doi:10.6041/j.issn.1000-1298.2020.08.004

基于 YOLOv3 目标检测的秧苗列中心线提取方法

张　勤[1]　王家辉[1]　李　彬[2]

(1. 华南理工大学机械与汽车工程学院, 广州 510641; 2. 华南理工大学自动化科学与工程学院, 广州 510641)

摘要: 针对秧苗列中心线的检测结果易受到水田中的浮萍、蓝藻以及水面反射、风速、光照情况等自然条件影响的问题,提出一种基于 YOLOv3 目标检测的秧苗列中心线检测算法。基于透视投影计算提取图像的 ROI(Region of interest)区域,采用 ROI 图像构建数据集,对 YOLOv3 模型进行训练,训练过程中通过减少 YOLOv3 模型的输出降低运算量,利用模型识别定位 ROI 内的秧苗,并输出其检测框,对同列秧苗的检测框进行自适应聚类。在对秧苗图像进行灰度化和滤波处理后,在同类检测框内提取秧苗 SUSAN(Smallest univalue segment assimilating nucleus)角点特征,采用最小二乘法拟合秧苗列中心线。试验结果表明,该算法对于秧苗的不同生长时期,以及在大风、蓝藻、浮萍和秧苗倒影、水面强光反射、暗光线的特殊场景下均能成功提取秧苗列中心线,鲁棒性较好,模型的平均精度为 91.47%,提取的水田秧苗列中心线平均角度误差为 0.97°,单幅图像(分辨率 640 像素 ×480 像素)在 GPU 下的平均处理时间为 82.6 ms,能够满足视觉导航的实时性要求。为复杂环境下作物中心线的提取提供了有效技术途径。

关键词: 水田秧苗; 中心线提取; 深度学习; ROI 图像; 自适应检测框聚类

中图分类号: TP24　　**文献标识码**: A　　**文章编号**: 1000-1298(2020)08-0034-10　　**OSID**:

Extraction Method for Centerlines of Rice Seedings Based on YOLOv3 Target Detection

ZHANG Qin[1]　WANG Jiahui[1]　LI Bin[2]

(1. *School of Mechanical and Automotive Engineering*, *South China University of Technology*, *Guangzhou* 510641, *China*
2. *School of Automation Science and Engineering*, *South China University of Technology*, *Guangzhou* 510641, *China*)

Abstract: In order to extract the centerlines of rice seedlings, a new method based on YOLOv3 target detection algorithm was presented, which can extract centerlines of different growth stages of rice seedlings in complex paddy field so as to provide guide lines for autonomous navigation of robot. Firstly, an industrial camera which was 1 m high above the ground with pitch angles of 45° to 60° was used to capture image of rice seedlings, and then the region of interest (ROI) of the crop image was determined in order to find the instructive guide lines. Because of the perspective projection, the rice seedlings rows were labeled in segments. Then, the ROI images dataset was built to train YOLOv3 model. After that, the best YOLOv3 model was used to detect the rice seedling in the ROI and output bounding boxes. Secondly, the bounding boxes of the same rice seedlings row was clustered. Thirdly, image segmentation was applied and the smallest univalue segment assimilating nucleus (SUSAN) feature points was extracted within the bounding box of the same cluster. Finally, the least square method was applied in the algorithm to extract the centerlines of rice seedling. For complex paddy field environment such as windy weather, dark light, rice seedlings shadow and light reflection on water surface, as well as the impacts like duckweed and cyanobacteria, the proposed algorithm successfully and accurately extracted the centerlines of rice seedlings. For 200 test images, the mean average precision of trained network reached 91.47%, the mean average angle errors of the extracted centerlines was 0.97° and the average runtime of one image (resolution: 640 pixels ×480 pixels) was 82.6 ms. Compared with another method for centerlines extracting, this algorithm had higher robustness, higher accuracy and faster runtime. The result showed that the method was real time and had application values.

Key words: rice seedlings; centerlines extraction; deep learning; image of region of interest; bounding boxes clustering

收稿日期: 2019-12-09　修回日期: 2020-01-05

基金项目: 广东省省级科技计划项目(2014A020208018)

作者简介: 张勤(1964—),女,教授,主要从事机器人及其应用研究,E-mail: zhangqin@scut.edu.cn

通信作者: 李彬(1979—),男,副教授,主要从事科学计算可视化、图像处理与模式识别研究,E-mail: binlee@scut.edu.cn

0　引言

在复杂的农田环境中，实时、准确提取机器人的导航线是实现机器人自主导航的重要保证[1-4]。由于作物个体差异、作物生长变化、叶片相互遮挡、光照不确定性等因素，经常导致图像中存在噪声和干扰，降低作物列中心线检测的准确率。在水田环境下，还经常会伴有大风、蓝藻、浮萍、秧苗倒影、水面反射、光照不均等环境因素的干扰，这些进一步增大了图像处理的难度，使水田秧苗列中心线的实时检测更为复杂，从而严重影响水田除草机器人自主导航的精度和效率。

为了解决这些问题，国内外学者做了大量研究工作，并取得了一定的成果。针对农田环境下自然光照变化对作物列中心线提取的影响，文献[5－7]选用不同的彩色模型，通过提取作物的颜色特征，有效地降低了光照变化对农作物图像分割的影响。考虑作物生长过程中形态变化造成的影响，文献[8]通过 8 邻域结构因子，对二值化后作物图像中的斑块噪声进行标记，剔除了作物由于生长变化形成的叶子尖端大面积噪声；文献[9]利用水稻秧苗的形态特征识别秧苗列的中心线，该方法对不同生长周期的秧苗具有较好的识别效果，解决了秧苗生长过程中叶片交叉遮挡问题；文献[10]提出一种改进遗传算法的作物行识别方式，从图像的顶部和底部随机选择 2 个点组成的直线编码为染色体，通过多次遗传进化，选择适应度最高的染色体作为作物的中心线，较好地提取到不同生长阶段的作物列中心线；文献[11]提出一种作物特征点提取方法，将作物图像分割成不同高度的 12 个子带，设置与子带等高度的窗口，对二值化后的白色作物像素进行跟踪，并求窗口内作物像素质心点，即为拟合作物线的特征点，该方法提取的特征点能较好地表征不同生长阶段的作物中心位置。在减小环境噪声的影响方面，文献[12]通过匹配水田中作物的红外图像和 R 分量图像，对其进行相减和二值化处理，减小了水面反射噪声对作物列中心线检测的干扰；文献[13]通过对农田图像灰度化，再利用小面积法进行图像滤波，去除了农田中土块等引起的颗粒状噪声；文献[14]采用双阈值的 Otsu 方法进行图像分割，成功剔除了作物田间的杂草噪声。

目前，针对南方复杂恶劣水田环境下的秧苗中心线提取，国内外相关研究文献很少。为了解决这个问题，文献[15－17]提出基于彩色模型和近邻法聚类的水田秧苗列中心线检测方法、基于 SUSAN 角点的秧苗列中心线提取算法，这些方法在图像中存在浮萍、蓝藻和秧苗倒影的情况下具有较高的鲁棒性，能够满足视觉导航的实时性要求，但对环境变化的适用性还有待进一步研究。

