

基于深度学习的南方野生鸟类识别系统的设计

赖小平

(广东交通职业技术学院信息学院, 广东 广州 510650)

✉laixiaoping1008@163.com



摘要: 针对鸟类的识别对人类和计算机视觉要求较高, 视频检测和人工识别等传统的识别方法均存在一定的困难, 提出一种基于深度学习的鸟类识别系统, 借助百度EasyDL深度学习平台, 以南方野生鸟类为实验对象进行数据标注、模型建立、模型训练, 经验证集测试验证, 识别精确率为94.2%。文章提出的模型可应用于已知鸟类的实时监测识别, 并辅助实现对未知鸟类的监测与发现, 为生物多样性保护提供技术支撑。

关键词: 深度学习; 百度EasyDL; 野生鸟类; 物体检测; 模型

中图分类号: TP391 **文献标识码:** A

Design of Southern Wild Bird Recognition System based on Deep Learning

LAI Xiaoping

(College of Information Technology, Guangdong Communication Polytechnic, Guangzhou 510650, China)

✉laixiaoping1008@163.com

Abstract: Recognition of birds requires a high level of human and computer vision, and traditional recognition methods such as video detection and manual recognition have certain difficulties. In view of these problems, this paper proposes a bird recognition system based on deep learning. With the help of Baidu EasyDL depth learning platform, data annotations, model building and model training are carried out by taking southern wild birds as experimental objects. Verification set test results show that the recognition accuracy rate is 94.2%. The proposed model can be applied to the real-time monitoring and recognition of the existing birds, and assist in the monitoring and discovery of unknown birds, providing technical support for biodiversity conservation.

Keywords: deep learning; Baidu EasyDL; wild birds; object detection; model

1 引言(Introduction)

野生鸟类是全人类的财富, 但随着人类活动范围的加大, 野生鸟类的种群和数量不断减少。分析野生鸟类资源日益减少的原因, 除环境因素外, 人类为经济利益大量捕捉、贩卖也是一大因素^[1]。为了保护鸟类和地球物种多样性, 增强人们保护野生鸟类的意识, 构建智能自动化的鸟类识别系统成为当务之急^[2]。鸟类识别技术能够帮助相关工作人员实现智能自动化的鸟类监测, 对于保护鸟类、了解鸟类生态系统和评测环境质量具有重要意义^[3], 对其准确分类一直为研究的热点。传统的识别方法有视频检测和人工识别方法, 视频检测成本高、检测目标范围有限; 而人工识别大多依靠专家经验, 效率低, 这种识别方式成本高, 识别过程受多种因素制

约, 实施困难。随着人工智能技术的快速发展, 越来越多先进的技术走入了人们的生活, 研究正式进入人工智能时代。本文拟采用人工智能技术, 基于百度EasyDL深度学习平台构建鸟类识别模型进行训练与识别, 以提高对鸟类识别的准确度和效率。

2 鸟类识别的主要研究成果(Major research achievements in bird recognition)

近年来, 科研人员对鸟类识别相关技术的研究仍处于起步阶段, 研究成果较少。在中国知网中输入检索条件: 主题为鸟类识别、文献发表时间为2012—2022年, 从图1中可知近10年来(2012年至今)与鸟类识别技术相关的文献数量不足百篇。

