

基于空间金字塔池化和深度卷积神经网络的作物害虫识别

张博¹, 张苗辉^{1,2*}, 陈运忠¹

(1. 河南省大数据分析处理重点实验室, 开封 475004; 2. 河南大学地理学博士后科研流动站, 开封 475004)

摘要: 为了减少因作物害虫姿态多样性和尺度多样性导致其识别精度相对较低的问题, 该文将空间金字塔池化与改进的YOLOv3深度卷积神经网络相结合, 提出了一种基于空间金字塔池化的深度卷积神经网络农作物害虫种类识别算法, 首先对测试图像上的害虫进行检测定位, 然后对检测定位出的害虫进行种类识别。通过改进YOLOv3的网络结构, 采用上采样与卷积操作相结合的方法实现反卷积, 使算法能够有效地检测到图片中体型较小的作物害虫样本; 通过对采集到的实际场景下20类害虫进行识别测试, 识别精度均值可达到88.07%。试验结果表明, 本文提出的识别算法能够有效地对作物害虫进行检测和种类识别。

关键词: 图像识别; 算法; 害虫分类; 深度卷积神经网络; 空间金字塔池化; 反卷积

doi: 10.11975/j.issn.1002-6819.2019.19.025

中图分类号: TP391.4 文献标志码: A 文章编号: 1002-6819(2019)-19-0209-07

张博, 张苗辉, 陈运忠. 基于空间金字塔池化和深度卷积神经网络的作物害虫识别[J]. 农业工程学报, 2019, 35(19): 209—215. doi:10.11975/j.issn.1002-6819.2019.19.025 <http://www.tcsae.org>

Zhang Bo, Zhang Miaohui, Chen Yunzhong. Crop pest identification based on spatial pyramid pooling and deep convolution neural network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(19): 209—215. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2019.19.025 <http://www.tcsae.org>

0 引言

农业是一个国家经济发展的基础, 粮食生产更是关系到国计民生。粮食生产受到诸多条件的影响, 其中, 病虫害是影响作物产量的重要原因之一。农业病虫害造成的作物减产覆盖了水稻、玉米和小麦等数十种农作物。因此, 作物害虫的识别是科学防治病虫害的重要依据^[1-4]。传统的作物害虫分类方法是专家通过观察害虫的外部特征, 结合已知的害虫图鉴对害虫的种类进行判别, 并对当季的各类害虫数量进行统计分析。该方法依赖于专家的经验水平和主观意识, 识别效率较低。

近年来, 借助于计算机视觉技术与模式识别的发展, 农作物害虫的分类识别方法也得到了广泛的研究。传统的害虫识别技术大多基于颜色、纹理、形状等特征, 通过稀疏表示的方法, 学习获得害虫的过完备字典, 从而对害虫进行分类^[5-14]。这一类方法在本质上使用人工预先设

计好的特征进行识别, 在一定程度上实现了智能化, 但在实际使用中容易受到害虫形态、背景干扰等限制。伴随深度学习与神经网络的发展, 对害虫的识别研究转向以卷积特征为核心的深度学习框架, 并在准确率和鲁棒性上获得了显著的提升^[15-18]。上述图像分类的方法对样本的输入有严格的要求, 依赖于人工对每个害虫样本进行选取和分离, 分类精度受背景干扰较大。为了进一步对害虫进行自动的精准定位和准确分类, 基于目标检测的害虫分类算法成为最优的解决方案。以Fast R-CNN (fast regions with CNN)和SSD(single shot multiBox detector)在害虫分类技术中的应用为例, 根据卷积层提取到的特征向量, 分别通过窗口分类和窗口回归, 得到每个候选区域中害虫的类别以及原始坐标。此类方法有效地提升了害虫种类判别的准确率, 并且能够对害虫目标进行定位^[19-23]。由于害虫目标的体型差异较大, 种类繁多, 浅层的卷积神经网络提取到的有效特征信息较少, 对多尺度输入的害虫样本泛化能力较差。

随着神经网络的发展, 基于深度卷积神经网络的分类识别技术展示出了超越人类视觉水平的表现^[24-31]。但对于作物害虫的分类识别问题, 传统的作物害虫分类算法^[7-8, 15]通常基于害虫的颜色、形态和纹理等表现特征, 通过浅层神经网络或者支持向量机(support vector machine, SVM)等分类器进行分类。这些算法除受害虫姿态多样性的因素影响外, 还受到尺度多样性的影响, 如体型较小的害虫样本在图片中所占像素较少, 使得传统的

收稿日期: 2019-04-26 修订日期: 2019-08-15

基金项目: 国家自然科学基金(61802111); 河南省青年骨干教师资助课题(2017GGJS019); 河南省教育厅科学技术研究重点项目(19A520002); 中国博士后面上基金(2015M582182); 河南省博士后基金(001703007)

作者简介: 张博, 主要研究方向为基于深度学习的目标检测与识别, 图像的分割, 行为识别。Email: zhangbo208@163.com

*通信作者: 张苗辉, 博士, 副教授, 主要研究方向为图像处理及应用、基于深度学习的目标检测与识别。Email: zhmh@henu.edu.cn

卷积结构难以提取样本的底层特征。本文基于 ResNet 的残差网络思想,使用更深层的基础网络,通过上采样与卷积操作相结合的方法,在 YOLOv3(you only look once version 3)^[32]网络中加入反卷积结构,用于检测出体型较小的害虫目标并分类;进一步将空间金字塔池化^[33]与改进后的 YOLOv3 相结合,使网络模型能够适应多尺度的样本输入,从而减小害虫体型差异较大对分类精度造成的影响。

