doi:10.6041/j.issn.1000-1298.2021.12.019

基于 IM - SSD + ACO 算法的整株大豆表型信息提取

宁　姗[1,2]　陈海涛[1]　赵秋多[1,2]　王业成[1]

(1. 东北农业大学工程学院, 哈尔滨 150030;
2. 黑龙江科技大学工程训练与基础实验中心, 哈尔滨 150022)

摘要: 为了减少检测整株大豆豆荚及茎秆时相互遮挡对精度造成的影响,提出了一种基于卷积神经网络的大豆豆荚及茎秆表型信息检测方法,根据大豆植株的生长特征和卷积网络的特点,对单次多框检测器(Single shot multibox detector, SSD)进行了改进。与传统 SSD 相比,改进 SSD(IM - SSD)具有更好的抗干扰能力和自学习能力。首先,通过大豆植株图像采集平台获取收获期的大豆植株图像,建立大豆植株 RGB 空间图像数据集,将数据集分为训练集、测试集和验证集,对训练集进行颜色变换、图像平移、旋转和缩放等方式实现数据的扩增,提高网络的泛化能力。其次,提出一种针对大豆植株图像中豆荚和茎秆的标注方法,仅对未被遮挡的部分进行标注,目的是降低遮挡产生的误判。IM - SSD 是在传统 SSD 结构的基础上增加 2 个残差层,使用低层特征图融合到高层特征图来增强对小目标的检测能力,提高网络的识别率,输入图像尺寸为 600 像素 ×300 像素,降低压缩变形带来的影响。对比试验结果表明,IM - SSD 的平均精度比 SSD300 高 7.79 个百分点,比 SSD512 高 3.83 个百分点。由于卷积神经网络获得的大豆植株茎秆定位是分段的,不能体现茎秆的真实特征,提出了一种基于蚁群优化(Ant colony optimization, ACO)算法的大豆植株茎秆提取方法,利用 ACO 结合 IM - SSD 的结果提取完整的大豆植株茎秆。最后,通过豆荚定位和大豆植株茎秆提取获得了大豆植株的部分表型信息,包括全株荚数、株高、有效分枝数、主茎与株型。

关键词: 大豆植株; 目标检测; 卷积神经网络; 蚁群优化

中图分类号: TP391.4　　**文献标识码**: A　　**文章编号**: 1000-1298(2021)12-0182-09　　**OSID**:

Detection of Pods and Stems in Soybean Based on IM - SSD + ACO Algorithm

NING Shan[1,2]　CHEN Haitao[1]　ZHAO Qiuduo[1,2]　WANG Yecheng[1]

(1. *College of Engineering*, *Northeast Agricultural University*, *Harbin* 150030, *China*
2. *Center of Engineering Training and Basic Experiment*, *Heilongjiang University of Science and Technology*, *Harbin* 150022, *China*)

Abstract: Soybean is an important crop in agriculture and plays an important role in the agricultural field of the world. With the increase of population and disasters caused by abnormal climate, how to cultivate more adaptive high-yield crop varieties has become a major problem faced by breeding experts. A soybean detection method was proposed based on deep learning to reduce the influence of illumination, growth difference and occlusion. In view of characteristics of soybean and accuracy of deep learning, single shot multibox detector (SSD) was improved. Compared with the SSD, the improved SSD (IM - SSD) had better anti-interference ability and self-learning ability. The first step was to build datasets by taking 3695 photos of harvested soybean plants under the fixed and defined light condition and blue background described. And the training set was randomly changed by images translation, rotation and scaling to expand the datasets. Then training datasets were used to train and construct network model. To study the effect of base network and the input size of networks, one contrast experiment was designed on SSD and IM - SSD. Through the analysis of the experimental results, the average accuracy of IM - SSD was 7.79 percentage points higher than that of SSD300 and 3.83 percentage points higher than that of SSD512, respectively. Compared with SSD, IM - SSD was improved in soybean pod and stem detection. Nevertheless, the location of the stem by IM - SSD was discontinuous. A method of stem extraction was proposed, which used IM - SSD results and ant colony optimization (ACO) algorithm to extract the whole

收稿日期: 2020 - 12 - 08　修回日期: 2021 - 03 - 13

基金项目: 国家重点研发计划项目(2016YFD0100201、2018YFD0201004)

作者简介: 宁姗(1978—),女,博士生,黑龙江科技大学讲师,主要从事机器视觉及农业机械化研究,E-mail: 4127756@qq.com

通信作者: 陈海涛(1962—),男,教授,博士生导师,主要从事农业机械化装备及生物质材料研究,E-mail: htchen@neau.edu.cn

stem. The experimental results showed that the IM－SSD and the stem extraction method could accurately locate pod and stem of soybean plants. Finally, some phenotypic information of soybean plants was obtained, including the number of pods of the whole plant, plant height, effective branch number, main stem and plant type.

