基于红外热成像与改进YOLOV3的夜间野兔监测方法

易 诗[1],李欣荣[1],吴志娟[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 成都 610059;2.电子科技大学电子科学与工程学院 成都 610054)

摘 要:随生态改善,野兔数量增多,对农田与林地的危害日益加重。野兔活动多为夜间,目标小,运动速度快,且出现环境较复杂,监控兔害,需要一种高效智能化的方法。针对野兔活动习性,该文提出了使用红外热成像实时监控,结合改进的YOLOV3目标检测方法对夜间野兔进行检测。根据YOLOV3目标检测网络基本结构提出了一种针对红外图像中野兔的实时检测的网络(infrared rabbit detection YOLO,IR-YOLO),该网络特征提取部分压缩YOLOV3特征提取网络深度,利用浅层卷积层特征以提高低分辨率红外小目标检测精度,降低运算量,网络检测部分使用基于CenterNet结构的检测方式以提高检测速度。使用热成像野外实时采集的夜间野兔图像作为数据集,包括不同距离,尺度,出现环境不同的野兔共计6 000幅红外图像制作训练集与测试集,比例为5:1。试验结果表明,IR-YOLO在红外热成像视频中复杂环境下出现的野兔检测率达75%,平均检测速度51帧/s,相对改进前YOLOV3检测率提高15个百分点,相对改进前YOLOV3检测速度提高5帧/s。相比其他目标检测算法各项检测指标更为优良,检测率方面相对Faster-RCNN与RFCN-RESNET101分别提高45个百分点与20个百分点,检测速度方面相对Faster-RCNN与RFCN-RESNET101分别提高30和与45帧/s。该方法可高效快速地对夜间复杂环境下出现的野兔进行检测,也可广泛应用于夜间对其他类型农业害兽的检测。

关键词:红外热成像;图像处理;野兔;检测;IR-YOLO;CenterNet检测结构;小目标检测

doi: 10.11975/j.issn.1002-6819.2019.19.027

中图分类号:TN919.5 **文献标识码**:A **文章编号**:1002-6819(2019)-19-0223-07

易 诗,李欣荣,吴志娟,朱竞铭,袁学松. 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):223－229. doi:10.11975/j.issn.1002-6819.2019.19.027 http://www.tcsae.org

Yi Shi, Li Xinrong, Wu Zhijuan, Zhu Jingming, Yuan Xuesong. 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): 223－229. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2019.19.027 http://www.tcsae.org

0 引 言

随近年来生态环境改善,野兔数量急剧增加,分布广泛,给林业、农业发展造成一定威胁,并有进一步恶化的趋势,随着中国对林业生产的不断重视,特别是退耕还林工程的实施,造林面积迅速扩大,林木数量大幅度增加,但随之出现的兔害造成了巨大的经济损失[1-2]。野兔活动习性多为夜间,零星出没,导致采用有光条件下的驱兔方法效果较差,而下网下夹、毒杀、电击、猎杀等方法非常危险[3],且野兔属于国家保护动物,因此,提出一种结合近年来人工智能技术的驱兔方法具有实际意义,此类方法中重要的环节在于夜间对野兔的检测。

野兔一般于秋冬季夜间9点至凌晨活跃于山间麦田,果林和人工经济林等区域,啃食作物和林木幼苗,普遍单只或成对出现,较少聚群。因此,对野兔的检测与预警,需解决如下几个问题:1)野外夜间的无光环境下探测。2)复杂地形,存在一定遮挡的条件下检测目标。3)小型目标检测。4)快速运动目标检测。

红外热成像系统成像原理为物体的温差,其不依赖光源,受天气影响小,探测距离远[4],适合于野外环境下对野生动物的检测,解决了无光环境下探测的问题,对于草木的一定遮挡,红外热成像能够有效探测目标泄露的红外光谱并成像,对复杂环境下目标探测具有较大优势。近年来,由于深度学习相关理论的快速发展及计算能力的提升,深度卷积网络在计算机视觉方面取得了很大的成功。在目标检测方面,基于深度学习的目标检测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2类,一类是基于区域生成的卷积网络结构,代表性的网络为RCNN、Faster-RCNN[5-6];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SSD、YOLO[7-9]等。熊俊涛等[10]采用更快速卷积神经网络(Faster-RCNN)实现了自然环境下绿色柑橘的识别。薛月菊等[11]采用YOLOV2实现了绿色芒果的识别。赵德安等[12]采用YOLOV3实

收稿日期:2019-05-20 修订日期:2019-07-18

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(201810616150)

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、嵌入式系统开发方面的研究。Email:549745481@qq.com