1 材料与方法

1.1 害虫图像样本的获取

为了保证该算法在实际环境中识别害虫的准确率,本文图像数据的采集依托河南省鹤壁市佳多集团在全国各地设置的 JDXC-1 型智能虫情测报仪。这些测报仪设置在田间或者林地,诱捕到的昆虫都是当地易发生病虫害的昆虫,可以根据捕捉到的害虫数量来预测接下来的一段时间内发生哪种病虫害的可能性最大,以便尽早做好预防措施。顶部的摄像头每间隔 10 min 采集 1 张害虫图像,通过 4G 网络信号传输至后台服务器,能够更真实的反应虫情动态。

由于不同种类的害虫分布具有季节性,本次害虫样本的图像采集工作自 2015 年 5 月 11 日起至 2016 年 9 月 30 日,分时段进行,主要集中在河南、山东和湖北等地。本次共采集到害虫样本图片 976 张,每张图片的分辨率为 2 592×1 944,害虫共 124 类,如图 1 所示。专家借助害虫图鉴和 LabelImg 数据集标注脚本,共计标注害虫样本 25 425 个。害虫数量分布如表 1 所示,表中的第 1 列为每个种类的害虫所标注的样本数量范围,第 3 列为样本量在该范围内的害虫种类编号,第 2 列为满足第 1 列样本数量的所有种类害虫的标注样本总数。其中,害虫种类编号为害虫图鉴中每一类害虫固定的编号值,如样本采集过程中 1 号、4 号、9 号等部分种类害虫并未出现。

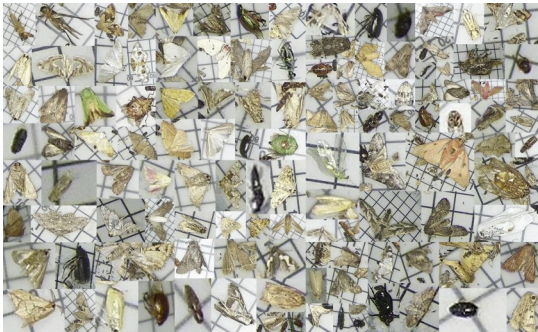


图 1 害虫样本示例
Fig.1 Examples of pest samples

1.2 改进的深度卷积神经网络

为了有效地提取到图像的细粒度特征,本文对 YOLOv3 作进一步改进。原始的 YOLOv3 由 106 层网络层组成,共使用步长为 2 大小为 3×3 的卷积核进行 5 次下采样操作,输出 3 种不同尺度的特征向量,每种尺度预测 3 个回归框的位置信息,如图 2a 所示。由于害虫目标在原始图像中所占比例较小,多次下采样会引起训练过程中卷积层提取到的小目标特征信息丢失,导致体型较小的害虫无法识别。因此,将网络输入尺度设定为 608×608,

表 1 害虫样本数量分布
Table 1 Quantity distribution of pests

每类样本数量 范围 Sample range	样本总个数 Total number of samples	害虫种类编号 Pest species number	害虫种类数量 Number of pest species
≥500	19 905	46、16、32、40、8、24、 39、41、42、11	10
≥100~500	4 217	44、48、43、13、54、7、6、 15、67、36、37、52、28、 53、45、98、64、69、47	19
≥50~100	618	5、63、60、89、65、58、 51、50、31	9
≥10~50	514	55、81、72、14、75、102、 78、95、85、17、87、93、 66、104、90、108、77、 49、84、29、114、92、88、 118、99、80、73	27
≥1~10	171	10、97、79、113、112、94、 25、82、56、117、83、68、 59、34、124、96、71、 123、111、110、109、57、 133、131、130、127、 125、122、107、103、 100、91、74、70、62、61、 35、13、3、135、134、 132、129、126、121、 120、119、116、115、106、 105、101、86、76、38、 27、20、12、2	59
总计 Total	25 425		124

下采样次数减少为 4 次,同时,使用上采样与卷积结合的方法,达到反卷积的效果。将网络首个全连接层的输出尺度提升至 38×38,并与下采样之前的网络输出相结合,尽可能提取出原始图片中体型较小的害虫特征,以提高小目标害虫检测的准确率,如图 2b 所示。本文使用双线性插值方法进行图像的上采样计算:

$$P_1 = \frac{f(x_1, y_0) - f(x_0, y_0)}{x_1 - x_0} \times u + f(x_0, y_0) \tag{1}$$

$$P_2 = \frac{f(x_1, y_1) - f(x_0, y_1)}{x_1 - x_0} \times u + f(x_0, y_1) \tag{2}$$

$$P = \frac{P_1 - P_2}{y_1 - y_0} \times v + P_1 \tag{3}$$

式(1)~(3)中, P 为上采样图像中 x, y 坐标的像素值,按采样比例对应至原图中 P' 的坐标为 $(i+u, j+v)$,其中 i, j 为整数部分, u, v 为小数部分。 P_1 和 P_2 分别是对应 y 方向和 x 方向的插值, $f(x_0, y_0)$ 和 $f(x_1, y_1)$ 为 P' 左上像素点和右下像素点的像素值。通过卷积与上采样结合的方法,实现了反卷积的操作,提升了 YOLOv3 输出特征的尺度,图 2b 为改进后的 YOLOv3 网络结构。

