

基于 YOLO v5s 和改进 SORT 算法的黑水虻幼虫计数方法

赵新龙¹ 顾臻奇¹ 李 军²

(1. 浙江理工大学信息科学与工程学院, 杭州 310018; 2. 台州学院智能制造学院, 台州 318000)

摘要: 目前农业环境下的无序目标的精确计数有很高的应用需求,这种计数对其生物量、生物密度管理起到了重要的指导作用。如黑水虻幼虫目标追踪过程中,追踪对象具有高速和非线性的特征,常规算法存在追踪目标速度不足和丢失目标后的再识别困难等问题。针对以上问题,本文提出了一种改进 SORT 算法,通过改进卡尔曼滤波模型的方式提升目标追踪算法的快速性和准确性,提升了计数的精度。另外,针对黑水虻幼虫目标识别过程中幼虫性状的多样性和混料导致的复杂背景问题,本文通过实验对比多种深度学习网络性能选定 YOLO v5s 算法提取图像多维度特征,提升了目标识别精度。实验结果表明:在划线计数方面,本文提出的改进 SORT 算法与原模型相比,平均精度从 91.36% 提升到 95.55%,提升 4.19 个百分点,通过仿真和实际应用,证明了本文模型的有效性;在目标识别方面,使用 YOLO v5s 模型在训练集上帧率为 156 f/s,mAP@0.5 为 99.10%,精度为 90.11%,召回率为 99.22%,综合性能优于其他网络。

关键词: 黑水虻幼虫; 目标识别; 目标追踪; 划线计数; YOLO v5s; SORT 算法

中图分类号: TP391.41 文献标识码: A 文章编号: 1000-1298(2023)07-0339-08

OSID:



Larvae of Black Soldier Fly Counting Based on YOLO v5s Network and Improved SORT Algorithm

ZHAO Xinlong¹ GU Zhenqi¹ LI Jun²

(1. School of Information Science and Engineering, Zhejiang Sci-tech University, Hangzhou 310018, China

2. Department of Intelligent Manufacturing, Taizhou University, Taizhou 318000, China)

Abstract: There is a high application demand for accurate counting of disordered targets in agricultural environments, and such counting plays an important guiding role in their biomass and biological density management. In the process of larvae of black soldier fly target tracking, the tracking object has the characteristics of high speed and non-linearity, and the conventional algorithm has the problems of insufficient speed of tracking target and difficulty of re-identification after losing the target. To address these problems, an improved SORT algorithm was proposed, which improved the speed and accuracy of the target tracking algorithm by improving the Kalman filter model, and enhanced the counting accuracy. In addition, for the complex background problem caused by larval trait diversity and mixing in the process of black gadfly larval target recognition, the target recognition accuracy was improved by experimentally comparing the performance of multiple deep learning networks, which selected YOLO v5s algorithm to extract multidimensional features of images. The experimental results showed that in terms of delineation counting, the improved SORT algorithm improved the average accuracy by 4.19 percentage points compared with the original model, from 91.36% to 95.55%, and the effectiveness of the model was proved through simulation and practical application. In terms of target recognition, using the YOLO v5s model on the training set achieved a frame rate of 156 f/s, mAP@0.5 value of 99.10%, accuracy of 90.11%, and recall rate of 99.22%. Its overall performance was better than other networks.

Key words: larvae of black soldier fly; target recognition; target tracking; scribe counting; YOLO v5s; SORT algorithm

收稿日期: 2022-12-15 修回日期: 2023-03-31

基金项目: 浙江省自然科学基金项目(LY22F030003)

作者简介: 赵新龙(1977—),男,教授,博士,主要从事精密驱动控制和智能控制研究,E-mail: zhaoxinlong@zstu.edu.cn

通信作者: 李军(1966—),男,教授,主要从事机电控制研究,E-mail: tzxyljun@126.com

0 引言

黑水虻幼虫又称“凤凰虫”，是一种全世界广泛分布的资源昆虫，其以禽畜粪便、生活废弃物为食，而其本身蛋白质、脂肪含量较高，能用于生产高价值的动物蛋白饲料，在全世界范围内被大量用于厨余垃圾、牲畜粪便等有机废物的无害化处理^[1-2]。有机废物转化效率受其生物密度影响，为达到效率的最大化，需要找到幼虫密度与有机废物处理速度之间的关系，因此需要对黑水虻幼虫进行计数^[3]。

过去常使用人工计数方式，近几年，随着机器视觉的广泛应用，基于机器视觉的目标计数成为一大研究方向。而黑水虻幼虫这类目标具有繁殖迅速、生物量巨大的特性，人工计数方式费时费力，所以基于机器视觉的目标计数方式有巨大的应用价值。在目前研究中，基于机器学习的检测计数有两种主要的实现方式：第 1 种为通过获得包含需求计数目标的图像数据集，将图像中的所有目标统计出来；第 2 种根据被识别目标的运动学特性实现计数。第 2 种方式自上而下需要经过目标识别和目标追踪两个过程，经过大量研究成果总结，第 1 种方式主要研究问题为：对大密度和重叠导致大量信息丢失的目标进行目标识别^[4-6]。第 2 种方式主要研究的问题为实现目标识别后的目标追踪任务^[7]。两种计数方式应用场景不同，各有优劣：第 1 种方式以张璐等^[8]对鱼群计数的研究为例，优点在于算法逻辑简单，但现实应用中存在图像过大导致处理速度下降的缺点。针对上述问题，本文主要进行第 2 种方式的研究。

