

Um Mapeamento Sistemático sobre Aprendizado de Máquina para a Classificação de Fotografias de Armadilhas Fotográficas

Diego T. Terasaka, Patricia C. de Souza

¹ Instituto de Computação
Universidade Federal de Mato Grosso (UFMT) – Cuiabá, MT – Brazil

diego.terasaka@sou.ufmt.br, patricia@ic.ufmt.br

Abstract. *Camera traps serve as valuable tools for biomonitoring, enabling the collection of large amounts of environmental data. However, this vast volume of data requires efficient processing to extract relevant information. This study presents a systematic mapping of recent publications that explore machine learning models for automating this task. Convolutional neural networks (CNN) based models dominate recent publications, with the YOLO series standing out as the most frequently employed. Additionally, the metadata associated with the images remains underutilized, despite its potential value.*

Resumo. *Armadilhas fotográficas atuam como ferramentas valiosas para o biomonitoramento, permitindo a coleta de grandes quantidades de dados ambientais. No entanto, esse grande volume de dados demanda um processamento eficiente para a extração de informações relevantes. Este estudo apresenta um mapeamento sistemático de publicações recentes que exploram modelos de aprendizado de máquina para a automação dessa tarefa. Modelos baseados em redes neurais convolucionais (CNNs) predominam publicações recentes, com a série YOLO se destacando como a mais frequentemente implementada. Observa-se ainda, um subaproveitamento de metadados potencialmente valiosos das capturas.*

1. Introdução

A utilização de armadilhas fotográficas para o biomonitoramento tem sido um valioso recurso para facilitar a aquisição de imagens e vídeos de animais em seus ambientes naturais de forma não invasiva. Essas imagens permitem o acompanhamento das influências humanas nos sistemas ecológicos [Buxton et al. 2018] e a coleta de dados taxonômicos pouco acessíveis [Oliver et al. 2023]. No entanto, sua implementação produz um grande volume de imagens a serem processadas para que a informação relevante seja extraída [Glover-Kapfer et al. 2019].

Para estas tarefas de classificação, a automação por meio de sistemas de visão computacional é capaz de acelerar esforços ecológicos de conservação diminuindo os custos e tornando a função mais acessível [Chalmers et al. 2023], além de possibilitar a expansão de escopo dos projetos de supervisão ambiental [Tuia et al. 2022].

Este mapeamento sistemático da literatura (MSL) tem como objetivo compilar publicações de anos recentes que explorem soluções baseadas em modelos de aprendizado

de máquina para tarefas de classificação de imagens obtidas por armadilhas fotográficas. Também são apresentadas as métricas alcançadas, a fim de fornecer um panorama sobre os resultados de métodos contemporâneos na automação dessa etapa em ações de biomonitoramento, além de destacar possíveis lacunas de opções não exploradas pelos trabalhos analisados.

Além desta introdução, três seções compõem este artigo: a seção 2 descreve a metodologia adotada para o MSL, detalhando o planejamento por trás dos parâmetros de busca e critérios de seleção. A seção 3 discute os dados obtidos, identificando os padrões notados. A seção 4 traz as conclusões do estudo, definindo o espaço para pesquisas futuras.

2. Metodologia

Como um MSL, este estudo procura reunir evidências a fim de estabelecer uma visão de alto nível sobre a área de conhecimento. Seguindo esta definição e estrutura apresentados por Kitchenham e Charters (2007), a metodologia deste estudo abre com a etapa de planejamento onde são definidas as questões de pesquisa e os critérios de seleção. Em seguida, esses critérios são aplicados às bases de dados para obter os trabalhos relevantes, que posteriormente serão analisados.

2.1. Planejamento

A elaboração das questões de pesquisa visa estabelecer as perguntas que orientarão todas as etapas subsequentes do estudo. A partir delas, são elaborados a *string* de busca e os critérios de seleção para garantir que os estudos analisados sejam relevantes para o escopo da pesquisa. Desta forma, este trabalho busca responder às seguintes questões:

1. Quais soluções baseadas em aprendizado de máquina têm obtido os melhores resultados na classificação de fotografias de armadilhas fotográficas?
2. O quão eficazes têm sido as soluções baseadas em aprendizado de máquina para a análise de fotografias de armadilhas fotográficas?

