

基于深度学习的无人机遥感小麦倒伏面积提取方法

申华磊¹ 苏歆琪^{1,2} 赵巧丽^{2,3} 周萌^{2,3} 刘栋^{1,4} 臧贺藏^{2,3}

(1. 河南师范大学计算机与信息工程学院, 新乡 453007; 2. 河南省农业科学院农业经济与信息研究所, 郑州 450002;
3. 农业农村部黄淮海智慧农业技术重点实验室, 郑州 450002;
4. 河南省教育人工智能与个性化学习重点实验室, 新乡 453007)

摘要: 为及时准确地提取小麦倒伏面积, 提出一种融合多尺度特征的倒伏面积分割模型 Attention_U²-Net。该模型以 U²-Net 为架构, 利用非局部注意力 (Non-local attention) 机制替换步长较大的空洞卷积, 扩大高层网络感受野, 提高不同尺寸地物识别准确率; 使用通道注意力机制改进级联方式提升模型精度; 构建多层级联合加权损失函数, 用于解决均衡难易度和正负样本不均衡问题。Attention_U²-Net 在自建数据集上采用裁剪方式提取小麦倒伏面积, 查准率为 86.53%, 召回率为 89.42%, F1 值为 87.95%。与 FastFCN、U-Net、U²-Net、FCN、SegNet、DeepLabv3 等模型相比, Attention_U²-Net 具有最高的 F1 值。通过与标注面积对比, Attention_U²-Net 使用裁剪方式提取面积与标注面积最为接近, 倒伏面积准确率可达 97.25%, 且误检面积最小。实验结果表明, Attention_U²-Net 对小麦倒伏面积提取具有较强的鲁棒性和准确率, 可为无人机遥感小麦受灾面积及评估损失提供参考。

关键词: 小麦; 无人机遥感; 倒伏面积提取; 深度学习; U²-Net

中图分类号: S126; S252 **文献标识码:** A **文章编号:** 1000-1298(2022)09-0252-09

OSID:



Extraction of Lodging Area of Wheat Varieties by Unmanned Aerial Vehicle Remote Sensing Based on Deep Learning

SHEN Hualei¹ SU Xinqi^{1,2} ZHAO Qiaoli^{2,3} ZHOU Meng^{2,3} LIU Dong^{1,4} ZANG Hecang^{2,3}

(1. College of Computer and Information Engineering, Henan Normal University, Xinxiang 453007, China
2. Institute of Agricultural Economy and Information, Henan Academy of Agricultural Sciences, Zhengzhou 450002, China
3. Huanghuaihai Key Laboratory of Intelligent Agricultural Technology, Ministry of Agriculture and Rural Affairs, Zhengzhou 450002, China
4. Henan Key Laboratory of Educational Artificial Intelligence and Personalized Learning, Xinxiang 453007, China)

Abstract: In order to extract the lodging area timely and accurately, a lodging area extraction model, namely Attention_U²-Net, was proposed. By integrating multi-scale features and based on U²-Net, Attention_U²-Net employed non-local attention mechanism to replace the hole convolution with large step size, expanded the receptive field of high-level network and improved the recognition accuracy of ground objects with different sizes, and utilized channel attention mechanism to improve the cascade mode and enhanced the accuracy. A multi-level joint weighted loss function was designed to balance the difficult and easy samples, and solve the challenge of imbalance between positive and negative samples. Patch-based pipelines were utilized to extract the lodging area. Experimental results on the self-built dataset showed effectiveness of Attention_U²-Net. The precision rate was 86.53%, the recall rate was 89.42%, and the F1 value was 87.95%, respectively. Compared with FastFCN, U-Net, U²-Net, FCN, SegNet and DeepLabv3, Attention_U²-Net achieved the highest F1 value and showed strong robustness and extraction accuracy. Compared with the labeled area, the extracted area obtained by Attention_U²-Net via cropping method was the closest one, and the accuracy rate of lodging area can reach 97.25%. Meanwhile, the false detection area of Attention_U²-Net was the smallest among all

收稿日期: 2022-04-16 修回日期: 2022-06-14

基金项目: 河南省科技攻关计划项目(212102110253、222102110244)、国家自然科学基金项目(62072160)、河南省农业科学院农业经济与信息研究所科技创新领军人才培养项目(2022KJCX02)和河南省农业科学院科技创新团队项目(2022TD14)

作者简介: 申华磊(1981—),男,副教授,博士,主要从事计算机视觉研究,E-mail: shenhualei@htu.edu.cn

通信作者: 臧贺藏(1983—),女,副研究员,博士,主要从事作物表型鉴定研究,E-mail: zanghechang@163.com

models. Experimental results showed that Attention_ U^2 -Net had strong robustness and high segmentation accuracy, which can be utilized as a valuable reference for UAV remote sensing of wheat affected area and loss assessment.

Key words: wheat; unmanned aerial vehicle remote sensing; lodging area extraction; deep learning; U^2 -Net

0 引言

小麦作为河南省主要粮食作物,连续 5 年播种面积稳定在 $5.67 \times 10^7 \text{ hm}^2$ 以上,占全国小麦种植总面积近 1/4,总产 $3.75 \times 10^{10} \text{ kg}^{[1-2]}$,肩负着我国粮食安全重任。倒伏是制约小麦品种产量的主要因素^[3],近年来由于台风天气偏多,暴风雨时有发生,对小麦产量影响极大,严重时减产达 50%^[4]。及时准确地提取小麦倒伏面积,可为灾后确定受灾面积及评估损失提供技术支撑^[5]。

目前,小麦倒伏面积的获取主要包括低通量的人工测量和高通量的遥感测量^[6-8]。人工测量法存在主观性强、随机性强、缺乏统一的标准,导致效率低下且费时费力,不能高效快速地提取倒伏面积。而遥感测量法是基于遥感影像中不同纹理^[9]、光谱反射率^[10]、颜色特征^[11]等进行特征融合,采用最大似然法对图像进行监督分类提取倒伏面积。随着深度学习在语义分割中的快速发展,国内外专家采用语义分割方法检测作物倒伏面积取得了突破性进展^[12-16]。这些研究主要采用遥感测量法进行特征分类,分割方法较为单一,未对不同特征筛选与分类方法进行组合优选,而深度学习方法存在无人机飞行高度较高的情况,只能实现粗略的倒伏区域分割。

深度学习的优势在于通过多层神经网络自动提取有效特征,这些特征不仅包括图像的局部细节特征,而且包括图像的高级语义特征。但由于计算量大、资源消耗的限制,特别是遥感高分辨率图像,内存约束要求必须对其进行下采样,或将其分割成多个块分别进行处理。然而,前一种方法会使图像失真,而后者则会由于缺乏全局信息造成误判。因此,本文移植并改进一种基于注意力机制的深层显著性网络 U^2 -Net^[17],对其进行轻量化,以对小麦倒伏面积进行信息提取和自动分割。同时,通过无人机拍摄图像并自建数据集,对该模型性能进行评价。

