基于增强型 Tiny-YOLOV3 模型的野鸡识别方法

易 诗¹,沈 练¹,周思尧¹,朱竞铭¹,袁学松²

(1. 成都理工大学信息科学与技术学院(网络安全学院、牛津布鲁克斯学院),成都 610059;2. 电子科技大学电子科学与工程学院,成都 610054)

摘 要:智慧农业病虫害检测技术发展迅猛,而对农作物具有危害的鸟类检测技术尚处于起步阶段,近年来由于生态改善,野鸡繁殖数量激增,其喜食小麦、玉米、红薯等农作物的种子与幼苗,对农业造成一定危害。该研究提出了一种适 宜于嵌入式系统部署的人工智能野鸡识别方法。由于在野外环境下移动平台上部署,需采用轻量级网络,同时保证检测 精度与实时性,因此,根据 Tiny-YOLOV3 轻量级目标检测网络基本结构,提出了一种针对野外复杂环境中出现野鸡的实 时检测网络-增强型轻量级目标检测网络(Enhanced Tiny-YOLO, ET-YOLO),该网络特征提取部分加深 Tiny-YOLOV3 特征提取网络深度,增加检测尺度以提高原网络目标检测精度,网络检测层使用基于 CenterNet 结构的检测方式以进一步 提高检测精度与检测速度。使用野外实地采集各种环境下出现的野鸡图像作为数据集,包括不同距离、角度、环境出现 的野鸡共计 6 000 幅高清图像制作数据集。试验结果表明,ET-YOLO 在视频中复杂环境下出现的野鸡平均检测精度达 86.5%,平均检测速度 62 帧/s,相对改进前 Tiny-YOLOV3 平均检测精度提高 15 个百分点,平均检测速度相对改进前 Tiny-YOLOV3 提高 2 帧/s,相对 YOLOV3、Faster-RCNN 与 SSD_MobileNetV2 主流代表性目标检测算法,平均检测精度 分别提高 1.5、1.1 与 18 个百分点,平均检测速度分别提高 38、47 与 1 帧/s。可高效实时地对复杂环境下出现的野鸡进行 识别,并且检测模型大小为 56 MB,适宜于在农业机器人,智能农机所搭载的嵌入式系统上部署。 关键词:农业;人工智能;嵌入式系统;野鸡识别;ET-YOLO; CenterNet 检测结构

doi: 10.11975/j.issn.1002-6819.2020.12.017

中图分类号: TN919.5

文章编号: 1002-6819(2020)-13-0141-07

易诗,沈练,周思尧,等.基于增强型 Tiny-YOLOV3 模型的野鸡识别方法[J].农业工程学报,2020,36(13):141-147. doi: 10.11975/j.issn.1002-6819.2020.13.017 http://www.tcsae.org

文献标志码: A

Yi Shi, Shen Lian, Zhou Siyao, et al. Recognition method of pheasant using enhanced Tiny-YOLOV3 model[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(13): 141-147. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2020.13.017 http://www.tcsae.org

0 引 言

随着中国退耕还林工程的实施山林面积大幅度增加,缺乏天敌控制,野鸡繁殖数量大增,南北方平原, 丘陵,山地分布普遍。由于野鸡喜食玉米、花生、小麦、 油菜等作物种子与果实,给农作物种植造成了巨大的经 济损失^[1]。农业生产中对于野鸡等危害农作物的鸟类传统 控制方法有架设假人模型,反光带,声音驱赶等^[2-3],该 类方法简单经济但效果有限。此外,下网、下夹、毒杀、 猎杀等方法非常危险,且野鸡为国家保护动物。近年来 智慧农业发展迅猛,发展一种结合人工智能的驱赶野鸡 方法有着实际意义,方法中最为重要的环节在于对野鸡 的发现与监测。

野鸡一般活跃于秋冬季与初春,06:00-09:00、16:00 -18:00,分布于田间、经济林、荒地等区域,季节性 单只或成群出现。对野鸡的检测与预警方法,主要需解 决如下几个问题:1)野鸡具有保护色,特别是雌性个

