Project - Literature review

Ting-Yu Hsu (24207795)
Yanjie Shao (24214602)

May 2025

1 Introduction

Wildlife image recognition plays a vital role in biodiversity monitoring and conservation research. In recent years, object detection models such as YOLO (You Only Look Once) [4] have achieved remarkable performance on well-lit, high-resolution datasets. However, these models often suffer from substantial performance degradation when the image quality is poor, such as when there are shadows or low lighting.

To overcome these limitations, several methods have been proposed. One notable direction is low-light image enhancement. For example, Zero-DCE (Zero-Reference Deep Curve Estimation) [2] is a deep curve estimation model that can adjust image brightness and contrast without relying on paired reference images, and has shown the potential to improve object visibility in dark environments. To better reflect real-world scenarios, recent datasets have also emerged to support research in this domain. The NTLNP dataset [1], introduced by [5], contains over 25,000 annotated nighttime infrared images of 17 wild animal species captured from camera traps in China's Northeast Tiger and Leopard National Park. Currently, several studies have benchmarked and improved animal detection models based on this dataset. For instance, Tan et al. [5] compared the performance of three mainstream object detection models, YOLOv5, Cascade R-CNN, and FCOS, on this dataset and found that the YOLOv5m model performed best in terms of mean Average Precision (mAP) and video classification accuracy. The YOLOv8-night model proposed by Wang et al. [7] improves the detection accuracy of animals in dark or shadowed environments by introducing a channel attention mechanism based on the YOLOv8 framework [6]. The model achieves an mAP of 0.854 on NTLNP nighttime images. Mulero-Pázmány et al. [3] further proposed a two-stage classification framework in which a global base model performs the initial classification and dedicated expert models refine the predictions within specific species clusters. This design effectively addresses issues such as class imbalance, background influence, and fine-grained species distinction, which are often encountered in wildlife camera trap images, and achieves an F1 score of 96.2% on over 1.3 million wildlife images.

This project aims to build on the above progress and develop a machine learning model that can classify animal species in night-vision images. The goal is to improve species recognition accuracy and contribute to the automation of ecological monitoring systems in low-light environments.

2 Benefits

1. **Enhanced image recognition accuracy in low-light environments.** – Improve the detection accuracy of wildlife species in night-vision images by combining low-light image enhancement techniques such as Zero-DCE with advanced object detection models such as YOLOv8-night.
2. **Improved efficiency in ecological monitoring.** – Automatically identify species from night-vision camera trap images, reducing the need for manual review and accelerating data analysis for ecological surveys.
3. **Support biodiversity research.** – Accurate and automated animal species identification enables long-term biodiversity tracking and facilitates population modeling.
4. **Broader applications of AI in low-light environments.** – It has reference value for other night imaging tasks (such as traffic monitoring and night face recognition).

3 Technologies and tools

- **NTLNP Dataset** – A public dataset of night-time wildlife images captured by infrared cameras, containing 17 major species.
- **YOLOv5 and YOLOv8** – YOLO is a real-time object detection model known for its high speed and high accuracy. YOLOv5 performs well on the NTLNP dataset, while the enhanced YOLOv8-night architecture incorporates an attention mechanism to improve its performance in low-light environments.
- **Zero-DCE** – A zero-reference low-light image enhancement algorithm based on depth curve estimation, which can adjust contrast and illumination without referring to ground truth, and is an effective preprocessing step for improving the visibility in infrared wildlife images.
- **OpenCV** – A powerful image library for image preprocessing, including resizing, denoising, and brightness enhancement.
- **Python** – The primary programming language used for data processing, model training, and evaluation.
- **PyTorch** – A widely used deep learning framework that supports flexible model development and efficient training for tasks such as object detection and image enhancement.

4 Expected steps

1. **Data collection** – Download the NTLNP dataset, examine its structure, and analyze the distribution of species and image conditions (lighting, resolution, etc.).
2. **Image preprocessing** – Apply image enhancement techniques (e.g. Zero-DCE) to improve visual quality under low-light conditions. Other preprocessing tech-

niques including denoising, normalization, and resizing using OpenCV can also be used.

3. **Model setup** – Design a new network architecture based on existing models (YOLOv5, YOLOv8-night, etc.), and adapt the network settings and training pipeline to suit the data characteristics.

4. **Model training and validation** – Train the model using PyTorch on the preprocessed dataset and validate its performance. Various methods (e.g. cross-validation, hold-out validation) can be employed to ensure generalization and avoid overfitting.

5. **Model evaluation** – Evaluate model performance using standard metrics such as mAP, precision, recall, and F1 score, and compare its results with other baseline methods to assess its effectiveness under low-light conditions.

References

[1] NTLNP dataset. Github repository, https://github.com/myyyyw/NTLNP.

[2] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1780–1789, 2020.

[3] Margarita Mulero-Pázmány, Sandro Hurtado, Cristóbal Barba-González, María Luisa Antequera-Gómez, Francisco Díaz-Ruiz, Raimundo Real, Ismael Navas-Delgado, and José F Aldana-Montes. Addressing significant challenges for animal detection in camera trap images: a novel deep learning-based approach. *Scientific Reports*, 15(1):1–18, 2025.

[4] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.

[5] Mengyu Tan, Wentao Chao, Jo-Ku Cheng, Mo Zhou, Yiwen Ma, Xinyi Jiang, Jianping Ge, Lian Yu, and Limin Feng. Animal detection and classification from camera trap images using different mainstream object detection architectures. *Animals*, 12(15):1976, 2022.

[6] Rejin Varghese and M Sambath. Yolov8: A novel object detection algorithm with enhanced performance and robustness. In *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, pages 1–6. IEEE, 2024.

[7] Tianyu Wang, Siyu Ren, and Haiyan Zhang. Nighttime wildlife object detection based on yolov8-night. *Electronics Letters*, 60(15):e13305, 2024.