

张善文, 邵 彧, 齐国红, 等. 基于多尺度注意力卷积网络的作物害虫检测[J]. 江苏农业学报, 2021, 37(3): 579-588.
doi:10.3969/j.issn.1000-4440.2021.03.005

基于多尺度注意力卷积网络的作物害虫检测

张善文, 邵 彧, 齐国红, 许新华
(郑州西亚斯学院电子信息工程学院, 河南 郑州 451150)

摘要: 田间作物害虫检测是精确防治虫害和减少农药使用量的前提。由于田间害虫种类多, 同种害虫个体间差异大, 田间同一只害虫的大小、颜色、姿态、位置和背景变化多样、无规律, 而且田间背景复杂、对比度低, 使得传统的作物害虫检测方法的性能不高。现有的基于深度学习的作物害虫检测方法需要大量高质量的标注训练样本, 而且训练时间长。在 VGG16 模型的基础上, 本研究提出一种基于多尺度注意力卷积网络(Multi-scale convolutional network with attention, MSCNA)的作物害虫检测方法。在 MSCNA 中, 多尺度结构和注意力模型用于提取多尺度害虫检测特征, 增强对形态较小害虫的检测能力; 在训练过程中引入二阶项残差模块, 减少网络损失和加速网络训练。试验结果表明, 该方法能较好地检测到农田中各种各样、大小不同的害虫, 检测平均准确率为 92.44%。说明该方法能够实现自然场景下作物害虫的精准检测, 可应用于田间作物害虫自动检测。

关键词: 作物害虫检测; 注意力机制; 卷积神经网络; 多尺度注意力卷积网络

中图分类号: TP391.41; S432 **文献标识码:** A **文章编号:** 1000-4440(2021)03-0579-10

Crop pest detection based on multi-scale convolutional network with attention

ZHANG Shan-wen, SHAO Yu, QI Guo-hong, XU Xin-hua

(School of Electronics and Information Engineering, Zhengzhou SIAS University, Zhengzhou 451150, China)

Abstract: Detection of crop pests in field is the prerequisite for accurate pest control and reduction of pesticide dosage. The performance of the traditional detection methods for crop pests is not high, due to the reasons such as various varieties of pests in the field, the difference between different pest individuals of the same variety is great. Besides, the size, color, posture, position and background of the same pest in the field are various and irregular, and the field background is complex and has low contrast. The existing crop pest detection methods based on deep learning require a large number of labeled training samples with high quality, and the training time is long. A multi-scale convolutional network with attention (MSCNA) method based on VGG16 model was proposed for crop pest detection. In MSCNA, the multi-scale structure and attention model were used to extract the detection features of pests on multi-scale and to enhance the ability in detecting smaller pests. Second-order term residual module was introduced in the training process to reduce network loss and accelerate network training. The experimental results showed that, the proposed method could detect various pests with different sizes in the farmland preferably, and the average detection accuracy was 92.44%. The results indicated that this method can detect crop pests accurately in natural scenes and can be applied in the automatic detection of crop pests in the field.

Key words: crop pest detection; attention mechanism; convolutional neural network (CNN); multi-scale convolutional neural network with attention (MSCNA)

收稿日期: 2020-09-19

基金项目: 国家自然科学基金项目(61473237); 河南省科技攻关项目(202102210157, 202102210386, 202102110278)

作者简介: 张善文(1965-), 男, 陕西西安人, 博士, 教授, 研究方向为模式识别及其在作物病虫害检测中的应用。(E-mail) wjd716@163.com

作物害虫检测和识别是害虫防治的一个重要步骤。目前已有很多作物害虫检测和识别方法。Martineau 等^[1]综述了 44 种昆虫分类方法, 并介绍了害虫图像采集、特征提取和测试数据集构建等。Yaakob

等^[2]提取昆虫图像的 shape 特征的 6 种不变矩,并验证了该方法的有效性。Fedor 等^[3]介绍了基于数字图像分析和人工神经网络系统的半自动害虫识别和监测工具,并提取害虫的头部、锁骨、翅膀、产卵器长度和宽度等形态特征,构成特征向量,最后由神经网络进行害虫识别。Wen 等^[4]提出了一种基于尺度不变特征描述子的昆虫分类方法,并在果园常见昆虫数据集上取得了较高的识别率。然后,提出了一种基于图像的昆虫自动分类方法^[5],该方法利用仿射不变量局部特征建立昆虫识别和分类的局部不变特征模型,建立了 54 个全局特征的昆虫识别分类模型和基于局部特征与全局特征的层次组合模型,最后利用组合模型进行田间昆虫图像分类。Boissard 等^[6]提出了一种基于视频信息的害虫检测系统,该系统能够对成熟阶段的白粉虱进行自动检测和计数。Zhu 等^[7]提出了一种基于局部平均颜色特征和支持向量机(SVM)的昆虫自动识别和分类方法,首先对采集的昆虫图像进行分割,去掉图像的背景,再进行分割,得到昆虫的 2 个翅膀图像,然后对它们的位置进行校准,最后从 2 个翅膀图像中提取 1 个 480 维的红绿蓝颜色特征向量,由 SVM 对昆虫进行分类。Fina 等^[8]将 K-均值聚类算法与自适应滤波器相结合,通过提取害虫与其栖息地(叶、茎)之间的不同特征属性,获得不同图像的相关峰值,利用自适应滤波器对作物害虫进行识别,结果表明该滤波器能够识别任意拍摄角度的害虫。Jayme^[9]提出了一种基于传统数字图像处理的大豆叶片白粉虱害虫识别方法,该方法易于实现和扩展应用到其他作物中。为了提高作物害虫分类精度,Xie 等^[10]开发了一个基于多任务稀疏表示和多核学习方法的昆虫识别系统,该系统结合多种昆虫种类特征来提高识别性能,采用稀疏编码直方图表示昆虫图像,而不是使用手工制作的描述符,由此能很好地量化颜色、形状和纹理等原始特征,通过共同优化核权值,有效地优化该学习模型。Zhang 等^[11]设计了 1 种田间害虫识别系统,该系统需要从二值化图像中提取害虫的 7 个形态特征,并进行归一化,根据 9 种害虫的特征均值和特征标准差,建立了标准向量模型库和隶属度函数。蓟马是为害草莓大棚的害虫之一,Ebrahimi 等^[12]采用差分核函数 SVM 方法对蓟马进行分类和检测,以大直径与小直径之比作为区域指标,以色调、饱和度、强化度作为颜色指标设计 SVM 结构,利用均方误差、均方根误差、平均绝对误差和平均百