关于目标识别，在应用于黑水虻幼虫的情境下，本文需要克服两个困难：饲料背景对幼虫目标的干扰以及幼虫聚集的情况。普通算法提取特征单一，无法考虑多维特征，但深度学习算法可以提取多维度特征，为此类情况的解决提供了方法。其中，基于卷积神经网络（Convolutional neural network, CNN）建立的目标识别算法，在农业采摘机械、光伏设备缺陷检测、各种机器人视觉中被大量应用^[9-13]。文献[14]在研究 Faster R-CNN（更快速区域卷积神经网络）检测方面的研究过程中，对比了 YOLO 网络和 Faster R-CNN 网络处理速度，相同条件下 YOLO 网络拥有比 Fast R-CNN 快一个数量级的速度优势。文献[15]通过改进 YOLO v3 网络提高了其对小目标检测的敏感性，使目标推理精度提高 2.36%，达到 92.43%。文献[16]使用 YOLO v3-tiny，为 YOLO 系列中结构最简单的网络，识别农田中的障碍物，精度和召回率也分别达到 95.5%

和 93.7%。最近，随着 Yolo 系列新网络模型 YOLO v6 和 YOLO v7 的推出，目标检测技术的研究和应用又取得了新的突破^[17-19]。其中，文献[20]证明了改进后的 YOLO v7 在小目标检测方面优于其他模型的效果。综上所述，YOLO 网络在目标识别的大量实际应用场景有较好的精度表现，对小目标具有很好的灵敏度，再加之其运行速度快，对应文中黑水虻幼虫个体目标小、数量多的特点，作为本研究的目标识别部分算法能达到较好的效果。

多目标追踪（Multiple object tracking, MOT）主要需要解决目标形态尺度变化、目标遮挡消失、重新出现后再次追踪这三大问题。由于其在实际应用中巨大的价值，在全世界有大量研究者，且提出了许多不同的方法解决这个问题，主要可以分为 3 种方向：基于排序算法（Simple online and realtime tracking, SORT）和 DeepSORT 的先检测后追踪的方式^[21-24]；基于 JDE 和 CenterTrack 的检测和追踪结合的方式；基于 TransTrack 和 TrackFormer 的注意力机制的方式^[25-26]。本文采用第 1 种目标追踪方式，其中 SORT 适用于处理高密度目标的情况，而 DeepSORT 是在 SORT 基础上发展而来的算法，它将深度学习技术应用于目标跟踪中，适用于需要进行多目标跟踪以及在遮挡和交叉行人等情况下进行目标追踪的应用场景。然而，针对黑水虻幼虫计数应用，其目标追踪在传送带上无遮挡，且由于幼虫密度高导致计数速率要求高，因此 SORT 算法更加适合当前的应用情况。文献[27]构建了标准 SORT 算法，在使用较小计算量的情况下，精度与最先进在线追踪器相当，同时，追踪器更新频率达到 260 Hz，为本文目标追踪实现提供了思路。

本文完成硬件平台搭建，根据硬件设备建立数据集。自上而下实现基于 YOLO v5s 深度学习网络黑水虻幼虫的目标识别任务和基于标准 SORT 目标追踪算法，结合两种算法实现上文的第 2 种计数方式。在目标识别算法方面，本文通过对比 YOLO v5s 与其他网络的性能差异，提供选择 YOLO v5s 的依据；在目标追踪方面，针对原始 SORT 算法目标追踪不稳定导致的计数精度不足问题，本文结合卡尔曼滤波原理，改进 SORT 目标追踪模型，通过仿真和应用实验，论证改进 SORT 模型在幼虫计数系统应用上的优越性。

1 材料与方法

1.1 硬件平台与实验数据

1.1.1 硬件平台搭建

为收集数据需要，设计并制造如图 1 所示设

备,设备主要由上料机构、传输机构、相机支架机构、支撑机构等 4 部分组成。使用深圳亿维 X1 - 32TD - E0 型可编程逻辑控制器(PLC)作为下位机控制机械部分动作,伺服电机选用禾川 SV - X3MM200A。使用个人计算机(PC)作为上位机进行图像处理,显卡为 NVIDIA RTX2060,显存 6 GB,CPU 型号为 Intel i7 - 11700 2.5 GHz,深度学习框架为 Pytorch (Python 3.9, Pytorch 1.10.2, torchvision 0.11.3)。



图 1 设备实物图

Fig.1 Physical drawing of equipment

1. 工业相机 2. 传输机构 3. 光源 4. 振动上料机构 5. 底座

1.1.2 实验数据收集

使用上述设备获取 5 段黑水虻幼虫计数视频作为本文计数结果测试数据集。另外,分别通过上述设备和手机拍摄的方式获取各 50 幅共 100 幅包含目标的图像作为训练集,构成中等规模的训练集用于 YOLO v5s 网络训练。

1.2 幼虫计数实现方法

为实现幼虫计数目标,需要研究基于 YOLO v5s 的目标识别、基于 SORT 的目标追踪和程序实现 3 个部分,各部分包含功能如图 2 所示。系统中首先使用 YOLO v5s 实现目标识别,再通过 SORT 算法实现被识别目标的身份标识号 (Identity document, ID) 指派,同一个 ID 目标分别有两帧画面在划线两侧则计数值增加,图 3 为划线计数逻辑示意图。

