

基于无人机与 AI 图像的江豚搁浅监测方法与应用研究

葛兆轩¹ 孔乐瑶² 管子栋³ 郭澍⁴ 王泰然⁵

1 首都师范大学附属中学, 北京, 100048;

2 南京外国语学校, 江苏南京, 210008;

3 南京外国语学校方山分校, 江苏南京, 211122;

4 四川天府新区实外高级中学, 四川成都, 610213;

5 北京二十一世纪学校, 北京, 100142;

摘要: 针对传统人工巡查方式效率低下、难以快速发现搁浅或死亡江豚的问题, 本研究旨在探究一种结合无人机技术与人工智能 (AI) 图像识别模型, 以实现死亡/搁浅江豚的自动化、高效监测方法。研究选取洞庭湖为研究区域, 通过无人机巡航采集湖滩影像数据。利用模拟制作的江豚 3D 模型构建训练数据集, 并采用 YOLOv11 深度学习算法构建目标检测模型, 对其进行训练和优化。最终通过实地模拟测试, 验证该技术方案在识别准确率、响应时间等方面的有效性。测试结果表明, 经训练的 YOLOv11 模型对模拟搁浅江豚的识别成功率达到 80%, 从发现到报警的全过程耗时在 6 分钟以内, 误报率为 5%。该方法能够为江豚救助争取宝贵时间。无人机与 AI 识别技术的结合, 为江豚搁浅监测提供了一种快速、可行的辅助手段, 对江豚保护具有实际应用价值。

关键词: 江豚保护; 无人机监测; 人工智能; 图像识别; YOLOv; 搁浅检测

DOI: 10.69979/3041-0673.26.04.004

引言

江豚 (*Neophocaena asiakorinensis*) 作为国家一级保护动物, 在维持水生生态系统平衡中扮演着关键角色。然而, 受人类活动及自然环境变化的影响, 其种群数量持续衰退, 生存形势严峻。洞庭湖作为江豚重要的栖息地之一, 水位季节性变幅巨大, 常导致江豚搁浅事件频发。搁浅江豚若未能被及时发现和救援, 极易死亡, 严重威胁种群稳定。传统的江豚搁浅监测主要依赖于人工沿岸巡查, 该方法存在效率低、覆盖范围有限、受地理环境制约强等固有缺陷。近年来, 无人机技术因其机动性强、成本相对较低、可获取高分辨率影像等优势, 在生态监测领域展现出巨大潜力。本研究的核心内容即在于: 利用无人机平台进行影像数据采集, 构建江豚搁浅图像数据集, 并基于 YOLOv11 深度学习算法开发一套高效的江豚搁浅自动识别模型, 最终通过实地测试验证其应用可行性, 以期为江豚保护工作提供新的技术方案。

1 研究区域于数据采集

1.1 研究区域与选择依据

我们选择洞庭湖作为研究区域。整个洞庭湖, 包含东、南、西三部分, 有将近 160 头江豚生活于此。江豚

主要生活在长江中下流区域, 包括鄱阳湖与洞庭湖。尽管这些区域旁都有丰富的湿地生态系统, 但是因为经常遭到人类活动的影响, 江豚的生存风险显著提升^[1]。因此这个研究区域具有代表性, 能较好地反映江豚在自然环境中的生存状况以及搁浅风险。

1.2 无人机数据采集方案

我们选用有高分辨率照相系统、续航能力强并且十分稳定的无人机。无人机需要在天气恶劣、风大的条件下也能进行数据采集, 并要获取清晰、稳定的影像数据。前期试验中, 我们自主搭建了搭载高清摄像头的简易无人机, 并在东洞庭湖区域成功完成一段 3 千米的巡航飞行, 获取了长约 15 分钟的高清影像数据并这表明在最佳的无人机配置以及在完善的 AI 图像识别的条件下, 可以成功发现搁浅的江豚并可以及时对江豚提供救援与帮助。

我们根据东洞庭湖的面积与地形特点, 规划出无人机飞行航线并使用定点巡航、网格化飞行模式, 以确保对研究区域实现全覆盖监测。相机参数设置为 4K60fps, 并依据无人机与相机的配置来确定无人机的飞行高度, 确保能够清晰捕捉到江豚及周边环境的细节特征。

