

单子叶作物叶片气孔自动识别与计数技术

孙壮壮, 姜东*, 蔡剑, 王笑, 周琴, 黄梅, 戴廷波, 曹卫星

(南京农业大学农业部小麦区域技术创新中心, 农业部南方作物生理生态重点开放实验室, 南京 210095)

摘要:为实现作物叶片气孔的自动识别与快速计数, 该研究采用卷积神经网络中高计算效率的YOLOv3算法, 开发了一种全自动气孔识别和计数解决方案。该算法优化了物体检测性能, 可准确识别显微图像中的气孔。其中, 对指甲油印迹法获得照片的气孔检测精确率、召回率和F1值分别为0.96, 0.98和0.97, 便携式显微镜拍摄法照片气孔检测精确率、召回率和F1值分别为0.95, 0.98和0.96, 具有很好的鲁棒性。该算法检测速度快, 可实现对30帧/s的视频文件进行快速气孔识别, 实现了实时检测。此外, 采用拍摄的小麦叶片照片进行训练得到的气孔识别模型, 还可同时实现对大麦、水稻和玉米等单子叶作物叶片气孔的识别, 其中, 大麦的检测精确率、召回率和F1值分别为0.94, 0.83和0.88; 水稻的检测精确率、召回率和F1值分别为0.89, 0.42和0.57; 玉米的检测精确率、召回率和F1值分别为0.91、0.76和0.83; 显示出模型良好的泛化能力。

关键词:卷积神经网络; 机器视觉; 模型; 单子叶作物; 气孔识别; 计数; 深度学习; 实时检测

doi: 10.11975/j.issn.1002-6819.2019.23.021

中图分类号: S127

文献标志码: A

文章编号: 1002-6819(2019)-23-0170-07

孙壮壮, 姜东, 蔡剑, 王笑, 周琴, 黄梅, 戴廷波, 曹卫星. 单子叶作物叶片气孔自动识别与计数技术[J]. 农业工程学报, 2019, 35(23): 170—176. doi: 10.11975/j.issn.1002-6819.2019.23.021 <http://www.tcsae.org>
Sun Zhuangzhuang, Jiang Dong, Cai Jian, Wang Xiao, Zhou Qin, Huang Mei, Dai Tingbo, Cao Weixing. Automatic identification and counting of leaf stomata of monocotyledonous crops[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(23): 170—176. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2019.23.021 <http://www.tcsae.org>

0 引言

气孔是陆生植物与外界进行水分和气体交换的主要通道, 是植物登陆并适应陆生环境所形成的关键器官之一^[1]。气孔也是协调作物叶片光合与水分散失平衡的重要通道, 在作物应对大气与土壤干旱、高温等逆境中发挥极为重要的作用^[2]。与大部分双子叶植物气孔可通过剥撕表皮细胞观察气孔不同, 单子叶植物叶肉层相对薄、表皮细胞剥离非常困难, 所以一般采用合适的固定剂固定气孔形态并结合显微镜观察, 或者直接通过显微镜观察其气孔形态^[3]。但是, 通过显微镜图像识别气孔并计数过程较为繁琐, 一般拍摄完图像后, 还需通过人工识别统计气孔的数目, 费时费力, 而且长时间通过肉眼观察也会增加出错的概率。随着技术的发展, 已实现通过Image-J等图像处理软件进行气孔识别半手工测量, 但需用户手动标记孔隙边界、气孔长度和宽度等兴趣点^[4]。这些软件可提供额外的插件, 以降低气孔识别的难度, 但仍需使用者手动调整每幅图像的参数, 才能获得较准确的结果。同时, 受观测时间等限制, 研究人员大多从捕获的显微

镜图像中只选择几个气孔进行取样人工测量, 导致只能采用较少的数据点进行相关的研究^[5-6]。目前, 也出现了很多有关气孔识别和计数的算法。如Duarte等用预处理后的图像通过颜色特征进行阈值分割以分离出气孔进行计数^[7], 但这种方法在采集照片的光照等环境发生变化时, 效率会明显下降。还有一些算法需要用户先手动分割气孔, 以为后面的气孔样本检测提供模板进行匹配^[8], 同时还需要用户输入一些模型的参数, 以适应具体的科研需求。总之, 上述研究未能达到自动化检测的效果。因此, 建立一种快速、全自动、精准的气孔识别与计数工具, 对提升相关领域的科研效能具有重要意义。

近年来, 卷积神经网络(CNN, convolutional neural networks)在目标检测、语义分割等领域获得广泛应用, 目标检测算法在精准农业领域也取得了很大的突破。如Madec等基于目标检测技术, 采用高分辨率RGB图像估算了小麦麦穗密度^[9]。目前, 较流行的目标检测算法有两类: 一类基于候选区域(region proposal)的R-CNN系算法(R-CNN, Fast R-CNN, Faster R-CNN), 该算法分两步进行, 首先使用启发式方法(selective search)或者CNN网络(RPN)产生候选区域, 然后再在候选区域上进行分类与回归。该类方法准确度较高, 但由于提取候选区域计算量较大, 降低了该类方法检测速度。第二类是YOLO(you only look once)算法, 该算法是一个端到端的实时目标检测系统, 将物体检测任务当做一个回归问题来处理, 使用一个神经网络, 直接从一整张图像来预测出后处理的坐标、框中包含物体的类别和置信度。通