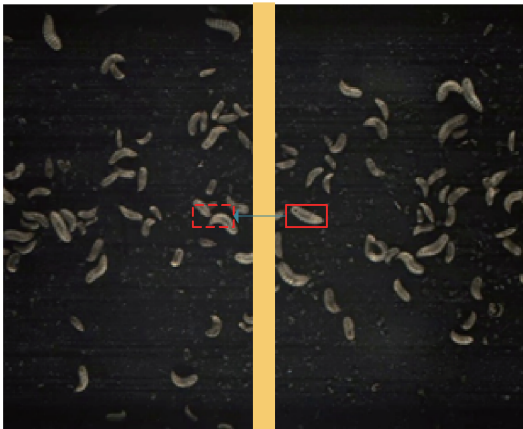


图 3 划线计数示意图

Fig.3 Schematic of line counting

1.2.1 基于 YOLO v5s 的目标识别实现

YOLO v5s 是 YOLO v5 网络中深度最小的网络,该网络在本文硬件配置情况下运行速度能够达到 156 f/s,目标识别的精度也能达到预期需求。YOLO v5s^[28]网络主要由 Backbone、Neck、Detect 网络组成,图 4 为其网络结构图。Backbone 由一系列卷积神经网络组成以提取图像特征,其第 1 层 Focus 层能减少模型的计算量,加快训练速度。Neck 是一系列图像特征混合组合的特征聚合层,将 Backbone 层传输来的特征降维排列,然后将输出的特征图传输到 Detect 层。Detect 层主要用于模型的最终检测,对上一层输出的特征图应用锚框,输出一个向量,包含被识别的目标位置、执行度以及类别,有 76 × 76、38 × 38、19 × 19 共 3 种尺寸的输出,分别用于对大、中、小 3 种尺度目标的

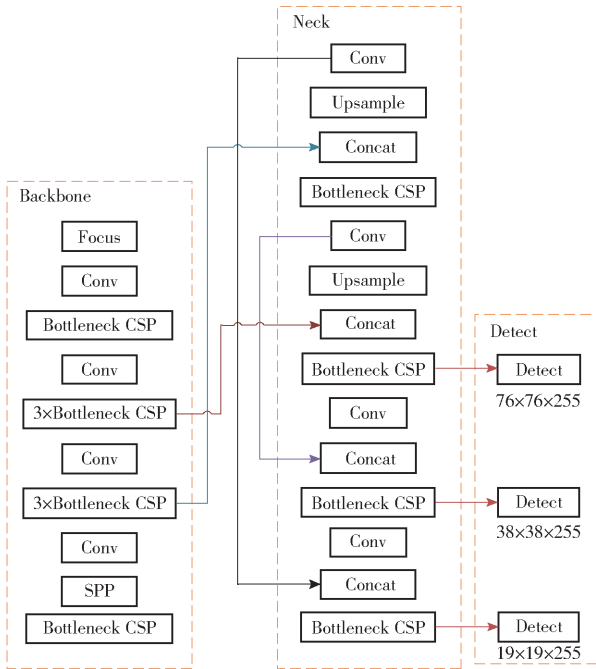


图 4 YOLO v5s 网络基本结构

Fig.4 Basic structure of YOLO v5s

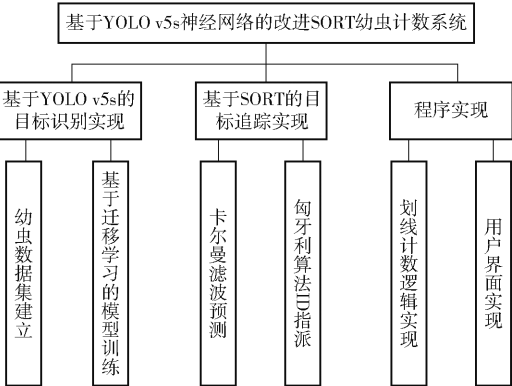


图 2 幼虫计数系统结构框图

Fig.2 Structure of larval counting system

识别,需提升对更小物体的敏感性可减少 19×19 输出层的大小。

1.2.2 基于 SORT 的目标追踪实现

图 5 为目标追踪算法实现过程的示意图。主要流程为:首先,获得 T_i 帧中目标位置;然后,使用卡尔曼滤波获得 T_i 帧目标在 T_{i+1} 帧时预测位置,如图 5 中 T_{i+1} 时刻的彩色框所示, T_{i+1} 时刻灰色目标为实际目标位置;最后,获得预测位置(彩色)与实际位置(灰色)的交并比(Intersection over union, IOU),通过匈牙利算法以上述交并比取非为代价矩阵得到前后两帧目标之间的对应关系。在主程序中实现过程:先通过目标识别算法获得包含 $[x1, y1, x2, y2, cls]$ 5 个特征的列表,然后将列表输入 SORT 算法后返回 $[x1, y1, x2, y2, ID]$ 列表,如前后两帧中相同 ID 的目标为同一目标。



