

DOI: 10.13332/j.1000-1522.20180007

基于全卷积神经网络的植物叶片分割算法

胡 静^{1,2} 陈志泊¹ 杨 猛¹ 张荣国² 崔亚稷³

(1. 北京林业大学信息学院, 北京 100083; 2. 太原科技大学计算机科学与技术学院, 山西 太原 030024;

3. 北京邮电大学国际学院, 北京 100876)

摘要:【目的】植物叶片分割旨在从背景中分割出叶片区域,去除背景对象干扰。这对植物病害识别和物种鉴定具有重大意义。【方法】本文设计了基于全卷积神经网络的植物叶片分割算法。首先,目标函数用对数逻辑函数代替复杂的 Softmax 多类预测函数,从而将分割任务转化为适合于植物叶片分割的二分类问题;其次,把批归一化技术引入全卷积神经网络,从而改善网络整体的收敛性。最后,针对当前植物叶片分割研究中缺乏评估指标的状况,设计了新的评估协议——受试者工作特征曲线,该曲线反映了不同阈值情况下植物叶片图像分割的召回率与误报率之间的变化情况。【结果】本文提出的算法降低了全卷积神经网络的参数复杂度,改善了网络的收敛性。实验结果表明,该方法比 Leafsnap 提到的基于颜色的分割方法更完整地分割了植物叶片区域;提出的 ROC 曲线能够充分评估植物叶片的分割性能。【结论】与传统方法相比,基于深度学习的植物叶片分割方法实现了输入图像的端对端处理,无需图像转换、噪声滤波和形态运算等预处理技术,因此在植物叶片分割上具有可行性。

关键词:深度学习;全卷积神经网络;植物叶片分割

中图分类号:TP391.41 文献标志码:A 文章编号:1000-1522(2018)11-0131-06

引文格式:胡静,陈志泊,杨猛,等. 基于全卷积神经网络的植物叶片分割算法[J]. 北京林业大学学报,2018,40(11):131-136. Hu Jing, Chen Zhibo, Yang Meng, et al. Plant leaf segmentation method based on fully convolutional neural network[J]. Journal of Beijing Forestry University, 2018, 40(11): 131-136.

Plant leaf segmentation method based on fully convolutional neural network

Hu Jing^{1,2} Chen Zhibo¹ Yang Meng¹ Zhang Rongguo² Cui Yaji³

(1. School of Information Science & Technology, Beijing Forestry University, Beijing 100083, China;

2. College of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, Shanxi, China;

3. International School, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: [**Objective**] Plant leaf segmentation aims to segment leaf regions from backgrounds for removing background object interferes, which is important for plant disease recognition and species identification. [**Method**] In this paper, a fully convolutional neural network (FCNN) was designed for plant leaf image segmentation. First, a log-logic function as objective function replaces the complex Softmax function, which transforms the segmentation task into a binary classification problem suitable for plant leaf segmentation. Second, the batch normalization (BN) technology was introduced into the FCNN, which improved the convergence of the whole FCNN. Finally, due to the lack of evaluation index in the research of plant leaf segmentation, receiver operating characteristic (ROC) curve, as a new evaluation protocol, was designed. It reflected the changes between recall rate and false alarm rate of

收稿日期: 2018-01-04 修回日期: 2018-09-10

基金项目: 国家自然科学基金项目(61402038), 中央高校建设世界一流大学(学科)和特色发展引导专项资金项目。

第一作者: 胡静, 博士生, 副教授。主要研究方向: 图像处理与模式识别。Email: 279641292@qq.com 地址: 030006 山西省太原市万柏林区窊流路 66 号太原科技大学计算机学院。

责任编辑: 陈志泊, 教授, 博士生导师。主要研究方向: 数据库技术、计算机软件与理论。Email: zhibo@bjfu.edu.cn 地址: 100083 北京市海淀区清华东路 35 号北京林业大学信息学院。

本刊网址: <http://j.bjfu.edu.cn>; <http://journal.bjfu.edu.cn>

plant leaf segmentation under different threshold settings. [Result] The method reduced the complexity of parameters and improved the convergence performance of FCNN. Experimental results showed that this method was more complete to segment the leaf image than the color-based method in Leafsnap. The proposed ROC curve adequately evaluated the performance of plant leaf segmentation. [Conclusion] Compared with traditional plant leaf segmentation methods, the proposed method based on deep learning realizes input image by end to end processing, and does not require pre-processing like image conversion, noise filter and morphological operations, etc. Therefore, this method can be used for leaf segmentation.