收稿日期: 2019-07-28 修订日期: 2019-10-22

基金项目: 重点研发计划项目 2016YFD0300107; 国家自然科学基金(U1803235、31771693); 国家现代小麦产业技术体系(CARS-03); 江苏省协同创新中心(JCIC-MCP); “111”引智项目(B16026)

作者简介: 孙壮壮, 主要从事作物表型技术研究。

Email: 2018101001@njau.edu.cn

*通信作者: 姜东, 教授, 主要从事作物生理生态与作物表型组学技术研究。

Email: jiangd@njau.edu.cn

过这种方式, YOLO 可实现 45 帧/s 的运算速度 (达到 24 帧/s, 人眼就认为是连续的), 完全满足实时观测的要求。YOLO 算法先后经历了 YOLOv1^[10]、YOLOv2^[11] 和 YOLOv3^[12] 3 个版本, YOLOv3 在保证前期版本高运算速度的基础上, 大幅度提高了识别精度, 在很多领域得以广泛使用^[13]。

为此, 本研究运用 YOLOv3 算法, 以单子叶作物叶片气孔照片为对象, 开发了单子叶作物叶片气孔快速精确检测与计数技术。以期对相关领域的科研工作者提供更好的研究方法。

1 材料与方法

1.1 试验设计

本研究通过两种方式进行照片采集: 指甲油印记法和便携式显微镜拍摄, 拍摄时期为灌浆期和幼苗期 (三叶一心期), 选用小麦品种扬麦 16 和苏麦 188 作为试验材料。

灌浆期植株于 2018—2019 年间种植在南京农业大学牌楼试验站, 使用盆栽种植, 土壤自然风干过筛后, 播种前与肥料充分混匀, 称取 7.5 kg 装入高 22 cm、直径 25 cm 的聚乙烯塑料盆, 每盆施用尿素 2 g, 磷酸二氢钾 2 g, 氮肥分两次施入, 基追比为 5:5, 追肥拔节期施入, 每桶选取饱满均匀一致的种子播种 21 粒, 三叶一心期定苗, 每桶留苗 7 株, 共计 50 盆。灌浆期时, 选取小麦顶展叶, 在叶背面中部用指甲油印迹法和便携显微镜分别拍摄气孔, 每盆拍摄 2~3 张。

幼苗期的小麦使用了水培试验方法。选取籽粒大小均匀一致的种子, 用质量分数为 15% H₂O₂ 溶液浸泡消毒 15~20 min, 经蒸馏水冲洗干净, 摆放在周转箱中, 将周转箱放置在光照培养箱中进行种子催芽, 待种子吸胀露白后, 选取均匀一致的种子摆入网筐中, 种子腹沟朝下, 按时补充水分。待小麦幼苗长至一叶一心时, 将长势一致的幼苗转移到盛有水培营养液的周转箱内。幼苗置于人工气候室培养, 气候室参数设置为昼夜温度 20 ℃ 和 14 ℃, 16 h 光照、光照强度为 300 $\mu\text{mol}/(\text{m}^2\cdot\text{s})$, 相对湿度 60%, 2 d 更换一次营养液。幼苗长至三叶一心时, 选取小麦顶展叶背面中部拍摄气孔。

1.2 图像获取

1.2.1 指甲油印迹法

将透明指甲油均匀涂抹在小麦等单子叶作物叶片表面, 自然晾干后, 用透明胶粘贴叶片表皮并轻轻撕下, 转移至载玻片上, 然后置于正倒置荧光显微镜下, 物镜倍数为 20 倍, 目镜倍数为 10 倍, 对制作的装片进行观察和显微摄影, 将采集的照片按顺序编号。

1.2.2 便携式显微镜拍摄法

该方法采用艾尼提便携式显微镜进行叶片气孔拍摄 (Anyty 3R-WM401WiFi/3R-WM601WiFi 系列, 以下简称艾尼提显微镜)。艾尼提显微镜与智能终端连接后, 安装相应软件, 即可在相应倍率下观测, 还可以进行拍照和录像。通过艾尼提显微镜, 直接对小麦等作物叶片气孔进行拍摄, 放大倍数为 400~600 倍, 每次拍摄放大倍数

需根据实际情况调节。2 个时期取样照片分布如表 1 所示。

表 1 取样照片分布情况
Table 1 Distribution of sample photos

方法 Method	三叶一心期照片数量 Number of photos when the third leaf occurred	灌浆期照片数量 Number of photos at grain filling stage	总计 Total
指甲油印迹法 Method1	49	89	138
便携式显微镜拍摄法 Method2	40	77	117

1.3 图像标注

使用 labelImg (<https://github.com/tzutalin/labelImg>) 工具对照片进行标注, 生成 label 文件。气孔用矩形框标注出, 矩形框以气孔最小的外接矩形为最佳, 位于照片边缘的气孔由人工判断, 气孔大小超过 2/3 的则计入整体个数。其中, 指甲油印迹法一共标注了 8 024 个气孔, 便携式显微镜拍摄法一共标注了 2 608 个气孔。

将照片中的所有气孔标出后, 点击保存, 生成与照片同名的 xml 格式文件, 文件中记录照片里每个气孔矩形框的位置坐标及通道数等信息, 由于训练时需要用 txt 格式的 label 文件, 所以用代码 xml-to-txt-updata.py (<https://github.com/shem123456/-/blob/master/xml-to-txt-updata.py>) 进行格式转化, 转化后生成 txt 格式文件, 作为训练的 label 文件。