近年来，深度学习在农业领域的应用日益广泛，深度学习方法具有检测精度高、适应性强的特点[18-19]。基于深度学习的方法主要分为两类：一类是通过提取候选区域，并对区域进行以深度学习为主的分类方案，称为 two-stage（两阶段）检测算法，典型的如 RCNN[20]、Fast RCNN[21]、Faster RCNN[22]等。另一类则是基于回归的方法，直接通过单个卷积深度神经网络遍历图像，回归出目标的类别和位置，称为 one-stage（一阶段）检测算法，典型算法如 YOLO[23]、YOLOv2[24]、YOLOv3[25]、SSD[26]。相对于 two-stage 检测算法，one-stage 检测算法检测速度更快，YOLOv3 深度神经网络还通过借鉴残差网络结构（ResNet）[27]和 SSD 中的多尺度预测，提升了模型的表现，在保证一定模型精度的同时，提升了检测速度。

本文提出基于 YOLOv3 目标检测的秧苗列中心线提取方法，采用 YOLOv3 模型对视频中感兴趣区域（Region of interest，ROI）内的秧苗进行识别定位，输出秧苗的检测框，然后在检测框区域内进行秧苗的特征提取和中心线的拟合。

1　材料与方法

1.1　算法概述

基于 YOLOv3 检测框区域的水田秧苗列中心线提取算法主要分为目标检测定位和中心线提取两部分，整体流程如图 1 所示。

（1）YOLOv3 目标检测：首先对于采集到的水田秧苗 RGB 图像，根据摄像头俯仰角与秧苗成像关系分析，构建图像中的 ROI，然后对 ROI 区域内的秧苗进行分段标注后，构建 YOLOv3 的训练数据集，寻找训练结束后最优的 YOLOv3 检测模型，利用模型定位秧苗并输出其检测框。

（2）基于同列秧苗检测框区域内的秧苗列中心线提取：对 YOLOv3 模型检测并输出的同列秧苗检测框进行自适应聚类；再基于归一化的 ExG（Excess green index）算子进行图像分割；在同类检测框区域内进行秧苗 SUSAN 角点的提取，最后利用最小二乘法对秧苗特征点进行中心线拟合。

1.2　数据采集

试验所用的图像均拍摄于华南农业大学国家优质稻新品种扩繁基地。图像分辨率为 1 280 像素 × 960 像素。拍摄过程中摄像头距离地面 1 m，摄像头光轴与水平线夹角为 45° ~ 60°。若小于 45°则图像

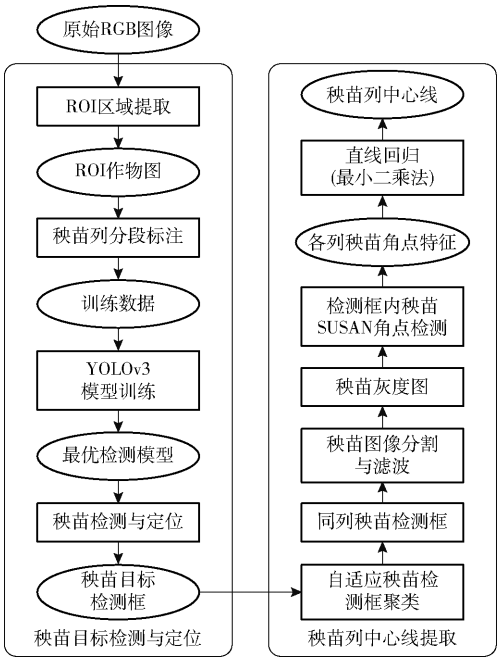


图 1 基于 YOLOv3 目标检测的秧苗列中心线提取方法流程图

Fig. 1 Flow chart of extraction method of seedling row centerline based on YOLOv3 target detection

中拍摄的秧苗列数过多,导致秧苗列在成像时距离过于密集,不利于后续导航线的检测,而如果大于 60°,则会使得拍摄到的秧苗列在图像中的长度过短,提取到的导航线与实际秧苗列中心线误差较大。从秧苗的 4 个生长阶段各采集了 500 幅图像,共采集到 2 000 幅图像,其中包括 5 种复杂场景(大风、蓝藻、浮萍和秧苗倒影、水面强光反射、暗光线)下的图像样本各 200 幅。从采集到的 4 个生长阶段的秧苗图像样本中随机选取各 400 幅图像用于制作训练集(1 200 幅)、验证集(200 幅)和测试集(200 幅),其中验证集只用于评估训练时模型的表现。

2 基于 YOLOv3 的秧苗目标检测和定位

2.1 YOLOv3 模型框架

YOLOv3 将目标检测转换为回归问题,基于一个端到端的网络,实现从原始秧苗图像的输入到最后秧苗检测框的输出。YOLOv3 模型先把秧苗输入图像统一为 416 像素 × 416 像素,然后再划分成 13 像素 × 13 像素的网格,每个网格负责检测 B 个边界框及其置信度,以及 C 个类别的概率。所以,每个边界框都包含了 5 个预测量:x、y、w、h、$C_{\text{onfidience}}$,其中(x,y)是秧苗检测框的左上角坐标,w、h 是秧苗检测框的宽、高,$C_{\text{onfidience}}$是秧苗的置信度。整个步骤通过输入秧苗图像经过一次 YOLO 网络,可以得到图像中秧苗的位置及其相应的置信概率。

YOLOv3 模型结构如图 2 所示,考虑到水田环境的复杂和多变容易导致秧苗识别困难,本文使用含有 53 个卷积层的 Darknet-53 作为特征提取网络来提取水田秧苗的特征。如图 2 所示,输入网络的图像,经过 Darknet-53 后输出 52 像素 × 52 像素(尺度 3)、26 像素 × 26 像素(尺度 2)、13 像素 × 13 像素(尺度 1)共 3 种尺度的特征图,每个特征图分配 3 种不同尺寸的锚点(Anchor)进行边界框预测,最后对 3 个输出张量进行合并处理,获得秧苗的预测信息。由于在秧苗识别过程中需要检测的目标只有秧苗这一类,所以通过修改 YOLOv3 的输出类别数目为 1,减少了网络的运算量,最终网络预测张量的长度为 18(即 3(锚点) × (4(边框坐标) + 1(边框置信度) + 1(对象类别数))),如图 2 中的输出张量所示。

