

基于 K - SSD - F 的东亚飞蝗视频检测与计数方法

李 林 柏 召 刁 磊 唐 詹 郭旭超

(中国农业大学信息与电气工程学院, 北京 100083)

摘要: 针对目前国内蝗虫监测主要以人工监测为主、监测效率低且计数不准确的问题,以 5 龄东亚飞蝗为实验对象,提出了一种蝗虫视频计数方法 K - SSD - F 算法。该方法可以实时、连续、自动监测蝗虫的数量。首先利用背景分离法中的 KNN 算法提取视频前后帧的时空特征;然后通过标注好的数据训练 SSD 模型,并对视频进行检测,提取视频的静态特征,二者结合以提高计数准确率;最后利用补帧算法识别因姿态变化导致的漏计数的帧。实验结果表明,蝗虫识别准确率为 97%,召回率为 89%,平均检测精度(mAP)为 88.94%,F1 值为 92.82%,且检测速度达到了 19.78 f/s。本文方法具有较好的鲁棒性,可以实现蝗虫的实时和自动计数,其精度优于其他模型,也可为其他种类的昆虫自动识别计数提供理论基础。

关键词: 东亚飞蝗; 视频目标计数; 背景分离法; SSD 算法; 补帧算法

中图分类号: TP391 **文献标识码:** A **文章编号:** 1000-1298(2021)S0-0261-07

Video Detection and Counting Method of East Asian Migratory Locusts
Based on K - SSD - F

LI Lin BAI Zhao DIAO Lei TANG Zhan GUO Xuchao

(College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China)

Abstract: At present, domestic locust monitoring is mainly based on manual monitoring, with low monitoring efficiency and inaccurate counting. In response to the above problems, the K - SSD - F algorithm, a video counting method of locusts, was proposed with the 5th instar migratory locust as the experimental object. This method can monitor the number of locusts in real time, continuously and automatically. Firstly, the KNN algorithm in the background separation method was used to extract the spatiotemporal features of the frames before and after the video; then the SSD model was trained through the labeled data, the video was detected, and the static features of the video were extracted, and the two were combined to improve the counting accuracy; finally, the frame compensation algorithm was used to recognize missing frames due to posture changes. The experimental results showed that the precision of locust identification was 97%, the recall rate was 89%, the average detection accuracy (mAP) was 88.94%, the F1 value was 92.82%, and the detection speed reached 19.78 f/s. The proposed method had good robustness, which can realize real-time and automatic counting of locusts, its accuracy was better than that of other models, and it can also provide a theoretical basis for automatic identification and counting of other kinds of insects.

Key words: East Asian migratory locust; video target counting; background separation method; SSD algorithm; frame compensation algorithm

0 引言

蝗虫是一种世界性的农业害虫,全世界约有

170 个国家或地区受到过蝗灾的威胁,其中以亚洲和非洲国家蝗灾发生最为频繁和严重^[1-2]。

目前,在中国最常用的蝗虫监测方法是人工调

查法^[3]。人工监测费时、费力、效率低、人工成本高;有些地区无法涉足,如湖泊和沼泽地,则难以实现人工监测^[4];部分工作人员对蝗虫特征了解不充分,对蝗虫的品系、龄期识别不准确,存在误计数的情况。近年来,计算机视觉技术发展迅速,在昆虫监测和鉴定中得到了广泛的应用^[5-9]。这些应用虽然降低了人工成本,但一般只估算昆虫数量,准确率较低。

随着深度卷积神经网络的发展,与传统计算机视觉方法相比,深度卷积神经网络的准确率有了很大的提高^[10-11],且在目标检测^[12-13]上取得了显著的进展。在农业病虫害检测领域已经有了广泛的应用。张银松等^[14]提出了一种基于改进 Faster R-CNN 模型的粘虫板图像昆虫识别与计数方法。符首夫^[15]提出了一种基于卷积神经网络的农作物昆虫识别与计数方法,完成了对田间采集的农作物昆虫图像自动识别与计数。YE 等^[16]利用 ResNet- Locust-BN 模型来识别蝗虫物种和龄期。XIN 等^[17]提出了基于迁移学习和深度卷积网络的大豆病虫害图像识别算法。AI 等^[18]提出了基于深度学习体系结构的野外飞蛾图像姿态估计相关识别方法。但这些传统神经网络算法侧重于检测静止图像中的目标^[19],无法长时间监测蝗虫。

