

# 控制与决策

Control and Decision

## 基于神经网络的多类别目标识别

赵静, 王弦, 王奔, 蒋国平, 谢非, 徐丰羽

引用本文:

赵静, 王弦, 王奔, 等. 基于神经网络的多类别目标识别[J]. 控制与决策, 2020, 35(8): 2037–2041.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.0577>

---

## 您可能感兴趣的其他文章

Articles you may be interested in

### 基于深度卷积网络的电熔镁炉欠烧工况在线识别

Online detection of semi-molten of fused magnesium furnace based on deep convolutional neural network

控制与决策. 2019, 34(7): 1537–1544 <https://doi.org/10.13195/j.kzyjc.2017.1646>

### 基于DropOut降噪自编码的磨矿系统故障诊断

DropOut denoising autoencoder-based fault diagnosis for grinding system

控制与决策. 2018, 33(9): 1662–1666 <https://doi.org/10.13195/j.kzyjc.2017.0597>

### 脉冲神经网络:模型、学习算法与应用

Spiking neural networks: Model, learning algorithms and applications

控制与决策. 2018, 33(5): 923–937 <https://doi.org/10.13195/j.kzyjc.2017.1444>

### 基于反卷积特征提取的深度卷积神经网络学习

Deep convolution neural network learning based on deconvolution feature extraction

控制与决策. 2018, 33(3): 447–454 <https://doi.org/10.13195/j.kzyjc.2017.0048>

### 基于火焰图像显著区域特征学习与分类器融合的回转窑烧结工况识别

Burning state recognition for rotary kiln sintering process based on burning salient zone image feature learning and classifiers fusion

控制与决策. 2017, 32(1): 187–192 <https://doi.org/10.13195/j.kzyjc.2015.1237>

## 基于神经网络的多类别目标识别

赵静<sup>1,2†</sup>, 王弦<sup>1</sup>, 王奔<sup>1</sup>, 蒋国平<sup>1,2</sup>, 谢非<sup>3</sup>, 徐丰羽<sup>1,2</sup>

(1. 南京邮电大学自动化学院、人工智能学院, 南京 210023; 2. 江苏省物联网智能机器人工程实验室, 南京 210023; 3. 南京师范大学电气与自动化工程学院, 南京 210023)

**摘要:** 随着智能化时代的到来,深度学习在图像处理领域的应用越来越广泛,为复杂的图像处理带来了新的解决方法。但是,深度学习在带来高准确推理结果的同时,往往会造成运算量、推理时间以及处理器内存的增加,受限于数据的缺失以及对高性能处理器的依赖。目前市场上有关深度学习技术的相关产品还未被广泛的应用。基于上述问题,提出一种具有广泛应用前景的基于神经网络的多类别目标识别方案,并在国产龙芯派平台下进行测试验证。首先,获取待处理图像数据集,进而在计算机平台下搭建并训练神经网络模型,利用得到的训练参数在龙芯派平台下建立优化后的神经网络结构,并对目标图像进行处理及识别,最后在用户界面显示目标图像所属类别。本系统利用深度学习能够自动从大数据中学习特征的优势,在龙芯派平台下实现上千类生活中常见对象的自动分类,识别准确率高达 96% 以上,识别速度在 3 s 以内,且该方案具有优秀的可扩展性。

**关键词:** 龙芯派; 图像识别; 深度学习; AlexNet 网络模型; 自动分类

中图分类号: TP273

文献标志码: A

## Multi-category target recognition based on neural network

ZHAO Jing<sup>1,2†</sup>, WANG Xian<sup>1</sup>, WANG Ben<sup>1</sup>, JIANG Guo-ping<sup>1,2</sup>, XIE Fei<sup>3</sup>, XU Feng-yu<sup>1,2</sup>

