

融合深度学习与无人机图像采集的野生鸟类 细粒度识别与监测方法研究

梁宇祥*, 李维平, 张明明, 翟学超[#]

黑龙江省野生动物研究所, 黑龙江 哈尔滨

收稿日期: 2025年7月29日; 录用日期: 2025年8月25日; 发布日期: 2025年9月22日

摘要

野生鸟类的细粒度识别与监测对于生态研究和保护工作具有重要意义, 但传统的识别方法在复杂自然环境中面临着泛化能力有限等挑战。为提高野生鸟类细粒度识别的准确性和实时性, 研究结合无人机图像采集技术构建并优化了一种基于改进卷积神经网络(Convolutional Neural Network, CNN)的识别模型, 通过引入残差连接和注意力机制, 结合数据增强和正则化技术, 显著提升了模型的性能。实验结果中, 改进CNN模型在Birdsnap数据集上的物种识别时间为15毫秒, 准确率为92.3%; 在CUB-200-2011数据集上的物种识别时间为14毫秒, 准确率为94.5%。结果表明, 改进CNN模型在识别速度和准确性方面均优于对照模型, 其具有更强的鲁棒性和更高的物种识别多样性。研究结果不仅为野生鸟类的细粒度识别提供了新的技术路径, 也为生态监测和保护工作提供了更高效、更准确的技术支持, 有助于推动鸟类生态学研究 and 保护工作的深入开展。

关键词

卷积神经网络, 注意力机制, 残差连接, 野生鸟类, 细粒度识别与监测

Fine-Grained Recognition and Monitoring of Wild Birds Based on Deep Learning and UAV Image Acquisition

Yuxiang Liang*, Weiping Li, Mingming Zhang, Xuechao Zhai[#]

Heilongjiang Institute of Wild Animal Research, Harbin Heilongjiang

Received: Jul. 29th, 2025; accepted: Aug. 25th, 2025; published: Sep. 22nd, 2025

*第一作者。

[#]通讯作者。

文章引用: 梁宇祥, 李维平, 张明明, 翟学超. 融合深度学习与无人机图像采集的野生鸟类细粒度识别与监测方法研究[J]. 环境保护前沿, 2025, 15(9): 1238-1250. DOI: 10.12677/aep.2025.159139

Abstract

Fine-grained identification and monitoring of wild birds are of great significance for ecological research and protection. However, traditional identification methods face challenges such as limited generalization ability in complex natural environments. To improve the accuracy and real-time performance of fine-grained recognition of wild birds, a recognition model based on the improved Convolutional Neural Network (CNN) was constructed and optimized by combining unmanned aerial vehicle (UAV) image acquisition technology. By introducing residual connection and attention mechanism, combined with data augmentation and regularization techniques, the performance of the model has been significantly improved. In the experimental results, the species recognition time of the improved CNN model on the Birdsnap dataset was 15 milliseconds, and the accuracy rate was 92.3%. The species identification time on the CUB-200-2011 dataset was 14 milliseconds and the accuracy rate was 94.5%. The results show that the improved CNN model is superior to the control model in both recognition speed and accuracy, and it has stronger robustness and higher species recognition diversity. The research results not only provide a new technical path for fine-grained identification of wild birds, but also offer more efficient and accurate technical support for ecological monitoring and protection work, which is conducive to promoting the in-depth development of bird ecology research and protection work.

Keywords

Convolutional Neural Network, Attention Mechanism, Residual Connection, Wild Birds, Fine-Grained Identification and Monitoring

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

野生鸟类的细粒度识别与监测对于生态研究、生物多样性保护以及环境评估具有重要意义。随着全球生态环境的变化,许多鸟类物种面临着栖息地丧失、气候变化和人类活动干扰等威胁[1][2]。例如在湿地生态系统中,鸟类的栖息地因人类活动而不断减少,导致某些物种的生存受到严重威胁。近年来,随着无人机技术和深度学习的快速发展,其为野生鸟类的细粒度识别与监测提供了新的技术手段[3][4]。深度学习技术通过自动分析鸟类图像,能够快速、准确地识别鸟类种类,为生态研究提供了新的视角。然而,现有的方法在处理复杂自然环境中的图像时,仍然面临诸多挑战,如图像噪声、光照变化和背景干扰等[5][6]。

当前已有学者采用深度学习技术针对鸟类的识别展开了研究, Farman H 等人针对鸟类物种识别与分类的现状,提出一种卷积神经网络(Convolutional Neural Network, CNN)模型,该模型引入了跳跃连接来增强特征提取能力。结果表明,该 CNN 模型的准确率达到 92%, 优于普通 CNN 的 90% 准确率, 显示出深度学习算法在识别不同栖息地鸟类方面的强大能力[7]。陈天华等人针对现有细粒度鸟类目标识别算法准确率不足的现状,提出了一种新型的基于注意力机制的鸟类目标检测算法。实验结果表明,该算法检测准确率达到 82.8%, 召回率达到 77.0%, 相比现有主流算法有明显提升, 在细粒度鸟类目标识别任务应用场景下具有更好的检测性能[8]。杨雪珂等人针对野生鸟类行为学自动识别研究的现状,提出一种基于

残差卷积神经网络的勺嘴鹬动作识别模型。结果表明,模型在勺嘴鹬动作识别上均表现出色,测试集准确率分别为 96.90%、96.94%和 96.90%,验证了基于残差卷积神经网络进行勺嘴鹬动作识别的可行性[9]。此外国内某研究团队构建了一套技术先进、功能完整的人工智能鹤类识别平台,以此实现了无人机监测与 AI 识别的深度融合,在技术研发、系统集成、交付服务等方面均达到预期目标。开发团队克服了数据标注、性能优化、设备兼容等挑战,展现了高效的问题解决能力与技术落地能力,为野生动物监测领域提供了可复制的智能化解决方案。

