

基于深度学习的青菜病害区域图像语义分割与定位

金 伦, 钱 莱 (浙江师范大学数学与计算机科学学院, 浙江金华 321000)

摘要 对青菜病害区域进行识别并将其与正常作物区分隔能够起到保护青菜的作用, 及时对相关灾害区域进行处理, 能够防止灾害的进一步蔓延。提出了一种基于深度学习的青菜灾害区域图像语义分割的方法, 通过 fine-tune FCN 以像素级精度分割出图像中作物灾害区进行识别, 并借助地面安置的图像定位标记判断出灾害在地面上的准确位置。由于目前暂无人机拍摄的公开青菜病害图像数据集, 通过专业无人机采集的方式自建青菜病害区域图像数据集以满足检测需要。结果表明, 在自建数据集下 fine-tune FCN 的 mIoU 为 53.2%, 对 4 种病害类型与健康情况的识别像素精度 PA 为 85.2%, 定位精确率为 96.8%, 能基本满足病害区域检测定位需求。同时, 对比了文章中提出的网络架构与 SDS, RCNN 与 FCN 在性能上的区别, 验证了该网络在病害区域的细粒度检测上拥有更好的鲁棒性。

关键词 青菜; 图像语义分割; 深度学习; 病害; 定位

中图分类号 S126

文献标识码 A

文章编号 0517-6611(2020)18-0235-04

doi: 10.3969/j.issn.0517-6611.2020.18.062

开放科学(资源服务)标识码(OSID):



Image Semantic Segmentation and Localization of *Brassica chinensis* Disease Area Based on Deep Learning

JIN Lun, QIAN Lai (College of Mathematics and Computer Science, Zhejiang Normal University, Jinhua, Zhejiang 321000)

Abstract Identifying *Brassica chinensis* disease areas and separating them from normal plant areas can play a role in protecting *B. chinensis*, and timely processing of related disaster areas can prevent further spread of the disaster. We proposed a method for semantic segmentation of *B. chinensis* disaster area images based on deep learning. The fine-tune FCN was used to segment the plant disaster area in the image with pixel-level accuracy for identification, and it was determined by the image positioning marks placed on the ground. As there was currently no public *B. chinensis* image data set taken by drones, a professional drone was used to collect self-built *B. chinensis* image data set to meet the detection needs. The results showed that in the self-built data set, the mIoU of fine-tune FCN was 53.2%, the pixel accuracy PA for identifying four types of diseases and health conditions was 85.2%, and the positioning accuracy was 96.8%, which could basically meet the detection and positioning of disease areas demand. Meanwhile, we compared the performance of the network architecture that we proposed with SDS, RCNN and FCN, and verified that it had better robustness in fine-grained detection of diseased areas.

Key words *Brassica chinensis*; Image semantic segmentation; Deep learning; Disease; Localization

青菜作为一种营养丰富的农作物, 能够提供多种人类生命活动所需的维生素与矿物质, 在农业的发展过程中起到了重要的作用。作为一种草本植物, 青菜在生产过程中容易受到不同种类病虫害的影响^[1]。当前研究的青菜病害识别方法主要是对植株本身进行识别, 从而进行植株病害的判断, 在实际的农业生产过程中, 用于识别病害的摄像头难以大量架设在农田中, 且固定架设在植物生长区域的光学镜片容易受到昆虫以及雨雪天气等情况影响, 导致破损或形变。

图像语义分割是指根据一定的相似性准则将图像划分成不同区域的过程, 是计算机视觉、图像处理等领域的基础性问题之一^[2]。近年来, 深度学习方法大大推动了计算机视觉领域的发展, 其中研究人员提出的 FCN^[3] 在 VOC2011 数据集上取得了 90.3% 的像素精度。笔者提出了一种基于 FCN 的对青菜种植区域航拍图像进行语义分割, 从而检测出病变区域的方法, 并在原网络的基础上进行 fine-tune, 改善网络效果; 借助无人机航拍获取青菜种植区域的俯视图并进行语义分割, 判断出病变区域的类别及其在图像中的相对位置; 最后使用 OCR^[4] 对病变区域图像进行扫描, 通过对地面各个固定位置事先安置好的不同字符标记进行识别并判断灾害区域所在位置的具体定位, 判断出灾害区在农田中的区域范围。