1 研究区概况与数据

1.1 研究区概况

研究区位于河南省农业科学院河南现代农业研究开发基地的小麦区域试验试验地,地处 $35^{\circ}0'44'' \text{ N}$, $113^{\circ}41'44'' \text{ E}$,如图 1 所示。气候类型属暖温带大陆

性季风气候,年平均气温为 14.4°C ,多年平均降水量为 549.9 mm ,全年日照时数 $2\,300 \sim 2\,600 \text{ h}$,冬小麦—夏玉米轮作为该地区的主要种植模式。

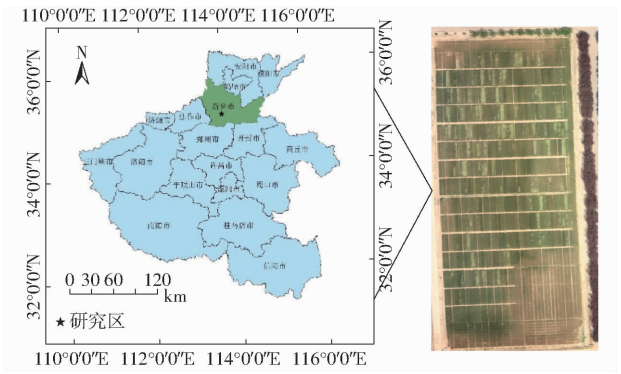


图 1 研究区位置
Fig. 1 Geographical location of study area

1.2 数据采集与预处理

根据国内外专家经验^[12-14,18]结合本研究实际情况,实验采用大疆精灵 4 Pro 型无人机,轴距 350 mm,相机像素为 2 000 万像素,影像传感器为 1 英寸 CMOS,镜头参数为 FOV 84° ,8.8 mm/24 mm(35 mm 格式等效),光圈 $f/2.8 \sim f/11$ 。搭载 GPS/GLONASS 双模定位,拍摄图像分辨率为 $5\,472 \text{ 像素} \times 3\,078 \text{ 像素}$,宽高比为 16:9。时间为 2020 年 5 月 14 日,此时研究区内小麦处于灌浆期。影像采集时间为 10:00,天气晴朗无云,垂直拍摄,飞行速度 3 m/s ,飞行时长 25 min,航向重叠度为 80%,旁向重叠度为 80%,相机拍照模式为等时间隔拍照,最终采集 700 幅原始图像。飞行采用大疆无人机自动规划的航线,共规划 5 条航线,航拍完成后采用自动返航的方式降落,如图 2 所示。

为实现小麦倒伏区域细粒度分割,使倒伏区域更加精确,本实验设置飞行高度为 30 m。低于 30 m,无人机可能与建筑碰撞,而高于 30 m,则无法获得较高分辨率图像。无论无人机飞行高度和天气条件等变量如何变化,只要在可控操作的环境下,通过合适的训练及参数调整,本模型技术均具有一定的有效性和准确性。图 3 为小麦倒伏图像 2 种分割策略,裁剪方法注重局部特征,下采样方法注重全局特征。

1.3 数据集构建和标注

实验使用原数据集 700 幅图像,通过对测试集



图 2 无人机自动规划航线图

Fig. 2 UAV automatic planning route map

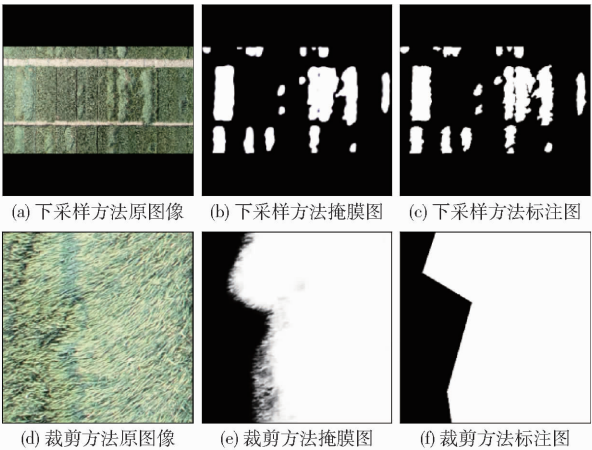


图 3 小麦倒伏图像分割策略

Fig. 3 Different methods for wheat lodging image segmentation

进行去重,训练集进行优选,最终筛选出 250 幅原始数据图像。深度学习通常需要大量数据,本实验采用的高通量数据分辨率为 5 472 像素×3 078 像素,深度学习通常使用数据分辨率仅为 512 像素×512 像素。如果使用滑动窗口进行裁剪,单幅图像可裁剪出 64 幅完全无重复图像,经过随机位置裁剪,单幅图像可得到 100 幅以上有效图像。250 幅原始图像经过数据处理后可得 25 000 幅有效图像,满足了深度学习的数据量要求。无人机飞行过程中,由于无人机拍摄角度和光影不同,不同航道拍摄相同位置图像会有差异,因此同样存在训练价值。为了均衡数据,选取第 1、2、3 号航线图像作为训练集,5 号航线图像作为测试集。本研究分下采样组和裁剪组,具体步骤为:

(1)筛选出无人机姿态平稳、拍摄清晰无遮挡数据,用于深度学习训练。

(2)人工标注:使用 Labelme 插件^[19]将小麦中度、重度倒伏区域标注为前景,其余区域标注为背景,并转换成二值图像作为训练集和测试集的标签。

(3)下采样组和裁剪组:下采样组将所有训练样本和测试样本等比例下采样至 342 像素×342 像素,之后通过背景填充将图像扩充至 512 像素×512 像素。裁剪组将测试样本裁剪为固定比例、边缘重叠和图像分辨率为 512 像素×512 像素,同时记录重叠区域的长和宽。

(4)数据增强:对下采样组训练样本进行无损变换,即水平或竖直随机旋转,以提高模型的鲁棒性。对裁剪组训练样本进行随机剪裁,剪裁区域尺寸为 512 像素×512 像素,以在每轮训练中生成不同的训练样本。

(5)图像拼接和恢复:将裁剪组掩膜图按记录的重叠区域长和宽进行合并,最终拼接成 5 472 像素×3 078 像素的分割结果图。将下采样组掩膜图裁剪为 342 像素×342 像素,并等比例放大复原。

(6)精度验证:对比分割结果 (Mask) 图和标注 (Ground truth) 图,计算模型指标。同时,通过地物关系与遥感图像映射,计算标注面积与分割面积,从而求出有效面积与准确率。

2 研究方法

2.1 U²-Net 模型

显著性目标检测^[20-21]主要用于人脸检测领域,通常旨在仅检测并分割场景中最显著的部分。中、重度小麦倒伏区域特征明显,U²-Net 是一种 2 层嵌套的 U 形结构的深度神经网络,用于显著性目标检测。该网络能够捕捉更多的上下文信息,并融合不同尺度的感受野特征,增加了网络深度,但没有显著提高计算代价。