2.2 基于透视投影计算的秧苗图像 ROI 提取

在机器人实际导航过程中,机器人的行进列

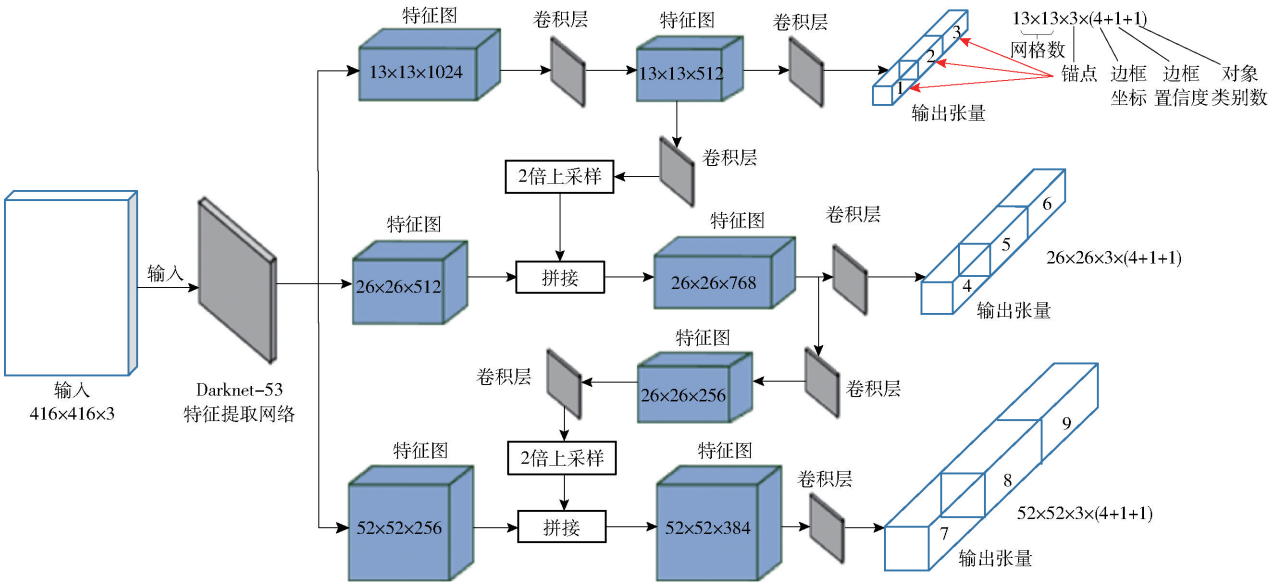


图 2 YOLOv3 模型结构图

Fig. 2 YOLOv3 model architecture

(一般是图像中部的秧苗列),以及距离机器人行进列最近的左右两侧秧苗列直接影响机器人导航的准确度。因此可以在图像中事先筛选出 ROI 区域,这样可以减少图像数据处理量,同时也减少图像边缘作物列的干扰,使得后续的图像处理步骤更加高效[28-30]。因此需要在训练 YOLOv3 网络前,构建秧苗图像的 ROI。然后通过标注后再对 YOLOv3 网络进行训练。

为了确定图像中秧苗列的 ROI,将图像在透视投影下简化为成像模型如图 3 所示,NM 为图像中部秧苗列中心线,EF、PQ 为左、右侧秧苗列中心线,四边形 $ACDB$ 内为 ROI 区域。

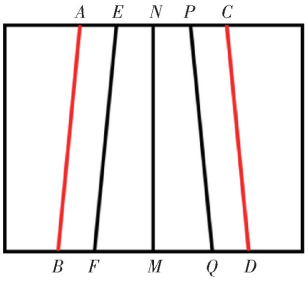


图 3　透视现象下的秧苗列成像模型

Fig. 3　Image model of rice seedlings rows under perspective

ROI 具体确定流程如下:

(1)在透视投影下,秧苗列中心线上下两端的像素距离 D_{NP} 与 D_{MQ} 有如下关系[31]

$$G_p=\frac{D_{NP}}{D_{MQ}}=\frac{\sin\left(\theta-\frac{\alpha}{2}\right)}{\sin\left(\theta+\frac{\alpha}{2}\right)} \tag{1}$$

式中　G_p——D_{NP} 与 D_{MQ} 的比值

θ——相机光轴与水平线夹角,45°~60°

α——相机的垂直视场角,(°)

(2)由摄像头成像[32]可得 M 点与 Q 点之间的像素距离为

$$D_{MQ}=\frac{2fd\sin\theta}{H(1+G_p)k} \tag{2}$$

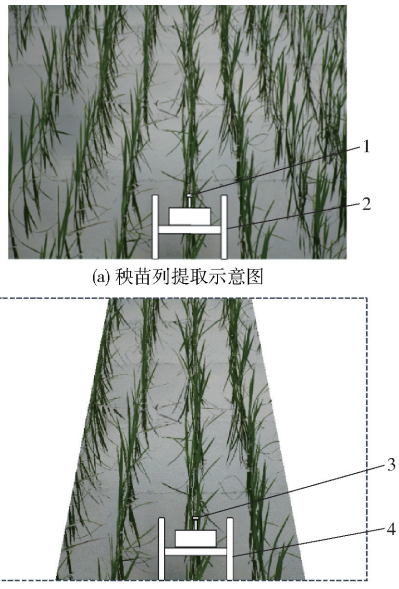
式中　d——秧苗列间的平均距离,mm

f——摄像头焦距,mm

H——摄像头光心离地高度,mm

k——图像中单个像素点的物理尺寸,mm

(3)在相机俯角最大($\theta=60°$)时,点 M 与点 Q 的像素距离 D_{MQ} 最大,把 $\theta=60°$ 代入式(1)、(2)可求得 D_{MQ}、D_{NP}。以此时的 D_{MQ}、D_{NP} 作为参考值,设置四边形 $ABDC$ 内的区域为 ROI,对图像非 ROI 的部分进行遮罩填充。其中 $D_{MD}=D_{MB}=cD_{MQ}$,$D_{NC}=D_{NA}=cD_{NP}$,c 为裕度系数($c=1.2$)。图 4a 的 ROI 提取效果如图 4b 所示。



(a) 秧苗列提取示意图

(b) 秧苗图像ROI的划分

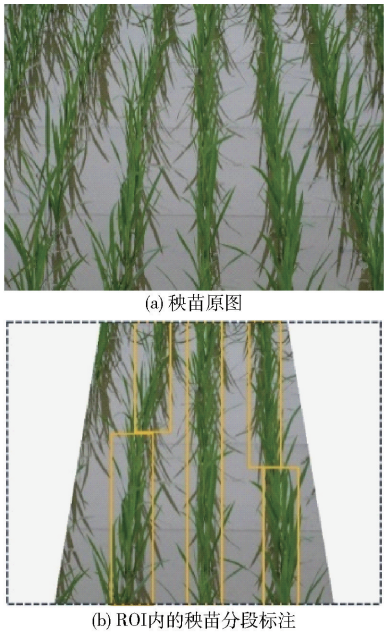
图 4　ROI 内的秧苗列

Fig. 4　Rice seedlings rows in ROI

1、3. 摄像头　2、4. 除草机器人

2.3　秧苗图像数据集的建立

在划分了秧苗图像的 ROI 后,再对 ROI 内的秧苗进行标注。由于在透视现象下,除了图像中部的秧苗列外,其他秧苗列成像时均呈现较大倾斜程度。此时在进行图像标注时,如果对这些倾斜程度较大的秧苗列进行整列标注,则会使得标注框内背景区域较多,不利于后续针对检测框区域进行秧苗中心线的提取。因此在保证背景区域较小的前提下,采用分段标注秧苗的方式,使得 YOLOv3 网络能更准确地识别与定位出秧苗列,有利于后续图像处理过程中减少背景噪声的干扰。标注方式如图 5 所示,



