

# 基于区域卷积神经网络的农业害虫检测方法

魏 杨 毕秀丽 肖 斌

(重庆邮电大学计算智能重庆市重点实验室 重庆 400065)

**摘 要** 当前农业害虫综合防治中,农业害虫检测主要通过专业人员手动收集和分类实地样本,这种手动分类方法既昂贵又耗时。现有的通过计算机实现的自动农业害虫检测对害虫所处背景环境的要求较高,并且无法实现农业害虫的定位。针对这些问题,文中基于深度学习的思想,提出了一种新的农业害虫自动检测方法,它由区域提取网络和 Fast R-CNN 两个部分组成。区域提取网络在任意大小且背景繁杂的图像上的某一个或多个区域进行特征提取,得到农业害虫的初步位置候选区;将农业害虫的初步位置候选区作为 Fast R-CNN 的输入,Fast R-CNN 通过学习农业害虫种类的种内差异和种间相似性,判定初步位置候选区中的目标类别并计算精准坐标。文中同时建立了一个已标注标签的实际场景的农业害虫数据库,将提出的农业害虫检测方法在此数据库上进行测试,识别精度的均值可达到 82.13%。实验结果表明,提出的方法能够有效地提升农业害虫类别判断的准确度,得到农业害虫的精准定位,优于以往的自动化农业害虫检测方法。

**关键词** 农业害虫数据库,精准定位,害虫分类

中图分类号 TP301 文献标识码 A

## Agricultural Insect Pest Detection Method Based on Regional Convolutional Neural Network

WEI Yang BI Xiu-li XIAO Bin

(Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

**Abstract** In the current integrated agricultural pest control, agricultural insect pests are detected primarily by professionals' sample collection and sorting manually, such manual classification method is both expensive and time consuming. Existing computer-aided automatic detection of agricultural pests has a high requirement on the background environment of pests and cannot locate agricultural pests. To solve these problems, this paper proposed a new method for automatic detection of agricultural pests based on the idea of the deep learning. It contains the region proposal network and the Fast R-CNN network. Region proposal network extracts feature in one or more regions of arbitrary size and complicated background images, then gets preliminary position of the candidate regions of agricultural pests. Preliminary position of the candidate regions of agricultural pests is taken as an input to Fast R-CNN. Fast R-CNN finally learns the classification of target in the preliminary location candidate area and calculates exact coordinates by studying the intraspecific differences and interspecies similarity of agricultural pests. Meanwhile, this paper also established a labeled actual scene tag agricultural pests database, and the proposed method was tested on this database, with the average precision up to 82.13%. The experimental results show that the proposed method can effectively enhance the accuracy of agricultural pests detection, and get accurate positions, and is superior to the previous automated agricultural pest detection methods.

**Keywords** Agricultural pest database, Precise positioning, Pest Classification

## 1 引言

农业害虫是已知的造成世界农业经济受损的主要因素之一<sup>[1]</sup>,作为昆虫的一种,其主要依赖于各种植物和农作物生存,给农、林、牧业都造成了不同程度的危害。农业害虫带来的经济影响遍及全世界,欧洲农业经济损失每年达到 28.2%,北美达到 31.2%,亚洲和非洲的经济损失高达 50% 以上<sup>[2]</sup>。以蝗虫为例,全世界的有害蝗虫共计 900 余种,在我国占 60 余种,蝗虫会严重损害牧草和农作物的生长并影响农牧业的

发展。自 20 世纪 60 年代以来,病虫害综合防治(IPM)<sup>[3]</sup>已成为主要的虫害防治模式,IPM 根据对不同种类害虫进行压力检测的结果,制定了有利于经济发展、维护生态和社会学意义的最佳农药建议。由此可见,虫害的精准识别和定位对于 IPM 来说至关重要。目前大多数检测方法是昂贵且耗时的,因为它们需要 IPM 专业人员手动收集和分类实地样本,这导致缺少 IPM 技术支持的发展中国家无法利用这些技术进行虫害防治。由此可见,在病虫害综合防治方面,急需一种快速、低成本的农业害虫自动检测方法。

