

基于深度学习的鸟声识别技术研究

——以北京翠湖国家城市湿地公园为例

Research on Bird Sound Recognition Technology Based on Deep Learning:
Taking Beijing Cuihu National Urban Wetland Park as an Example

王一宇^{1,2} 夏 舫³ 刘 松³ 彭 涛³ 郭慧敏¹ 鉴海防^{1*}
WANG Yiyu^{1,2} XIA Fang³ LIU Song³ PENG Tao³ GUO Huimin¹ JIAN Haifang^{1*}

(1.中国科学院半导体研究所, 北京 100083; 2.山东师范大学物理与电子科学学院, 济南 250358; 3.北京市海淀区湿地和野生动植物保护管理中心, 北京 100094)

(1. Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China, 100083; 2. School of Physics and Electronics, Shandong Normal University, Jinan, Shandong, China, 250358; 3. Haidian District Wetland and Wildlife Conservation and Management Center, Beijing, China, 100094)

文章编号: 1000-0283(2024)04-0019-08
DOI: 10.12193/j.laing.2024.04.0019.003
中图分类号: TU986
文献标志码: A
收稿日期: 2024-01-09
修回日期: 2024-02-28

摘要

鸟类是生态系统的重要组成部分, 在评估生态系统状态方面具有不可忽视的作用, 因此鸟类的调查和监测对于保护生态环境和维护生物多样性至关重要。传统的鸟类监测手段需要消耗大量人力物力, 并且结果的准确性和可靠性有限。针对这一问题, 提出一种基于深度学习的鸟类声音数据分析识别系统, 并在北京翠湖国家城市湿地公园进行应用示范, 以验证系统的性能和准确性。该系统首先使用传统声音信号处理方法对前端拾音器采集的音频进行预处理, 然后使用残差神经网络提取音频特征并进行分类, 从而实现对目标声音所包含物种信息的自动识别。在系统运行期间, 成功监测到有效鸟类声音片段共计 200 044 条, 其识别准确率达到 93%。系统共识别出野生鸟类 9 目 16 科 52 种, 其中, 属于国家 II 级重点保护野生动物有 6 种, 分别是大天鹅、鸿雁、红角鸮、纵纹腹小鸮、蓝喉歌鸮、游隼; 属于北京市重点保护野生动物共计 22 种, 分别是苍鹭、夜鹭、普通鸬鹚、绿头鸭等。监测到音频片段中相对多度较高的前 5 个物种依次是苍鹭 (26%)、鸿雁 (16%)、夜鹭 (13%)、白头鸭 (11%)、普通鸬鹚 (8%)。实验结果表明, 该系统实现了对鸟类声音的自动采集和分析, 显著提高了鸟类监测的效率, 为风景园林的合理规划、生态节点的保留以及景观可持续性的提升提供了有力支持。

关键词

翠湖国家城市湿地公园; 鸟类监测; 深度学习; 鸟类声音识别; 鸟类多样性; 北京

Abstract

Birds are an important part of the ecosystem and play an indispensable role in assessing the state of the ecosystem, so the survey and monitoring of birds are crucial for protecting the ecosystem and preserving biodiversity. Traditional means of bird monitoring consume a lot of workforce and resources, and the accuracy and reliability of the results could be improved. To address this problem, a deep learning-based bird sound data analysis and recognition system is proposed, and an application demonstration is carried out in the Cuihu National Urban Wetland Park in Beijing to verify the performance and accuracy of the system. Using traditional sound signal processing methods, the system first preprocesses the audio captured by the front-end pickups. Then, it extracts audio features and classifies them using residual neural networks to achieve automatic identification of the species' information contained in the target sounds. During the system's operation, 200 044 effective bird sound clips were successfully monitored, and its recognition accuracy reached 93%. The system identified a total of 52 species of wild birds in 9 orders and 16 families, among which there are 6 species of national class II wildlife under key protection, namely, *Cygnus cygnus*, *Anser cygnoides*, *Otus sunia*, *Athene noctua*, *Luscinia svecica*, and *Falco peregrinus*; and a total of 22 species of Beijing's key wildlife under key protection, namely, *Ardea cinerea*, *Nycticorax nycticorax*, *Phalacrocorax carbo*, and *Anas platyrhynchos*, etc. The system was able to automatically identify the species of wild birds in 9 orders and 16 families. The top 5 species with a higher relative plurality in the monitored audio clips are *Ardea cinerea* (26%), *Anser cygnoides* (16%), *Nycticorax nycticorax* (13%), *Pycnonotus sinensis* (11%), and *Phalacrocorax carbo* (8%), in that order. The experimental results show that the system achieves automatic collection and analysis of bird sounds, significantly improves the efficiency of bird monitoring, and provides strong support for the rational planning of landscape gardens, the retention of ecological nodes, and the enhancement of landscape sustainability.