(a) 秧苗原图

(b) ROI内的秧苗分段标注

图 5　ROI 内的秧苗分段标注

Fig. 5　Segmental labeling of rice seedlings rows in ROI

即在标注框所含背景区域较少的情况下,标注出秧苗的主干部分即可,由图 5b 可以看出标注框与秧苗主体较为紧凑。

标注后的图像采用 YOLO 数据集的格式保存为 txt 文件用于网络训练。每个 txt 文件内包含秧苗的类别数量(C)和标注秧苗边界框的左上角坐标信息(x,y)以及框的宽与高(w,h),在本文中只需要检测秧苗这一类目标,所以本文类别总数 C 取 1。

2.4 模型参数的设置与训练

训练时一轮迭代(Iteration)处理 64 个样本,图像训练前被降采样到 640 像素 ×480 像素,使用 BN(Batch normalization)在每次权值更新时进行正则化,权值衰减(Decay)设置为 0.000 5,动量衰减(Momentum)设置为 0.9,初始学习率(Learning rate)设置为 0.001,并在迭代 4 000、6 000 次时衰减 10 倍,为了在初始迭代阶段使得模型迭代较为稳定,设置一个学习率变化点 $b_{\text{urn_in}}$ 的值为 1 200,在前 1 200 轮迭代过程中,学习率计算式为

$$l_{\text{rate}} \Leftarrow l_{\text{rate}} \left(\frac{i_{\text{teration}}}{b_{\text{urn_in}}}\right)^4 \tag{3}$$

式中 l_{rate}——学习率

i_{teration}——迭代次数

⇐——赋值运算符

在训练前通过 K-means 聚类算出训练集中锚点的初始值,本文中设置 9 个锚点分别应用于 3 个尺度的特征图,9 个锚点的尺寸分别为 35 ×104、35 ×174、49 ×181、39 ×244、51 ×266、62 ×221、70 ×290、46 ×456、67 ×461。

对于深度神经网络训练所需的数据量而言,1200 幅图像远远不足。所以本文在训练的时候采用数据增强方法,通过改变图像的色调(变化范围 1 ~1.5 倍)、曝光度(变化范围 1 ~1.5 倍)、色量化(变化范围 0.8 ~1.1 倍),设置数据抖动值,即随机改变输入图像尺寸和宽高比值(变化范围为[0.4,1.6])来扩充训练集。在每一轮迭代过程中,随机使用上述一种方法应用于 64 幅图像样本。

最终模型共训练 9 000 次,使用了 576 000 幅图像,其损失值变化如图 6 所示。

从图 6 可以看出,在最初的 1 200 次迭代中拟合速度较快,在达到 4 000 次迭代时由于此时学习率衰减 10 倍,模型的训练步长减小,拟合能力上升,损失值进一步下降。在 6 000 次迭代时学习率再衰减 10 倍后,由于此时模型逐渐稳定,所以损失值振荡较小,最终损失值稳定在 0.42。

2.5 最优模型的选择

在 9 000 次训练迭代过程中,从 $b_{\text{urn_in}}$ 值(1 200)后每隔 100 次迭代输出一次权值(Weight),最终共得到 78 个模型,而模型并不是随着迭代次数越多而表现越好,过多的迭代次数可能会导致模型过拟合,为了后续更好地执行检测任务,需要对模型进行评估,寻找出最优的检测模型。模型的评价指标有准确率(Precision)、召回率(Recall)、平均精度均值(Mean average precision,mAP)、交并比(Intersection over union,IOU)、F1 值。

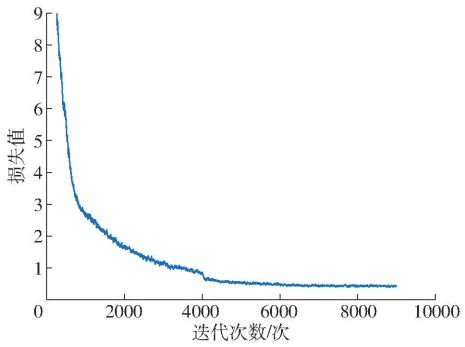


图 6 损失值随迭代次数的变化曲线

Fig. 6 Changing curve of loss value with iterations

$$P = \frac{T_P}{T_P + F_P} \times 100\% \tag{4}$$

$$R_0 = \frac{T_P}{T_P + F_N} \times 100\% \tag{5}$$

$$mAP = \frac{1}{C} \sum_{i=1}^{N} P(i) \Delta R_0(i) \tag{6}$$

$$IOU = \frac{S_{\text{overlap}}}{S_{\text{union}}} \tag{7}$$

$$F_1 = \frac{2PR_0}{P + R_0} \times 100\% \tag{8}$$

式中 T_P——被正确划分为正样本的数量

F_P——被错误划分为正样本的数量

F_N——被错误划分为负样本的数量

P——准确率 F_1——F1 值

R_0——召回率 i——识别序号

S_{overlap}——预测的边界框与真实的边界框之间的交集区域

S_{union}——预测的边界框与真实的边界框之间的并集区域

IOU——交并比

mAP——平均精度均值

本文使用 mAP 来评估模型的整体性能。计算所得的每个模型的 mAP 值如图 7 所示。由图 7 可以看到,当迭代次数达到 6 000 次时 mAP 值已经稳定。因此最优模型的选取从迭代区间 6 000 ~9 000 当中选择,其中 mAP 在迭代次数为 6 800 时达到最大值 91.47%,因此选择此模型作为最优模型。

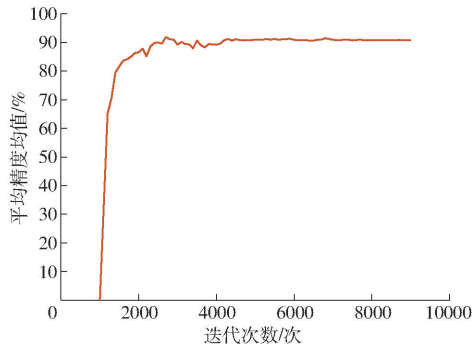


图 7　平均精度均值随迭代次数的变化曲线

Fig. 7　Changing curve of mean average precision with iterations

3　基于检测框区域的秧苗列中心线提取

3.1　自适应同列秧苗检测框聚类

由于采用了分段标注秧苗的方式，在使用训练好的 YOLOv3 模型进行检测时，定位输出秧苗列的检测框也是分段的，因此需要对同列秧苗的检测框进行聚类，后续再基于同类检测框区域进行秧苗列中心线提取。

自适应检测框聚类算法流程如下：

(1)定义网络输出的检测框结构体 $box=\{x,y,w,h\}$。

(2)创建两个一维向量 vector〈rightbox〉，vector〈leftbox〉。如果 $box.x$ < picture. weight/2，则 box 归入 vector〈leftbox〉；反之则归入 vector〈right〉。

