

DOI: 10.20103/j.stxb.202405071019

庄鸿飞, 侯金, 王守强, 高宇, 张朝晖, 王宗灵, 赵林林, 何育欣, 周庆杰, 鹿志创, 邢衍阔, 杜光迅. 基于面部特征的斑海豹 (*Phoca largha*) 个体识别研究. 生态学报, 2025, 45(13): - .

Zhuang H F, Hou J, Wang S Q, Gao Y, Zhang Z H, Wang Z L, Zhao L L, He Y X, Zhou Q J, Lu Z C, Xing Y K, Du G X. Facial feature-based individual identification of spotted seals (*Phoca largha*). Acta Ecologica Sinica, 2025, 45(13): - .

基于面部特征的斑海豹 (*Phoca largha*) 个体识别研究

庄鸿飞¹, 侯金², 王守强¹, 高宇¹, 张朝晖¹, 王宗灵¹, 赵林林¹, 何育欣^{3,*}, 周庆杰⁴, 鹿志创⁵, 邢衍阔⁵, 杜光迅⁶

1 自然资源部第一海洋研究所, 渤海海峡生态通道野外科学观测研究站, 青岛 266061

2 北京师范大学生命科学学院, 北京 100875

3 川北医学院, 南充 637100

4 青岛极地海洋世界有限公司, 青岛 266061

5 辽宁省海洋水产科学研究院, 大连 116023

6 自然资源部北海发展研究院, 青岛 266102

摘要: 针对渤海斑海豹种群传统监测方法中高成本、低效率个体识别难题, 本文提出一种基于面部特征的斑海豹个体识别模型。该模型基于卷积神经网络架构, 结合注意力机制和多尺度特征处理技术, 提高了对面部关键特征的识别能力, 并通过深度可分离卷积降低了参数量和计算需求。实验结果显示, 融合 SlimC2F、CBAM 和 SPPF 模块可显著提高网络识别性能, 本文方法 (SlimC2F-CBAM+SPPF) 取得最佳识别效果。在自建背景数据集与 Caltech-256 公共数据集上, 模型识别准确率达 98.35%、80.70%, 优于目前主流模型。在迁移学习策略的应用下, 模型对不同背景的适应性增强, 识别准确率提升至 98.70%。此外, 即使在较小数据集上, 模型仍能保持较高的识别准确率。最后, 可视化分析展示了模型主要关注眼睛和鼻子等关键部位, 进一步验证了模型的有效性和可解释性, 为实现斑海豹个体的高效监测提供了新的技术手段。

关键词: 斑海豹; 个体识别; 面部识别; 深度学习; 卷积神经网络

Facial feature-based individual identification of spotted seals (*Phoca largha*)

ZHUANG Hongfei¹, HOU Jin², WANG Shouqiang¹, GAO Yu¹, ZHANG Zhaohui¹, WANG Zongling¹, ZHAO Linlin¹, HE Yuxin^{3,*}, ZHOU Qingjie⁴, LU Zhichuang⁵, XING Yankuo⁵, DU Guangxun⁶

1 Observation and Research Station of Bohai Strait Eco-Corridor, First Institute of Oceanography, Ministry of Natural Resources, Qingdao 266061, China

2 College of Life Sciences, Beijing Normal University, Beijing 100875, China

3 North Sichuan Medical College, Nanchong 637100, China

4 Qingdao Polar Ocean World Co., Ltd., Qingdao 266061, China

5 Liaoning Ocean and Fisheries Science Research Institute, Dalian 116023, China

6 North China Sea Development Research Institute, Ministry of Natural Resources, Qingdao 266102, China

Abstract: In response to the high cost and low efficiency of traditional monitoring methods for the population of spotted seals in the Bohai and Yellow Seas, this paper proposes a facial feature-based individual recognition model for spotted seals. This model is based on a convolutional neural network architecture, which integrates attention mechanism and multi-scale feature processing technology to enhance the recognition ability of key facial features. It also utilizes depthwise separable

基金项目: 国家自然科学基金项目 (32201433); 青岛博士后应用研究项目 (QDBSH202205); 国家重点研发计划项目 (2022YFF0802204); 川北医学院校级科技发展基金青年项目 (CBY20-QA-Y26)

收稿日期: 2024-05-07; **网络出版日期:** 2025-00-00

* 通讯作者 Corresponding author. E-mail: hyx_nsmc@163.com

convolution to reduce model parameters and computational requirements. The results showed that integrating SlimC2F, CBAM, and SPPF modules significantly enhances network recognition performance, with the proposed method (SlimC2F-CBAM+SPPF) achieving the best results. On the self-built dataset with backgrounds and the Caltech-256 public dataset, the model recognition accuracy reaches 98.35% and 80.70%, outperforming current mainstream methods. With the application of transfer learning strategies, the model's adaptability to different backgrounds has been enhanced, with an increase in identification accuracy to 98.70%. In addition, the model maintains high identification accuracy even on smaller datasets. Finally, visualization analysis demonstrated that the model primarily focused on capturing key areas such as eyes and nose for spotted seal facial recognition, further validating the effectiveness and interpretability of this method while providing new technical means for efficient monitoring of spotted seals.

Key Words: spotted seals; individual recognition; facial recognition; deep learning; convolutional neural network

邻近我国的渤海斑海豹种群位于该物种全球分布最南端,种群规模最小,属于典型的边缘小种群^[1]。自 20 世纪 40 年代以来该种群在人类活动和气候变化的影响下不断衰退,面临局部灭绝风险^[2-3]。我国相关管理部门为此采取了一系列保护行动,在就地保护方面,建立了大连斑海豹国家级自然保护区、辽宁辽河口国家级自然保护区、庙岛群岛斑海豹省级自然保护区等自然保护地;在政策制定方面,制定并实施了《斑海豹保护行动计划(2017—2026 年)》,经过多年努力,渤海斑海豹种群保护取得了重要进展^[4]。尽管如此,目前该物种的受威胁状态仍不容乐观,属于中国易危物种,并于 2021 年升级为国家一级保护野生动物^[5-6]。未来保护应考虑为这一边缘小种群“建档立卡”,摸清底数与种群社会结构,监测种群动态^[7]。

斑海豹种群监测方法的选用需权衡成本、时效性和准确性。上世纪 80 年代我国就开始依托破冰船、飞机对斑海豹种群开展人工调查,这种方法调查成本高,调查时间和调查路线设定受限于地理环境情况与先验知识,且无法对个体进行识别^[8-9]。2010 年,开始对斑海豹个体进行卫星信标跟踪调查,获得斑海豹个体在一个时间段内的连续活动情况,但信标大都在 1—3 月脱落,并且野外释放斑海豹的机会具有偶然性,在实际大规模监测中的可操作性不强^[10-12]。基于血样和环境中的 DNA 检测也被用于斑海豹鉴定,这种方法的个体识别率高,但成本高、遇见率低,难以广泛应用于野外种群个体识别与监测^[13-14]。综上所述,过去虽然围绕斑海豹种群(个体)调查监测进行了各种尝试,但这些方法不能实现种群水平上斑海豹个体的有效鉴别。亟需探索高效、便捷、低成本和可操作性的种群个体识别和监测方法。

近年来,基于面部特征的人脸识别启发了野生动物个体识别的新思路。在计算机图像识别领域,深度学习技术的发展为计算机图像特征学习提供了有力支撑,这一交叉技术的应用已经在许多动物研究中开展。据现有研究报道,大熊猫、中华白海豚、港海豹等珍稀濒危物种已经成功基于图像资料和深度学习模型进行了个体识别,该技术具有突破传统斑海豹监测识别方法局限性的潜力^[15-20]。