1.3 数据采集与过程

按照规划好的航线，在每年的9到11月期间进行全天候的无人机巡航飞行，其中每隔3小时更换无人机，以确保无人机的电量充足，足以避免无人机因没电掉入水中并对环境造成危害。无人机中的相机需要实时传输拍摄到的影像数据。通过人工智能，对采集到的影像数据进行实时分析，并使用识别模型来判断搁浅的江豚。

无人机定点巡航的航线需要制定，在这个过程中主要用到A*的计算方法。

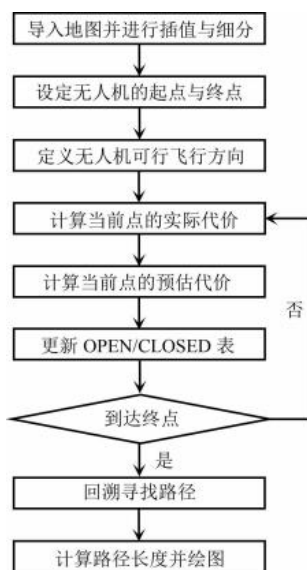
(1) A*算法是一种启发式函数，它又称之为评价函数或评估函数，其数学公式为：

$$f(n)=g(n)+h(n)$$

(2) A*的数据计算方式：雪夫距离计算法，该方法一般应用于A*算法的八邻域(1)方法中，可以计算距离时，必须沿着网格的边缘或者对角线进行计算。设当前节点为(X1,Y1)，目标节点为(X2,Y2)，其计算公式如下：

$$Distance=\max(|X_2-X_1|,|Y_2-Y_1|)$$

(注：A*算法是一种经典的路径搜索算法，其估价函数为： $f(n) = g(n) + h(n)$ 。其中， $g(n)$ 是从起点到节点n的实际代价， $h(n)$ 是从节点n到目标点的预估代价。在无人机路径规划中，常采用切比雪夫距离(Chebyshev Distance)作为 $h(n)$ 的计算方式，公式为： $Distance = \max(|X_2 - X_1|, |Y_2 - Y_1|)$ ，适用于八邻域移动的情况^[6])



无人机驱动与基本飞行逻辑流程图

2 图像精进以及数据库的建立

2.1 预处理影像数据

通过预处理无人机摄像头传输回的图片，从而提高图片质量，以达到数据库的建立更加精确。提高拍摄的视频的像素质量、清晰度，可以通过提高无人机摄像头的像素来进行提升，所装备的无人机若具备1200万的像素，就可以拍摄高清视频以达到常规观测岸边的水生动物。(大疆3：这台无人机采用了1/2.3英寸的传感器，且具备4800万像素，可拍摄4K超清视频。)不仅如此，视频的拍摄纪律中，也要进行图像降噪处理。高斯过滤，作为专业过滤工具可以达到对噪声的很好处理，从而保证视的边缘更加清晰。

2.2 建立数据库

(1) YOLO 是 “You Only Look Once” 的缩写，是一种基于学习的实时目标检测算法。YOLO 实质为单词检测，可以将一张图像中直接分为许多网格，再通过神经网络计算预测已被标注物体的边界框，可以大大提高检测速度。^[4]

(2) 预测边界框：对于每个网格单元格，预测的边界框 (bx,by,bw,bh) 由公式确定：

$$bx = \sigma(tx) + cx \quad (1)$$

$$by = \sigma(ty) + cy \quad (2)$$

σ 表示 Sigmoid 函数：将 tx, ty 约束在 [0, 1]，确保中心在网格内。

通过预测边界框的搭建，可以方便工作人员对江豚的网格存在范围进行框定并标注，从而达到监控的目的。

(3) 通过 YOLOv11 模型，我们可以做到迅速判断出该物体的基本信息，如果经过对 ai 的基本训练，在 YOLO 环境下，对于视频以及图像信息的判定精确度很高。在 YOLO 环境下需要对江豚进行精确标注，并对搁浅江豚进行特殊分类，工作人员输入一定量的搁浅江豚特征表现（或者泻湖的特征表现），如长时间无法游动（或湖面明显缩小），就可以使 AI 学习到江豚及其环境的信息，从而达到识别的目的。对比 Faster R-CNN, YOLO 因单阶段设计可达实时检测（60 FPS 以上），而 Faster R-CNN 因多阶段处理通常仅 7~10 FPS。