(3)根据 box 的 y 坐标最小原则，找出所有一维向量中离图像上边界最近的初始聚类检测框 $box0=\{x_0,y_0,w_0,h_0\}$。

(4)找到初始聚类检测框后，按照如下准则在图像中从上往下聚类下一层邻近检测框(定义为 $box1=\{x_1,y_1,w_1,h_1\}$)。如果 $0\leqslant|x_0-x_1|<w_1+T_1$ 且 $h_0-T_2<y_1-y_0<h_0+T_2$，则 $box0$ 与 $box1$ 归为同一类。其中 $T_1=W/64$，$T_2=H/16$，W 为图像宽度，H 为图像高度，T_1 与 T_2 的取值将在后续中讨论。

(5)把 $box1$ 设为新的初始聚类框 $box0$，继续依照步骤(4)中的条件判断往下查找邻近的同列秧苗检测框，直至完成同列秧苗的检测框聚类为止。

同列秧苗的检测框聚类完成后，同颜色的检测框即为同一类(即属于同列秧苗)。

3.2　基于归一化 ExG 算子的秧苗图像分割

由于水田背景噪声的多变和秧苗的生长形态不同，需要从复杂背景中提取出秧苗特征。利用 ExG 算子，其能从背景中较好地分割出绿色作物。在进行灰度化前，先对图像中的颜色分量 R、G、B 进行归一化，以减小光照变化对图像带来的影响，计算公式为

$$\begin{cases}r=G/(R+G+B)\\g=R/(R+G+B)\\b=B/(R+G+B)\end{cases}\tag{9}$$

式中　R、G、B——图像的红、绿、蓝 3 个颜色通道的数值

r、g、b——归一化后的红、绿、蓝数值

$$I(x,y)=\begin{cases}0 & (2g-r-b\leqslant 0)\\2g-r-b & (0<2g-r-b<255)\\255 & (2g-r-b\geqslant 255)\end{cases}\tag{10}$$

式中　$I(x,y)$ ——坐标(x,y)处的灰度

3.3　基于检测框区域内的 SUSAN 角点特征提取

经过灰度化处理后的图像，除去了大部分图像的背景噪声，但是存在蓝藻或者浮萍的情况下，会产生很强的椒盐噪声。使用中值滤波去除背景的椒盐噪声，对比均值滤波和高斯滤波，中值滤波在很好地保护图像细节的同时，抑制图像中大部分的椒盐噪声，再通过 7×7 的类圆形 USAN 模板对灰度化后的图像中的检测框区域提取秧苗 SUSAN[35] 角点特征，最后使用 5×5 的正方形模板进行非极大值抑制，以剔除伪角点。

3.4　基于最小二乘法的秧苗列中心线拟合

由于基于 YOLOv3 的秧苗预测的方法，可以预先定位秧苗列区域的大致位置，所以在基于检测框区域对秧苗的特征提取时，避免了在图像全局区域进行秧苗的特征扫描提取，能较大程度避免背景噪声对特征提取的干扰，而同时图像经过中值滤波后，抑制了大部分水田中的浮萍和蓝藻带来的椒盐噪声，所以本文在中心线的提取上，采用拟合效果较理想的、计算速度较快的最小二乘法。

4　结果与分析

4.1　试验平台

本文训练和测试均在同一台计算机进行，计算机主要配置为 AMD Ryzen 5 CPU，主频 3.20 GHz、11 GB的 GPU GeForce GTX 1080Ti 和 16 GB 的运行内存。开发环境为 Windows 10(64 位)系统，VS2015，开发语言为 C++。

4.2　检测框聚类的距离阈值确定

T_1 和 T_2 分别为属于同列秧苗的上下相邻检测框之间的横向距离阈值与纵向距离阈值。T_1 和 T_2 阈值的选择会影响检测框聚类的准确性，而聚类的效果只能依靠聚类处理后的图像进行判断，没有定量的指标来衡量其效果。本文通过从秧苗 4 个生长阶段图像集中随机均匀选取 100 幅图像进行试验，

得出:当 T_1 的取值在图像宽度 W 的 1/64 ~ 1/8 之间时,T_2 的取值在图像高度 H 的 1/24 ~ 1/12 之间,均能得到较好的聚类效果。

通过试验发现,由于不同秧苗列间的检测框存在一定的横向距离,同列秧苗列的相邻(上下相邻)检测框存在较小的横向距离,T_1 的取值应不大于图像宽度 W 的 1/8,过大则容易使得属于不同列的秧苗检测框误聚类;T_1 的取值如果小于图像宽度 W 的 1/64,则容易使得属于同列的相邻(上下相邻)检测框聚类失败。另一方面 T_2 在大于图像高度 H 的 1/12 时,在秧苗列较为密集的情况下,不同列间的检测框距离较小,容易产生误聚类;当 T_2 小于图像高度 H 的 1/24 时,则容易产生属于同列的且有重叠部分(或者非重叠、有一定纵向距离)的相邻(上下相邻)检测框聚类失败。因此在初定 $T_1 \in [W/64, W/8]$、$T_2 \in [H/24, H/12]$ 后,本文选取 T_1 取值范围内的 4 个典型值 $W/64$、$W/32$、$W/16$、$W/8$,T_2 的 3 个典型值 $H/24$、$H/16$、$H/12$,两者进行组合,对所选的图像进行秧苗检测框聚类,并进行聚类准确性的统计,如表 1 所示。由表 1 可以看出,T_1 与 T_2 取($W/64$,$H/16$)时,各阶段秧苗图像的检测框聚类准确率均达到最大值,平均准确率为 95%,在秧苗生长的初期,秧苗列稀疏,间距较大,网络输出的不同秧苗列间的检测框区分度较大,不易产生误聚类的情况,因此聚类的准确率较高。而在秧苗生长的中后期,由于生长形态发生变化,导致不同秧苗列间的检测框间距较小,增加了聚类的难度,导致聚类准确率下降。

表 1 T_1 与 T_2 不同取值下检测框聚类的准确率

Tab. 1 Accuracy of clustering bounding boxes with different values of T_1 and T_2 %

(T_1,T_2)	插秧第 1 周	插秧第 2 周	插秧第 3 周	插秧第 4 周
($W/64$,$H/24$)	96	95	93	91
($W/64$, $H/16$)	98	97	94	91
($W/64$, $H/12$)	93	92	90	88
($W/32$, $H/24$)	93	91	91	88
($W/32$, $H/16$)	95	93	93	87
($W/32$, $H/12$)	91	90	90	85
($W/16$, $H/24$)	93	91	89	83
($W/16$, $H/16$)	91	90	88	80
($W/16$, $H/12$)	86	83	82	77
($W/8$, $H/24$)	84	79	73	76
($W/8$, $H/16$)	83	80	76	73
($W/8$, $H/12$)	79	74	69	68