1 研究方法

1.1 图像获取

由于目前暂无通过旋翼无人机航拍获取的

公开青菜病害图像数据集, 该研究采用了自建数据集进行学习的方式^[5]。通过专业四旋翼无人机在浙江金华东阳地区黄田畈农田进行相关的作物图像采集^[6], 如图 1A 所示; 并在地面放置了字符标记, 如图 1B 所示。图像采集工作均通过四旋翼无人机在距离作物 2~3 m 的高度进行, 通过保证地面图像清晰度以更好地实现病害的识别与定位。最终在航拍图片中筛选出 600 幅作为训练与测试的基本数据, 其中包含了昆虫啃食、软腐病霜霉病和病毒病^[7-10] 的图像分别为 412、237、234 和 103 张, 分割后像素 RGB 值分别为 (255, 192, 203)、(255, 0, 255)、(75, 0, 130) 和 (255, 215, 0)。图像分辨率为 1 920×1 080, 筛选准则为图像中阴影区域较少^[11] 且图像中存在较明显的病害区域。为了扩充数据集, 采用旋转翻转等方式将数据集图像泛化为 3 660 张, 同时为了避免字符标记对病害识别训练的影响, 训练集中图像均为无字符标记农田图像。

1.2 图像语义分割

1.2.1 图像预处理。 为了使病变区域的色彩变化更加明显、色彩区域之间存在更加明显的梯度变化, 图像语义分割前按照图 2 通过直方图均衡化对图像进行处理, 提高图像对比度并增强了图像细节, 使得不同区域之间的色彩分割更加明显^[12]。

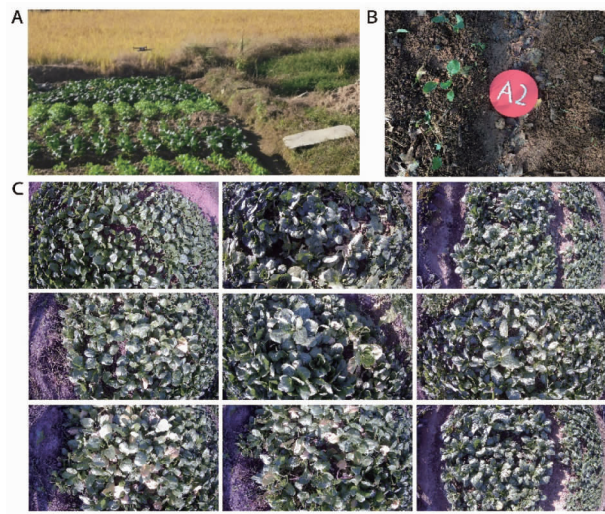
从图 2 可以看出, 在直方图均衡化处理过后, 图像的对比度有所提升, 病变区域与正常青菜区域间的色彩变化更明显。

1.2.2 构建网络。 该研究使用了 FCN 作为图像语义分割的

作者简介 金伦 (1997—), 男, 浙江温州人, 从事计算机视觉和人工智能研究。

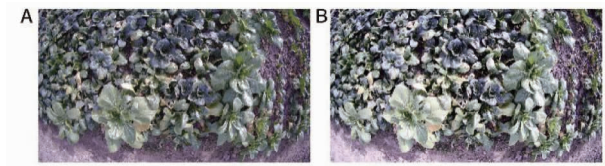
收稿日期 2020-03-12

基本网络结构,并在原网络基础上进行 fine-tune 使之更好地适应该数据集。修改后的网络通过将每个预测模块输出纬度中的最后一维调整为 7,分别对应着普通路面背景、泥土背景、绿色作物区以及 4 种不同的青菜病害分割区。FCN 实现了对作物区域像素级的分类,不同于传统的检测模型,不需要使用 Bounding-Box^[13] 来进行一个矩形区域的分割,对于无规则形状病害区域具有更好的检测效果。