如图 2c 所示, DBL 结构由卷积层、批标准化(batch normalization, BN)和非线性激活函数(leaky relu)组成,构成 YOLOv3 网络结构的基本组件。传统的卷积神经网络在信息传递时会存在一定的信息丢失和损耗,造成梯度消失或者梯度爆炸等问题,随着网络的加深,出现训练过程中准确率下降,甚至无法训练的现象。本文使用 Resblock 残差结构,如图 2d 所示,能够在避免上述问题

的情况下大幅度的加深网络层数,使神经网络提取到更深层次的目标信息,提升目标定位和分类的准确率。在图2a和图2b中,Res2表示由2个Resblock结构叠加组成。本文对于回归框的预测采用logistic回归方法:

$$b_x = \sigma(t_x) + c_x \quad (4)$$

$$b_y = \sigma(t_y) + c_y \quad (5)$$

$$b_w = p_w e^{t_w} \quad (6)$$

$$b_h = p_h e^{t_h} \quad (7)$$

式中 c_x, c_y 是坐标偏移量, p_w, p_h 是预设的候选框边长。最终得到的边框坐标值是 $b_{x,y,w,h}$, 网络的学习目标是 $t_{x,y,w,h}$ 。在3种不同尺度生成的9个先验框中确定出目标存在可能性得分最高的候选框,从而对该候选框中的目标进行分类。该方法确定了待检测图像中先验框的数量,检测速度快、背景误检率低、对于非自然图像物体的检测和分类的准确率远远高于传统R-CNN (regions with CNN)等系列的目标检测方法^[32]。

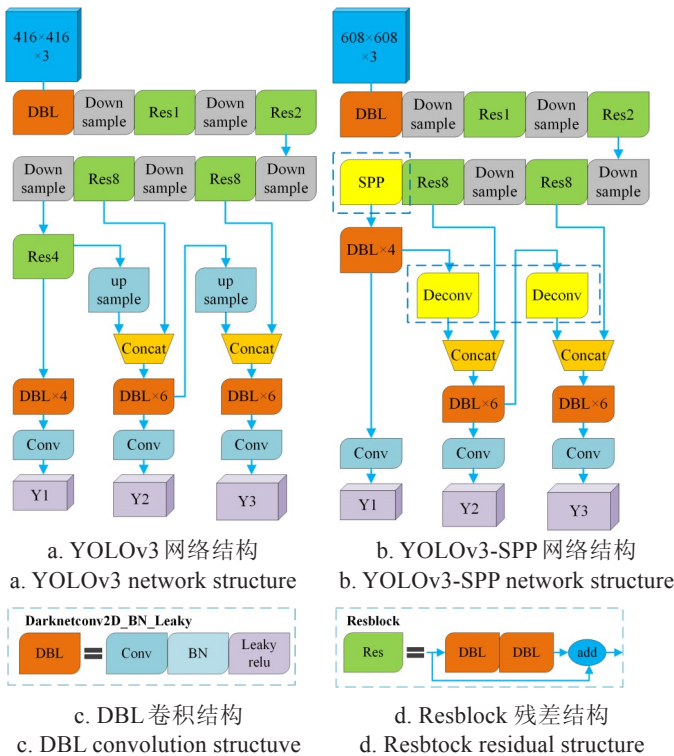


图2 YOLOv3 spp与YOLOv3网络结构对比

Fig.2 Comparison of YOLOv3-SPP and YOLOv3 network structure

1.3 空间金字塔池化

空间金字塔池化 (spatial pyramid pooling, SPP) 是一种将局部特征映射到不同维度空间并将其融合的池化方法,可以产生固定大小的特征向量并且使卷积神经网络结构适应不同比例、多尺寸的图像输入,并且可以有效地提取到害虫的多尺度特征信息。空间金字塔池化的原理如图3a所示。

在YOLOv3的首个全连接输出之前加入SPP结构,该结构的输入尺度为 $38 \times 38 \times 512$,如图3b所示,将3次最大池化的结果相融合,得到 $38 \times 38 \times 512$ 的固定输出用于首个全连接层的输入。该方法将3种不同尺度的特征进行融合,使卷积核感受的视野范围更大。在SPP的结构中,从上一个卷积层传递的特征保留了原始输入中的害

虫特征,这部分害虫特征大多为体型适中的害虫类别,通过下采样,将体型较大的害虫进行特征重提取,通过反卷积提取到体型更小的害虫特征,将3种特征相互融合,用于消除因害虫体型差异较大而导致有效特征信息不一致的影响,有助于提升体型差异明显的害虫分类准确率。