图 5 SORT 算法追踪示意图

Fig. 5 Graphical representation of SORT tracking

(1)标准卡尔曼滤波跟踪模型实现
在 SORT 算法中使用标准卡尔曼滤波模型,其状态量公式为

$$\mathbf{X} = (x, y, s, r, \dot{x}, \dot{y}, \dot{s}) \tag{1}$$

式中 x ——目标框长轴中心位置
 y ——目标框宽轴中心位置
 s ——目标框面积
 r ——目标框长宽比
 $\dot{x}$ ——长轴中心位置变化率
 $\dot{y}$ ——宽轴中心位置变化率
 $\dot{s}$ ——面积变化率
 $\mathbf{X}$ ——状态向量
使用通用运动学模型离散化后得到

$$\mathbf{X}_k = \mathbf{F} \times \mathbf{X}_{k-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \times \mathbf{X}_{k-1} \tag{2}$$

式中 $\mathbf{F}$ ——状态转移矩阵
 $\mathbf{X}_{k-1}$ ——第 $k-1$ 时刻的状态向量
 $\mathbf{X}_k$ ——第 k 时刻的状态向量
观测量为卡尔曼预测输出结果,只需要状态量中前 4 个变量,得到卡尔曼滤波观测方程为

$$\mathbf{Z}_k = \mathbf{H} \times \mathbf{X}_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \times \mathbf{X}_k \tag{3}$$

式中 $\mathbf{H}$ ——观测矩阵
 $\mathbf{Z}_k$ ——第 k 时刻的观测向量
(2)交并比计算实现

两个目标框交集面积与并集面积的比值可以用于评定两个目标框重合的程度,即

$$I_{ou}(C, G) = \frac{A(C) \cap A(G)}{A(C) \cup A(G)} \tag{4}$$

式中 $A(C)$ ——目标框 C 面积
 $A(G)$ ——识别框 G 面积
 $I_{ou}(C, G)$ ——目标框 C 与识别框 G 的交并比

重合度与 IOU 的关系如图 6 所示。

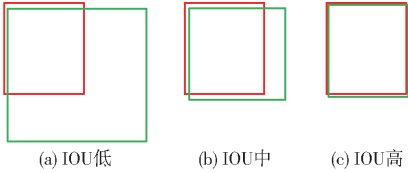


图 6 交并比计算示意图

Fig. 6 Illustrations of IOU calculation

(3)使用匈牙利算法解决 ID 指派问题

在卡尔曼滤波器正常工作的前提下,预测框与目标框的重合度越高,则 IOU 越大;反之,重合度越低,则 IOU 越小。匈牙利算法能够解决 $n \times m$ 维矩阵的最小代价匹配问题,使用 $\mathbf{M}_{iou}$ 作为匈牙利算法的代价矩阵,来匹配目标框和与其 IOU 最小的追踪框,实现两个框之间的最高重合度匹配。交并比矩阵 $\mathbf{M}_{iou}$ 为

$$\begin{aligned} \mathbf{M}_{IOU} = & \\ \begin{bmatrix} -I_{ou}(d_0, t_0) & -I_{ou}(d_0, t_1) & \cdots & -I_{ou}(d_0, t_m) \\ -I_{ou}(d_1, t_0) & -I_{ou}(d_1, t_1) & \cdots & -I_{ou}(d_1, t_m) \\ \vdots & \vdots & & \vdots \\ -I_{ou}(d_n, t_0) & -I_{ou}(d_n, t_1) & \cdots & -I_{ou}(d_n, t_m) \end{bmatrix} \end{aligned} \quad (5)$$

式中 $I_{ou}(d_n, t_m)$ ——第 n 个目标框和第 m 个识别框的交并比

1.2.3 改进卡尔曼滤波模型设计

图 7 为同时考虑运动模型预测位置 and 实际识别位置获得真实位置 $\hat{x}_k$ 的示意图。对于运动模型预测来说,本文假设真实位置 $\hat{x}_k$ 出现在预测位置 x_k 周围空间的概率满足高斯分布,因此加入满足高斯分布的噪声 w 。同时,本文认为实际检测位置受帧率不稳定、识别框抖动干扰而不稳定,也需要引入满足高斯分布的噪声 u 。运动模型预测位置 x_k 与实际识别位置 Z_k 两者高斯分布的重合部分就是真实位置 $\hat{x}_k$ 出现位置的分布区域,且满足高斯分布,如图中灰线所示。噪声 w 和 u 满足高斯分布

$$\begin{cases} p(w) \sim N(0, Q) \\ p(u) \sim N(0, R) \end{cases} \quad (6)$$

式中 $N(0, Q)$ ——均值为 0、协方差为 Q 的正态分布
 $N(0, R)$ ——均值为 0、协方差为 R 的正态分布
 $p(w)$ —— w 噪声概率
 $p(u)$ —— u 噪声概率

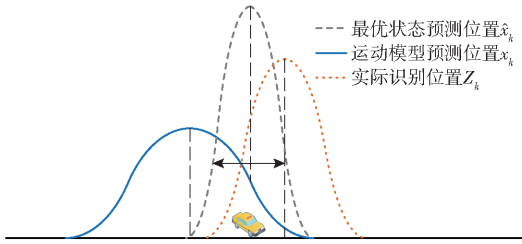


图 7 可能位置概率图
Fig. 7 Probability diagram of locations

根据卡尔曼滤波原理,每个预测循环中进行的滤波器更新公式为

$$\begin{cases} \mathbf{P}'_k = \mathbf{A}\mathbf{P}_{k-1} \times \mathbf{A}^T + \mathbf{Q} \\ K_k = \frac{\mathbf{P}'_k}{\mathbf{P}'_k + \mathbf{R}} \\ \hat{x}_k = x_k + K_k(Z_k - \mathbf{H}x_k) \\ \mathbf{P}_k = (1 - K_k)\mathbf{P}'_k \end{cases} \quad (7)$$