Keywords

Cuihu National Urban Wetland Park; bird monitoring; deep learning; bird sound recognition; bird diversity; Beijing

王一宇

2000年生/女/内蒙古包头人/在读硕士研究生/研究方向为智能信号处理

鉴海防

1978年生/男/山东东营人/博士/研究员、博士生导师/研究方向为高性能专用集成电路设计、智能信息处理算法与系统

*通信作者 (Author for correspondence)
E-mail: jhf@semi.ac.cn

在城市的绿化区域中,鸟类的群落结构与植被类型的多样性紧密相连。由于各种鸟类对生活环境的需求存在差异,因此它们对栖息地的要求也不尽相同。茂密的树冠和较为安静的环境为鸟类提供了理想的栖息地,而低矮的灌木丛则为它们提供了获取食物的场所^[1]。在国内,园林作为城市中较少展现森林景观的地方,在保护城市鸟类和维护城市绿地系统方面扮演着关键角色。园林绿地植被不仅增加了城市绿化覆盖率,还为鸟类提供了重要的生态环境空间。

然而,随着中国城市化快速推进,土地利用形式和性质发生了急剧变化,导致鸟类原始栖息地环境遭到破坏,城市生态环境逐渐恶化。尽管城市规划结构为人类提供了方便,其却对鸟类的稳定性和物种多样性带来了不利影响。城市快速发展导致短期内生态环境的剧变,给鸟类带来了更大的环境适应压力,使它们面临更严峻的生存挑战^[1]。在城市中,鸟类以城市公园绿地为主要栖息地,它们在这里直接或间接地获取食物。《园冶》强调了鸟类在构成中国园林中的重要作用,从中国古典园林美学的角度来看,动物景观作为越发被重视的组成部分,声音美被提到了一个重要层面,鸟类发出的鸣声正满足了这一层面的需求^[2]。通过鸟类监测,研究人员可以评估园林修复计划对多样性鸟类群落的吸引力,从而提高生态系统中物种的多样性;通过了解鸟类在不同环境中的栖息偏好,可以更有效地设计和改进园林修复计划,创造更加适宜的生态环境;同时鸟类对环境变化十分敏感,它们的存在和数量可作为环境健康的指标,通过监测鸟类的活动,能更好地了解环境中潜在问题,如污染和栖息地破坏,为环境管理提供有力的数据支持^[3]。

目前,对于园区鸟类的保护监测,通常

采用传统的人工监测手段,多数鸟类在日间活动,可通过望远镜直接观察,辨识种类和技术;习性较为隐秘的鸟类可通过声音侦测并记录,鸟类通常在繁殖前期和每天日出前会发出复杂的鸣声,研究人员可在这些时段辨别种类;一些外观相似,鸣声差异大的鸟类也可以通过声音轻松识别;还有一些鸟类可以通过飞行姿态和觅食行为来鉴定^[4]。目前,常规检测中常用的方法是样线法,调查人员沿着一段样线匀速前进时记录见到或听到的沿线两侧的鸟类及其据样线的垂直距离然后通过距离取样原理,记录该生境中鸟类的密度^[5]。由于有的地形较为崎岖,可采用样点法,将样线换成样点,调查人员依次到达事先设定好的样点,记录一段时间内鸟类种类和距离以及数量^[6]。但是这些方法不仅需要大量的人力物力,而且需要具备专业知识的鸟类专家才能进行辨别^[4],难以实现高效的鸟类监测。随着观鸟活动的兴起,公民科学也成为了重要的鸟类监测方式,但是这种方式取决于所在地区的到达难易程度和当地观鸟者的数量。同时观鸟者倾向于记录更为稀有的鸟类,对常年的鸟类记录不准确,容易导致偏差。与此同时,保护园区的视频监控系统主要面向安防监控需求,监控数据每隔一段时间就会被覆盖,导致无法进行有效数据的存放与分析。

近年来,随着以深度学习为代表的人工智能技术的不断发展,基于深度学习的声音识别技术取得了显著的进步,并在语音处理^[7]、语音识别^[8,9]等领域得到了广泛应用。受此启发,一些研究者尝试将该类技术应用在鸟类声音识别任务中^[10],通过特征提取和分类训练实现鸟类物种信息的自动识别(图1)。2021年,康奈尔大学鸟类学实验室通过开发一种名为BirdNET的深度神经网络,能够识别出北美和欧

洲的大部分鸟类^[11];2022年,广州大学电子与通信工程学院的团队在广州市白云山公园尝试运用无监督的音节聚类分析提取鸟鸣音节并对其进行推断分类,实现了对鸟类声音的识别^[12]。在鸟类识别软硬件系统方面,科研人员、研究机构也同样开展了许多研究。例如,Hüppo等^[13]通过布设自动录音装置录制不同天气条件下的鸟声音频数据,能够分析得出鸟类迁徙活动与天气之间的关系,进而评估风力发电厂对欧洲候鸟种群的潜在影响;2016年,加拿大自然资源部的科学家基于计算机视觉技术和机器学习算法开发了鸟类视频监测系统,其能够实现鸟类巢穴和繁殖情况的有效监测;2019年,西南林业大学基于人工智能技术构建了云南野生鸟类自动图像识别系统,以识别出云南省全境鸟类并给出该种类个体对应的目科属等相关信息^[14]。综上所述,目前常见的野外声音监测设备在野外放置一段时间进行录音后带回交由专家分析,无法实时录音识别。现有的野外实时监测设备,能够进行鸟类自动识别的软硬件系统大多基于图像数据进行分析。这类系统通常需要足够的光照条件才能正常工作,阴天、黑夜或恶劣的天气条件下会影响图像的质量和稳定性,难以实现全天候的鸟类监测。

针对以上问题,本研究基于深度学习和信号处理技术,构建鸟类声音识别系统,对采集到的声音数据进行实时自动识别,不仅能够高效获取不同时间段鸟类多样性的变化情况,还能够自动识别记录该保护区出现的新鸟类物种,可以为风景园林的景观规划提供重要信息。