1.4 数据集划分

整个数据包括照片及其所对应的 txt 文件, 数据集分为训练集和测试集, 其中, 80% 的照片和与其对应的 txt 文件放入训练集 (指甲油法和便携显微镜拍摄法各 111 和 94 个), 用于模型训练; 20% 的照片和与其对应的 txt 文件放入测试集 (指甲油法和便携式显微镜拍摄法各 27 和 23 个)。

2 数据处理

2.1 处理流程

整个处理流程如图 1 所示, YOLOv3 代码目前已经开源, 可从网站 (<https://pjreddie.com/darknet/yolo/>) 下载源代码。

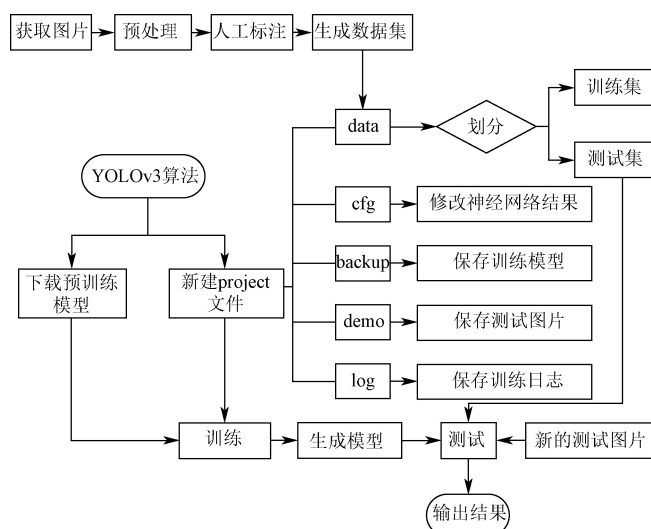


图 1 系统总体工作流程

Fig.1 General work-flow of the proposed system

参照官网的要求, 下载 darknet53.conv.74 预训练权重文件 (YOLO 作者在 ImageNet 上预训练的权重)。并新建一个名为 project 的文件夹, 在里面建 5 个文件夹, 分别为 data、cfg、backup、demo、log。其中: data 文件夹存放数据集, 包括训练、测试集, 以及相应的 txt 文件; cfg 文件夹存放 darknet 描述神经网络结构的文件; backup 存放训练的模型, 同时存放训练过程中的备份模型, 可以作为断点备份继续训练; demo 用于保存测试图片; log 用于保存训练日志。

2.2 模型训练

本文采用深度学习中的 YOLOv3 算法进行模型构建和训练, YOLOv3 神经网络结构为 Darknet-53, 如图 2 所示。整个网络一共有 53 层, 前 52 层主要由一系列的 1×1 和 3×3 的卷积层组成, 进行了 5 次下采样, 每次采样步长为 2, 在 3 个特征层进行了边框的预测, 分别是在 32 倍下采样, 16 倍下采样和 8 倍下采样的特征图, 在 3 个尺度的融合特征图上分别独立做检测, 这样有助于让网络同时学习深层和浅层特征, 表达效果更好, 同时保留细粒度特征, 增强小目标的检测效果。第 53 层是全连接层, 通过 Softmax 函数进行目标的分类。Convolutional 由 Conv2d Layer、BN Layer 和 LeakyReLU Layer 等 3 个模块组成, 整个网络中使用大量残差结构 (Residual), 这种残差结构可以使网络结构在很深的情况下, 仍能收敛, 而残差结构中 1×1 的卷积, 极大减少了每次卷积的通道数, 有利于减少参数量 (参数量越大, 保存的模型越大), 同时也在一定程度上也减少计算量。

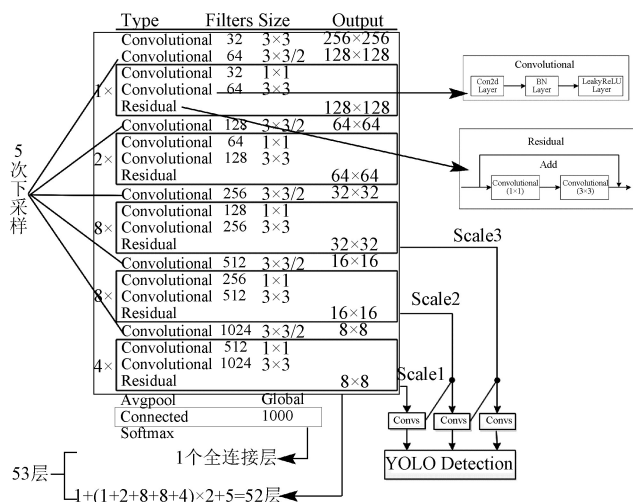


图 2 YOLOv3 模型结构图

Fig.2 Model structure of YOLOv3

训练时使用的电脑系统为 Ubuntu16.04, 采用 GPU (NVIDIA GeForce GTX 1050Ti) 加速。在训练之前, 使用预训练网络 (darknet53.conv.74) 的模型, 这是一种常用且高效的迁移学习方法, 因为这个预训练网络的原始数据集足够大且通用, 所以它学到的特征空间层次结构可以有效地作为人眼视觉世界的通用模型。