现了水下摄像头采集的河蟹图像的识别。其中YOLOV3目标检测框架具备对小目标检测精度高，速度快的优点，但对夜间红外图像中野兔检测存在以下问题：1) 低分辨率红外小目标的检测能力需要提高。2) 对运动目标检测速度需进一步提升。3) 对存在一定遮挡下的目标检测能力需要优化。

因此，本文针对红外图像中野兔检测的特点，在YOLOV3网络的基础结构上提出了一种红外图像中检测快速运动小目标的检测网络（infrared rabbit detection YOLO）IR-YOLO。该网络根据夜间野兔多为数量较少的快速运动小目标，红外热成像图像分辨率低，细节模糊，缺乏色彩特征的特点，在YOLOV3网络的基础结构上压缩特征提取网络深度，提取浅层卷积层特征，提高低分辨率红外小目标检测精度，降低运算量，在网络检测层采用CenterNet结构[13]，提高检测速度与实时性。在训练样本中加入存在一定比例草木遮挡条件下的野兔目标样本，优化存在遮挡情况下野兔识别的准确率。

1 红外热成像夜间野兔检测模型

1.1 YOLOV3网络结构

YOLOV3实时目标检测模型的特征提取网络Darknet53由53个卷积层，24个残差层构成。后20层为YOLO网络的特征交互层，分为3个尺度，每个尺度内，通过卷积核的方式实现局部的特征交互，作用类似于全连接层，通过卷积核（3×3和1×1）的方式实现特征图（feature map）间的局部特征交互，全连接层进行全局的特征交互[14-19]。卷积层提取图像特征，全连接层预测图像位置与类别估计概率值。YOLOV3根据输入的图像数据，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。

YOLOV3在3个检测尺度S上（S=13,26,52），将输入图像分成$S\times S$个单元格，每个单元格的神经元负责检测落入该单元格的对象，最多可包括两个预测对象的滑动窗口。T为目标置信度，它反映当前滑动窗口是否包含检测对象及其预测准确性的估计概率，计算如下：

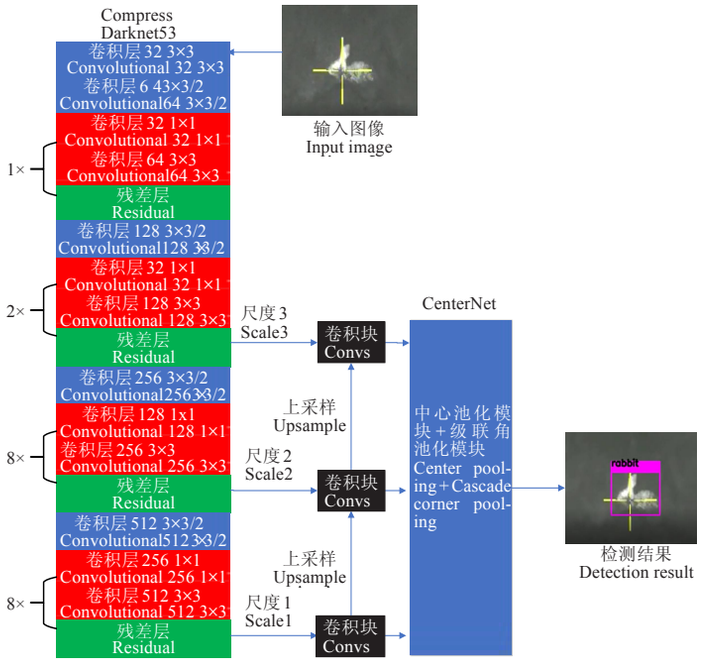
$$T=P_0\times P_1 \tag{1}$$

式中P_0表示滑动窗口包含检测对象的概率，P_1表示滑动窗口与真实检测对象区域的重叠面积（像素单位）。若B为每个单元格可以用于预测对象的滑动窗口数量，C为类别总数，则YOLOV3的全连接层的输出维度N计算如下：

$$N=S\times S\times(B\times 5+C) \tag{2}$$

1.2 IR-YOLO神经网络结构

为提高YOLOV3网络对红外图像中小目标的检测能力，本文提出了IR-YOLO神经网络结构，红外小目标的特点主要为分辨率低（红外小目标普遍20×20像素），细节模糊，缺乏色彩特征，因此，网络结构改进的基本思想为压缩特征提取网络深度，利用浅层卷积层特征，浅层的卷积特征感受视野包含的背景噪声小，适宜于提取低分辨率小目标语义特征，对红外目标具有更好的表征能力。深层卷积层处理高分辨率细节特征更为合适，而对红外图像这类低分辨率图像特征感受视野包含的背景噪声较多，可利用有效信息量较少，因此可进行压缩。同时，为进一步提高检测实时性，在YOLOV3原有的检测层采用无锚点CenterNet结构。IR-YOLO神经网络结构如图1所示。



注：2×、8×分别指该层在模型中通过级联的方式重复级联2次与8次；1×1、3×3、3×3/2分别指该层采用卷积核大小为1×1、3×3以及3×3再上采样2次。
Note: 2× and 8× mean that the layer is cascaded 2 and 8 times in the model by cascade;1×1, 3×3,3×3/2 mean that the layer with convolution core size of 1×1, 3×3; 3×3 /2 means that the layer is up sampled twice with convolution core size of 3×3.