从深度学习技术的角度来说,用于计算机学习的图像需要达到一定的数量才能够满足识别网络的训练要求,这也是其“大数据”的特征之一^[21]。然而,斑海豹虽然是全国海洋生物多样性保护的旗舰物种^[22],但与中华白海豚和大熊猫等物种相比,其调查基础相对薄弱,加之该物种对人类活动十分警觉,目前积累的野外影像资料的数量和质量不足以精准构建斑海豹个体识别模型以及进行最优模型探索^[23]。

鉴此,考虑到斑海豹个体识别技术开发的重要性和图像特征学习的可迁移性,本研究以青岛海昌极地海洋世界的 14 只豢养斑海豹作为研究对象,结合深度学习技术,探索和构建针对斑海豹的个体识别方法,以支撑渤海野外斑海豹种群的监测和保护管理工作,同时希望为其他海洋动物的个体识别提供方法和案例参考。

1 研究方法 with 数据来源

本节将系统介绍基于斑海豹面部特征的个体识别研究方法。首先,对数据集的构建过程进行了详细描

述,包括数据的采集与处理流程。其次,介绍了网络模型的结构设计、优化策略以及实验方案。研究流程如图 1 所示。

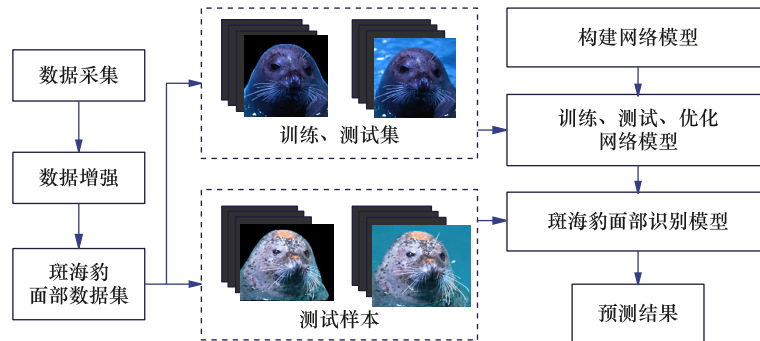


图 1 研究流程图

Fig.1 Research process diagram

1.1 数据采集

本研究所采用的斑海豹面部数据集来自青岛海昌极地海洋世界。通过采用便携手机终端,统一拍摄的部位、角度和时长等标准,在多个时间段对多个斑海豹进行了视频拍摄。为了确保数据集的质量,将拍摄的视频逐帧转换为静态图像,并从中筛选出清晰展现斑海豹完整面部特征的照片。在筛选过程中,排除图像模糊、面部不完整及拍摄角度不佳(如俯视、仰视或正面旋转角度超过 45°)的图像。最终,得到了包含 14 只斑海豹个体的共计 3914 张有效图像数据,每只斑海豹的面部图像数量为 100—300 张。图 2 展示了部分斑海豹面部图像。

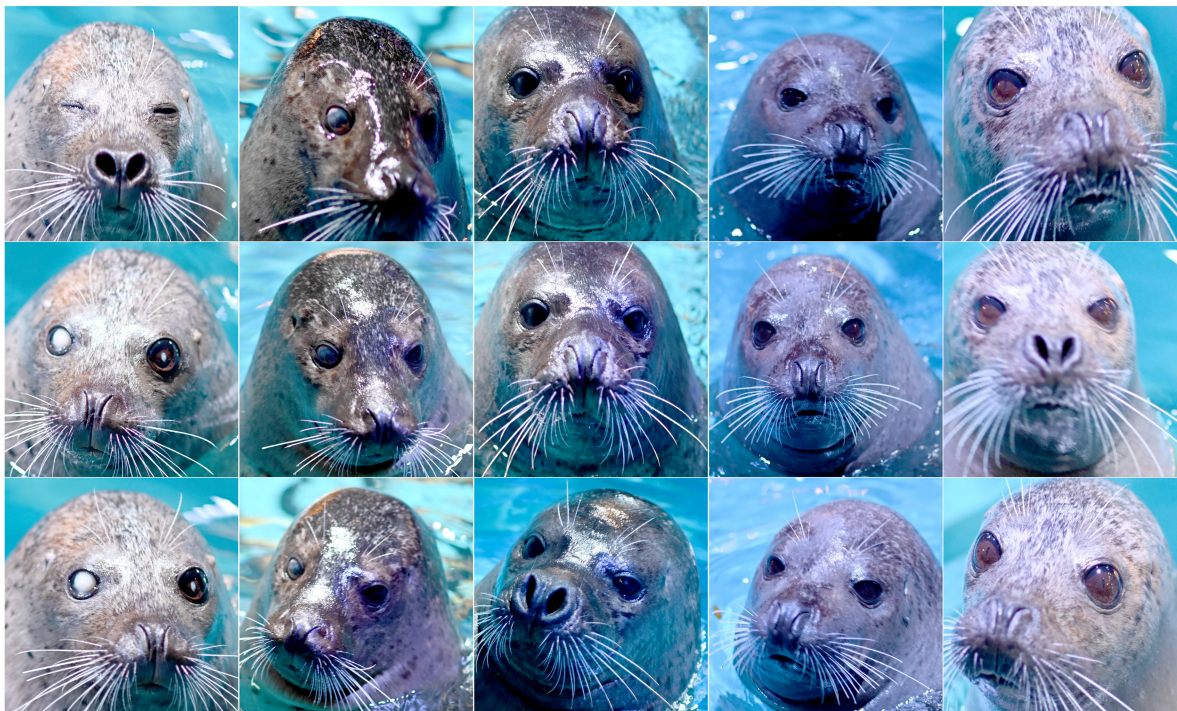


图 2 部分斑海豹面部图像

Fig.2 Partial facial images of spotted seals

1.2 数据增强

为了增强模型识别的泛化能力与鲁棒性,本研究采用多种数据增强技术对原始图像进行随机增强处理,比例设置为 1:4。具体增强方法包括旋转、平移、镜像反转、缩放及添加噪声等,以模拟斑海豹在现实环境中可能遇到的多种变化,例如光照条件的差异、面部姿态的变化等。这些操作丰富了训练数据的多样性,增强了模型对复杂环境的适应能力,从而在实际应用场景中,能够保持较高的识别准确性和稳定性。部分数据增强效果如图 3 所示。

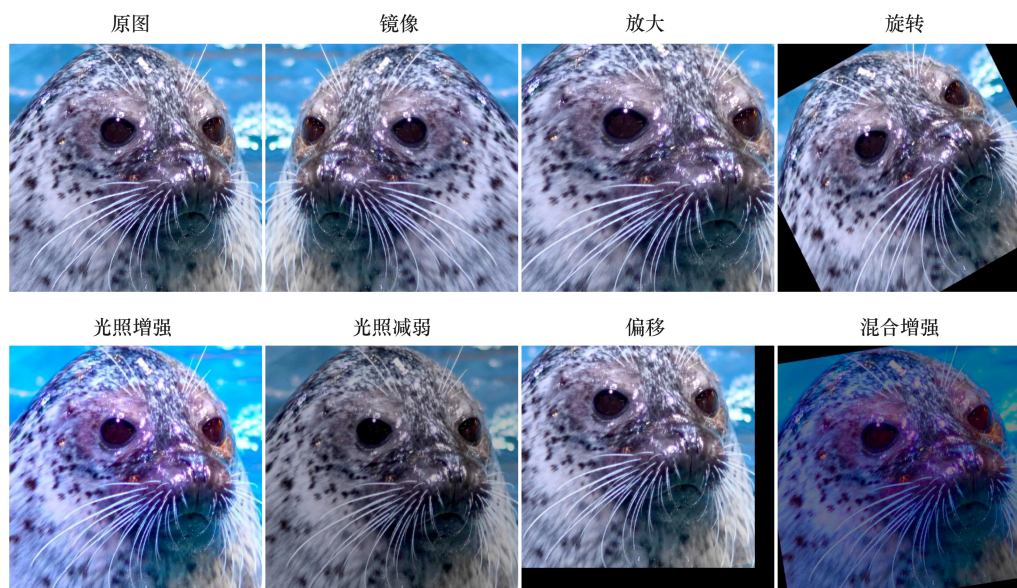


图 3 数据增强

Fig.3 Data Augmentation

1.3 斑海豹面部识别模型构建

考虑到斑海豹个体间的面部特征具有较大的相似性,本研究设计了一种基于卷积神经网络 (Convolutional Neural Network) 的斑海豹面部识别模型,该模型以 Darknet 网络模型结构^[24]为基础,融合了注意力机制以及多尺度的特征处理技术,实现对不同斑海豹个体的面部识别。网络结构如图 4 所示。