(1. College of Automation & College of Artificial Intelligence, Nanjing University of Posts and Telecommunication, Nanjing 210023, China; 2. Jiang Su Engineering Laboratory for Internet of Things and Intelligent Robotics, Nanjing 210023, China; 3. College of Electrical and Automation Engineering, Nanjing Normal University, Nanjing 210023, China)

**Abstract:** With the coming of intelligent era, deep learning algorithm has been applied more and more widely in the field of image processing, which brings new solutions to complex image processing. However, while deep learning can produce highly accurate inference results, it often leads to computational complexity, the increase of inference time and processor memory. Due to the lack of data and reliance on high-performance processors, deep learning related products have not been widely used in real life scenarios. Based on the above questions, this paper proposes a multi-category image recognition scheme based on deep learning with broad application prospects and carries out test verification under the Loongson platform. Firstly, the image data set of the category to be identified is obtained, and then the neural network model is built and trained under the computer platform. The optimized neural network structure is established under the Loongson platform using the obtained training parameters, and then the target image is processed and recognized. Finally, the category of the target image is displayed in the user interface. The system utilizes the advantages of deep learning to automatically learn big data features and realizes automatic classification of thousands of common objects in life under the Loongson platform. The recognition accuracy can reach more than 96%, and the recognition time is limited to 3s or less. The scheme also demonstrates excellent scalability.

**Keywords:** Loongson Pi; image recognition; deep learning; AlexNet; automatic classification

## 0 引言

卷积神经网络是神经网络算法中的一种,作为深度学习重要模型之一的深度卷积神经网络(deep convolutional neural network, DCNN)在图像分类<sup>[1]</sup>、

目标检索<sup>[2]</sup>、行为识别<sup>[3]</sup>和其他视觉任务<sup>[4]</sup>等领域中均得到了成功应用<sup>[5]</sup>。卷积神经网络的优点是能够直接与图像像素进行卷积,从图像像素中提取图像特征,这种处理方式更加接近人类大脑视觉系统的处理

收稿日期: 2019-04-30; 修回日期: 2019-08-14.

基金项目: 国家自然科学基金项目(61601228); 江苏省自然科学基金项目(BK20161021); 江苏省六大人才高峰项目(JY-081).

责任编辑: 薛建儒.

†通讯作者. E-mail: zhaojing@njupt.edu.cn.

方式. 在图像分类任务中,采用深层结构的卷积神经网络(convolutional neural network, CNN)的效果大大超越了传统方法<sup>[6]</sup>.

如今,卷积神经网络算法在学术界和工业界越来越受到重视,尤其是在计算机视觉方面,卷积神经网络算法可以直接对原始输入图像做处理,不需要对图像先做预处理,跳过繁杂的特征提取过程,因此被更多人所接受和采纳. 迄今为止,卷积神经网络已经在众多领域上得到应用,第一个成功的案例是将卷积神经网络算法应用到手写数字识别上,该算法在此问题中的分类精度高达99.77%,远远领先其他传统的图像分类方法和人工分类. 此外,卷积神经网络还在航空安检、行人跟踪、物体识别等其他问题中都得到了很好的分类结果. 在当今“大数据时代”,神经网络具有强大的特征提取与抽象能力,能够整合多源信息、处理异构数据、捕捉变化动态,是大数据实现价值转化的桥梁. 图像也是信息的重要来源之一,包括机器视觉、医学医疗、人工智能等领域都离不开图像处理技术. 而在日常生活中,图像识别技术的应用也极大地便利了人们的生活. 2018年,张雪芹等<sup>[7]</sup>基于卷积神经网络中的AlexNet网络模型,提出一套植物识别算法,有效地解决了植物图像识别在种类数量、准确度和速度上的难点<sup>[8]</sup>. 目前,计算机平台下图像识别的方法与应用研究取得了丰硕的成果<sup>[8-13]</sup>,但是基于国产芯片以及嵌入式平台的图像识别应用的研究成果却很少. 龙芯作为自主研发的国产芯片,目前的生态亟待完善,在图像处理领域相关的应用也是空白. 传统的图像识别算法仅在识别简单图像时比较有效,且对环境变化的适应性差,相比之下神经网络在识别此类对象时具有明显优势. 此外,采用多层神经网络的深度学习方法具有较大的计算量,使用高性能处理器的成本是难以接受的,而使用嵌入式设备和低性能处理器又会导致计算时间的增加,无法解决实际问题.