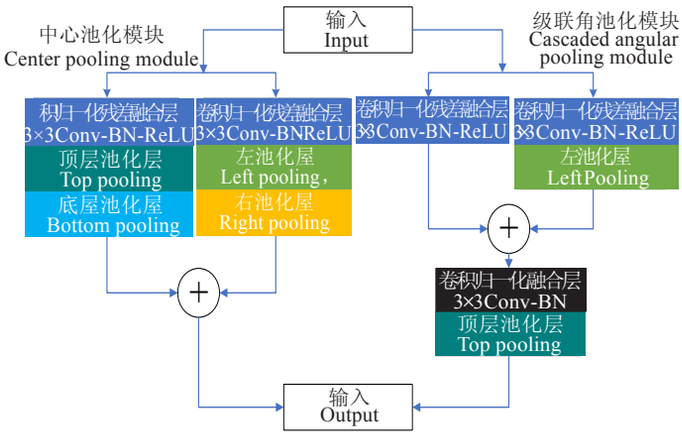
图1 IR-YOLO网络结构图
Fig.1 IR-YOLO network structure

IR-YOLO神经网络结构中，剪裁了YOLOV3主干特征提取网络（Darknet53）中44到53共9个卷积层以及最后4个残差层，使原有主干特征提取网络由74层降低至61层，形成压缩网络结构（Compress Darknet53）。最小尺度输出层，检测尺度S=13，输入512通道，输出75通道。中尺度输出层，检测尺度S=26，与48层的特征图合并，输入256通道，输出75通道。大尺度输出层，检测尺度S=52，与23层的特征图合并，输入128通道，输出75通道。由于红外图像分辨率较低，细节模糊，缺乏高分辨率语义特征，因此44层以上的深层卷积层提取的红外图像特征中包含有效红外目标语义特征信息较少，同时容易引入感受视野噪声，造成识别精度的降低[20-25]，而23层开始的浅层卷积层包含较多红外目标语义特征信息，存在感受视野噪声较少，因此，在主干特征提取网络上进行深层卷积层的剪裁，检测输出层的向浅层移动的压缩改进方式将更好利用有效表征红外图像目标语义特征信息的浅层卷积层特征图，提高红外小目标识别率，同时压缩网络深度将降低运算量，提高实时性。

检测部分3个检测尺度上均使用CenterNet结构代替原有的YOLO层，CenterNet是CVPR2019大会上提出的最新目标检测方法，其基本思想为：基于锚点（anchor box）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YOLOV3在检测层中需要大量的anchor box进行目标检测。而CenterNet是一个

单阶段的关键点检测模型，将检测每个目标物看作3个关键点（中心点，上下对角点），避免产生大量anchor box，降低运算量提高实时性的同时提高检测准确率及召回率。

IR-YOLO所使用的CenterNet由Cascade corner pooling及Center pooling构成，从而获得从左上角及右下角的丰富信息，同时在中间区域获得更多的识别信息。检测部分结构如图2所示。



注：卷积核大小为3×3。
Note: The convolution core size is 3×3 in this layer.

图2　CenterNet结构
Fig. 2　CenterNet architecture

CenterNet网络中的Center pooling模型由2个卷积归一化残差融合层（conv-bn -relu），1个左池化层（left pooling），1个右池化层（right pooling），1个顶层池化层（top pooling），1个底层池化层（bottom pooling）组成，作用在于用于预测中心关键点的分支，有利于中心获得更多目标物的中心区域，进而更易感知建议区域的中心位置，通过取中心位置横向与纵向响应值的和的最大值实现此方法。Cascade corner pooling模型由2个卷积归一化残差融合层（conv-bn-relu），1个左池化层（left pooling），1个卷积归一化融合层（conv-bn），1个顶层池化层（top pooling）构成，作用在于增加原始的corner pooling感知内部信息的功能。结合了特征图（feature map）中目标物内部及边界方向的响应值和的最大值以预测左上角及右下角的2个角点。中心池化模块(center pooling)及级联角池化模块(cascade corner pooling)的输出融合后即可得到准确的目标预测位置。