本文受国家自然科学基金(61572092),国家自然科学基金-广东联合基金(U1401252),国家重点研发计划(2016YFC1000307-3)资助。

魏 杨(1993-),男,硕士生,主要研究方向为图像篡改检测;毕秀丽(1982-),女,博士,讲师,主要研究方向为数字图像处理、多媒体信息安全及数字水印等,E-mail:bixl@cqupt.edu.cn;肖 斌(1982-),男,博士,教授,主要研究方向为图像处理、模式识别及数字水印等。

Mundada 等<sup>[4]</sup>在 SVM<sup>[5]</sup>的基础上实现了农业害虫检测,该方法使用背景单一、图片大小固定且经过预处理的数据作为训练集,在两种类别的害虫检测中取得了较好的结果。深度学习是一个近十余年来快速发展的理论,在特征提取、模型泛化与拟合方面的性能优越,受到了越来越多研究者的关注。由于训练数据的不断扩张和完善,随着硬件性能的提升,深度学习得到了进一步发展,至今在计算机视觉、语义分析、自然语言处理、语音识别等多个领域成效卓著。神经网络作为当前深度学习的主要形式,衍生出一系列优秀网络结构,如卷积神经网络(Convolutional Neural Network, CNN<sup>[6]</sup>)、循环卷积神经网络(Recurrent Neural Networks, RNN<sup>[7]</sup>)、生成式对抗网络(Generative Adversarial Networks, GAN<sup>[8]</sup>)、区域卷积神经网络(Regional Convolutional Neural Network, RCNN<sup>[9]</sup>)等。最近, Liu 等<sup>[10]</sup>基于原生“AlexNet”CNN 网络提出了一种农业害虫检测模型,此模型可进行多类别害虫检测,但需要先对输入图片进行显著区域提取,再将提取的区域归一化到统一大小进行深度学习分类,该方法在显著区域提取时受场景背景复杂度的影响较大,并且只能实现单一目标的检测。现实农业害虫分类的实际环境非常复杂,为了实现更有效且适用范围更广的农业害虫检测技术,满足使用最少、最便捷的操作完成专家级的农业害虫检测的需求,本文将深度学习和农业害虫检测相结合,提出了基于区域卷积神经网络的农业害虫检测方法,在农业害虫训练集不足和使用真实数据(未进行预处理的背景繁杂的数据)的情况下,该方法都获得了很好的检测效果。在使用真实数据进行害虫检测时,本文较 Mundada 等和 Liu 等提出的方法在农业害虫检测准确率上有很大提升,同时能实现图像中害虫位置的精准定位。

## 2 基于区域卷积神经网络的农业害虫检测

基于区域卷积神经网络的农业害虫检测方法由两部分组成:区域提取网络和 Fast R-CNN。在区域提取网络中,通过 ImageNet 数据库训练好的 VGG-16 网络<sup>[11]</sup>(抛弃掉最后两层全连接层)提取输入图像  $I$  的原始特征  $F$  (VGG-16 网络输出 512 幅特征图),在原始特征  $F$  上滑动一个窗口进行卷积操作,同时在原始特征  $F$  的每一个位置上取得若干候选区,根据卷积得到的特征向量分别通过窗口分类层和窗口回归层,得到每个候选区的分类(目标或非目标)以及原始坐标。区域提取网络部分的目的是根据候选区集合中输出分类为目标候选区集合获得图像中农业害虫的初步位置候选区。在 Fast R-CNN 网络中,农业害虫的初步位置候选区的原始坐标会被映射到原始特征  $F$  上,并得到目标类候选区特征,目标类候选区特征经过卷积处理后输入到分类层和位置精修层,最后得到害虫分类结果和其在图像中的位置。2.1 节对区域提取网络的操作流程进行了说明,2.2 节详细描述了 Fast R-CNN 的流程,2.3 节描述了提出的检测方法的训练方法和测试方法。