本文旨在以2K 1000龙芯派为验证平台,提出一种图像识别应用的开发方案,主要创新性在于其应用于龙芯以及嵌入式平台的可行性和适用性. 基于此原因,本文提出以AlexNet算法为核心的实施方案,开发一款可完成多类别图片识别与分类的应用程序,且设计便于用户使用的人机交互界面,在龙芯派平台下处理多类别复杂图像,能够保证良好的识别率和识别速度.

## 1 神经网络架构及系统搭建

### 1.1 搭建方案

本文设计以典型的深度学习模型AlexNet网络模型为主体,利用了其处理图像分类问题的优势. 该网络的特点是创建对局部神经元的活动竞争机制,使其中响应比较大的值变得相对更大,并抑制其他反馈较小的神经元,增强了模型的泛化能力,有助于快速收敛. 采用Dropout策略(即在网络循环中,按一定概率随机选择神经层的一些单元将其隐藏,不参加前向传播和反向传播),有效减少了过拟合. 此外,考虑到GPU不可用的情况,通过优化网络结构和参数,降低了运算量,以便减轻处理器的运算负担. 图像识别程序使用C++进行编写,调用了Opencv的相关库文件,分别用于图像的读取、处理以及神经网络框架的建立,并且构建了图形界面以模拟应对实际需求场景. 为了在Qt下编译的程序可以在不同平台下直接运行,而不用过于依赖运行环境(Opencv, Qt等境的配置),本设计利用动态链接库方法,将程序运行所用到的库文件进行打包,并建立对应库的连接,即可在不同平台下运行该程序. 本文的设计以识别的快速性、准确性以及图片涉及对象的多样性为目标,能够排除干扰、识别不同品质的图片,满足不同人群对相关数据的需求.

### 1.2 系统整体架构

系统运行的主要流程包括3个部分: 1) AlexNet神经网络模型的搭建和参数的训练,此部分主要使用Opencv的DNN模块,加载prototxt模型配置文件; 2) 目标图片的读取和识别过程,此部分主要在网络中对图像数据进行计算和分类; 3) 运行结果的处理与显示.

### 1.3 系统硬件搭建

本系统基于2K龙芯派进行开发,其集成存储、显示、音频、网络等功能,功耗较低. 2K龙芯派采用的处理器为龙芯最新一代2号双核64位处理器2K 1000,该处理器面向网络安全领域及移动智能终端领域,主频为1 GHz,峰值运算速度为8 Gflops. 对龙芯开发板进行配置,选择基于Fedora内核的Loongson操作系统作为运行环境,配备U盘作为系统载体,相关环境的配置过程不做赘述. 通过mini HDMI线连接LCD显示屏作为输出设备,使用TTL将板子与电脑相连. 为了方便电脑与龙芯之间文件的传输,采用SSH协议进行通信.

## 1.4 系统算法与软件实现

### 1.4.1 图像数据集的获取

在本文的技术方案中,图像数据集包括训练集和测试集. 图像训练过程在PC端完成,选择ImageNet作为图像数据集,由于训练集数据量过大,受限于计算机的处理性能,需要对图像数据集进行二次整理. 筛选1000类对象(涉及生活用品、动植物、交通工具等),每类对象100张图片,大大减少训练量. 保证样本数据涵盖每一类对象不同形态、角度、清晰度的图像,建立txt格式文本,对每一类对象添加相应的标签信息. 训练集的数据量大小会影响最终识别结果,通常数据量越大识别率会越高,但是随之带来的训练时间和训练成本也会增加.