在使用CenterNet进行检测时，中心区域的大小影响边界框的检测结果。因此，采取尺度敏感区域以适应不同尺寸大小的目标物。其将生成相对小目标较大，相对大目标较小的中心区域，例如，需要判断一个边界框I是否需要被保留，tl_x，tl_y代表框左上角的点，br_x，br_y代表框右下角的点。定义一个中心区域j，定义左上角的点的坐标为(ctl_x,ctl_y)右下角点(cbr_x，cbr_y)。这些参数定义满足：

$$ctl_x=\frac{(n+1)tl_x+(n-1)br_x}{2n}$$
$$ctl_y=\frac{(n+1)tl_y+(n-1)br_y}{2n} \tag{3}$$
$$cbr_x=\frac{(n-1)tl_x+(n+1)br_x}{2n}$$
$$cbr_y=\frac{(n+1)tl_y+(n-1)br_y}{2n}$$

式中n为奇数，代表中心区域j的大小。对于边界框小于150时n=3，大于150时n= 5。

2　红外热成像夜间野兔检测试验

2.1　试验平台

实验硬件数据采集平台为普雷德G35氧化钒384×288分辨率机芯，35 mm镜头的户外热成像探测仪，输出为单通道16bit数据。模型训练使用硬件平台为Core i7-8750H 2.2Ghz处理器，16GB内存，Geforce GTX 1080 8GB显卡。软件平台使用Win10，tensorflow1.9.0，CUDA9.2，VS2017，opencv4.0。

2.2　试验数据

由于红外热成像开放数据集较少，而且针对夜间野外活动野兔的数据集缺乏。因此，本文采用热成像夜间野外拍摄野兔在各种环境，距离下出现的视频制作自建数据集，数据集的采集时间为野兔夜间活跃的秋冬季节晚9点至凌晨2点，采集地点为中国四川省境内农业大县梓潼，盐亭，剑阁。数据集中，野兔出现环境包括农田，果园，林地，荒地，草丛等，出现距离包括近距离（10~20 m），中距离（20~50 m）与热成像所能识别的较远距离（50~100 m），野兔形态包括正面，侧面，背面等，特别采集了一定数量存在草木遮挡情况的野兔红外图像以训练提升存在遮挡情况下的野兔目标检测能力。数据集包括6 000幅野兔红外图像，训练集与测试集按5:1比例划分，数据集中部分样本如图3所示。

a. 10~20 m　　b. 20~50 m　　c. 50~100 m

图3　不同拍摄距离的野兔红外检测样本
Fig. 3　Rabbit infrared detection samples with different shooting distances

2.3　模型训练

在模型训练环节，使用迁移学习的方法以节省时间，提高训练速度。首先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100幅图像为一个批次进行小批量训练，每训练一批图像，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IR-YOLO网络进行20 000次迭代，每间隔2 000次迭代后保存一次模型，最终选取精度最高的模型。

整个训练的目标损失函数计算公式如下：

$$L=L_{det}^{co}+L_{det}^{ce}+L_d \tag{4}$$

式中L_{det}^{co}与L_{det}^{ce}分别为边缘关键点与中心点的散焦损失，用于训练检测目标边缘与中心点。L_d为向量距离损失函数，计算如下式：

$$L_d=\alpha L_{pull}^{co}+\beta L_{push}^{co}+\gamma(L_{off}^{ce}+L_{off}^{co}) \tag{5}$$

式中L_{pull}^{co}为边缘点回拉损失，用于将同一目标内嵌向量最小化，L_{push}^{co}为边缘点外推损失，用于将不同目标内嵌向量最大化，L_{off}^{ce}与L_{off}^{co}分别为中心关键点与边缘关键点的偏移损失，用于预测二者的偏移值。α，β，γ为权重系数，本文参考文献[13]的思路3个参数分别取0.1,0.1,1。经过20 000次迭代，模型平均损失降至0.2以下。

3 结果与分析

3.1 评价指标

测试环节首先通过在测试样本集中分析模型改进前后的检测精度、检测速度以对比模型的性能。将测试图像输入训练好的网络模型，将夜间红外图像中的野兔目标进行检测并记录结果，当模型预测的目标边界与测试集对应的标注数据中的边界框的交并集比(IOU)大于等于设定的阈值时，则认为检测结果正确，否则视为检测错误。检测精度的评价指标选取准确率P，召回率R，调和均值F以及模型处理速度指标M_o（帧/s）。各平价指标定义如下式：