### 2.1 区域提取网络

区域提取网络主要根据原始特征  $F$ , 得到农业害虫在输入图像  $I$  中的初步位置候选区。区域提取网络的详细说明如图 1 所示。以  $3 \times 3$  的窗口在原始特征  $F$  上进行滑动,在每个滑窗内进行卷积核大小为  $3 \times 3$  的卷积操作,得到一个 512 维的特征向量,同时以每个滑窗的中心为中心选择 9 个不同大

小的候选区,每个特征向量分别经过  $1 \times 1$  的卷积核进行卷积操作后输入到窗口分类层和窗口回归层,最后得到每个候选区的类别(目标或非目标)以及坐标结果。

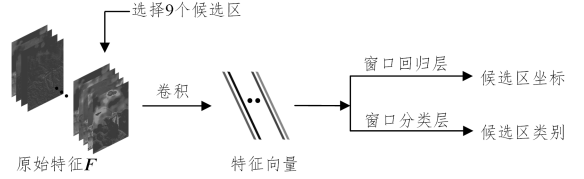


图 1 区域提取网络的流程图

在区域提取网络中,以  $3 \times 3$  的滑动窗口中心为中心选取 9 个大小不同的候选区,选择候选区的方式如图 2 所示。为了避免输入图像大小变化的影响,9 个候选区映射到原始图像上的尺寸需要保证为:

$$\{128 \times 128, 256 \times 256, 512 \times 512\} \times \{1:1, 1:2, 2:1\}$$

利用窗口分类层输出每一个位置上 9 个候选区分别属于目标或非目标的概率,并且窗口分类层采用一个 softmax 函数进行分类。使用的 softmax 函数如下:

$$\begin{bmatrix} P(t=1|z) \\ P(t=2|z) \end{bmatrix} = \frac{1}{\sum_{d=1}^2 e^{z_d}} \begin{bmatrix} e^{z_1} \\ e^{z_2} \end{bmatrix} \quad (1)$$

其中,  $z$  表示输入的特征向量,  $\sum_{d=1}^2 e^{z_d}$  项对所有类别的概率分布进行归一化,使得所有类别的概率和为 1,  $P(t=1|z)$  表示属于类别 1(即目标)的概率,  $P(t=2|z)$  表示属于类别 2(即非目标)的概率。窗口回归层会输出目标类候选区在原始图像坐标上的中心坐标  $(x, y)$  以及宽高  $(w, h)$ , 通过式(2)计算得到窗口回归层输出的向量化后的目标类候选区的中心坐标、宽度及高度  $(t_x, t_y, t_w, t_h)$ 。

$$\begin{aligned} t_x &= (x - x_a) / w_a, t_y = (y - y_a) / h_a \\ t_w &= \log(w / w_a), t_h = \log(h / h_a) \end{aligned} \quad (2)$$

其中,目标类候选区对应的滑动窗口的中心坐标为  $(x_a, y_a)$ , 宽高为  $(w_a, h_a)$ 。输入图像中农业害虫所在区域的坐标及宽高为  $(x^*, y^*, w^*, h^*)$ , 通过式(3)可以得到向量化后的农业害虫真实区域的坐标编码  $(t_x^*, t_y^*, t_w^*, t_h^*)$ 。

$$\begin{aligned} t_x^* &= (x^* - x_a) / w_a, t_y^* = (y^* - y_a) / h_a \\ t_w^* &= \log(w^* / w_a), t_h^* = \log(h^* / h_a) \end{aligned} \quad (3)$$

最后,通过最小化代价函数来对窗口分类层和窗口回归层输出的结果进行优化,其代价函数如下:

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (4)$$

其中,  $L_{cls}(p_i, p_i^*)$  和  $L_{reg}(t_i, t_i^*)$  分别为窗口分类层和窗口回归层的代价函数,两者分别如式(5)和式(6)所示:

$$L_{cls}(p, u) = -\log p_u \quad (5)$$

$$L_{reg}(t, t^*) = \sum_{j \in \{x, y, w, h\}} \text{smooth}_{L_1}(t_j - t_j^*) \quad (6)$$

式(4)中,  $cls$  代表窗口分类层,  $reg$  代表窗口回归层;  $i$  表示候选区的索引;  $p_i$  是候选区  $i$  预测为目标概率;若候选区  $i$  是目标,则真实样本标记  $p_i^*$  为 1, 否则为 0;  $t_i$  为预测的目标候选区的坐标向量化的结果;  $t_i^*$  为真实目标候选区的坐标向量化的结果;  $N_{cls}$  是  $cls$  项的归一化值,每一个批次的数量为 256;  $N_{reg}$  为  $reg$  项的归一化值,候选区位置的数量为 2400。式(5)中,  $p$  指代  $p_i$ ,  $u$  指代  $p_i^*$ 。差异函数  $\text{smooth}_{L_1}(x)$ <sup>[12]</sup> 的定义如式(7)所示:

$$\text{smooth}_{L1}(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \quad (7)$$

在区域提取网络中,因为端到端的过程能使检测实时进行并且耗费的资源更少,本文根据反向传播算法<sup>[13]</sup>和随机梯度下降<sup>[14]</sup>来对代价函数进行优化,优化过程如式(8)所示:

$$\delta_i^2 = \frac{\partial L}{\partial s_i^2} = \frac{\partial \sum_{i=0}^k \frac{1}{2} (\overline{y_i} - y_i)^2}{\partial s_i^2} = m_i \cdot \theta'(s_i^2) \quad (8)$$

其中, $L$ 表示代价函数, $s_i^2$ 表示输出层第*i*个神经元的输入, $k$ 为神经元的个数, $\overline{y_i} - y_i$ 表示目标*i*的预测值与真实值之间的误差, $m_i$ 为 $\overline{y_i} - y_i$ 的简写, $\theta$ 指代激活函数。计算得到 $\delta_i^2$ 后,再根据式(9)计算权重梯度 $\partial L / \partial \omega_{ij}^2$ ,进而得到新的权重,从而更新各个隐层的参数,达到训练区域提取网络的目的。

$$\frac{\partial L}{\partial \omega_{ij}^2} = \frac{\partial L}{\partial s_j^2} \cdot \frac{\partial s_j^2}{\partial \omega_{ij}^2} = \delta_j^2 \cdot \theta(s_j^2) \quad (9)$$

其中, $\omega_{i,j}$ 表示神经元的权重, $s_i^j$ 表示中间层第*i*个节点的输入。

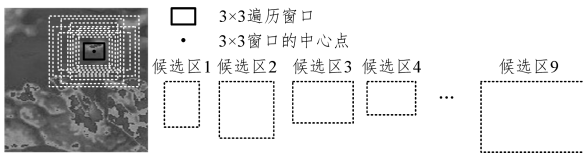


图2 候选区示例

## 2.2 Fast R-CNN

Fast R-CNN主要是对区域提取网络输出的目标类候选区进行精修得到图像中农业害虫区域的准确中心坐标及宽高,同时判断出区域内目标的害虫类别。Fast R-CNN的结构如图3所示,在原始特征 $F$ 映射出目标类候选区对应的候选区特征,通过感兴趣区域校准池化操作(ROIAlign Pooling<sup>[15]</sup>)将候选区特征统一到相同长度的特征向量,再将特征向量输入给两个并行的全连接层:位置精修层和分类层,最后得到图片中农业害虫的准确位置和害虫的类别。下面将详细叙述检测方法中Fast R-CNN的操作过程。

