

Hệ thống nhận diện động vật sử dụng YOLO và ứng dụng web

Trường Đại học Tân Tạo
Khoa Công nghệ Thông tin

GVHD: Tiến sĩ Trần Ngọc Anh

SV: Phan Văn Tài - MSSV: 2202081

Đồ án
môn học

Nội dung trình bày

- | | | | |
|---|------------------------|---|--------------------|
| 1 | Giới thiệu đề tài | 5 | Huấn luyện mô hình |
| 2 | Vấn đề và mục tiêu | 6 | Kết quả |
| 3 | Cơ sở lý thuyết (YOLO) | 7 | Demo ứng dụng web |
| 4 | Xử lý dữ liệu | 8 | Kết luận |

Giới thiệu đề tài

Bối cảnh

- Phát hiện đối tượng là bài toán quan trọng trong Computer Vision
- Ứng dụng: Bảo tồn, nghiên cứu, quản lý động vật

Đề tài

- Xây dựng hệ thống nhận diện 80 lớp động vật
- Sử dụng YOLOv8 – mô hình one-stage detection hiện đại
Triển khai ứng dụng web với React + FastAPI

Vấn đề và mục tiêu

Vấn đề chính

Mục tiêu

- ① Dataset mất cân bằng nghiêm trọng
 - Imbalance ratio: 321:1
 - Butterfly: 2.045 ảnh vs Squid: 28 ảnh
- ② Đa lớp phức tạp: 80 classes động vật
- ③ Yêu cầu độ chính xác cao: mAP50 ≥ 0.75

- Cân bằng dataset và cải thiện chất lượng Đạt mAP50 ≥ 0.75
- Xây dựng ứng dụng web dễ sử dụng

Cơ sở lý thuyết – YOLO

YOLO (You Only Look Once)

- One-stage detector: Nhanh, phù hợp real-time
- YOLOv8 (2023): Phiên bản mới nhất
 - Anchor-free architecture
 - C2f module (gradient flow tốt hơn)
 - Decoupled head

Kiến trúc YOLOv8

Backbone (CSPDarknet) → Neck (PANet) → Head (Decoupled)

Ưu điểm

- Tốc độ nhanh (~3.7ms/ảnh)
- Độ chính xác cao
- Dễ triển khai

Xử lý dữ liệu – Tổng quan

Dataset

- Nguồn: Kaggle – Animals Detection Images Dataset
- Tổng số ảnh: 29.071
- Số classes: 80
- Vấn đề: Mất cân bằng nghiêm trọng

Pipeline xử lý

Raw Dataset → Analysis → Validation → Balancing → Split → YOLO Format

Xử lý dữ liệu – Phân tích

Thống kê ban đầu

Metric	Giá trị
Tổng ảnh	29.071
Số classes	80
Max/class	2.045 (Butterfly)
Min/class	28 (Squid)
Trung bình	363,4 ảnh/class
Imbalance ratio	73:1

Phân loại classes

- Rất ít (<15): 0 classes
- Ít (15–30): 2 classes (Squid, Turtle)
- Tốt (≥ 30): 78 classes

Xử lý dữ liệu – Validation

Image Validation

- Format: .jpg, .jpeg, .png, .bmp
- Kích thước: $32 \times 32 \leq \text{size} \leq 10000 \times 10000$ Kiểm tra
- corrupt files

Bounding Box Validation

- Swap nếu $xmin > xmax$
- Clamp về $[0, \text{img_width}/\text{height}]$
- Loại bỏ nếu:
 - $\text{width}/\text{height} < 5 \text{ pixels}$
 - Area $< 0.05\%$ hoặc $> 98\%$ ảnh

Kết quả: Loại bỏ các samples không hợp lệ

Xử lý dữ liệu – Cân bằng

Chiến lược

- ① Loại bỏ: Classes < 15 ảnh (0 classes)
- ② **Oversampling:** Classes 15–30 → 30 ảnh
 - Squid: 22 → 30 (+8)
 - Turtle: 27 → 30 (+3)
- ③ Giữ nguyên: Classes ≥ 30 ảnh