收稿日期: 2020-02-04 修订日期: 2020-05-27

体。2)野外复杂环境,野鸡出现区域普遍存在一定草木的遮挡。3)野鸡出现的距离和目标大小随时变化。

近年来,在目标检测方面,基于深度学习的方法准 确率大大超过了传统的人工设计特征的检测方法。基于 深度学习的目标检测主要包括两类,一类是基于区域生 成的卷积网络结构,代表性的网络为区域卷积网络 Region-CNN (RCNN) 、 更 快 速 卷 积 神 经 网 络 Faster-Region-CNN(Faster-RCNN)^[4-5]; 一类是把目标 位置的检测视作回归问题,直接利用 CNN 网络结构图像 进行处理,同时预测出目标的类别和位置,代表性的网 络有单次多框检测器 Single Shot MultiBox Detector (SSD)、单次检测器 You Only Look Once (YOLO)^[6-8] 等。智慧农业目标检测方面,熊俊涛等^[9]采用Faster-RCNN 实现了自然环境下绿色柑橘的识别。赵德安等[10]采用 YOLOV3 实现了水下摄像头采集的河蟹图像的识别。吕 石磊等^[11]采用改进 YOLOv3-LITE 轻量级神经网络进行 柑橘识别。燕红文等^[12]采用改进 Tiny-YOLO 模型进行 群养生猪脸部姿态检测。易诗等[13]采用红外热成像与改 进的 YOLOV3 进行夜间野兔识别。其中 Tiny-YOLOV3 轻量级目标检测网络具备对目标检测精度高,速度快, 模型精简合适与嵌入式计算平台部署的优点。但对野外 复杂环境中出现的野鸡识别存在以下问题:1)遮挡条件

基金项目:国家自然科学基金项目(61771096);国家大学生创新创业项目(201910616129)

作者简介:易诗,高级实验师,主要从事人工智能,红外图像处理,信号处理,嵌入式系统开发。Email: 549745481@qq.com

下目标的检测精度需提高。2)尺度变化的运动目标检测 精度需进一步提升。3)目标检测模型的精度,速度与体 量需进一步调整平衡。

因此,本研究针对野外复杂环境中进行野鸡识别的 特点,在Tiny-YOLOV3 轻量级目标检测网络的基础结构 上提出了一种增强型轻量级目标检测网络(Enhanced Tiny-YOLO,ET-YOLO)。该网络根据野外野鸡活动的 特点,在Tiny-YOLOV3 网络的基础结构上加深特征提取 网络深度,增加检测尺度以提高对运动中尺度与姿态变 化目标的检测精度,在网络检测层根据参考文献[14]采用 CenterNet结构替代原网络的YOLO检测层以提高检测速 度,增强特征检测能力提高对部分遮挡目标(半身、背 部、头部)的检测精度。经试验验证,增强的轻量级目 标检测网络经训练生成的检测模型适宜于部署嵌入式计 算平台,在野外环境下高效地进行野鸡目标识别。

1 试验数据

1.1 试验数据采集

由于公开野鸡监测数据集的缺乏,开源鸟类分类数据集不适宜于针对野鸡的监测。本文采用野外拍摄野鸡在各种环境,距离下出现的视频制作野鸡监测数据集,数据集的采集时间为野鸡活跃的秋冬季及初春 06:00-09:00、16:00-18:00,采集地点为中国四川省绵阳市境内农业大县梓潼、盐亭、剑阁,采集环境包括麦田、玉米地、荒地、树林等野鸡常见出没区域。由于野鸡性情谨慎,出没路线无规律,因此,采集平台采用加载野鸡叫声的引鸟放音器引诱野鸡出现,配合运动相机远程录制 2 K 高清视频,放置于杂草丛边或悬挂于小树枝上进行图像采集。

实地采集视频制作的野鸡监测数据集共计 6000 幅 高清野鸡图像,根据智能农机及农业机器人搭载摄像头 拍摄能力与深度学习目标检测网络检测能力,采集距离 为5~50 m,包括正面、侧面、背面,各种姿态与尺度的 野鸡样本,同时采集实际情况下存在遮挡环境与保护色 融入的野鸡样本,提高对部分遮挡情况下野鸡目标的监 测精度。数据集中部分样本如图1所示。







头部 Head



半身 Half body



背部 Back b. 不同采集姿态 b. Different acquisition postures 图 1 野鸡监测样本

Fig.1 Monitoring samples of pheasant

1.2 数据集制作

使用 OpenCV 工具对原始数据集进行增强,包括水 平方向上正负 45°的随机旋转、随机水平/垂直旋转、剪裁 与缩放,对每张图片随机采用以上一种方式进行扩增, 最终得到 12 000 幅图像的增强数据集。

对增强后数据集使用 labelImg 工具对检测目标进行标注。按 5:1 划分训练集与测试集,样本中野鸡出现各类情况如表 1 所示。

表1	野鸡样本数量

Table I Pheasant sample quantity								
数据集 Dataset	正常姿态	特殊姿态	遮挡隐藏					
	Normal posture	Special posture	Occlusion concealment					
训练集 Training set	6 563	938	2 500					
测试集 Test set	1 312	187	500					
总计 Total	7 875	1 125	3 000					