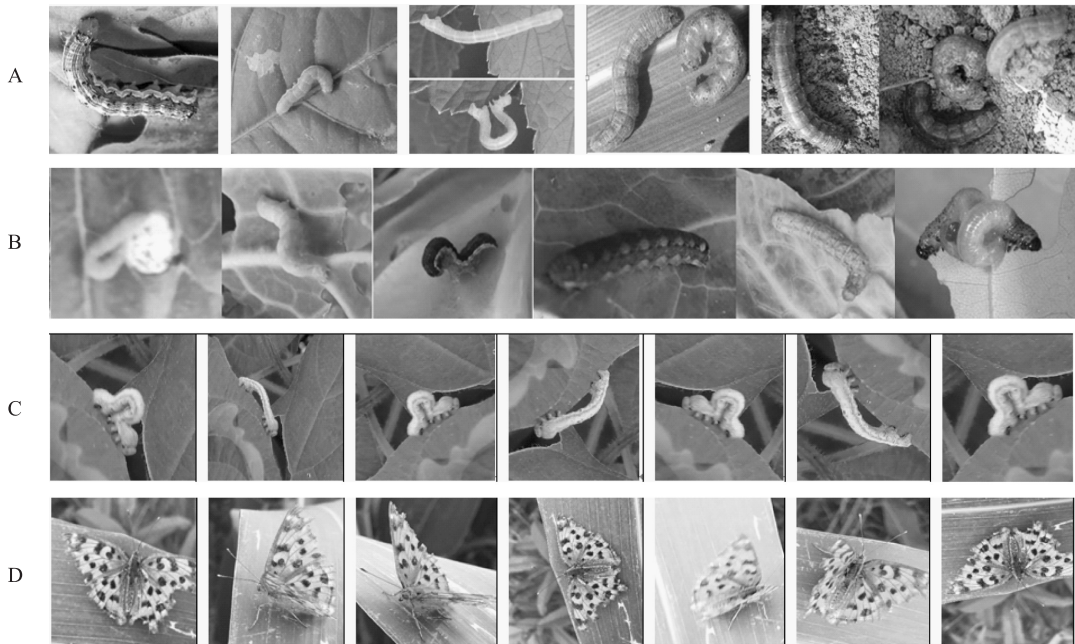
分比误差评估分类效果。为了提高害虫图像分割的准确性和稳定性,Wang 等^[13]提出了一种分割害虫图像的方法:首先对害虫图像进行分块,然后采用自适应学习算法精确选择初始聚类中心,再利用 K-均值聚类得到初步的分割结果,最后利用 3 个数字形态特征分割害虫图像。

上述传统的害虫检测和识别算法研究取得了较好的结果,但这些算法都有一定的局限性,比如他们的检测与识别性能特别依赖于预先设计的手工特征提取效果以及选择的分类器的性能。由于同种或异种作物害虫在不同时期甚至不同时刻具有不同的表型,如图 1 所示,害虫的大小、颜色和形状各异,姿态和位置变化多样,背景复杂,所以利用传统的模式识别方法很难准确描述害虫图像。

卷积神经网络(Convolutional neural network, CNN)具有很强的图像分类特征学习能力,能够从输入图像中学习到深度高阶特征,在复杂图像的目标检测、分割和分类等方面具有明显优势^[14-15],能够自动学习复杂图像的形状、颜色、纹理等多层次特征。农作物害虫图像复杂无规律,因此 CNN 可能适合于作物害虫图像特征提取^[16]。Bhatt 等^[17]将 3 种广泛使用的 CNN 模型应用于昆虫识别,并比较了 3 种模型的性能。Nanni 等^[18]提出了一种显著性模型与 CNN 相结合的自动分类器,其中 3 种显著性模型用于图像预处理,突出显示图像中最相关的像素。Witenberg 等^[19]提出了一种基于改进残差 CNN 模型的害虫识别方法,并在原始害虫图像和增强图像数据集上进行了试验。Xia 等^[20]利用 CNN 模型解决农作物昆虫的多分类问题,采用区域建议网络代替传统的选择性搜索技术,生成较少的建议窗,用于分类特征提取,提高了预测精度和加速了网络训练。Liu 等^[21]提出了一种基于 CNN 的大规模害虫检测分类方法,该方法利用害虫图像的位置敏感得分图代替 CNN 的全连接层,极大提高了 CNN 的分类性能。为了检测自然场景下的水稻害虫,钱蓉等^[22]提出了一种基于权重参数为 16 层的经典网络(VGG16)的水稻害虫智能识别方法,该方法根据水稻害虫的个体特征和自然场景对 VGG16 的卷积层局部调整,优化模型的主要参数,实现了水稻害虫的智能识别。注意力机制推动了深度学习的发展,已成为深度学习的重要组成部分,在自然语言处理、统计学习、图像识别及语音识别等领域得到了充分研

究和成功应用^[23-24]。梁斌等^[25]提出了一种基于多注意力卷积神经网络的特定目标情感分析方法,该网络能够接收平行化输入的文本信息,大大降低网络模型的训练时间,有效弥补仅依赖内容层面注意

力机制的不足。乐毅等^[26]提出了一种分段 CNN 与多层注意力机制相结合的远程监督关系抽取方法,并应用于作物病虫害检测中,取得了较高的识别率。



A:不同姿态、颜色、形状、大小和旋转方式的不同种作物害虫;B:各种姿势、颜色、形状、大小和旋转的卷心菜害虫;C:一只斑粉蝶幼虫不同姿态、颜色、形状、大小和旋转方式的图像;D:一只斑粉蝶成虫在不同姿态、颜色、形状、大小和旋转方式的图像。

图1 作物害虫在大田中的图像

Fig.1 Images of crop pests in the field

在多尺度 CNN 和注意力机制的启发下^[27],本研究提出一种多尺度注意力卷积网络 (Multi-scale convolutional network with attention, MSCNA) 模型,应用于多尺度害虫检测,并验证该模型的有效性。