具体而言,U²-Net 是一个 2 层嵌套的 U 型网络架构,其外层是由 11 个基本模块组成的 U 型结构,由 6 级编码器、5 级解码器和显著图融合模块组成,其中每个模块由一个基于残差的 U-Net 块填充。因此,嵌套的 U 型网络结构可以更有效提取每个模块内的多尺度特征和聚集阶段的多层次特征。

虽然原始的 U²-Net 已经具备优异的性能,但是为了对高通量小麦倒伏面积的特征特异性进行提取,对 U²-Net 做出进一步改进:引入通道注意力机制和一种 Non-local 注意力机制,构建一种新的小麦倒伏面积分割模型——Attention_U²-Net。该模型在进一步挖掘现有语义特征的同时,优化了网络结构。

2.2 基于注意力机制的 Attention_U²-Net 语义分割模型

如图 4 所示,Attention_U²-Net 由 2 层嵌套的 U 型结构组成。本文改进了 U²-Net 中的 RSU

(ReSidual U-blocks) 块, 使用了基于通道注意力机制的级联代替了 U^2 -Net 本身的级联, 在每个 Block 层使用 Non-local^[22] 机制代替 U^2 -Net 中的空洞卷积^[23], 并使用改进的 Multi focal loss 缓解训练样本难易程度不均和类别不平衡问题。

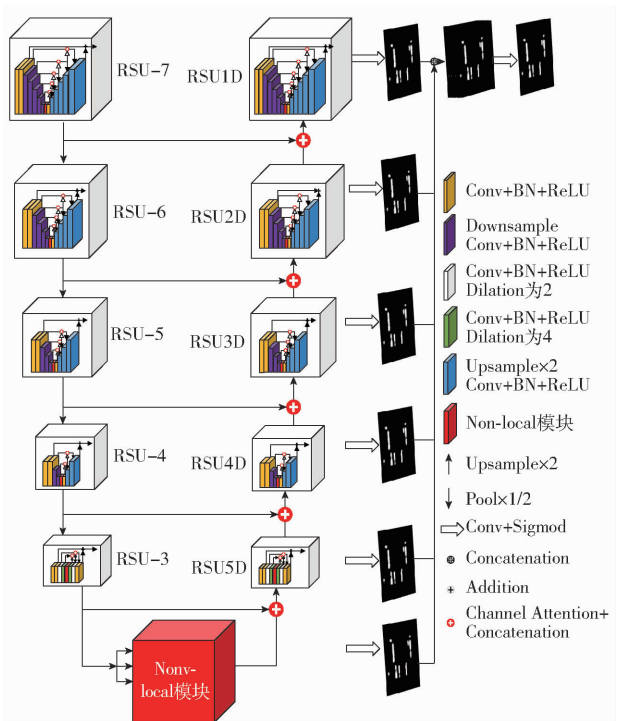


图4 Attention_ U^2 -Net 网络结构图

Fig. 4 Attention_ U^2 -Net architecture

U^2 -Net 使用了大量的空洞卷积, 在尽量不损失特征信息的前提下, 增加感受野。对于显著性目标需要大感受野, 而裁剪后数据语义混乱且倒伏面积随机。由于空洞卷积的卷积核不连续, 导致特征空间上下文信息可能丢失; 频繁使用大步长空洞卷积可能增加小麦倒伏区域边缘识别难度。同时, 空洞卷积使得卷积结果之间缺乏相关性, 从而产生局部信息丢失。

Non-local 机制 (图 5a) 是一种 Self-attention^[24] 机制, 原理式为

$$y_i = \frac{1}{C(x)} \sum_j f(x_i, x_j) g(x_j) \quad (1)$$

式中 x ——输入特征图

y_i ——特征图 i 对应位置的值

i ——输出位置的响应

j ——全局位置的响应

$f()$ ——计算特征图在 i 和 j 位置的相似度

$g()$ ——计算特征图在 j 位置的表示

$C()$ ——归一化函数, 保证变换前后信息不变

Non-local 可以通过计算任意 2 个位置之间的交互直接捕捉远程依赖, 而不用局限于相邻点, 相当于

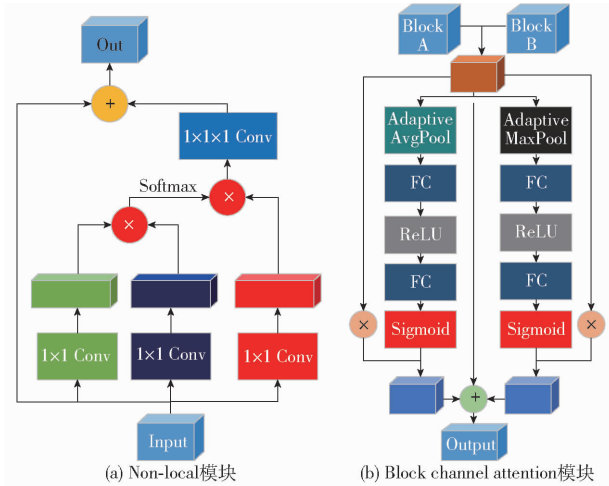


图5 分支结构图

Fig. 5 Branch structure diagrams

构造了一个和特征图谱尺寸一样大的卷积核, 从而保留更多信息。Attention_ U^2 -Net 在每个 RSU 块里保留了少量扩张率低的空洞卷积用于提取上下文信息特征, 广泛使用 Non-local 模块替换了扩张率大的空洞卷积, 同时 Non-local 模块也替代了整个 U^2 -Net 网络底层, 增强了网络模型的特征提取能力, 同时减少了计算量。

U^2 -Net 采用级联的方式将上采样 Block 和下采样 Block 结合, 产生多个通道, 通过 Block channel attention (图 5b) 使神经网络能够自动为融合后多个 Block 自动分配合适的权重。本文采用了全局平均池化和最大池化 2 种方式, 分别获取 Block 不同的语义特征, 并设计一个残差结构进行信息融合。

经过随机裁剪的样本, 可能存在样本难易度和类别分配不均衡问题, 从训练组数据每个航道随机抽取 144 幅裁剪后图像用于类别统计, 如表 1 所示。单幅图像倒伏面积大于 30% 样本约占总样本的 24%, 以致于大部分裁剪图像中无倒伏面积, 正负样本比例失衡, 样本难易度同样存在比例不均衡问题。由表 2 可以看出, 将倒伏面积小于 10% 的样本以及边缘特征不明显的样本定义为高难度样本, 其它倒伏样本定义为低难度样本。虽然高难度样本总占比约为 9.31%, 但倒伏样本中高难度样本占比高达 27.56%, 这并不意味着能够抛弃高难度样本而专注于提升低难度样本的分割准确率。本实验基于 U^2 -Net 的 Multi bce loss 和 Focal loss^[25] 提出了一种适用于小麦倒伏面积分割的损失函数: Multi focal loss, 计算式为

$$L = \sum_{m=1}^M w_s^{(m)} \xi_s^{(m)} + w_f \xi_f \quad (2)$$

式中 L ——Multi focal loss 损失函数值

M ——嵌套 U -Net 层数

m ——当前嵌套数
 $w_s^{(m)}$ ——第 m 层 loss 项对应权重
 $\xi_s^{(m)}$ ——第 m 层 loss 值
 ξ_f ——特征融合后多掩膜图的 loss 值
 w_f ——特征融合后 loss 项对应权重