2 检测模型

2.1 Tiny-Y0L0V3 网络模型

Tiny-YOLOV3目标检测模型为YOLO系列中针对移动端与边缘计算设计的轻量级目标检测网络,特点为精简,高实时性,检测精度较高。

Tiny-YOLOV3 主干网络为 YOLOV3 的简化,由 13 个卷积层,6个最大池化层,1个连接层构成,结构如图 2 所示。



注: Conv: 卷积操作: Maxpool: 最大池化操作; Concatenate: 连接操作; Upsample: 上采样操作; YOLO detection: YOLO 检测层; NMS: 非极大值 抑制。416×416、208×208、26×26、13×13 分别指输入图像与特征图分辨率, 16、256、128、24 为卷积层深度。下同。

Note: Conv: Convolution operation; Maxpool: Max pooling operation; Concatenate: Concatenate operation, Upsample: Upsampling operation, YOLO detection: YOLO detection layer; NMS: Non-maximum suppression. 416×416, 208×208, 26×26 and 13×13 refer to the resolution of input image and feature map respectively. 16, 256, 128 and 24 as convolution layer depth. Same as below.

图 2 Tiny-YOLOV3 网络结构图 Fig.2 Tiny-YOLOV3 network structure diagram Tiny-YOLOV3 目标检测网络输入图像分辨率 416×416,经一系列卷积,最大池化,连接,上采样操作 对图像进行特征提取,YOLO 检测层根据输入的图像数 据,分别在13×13,26×26,这2个检测尺度上运用回归 分析法输出图像数据的多个滑动窗口位置及该窗口中检 测到的目标类别,设置多个锚点框以检测目标,完成目 标检测后引入非极大值抑制^[15]剔除冗余的检测框以使得 对于每个目标均有唯一检测框,使其位置信息更准确, 置信度更高。

由于 Tiny-YOLOV3 目标检测网络较为精简, 虽实时 性高,适宜于嵌入式边缘计算平台部署,但其特征提取 操作精简,检测尺度较少,难以准确检测运动中呈现多 种姿态,保护色掩护,具有一定隐藏性的野鸡目标。因 此应对其做出改进与增强,使其适宜于复杂环境下出现 的野鸡目标的检测。

2.2 ET-YOL0 网络模型

为提高 Tiny-YOLOV3 目标检测网络对野外复杂环 境中具有保护色,遮挡以及运动中姿态变化的野鸡目标 检测精度同时保证模型的精简性与检测的实时性,本文 提出的 ET-YOLO 增强型轻量级目标检测网络,在 Tiny-YOLOV3 目标检测网络基础上加深主干特征提取网 络,增强检测尺度,改进网络检测层,提升其目标检测 性能,提升实时性。

根据文献[16]提出的主干特征提取网络改进方法,能 够在 Tiny-YOLOV3 的特征提取主干网络上增强多尺度 目标的检测能力,由于 Tiny-YOLOV3 目标检测网络浅层 卷积 Conv4 层能够较为有效地表征在隐藏环境下野鸡部 分身体出现所呈现的小目标的语义信息,而 Conv3 层特 征感受野太小, Conv5 层则特征感受野太大, 包含一定 的背景噪声干扰。因此,为提升遮挡情况下小目标检测 能力, ET-YOLO 目标检测网络将 Tiny-YOLOV3 中 Maxpool3 层替换为 Conv4 层, 增加 Conv5 层将 Conv4 层的特征通道维度压缩以减少无效参数;同时在 Tiny-YOLOV3 模型的基础上增加上采样层 Upsample2, 将 Conv5 层 和 Upsample2 在通道维度上进行连接操作, Conv7 通过 Upsample2 进行升采样操作,形成特征金字 塔 增 加 的 一 个 特 征 图 层 , 特 征 金 字 塔 网 络 由 Tiny-YOLOV3 的 13×13 和 26×26 像素提升为 13×13、 26×26 和 52×52 像素的 3 个检测尺度结构以更好地检测多 尺度目标。在检测网络部分,使用 CenterNet 结构代替原 有 YOLO 层,优化检测方法,增强特征检测能力以提高 对部分遮挡目标的检测精度,降低误检率,降低运算量, 提高检测实时性。ET-YOLO网络结构如图 3 所示。

为提高对隐藏环境中野鸡目标的检测精度,提高检测实时性,降低运算量,检测网络部分使用 CenterNet 结构,用无锚点方式代替原有基于锚点的 YOLO 检测层。 使改进后的目标检测网络检测速度与特征检测能力进一步提升。

