

改进的基于端到端学习的蜻蜓目昆虫识别算法^{*}

彭明杰^{1,2,3}, 唐万梅¹, 皮家甜^{1,2,3}, 于 昕⁴, 吴至友³, 黄芳婷⁴

(1. 重庆师范大学 计算机与信息科学学院; 2. 重庆师范大学 重庆市数字农业服务工程技术研究中心;
3. 重庆师范大学 智慧金融与大数据分析重庆市重点实验室; 4. 重庆师范大学 生命科学学院, 重庆 401331)

摘要:【目的】传统的蜻蜓目(Odonata)昆虫的鉴别算法是在标本图片上进行人工的特征提取来训练分类器,此类方法所用的数据集包含的蜻蜓目昆虫种类和数量均较少,并且样本背景单一,导致识别率低且只能识别标本图片。针对这些问题制作了一个同时包含蜻蜓目昆虫生态图片和标本图片的数据集,提出一种基于深度学习的蜻蜓目昆虫的识别算法。【方法】采用具备端到端学习优势的网络框架,在上述数据集上,使用在 ImageNet 上迁移学习的 ResNet50 进行特征提取,使用新的区域建议网络 Imp-RPN 进行特征筛选,再使用改进的损失函数来解决样本分布不均的问题。【结果】所提出的识别算法在精确到种的 46 类分类任务中识别率达到了 96.2%,在精确到种并包含性别信息的 90 类分类任务中识别率则达到了 88.7%。【结论】端到端的深度学习网络框架免除了人工标注样本的时间成本,提高了识别准确率,更能满足物种鉴别任务的需求。

关键词:蜻蜓目识别;细粒度识别;深度学习;端到端学习;NTS-Net

中图分类号:TP391.4

文献标志码:A

文章编号:1672-6693(2021)03-0094-06

蜻蜓是节肢动物门(Arthropoda)昆虫纲(Insecta)蜻蜓目(Odonata)昆虫的通称。该目均为捕食性昆虫,且食物主要为蝇蚊类、叶蝉类、虻蠓类、小型蝶蛾类等多种农林牧业害虫,是一类重要的天敌昆虫。蜻蜓还是重要的环境指示昆虫,可以用来监测环境污染,在农业和环境生态系统中有着重要作用^[1]。同时,亦因它们五光十色的外表,翩翩的舞姿而具有相当的观赏价值,自古便受到人类的青睐,也是著名的文化昆虫。目前中国的已知蜻蜓种类大约有 700 种^[2-3],是世界上蜻蜓种类最丰富的国家之一。然而,由于存在“雌雄异型、近缘种表型趋同”等现象,人工鉴别蜻蜓种类需要经过长期的专业训练和经验积累,尤其有一些常见但容易混淆的种类,更使得与蜻蜓相关的应用研究的前期成本过高。为了满足蜻蜓目昆虫研究中的物种鉴别任务需求,提高鉴别效率,推进相关研究工作,同时也为了方便昆虫爱好者以及大众了解蜻蜓,本研究将制作包含大量生态图片和部分标本图片的中国蜻蜓目昆虫数据集(仅包括中国本土物种),提出改进的基于 NTS-Net 的深度学习网络模型来构建蜻蜓种类自动识别系统。

近些年来,深度学习技术在生物识别上的作用受到了越来越多的研究者的关注,并出现了一些昆虫识别的应用研究,如谢娟英等人^[4]提出的蝴蝶检测和识别方法等。此类研究大都采取目标检测到目标识别两步走的架构来构建深度学习网络。其中,目标检测部分常用的深度学习目标识别方法有 Girshick 等人^[5-6]提出的目标检测框架 R-CNN 和 Fast R-CNN; Redmon 等人^[7-8]提出的 YOLO 目标检测框架及后续改进模型 YOLOv3、YOLOv4; Liu 等人^[9]提出的 SSD 检测框架等。此类方法能够高效地进行目标检测,对不同种类的目标有不错的识别率,但在更加精细区分的细粒度识别任务上往往识别率很低,同时该类方法的训练数据需要匹配大量的人工标注信息,颇为耗时。针对以上缺点,本文在运用了区域建议网络(Region proposal networks, RPN)思想的网络框架 NTS-Net^[10]的基础上,以新的区域建议网络 QT-RPN 作为特征筛选网络,根据数据集的分布特点改进了损失函数,提出一种新的识别算法。

蜻蜓目的自动识别研究当前主要面临以下 3 个问题。1) 现有的识别算法^[11]仍使用传统的人工设计的特征

^{*} 收稿日期:2020-10-11 修回日期:2021-05-04 网络出版时间:2021-05-20 09:25

资助项目:科学技术部科技基础性工作专项重点项目(No. 2015FY210300);重庆市教育委员会科技项目青年项目(No. KJQN201800521);重庆市基础研究与前沿探索项目(No. cstc2018jcyjAX0470);重庆市研究生科研创新项目(No. CYS20247)