Key words: deep learning; fully convolutional neural network(FCNN); plant leaf segmentation

在实践中,植物鉴定主要根据叶和花的形状来进行判断。植物学家普遍认为植物叶片特征的变化是作为植物病害识别研究的重要依据^[1-2]。此外,与花朵相比,叶子特征分析往往更容易,因为叶子通常在形状上几乎都是二维的,而花朵却常常是三维的。因此,基于计算机视觉的植物物种识别往往基于植物叶片图像进行,而且首先要从图像中分割出叶片区域。

植物叶片分割通常用于确定图像中的叶片区域和边界(见图1),以减少背景区域的干扰,这是植物种类识别和植物病害识别中的重要步骤^[3-5]。已经有许多植物图像分割方法被提出,它们大致可以分为基于非学习的分割方法和基于深度学习的分割方法两类。基于非学习的分割方法包括图像处理^[6]、图形剪切^[7]、归一化剪切^[8]、马尔可夫随机场^[9]、模糊理论^[10]、聚类^[11]等。对于基于深度学习的分割方法,主要有全卷积神经网络^[12-13]、反卷积网络^[5,14]和全连接条件随机场(Conditional random field, CRF)^[15]等。

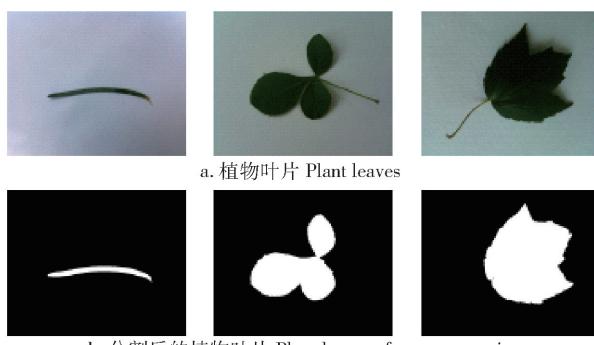


图1 分割前后的植物叶片图像

Fig. 1 Plant leaf images before and after segmentation

目前,大多数植物叶片分割方法是基于非学习的分割方法。例如,Kumar等^[3]通过期望最大化方法估计出前景和背景的颜色分布,并使用这些分布对每个像素进行独立分类。Valliammal等^[10]使用模糊阈值和聚类技术检测植物叶片图像中的均匀叶片区域。此外,Valliammal等^[11]还提出了基于离散小波变换(Discrete wavelet transform, DWT)和

K-means聚类的植物叶片图像分割方法。虽然这些传统的分割方法能够从彩色或灰度图像中分割出叶片区域,但它们对于去除围绕在叶片周围的一些干扰,如小气泡、小石块等,分割结果往往不尽如人意。另一方面,基于深度学习的图像分割方法取得了长足的进步,例如,全卷积神经网络^[12]的语义分割是第一个基于深度学习的通用分割模型,在PASCAL VOC 2011数据集中的分割效果令人鼓舞。在反卷积网络^[14]中,有多个反卷积层被学习得到,用以更好地对卷积神经网络(Convolutional neural network, CNN)学习特征映射图(Feature map, FM)进行上采样,获得了更精确的分割结果。在全连接的条件随机场^[15]中,通过将最后一个反卷积层的响应与全连接条件随机场进行结合来提高深度网络的前景分割能力。在这些进展的启发下,本文设计了基于深度学习的植物叶片分割方法,以提高植物叶片分割的准确率,同时还设计了ROC曲线评估协议用来评估植物叶片的分割性能,通过实验对比分析验证所提方法的优越性。