上述研究虽取得了较好的鸟类识别效果,但数据标注效率与精度平衡较差,在复杂场景下鸟类重叠、姿态遮挡导致的标注困难,传统人工标注耗时耗力,且不同标注员存在标准不一致问题;同时无人机图传视频流的高帧率与模型推理速度的矛盾,导致存在延迟卡顿现象,影响监测实时性。基于此研究提出一种结合无人机图像采集技术与改进 CNN 模型的野生鸟类细粒度识别与监测方法。研究旨在通过融合深度学习与无人机图像采集技术,提高野生鸟类细粒度识别与监测的准确性和效率。研究的创新点在于引入残差连接和注意力机制,结合数据增强和正则化技术,显著提高了模型的识别性能和鲁棒性;同时通过优化模型结构和训练策略,使得模型在野外环境中能够快速提供识别结果。

2. 野生鸟类细粒度识别与监测的技术实现

2.1. 无人机图像采集技术

在野生鸟类细粒度识别与监测任务中,高质量的图像数据是实现准确识别的基础。因此,研究首先聚焦于无人机图像采集技术,旨在通过优化无人机的飞行控制、图像传感器配置以及数据采集策略,获取清晰、稳定且信息丰富的鸟类图像数据。具体的,研究通过下述方式优化无人机图像采集技术。首先采用网格化路径规划方法,将目标区域划分为多个网格,无人机按照预设的网格路径进行飞行,确保每个网格均被覆盖;其次采用比例-积分-微分(Proportion Integration Differentiation, PID)控制器对无人机飞行进行控制,确保无人机在飞行过程中保持稳定,并能够根据预设的路径进行精确导航;最后选择高分辨率的互补金属氧化物半导体(Complementary Metal-Oxide-Semiconductor, CMOS)和电荷耦合器件(Charge-Coupled Device, CCD)传感器,根据鸟类活动特征调整数据的采集、存储以及压缩方式。无人机的飞行控制和路径规划是获取高质量图像数据的关键步骤。飞行路径规划的目标是使无人机能够覆盖目标区域的所有关键位置,同时避免不必要的飞行路径,以节省时间和能源。研究采用网格的方式,即将目标区域划分为多个网格,无人机按照预设的网格路径进行飞行,以此确保每个网格均被覆盖[10] [11]。路径规划模型的数学表达如公式(1)所示。

$$Path = \{(x_i, y_i) | i = 1, 2, \dots, n\} \quad (1)$$

公式(1)中, (x_i, y_i) 表示路径上的第 i 个坐标, n 表示路径上的总点数。在飞行策略中,其需要确保无人机能够在飞行过程中保持稳定,并能够根据预设的路径进行精确导航。研究采用 PID 控制器的方法对无人机飞行进行控制,其控制方程如公式(2)所示。

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt} \quad (2)$$

公式(2)中, $u(t)$ 表示控制输入, $e(t)$ 表示误差信号, K_p 、 K_i 和 K_d 分别表示比例、积分和微分增益。受鸟类种类以及形态特征的影响,其图像种类较为繁杂,清晰且细节丰富的鸟类图像能够为细粒度识别提供可靠的数据支持,相关示意如图 1 所示。

为获取图 1 所示的高质量鸟类图像,图像传感器的选择和配置至关重要。研究选择的图像传感器包括 CMOS 传感器和 CCD 传感器。镜头的焦距会影响图像的视角和放大倍率,选择合适的焦距可以确保

鸟类在图像中占据适当的比例。光圈大小和 ISO 值的设置则会影响图像的曝光和噪声水平。例如，较大的光圈可以增加进光量，但会减少景深；较高的 ISO 值可以提高图像亮度，但会增加噪声[12]。因此，需要根据实际环境和需求进行合理的配置。数据采集策略是无人机图像采集技术中的重要组成部分，其直接影响到图像数据的质量、存储和传输效率。数据采集策略主要涉及图像的采集频率、存储格式以及压缩方法，相关示意如图 2 所示。



Figure 1. Schematic diagram of birds
图 1. 鸟类示意图

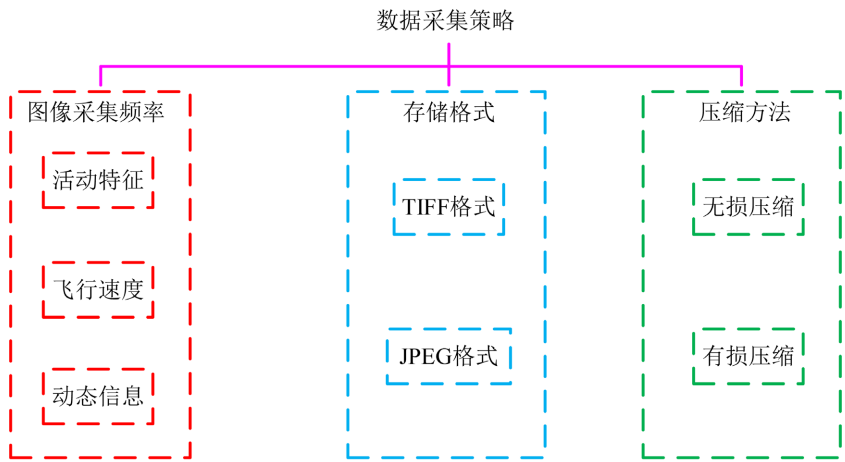


Figure 2. Data collection strategy
图 2. 数据采集策略

图 2 中，在确定图像采集频率时需要综合考虑目标鸟类的活动特性和无人机的飞行速度。例如，对于飞行速度较快的鸟类，如蜂鸟，我们设定较高的采集频率(如每秒 30 帧)，以确保捕捉到足够的动态信息，避免因采样不足而丢失关键的识别特征。相反，对于飞行速度较慢或主要在栖息状态下的鸟类，如猫头鹰，可以适当降低采集频率(如每秒 5 帧)，以减少数据量并提高存储效率。存储格式上主要选择 TIFF