Key words: soybean plant; target recognition; convolution neural network; ant colony optimization

0　引言

种质资源是大豆品质和产量的基础，大豆考种是大豆作物表型组学研究中种质资源鉴定与筛选的重要方法之一。作物表型组学研究能够获得高质量、高精度和可重复性的植物表型数据[1－2]。然而，传统大豆考种工作主要依靠人工，在考种过程中需要消耗大量的人力物力。同时，人工操作还存在人为主观误差、效率低、时间周期长等问题。人工考种的方式已经制约了植物生物学研究的发展，随着机器视觉技术的飞速发展，采用高效率、低误差、低成本的自动化表型信息检测平台获取植物表型信息已成为农业领域研究热点且具有巨大潜力[3]。植物表型采集平台按照搭载方式可分为台式、传送带式、车载式、自走式、门架式、悬索式以及无人机式植物表型平台；按照使用环境分为室内植物表型平台和田间植物表型平台。大豆考种是对成熟后的大豆植株和籽粒性状调查的田间试验环节，目的是了解不同品种或不同处理因素对植株性状及籽粒的影响，从而分析掌握不同品种、不同处理因素的作用。其中，室内考种通常包括植株考种和籽粒考种两部分[4]。本文研究的是室内植株考种，其调查项目包括株高、有效分枝数、主茎、株型及全株荚数。

近年来，基于深度学习的目标检测在农业领域得到了广泛的应用，如大豆豆荚识别[5]、杂草的田间识别[6]、番茄关键器官识别[7－8]、猕猴桃图像识别[9－10]、蛋类识别[11]等。深度卷积神经网络[12－16]（Deep convolution neural network, DCNN）是深度学习中较为常用的网络模型，在目标检测方面的应用非常广泛并显现出巨大的优越性，它主要包括两大类：一类是基于区域生成的检测方法，即根据生成可能包含目标的候选区域采用卷积神经网络（Convolution neural network, CNN）对每个候选区域进行分类，如 RCNN[17]、Fast R－CNN[18]、Faster R－CNN[19]；另一类是基于回归的检测方法，采用 CNN 对整个图像进行处理，在实现目标定位的同时预测目标类别，如 YOLO[20]、YOLO v2[21]、SSD[22]。

LIU 等[22]对 Faster R－CNN、YOLO、SSD300、SSD512 进行了比较，YOLO 的实时性较好，但准确率相对 SSD 低。REDMON 等[20]将 YOLO v2 与 SSD 进行比较，它们有较相似的整体框架，但在低分辨率图像上 SSD 精度略高，SSD 对 GPU 资源占用率较低。SSD 通过 anchor boxes 机制，在不同尺度感受野的特征层上提取特征，回归计算得到目标的空间信息以及分类结果，进而提升准确率。然而，如果将 SSD 直接应用于大豆植株的豆荚与茎秆的定位与识别，其检测精度受到豆荚与豆荚、豆荚与茎秆之间相互遮挡影响较大，在检测重叠、遮挡等特殊情况时易出现漏检、误检等情况。针对传统 SSD 的缺陷，本文提出一种增加残差结构的 IM－SSD 模型，在网络的中低层部分引入残差结构，将低层的细节传到高层进行融合，调整基础网络结构，提高识别准确率。

在植物茎秆定位与识别研究中，深度学习算法常常需要与其它算法相结合来准确提取茎秆[23]，本文采用 IM－SSD 模型对大豆植株茎秆分段识别和定位，再结合蚁群优化（Ant colony optimization, ACO）算法提取大豆植株的完整茎秆，根据豆荚与茎秆的定位获取全株荚数、株高、主茎、有效分枝数与株型。

1　材料与方法

1.1　数据采集

试验大豆样本为收获期的大豆植株，包括东农 52、东农 251、东农 252 和东农 253 共 4 个品种，均来自于东北农业大学向阳农场试验基地，分别在 2017 年 9 月底和 2018 年 10 月初分两次采集。试验样本图像采集采用大豆植株图像表型检测平台，如图 1 所示，该检测平台采用佳能 5D Mark Ⅱ型相机搭配佳能 EF 24～105 mm 1∶4 L IS USM 型变焦镜头，垂直固定于距放置植株的采集平台 2 000 mm 处，照明系统采用高角度照明方式，为 DPX－400 型工作室闪光灯，指数为 GN66，色温为 5 500 K，输出量为 0.375，柔光箱长、宽、高为 1 200、800、450 mm。拍摄图像的分辨率为 5 616 像素×3 744 像素，焦点为 55 mm，光圈值 f/11，曝光时间 0.005 s，ISO 为 100，曝光补偿为 0。在植株表型检测中，叶片不作为检测目标，为避免叶片可能产生的误识别及遮挡等问题，样本采集前，将大豆植株上未脱落的叶片清理干净。同时为减少复杂背景的干扰，植株采集平台以蓝色无纺布为背景，采集时将大豆植株平放并将植

株子叶结与采集平台的子叶结基准线对齐。最后，批量剪切图像尺寸为 5 400 像素 ×2 700 像素，图像背景处理采用 RGB 彩色模型并根据每个像素的 R、G、B 分量的亮度来判定是否为背景，如果 B 分量的亮度大于其它两个分量，则视该像素为背景，如图 2 所示，图像保存为. jpg 格式。

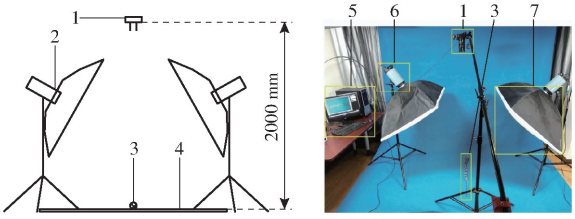


图 1 大豆植株数据采集装置

Fig. 1 Soybean plant data acquisition device

1. 相机与镜头 2. 照明系统 3. 大豆植株 4. 蓝色背景采集平台 5. 数据处理系统 6. 闪光灯 7. 柔光箱



图 2 处理后的整株大豆植株样本

Fig. 2 A sample of processed soybean plant

1.2 数据集标注

为实现对 IM - SSD 模型的训练，本文使用图像标注工具 Windows v1. 6. 0 对植株上的豆荚和茎秆分别进行定位及标记，豆荚为“pod”类，茎秆为“stem”类。由于豆荚与豆荚、豆荚与茎秆的相互遮挡，对豆荚和茎秆的识别和定位影响较大，而豆荚尖部的相互遮挡概率相对较小，为了减少相互遮挡的影响，在真实框标注中仅标注豆荚尖部即远离茎秆的 1/3 至 1/2 部分，具体标注范围视被遮挡情况而定。茎秆标注采用分段标注，仅标注未被遮挡的部分，因为未被遮挡的茎秆易于识别与定位，大豆植株标注如图 3 所示。