第一作者简介:彭明杰,男,研究方向为目标检测与识别,E-mail:812623090@qq.com;通信作者:皮家甜,男,讲师,博士,E-mail:pjijiatian@cqnu.edu.cn

网络出版地址:https://kns.cnki.net/kcms/detail/50.1165.N.20210519.1537.018.html

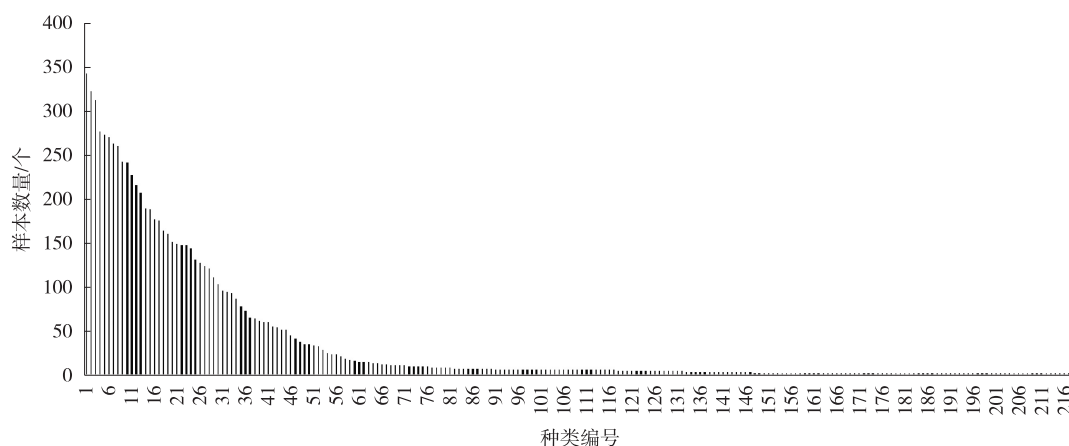
提取方式,识别率很低且只能识别标本图片;2) 训练深度学习模型所需的数据量巨大,但找不到可用的公共数据集;3) 一些稀有的蜻蜓种类在自然状态下的照片太少,无法直接用作训练集。鉴于以上情况,本研究制作了一个较为丰富的蜻蜓目昆虫图片数据集 QT216-CQNU,包含中国现有的蜻蜓目昆虫种类中的 216 种,其中一部分是蜻蜓标准图片(白色背景标本图片),大部分则是蜻蜓在生态环境中的图片,并在时间推移过程中不断完善该数据集。

1 数据集

1.1 中国本土蜻蜓目昆虫数据集

本研究制作的蜻蜓目数据集中的蜻蜓图片不仅包含实验室中的蜻蜓标本图片,更包含在生态环境中的蜻蜓图片。其中前者称为标准照片,在所有的蜻蜓图片中占比较少,主要拍摄于实验室或从他处扫描而得;后者称为生态照,占绝大部分。

数据集共包含 8 056 张蜻蜓目昆虫的图片,共 216 种,其中生态照样本如封二彩图 1 所示。蜻蜓生态照的特点是每张图片中的蜻蜓数量不确定,但绝大多数为 1 只,并且由于蜻蜓自身的生物习性,使之总是倾向于处在有利于隐藏自身的环境中,让观察者辨别困难。数据集的种类分布信息如图 2 所示。每个种类至少包含 1 个样本,最多的种类含有 342 个样本,总体呈长尾分布。



注:为便于浏览,图中的线柱采用深色与浅色交替出现的风格

图 2 数据集分布图

Fig. 2 Dataset distribution map

1.2 训练数据集

模型使用的数据集无需进行人工标注,均为原图片本身。因数据集中各种类对应的样本数量呈长尾分布,为满足深度学习训练要求,先选取了样本数量排在前 46 的种类共 6 730 张进行训练,该数据集名称为 QT46-CQNU,其中训练集和测试集各 3 365 张;又因蜻蜓目昆虫具有“雌雄异型”(封二彩图 3)的特点,故又选取了 90 种不同型体的精确到性别的蜻蜓目物种进行训练,该数据集名称为 QT90-CQNU,其中训练集和测试集各 3 587 张,共 7 174 张。

2 网络结构

2.1 NTS-Net 框架

NTS-Net 是由 Yang 等人^[10]提出的一种基于区域建议网络思想的细粒度分类网络设计框架。主要由 3 部分组成:Navigator 网络、Teacher 网络和 Scrutinizer 网络。该框架特点在于利用区域建议网络 Navigator 网络来寻找输入图像中的高信息区域,经 Teacher 网络筛选后作为 Scrutinizer 网络的输入,以此来降低图片中的复杂背景产生的噪声从而提高模型的准确度。