格式进行无损压缩存储,同时针对有损压缩方法,如 JPEG 格式,通过丢弃一些图像数据来实现更高的压缩比,适用于对图像质量要求不是特别高的应用场景。最后通过优化无人机的飞行控制、路径规划以及图像传感器的配置和数据采集策略,研究可有效提高无人机采集的野生鸟类图像的质量。这些优化措施为后续的细粒度识别和监测任务提供了坚实的数据基础,确保了识别结果的准确性和可靠性。

2.2. 基于改进 CNN 的野生鸟类细粒度识别模型构建与优化

在成功获取到高质量的野生鸟类图像数据后,为实现对不同鸟类种类及其细微特征的精准区分,研究构建并优化了一种基于改进 CNN 的识别模型。传统 CNN 模型包括卷积层、池化层以及全连接层,该模型在处理复杂图像时存在梯度消失和特征提取不足的问题。为了解决这些问题,研究引入残差连接和注意力机制对 CNN 模型进行了改进。具体的通过将输入直接连接到后面的层,解决了深层网络中的梯度消失问题。残差连接通过引入一个跳跃连接,将输入直接加到输出上,从而允许梯度直接传播,缓解了梯度消失问题;其次通过加权不同特征区域,使模型更加关注于图像中的关键部分,如羽毛纹理、喙部形状等。注意力机制可以帮助模型自动聚焦于这些关键区域,从而提高识别的准确性。改进后的 CNN 模型网络结构如图 3 所示。

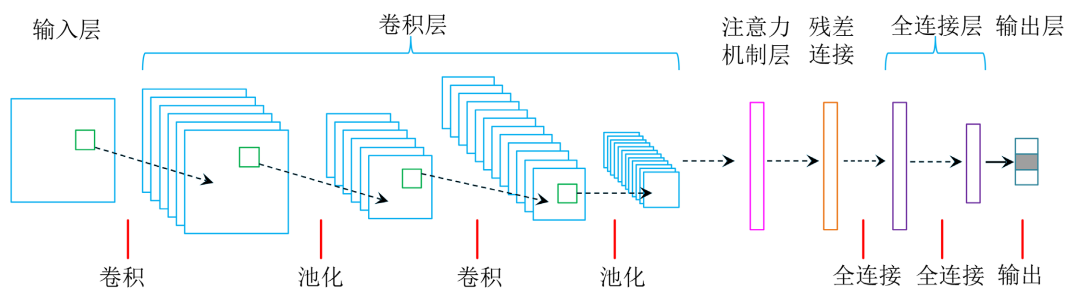


Figure 3. Schematic diagram of the improved CNN model network structure

图 3. 改进 CNN 模型网络结构示意图

图 3 中,残差连接通过将输入直接连接到后面的层,解决了深层网络中的梯度消失问题。在 CNN 模型中,随着网络层数的增加,梯度在反向传播过程中可能会逐渐消失或爆炸,导致网络难以训练。残差连接通过引入一个跳跃连接,将输入直接加到输出上,从而允许梯度直接传播,缓解了这一问题。相关数学表达如公式(3)所示。

$$y = f(x) + x \quad (3)$$

公式(3)中, $f(x)$ 表示卷积层的输出, x 表示输入, y 表示残差连接的输出。注意力机制通过加权不同特征区域,使模型更加关注于图像中的关键部分。在野生鸟类细粒度识别任务中,鸟类的某些特征区域(如羽毛纹理、喙部形状等)对于识别至关重要。基于 Grad-CAM 生成的注意力特征如图 4 所示。

图 4 中,引入注意力机制后,鸟类的特征部位被突出展示,其明显特征集中在头部。注意力机制可以帮助模型自动聚焦于这些关键区域,从而提高识别的准确性。相关数学表达如公式(4)所示。

$$y = \alpha \cdot f(x) + (1 - \alpha) \cdot x \quad (4)$$

公式(4)中, α 表示权重系数,其通常通过一个学习过程来确定,使得模型能够自动调整对不同特征区域的关注程度。为了提高模型的性能,研究采用数据增强和正则化两种优化方式对 CNN 模型进行了优化改进。数据增强通过随机变换图像,增加数据的多样性,防止过拟合。常用的数据增强方法包括随机旋转、翻转、颜色调整等,相关数学表达如公式(5)所示。

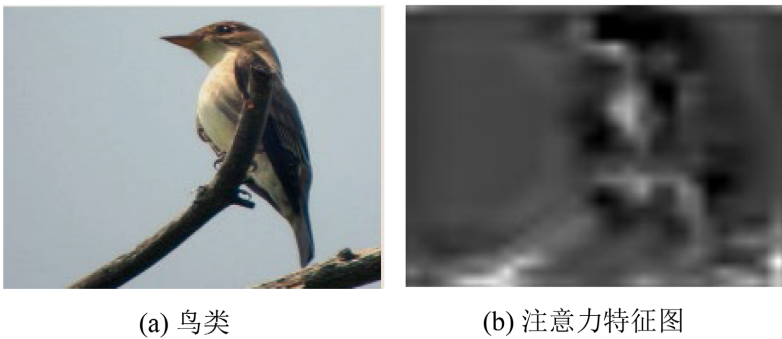


Figure 4. Attention feature map
图 4. 注意力特征图

$$\tilde{x} = T(x) \tag{5}$$

公式(5)中， $\tilde{x}$ 表示变换后的图像， T 表示变换后的图像，变换后可生成更多的训练样本，以此提高模型的泛化能力。正则化通过在损失函数中添加正则项，以此限制模型的复杂度，进而防止过拟合现象，相关数学表达如公式(6)所示。

$$L = L_0 + \lambda \cdot R(W) \tag{6}$$

公式(6)中， L_0 表示原始损失函数， $R(W)$ 表示正则项， λ 表示正则化系数， W 表示模型的权重。最后为了进一步提高模型在野生鸟类细粒度识别任务中的性能，研究采用了迁移学习技术。通过在大规模图像数据集上预训练模型，然后将预训练得到的权重作为初始化权重，再在特定的鸟类图像数据集上进行微调。迁移学习可以显著减少训练时间和计算资源，同时提高模型的性能。预训练模型在大规模数据集上学习到的特征具有一定的通用性，能够为特定任务提供良好的初始特征表示，相关示意如图 5 所示。