1 研究方法

1.1 研究区概况

北京翠湖国家城市湿地公园(以下简称

北京翠湖) 位于北京市海淀区上庄镇上庄水库北岸, 坐标为40° 5' 36" S-40° 06' 39" N, 116° 10' 27" W-116° 11' 28" E, 是北京最大或也是仅有的一块自然湿地^[15]。北京翠湖植被茂盛, 鸟类物种丰富, 在春秋迁徙季是鸟类重要补给地和栖息地。到目前为止, 在北京翠湖共观察到野生鸟类21目64科281种中, 属于国家级重点保护野生动物(2021)共63种, 其中I级保护鸟类13种, II级保护鸟类50种; 属于北京市保护野生动物(2021)共112种, 其中I级保护鸟类17种, II级保护鸟类95种^[16]。

1.2 平台简介

鸟类声音识别系统由北京翠湖与中国科学院半导体研究所合作建设, 监测对象为北京翠湖候鸟。系统架构包括设备端、数据终端和数据展示端。具体工作流程见图2, 前端拾音器设备负责周围声音数据的实时采集; 后端边缘设备通过内置的声音信号端点检测算法与鸟类声音识别算法完成对采集声音片段的自动分析处理; 在获取到检测结果后, 通过无线网络传输模块将识别结果与对应的音频数据传输到数据终端进行可视化展示。

系统从数据终端获取到监测结果后, 会自动在监测界面(图3)中设备安装点位弹出相应的鸟类图像并播放声音。与此同时, 当天监测到的所有鸟类物种信息也会在右下角的监测记录中滚动播放, 便于管理者及时了解园内的鸟类信息。

1.3 硬件系统架构

(1) 声音拾取设备(拾音器)。本系统采用的声音拾取设备为烽火通信科技股份有限公司生产的烽火FH-300 N室外防水防爆烽火

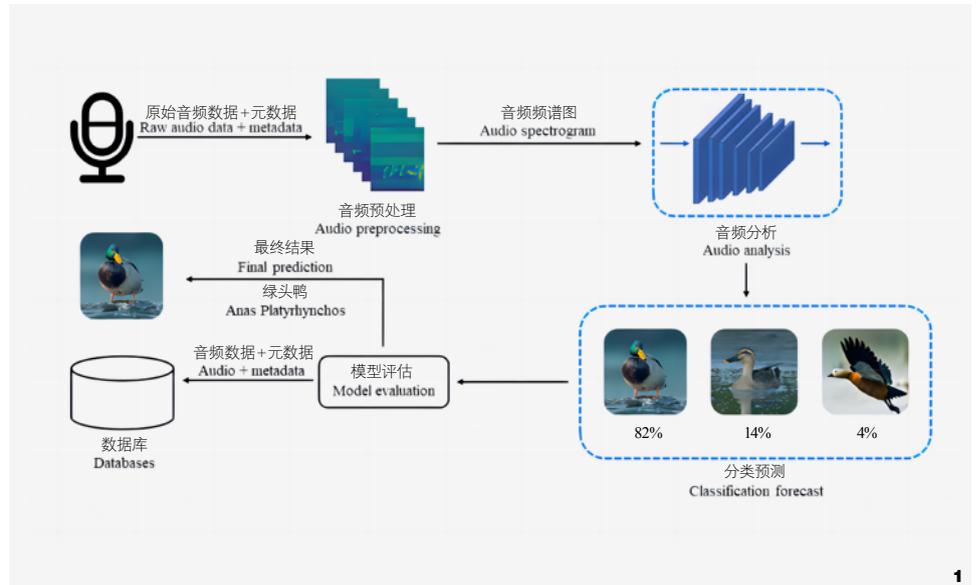


图1 鸟类声音识别流程图
Fig. 1 Flowchart of bird voice recognition model

图3 鸟类声音识别系统监测界面
Fig. 3 Monitoring interface of intelligent bird sound recognition system

图2 鸟类声音识别系统流程图
Fig. 2 Flowchart of bird sound recognition system

拾音器。该拾音器的采样率为36 KHz, 指向性为全向, 灵敏度为-36±3 dBV, 最大承受声压为128 dB SPL, 信噪比为65 dB。其同时具有噪声抑制功能, 能够实现保护区内声音数据的实时采集及噪声抑制。

(2) 鸟类声音识别边缘设备。鸟类声音识别边缘设备是一种集成鸟声识别算法与计算模块的嵌入式系统设备, 能够在设备本地进行声音信号的处理和分析, 实现对鸟类声音的不间断监测与实时识别, 并通过宽带网络或4G/5G无线网络将声音样本和识别结果按照不同的主题(Topic)通过分布式消息队列(Kafka)发送到数据终端平台。

(3) 数据终端平台。数据终端平台从Kafka接收音频数据和对应识别结果之后, 将音频保存在本地, 并将音频位置和识别结果写入数据库。

(4) 开发用户界面(UI)。通过UI, 在有新数据存入数据库之后, 自动进行数据获取, 并在监测界面中设备安装点位弹出相应的鸟类图像并播放声音。

1.4 软件系统构建

1.4.1 信号预处理

鸟类声音识别设备采集到声音数据后, 由于原始声音信号中含有大量的环境噪声和静音片段, 需要对其进行预处理以提高识别效果。为此, 可以使用声音信号降噪算法和声音信号端点检测算法, 对声音信号进行分析和处理。具体为, 首先通过声音信号降噪算法去除环境噪声对声音信号的影响, 以获得更清晰的声音信号; 然后, 通过声音信号端点检测算法, 去除无效的静音片段, 以提取有效的声音片段。