在训练过程中, 可以通过代码将每次迭代完返回的信息写入日志中, 并存放在 log 文件中。通过 python 代码 extract_log.py, 解析日志文件, 再使用

train_loss_visualization.py 绘制损失变化曲线, 进行训练过程的可视化, 效果如图 3 所示。其中, 横坐标代表训练次数, 纵坐标代表损失值, loss 和 avg loss 分别表示损失和平均损失。当返回的 loss 和 avg loss 数值很低, 且趋于稳定时, 即可停止训练, 保存训练模型。从训练损失变化曲线可知, 训练次数在 100 次之后, loss 和 avg loss 会迅速下降, 到 400 次左右时, 下降的幅度又会变得缓慢, 最终, 选择在 1200 次左右时, 停止训练, 并对模型进行保存。以上的 python 代码均可在 <https://github.com/shem123456/-/blob/master> 中下载。

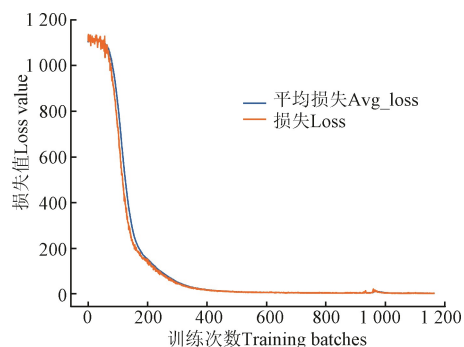
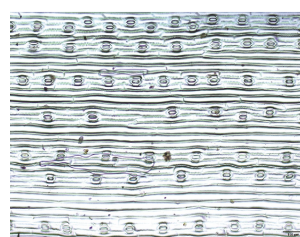


图 3 平均损失和损失变化曲线

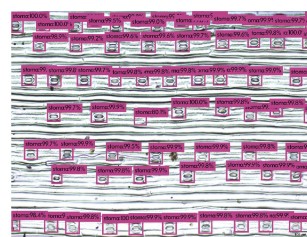
Fig.3 Average loss and loss curves

2.3 模型测试

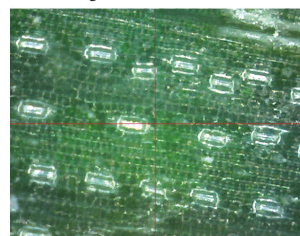
对 2 个数据集分别进行训练后, 再用训练后得到的模型对单张照片进行测试。如图 4 所示, 算法会自动预测出每个气孔的位置, 并用矩形框对其大小进行标注, 同时它给出该物体类别的置信度 (为了减少遮挡, 结果与讨论部分的图片隐藏了置信度的显示), 算法默认的置信度阈值为 50%, 即只会标出置信度大于 50% 的物体, 并在图片中显示出置信度。



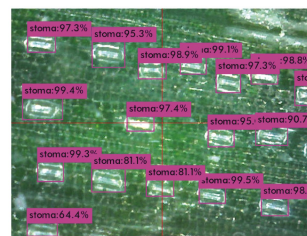
a. Method 1 原始图片 (放大倍数: 200 倍)
a. Original image from Method 1 with a magnification of 200 times



b. Method 1 检测结果
b. Detection result from Method 1



c. Method 2 原始图片 (放大倍数: 600 倍)
c. Original image from Method 2 with a magnification of 600 times



d. Method 2 检测结果
d. Detection result from Method 2

注: Method 1 为指甲油印迹法, Method 2 为便携显微镜拍摄法。
Note: Method 1 is the method of nail polish printing, and Method 2 is the method of portable microscopy.

图 4 单张图片测试

Fig.4 Test result using single image

2.4 模型评估

用 P (Positive) 和 N (Negative) 代表模型的判断结果, T (True) 和 F (False) 代表评价模型的判断结果是否正确, 这样所有结果可以分为 4 类: FP (假正类)、FN (假负类)、TP (真正类) 和 TN (真负类)。根据这 4 种分类, 采取如下 3 个指标来评估模型的性能。

召回率 (Recall), 针对数据集集中的所有正类 (TP+FN) 而言, 模型正确判断出的正类 (TP) 占数据集中所有正类的比例。召回率也叫查全率, 以气孔检测为例, 召回率高代表着模型可以找出图片中更多的气孔。

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

精确率 (Precision), 也叫查准率, 表示模型检测出的物体为真实物体的比例, 指针对模型判断出的所有正类 (TP+FP) 中, 真正类 (TP) 占的比例。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

F1 值 (F1 Score), 是统计学中用来衡量二分类模型精确度的一种指标。它同时兼顾了分类模型的精确率和召回率。F1 值可以看作是模型精确率和召回率的一种加权平均, 最大值为 1, 最小值为 0。

$$\text{F1} = 2 \times \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

3 结果与分析

3.1 指甲油印迹法在图片清晰度和模型性能方面均要优于便携式显微镜拍摄法

2 种气孔图片采集方式相比, 指甲油印迹法得到的照片中气孔形态和轮廓更加清晰, 而且采样的面积也更大, 单张照片中能观测到气孔数量更多。受镜头物理倍率及软件功能的限制, 艾尼提显微镜拍摄照片的最高分辨率仅为 11.8 像素/cm, 但其优点是拍摄的照片更加真实, 且操作方便、快捷, 不受观察时间、地点限制, 更重要的是不需要对叶片进行破坏性处理, 可实现连续观测。