Com base nestas questões formuladas, os parâmetros de busca e a seleção subsequente foram definidos para garantir a inclusão de artigos que atendam aos seguintes critérios:

- (i) Referente à classificação de fotografias.
- (ii) Referente a fotografias de animais.
- (iii) Referente a modelos de aprendizado de máquina.
- (iv) Apresenta resultados de implementações dos métodos de aprendizado de máquina.
- (v) Publicado em um período recente.

A *string* de busca foi então composta por palavras-chave como parâmetros pertinentes aos critérios (i), (ii), (iii) e (iv) como listados na Tabela 1, enquanto o critério (v) foi efetivado através de filtros nas ferramentas de busca de cada fonte de dados. Os componentes “detection” e “model” são adicionados à *string* em casos onde os parâmetros originais não foram suficientes para uma filtragem robusta.

Tabela 1. Componentes da string de busca

Componentes definidos	CrITÉrios atendidos	<i>String</i> de busca base
"machine learning", "artificial intelligence"	(iii)	("machine learning" OR "artificial intelligence") AND ("trail camera" OR "camera trap" OR "camera trapping")
"trail camera", "camera trap", "camera trapping"	(i) e (ii)	
“detection”	(iii) e (iv)	
“model”	(iii) e (iv)	

2.2. Seleção

As buscas por publicações foram realizadas nas bibliotecas digitais IEEE Xplore utilizando a string de busca base e ScienceDirect utilizando a string com os dois componentes adicionais, resultando em 41 e 187 artigos respectivamente, para um total de 228 como ilustrado na Figura 1.

Após esta busca inicial, os resultados foram filtrados por ano de publicação conforme o critério de recência (v) estabelecido no planejamento. O recorte de tempo entre janeiro de 2022 e fevereiro de 2025 preservou os 165 artigos totais publicados nos últimos três anos. Por fim, uma triagem manual de acordo com os demais critérios de inclusão resultou na seleção de um total de 26 artigos como produtos finais.

3. Resultados

A Tabela 2 apresenta o conjunto final de artigos. Dos estudos selecionados foram extraídos os modelos de maior eficácia segundo as métricas definidas em cada trabalho juntamente às suas respectivas performances. Para os casos onde múltiplas métricas são providenciadas, priorizou-se a métrica de precisão média por classe (mean average precision - mAP) por ser a mais predominante.

Os métodos selecionados foram categorizados conforme sua finalidade na tarefa de classificação. Denotadas como classificação de “presença” estão as implementações responsáveis por detectar a presença das espécies-alvo nas imagens de entrada, enquanto as implementações pertinentes à classificação de “espécie” têm como tarefa adicional classificar o indivíduo detectado entre um conjunto de classes conhecidas.

Para análises voltadas às questões de pesquisa estabelecidas, as Figuras 2 e 3 visualizam contagens dos modelos selecionados dentre os dados extraídos. Note que como de cada trabalho foi extraído apenas o modelo de maior resultado, estudos que avaliaram múltiplos métodos não têm todos eles representados na contagem. Esses levantamentos oferecem uma visão panorâmica da distribuição dos métodos de classificação.