NTS-Net 框架的设计结构如封二彩图 4 所示。它的整体训练方法为:对于输入图像,提取深层特征图后,输入 Navigator 网络计算所有区域的信息量。然后通过 NMS 选出信息量排名靠前的 M 个区域(每个区域的信息表示为 $\{I_1, I_2, \dots, I_M\}$),从完整的图像中裁剪出来,将它们调整为指定的大小,然后输入到 Teacher 网络中,得到这些区域的置信度 $\{C_1, C_2, \dots, C_M\}$ 。并通过优化使得 $\{I_1, I_2, \dots, I_M\}$ 和 $\{C_1, C_2, \dots, C_M\}$ 的顺序相同。最后将这

M 个区域与原图进行级联后输入 Scrutinizer 网络进行分类。该框架的特征提取网络使用的是 ResNet-50^[12]。

2.2 RPN

Zhou 等人^[13]指出,卷积神经网络中的卷积核本身已拥有图像定位的作用,这一作用在网络末端使用全连接层分类时则消失,故只需将此处的全连接层替换为全局池化层,再利用类激活映射(Class activation mapping, CAM)技术反向映射卷积特征即可直接观察到卷积神经网络从图像中提取的高信息区域。这一结果给卷积神经网络的可解释性和区域建议网络(Region proposal networks, RPN)思想的发展提供了技术支撑。RPN 在思想和方法论上与注意力机制异曲同工,都是期望得到更具判别价值的输入特征,不同的是:前者在空间维度上学习特征权重后,使用不同尺度的候选框来对输入图片进行区域筛选,而后者试图从不同的维度(如通道、空间、时间等)学习权重分布,再将此权重施加于原图进行加权求和,使得网络自行关注权重高的部分。此后 Ren 等人^[14]率先在他们提出的 Faster R-CNN 中采用这一思想,通过利用卷积神经网络寻找图像中的多个指定大小的高信息区域,代替原图作为后续分类网络的输入,有效地提升了模型的准确度。在蜻蜓数据集上的高信息区域定位结果如封三彩图 5 所示,可以发现其中有些样本的高信息区域出现在了图片的边沿,进而对 RPN 进行了改进。

2.3 QT-RPN

2.3.1 池化与反池化 在卷积神经网络中,池化层因具有平移、旋转和尺度不变性而能在使图片降维的同时仍保留图片的主要特征。当使池化层的填充(Padding)小于随后的卷积层的填充时,获取的特征图可以一定程度上剔除图片的边缘信息。在本文的数据集和平常的生态环境中,蜻蜓通常不会靠在图片的边缘,故而可以通过使得 RPN 忽略一些边缘信息来提高所筛选的高信息区域的置信度。

2.3.2 深度可分离卷积 深度可分离卷积^[15-16](Depthwise separable convolution, DSC)将完整的卷积运算分解为两个步骤,即深度卷积(Depthwise convolution)和点卷积(Pointwise convolution)。深度卷积过程中,每个通道都与独立的卷积核进行运算,运算完成后输出的通道数与输入相同,无法扩展特征图,同时也忽略了各个通道在相同空间位置上的特征信息。因此,引入点卷积来生成新的特征图并组合这些信息以得到跟标准卷积运算相同的结果。这样就在达到了卷积效果的同时大量地减少了网络的参数量。

2.3.3 新的区域建议网络 考虑到因生态环境的多样性而导致的蜻蜓图片背景复杂的特点和端到端模型存在的网络参数量较大的情况,本文基于上述模块,搭建了全新的区域建议网络 QT-RPN。首先对输入图片进行 3×3 的深度卷积加 1×1 的点卷积,次接 3×3 的平均池化,再接 3×3 的最大反池化(从此处单分一支 1×1 的卷积),随后又是一组深度可分离卷积(此处单分一支 1×1 的卷积),后接 3×3 标准卷积,紧接 1×1 卷积,最后与两支路级联。具体网络结构如封三彩图 6 所示。

2.4 改进损失函数

由于本文的数据集中的各类蜻蜓数量呈长尾分布,这会导致数量较大的种类在网络中的权重会优于数量小的种类。故而将原 Scrutinizer 网络的损失函数由交叉熵损失(Entropy loss)替换为焦点损失(Focal loss)^[17],并引入新的超参数 μ 来控制作用范围。当 Navigator 网络给出高信息区域 $\{R_1, R_2, \dots, R_K\}$ 后,Scrutinizer 网络给出原样本 X 与这些特征级联后的特征的识别结果 $P = S(X, R_1, R_2, \dots, R_K)$ 。Scrutinizer 网络的原损失函数为 $L_s = -\log_{10} S(X, R_1, R_2, \dots, R_K)$,新损失函数为 $\hat{L}_s = -\alpha(1-P)^\gamma \log_{10}(P)$,最后得到最终的总损失函数为: $L = L_1 + \lambda L_2 + \mu \hat{L}_s$ 。其中: L_1 为 Navigator 网络的损失函数; L_2 为 Teacher 网络的损失函数; α 和 γ 是 Focal loss 的超参数,因为在多分类任务中, α 的取值不影响结果,此处恒定为 0.5, γ 的取值为 2; $\lambda = 1$ 。