如图 5 所示,下采样卷积层 (Conv3×3_Downsampling) 采用 3×3 大小、步长为 2 的卷积核,用于实现特征提取与降维。特征强化层 (P-AC) 通过 SlimC2F 模块连接多个卷积层,对特征进行提取并融合,同时结合注意力机制,强化模型对于关键面部特征的识别能力。多尺度特征融合层 (SPPF) 通过整合不同尺度的特征信息,增强模型对尺度变化的适应性。全局平均池化层 (GAP) 对特征图每个通道进行平均池化,生成最终的特征向量。最后,将特征向量输入全连接层 (FC) 进行推理,输出对应的类别。

1.3.1 特征强化层

本研究提出一种基于深度学习的 P-AC (Post-AttentionCSP) 特征强化层,其结构如图 4 所示。P-AC 首先使用经过轻量化改良的 C2F 结构 (CSP Bottleneck with 2 Convolutions Fast)^[24] 对斑海豹面部图像的特征进行提取与融合,再引入 CBAM 模块 (Convolutional Block Attention Module)^[25] 进一步增强对斑海豹面部关键特征区域的识别。

1.3.2 高效特征提取与模型轻量化

本研究基于分割跨阶段处理方法^[26],设计了一种轻量化特征提取模块 (SlimC2F),其结构如图 4 所示。SlimC2F 将输入特征图按照通道数平均分割为两部分,一部分直接传递,另一部分进行残差学习^[27] 后与直接传递部分在通道维度上进行融合。这种处理方式加速了信息的传递,增强了模型的特征提取能力。此外,为

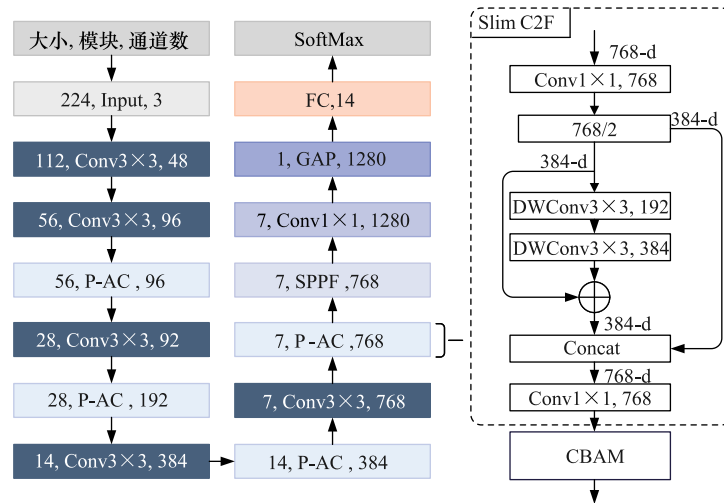


图 4 DarkNet 改进网络结构

Fig.4 Improved network structure of DarkNet

Conv3×3:卷积核大小为 3×3 的标准卷积层 3×3 Convolution Layer;P-AC:特征强化层 Post-AttentionCSP;SPPF:快速空间金字塔池化 Spatial Pyramid Pooling-Fast;Conv1×1:卷积核大小为 1×1 的标准卷积层 1×1 Convolution Layer;GAP:全局平均池化层 Global Average Pooling;FC:全连接层 Fully Connected Layer;DWConv3×3:卷积核大小为 3×3 的深度可分离卷积层 3×3 Depthwise Separable Convolution ;CBAM:卷积块注意力模块 Convolutional Block Attention Module;768-d:768 个通道数 768-Dimensional;Concat:拼接 Concatenation

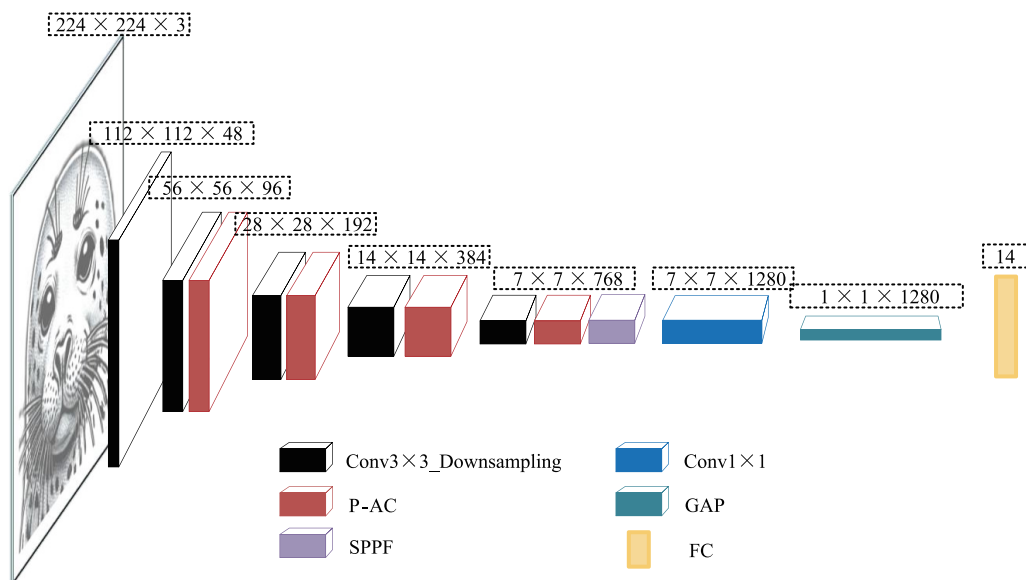


图 5 网络结构可视化

Fig.5 Visualization of network structure

Conv3×3_Downsampling:卷积核大小为 3×3 的下采样卷积层 3×3 Convolution with downsampling

了进一步降低计算复杂度,残差块中采用 3×3 深度可分离卷积 (Depthwise Separable Convolution)^[28],其结构如图 6 所示。深度可分离卷积将标准卷积拆分为深度卷积 (针对每个输入通道进行卷积) 和逐点卷积 (1×1 卷积,用于组合来自深度卷积的特征) 两个步骤。以图 4 中的 SlimC2F 结构为例,从输入特征图的 384 通道到输出特征图的 192 通道,采用标准卷积的参数量为 $(3 \times 3 \times 384) \times 192 = 663552$ 。而采用深度可分离卷积所需参数量为 $(3 \times 3 \times 1) \times 384 + (1 \times 1 \times 384) \times 192 = 77184$,极大降低了模型的参数量和计算需求。