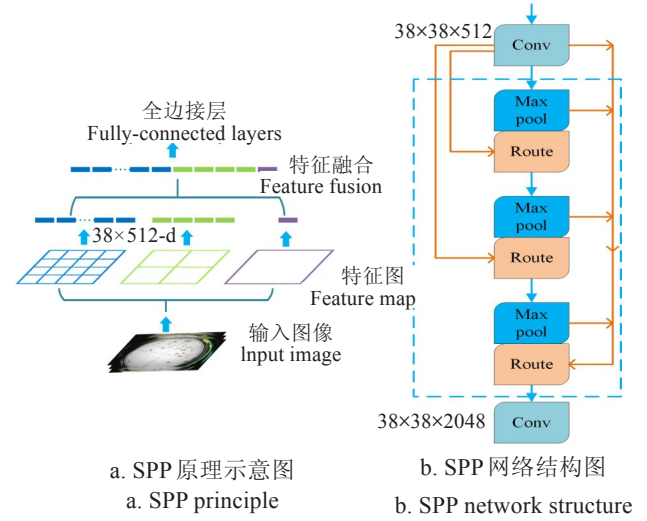


图3 SPP原理及网络结构示意图

Fig.3 SPP principle and network structure

为了进一步提取到害虫图像的细粒度特征,本文对YOLOv3进行改进,提升了YOLOv3网络的特征输出尺度,使用Resnet残差网络结构和反卷积操作,使多尺度卷积特征相融合;为了减小害虫体型差异较大对分类精度造成的影响,将空间金字塔池化与改进后的YOLOv3相融合,使网络模型能够适应多尺度的样本输入,设计一种基于端到端的害虫种类识别网络YOLOv3-SPP,用于害虫定位与害虫种类识别。

2 结果与分析

2.1 数据集与评价

从表1中害虫样本数量分布可以看出,样本数在50以下的害虫种类共有86类,呈现出长尾效应。这是由于害虫样本的采集工作集中在5月至9月,且分布在多个地域,导致各类害虫数量不均衡,出现长尾效应。为了保证训练样本的相对均衡,本次试验从样本数量不低于50的害虫种类中随机选取20类,试验数据集分布如表2所示,20类害虫总共包含10462个标注样本,分布在害虫样本的976张图片中。试验随机选取196张图片为验证集,其余780张为训练集,训练集与验证集均包含选取的20类害虫。其中,验证集只用于模型精度的测试,不参与模型的训练。

本文所有的试验都是在GeForce GTX Titan X 12GB GPU, Ubuntu16.04操作平台上实现的,为了充分利用硬件资源,以GPU版本的Darknet深度学习架构为实现框架。评价参数采用每一类的平均精度值 (average precision, AP) 和总的平均精度均值 (mean average precision, mAP), mAP的计算如式(8)~(10)所示。

$$Pr = \frac{TP}{TP + FP} \quad (8)$$

$$AP(i) = \frac{Pr_c}{N_c} \quad (9)$$

$$\text{mAP} = \frac{\sum_{i=1}^Q \text{AP}(i)}{Q} \times 100\%$$

(10)

式(8)~(10)中Pr为准确率,TP为真实的害虫正样本数量,FP为虚假的正样本数量,C为害虫类别, N_c 为包含类别C的所有样本数量,AP(i)表示第i类害虫的平均精度值,Q为所有害虫类别数。mAP的值越高,表示该方法在数据集上预测害虫的位置和类别越准确。

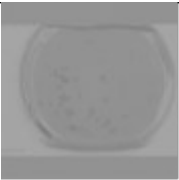





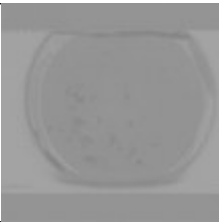





表2 20类害虫数据集分布

Table 2 Sample distribution of 20 species of pests

害虫种类 编号 Pest species No.	训练样本 数 Number of training samples	测试样 本数 Number of test samples	害虫种 类编号 Pest species No.	训练样本 数 Number of training samples	测试样 本数 Number of test samples
16	1 670	861	6	152	79
32	905	466	15	143	73
40	690	356	37	116	60
8	608	313	52	108	56
24	570	293	28	104	54
39	537	277	69	77	39
11	336	173	5	59	30
48	272	140	63	55	29
43	221	114	51	40	20
54	206	106	50	36	18

表3 YOLOv3与YOLOv3-SPP的部分网络特征重构

Table 3 Parts of network feature with YOLOv3 and YOLOv3-SPP

	Conv1 输出 Conv1 output	第2次降采样输出 2 nd downsampled output	第3次降采样输出 3 rd downsampled output	第4次降采样输出 4 rd downsampled output	第1次上采样输出 1 st upsampled output	第2次上采样输出 2 nd upsampled output
YOLOv3						
特征尺寸 Feature size	416×416	104×104	52×52	26×26	26×26	52×52
YOLOv3-SPP						
特征尺寸 Feature size	608×608	152×152	76×76	38×38	76×76	152×152

2.3 不同方法对20类害虫的识别准确率测试

为了验证本文害虫识别方法的有效性,在20类害虫数据集上,分别利用不同的目标检测方法对害虫样本进行训练并测试识别精度,数据集及识别结果如表4所示。

从表4可以看出,20类害虫按样本数量的大小降序排列。结合HOG(histogram of oriented gridients)特征与SVM的传统识别方法,识别率仅为68.46%,明显低于其余神经网络的识别方法,且识别速度较慢。由此可以看出,基于神经网络的目标检测算法在害虫识别中更具有优势。