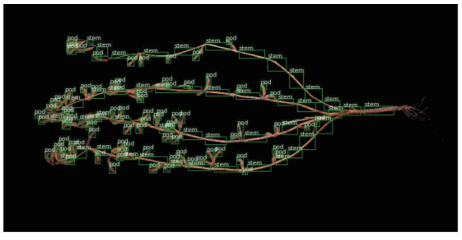


图 3 大豆植株标注

Fig. 3 Annotation of soybean plant

1.3 数据集处理

图像样本总量为 3 695 幅，按 7∶2∶1 分成训练集、测试集和验证集，训练集样本 2 600 幅、测试集样本 735 幅、验证集样本 360 幅。训练集和测试集分别用于训练和评估模型。在模型训练过程中，利用验证集对超参数进行优化。为保证样本集的均匀分布及评价的可靠性，所有样本集均为随机抽取。

为降低颜色变化、拍摄角度、生长状态等对识别的影响，提高检测精度[24-25]，实现模型的鲁棒性，通过以下几种方式对训练数据进行扩增。采用 RGB 颜色 3 通道对训练集图像进行随机调整，以提高图像颜色的多样性。对训练集图像进行水平、垂直和对角翻转、随机旋转和平移操作。将所有变换后的训练集样本进行整理，如果变换后样本图像的标注框超出边界，该标注框将被丢弃，如果变换后图像丢弃标注后没有标注信息，则删除图像。

1.4 IM - SSD 卷积神经网络模型

SSD 网络是在 VGG16[26] 网络的基础上进行修改，在 VGG16 网络结构上增加了 5 个特征提取层，同时分别在 6 个不同尺度的卷积层输出特征图，目的是对不同尺度的目标进行检测。传统 SSD 网络主要包括 SSD300 和 SSD512，其输入图像分别被压缩为 300 像素 ×300 像素、512 像素 ×512 像素，这种压缩比对于植株图像的压缩变形大，为降低图像变形带来的影响，IM - SSD 的输入图像被压缩为 300 像素 ×600 像素。另外，传统 SSD 在检测重叠、遮挡等情况下的豆荚时易出现漏检、误检。针对传统 SSD 的缺点，IM - SSD 在网络的中低层部分引入残差结构，将低层的细节传到高层进行融合，调整基础网络，提高识别率。

IM - SSD 结构在 SSD 深度卷积神经网络结构的基础上增加了 2 个卷积层：Add1 层，由 Maxpool2 和 Conv3_2 按位相加并归一化（Batch normalization，BN）[27]；Add2 层，由 Maxpool3 和 Conv4_2 按位相加并归一化，其结构如图 4 所示，红色的层为卷积核 3 ×3 的卷积层，绿色的层为卷积核 1 ×1 的卷积层，蓝色的层为最大池化层，步长为 2。增加残差结构的卷积层能够提高网络对小目标的检测能力，从而提升被遮挡或重叠豆荚的识别率。

IM - SSD 采用特征金字塔检测目标，分别对 Conv4_3 层、Conv7 层、Conv8_2 层、Conv9_2 层、Conv10_2 层和 Conv11_2 层生成比例不同的默认框，每一类共生成 17 236 个检测框，使用非极大值抑制（Non-maximum suppression，NMS）[28] 过滤重复的预测框。在多个特征图上同时进行识别和定位回归，损失值 L 由置信损失值（L_{conf}）和定位损失值（L_{loc}）的加权组成[22]，计算式为

$$L(x,c,l,g)=\frac{1}{N}(L_{conf}(x,c)+\alpha L_{loc}(x,l,g)) \quad (1)$$

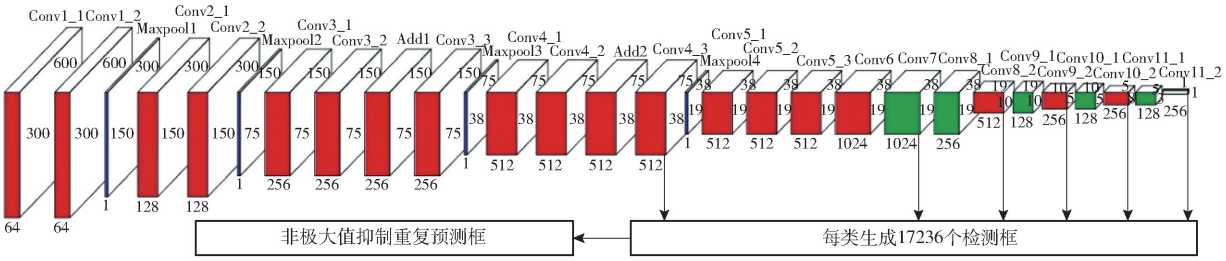


图 4　IM－SSD 卷积神经网络模型结构图

Fig. 4　Architecture of IM－SSD convolution neural network

其中　$L_{conf}(x,c) = -\sum_{i\in Pos}^{N} x_{ij}^{p}\lg\hat{c}_i^{p} - \sum_{i\in Neg}\lg\hat{c}_i^{0}$　（2）

$$L_{loc}(x,l,g) = \sum_{i\in Pos}^{N}\sum_{m\in\{c_x,c_y,w,h\}} x_{ij}^{p} f(l_i^m - \hat{g}_j^m) \quad (3)$$