使用精确度、召回率和 F1 值对 2 个模型的性能进行评估。首先对 Method 1 和 Method 2 测试集中每张图片进行单张图片测试, 然后统计每张图片中的 FP、FN、TP 和 TN, 最后计算出每张图片的精确率、召回率和 F1 值 (图 5)。图中横坐标代表精确率, 纵坐标代表召回率, 由于气泡的大小较为接近, 在图的右侧添加了颜色棒, 便于区分气泡的大小, 由深到浅代表气泡大小逐渐变小。由图可知, 2 种方法生成模型的性能都不错。在测试集检测的结果中, Method 1 平均的精确率、召回率和 F1 值分别为: 0.96, 0.98 和 0.97; Method 2 平均的精确率、召回率和 F1 值分别为: 0.95, 0.98 和 0.96。

3.2 YOLOv3 算法在测试集中表现出良好的鲁棒性

YOLOv3 算法在 2 种数据集上都表现出良好的性能, 而且两种方法均可实现准确气孔计数, 以置信度 50% 为阈值, 气孔真实个数与预测个数线性回归 R^2 分别为 0.980 1 和 0.962 5。用两种方法的测试集作为输入, 进行批量测试, 置信度阈值为 50%, 得到每张图片中预测气

孔个数, 将每张图片实际标注的气孔个数与预测的气孔个数进行线性回归分析, 发现 2 种方法实际标注的气孔个数与预测的气孔个数均有很强的相关性 (图 6)。

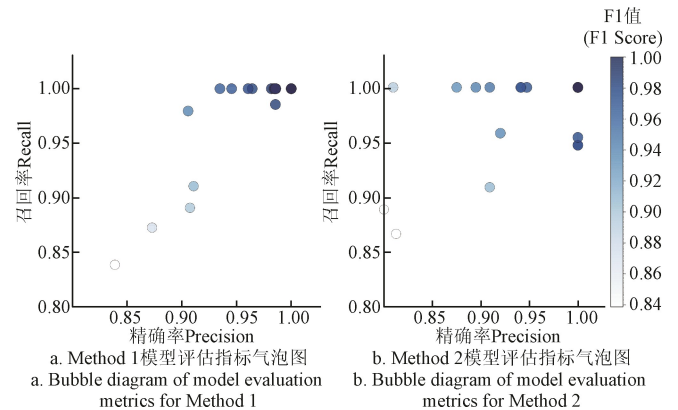


图 5 模型评估指标气泡图

Fig.5 Bubble diagram of model evaluation metrics

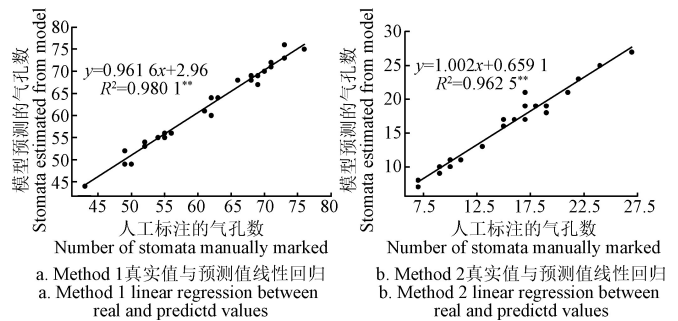


图 6 实际标注与预测气孔数目间的线性回归关系

Fig.6 Linear regression relationship between the labeled stomata number and the predicted number

3.3 YOLO 算法可实现气孔的实时检测

如前所述, YOLO 是一个端到端的实时的目标检测系统, 它将物体检测任务当做一个回归问题来处理, 使用一个神经网络, 直接从一整张图像来预测出边界框的坐标、边界框中包含物体的类别和置信度。通过这种方式, 可实现较高的运算速度, 能满足实时性要求。

本研究用指甲油印迹法制作一样本, 将其置于倒置荧光显微镜下, 调整好放大倍数及焦距至影像清晰, 打开配套的软件, 点击录制视频, 然后缓缓转动载物台旋钮, 使观测的视野移动, 同时保证图像的清晰, 录制视频文件, 得到的视频文件帧速率为 30 帧/s。

将录制的视频文件作为测试对象, 输入到 YOLOv3 算法中, 算法可自动识别检测每一帧影像中的气孔, 并标识其位置, 最后输出一个标记后的视频文件。如图 7, 显示了气孔个数实际值与预测值随时间变化的曲线, 表明在检测视频文件时, 算法也具有很好的鲁棒性。

3.4 该方法对其他单子叶作物的适用性

区别于双子叶植物, 小麦、大麦、玉米、水稻等单子叶作物的气孔具有较为相似的形态特征^[14]。由此, 选用由小麦叶片气孔照片训练构建的模型, 测试该模型是否也适于其他单子叶作物叶片的气孔识别。为此, 采用水培法培育了大麦、水稻、玉米等幼苗, 利用便携式显

显微镜采集了作物叶片气孔的照片,以此来测试模型对其他单子叶作物的识别效果。

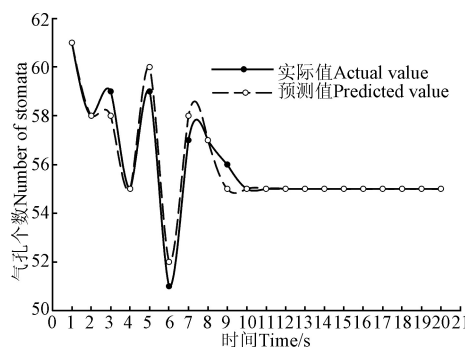


图7 气孔个数实际值与预测值随时间变化曲线

Fig.7 Curve of actual value and the predicted value of number of stomata with time