在同一数据集下,YOLOv3和YOLOv3-SPP获得较高的平均识别率,分别为85.27%和88.07%,改进后的

2.2 YOLOv3与YOLOv3-SPP对比试验

为了验证SPP结构对提高害虫检测准确率的有效性,试验分别采用YOLOv3和加入SPP结构的YOLOv3-SPP网络,对20类害虫进行50 200次的迭代训练。模型初始化参数使用基础网络DarkNet-53的预训练模型,设定初始学习率为0.001,采用随迭代次数改变的学习率下降策略,当迭代次数到达40 000次时,损失值基本稳定。为了避免学习率过大,网络模型在最优值附近震荡,因此设定steps为40 000和45 000,当迭代次数达到40 000次和45 000次时,学习率下降10倍,使得网络模型能够逐渐逼近最优值,确保训练能更进一步的收敛。由于GPU的显存大小和计算能力不同,本次试验设定batch为64, subdivisions为16,其中batch表示每批训练的图片数目是64张,分为16个批次,依次送入GPU进行训练。

表3为YOLOv3与YOLOv3-SPP训练模型的部分网络特征重构图。由图2a可知,YOLOv3的3个输出结果是通过3次降采样与2次上采样的特征相融合所得。从表3中对比YOLOv3与YOLOv3-SPP的特征重构结果可以看出,YOLOv3所输出的特征尺寸较小,特征信息不明显,难以对小目标害虫进行检测与分类。而改进后的YOLOv3-SPP提升了网络模型的输入尺寸,减少了下采样次数,在特征重构的结果中保留了大量害虫样本的原始信息,能够提高害虫目标的检测与分类的准确率。

YOLOv3-SPP较YOLOv3的识别率提升2.8个百分点;Faster R-CNN作为2阶段(two stage)的目标检测算法,虽然在各别害虫种类上获得了较好的测试精度,但平均识别率较低,为78.43%。这是由于Faster R-CNN的网络结构不能有效地提取害虫的多尺度信息,对于体型较小的害虫目标识别率较差。由此可以看出,融合了空间金字塔池化的深度卷积神经网络YOLOv3-SPP在保证26 帧/s识别速度的同时,具有更好的害虫识别能力,较HOG+SVM的传统识别算法高出19.61个百分点,较Faster RCNN高出9.64个百分点。

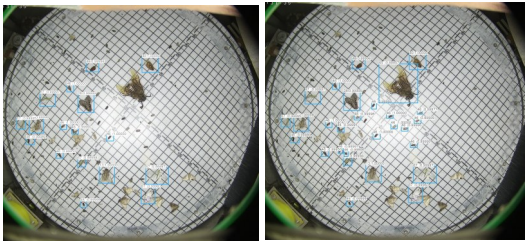
图4为YOLOv3与YOLOv3-SPP的识别结果,由图可

以看出, YOLOv3-SPP 比 YOLOv3 检测到的害虫目标更多, 即漏检率更低, 在 YOLOv3-SPP 的检测结果中, 除数据集未标注的昆虫外, 其余害虫全部被识别。对 2 种方法的识别结果细节区域进行对比, 如图 4c 和 4d 所示。YOLOv3 对于体型较大和体型较小的害虫目标漏检较多, 对

应图 4d 中的 65 号和 46 号害虫, 而 YOLOv3-SPP 能够识别出图中所有的害虫目标, 且分类精度较高。由此可见, 本文提出的 YOLOv3-SPP 害虫识别算法对于害虫体型差异较大, 特别是害虫体型较小的情况下识别效果明显优于 YOLOv3。

表 4 20 类害虫的测试精度
Table 4 Test accuracy of 20 species pest

害虫种类编号 Pest species No.	种类名称 Species name	训练样本 Training sample	测试样本 Test sample	识别率 Recognition rates /%			
				HOG+SVM	Faster R-CNN	YOLOv3	YOLOv3-SPP
16	三化螟 <i>Tryporyza. incertulas</i>	1670	861	80.37	89.36	90.62	95.1
32	暗黑鳃金龟 <i>Holotrichia parallela</i>	905	466	78.26	87.15	96.28	97.22
40	铜绿丽金龟 <i>Anomala. corpulenta motschulsky</i>	690	356	70.35	86.29	93.74	95.48
8	黄菜蛾 <i>Plutella maculipennis</i>	608	313	73.68	75.72	84.7	89.12
24	稻纵卷叶螟 <i>Cnaphalocrocis medinalis</i>	570	293	67.38	84.32	90.19	92.94
39	二化螟 <i>Chilo suppressalis</i>	537	277	65.13	79.48	82.44	84.6
11	甜菜夜蛾 <i>Beet armyworm</i>	336	173	71.29	84.15	86.99	88.37
48	龙虱 <i>Predaci diving beetle</i>	272	140	68.55	82.04	85.18	89.74
43	玉米螟 <i>Pyrausta nubilalis</i>	221	114	64.79	77.15	78.21	85.44
54	棉铃虫 <i>Helicoverpa armigera</i>	206	106	79.68	74.36	88.32	89.81
6	棉水螟 <i>Nymphula interruptalis</i>	152	79	72.43	78.69	84.55	85.98
15	地老虎 <i>Agrotis ypsilon</i>	143	73	67.32	77.27	78.98	84.48
37	蝼蛄 <i>Gryllotalpa spp</i>	116	60	68.96	69.29	77.95	82.74
52	小麦椿象 <i>Aelia sibirica reuter</i>	108	56	74.18	76.61	89.9	91.98
28	旱柳原野螟 <i>Proteuclasta stotzneri</i>	104	54	69.23	73.47	87.31	89.97
69	二点委夜蛾 <i>Athetis lepigone</i>	77	39	63.82	72.37	81.62	82.92
5	菜螟 <i>Hellula undalis Fabricius</i>	59	30	60.15	78.48	82.02	82.41
63	黄环绢野螟 <i>Palpita annulata</i>	55	29	61.23	74.31	75	80.86
51	毛首夜蛾 <i>Craniophora inquieta draudt</i>	40	20	57.68	73.17	86.25	86.08
50	粘虫 <i>Mythimna separata</i>	36	18	54.62	74.92	85.19	86.33
平均 Average				68.46	78.43	85.27	88.07
识别速度 Recognition speed/(帧·s ⁻¹)				21	6	32	26