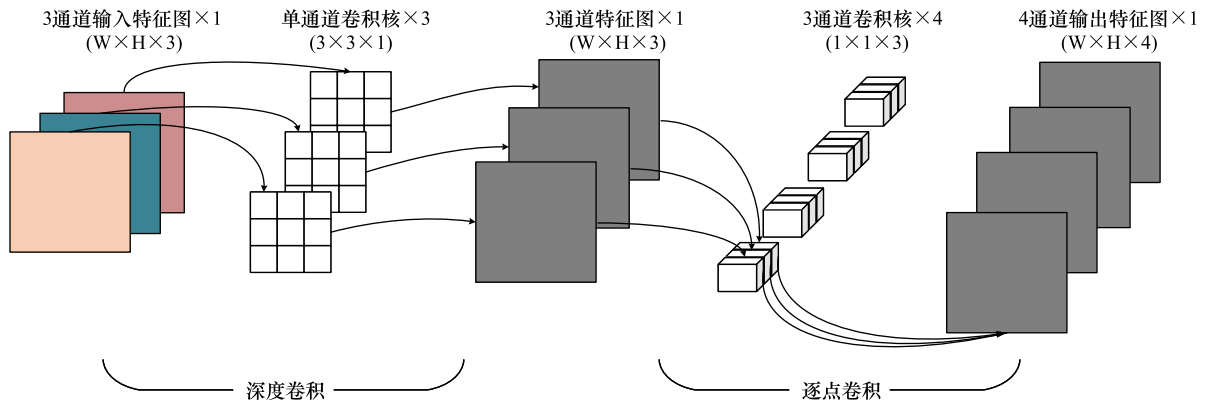


图6 深度可分离卷积结构

Fig.6 Depthwise separable convolution structure

W×H×3; 宽×高×通道数 3 Width×Height×3

1.3.3 通道与空间注意力增强

在斑海豹面部识别任务中,准确定位具备辨识度的局部特征信息对于提升模型性能至关重要。为此,模型中引入卷积块注意力模块 CBAM。如图 7 所示,CBAM 模块通过在通道和空间维度上对输入特征图施加注意力机制,强化了对关键通道特征的响应,突出了输入特征图中的重要区域,从而有效提高了模型对关键辨识特征的提取能力。

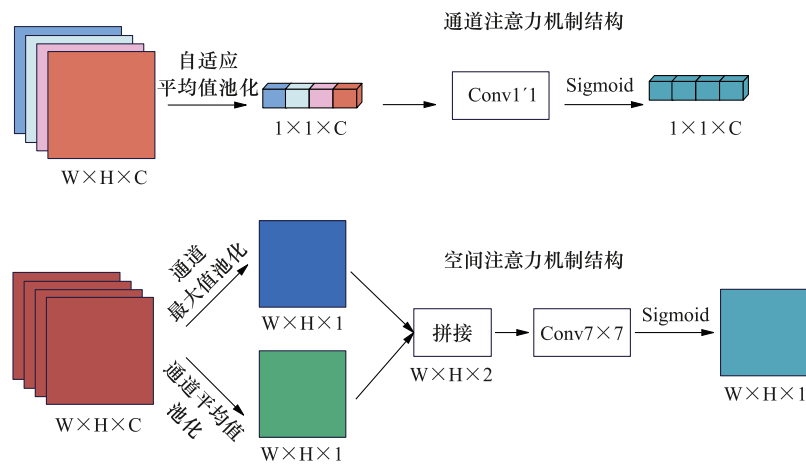


图7 通道和空间注意力机制结构

Fig.7 Channel and spatial attention mechanism structure

Sigmoid: 一种非线性函数

如图 8 所示,给定输入特征图 $F \in R^{W \times H \times C}$,会依次与 CBAM 模块生成的通道注意力权重 $M_c \in R^{1 \times 1 \times C}$ 、空间注意力权重 $M_s \in R^{W \times H \times 1}$ 进行逐元素相乘操作,其过程可以总结为以下表达式:

$$F' = M_c(F) \otimes F$$

$$F'' = M_s(F') \otimes F'$$

式中, $\otimes$ 表示对应位置逐元素相乘操作, F'' 是注意力模块的最终输出。

1.3.4 多尺度特征融合

快速空间金字塔池化 SPPF (Spatial Pyramid Pooling-Fast)^[29],通过不同尺寸的池化窗口捕捉图像中的多尺度特征,进一步增强模型鲁棒性。如图 9 所示,针对最终 P-AC 模块输出的 7×7 大小特征图,为了更好地匹

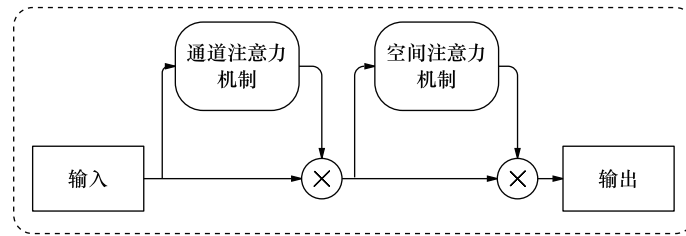


图8 卷积块注意力模块

Fig.8 CBAM Module

配斑海豹面部识别任务中特征图的多尺度信息,将传统SPPF结构中的池化核大小调整为 3×3 ,并串联三个步长为1的池化层,实现 1×1 、 3×3 、 5×5 、 7×7 的逐级增长感受野。通过该方法,模型不仅能够更有效地捕获局部细节特征(1×1 、 3×3 感受野),而且能够融合中等范围的上下文信息(5×5 感受野)和理解特征图的全局布局(7×7 感受野),提升了模型对图像多尺度特征的捕捉能力。

1.4 实验设计与评估方法

1.4.1 实验环境

实验在配置有 Intel Core i5-13600KF 处理器、NVIDIA GeForce RTX 4080 (16G) 显卡电脑上进行,操作系统为 Windows 10,编程语言为 Python,并基于 Pytorch1.13 深度学习框架。

1.4.2 实验参数

表1是模型训练时的超参数设置,训练前将数据集按照8:1:1比例随机划分为训练集、验证集、测试集。在训练过程中,学习率的调节采用余弦衰减^[30]。

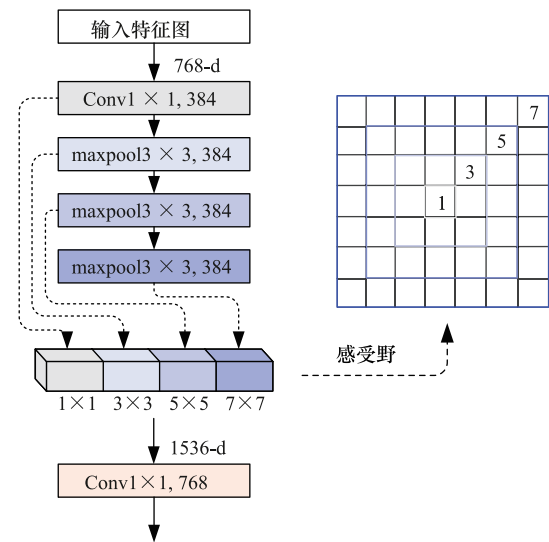


图9 改进快速空间金字塔池化结构

Fig.9 Improved SPPF structure

Maxpool 3×3 ; 3×3 最大池化层 3×3 Max Pooling

表1 超参数设置

Table 1 Hyperparameter Settings

参数 Hyperparameter	初始值 Initial value	参数 Hyperparameter	初始值 Initial value
学习率 Learning rate	0.001	批次样本量 Batchsize	128
动量 Momentum	0.937	优化器 Optimizer	SGD
权重衰减系数 Weight decay	0.0005	输入图像大小 Imagesize	$224\times 224\times 3$

1.4.3 评价指标

模型性能评估使用准确率 (Accuracy)、精准率 (Precision)、召回率 (Recall) 和 F1 分数 (F1-score, 精确率和召回率的调和平均数) 等评价指标。计算方式如下:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