$$\hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)} \quad (4)$$

式中　N——匹配的默认框数

x——类别的真实值

c——类别置信度的预测值

l——预测框位置

g——真实框位置

α——加权因子

x_{ij}^{p}——第 i 个预测框与第 j 个真实框是类别 p 则取值为 1，否则为 0

$\hat{c}_i^p$——类别 p 的第 i 个预测框的置信度

c_x——预测框中心 x 坐标

c_y——预测框中心 y 坐标

w——预测框宽度

h——预测框高度

$f(\cdot)$——平滑函数

N_{eg}——负样本集

$\hat{c}_i^0$——背景的第 i 个预测框的置信度

c_i^p——类别 p 的第 i 个真实框的置信度

P_{os}——正样本集

l_i^m——第 i 个位置为 m 的预测框

$\hat{g}_j^m$——第 j 个位置为 m 的真实框

置信度的损失值是多类置信度的 softmax 损失，定位损失是 l 与 g 之间的损失值。$f(x)$ 是一个分段平滑函数，用 σ 来控制平滑度。利用平移和尺度因子 d_i^{cx}、d_i^{cy} 得到实际标签 d_i^w、d_i^h 的近似回归预测框。

$$f(x) = \begin{cases} 0.5(\sigma x)^2 & \left(|x| < \dfrac{1}{\sigma^2}\right) \\ |x| - \dfrac{0.5}{\sigma^2} & \left(|x| \geqslant \dfrac{1}{\sigma^2}\right) \end{cases} \quad (5)$$

$$\begin{cases} \hat{g}_j^{cx} = (g_j^{cx} - d_i^{cx})/d_i^w \\ \hat{g}_j^{cy} = (g_j^{cy} - d_i^{cy})/d_i^h \end{cases} \quad (6)$$

$$\begin{cases} \hat{g}_j^{w} = \lg\dfrac{g_j^w}{d_i^w} \\ \hat{g}_j^{h} = \lg\dfrac{g_j^h}{d_i^h} \end{cases} \quad (7)$$

式中　g_j^{cx}——第 j 个真实框的 x 坐标

g_j^{cy}——第 j 个真实框的 y 坐标

σ——光滑度控制系数

d_i^{cx}——第 i 个预测框 x 轴方向的距离因子

d_i^{cy}——第 i 个预测框 y 轴方向的距离因子

d_i^w——第 i 个预测框宽度

d_i^h——第 i 个预测框高度

g_j^w——第 j 个真实框宽度

g_j^h——第 j 个真实框高度

$\hat{g}_j^{cx}$——第 j 个转换后预测框的 x 坐标

$\hat{g}_j^{cy}$——第 j 个转换后预测框的 y 坐标

$\hat{g}_j^{w}$——第 j 个转换后预测框的宽度

$\hat{g}_j^{h}$——第 j 个转换后预测框的高度

1.5　基于 ACO 算法的大豆茎秆提取

IM－SSD 对植株茎秆的定位是不连续的，不能完全反映真实的茎秆形态。本文提出了一种基于 ACO 算法的完整茎秆提取方法，该方法利用 IM－SSD 已定位茎秆的数据，以及茎秆的连续性特点，运用 ACO 算法获取最短路径将分段茎秆连接起来最终获得完整的茎秆。ACO 算法对路径搜索空间有一定要求，如果搜索空间大、复杂度高，蚂蚁易陷入非可行点或规划路径迂回等问题。通过 IM－SSD 的茎秆定位，根据已定位茎秆的坐标数据对二值图像进行剪切，为减小栅格地图的规模，将剪切图像分为 3 部分，如图 5 所示。3 部分图像运用等间隔像素采样的方法将图像压缩成 $z\times z$ 栅格并形成栅格地图，如图 6 所示。此外，每个部分都根据已定位茎秆的位置设置多个起点和终点，为降低算法的复杂度，将信息素存储在栅格中。

在 ACO 的栅格模型中采用二进制，其中“1”表示非植株栅格，“0”表示植株栅格。每个栅格按照从上至下、从左到右的顺序进行编号。

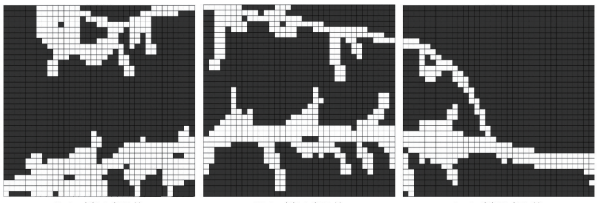
栅格地图与植株二值图像的二维坐标 (x',y')



(a) 第1部分 (b) 第2部分 (c) 第3部分

图 5 分割为 3 部分的大豆植株图像

Fig. 5 Three parts of soybean plant image



(a) 第1部分 (b) 第2部分 (c) 第3部分

图 6 分割为 3 部分的栅格地图

Fig. 6 Grid maps of three parts

转换式为

$$\begin{cases} x' = r(\mathrm{mod}(n,R) - 0.5) \\ y' = r\left(R + 0.5 - \mathrm{ceil}\left(\dfrac{n}{R}\right)\right) \end{cases} \tag{8}$$

式中 r——栅格包含的像素行数,列数和行数相等

n——栅格数目

R——当前行数

mod(·)——取余函数

ceil(·)——取整函数

ACO 是一种进化智能启发式搜索算法,用于寻找蚂蚁从居住地到食物源的最短路径[29-30]。蚂蚁 k 从栅格 i 移动到下一个栅格 j 是基于距离和未被访问栅格的信息素浓度。