如表2所示,大麦的检测精确率、召回率和F1值分别为0.94, 0.83和0.88,水稻的检测精确率、召回率和F1值分别为0.89, 0.42和0.57,玉米的检测精确率、召回率和F1值分别为0.91、0.76和0.83。如图8所示,仅用小麦气孔照片进行训练得到的模型,也可实现对大麦、水稻和玉米等单子叶作物叶片气孔的识别。原因在于卷积神经网络在提取特征时,不仅提取了颜色、形状等单一的特征,而是更为复杂和丰富的抽象结构^[15]。

表2 其他单子叶作物的检测精度评价

Table 2 Evaluation of detection accuracy of other monocotyledons

作物 Crops	气孔个数 Number of stomata			精确率 Precision	召回率 Recall	F1 值 F1 Score
	TP	FP	FN			
大麦	300	19	60	0.94	0.83	0.88
水稻	160	20	220	0.89	0.42	0.57
玉米	260	27	80	0.91	0.76	0.83

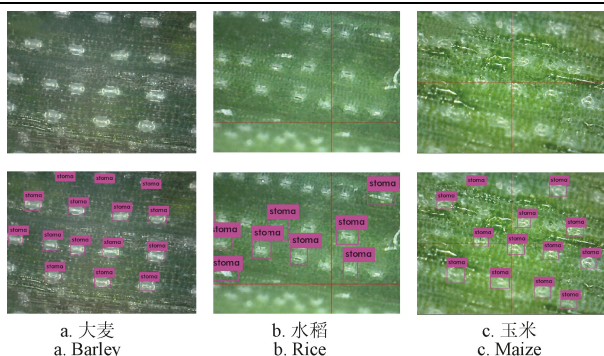


图8 其他单子叶作物

Fig.8 Leaf stomata identifying and counting of barley, rice and maize

在一般的深度神经网络中,刚开始提取的特征是各种边缘形状的集合,而随着层数的加深,提取的特征变得越来越抽象,并且越来越难以直观理解。深度神经网络可以有效地作为信息蒸馏管道,反复地对原始数据(RGB图像)进行处理,过滤掉无关信息,并放大和细化有用信息^[16]。从直观上看,能发现大麦作为小麦的近源种,其气孔在形态结构与小麦最相近,与水稻和玉米气孔形态差异较大。

4 讨论

本文提出了一种快速、全自动、精准的气孔识别与计数工具。与常规的基于阈值分割气孔的方法相比,基于卷积神经网络的算法具有更好的泛化能力,而且本文提出的方法不需要用户再手动输入参数进行调整,实现了自动检测。与Fetter等使用深度卷积神经网络提出的气孔自动计数系统StomataCounter相比^[17],本文提出了以视频为输入的气孔动态识别与计数,为全方位地监测作物气孔的动态变化提供可能。

虽然本文的方法在测试集上表现出良好的性能,但对模型的泛化能力仍需进行更多的测试和探讨。我们使用支架固定便携式显微镜,将小麦叶片置于载物台上,放大倍数为600倍,对叶片气孔进行24h的视频录制,并通过算法检测其气孔的数目。由于原视频文件较大,通过压缩工具将原视频从10.7 GB压缩到了52.6 MB,如图9所示,截取了视频中20:00,次日1:00和次日7:00等3个时间点的画面,其中20:00的检测精确率、召回率和F1值分别为1.00, 1.00和1.00;次日1:00的检测精确率、召回率和F1值分别为1.00, 0.92和0.96;次日7:00的检测精确率、召回率和F1值分别为1.00, 0.92和0.96。通过视频,能较完整地观测到叶片气孔由闭到开的过程,同时也验证了昼夜环境变化和气孔开闭程度差异等因素并不会影响模型的检测性能。

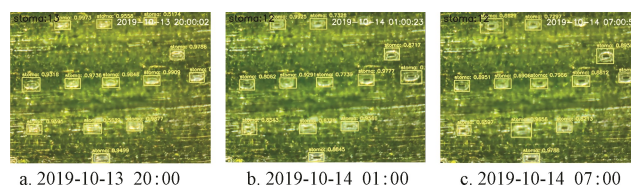


图9 昼夜变化下的气孔检测

Fig.9 Stomatal detection under diurnal variation

在试验过程中,还关注了干旱、渍水、低温等在内的非生物胁迫环境下的模型性能。采用在人工气候室中水培幼苗的方法,模拟了小麦在干旱、渍水、低温等非生物胁迫下的生长环境,在这些环境下,模型仍能精确识别出叶片中的气孔,其中,干旱环境下的气孔检测精确率、召回率和F1值分别为0.94, 0.99和0.97,渍水环境下的气孔检测精确率、召回率和F1值分别为0.95, 1.00和0.97,低温环境下的气孔检测精确率、召回率和F1值分别为0.96, 0.99和0.98。从目前的试验中,模型表现出了良好的泛化能力,但仍不能保证该模型适用于所有场景下的气孔识别,如患有赤霉病、白粉病等生物胁迫下小麦叶片的气孔。由于生物胁迫下小麦叶片会出现病斑等噪声干扰,可能会出现假正类(FP),这种情况下模型的性能有待进一步验证。