1 研究方法

1.1 叶片分割目标函数

全卷积神经网络(Fully convolutional neural network, FCNN)^[12]最初用于实现图像语义分割。图像语义分割是一个复杂的多分类问题,需要为隶属不同类别物体的像素配置不同的标签。为了实现这种精细的语义分割,全卷积神经网络采用的是 $(C+1)$ 类的Softmax函数,其中 C 是前景物体类别数量,多出一类是背景。基于Softmax这种需要强监督信息(物体类别标签)作为目标函数的网络模型,随着类别个数的增加,在实践中一般需要大规模训练数据集来驱动网络学习,否则容易出现模型过拟合现象。

植物叶片分割是典型的二分类任务,其核心目标是将植物图像分割成二值化图像,将前景像素和背景像素分别标记为1和-1。对于这种二分类任务,无需用到复杂的Softmax函数。因此,

本文用对数逻辑函数来代替复杂的 Softmax 多类预测函数作为分割目标函数,从而降低全卷积神经网络参数复杂度。基于对数逻辑函数的分割目标函数(L_s):

$$L_s = \frac{1}{K} \sum_{k=1}^K \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W \ln \left(1 + e^{-Y_{h,w}^k X_{h,w}^k} \right) \quad (1)$$

式中: $X \in R^{H \times W \times K}$ 是卷积神经层提供的预测分数; $Y \in R^{H \times W \times K} \in \{1, -1\}$ 是植物前景和背景标定的真实数据(Ground Truth),1 和 -1 分别表示前景和背景像素; H 、 W 和 K 分别表示高度、宽度和样本数; $h=1, 2, \dots, H; w=1, 2, \dots, W; k=1, 2, \dots, K$ 。

实际训练过程中,本文采用随机梯度下降算法^[16]对分割目标函数进行优化。为了使分割目标函数 L_s 最小化,梯度下降算法会自动调整,尽量加

大正样本的预测分数,同时减少负样本的预测分数。公式(1)的反向传播计算公式如下:

$$\frac{\partial L_s}{\partial X_{h,w}^k} = -\frac{1}{K} \frac{1}{HW} \frac{Y_{h,w}^k e^{-Y_{h,w}^k X_{h,w}^k}}{1 + e^{-Y_{h,w}^k X_{h,w}^k}} \quad (2)$$

1.2 卷积神经网络模块

全卷积神经网络由卷积神经网络模块和反卷积(Deconvolution)模块组成(图 2)。首先,借助卷积神经网络强大的特征学习能力和端到端的优势,卷积神经网络模块可以自动学习到有效的特征映射。其次,反卷积模块用于实现上采样和像素对齐,以确保分割后的二值图像和输入图像尺寸大小相等。再配合公式(1)中的目标函数,尽可能增大正样本的预测分数,同时减少负样本的预测分数,从而有利于实现植物叶片分割。

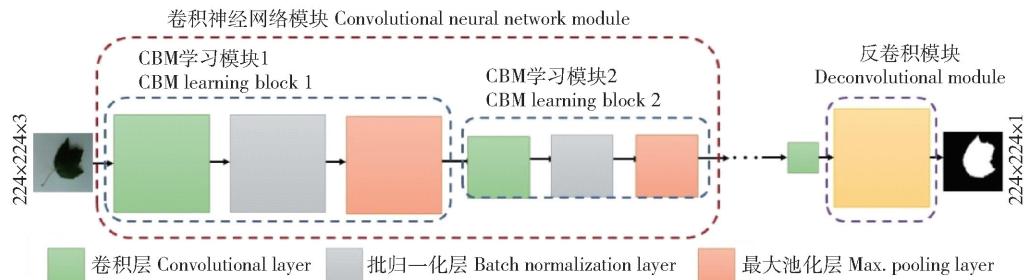


图 2 基于全卷积神经网络植物叶片图像分割方法框架

Fig. 2 Framework of plant leaf image segmentation method based on fully convolutional neural network

