

CẢI THIỆN CHUYÊN SÂU HỌ R-CNN TRONG BÀI TOÁN PHÁT HIỆN ĐỐI TƯỢNG DỰA TRÊN TÀI LIỆU TIẾNG VIỆT DẠNG ẢNH

Lê Thị Phương Vy - 20520355

Lê Trương Ngọc Hải - 20520481

Nguyễn Nhật Trường - 20522087

Tóm tắt

- Lớp: CS519.N11
- Link Github của nhóm:
<https://github.com/truong11062002/CS519.N11--PhuongPhapLuanNCKH>
- Link YouTube video: <https://www.youtube.com/watch?v=Jq76H0KSZtM>



Nguyễn Nhật Trường



Lê Trương Ngọc Hải



Lê Thị Phương Vy

Giới Thiệu

- Động lực:
 - Phát hiện đối tượng trên trang văn bản giúp việc số hóa tài liệu dễ dàng hơn.
 - Các ngôn ngữ ít phổ biến như Tiếng Việt chưa có một mô hình đủ tốt.
- Giới thiệu bài toán:
 - Input: ảnh dạng tài liệu tiếng Việt.
 - Output: vị trí của các đối tượng trong ảnh được thể hiện bằng các hộp giới hạn tối thiểu với các màu sắc khác nhau cho từng lớp tương ứng.

Mục tiêu

- Khảo sát và phân tích bộ dữ liệu tài liệu dạng ảnh tiếng Việt UIT-DODV.
- Tìm hiểu và thực nghiệm đánh giá các phương pháp cải thiện hiệu suất của R-CNN: Double-Head R-CNN, Libra R-CNN và Guided Anchoring.
- Đề xuất phương pháp cải thiện họ RCNN: Guided Anchoring Cascade R-CNN.

Nội dung và Phương pháp

A. NỘI DUNG

- Khảo sát và phân tích những đặc trưng của bộ dữ liệu dạng ảnh tiếng Việt UIT-DODV để tiếp cận gần hơn với bài toán POD (Page Object Detection).
- Nghiên cứu về cơ sở lý thuyết của R-CNN để nắm được những đặc tính của phương pháp.
- Tìm hiểu về các phương pháp gia tăng hiệu suất của họ R-CNN.
- Giả thiết: **“Liệu có cách cải tiến hiệu suất phát hiện đối tượng trên trang tài liệu dạng ảnh của họ R-CNN?”**
- Huấn luyện và đánh giá.

Nội dung và Phương pháp

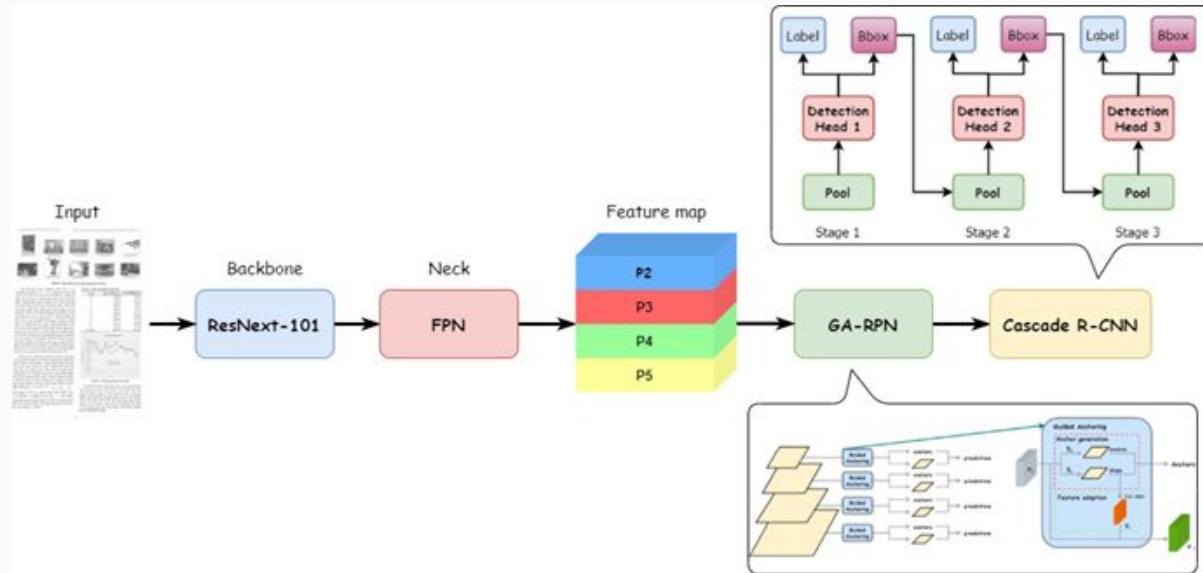
B. PHƯƠNG PHÁP

- Tìm hiểu và phân tích tổ chức của bộ dữ liệu UIT-DODV [1].
- Tìm hiểu và đánh giá các phương pháp cải thiện hiệu suất của R-CNN gồm: Double-Head R-CNN [2], Libra R-CNN [3] và Guided Anchoring [4].
- So sánh hiệu suất giữa Faster R-CNN và Cascade R-CNN [5] và quyết định chọn phương pháp nào để làm detector chính cho phương pháp đề xuất.

Nội dung và Phương pháp

B. PHƯƠNG PHÁP

- Chúng tôi tiến hành thử nghiệm kết hợp mô hình dựa trên kiến trúc Cascade R-CNN kết hợp với Guided Anchoring.



Kết quả dự kiến

- Nếu kết quả thực nghiệm đạt kết quả như mong đợi tiến hành đề xuất một phương pháp mới cho bài toán phát hiện đối tượng trang trong một bài báo SCOPUS/tạp chí Q3.
- Xây dựng bảng đánh giá và phân tích kết quả thực nghiệm trên bộ dữ liệu UIT-DODV.

Tài liệu tham khảo

- [1]. Linh Truong Dieu, Thuan Trong Nguyen, Nguyen D. Vo, Tam V. Nguyen, Khang Nguyen: Parsing Digitized Vietnamese Paper Documents. CAIP (1) 2021: 382-392
- [2]. Yue Wu, Yinpeng Chen, Lu Yuan, Zicheng Liu, Lijuan Wang, Hongzhi Li, Yun Fu: Rethinking Classification and Localization for Object Detection. CVPR 2020: 10183-10192
- [3]. Jiangmiao Pang, Kai Chen, Jianping Shi, Huajun Feng, Wanli Ouyang, Dahua Lin: Libra R-CNN: Towards Balanced Learning for Object Detection. CVPR 2019: 821-830
- [4]. Jiaqi Wang, Kai Chen, Shuo Yang, Chen Change Loy, Dahua Lin: Region Proposal by Guided Anchoring. CVPR 2019: 2965-2974
- [5]. Zhaowei Cai, Nuno Vasconcelos: Cascade R-CNN: Delving Into High Quality Object Detection. CVPR 2018: 6154-6162