4.3 算法性能分析

为了进一步检测本文算法在南方复杂水田环境中的性能,将本文提出的算法与文献[16]中的在相同应用环境下的算法在运行速度和识别误差上进行对比。在 4.1 节中所述的硬件平台上,检测 1 幅 640 像素 ×480 像素的彩色图像,文献[16]中的算法平均耗时为 256 ms,本文算法平均耗时 82.6 ms。对于 2 种算法的检测误差比较,本文通过 3 人手动绘制水田秧苗图像 ROI 范围内的秧苗列中心线,并取其平均值代表秧苗列的参考中心线,计算实际检测出来的直线与手动绘制直线的夹角,夹角越大,则提取的直线误差越大。经过对测试集中 200 幅图像进行试验,文献[16]算法提取中心线的平均夹角误差为 3.50°,标准差为 3.28°,本文算法的平均夹角误差为 0.97°,标准差为 0.73°。可见,本文提出的算法不仅计算速度得到提高,而且还具有较高的识别精度。

4.4 不同生长阶段秧苗列中心线提取对比试验

从采集到的图像中随机选取秧苗每个生长阶段各 50 幅图像,使用本文算法和文献[16]算法分别进行秧苗中心线提取,2 种算法的对比效果如图 8 所示。由图 8 可以看出,在秧苗生长前期(第 1、2 周内),由于此时水田背景噪声较少,水田环境相对简单,两种方法提取中心线的效果大体相当。当在插秧第 3 周时,由于秧苗生长导致形态发生变化,开始出现秧苗侧偏、叶片相互遮挡,而在插秧第 4 周时情况更加明显,此时水田中还伴有蓝藻、浮萍等背景噪声的影响,由于文献[16]中基于滑动窗口聚类的方法受背景噪声、秧苗生长形态的影响较大,所以在提取中心线的过程中均出现了明显偏差。而本文算法基于秧苗检测框区域进行秧苗特征提取,避免对整幅图像进行基于窗口扫描的特征提取,较大程度地避免了背景噪声的干扰。另一方面 YOLOv3 模型检测定位出的是秧苗的主干部分,减少了秧苗生长过程中由于形态变化所造成秧苗特征提取容易偏离作物中心的影响。

2 种算法的试验统计结果如表 2 所示。对于秧苗的 4 个生长阶段,本文算法所提取的秧苗列中心线角度误差均小于文献[16]算法,本文算法提取的秧苗中心线平均角度误差为 0.93°,文献[16]算法提取的秧苗中心线平均角度误差为 2.92°。由此可见本文算法对于不同生长阶段的秧苗适应性更好,精度更高。

4.5 复杂水田环境下秧苗列中心线提取对比试验

对于一些较为复杂的水田环境,如大风、蓝藻、浮萍和秧苗倒影、水面强光反射、暗光线的特殊场景,秧苗列中心线提取效果对比如图 9 所示。由图 9 可以看到在这些特殊场景下,由于环境变化使

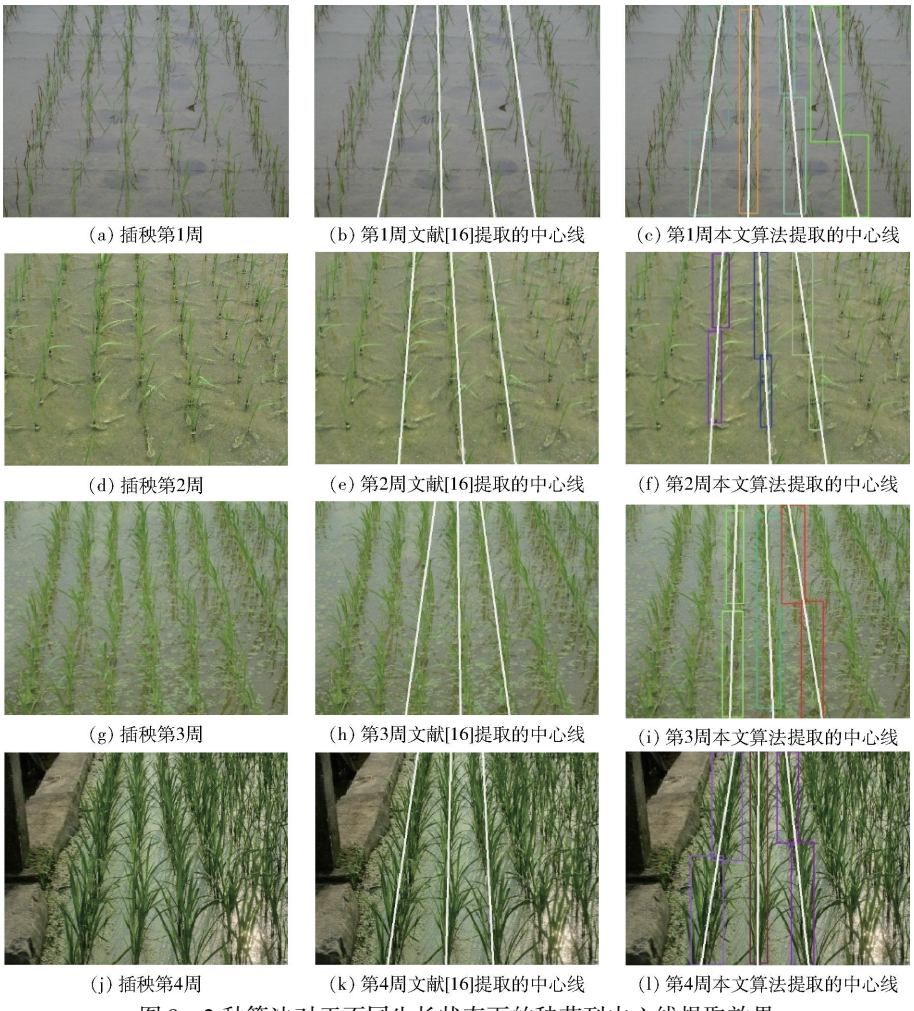


图 8　2 种算法对于不同生长状态下的秧苗列中心线提取效果

Fig. 8　Two algorithms for extracting centerlines of rice seedlings under different growth stages

表 2　2 种算法对不同生长阶段下秧苗列中心线提取角度误差对比

Tab. 2　Error comparison of extracting centerlines of rice seedlings in different growth stages by two algorithms

(°)

算法	插秧第 1 周	插秧第 2 周	插秧第 3 周	插秧第 4 周	平均值
本文算法	0. 62	0. 77	1. 03	1. 31	0. 93
文献[16]算法	2. 27	2. 52	3. 20	3. 70	2. 92

得秧苗角点特征分布更加分散,此时文献[16]算法所提取的中心线误差较大。

而本文基于同列秧苗检测框区域的中心线提取算法,较大减少了环境变化对中心线提取的干扰。基于 YOLOv3 算法能较好检测到秧苗列的主体部分,基于检测框区域所提取的角点特征能很好地表征秧苗列的中心位置,使得拟合的直线更加接近秧苗列的中心线。