在准备数据集时按各类所含样本数多少的降序来给予编号(n_{id}),即第一类数量最多,最后一类数量最少。当当前训练样本的标签在总标签数($V_{classes}$)的前一半时, $\mu = 0.5$,反之 $\mu = 1$ 。从而 μ 的取值可以表示为:

$$\mu = \begin{cases} 0.5, n_{id} < \frac{V_{classes}}{2} \\ 1, n_{id} > \frac{V_{classes}}{2} \end{cases}。$$

3 实验结果与分析

3.1 设备与参数

实验所使用的 GPU 为 TITAN XP 单卡,操作系统为 Ubuntu 16.04。将图像的输入大小(Input size)设置

为 448×448 像素, M 为 RPN 筛选的高信息区域数量, K 为 Teacher 网络从这些区域中选出的最终作为判别网络的输入数量,且 M 大于等于 K 。最终在硬件条件允许的范围内,在 M 和 K 均为 6 的情况下取得了最好的效果。实验使用在 ILSVRC2012 数据集^[18]上预训练过的 ResNet-50 作为特征提取器,使用 Batch Normalization (BN)来实现正则化。模型训练的 Batch size 设置为 8,初始学习率(Learning rate, LR)为 0.001,优化器为 Momentum SGD。NMS 阈值为 0.25,权重衰减(Weight decay)为 0.000 1,迭代 epoch 为 200。具体超参数设置如表 1 所示。

3.2 数据预处理

由于数据集中每张图片的大小不一致,所以先将所有图片重新设定大小到 600×600 像素,然后对训练集使用随机裁剪和概率为 0.5 随机水平翻转,对测试集使用中心裁剪,裁剪的大小均为 Input size。

3.3 结果分析

实验使用 top1 准确率(即模型对测试图片返回的预测标签中概率最大者为正确标签的概率)作为模型的评估指标。由于此网络无须数据集的标注信息,所以无法跟需要数据集标注信息的网络进行比较,本文对比了同样无须数据集标注信息的其他细粒度分类网络。分别给出了 WS-DAN^[19]、DCL^[20]、RA-CNN^[21]、Bilinear-CNN^[22]、DT-RAM^[23]、DFL-CNN^[24]及原 NTS-Net 网络在数据集 QT46-CQNU 和 QT90-CQNU 上的实验结果。在数据集 QT46-CQNU 上,经过改进的 ImpNTS-Net 在 top1 准确率上达到了 96.2%,较原网络提高了 1.3%,较其他网络提高了 1%~15%(见表 2)。在数据集 QT90-CQNU 上准确率达到 88.7%,较原网络提高了 1.5%,较其他网络提高了 0.6%~17.8%。

在新损失函数下,在样本数量较低的类别中分类结果的混淆矩阵更为稀疏(封三彩图 7)。在封三彩图 7 中的混淆矩阵中,列为真实标签,行为预测标签,元素值为真实标签预测为预测标签的样本个数。可以看出在新的损失函数下混淆矩阵的稀疏性和收敛性较好。

表 2 在 QT46-CQNU 数据集上的结果
Tab.2 Results on the QT46-CQNU dataset

网络模型	Top1 准确度/%
WS-DAN	95.2
DCL	93.2
RA-CNN	91.8
Bilinear-CNN	92.4
DT-RAM	92.4
DFL-CNN	81.2
NTS	94.9
ImpNTS	96.2

表 3 在 QT90-CQNU 数据集上的结果
Tab.3 Results on the QT90-CQNU dataset

网络模型	Top1 准确度/%
WS-DAN	88.1
DCL	83.5
RA-CNN	84.5
Bilinear-CNN	82.6
DT-RAM	84.7
DFL-CNN	70.9
NTS	87.2
ImpNTS	88.7

3.4 蜻蜓自动识别系统

将在 QT46-CQNU 上训练的模型称为 Mod46,将在 QT90-CQNU 上训练的模型称为 Mod-gender。基于此二者,本文结合 Python 语言及该语言的 UI 交互界面包 QT5 完成了蜻蜓目自动识别系统的搭建。该系统主要分为两个模块,即输入模块与输出模块。输入模块是指用户选择需要鉴别的蜻蜓目昆虫图片传入系统,程序会首先对输入的图片进行预处理,随后交由识别模型进行判别;输出模块用来显示模型给出的判别结果,返回的结果主要有蜻蜓的种类名称,鉴别为该种类的概率以及一张该类蜻蜓的示例图,如封三彩图 8 所示。当搭载模型为 Mod46 时,结果精确到种,当搭载模型为 Mod-gender 时,结果精确到包含性别的种。