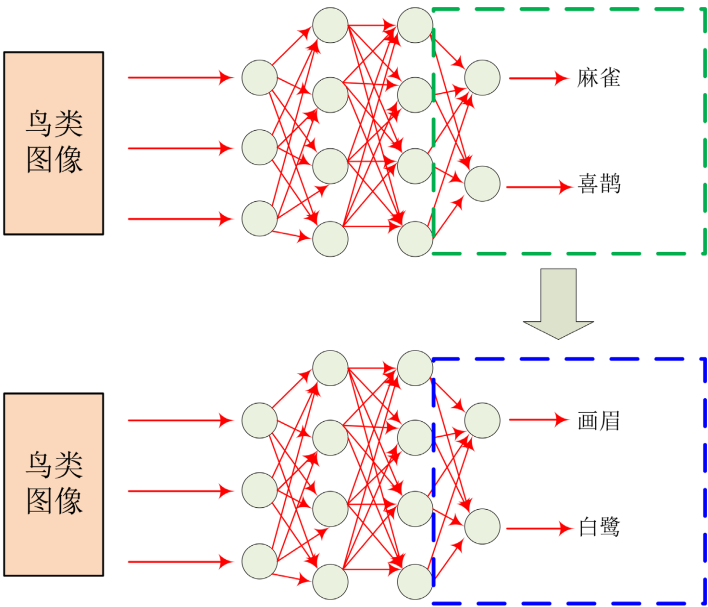


Figure 5. Schematic diagram of transfer learning
图 5. 迁移学习示意图

图 5 所示迁移学习中，其微调过程如下所述。首先将预训练模型在数据集上训练得到的权重并加载

至模型中；其次选择冻结预训练模型的部分层；紧接着为了适应野生鸟类细粒度识别任务，在预训练模型的基础上添加了自定义的全连接层。最后使用鸟类图像数据集对模型进行训练。训练过程中，使用较小的学习率，以避免对预训练权重进行过大的调整。最后研究中关于野生鸟类细粒度识别与监测的流程如图 6 所示。

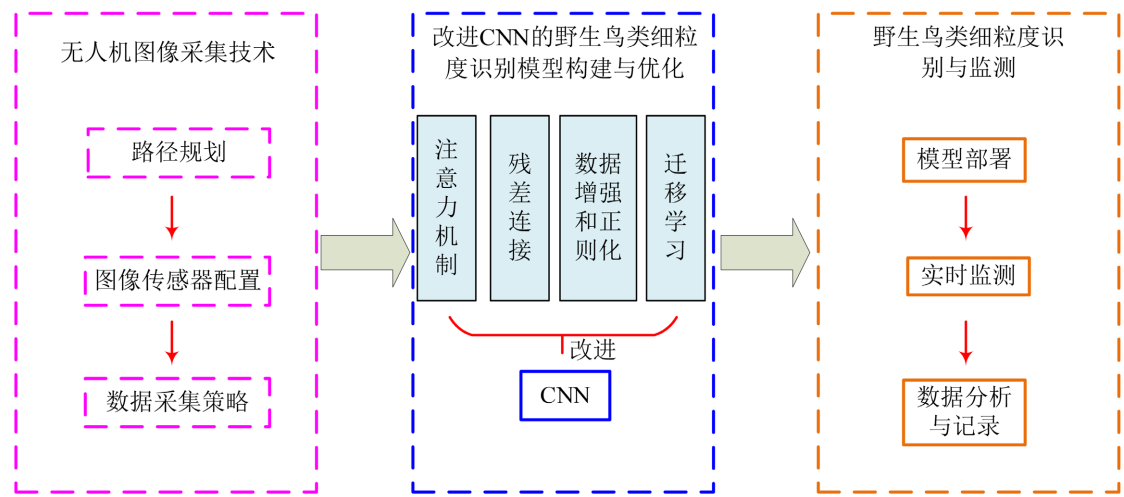


Figure 6. Wild bird fine-grained identification and monitoring process
图 6. 野生鸟类细粒度识别与监测流程

图 6 中，在无人机图像采集技术方面，研究采用了下述具体措施来优化飞行路径规划和图像传感器配置。首先将目标监测区域细分为多个网格单元，每个网格单元代表无人机需要覆盖的特定区域。无人机按照预设的网格路径进行系统化飞行，确保对整个监测区域的全面覆盖，从而提高数据采集的效率和覆盖率；其次利用 PID 控制器对无人机的飞行姿态和路径进行精确控制。PID 控制器通过调整比例、积分和微分增益，实现对无人机飞行稳定性和路径跟踪精度的优化。紧接着选择高分辨率的 CMOS 和 CCD 传感器，以捕捉鸟类的细微特征。根据鸟类的活动特性(如飞行速度、栖息行为等)，调整镜头的焦距、光圈大小和 ISO 值，以优化图像的清晰度和细节表现。最后根据鸟类的活动模式和无人机的飞行速度，调整图像的采集频率，确保捕捉到足够的动态信息。同时，选择适当的图像存储格式和压缩方法，以平衡图像质量和存储效率。

3. 改进 CNN 模型性能验证及鸟类细粒度识别与监测

3.1. 改进 CNN 模型的性能验证

为了验证改进后的 CNN 模型在野生鸟类细粒度识别任务中的性能，研究设计了一系列实验旨在评估模型的准确性、鲁棒性和泛化能力。研究中实验环境以及改进 CNN 模型的参数设置如表 1 所示。