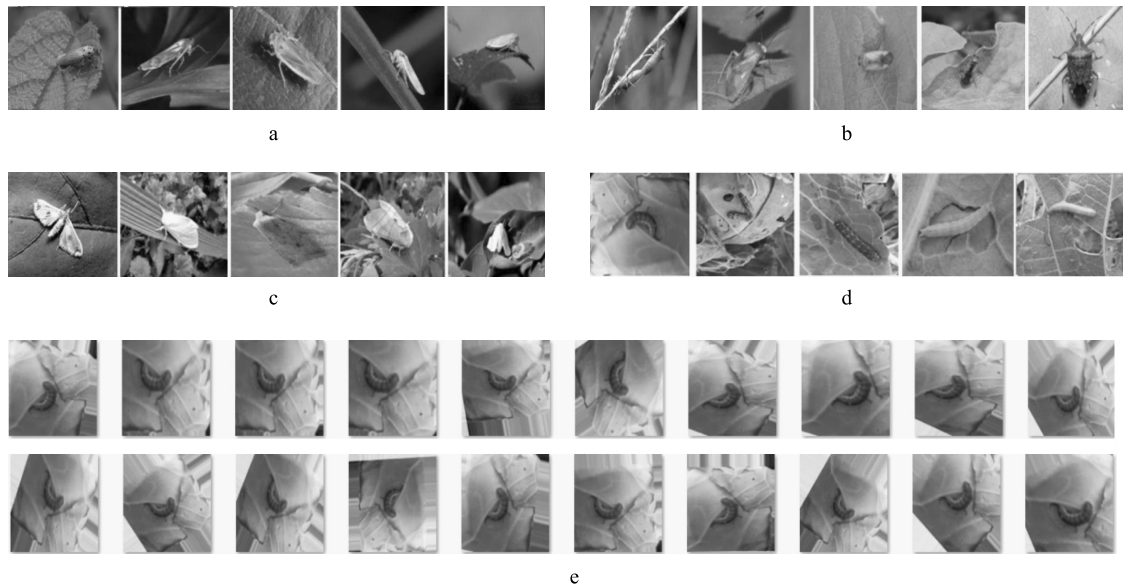
1 材料与方法

1.1 图像采集与扩充

叶蝉、盲蝽、飞蛾和夜蛾幼虫是农业生产中常见的4种害虫,本研究选择这4种害虫的图像为研究对象。所有图像均在大田自然背景下采集,为后期实际推广应用奠定基础。图像采集设备为 iPhone7、华为 P10、WIFI 控制摄像头和物联网等。构建1个原始的害虫图像数据集,包含1 000幅不同大小的害虫图像,每一种害虫 250 幅图像,每幅图像分辨率约为4 928×3 264像素。由于 iPhone 通常能够捕捉更清晰的视频,所以大部分害虫视频图像由 iPhone7 拍摄。部分作物害虫图像如图2所示,从图2可以

看出,同一种害虫甚至同一只害虫的形状、颜色、大小、姿态和所处位置以及背景环境等多种多样,而且在图像中害虫尺寸相对较小。

训练数据集规模对训练网络的性能影响很大。当样本的特征空间维度大于训练样本数目时,模型容易出现过拟合现象。为了增强网络的鲁棒性和泛化能力,一般通过对有限的训练样本进行扩充,增加训练样本数目。常用的扩充方法包括:图像平移,使得网络学习到平移不变的特征;图像旋转,使得网络学习到旋转不变的特征。一般情况下,害虫图像有多种不同的姿态,图像旋转可以弥补样本中姿态较少的问题;图像镜像,与旋转的功能类似;图像加噪声,可以用不同的模板卷积产生含噪图像,使得网络学习到含噪图像的特征;图像亮度变化,使得网络学习到不同亮度条件下的图像的特征;裁剪、缩放等操作,模拟多种实际情况下的样本。



a: 叶蝉; b: 盲蝽; c: 飞蛾; d: 夜蛾幼虫; e: 夜蛾幼虫第一幅图像的 20 幅扩充图像。

图 2 4 种作物害虫的图像

Fig.2 Images of four kinds of crop pests

采用以上扩充方法,可以将每幅图像扩充到 20 幅。图 2E 为 1 幅图像的 20 幅扩充图像。尽管这些害虫图像具有一定误差且图像质量较差,但能够增强网络的鲁棒性和泛化能力。最后得到一个包含 21 000 幅害虫图像的数据集。

1.2 试验方法

1.2.1 卷积神经网络(CNN) 与人工神经网络相比,CNN 具有局部(稀疏)连接和权值(参数)共享的特点,从而大大降低了网络的参数数量。CNN 直接以原始图像作为输入,无需进行复杂的图像预处理。与前馈神经网络的全连接不同,CNN 中卷积层的神经元仅与其相邻层的神经元连接,而不是与全部神经元连接。CNN 的局部连接能够提高网络结构的稳定性和泛化能力,避免过拟合问题,减少权重参数的总量,加速网络训练,且在计算时减少了内存开销^[20]。图 3 为经典的 VGG-16 结构^[22],包括 13 个卷积层、5 个池化层、3 个全连接层和 1 个分类层,具有局部连接、权值共享、池化等特点。其结构比较简单,但它的权重数目很大,导致训练时间过长,计算量很大,而且需要较大的储存空间,不利于部署到算力受限的环境。

1.2.2 注意力机制 在深度 CNN 中,高层次特征图的判别能力强,但目标模糊;低层次特征图的目标定位比较准确,但分类能力较差,虚拟目标信息

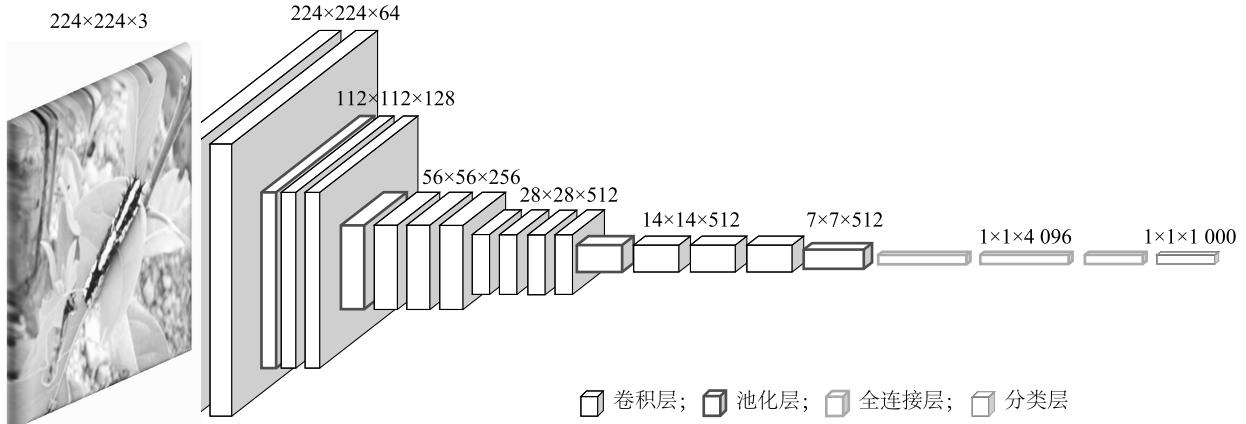
和场景内容较多。为了整合高层次特征图和低层次特征图的优点,在 CNN 中加入注意力机制,生成粗略的注意图^[23]。在 CNN 中,注意力一般可分为通道注意力和空间注意力模块,如图 4 所示,其中通道注意力能够对通道进行选择,表示图像不同的特征信息,空间注意力用于选取图像特征图中所注意的区域。