现有蝗虫计数方法存在以下问题:①传统人工计数法的计数误差较大,且耗费巨大的人力物力。②基于计算机视觉、深度卷积神经网络的蝗虫计数方法,大多是以静态图像进行识别计数,只能检测某一时刻的蝗虫图像,不能长时间监测蝗虫,无法实时监测蝗虫的数量,不具有连续性。③对运动过程中出现运动模糊、部分遮挡情况下的蝗虫无法计数。

为了解决上述问题,本文提出一种基于视频的东亚飞蝗计数方法。首先利用背景分离法(KNN)提取前后帧的时空特征,并与 SSD 算法提取的静态特征结合。然后利用提出的补帧算法解决蝗虫运动过程中因姿态变化导致的个别视频帧无法计数的问题。最后对所识别的蝗虫目标进行计数。

1 数据和方法

1.1 东亚飞蝗的形态学特征

5 龄的东亚飞蝗分为群居型和散居型 2 种,群居型体色红褐色或黑褐色,无光泽,前胸背板背面具有黑色纵纹,背板镶有狭波状的黄色边缘,中胸及后胸背板微凸,如图 1 所示。散居型体色常为绿色、黄绿色或淡褐色。体长 25~40 mm,其中雄性体长较短(25~36 mm),雌性体长较长(30~40 mm)。其

外形特征翅芽较前胸背板长或等长,翅芽长度可到达腹部第 4、5 节。识别蝗虫的关键是头部特征、背板特征、尾部特征,通过眼睛、触角、大腿、前足、中足、尾部可以判定为蝗虫,然后通过翅芽、体长可以判定蝗虫的龄期。在自然环境中,存在很多类似于蝗虫的物体,比如枯草颜色类似于蝗虫的大腿,潮湿的土壤类似于蝗虫的背板,单独依靠当前帧的形态学特征,容易造成误检测的情况。单独利用深度模型单帧检测蝗虫图像效果较差,所以本文在 SSD 模型上,加入 KNN 算法和补帧算法,融合更多的时空特征,弥补在训练或者检测中单帧特征不足的问题。

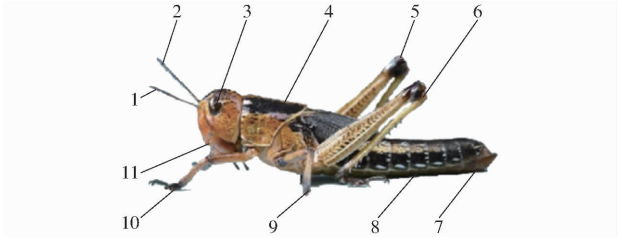


图 1 5 龄东亚飞蝗形态结构
Fig. 1 Morphological structure of the fifth East Asian migratory locust
1、2. 触角 3. 眼睛 4. 背板 5、6. 大腿 7. 尾部 8. 气门 9. 中足 10. 前足 11. 口器

1.2 数据集

为了增加蝗虫检测的准确性和鲁棒性,真实反映实际应用场景,需要在不同的实验条件下,采用大量的数据进行模型训练。然而,目前符合本文研究要求的公开可用数据极度缺乏。因此,本研究在中国农业大学植保学院蝗虫养殖基地采集数据,选取 5 龄东亚飞蝗为实验对象,设置每平方米的蝗虫数量为 40 只。通过 Canon EOS 5D Mark II 型相机采集了 200 GB 东亚飞蝗视频数据。对光源、植物环境、视频分辨率等做了不同设置,以丰富数据集,增加模型的鲁棒性。采用 60 W 白炽灯模拟自然光,将灯光置于实验箱侧面 0.5 m 处模拟早晨的阳光、正上方 1 m 处模拟室外中午时刻的阳光、侧面 1 m 处模拟傍晚时刻的阳光、无白炽灯补光模拟阴天时刻的阳光,每种场景下的视频数据各占 1/4。