基于表 1 的实验环境与参数设置情况，数据集选取 Birdsnap 数据集和 CUB-200-2011 数据集。其中 Birdsnap 数据集包含 500 种不同的鸟类，总共超过 47000 张图像，每张图像都提供了鸟类的种类标签和边界框信息；CUB-200-2011 数据集包含 200 种不同的鸟类，总共 11788 张图像，每张图像都提供了鸟类的种类标签、边界框信息以及关键点标注。同时按照 6:2:2 的比例将数据集划分为训练集、验证集和测试集。研究同时选取 VGG16、ResNet-50 以及 Inception-V3 模型做对照实验，通过对比损失函数曲线、准确率曲线、平均精度均值(mean Average Precision, mAP)和模型大小等指标，以评估改进 CNN 模型的性能。研究首先对比了四种模型的损失函数曲线以及准确率曲线，结果如图 7 所示。

Table 1. Experimental environment and model parameter settings
表 1. 实验环境及模型参数设置

组件	具体信息	参数	数值
CPU	Intel Core i9-12900K (16 核, 32 线程)	优化器	Adam, 学习率: 0.0001, 衰减率: 0.9
GPU	NVIDIA GeForce RTX 3090 (24GB GDDR6X)	损失函数	交叉熵损失
内存	64GB DDR4 3200MHz	批量大小	32
存储	1TB NVMe SSD + 4TB HDD	训练周期	50 个 epoch
主板	ASUS ROG Maximus Z690 Hero	数据增强	随机旋转($\pm 15^\circ$), 随机翻转(水平和垂直), 随机裁剪(80%)
电源	Corsair RM1000x (1000W)	正则化	L2 正则化, 正则化系数: 0.0005
散热	NZXT Kraken X73 (360mm AIO)	迁移学习	冻结前 15 层, 添加自定义全连接层
操作系统	Ubuntu 20.04 LTS	微调学习率	0.00001
深度学习框架	TensorFlow 2.10.0	/	/
Python 版本	Python 3.8.10	/	/

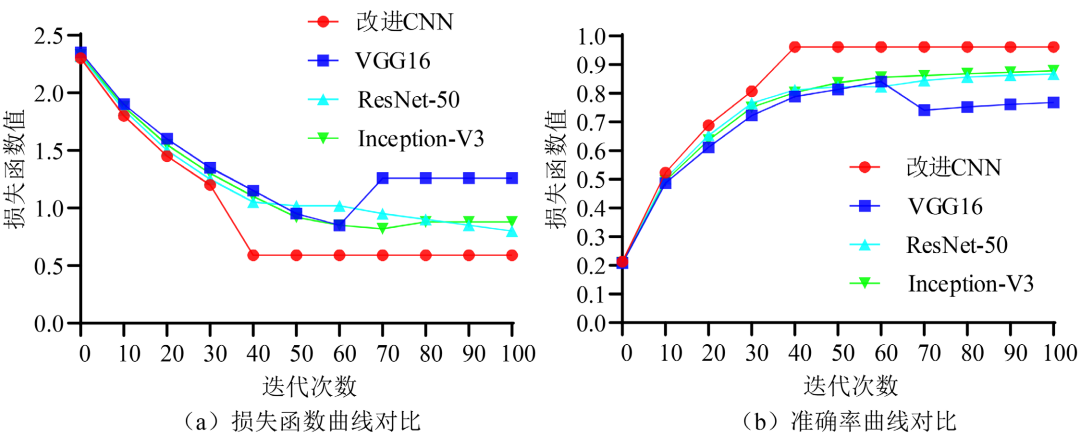


Figure 7. Comparison of loss curve and accuracy curve
图 7. 损失曲线和准确率曲线对比

图 7(a)为四种模型损失曲线对比。结果表明改进 CNN 模型在 40 次迭代时, 损失值迅速下降至 0.59, 并在后续迭代中保持稳定, 显示出改进 CNN 模型的快速收敛能力; VGG16 在 70 次迭代时, 损失值下降至 1.26, 并在后续迭代中保持稳定, 收敛速度较慢。ResNet-50 与 Inception-V3 模型分别在 60 和 80 时收敛, 最终损失值分别趋于 1.02 和 0.88。图 7(b)为四种模型准确率曲线对比。结果表明改进 CNN 在 40 次迭代时, 准确率迅速上升至 0.962, 并在后续迭代中保持稳定。VGG16 存在一定的波动, 其准确率最终趋于 0.741。ResNet-50 与 Inception-V3 模型收敛后准确率分别趋于 0.823 和 0.869。研究进一步对比了四种模型的 mAP 和模型大小。结果如图 8 所示。

图 8(a)为平行实验 1 mAP 值与模型大小对比, 图 8(b)为平行实验 2 mAP 值与模型大小对比。结果表明改进 CNN 模型的 mAP 平均为 0.9285, 表明改进 CNN 模型在两次平行实验中均表现出较高的识别精度, 且精度较为稳定; VGG16 模型的 mAP 平均为 0.8495, 其数值最低, 表明其在识别精度方面存在一定的局限性, 且精度波动较大。模型大小方面改进 CNN 平均值为 123 MB, 表明改进 CNN 模型在保持高精度的同时, 模型大小适中, 具有较好的存储效率。VGG16 模型大小平均值为 151 MB, 表明 VGG16