CenterNet 是 Zhou 等^[17]于 2019 年提出的目标检测最 新方法,其基本思想为:基于锚点的方法经常出现大量 不正确的边界框,由于缺乏对相关剪裁区域的额外监督 所造成,因此原 Tiny-YOLOV3 在 YOLO 检测层中需大量锚点框进行目标检测。而 CenterNet 检测层的优势在于: 1) CenterNet 基于目标中心点的预测,无锚点框概念,无 框重叠阈值门限以确定正锚点与负锚点,同时无需要区 分锚点框是物体还是背景。因为每个目标只对应一个中 心点,这个中心点是通过热力图预测,因此后端无需极 大值抑制再进行筛选。2) CenterNet 结构简单,直接检测 目标的中心点和大小,为真正意义上的无锚点检测。因 此在降低运算量,提高实时性的基础上同时提高了目标 特征检测能力,提高了检测精度。



注: 52×52 指输入图像与特征图分辨率。

Note: 52×52 refers to the resolution of input image and feature map respectively.

图 3 ET-YOLO 网络结构图 Fig.3 ET-YOLO network structure diagram

CenterNet 检测原理为:输入图像 $I \in R^{W \cdot H \times 3}$,其中 R为图像缩放比,W、H为图像宽与高,检测目标首先在于生成关键点热力图 Y如式 (1)所示:

$$Y \in [0,1]^{\frac{p}{R-R}C} \tag{1}$$

式中 C 为关键点类型数(即输出特征图通道数)对于野鸡监测这个单目标检测问题 C 为1。热力图中,中心点 p 计算如式(2)所示:

$$p = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right) \tag{2}$$

式中 x_1 、 x_2 、 y_1 、 x_2 分别为标注目标的横纵坐标边界点。 对于下采样后的中心点确定为 $p_c = \frac{p}{R}$ 使用高斯函数将关 键点分布于特征图上,如式(3)所示:

$$Y_{xyc} = \exp(-\frac{(x - p_{cx})^2 + (y - p_{cy})^2}{\sigma_p^2})$$
(3)

式中 *Y_{xyc}* 为热力图中关键点分布函数, *x* 为横坐标取值, *y* 为纵坐标取值, *p_{cx}* 为横坐标中心点, *p_{cy}* 为纵坐标中心

点, σ_n为中心点标准差。

CenterNet 检测层采用焦点损失的思想,在训练中,中心点的周围其他点的损失则是经过衰减后的损失,而目标的长和宽是经过对应当前中心点的长和宽回归计算所得,中心点的损失函数 *L*_k 如式(4)所示:

$$L_{k} = -\frac{1}{N} \cdot \begin{cases} (1 - \frac{Y_{xyc}}{R})^{\alpha} \cdot \ln(\frac{Y_{xyc}}{R}) & Y_{xyc} = 1\\ (1 - Y_{xyc})^{\beta} \cdot (\frac{Y_{xyc}}{R})^{\alpha} \cdot \ln(1 - \frac{Y_{xyc}}{R}) & Y_{xyc} \neq 1 \end{cases}$$
(4)

式中 α 和 β 是焦点损失的超参数,根据参考文献[17]分别 取 2 与 4, *N*是图像*I*的关键点数量,用于将所有的正焦 点损失标准化为 1。这个损失函数是焦点损失的改进型, 适用于 CenterNet 检测。

假设 $(x_1^{(k)}, y_1^{(k)}, x_2^{(k)}, y_2^{(k)})$ 为目标 k 的边界角点 坐标 ,目标大小的损失的中心点 p_k 计算如式 (5) 所示:

$$p_{k} = \left(\frac{x_{1}^{(k)} + x_{2}^{(k)}}{2}, \frac{y_{1}^{(k)} + y_{2}^{(k)}}{2}\right)$$
(5)

使用关键点 $\frac{Y_{xyc}}{R}$ 预测去预测所有的中心点。然后对每个目标 k 的大小进行回归计算,回归函数 S_{k} =

($x_2^{(k)} - x_1^{(k)}, y_2^{(k)} - y_1^{(k)}$),该值在训练前提前计算,即进行 了下采样之后的长宽值。因此目标大小损失函数 L_{size} 采用 一点绝对损失函数 (L1 loss)监督长与宽的回归,计算如 式 (6) 所示:

$$L_{size} = \frac{1}{N} \cdot \sum_{k=1}^{N} \left| \frac{S_k}{R} - S_k \right| \tag{6}$$

CenterNet 检测层整体的损失函数为中心点损失与大 小损失之和,每个损失都有相应的权重。计算如式(7) 所示。

$$L = L_k + 0.1L_{size} \tag{7}$$

式中L为CenterNet检测层整体损失函数。

综上所述,使用 CenterNet 检测层的 ET-YOLO 野鸡 检测网络首先根据热力图中关键点分布回归计算确定野 鸡目标的中心点,再根据中心点位置回归计算确定目标 大小范围,最终确定野鸡目标。