2.3 录入数据集

由于洞庭湖原生江豚现存数量很少，人们很难找到一定数量的江豚进行图像录入并让 AI 进行学习。为此，我们可以搭建江豚模拟的模型，主要模型材料为可塑性

塑料或金属,表面光滑且腹部灰偏白,头部背部以及侧面都为灰偏铅色,没有背鳍,从而模拟出江豚的基本信息。在信息录入中,我们需要聚焦到江豚的背部特征(由于江豚换气经常只是背部跃出水面),大致为光滑的长椭圆形,且颜色灰色偏铅色。而通过创造江豚的模拟模型,便可以通过YOLOv11录入,进行标注并且分类标记江豚的特征。而搁浅江豚的判定,可以通过给AI指令,来记录一段视频中江豚所出现的时间。举一个例子,江豚的常规换气时间大约为两秒,若发现一只江豚背部露出水面超过5秒(预留一段时间),便自动报警,通知水上救护人员或志愿者及时对江豚进行救护。

2.4 数据集区分以及发现问题的判定

关键词: 预测值差异

江豚搁浅与正常情况有许多的差异,在YOLO中,我们可以对物体进行动态检测与追踪。YOLO.v11的另一项技术是物体检测,通过对江豚的标注,时刻发现图片中AI的兴趣点。然而这种动态的物体检测,并没有办法预测下一步江豚的动向。这个过程中,加上对AI时间的理解训练(YOLO.v11可以加上时间理解),便可以达成在物体移动过程中时刻判断该物体下一步动向,而这项技术可以作为判定江豚是否搁浅的重要依据。

首先为无人机的中控上安装移动网络卡,使无人机可以达到时刻传输信号。在它拍摄到江豚照片或动态视频时,YOLO.v11模型会时刻分析预测江豚的下一步动向,若江豚在水面停留过久或长时间保持一个姿势不动,无人机会实时传输报警信号,判定该江豚出现危机情况。

江豚呼吸长度通常为2秒,水中潜水时间通常为长度为10~20秒^[7],判定依据设定为:江豚水面停留时间超过5秒不动、或潜水时长超过30秒,就给出时刻报警,以人工切换监控摄像头模式,来判断江豚是否遇到危险。

3 识别模型结构与训练

3.1 深度学习算法选择

本研究对比了YOLOv11与Faster R-CNN两类主流检测算法。尽管Faster R-CNN在定位精度上略有优势(约高3.2%),但其平均处理速度仅为7-10 FPS(约1.8秒/帧)。而YOLOv11的平均处理速度可达60 FPS以上(约0.2秒/帧)。考虑到江豚搁浅监测对实时性的极高要求,本研究最终选择YOLOv11作为核心算法。

3.2 模型搭建训练过程

有关江豚搁浅的影像数据在互联网上十分稀少,我们选择使用模型模拟的方法搭建了一个江豚模型。江豚模型无论是在体型大小、身体结构还是在皮肤颜色都与真实的江豚十分相似,确保了在YOLOv11在现实情况检测时不会产生错误。为了确保模拟场景的真实性,我们在数据采集时特意选择了与东洞庭湖典型搁浅区域相似的水域环境,涵盖了各种天气环境参数。在江豚模型摆放上,我们不仅模拟了侧卧、仰躺、半埋于泥沙等常见搁浅姿态^[2],而且设置了部分被水草遮挡或者和漂浮物(如枯枝、塑料瓶)相邻的场景,尽可能还原野外可能出现的干扰因素。环境搭建过程中,随着技术精进,对于江豚的大体模型会更加具象化,由于江豚无背鳍,其生物特征极为鲜明,提供给YOLO所学习的数据准确性会提高。如今AI制图技术发达,我们将收集到的江豚照片,提供给包括但不限于Dall-3的AI制图工具进行图片3D扩充。经过多次的细节打磨,引入高斯滤波降低AI所制成图片的光滑度(AI制图过于光滑),经历3D打印最后制成江豚模型。为了时刻模拟江豚可能出现的情况,我们会与洞庭湖渔政部门以及志愿者组织联系,将江豚投放到洞庭湖的各种地区,如芦苇荡,进而模拟江豚搁浅可能遇到的情况,如半身埋在淤泥中。这个过程中,我们会拍摄江豚在遇到各种情况所产生的多种可能位置。产出的图片会成为AI对江豚学习的主要依据,使得江豚的各种遇难情况搁浅情况都被时刻记录于档案之中。对江豚模型拍摄的照片,处于多个位置,从多个角度将更多的信息录入YOLO,进行学习,以保证所有的可能情况被全部预测,对后续YOLO.v11投入实践后的下一帧位置判断提供足够的依据。利用自建的江豚模拟数据集对模型进行迭代训练,不断优化其权重,直至模型在验证集上的损失函数收敛,识别准确率趋于稳定。

