalized cross-entropy loss[C]//24th Iranian Conference on Electrical Engineering (ICEE), 2016:1574-1579.
- [25] HEIDLER K, MOU L C, BAUMHOER C, et al. HED-UNet: a multi-scale framework for simultaneous segmentation and edge detection[C]//International Geoscience and Remote Sensing Symposium (IGARSS), 2021:3033-3040.
- [26] HE K, ZHANG X, REN S, et al. Delving deep into rectifiers: surpassing human-level performance on imagenet classification[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015:1026-1034.
- ~~~~~

(上接第 169 页)

- [19] BATRA A, SINGH S, PANG G, et al. Improved road connectivity by joint learning of orientation and segmentation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.
- [20] ZHANG X, HAN X, LI C, et al. Aerial image road extraction based on an improved generative adversarial network[J]. Remote Sensing, 2019,11(8):930.
- [21] DAI J, MA R, AI H. Semi-automatic extraction of rural roads from high-resolution remote sensing images based on a multifeature combination[J]. IEEE Geoscience and Remote Sensing Letters, 2022,19:1-5.
- [22] TAN H, SHEN Z, DAI J. Semi-automatic extraction of rural roads under the constraint of combined geometric and texture features[J]. ISPRS International Journal of Geo-Information, 2021,10(11):754.
- [23] 韩文霆, 张立元, 张海鑫, 等. 基于无人机遥感与面向对象法的田间渠系分布信息提取[J]. 农业机械学报, 2017, 48(3):205-214.
HAN Wenting, ZHANG Liyuan, ZHANG Haixin, et al. Extraction method of sublateral canal distribution information based on UAV remote sensing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017,48(3):205-214. (in Chinese)
- [24] 李前景, 刘珺, 米晓飞, 等. 面向对象与卷积神经网络模型的 GF-6 WFV 影像作物分类[J]. 遥感学报, 2021,25(2):549-558.
LI Qianjing, LIU Jun, MI Xiaofei, et al. Object-oriented crop classification for GF-6 WFV remote sensing images based on convolutional neural network[J]. National Remote Sensing Bulletin, 2021,25(2):549-558. (in Chinese)
- [25] 王斌, 陈占龙, 吴亮, 等. 兼顾连通性的 U-Net 网络高分辨率遥感影像道路提取[J]. 遥感学报, 2020,24(12):1488-1499.
WANG Bin, CHEN Zhanlong, WU Liang, et al. Road extraction of high-resolution satellite remote sensing images in U-Net network with consideration of connectivity[J]. Journal of Remote Sensing, 2020,24(12):1488-1499. (in Chinese)
- [26] 韩振浩, 李佳, 苑严伟, 等. 基于 U-Net 网络的果园视觉导航路径识别方法[J]. 农业机械学报, 2021,52(1):30-39.
HAN Zhenhao, LI Jia, YUAN Yanwei, et al. Path recognition of orchard visual navigation based on U-Net[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021,52(1):30-39. (in Chinese)
- [27] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[J]. arXiv:1502.03167, 2015.
- [28] SOBEL I. Neighborhood coding of binary images for fast contour following and general binary array processing[J]. Computer Graphics and Image Processing, 1978,8(1):127-135.
- [29] KINGMA D P, BA J. Adam: a method for stochastic optimization[C]//International Conference on Learning Representations (ICLR), 2015.
- [30] HEIPKE C, MAYER H, WIEDEMANN C. Evaluation of automatic road extraction[J]. International Archives of Photogrammetry and Remote Sensing, 1997,32(3):151-160.