式中, TP 、 FP 、 FN 和 TN 分别代表正确识别的正样本数、错误识别的正样本数、错误识别的负样本数和正确识别的负样本数。

1.4.4 实验方案

为评估所构建的斑海豹面部识别模型的有效性及其性能,本研究设计并进行了多组实验。初步实验采用消融研究方法,通过移除或替换模型中的关键组件来评估各组件对模型整体性能的贡献,从而确定模型的最优配置。随后,将最优配置模型与当前领域内的主流模型进行性能对比,探究其性能特点及有效性。接下来,实验进一步聚焦于背景信息对识别效果的影响,分析模型在不同背景条件下的性能变化。最后,通过系统地缩减训练数据集规模,研究数据量对模型性能的影响,以验证该模型在数据有限场景中的适用性。

2 结果与分析

本节经过一系列实验验证,成功构建了一个高效的斑海豹面部识别模型。该模型的识别准确率达到 98.35%,其精确率、召回率以及 F1 分数达到 98.30%、98.31%、98.31%,展现了该模型在相关领域内的实际应用潜力。

2.1 不同模块影响分析

为评估模型结构中不同模块对整体识别效果的影响,本节进行了多组消融实验,具体结果如表 2 所示。

表 2 不同模块对识别结果的影响

Table 2 Impact of different modules on recognition results

方法 Method	轻量化特征 提取结构 SlimC2F	卷积块注意力 结构+轻量化特征 提取结构 CBAM+SlimC2F	轻量化特征提 取结构+卷积块 注意力结构 SlimC2F+CBAM	快速金字塔 池化结构 SPPF	准确率 Accuracy/%	精确率 Precision/%	召回率 Recall/%	F1 分数 F1-score/%
①	√				87.75	87.81	90.02	88.90
②		√			91.78	92.07	92.98	92.52
③			√		96.35	96.40	96.45	96.43
④	√			√	90.97	91.19	92.25	91.72
⑤		√		√	94.17	93.86	94.57	94.21
⑥			√	√	98.35	98.30	98.31	98.31

从表 2 可以看出,在 SlimC2F 模块的基础上引入 CBAM 和 SPPF 模块明显提升了模型的性能。特别是 SlimC2F-CBAM+SPPF 配置(本文方法)取得了最佳识别效果,四项指标分别达到 98.35%、98.30%、98.31%、98.31%。与仅使用 SlimC2F 模块的实验①相比,实验②(SlimC2F 前接 CBAM)和实验③(SlimC2F 后接 CBAM)均显示出 CBAM 模块的引入提高了模型的识别准确率,且后接 CBAM 表现更好。实验③(SlimC2F+CBAM)与实验⑥(SlimC2F-CBAM+SPPF)对比验证了 SPPF 模块的正向效应。综上所述,通过融合 SlimC2F、CBAM 和 SPPF 模块,有效地提高了网络的识别性能。

2.2 对比实验

2.2.1 不同模型性能对比分析

为充分评估所提出模型的现实性能表现,本研究选取了深度学习领域内多个主流模型进行对比实验,包括 Swin Transformer^[31]、EfficientNetV2^[32]、MobileNetV2^[33] 和 VGG16^[34] 等。对比实验将分别采用自建数据集与公共数据集,均在相同的实验条件下进行。

如图 10 所示,在自建数据集上,各模型训练达到收敛后,本研究提出的模型在识别准确率与模型大小方

面表现出明显优势。

表 3 展示了各模型在测试集上的性能指标。从表中可以看出,本研究提出的模型在保持较低参数量的同时,准确率、精确度、召回率以及 F1 分数等多个评价指标上均表现优异,在一定程度上验证了该模型的有效性。

为进一步验证模型性能,继续采用 Caltech-256 公共数据集(256 类,30607 张图像)对模型进行训练,以各模型达到收敛时的 Top-1 准确率和 Top-5 准确率作为评估指标。如表 4 所示,本研究提出的模型在 Top-1 和 Top-5 准确率上均取得领先,进一步证明了其有效性。

2.2.2 有背景与无背景数据集对比分析

在实际应用中,斑海豹面部识别能力往往受到多种因素的影响,其中背景信息是一个常见且重要的影响因素。本研究所使用的斑海豹面部数据集多以水中或地面自然场景为背景,这些背景可能会引入额外的干扰,影响模型对斑海豹面部特征的识别能力。为了更准确地评估背景信息对模型性能的影响,本研究构建了一个无背景的斑海豹面部数据集,并进行了对比实验,部分数据集如图 11 所示。

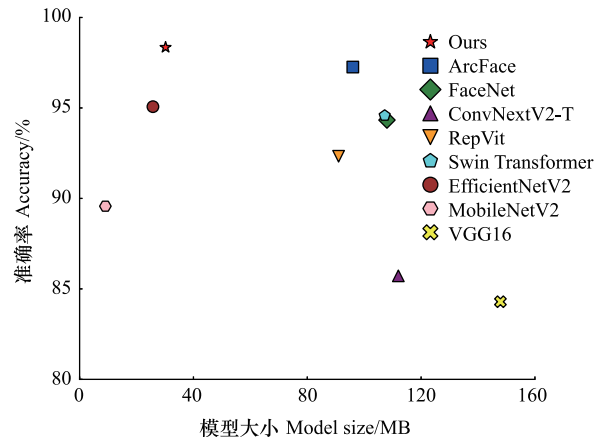


图 10 各模型在自建数据集上的识别准确率

Fig.10 Recognition accuracy of different models on the self-built dataset

图例为各模型名称

表 3 不同模型性能对比(自建数据集)

Table 3 Performance comparison of different models (self-built dataset)

序号 Number	模型 Model	准确率 Accuracy/%	精确率 Precision/%	召回率 Recall/%	F1 分数 F1-score/%	模型大小 Model size/MB
①	VGG16	84.29	84.66	89.29	86.91	147.87
②	MobileNetV2	89.57	89.84	91.24	90.53	9.05
③	EfficientNetV2	95.07	94.94	95.31	95.12	25.76
④	Swin Transformer	94.58	94.34	94.93	94.63	107.2
⑤	RepVit	92.33	92.40	92.21	92.30	91
⑥	ConvNextV2-T	85.72	85.91	85.63	85.77	112
⑦	FaceNet	94.33	94.50	94.20	94.35	108
⑧	ArcFace	97.26	97.40	97.15	97.27	96
⑨	Ours	98.35	98.30	98.31	98.31	30.2

表 4 不同模型性能对比(公共数据集)

Table 4 Performance comparison of different models (public dataset)

序号 Number	模型 Model	Top-1 准确率 Top-1 accuracy/%	Top-5 准确率 Top-5 accuracy/%	序号 Number	模型 Model	Top-1 准确率 Top-1 accuracy/%	Top-5 准确率 Top-5 accuracy/%
①	VGG16	68.02	85.75	⑥	ConvNextV2-T	67.19	84.85
②	MobileNetV2	71.08	87.37	⑦	FaceNet	76.60	90.09
③	EfficientNetV2	78.64	91.12	⑧	ArcFace	79.23	91.25
④	Swin Transformer	76.29	89.90	⑨	Ours	80.70	92.63
⑤	RepVit	75.44	89.34				

对比实验分别使用有背景和无背景的斑海豹数据集进行模型训练,并在相反的测试集上进行性能评估。为确保实验的公正性和准确性,采用本研究所提出的模型架构和统一的训练策略,唯一变量为数据集的背景信息。测试结果如表 5 所示。



图 11 部分无背景斑海豹面部图像
Fig.11 Partial facial images of spotted seals without background

表 5 基于不同数据集的模型性能对比
Table 5 Model performance comparison based on different datasets