在一般的图像分类任务中,为了测试一个模型的性能,在计算机视觉领域往往使用一些公共的、标准的大型数据集,如主要用于图像分类、目标检测和语义分割的ImageNet和Pascal VOC数据集^[18],以及主要关注实例分割的COCO数据集^[19]。这些大型数据集由专业的数据

公司整理标注, 体量十分庞大, 在这样的数据集进行测试更能衡量一个模型的性能。但是, 目前在植物表型领域这种标准化的数据集还比较少, 表型数据的整合和标准化还需多个科研团队之间的学科交叉和长期合作^[20]。因此, 本文对算法进行了较完整的封装, 并开源了原始的 Python 代码, 可在 GitHub 中下载 (<https://github.com/shem123456/>), 给相关领域的科研人员提供了一个可以测试气孔照片或视频的接口, 以期能给相关科研人员提供帮助或参考, 并希望能进一步完善数据集, 增加模型的泛化能力。

5 结 论

本文先采用两种常用的图像采集方法获取了叶片气孔图像, 然后运用深度学习中的 YOLOv3 算法, 分别对 2 种方法采集的气孔图像进行识别与计数建模。结果表明, 该算法建立的模型能够正确地识别出 2 种采集方法中的气孔, 其中, 对指甲油印迹法获得照片的气孔检测精确率、召回率和 F1 值分别达到 0.96, 0.98 和 0.97, 便携式显微镜拍摄法照片气孔检测精确率、召回率和 F1 值分别达到 0.95, 0.98 和 0.96, 具有很好的鲁棒性。该方法能准确地估算出其气孔数目, 通过将测试集的图片上的气孔真实个数与预测个数进行线性回归, 发现两种方法下实际标注的气孔个数与预测的气孔个数均有很强的相关性, 线性回归 R^2 分别为 0.980 1 和 0.962 5。同时, YOLOv3 算法是一个端到端的实时目标检测系统, 在检测速度上具有很大的优势, 以视频文件 (帧速率: 30 帧/s) 为输入时, 达到实时检测的效果。本模型虽然是基于小麦气孔图像进行标注训练的, 但是对于大麦、水稻和玉米等单子叶作物的气孔, 同样具有很好的检测效果, 其中, 大麦的检测精确率、召回率和 F1 值分别为 0.94, 0.83 和 0.88, 水稻的检测精确率、召回率和 F1 值分别为 0.89, 0.42 和 0.57, 玉米的检测精确率、召回率和 F1 值分别为 0.91、0.76 和 0.83。

参 考 文 献

- [1] 王瑞丽, 于贵瑞, 何念鹏, 等. 气孔特征与叶片功能性状之间关联性沿海拔梯度的变化规律: 以长白山为例[J]. 生态学报, 2016, 36(8): 2175—2184.
Wang Ruili, Yu Guirui, He Nianpeng, et al. Attitudinal variation in the covariation of stomatal traits with leaf functional traits in Changbai Mountain[J]. Acta Ecologica Sinica, 2016, 36(8): 2175—2184. (in Chinese with English abstract)
- [2] 成雪峰, 张凤云, 柴守玺. 春小麦对不同灌水处理的气孔反应及其影响因子[J]. 应用生态学报, 2010, 21(1): 36—40.
Cheng Xuefeng, Zhang Fengyun, Chai Shouxi. Stomatal response of spring wheat and related affecting factors under different irrigation treatment[J]. The Journal of Applied Ecology, 2010, 21(1): 36—40. (in Chinese with English abstract)
- [3] 韦存虚, 谢佩松, 周卫东, 等. 凹凸不平的植物叶片表皮制片方法的观察比较[J]. 生物学杂志, 2008, 25(2): 63—66.
- [4] Wei Cunxu, Xie Peisong, Zhou Weidong, et al. Comparison of preparing slice techniques of convexo-concave leaf epidermis[J]. Journal of Biology, 2008, 25(2): 63—66. (in Chinese with English abstract)
- [4] Cheng Yan, Cao Ling, Wang Sheng, et al. Analyses of plant leaf cell size, density and number, as well as trichome number using cell counter plugin[J]. Bio-protocol, 2014, 4(13): 54—57.
- [5] Jayakody H, Liu S, Whitty M, et al. Microscope image based fully automated stomata detection and pore measurement method for grapevines[J]. Plant Methods, 2017, 13(1): 94—98.
- [6] Peel J R, Mandujano Sánchez M C, López Portillo J, et al. Stomatal density, leaf area and plant size variation of *Rhizophora mangle* (Malpighiales: Rhizophoraceae) along a salinity gradient in the Mexican Caribbean[J]. Revista de Biología Tropical, 2017, 65(2): 701—712.
- [7] Duarte K, Carvalho M, Martins P. Segmenting high-quality digital images of stomata using the wavelet spot detection and the watershed transform[C]// Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Setúbal: Science and Technology Publications, 2017: 540—547.
- [8] Laga H, Shahinnia F, Fleury D. Image-based plant stomata phenotyping[C]. 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV). IEEE, 2014: 217—222.
- [9] Madec S, Jin Xiuliang, Liu Shouyang, et al. Ear density estimation from high resolution RGB imagery using deep learning technique[J]. Agricultural and Forest Meteorology, 2019, 264: 225—234.
- [10] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society: Piscataway, NJ. 2017: 6517—6525.
- [11] Redmon J, Farhadi A. YOLO9000: Better, faster, stronger[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 7263—7271.
- [12] Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. 2018, arXiv: 1804.02767.
- [13] Braun M, Krebs S, Flohr F, et al. Euro City Persons: A novel benchmark for person detection in traffic scenes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(8): 1844—1861.
- [14] 侯慧姣, 韩鸿鹏. 单、双子叶植物气孔发育调控差异研究进展[J]. 分子植物育种官方网站, 2016, 14(3): 896—903.
Hou Huijiao, Han Hongpeng. Research progress on the difference of stomata development and their regulation between monocotyledon and dicotyledon[J]. Molecular Plant Breeding, 2016, 14(3): 896—903. (in Chinese with English abstract)
- [15] 刘军, 后士浩, 张凯, 等. 基于增强 Tiny YOLOV3 算法的车辆实时检测与跟踪[J]. 农业工程学报, 2019, 35(8):