式中 $\mathbf{A}$ ——状态转移矩阵
 $\mathbf{P}_{k-1}$ ——第 $k-1$ 时刻的状态估计误差协方差矩阵
 $\mathbf{P}_k$ ——第 k 时刻的状态估计误差协方差矩阵
 $\mathbf{P}'_k$ ——第 k 时刻的协方差矩阵计算过渡矩阵

$\mathbf{Q}$ ——系统噪声协方差矩阵
 $\mathbf{R}$ ——测量噪声协方差矩阵
 K_k ——第 k 时刻的卡尔曼增益

下一帧位置更新公式为

$$x_{k+1} = \mathbf{A}\hat{x}_k + vt \quad (8)$$

式中 x_{k+1} ——第 $k+1$ 时刻的运动模型预测位置

针对实际应用中面对的匀速直线运动情况,传统的卡尔曼滤波模型可以综合考虑多维的运动变量来预测目标,但对于当前的应用对象,引入其他维度的预测信息只会对 IOU 产生负面影响,从而降低 ID 指派成功率。因此,本文建立一维运动模型,将状态转移矩阵 $\mathbf{A}$ 设置为 1,将观测矩阵 $\mathbf{H}$ 设置为 1,从而提高了卡尔曼滤波的目标追踪效果。

此外,随着传送带运动速度的增加,目标可能仅在画面中出现 2、3 帧,因此需要提高卡尔曼滤波器的响应速度。 $\mathbf{Q}$ 、 $\mathbf{R}$ 、 $\mathbf{P}$ 矩阵分别描述系统状态的随机扰动、观测值的噪声和状态估计误差的不确定性,降低这些矩阵的值可以提高滤波器对观测值的信任程度,从而使预测框更快地追踪上目标框。但在实际应用中,除了需要满足快速性外,当目标框的位置跳动较大时,卡尔曼滤波的输出也不能过于接近观测值。因此,需要通过实验和仿真确定一个折中的值,使预测框既能快速跟上目标框,又不会受到过多目标框位置跳动的影响。本文最终将 $\mathbf{Q}$ 、 $\mathbf{R}$ 、 $\mathbf{P}$ 的初值分别设置为 10、10、800。

2 实验结果与分析

2.1 YOLO v5s 网络模型与其他网络模型对比

在当前实验应用的前提下,假定配置环境相同且使用相同训练参数,YOLO v5s 和其他网络模型的训练结果如表 1(表中 mAP@ 0.5 表示 IOU 重叠阈值大于 0.5 的识别对象的平均精度)所示,本文使用的 YOLO v5s 网络经过训练后运行帧率达到 156 f/s, mAP@ 0.5 为 99.10%,精度(Precision, P)为 90.11%,召回率(Recall, R)为 99.22%。对于真实应用场景中的图像,选取 3 种 YOLO 网络模型进行推理,并对其效果进行比较,结果如图 8 所示。

表 1 不同网络模型实验结果对比

Tab. 1 Comparison of experimental results of different network models

网络模型	输入图像尺寸/ (像素×像素)	帧率/ (f·s ⁻¹)	mAP@ 0.5/%	精度/ %	召回 率/%
Faster R-CNN	640×640	9	88.06	80.19	88.05
YOLO v5s	640×640	156	99.10	90.11	99.22
YOLO v6s	640×640	145	93.50	92.43	94.90
YOLO v7	640×640	110	98.36	98.15	97.26

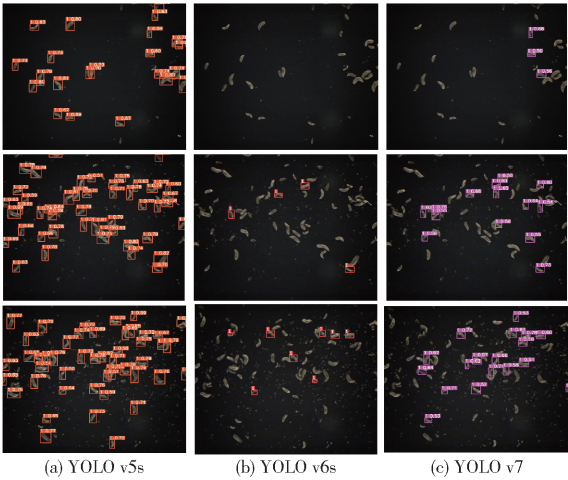


图 8 不同 YOLO 模型应用效果对比

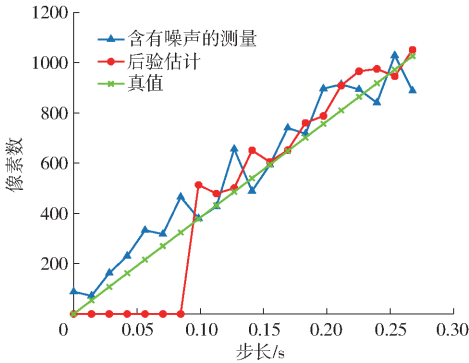
Fig. 8 Comparison of effect of different YOLO model applications