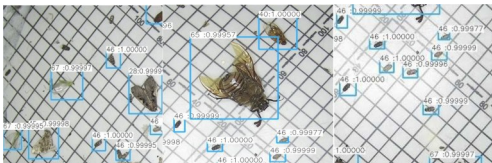


a. YOLOv3 识别结果
a. Recognition results of YOLOv3

b. YOLOv3-SPP 识别结果
b. Recognition results of YOLOv3-SPP



c. YOLOv3 识别细节
c. Recognition details of YOLOv3



d. YOLOv3-SPP 识别细节
d. Recognition details of YOLOv3-SPP

图 4 不同方法的害虫种类识别结果

Fig.4 Identification results of pest species by different methods

3 结 论

本文提出了一种基于空间金字塔池化与深度卷积神

经网络的作物害虫识别算法, 试验表明该算法能够有效地检测小目标害虫, 适用于多尺度的害虫定位以及种类识别, 平均识别准确率达到 88.07%, 检测速度为 26 帧/s。

1) 为了解决因害虫尺度多样性导致其识别精度相对较低的问题, 本文使用残差网络结构和反卷积对 YOLOv3 网络进行改进, 同时将空间金字塔与改进后的 YOLOv3 相结合, 设计了深度卷积神经网络模型 YOLOv3-SPP, 与 YOLOv3 相比在识别准确率上提升了 2.8 个百分点。

2) 对比了 HOG+SVM、Faster R-CNN、YOLOv3 在相同的试验环境下害虫识别的准确率, 本文算法较 HOG+SVM 的传统识别算法高出 19.61 个百分点, 较 Faster R-CNN 高出 9.64 个百分点。

3) 基于神经网络的识别算法依赖于数据集的规模及标注的准确性, 同时较多的网络参数需要大量的计算能力, 在将来的研究中应尽可能减少网络参数和训练模型对数据量的依赖。

[参 考 文 献]

[1] 姚青, 吕军, 杨保军, 等. 基于图像的昆虫自动识别与计数研究进展[J]. 中国农业科学, 2011, 44(14):2886—2899.
Yao Qing, Lü Jun, Yang Baojun, et al. Progress in research on digital image processing technology for automatic insect identification and counting[J]. Scientia Agricultura Sinica, 2011, 44(14): 2886—2899. (in Chinese with English abstract)