4 结论

本文制作了一个数量比较丰富的可用于深度学习中国本土蜻蜓目昆虫数据集,弥补了现有蜻蜓自动识别的研究中使用的数据集中蜻蜓种类少、数据量小和缺乏自然环境下蜻蜓数据的缺陷。在此基础上提出改进的端到端学习的蜻蜓识别算法。在端到端学习网络框架 NTS-Net 的基础上,通过迁移学习的预训练模型提取特征,使用新的区域建议网络 QT-RPN 筛选特征,再使用改进的 Focal loss 损失函数降低样本数量分布不均的影响。最后搭建了一个能对自然环境中的蜻蜓目昆虫进行物种鉴定的蜻蜓识别系统。实验结果表明:针对处于复杂生态背景中的蜻蜓目昆虫,当精确到带有性别信息的某一种时,该系统有较高的识别精度;就精确到某一种而言,该系统有很高的识别精度。因此,该系统可以协助蜻蜓研究人员或爱好者进行蜻蜓种类的鉴别。在后续的工作中,将考虑从两个方面进行研究:1) 持续收集新的种类的蜻蜓目昆虫样本以及补充现有种类的样本数量;2) 在保证较高识别率的基础上,使得神经网络更加轻量化,以便于在移动设备中使用。

参考文献:

- [1] 于昕,卜文俊,朱琳.应用蜻蜓目昆虫进行生态环境评价的研究进展[J].生态学杂志,2012,31(6):1585-1590.
YU X,BU W J,ZHU L. Research advances in eco-environment assessment using dragonfly as a bioindicator[J]. Chinese Journal of Ecology,2012,31(6):1585-1590.
- [2] 于昕. 蜻蜓学研究[EB/OL]. [2020-10-11]. <http://www.china-odonata.top>.
YU X. Odonata research[EB/OL]. [2020-10-11]. <http://www.china-odonata.top>.
- [3] 王治国. 中国蜻蜓分类名录(蜻蜓目)[J]. 河南科学,2017,35(1):48-77.
WANG Z G. Systematic dragonflies names of China (odonata)[J]. Henan Science,2017,35(1):48-77.
- [4] 谢娟英,侯琦,史颖欢,等. 蝴蝶种类自动识别研究[J]. 计算机研究与发展,2018,55(8):1609-1618.
XIE J Y,HOU Q,SHI Y H,et al. Automatic recognition of butterfly species[J]. Computer Research and Development,2018,55(8):1609-1618.
- [5] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Washington D. C. : IEEE Computer Society,2014:580-587.
- [6] GIRSHICK R. Fast R-CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision, Piscataway, New Jersey, USA:IEEE,2015:1440-1448.
- [7] REDMON J,FARHADI A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2020-10-11]. <https://arxiv.org/abs/1804.02767>.
- [8] BOCHKOVSKIY A,WANG C Y,LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2020-10-11]. <https://arxiv.org/abs/2004.10934>.
- [9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]//Proceedings of the 2016 European Conference on Computer Vision, Cham,Switzerland;Springer,2016:21-37.
- [10] YANG Z,LUO T,WANG D,et al. Learning to navigate for fine-grained classification[C]//Proceedings of the 2018 European Conference on Computer Vision, Cham,Switzerland;Springer,2018:438-454.
- [11] OŽANA S,BURDA M,HYKEL M,et al. Dragonfly hunter CZ: mobile application for biological species recognition in citizen science[J]. PLoS ONE,2019,14(1):e0210370.
- [12] HE K,ZHANG X,REN S,et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway, USA:IEEE,2016:770-778.
- [13] ZHOU B,KHOSLA A,LAPEDRIZA A, et al. Learning deep features for discriminative localization[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway, USA:IEEE,2016:2921-2929.
- [14] REN S,HE K,GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//Proceedings of 28th Annual Conference on Neural Information Processing Systems. Cambridge, USA: MIT Press,2015:91-99.
- [15] NOH H,HONG S, HAN B, et al. Learning deconvolution network for semantic segmentation[C]//Proceedings of IEEE Conference on International Conference on Computer Vision, Piscataway, USA:IEEE,2015:1520-1528.
- [16] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway, USA:IEEE,2017:1800-1807.

- [17] LIN T, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of IEEE Conference on International Conference on Computer Vision, Piscataway, USA:IEEE, 2017:2999-3007.
- [18] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet large scale visual recognition challenge[EB/OL]. (2014-09-01)[2020-10-11]. <https://arxiv.org/abs/1409.0575v2>.
- [19] HU T, QI H. See better before looking closer; weakly supervised data augmentation network for fine-grained visual classification[EB/OL]. (2019-03-23)[2020-10-11]. <https://arxiv.org/abs/1901.09891>.
- [20] CHEN Y, BAI Y, ZHANG W, et al. Destruction and construction learning for fine-grained image recognition[C]//Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway, USA:IEEE, 2019:5157-5166.
- [21] FU J, ZHENG H, MEI T, et al. Look closer to see better; recurrent attention convolutional neural network for fine-grained image recognition[C]//IEEE Conference on Computer Vision & Pattern Recognition, Piscataway, USA:IEEE, 2017:4476-4484.
- [22] LIN T, ROYCHOWDHURY A, MAJI S, et al. Bilinear CNN models for fine-grained visual recognition[C]//Proceedings of the 2015 IEEE Conference on International Conference on Computer Vision, Piscataway, USA:IEEE, 2015:1449-1457.
- [23] LI Z, YANG Y, LIU X, et al. Dynamic computational time for visual attention[C]//Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway, USA:IEEE, 2017:1199-1209.
- [24] WANG Y, MORARIU V, DAVIS L. Learning a discriminative filter bank within a CNN for fine-grained recognition[C]//Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway, USA:IEEE, 2018:4148-4157.