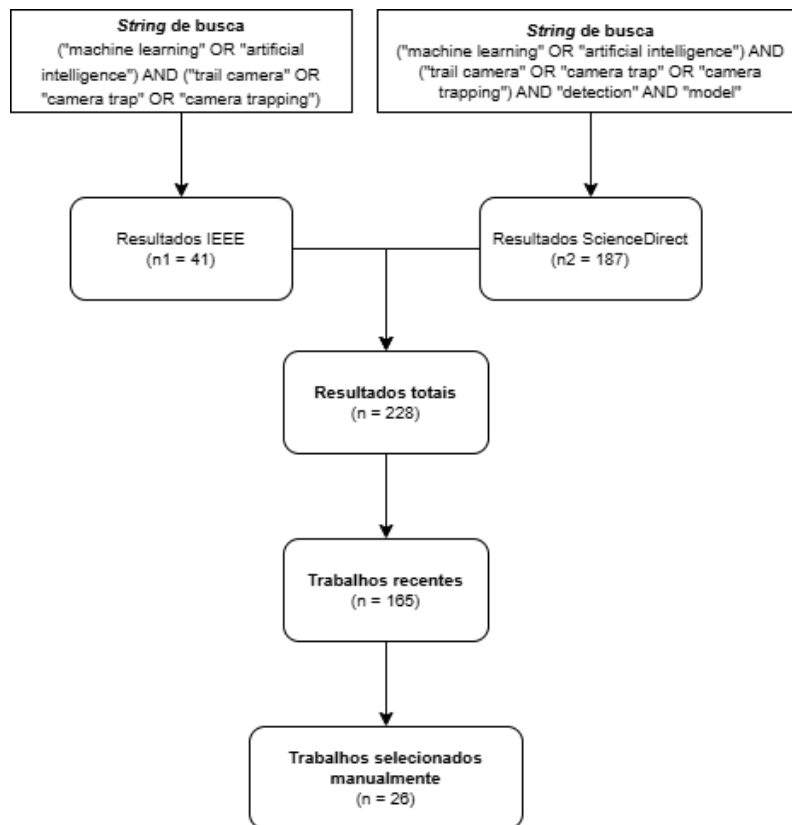


Figura 1. Processo de seleção de trabalhos

Redes neurais convolucionais (CNNs) são a base da maioria dos métodos implementados. Com exceção de um único trabalho [Kumar et al. 2022] que emprega regressão logística, todos os demais utilizam arquiteturas baseadas em CNNs, consolidando essa tecnologia como a abordagem padrão para objetivos como esse.

Entre essas abordagens, a série YOLO (You Only Look Once) predomina ambas as tarefas analisadas, com versões e modificações do modelo apresentando o melhor desempenho em 63% dos estudos de detecção e 53% de classificação, seguido pelo uso do MegaDetector para detecção e de modelos baseados na arquitetura ResNet para classificação de espécies.

Considerando os modelos mais comuns, YOLOv8 apresenta precisão média entre 92% [Liu et al. 2024] e 99% [Trevathan et al. 2025] para detecção, enquanto YOLOv5 alcança de 79% [Njathi et al. 2023] a 95% [Ansari et al. 2024; Balkana et al. 2024] para classificação. A diversidade de tarefas e métricas impossibilita cálculos ou comparativos abrangentes destas performances, mas os resultados explorados em trabalhos publicados estabelecem um referencial para futuras investigações na área.

Muitos dos trabalhos experimentam com diferentes algoritmos [Boukar et al. 2024, Roy et al. 2023, Bakana et al. 2024] ou versões [Trevathan et al. 2025, Huang et al. 2023] para avaliações de performance, mas nenhum explora o prospecto de abordagens baseadas em combinação de resultados (*ensemble approaches*) entre diferentes modelos, mesmo em contextos onde a velocidade de operação em tempo real não é um requisito.