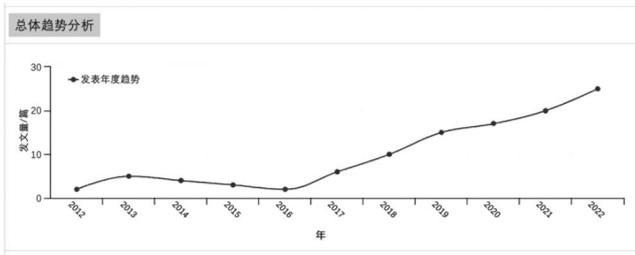


图1 研究趋势分析

Fig.1 Research trend analysis

主要研究成果如下：(1)采用基于部位的思想将识别任务划分为四个模块：目标与部位检测、数据扩增、特征提取和分类决策，用于解决鸟类识别中的难点问题^[2]。(2)利用级联分类器和卷积神经网络人工智能识别技术对鸟类进行监测、识别^[4]。(3)采用一种基于小波变换和霍特林变换(KL)进行候鸟特征提取，并以BP神经网络作为分类器研究设计出一套适用于闽江口湿地复杂环境的候鸟识别方案^[5]。(4)两种基于弱监督信息图像识别方法用于细粒度图像分类(FGVC)。一种是联合残差网络和Inception网络，通过优化卷积神经网络的网络结构提高捕捉细粒度特征的能力；另一种是对双线性卷积神经网络模型(CNN)进行改进，特征提取器选取Google提出的Inception-v3模组和Inception-v4模组，最后把不同的局部特征汇集起来进行分类^[6]。(5)一种基于跨层精简双线性池化的深度卷积神经网络模型。首先，根据Tensor Sketch算法计算多组来自不同卷积层的精简双线性特征向量；其次，将归一化后的特征向量级联送至Softmax分类器；最后，引入成对混淆交叉熵损失函数进行正则化以优化网络^[7]。(6)一种新的无完全连接层的掩码CNN模型。基于部位标注，提出的模型由一个完全卷积网络组成，用于定位识别部位(如头部和躯干)，更重要的是生成加权对象/零件掩码，用于选择有用且有意义的卷积描述符。由于摒弃了参数冗余的全连接层，与其他细粒度方法相比，此Mask CNN具有较小的特征维数和高效的推理速度^[8]。

3 鸟类识别系统的设计(Design of bird recognition system)

本课题结合传统检测算法和深度学习算法，基于百度EasyDL开发平台，在实际鸟类观测中利用训练样本集数据进行训练，并检测识别模型。经训练后，进行测试验证，如达到工程可用的程度(准确率>85%)，就能应用于已知鸟类的识别，并可辅助工作人员发现未知鸟类。

3.1 百度EasyDL

EasyDL是百度大脑推出的零门槛人工智能(AI)开发平台，具有从数据采集、标注、清洗到模型训练、部署的一站式AI开发能力。EasyDL设计简约，极易理解，采集到的原始图片、文本、音频、视频、文字识别(OCR)、表格等数据，经过EasyDL加工、学习、部署后，可通过公有云应用程序界面(API)调用，或部署在本地服务器、小型设备、软硬一体方案的专项适配硬件上，通过离线软件开发工具包(SDK)或私有API进一步集成。开发流程如图2所示。

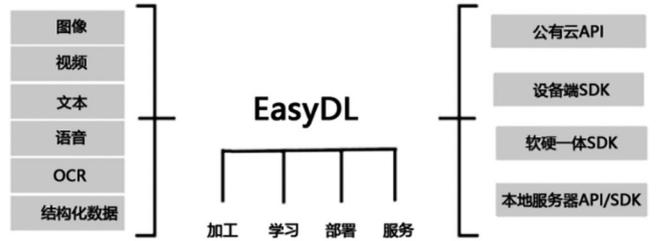


图2 EasyDL开发流程

Fig.2 EasyDL development process

操作流程大致分为以下四步：第一步：创建模型，即确定模型名称，可添加模型描述，便于后续模型迭代管理。第二步：上传并标注数据。上传数据后，根据不同模型类型的的数据要求进行标注，如果有本地已标注的数据，也可以直接上传。通常，需要对训练集、测试集和验证集三类数据集进行数据标注。第三步：训练模型并校验效果。选择算法类型、配置训练任务相关参数完成训练任务启动。模型训练完毕后支持可视化查看模型效果评估报告，也支持通过模型校验功能在线上上传实测数据验证模型效果。第四步：部署模型。根据业务场景，支持将模型部署为公有云API实现在线调用，或部署在本地服务器/小型设备/软硬一体方案的专项适配硬件上，通过API/SDK集成离线应用。

3.2 创建模型

EasyDL支持六大技术方向，每个方向包括不同的模型类型：(1)EasyDL图像：图像分类、物体检测、图像分割；(2)EasyDL文本：文本分类-单标签、文本分类-多标签、文本实体抽取、情感倾向分析、短文本相似度；(3)EasyDL语音：语音识别、声音分类；(4)EasyDL OCR：文字识别；(5)EasyDL视频：视频分类、目标跟踪；(6)EasyDL结构化数据：表格预测。

本系统采用的是EasyDL图像的物体检测模型，创建物体检测模型的方法：百度EasyDL开发平台→选择模型类型为【物体检测】→登录百度账号提交基础信息→点击【创建模型】，填写模型名称、说明等信息，完成模型创建。