方法 Method	训练集 Training set	测试集 Test set	准确率 Accuracy/%	精确率 Precision/%	召回率 Recall/%	F1 分数 F1-score/%
①	有背景	有背景	98.35	98.30	98.31	98.31
②	无背景	无背景	98.84	98.83	98.80	98.82
③	有背景	无背景	91.57	91.92	92.53	92.22
④	无背景	有背景	95.80	95.91	95.83	95.87

从表 5 中可以发现,背景信息对模型识别性能有较大影响。当训练集和测试集背景一致时,模型识别性能较高;背景信息不一致时,性能则下降。特别是实验③从有背景向无背景转换的测试场景中,性能下降最为显著。

为了解决背景不一致导致的识别性能下降问题,本研究采用迁移学习方法^[35]对网络进行交叉训练。首先采用有背景的数据集进行训练,并将训练后的参数迁移至无背景数据集进行进一步训练。从表 6 可以看到,采用迁移学习的训练方法弥补了模型在处理不同背景信息时的不足,进一步提升了斑海豹面部识别的准确性与鲁棒性。

表 6 基于混合数据集的模型性能
Table 6 Model performance based on mixed dataset

方法 Method	训练集 Training set	测试集 Test set	准确率 Accuracy/%	精确率 Precision/%	召回率 Recall/%	F1 分数 F1-score/%
①	混合数据集	有背景	98.70	98.65	98.68	98.67
②	混合数据集	无背景	98.55	98.50	98.53	98.52

2.2.3 数据集大小对比分析

为探究斑海豹面部识别模型在不同数据集规模下的性能变化,本研究设计了一项最小数据集测试实验。在当前的数据集中,每只斑海豹的平均面部图像数量约为 279 张。然而在实际应用中可能面临数据集规模更小或数据获取成本较高的情况。因此,本实验旨在确定当每只斑海豹的面部图像数量减少时,模型识别准确率能维持稳定的临界点。

实验采用本文提出的模型和训练策略,仅改变数据集的数量作为变量。实验前,将有背景数据集按照 9:1 比例随机划分为训练集、测试集。在之后的每次训练中,训练集的图像数量按个体等比例减少 10%,并重新进行训练与测试。最终测试结果如图 12 所示。

实验结果表明,随着训练数据集数量的逐渐减少,模型的测试准确率呈现出逐步下降的趋势。然而,在达到 40%这一临界点之前(平均每只斑海豹约为 100 张图片),即使减少了图片数量,模型的性能也能保持相对稳定。从本实验可以得出,在一定程度的减少训练数据,模型能够保持较高的识别准确率。但数据集规模的过度缩减,将对模型性能产生显著影响。这一结果为实际应用中在有限数据集下优化模型提供了重要参考。

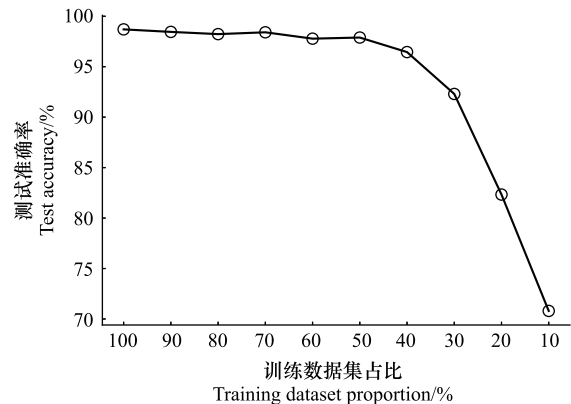


图 12 训练数据集占比与测试准确率曲线

Fig.12 The curve of training dataset proportion and test accuracy

2.3 斑海豹面部特征可视化

为了更深入地理解模型在进行斑海豹面部识别时所依赖的关键特征,本研究采用梯度加权分类激活映射(Grad-CAM)技术^[36],通过生成热力图实现特征的可视化。热力图直观地揭示了模型在识别斑海豹面部图像时关注的重点区域,从而显著增强了模型的可解释性。部分可视化结果如图 13 所示。

从热力图中可以观察到,不同斑海豹个体的面部识别特征存在差异,但主要集中在眼睛与鼻子等部位,这表明在斑海豹面部识别的过程中,眼睛和鼻子等部位的位置和形状可能是模型判断个体身份的重要依据,这为进一步优化模型和提升识别准确度提供了思路与方向。

3 结论与讨论

3.1 结论

本研究以青岛海昌极地海洋世界的 14 只斑海豹作为研究对象,针对斑海豹种群特征的多样性和自然环境中的各种挑战,使用卷积神经网络实现面部图像识别。通过采用多种策略优化模型,有效提升了识别准确率和模型的鲁棒性。实验结果显示,斑海豹面部识别模型在自建背景数据集上识别准确率达 98.35%,在混合数据集上的识别准确率达 98.70%。

就从事斑海豹保护和研究的相关工作人员而言,通过高效准确的个体识别技术,能够帮助相关人员进行斑海豹个体跟踪、行为研究以及种群管理等方面节省时间和人力资源。同时,该模型在斑海豹研究上的有效应用,也为其他动物个体识别提供新的技术手段和研究思路,有望在保护生态学的实践中得到更广泛的应用和研究。

3.2 讨论

3.2.1 技术方法

目前,在动物个体识别领域,基于卷积神经网络技术(CNN)的应用日益广泛。Kumar 等人^[37]通过结合卷积神经网络和深度置信网络(DBN)实现了对奶牛鼻口特征的识别,个体识别准确率达 98.99%,但精准捕捉到