Tabela 2. Trabalhos selecionados

Título	Espécies-alvo	Método	Performance	Classificação	Referência
Animal species detection and classification framework based on modified multi-scale attention mechanism and feature pyramid network	Rinocerontes, zebras, búfalos, elefantes	ResNet-50	0.87 mAP	Espécie	Ukwuoma et al. (2022)
Automated detection and classification of southern African Roman seabream using mask R-CNN	Peixes	Mask R-CNN	0.8 mAP	Espécie	Conrady et al. (2022)
Automatic detection for the world's rarest primates based on a tropical rainforest environment	Gibões	YOLOv3	0.8 mAP	Presença	Wang et al. (2022)
Human vs. machine: Detecting wildlife in camera trap images	Renas, aves, roedores, raposas	MegaDetector	0.97 F1	Presença	Leorna e Brinkman (2022)
Identification of Endangered Animal species of Pakistan using Classical and Ensemble Approach	Cabras, Ursos, Leopardos	Regressão Logística	0.92 F1	Espécie	Kumar et al. (2022)
Shark detection and classification with machine learning	Tubarões	VGG16	0.91 F1	Presença	Jenrette et al. (2022)
An Accurate and Fast Animal Species Detection System for Embedded Devices	Ursos, pumas, cervos, cabras	YOLOv4	0.89 mAP	Espécie	Ibraheam et al. (2023)
Automated wildlife image classification: An active learning tool for ecological applications	Lebres, raposas, cervos, etc.	MegaDetector	0.98 mAP	Presença	Bothmann et al. (2023)
Camera with Artificial Intelligence of Things (AIoT) Technology for Wildlife Camera Trap System	Cervos	YOLOv3	0.94 acurácia	Espécie	Huang, Lin e Liu (2023)
Comparison of two individual identification algorithms for snow leopards (Panthera uncia) after automated detection	Leopardos	YOLOv2	0.94 AP	Presença	Bohnett et al. (2023)
DeepWILD: Wildlife Identification, Localisation and estimation on camera trap videos using Deep learning	Cabras, cervos, raposas, etc.	MegaDetector	0.73 mAP	Espécie	Simões, Bouveyron e Precioso (2023)
Efficient Camera Trap Image Annotation Using YOLOv5	Ímpalas	YOLOv5	0.79 mAP	Espécie	Njathi et al. (2023)
Human Wildlife Conflict Mitigation Using YOLO Algorithm	Rinocerontes, Elefantes, Tigres, etc.	YOLOv4	0.94 mAP	Espécie	Leonid et al. (2023)
Wild Animal Detection using YOLOv8	Leões, tigres, leopardos, ursos	YOLOv8	0.94 mAP	Presença	Dave et al. (2023)
WillDect-YOLO: An efficient and robust computer vision-based accurate object localization model for automated endangered wildlife detection	Ursos, pinguins, elefantes, etc.	YOLOv4	0.96 mAP	Presença	Roy et al. (2023)
Wildlife Species Recognition Using Deep Learning	Gado, Pássaros, Cervos, etc.	ResNet-34	0.7 mAP	Espécie	Salomón et al. (2023)
Adaptive image processing embedding to make the ecological tasks of deep learning more robust on camera traps images	Ursos, pumas, javalis, etc.	YOLOv3	0.86 mAP	Espécie	Yang, Tian e Zhang (2024)
Animal Identity Recognition using Object Detection Techniques	Macacos	YOLOv5	0.95 mAP	Espécie	Ansari et al. (2024)
Benchmarking wild bird detection in complex forest scenes	Aves	Cascade R-CNN	0.70 mAP	Presença	Song et al. (2024)
Camouflage detection: Optimization-based computer vision for Alligator sinensis with low detectability in complex wild environments	Jacarés	YOLOv8	0.92 mAP	Presença	Liu et al. (2024)
Improving the Accuracy of Animal Species Classification in Camera Trap Images Using Transfer Learning	Búfalos, felinos, cervos, etc.	GoogLeNet	0.86 F1	Espécie	Boukar et al. (2024)
Object classification and visualization with edge artificial intelligence for a customized camera trap platform	Pássaros, lincos, gambás, etc.	MobileNetV2	0.91 F1	Espécie	Nazir e Kaleem (2024)
Vehicle-camel collisions in Saudi Arabia: Application of single and multi-stage deep learning object detectors	Camelos	YOLOv3	0.98 mAP	Presença	Alghamdi, Algethami e Tan (2024)
WildARE-YOLO: A lightweight and efficient wild animal recognition model	Rinocerontes, guepardos, leões, etc.	YOLOv5	0.95 mAP	Espécie	Bakana, Zhang e Trala (2024)
A computer vision enhanced IoT system for koala monitoring and recognition	Coalas	YOLOv8	0.99 mAP	Presença	Trevathan et al. (2025)
Stray Dog Detection System using YOLOv5	Cães	YOLOv5	0.93 mAP	Espécie	Bhosale et al. (2025)

Sendo os modelos mais empregados voltados somente à análise da imagem, os métodos implementados deixam de explorar os metadados das capturas capazes de providenciar contexto temporal e espacial, informações potencialmente valiosas para o monitoramento de espécies selvagens não nômades [Powell 2000].

4. Conclusão

Os procedimentos e resultados deste mapeamento sistemático da literatura visam identificar as soluções baseadas em aprendizado de máquina empregadas na classificação de imagens de armadilhas fotográficas e sua eficácia. Os resultados salientam o domínio de arquiteturas baseadas em CNNs, com destaque à série YOLO que apresenta maior desempenho na maioria dos estudos.

Ainda que a falta de padronização das métricas de avaliação tenha limitado comparações quantitativas diretas, os achados deste trabalho fornecem um panorama abrangente sobre as abordagens predominantes na área. Além disso, observou-se uma ausência de estratégias baseadas em combinação de resultados (*ensemble*) e um subaproveitamento de metadados das capturas capazes de contextualizar as entradas de forma valiosa para a detecção.

Este trabalho presta então referência para futuros estudos que apliquem os métodos identificados em novos conjuntos de dados, comparem resultados ou ampliem o escopo explorando abordagens híbridas e contextuais na automação da análise de armadilhas fotográficas.