- 118—125.
- Liu Jun, Hou Shihao, Zhang Kai, et al. Real-time vehicle detection and tracking based on enhanced Tiny YOLOV3 algorithm[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(8): 118—125. (in Chinese with English abstract)
- [16] Francois C. Deep learning with Python[M]. 178 South Hill Drive, Westampto, NJ 08060 USA. Manning Publications. 2018.
- [17] Fetter K C, Eberhardt S, Barclay R S, et al. StomataCounter: a neural network for automatic stomata identification and counting[J]. New Phytologist, 2019, 223(3): 1671—1681.
- [18] Hoiem D, Divvala S K, Hays J H. Pascal VOC 2008 challenge[C]//PASCAL challenge workshop in ECCV. 2010.
- [19] Veit A, Matera T, Neumann L, et al. Coco-text: Dataset and benchmark for text detection and recognition in natural images[J]. arXiv preprint arXiv:1601.07140, 2016.
- [20] 周济, Tardieu F, Pridmore T, 等. 植物表型组学: 发展、现状与挑战[J]. 南京农业大学学报, 2018, 41(4): 580—588.
- Zhou Ji, Tardieu F, Pridmore T, et al. Plant phenomics: history, present status and challenges[J]. Journal of Nanjing Agricultural University, 2018, 41(4): 580—588. (in Chinese with English abstract)

Automatic identification and counting of leaf stomata of monocotyledonous crops

Sun Zhuangzhuang, Jiang Dong*, Cai Jian, Wang Xiao, Zhou Qin, Huang Mei, Dai Tingbo, Cao Weixing

(Regional Technique Innovation Center for Wheat Production, Ministry of Agriculture, Key Laboratory of Crop Physiology and Ecology in Southern China, Ministry of Agriculture, Nanjing Agricultural University, Nanjing 210095, China)

Abstract: Stomata is the dominant gate for water and gas exchange for plant leaves, and thus plays key roles in plants in response to the fluctuations of the environmental variables. Observation and counting stomata amounts is generally one of the obligatory determinants in the research of plant ecology and physiology. The classic measurement protocol of leaf stomata usually includes the capture of leaf stomata by a microscope, followed by manually identifying and counting of the target stomata. This method is well-known in disadvantages of both time and labor consuming, and of low accuracy. Although some algorithms for stomata recognition have been proposed at present, their recognition abilities showed limitation, and they could not realize the full effect of automatic recognition. Thereafter, we developed an automatic identification and counting technique based on YOLOv3, one of the high speed convolutional neural networks (CNN) algorithm in the present study. We acquired pictures of leaf stomata after the third leaf occurred and during grain filling stage of wheat (*Triticum aestivum*), in which, 138 pictures were taken from the method of nail polish printing, and another 117 pictures were taken from the method of portable microscopy. After that, we created separate data sets and then trained the corresponding models respectively. During the training process, we visualized the loss and average loss which were the most important training parameters, and finally stopped the training at 1 200 times. To better describe the parameters of both models, we used the key metrics to evaluate the models, such as precision, recall and F1. The precision, recall and F1 reached 0.96, 0.98 and 0.97 in the method of nail polish printing, whereas reached 0.95, 0.98 and 0.96 in the method of portable microscopy. Secondly, this algorithm could count stomata amounts accurately, and showed excellent robustness. By linear regression between the labeled and predicted stomata amounts in pictures from test sets, we found that this algorithm showed strong correlation, R^2 were 0.980 1 and 0.962 5, respectively. What's more, this algorithm also showed high performance in high-throughput and real-time, since it identified stomata with a speed of 30 frames per second. With this technique, we optimized the objective identifying performance, which conferred accurate identification performance of stomata in the microscope pictures of leaf stomata. Firstly, compared with the method of nail polish printing, the method of portable microscopy showed low precision and F1, but was harmless to samples. Secondly, YOLOv3 algorithm exhibited the merits of accuracy, high efficiency, as well as real-time, long-time and dynamic detection. Thirdly, this technique was high compatible due to its power in accurately identifying stomata of other monocotyledonous crops such as barley (*Hordeum vulgare*), rice (*Oryza sativa*) and maize (*Zea mays*). Lastly, in order to facilitate the use of more researchers, we not only opened source of the detailed Python code, but also encapsulated the method in a relatively complete way. It could provide an interface for relevant researchers to detect their stomata photo or video files. The files in our stomata project could be consulted and downloaded in Github (<https://github.com/shem123456/>).

Keywords: convolutional neural network; machine vision; models; monocotyledonous crop; stomata identification; stomata counting; deep learning; real-time detection