3.3 上传并标注数据

我们拍摄整理了5种鸟类共300多张照片，创建了三类数据集：即训练集(数量为254)、测试集(数量为73)和验证集(数量为50)。

标注的过程如图3所示，先创建5个标签，即对应5种鸟类(乌鸫、铜蓝鹟、普通翠鸟、红头长尾雀、叉尾太阳鸟)，然后对每张图片进行矩形标注，最后保存，即可完成标注。



图3 数据标注

Fig.3 Data annotations

3.4 训练模型并校验效果

选择待训练的模型，设置训练参数，主要是选择训练

集、测试集和验证集，然后开始进行训练，支持对同一模型多次迭代；不同的模型完成训练的时间不一样，训练完成后，账号对应的手机号会接收到短信信息通知，后续可进行模型效果校验。

上传校验图片，模型校验结果如图4所示。物体检测模型存在一个可调节的阈值(*threshold*)，是正确结果的判定标准，例如阈值是0.6，置信度大于0.6的识别结果会被当作正确结果返回。每个物体检测模型训练完毕后，可以在模型评估报告中查看推荐阈值。校验模型准确度可根据需求调整阈值，本模型设置阈值为0.8时，效果最佳，*mAP*平均精度为95.38%，当前图片检测置信度为98.37%。



图4 模型校验

Fig.4 Model verification

3.5 部署模型

训练完成的模型即可申请模型上线，上线后在“我的模型”查看“上线详情”，获得专有的定制图像识别API。后续还可创建应用，采用Python等进行应用开发。

百度Easy DL平台提供了H5体验，用手机等终端扫描系统提供的二维码即可进行体验，在图5(a)中点击上传一张鸟类图片，即可查看识别结果如图5(b)所示。



(a)首界面



(b)识别结果

图5 H5 体验

Fig.5 H5 experience

同时，Easy DL平台提供了编码调用模型进行测试，参考代码如下：

```
# 目标图片的本地文件路径，支持jpg/png/bmp格式
IMAGE_FILEPATH="[检测图片地址]"
# 可选的请求参数
# threshold：默认值为建议阈值，请在我的模型-模型效果-完整评估结果-详细评估查看建议阈值
PARAMS={"threshold": 0.3}
# 服务详情中的接口地址
MODEL_API_URL="[服务接口地址]"
# 调用API需要ACCESS_TOKEN。若已有ACCESS_TOKEN则于下方填入该字符串
# 否则，留空ACCESS_TOKEN，于下方填入该模型部署的API_KEY以及SECRET_KEY，会自动申请并显示新ACCESS_TOKEN
ACCESS_TOKEN=""
API_KEY="[API_KEY]"
SECRET_KEY="[SECRET_KEY]"
print("1. 读取目标图片{}".format(IMAGE_FILEPATH))
with open(IMAGE_FILEPATH, 'rb') as f:
    base64_data=base64.b64encode(f.read())
    base64_str=base64_data.decode('UTF8')
print("将BASE64编码后图片的字符串填入PARAMS的'image'字段")
PARAMS["image"]=base64_str
if not ACCESS_TOKEN:
    print("2. ACCESS_TOKEN为空，调用鉴权接口获取TOKEN")
    auth_url="https://aip.baidubce.com/oauth/2.0/token?grant_type=client_credentials" + "&client_id={}&client_secret={}".format(API_KEY, SECRET_KEY)
    auth_resp=requests.get(auth_url)
    auth_resp_json=auth_resp.json()
    ACCESS_TOKEN=auth_resp_json["access_token"]
    print("新ACCESS_TOKEN: {}".format(ACCESS_TOKEN))
else:
    print("2. 使用已有ACCESS_TOKEN")
print("3. 向模型接口'MODEL_API_URL'发送请求")
request_url="{}?access_token={}".format(MODEL_API_URL, ACCESS_TOKEN)
response=requests.post(url=request_url, json=PARAMS)
response_json=response.json()
response_str=json.dumps(response_json, indent=4, ensure_ascii=False)
print("结果:\n{}".format(response_str))
```

3.6 模型评估

与模型效果相关的主要指标有以下几类。

(1)准确率(*Accuracy*)。准确率是图像分类/文本分类/声音分类等分类模型的最简单、最直观的衡量指标，即正确分类的样本数与总样本数之比，比值越接近1的模型，识别效果越好。

(2)精确率(*Precision*)。精确率又叫查准率，容易与准确率弄混，精确率代表对正样本结果中的预测准确程度，而准确率则代表整体的预测准确程度，既包括正样本，也包括负样本。模型的精确率为正确预测某一类别的样本数与预测为该类别的总样本数之比。

(3)召回率(*Recall*)。召回率又叫查全率，它是针对原样本而言的，即在实际为正的样本中被预测为正样本的概率。模型召回率为正确预测某一类别的样本数与该类别的总样本数之比。