对于每一项,使用 focal loss 来计算损失

$$\xi = -\alpha_i(1 - p_i)^{\gamma} \lg p_i \tag{3}$$

式中 p_i ——每个类别分类概率
 r ——样本难易程度加权值,用于控制难易程度不均衡
 α_i ——正负样本加权值,用于控制正负样本不均衡

表 1 随机抽样正负样本分布

Tab.1 Distribution of randomly selected positive and negative samples

航道	倒伏面积小于 30% 样本数	倒伏面积大于 等于 30% 样本数	大面积倒伏样本 占比/%
1	120	34	22.08
2	103	51	33.12
3	127	27	17.53
合计	350	112	24.24

表 2 随机抽样样本难易度分布

Tab.2 Statistics on difficulty of randomly selected samples

航道	难度高 样本数	难度低 样本数	完全无倒 伏样本数	高难度样本 总占比/%	高难度倒伏 样本占比/%
1	10	53	91	6.49	18.87
2	14	62	78	9.09	22.58
3	19	41	94	12.34	46.34
合计	43	156	263	9.31	27.56

使用 Focal loss 可以通过设置不同权重以抑制简单样本并解决正负样本比例严重失衡问题。Multi focal loss 降低了大量简单负样本在训练中所占权重,极大程度上抑制了裁剪带来的噪声;该损失函数控制了难易分类样本权重,并将每层掩膜图叠加,从而提高了模型鲁棒性,使其更适合用于小麦倒伏面积提取。

模型的输入图像分辨率为 512 像素 $\times$ 512 像素,输出为单通道掩膜图像。Attention_U²-Net 沿用了 U²-Net 的编、解码结构,由 6 层编码器、5 层解码器和掩膜图融合模块组成。在前 5 个编码阶段,Attention_U²-Net 同 U²-Net 将其分别标记为 RSU-7、RSU-6、RSU-5、RSU-4 和 RSU-3;其中“7”、“6”、“5”、“4”和“3”表示 RSU 块的高度 (H),对于高度和宽度较大的特征图,上层使用较大的 H 来捕获更大尺度的信息。RSU-4 和 RSU-3 中的特征

图的分辨率相对较小,进一步降低这些特征图的采样会导致裁剪区域上下文信息丢失。底层使用 Non-local 结构替换 U²-Net 大步长串联空洞卷积,降低了模型深度的同时,使其拥有更大的感受野能更好地识别边缘信息。在后 5 个解码阶段,Attention_U²-Net 使用线性插值进行上采样,解码模块同编码器结构保持一致,但对输入特征向量进行了处理,通过级联上一层特征与同一层相同分辨率特征,经过改进的通道注意力机制进行特征融合后输入上采样块,可以更有效地保证语义信息的完整性。

通过替换空洞卷积为 Non-local 结构,可以提升分割精度,但同样带来了巨大的参数量。Attention_U²-Net 只对大步长的空洞卷积进行了替换,在每个 RSU 块中,Attention_U²-Net 使用 Non-local 结构替换了大步长的空洞卷积,从而权衡模型速度和精度。掩膜图融合阶段,生成掩膜图概率映射,通过 3 $\times$ 3 卷积和线性插值生成每一阶段相同分辨率的掩膜图。将 6 个阶段的掩膜图并在一起,之后通过 1 $\times$ 1 卷积层和 Sigmoid 函数输出最终的掩膜图。

2.3 模型训练

实验选用 Intel(R) Core(TM) i7-10600 CPU,主频 2.90 GHz, GPU 选择 NVIDIA GeForce RTX3090,显存 24 GB,使用 PyTorch 作为深度学习框架。

实验将训练集和测试集分为多个批次,遍历所有批次后完成一次迭代。优化器选择 Adam,设置初始学习率为 0.001,随着迭代次数提升降低学习率至 0.000 1。

2.4 评价指标

采用查准率 (Precision)、召回率 (Recall)、F1 值 (F1-Score) 和 IoU (Intersection over Union) 指数评估模型性能,使用准确率量化倒伏面积提取能力。其中查准率指预测为倒伏面积占实际倒伏面积的比例;召回率表示预测倒伏面积占实际倒伏面积的比例。F1 值为查准率和召回率二者的调和均值;IoU 指数为倒伏面积预测面积和实际倒伏面积的重叠率;准确率指识别有效面积与提取总面积的比值。以上指标取值在 0 ~ 1 之间,值越大,表明评估效果越好。本文定义了一种用于量化倒伏面积准确率的公式

$$P_s = \frac{L_t + N_t}{L_t + N_t + L_f + N_f} \times 100\% \tag{4}$$

式中 L_t ——正确识别为倒伏小麦面积
 N_t ——正确识别非倒伏小麦面积
 L_f ——误把倒伏小麦识别为非倒伏小麦面积
 N_f ——未正确识别出倒伏小麦面积

P_s ——倒伏面积预测准确率

3 结果与分析

3.1 不同分割模型训练结果

基于测试样本数据,对比了 Attention_U²-Net、U²-Net 和主流模型 FastFCN^[26] (预训练网络 ResNet^[27])、U-Net^[28]、FCN^[29] (预训练网络 VGG^[30])、SegNet^[31]、DeepLabv3^[32] 的分割性能,图6为训练图像可视化进行了指数平滑。采用下采样所得样本训练神经网络收敛速度快、准确率高,而采用裁剪所得训练样本的训练收敛速度慢。由于正负样本不均衡导致裁剪后训练难度增大,部分模型的决策边界偏向数量多的负样本,使得准确率波动不明显并偏高。

