

DETECTION AND RECOGNITION OF INSECTS BY TAXONOMIC ORDERS USING DEEP LEARNING

Huỳnh Anh Kiệt - 240101014

Tóm tắt

- Lớp: CS2205.CH183
- Link Github của nhóm: <https://github.com/favkiet/huynhanhkiem-CS2205.CH183.git>
- Link YouTube video: <https://youtu.be/crim5RYx0yU>
- Ảnh + Họ và Tên của các thành viên:

Huỳnh Anh Kiệt - 240101014



Giới thiệu

- Tầm quan trọng: Côn trùng có vai trò sinh thái quan trọng, nhưng cũng gây hại cho nông nghiệp. Nhận dạng chính xác giúp nghiên cứu và kiểm soát dịch hại hiệu quả.
- Thách thức: Hơn 1,3 triệu loài côn trùng, bài toán nhận dạng phức tạp.
- Công nghệ:
 - CNN: Hiệu quả trong trích xuất đặc trưng ảnh.
 - Transformer & Attention: Tiềm năng trong xử lý ảnh đa phương thức.
- Phương pháp tiếp cận:
 - Xác định họ trước, sau đó nhận dạng loài.
 - Giảm độ phức tạp, tận dụng hệ thống phân loại sinh học.
- Thực nghiệm:
 - So sánh Faster R-CNN ResNet50, YOLOv10, Florence-2.
 - Bộ dữ liệu ArTaxOr với nhãn hộp giới hạn.

Mục tiêu

Mục tiêu:

- Đánh giá **Faster R-CNN ResNet50, YOLOv10, Florence-2** trên bộ dữ liệu **ArTaxOr**.
- Phân tích kết quả để đề xuất khung phương pháp nhận dạng hiệu quả.

Phân tích kết quả:

- So sánh độ chính xác, tốc độ, khả năng tổng quát hóa.
- Xác định mô hình tối ưu cho bài toán thực tế.

Ứng dụng & Đóng góp:

- Hỗ trợ **quản lý dịch hại** trong nông nghiệp.
- Nâng cao **hiệu quả sản xuất & bảo vệ cây trồng**.

Nội dung và Phương pháp

Nội dung nghiên cứu

- **Mục tiêu:** Nhận dạng côn trùng theo bậc phân loại bằng mô hình học sâu.
- **Các bước chính:**
 - **Tổng quan bài toán:** Ứng dụng trong nông nghiệp.
 - **Xây dựng & xử lý dữ liệu:** Bộ dữ liệu **ArTaxOr** (JPEG, JSON).
 - **Triển khai mô hình:**
 - **Faster R-CNN ResNet50:** Mạng CNN cho phát hiện đối tượng.
 - **YOLOv10:** Cân bằng tốc độ & độ chính xác.
 - **Florence-2:** Transformer, tiếp cận đa phương thức.
 - **Thực nghiệm & đánh giá:** So sánh hiệu suất theo mAP.
 - **Đề xuất ứng dụng:** Hỗ trợ kiểm soát dịch hại.

Nội dung và Phương pháp

Phương pháp nghiên cứu

- **Xử lý dữ liệu:**
 - Chuyển đổi sang định dạng **COCO, YOLO, Florence-2**.
 - Chia tập **Train (80%) - Test (20%)**.
- **Huấn luyện mô hình:**
 - **Image Augmentation:** Scaling, Rotation, Flip, Mosaic, ...
 - Dùng **PyTorch** trên **2 GPU Kaggle**.
- **Đánh giá mô hình:**
 - **mAP:** Độ chính xác nhận dạng.
- **Thử nghiệm ứng dụng:**
 - Khả năng triển khai vào hệ thống thực tế.
 - Đề xuất mô hình tối ưu cho giám sát côn trùng.

Kết quả dự kiến

- Kết quả thực nghiệm, độ chính xác (mAP) của các mô hình.
- Kết quả đánh giá và so sánh điểm nổi bật và độ chính xác của các mô hình đề xuất trên cùng một tập dữ liệu.
- Lựa chọn mô hình có độ chính xác cao nhất để có thể đưa vào ứng dụng thực tế.

Tài liệu tham khảo

- [1] Ao Wang, Hui Chen, Lihao Liu, Kai Chen, Zijia Lin, Jungong Han, Guiguang Ding: YOLOv10: Real-Time End-to-End Object Detection: arXiv 2024:2405.14458
- [2] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks: IEEE 2016: 10.1109/TPAMI.2016.2577031
- [3] Bin Xiao, Haiping Wu, Weijian Xu, Xiyang Dai, Houdong Hu, Yumao Lu, Michael Zeng, Ce Liu, Lu Yuan: Florence-2: Advancing a Unified Representation for a Variety of Vision Tasks: CVPR 2024
- [4] Jun Li, Zhiqiang Shen, Jian Zhang: Large-image Object Detection for Fine-grained Recognition: arXiv 2020:2501.12489
- [5] Z. Liu, M. Wu, L. Wang: YOLOv5: A New Approach to Real-Time Object Detection with Improved Performance and Efficiency: arXiv 2020:2104.10934
- [6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun: Deep Residual Learning for Image Recognition: CVPR 2016.
- [7] Spiking-YOLO: Spiking Neural Network for Real-time Object Detection: Seijoon Kim, Seongsik Park, Byunggook Na, Sungroh Yoon: arXiv:1903.06530v2
- [8] Mingxing Tan Ruoming Pang Quoc V. Le: EfficientDet: Scalable and Efficient Object Detection: CVPR 2020