迭代中,当每只蚂蚁到达终点或出现死锁时,搜索将停止。完成搜索后,保存所有到达终点蚂蚁的搜索路径长度。信息素浓度计算式为

$$\tau_{ij}(s_t + 1) = (1 - \rho)\tau_{ij}(s_t) + \sum_{k=1}^{m} \Delta\tau_{ij}^{k} \tag{9}$$

其中

$$\Delta\tau_{ij}^{k} = \begin{cases} \dfrac{Q}{L_k} & (\text{蚂蚁 } k \text{ 到达}(i,j)) \\ 0 & (\text{蚂蚁 } k \text{ 未到达}(i,j)) \end{cases} \tag{10}$$

$$L_k = \sum_{i=1}^{N} d_{O_iO_{i+1}} \tag{11}$$

式中 $\tau_{ij}(s_t + 1)$——下一时刻从栅格 i 到下一个栅格 j 的信息量

m——蚂蚁数量

ρ——信息挥发系数

s_t——当前时刻

$\Delta\tau_{ij}^{k}$——蚂蚁 k 从栅格 i 到下一个栅格 j 留下的信息量

Q——常量

L_k——蚂蚁搜索的路径长度

$d_{O_iO_{i+1}}$——蚂蚁一次移动的长度

$R_k = \{O_1, O_2, \cdots, O_{n+1}\}$ 是组成路径的一组栅格,$O_1 = S$,$O_{n+1} = E$。从初始点到终点有多条可选路径,路径规划可以得到一条最优或次优的可行路径,最短路径 L_{opt} 的数学模型为

$$L_{opt} = \min(L_k) \tag{12}$$

最短路径与大豆茎秆坐标转换公式为

$$R_{opt} = \min(R_k) \tag{13}$$

$$F(R_{opt}) = (k_1 O_{Ni}, k_2 O_{Nj}) \tag{14}$$

式中 O_{Ni}——横坐标值

k_1、k_2——横坐标转换系数

O_{Nj}——纵坐标值

R_{opt}——最优路径

F——路径转换函数

2 试验与结果分析

2.1 试验环境与参数

本研究深度学习运行硬件环境为 Intel(R) Core TM i7-9700K CPU,3.6 GHz,16 GB 内存,3 TB 硬盘,256 GB 固态硬盘,NVIDIA GTX1080GPU 显卡,操作环境为 Ubuntu 16.04 系统、Anaconda 4.2.0、Tensorflow 1.4.0 和 Keras 2.0.8。

由于采集的训练样本数量有限,不足以训练全新的网络,同时也为了减少训练时间,采用迁移学习[31-32]方法,使用 Pascal-VOC"07+12"[33] 20 类数据集的网络权重剪裁后对训练集进行训练。用随机梯度下降算法对网络分类层进行微调。加权因子 α 为 0.8 时,可降低局部损失因子对全局损失函数的影响。批处理量设置为 32,动量衰减设置为 0.9,初始学习率为 0.01,每 20 次迭代验证一次。根据平均损失的变化率,每次将学习率降低为原来的 1/10,直至网络不再收敛。

2.2 模型评价指标

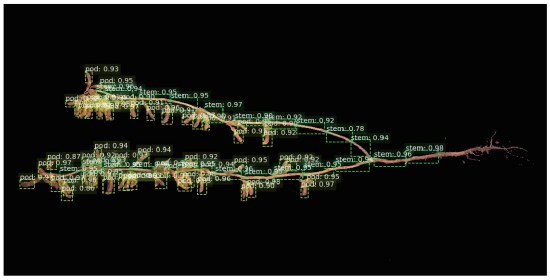
采用 3 个全局指标和 1 个局部指标来评估网络模型的稳定性和分类性能,分别为精确率(Precision,P)、假阴性率(False,F)、召回率(Recall,R)和交并比(Intersection over union,IoU)。其中,使用精确率和假阴性率来获得测试集中的性能指标,召回率代表预测框与所有真实框匹配的比例,IoU 评价每个目标检测结果的指标。

在确定默认框的情况下,正样本为 IoU 大于 0.4 的默认框。精确率是预测样本为阳性样本占所有阳性样本的比例。假阳性是指被误判为阴性的样本。假阴性率是指被误判的阳性样本占所有阳性样本的比例。

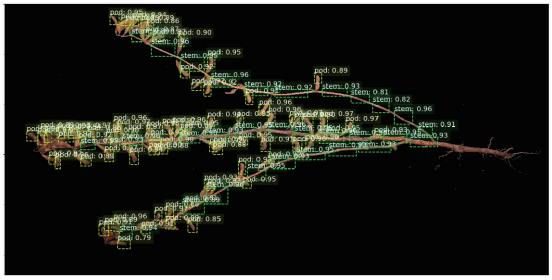
2.3 基于 IM-SSD 的豆荚和茎秆识别与定位

根据 2.1 节的模型参数,通过随机抽取训练样

本和验证样本的方式对 IM－SSD 模型进行训练和性能评估，以判断模型的可靠性和稳定性，经过 4 200 次迭代后收敛，其试验结果如图 7 所示。



(a) 大豆植株样本Ⅰ



(b) 大豆植株样本Ⅱ

图 7　IM－SSD 的检测结果

Fig. 7　Detection results of IM－SSD

大豆植株样本Ⅰ通过 IM－SSD 模型识别豆荚和茎秆的预测框分别为 52、21 个，如图 7a 所示。其中，在豆荚检测中有 4 个未被检出，在茎秆检测中多检出 2 个、少检出 1 个茎秆。大豆植株样本Ⅱ的豆荚和茎秆预测框分别为 75、31 个，如图 7b 所示。其中，在豆荚检测中有 7 个未被检出，在茎秆的检测中多检出 3 个、少检出 4 个茎秆。从试验结果可以看出，在豆荚的识别中，未重叠豆荚的识别率最高，相邻豆荚次之，重叠豆荚识别率较差，被遮挡豆荚识别率最差。当豆荚由于拍摄或摆放角度产生变形时，容易被漏检。因此，在全株豆荚数检测时，增加全株豆荚数补偿因子，全株豆荚数计算公式为