Improved End-to-End Learning Odonata Recognition Algorithm

PENG Mingjie^{1,2,3}, TANG Wanmei¹, PI Jiatian^{1,2,3}, YU Xin⁴, WU Zhiyou³, HUANG Fangting⁴

(1. College of Computer and Information Sciences, Chongqing Normal University;

2. Chongqing Center of Engineering Technology Research on Digital Agriculture Service, Chongqing Normal University;

3. Chongqing Key Laboratory of Intelligent Finance and Big Data Analysis, Chongqing Normal University;

4. College of Life Sciences, Chongqing Normal University, Chongqing 401331, China)

Abstract: [Purposes] The traditional Odonata insect identification algorithm is to perform artificial feature extraction on specimen pictures to train the classifier. The data set used by this method contains fewer types and numbers of Odonata insects, and the sample background is single, resulting in low recognition rate and can only identify the specimen pictures. In response to these problems, a data set containing both ecological pictures and specimen pictures of Odonata insects was produced, and a recognition algorithm based on deep learning was proposed. [Methods] Using a network framework with end-to-end learning advantages, on the above data set, use ResNet50 migrated and learned on ImageNet for feature extraction, use the new region suggestion network Imp-RPN for feature selection, and then use an improved loss function to solve the problem of uneven sample distribution. [Findings] Experimental results show that the recognition rate of the recognition algorithm reaches 96.2% in 46 classification tasks that are accurate to species, and 88.7% in 90 classification tasks that are accurate to species and contain gender information. [Conclusions] The end-to-end deep learning network framework eliminates the time cost of manually labeling samples, improves the recognition accuracy, and can better meet the needs of species identification tasks.

Keywords: Odonata recognition; fine-grained recognition; deep learning; end-to-end learning; NTS-Net

(责任编辑 许 甲)

(接正文53页)

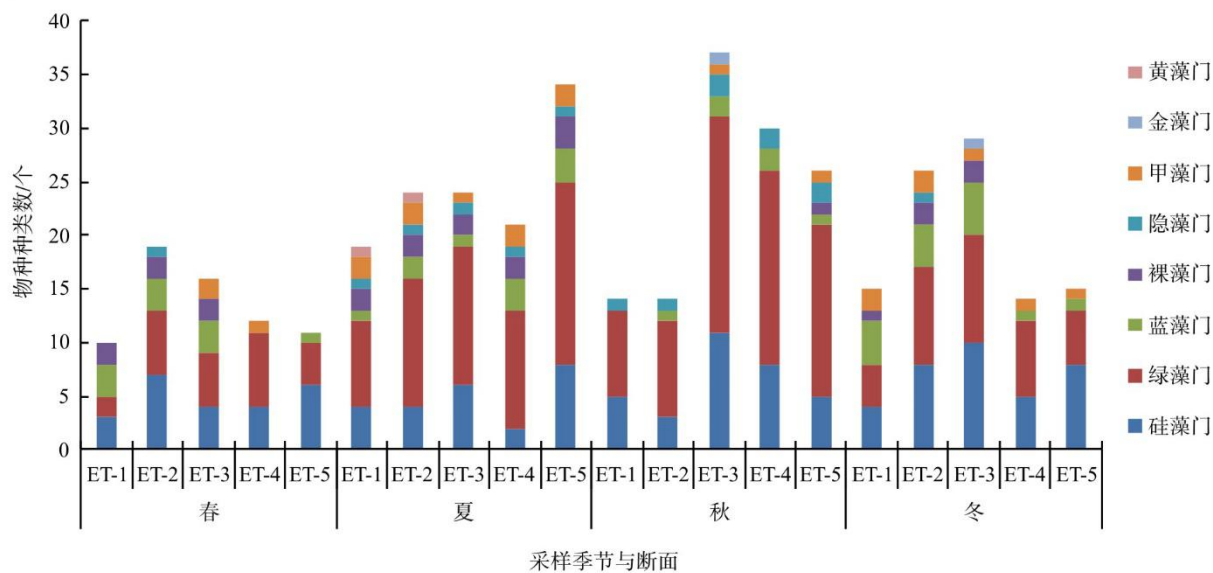


图3 二滩水库浮游植物种类组成
Fig. 3 Species composition of phytoplankton in Ertan reservoir

(接正文95页)