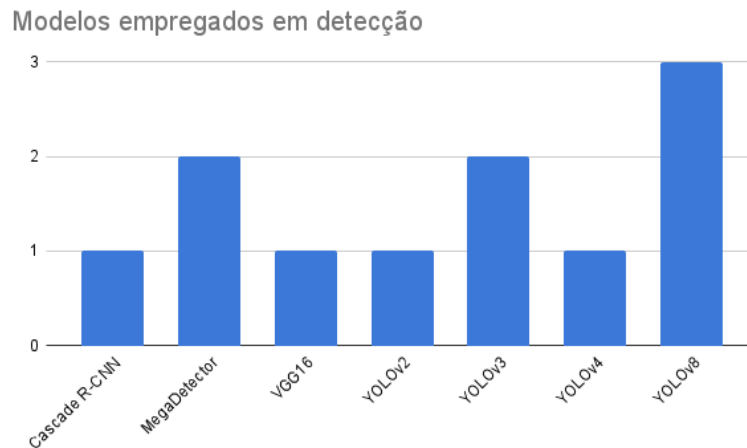


Figura 2. Gráfico de barras da contagem de modelos empregados na detecção



Figura 3. Gráfico de barras da contagem de modelos empregados na classificação

Referências

- Alghamdi, S., Algethami, A., and Tan, T. (2024). Vehicle-camel collisions in saudi arabia: application of single and multi-stage deep learning object detectors. *Ain Shams Engineering Journal*, 15(1):102328.
- Ansari, M. Z., Ahmad, F., Taheri, E. N., Reddy, R. G. J., and Mabood, F. (2024). Animal identity recognition using object detection techniques. *Procedia Computer Science*, 233:651–659.
- Bakana, S. R., Zhang, Y., and Twala, B. (2024). Wildare-yolo: A lightweight and efficient wild animal recognition model. *Ecological Informatics*, 80:102541.
- Bhosale, A., Shinde, P., Firke, Y., Patil, S., Mitake, P., and Shinde, S. (2025). Stray dog detection system using yolov5. *Procedia Computer Science*, 252:806–813.
- Bohnett, E., Holmberg, J., Faryabi, S. P., An, L., Ahmad, B., Rashid, W., and Ostrowski, S. (2023). Comparison of two individual identification algorithms for snow leopards

- (panthera uncia) after automated detection. *Ecological Informatics*, 77:102214.
- Bothmann, L., Wimmer, L., Charrakh, O., Weber, T., Edelhoff, H., Peters, W., and Menzel, A. (2023). Automated wildlife image classification: an active learning tool for ecological applications. *Ecological Informatics*, 77:102231.
- Boukar, M. M., Mahamat, A. A., Djibrine, O. H., and Abubakar, U. B. (2024). Improving the accuracy of animal species classification in camera trap images using transfer learning. In *2024 International Conference on Artificial Intelligence, Computer, Data Sciences and Applications (ACDSA)*, pages 1–6. IEEE.
- Buxton, R. T., Lendrum, P. E., Crooks, K. R., and Wittemyer, G. (2018). Pairing camera traps and acoustic recorders to monitor the ecological impact of human disturbance. *Global Ecology and Conservation*, 16:e00493.
- Chalmers, C., Fergus, P., Wich, S., Longmore, S. N., Walsh, N. D., Stephens, P. A., and Nuseibeh, A. (2023). Removing human bottlenecks in bird classification using camera trap images and deep learning. *Remote Sensing*, 15(10):2638.
- Conrady, C. R., Er, , Attwood, C. G., Roberson, L. A., and de Vos, L. (2022). Automated detection and classification of southern african roman seabream using mask r-cnn. *Ecological Informatics*, 69:101593.
- Dave, B., Mori, M., Bathani, A., and Goel, P. (2023). Wild animal detection using yolov8. *Procedia Computer Science*, 230:100–111.
- Glover-Kapfer, P., Soto-Navarro, C. A., and Wearn, O. R. (2019). Camera-trapping version 3.0: current constraints and future priorities for development. *Remote Sensing in Ecology and Conservation*, 5(3):209–223.
- Huang, H. C., Lin, C. H., and Liu, J. (2023). Camera with artificial intelligence of things (aiot) technology for wildlife camera trap system. In *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 0252–0258. IEEE.
- Ibraheam, M., Li, K. F., and Gebali, F. (2023). An accurate and fast animal species detection system for embedded devices. *IEEE Access*, 11:23462–23473.
- Jenrette, J., Liu, Z. C., Chimote, P., Hastie, T., Fox, E., and Ferretti, F. (2022). Shark detection and classification with machine learning. *Ecological Informatics*, 69:101673.
- Kitchenham, B. and Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. Technical Report 2.3, EBSE Technical Report.
- Kumar, R., Bhatti, S., Ruk, S. A., and Pathan, N. (2022). Identification of endangered animal species of pakistan using classical and ensemble approach. In *2022 2nd International Conference on Computing and Machine Intelligence (ICMI)*, pages 1–5. IEEE.
- Leonid, T. T., Kanna, H., VJ, C. C., AS, H., and Lokesh, C. (2023). Human wildlife conflict mitigation using yolo algorithm. In *2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*, pages 1–7. IEEE.
- Leorna, S. and Brinkman, T. (2022). Human vs. machine: Detecting wildlife in camera trap images. *Ecological Informatics*, 72:101876.