2.3 模型的训练与测试

2.3.1 试验平台

模型训练与测试所使用硬件平台为 Core i7-8750H 2.2 Ghz 处理器+16 GB 内存+Geforce GTX 1080 8 GB 显 卡, 软件平台为 Win10+tensorflow1.9.0+CUDA9.2+ VS2017+opencv4.0。

2.3.2 评价指标

对于野外复杂环境下野鸡目标的识别,需考虑检测 网络的精度与实时性,同时由于需要部署于智能农机或 农业机器人的嵌入式边缘计算平台,要求检测模型轻量化。

本文采用平均精度(Mean Average Precision, mAP) 作为模型检测精度的评价指标,(Average Precision, AP) 作为每一类别的检测精度的评价指标^[18]。mAP 和 AP 与 准确率(Precision, *P*)、召回率(Recall, *R*)有关,准 确率和召回率的计算如下:

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{8}$$

$$R = \frac{\text{FP}}{\text{FP+TN}} \tag{9}$$

式中 TP 代表被正确划分到正样本的数量,FP 代表被错误划分到正样本的数量,FN 代表被错误划分到负样本的数量。通过计算所得准确率与召回率可以绘制出准确率-召回率曲线,该曲线以召回率为横坐标,以准确率为纵 坐标,代表某一类别的准确率与召回率情况,对该曲线 进行积分所得。野鸡识别为单目标检测 mAP 计算如式 (10) 所示:

$$mAP = \int_0^1 (P \cdot R) dR \tag{10}$$

平均运算速度评价指标,模型权重大小评价指标由 实际训练与测试得到。

2.3.3 ET-YOLO 网络训练

在 ET-YOLO 网络模型训练环节,优化器使用 SGD (Stochastic Gradient Descent)^[19]算法,将数据集中的图像 样本全部转换为 416×416 像素的图像,训练时以 24 幅图 像为一个批次进行小批量训练,训练一批图像,权值进 行一次更新。权值的衰减速率设为 0.000 5,动量设置为 0.9,初始学习率设为 0.001,对 ET-YOLO 网络进行 5 000 次迭代,每间隔 500 次迭代后保存一次模型,最终选取 精度最高的模型^[20]。

3 结果与分析

为验证本研究提出的增强型轻量级 ET-YOLO 野鸡 检测网络对野外复杂环境下出现的野鸡目标识别的有效 性与优势,在对比改进 Tiny-YOLOV3 前后各项指标的同 时,使用最新轻量级目标检测网络 SSD_MobileNetV2, 以及目前具有代表性的目标检测网络 YOLOV3、 Faster-RCNN采用相同数据集与训练参数进行模型训练, 并在测试集上进行对比测试^[21-24]。

具有保护色融入环境以及隐藏遮挡的野鸡目标在野 外复杂环境中出现较为常见,因此首先选取测试集中在 此类情况下出现的野鸡目标进行检测识别,检测识别结 果对比如图4所示。

由结果可知,草丛中隐藏具备保护色掩护的野鸡目标共计4个,Faster-RCNN与本文提出的ET-YOLO目标检测网络检出率最高,识别出全部4个野鸡目标。而Tiny-YOLOV3、SSD_MobileNetV2与YOLOV3仅识别出隐藏情况下4个野鸡目标中的3个。

在野外实际环境中,野鸡目标常以不同的姿态出现,存在保护色,对检测识别造成一定难度,因此再选取测 试集中特殊姿态下出现的野鸡目标进行检测,检测识别 结果对比如图 5 所示。

由结果可知,在该测试图像中存在 2 个野鸡目标, 其中1个为无遮挡正常姿态,1个俯卧于地面,这种姿态 加以保护色掩护,对检测识别造成一定难度。所有检测 网络均可成功识别无遮挡正常姿态的野鸡目标, YOLOV3、Tiny-YOLOV3 与本文提出的 ET-YOLO 可正

确识别出特殊姿态的野鸡目标,而 Faster-RCNN、 SSD MobileNetV2 无法成功识别出特殊姿态的野鸡 目标。

在野外农田中时常出现其他对农作物危害较小或者







图 4 隐藏情况下野鸡目标识别 Fig.4 Object recognition of pheasant in hidden condition

c Faster-RCNN

无危害的鸟类,而其特征与野鸡目标相似,因此检测网 络不能误识别该类目标。最后一步测试,选取采集视频 中同一环境下出现的其他鸟类,对比测试各目标检测网 络的误识别性能,识别结果对比如图6所示。





e ET-YOLO

注: SSD_MobileNetV2、YOLOV3、Faster-RCNN、Tiny-YOLOV3、ET-YOLO为5种目检测网络。下同。 Note: SSD_MobileNetV2, YOLOV3, Faster-RCNN, Tiny-YOLOV3, ET-YOLO are five object detection networks. Same as below.