由于鸟类声音信号中高频成分较少, 且高频信号的衰减率比低频高, 因此在提取特

征前要进行预加重, 补偿高频分量在传输中衰减的部分, 以增强声音信号中高频部分的分辨率。在此, 运用一阶有限长单位冲激响应高通数字滤波器(HR)对信号进行预加重^[17], 具体见公式(1)。

$$y(n) = x(n) - \alpha x(n-1) \quad (1)$$

式中, $x(n)$ 为输入信号, n 为时间点, $y(n)$ 为预加重后的输出信号。 α 为预加重系数, 一般取值范围是 $0.9 < \alpha < 1$, 在本研究中, 将其设置为0.97。

鸟声信号多为非平稳信号^[18], 因此, 在对长序列声音信号做傅里叶变换时, 需要先通过窗函数 $h(n)$ 对原始信号进行分帧, 以确保局部信号的平稳性; 再逐段进行傅里叶变换得到信号的局部“频谱”; 最后, 将多个局部“频谱”按时间维度进行堆叠, 生成完整的声音谱图, 相应的短时傅里叶变换见公式(2)。式中, $STFT(t, \omega)$ 表示在时刻 t 和频率 ω 处的短时傅里叶变换结果。

$$STFT(t, \omega) = \sum_{n=-\infty}^{\infty} x(n) h(n-t) e^{-j\omega n} \quad (2)$$

1.4.2 梅尔滤波器

根据接收信号频率的不同, 人耳耳蜗在进行声音信号获取时, 对不同频率的感知能力是非线性的, 通常对低频信号敏感, 对高频信号不敏感。因此, 在进行声音信号的采集时, 需要考虑到人耳的感知特征, 为了模拟人耳这种效应, 本研究将使用梅尔滤波器对每帧信号进行滤波, 通过梅尔滤波器组的输出公式见(3)。

$$Y_t(m) = \sum_{k=1}^N H_m(k) |X_t(k)|^2 \quad (3)$$

式中, k 表示信号, $x(t)$ 经过FFT变换后的编号, m 表示梅尔滤波器的编号。其中, 梅尔滤波器组 $H_m(k)$ 是一系列等高的三角形滤波

器, 具体见公式(4)。式中, $f(m)$ 为中心频率。

$$H_m(k) = \begin{cases} 0 & k < f(m-1) \\ \frac{k-f(m-1)}{f(m)-f(m-1)} f(m-1) & f(m-1) \leq k < f(m) \\ 1 & k = f(m) \\ \frac{f(m+1)-k}{f(m+1)-f(m)} f(m) & f(m) < k \leq f(m+1) \\ 0 & k > f(m+1) \end{cases} \quad (4)$$

1.4.3 网络结构

卷积神经网络(CNN)^[19]是实现声音信号特征提取及分类的主流方式之一, 主要由卷积层、池化层、全连接层等构成, 例如VGGNet^[20]、GoogLeNet^[21]、ResNet^[22]。VGGNet网络的结构非常简洁, 整个网络都使用了同样大小的卷积核尺寸(3×3)和最大池化尺寸(2×2), 其中VGG19比VGG16增加了三层卷积核, 但由于缺乏残差连接块, 导致模型在训练和推理过程中需要更多的计算资源。GoogLeNet由多个Inception模块堆叠而成, 这些模块在同一层级上使用不同大小的卷积核并行运行, 然后将结果合并, 以提高特征表征的丰富度和计算效率。相较于同等深度的VGGNet, GoogLeNet使用了更少的参数。然而, 随着模型深度的不断增加, 仍然会在训练过程中遇到梯度消失的问题。针对深层模型难以训练的问题, ResNet引入了残差连接(residual connection)的概念。这种连接方式允许网络学习残差, 即目标函数与输入的差异, 而不是学习整个函数。通过将前一层的输出直接与后续层的输入相加, ResNet解决了深度网络训练中的梯度消失或爆炸的问题。这种设计使得网络更易于训练, 可以训练更深的网络, 从而在图像识别等任务中取得了显著的性能提升。

表1为不同网络模型在ImageNet验证集上的错误率^[22]。由表1可知, ResNet相较于其他两种常用算法错误率最低, 考虑到音频分类

任务与图像分类任务类似，因此本系统也同样采用图4中的ResNet50结构提取音频特征并分类。该ResNet50网络包括48个卷积层、1个最大池化层和1个平均池化层，该网络将常规的残差模块替换为瓶颈结构 (bottleneck)，在保证识别精度的同时，还大幅地减少了网络参数量。

2 实验设置

2.1 数据集组成

本研究前期通过Xeno^①进行鸟类声音样本的收集，并基于这些样本建立一个用于训练鸟类声音识别模型的先验数据集，具体来说，本研究根据北京翠湖鸟类名录从Xeno上获得相应的音频数据，对于每条数据，按质量分数进行筛选，共提取出6 200条音频。在获得原始音频后，通过运用双门限法的声音端点检测算法^[23]滤除每个音频信号中的静音片段。在Kahl等^[11]的研究中提到大量物种的发声平均时长为1.94 s，因此每个音频片段被剪辑成3 s。本研究也采取同样的剪辑的方式，使用3 s的剪辑窗口将处理后的音频分割成块，共得到17 056个鸟声剪辑片段，最后这些片段以7 : 1 : 2的比例分为训练集、验证集和测试集，形成一个完整的鸟声数据集。