与文献[12]的全卷积神经网络相比,除了目标函数的设计不同之外,本文在卷积神经网络模块的设计上与文献[12]亦有不同之处。本文的卷积神经网络模块采用了批归一化(Batch normalization, BN)层^[17],以改善全卷积神经网络整体的收敛性。为了使描述简洁明确,将卷积层(Convolutional layer)、BN 层和最大池化层(Max pooling layer)合并简称为 CBM 学习块。所提出的卷积神经网络模块中,有 5 个 CBM 学习块(CBM learning block)。每个 CBM 学习块的 BN 层后使用 ReLU 作为神经元激活函数,并在卷积层和最大池化层中应用补零填充操作,从而保证预测分数和标签数据的对准更精确。在所提出的卷积神经网络模块中,除了 5 个 CBM 学习块之外,还应用 conv 6 卷积层来调整第 5 个 CBM 学习块输出特征图的维数。值得注意的是,conv 6 卷积层中采用 1×1 大小的滤波器,并且不使用 ReLU 神经激活函数,以避免丢失太多信息。

表 1 给出了设计的卷积神经网络模块的参数配置,包括卷积层的通道数、滤波器大小、池化窗口的大小等。

1.3 反卷积模块

利用卷积神经网络模块,将输入图像转换为

$7 \times 7 \times 1 \times 1024$ 维特征映射(Feature map),然后对反卷积模块进行上采样操作;将此特征映射进一步转化为与输入图像分辨率大小相等的分割预测分数,从而实现像素级别的植物叶片分割。

图 2 中的卷积神经网络模块,一共进行 5 次长宽各自降 2 采样,总体上相当于进行了长宽各自降 32 采样,因此必须设计一个升 32 插值的反卷积模块。为此,本文设计了插值滤波器 $F \in R^{64 \times 64}$ 。为了避免插值滤波器 $F \in R^{64 \times 64}$ 的过渡学习,又设计了正则化项来限制插值滤波器和双线性插值滤波器之间的差异。

$$L = L_s + \frac{1}{2} \lambda \| F - F_0 \|_2^2 \quad (3)$$

式中: L 为目标函数, $\lambda \geq 0$ 用于控制正则化项贡献的超参数, F 为插值滤波器, F_0 为双线性插值滤波器。

式(3)中双线性插值滤波器 F_0 的计算公式为:

$$F_0(i,j) = \left(1 - \frac{|i-32|}{32} \right) \left(1 - \frac{|j-32|}{32} \right) \quad (4)$$

式中: $i, j \in \{1, 2, \dots, 64\}$ 和 $F_0(i,j)$ 表示排列在第 i 行和第 j 列的元素。反卷积模块的反向传播过程中也应该考虑正则化项的作用,其计算公式如下:

表1 本文提出的卷积神经网络的参数配置

Tab. 1 Parameter configuration of the proposed fully convolutional neural network

CBM 学习块	名称	激活函数	滤波器($h \times w \times c \times g$)	步进	池化操作
CBM learning block	Name	Activiate function	Filter ($h \times w \times c \times g$)	Step	Pooling operation
模块 1 Block 1	卷积层 1 Convolutional layer 1 (conv 1)		$3 \times 3 \times 3 \times 64$	1	
	批归一化层 1 Batch normalization layer 1	ReLU		1	
	最大池化层 1 Max. pooling layer 1			2	3×3
模块 2 Block 2	卷积层 2 Convolutional layer 2 (conv 2)		$3 \times 3 \times 64 \times 128$	1	
	批归一化层 2 Batch normalization layer 2	ReLU		1	
	最大池化层 2 Max. pooling layer 2			2	3×3
模块 3 Block 3	卷积层 3 Convolutional layer 3(conv 3)		$3 \times 3 \times 128 \times 256$	1	
	批归一化层 3 Batch normalization layer 3	ReLU		1	
	最大池化层 3 Max. pooling layer 3			2	3×3
模块 4 Block 4	卷积层 4 Convolutional layer 4 (conv 4)		$3 \times 3 \times 256 \times 512$	1	
	批归一化层 4 Batch normalization layer 4	ReLU		1	
	最大池化层 4 Max. pooling layer 4			2	3×3
模块 5 Block 5	卷积层 5 Convolutional layer 5 (conv 5)		$3 \times 3 \times 512 \times 512$	1	
	批归一化层 5 Batch normalization layer 5	ReLU		1	
	最大池化层 5 Max. pooling layer 5			2	3×3
卷积层 6 Convolutional layer 6 (conv 6)				$1 \times 1 \times 512 \times 1024$	