植物环境:蝗虫生存环境主要为株高 3~10 cm 的小麦。

视频分辨率:共采集了 3 种不同分辨率的视频。分别为 640 像素×480 像素,帧速率 25 f/s;1 280 像素×720 像素,帧速率 50 f/s;1 920 像素×1 080 像素,帧速率 25 f/s。

1.3 数据标注方法

将采集的视频数据按照 9:1 划分,其中 180 GB 视频作为训练集,20 GB 视频作为测试集。对训练

集中的视频每隔 5 s 截取一帧,并对其进行预处理,即剪裁或者缩放尺寸较大的图像,同时删除一些无蝗虫的图像,最后得到 5 000 幅图像。采用人工标注与半自动标注结合的方法标注数据。人工标注方法是通过 lableImg 软件对蝗虫位置进行标框, lableImg 将标注框的种类和对角坐标 ($X_{\min}$ 、 $Y_{\min}$ 、 $X_{\max}$ 、 $Y_{\max}$) 按固定格式存储至对应的文件,如图 2 所示。半自动标注方法利用 2 500 幅已经人工标注好的数据作为训练集,训练得到标注数据模型,利用此标注数据模型标注余下 2 500 幅图像数据,并按照 lableImg 标注数据的存储格式自动生成蝗虫框,最后通过 lableImg 软件人工核对并更正识别框的准确性。人工标注 2 500 幅图像耗费了 50 h,半自动标注 2 500 幅图像仅花费了 10 h,效率提高了 4 倍。

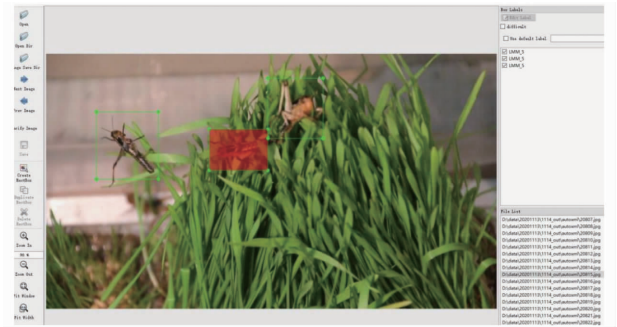


图 2 lableImg 标注结果
Fig. 2 lableImg labeling results

1.4 数据增强

深度学习的基础在于数据,数据的丰富程度决定了训练网络的精度和收敛性^[20]。本文将数据集分为简单样本和困难样本两种,简单样本指图像清晰、无干扰、无遮挡的蝗虫样本,见图 3a;困难样本指图像模糊、存在噪声干扰、存在遮挡的样本,见图 3b~3d。本数据集中简单样本较多,而困难样本的数据较少,所以对标注好的 5 000 幅图像进行数据增强处理。通过高斯噪声、调节图像亮度两种方法将原数据集扩展为 15 000 幅图像。

采用 OpenCV 中的 addweighted() 函数调节图像的亮度和对比度,公式为

$$I_m = \alpha s_{rc1} + \beta s_{rc1} + \gamma \tag{1}$$

式中 I_m ——最终输出的图像矩阵
 s_{rc1} ——输入的图像矩阵
 α 、 β 、 γ ——图像矩阵的权重

处理后的图像如图 3b 所示。

高斯噪声是指噪声的概率密度函数服从正态分布的一类噪声,它的幅度分布服从高斯分布,而它的功率谱密度又是均匀分布,因此称为高斯白噪声。其计算公式为



图 3 东亚飞蝗的不同样本
Fig. 3 Different samples of East Asian migratory locust

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\bar{z})^2}{2\sigma^2}} \tag{2}$$

式中 σ —— z 的标准差 $\bar{z}$ ——像素均值
 z ——图像像素 p ——概率密度
处理后的图像如图 3c 所示。经过图像增强后的图像不需要重新标注,利用原图的标注文件自动生成图像增强后的图像标注文件。