$$P_s = P_f + P_b \tag{15}$$

其中

$$P_b = 0.08 P_f \tag{16}$$

式中　P_f——IM－SSD 模型已检测到的豆荚数目

P_b——补偿因子

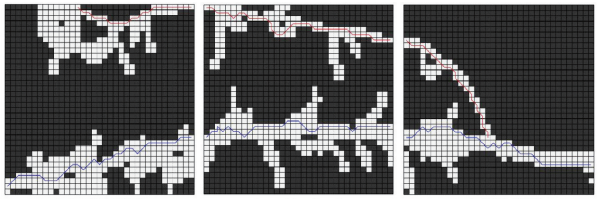
P_s——全株豆荚数

计算得到样本Ⅰ的全株豆荚数检测值为 56 个，样本Ⅱ的全株豆荚数检测值为 81 个。而在茎秆的识别过程中，中间部分多检或少检对茎秆的完整提取影响较小。IM－SSD 模型对豆荚和茎秆的检测精确率分别为 89.68% 和 81.99%，平均检测精确率为 85.84%，豆荚的识别精确率高于茎秆。另外，通过增加豆荚数补偿因子使全株豆荚数的统计精确率达到 92.76%。

2.4　基于 ACO 的大豆植株茎秆提取

在茎秆完整提取试验中，为了验证 ACO 算法提取大豆茎秆的有效性，以其中一个验证集样本为例，将其分为 3 部分，3 部分压缩尺寸均以 35×35 的栅格作为蚁群规划地图。参数设置为：种群规模 n = 50，最大迭代 T = 200，α' = 1，β' = 7，ρ = 0.3，Q = 1。

在蚁群地图中，将已定位栅格的最左侧中“0”栅格作为起始栅格 S，将最右侧与起始栅格相对应的“0”栅格作为结束栅格 E，分别对植株的每个分枝进行基于蚁群算法的路径规划，并根据路径规划结果获得完整的植株茎秆定位，图 8 为最优路径规划结果，图 9 为最优路径的收敛曲线。最终获得的完整大豆植株茎秆的预测结果如图 10 所示。



(a) 第1部分　(b) 第2部分　(c) 第3部分

图 8　最优路径规划结果

Fig. 8　Results of optimal path planning

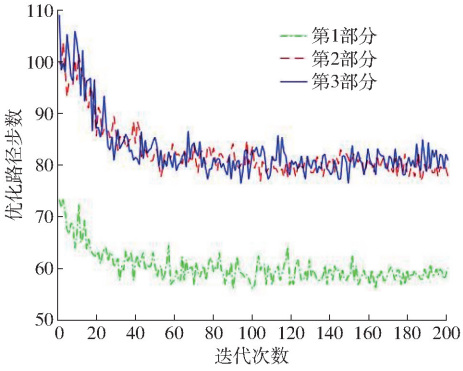


图 9　最优路径的收敛曲线

Fig. 9　Optimal path convergence curve

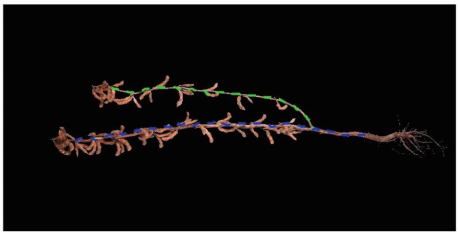


图 10　大豆茎秆的预测结果

Fig. 10　Prediction results of soybean stem

基于 ACO 算法的大豆植株茎秆提取同样适用于具有较复杂分枝结构的大豆植株，如图 11 所示。有效分枝数是指成熟期植株上有两个以上茎节并有一个以上成熟豆荚的第 1 级分枝数。试验中通过拟合的分枝个数确定植株的有效分枝数。在多分枝大豆植株中，与其它分枝相比主茎较粗，将平均茎粗最大的茎作为主茎，用不同颜色标出不同的分枝，其中蓝色虚线标出主茎，图 10 共 2 个有效分枝、图 11 共 4 个有效分枝。株型的判断是通过成熟期下部分枝

与主茎的收敛程度确定,如图 12 所示,下部分枝的生长方向与主茎的自然夹角为 α_1、α_2、α_3,当所有夹角的平均值小于 30°则为收敛型,大于 30°小于 60°为半开张型,大于 60°为开张型,其中图 10 为收敛型,图 11 为半开张型植株。株高是成熟期从子叶节到植株生长点的高度,如图 12 所示,株高的计算是通过最顶端茎秆像素点到子叶结基准线的垂直距离乘以 0.05,其株高的单位为 mm,其中图 10 的株高为 992 mm,图 11 的株高为 875 mm。