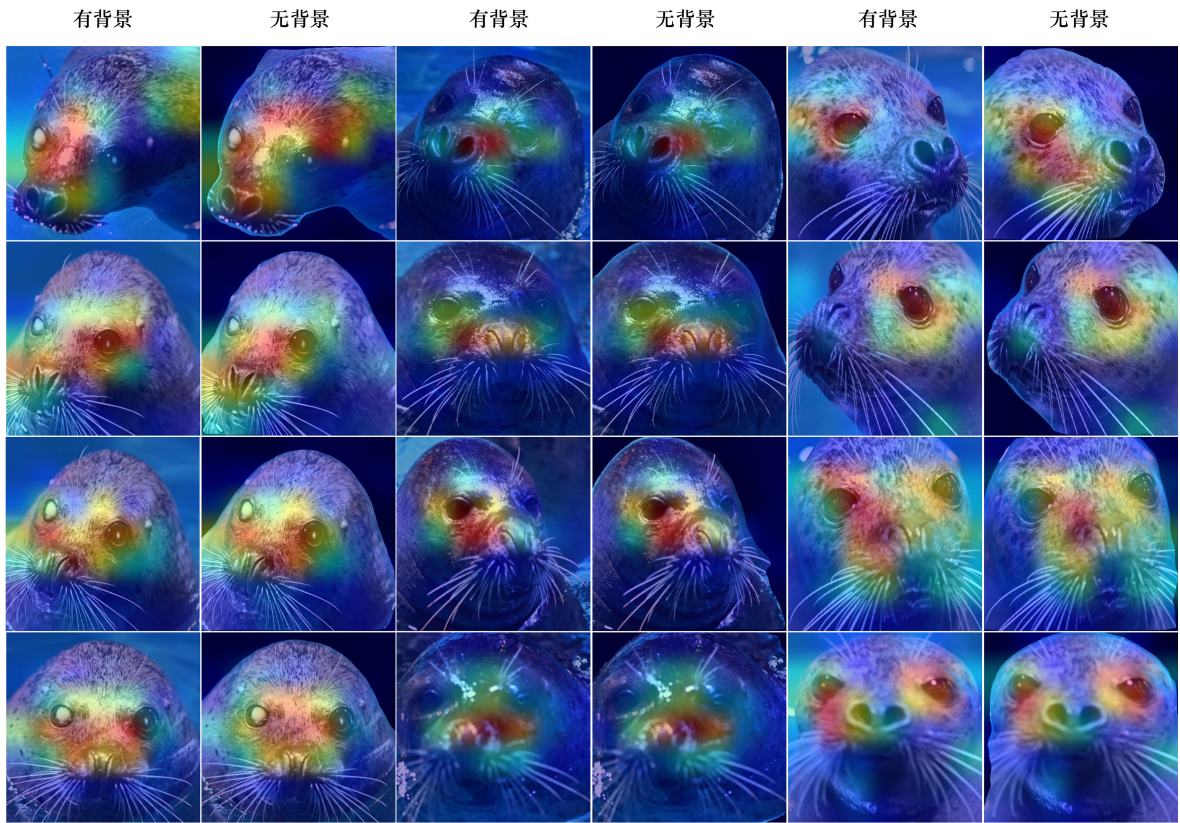


图 13 斑海豹面部识别可视化
Fig.13 Visualization of facial recognition for spotted seals

奶牛的鼻口图像存在较大难度。史春妹等人^[38]使用 ResNet、VGGNet 等网络实现了对东北虎右侧躯干条纹的识别,个体识别准确率达 97.01%,但在野生环境下,东北虎身体被污物覆盖、条纹模糊、树木遮挡或不利的拍摄角度等影响,获取稳定的右侧躯干条纹图像成为一大挑战。侯金等人^[16]使用 VGGNet 网络对大熊猫面部进行个体识别,识别准确率达 95%,但 VGGNet 由于其庞大的网络结构参数和高计算需求,在实际应用中,特别是在资源受限的环境中,可能会面临效率低下的问题。

本研究提出的网络结构,在保持高识别准确率的同时,具有更小的参数量,更贴合于移动设备和边缘计算环境,极大地提高了在户外工作条件下的实用性和可操作性。此外,本研究进一步为斑海豹个体自动化识别提供了可行性方案。如图 14 所示,通过进一步引入斑海豹面部检测模块,可以构建一个完整的自动化识别系统,从而极大地提升识别过程的效率和便捷性。

3.2.2 技术应用

斑海豹是典型洄游海洋动物,通过基于图像深度学习的个体识别技术在关键繁殖栖息地和洄游通道进行常态化定点监测,有助于揭示斑海豹洄游的主要驱动力(生殖或索饵),社会结构和社会行为特征,种群动态(如数量不变、个体变化,则种群稳定、健康),全球八大繁殖区间的种群交流情况,以及气候变化影响。该技术属于非入侵、低成本、易开展的海洋监测调查方法。通过同步选择时间段进行定点拍摄识别,可以避免重复计数,确定同时间个体数量,对于摸清渤海黄海斑海豹种群这一边缘小种群底数具有重要意义。此外,在未来进行野外实际监测中,该方法可以配合现有条件进行实际应用。一方面,自 2021 年以来,已有超过 100 只斑海豹,包括在野外受伤得到救助并康复的个体及人工繁育的幼体,被放归自然海域,其中部分个体安装了卫星信标跟踪器,这使得我们可以在后期进行跟踪监测,并积累大量的图像数据。此外,众多海洋保护区(如辽东湾(斑海豹的主要繁殖区域)和长岛(斑海豹的重要停歇地)等)已经配备了远程高分辨率监控摄像头。今后,研

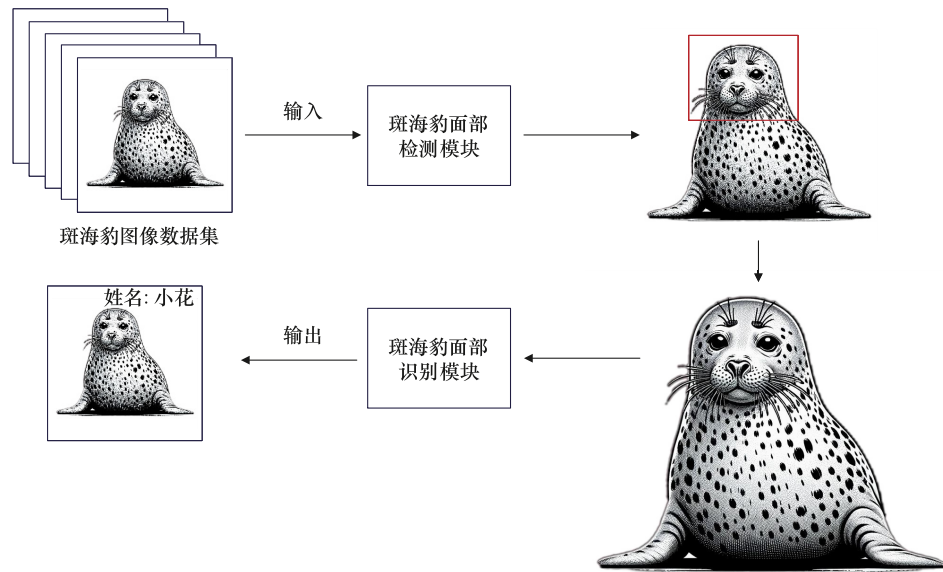


图 14 斑海豹面部检测与识别

Fig.14 Spotted seal facial detection and recognition

研究者可以依托该研究成果,与渔业渔政部门及野生动物保护救助机构等携手,对放归的斑海豹个体进行精确识别,并建立详细的图像数据库,对其生存状况及其生命周期行为进行详细分析。最后,通过与韩国、日本、俄罗斯等国家的合作,能够通过长期的监测和个体识别技术,精确估算斑海豹的种群规模、结构,并对其种群动态进行分析和评估,进而提高斑海豹的保护与管理效率。

本研究的斑海豹个体识别技术转化应用于野外自然环境场景面临一定挑战,主要体现在数据采集难度显著增加,1)是野外光照条件复杂,可能会导致面部图像质量不稳定;2)是背景环境更加多样,例如水体反射、海浪干扰等都可能增加数据的噪声;3)是斑海豹在野外环境中的运动更加随机,持续收集同一个体的高质量图像具有较大难度。针对上述问题,本研究下一步将从以下几方面展开工作,以提高模型在野外个体识别场景中的实用性:

(1) 丰富野外场景数据集:通过部署采集设备(如水下摄像机或无人机)采集更多野外斑海豹的面部图像。同时,结合数据增强技术模拟野外环境中的变化(如光照变化、动态模糊等),提升模型的泛化能力。

(2) 建立统一的特征空间:利用豢养数据和新增的野外数据,通过提取斑海豹面部嵌入特征(Embedding),训练一个统一的特征空间,使模型能够将新个体的特征映射到该空间中。基于特征间的相似性(如余弦相似度),模型可判断新个体是否与已有个体匹配,或识别其为全新的个体。该方法有效支持零样本学习和少样本学习场景,减少对大量标注数据的依赖。

(3) 动态特征更新:利用已有数据对模型进行初始训练,并在新增数据场景下动态更新分类器和特征提取模块。定期对模型进行微调,确保模型既能够适应新数据,又保持对旧类别的良好识别性能。结合迁移学习,将圈养数据训练的初始模型迁移到野外场景,并通过少量野外标注数据微调优化。

(4) 多模态特征融合:除面部特征外,引入其他生物特征(如斑点特征)进行多模态信息融合,进一步提升模型在野外环境中的识别能力。

参考文献(References):

- [1] Zhuang H F, Shao F, Zhang C, Xia W C, Wang S Q, Qu F Y, Wang Z L, Lu Z C, Zhao L L, Zhang Z H. Spatial-temporal shifting patterns and *in situ* conservation of spotted seal (*Phoca largha*) populations in the Yellow Sea ecoregion. Integrative Zoology, 2024, 19(2): 307-318.
- [2] Murray N J, Ma Z J, Fuller R A. Tidal flats of the Yellow Sea: a review of ecosystem status and anthropogenic threats. Austral Ecology, 2015, 40(4): 472-481.