对于给定的 $F \in R^{H \times W \times C}$,其中 H 、 W 和 C 分别表示特征图的长度、宽度和通道数,经过通道和空间卷积注意力模块后分别得到通道注意力(M_c)的特征图 and 空间注意力(M_s)的特征图:

$$\begin{cases} F' = M_c(F) \otimes F \\ F'' = M_s(F') \otimes F' \end{cases} \quad (1)$$

其中, F 、 F' 和 F'' 分别为输入特征图、通道注意力特征图 and 空间注意力特征图, \otimes 表示注意力图与输入的特征图相乘,用于特征的自适应学习。

通过通道注意力和空间注意力模块可以使 CNN 聚焦于最具有判别性的局部特征图,同时结合残差连接提高网络模型的判别能力。注意力机制可以应用到 CNN 中的原始图像、特征图甚至特征图的每个像素上。将注意力机制应用于图像的空间尺度,可以对不同的空间区域赋予权重;将其应用于通道尺度,可以对不同的通道特征赋予权重。



图中数字表示输入图像或中间隐层节点得到的特征图的尺寸大小。

图3 VGG16 结构

Fig.3 Architecture of VGG16

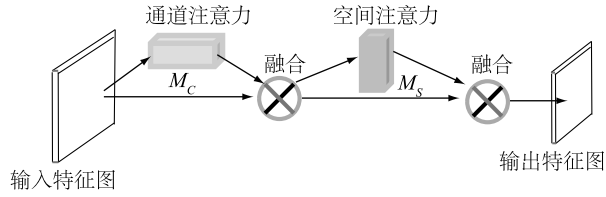


图4 CNN 中的注意力模块

Fig.4 Attention module in CNN

1.2.3 多尺度注意力卷积网络 为了提高基于VGG16的作物害虫检测方法的识别率,在分层多尺度卷积特征提取方法^[27]的启发下,在VGG16中引入通道和空间注意力机制以及残差模型,构建一种多尺度注意力卷积网络(MSCNA)模型,其结构如图5所示。

多尺度注意力卷积网络模型的训练具体步骤如下:

(1)分别使用 3×3 、 5×5 和 7×7 大小的3个卷积核对原始害虫图像进行卷积,提取不同尺度下的特征图,然后利用文献[28]中的整合方法 $F_1 = \max[0, \text{concat}(F_{11}, F_{12}, F_{13})]$ (其中 F_{11} 、 F_{12} 、 F_{13} 分别为3个不同卷积核得到的卷积图,concat为连接运算)合并三通道特征图,再进行卷积、池化、卷积。

(2)在注意力机制模块中,沿通道维度分别计算每幅特征图的全局平均池化及全局最大池化操作,然后相加得到通道注意力参数和特征图(F),再经过通道注意力(M_c)和空间注意力(M_s),由公式(1)得到特征图 F' 和 F'' ;

(3)在注意力机制模块前后的第3和第4个卷积层之间,引入残差连接模块,能够融合多尺度特征

和注意力机制模块的输入信息,从而在一定程度上对损失的信息进行补充,解决了特征信息传输过程中的信息丢失现象,进一步增强特征融合。

(4)第4个卷积层后的结构与VGG16基本相同^[22,29],不同之处是将VGG16的3个全连接层用1个全局池化层代替,以极大地减少网络参数。

在MSCNA的注意力模块中,平均池化操作可以滤掉目标对象的全局背景信息,而最大池化可以突出目标对象的显著性。设输入的特征图为 $X = [x_1, x_2, \dots, x_c]$,其中 x_c 为第 c 个灰度值,则平均池化和最大池化操作的输出分别为 $\text{Avgpool}(F) \in \mathbb{R}^{1 \times 1 \times C}$ 和 $\text{Maxpool}(F) \in \mathbb{R}^{1 \times 1 \times C}$,计算如下:

$$\text{Avgpool}(F) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (2)$$

$$\text{Maxpool}(F) = \max \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (3)$$

在公式(1)中使用1个共享的多层感知机(MLP)进行注意力推断来保存参数,则通道注意力(M_c)和空间注意力(M_s)分别表示为:

$$M_c(F) = \sigma \{ \text{MLP}[\text{Avgpool}(F)] + \text{MLP}[\text{Maxpool}(F)] \} \quad (4)$$

$$M_s(F) = \sigma \{ f^{7 \times 7}[\text{Avgpool}(F); \text{Maxpool}(F)] \} \quad (5)$$

其中, $\text{Avgpool}(F)$ 和 $\text{Maxpool}(F)$ 为在空间维度上使用的平均池化和最大池化操作, $f^{7 \times 7}$ 为 7×7 大小的卷积操作, σ 为Sigmoid函数,网络模型中的参数都通过Sigmoid来进行归一化。

