

CẢI THIỆN ĐỘ CHÍNH XÁC VÀ TỐC ĐỘ CỦA MÔ HÌNH

EFFICIENTDET

Nhóm HNH

Bùi Đào Gia Huy	18520818
Nguyễn Trần Trung Hiếu	16520405
Nguyễn Thanh Hoàng	16520442

Tóm tắt

- Lớp: CS519.M11
- Link Github của nhóm: <https://github.com/BuiDaoGiaHuy/CS519.M11>
- Link YouTube video:
- Ảnh + Họ và Tên của các thành viên



Bùi Đào Gia Huy
18520818

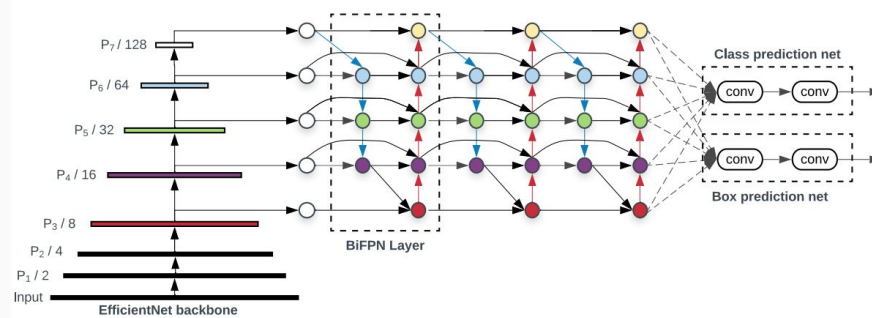


Nguyễn Trần Trung Hiếu
16520405



Nguyễn Thanh Hoàng
16520442

Giới thiệu



- EfficientDet[3] là một họ các mô hình thuộc loại one-stage object detector.
- Mô hình được thiết kế để hoạt động hiệu quả trên phần cứng ràng buộc về tài nguyên tính toán(FLOPs). Tuy nhiên điều này là chưa đủ, đặc biệt đối với các ứng dụng đòi hỏi tính toán trong thời gian thực, tốc độ đóng một vai trò quan trọng.
- Do đó, cải thiện thời gian xử lý của mô hình nhưng vẫn đảm bảo tính hiệu quả ban đầu là điều cần thiết.

Mục tiêu

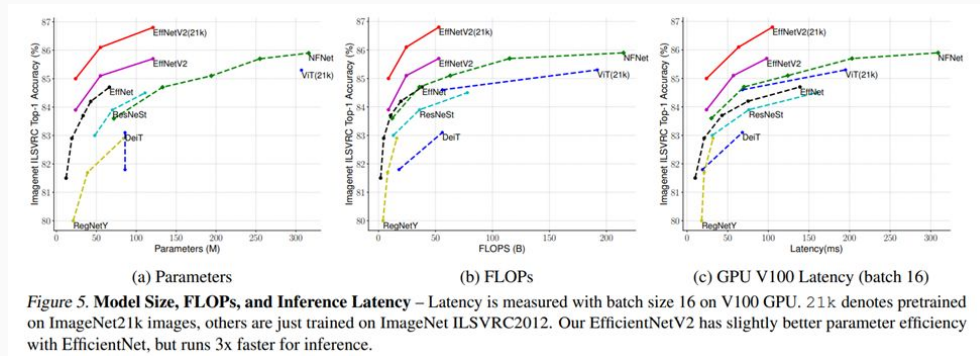
- **Mục tiêu:** Cải thiện độ chính xác và tốc độ của mô hình EfficientDet.
- **Giả thuyết:** Để cải thiện hiệu suất tổng thể mô hình, các đặc trưng của mô hình cần được lựa chọn và tối ưu hợp lý.

Để đạt được mục tiêu, nhóm đề xuất một vài thay đổi trong mô hình:

1. Backbone. (ảnh hưởng phần lớn đến tốc độ của mô hình)
2. Hàm loss cho dự đoán bounding box. (tăng độ chính xác)
3. Sử dụng các thuật toán NMS với cơ chế tính toán song song. (tăng tốc độ mô hình)

Nội dung và Phương pháp

1. Backbone: EfficientNetV2[4]



Sử dụng EffNetV2-(S/M) với bộ tham số **pretrained ImageNet1K** và **pretrained ImageNet21K finetuned trên ImageNet1K** cho backbone.

Dựa trên quan sát về thiết kế các mô hình object detection trước đây: Các pretrained model đạt độ chính xác cao ở bài toán classification thường mang lại hiệu suất tốt ở bài toán object detection.[2,3]

Nội dung và Phương pháp

2. Optimizing bounding box regression: CIOU loss[6]

Hàm Smooth-L1 Loss được sử dụng trong mô hình có một vài hạn chế:

- Bỏ qua sự tương quan giữa các biến (x, y, w, h)
- Ảnh hưởng bởi các bounding box lớn (do dạng chưa chuẩn hóa)

CIOU Loss[6], biến thể của IOU Loss, có thể khắc phục được nhược điểm bên trên do tính bất biến theo tỉ lệ.

CIOU Loss tối ưu bounding box theo 3 yếu tố hình học quan trọng: overlap area, khoảng cách chuẩn hóa của 2 tâm và aspect ratio.

Nội dung và Phương pháp

3. Non maximum Suppression: Matrix NMS[5] và Fast-DIOU-NMS

Thử nghiệm 2 thuật toán sau:

- **Matrix NMS[5]**: giảm score của những bounding box chồng lẫn nhau tương tự Soft-NMS(thuật toán đang được sử dụng trong mô hình ban đầu) nhưng khác với các phương pháp NMS trước đây thực hiện loại bỏ các bounding box dư thừa một cách tuần tự, Matrix NMS thực hiện điều này một cách song song khiến cho tốc độ được cải thiện đáng kể.
- Dựa trên ý tưởng Fast-NMS[1], chúng tôi đề xuất kết thay thế IOU bằng DIOU[6] vào ma trận tính toán theo từng cặp candidates, gọi là **Fast-DIOU-NMS**. Chúng tôi mong muốn tạo ra thuật toán mới vừa có độ chính xác xấp xỉ DIOU-NMS vừa có tốc độ thực thi nhanh như Fast-NMS.

Kết quả dự kiến

Đạt được mục tiêu đề ra là cải thiện độ chính xác(trên bộ dữ liệu thử nghiệm MS COCO) và thời gian thực thi của mô hình EfficientDet dựa trên sự kết hợp từ các đề xuất trên.

Tài liệu tham khảo

- [1] Daniel Bolya, Chong Zhou, Fanyi Xiao, Yong Jae Lee: *YOLACT: Real-Time Instance Segmentation*. ICCV 2019: 9156-9165
- [2] Golnaz Ghiasi, Tsung-Yi Lin, Quoc V. Le: NAS-FPN: Learning Scalable Feature Pyramid Architecture for Object Detection. CVPR 2019: 7036-7045
- [3] Mingxing Tan, Ruoming Pang, Quoc V. Le: EfficientDet: Scalable and Efficient Object Detection. CVPR 2020: 10778-10787
- [4] Mingxing Tan, Quoc V. Le: *EfficientNetV2: Smaller Models and Faster Training*. ICML 2021: 10096-10106
- [5] Xinlong Wang, Rufeng Zhang, Tao Kong, Lei Li, Chunhua Shen: *SOLOv2: Dynamic and Fast Instance Segmentation*. NeurIPS 2020
- [6] Zhaohui Zheng, Ping Wang, Wei Liu, Jinze Li, Rongguang Ye, Dongwei Ren: Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. AAAI 2020: 12993-13000