- Liu, Y., Che, S., Ai, L., Song, C., Zhang, Z., Zhou, Y., and Xian, C. (2024). Camouflage detection: Optimization-based computer vision for alligator sinensis with low detectability in complex wild environments. *Ecological Informatics*, 83:102802.
- Nazir, S. and Kaleem, M. (2024). Object classification and visualization with edge artificial intelligence for a customized camera trap platform. *Ecological Informatics*, 79:102453.
- Njathi, Y., Wanjiku, L., Mugambi, L., Kabi, J. N., Kiarie, G., and wa Maina, C. (2023). Efficient camera trap image annotation using yolov5. In *2023 IEEE AFRICON*, pages 1–6. IEEE.
- Oliver, R. Y., Iannarilli, F., Ahumada, J., Fegraus, E., Flores, N., Kays, R., and Jetz, W. (2023). Camera trapping expands the view into global biodiversity and its change. *Philosophical Transactions of the Royal Society B*, 378(1881):20220232.
- Powell, R. A. (2000). Animal home ranges and territories and home range estimators. *Research techniques in animal ecology: controversies and consequences*, 442:65–110.
- Roy, A. M., Bhaduri, J., Kumar, T., and Raj, K. (2023). Wildect-yolo: An efficient and robust computer vision-based accurate object localization model for automated endangered wildlife detection. *Ecological Informatics*, 75:101919.
- Salomón, S., Bringas, S., Duque, R., Montaña, J. L., and González, A. J. (2023). Wildlife species recognition using deep learning. In *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 221–227. IEEE.
- Simões, F., Bouveyron, C., and Precioso, F. (2023). Deepwild: Wildlife identification, localisation and estimation on camera trap videos using deep learning. *Ecological Informatics*, 75:102095.
- Song, Q., Guan, Y., Guo, X., Guo, X., Chen, Y., Wang, H., and Bao, L. (2024). Benchmarking wild bird detection in complex forest scenes. *Ecological Informatics*, 80:102466.
- Trevathan, J., Tan, W. L., Xing, W., Holzner, D., Kerlin, D., Zhou, J., and Castley, G. (2025). A computer vision enhanced iot system for koala monitoring and recognition. *Internet of Things*, 29:101474.
- Tuia, D., Kellenberger, B., Beery, S., Costelloe, B. R., Zuffi, S., Risse, B., and Berger-Wolf, T. (2022). Perspectives in machine learning for wildlife conservation. *Nature Communications*, 13(1):792.
- Ukwuoma, C. C., Qin, Z., Yussif, S. B., Happy, M. N., Nneji, G. U., Urama, G. C., and Agobah, H. (2022). Animal species detection and classification framework based on modified multi-scale attention mechanism and feature pyramid network. *Scientific African*, 16:e01151.
- Wang, X., Wen, S., Niu, N., Wang, G., Long, W., Zou, Y., and Huang, M. (2022). Automatic detection for the world’s rarest primates based on a tropical rainforest environment. *Global Ecology and Conservation*, 38:e02250.
- Yang, Z., Tian, Y., and Zhang, J. (2024). Adaptive image processing embedding to make the ecological tasks of deep learning more robust on camera traps images. *Ecological Informatics*, 82:102705.