[2] Cui S, Ling P, Zhu H, et al. Plant pest detection using an artificial

- nose system: a review[J]. *Sensors*, 2018, 18(2): 378.
- [3] 李文勇, 李明, 陈梅香, 等. 基于机器视觉的作物多姿态害虫特征提取与分类方法[J]. *农业工程学报*, 2014, 30(14): 154—162.
Li Wenyong, Li Ming, Chen Meixiang, et al. Feature extraction and classification method of multi-pose pests using machine vision[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2014, 30(14): 154—162. (in Chinese with English abstract)
 - [4] 封洪强, 姚青. 农业害虫自动识别与监测技术[J]. *植物保护*, 2018, 44(5): 127—133, 198.
Feng Hongqiang, Yao Qing. Automatic identification and monitoring technologies of agricultural pest insects[J]. *Plant Protection*, 2018, 44(5): 127—133, 198. (in Chinese with English abstract)
 - [5] 胡永强, 宋良图, 张洁, 等. 基于稀疏表示的多特征融合害虫图像识别[J]. *模式识别与人工智能*, 2014, 27(11): 985—992.
Hu Yongqiang, Song Liangtu, Zhang Jie et al. Pest image recognition of multi-feature fusion based on sparse representation[J]. *Pattern Recognition and Artificial Intelligence*, 2014, 27(11): 985—992. (in Chinese with English abstract)
 - [6] 谢成军, 李瑞, 董伟, 等. 基于稀疏编码金字塔模型的农田害虫图像识别[J]. *农业工程学报*, 2016, 32(17): 144—151.
Xie Chengjun, Li Rui, Dong Wei, et al. Tianjiao. Recognition for insects via spatial pyramid model using sparse coding[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2016, 32(17): 144—151. (in Chinese with English abstract)
 - [7] 张超凡, 王儒敬, 谢成军. 基于多特征字典学习的害虫图像自动分类方法[J]. *计算机应用与软件*, 2017, 34(3): 142—147, 180.
Zhang Chaofan, Wang Rujing, Xie Chengjun, et al. Automatic classification method for pest image based on multi-feature dictionary learning[J]. *Computer Applications and Software*, 2017, 34(3): 142—147, 180. (in Chinese with English abstract)
 - [8] 张永玲, 姜梦洲, 俞佩仕, 等. 基于多特征融合和稀疏表示的农业害虫图像识别方法[J]. *中国农业科学*, 2018, 51(11): 2084—2093.
Zhang Yongling, Jiang Mengzhou, Yü Peishi, et al. Agricultural pest identification based on multi-feature fusion and sparse representation[J]. *Scientia Agricultura Sinica*, 2018, 51(11): 2084—2093. (in Chinese with English abstract)
 - [9] Ebrahimi M A, Khoshtaghaza M H, Minaei S, et al. Vision-based pest detection based on SVM classification method[J]. *Computers and Electronics in Agriculture*, 2017, 137: 52—58.
 - [10] Sriwastwa A, Prakash S, Swarit S, et al. Detection of pests using color based image segmentation[C]//2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT). IEEE, 2018: 1393—1396.
 - [11] 程鲁玉, 孟小艳, 达新民. 关于果林中果害虫图像特征高效分类识别仿真[J]. *计算机仿真*, 2018, 35(2): 425—428.
 - [12] Cheng Luyu, Meng Xiaoyan, Da Xinmin, et al. Simulation of efficient classification for image characteristics of fruit in fruit pest[J]. *Computer Simulation*, 2018, 35(2): 425—428. (in Chinese with English abstract)
 - [13] 田冉, 陈梅香, 董大明, 等. 红外传感器与机器视觉融合的果树害虫识别及计数方法[J]. *农业工程学报*, 2016, 32(20): 195—201.
Tian Ran, Chen Meixiang, Dong Daming, et al. Identification and counting method of orchard pests based on fusion method of infrared sensor and machine vision[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2016, 32(20): 195—201. (in Chinese with English abstract)
 - [14] 潘春华, 肖德琴, 林探宇, 等. 基于 SVM 和区域生长结合算法的南方主要蔬菜害虫分类识别[J]. *农业工程学报*, 2018, 34(8): 192—199.
Pan Chunhua, Xiao Deqin, Lin Tanyu, et al. Classification and recognition for major vegetable pests in Southern China using SVM and region growing algorithm[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2018, 34(8): 192—199. (in Chinese with English abstract)
 - [15] 张苗辉, 李俊辉, 李佩琛. 基于深度学习和稀疏表示的害虫识别算法[J]. *河南大学学报自然科学版*, 2018, 48(2): 207—213.
Zhang Miaohui, Li Junhui, Li Peichen, et al. Algorithm on pest identification based on depth learning and sparse representation[J]. *Journal of Henan University Natural Science*, 2018, 48(2): 207—213. (in Chinese with English abstract)
 - [16] 杨国国, 鲍一丹, 刘子毅. 基于图像显著性分析与卷积神经网络的茶园害虫定位与识别[J]. *农业工程学报*, 2017, 33(6): 156—162.
Yang Guoguo, Bao Yidan, Liu Ziyi. Localization and recognition of pests in tea plantation based on image saliency analysis and convolutional neural network[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2017, 33(6): 156—162. (in Chinese with English abstract)
 - [17] 梁万杰, 曹宏鑫. 基于卷积神经网络的水稻虫害识别[J]. *江苏农业科学*, 2017, 45(20): 241—243, 253.
Liang Wanjie, Cao Hongxin. Identification of rice insect pests based on CNN model[J]. *Jiangsu Agricultural Sciences*, 2017, 45(20): 241—243, 253. (in Chinese with English abstract)
 - [18] Alfarisy A A, Chen Q, Guo M. Deep learning based classification for paddy pests & diseases recognition[C]//Proceedings of 2018 International Conference on Mathematics and Artificial Intelligence. ACM, 2018: 21—25.
 - [19] 魏杨, 毕秀丽, 肖斌. 基于区域卷积神经网络的农业害虫检测方法[J]. *计算机科学*, 2018, 45(S2): 226—229, 233.
Wei Yang, Bi Xiuli, Xiao Bin. Agricultural insect pest detection method based on regional convolutional neural network[J]. *Computer Science*, 2018, 45(S2): 226—229, 233. (in Chinese with English abstract)
 - [20] 孙钰, 周焱, 袁明帅, 等. 基于深度学习的森林虫害无人机实时监测方法[J]. *农业工程学报*, 2018, 34(21): 74—81.
Sun Yu, Zhou Yan, Yuan Mingshuai, et al. UAV real-time monitoring for forest pest based on deep learning[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2018, 34(21): 74—81. (in Chinese with English abstract)
 - [21] Yue Y, Cheng X, Zhang D, et al. Deep recursive super resolution network with Laplacian Pyramid for better agricultural pest surveillance and detection[J]. *Computers and electronics in agriculture*, 2018, 150: 26—32.
 - [22] 赵德安, 吴任迪, 刘晓洋, 等. 基于 YOLO 深度卷积神经网络的复杂背景下机器人采摘苹果定位[J]. *农业工程学报*, 2019, 35(3): 164—173.
Zhao Dean, Wu Rendi, Liu Xiaoyang, et al. Apple positioning based on YOLO deep convolutional neural network for picking robot in complex background[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2019, 35(3): 164—173. (in Chinese with English abstract)
 - [23] 程曦, 吴云志, 张友华, 等. 基于深度卷积神经网络的储粮害虫图像识别[J]. *中国农学通报*, 2018, 34(1): 154—158.
Cheng Xi, Wu Yunzhi, Zhang Youhua, et al. Image recognition of stored grain pests: Based on deep convolutional neural network[J]. *Chinese Agricultural Science Bulletin*, 2018, 34(1): 154—158. (in Chinese with English abstract)