由表 1 和图 8 可知 YOLO 系列的网络性能优于 Faster - R CNN; 另一方面, 即使 YOLO v5s 精度低于其他两个 YOLO 网络, 但其帧率、mAP@ 0.5 和召回率却高于其他两个网络。在帧率方面, 由于本文提出的计数方法需要应用于工业现场, 实时性要求高, 因此 YOLO v5s 具有更大的优势。在召回率方面, 由于目标识别过程中无法被召回而丢失的个数将为计数过程引入绝对误差, 因此 YOLO v5s 具有更高的召回率。在精度和 mAP@ 0.5 方面, 其他两个网络的该值低于 YOLO v5s, 但它们全部识别目标的平均精度高于 YOLO v5s, 说明它们在训练后赋予了过高的精度于 IOU 重叠阈值较低的个体, 这表明模型的泛化能力较差。这与当前应用条件下训练集数量过少, 以及 YOLO v6s 和 YOLO v7 模型相对于 YOLO v5s 模型更加复杂有关。从图 8 的实际应用效果比较中, 也可以得出这一结论。综上所述, 在当前应用条件下, 相比其他网络模型, YOLO v5s 网络具有更好的应用效果。

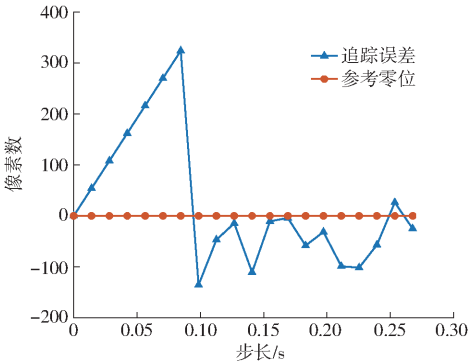
2.2 改进卡尔曼滤波追踪效果仿真与实际追踪效果

为了直观显示追踪效果, 假设目标运动的真实运动模型为从零点位置以速度 3 834 像素/s 匀速运动, 步长选择视频帧率的倒数 1/71 s。由于实际应用的需求, 需要在较少个步长内追上识别框, 因此在仿真实验中只需仿真 20 个循环。同时, 相机拍摄为离散情况, 导致物体出现第一帧不在坐标原点, 因此, 在仿真的第 9 个循环开始加入卡尔曼滤波追踪。

图 9a 为预测开始之后预测位置、真实位置、加入干扰的观测位置随时间的变化, 图 9b 为预测位置与真实位置的差值, 可以看到, 预测能快速逼近真实位置, 误差像素能稳定在系统正常运行允许的区间内。



(a) 卡尔曼滤波预测仿真结果



(b) 预测位置与真实位置误差

图 9 仿真实验结果

Fig. 9 Simulation experiment results

将通过仿真的改进卡尔曼滤波模型应用于 SORT 算法中, 实际黑水虻幼虫目标追踪效果如图 10 所示, 其中红色为预测框, 绿色为目标框。可见改进卡尔曼滤波模型实际追踪效果达到仿真预期: 预测框能快速跟上目标框, 同时, 在第 $k + 2$ 帧追踪目标位置发生跳变的情况下, 自动进行误差补偿, 在长时间内保持了稳定的追踪效果。

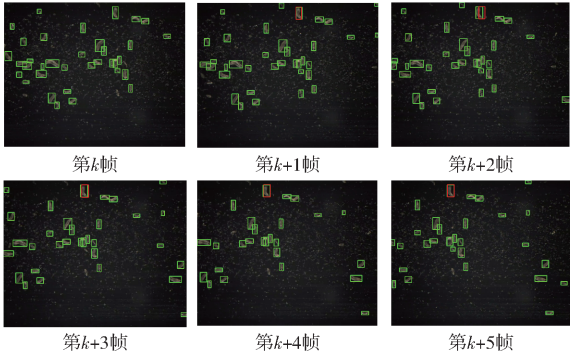


图 10 改进卡尔曼滤波预测效果

Fig. 10 Improved Kalman filter prediction

2.3 计数系统精度对比结果

使用同样的 YOLO v5s 目标识别网络, 将匈牙利算法的 3 个参数同样设置为 $\text{max_age} = 1$ 、 $\text{min_hits} = 0$ 、 $\text{iou_threshold} = 0.1$ 。由于上料密度会改变料的堆积、并排等情况, 导致目标识别准确度下降影响精度, 因此, 采集 5 段测试用数据集, 包含不同上料密度情况, 对比本文方法与基于标准 SORT 算法

的划线计数算法计数精度,具体结果如表 2、3 所示。

表 2 YOLO v5s + 标准 SORT 计数方式实验结果
Tab.2 YOLO v5s + standard SORT counting method

参数	实验序号				
	1	2	3	4	5
有效帧数	613	1 251	1 142	568	604
实际数量	1 030	1 030	1 030	1 030	1 030
计数数量	921	989	993	914	888
精度/%	89.42	96.02	96.41	88.74	86.21
平均精度/%	91.36				