注: h 、 w 、 c 、 g 分别表示高度、宽度、通道和组的大小。Notes: h , w , c , g indicate height, width, channel and group size, respectively.

$$\frac{\partial L}{\partial F} = -\frac{\partial L_s}{\partial F} - \lambda (F - F_0) \quad (5)$$

1.4 评估协议

针对植物叶片分割中正、负样本不对称分类问题,准确率不完全适合评估分割性能。为了能充分评估植物叶片分割的性能,本文提出受试者工作特征(Receiver Operating Characteristic, ROC)曲线。该曲线通过不同阈值情况下召回率(Recall Rate, RR)与误报率(False Alarm Rate, FAR)之间的变化评估分割性能,召回率和误报率的计算公式分别为:

$$RR = \frac{TP}{TP + FN} \quad (6)$$

$$FAR = \frac{FP}{FP + TN} \quad (7)$$

式中:TP 表示被正确检测出来的正样本数量,TN 表示被正确检测出来的负样本数量, FN 表示被错误检测出来的负样本的正样本数量, FP 表示被错误检测出来的正样本的负样本数量。

2 实验与分析

本文在Leafsnap 数据库^[3]上评估了基于全卷积神经网络(FCNN)的植物叶分割方法,并将其与已有的植物叶分割方法进行比较,验证所提出的方法的优越性。

2.1 Leafsnap 植物叶片图像数据库

Leafsnap 植物叶片图像数据库涵盖了美国东北部的 185 种树种(图 3),由移动设备在野外拍摄的 7 719 张图像组成。该数据库中的植物图像含有不同程度的模糊、视角变化和光照变化。

2.2 训练配置

Leafsnap 数据库中所有图像都被随机裁剪为 224×224 像素,并通过水平镜像和随机旋转进行数据增强。比如 $[-3^\circ, 0^\circ]$ 、 $[0^\circ, 3^\circ]$ 就是随机旋转图像。每层权重使用正态分布 $N(0, 0.01)$ 初始化,并将偏置项初始化为 0。式(3)中正则化权重设置为 0.5。动量(momentum)设置为 0.9。随机梯度下降算法中,小批(mini-batch)数据量是 128。随着训练的进行,学习率从 0.01 逐渐降低。如果目标函数在某个阶段收敛,则学习率降低到当前值的 1/10,最小小学习率为 0.0001。在实验中,图像随机地被分为训练集、验证集和测试集,3 个集合的比例为 0.6:0.2:0.2。

2.3 与现有植物叶片分割方法的比较

由于目前很多植物叶片分割算法都未提供源代码、可供复现的技术细节和可供比较的分割实验结果,尤其是定量的实验数据,所以给植物叶片分割算法性能的对比造成了较大的障碍。因此,本文在Leafsnap 数据库上,与文献[3]中提出的基于颜色植物叶片分割方法进行定性比较。从图 4 中可以看



此图源自文献[3] This image is derived from reference [3].