a. SSD_MobileNetV2







c. Faster-RCNN

Object recognition of pheasant in special posture

图 5 特殊姿态野鸡目标识别



e. ET-YOLO





Fig.5

h YOLOV3



d. Tiny-YOLOV3



e. ET-YOLC

图 6 其他鸟类出现情况下的目标识别 Fig.6 Object recognition under the occurrence of other birds

由结果可知,测试图像中出现目标为1只画眉,其 与野鸡同为农田环境下容易出现的鸟类,特征具有一定 相似性,YOLOV3、Faster-RCNN将其误识别为野鸡,而 Tiny-YOLOV3、SSD_MobileNetV2 以及本文提出的 ET-YOLO 没有产生误识别^[25-30]。

经完整测试后,对各个目标检测网络检测野鸡目标 的准确率 P, 召回率 R, 平均精度 mAP, 平均运算速度 (Mean Operation rate, MO),模型大小的统计如表2所示。

	表 2	野鸡不同检测网络的综合对比		
Table 2	Comprehensive comparison of different detect			
		networks for nheasant		

networks for pheasant									
检测网络 Detection network	准确率 Precision /%	召回率 Recall /%	平均精度 Mean average precision/%	平均运算速度 Mean operation rate/(帧·s ⁻¹)	模型大小 Model size/MB				
SSD_MobileNetV2	84.0	81.5	68.5	61	28				
YOLOV3	93.5	90.5	85.0	24	234				
Faster-RCNN	93.8	91.5	85.4	15	265				
Tiny-YOLOV3	86.0	83.0	71.5	60	34				
ET-YOLO	94.5	92.5	86.5	62	56				

根据综合对比测试统计可见,本文提出的 ET-YOLO 野鸡检测网络,对于野鸡目标的检测在平均精度上高于 增强前 Tiny-YOLOV3 轻量级目标检测网络 15 个百分点, 分别高于 SSD MobileNetV2、YOLOV3 与 Faster-RCNN 目标检测网络18, 1.5 与1.1 个百分点, 对复杂环境下出 现的野鸡目标检测精度最为优良。在检测实时性上,由 于在检测层使用了基于中心点的 CenterNet 结构,降低了 计算量,后端无需极大值抑制处理,因此在主干特征提 取网络增加了标准卷积层和上下融合层的情况下平均检 测速度相对增强前 Tiny-YOLOV3 轻量级目标检测网络 提高 2 帧/s, 相对 SSD MobileNetV2、YOLOV3 与 Faster-RCNN 目标检测网络分别提高 1、38 与 47 帧/s, 对于实时检测移动的野鸡目标效果最为优良。在模型大 小方面,由于为提高检测精度增加了特征提取网络深度 与检测尺度,因此 ET-YOLO 生成检测模型大小 56 MB, 高于 Tiny-YOLOV3 与 SSD MobileNetV2 轻量级目标检 测网络模型 22 与 28 MB, 但相对 YOLOV3 与 Faster-RCNN 深度目标检测网络 200 MB 以上的模型大 小, ET-YOLO 在保持最优野鸡目标检测精度的前提下模型大小控制在 100 MB 以内,仍属于轻量级目标检测网络,适宜于在智能农机与农业机器人的嵌入式边缘计算平台上部署。

4 结 论

1)本文所提出的针对野外复杂环境中出的野鸡目标 检测网络 ET-YOLO 对野鸡目标识别具备较高的准确性, 试验结果表明,模型平均精度达到了 86.5%,平均检测速 度达 62 帧/s,模型大小 56 MB。

2)针对野外复杂环境中出的野鸡目标特点对 Tiny-YOLOV3轻量级目标检测网络结构做出改进,加深 特征提取网络深度,增加检测尺度,提高小目标,隐藏 目标检测精度,在网络检测部分采用 CenterNet 层检测结 构替原有 YOLO 层检测结构,提高了检测速度与精度。

3) 对实际复杂环境下出现的野鸡目标检测测试中, 本文方法表现出较高检测精度性与实时性,平均检测率 高于增强前 Tiny-YOLOV3 轻量级目标检测网络 15 个百 分点,分别高于 SSD_MobileNetV2,YOLOV3 与 Faster-RCNN 目标检测网络 18、1.5 与 1.1 个百分点,检 测速度方面相对增强前 Tiny-YOLOV3 轻量级目标检测 网络提高 2 帧/s,相对 SSD_MobileNetV2、YOLOV3 与 Faster-RCNN 目标检测网络分别提高 1、38 与 47 帧/s,且 检测模型大小控制在 100 MB 以内,适宜于嵌入式边缘计算 平台的部署。