- [3] Yan H K, Wang N, Wu N, Lin W N. Abundance, habitat conditions, and conservation of the largha seal (*Phoca largha*) during the past half century in the Bohai Sea, China. *Mammal Study*, 2018, 43(1): 1-9.
- [4] 农业部. 农业部关于印发《斑海豹保护行动计划(2017—2026 年)》的通知. https://www.moa.gov.cn/nybg/2017/djq/201802/t20180202_6136361.htm, 2017-08-31.
- [5] 生态环境部, 中国科学院. 关于发布《中国生物多样性红色名录—脊椎动物卷(2020)》和《中国生物多样性红色名录—高等植物卷(2020)》的公告. https://www.mee.gov.cn/xxgk/2018/xxgk/xxgk01/202305/t20230522_1030745.html, 2023-05-19.
- [6] 国家林业和草原局, 农业农村部. 国家林业和草原局 农业农村部公告(2021 年第 3 号)(国家重点保护野生动物名录). <https://www.forestry.gov.cn/main/5461/20210205/122418860831352.html>, 2021-02-05.
- [7] Thompson P M, Wheeler H. Photo-ID-based estimates of reproductive patterns in female harbor seals. *Marine Mammal Science*, 2008, 24(1): 138-146.
- [8] 王丕烈. 西太平洋斑海豹在黄渤海的分布、生态和资源保护. *海洋学报*, 1985, 7(2): 205-211.
- [9] 王丕烈. 渤海斑海豹的分布调查. *水产科学*, 1988, 7(4): 7-11.
- [10] Stewart B S, Leatherwood S, Yochem P K, Heide-Jørgensen M P. Harbor seal tracking and telemetry by satellite. *Marine Mammal Science*, 1989, 5(4): 361-375.
- [11] Vincent C, McConnell B J, Ridoux V, Fedak M A. Assessment of Argos location accuracy from satellite tags deployed on captive gray seals. *Marine Mammal Science*, 2002, 18(1): 156-166.
- [12] 韩家波, 鹿志创, 田甲申, 马志强, 王召会, 杨勇, 王勤国, 宋新然, 彭志平. 基于卫星信标跟踪的斑海豹放流效果研究. *兽类学报*, 2013, 33(4): 300-307.
- [13] 孙凡越, 高祥刚, 韩家波, 赫崇波. 利用线粒体 DNA 序列进行斑海豹鉴定的研究. *水产科学*, 2008, 27(5): 254-256.
- [14] 史雪威, 张晋东, 欧阳志云. 野生大熊猫种群数量调查方法研究进展. *生态学报*, 2016, 36(23): 7528-7537.
- [15] Birenbaum Z, Do H, Horstmyer L, Orff H, Ingram K, Ay A. SEALNET: Facial recognition software for ecological studies of harbor seals. *Ecology and Evolution*, 2022, 12(5): e8851.
- [16] Hou J, He Y X, Yang H B, Connor T, Gao J, Wang Y J, Zeng Y C, Zhang J D, Huang J Y, Zheng B C, Zhou S Q. Identification of animal individuals using deep learning: a case study of giant panda. *Biological Conservation*, 2020, 242: 108414.
- [17] Kim H W, Uh N, Lee S. First satellite tracking of a free-ranging spotted seal (*Phoca largha*) from the Baengnyeongdo Island. *Fisheries and Aquatic Sciences*, 2023, 26(11): 669-677.
- [18] Zheng X, Owen M A, Nie Y, Hu Y, Swaisgood R R, Yan L, Wei F. Individual identification of wild giant pandas from camera trap photos: a systematic and hierarchical approach. *Journal of Zoology*, 2016, 300(4): 247-256.
- [19] 侯金, 严淋露, 黎亮, 李玉杰, 廖玉杉, 张晋东. 野生大熊猫行为谱及 PAE 编码系统. *兽类学报*, 2020, 40(5): 446-457.
- [20] 钟铭鼎. 厦门湾中华白海豚种群生态学[D]. 厦门: 自然资源部第三海洋研究所, 2021.
- [21] 李响, 刘明, 刘明辉, 姜庆, 曹扬. 基于样本个体差异性的深度神经网络训练方法. *软件学报*, 2022, 33(12): 4534-4544.
- [22] 国家林业和草原局调查规划设计院, 国家林业和草原局国家公园管理办公室, 清华大学, 等. 国家公园设立规范. 国家市场监督管理总局; 国家标准化管理委员会. 2021.
- [23] Crouse D, Jacobs R L, Richardson Z, Klum S, Jain A, Baden A L, Tecot S R. LemurFaceID: a face recognition system to facilitate individual identification of lemurs. *BMC Zoology*, 2017, 2(1): 2.
- [24] Ultralytics. YOLOv8. (2023-1-10) [2024-12-11]. <https://github.com/ultralytics/ultralytics>.
- [25] Woo S, Park J, Lee J Y, Kwon I S. CBAM: convolutional block attention module. *Computer Vision-ECCV 2018*. ACM, 2018: 3-19.
- [26] Wang C Y, Mark Liao H Y, Wu Y H, Chen P Y, Hsieh J W, Yeh I H. CSPNet: A New Backbone that can Enhance Learning Capability of CNN. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). June 14-19, 2020. Seattle, WA, USA. IEEE, 2020: 390-391.
- [27] He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [28] Chollet F. Xception: deep learning with depthwise separable convolutions. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 1800-1807.
- [29] Ultralytics. YOLOv5. (2020-5-27) [2024-12-11]. <https://github.com/ultralytics/yolov5>.
- [30] Loshchilov I, Hutter F. SGDR: stochastic gradient descent with warm restarts. 2016: 1608.03983. <https://arxiv.org/abs/1608.03983v5>.
- [31] Liu Z, Lin Y T, Cao Y, Hu H, Wei Y X, Zhang Z, Lin S, Guo B N. Swin transformer: hierarchical vision transformer using shifted windows. 2021 IEEE/CVF International Conference on Computer Vision (ICCV). October 10-17, 2021, Montreal, QC, Canada. IEEE, 2021: 9992-10002.
- [32] Tan M X, Le Q. Efficientnetv2: Smaller models and faster training. In: JMLR. The 38 International conference on machine learning. Vienna, Austria; JMLR, 2021: 10096-10106.
- [33] Sandler M, Howard A, Zhu M L, Zhmoginov A, Chen L C. MobileNetV2: inverted residuals and linear bottlenecks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 4510-4520.
- [34] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014: 1409.1556. <https://arxiv.org/abs/1409.1556v6>.
- [35] Zhuang F Z, Qi Z Y, Duan K Y, Xi D B, Zhu Y C, Zhu H S, Xiong H, He Q. A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 2021, 109(1): 43-76.
- [36] Selvaraju R R, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. *International Journal of Computer Vision*, 2020, 128(2): 336-359.
- [37] Kumar S, Pandey A, Sai Ram Satwik K, Kumar S, Singh S K, Singh A K, Mohan A. Deep learning framework for recognition of cattle using muzzle point image pattern. *Measurement*, 2018, 116: 1-17.
- [38] 史春妹, 谢佳君, 顾佳音, 刘丹, 姜广顺. 基于目标检测的东北虎个体自动识别. *生态学报*, 2021, 41(12): 4685-4693.