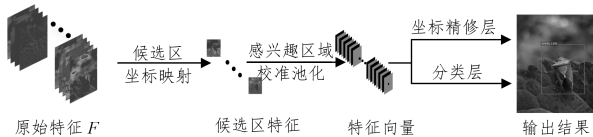


图3 Fast R-CNN的流程

将区域提取网络输出的目标类候选区映射到原始特征 $F$ ,从而得到一系列的目标类候选区特征,然后通过感兴趣区域校准池化层(ROIAlign Pooling<sup>[15]</sup>)统一每个目标类候选区特征的大小。ROIAlign Pooling 移除了区域卷积神经网络<sup>[9]</sup>中ROI Pool粗糙的量化,从原图到特征图之间的感兴趣区域映射直接使用双线性插值,不再使用取整操作,以尽可能地减小误差。每个特征候选区被均匀分成 $m \times n$ 块,对每一块进行最大池化(max pooling)操作,这样每一个目标类候选区特征都会得到大小统一的特征向量。之后将得到的特征向量分别输入到两个并行的全连接层(分类层和位置精修层)中,分类层连接softmax函数输出分类结果,位置精修层输出窗口坐标。为了降低预测的分类结果和窗口坐标与真实样本标记的误差,本文对预测的分类结果和窗口坐标与真实样本标记进行误差计算,通过代价函数分别对预测的分类结果和窗口

坐标进行优化,整体的代价函数如式(10)所示:

$$L(p, u, t^u, v) = L_{cls}(p, u) + \lambda [u \geq 1] L_{loc}(t^u, v) \quad (10)$$

其中, $L_{cls}(p, u)$ 和 $L_{loc}(t^u, v)$ 分别为分类层和位置精修层的代价函数,如式(11)、式(12)所示:

$$L_{cls}(p, u) = -\log p_u \quad (11)$$

$$L_{loc}(t^u, v) = \sum_{i \in \{x, y, w, h\}} \text{smooth}_{L1}(t_i^u - v_i) \quad (12)$$

式(11)中, $p_u$ 为农业害虫类别*u*的概率。式(12)中, $i \in \{x, y, w, h\}$ 代表在进行位置代价函数计算时,需要对预测窗口坐标 $t^u = (t_x^u, t_y^u, t_w^u, t_h^u)$ 与实际窗口坐标 $v = (v_x, v_y, v_w, v_h)$ 的中心点 $(x, y)$ 、窗口的宽高 $(w, h)$ 都进行累计,且差异函数 $\text{smooth}_{L1}(x)$ 的定义如式(7)所示。在式(10)中,类别 $u > 0$ 代表农业害虫,而类别 $u = 0$ 表示背景类。最后根据反向传播算法和随机梯度下降来对代价函数进行优化,更新Fast R-CNN的权值参数从而达到最终收敛。

## 2.3 基于区域卷积神经网络的农业害虫检测方法的训练和测试

本文将农业害虫训练数据输入至本文提出的检测模型中,权重在迭代更新(至少4万次)后,整体代价函数趋于平稳直至收敛。根据不同训练参数的实验结果对比分析,最终确定学习率为0.001,动量为0.95,权重衰减为0.0005。由于VGG-16卷积层数较深,为了避免梯度消失,加入批量归一化(Batch Normalization),使网络模型更好地收敛。图4给出了随机从测试数据集中选取3幅图像数据进行测试的结果,矩形框即为检测结果,左上角的英文及数字分别是候选框对应的害虫类别和预测概率。图4(a)为叶蝉类农业害虫的检测结果,图4(b)为天牛类农业害虫的检测结果,图4(c)为金龟子类农业害虫的检测结果。



图4 基于区域卷积神经网络的农业害虫检测方法的检测结果实例

## 3 实验结果

### 3.1 实验数据

因为当前还没有针对农业害虫检测的公共数据可以使用,所以本文从网络上收集农业害虫图片数据,最终整理出包含10个类别的农业害虫数据库,如表1所列,平均每类图像包含约250幅图像,共计2472幅图像。