对采集到的图像中随机选取大风、蓝藻、浮萍和秧苗倒影、水面强光反射、暗光线的特殊环境下各 50 幅图像分别使用本文算法和文献[16]算法进行秧苗中心线提取,误差对比如表 3 所示。在大风、蓝藻、浮萍和秧苗倒影、水面强光反射、暗光线的复杂水田环境下,本文提取的秧苗列中心线角度误差均小于文献[16]算法,本文算法提取的秧苗中心线平均角度误差为 0. 92°,文献[16]算法提取的秧苗中心线平均角度误差为 3. 11°,由此可见本文算法面对复杂的水田环境时的鲁棒性更强,精度更高。

5　结论

(1)针对我国南方复杂的水田环境,提出了基于 YOLOv3 目标检测的水田秧苗列中心线提取算法。构建图像的 ROI 区域,以减小图像边缘作物列对后续图像处理的干扰;构建秧苗识别深度神经网络,提高了复杂场景下提取秧苗列中心线的精度;基于同列秧苗检测框聚类和检测框内的秧苗特征进行提取,避免针对图像全局的特征提取,减小了图像处理的数据量和环境噪声的干扰。研究和试验表明,所提出的方法适用于秧苗的不同生长阶段和复杂的水田环境,如大风、蓝藻、浮萍与秧苗倒影、水面强光

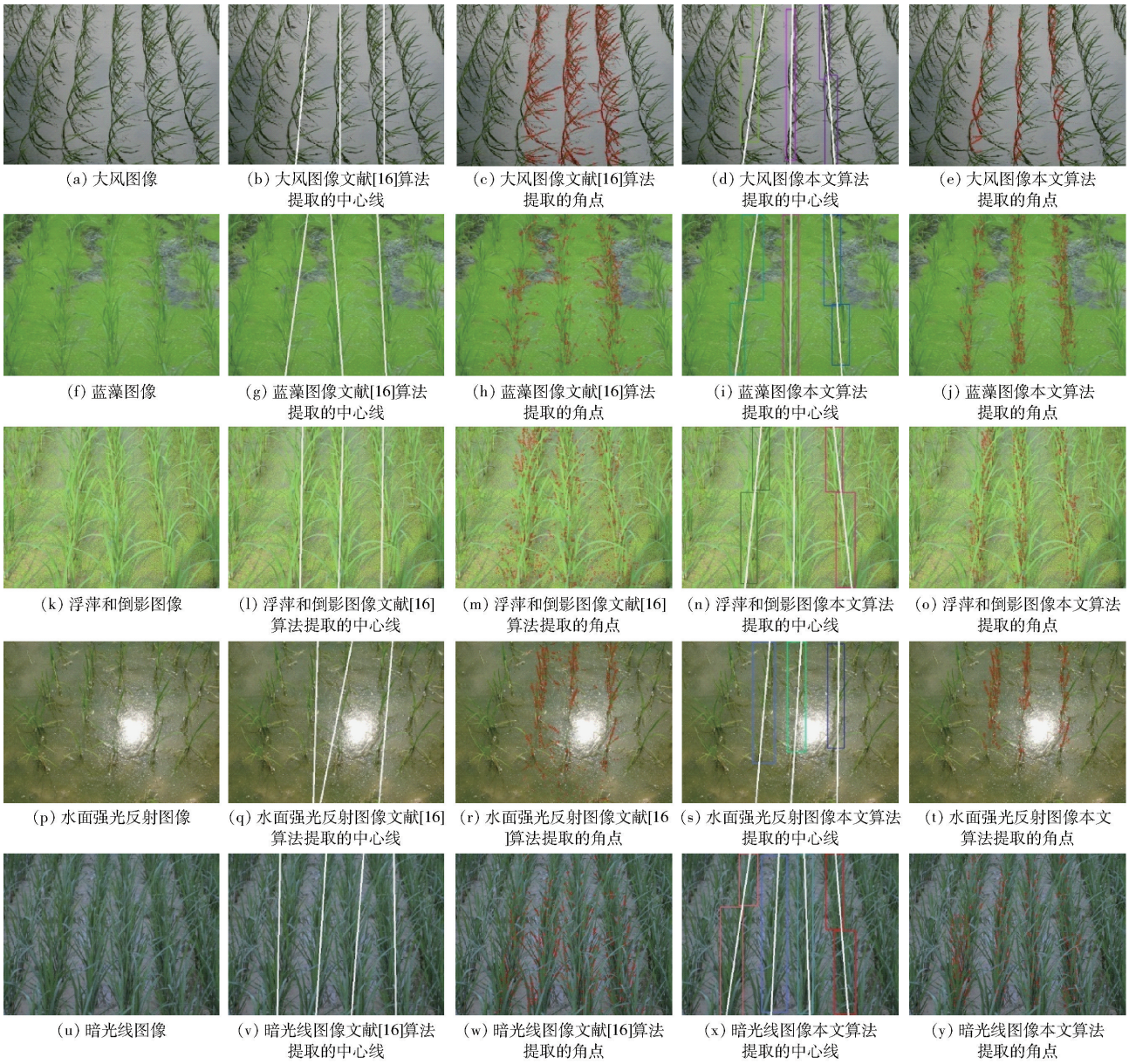


图 9 复杂水田环境下 2 种算法的秧苗列中心线提取效果

Fig. 9 Centerlines extraction of rice seedlings based on two algorithms in complex paddy environment

表 3 2 种算法在复杂环境下提取秧苗列中心线误差对比

Tab. 3 Error comparison of centerlines of rice seedlings rows extracted in complex environment by two algorithms

(°)

算法	大风	蓝藻	浮萍和秧苗倒影	水面强光反射	暗光线	平均值
本文算法	1.06	0.59	1.33	0.72	0.90	0.92
文献[16]算法	3.46	3.30	4.06	2.04	2.70	3.11

反射、暗光线等。

(2)采用本文算法提取的秧苗列中心线平均角度误差为 0.97°, YOLOv3 模型的平均精度为 91.47%,单幅图像(分辨率 640 像素 ×480 像素)在 GPU 下的平均处理时间为 82.6 ms。与相同应用背景下的文献[16]算法相比,本文算法在复杂的水田环境下鲁棒性更好、检测速度更快、精度更高,提高了机器人视觉导航的实时性。

参 考 文 献

[1] 胡静涛,高雷,白晓平,等. 农业机械自动导航技术研究进展[J]. 农业工程学报, 2015,31(10):1 - 10.
HU Jingtao, GAO Lei, BAI Xiaoping, et al. Review of research on automatic guidance of agricultural vehicles[J]. Transactions of the CSAE,2015,31(10):1 - 10. (in Chinese)

[2] HIREMATH S, VAN EVERT F K, BRAAK C J, et al. Image-based particle filtering for navigation in a semi-structured agricultural environment[J]. Biosystems Engineering, 2014, 121: 85 - 95.

[3] ZHANG B, ZHANG X, LI X, et al. Automated robust crop-row detection in maize fields based on position clustering algorithm and shortest path method[J]. Computers and Electronics in Agriculture, 2018,154:165 - 175.

[4] BASSO M, DE FREITAS E P. A UAV guidance system using crop row detection and line follower algorithms[J]. Journal of Intelligent & Robotic Systems, 2020,97(3):605 - 621.