模型在存储效率方面表现不佳，模型较大。研究最后对比了四种模型的训练时间和推理时间，结果如表 2 所示。

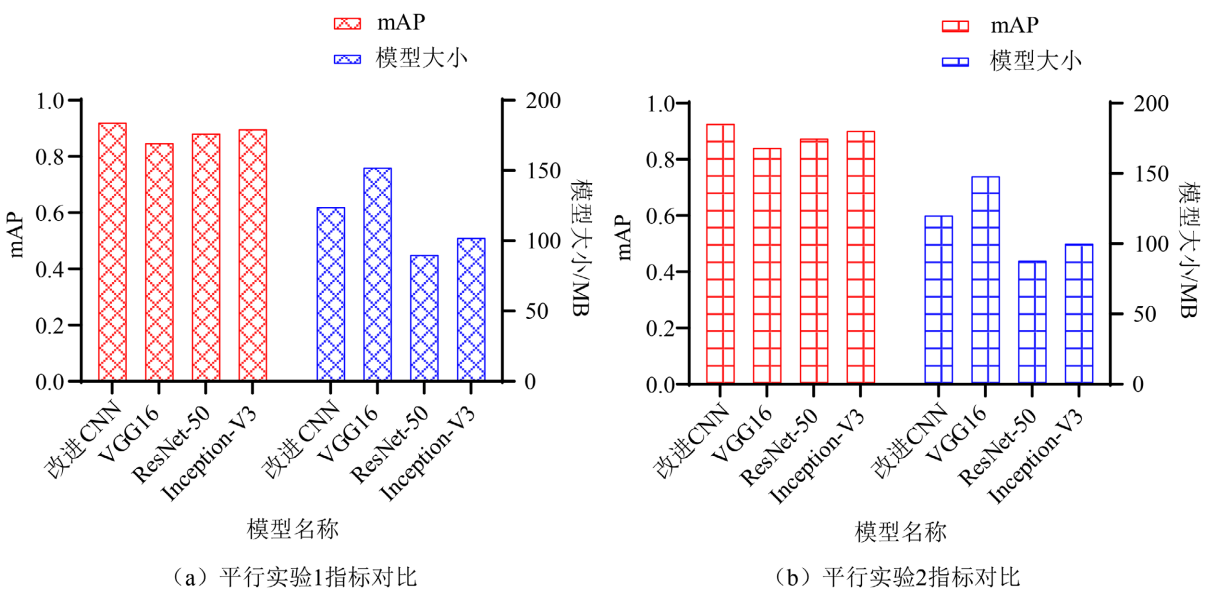


Figure 8. mAP value and model size comparison
图 8. mAP 值与模型大小对比

Table 2. Comparison of model training time and inference time
表 2. 模型训练时间和推理时间对比

平行实验	模型	训练时间/h	推理时间/ms
平行实验 1	改进 CNN	2.7	17
	VGG16	3.2	22
	ResNet-50	3.0	20
	Inception-V3	3.4	18
平行实验 1	改进 CNN	2.3	13
	VGG16	2.8	18
	ResNet-50	2.6	16
	Inception-V3	3.0	14

表 2 中，改进 CNN 模型的训练时间平均为 2.5 h，表明改进 CNN 模型在训练效率方面表现较好，能够在较短的时间内完成训练；同时其推理时间平均为 15 ms，表明改进 CNN 模型在推理效率方面表现优秀，能够在较短的时间内完成单次预测，适合实时应用。VGG16 的平均训练时间与推理时间均最长，数值分别为 3.0 h 和 20 ms，表明 VGG16 模型在训练效率和推理效率上表现较差。ResNet-50 和 Inception-V3 模型训练时间和推理时间适中，适合野生鸟类细粒度识别的实时应用。

3.2. 野生鸟类细粒度识别与监测

对改进 CNN 模型性能进行验证后，研究进一步对比了四种模型对野生鸟类细粒度识别与监测的相关指标，包括物种识别时间、识别稳定性、物种识别多样性和识别置信度。首先四种模型的物种识别时

间对比如图 9 所示。

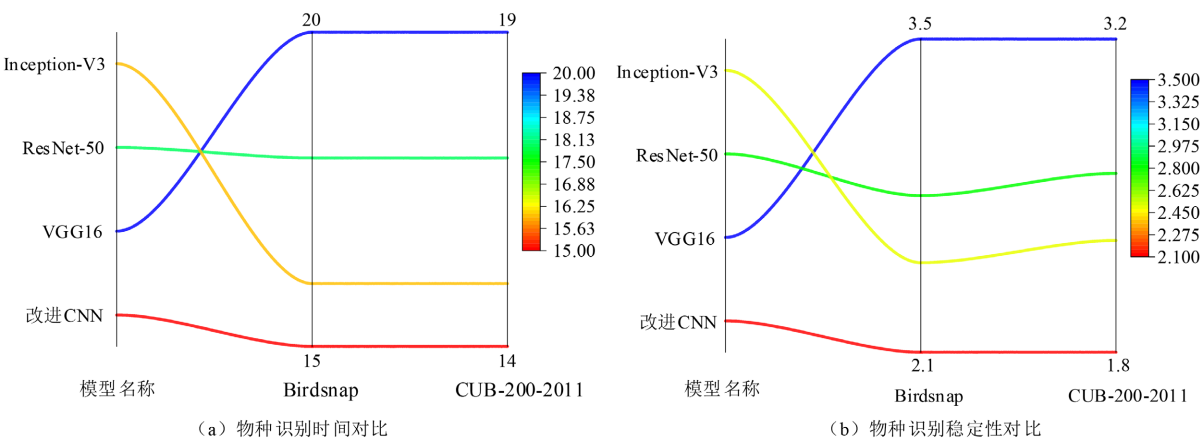


Figure 9. Species identification time comparison
图 9. 物种识别时间对比

图 9(a)为四种模型物种识别时间对比。结果表明改进 CNN 模型在 Birdsnap 数据集和 CUB-200-2011 数据集上分别为 15 ms 和 14 ms，其数值显著低于其余对照模型，表明改进 CNN 模型具有较快的识别速度，适合实时应用。VGG16 模型物种识别时间最长，分别为 20 ms 和 19 ms，表明其识别速度较慢。图 9(b)为四种模型物种识别稳定性对比。结果表明改进 CNN 模型的准确率变化范围为 $\pm 2.1\%$ 和 $\pm 1.8\%$ ，表明改进 CNN 模型在不同环境条件下的识别结果较为稳定，具有较高的鲁棒性。VGG16 模型的准确率变化范围最大，其变化范围为 $\pm 3.5\%$ 和 $\pm 3.2\%$ ，表明其不同环境条件下的识别结果波动较大，稳定性较差。研究进一步对比了四种模型的鸟类物种识别多样性和识别置信度，结果如图 10 所示。