表1 农业害虫数据库中的害虫类别及数量

类别	数量(2472)	比例/%
蚜虫(aphid)	245	9.91
黏虫(armyw)	176	7.12
金龟子(chaffer)	284	11.49
小菜蛾(dbmoth)	109	4.41
叶蝉(leafho)	256	10.36
天牛(lhbeetle)	310	12.54
蝗虫(locust)	288	11.65
蝼蛄(molec)	241	9.75
飞虱(plantho)	258	10.44
椿(stinkbu)	305	12.34

为了使检测模型能更好地学习图像特征,同时全面验证检测模型的真实效果,数据库包含了简单和复杂的真实背景图片、不同角度及被部分遮挡需检测的目标图片、明和暗的图片、不同分辨率的图片等。数据库中每幅图像都标注了相对应的标签,以便输入检测模型进行监督训练和测试。

### 3.2 实验结果及分析

本文中所有实验都是在 Intel XeonE5-2603 v4 六核 (X2) CPU,32 GB 内存,Ubuntu16.04 操作系统的平台上实现的。为了充分利用硬件资源,以 Google 的 TensorFlow 深度学习架构为框架。评价参数采用平均准确率均值 (mean Average Precision, mAP),即每个类别经过检测后的平均准确率的均值,如式(3)所示:

$$mAP = \frac{\sum_{q=1}^Q AP(q)}{Q} \times 100\% \quad (13)$$

其中,  $Q$  是害虫类别总数,  $AP(q)$  表示类别  $q$  的检测结果的平均准确率。在利用农业害虫数据库检测模型训练时,为了保证所有样本都能被用作训练和测试,并且避免偶然性,本文采用  $K$  折交叉验证 ( $K$ -fold cross-validation<sup>[16]</sup>),其中  $k$  选取 10,每次选择 9 个子集作为训练数据(大约 2250 张),1 个子集作为测试数据(大约 250 张)。图 5 给出了每类农业害虫的平均准确率以及最终的平均准确率均值(图中黑色虚线),由图 5 可知,大部分害虫类别的检测平均准确率较高,平均准确率均值  $mAP$  达到 82.13%。其中,aphid(蚜虫)的平均准确率相较于其他几类最低,只有 49.13%,拉低了近 3%的  $mAP$ 。随机从测试集的蚜虫类中取 2 张图片进行检测,并进一步分析其原因,检测结果如图 6 所示。从图 6(a)可观察到,检测模型能正常检测出图像中的蚜虫,同时也可得到蚜虫在图像中的位置。图 6(b)所示为检测失败的结果,将图 6(b)与图 6(a)进行比较可以发现,图 6(b)中的蚜虫部分较密集,且目标较小,这极大地增加了检测的难度。进一步对数据库中的蚜虫类进行分析,发现包含蚜虫类的图片中蚜虫目标小且分布密集。因此,本文提出的检测模型受候选区域特性(具有一定大小的尺度限制)的影响,对小目标的特征学习有一定的局限性,这造成了检测方法在对小目标进行检测时容易失效。

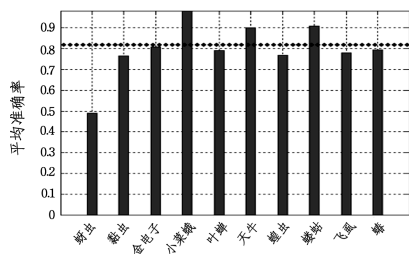
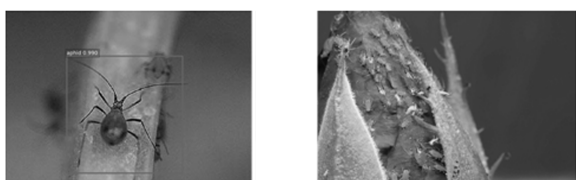