表 3 YOLO v5s + 改进 SORT 计数方式实验结果
Tab.3 YOLO v5s + improved SORT counting method

参数	实验序号				
	1	2	3	4	5
有效帧数	613	1 251	1 142	568	604
实际数量	1 030	1 030	1 030	1 030	1 030
计数数量	997	1 027	1 010	964	923
精度/%	96.80	99.71	98.06	93.59	89.61
平均精度/%	95.55				

本文采用视频中包含幼虫目标的有效帧总数来定量判断实际上料密度,由于上料总数不变,有效帧越多,则单位帧平均幼虫数量越少,上料密度越小;反之,有效帧越少,则单位帧平均幼虫数量越多,上料密度越大。由表 2、3 可知,较大密度上料的计数精度往往差于较小密度上料的计数精度。同时,使

用本文改进卡尔曼滤波的 SORT 目标追踪计数方式不论大密度或是小密度上料的计数精度都高于标准 SORT 目标追踪计数方式。本方法计数最高精度达到 99.71%,平均精度达到 95.55%,满足系统设计要求。

3 结论

(1)针对黑水虻幼虫目标识别过程中需要克服的饲料与幼虫混合的干扰以及幼虫堆料,本文对比多种目标识别模型效果,最终选用 YOLO v5s 深度学习网络提取图像多尺度特征,使用中型规模的数据集训练出可用的目标识别网络,实验结果表明,其帧率为 156 f/s, mAP@0.5 为 99.10%,精度为 90.11%,召回率为 99.22%,达到本文需求的目标识别要求。

(2)通过实际实验,验证了使用标准卡尔曼滤波配合匈牙利算法实现 SORT 目标追踪为划线计数分配 ID 方式的可行性。

(3)针对标准卡尔曼滤波预测效果不佳导致计数精度较差的问题,本文提出了改进卡尔曼滤波模型。通过仿真和实际实验发现,基于本文模型的计数最高精度为 99.71%,平均精度为 95.55%,均高于使用标准卡尔曼滤波预测模型。因此,相比于标准卡尔曼滤波模型,本文提出的模型更适用于黑水虻幼虫划线计数系统。

参 考 文 献

[1] 喻国辉,陈燕红,喻子牛,等. 黑水虻幼虫和预蛹的饲料价值研究进展[J]. 昆虫知识,2009,46(1):41-45.
YU Guohui, CHEN Yanhong, YU Ziniu, et al. Research progress on the feed value of larvae of black soldier fly and pre-pupae [J]. Insect Knowledge, 2009, 46(1): 41-45. (in Chinese)

[2] 毛雪,包洁,赵前松,等. 黑水虻幼虫对餐厨垃圾无害化处理分析研究[J]. 现代农业研究,2022,28(7):108-110.
MAO Xue, BAO Jie, ZHAO Qiansong, et al. Analysis on harmless treatment of kitchen waste by black water fly larvae[J]. Modern Agriculture Research, 2022, 28(7): 108-110. (in Chinese)

[3] 王成成,贾昭炎,刘洋,等. 黑水虻转化有机生活废弃物相关研究进展[J]. 生物加工过程,2021,19(4):432-439.
WANG Chengcheng, JIA Zhaoyan, LIU Yang, et al. Progress in transformation of organic domestic waste by black soldier fly [J]. Chinese Journal of Bioprocess Engineering, 2021, 19(4): 432-439. (in Chinese)

[4] 张领先,陈运强,李云霞,等. 基于卷积神经网络的冬小麦麦穗检测计数系统[J]. 农业机械学报,2019,50(3):144-150.
ZHANG Lingxian, CHEN Yunqiang, LI Yunxia, et al. Detection and counting system for winter wheat ears based on convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(3): 144-150. (in Chinese)

[5] 田月媛,邓森磊,高辉,等. 基于深度学习的人群计数算法综述[J]. 电子测量技术,2022,45(7):152-159.
TIAN Yueyuan, DENG Miaolei, GAO Hui, et al. A review of crowd counting algorithms based on deep learning[J]. Electronic Measurement Technology, 2022, 45(7): 152-159. (in Chinese)

[6] 贾云舒. 基于深度学习的人群计数方法研究[D]. 成都:电子科技大学,2022.
JIA Yunshu. Research on crowd counting method based on deep learning[D]. Chengdu: University of Electronic Science and Technology, 2022. (in Chinese)

[7] 姚娟. 多目标跟踪算法及其在车辆计数中的应用研究[D]. 武汉:华中科技大学,2020.
YAO Juan. Multiple object tracking algorithm and its application in vehicles counting[D]. Wuhan: Huazhong University of Science and Technology, 2020. (in Chinese)

[8] 张璐,黄琳,李具备,等. 基于多尺度融合与无锚点 YOLO v3 的鱼群计数方法[J]. 农业机械学报,2021,52(增刊): 237-244.