第3和第4个卷积层之间的残差模块中,将第3个卷积层的特征发送到第4个卷积层,能够最大

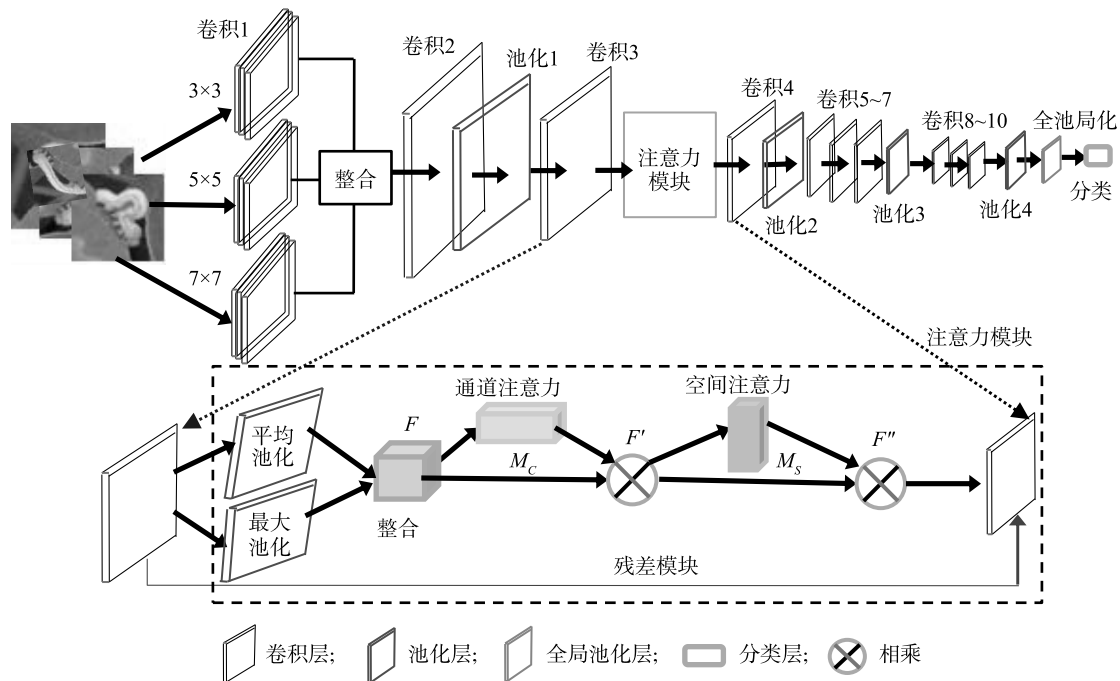
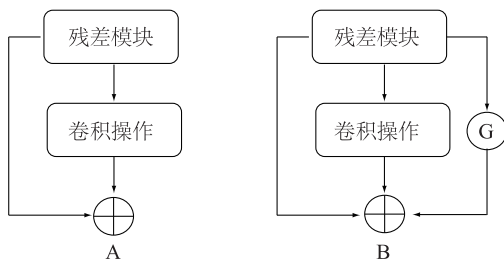


图 5 多尺度注意力卷积网络模型结构

Fig.5 Architecture of multi-scale convolutional network with attention (MSCNA) model

程度地保留浅层的全局特征,在训练时有助于保持反向传播的梯度稳定性。传统残差网络(Residual networks, ResNet)的连接方式是一种线性连接,限制了网络对特征空间复杂分布的学习能力。MSCNA 的特征提取操作可以看作一个深度全卷积网络(Fully convolutional networks, FCN),FCN 利用大量的残差连接引导浅层特征的传递和梯度的流动,考虑到注意力模块产生的特征信息更有意义,在残差连接上增加二阶项和偏置项,提高网络的非线性程度,改进的二阶项连接残差网络如图 6 所示^[30-31]。



A: 传统的残差连接; B: 二阶项残差。

图 6 残差连接结构

Fig.6 Architecture of residual connection

假设第 3 个卷积层得到的特征图为 x , 第 4 个卷积层得到的特征图为 x' , 则输出为:

$$f = X + X' + G'' \quad (6)$$

其中, $G'' = g(X \odot X' + \xi)$, \odot 为矩阵间元素的乘法, ξ 为调优偏移参数, g 为可微函数。

采用平方根形式对 g 进行梯度推导, 反向更新传播参数。改进残差连接计算如下:

$$f = X + X' + \sqrt{\text{ReLU}(X) \odot \text{ReLU}(X') + \xi} \quad (7)$$

其中, g 取偏移量 $\xi = 0.0001$ 用以保持反向传播过程中的梯度稳定性。

对经过 2 种注意力和残差重组后的特征图进行元素归一化后再相加, 其结果表示为:

$$F''' = F'' + f \quad (8)$$

由于在矩阵和特征矩阵中的元素可能为负, 计算前需要进行非负处理。为了有效地融合多层卷积特征, 利用双线性插值将所有的特征映射调整为固定大小。本研究将特征图的大小固定为 224×224 。基于 MSCNA 的田间害虫检测方法的检测流程图见图 7。

具体步骤如下: (1) 采用不同尺度的滑动窗口扫描并截取图像, 找到待检测图像中可能存在害虫的候选窗口; (2) 将候选窗口中的图像归一化为 32×32 像素大小; (3) 将归一化后的图像作为训练后 MSCNA 的输入, 计算网络输出; (4) 删除步骤 (3) 中判定为非害虫的候选窗口, 剩余的候选窗口为图像中害虫的位置信息。

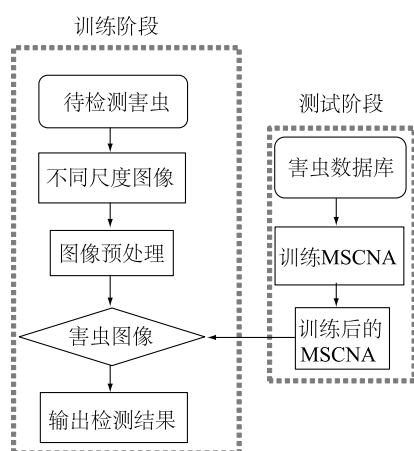


图7 作物害虫检测流程图

Fig.7 Flowchart of crop pest detecting

2 结果与分析

在实际害虫图像数据库上验证基于 MSCNA 的作物害虫图像的检测方法,并将基于局部特征和神经网络 (Local features and neural network, LFNN)^[3]、基于局部不变特征和支持向量机 (Local invariant features and supported vector machines, LFS-VM)^[4]、基于深度学习模型的玉米害虫检测识别 (Deep learning, DL)^[19]、基于改进 CNN 的作物害虫检测与分类方法 (Improved CNN, ICNN)^[20] 和基于 VGG16 的害虫检测 (VGG16)^[22] 5 种害虫图像识别方法与其进行比较。其中,LFNN 和 LFSVM 为 2 种传统的害虫图像识别方法,DL、ICNN 和 VGG16 为 3 种基于深度学习的害虫检测方法。所有试验实现环境为操作系统 Win10 64bit、处理器 Intel Xeon E5-2643v3@3.40 GHz CPU、内存 64 GB、NVidia Quadro M4000 GPU、显存 8 GB、CUDA Toolkit 9.0、CUDNN V7.0、Python 版本 3.6.4、Tensorflow-GPU 1.8.0 框架。