2.2 监测点位布置

考虑到靠近水边的树木、灌丛、湿地等区域是鸟类繁殖、觅食和栖息的主要场所，通常会有更多种类的鸟类活动，且湖边环境开阔可以减少噪声的干扰。因此，本研究选择在北京翠湖荷花塘附近布设一个鸟类音频监测点位 (图5)，该地点地处水域边缘，环

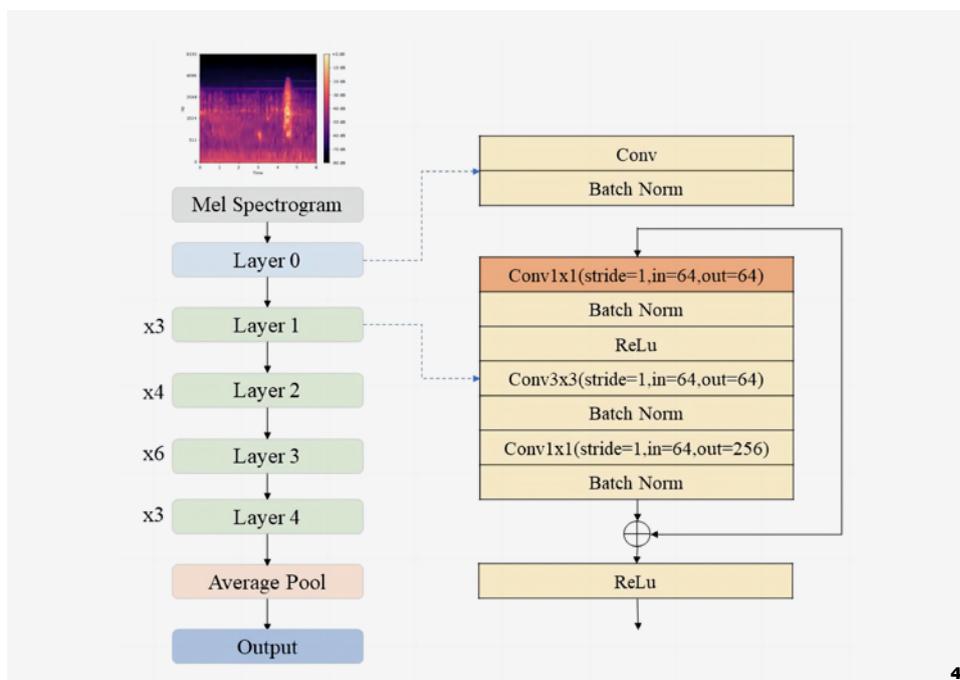


图4 Resnet50网络结构示例
Fig. 4 Example network architectures for Resnet50

境优美，拥有丰富的树木和植被，为鸟类提供了丰富的栖息和觅食资源。

2.3 研究结果

对于部署在北京翠湖的监测设备，拾音器捕获到鸟声信号后，声音数据经过声音降噪和端点检测算法去除噪声和静音片段，经过分帧加窗后使用梅尔滤波器提取出鸟声的梅尔频谱图，原始波形图和梅尔声谱图如图6所示。

然后，梅尔声谱图输入到已经加载预训练模型参数的鸟声识别模型，鸟声经过模型选取预测概率最大的鸟类作为最后的预测结果。

部署在北京翠湖荷花塘的音频监测系统自2023年2月持续运行至12月，累计工作10个月，共获得209 686条音频数据。这些数据完整地记录了北京区域不同季节鸟类的种类以及叫声频次变化，其中，鸟类鸣叫片段200 044条，人类干扰声音3 741条，其他环境噪声5

表1 ImageNet 验证集上单一模型结果的错误率 (单位: %)
Tab. 1 Error rates (%) of single-model results on the ImageNet validation set

网络模型 Network model	Top-3错误率 top-3 err	Top-5错误率 top-5 err
VGG	24.4	7.1
GoogLeNet	—	7.89
ResNet50	20.74	5.25

① Xeno 网址为 <https://xeno-canto.org/>。



图5 翠湖荷花塘音频监测点位
Fig. 5 Audio monitoring point of lotus pond in Cuihu Lake

123条。通过专家的校验，系统识别鸟声准确率为93%。在识别的鸟类鸣声片段中，共涵盖了9目16科52种鸟类。图7为2023年2-12月的声音频次统计图，其中，相对多度较高的前5个物种为苍鹭(26%)、鸿雁(16%)、夜鹭(13%)、白头鹤(11%)和普通鸬鹚(8%)；属于国家Ⅱ级重点保护野生动物有6种，分别是大天鹅、鸿雁、红角鸮、纵纹腹小鸮、蓝喉歌鸲和游隼，北京市重点保护野生动物苍鹭、夜鹭、普通鸬鹚、绿头鸭等22种，详见表2。

3 讨论

本研究通过采用鸟类声音识别系统，实现了北京翠湖鸟类多样性的自主监测，相较于传统方法，采用该系统不仅能够实现鸟类声音数据的长期获取及自动分析，还能够形成结构化的识别数据，便于鸟类的监测和保护，为北京翠湖鸟类种群和生态多样性的研究提供了丰富的鸟类声音数据支撑。