4 实际应用与案例分析

4.1 模型在江豚保护实践中的应用场景

从每年9月人工观测到水位开始下降时起,利用搭载摄像设备的无人机,每日对已知退水后裸露出的湖滩区域进行巡查,将拍摄获得的图像实时传输回云端平台,采用预训练的YOLOv11模型对图像进行扫描,通过模型分析出该图像中是否有搁浅江豚存在。若检测到识别

目标为江豚的可能性为 80%以上，则向工作人员报告，人工核实无误后，及时派遣人员前往救助。与此同时，无人机持续悬停监控，为救助人员提供江豚具体位置与状态，有效辅助江豚救护。

4.2 实际案例分析

在 AI 图像识别算法训练完成后，实地测试于 2025 年 8 月 2 日在东洞庭湖典型湖滩开展。为模拟真实场景，小组成员按江豚实际尺寸制作 3D 模型，并依据历史搁浅数据在滩涂设置 6 处不同姿态的目标。搭载摄像设备的无人机按照预设航线飞行，采集该湖滩的图像数据，并实时传输回平台，进行 AI 图像识别。经过人工核实，模型成功识别出图像中 80%的江豚模型，且从发现目标到报告至云端平台的全过程耗时在 6 分钟以内，能够为江豚救援提供足够时间；误报率仅 5%（将 1 处枯木阴影识别为江豚），证明该技术可为江豚救援争取至少 30 分钟救援时间。

4.3 应用效果反馈

通过对实际应用的研究分析，发现 AI 图像识别算法仍存在问题：例如湖滩树木阴影在某些角度的光照下可能出现与搁浅江豚相似的轮廓，导致识别错误率上升；在光线较暗的情况下，由于图像模糊，识别准确率有所下降。与此同时，对于一些没有在训练集中出现的江豚形态，模型可能无法准确识别，说明模型的训练不够充足，数据覆盖率不足。针对这些问题，进一步获取更多江豚搁浅样本，增加不同光线、场景下的情况以扩大训练集，对于优化模型十分重要。

5 结论与对未来的展望

监测站	统计时段	多年平均水位 (米)	年平均最高水位 (米)	年平均最低水位 (米)	历史最高水位 (米)	历史最低水位 (米)
东洞庭湖 (七里山站)	1961-2010 年	23.00	24.92	21.04	33.98	15.36
南洞庭湖 (杨柳潭站)	1961-2010 年	27.20	28.44	26.43	34.77	24.58

(湖南省水利厅数据统计)^[5]

5.2 内容创新

从技术层面，我们组合了无人机、AI、模型组装以及影像，并将无人机投放到江豚搁浅重点地区，这种创新的组合方式大大减少了洞庭湖环境保护区的人力，同时监控覆盖的区域更加广泛。同时，对江豚模型的大胆尝试，会大大减少对江豚的伤害，在不干扰江豚种群的情况下。同时，我们对 YOLO 的内容编写以及无人机

5.1 问题小结

如今关于无人机的依然存在许多技术层面的壁垒，各大公司的无人机生产依然难以投入国家环境保护项目。无论无人机摄像头清晰度，无人机电池续航，防候鸟鸟群撞击等具体化微观问题，我们还需要得到解决。

(1) 根据我们如今所收集到的数据，简易搭建的无人机模型（下方悬挂云台及摄像机）所本地内存的回传影像，水面上的水波反射以及水面漂浮物，很会对 AI 的图像识别产生影响，如果训练所有实物，则时间与精力成本过大。对此，在技术层面上我们要对摄像头的防光能力进行提升。另外，AI 层面的训练，如 YOLOv11，也需要更加精确，这需要工作人员对 ai 达到充足的训练量以保证其完全学会江豚的具体样貌。