### 1.4.2 训练网络的建立

本文建立AlexNet训练网络的主要过程如下:首先,数据集中的图像先经过预处理和大小归一化处理,将所获取的图像数据集中所有图像都归一化为 $256 \times 256$ 的大小(如果图像数据集中有长方形的图像,则先将短边长度改为256,再裁剪图片中心对应大小的部分),并且对每张图像进行减均值处理. 其次,在Caffe深度学习框架下建立初始AlexNet网络模型,设定网络模型为8层. 前5层为卷积层,分别记为Conv 1~Conv 5,用于提取图像特征,卷积层数不足或过多都会导致性能的下降;后3层为全连接层,分别记为Conv 6~Conv 8,用于综合前面所提取到的特征,进行逻辑推断. 在卷积层Conv 1与Conv 2之间、卷积层Conv 2与Conv 3之间设立池化层,底层网络中特征的映射通常较大,意味着参数量也较大,采用池化融合特征,并对图像进行降维,减少参数量以防止数据过拟合. 同时,卷积层Conv 5与全连接层Conv 6之间也设立池化层,此处的池化是为了保证提取到的高级特性在图像任意位置的一致性. 池化层采用的是重叠最大池化,池化窗口大小为 $3 \times 3$ ,步长为2,有效地降低了识别错误率. AlexNet网络模型采用的激活函数均为ReLU函数,其表达式为

$$f(x) = \max(0, x), \quad (1)$$

其中 $x$ 在卷积层中代表RGB某个通道的某个像素,在全连接层中代表某个神经元的输出值. 相较于传统Tanh或Sigmoid这样的饱和和激活函数,输入为正时,ReLU函数输出值取其本身,此时关于输入 $x$ 的导数为1,是一个常数,避免了消失的梯度问题;当输入小于0时,输出为0,引进了稀疏性,训练速度得以加快. 最后确立训练网络模型,将预处理过的图像数据输入到所建立的AlexNet网络模型中进行训练,在训

练过程中,该网络模型的结构参数不断优化,最终训练参数被保存为二进制文件.

### 1.4.3 多类别图像识别算法的实现

在龙芯派平台下实现多类别深度学习图像识别程序,所采用的龙芯派平台以龙芯2K1000芯片为处理器,在操作系统中搭建好AlexNet网络模型后,从prototxt类型文件读取网络基本结构,从caffemodel类型文件读取训练参数,得到AlexNet网络模型, AlexNet网络模型具有5层用来对图像进行特征提取的卷积层以及3层对提取到的特征进行融合的全连接层. 主函数从指定系统文件夹位置读取图像文件,图像文件为任意图像格式,通过OpenCv环境下的imread函数,将输入图像转换成RGB数据的形式. 利用AlexNet网络模型对图像数据进行处理, AlexNet网络模型中最后一层网络(即最后一层全连接层)的输出通过如下softmax函数处理:

$$\sigma(z)_j = e^{z_j} / \sum_{k=1}^K (e^{z_k}),$$

$$j = 1, 2, \dots, K, \sum_{j=1}^k \sigma(z)_j = 1. \quad (2)$$

输出为1000维的向量,向量的每一个分量值均在 $0 \sim 1$ 之间,表征某一分类的概率值,主函数读取AlexNet网络模型的输出和标签文件,匹配输出概率最大的分量所对应的标签,并输出标签文件信息.

## 2 实验与结果分析

### 2.1 图片的采集与处理

利用Qt环境的widget方法组合各个控件,形成图形用户界面,实现选取文件夹、选取图片到识别结果显示的可视化过程. 设计图形用户界面,实现由用户选取待识别图像,自动显示目标图像所属类别.