图3 来自Leafsnap数据库的所有184种树种的缩略图

Fig. 3 Thumbnail of all 184 tree species from the Leafsnap database

出,本文提出的分割方法比文献[3]中基于颜色的分割方法更能完整地分割植物叶片区域。

在实验中,由于负样本(背景像素)远大于正样本(植物叶片像素),植物叶片分割被视为不平衡的二进制分类问题,准确率不完全适合评估分割性能,因此本文提出ROC评估协议来评估分割性能。为了让未来的植物叶片分割算法能够更方便地与本文提出的分割算法进行性能比较,根据本文提出的

ROC评估协议给出了ROC曲线(图5)。从图5可以看出,一个阈值对应ROC曲线上一个点,一条ROC曲线代表不同阈值下召回率(Recall Rate, RR)与误报率(False Alarm Rate, FAR)的对应情况。

表2中列出本文算法在阈值默认为0的情况下误报率对应的召回率。例如,当误报率设置为1%时,对应的召回率为96.62%;当误报率设置为

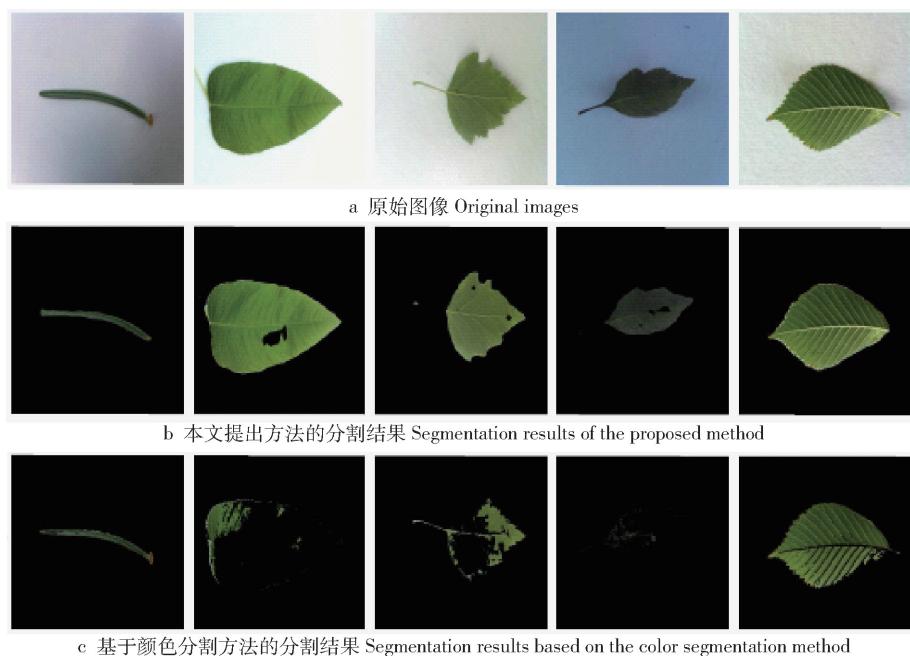


图4 基于全卷积神经网络的植物叶分割方法与基于颜色的分割方法的比较

Fig. 4 Comparison between plant leaf segmentation method based on fully convolutional neural network and that based on the color segmentation method

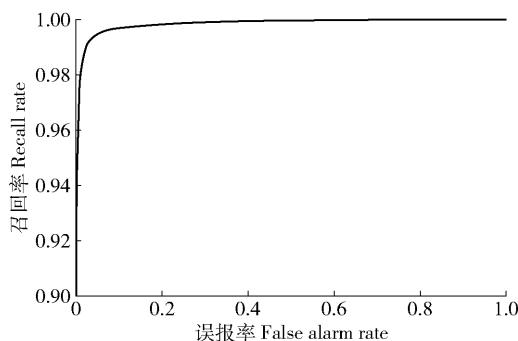


图5 基于全卷积神经网络的植物叶片图像分割方法所产生的ROC曲线

Fig. 5 ROC curve generated by plant leaf segmentation method based on fully convolutional neural network

0.67%时,对应的召回率为95.80%。由此可见,对于整个Leafsnap数据库中的叶片像素来说,本文算法基本都能较为完整地将叶片分割出来。

表2 阈值为0时误报率对应的召回率

Tab. 2 Recall rate corresponding to false alarm rate when threshold is 0

误报率 False alarm rate	0.1%	0.67%	1%	10%
召回率 Recall rate	80.45%	95.80%	96.62%	99.68%