1.5 K-SSD-F 算法

针对困难样本中出现图像模糊、噪声干扰、存在遮挡、姿态变化引起的计数不准确的问题,提出了 K-SSD-F 算法。首先利用背景分离法中的 KNN 算法提取视频前后帧的时空特征;然后与 SSD 算法提取的静态特征结合,对置信度按比例加权,对困难样本的置信度进行增强,解决图像模糊、噪声干扰、存在遮挡导致的识别失败问题;最后利用补帧算法对姿态变化引起的个别蝗虫识别失败问题进行处理,提高检测的准确度。K-SSD-F 算法流程图如图 4 所示,K-SSD-F 算法检测步骤为:

(1)将处理好的数据集放入卷积神经网络中训练 SSD 模型,利用训练好的模型对蝗虫视频进行检测,筛选置信度大于 0.05 的蝗虫检测框,并将筛选后的目标框记为 SSDbox,目标置信度记为 SSDscore。

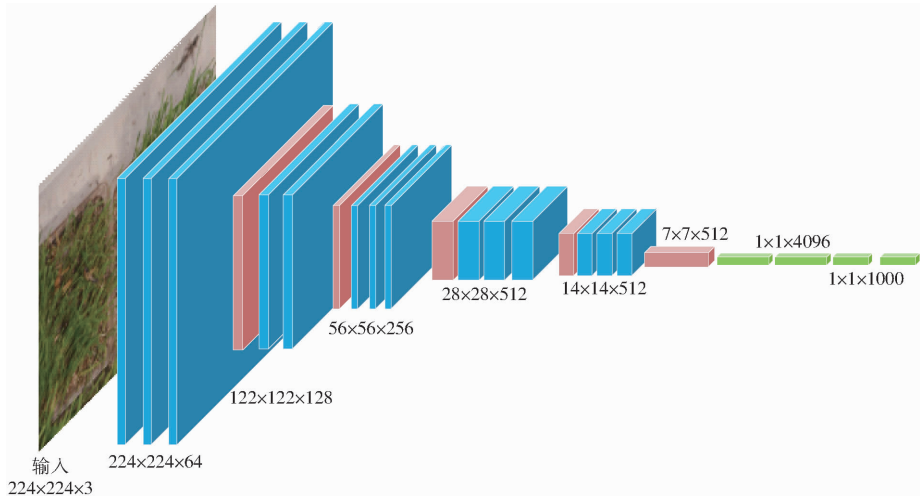


图 6 VGG16 网络结构图
Fig. 6 VGG16 network structure diagram

接层。底层是颜色、轮廓、纹理等基础结构,显然大部分问题都由这些相同的基础结构组成,所以可以冻结这些层,加快训练速度。调整学习率为 10^{-4} ,降低学习率防止出现梯度震荡,保证模型收敛加快。

训练完成后使用模型对蝗虫视频进行检测,将检测框的坐标记为 SSDbox,目标概率记为 SSDscore。设置 SSDscore 阈值为 0.05,筛选所有 SSDscore 大于 0.05 的框为 SSDbox。

1.5.3 补帧算法

蝗虫在运动过程中姿态一直变化,导致部分视频帧出现漏检测的情况。补帧算法充分利用前后帧的关系,利用前后帧同时存在的框来补全当前帧丢失的框。主要算法流程如下:首先设置一个长度为 3 的滑动窗口,用来存储连续 3 帧视频检测结果;然后输入视频,利用训练好的模型连续识别并存储连续 3 帧蝗虫目标框的位置和置信度;其次求出第 1 帧和第 3 帧所有框和框对应的置信度并集分别记为 union_box 和 union_score;最后将 union_box 和 union_score 与第 2 帧对比,判断第 2 帧是否出现框丢失的情况,如果存在则将第 2 帧与并集框取交集,补充第 2 帧的框和对应的置信度。经过补帧算法处理后,对识别到的蝗虫按照先 x 轴后 y 轴顺序排序,然后依次对每只蝗虫进行标号,最后统计当前帧的所有蝗虫数量。图 7 为补帧算法检测结果对比。

2 实验结果与分析

本文将计算机视觉技术与深度学习技术相结合,构建基于 K-SSD-F 算法的蝗虫监测模型,对视频中 5 龄东亚飞蝗进行识别和检测。为了验证算法的有效性,将本文算法与 SSD^[21]、YOLO v4^[22]、YOLO v3^[23]、Faster R-CNN^[14] 常用的深度学习模型进行对比。本文对所有测试集的视频每隔 1 s 截