[5] 孟庆宽,张漫,杨耿煌,等. 自然光照下基于粒子群算法的农业机械导航路径识别[J/OL]. 农业机械学报,2016,47(6):11 - 20.
MENG Qingkuan, ZHANG Man, YANG Genghuang, et al. Guidance line recognition of agricultural machinery based on particle swarm optimization under natural illumination[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016,47(6):11 - 20. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no = 20160602&flag = 1&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2016.06.002. (in Chinese)
[6] 胡炼,罗锡文,曾山,等. 基于机器视觉的株间机械除草装置的作物识别与定位方法[J]. 农业工程学报,2013,29(10):12 - 18.
HU Lian, LUO Xiwen, ZENG Shan, et al. Plant recognition and localization for intra-row mechanical weeding device based on machine vision[J]. Transactions of the CSAE, 2013, 29(10):12 - 18. (in Chinese)
[7] 张志斌,罗锡文,臧英,等. 基于颜色特征的绿色作物图像分割算法[J]. 农业工程学报,2011,27(7):183 - 189.
ZHANG Zhibin, LUO Xiwen, ZANG Ying, et al. Segmentation algorithm based on color feature for green crop plants[J]. Transactions of the CSAE, 2011,27(7):183 - 189. (in Chinese)
[8] 宋宇,刘永博,刘路,等基于机器视觉的玉米根茎导航基准线提取方法[J/OL]. 农业机械学报,2017,48(2):38 - 44.
SONG Yu, LIU Yongbo, LIU Lu, et al. Extraction method of navigation baseline of corn roots based on machine vision[J/OL]. Transactions of the Chinese Society of Agricultural Machinery, 2017,48(2):38 - 44. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no = 20170205&flag = 1&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2017.02.005. (in Chinese)
[9] CHOI K H, HAN S K, HAN S H, et al. Morphology-based guidance line extraction for an autonomous weeding robot in paddy fields[J]. Computers and Electronics in Agriculture, 2015,113:266 - 274.
[10] MENG Q, HAO X, ZHANG Y, et al. Guidance line identification for agricultural mobile robot based on machine vision[C]// 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). IEEE, 2018: 1887 - 1893.
[11] GARCÍA-SANTILLÁN I D, MONTALVO M, GUERRERO J M, et al. Automatic detection of curved and straight crop rows from images in maize fields[J]. Biosystems Engineering, 2017,156: 61 - 79.
[12] KAIZU Y, IMOU K. A dual-spectral camera system for paddy rice seedling row detection[J]. Computers and Electronics in Agriculture, 2008,63(1):49 - 56.
[13] 乔智利,崔彦平,邹新光. 农田视觉导航基准线的识别与提取方法研究[J]. 农机化研究,2016,38(7):124 - 128.
QIAO Zhili, CUI Yanping, ZOU Xinguang. Study on the method of visual navigation baseline identification and extraction of farmland[J]. Journal of Agricultural Mechanization Research, 2016,38(7):124 - 128. (in Chinese)
[14] MONTALVO M, PAJARES G, GUERRERO J M, et al. Automatic detection of crop rows in maize fields with high weeds pressure[J]. Expert Systems with Applications, 2012,39(15):11889 - 11897.
[15] 张勤,黄小刚,李彬. 基于彩色模型和近邻法聚类的水田秧苗列中心线检测方法[J]. 农业工程学报,2012,28(17):163 - 171.
ZHANG Qin, HUANG Xiaogang, LI Bin. Detection of rice seedlings rows' centerlines based on color model and nearest neighbor clustering algorithm[J]. Transactions of the CSAE, 2012,28(17):163 - 171. (in Chinese)
[16] 张勤, 陈少杰, 李彬. 基于 SUSAN 角点的秧苗列中心线提取方法[J]. 农业工程学报,2015,31(20):165 - 171.
ZHANG Qin, CHEN Shaojie, LI Bin. Extraction method for centerlines of rice seedlings based on SUSAN corner[J]. Transactions of the CSAE, 2015, 31(20): 165 - 171. (in Chinese)
[17] ZHANG Q, CHEN M E S, LI B. A visual navigation algorithm for paddy field weeding robot based on image understanding [J]. Computers and Electronics in Agriculture, 2017, 143: 66 - 78.
[18] 孙红,李松,李民赞,等. 农业信息成像感知与深度学习应用研究进展[J/OL]. 农业机械学报,2020,51(5):1 - 17.
SUN Hong, LI Song, LI Minzan, et al. Research progress of image sensing and deep learning in agriculture[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020,51(5):1 - 17. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20200501&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2020.05.001. (in Chinese)
[19] 孙哲, 张春龙, 葛鲁镇, 等. 基于 Faster R - CNN 的田间西兰花幼苗图像检测方法[J/OL]. 农业机械学报, 2019, 50(7):216 - 221.
SUN Zhe, ZHANG Chunlong, GE Luzhen, et al. Image detection method for broccoli seedlings in field based on Faster R - CNN[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2019,50(7):216 - 221. http://www.jcsam.org/jcsam/ch/reader/view_abstract.aspx? file_no = 20190723&flag = 1&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2019.07.023. (in Chinese)
[20] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580 - 587.
[21] GIRSHICK R. Fast R - CNN[C]// Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440 - 1448.
[22] REN S, HE K, GIRSHICK R, et al. Faster R - CNN: towards real-time object detection with region proposal networks[C]// Advances in Neural Information Processing Systems, 2015: 91 - 99.
[23] 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.
[24] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263 - 7271.
[25] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. arXiv Preprint arXiv:1804.02767, 2018.
[26] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]// European Conference on Computer Vision, 2016:21 - 37.
[27] HE K, ZHANG X, REN S, et al. Identity mappings in deep residual networks[C]// European Conference on Computer Vision, 2016: 630 - 645.
[28] GUERRERO J M, RUZ J J, PAJARES G. Crop rows and weeds detection in maize fields applying a computer vision system based on geometry[J]. Computers and Electronics in Agriculture, 2017,142:461 - 472.
[29] ARAVIND R, DAMAN M, KARIYAPPA B S. Design and development of automatic weed detection and smart herbicide sprayer robot[C]//2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS). IEEE, 2015: 257 - 261.
[30] JIANG G, WANG Z, LIU H. Automatic detection of crop rows based on multi-ROIs[J]. Expert Systems with Applications, 2015,42(5):2429 - 2441.
[31] 李金屏,修择,吕玉彬,等. 基于虚拟定标线的巡检机器人视觉导航[J]. 中国体视学与图像分析,2011,16(1):50 - 56.
LI Jinping, XIU Ze, LÜ Yubin, et al. A patrol robot visual navigation method based on virtual calibration line[J]. Chinese Journal of Stereology and Image Analysis, 2011,16(1):50 - 56. (in Chinese)
[32] SMITH S M, BRADY J M. SUSAN—a new approach to low level image processing[J]. International Journal of Computer Vision, 1997,23(1):45 - 78.