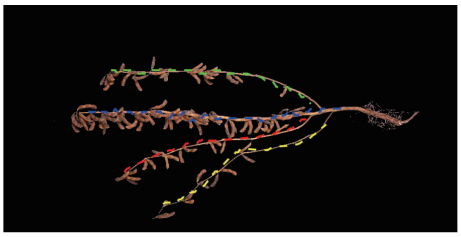


图 11 复杂分枝结构的茎秆预测结果

Fig. 11 Soybean stem prediction of complex branching structure

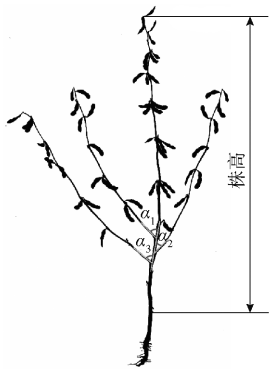


图 12 株型及株高

Fig. 12 Plant type and height

2.5 不同网络模型的对比试验

利用大豆植株训练集对 3 种网络模型(SSD300、SSD512、IM - SSD)的性能进行评价,所有默认框的 IoU 大于 0.4 即为正样本。经过 4 200 次迭代,损失曲线如图 13 所示。在相同的网络参数下,SSD300 模型的网络训练初始损失误差小,收敛速度快。而 SSD512 模型的初始损失误差最大,收敛速度较快,但 IM - SSD 模型在 3 种模型中经过 2 000 次迭代后损失值最小。

将 3 种模型与植株真实值进行对比,其样本均来自于验证集,其中 3 种模型分别对豆荚和茎秆通过分段定位进行预测。由于人工检测可能存在主观误差,为获得植株的真实值,采用多次人工选荚和人工框选取茎秆的方法,即对所有验证集样本进行 5 次重复检测,真实值为 5 次观测选荚的数量结果中出现概率最多的数值,茎秆框选真实位置为 5 次框选中IoU最大的一次框选,结果如表1所示。从试

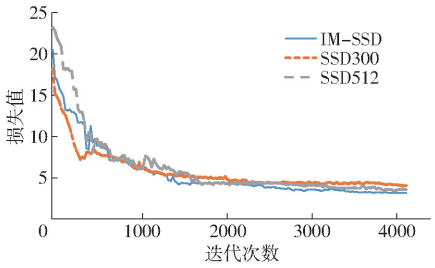


图 13 训练样本损失值曲线

Fig. 13 Loss value curves of training samples

表 1 3 种模型与人工检测结果的比较

Tab. 1 Comparison of three models and manual detection results %

检测方法	类别	*P*	*F*	*R*	F1	平均 *P*
SSD300	豆荚	83.49	6.90	93.10	88.29	78.05
	茎秆	72.61	6.59	93.41	83.01	
SSD512	豆荚	87.52	5.57	94.43	90.98	82.01
	茎秆	76.49	5.39	94.61	85.55	
IM - SSD	豆荚	89.68	6.55	93.45	91.57	85.84
	茎秆	81.99	4.95	95.05	88.52	
人工检测	豆荚	95.23	2.58	97.42	96.32	95.72
	茎秆	96.21	2.50	97.50	96.86	

验结果可以看出,人工对豆荚和茎秆的检测与定位的精确率最高,但人力和时间耗费较大,而文中模型通过样本训练后可自动对豆荚及茎秆定位,无需人工操作。另外,与 SSD300 网络和 SSD512 网络相比,IM - SSD 平均检测精确率分别提高了 7.79、3.83 个百分点,在 3 种模型中 IM - SSD 模型对大豆植株豆荚与茎秆的检测更稳定、有效。

2.6 基于 ACO 算法的茎秆表型特征数据分析

由 IM - SSD 模型得到的植株茎秆定位是不连续的非完整茎秆。本文提出了基于 ACO 算法并运用 IM - SSD 获得茎秆定位来提取完整茎秆的方法能够准确定位茎秆,并弥补 IM - SSD 的不足。同时,根据完整茎秆定位,提取出大豆植株的株高、有效分枝数和株型。将人工 5 次重复测量株高及有效分枝数、株型中出现次数最多的值作为真实值。人工检测的单次株高、有效分枝数和株型的精确率分别为 98.65%、99.12%、93.21%,运用完整茎秆定位提取的精确率分别为 95.56%、96.12%、90.28%,其精确率均略低于人工测量。株型判断的准确率较低的原因是受拍摄角度的影响,单一角度很难准确获取复杂结构的植株分支夹角。人工检测虽然有较高的精确率但消耗大量的人力物力且劳动强度高、时间周期长、对数量级也有一定要求,而本文检测模型通过样本训练后可自动对豆荚及茎秆定位,无需人工操作,减少了主观误差、降低了劳动强度、缩短了检测周期,适于数据量较大的样本检测,

更具有客观性，精确率均在 90% 以上，基本满足植株表型平台自动化的要求。

3　结论

（1）提出了一种基于 IM－SSD 模型的大豆植株豆荚与茎秆表型信息检测方法。试验结果表明，该模型具有较高的检测精度，对遮挡和重叠干扰有较强的鲁棒性。

（2）在整株大豆植株样本训练集上，对已标注训练样本进行随机处理，实现图像扩增，提高网络的特征提取能力。在真实框标注时，只标注一部分特征区域，有效降低了相互遮挡产生的误判率。

（3）提出了基于 SSD 结构的残差网络，将低层特征图融合到高层特征图，增强对小目标检测能力，提高网络的识别率。与 SSD300 网络和 SSD512 网络相比，IM－SSD 平均检测精确率分别提高了 7.79、3.83 个百分点。同时，增加全株豆荚数补偿，使全株豆荚数统计的精确率提高到 92.76%。IM－SSD 模型与传统 SSD 相比，检测精度得到了提高，同时也为其他植株检测的研究提供了思路。

（4）采用 ACO 算法弥补 IM－SSD 分段提取茎秆的不足，从而获取大豆植株的株高、有效分枝数和株型。

参　考　文　献

[1] TESTER M, LANGRIDGE P. Breeding technologies to increase crop production in a changing world[J]. Science, 2010, 327(5967): 818－822.

[2] RIBAUT J M, VICENTE M D, DELANNAY X. Molecular breeding in developing countries: challenges and perspectives[J]. Curr. Opin. Plant Biol., 2010, 13(2): 213－218.