- ZHANG Lu, HUANG Lin, LI Beibei, et al. Fish school counting method based on multi-scale fusion and no anchor YOLO v3[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(Supp.): 237–244. (in Chinese)
- [9] 宗泽, 刘刚. 基于机器视觉的玉米定位施肥控制系统设计与试验[J]. 农业机械学报, 2021, 52(增刊): 66–73.
ZONG Ze, LIU Gang. Design and experiment of maize fertilization control system based on machine vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(Supp.): 66–73. (in Chinese)
- [10] 向宽, 李松松, 栾明慧, 等. 基于改进 Faster RCNN 的铝材表面缺陷检测方法[J]. 仪器仪表学报, 2021, 42(1): 191–198.
XIANG Kuan, LI Songsong, LUAN Minghui, et al. Improved Faster RCNN-based surface defect detection method for aluminum[J]. Journal of Instrumentation, 2021, 42(1): 191–198. (in Chinese)
- [11] 陈立功, 李柯达, 张泽宇, 等. 基于计算机视觉的目标测距算法[J]. 弹箭与制导学报, 2020, 40(2): 93–96.
CHEN Ligong, LI Keda, ZHANG Zeyu, et al. Computer vision-based target ranging algorithm[J]. Journal of Projectiles, Rockets, Missiles and Guidance, 2020, 40(2): 93–96. (in Chinese)
- [12] 柳锋, 李俊峰, 戴文战. 基于深度学习语义分割的导光板缺陷检测方法[J]. 计算机系统应用, 2020, 29(6): 29–38.
LIU Feng, LI Junfeng, DAI Wenzhan. A deep learning semantic segmentation-based method for detecting defects in light guide panels[J]. Computer Systems and Applications, 2020, 29(6): 29–38. (in Chinese)
- [13] 李琦, 尚锋岚, 李宝山. 基于头部图像特征的草原羊自动计数方法[J]. 中国测试, 2020, 46(11): 20–24.
LI Qi, SHANG Jianglan, LI Baoshan. Automatic counting method for grassland sheep based on head image features[J]. Chin. Measur. Test. Technol., 2020, 46(11): 20–24. (in Chinese)
- [14] 曹之君, 张良. 基于 Faster-RCNN 的快速目标检测算法[J]. 航天控制, 2020, 38(4): 49–55.
CAO Zhijun, ZHANG Liang. Fast target detection algorithm based on Faster-RCNN[J]. Aerospace Control, 2020, 38(4): 49–55. (in Chinese)
- [15] 岳晓新, 贾君霞, 陈喜东, 等. 改进 YOLO V3 的道路小目标检测[J]. 计算机工程与应用, 2020, 56(21): 218–223.
YUE Xiaoxin, JIA Junxia, CHEN Xidong, et al. Improving YOLO V3 for road small target detection[J]. Computer Engineering and Applications, 2020, 56(21): 218–223. (in Chinese)
- [16] 陈斌, 张漫, 徐弘祯, 等. 基于改进 YOLO v3-tiny 的全景图像农田障碍物检测[J]. 农业机械学报, 2021, 52(增刊): 58–65.
CHEN Bin, ZHANG Man, XU Hongzhen, et al. Farmland obstacle detection in panoramic image based on improved YOLO v3-tiny[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(Supp.): 58–65. (in Chinese)
- [17] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J]. arXiv Preprint, arXiv:2207.02696, 2022.
- [18] LI C, LI L, GENG Y, et al. YOLOv6 v3.0: a full-scale reloading[J]. arXiv Preprint, arXiv:2301.05586, 2023.
- [19] LI C, LI L, JIANG H, et al. YOLOv6: a single-stage object detection framework for industrial applications[J]. arXiv Preprint, arXiv:2209.02976, 2022.
- [20] 戚玲珑, 高建瓴. 基于改进 YOLOv7 的小目标检测[J]. 计算机工程, 2023, 49(1): 41–48.
QI Linglong, GAO Jianling. Small object detection based on improved YOLOv7[J]. Computer Engineering, 2023, 49(1): 41–48. (in Chinese)
- [21] KAPANIA S, SAINI D, GOYAL S, et al. Multi object tracking with UAVs using deep SORT and YOLOv3 RetinaNet detection framework[C]//Proceedings of the 1st ACM Workshop on Autonomous and Intelligent Mobile Systems, 2020: 1–6.
- [22] PUNN N S, SONBHADRA S K, AGARWAL S, et al. Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and Deepsort techniques[J]. arXiv Preprint, arXiv:2005.01385, 2020.
- [23] 张宏鸣, 汪润, 董佩杰, 等. 基于 DeepSORT 算法的肉牛多目标跟踪方法[J]. 农业机械学报, 2021, 52(4): 248–256.
ZHANG Hongming, WANG Run, DONG Peijie, et al. Beef cattle multi-target tracking based on DeepSORT algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4): 248–256. (in Chinese)
- [24] 涂淑琴, 刘晓龙, 梁云, 等. 基于改进 DeepSORT 的群养生猪行为识别与跟踪方法[J]. 农业机械学报, 2022, 53(8): 345–352.
TU Shuqin, LIU Xiaolong, LIANG Yun, et al. Behavior recognition and tracking method of group housed pigs based on improved DeepSORT algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(8): 345–352. (in Chinese)
- [25] KOCUR V, FTACNIK M. Multi-class multi-movement vehicle counting based on CenterTrack[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 4009–4015.
- [26] SUN P, CAO J, JIANG Y, et al. Transtrack: multiple object tracking with transformer[J]. arXiv Preprint, arXiv:2012.15460, 2020.
- [27] BEWLEY A, GE Z, OTT L, et al. Simple online and realtime tracking[C]//2016 IEEE International Conference on Image Processing (ICIP). IEEE, 2016: 3464–3468.
- [28] YAN B, FAN P, LEI X, et al. A real-time apple targets detection method for picking robot based on improved YOLOv5[J]. Remote Sensing, 2021, 13(9): 1619.