图 5 基于区域卷积神经网络的农业害虫检测结果



(a) 检测成功的蚜虫图片及定位结果

(b) 检测失败的蚜虫图片

图 6 基于区域卷积神经网络的农业害虫检测方法的检测结果

### 3.3 实验对比

本文选择了 3 种检测方法进行对比。第一种是 Mundada

等<sup>[4]</sup>提出的基于 SVM 进行农业害虫检测的方法,为方便后文叙述,暂且将其记为 SVM,该方法对训练数据进行了预处理,但为了保证对比实验的统一性,省去了预处理步骤。第二种是 Liu 等<sup>[10]</sup>提出的在原生 AlexNet 网络结构上进行改动后的害虫检测框架(此框架需要固定输入图片的大小),为方便后文叙述,暂且将其称为 AlexNet。由于本文的害虫数据库的图片数据的大小不固定,因此需在 AlexNet 的基础上稍作改动,将 AlexNet 中的全连接层换成全局平均池化 (Global Average Pooling, GAP<sup>[15]</sup>),GAP 可以保证输入的图像数据的大小无需固定,同时降低过拟合,更有益于模型的收敛及泛化。第三种是使用 ROI Pool 操作的原始 Faster R-CNN。使用相同测试数据,分别以本文提出的方法和 SVM、AlexNet 框架以及原始 Faster R-CNN 进行测试,对比结果如表 2 所列。可以发现,本文提出的检测模型的分类精度相较于 AlexNet 提高了 20% 左右,整体检测分类精度有较大的提升,同时本文提出的检测模型还能对图像中害虫所在区域进行准确定位;相较于原始 Faster R-CNN 提高了 2% 左右,且检测框的定位更加准确。

表 2 检测结果对比表

方法	准确率均值/%
SVM <sup>[4]</sup>	54.59
AlexNet <sup>[10]</sup>	71.77
Faster R-CNN(ROI Pool)	80.59
本文方法	82.13

**结束语** 本文受深度学习中运用卷积神经网络进行目标检测的启发,提出了一种基于区域卷积神经网络的农业害虫检测方法。该方法首先利用原始特征提取网络 VGG-16 来提取原始特征,然后在原始特征上再进行卷积,并在原始特征的每一个位置上取 9 个候选区域,计算每个候选区域属于目标类的概率值及坐标编码,同时计算概率值和坐标输入代价函数,利用反向传播和随机梯度算法来对代价函数进行优化,得到最终用于农业害虫检测的模型。实验结果表明,本文提出的方法能较好地农业害虫检测,其检测结果准确,检测速度快。但是,本文方法也存在一些不足,例如方法受到区域提取网络中候选区域大小的显示,对于小目标的检测不够理想。

### 参考文献

- [1] ESTRUCH J J, CAROZZI N B, DESAI N, et al. Transgenic plants: an emerging approach to pest control [J]. Nature Biotechnology, 1997, 15(2): 137-141.
- [2] LI Y, XIA C, LEE J. Detection of small-sized insect pest in greenhouses based on multifractal analysis [J]. Optik-International Journal for Light and Electron Optics, 2015, 126(19): 2138-2143.
- [3] PARSA S, MORSE S, BONIFACIO A, et al. Obstacles to integrated pest management adoption in developing countries [J]. Proceedings of the National Academy of Sciences, 2014, 111(10): 3889-3894.
- [4] MUNDADA R G, GOHOKAR V. Detection and classification of pests in greenhouse using image processing [J]. IOSR Journal of Electronics and Communication Engineering, 2013, 5(6): 57-63.
- [5] CORTES C, VAPNIK V. Support-vector networks [J]. Machine Learning, 1995, 20(3): 273-297.