注:A.无人机采集病害区域图像过程;B.地面的字符标记;C.收集到的部分航拍图像数据集
Note: A. The process of acquiring the image of disease area by UAV; B.Character mark on the ground;C.Part of the aerial image data set collected

图 1 作物图像采集
Fig.1 Crop image collection



注:A.农田采集原图像;B.直方图均衡化之后的农田图像
Note: A. Original image of farmland collected;B. farmland image after histogram equalization

图 2 农田图像
Fig.2 Farmland image

网络采用全卷积层加池化层的方式,将 CNN 中的全连接层替换成卷积层。网络的前 5 层保留 CNN 的结构,其中第 7 层采用了大小为(4 096,1,1)的卷积,为了使得在学习过程中能够提取更多的细节特征,在第 2 组卷积池化网络后又添加了 2 层卷积层,网络在最后 1 层使用了大小为(7,1,1)的卷积,分别对应着背景和作物区的 6 种识别结果,网络最终通过反卷积组合输出 1 张已经完成标记的图,其中 fine-tune FCN 的网络结构如图 3 所示。

同时,单层的输出结果会导致冗余过少和结果太过粗糙,为了优化语义分割结果,在网络结构中设计了 Skip Architecture^[14-15],对不同池化层的结果上采样,最终优化输出,

分别得到 fine-tune FCN-32s 与 fine-tune FCN-16s 为中间结果。图 4 分别给出了 Ground truth 与通过不同池化层向上反卷积后叠加最终得到的 fine-tune FCN-8s、fine-tune FCN-16s、fine-tune FCN-32s 的效果,最终输出效果为 fine-tune FCN-8s 显示效果。

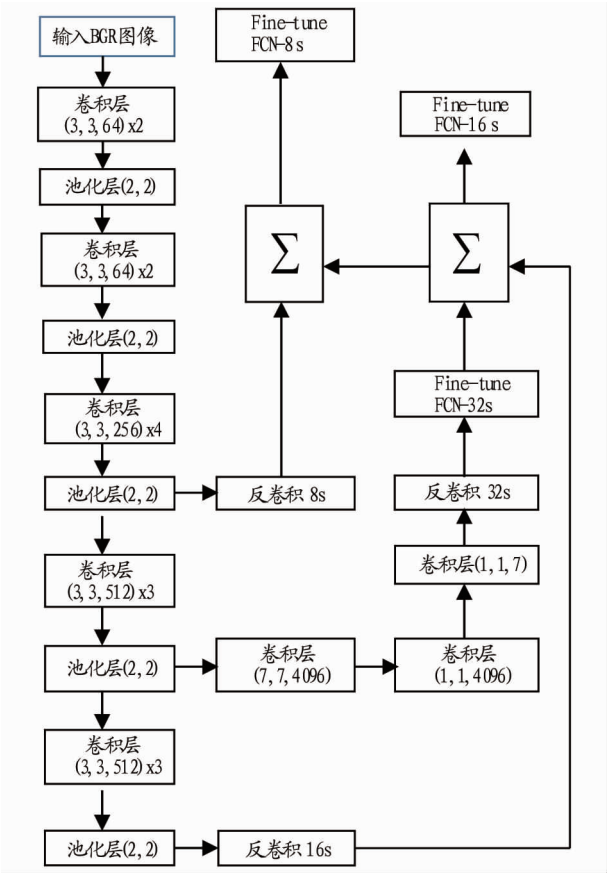
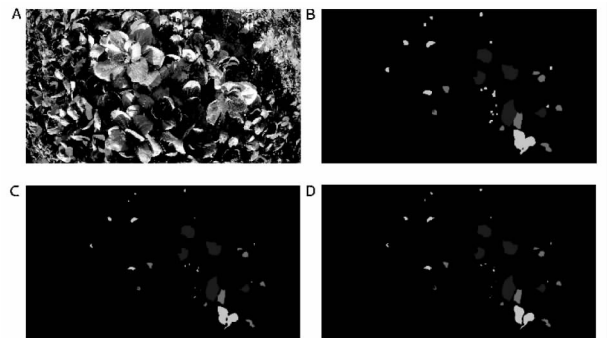