$$P=\frac{T_p}{T_p+F_p}\times 100\% \tag{6}$$

$$R=\frac{T_p}{T_p+F_N}\times 100\% \tag{7}$$

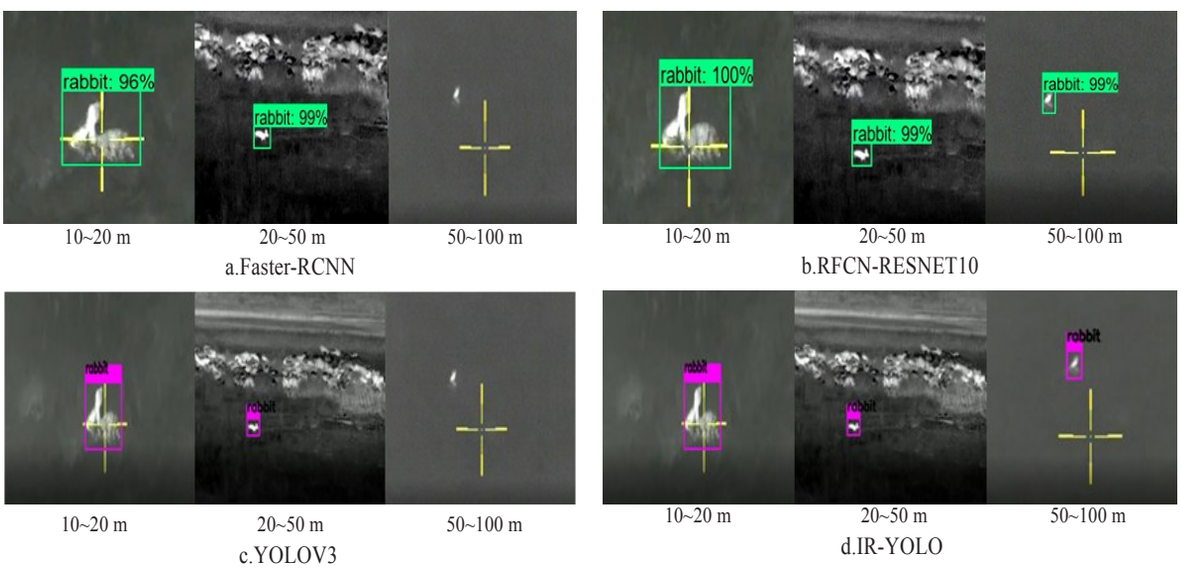
$$F=\frac{2PR}{P+R}\times 100\% \tag{8}$$

式中T_p为正确检测到野兔的数量，F_p为误将非野兔目标检测为野兔的数量，F_N为误将野兔检测为背景的数量。

此外，应进行模型在夜间野外复杂环境中的鲁棒性测试，该测试环节采用复杂环境下对红外热成像视频中野兔目标的平均检测率，与平均检测速度作为评价指标。

3.2 夜间野兔识别结果与分析

实际测试时选取默认IOU为0.5，大于该阈值认为在红外图像中检测到了野兔目标。测试选用近距离、中等距离与远距离不同尺度的红外视频中野兔目标进行测试，为验证测试方法的可靠性与准确性，在对比改进YOLOV3前后各项指标的同时，使用Faster-RCNN与RFCN-RESNET101方法采用相同数据集与训练参数进行模型训练，并进行对比测试，结果如图4所示。



注：图中百分比表示各检测方法对目标野兔的判定阈值分数。

Note: The percentage in the figure is the threshold score of each detection method for target rabbits determination.

图4 不同神经网络对不同距离的目标野兔识别结果对比

Fig.4 Comparison of recognition results of target rabbits with different distances by different neural networks

图4可见，红外热成像图像中，近距离野兔与中距离野兔几种检测网络均可识别，而远距离野兔仅深度目标检测网络RFCN-RESNET101与本文提出的IR-YOLO网络可成功识别。

对测试样本集中的各个目标检测网络的准确率(P)、召回率(R)、调和均值(F)和平均处理速度的统计如表1所示。

表1 网络评价指标分析

Table 1 Analysis of network evaluation index

模型Model	Accuracy rate$P$/%	Recall rate R/%	F/%	Prolessing speed M_0 /(帧·s^{-1})
Faster-RCNN	85	87	86	21
RFCN-RESNET101	89	90	89.5	6
YOLOV3	93.5	89.3	91.5	46
IR-YOLO	97.8	92.5	93.2	51