[参考文献]

- 卢振宇,孙连新,黄志银,等.野鸡对春玉米的危害及综防策略研究[J].农业科技通讯,2017(7):65-67. Lu Zhenyu, Sun Lianxin, Huang Zhiyin, et al. Study on the harm of pheasant to spring corn and its integrated control strategy[J]. Agricultural Science and Technology Communication, 2017(7): 65-67. (in Chinese with English abstract)
- [2] 国家森防总站. 2004 年全国林业有害生物发生趋势预测
 [EB/OL]. 2004-02-19 [2019-05-23]. http://gssfz. nease. net/4. html.

orchard[J]. Northern Horticulture, 2012(3): 134-135. (in Chinese with English abstract)

- [4] He K M, Zhang X Y, Ren S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [5] Girshick R. Fast R-CNN [C]//IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 1440-1448.
- [6] 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. Columbus: IEEE, 2014: 580-587.
- [7] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//European Conference on Computer Vision. Cham: Springer, 2016: 21-37.
- [8] Redmon J, Farhadi A. Yolov3: An incremental improvement [EB/OL]. (2018-04-08) [2018-09-07]. org/abs/1804. 02767.
- [9] 熊俊涛,刘振,汤林越,等.自然环境下绿色柑橘视觉检测技术研究[J].农业机械学报,2018,49(4):45-52.

Xiong Juntao, Liu Zheng, Tang Linyue, et al. Research on visual inspection technology of green citrus in natural environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(4): 45-52. (in Chinese with English abstract)

- [10] 赵德安,刘晓洋,孙月平,等.基于机器视觉的水下河蟹 识别方法[J]. 农业机械学报,2019,50(3):151-158.
 Zhao Dean, Liu Xiaoyang, Sun Yueping, et al. Detection of under water crabs based on machine vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(3):151-158. (in Chinese with English abstract)
- [11] 吕石磊,卢思华,李震,等. 基于改进 YOLOv3-LITE 轻量级神经网络的柑橘识别方法[J]. 农业工程学报,2019,35(17):205-215.
 Lü Shilei, Lu Sihua, Li Zhen, et al. Orange recognition method using improved YOLOv3-LITE lightweight neural network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(17):205-215. (in Chinese with English abstract).
- [12] 燕红文,刘振宇,崔清亮,等.基于改进 Tiny-YOLO 模型的群养生猪脸部姿态检测[J].农业工程学报,2019,35(18): 169-180.
 Yan Hongwen, Liu Zhenyu, Cui Qingliang, et al. Detection of facial gestures of group pigs based on improved Tiny-YOLO[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(18): 169-180. (in Chinese with English abstract)
- [13] 易诗,李欣荣,吴志娟,等.基于红外热成像与改进 YOLOV3 的夜间野兔监测方法[J]. 农业工程学报,2019, 35(19): 215-222.
 Yi Shi, Li Xinrong, Wu Zhijuan, et al. Night hare detection method based on infrared thermal imaging and improved YOLOV3[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(19): 215-222. (in Chinese with English abstract)
- [14] Duan Kaiwen, Bai Song, Ling Xixie, et al. CenterNet: Keypoint triplets for object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition. United States: IEEE, 2019: 3668-3685.
- [15] Neubeck A, Van Gool L. Efficient non-maximum suppression[C]//18th International Conference on Pattern Recognition (ICPR). Springer: Berlin, German, 2006, 3: 850-855.
- [16] 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):118-126.
 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): 118-126. (in Chinese with English abstract).
- [17] Zhou Xingyi, Wang Dequan, Krähenbühl P, et al. Objects as Points [C]//IEEE Conference on Computer Vision and Pattern Recognition. United States: IEEE, 2019: 2678-2689.
- [18] 李善军, 胡定一, 高淑敏, 等. 基于改进 SSD 的柑橘实时分类检测[J]. 农业工程学报, 2019, 35(24): 307-314. Li Shanjun, Hu Dingyi, Gao Shumin, et al. Real-time classification and detection of citrus based on improved single short multibox detecter[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(24): 307-314. (in Chinese with English abstract)
- [19] Loshchilov I, Hutter F. SGDR: stochastic gradient descent with warm restarts[J]. arXiv: 2016, 1608. 03983.
- [20] Mehta S, Rastegari M, Shapiro L, et al. Espnetv2: A light-weight, power efficient, and general purpose convolutional neural network[C]//Proceedings of the IEEE

Conference on Computer Vision and Pattern Recognition. United States: IEEE, 2019: 9190-9200.