- [24] He K, Zhang X, Ren S, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification[C]//Proceedings of the IEEE international conference on computer vision. 2015: 1026—1034.
- [25] Baker N, Lu H, Erlikhman G, et al. Deep convolutional networks do not classify based on global object shape[J]. PLoS computational biology, 2018, 14(12): e1006613.
- [26] Ma B, Xia Y. Autonomous deep learning: a genetic DCNN designer for image classification[J / OL]. 2018. <https://arxiv.org/pdf/1807.00284.pdf>. [2019-04-24].
- [27] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770—778.
- [28] 袁培森, 黎薇, 任守纲, 等. 基于卷积神经网络的菊花花型和品种识别[J]. 农工业工程学报, 2018, 34(5):152—158.
- Yuan Peisen, Li Wei, Ren Shougang, et al. Recognition for flower type and variety of chrysanthemum with convolutional neural network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2018, 34(5):152—158. (in Chinese with English abstract)
- [29] 王璨, 武新慧, 李志伟. 基于卷积神经网络提取多尺度分层特征识别玉米杂草[J]. 农工业工程学报, 2018, 34(5):144—151.
- Wang Can, Wu Xinhui, Li Zhiwei. Recognition of maize and weed based on multi-scale hierarchical features extracted by convolutional neural network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2018, 34(5): 144—151. (in Chinese with English abstract)
- [30] 马浚诚, 刘红杰, 郑飞翔, 等. 基于可见光图像和卷积神经网络的冬小麦苗期长势参数估算[J]. 农工业工程学报, 2019, 35(5):183—189.
- Ma Juncheng, Liu Hongjie, Zheng Feixiang, et al. Estimating growth related traits of winter wheat at seedling stages based on RGB images and convolutional neural network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(5):183—189. (in Chinese with English abstract)
- [31] Backsan M. Pest control system using deep learning image classification method[J]. Journal of the Korea Society of Computer and Information, 2019, 24(1): 9—23.
- [32] Redmon J, Farhadi A. Yolov3: An incremental improvement[J/OL]. <https://arxiv.org/pdf/1804.02767.pdf>. [2019-04-24].
- [33] He K, Zhang X, Ren S, 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.

Crop pest identification based on spatial pyramid pooling and deep convolution neural network

Zhang Bo¹, Zhang Miaohui^{1, 2*}, Chen Yunzhong¹

(1. Henan Key Laboratory of Big Data Analysis and Processing, Kaifeng 475000, China;

2. Postdoctoral Research Station of Geography, Kaifeng 475000, China)

Abstract: Traditional methods for classifying crop pests are based on the appearance of the pests such as their color, morphology and texture using algorithms such as back propagation neural network and support vector machine. These methods are sensitive to the environments where the pests appear. Also, the imbalance between training sample numbers, difference in scales at which the training samples are taken, along with that the pests might seasonally change their color and shape, make it difficult for these algorithms to detect and recognize the pests. In order to improve the accuracy of detecting crop pests appearing in diverse environments, this paper proposed a deep convolutional neural network method by combining the spatial pyramid pooling with the improved YOLOv3 deep convolutional neural network algorithm. The proposed method first located the pests on a test image and then identified the species they belongs to. Using the improved network structure in YOLOv3, the proposed method used up-sample and convolution operations to calculate the de-convolution. Bilinear interpolation was used to enhance the output of the network, and the depth characteristics of the pests were extracted by the depth residual neural network. As a result, the improved network can effectively detect and recognize small pests in the images. By fusing with the spatial pyramid pooling, the Yolov3-SPP network can map the extracted eigenvectors to different spatial dimensions. The network can be used at various scales to detect pests of different sizes. Results from identifying and testing 20 types of pests collected by traps showed that the average accuracy of the proposed method was 88.07%, with a detection speed of 26 frames/s. Compared with the typical YOLOv3 algorithm, the proposed methods improved accuracy by 2.8 percentage points. We also reconstructed the maps of the network at each stage and compared them with the typical YOLOv3 algorithm in attempts to demonstrate that the features extracted by the proposed method was more recognizable. The results revealed that the typical YOLOv3 algorithm could miss the targeted objects, while the proposed method is not only able to detect pests of different sizes, it also has high recognition accuracy. Using the same dataset, we compared the proposed method with other detection algorithms such as HOG+SVM, Faster R-CNN, YOLOv3. The comparison showed that the proposed method was 19.61 percentage points higher than that of HOG + SVM, and 9.64 percentage points higher than that of Faster R-CNN, Faster R-CNN was unable to extract small pests, and its recognition accuracy was only 78.43%.

Keywords: image recognition; algorithms; pest classification; deep convolutional neural network; space pyramid pooling; deconvolution