图1 蜻蜓生态照样本
Fig. 1 Samples of natural dragonfly images

图3 雌雄异型示例
Fig. 3 An example of hermaphroditism

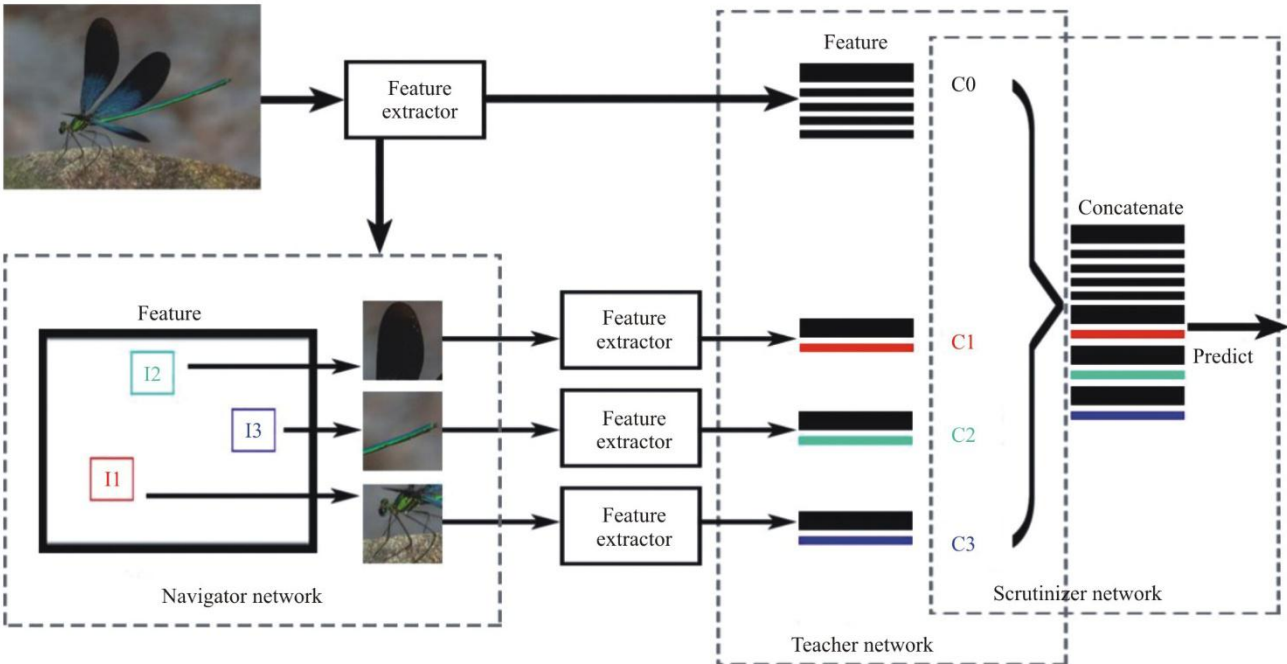


图4 NTS网络设计结构
Fig. 4 Design structure of NTS-Net

(接正文96~97页)