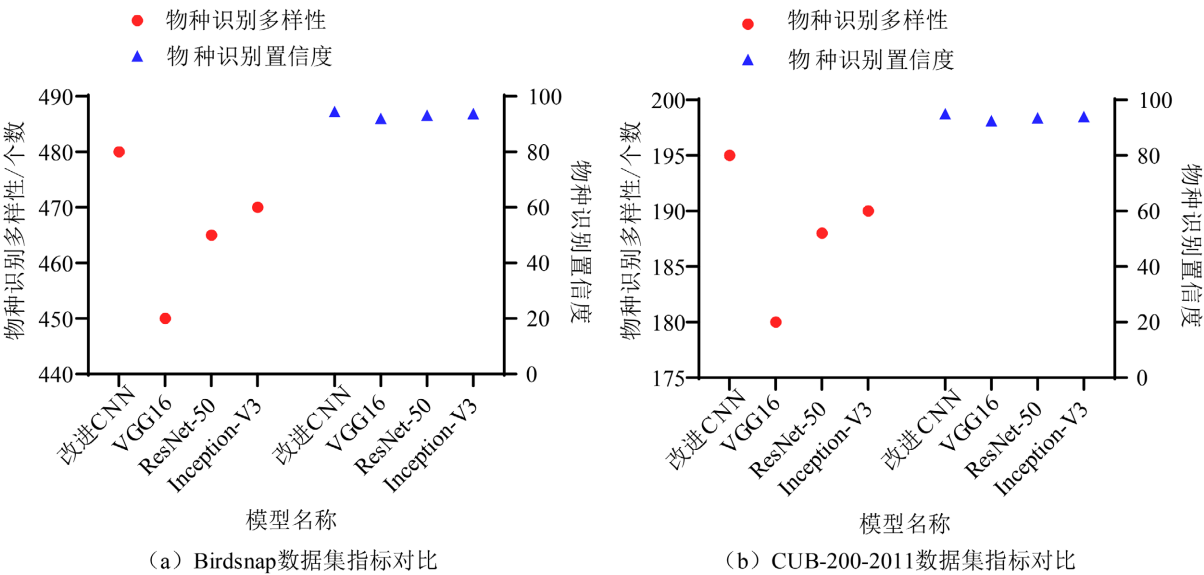


Figure 10. Comparison of bird species identification diversity and identification confidence
图 10. 鸟类物种识别多样性和识别置信度对比

图 10(a)为 Birdsnap 数据集指标对比，图 10(b) CUB-200-2011 数据集指标对比。结果表明改进 CNN 模型在 Birdsnap 数据集上识别了 480 种鸟类，在 CUB-200-2011 数据集上识别了 195 种鸟类，表明改进

CNN 模型在两个数据集上均具有较高的物种识别多样性，能够识别更多的鸟类种类；同时其平均置信度分别为 94.5%和 95.0%，表明其对识别结果的置信度较高，识别结果较为可靠。VGG16 模型识别鸟类数目最少，分别为 450 种和 180 种，表明在物种识别多样性方面稍逊一筹，识别的鸟类种类较少；同时其平均置信度分别为 92.0%和 92.5%，表明其对识别结果的置信度较低，识别结果的可靠性稍差。研究最后对比了三种不利条件下的物种识别鲁棒性，结果如表 3 所示。

Table 3. Species recognition robustness comparison
表 3. 物种识别鲁棒性对比

不利条件	模型	光照变化下的准确率(%)	背景噪声下的准确率(%)	部分遮挡下的准确率(%)
光照变化	改进 CNN 模型	92.3	90.7	89.2
	VGG16	87.6	85.9	84.3
	ResNet-50	90.4	88.2	86.7
	Inception-V3	91.2	89.5	87.8
背景噪声	改进 CNN 模型	91.4	89.8	88.3
	VGG16	86.7	85.0	83.5
	ResNet-50	88.9	87.1	85.4
	Inception-V3	90.1	88.4	86.6
部分遮挡	改进 CNN 模型	90.6	88.9	87.4
	VGG16	85.8	84.1	82.6
	ResNet-50	87.6	85.8	84.2
	Inception-V3	88.7	86.9	85.3

表 3 中，改进 CNN 模型在不同不利条件下均表现出较高的准确率，其数值分别为 92.3%、90.7%和 89.2%，数值显著高于其余模型，表明其鲁棒性能较好。VGG16 模型在不同条件下的准确率分别为 87.6%、85.9%和 84.3%，表明其鲁棒性较差。ResNet-50 模型和 Inception-V3 模型的准确率适中，其数值在不同条件下介于 VGG16 与改进 CNN 模型之间。综上，改进 CNN 模型在光照变化、背景噪声和部分遮挡条件下均表现最佳，表明其在野生鸟类细粒度识别任务中具有较高的鲁棒性，能够在不利条件下保持较高的识别准确率。

4. 讨论

为提高野生鸟类细粒度识别与监测的准确性和效率，研究构建并优化了一种基于改进 CNN 的识别模型，并在 Birdsnap 数据集和 CUB-200-2011 数据集上进行了广泛的实验验证。实验结果中，改进 CNN 模型在多个关键性能指标上表现出色，其识别时间为 15 ms 和 14 ms，显著优于 VGG16 模型(20 ms 和 19ms)、ResNet-50 (20 ms 和 16 ms)和 Inception-V3 (18 ms 和 14 ms)，同时准确率最终趋于 0.962，表明改进 CNN 模型具有更快的识别速度，适合实时应用，能够在野外环境中快速提供识别结果，该结果与林海等人的结果一致[13]。林海等人针对传统水鸟监测方法耗时耗力且存在主观偏差的现状，提出了一种基于 CNN 的无人机遥感影像水鸟自动识别模型。结果表明，该模型召回率达到 93.00%，精度达到 90.83%，在种群监测工作中具有应用潜力。在不同环境条件下的识别稳定性方面，改进 CNN 模型表现出显著优势。在光照变化、背景噪声和部分遮挡条件下，改进 CNN 模型的准确率分别为 92.3%、90.7%和 89.2%，显著优于对照模型。该结果与 Wang Q 等人的研究存在一定的差异[14]。Wang Q 等人针对传统鸟类声学