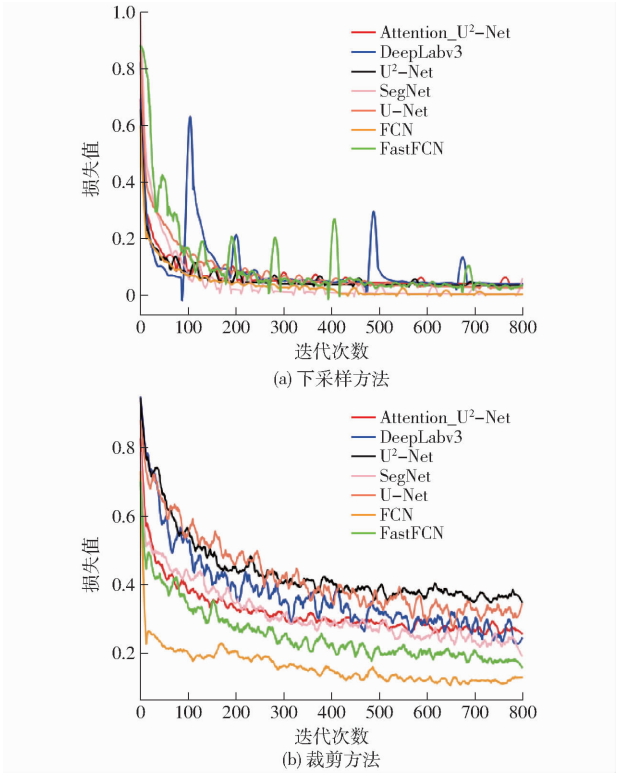


图6 训练损失
Fig.6 Train loss

3.2 不同分割模型性能对比分析

由表3可以看出,Attention_U²-Net 的分割效果最佳。U-Net、FCN、FastFCN、SegNet 等在下采样方法上性能相差较小,但在裁剪方法上多尺度适应性优势未能体现,识别准确率较低。U-Net、SegNet 等浅层网络对于裁剪出的 512 像素 × 512 像素掩膜图误检率较高,而 DeepLabv3 的整体分割效果较好。说明深层网络在下采样图像上的性能和浅层网络相似,浅层网络模型对解决许多简单并有良好约束的问题非常有效。深层网络训练速度慢,内存占用大,因此可以携带更多的数据,能够实现更复杂的数据

关系映射。由图7可以看出,对比下采样方法和裁剪方法,严重倒伏区域有显著的纹理和颜色特征,易于分割;小范围或轻微倒伏区域的纹理和颜色特征不明显,采用下采样后的分割效果较差。通过裁剪得到的边缘特征较为明显,能够识别难度较高样本,但模型收敛速度慢,算力需求高。

表3 不同分割模型在提取小麦倒伏面积时的评价指标
Tab.3 Evaluation indexes of wheat lodging area extraction by different segmentation models

模型	预训练模型	方法	评价指标/%			
			查准率	召回率	F1 值	IoU
FastFCN	ResNet	下采样	67.86	82.3	74.39	58.08
		裁剪	76.56	82.52	79.43	65.70
U-Net	无	下采样	72.07	87.21	78.92	65.09
		裁剪	73.57	79.00	76.19	60.72
FCN	VGG16	下采样	74.34	77.80	76.03	60.76
		裁剪	78.30	78.34	78.32	64.14
SegNet	无	下采样	74.43	73.99	74.21	58.37
		裁剪	78.29	78.13	78.21	63.67
DeepLabv3	ResNet	下采样	76.62	83.96	80.12	66.37
		裁剪	81.18	81.71	81.44	68.52
U ² -Net	无	下采样	79.53	85.24	82.29	69.59
		裁剪	86.46	82.24	84.30	72.93
Attention_U ² -Net	无	下采样	77.62	86.51	81.82	68.82
		裁剪	86.53	89.42	87.95	78.43

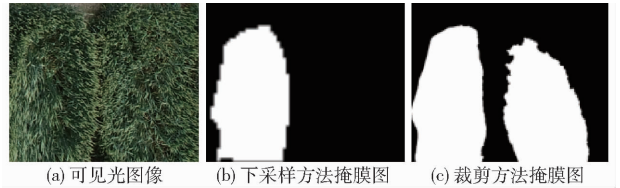


图7 下采样和裁剪方法对小范围倒伏区域分割效果

Fig.7 Segmentation of down sampling method and cropping method on small-scale lodging area

采用裁剪方法处理纹理细节时,深层网络能获得更充分的上下文信息,而浅层网络采用裁剪方法时,分割结果较差。采用下采样方法分辨率损失严重,甚至无法辨别轻、中度小麦倒伏面积。由于数据集中严重倒伏面积的占比较大,轻微倒伏面积占比较少,从而导致基于下采样方法的评价指标偏高。由于人工标注误差,整体准确率偏低。移植后 U²-Net 的整体性能略高于 DeepLabv3,同时处理裁剪图像时的性能较其它模型有较大提升。由于 Attention_U²-Net 基于裁剪方式进行改进,从而更关心局部特征,使用下采样方法处理数据不能很好地提取全局特征导致模型效能较差。使用裁剪方法时 Attention_U²-Net 的计算成本略高于原 U²-Net,但极大地增强了特征提取能力和泛化能力,F1 值提高 3.65 个百分点,识别效能有效提高。

3.3 不同模型面积提取效能分析

为了通过掩膜图像计算实际区域倒伏面积,以实地测量方式测得一个小区面积为 $8\text{ m} \times 1.5\text{ m}$,对应区域遥感图像像素数为 356 400 个。通过计算可得 29 700 个像素对应实际面积为 1 m^2 ,从而求出标注面积与提取面积。

如表 4 与表 5(结果保留两位小数)所示,为了对模型实际性能进行评估,对标记数据进行地物关系映射,测得标注倒伏面积为 0.40 hm^2 ,非倒伏面积为 3.0 hm^2 。非倒伏区域面积大、识别难度低、误检率低;倒

伏区域面积小,但部分倒伏区域识别难度大、误检率高。大部分模型使用裁剪方法提取倒伏面积效能较下采样方法有所提升。其中使用裁剪方法时 Attention_U²-Net 检测倒伏面积为 0.42 hm^2 ,准确面积为 0.37 hm^2 ;检测非倒伏面积为 2.98 hm^2 ,其中准确面积为 2.94 hm^2 , P_s 为 97.25%。Attention_U²-Net 提取倒伏区域有效面积最接近标注面积,拥有最高的 P_s 值,且误检面积较其它方法最低,能够检测出其它模型无法检测出的异常样本,体现出在复杂大田环境下准确判断倒伏区域的有效性,具有更高的实用价值。

表 4 裁剪方法各模型提取面积准确率对比
Tab.4 Comparison of area extraction accuracy of each model of cropping method

参数	FastFCN		U-Net		FCN		SegNet		DeepLabv3		U ² -Net		Attention_U ² -Net	
	非倒伏	倒伏	非倒伏	倒伏	非倒伏	倒伏	非倒伏	倒伏	非倒伏	倒伏	非倒伏	倒伏	非倒伏	倒伏
标注面积/ hm^2	3.00	0.40	3.00	0.40	3.00	0.40	3.00	0.40	3.00	0.40	3.00	0.40	3.00	0.40
像素数量	8.62×10^8	1.48×10^8	8.56×10^8	1.51×10^8	8.91×10^8	1.20×10^8	8.63×10^8	1.48×10^8	8.73×10^8	1.38×10^8	8.96×10^8	1.15×10^8	8.85×10^8	1.25×10^8
提取面积/ hm^2	2.90	0.50	2.89	0.51	3.00	0.40	2.91	0.50	2.94	0.46	3.02	0.39	2.98	0.42
准确面积/ hm^2	2.86	0.36	2.83	0.34	2.92	0.32	2.85	0.35	2.89	0.36	2.95	0.34	2.94	0.37
误检面积/ hm^2	0.04	0.14	0.06	0.17	0.08	0.08	0.05	0.14	0.05	0.11	0.07	0.05	0.04	0.06
$P_s/\%$	94.67		93.22		95.34		94.28		95.43		96.62		97.25	