### 2.2 多类别目标识别

在PC平台下测试了800张图片,PC使用的CPU为Intel酷睿i77700k,系统为微软Windows 10,未启用GPU加速. 在龙芯派平台下,由于测试平台所限,仅筛选200张图片进行测试,测试图片内容包含动物、交通工具、日用品等数10个大类,用来验证识别的可靠性. 部分测试结果如图1所示. 经过测试,目标图片在PC平台的识别率为96.5%,接近理论平均值96%;而在龙芯派平台下,全部测试图片中只有4张图片的识别出现了错误,有2张图片由于多个目标对象同时出现而识别结果不准确,总体识别率成功率达到97%,由于样本量的不足,识别率略高于理论平均值.

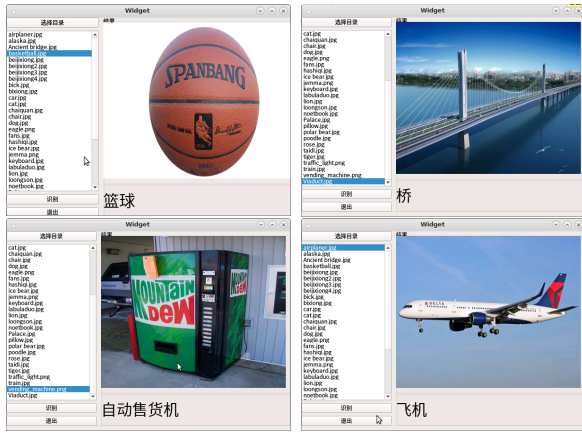


图1 多类别识别效果



图4 北极熊多姿态识别效果

### 2.3 单类别目标识别

本实验中选取类别相同但是相似程度不同的对象进行分类测试,记录识别结果,见图2~图4.选取犬类作为例子,实验结果如图2所示,表明此方法不仅可以识别出对象为犬类,且能具体识别出其品种,即使是让人容易混淆的阿拉斯加雪橇犬和哈士奇,其识别结果也相当精准.图3所示为对不同类型的车辆进行识别,也有非常好的识别效果.进一步,考虑多因素来测试方法的有效性.在不同背景、不同姿态下,对识别目标进行多次测试,如图4所示,考虑不同因素

下的识别速度和准确率依然不会受到任何影响.

### 2.4 算法性能分析及对比

本文所提出的方案核心为AlexNet网络,为了更好地体现本方案的可行性及优势,采用本文所提出的龙芯派平台下图像识别系统搭建方案,更换另3种常用神经网络模型VGG-16、GoogleNet以及MobileNet-V2.为了保证实验条件的统一,使用ImageNet作为训练数据集,保证数据集样本量数目的充足以及统一,共1000类图像,每类图像1000张.受限于设备的性能以及时间,该对比实验训练参数来源于他人训练结果.测试集包含训练样本所覆盖类别的200张图像.

测试结果如表1所示,所采用的4种网络中,VGG-16的识别准确率最高,但是由于参数量大,导致计算时间过长.GoogleNet和MobileNet是比较热门的轻量化网络,其参数量较小,因此识别速度很快,识别时间不到1s,但是识别准确率相比之下较低;采用AlexNet网络能保证识别时间在3s左右,且具有97%的高识别准确率.

表1 4种网络性能对比

网络模型	测试图片数量/张	识别成功数量/张	识别成功率/%	平均识别时间/s
VGG-16	200	198	99.0	20.3
AlexNet	200	194	97.0	3.1
GoogleNet	200	177	88.5	<1
MobileNet-V2	200	169	84.5	<1