[3] 张慧春, 周宏平, 郑加强, 等. 植物表型平台与图像分析技术研究进展与展望[J/OL]. 农业机械学报, 2020,51(3): 1－17.
ZHANG Huichun, ZHOU Hongping, ZHENG Jiaqiang, et al. Research progress and prospect in plant phenotyping platform and image analysis technology[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020,51(3): 1－17. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20200301&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2020.03.001. (in Chinese)

[4] 邱丽娟. 大豆种质资源描述规范和数据标准[M]. 北京:中国农业出版社, 2006.

[5] 闫壮壮, 闫学慧, 石嘉, 等. 基于深度学习的大豆豆荚类别识别研究[J]. 作物学报, 2020, 46(11): 1771－1779.
YAN Zhuangzhuang, YAN Xuehui, SHI Jia, et al. Classification of soybean pods using deep learning[J]. Acta Agronomica Sinica, 2020,46(11): 1771－1779. (in Chinese)

[6] 姜红花,张传银,张昭,等. 基于 Mask R－CNN 的玉米田间杂草检测方法[J/OL]. 农业机械学报,2020,51(6):220－228,247.
JIANG Honghua, ZHANG Chuanyin, ZHANG Zhao, et al. Detection method of corn weed based on Mask R－CNN[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2020,51(6): 220－228, 247. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20200623&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2020.06.023. (in Chinese)

[7] YUAN Ting, LÜ Lin, ZHANG Fan, et al. Robust cherry tomatoes detection algorithm in greenhouse scene based on SSD[J]. Agriculture, 2020, 10(5): 160.

[8] 周云成,许童羽,邓寒冰,等. 基于双卷积链 Fast R－CNN 的番茄关键器官识别方法[J]. 沈阳农业大学学报,2018,49(1):65－74.
ZHOU Yuncheng, XU Tongyu, DENG Hanbing, et al. Recognition method of tomato key organs based on dual convolution Fast R－CNN[J]. Journal of Shenyang Agricultural University, 2018, 49(1):65－74. (in Chinese)

[9] FU L, WANG B, CUI Y, et al. Kiwifruit recognition at nighttime using artificial lighting based on machine vision[J]. International Journal of Agricultural and Biological Engineering, 2015, 8(4): 52－59.

[10] 傅隆生,冯亚利,ELKAMIL Tol,等. 基于卷积神经网络的田间多簇猕猴桃图像识别方法[J]. 农业工程学报,2018,34(2):205－211.
FU Longsheng, FENG Yali, ELKAMIL Tol, et al. Image recognition method of multi-cluster kiwifruit in field based on convolutional neural networks[J]. Transactions of the CSAE,2018, 34(2):205－211. (in Chinese)

[11] 李庆旭,王巧华,肖仕杰,等. 基于深度学习的群体种鸭蛋受精信息检测方法[J/OL]. 农业机械学报,2021,52(1):193－200.
LI Qingxu,WANG Qiaohua,XIAO Shijie, et al. Detection method for fertilizing information of group duck eggs based on deep learning[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(1):193－200. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20210122&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2021.01.022. (in Chinese)

[12] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436－444.

[13] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//

International Conference on Neural Information Processing Systems, 2012: 1097 - 1105.
[14] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2016: 770 - 778.
[15] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1 - 9.
[16] HUANG Gao, LIU Zhuang, WEINBERGER K Q, et al. Densely connected convolutional networks[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2017.
[17] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580 - 587.
[18] GIRSHICK R. Fast R - CNN[C]//IEEE International Conference on Computer Vision, 2015: 1440 - 1448.
[19] REN S, HE K, GIRSHICK R, et al. Faster R - CNN: towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2015,28:91 - 99.
[20] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2016:779 - 788.
[21] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6517 - 6525
[22] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//European Conference on Computer Vision. Springer, Cham, 2016: 21 - 37.
[23] 刘慧力,贾洪雷,王刚, 等. 基于深度学习与图像处理的玉米茎秆识别方法与试验[J/OL]. 农业机械学报,2020,51(4): 207 - 215.
LIU Huili, JIA Honglei, WANG Gang, et al. Stems recognition based on deep learning and image processing[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2020,51(4):207 - 215. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no = 20200424&flag = 1&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2020.04.024. (in Chinese)
[24] 张德丰. 数字图像处理[M]. 北京:机械工业出版社, 2012.
[25] 吴娱. 数字图像处理[M]. 北京:北京邮电大学出版社, 2017.
[26] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//International Conference on Learning Representations, 2015:1 - 14.
[27] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//International Conference on Machine Learning. PMLR, 2015: 448 - 456.
[28] NEUBECK A, GOOL L J V. Efficient non-maximum suppression[C]//International Conference on Pattern Recognition. IEEE Computer Society, 2006: 850 - 855.
[29] KHALED A, FARID K. Mobile robot path planning using an improved ant colony optimization[J]. International Journal of Advanced Robotic Systems, 2018, 15(3):1 - 5.
[30] WANG L. Path planning for unmanned wheeled robot based on improved ant colony optimization[J]. Measurement and Control, 2020(4):1014 - 1020.
[31] WEISS K, KHOSHGOFTAAR T M, WANG D D. A survey of transfer learning[J]. Journal of Big Data, 2016, 3(1):1 - 40.
[32] PAN S J, QIANG Y. A survey on transfer learning[J]. IEEE Trans. Knowledge Data Engin., 2010, 22(10): 1345 - 1359.
[33] EVERINGHAM M, GOOL L V, WILLIAMS C K I, et al. The Pascal visual object classes (VOC) challenge[J]. International Journal of Computer Vision, 2010, 2: 303 - 338.