表 5 下采样方法各模型提取面积准确率对比
Tab.5 Comparison of area extraction accuracy of each model of down sampling method

参数	FastFCN		U-Net		FCN		SegNet		DeepLabv3		U ² -Net		Attention_U ² -Net	
	非倒伏	倒伏	非倒伏	倒伏	非倒伏	倒伏	非倒伏	倒伏	非倒伏	倒伏	非倒伏	倒伏	非倒伏	倒伏
标注面积/ hm^2	3.00	0.40	3.00	0.40	3.00	0.40	3.00	0.40	3.00	0.40	3.00	0.40	3.00	0.40
像素数量	8.89×10^8	1.21×10^8	8.99×10^8	1.12×10^8	9.03×10^8	1.08×10^8	9.08×10^8	1.03×10^8	8.95×10^8	1.16×10^8	9.02×10^8	1.09×10^8	9.02×10^8	1.08×10^8
提取面积/ hm^2	2.99	0.41	3.03	0.38	3.04	0.36	3.06	0.35	3.01	0.39	3.04	0.37	3.04	0.36
准确面积/ hm^2	2.89	0.30	2.88	0.26	2.92	0.28	2.92	0.26	2.92	0.31	2.93	0.30	2.92	0.28
误检面积/ hm^2	0.11	0.11	0.14	0.12	0.12	0.08	0.14	0.08	0.10	0.08	0.11	0.07	0.12	0.08
$P_s/\%$	93.49		92.31		93.94		93.49		94.70		94.71		93.96	

3.4 不同分割模型定性比较

如图 8 和图 9 所示,预测图中白色区域为判断倒伏小麦的高权重区域,黑色为低权重区域。从图 8 可看出,U²-Net 和 Attention_U²-Net 可以更好地实现裁剪后小麦倒伏面积提取,其中 Attention_U²-Net 的验证结果更接近标注图,U-Net 和 SegNet 的验证结果较差;Attention_U²-Net、U²-Net 和浅层网络训练结果差距不大,但算力消耗较大。综合图 8、9 分割结果,采用下采样方法进行小麦倒伏面积分割结果较裁剪方法的误差大;而使用裁剪方法的训练难度高。

4 结论

(1)采用下采样和裁剪两种策略对无人机遥感小麦倒伏区域进行了分割。为了提高困难样本的检

测率,提出了一种深层神经网络 Attention_U²-Net。首先移植了 U²-Net 网络并使用改进注意力机制优化了级联模式,并使用 Non-local 替代了大步长的空洞卷积,使模型能从深层和浅层捕获更多的局部细节信息和全局语义信息;然后融合所有中间层的 Focal 损失,能在每层上更好地梳理样本分配不均和难易不平衡问题,进一步提高网络分割精度。

(2)基于无人机高通量倒伏区域识别方法精度高,能识别细微倒伏区域,移植后的网络通过采用裁剪方式,对小麦倒伏数据集的语义分割 F1 值为 84.30%。改进后的 Attention_U²-Net 分割小麦倒伏区域,其 F1 值可达 87.95%。为了对模型实际性能进行评估,本实验对倒伏区域进行人工标注并进行地物关系映射,测得标注倒伏面积为 0.40 hm^2 ,非倒伏面积为 3.0 hm^2 。Attention_U²-Net 检测倒伏面

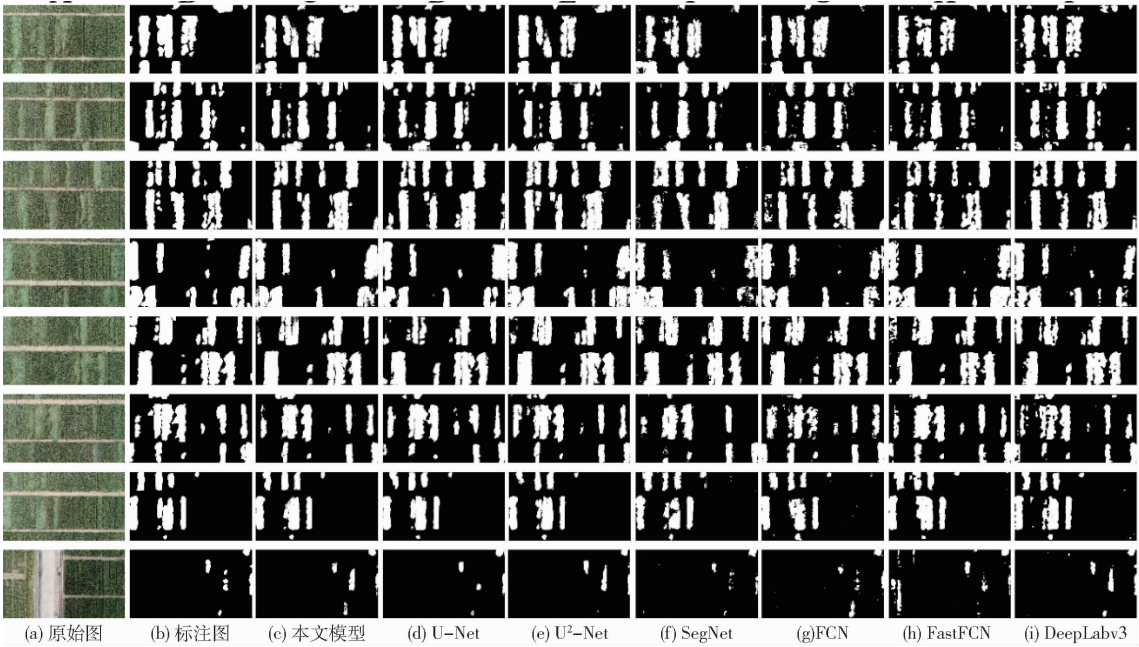


图 8 下采样方法实验结果定性比较

Fig. 8 Qualitative comparison of crop experimental results of down sampling method