在本系统的监测下，共记录到52种野生鸟类，其中包括6种国家Ⅱ级重点保护野

生动物，通过与专家校验对比，研究发现本系统的识别准确率达到93%。此外，还基于这些实采数据对比了不同网络模型的性能。从表3可以看出，本研究所采用的ResNet50模型相较于VGG16和VGG19模型，在准确率上分别提升了1%和0.8%。同时，其参数量分别减少了112.5 MB和118.5 MB。这一优势极大地降低了模型在算力受限的边缘设备上的部署难度，并且提升了模型的推理效率。这种显著的改进主要源于ResNet50采用的残差结构。这种结构使得网络更深，同时减少了梯度消失的问题，从而提高了模型的准确率。此外，ResNet50的卷积层中使用了1x1的卷积核，这一设计有效地减少了模型的计算量和参数量。

本研究所提出的鸟类声音识别系统与传统方法相比有效提升了鸟类监测的效率，并实现了较高的识别精度。但性别、生长阶段、栖息环境和繁殖状况等因素会对鸟类的发声产生影响，从而影响系统的识别性能。此外，鸟类在繁殖、交流、寻找食物和保护领土时表现出丰富多样的声音，这种声音不仅是生物

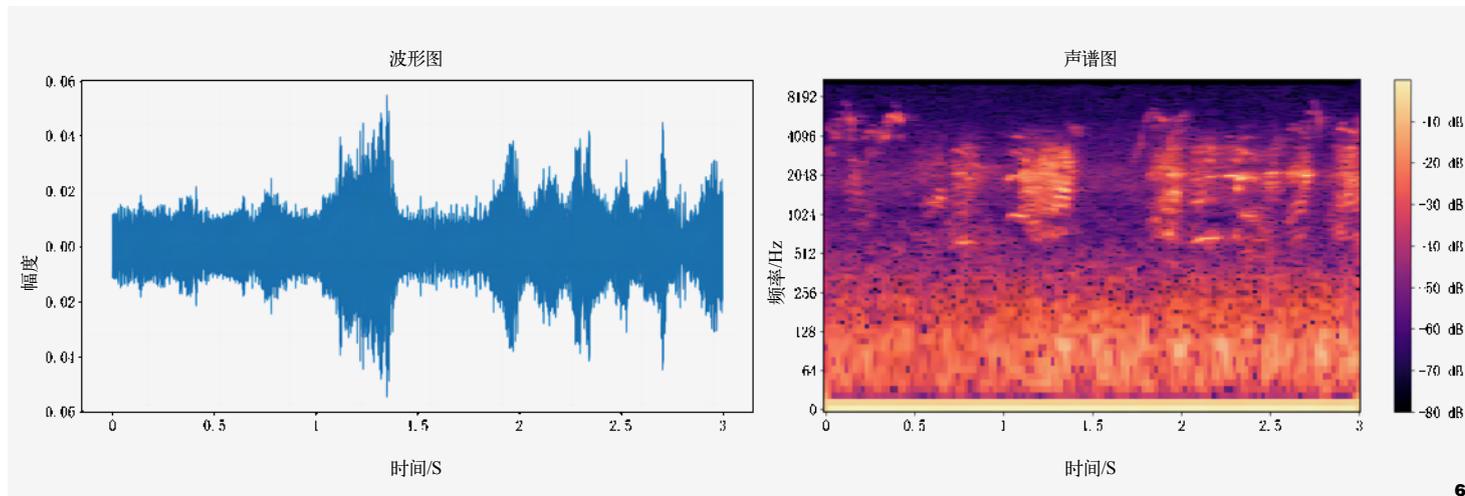


图6 鸿雁叫声波形图和声谱图
Fig. 6 Waveforms and spectrograms of honkers sounds

学特征，更是风景园林的声音调色板。因此，对鸟类声音与不同行为之间的深入分析对于理解园林生态的演变具有非常重要的意义。

后续研究将通过收集处于不同生长阶段、位于不同地理位置以及不同性别的鸟类数据，建立一个更加全面的鸟类声音数据库，为系统识别性能的提升和鸟类多样性的研究提供更为有利的数据支撑；同时，基于深度学习技术，建立鸟类声音与具体行为之间的关联，以实现对于鸟类个体行为的准确检测，从而为风景园林规划和管理提供更深层次的生态学信息。这不仅有助于提高园林的生物多样性，同时也为打造更具生态平衡和可持续性的园林环境提供了科学支持。

4 结语

鸟类作为生态系统中不可或缺的组成部分，城市生态的完善程度很大程度上可通过鸟类的种类和数量反映。在城市园林修复中，鸟类监测起到了不可替代的重要作用。通过在典型生态区域中部署鸟类声音识别系统，能够实现鸟类信息的自动监测分析，并为研究人员提供详实的统计报告和监测数据，有助于深入了解它们在城市环境中的生存状况。这些监测结果不仅是城市园林修复的科学依据，更是保护生物多样性、维护生态平衡的有力工具。此外，在城市园林修复过程中，鸟类监测可以帮助评估修复效果，指导相关工作的调整和改进，以确保修复项目的顺利实施和生态效益的最大化。只有制定出更加精准、可持续的修复策略，才能为城市创造出更宜人、更生态友好的绿色空间。因此，深刻认识并充分利用鸟类监测的重要性，将为城市园林的未来发展注入新的活力，实现城市与自然的和谐共生。