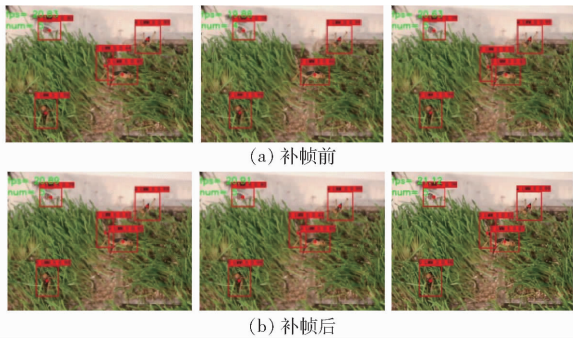


图 7 补帧算法检测结果对比

Fig. 7 Detection results of frame complementing algorithm

取一帧,人工标注每帧中的蝗虫,标注后的数据作为正确框(ground true)。利用训练好的模型对测试视频进行检测,同样每隔 1 s 保存一帧检测结果。将算法检测的目标框与人工标注框求交并比 (IOU),交并比大于 50% 视为检测正确。采用精确度 (Precision)、召回率 (Recall)、平均精度均值 (mAP)、F1 值、检测速度对模型进行综合评判。K-SSD-F 算法的平均检测精度达到了 97%,平均召回率达到了 89%,mAP 达到了 88.94%,F1 值为 92.82%。实验环境为:主板为 Inter(R) Core(TM) i7-9700K,CPU 内存为 32 GB,显卡为 GTX 1660Ti,操作系统为 Windows 10。

2.1 实验结果

对测试集数据进行测试,为了提高人工标注的效率,对视频每隔 1 s 保存一次预测结果,然后将预测结果与人工标注结果进行对比。如表 1 所示,K-SSD-F 算法相对于 SSD 算法,在保证准确率的前提下,召回率提高了 1 个百分点,说明本优化算法降低了蝗虫的漏检率,mAP 提高了 1.01 个百分点,说明模型能够有效提高检测精度;F1 提高了 0.54 个百分点,说明本算法检测效果较好。综合对比,K-SSD-F 算法提高了蝗虫检测精度,性能最优。

表 1 不同模型的实验结果对比

Tab.1 Comparison of experimental results of different models					
模型	<i>P</i> /%	<i>R</i> /%	mAP/%	F1/%	检测速度/ ($\text{f}\cdot\text{s}^{-1}$)
K-SSD-F	97	89	88.94	92.82	19.78
SSD	97	88	87.93	92.28	20.01
YOLO v4	85	96	84.58	90.17	14.50
YOLO v3	91	82	86.58	86.27	16.00
Faster R-CNN	94	72	81.18	81.54	1.80

表 2 为 4 段视频部分计数准确率。对每段测试视频每隔 1 s 截取 1 帧图像,保存预测框的坐标位置;然后人工标注每个蝗虫位置,对比预测框和人工标注框的位置,对整段视频中所有预测正确的框计数。虽然 Faster R-CNN 模型计数准确率较高,但是综合表 1 中的 *R* 和检测速度,其误检率较高而且速度较低。K-SSD-F 算法相比于其他目标检测模型,其计数准确率最高,检测效果最优。

表 2 不同算法计数准确率对比

Tab.2 Comparison of counting accuracy of different algorithms			
视频序号	检测方法	数量/只	蝗虫检测 准确率/%
1	人工计数	1 085	
	K-SSD-F	971	89.49
	SSD	958	88.29
	YOLO v4	875	80.64
	YOLO v3	889	81.93
	Faster R-CNN	964	88.85
2	人工计数	1 355	
	K-SSD-F	1 122	82.80
	SSD	1 020	75.28
	YOLO v4	956	70.55
	YOLO v3	1 029	75.94
	Faster R-CNN	1 105	81.54
3	人工计数	292	
	K-SSD-F	275	94.18
	SSD	270	92.47
	YOLO v4	243	83.22
	YOLO v3	240	82.19
	Faster R-CNN	284	97.26
4	人工计数	1 142	
	K-SSD-F	876	76.71
	SSD	876	76.64
	YOLO v4	804	70.28
	YOLO v3	783	68.38
	Faster R-CNN	902	78.71