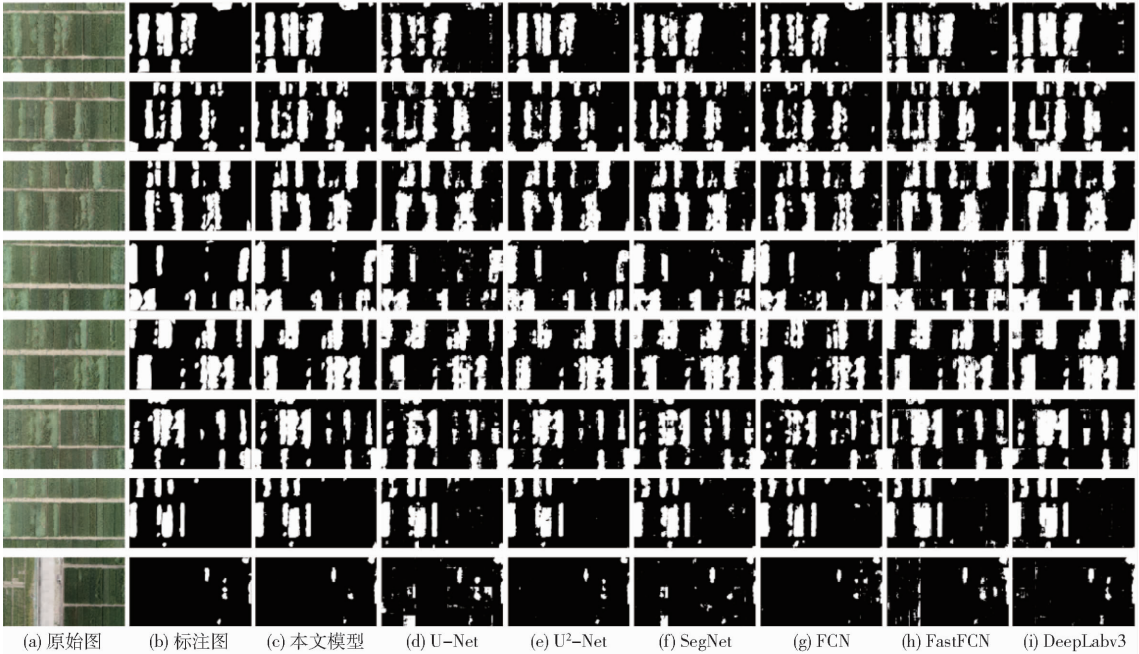


图 9 裁剪方法实验结果定性比较

Fig. 9 Qualitative comparison of crop experimental results of cropping method

积为 0.42 hm²,其中准确面积为 0.37 hm²;检测非倒伏面积为 2.98 hm²,其中准确面积为 2.94 hm², P_v 为 97.25%。通过与 FastFCN、U-Net、FCN、SegNet、DeepLabv3 主流神经网络模型对比, Attention_U²-Net 具有最高的准确率及 F1 值,表明本文模型在小麦倒伏区域检测应用中的准确性和有效性。

(3)实验结果表明,采用裁剪方法处理小麦倒伏数据,可能导致小麦倒伏区域的语义信息丢失,且

训练难度大;采用下采样方法通过浅层网络可以兼顾训练速度和训练效果,但只能适用于区域大、倒伏程度严重的情况,准确率较裁剪方法整体偏低。本文提出的 Attention_U²-Net 采用裁剪方法可以完成高难度训练任务且不显著占用计算资源,能够准确提取出小麦倒伏面积,可以满足麦田环境下的高通量作业需求,为后续确定受灾面积及评估损失提供技术支撑。

参 考 文 献

[1] 臧贺藏,曹廷杰,张杰,等.不同生态条件下小麦新品种产量的基因型与环境互作分析[J].华北农学报,2021,36(6):88-95.

- ZANG Hecang, CAO Tingjie, ZHANG Jie, et al. Genotype and environment interaction effect on yield of new wheat cultivars under different ecological conditions[J]. ACTA Agriculturae Boreali-sinica, 2021,36(6):88–95. (in Chinese)
- [2] 河南省统计局,国家统计局河南调查总队.河南统计年鉴[M].北京:中国统计出版社,2021.
- [3] 胡卫国,曹廷杰,杨剑,等.小麦新品种(系)抗倒性及产量构成因素评价[J].种子,2021,40(2):110–115.
- HU Weiguo, CAO Tingjie, YANG Jian, et al. Evaluation of lodging resistance and yield components of new wheat varieties (lines)[J]. Seed, 2021,40(2):110–115. (in Chinese)
- [4] 王芬娥,黄高宝,郭维俊,等.小麦茎秆力学性能与微观结构研究[J].农业机械学报,2009,40(5):92–95.
- WANG Fen'e, HUANG Gaobao, GUO Weijun, et al. Mechanical properties and micro-structure of wheat stems[J]. Transactions of the Chinese Society for Agricultural Machinery, 2009,40(5):92–95. (in Chinese)
- [5] CHAUHAN S, DARVISHZADEH R, BOSCHETTI M, et al. Remote sensing-based crop lodging assessment: current status and perspectives[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2019,151(5):124–140.
- [6] 刘建刚,赵春江,杨贵军,等.无人机遥感解析田间作物表型信息研究进展[J].农业工程学报,2016,32(24):98–106.
- LIU Jian'gang, ZHAO Chunjiang, YANG Guijun, et al. Review of field-based phenotyping by unmanned aerial vehicle remote sensing platform[J]. Transactions of the CSAE, 2016,32(24):98–106. (in Chinese)
- [7] 史舟,梁宗正,杨媛媛,等.农业遥感研究现状与展望[J].农业机械学报,2015,46(2):247–260.
- SHI Zhou, LIANG Zongzheng, YANG Yuanyuan, et al. Status and prospect of agricultural remote sensing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015,46(2):247–260. (in Chinese)
- [8] HEIN N T, CIAMPITTI L A, JAGADISH S V K. Bottlenecks and opportunities in field-based high-throughput phenotyping for heat and drought stress[J]. Journal of Experimental Botany, 2021,72(14):5102–5116.
- [9] 刘龙飞,陈云浩,李京.遥感影像纹理分析方法综述与展望[J].遥感技术与应用,2003,18(6):441–447.
- LIU Longfei, CHEN Yunhao, LI Jing. Texture analysis methods used in remote sensing images[J]. Remote Sensing Technology and Application, 2003,18(6):441–447. (in Chinese)
- [10] 冯书谊,张宁,沈霁,等.基于反射率特性的高光谱遥感图像云检测方法研究[J].中国光学,2015,8(2):198–204.
- FENG Shuyi, ZHANG Ning, SHEN Ji, et al. Method of cloud detection with hyperspectral remote sensing image based on the reflective characteristics[J]. Chinese Optics, 2015,8(2):198–204. (in Chinese)
- [11] 李宇昊,石田.利用航空数码影像的色彩特征估测森林郁闭度[J].西北林学院学报,2014,29(1):148–154.
- LI Yuhao, SHI Tian. Estimating the canopy density through color features of aerial digital images[J]. Journal of Northwest Forestry University, 2014,29(1):148–154. (in Chinese)
- [12] ZHAO X, YUAN Y, SONG M, et al. Use of unmanned aerial vehicle imagery and deep learning U-Net to extract rice lodging[J]. Sensors, 2019,19(18):3859.
- [13] ZHANG Z, FLORES P, IGATHINATHANE C, et al. Wheat lodging detection from UAS imagery using machine learning algorithms[J]. Remote Sensing, 2020,12(11):1838.
- [14] MDYA B, JGB C, HUI P, et al. Adaptive autonomous UAV scouting for rice lodging assessment using edge computing with deep learning EDANet[J]. Computers and Electronics in Agriculture, 2020,179(10):5817.
- [15] YANG M D, TSENG H H, HSU Y C, et al. Semantic segmentation using deep learning with vegetation indices for rice lodging identification in multi-date UAV visible images[J]. Remote Sensing, 2020,12(4):633.
- [16] MARDANISAMANI S, MALEKI F, KASSANI S H, et al. Crop lodging prediction from UAV-acquired images of wheat and canola using a DCNN augmented with handcrafted texture features[J]. arXiv preprint arXiv:1906.0071,2019.
- [17] QIN X, ZHANG Z, HUANG C, et al. U²-Net: going deeper with nested u-structure for salient object detection[J]. Pattern Recognition, 2020,106:107404.
- [18] 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.
- [19] RUSSELL B C, TORRALBA A, MURPHY 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.
- [20] 李岳云,许悦雷,马时平,等.深度卷积神经网络的显著性检测[J].中国图象图形学报,2016,21(1):53–59.
- LI Yueyun, XU Yuelei, MA Shiping, et al. Saliency detection based on deep convolutional neural network[J]. Journal of Image and Graphics, 2016,21(1):53–59. (in Chinese)
- [21] BORJI A, CHENG Mingming, HOU Qibin, et al. Salient object detection: a survey[J]. Computational Visual Media, 2019,5(2):117–150.
- [22] WANG X, GIRSHICK R, GUPTA A, et al. Non-local neural networks[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. United States:IEEE, 2018:7794–7803.
- [23] FISHER YU, KOLTUN V. Multi-scale context aggregation by dilated convolutions[J]. arXiv preprint arXiv:1511.07122, 2016.
- [24] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. arXiv preprint arXiv:1706.03762,2017.
- [25] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017,99:2999–3007.