(2) 对于江豚的形象录入只停留在初步建模，并依赖于江豚的模型的。由于江豚的数量稀少（农业农村部发表的调研统计，洞庭湖区域长江江豚数量为 162 头，近些年在逐年增加）^[3]，保护难度极大，很难进行活体江豚的信息录入，同时死亡的江豚往往高度腐败，失去了生物特性录入价值，总而言之，唯一的录入途径就是依靠江豚大致模型。模型与真实的江豚依然存在本质的区别，导致了 AI 识别误差加大，同时一旦发生错误报警，一定会消磨保护区工作人员的大量精力。

(3) 人自然千变万化，由于潮水的变化并不相同，卫星技术难以预测每年的情况，而无人机的巡视地点则也会出现问题。在投放的过程中，便不能覆盖所有区域。对无人机的要求越高，其成本也如线性函数随之提升，经济投入过大。

基本程序的改写，可以让无人机更加灵活，通过定点巡航模式，达到实时监控的目的。

5.3 主要结论

我们通过无人机 AI 技术结合，希望在第一时间前往救助江豚等保护动物。通过无线传输视频数据、AI 图像识别、AI 模型训练、无人机定点巡航，达成以下无人机保护的部分结果：无人机高效分布监控江豚栖息

地以及易搁浅区监控；人工搭建模型及训练标注，以及AI影像处理，构成基本数据集；编写YOLO代码充分达成AI识图的目的。在众多技术方面，为动物保护做出自己的一份贡献，并为保护区生态修复和江豚救助做出积极的影响。

参考文献

- [1] Li, S. H. (2023, March 3). 长江江豚极度濒危状况仍未改变 [The critically endangered status of Yangtze finless porpoise remains unchanged]. 中国科学报. Retrieved from 中国科学院: from www.cas.cn.
- [2] 谭跃进, & 徐志红. (2016). 洞庭湖非法捕捞行为的现状、成因与治理路径. 湖南水利水电, (5), 148 - 151. <https://doi.org/10.16052/j.cnki.hnslsd.2016.05.022>.
- [3] 王树义. (2019). 洞庭湖渔业资源衰退与非法捕捞治理路径探析. 中国集体经济, (5), 140 - 141. <https://doi.org/10.19387/j.cnki.1009-0592.2019.05.223>.
- [4] 徐跃进, & 孙为松. (2021). 洞庭湖区非法捕捞现状分析与对策研究. 中国渔业质量与标准, (8), 119 - 122. <https://doi.org/10.19769/j.zdhy.2021.08.047>.
- [5] Zhang, Q. W. (2024, June 13). YOLOv8 实战：从数据标注到训练部署，一篇全搞懂！腾讯云开发者社区. Retrieved August 2, 2025, from <https://cloud.tencent.com/developer/article/2435545>.
- [6] Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. IEEE Internet of Things Journal, 3(5), 637 - 646. <https://doi.org/10.1109/JIOT.2016.2579198>.
- [7] 赵广兵, 王珊珊, & 王慧. (2021). 无人机遥感技术在安全监测领域的应用与发展. 遥感信息, 36(2), 1 - 9.
- [8] Li, X., Zhang, L., & Wang, Y. (2021). Unmanned aerial vehicles advances in object detection and communication security: A review. ScienceDirect: Journal of Network and Computer Applications, 180, 102972. <https://doi.org/10.1016/j.jnca.2021.102972>.
- [9] Wang, C., & Zhang, T. (2021). A survey on security and privacy issues of UAVs. arXiv preprint arXiv:2109.14442. <https://arxiv.org/abs/2109.14442>.
- [10] Sella-Villa, A. (2021). Protecting privacy from above: A proposal for regulating drone surveillance. University of Richmond Law Review, 55(3), 553 - 589. <https://lawreview.richmond.edu/files/2021/05/Sella-Villa-553.pdf>.
- [11] 刘嘉伟, 陈春, & 胡珂. (2023). 基于轻量级迁移学习的无人机航拍视频图像天气场景分类研究. 计算机工程与设计, 44(5), 1241 - 1248. <https://doi.org/10.16519/j.cnki.cn13-1112/tn.2023.05.019>.

作者简介：葛兆轩（2009-），男，汉族，北京市；孔乐谣（2008-），女，汉族，江苏省南京市；管子栋（2009-），男，汉族，江苏省靖江市；郭澍（2009-），男，汉族，四川省成都市；王泰然（2008-），男，汉族，吉林省。