注：文中图表均由作者绘制。

表2 翠湖鸟类声音监测名录
Tab. 2 Bird sound monitoring list in Cuihu Lake

序号 No.	中文名 Chinese name	学名 Scientific name	保护级别 (国家/市) Protection level (Nation/City)	频次总计 Total frequency
1	苍鹭	<i>Ardea cinerea</i>	—/二	53 106
2	鸿雁	<i>Anser cygnoides</i>	二/—	32 259
3	夜鹭	<i>Nycticorax nycticorax</i>	—/二	26 823
4	白头鹮	<i>Pycnonotus sinensis</i>	—/二	22 341
5	普通鸬鹚	<i>Phalacrocorax carbo</i>	—/二	16 067
6	绿头鸭	<i>Anas platyrhynchos</i>	—/二	13 880
7	喜鹊	<i>Pica pica</i>	—/—	13 357
8	赤膀鸭	<i>Mareca strepera</i>	—/二	4 318
9	灰喜鹊	<i>Cyanopica cyanus</i>	—/二	3 912
10	棕头鸦雀	<i>Sinosuthora webbiana</i>	—/二	3 129
11	黑翅长脚鹬	<i>Himantopus himantopus</i>	—/二	1 446
12	金眶鸻	<i>Charadrius dubius</i>	—/—	1 211
13	四声杜鹃	<i>Cuculus micropterus</i>	—/二	896
14	北红尾鸲	<i>Phoenicurus auroreus</i>	—/—	895
15	蓝喉歌鸲	<i>Luscinia svecica</i>	二/—	792
16	游隼	<i>Falco peregrinus</i>	二/—	549
17	达乌里寒鸦	<i>Coloeus dauuricus</i>	—/—	435
18	秃鼻乌鸦	<i>Corvus frugilegus</i>	—/—	390
19	黄眉柳莺	<i>Phylloscopus inornatus</i>	—/二	367
20	灰雁	<i>Anser anser</i>	—/二	364
21	红角鸮	<i>Otus sunia</i>	二/—	335
22	山斑鸠	<i>Streptopelia orientalis</i>	—/—	325
23	大嘴乌鸦	<i>Corvus macrorhynchos</i>	—/—	324
24	沼泽山雀	<i>Poecile palustris</i>	—/二	272
25	白额燕鸥	<i>Sternula albifrons</i>	—/—	270
26	铁爪鹀	<i>Calcarius lapponicus</i>	—/—	216
27	灰头绿啄木鸟	<i>Picus canus</i>	—/—	208
28	斑嘴鸭	<i>Anas zonorhyncha</i>	—/二	182
29	绿翅鸭	<i>Anas crecca</i>	—/二	175
30	极北柳莺	<i>Phylloscopus borealis</i>	—/二	121
31	金翅雀	<i>Chloris sinica</i>	—/二	108
32	黄腹鹪	<i>Anthus rubescens</i>	—/—	98
33	大天鹅	<i>Cygnus cygnus</i>	二/—	95
34	红喉姬鹀	<i>Ficedula albicilla</i>	—/二	82
35	小嘴乌鸦	<i>Corvus corone</i>	—/—	71
36	树鹀	<i>Anthus hodgsoni</i>	—/—	57
37	灰斑鸠	<i>Streptopelia decaocto</i>	—/—	53
38	银喉长尾山雀	<i>Aegithalos glaucogularis</i>	—/二	52
39	灰鹡鸰	<i>Motacilla cinerea</i>	—/—	51
40	纵纹腹小鸮	<i>Athene noctua</i>	二/—	50
41	白腰草鹀	<i>Tringa ochropus</i>	—/—	44
42	叽喳柳莺	<i>Phylloscopus collybita</i>	—/—	43
43	星头啄木鸟	<i>Yungipicus canicapillus</i>	—/—	42
44	褐柳莺	<i>Phylloscopus fuscatus</i>	—/二	38
45	淡眉柳莺	<i>Phylloscopus humei</i>	—/—	37
46	普通翠鸟	<i>Alcedo atthis</i>	—/—	28
47	赤麻鸭	<i>Tadorna ferruginea</i>	—/二	26
48	红胁蓝尾鸲	<i>Tarsiger cyanurus</i>	—/—	26
49	灰背鸲	<i>Turdus hortulorum</i>	—/—	23
50	黄雀	<i>Spinus spinus</i>	—/二	21
51	大斑啄木鸟	<i>Dendrocopos major</i>	—/—	17
52	褐头山雀	<i>Poecile montanus</i>	—/二	17