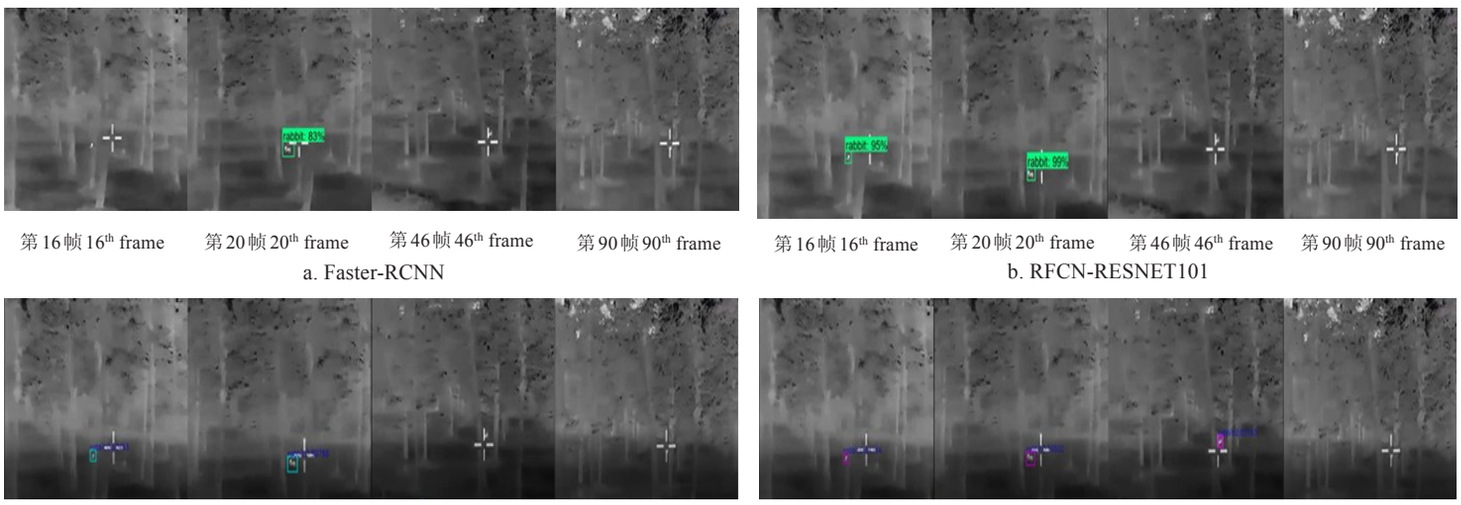
根据表1可知改进后的IR-YOLO目标检测模型相对YOLOV3目标检测模型对夜间野兔检测的准确率和召回率比分别提高了4.3个百分点和3.3个百分点，调和均值提高1.7个百分点，平均处理速度提高5帧/s，而另外2个目标检测网络由于在测试样本集中误将非野兔目标检测为野兔的数量较高，影响了总体精度，平均处理速度方面远低于YOLOV3与本文提出的IR-YOLO。

实际野外环境中，夜间野兔常出现于林地，杂草较多的田地，存在遮挡，目标小，运动速度快的特点，因此，使用在一片较为茂密的经济林中夜间较远距离拍摄（80 m）的100帧快速运动的野兔红外热成像视频以测试本文提出方法与Faster-RCNN,RFCN-RESNET101, YOLOV3在鲁棒性上的对比。

首先，在100帧视频中随机提取4帧，测试结果如图5所示。

随机提取的4帧图像分别为第16帧，第20帧，第46帧与第90帧。其中第16帧存在快速运动与部分遮挡，Faster-RCNN不能成功识别，其余目标检测网络均可成功识别，第20帧目标相对静止，遮挡程度小，所有检测网络均可成功识别，第46帧存在快速运动与较大程度的遮挡，仅本文提出的IR-YOLO网络成功识别，其余网络均无法识别，第90帧目标快速运动，同时遮挡面积非常严重，所有网络均无法识别。由上述对比测试结果可见本文提出的IR-YOLO网络在夜间野外复杂环境下对红外野兔目标识别优势最大，同时平均处理速度最高，达50帧/s以上。

为更为可靠的验证检测网络鲁棒性，对整个测试视频100帧中目标平均检测率，漏检率，误检率，平均检测速度做出统计，统计结果如表2所示。



第16帧16th frame　第20帧20th frame　第46帧46th frame　第90帧90th frame
a. Faster-RCNN

第16帧16th frame　第20帧20th frame　第46帧46th frame　第90帧90th frame
b. RFCN-RESNET101

第16帧16th frame　第20帧20th frame　第46帧46th frame　第90帧90th frame
c. YOLOV3

第16帧16th frame　第20帧20th frame　第46帧46th frame　第90帧90th frame
d. IR-YOLO

图5 复杂环境下不同神经网络的检测性能对比

Fig.5 Detection performance comparison of different neural networks in complex environments

表2 复杂环境下不同神经网络的鲁棒性对比

Table 2 Robustness comparison of different neural networks in complex environments

模型 Model	目标出现总帧数 Total number of frames for target occurrence	目标检测总帧数 Total number of frames for target detection	目标漏检总帧数 Total number of frames for target missed detection	目标错误检测总帧数 Total number of frames for target error detection	平均检测率 Average detection rate /%	平均检测速度 Average detection speed/ (帧·s^{-1})
Faster-RCNN	72	36	36	6	30	21
RFCN-RESNET101	72	50	22	10	55	6
YOLOV3	72	43	29	0	60	46
IR-YOLO	72	54	18	0	75	51