在 Matlab 平台上完成自动裁切工作,得到裁切后的图像。将所有图像集按照 6:4 随机分成训练子集和测试子集,其中 12 600 幅图像用于训练模型,8 400 幅图像用于测试模型。采用 LabelImg 工具标注训练图像,用于深度学习训练。6:4 划分试验重复 100 次,得到稳定可靠的检测平均结果。在试验中,分多批进行网络训练,每批使用 35 幅图像,批处理设置为 1,最大迭代次数为 10 000,权重衰减指数设置为 0.000 1,动量参数设置为 0.9,初始学习率设置为 0.01,在 1 000 次、2 000 次、4 000 次、6 000 次、

8 000 次迭代时,权重指数衰减降低到 0.000 1。引入权值衰减,使权值衰减到一个较小的值,从而缓解过拟合问题。在对比试验中,将参数设置为合适的值。在测试时,采用 5 折交叉验证方法将样本集随机分为 5 等份,轮流选取其中 1 份组成测试集,剩余的组成训练集,重复 5 次交叉验证,将 5 次交叉验证结果的平均值作为 MSCNA 最终的识别率。

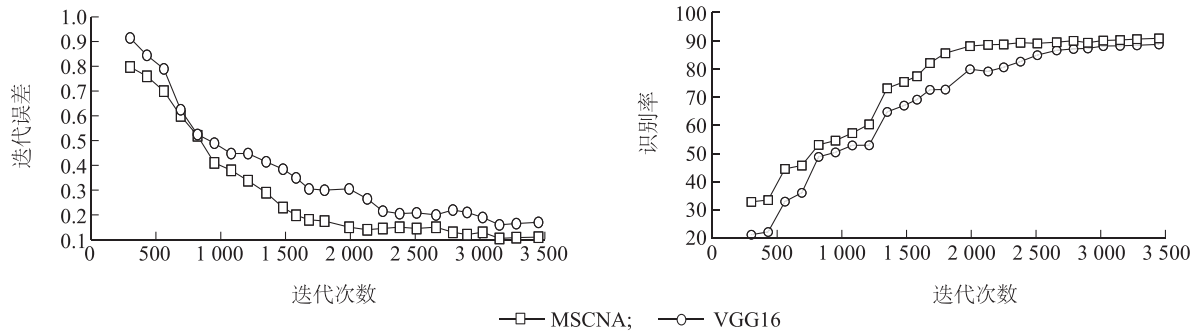
为了分析训练性能,采用相同的试验参数对 3 个用于比较的网络模型进行训练,以保证比较结果的可靠性。图 8 为 MSCNA 和 VGG16 的迭代误差和识别率与迭代次数的关系。由图 8A 可以看出, MSCNA 的结果优于 VGG16,在 1 500 次迭代前,其训练过程相对稳定,并且在 1 500 次迭代后, VGG16 的下降速度比 MSCNA 慢很多。由图 8B 可以看出,在整体上 MSCNA 的识别率高于 VGG16。结果表明, MSCNA 的收敛性优于 VGG16,在 2 500 次迭代后, MSCNA 的训练性能有了很大提高。主要原因是 MSCNA 中利用了残差的密集模块和注意力机制。

在训练过程中,将害虫训练图像输入到 MSCNA 中,通过几个卷积层和池化层逐步提取特征图。为了测试 MSCNA 的检测性能,从数据集中随机选取 1 幅夜蛾幼虫图像,将检测结果可视化 (图 9)。通过图 9 中的可视化特征图可以发现, MSCNA 能够捕捉到害虫图像的精细节,较浅的卷积核可提取出害虫的轮廓,较深的卷积核可提取出害虫的纹理和细节特征。从图 9E 中可以发现,卷积特征图没有明显的锐化边缘,并且逐渐褪色,这是因为模型中引入注意力机制,更多关注害虫图像区域,而不是害虫图像的边缘。从图 9F 可以看出,第一幅特征图靠近输出层,害虫图像区域更加集中。

由图 9 可以看出,害虫图像的低阶卷积特征包含更多的细节信息,高层卷积特征包含害虫的关键信息。第一层的卷积核比较大,包含稀疏的正负响应的平面,从而增加包含适当特征的机会;第二层的卷积核比较小,容易训练,梯度较低,在容量方面也比较理想。进一步分析得出,在浅层卷积层中大量特征图受到背景特征影响,只有少部分能得到害虫特征图。随着层次的深入,卷积特征图忽略不相关的背景,突出更抽象的高维分类特征。由图 9 还可以观察到,底层特征包含一些非害虫信息,低层卷积特征的背景信息分布较为分散,这就导致了低层卷积层很难直接去除非害虫信息。因此,为了有效地将低层卷积特征与高

层卷积特征连接起来,获得增强鉴别能力的卷积特征,我们采用注意力机制选择害虫的特征,去除低层卷积特征的非害虫信息。表 1 为 LFNN^[3]、LFS-

VM^[4]、DL^[19]、ICNN^[20]、VGG16^[22] 和本研究提出的 MSCNA 模型对 4 种害虫图像的检测结果。图 10 为 6 种方法对 4 种害虫的检测准确率。



A: 迭代误差与迭代次数的关系; B: 检测精度与迭代次数的关系。

图 8 基于 MSCNA 检测方法和基于 VGG16 检测方法的迭代误差和检测精度与迭代次数的关系

Fig.8 Relationship between iterative error, detection accuracy and iterative times based on detection methods of multi-scale convolutional network with attention (MSCNA) and VGG16