2.2 结果分析

本文算法与 SSD 算法处理效果对比如图 8 所示。左上位置输出当前帧的蝗虫数量 (num);每个框上面对应蝗虫的标号、框的种类 (5 龄东亚飞蝗 LMM_5) 和置信度。在图 8a 中,无遮挡的简单样本 2 个,识别标号为 1、2;存在遮挡、模糊的困难样本 2 个。图 8a 中只检测到了 3 只蝗虫,其中左上困难样本因被遮挡、模糊导致 SSD 算法未检测到。图 8b 中 K-SSD-F 算法对置信度加权后为 0.74,成功检测出困难样本。

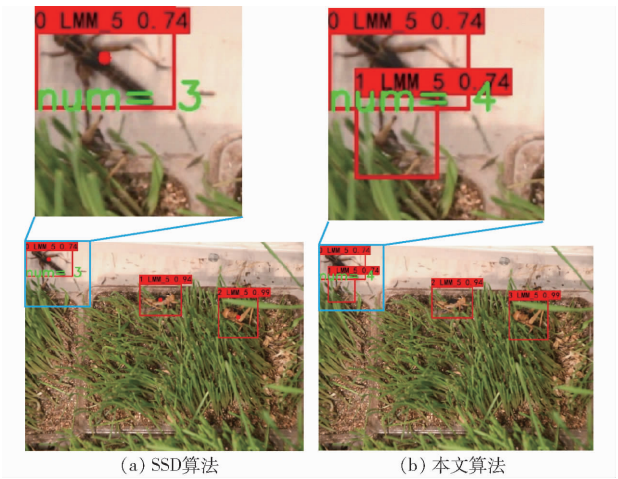


图 8 K-SSD-F 与 SSD 算法处理效果对比

Fig.8 Comparison between K-SSD-F and SSD algorithms

3 结束语

提出了一种基于 K-SSD-F 算法的东亚飞蝗视频目标计数方法,利用 K-SSD-F 算法解决了因图像模糊、噪声干扰、蝗虫遮挡、姿态变化导致的计数不准确的问题。利用背景分离法中的 KNN 算法提取视频前后帧的时空特征并与 SSD 算法提取的静态特征结合,对置信度按比例加权,解决图像模糊、噪声干扰、存在遮挡导致的识别失败问题;利用补帧算法,解决了蝗虫运动过程中因姿态变化导致的个别视频帧无法计数的问题。本文方法检测速度可以达到 19.78 f/s,可以实时识别蝗虫,并且自动对每帧视频的蝗虫记数。通过 20 GB 的东亚飞蝗视频验证了本文算法的鲁棒性、泛化性。蝗虫视频目标检测实验表明,与 SSD、YOLO v4、YOLO v3、Faster R-CNN 相比,综合 *P*、*R*、mAP、F1 值、平均速度 5 项指标可得 K-SSD-F 算法对蝗虫识别计数最优。本文所提出的 K-SSD-F 算法检测和计数蝗虫时表现出较优的性能和准确性,对其他昆虫的视频目标检测具有良好的借鉴意义。