(4)*mAP*(mean Average Precision)是物体检测(Object Detection)算法中衡量算法效果的指标。对于物体检测任务，每一类object都可以计算出其精确率和召回率，在不同阈值下多次计算/试验，每个类都可以得到一条P-R曲线，曲线下的面积就是average。

(5)混淆矩阵(Confusion Matrix)。混淆矩阵也被称为错误矩阵(Error Matrix)，是机器学习中总结分类模型预测结果的情形分析表，以矩阵形式将数据集中的记录按照真实的类别与分类模型预测的类别进行汇总。矩阵的每一列表达了分类器对样本的类别预测，每一行则表达了版本所属的真实类别。通过混淆矩阵能够很容易地看到机器学习有没有将样本的类别混淆。

本文构建的模型，对南方五种野生鸟类，即叉尾太阳鸟、红头长尾山雀、普通翠鸟、铜蓝鹟和乌鸫进行识别检测，经过三次迭代，训练时间与数据量大小有关，本次训练大约耗时15 min，整体评估如图6所示。



图6 模型评估

Fig.6 Model evaluation

模型评估中可以看到模型训练整体的情况说明，包括基本结论、*mAP*(96.9%)、精确率(94.7%)、召回率(93.5%)。当数据量较少时，得出的模型评估报告结果仅供参考，无法完全准确体现模型效果。一般要求每种类别的数据量不少于30，以50—100为宜。本鸟类识别系统模型的混淆矩阵如表1所示。

表1 混淆矩阵

Tab.1 Confusion matrix

类别	铜蓝鹟	红头长尾山雀	普通翠鸟	叉尾太阳鸟	乌鸫	[default]	精确率/%	测试集数量
铜蓝鹟	15	0	0	0	0	0	100	15
红头长尾山雀	0	14	0	3	0	1	78	18
普通翠鸟	0	0	15	1	0	0	94	16
叉尾太阳鸟	0	0	0	15	0	0	100	15
乌鸫	0	0	0	0	13	0	100	13

从表1中可知，有3个【红头长尾山雀】和【叉尾太阳鸟】的检测框发生了混淆，其中【红头长尾山雀】分到【叉尾太阳鸟】为3个，有1个【红头长尾山雀】漏检；有1个【普通翠鸟】和【叉尾太阳鸟】的检测框发生了混淆，其中【普通翠鸟】分到【叉尾太阳鸟】为1个。

4 结论(Conclusion)

深度学习算法需要大量数据进行训练建模，而鸟类资源保护尤其是珍稀濒危鸟类保护面临的现实问题就是原始数据严重不足。本文对南方野生鸟类进行建模、训练，经验证集测试验证，识别精确率为94.2%。本文提出的模型可应用于已知鸟类的实时监测识别，并辅助实现对未知鸟类的监测与发现，为生物多样性保护提供技术支持。本模型中用到的数据集较少，检测准确率还有提升空间。后续需要加大对鸟类观测数据的采集力度，积累更多样本数据，同时加大模型训练的力度，对模型参数调整优化，进行多次迭代，预期可以进一步提升识别准确率，最终实现在前端设备上的现场在线检测识别。

参考文献(References)

- [1] 林英华,顾海军.野生鸟类疾病的研究现状[J].动物学杂志,1999(06):45-48.
- [2] 智绪晔.基于深度学习的鸟类图像识别研究[D].大连:大连理工大学,2019.
- [3] 王越,冯振.基于CAM与双线性网络的鸟类图像识别方法[J].重庆理工大学学报(自然科学),2021,35(11):136-141,239.
- [4] 蒋敏,田元,吴伟志,等.鸟类全景观测系统中的人工智能识别技术[J].浙江林业科技,2021,41(04):108-113.
- [5] 林茜,邢婷婷,吴帆旭.基于BP神经网络的闽江口湿地鸟类物种识别的研究[J].网络安全技术与应用,2019(05):37-39.
- [6] 朱阳光,刘瑞敏,黄琼桃.基于深度神经网络的弱监督信息细粒度图像识别[J].电子测量与仪器学报,2020,34(02):115-122.
- [7] 蓝洁,周欣,何小海,等.基于跨层精简双线性网络的细粒度鸟类识别[J].科学技术与工程,2019,19(36):240-246.
- [8] WEI X S, XIE C W, WU J X, et al. Mask-CNN: Localizing parts and selecting descriptors for fine-grained bird species categorization[J]. Pattern Recognition, 2018(76):704-714.

作者简介:

赖小平(1979-),女,硕士,副教授.研究领域:人工智能,软件技术.