图 3 Fine-tune FCN 网络结构
Fig.3 The fine-tune FCN network structure



注:A.原图;B. fine-tune FCN-32s 效果;C. fine-tune FCN-16s 效果;D. fine-tune FCN-8s 效果
Note: A. Original drawing;B. fine-tune FCN-32s effect;C.C. fine-tune FCN-16s effect;D. fine-tune FCN-8s effect

图 4 不同池化层向上反卷积采样叠加后结果
Fig.4 Results of up deconvolution sample stacking at different pooling layers

1.2.3 损失函数与训练。在损失函数的设计上,采用 MSE 作为网络训练用的损失函数,每个像素点通过 softmax 输出维度为 7 的结果,将 1 张图片上的所有像素点分类误差的平

方累加,求得最终损失函数为

$$\text{loss}(W^*) = \sum_{c=1}^N \sum_{r=1}^M \sum_{i=1}^7 (\text{softmax}(\text{predict}_{c,r,i}) - \text{GrandTruth}_{c,r,i})^2 \quad (1)$$

在训练过程中,采用 SGD^[16] 为优化器进行损失函数的下降,设置训练过程中的 Batch Size 为 20,同时对比了 FCN-32s 第一轮迭代期间,初始学习率分别为 10^{-5} 、 10^{-6} 下的损失函数下降效果,结果如图 5。

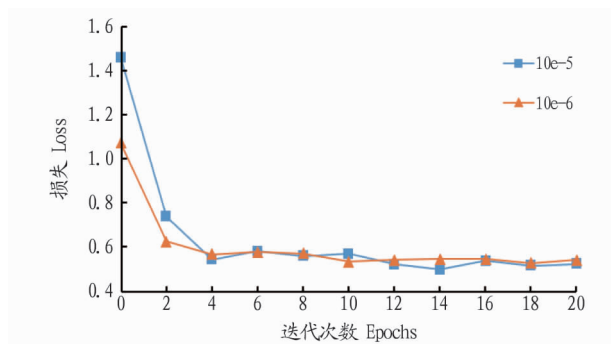


图 5 FCN-32s 不同初始学习率下的损失函数收敛曲线

Fig.5 Convergent curve of FCN-32s loss function under different initial learning rates

最终选择了 10^{-5} 为训练的初始学习率,并采用了 5^{-2} 的衰减率,以指数衰减的方式提高损失函数收敛精度,同时将泛化后的数据集进行划分,以 7:3 的比例划分训练集和测试集,其中训练集图像为 2 563 张,测试集中为 1 097 张。

在参数初始化的过程中,选择了直接迁移 FCN 作者训练的前 5 层参数以提取图像特征,并对其他层进行了归零初始化。在训练过程中发现,随机初始化并不能取得很好的收敛效果,同时在分类网络部分训练中采用了 0.5 的 dropout 以防止过拟合。

1.3 病害区域定位 为了确定病变区域的位置,首先需要确定标记所在位置,如果直接使用 Adaboost 来进行目标检测从而来确定目标的相对位置,则需要提高无人机高度,以保证地面定位目标在无人机视场角范围内,通过多个地面定位目标来判断当前病变区域所在位置。这种做法不但在地面定位目标安装的时候会有很高的精度要求,同时也拉高了无人机摄像头的拍摄高度,对作物区域进行视觉检测的时候会因为距离过远导致过多的图像细节遗漏。

为此,采取了一种基于 OCR 的病变区域定位方式。该研究使用的专业无人机航摄像头视场角为长边 120° ,为了适应更加普遍的情况,设该摄像头的视场角为 θ ,无人机距离地面高度为 $H \pm \delta$,为了使病变区域在单个目标标记范围内且目标标记区域不重合,在地面安置 2 个标记的时候,需要使得 2 个标记的距离 D 略大于目标方形区域范围(图 6),以实现在同一时刻场景中只存在 1 个定位目标。设在地面上标记的第 i 个定位目标为 a_i ,属于标记集合 A ,当摄像头范围内出现被 OCR 识别后属于集合 A 的光学字符,且该区域图像经语义分割后病害像素点个数大于 0 后,判定该标记对应的位置区间为病害区域。关系如下:

$$D \geq 2H \tan \theta + 2\delta \tan \theta \quad (2)$$

$$\begin{aligned} & \forall \text{argmax}(\text{softmax}(\text{predict}_{i,j})) \neq 1, 1 \leq i \leq N, 1 \leq j \leq M \\ & \wedge \exists x \in A \wedge x = a_i \Rightarrow \text{Result} = a_i \end{aligned} \quad (3)$$