参 考 文 献

- [1] MANGEON S, SPESSA A, DEVESON E, et al. Daily mapping of Australian plague locust abundance[J]. Scientific Reports, 2020,10:151–169.
- [2] ZHAO L, HUANG W, CHEN J, et al. Land use/cover changes in the oriental migratory locust area of China: implications for ecological control and monitoring of locust area[J]. Agriculture Ecosystem & Environment, 2020, 303:107–110.
- [3] JI R, XIE B, LI Y, et al. Use of modis data to monitor the oriental migratory locust plague[J]. Agriculture Ecosystems & Environment, 2004, 104(3): 615–620.
- [4] YE S, ZHU D, YAO X, et al. Development of a highly flexible mobile GIS-based system for collecting arable land quality data [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014,7(11):4432–4441.
- [5] SU J, GU J, GUO X, et al. A method for recognition of locust age with few observations based on spectral analysis combined with chemical analysis[J]. Computers and Electronics in Agriculture, 2019, 162:450–458.
- [6] 刘德营,丁为民,陈坤杰. 野外环境昆虫图像自动采集装置[J]. 农业机械学报,2011,42(6):184–187.
LIU Deying, DING Weimin, CHEN Kunjie. Automatic acquisition system for insects images in field environment [J]. Transactions of the Chinese Society for Agricultural Machinery,2011,42(6):184–187. (in Chinese)
- [7] 肖德琴,张玉康,范梅红,等. 基于视觉感知的蔬菜害虫诱捕计数算法[J]. 农业机械学报,2018,49(3):51–58.
XIAO Deqin, ZHANG Yukang, FAN Meihong, et al. Vegetable pest counting algorithm based on visual perception [J]. Transactions of the Chinese Society for Agricultural Machinery,2018,49(3):51–58. (in Chinese)
- [8] 邱白晶,王天波,李娟娟,等. 黄瓜蚜虫的图像识别与计数方法[J]. 农业机械学报,2010,41(8):151–155.
QIU Baijing, WANG Tianbo, LI Juanjuan, et al. Image recognition and counting for glasshouse aphis gossypii[J]. Transactions of the Chinese Society for Agricultural Machinery,2010,41(8):151–155. (in Chinese)
- [9] 王茂林,荣二花,张利军,等. 基于图像处理的蓟马计数方法研究[J]. 山西农业科学,2020,48(5):812–816.
WANG Maolin, RONG Erhua, ZHANG Lijun, et al. Research on thrips counting method based on image processing [J]. Shanxi Agricultural Sciences,2020,48(5):812–816. (in Chinese)
- [10] SILVER D, HUANG A, MADDISON C, et al. Mastering the game of Go with deep neural networks and tree search[J]. Nature,2016,529:75–87.
- [11] KRIZHEVSKY A, SUTSKEVER I, HINTON G. Imagenet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84–90.
- [12] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017,39:1137–1149.
- [13] WEN C, WU D, HU H, et al. Pose estimation-dependent identification method for field moth images using deep learning architecture[J]. Biosystems Engineering, 2015,136:117–128.
- [14] 张银松,赵银娣,袁慕策. 基于改进 Faster R-CNN 模型的粘虫板图像昆虫识别与计数[J]. 中国农业大学学报,2019, 24(5):115–122.
ZHANG Yinsong, ZHAO Yindi, YUAN Muce. Insect recognition and counting of sticky insect board images based on improved Faster R-CNN model[J]. Journal of China Agricultural University,2019,24(5):115–122. (in Chinese)
- [15] 符首夫. 基于卷积神经网络的农作物昆虫识别与计数[D]. 镇江:江苏科技大学,2020.
FU Shoufu. Crop insect recognition and counting based on convolutional neural network[D]. Zhenjiang:Jiangsu University of Science and Technology, 2020. (in Chinese)
- [16] YE S, LU S, BAI X, et al. ResNet-Locust-BN network-based automatic identification of East Asian migratory locust species and instars from RGB images[J]. Insects, 2020,11(8):458.
- [17] XIN Mingyuan, WANG Yong. An image recognition algorithm of soybean diseases and insect pests based on migration learning and deep convolution network[C]//2020 International Wireless Communications and Mobile Computing,2020.
- [18] AI Y, SUN C, TIE J, et al. Research on recognition model of crop diseases and insect pests based on deep learning in harsh environments[J]. IEEE ACCESS, 2020, 171(8):686–693.
- [19] KANG K, OUYANG W L, LI H S, et al. Object detection from video tubelets with convolutional neural networks[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2016: 817–825.
- [20] 张霄峰. 被遮挡目标的检测与跟踪算法研究[D]. 济南:山东大学,2019.
ZHANG Xiaofeng. Research on the detection and tracking algorithm of occluded targets[D]. Jinan:Shandong University, 2019. (in Chinese)
- [21] WEI L, DRAGOMIR A, DUMITRU E, et al. SSD: single shot multibox detector[J]. Lecture Notes in Computer Science, 2016, 9905(2):21–37.
- [22] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. arXiv: 1804.02767.
- [23] BOCHKOVSKIY A, WANG C, LIAO H M, et al. YOLOv4: optimal speed and accuracy of object detection[J]. arXiv: 2004.20934.