Kết quả

	Trước	Sau
Tổng samples	29.071	28.184
Imbalance ratio	73:1	73:1
Oversampled	—	11 samples

Cải thiện: Tất cả classes đều có ≥ 30 samples

Xử lý dữ liệu – Chia dữ liệu

Stratified Split (80/20)

- Train: 22.518 samples (80%)
- Validation: 5.666 samples (20%)

YOLO Format Conversion

$$x_{center} = \frac{(x_{min} + x_{max})/2}{img_width}$$

$$y_{center} = \frac{(y_{min} + y_{max})/2}{img_height}$$

$$width = \frac{x_{max} - x_{min}}{img_width}$$

$$height = \frac{y_{max} - y_{min}}{img_height}$$

Cấu trúc cuối cùng: yolo_dataset_pro/ → images/train, val + labels/train, val + data.yaml

Huấn luyện mô hình

Tham số	Giá trị
Model	YOLOv8n (nano)
Epochs	100
Batch size	32
Image size	640×640
Optimizer	SGD
Learning rate	0.002 → 0.0001
Hardware	Tesla P100 (16GB)

Thời gian

- Training time: 8 giờ 21 phút
- Tốc độ: ~2.6–2.9 it/s

Kết quả – Metrics

Metrics tổng hợp

Metric	Giá trị
mAP50	0.7565 (75.65%)
mAP50-95	0.6322 (63.22%)
Precision	0.7140
Recall	0.7469
F1-Score	0.7301

So sánh với baseline

Dataset	mAP50	Improvement
Imbalanced	0.6925	–
Balanced	0.7565	+9.2%

Kết luận: Cân bằng dữ liệu cải thiện đáng kể!

Kết quả – Top classes

Top 5 classes tốt nhất

Class	mAP50	Samples
Woodpecker	0.991	41
Ladybug	0.975	86
Eagle	0.963	179
Zebra	0.965	39
Polar bear	0.951	56

Classes cần cải thiện

Class	mAP50	Samples
Turtle	0.076	6
Squid	0.172	6
Goose	0.381	65

Nhận xét: Classes có ít samples có hiệu năng thấp

Tech Stack: React 18.2.0 – FastAPI 0.104.1 – Model: YOLOv8n

- /api/detect
- /api/batch; [draw,below of=be,fill=orange!10] (model) YOLOv8 Model
(best.pt); [->,thick] (fe) – (be); [->,thick] (be) – (model);

Demo ứng dụng web

Tính năng chính

1 Upload ảnh

- Single image (drag & drop)
- Batch processing (tối đa 20 ảnh)
- File validation

2 Nhận diện

- Single detection
- Batch processing
- Tùy chỉnh thresholds (Confidence 0.0–1.0, IoU 0.0–1.0)

3 Hiển thị kết quả

- Ảnh với bounding boxes
- Bảng detections (sortable)
- Thống kê chi tiết

Tốc độ API: Single ~100–200ms | Batch(20) ~2–4 giây

Kết luận

Kết quả đạt được

- ① Pipeline xử lý dữ liệu chuyên nghiệp
 - Cân bằng dataset: **321:1 → 73:1**
 - Validation và cleaning
- ② Model hiệu quả
 - mAP50 = 0.7565 (gần mục tiêu 0.78–0.82)
 - Cải thiện **+9.2%** so với baseline
 - Tốc độ: ~3.7ms/ảnh
- ③ Ứng dụng web hoàn chỉnh
 - React + FastAPI
 - Hỗ trợ single và batch
 - Giao diện thân thiện

Hướng phát triển

Cải thiện dữ liệu

- Thu thập thêm dữ liệu cho classes yếu
- Tăng số lượng samples

Cải thiện model

- Thử YOLOv8s/m
- Train thêm epochs (120–150)
- Fine-tuning

Tính năng mới

- Video detection (real-time)
- Object tracking
- Export kết quả (JSON, CSV)
- Mobile app

THANKS