式中, N 为图像的行数, M 为图像的列数, x 为图像范围内被识别出的字符,Result 为最终识别出病害所发生的区域。

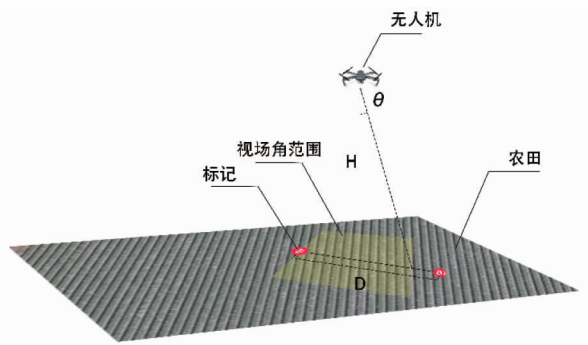


图 6 无人机探测病变区域的位置

Fig.6 Position of lesion area detected by UAV

2 结果与分析

2.1 网络效果验证 为了对比调整网络结构后的 fine-tune FCN 的效果,更好地评价算法的鲁棒性,同时对比了 FCN、SDS、R-CNN 与 fine-tune FCN 在自建数据集下的分割效果。为了保证对比结果的有效性,采用了 PASCAL VOC 计算机视觉竞赛数据集下 4 种进行语义分割和场景解析评估的常用指标来进行模型的评价,分别为像素精度(PA)、均像素精度(MA)、均交并比(MIoU)和频权交并比(FWIoU)^[17],试验结果表 1 所示。

表 1 不同图像语义分割模型在自建数据集下的性能对比

项目 Item	PA	MA	MIoU	FWIoU
SDS	76.2	52.8	32.7	70.1
R-CNN	78.1	64.3	37.2	71.5
FCN	84.6	71.4	50.6	79.4
fine-tune FCN	85.2	73.6	53.2	82.3

其中,

$$PA = \sum_i n_{ii} / \sum_i t_i \quad (4)$$

$$MA = (1 - n_{cl}) \sum_i n_{ii} / t_i \quad (5)$$

$$MIoU = (1 - n_{cl}) \sum_i n_{ii} / [t_i + \sum_j n_{ji} - n_{ii}] \quad (6)$$

$$FWIoU = [\sum_k t_k]^{-1} \sum_i t_i n_{ii} / (t_i + \sum_j n_{ji} - n_{ii}) \quad (7)$$

式中, n_{ji} 属于类 i 的像素被预测为类 j 的像素的个数, n_{cl} 为数据集中类的数量, t_i 为所有属于类 i 的像素个数之和。

2.2 定位结果准确度验证 该研究的定位方式属于区域定位,以命中率为定位结果准确度的评价指标,以通过图像语义分割与 OCR 字符判断计算得到的病害所在标记区域是否为该病害真实所在标记区域,从而获取相应评价关系(表 2)。

由于在实际生产中,仅凭借判断预测出的病害标记区域是否与真实标记区域一致,难以对治理进行协助,无法完全作为一个病害预测系统的评价标准,因此采用定位准确率

(Accuracy)和定位精确率(Precision)作为命中率的评价标准,并同时对比了FCN、SDS和R-CNN在相同工作下的效果,结果如表3所示。

表 2 定位命中率的评价指标

Table 2 Evaluation index of location hit ratio

项目 Item	Estimate: Within range	Estimate: Out of range
Actual: Within range	hit(TP)	missing(FP)
Actual: Out of range	missing(FN)	hit(TN)

表 3 不同图像语义分割模型的定位命中率比较

Table 3 Comparison of location hit ratios of different image semantic segmentation models %

项目 Item	准确率 Accuracy	定位精确率 Precision
SDS	87.1	84.8
R-CNN	88.1	87.9
FCN	91.8	92.8
fine-tune FCN	94.8	96.8