识别方法中忽视鸟类种群内在关系导致生物信息丢失的现状,提出了一种基于系统发育视角改进 CNN 模型。在鲁棒性结果中,改进 CNN 模型较 Inception-V3 较差,其可能的原因为 Inception-V3 通过在大规模数据集上进行预训练,使得其能够学习到更鲁棒的特征。此外改进 CNN 模型在物种识别多样性和置信度方面也表现出色。在 Birdsnap 数据集上,改进 CNN 模型识别了 480 种鸟类,平均置信度为 94.5%;在 CUB-200-2011 数据集上,识别了 195 种鸟类,平均置信度为 95.0%。相比之下,VGG16 模型仅识别了 450 种和 180 种鸟类,平均置信度分别为 92.0%和 92.5%。这表明改进 CNN 模型能够识别更多种类的鸟类,并且对识别结果的置信度更高,识别结果更为可靠。综上改进 CNN 模型在野生鸟类细粒度识别与监测任务中具有显著的优势。其快速的识别速度、较高的鲁棒性、丰富的物种识别多样性和高置信度的识别结果,使其能够更好地适应野外复杂环境,为野生鸟类的监测和研究提供有力的技术支持。

5. 结论

研究通过融合深度学习与无人机图像采集技术,成功构建并优化了一种基于改进 CNN 的野生鸟类细粒度识别与监测方法。实验结果表明,改进 CNN 模型在物种识别时间、鲁棒性、多样性和置信度等多个关键性能指标上均表现出色,显著优于 VGG16、ResNet-50 和 Inception-V3 模型。结果表明,改进 CNN 模型能够有效提高野生鸟类细粒度识别与监测的准确性和效率,为野生鸟类的监测和研究提供了有力的技术支持。然而研究仍存在一些不足之处,模型的训练和推理过程对计算资源的需求较高,限制了其在资源受限的嵌入式设备上的应用。未来的研究可探索更高效的算法,以降低计算资源需求,使其更适合在移动设备和边缘计算环境中应用。同时开发移动端应用程序实现监测结果实时推送,集成地理信息系统展示鸟类活动轨迹,增加异常行为(如伤病、集群迁徙)智能预警功能,进而提升用户体验。此外探索多无人机协同监测的分布式算法,以此提升复杂场景下的监测精度与效率。

基金项目

黑龙江省省属科研院所科研业务费项目(CZKYF2024-1-B037)。

参考文献

- [1] 王洪昌, 夏舫, 张渊媛, 等. 基于深度学习算法的鸟类及其栖息地识别——以北京翠湖国家城市湿地公园为例[J]. 生态学杂志, 2024, 43(7): 2231-2238.
- [2] Chen, X. (2023) Vehicle Feature Recognition via a Convolutional Neural Network with an Improved Bird Swarm Algorithm. *Journal of Internet Technology*, **24**, 421-432.
- [3] 邓抒憧, 陈爱斌, 戴子健. 基于多路激励和金字塔切分注意力的鸟类行为识别[J]. 应用科学学报, 2025, 43(1): 154-168.
- [4] 王蕊, 史玉龙, 孙辉, 等. 基于轻量化的高分辨率鸟群识别深度学习网络[J]. 华中科技大学学报: 自然科学版, 2023, 51(5): 81-87.
- [5] Xu, X., Yang, C.C., Xiao, Y., et al. (2023) A Fine-Grained Recognition Neural Network with High-Order Feature Maps via Graph-Based Embedding for Natural Bird Diversity Conservation. *International Journal of Environmental Research and Public Health*, **20**, 4924-4639. <https://doi.org/10.3390/ijerph20064924>
- [6] 韩冰, 王红昌, 苏志刚, 等. 针对小型鸟类目标的基于门控循环单元的扩展卡尔曼跟踪方法[J]. 信号处理, 2024, 40(5): 944-956.
- [7] Farman, H., Ahmed, S., Imran, M., et al. (2023) Deep Learning Based Bird Species Identification and Classification Using Images. *Journal of Computing & Biomedical Informatics*, **6**, 79-96.
- [8] 陈天华, 朱家煊, 印杰. 基于注意力机制的鸟类识别算法[J]. 计算机应用, 2024, 44(4): 1114-1120.
- [9] 杨雪珂, 蒙金超, 冯悦恒, 等. 基于残差卷积神经网络模型的勺嘴鹬动作识别[J]. 热带生物学报, 2023, 14(5): 481-489.
- [10] 谢威宇, 张强. 基于深度学习的图像中无人机与飞鸟检测研究综述[J]. 计算机工程与应用, 2024, 60(8): 46-55.

- [11] 邱浩楠, 何明, 韩伟, 等. 一种仿鸟群行为的无人机集群相变控制方法[J]. 现代防御技术, 2025, 53(1): 11-22.
- [12] 段海滨, 尤灵辰, 范彦铭, 等. 仿鸟群自推进机制的无人机集群相变控制[J]. 自动化学报, 2025, 51(5): 960-971.
- [13] 林海, 高大中, 张童, 等. 基于卷积神经网络的无人机遥感影像水鸟自动识别[J]. 动物学杂志, 2024, 59(3): 450-459.
- [14] Wang, Q., Song, Y., Du, Y., Yang, Z., Cui, P. and Luo, B. (2024) Hierarchical-Taxonomy-Aware and Attentional Convolutional Neural Networks for Acoustic Identification of Bird Species: A Phylogenetic Perspective. *Ecological Informatics*, **80**, Article 102538. <https://doi.org/10.1016/j.ecoinf.2024.102538>