(a)本文算法 (b)KCF (c)TLD (d)CSK

图 6 Tiger2(#112 和 #257)

**结束语** 在传统 KCF 的目标跟踪的基础上,提出了一种基于 KCF 的多模板匹配目标跟踪方法。首先通过多模板匹配得到多个样本,然后通过改进的 k-means 方法得到目标位置,最后通过随机模板更新策略更新参数和模板。该方法能在跟踪过程中准确地检测到目标的正确位置,因此可以获得更多的有效目标信息。实验结果表明,在距离精度中,本文方法比传统 KCF 高了 8.41%。本文方法在复杂的场景中获得了更好的鲁棒性和自适应性,提高了跟踪的成功率。因此,改进模板选取、提高特征多样性、采用多模板匹配的 KCF 是我们未来的研究内容。

### 参 考 文 献

- [1] MARCENARO L, FERRARI M, MARCHESOTTI L, et al. Multiple object tracking under heavy occlusions by using Kalman filters based on shape matching[C]// International Conference on Image Processing, 2002:341-344
- [2] ARULAMPALAM M S, MASKELL S, GORDON N, et al. A

tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking[J]. IEEE Transactions on Signal Processing, 2002, 50(2):174-188.

- [3] COMANICIU D, RAMESH V, MEER P. Real-Time tracking of non-rigid objects using mean shift[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2002: 2142.
- [4] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7):1409-1022.
- [5] RUI C, MARTINS P, BATISTA J. Exploiting the Circulant Structure of Tracking-by-Detection with Kernels[C]// Proceedings of European Conference on Computer Vision. ECCV, 2012:702-715.
- [6] BABENKO B, YANG M H, BELONGIE S. Robust object tracking with online multiple instance learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(8):1619-1632.
- [7] HENRIQUES J F, RUI C, MARTINS P, et al. High-speed tracking with kernelized correlation filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3):583-596.
- [8] 张雷,王延杰,刘艳涯,等.基于相关滤波器的视觉目标跟踪方法[J].光电子激光,2015,26(7):1349-1357.
- [9] 喻金平,郑杰,梅宏标.基于改进人工蜂群算法的 K 均值聚类算法[J].计算机应用,2014,34(4):1065-1069.
- [10] 刘梦娇,吴成茂.一种改进的局部模糊 C-均值聚类分割算法研究[J].计算机科学,2015,42(6):190-191.

(上接第 229 页)

- [6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]// Proceedings of the Advances in Neural Information Processing Systems, 2012.
- [7] OORD A V D, KALCHBRENNER N, KAVUKCUOGLU K. Pixel recurrent neural networks [J/OL]. <https://cn.arXiv.org/abs/1601.06759>.
- [8] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks [J/OL]. <https://arxiv.org/abs/1511.06434>.
- [9] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]// Proceedings of the Advances in Neural Information Processing Systems, 2015.
- [10] LIU Z, GAO J, YANG G, et al. Localization and Classification of Paddy Field Pests using a Saliency Map and Deep Convolutional Neural Network [J]. Scientific reports, 2016, 6:20410.
- [11] METZ A R L, RESEARCH I, BOSTON M, et al. Very deep

convolutional networks for large-scale image recognition [J/OL]. [https://www.researchgate.net/publication/294284631\\_Localization\\_and\\_Classification\\_of\\_Paddy\\_Field\\_Pests\\_using\\_a\\_Saliency\\_Map\\_and\\_Deep\\_Convolutional\\_Neural\\_Network](https://www.researchgate.net/publication/294284631_Localization_and_Classification_of_Paddy_Field_Pests_using_a_Saliency_Map_and_Deep_Convolutional_Neural_Network).

- [12] GIRSHICK R. Fast r-cnn[C]// Proceedings of the IEEE International Conference on Computer Vision, 2015.
- [13] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors [J]. Cognitive modeling, 1988, 5(3):1.
- [14] CARPENTER B. Lazy sparse stochastic gradient descent for regularized multinomial logistic regression [J/OL]. <https://www.mendeley.com/research-papers/lazy-sparse-stochastic-gradient-descent-regularized-multinomial-logistic-regression>.
- [15] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP(99):1.
- [16] KOHAVI R. A study of cross-validation and bootstrap for accuracy estimation and model selection[C]// Proceedings of the Ljcai. Stanford, CA, 1995.