根据鲁棒性测试结果可见，本文提出的IR-YOLO目标检测模型在复杂环境中的鲁棒性良好，对快速运动，存在遮挡，距离较远的野兔目标检测率相对YOLOV3目标检测网络提高15个百分点，检测速度提高5帧/s，检测率方面相对Faster-RCNN与RFCN-RESNET101分别提高45个百分点与20个百分点，检测速度方面相对Faster-RCNN与RFCN-RESNET101分别提高30与45帧/s。适宜于夜间复杂环境下对野兔的检测。

4 结 论

1)本文所提出的针对红外热成像中出现环境复杂，快速运动的红外小目标检测网络IR-YOLO对夜间野兔目标具备较高的准确性，试验结果表明，模型准确率和召回率分别达到了97.8%和92.5%，平均检测速度达51帧/s。

2)针对红外图像小目标特点对YOLOV3网络结构做出改进，压缩特征提取网络深层卷积层，向浅层卷积层平移检测输出，提取利用较好反映低分辨率红外图像特征的浅层卷积层语义特征，提高红外小目标检测精度，降低运算量，在网络检测部分采用CenterNet检测结构代替原有YOLO层，进一步提高检测速度与精度。

3)对实际夜间复杂环境下出现的较远距离快速运动野兔目标的鲁棒性测试中，本文方法表现出较强鲁棒性与实时性，检测率较YOLOV3目标检测网络改进前提升15个百分点，检测速度提升5帧/s。检测率方面相对Faster-RCNN与RFCN-RESNET101分别提高45个百分点与20个百分点，检测速度方面相对Faster-RCNN与RFCN-RESNET101分别提高30和45帧/s。

[参 考 文 献]

[1] 宋晓斌,韩恩贤, 薄颖生,等.兔害对退耕还林的危害调查与防治对策[J]. 陕西林业科技, 2002 (1): 55—57.
Song Xiaobin, Han Enxian, Bo Yingsheng, et al. Harm investigation and control countermeasures of rabbit damage on returning

farmland to forestry [J]. Shaanxi Forestry Science and Technology, 2002 (1): 55—57. (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: //gs sf z. neas e. n et/4. h tml.

[3] 张耀,申世友 . 陕北榆林幼林地野兔危害调查与防治对策[J] . 防护林科技, 2003(4):71—72 .
Zhang Yao, Shen Shiyou. Harm investigation and control countermeasures of hares in young forest land of Yulin, northern Shaanxi [J]. Shelterbelt Science and Technology, 2003 (4): 71—72. (in Chinese with English abstract).

[4] 崔美玉 . 论红外热像仪的应用领域及技术特点[J]. 中国安防, 2014(12): 90—93.
Cui Meiyu. on the application field and technical characteristics of Infrared thermal Imager[J]. China Security, 2014 (12): 90—93. (in Chinese with English abstract).

[5] 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.

[6] Girshick R. Fast R-CNN [C]. IEEE International Conference on Computer Vision, 2015: 1440—1448.

[7] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision, 2016: 21—37.

[8] Redmon J, Farhadi A. YOLO9000: Better, faster, stronger [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6517—6525.

[9] Redmon J, Farhadi A. Yolov3: An incremental improvement[EB / OL]. (2018 -04-08)[2018-09-07]. org/abs/1804. 02767.

[10] 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]. Journal of Agricultural Machinery, 2018, 49 (4): 45—52. (in Chinese with English abstract).

[11] 薛月菊，黄宁，涂淑琴，等．未成熟芒果的改进YOLOV2识别方法[J]．农业工程学报，2018，34(7)：173—179.
Xue Yueju, Huan Ning, Tu Shuqin, et al. Improved YOLOV2 identification method for immature mango [J].Transactions of the Chinese Society for Agricultural Machinery, 2018, 34 (7): 173—179. (in Chinese with English abstract).

[12] 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 underwater crabs based on machine vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(3): 151—158. (in Chinese with English abstract).

[13] Kaiwen Duan, Song Bai, Ling Xixie, et al. CenterNet: Keypoint triplets for object detection [C] //IEEE Conference on Computer Vision and Pattern Recognition, 2019: 3668—3685.

[14] Ren S Q, He K M, Girshick R, et al. Faster r-cnn: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(6): 1137—1146.

[15] Kurtulmus F, Lee W S, Vardar A. Green citrus detection using ‘eigenfruit’, color and circular Gabor texture features under natural outdoor conditions[J]. Computers and Electronics in Agriculture, 2011, 78(2): 140—149.