- [21] Sandler M, Howard A, Zhu M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]//IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 4510-4520.
- [22] Lu J, Lee W S, Gan H, et al. Immature citrus fruit detection based on local binary pattern feature and hierarchical contour analysis[J]. Biosystems Engineering, 2018, 171: 78-90.
- [23] Zhao C, Lee W S, He D. Immature green citrus detection based on colour feature and sum of absolute transformed difference (SATD) using colour images in the citrus grove[J]. Computers and Electronics in Agriculture, 2016, 124: 243-253.
- [24] Liu S, Yang C, Hu Y, et al. A method for segmentation and recognition of mature citrus and branches-leaves based on regional features[C]//Chinese Conference on Image and Graphics Technologies. Singapore: Springer, 2018: 292-301.
- [25] Lin Tsung Yi, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE

Conference on Computer Vision and Pattern Recognition, Honolulu, Hawaii, USA: IEEE, 2017: 2117-2125.

- [26] Liang Zhen, Shao Jie, Zhang Dongyang, et al. Small object detection using deep feature pyramid networks[C]//Pacific Rim Conference on Multimedia, Hefei, China: Springer, 2018: 554-564.
- [27] Microsoft. PASCAL-VOC2012 [DB/OL]. (2012-02-20) [2019-08-02]. http://host.robots.ox.ac.uk/pascal/VOC/voc2012.
- [28] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [29] Kirillov A, Girshick R, He Kaiming, et al. Panoptic feature pyramid networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, California: IEEE, 2019: 6399-6408.
- [30] Liu Jun, Li Pengfei. A mask R-CNN model with improved region proposal network for medical ultrasound image[C]//Proceedings of the International Conference on Intelligent Computing, Wuhan: Springer, 2018: 26-33.

Recognition method of pheasant using enhanced Tiny-YOLOV3 model

Yi Shi¹, Shen Lian¹, Zhou Siyao¹, Zhu Jingming¹, Yuan Xuesong²

(1. College of Information Science and Technology (College of Cyber Security, College of Oxford Brookes), Chengdu University of Technology, Chengdu 610059, China; 2. College of Electronic Science and Engineering, University of Electronic Science and Technology, Chengdu 610054, China)

Abstract: The increase of pheasants has posed a threaten to crops as the advancement of ecology. However, most conventional methods of bird repellent have inherent deficiencies in terms of efficiency and danger. An efficiency monitoring method for pheasant is necessary to combine with artificial intelligence, in order to provide early warning and expulsion of pheasants. Normally, pheasant activities are mostly in the early morning and dusk under complex environment with protective color or habit of hiding. This behavior has made monitoring methods much more challenge. In this paper, a novel recognition method for pheasant has been proposed on the deployment of embedded computing platform, combined with the enhanced Tiny-YOLOV3 target detection network, particularly on considering the behavior of pheasant and specific living conditions. A lightweight network is required to ensure the accuracy and real-time monitoring due to the deployment on a mobile platform in the field environment. A real-time monitoring network ET-YOLO has also been established for the emergence of pheasants in a complex field environment, according to the basic structure of the Tiny-YOLOV3 lightweight target detection network. The feature extraction can deepen the net depth of Tiny-YOLOV3, and thereby increase the detection scale to improve the detection accuracy of original net target. CenterNet structure was used in the net detection layer to further enhance the detection accuracy and speed. The dataset of pheasant monitoring was produced after augmentation using the field collection of images in various environments, including 6000 high resolution images of pheasant in different distances, angles and environments. The indicators of experimental evaluation were mainly tested in terms of accuracy, real-time performance, and model size. Specifically, the average detection accuracy, average detection speed, and detection model size of the pheasant were used for evaluation. The experimental results showed that the average detection accuracy of ET-YOLO in the complex field environment was 86.5%, and the average detection speed was 62 frames/s, 15% higher than that of initial Tiny-YOLOV3. The average detection accuracy was higher than that of YOLOV3, Faster-RCNN and SSD_MobileNetV2 by 1.5%, 1.1% and 18%, respectively. The average detection speed was 38 frames /s, 47 frames /s and 1 frame/s higher than that of YOLOV3, Faster-RCNN and SSD MobileNetV2, respectively, when the detection model size of 56 MB. The proposed method can be suitable for the deployment on embedded computing platforms equipped with agricultural robots and intelligent machines in terms of recognition accuracy, real-time performance, and model size, particularly recognizing pheasants in complex environments.

Keywords: agriculture; artificial intelligence; embedded system; pheasant recognition; ET-YOLO; CenterNet detection structure