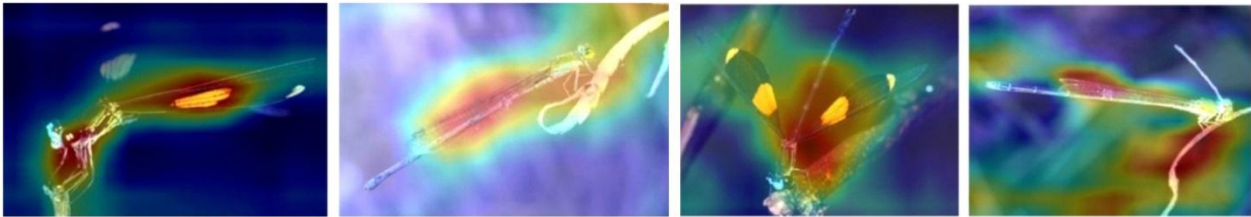


图5 蜻蜓数据集集中的高信息区域
Fig. 5 High information area in the dragonfly dataset

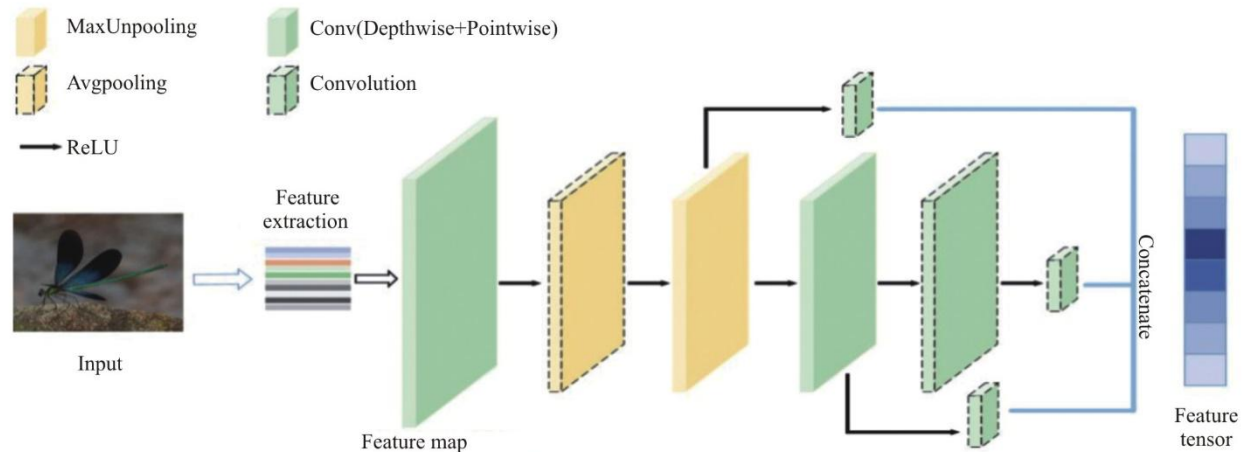


图6 QT-RPN结构图
Fig. 6 QT-RPN structure diagram

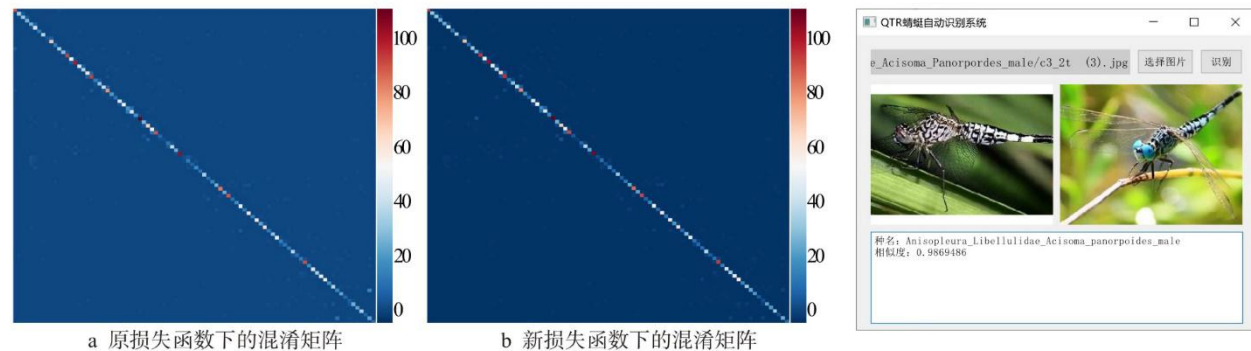


图7 不同损失函数下的测试结果混淆矩阵
Fig. 7 Confusion matrix of test results under different losses

图8 蜻蜓目识别GUI
Fig. 8 Odonata recognition GUI

(接正文104页)

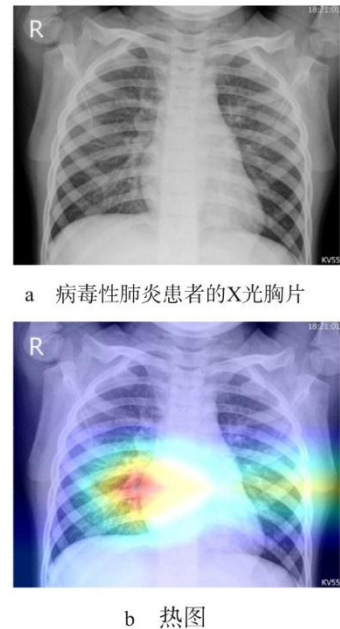


图3 G-Module可视化
Fig. 3 G-Module visualization

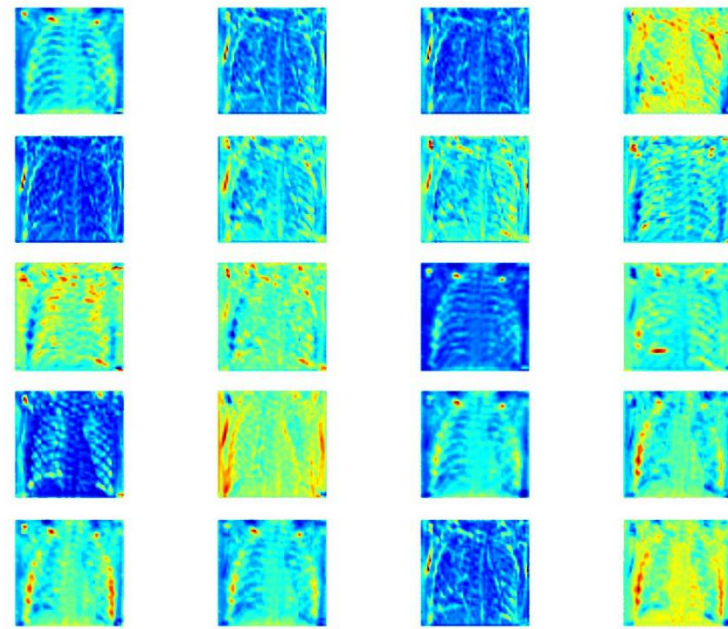


图4 第6个DenseNetLayer后经过1×1卷积后的特征图
Fig. 4 Feature map after 1×1 convolution after the 6th DenseNetLayer