[16] Pan S J, Yang Q. A survey on transfer learning[J]. IEEE Transactions on Knowledge & Data Engineering, 2010, 22(10):166—169

[17] 魏丽, 丁萌, 曾丽君, 等 . 红外图像中基于似物性与稀疏编码的行人检测[J]. 红外技术, 2016, 38(9): 752—757.
Wei Li, Ding Meng, Ceng Lijun, et al. Pedestrian detection based on quasi physical properties and sparse coding in infrared images [J].Infrared Technology, 2016, 38(9): 752—757. (in Chinese with English abstract).

[18] 张科, 刘彦 . 改进的基于背景预测的红外弱小目标检测方法[J]. 火力与指挥控制,2008, 33(11): 22—24.
Zhang Ke, Liu Yan. An improved small target detection method command control[J], Fire Control & Command Control. 2008, 33 (11): 22—24. (in Chinese with English abstract).

[19] 杨阳, 杨静宇 . 基于显著性分割的红外行人检测[J]. 南京理工大学学报: 自然科学版, 2013, 37(2): 251—256.
Yang Yang, Yang Jingyu. Infrared pedestrian detection based onsaliency segmentation[J]. Journal of Nanjing University of Science and Technology, 2013, 37(2): 251—256. (in Chinese with English abstract).

[20] Wangz H, Fanb, Wu F C. Local intensity order pattern for feature description[C]//International Conference on Computer Vision, 2011: 6—13.

[21] Linker R, Cohen O, Naor A. Determination of the number of green apples in RGB images recorded in orchards[J]. Computers and Electronics in Agriculture, 2012, 81: 45—57.

[22] Arivazhagan S, Shebiah R N, Nidhyanandhan S S, et al. Fruit recognition using color and texture features[J]. Journal of Emerging Trends in Computing and Information Sciences, 2010, 1(2): 90—94.

[23] Xu Y, Imou K, Kaizu Y, et al. Two-stage approach for detecting slightly overlapping strawberries using HOG descriptor[J]. Biosystems Engineering, 2013, 115(2): 144—153.

[24] 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.

[25] 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.

[26] Lecun Yann, Bengio Yoshua, Hinton Geoffrey. Deep learning[J]. Nature, 2015, 521: 436－444.

[27] Smeulders A, Chu D, Cucchiara R, et al. Visual tracking: an experimental survey[J]. IEEE Transactions on Pattern Analysis& Machine Intelligence,2014,36(7):1442—1468.

[28] Joseph Redmon, Divvala Santosh, Girshick Ross, et al. You only look once: Unified, real-Time object detection[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2015: 779－788.

[29] Long Jonathan, Shelhamer Evan, Darrell Trevor. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 39(4): 640—651.

[30] Bargoti S, Underwood J. Deep fruit detection in orchards[C]// IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2017: 3626－3633.

Night hare detection method based on infrared thermal imaging and improved YOLOV3

Yi Shi[1], Li Xinrong[1], Wu Zhijuan[1], Zhu Jingming[1], Yuan Xuesong[2]

(1.*College of Information Science and Technology, Chengdu University of Technology, Chengdu* 610059, *China*)

(2.*College of Electronic Science and Engineering, University of Electronic Science and Technology, Chengdu* 610054, *China*)

Abstract: Restoration of and improvement in ecosystems has led to a growth in hare numbers to the detriment of farmland and woodland where they appear. Understanding forage of the hare is imperative to better manage them but proves difficult because most hares forage fruit forests and mountain farmlands in night. In this paper we present a method to monitor real-time forage of the hares using infrared thermal imaging and the improved YOLOV3 target detection method - the real-time infrared hare detection network - YOLO (IR-YOLO) network. The feature-extraction component in the proposed network compressed the depth of the feature-extract in the YOLOV3 in attempts to improve the accuracy of detecting small targets from low-resolution infrared images by using the shallow-layer convolution method. In order to simplify computation and improve detection efficiency and accuracy of small targets, the detection component in the network was based on the CenterNet structure. Considering that hares are more active in autumn, real-time hare forage in the infrared images taken from wheat and rape fields were used to validate the proposed method. The datasets include 6 000 infrared images taken at different scales and environments, and the ratio of the datasets used for training to the sets used for validating was 5:1. The results show that the proposed method identified 75% of hares based on the infrared thermal image videos. The average processing speed of the proposed method is 51 frames/s, the detection rate is 15 percentage points higher than that of YOLOV3, Compared with the Faster-RCNN and RFCN-RESNET101, the proposed method increased processing speed by 45 and 20 percentage points respectively. In summary, the proposed method can quickly and accurately detect the forage of hares in complex environment in night, and it can be used to monitor night-movement of other animals appearing in other ecosystems.

Keywords: infrared thermal imaging; image processing; hare detection; IR-YOLO; Center Net structure; small target detection