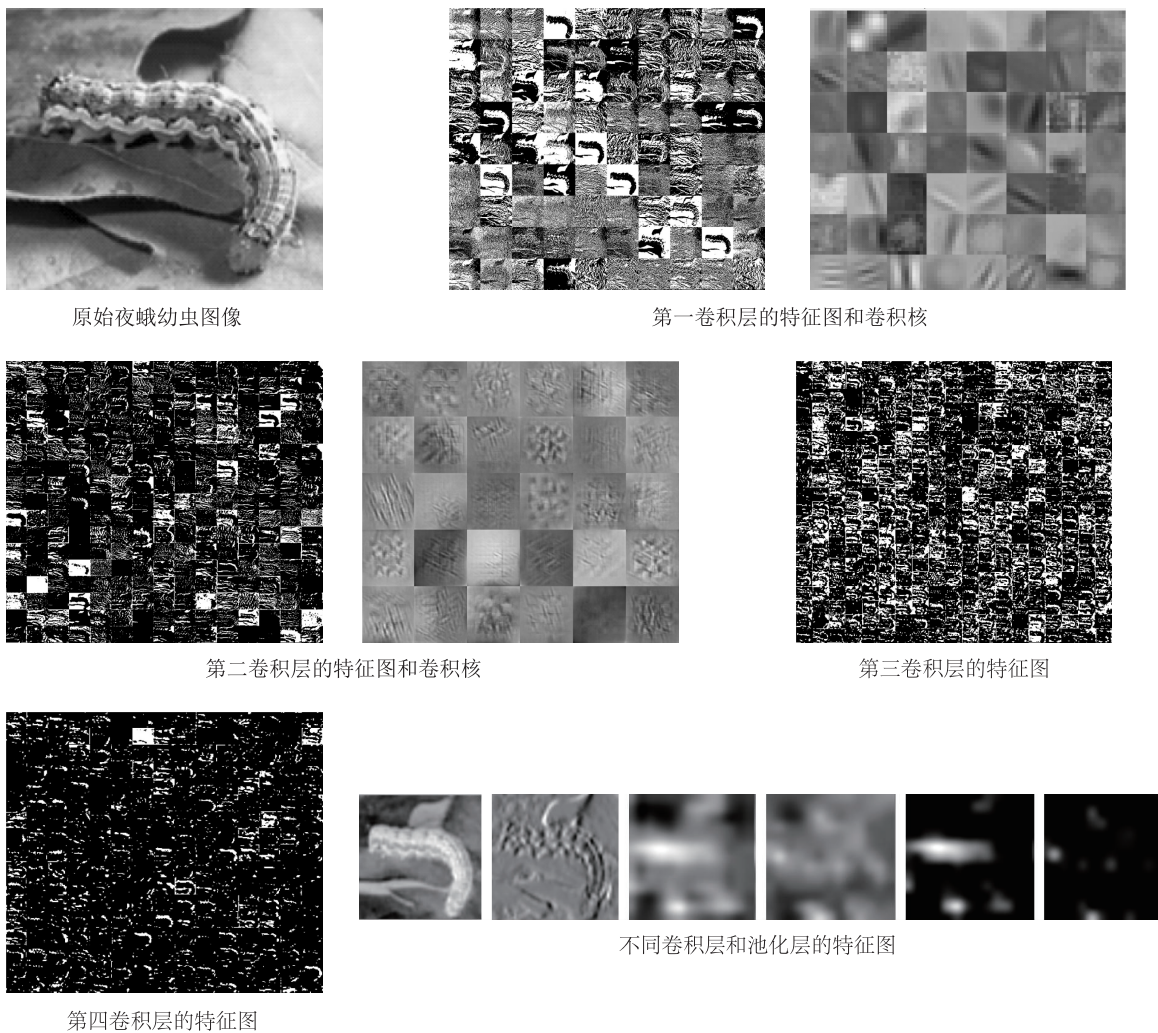


图 9 不同卷积层的特征图和卷积核

Fig.9 Feature maps and convolutional kernels of different convolutional layers

表 1 6 种方法对 4 种害虫的检测结果

Table 1 Detection results of four pest species by six methods

项目	检测方法					
	LFNN	LFSVM	DL	ICNN	VGG16	MSCNA
精度 (%)	56.25	61.76	86.34	88.15	89.37	92.44
训练时间(h)	0.84	0.47	17.36	15.74	21.16	9.55
测试时间(s)	0.65	0.38	0.72	0.49	0.37	0.28

LFNN:基于局部特征和神经网络的检测方法;LFSVM:基于局部不变特征和支持向量机的检测方法;DL:基于深度学习模型的玉米害虫检测识别方法;ICNN:基于改进 CNN 的作物害虫检测与分类方法;VGG16:基于权重参数为 16 的经典网络的检测方法;MSCNA:基于多尺度注意力卷积网络检测方法。

从表 1 看出,基于 DL、ICNN、VGG16 和本研究模型(MSCNA)的 4 种深度学习方法的检测精度远大于 LFNN 和 LFSVM 2 种传统方法,其原因是深度学习能够从复杂图像中自动学习图像的分类特征,而传统方法很难提取到害虫的鲁棒特征,因为在田间即使是同一只害虫每时每刻也变化多样。一般传统方法都需要对害虫图像进行预处理,再进行害虫图像分割,表 1 的 6 种方法都采用原始图像及其扩展图像进行试验,由于 2 种传统方法不能提取有效的分类特征,所以他们的识别率很低。4 种基于深度学习的方法的训练时间远长于 2 种传统方法,其原因是,与 LFNN 和 LFSVM 相比,深度学习需要花大量时间训练模型中的参数。对于训练好的 LFNN、LFSVM 和 4 种深度学习模型,测试时间差异不大。本研究方法(基于 MSCNA 检测方法)优于其他方法,检测率最高,检测时间最少。

综合来看,基于 MSCNA 方法在少量数据的基础上进行有监督学习,避免了复杂的图像处理过程,在准确率、训练和测试时间方面相比其他方法较好,说明 MSCNA 具有很好的分类性能。主要原因是多尺度和注意力机制引入到 MSCNA,在特征提取阶段,将卷积后的加权特征向量替换原来的特征向量用于残差融合,再利用二阶项在训练过程中减少信息损失和加快网络训练。

3 讨论

及时、准确地检测到作物害虫,是作物害虫精准防治的前提^[32-33]。由于大田背景复杂,大田害虫种类多,同一只害虫的姿态和形态多种多样,所以作物害虫检测是一项重要且具有挑战性的研究^[34-35]。本研究将注意力机制、残差模型与 VGG16 相结合,提出了一种 MSCNA 模型,并应用于多尺度害虫检测中。该模型在特征提取过程中引入注意力机制,

然后用滤波后的加权特征向量代替原始特征向量,再进行残差融合,最后在训练过程中引入二阶残差项,由此减少信息损失,加快网络训练速度。在害虫图像数据集上的试验结果表明,MSCNA 是一种有效的害虫检测模型,其检测精度达到 90% 以上。试验中的害虫图像均在大田条件下采集,而不是在理想条件下采集,这使得提出的模型具有较强的抗干扰能力。基于 MSCNA 的害虫识别方法可以在图像中定位害虫,而其他方法主要实现害虫图像分类。所以,本研究方法能有效地缓解数据处理中人为因素的干扰和人为负担。该方法可应用于作物害虫自动检测与识别。未来的工作重点主要有:害虫数据库不足以反映所提方法的可行性,需要扩充,后期利用物联网自动采集害虫图像,扩大数据集。在害虫防治过程中,害虫分类要求更加细致,害虫的生长周期需要划分,需要根据害虫生长的不同时期识别害虫类别,由此采取不同的防治措施。

参考文献:

- [1] MARTINEAU M, CONTE D, RAVEAUX R, et al. A survey on image-based insect classification [J]. Pattern Recognition, 2017, 65:273-284.
- [2] YAAKOB S N, JAIN L. An insect classification analysis based on shape features using quality threshold ARTMAP and moment invariant [J]. Applied Intelligence, 2012, 37(1):12-30.
- [3] FEDOR P, JAROMIR V, HAVEL J, et al. Artificial intelligence in pest insect monitoring [J]. Systematic Entomology, 2009, 34(2):398-400.
- [4] WEN C, GUYER D E, LI W. Local feature-based identification and classification for orchard insects [J]. Biosystems Engineering, 2009, 104(3):299-307.
- [5] WEN C, GUYER D. Image-based orchard insect automated identification and classification method [J]. Computers & Electronics in Agriculture, 2012, 89:110-115.
- [6] BOISSARD P, MARTIN V, MOISAN S. A cognitive vision ap-

- proach to early pest detection in greenhouse crops[J]. Computers & Electronics in Agriculture, 2008, 62(2):81-93.
- [7] ZHU L Q, ZHANG Z. Automatic insect classification based on local mean colour feature and supported vector machines [J]. Oriental Insects, 2012, 46(3/4):260-269.
- [8] FINA F, BIRCH P, YOUNG R, et al. Automatic plant pest detection and recognition using *k*-means clustering algorithm and correspondence filters [J]. International Journal of advanced Biotechnology & Research, 2013, 4:189-199.
- [9] JAYME G A. Using digital image processing for counting whiteflies on soybean leaves [J]. Journal of Asia Pacific Entomology, 2014, 17(4):685-694.
- [10] XIE C, ZHANG J, LI R, et al. Automatic classification for field crop insects via multiple-task sparse representation and multiple-kernel learning [J]. Computers & Electronics in Agriculture, 2015, 119:123-132.
- [11] ZHANG H T, HU Y X, ZHANG H Y. Extraction and classifier design for image recognition of insect pests on field crops [J]. Advanced Materials Research, 2013(756/759):4063-4067.
- [12] EBRAHIMI M A, KHOSHAGHAZA M H, MINAEI S, et al. Vision-based pest detection based on SVM classification method [J]. Computers & Electronics in Agriculture, 2017, 137:52-58.
- [13] WANG Z B, WANG K Y, LIU Z Q, et al. A cognitive vision method for insect pest image segmentation [J]. IFAC-Papers On Line, 2018, 15(17):85-89.
- [14] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6):1229-1251.
- [15] BERNAL J, KUSHIBAR K, ASFAW D S, et al. Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review[J]. Artificial Intelligence in Medicine, 2019, 95:64-81.
- [16] TRKOLU M, HANBAY D. Plant disease and pest detection using deep learning-based features [J]. Turkish Journal of Electrical Engineering and Computer, 2019, 27(3):1636-1651.
- [17] BHATT N, PATEL D. Insect identification among deep learning's meta-architectures using tensorflow [J]. International Journal of Engineering and Advanced Technology, 2019, 9(1):1910-1914.
- [18] NANNI L, MAGUOLO G, PANCINO F. Insect pest image detection and recognition based on bio-inspired methods [J]. Ecological Informatics, 2020, 57:101089.
- [19] WITENBERG S R, ADAO N A, D'IBIO L B. A deep learning model for recognition of pest insects in maize plantations [C] // FANTI M P, ZHOU M C. IEEE International Conference on Systems, Man and Cybernetics (SMC). Bari Italy: IEEE Press, 2019.
- [20] XIA D, CHEN P, WANG B. Insect detection and classification based on an improved convolutional neural network [J]. Sensors, 2018, 18(12):4169.
- [21] LIU L, WANG R, XIE C, et al. Pestnet: an end-to-end deep learning approach for large-scale multi-class pest detection and classification [J]. IEEE Access, 2019, 7:45301-45312.
- [22] 钱蓉, 孔娟娟, 朱静波, 等. 基于 VGG-16 卷积神经网络的水稻害虫智能识别研究[J]. 安徽农业科学, 2020, 48(5):235-238.
- [23] 徐诚极, 王晓峰, 杨亚东. Attention-YOLO: 引入注意力机制的 YOLO 检测算法 [J]. 计算机工程与应用, 2019, 55(6):19-29, 131.
- [24] BARROS P, PARISI G I, WEBER C, et al. Emotion-modulated attention improves expression recognition: a deep learning model [J]. Neurocomputing, 2017, 253(30):104-114.
- [25] 梁斌, 刘全, 徐进. 基于多注意力卷积神经网络的特定目标情感分析[J]. 计算机研究与发展, 2017, 54(8):1724-1735.
- [26] 乐毅, 王文宇, 张凯, 等. 基于多层注意力机制的农业病虫害远程监督关系抽取研究[J]. 安徽农业大学学报, 2020, 47(4):189-193.
- [27] 孙皓泽, 常天庆, 王全东, 等. 一种基于分层多尺度卷积特征提取的坦克装甲目标图像检测方法[J]. 兵工学报, 2017(9):1681-1691.
- [28] MUSTAFA H T, YANG J, ZAREAPOOR M. Multi-scale convolutional neural network for multi-focus image fusion [J]. Image and Vision Computing, 2019, 85(5):26-35.
- [29] LIU Z, WU J, FU L, et al. Improved kiwifruit detection using pre-trained VGG16 with RGB and NIR information fusion [J]. IEEE Access, 2020, 8(1):2327-2336.
- [30] WANG X, GU Y, GAO X, et al. Dual residual attention module network for single image super resolution [J]. Neurocomputing, 2019, 364:269-279.
- [31] LIU Z, HUANG J, ZHU C, et al. Residual attention network using multi-channel dense connections for image super-resolution [J]. Applied Intelligence, 2020(1):1-15.
- [32] 陶震宇, 孙素芬, 罗长寿. 基于 Faster-RCNN 的花生害虫图像识别研究[J]. 江苏农业科学, 2019, 47(12):247-250.
- [33] 邢鲲, 曹俊宇, 王媛媛, 等. 设施蔬菜昆虫群落结构与时序动态[J]. 江苏农业学报, 2019, 35(3):564-574.
- [34] 梁勇, 赵健, 林营志, 等. 基于红外传感器的实蝇类害虫实时监测装置的设计[J]. 江苏农业科学, 2020, 48(4):230-234.
- [35] 马林, 林金盛, 陆娜, 等. 江浙地区秀珍菇双翅目害虫鉴定及防治[J]. 南方农业学报, 2019, 50(1):68-73.

(责任编辑:陈海霞)