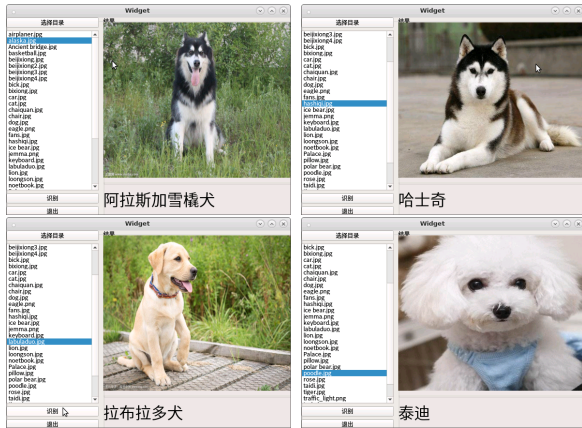


图2 犬类别识别效果

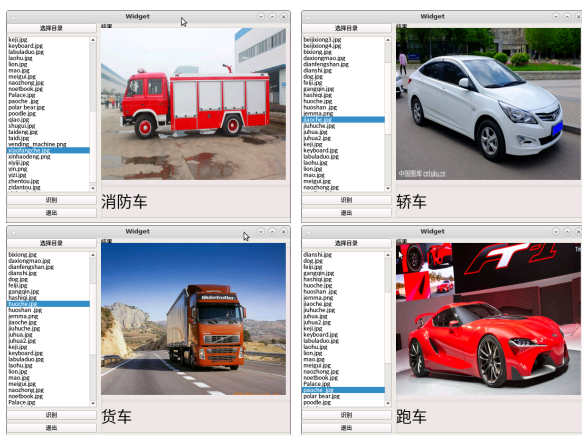


图3 车类别识别效果

### 2.5 扩展功能

除了图像识别功能,该应用还具有可扩展特性,通过替换网络模型和训练参数,可以实现图像处理领域内各类实用功能.另外,在本文所提出方案的基础上做了一些扩展尝试,比如实现人体骨骼检测(openpose算法)、目标检测、手势识别等,通过神经网络的运算,可以分析出人体骨骼分布,但是该实验的响应时间仍比较长,需要后续对算法进行持续改进.

### 3 结论

本文基于神经网络模型 AlexNet 具有优秀的识别率和识别速度的优点,设计了一种具有广泛应用前景的基于深度学习的多类别目标识别方案,并在国产龙芯派平台下进行了测试验证,解决了传统的图像识别种类少,外界环境因素变化适应性差等应用缺陷。

本文所做工作主要有以下特点: 1) 识别的多类别多目标性. 所设计的多类别目标识别方案对各种格式、各种品质的图像皆有良好的识别效果,囊括了各种常见的目标,理论上可识别近千种对象,具有普适性. 2) 适用性强. 利用动态编译的方式,并将训练与识别过程分开,不依赖处理器的高性能,在各种平台下都能完美适用. 3) 扩展性. 在应用方面,本设计具有良好的扩展功能. 当需要完成其他识别任务或是处理其他相关问题时,只需遵循本方案的设计步骤,通过修改网络结构、重新训练参数,能够快速完成基于龙芯派的图像处理应用开发。