其中

Precision=TP/(TP+FP) (8)

Accuracy=(TP+TN)/(TP+TN+FP+FN) (9)

3 结论

该研究提出了一种基于深度学习进行青菜病害区域图像语义分割与定位的方法,能够通过 fine-tune FCN 对青菜农田图像进行图像语义分割,将病害区域从图像中分离出来,通过识别地面固定距离摆放的标记来进行区域识别,能够实现对病害区域的定位。同时,为了提高 FCN 在病害区域识别过程中的精确度,对 FCN 网络进行了 fine-tune,在第3层池化层前加入了2层3×3×256的卷积层,并将最后1层卷积核修改为1×1×7,从而更好地进行了抽象特征的提取。算法最终在通过专业无人机采集的自建数据集上进行验证,并与SDS、FCN、R-CNN进行对比,在自建数据集下 fine-tune FCN 的 mIoU 为 53.2%,对4种病害类型与健康情况的识别

像素精度 PA 为 85.2%,最终成功定位正确率为 96.8%,因此具有更好效果。综上所述, fine-tune FCN 在青菜病害区域图像语义分割上有较好的性能,结合地面标记识别能够有效实现青菜病害区域的识别与定位,协助农业生产以保证青菜的正常生长。

参考文献

[1] 郭萍.小青菜病虫害防治[J].北京农业,2012(6):104.
[2] 姜枫,顾庆,郝慧珍,等.基于内容的图像分割方法综述[J].软件学报,2017,28(1):160-183.
[3] LONG J,SHELHAMER E,DARRELL T.Fully convolutional networks for semantic segmentation[J].IEEE Transactions on Pattern Analysis & Machine Intelligence,2014,39(4):640-651.
[4] CHAUDHURI B B,PAL U.A complete printed Bangla OCR system[J].Pattern recognition,1998,31(5):531-549.
[5] 秦丰,刘东霞,孙炳达,等.基于深度学习和支持向量机的4种苜蓿叶部病害图像识别[J].中国农业大学学报,2017,22(7):123-133.
[6] SUN M X,YANG W C,HIGASHIYAMA T.Special issue on plant reproduction research in Asia[J].Plant reproduction,2018,31(1):1-2.
[7] 郝变青,马利平,朱九生,等.山西、河北蔬菜生产质量安全调查[J].山西农业科学,2013,41(10):1116-1118.
[8] 史婵,陈薇兰,王静,等.三份重要水稻资源的稻瘟病抗性鉴定及不同评价指标间的相关性分析[J].植物保护学报,2014,41(4):390-395.
[9] 宋凯,孙晓艳,纪建伟.基于支持向量机的玉米叶部病害识别[J].农业工程学报,2007,23(1):155-157.
[10] 田世平,范青.控制果蔬采后病害的生物学技术[J].植物学通报,2000,17(3):211-217.
[11] 张军国,韩欢庆,胡春鹤,等.基于无人机多光谱图像的云南松虫害区域识别方法[J].农业机械学报,2018,49(5):249-255.
[12] 王彦翔,张艳,杨成娅,等.基于深度学习的农作物病害图像识别技术进展[J].浙江农业学报,2019,31(4):669-676.
[13] 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(1/2/3):157-173.
[14] HARIHARAN B,ARBELÁEZ P,GIRSHICK R,et al.Simultaneous detection and segmentation[J].Lecture notes in computer science,2014,8695:297-312.
[15] LECUN Y,BENGIO Y.Convolutional networks for images, speech, and time series[M]//ARBIB M A.The handbook of brain theory and neural networks.Cambridge:MIT Press,1995:255-258.
[16] BOTTOU L.Large-scale machine learning with stochastic gradient descent[C]//LECHEVALLIER Y,SAPORTA G.Proceedings of COMPSTAT'2010.Heidelberg:Physica-Verlag HD,2010:177-186.
[17] GIRSHICK R,DONAHUE J,DARRELL T,et al.Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014IEEE conference on computer vision and pattern recognition.Ohio,USA:IEEE,2014:580-587.