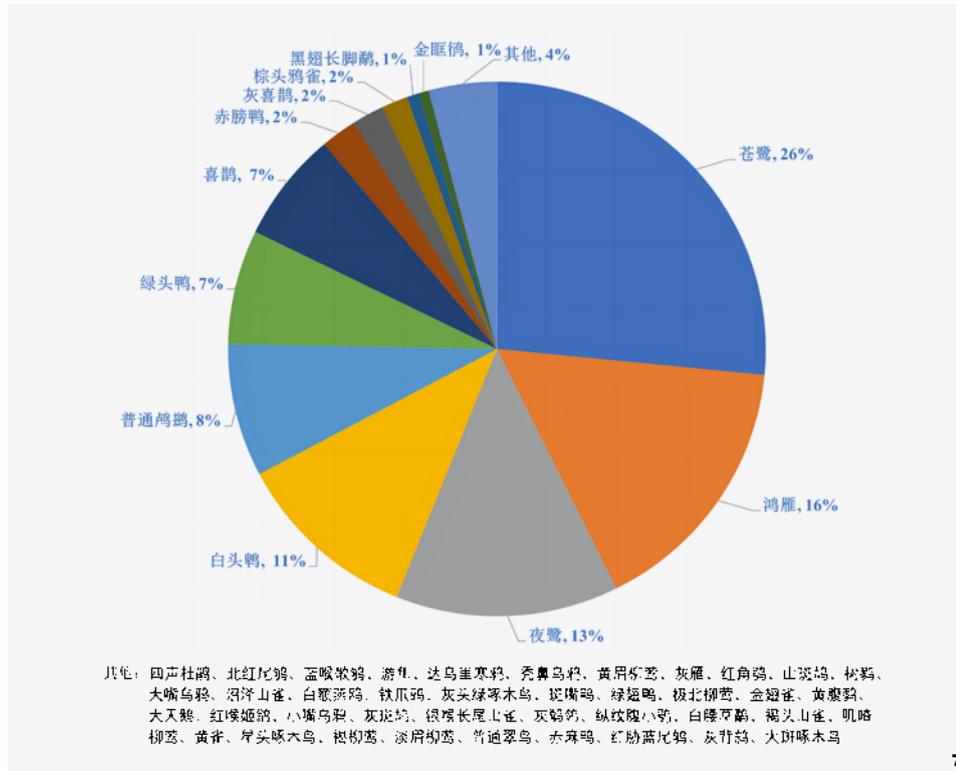


图7 翠湖鸟类物种声音频次统计图 (2023.02-2023.12)
Fig. 7 Statistical chart of bird species sound frequency in Cuihu Lake (2023.02-2023.12)

表3 ResNet50与目前已有算法比较
Tab. 3 Comparison between ResNet50 and existing algorithms

网络模型 Network model	参数量/MB Parameter quantity	准确率/% Accuracy
VGG16	138	92
VGG19	144	92.2
ResNet50	25.5	93

参考文献

[1] 李晓, 王婷婷, 马敬民, 等. 城市园林引鸟设计的原则与措施[J]. 绿色科技, 2016(03): 67-68.
[2] 袁晓梅, 吴硕贤. 中国古典园林声景观的三重境界[C]// 第十届全国建筑物理学术会议论文集. 广州: 华南理工大学出版社, 2008.
[3] 魏筱. 浅谈城市园林绿化对鸟类栖息环境的影响和对策[J]. 科技中国, 2015.
[4] 孙戈, 曾立雄, 钱法文, 等. 鸟类监测技术现状与发展趋势[J]. 地理信息世界, 2022, 29(6): 26-29.
[5] BARRY S C, WELSH A H. Distance Sampling Methodology[J]. Journal of the Royal Statistical

Society: Series B (Statistical Methodology), 2001, 63(1): 23-31.
[6] ZURITA G, PE' ER G, BELLOCQ M I, et al. Edge Effects and Their Influence on Habitat Suitability Calculations: A Continuous Approach Applied to Birds of the Atlantic Forest[J]. Journal of Applied Ecology, 2012, 49(2): 503-512.
[7] PURWINS H, LI B, VIRTANEN T, et al. Deep Learning for Audio Signal Processing[J]. IEEE Journal of Selected Topics in Signal Processing, 2019: 1.
[8] CHACHADA S, KUO C C J. Environmental Sound Recognition: A Survey[J]. APSIPA

Transactions on Signal and Information Processing, 2014, 3: e14.
[9] DESPLANQUES B, THIENPOND T J, DEMUYNCK K. ECAPA-TDNN: Emphasized Channel Attention, Propagation and Aggregation in TDNN Based Speaker Verification[J]. Znterspeech, 2020: 25-29.
[10] 张晓伟, 王建凯, 郇希墨. 人工智能在鸟类保护系统中的应用[J]. 湿地科学与管理, 2021, 17(2): 61-64.
[11] KAHL, STEFAN W, CONNOR M E, et al. BirdNET: A Deep Learning Solution for Avian Diversity Monitoring[J]. Ecological Informatics: An International Journal on Ecoinformatics and Computational Ecology, 2021, 61(1): 10.
[12] 吴科毅, 阮文达, 周隽锋, 等. 基于音节聚类分析的被动声学监测技术及其在鸟类监测中的应用[J]. 生物多样性, 2023, 31(1): 126-136.
[13] HÜPPOP O, WINKEL W. Climate Change and Timing of Spring Migration in the Long-Distance Migrant Ficedula Hypoleuca in Central Europe: The Role of Spatially Different Temperature Changes Along Migration Routes[J]. Journal of Ornithology, 2006, 147: 344-353.
[14] 赵毅力, 李禹成, 陈皓. 云南野生鸟类图像自动识别系统[J]. 计算机应用研究, 2020, 37(S01): 423-425.
[15] 闫亮亮, 王博宇, 于方, 等. 翠湖国家城市湿地公园鸟类环志研究[J]. 湿地科学与管理, 2021, 17(2): 27-30.
[16] 国家林业和草原局, 农业农村部. 国家重点保护野生动物名录[J]. 野生动物学报, 2021, 42(2): 605-640.
[17] 樊昌信. 通信原理(第5版)[M]. 北京: 国防工业出版社, 2001.
[18] 吕坤朋, 孙斌, 赵玉晓. 基于鸟鸣声及深度学习的鸟类识别方法研究[J]. 科技通报, 2021, 37(10): 24-30.
[19] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet Classification with Deep Convolutional Neural Networks[J]. Advances in Neural Information Processing Systems, 2017(06): 84-90.
[20] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014(09): 1556.
[21] SZEGEDY C, LIU W, JIA Y, et al. Going Deeper with Convolutions[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1-9.
[22] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 770-778.
[23] LI X, LI G, LI X. Improved Voice Activity Detection Based on Iterative Spectral Subtraction and Double Thresholds for CVR[J]. IEEE Computer Society, 2008(08):2-3.