[9] LIN D, ZHANG L, XIA X. Hierarchical model predictive control of Venlo-type greenhouse climate for improving energy efficiency and reducing operating cost[J]. Journal of Cleaner Production, 2020, 264:121513.

[10] CHEN Weihai, YOU Fengqi. Efficient greenhouse temperature control with data-driven robust model predictive[C] // Proceedings of the American Control Conference, 2020.

[11] 邓璐娟,张侃愉,龚幼民,等. 温室环境多级控制系统及优化目标值设定的初步研究[J]. 农业工程学报,2005,21(5): 119-122.

DENG Lujuan, ZHANG Kanyu, GONG Youmin, et al. Preliminary study on hierarchical greenhouse environment control system and setting of the optimized target values[J]. Transactions of the CSAE, 2005, 21(5): 119-122. (in Chinese)

[12] 戴剑锋,罗卫红,乔晓军,等. 基于模型的温室加温控制目标优化系统研究[J]. 农业工程学报,2006,22(11):187-191.

DAI Jianfeng, LUO Weihong, QIAO Xiaojun, et al. Model-based decision support system for greenhouse heating temperature set point optimization[J]. Transactions of the CASE, 2006, 22(11): 187-191. (in Chinese)

[13] XU Dan, DU Shangfeng, WILLIGENBURG G V. Double closed-loop optimal control of greenhouse cultivation[J]. Control Engineering Practice, 2019, 85: 90-99.

[14] XU Dan, DU Shangfeng, WILLIGENBURG G V. Optimal control of Chinese solar greenhouse cultivation[J]. Biosystems Engineering, 2018, 171(1):205-219.

[15] 李康吉,张世通,孟凡跃,等. 考虑时空变异特性的温室多环境因子优化策略[J]. 农业机械学报,2021,52(11):343-350.

LI Kangji, ZHANG Shitong, MENG Fanyue, et al. Optimization strategy of greenhouse multiple environmental factors considering temporal and spatiotemporal variability[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(11): 343-350. (in Chinese)

[16] 徐立鸿,徐赫,蔚瑞华. 温室蓝莓光温协调优化模型与控制策略研究[J]. 农业机械学报,2022,53(1):360-369.

XU Lihong, XU He, WEI Ruihua. Optimal model of blueberry greenhouse light and temperature coordination[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(1): 360-369. (in Chinese)

[17] 裴雪. 日光温室温度预测模型与卷帘智能控制方法研究[D]. 杨凌:西北农林科技大学,2018.

PEI Xue. Temperature prediction model of greenhouse and the design of intelligent control method of rolling quilt[D]. Yangling: Northwest A&F University, 2018. (in Chinese)

[18] 顾寄南,毛罕平,李萍萍. 温室系统综合动态模型的研究[J]. 农业工程学报,2001,17(4):79-82.

GU Ji'nan, MAO Hanping, LI Pingping. Comprehensive dynamic models of greenhouse system[J]. Transactions of the CSAE, 2001, 17(4): 79-82. (in Chinese)

[19] 孙润峰. 北方地区冬季日光温室黄瓜栽培生产技术[J]. 吉林蔬菜,2017(9):10-11.

SUN Runfeng. Cultivation and production techniques of cucumber in solar greenhouse in winter in northern China[J]. Jilin Vegetable, 2017(9): 10-11. (in Chinese)

[20] 王康丽,刘湘伟. 智能连栋玻璃温室番茄高效生产管理技术[J]. 农业工程技术,2020,40(34):12-14.

WANG Kangli, LIU Xiangwei. Efficient production management technology of tomato in intelligent multi span glass greenhouse[J]. Agriculture Engineering Technology, 2020, 40(34): 12-14. (in Chinese)

[21] 顾小小,朱春燕. 水培蔬菜连栋温室温湿度变化特征的研究[J]. 农业工程技术,2019,39(7):56-60.

GU Xiaoxiao, ZHU Chunyan. Study on temperature and humidity variation characteristics of hydroponic vegetable multi span greenhouse[J]. Agriculture Engineering Technology, 2019, 39(7): 56-60. (in Chinese)

(上接第260页)

[26] WU H, ZHANG J, HUANG K, et al. FastFCN: rethinking dilated convolution in the backbone for semantic segmentation[J]. arXiv preprint arXiv:1903.11816,2019.

[27] HE Kaiming, ZHANG Xiangyu, REN Shaoping, et al. Deep residual learning for image recognition[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.

[28] RONNEBERGER O, FISCHER P, BROX T. U Net: convolutional networks for biomedical image segmentation[J]. arXiv preprint arXiv:1505.04597,2015.

[29] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015,39(4):640-651.

[30] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv: 1409.1556,2014.

[31] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017,39(12):2481-2495.

[32] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[J]. arXiv preprint arXiv:1706.05587,2017.