#### 参考文献(References)

- [1] LeCun Y. Learning invariant feature hierarchies[J]. Lecture Notes in Computer Science, 2012, 7583(1): 496-505.
- [2] Kabani A W, El-Sakka M R. Object detection and localization using deep convolutional networks with softmax activation and multi-class log loss[C]. International Conference on Image Analysis and Recognition. Portugal, 2016: 358-366.
- [3] Karpathy A, Toderici G, Shetty S, et al. Large-scale video classification with convolutional neural networks[C]. Proceeding of the IEEE Confences on Computer Vision and Pattern Recognition. Piscataway: IEEE Computer Society, 2014: 1725-1732.
- [4] Pfister T, Simonyan K, Charles J, et al. Deep convolutional neural networks for efficient pose estimation in gesture videos[J]. Lecture Notes in Computer Science, 2015, 9003(1): 538-552.
- [5] 罗可, 周安众, 罗潇. 一种利用知识迁移的卷积神经网络训练策略[J]. 控制与决策, 2019, 34(3): 511-518. (Luo K, Zhou A Z, Luo X. Convolutional neural network training strategy using knowledge transfer[J]. Control and Decision, 2019, 34(3): 511-518.)
- [6] 吕恩辉, 王雪松, 程玉虎. 基于反卷积特征提取的深度卷积神经网络学习[J]. 控制与决策, 2018, 33(3): 447-454. (Lv E H, Wang X S, Cheng Y H. Deep convolution neural network learning based on deconvolution feature extraction[J]. Control and Decision, 2018, 33(3): 447-454.)
- [7] 张雪芹, 陈嘉豪, 诸葛晶晶, 等. 基于深度学习的快速植物图像识别[J]. 华东理工大学学报: 自然科学版, 2018, 44(6): 887-895. (Zhang X Q, Chen J H, Zhuge J J, et al. Deep learning based first plant image recognition[J]. Journal of East China University of Science and Technology: Natural Science Edition, 2018, 44(6): 887-895.)
- [8] 马永杰, 李雪燕, 宋晓凤. 基于改进深度卷积神经网络的交通标志识别[J]. 激光与光电子学进展, 2018, 55(12): 250-257. (Ma Y J, Li X Y, Song X F. Traffic sign recognition based on improved deep convolution network[J]. Laser and Optoelectronics Progress, 2018, 55(12): 250-257.)
- [9] 张敏, 刘杰, 蔡高勇. 基于卷积神经网络的柑橘溃疡病识别方法[J]. 计算机应用, 2018, 38(S1): 48-52. (Zhang M, Liu J, Cai G Y. Recognition method of citrus canker disease based on convolution neural network[J]. Journal of Computer Applications, 2018, 38(S1): 48-52.)
- [10] 孙星, 金鑫, 张晓昆, 等. 基于卷积神经网络的本征图像分解的实现[J]. 北京电子科技学院学报, 2017(4): 74-80. (Sun X, Jin X, Zhang X K, et al. Implementation of eigen image decomposition based on convolutional neural network[J]. Journal of Beijing Electronic Science and Technology Institute, 2017(4): 74-80.)
- [11] 袁公萍, 汤一平, 韩旺明, 等. 基于深度卷积神经网络的车型识别方法[J]. 浙江大学学报: 工学版, 2018, 52(4): 694-702. (Yuan G P, Tang Y P, Han W M, et al. Vehicle category recognition based on deep convolutional neural network[J]. Journal of Zhejiang University: Engineering Science, 2018, 52(4): 694-702.)
- [12] 刘婷婷, 王婷, 胡林. 基于卷积神经网络的水稻纹枯病图像识别[J]. 中国水稻科学, 2019, 33(1): 90-94. (Liu T T, Wang T, Hu L. Rhizocotonia solani recognition algorithm based on convolutional neural network[J]. Chinese Journal of Rice Science, 2019, 33(1): 90-94.)
- [13] 闫河, 王鹏, 董莺艳, 等. 改进的卷积神经网络图片分类识别方法[J]. 计算机应用与软件, 2018, 35(12): 193-198. (Yan H, Wang P, Dong Y Y, et al. Image classification and identification method based on improved convolutional neural network[J]. Computer Applications and Software, 2018, 35(12): 193-198.)

#### 作者简介

赵静(1983—), 女, 副教授, 博士, 从事故障诊断与容错控制、智能控制等研究, E-mail: zhaojing@njupt.edu.cn;  
 王弦(1996—), 男, 硕士生, 从事故障诊断、智能控制以及嵌入式系统的研究, E-mail: 1017051322@njupt.edu.cn;  
 王奔(1998—), 男, 本科生, 从事智能电网、智能控制的研究, E-mail: 18851981069@163.com;  
 蒋国平(1966—), 男, 教授, 博士生导师, 从事混沌系统同步与控制、复杂网络理论及其应用等研究, E-mail: jianggp@njupt.edu.cn;  
 谢非(1983—), 男, 副教授, 博士, 从事视觉与图像处理、多源导航与定位的研究, E-mail: xiefei@njnut.edu.cn;  
 徐丰羽(1979—), 男, 副教授, 博士, 从事机器人及自动化、机电一体化技术的研究, E-mail: xufengyu598@163.com.