3 结 论

本文设计了基于全卷积神经网络的植物叶片图像分割方法。首先,分割的目标函数用对数逻辑函数代替复杂的Softmax多类预测函数,从而将植物叶片分割任务转化为二分类问题,降低了全卷积神经网络参数复杂度。其次,把批归一化技术引入全卷积神经网络,改善网络整体的收敛性。实验结果表明,与现有的植物叶片分割方法相比,本文所提出的方法能更好地分割植物叶片图像;同时,提出的ROC曲线能够充分地评估植物叶片的分割性能。

参 考 文 献

- [1] Cope J S, Corney D, Clark J Y, et al. Plant species identification using digital morphometrics: a review [J]. *Expert Systems with Applications*, 2012, 39(8):7562–7573.
- [2] Clarke J, Barman S, Remagnino P, et al. Venation pattern analysis of leaf images [C] // Bebis G, Boyle R, Parvin B, et al. *Advances in visual computing: part II*. Berlin: Springer, 2006:427–436.
- [3] Kumar N, Belhumeur P N, Biswas A, et al. Leafsnap: a computer vision system for automatic plant species identification [C] // Fitzgibbon A, Lazebnik S, Perona P, et al. *Computer vision: ECCV 2012*. Berlin: Springer, 2012:502–516.
- [4] Wang X F, Huang D S, Du J X, et al. Classification of plant leaf images with complicated background [J]. *Applied Mathematics & Computation*, 2008, 205(2):916–926.
- [5] Lee S H, Chan C S, Wilkin P, et al. Deep-plant: plant identification

with convolutional neural networks [C] // *Proceedings of IEEE International Conference on Image Processing*. Piscataway: IEEE Signal Processing Society, 2015:452–456.

- [6] Kavitha P, Ananthi B. Segmentation of unhealthy region of plant leaf using image processing techniques: a survey [J]. *International Journal of Research in Engineering and Technology*, 2014, 11(3):24–27.
- [7] Boykov Y, Funka-Lea G. Graph cuts and efficient N-D image segmentation [J]. *International Journal of Computer Vision*, 2006, 70(2):109–131.
- [8] Shi J, Malik J. Normalized cuts and image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(8):888–905.
- [9] Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation–maximization algorithm [J]. *IEEE Transactions on Medical Imaging*, 2001, 20(1):45–57.
- [10] Valliammal N, Geethalakshmi S N. A novel approach for plant leaf image segmentation using fuzzy clustering [J]. *International Journal of Computer Applications*, 2012, 44(13):10–20.
- [11] Valliammal N, Geethalakshmi S N. Leaf image segmentation based on the combination of wavelet transform and k means clustering [J]. *International Journal of Advanced Research in Artificial Intelligence*, 2012, 1(3):25–39.
- [12] Shelhamer E, Long J, Darrell T, et al. Fully convolutional networks for semantic segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4):640–651.
- [13] 张帅,淮永建.基于分层卷积深度学习系统的植物叶片识别研究[J].北京林业大学学报,2016,38(9):108–115.
Zhang S, Huai Y J. Leaf image recognition based on layered convolutions neural network deep learning [J]. *Journal of Beijing Forestry University*, 2016, 38(9):108–115.
- [14] Noh H, Hong S, Han B. Learning deconvolution network for semantic segmentation [C/OL] // *Proceedings of IEEE International Conference on Computer Vision*. Washington D C: IEEE Computer Society, 2015 [2017-10-10]. <https://arxiv.org/pdf/1505.04366.pdf>.
- [15] Chen L C, Papandreou G, Kokkinos I, et al. Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4):834–848.
- [16] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C] // Pereira F, Burges C J C, Bottou L, et al. *Advances in neural information processing systems*. New York: Curran Associates Inc., 2012:1097–1105.
- [17] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C/OL]. *ICML'15 Proceedings of the 32nd International Conference on Machine Learning*, 2015, 37:448–456 [2017-09-20]. <https://arxiv.org/pdf/1502.03167.pdf>.