

BÁO CÁO ĐỒ ÁN CUỐI KỲ

Môn học: CS519 - PHƯƠNG PHÁP LUẬN NCKH

Lớp: CS519.N11

GV: PGS.TS. Lê Đình Duy

Trường ĐH Công Nghệ Thông Tin, ĐHQG-HCM



TĂNG CƯỜNG ĐỘ CHÍNH XÁC CỦA NHẬN DẠNG VĂN BẢN CẢNH BẰNG PHƯƠNG PHÁP HƯỚNG DẪN TỪ ĐIỂN

**Ngô Thành Phát -
19521994**

Tóm tắt

- Link Github của nhóm: <https://github.com/Pakm19>
- Link YouTube video: <https://www.youtube.com/watch?v=hSgUlujSHYE>
- Ảnh + Họ và Tên



Ngô Thành Phát

Giới thiệu

- Bài toán nhận dạng văn bản cảnh (scene text recognition) là một bài toán trong lĩnh vực thị giác máy tính (computer vision) nhằm mục đích nhận dạng và giải mã các văn bản xuất hiện trong các cảnh hình ảnh
- Mục tiêu của chúng ta là phát hiện và trích xuất các ký tự từ các hình ảnh có chứa văn bản

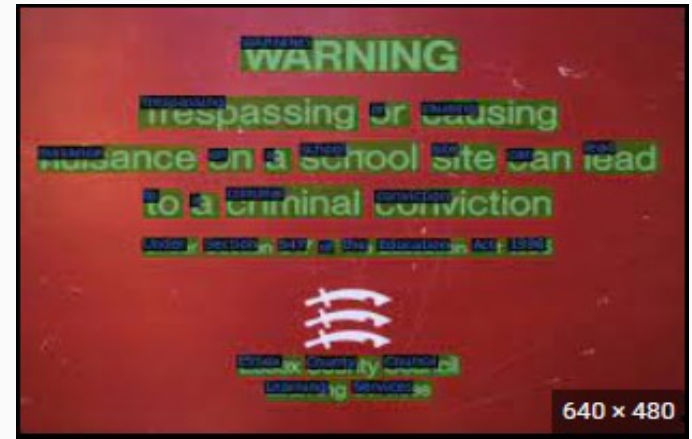


Giới thiệu

- Mô tả input và output



Input



Output

Mục tiêu

- Nghiên cứu và đánh giá các hướng tiếp cận truyền thống của bài toán Scene Text Recognition
- Đề xuất bộ dữ liệu tiếng Việt mới cho bài toán Scene Text Recognition
- Đề xuất phương pháp tăng cường độ chính xác cho bài toán Scene Text Recognition

Nội dung và Phương pháp

Nghiên cứu và đánh giá các hướng tiếp cận truyền thống của bài toán Scene Text Recognition

Recognition

- Chúng tôi tìm kiếm các công trình nghiên cứu gần đây nhất đã được công bố trên các diễn đàn về Thị giác máy tính như CVPR, ICCV, ECCV,... để tìm kiếm các bài báo, luận văn, và tài liệu khác liên quan đến bài toán này.
- Chúng tôi xác định các phương pháp tiếp cận tiên tiến nhất và có kết quả tốt nhất trong bài toán này.
- Chúng tôi phân tích và đánh giá ưu nhược điểm của từng phương pháp, bao gồm hiệu suất, độ phức tạp tính toán, thời gian huấn luyện và suy luận.



Nội dung và Phương pháp

Đề xuất bộ dữ liệu tiếng Việt mới cho bài toán Scene Text Recognition

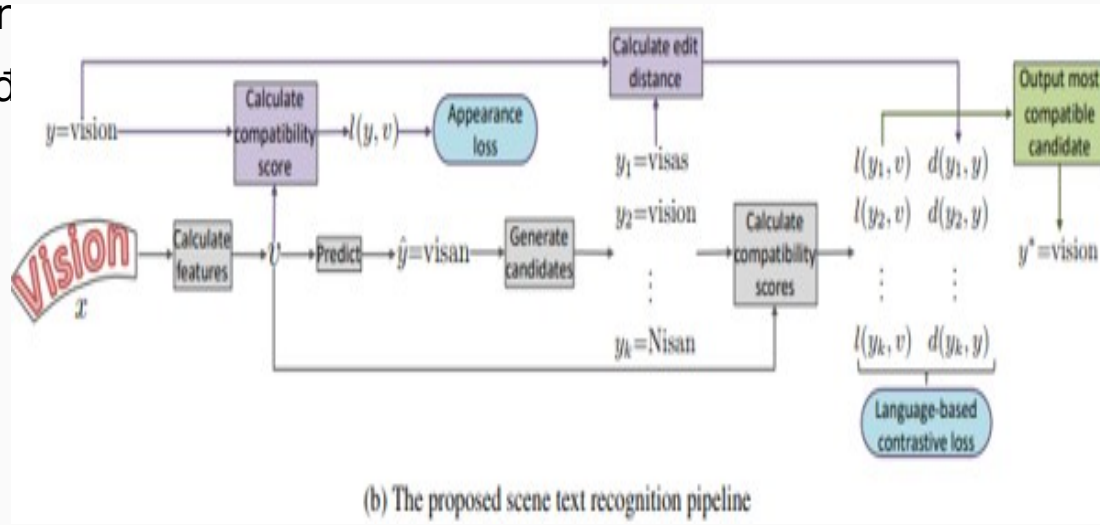
- Chúng tôi xây dựng một đội ngũ riêng cho việc sưu tập dữ liệu, tập hợp đa dạng của các văn bản cảnh được bắt gặp trong cuộc sống hàng ngày ở Việt Nam
- Chúng tôi xây dựng một đội ngũ cho việc xử lý dữ liệu thô đầu vào trở thành dữ liệu phục vụ cho bài toán Scene Text Recognition



Nội dung và Phương pháp

Đề xuất phương pháp tăng cường độ chính xác cho bài toán Scene Text Recognition

- Trọng tâm kỹ thuật chính trong công việc của chúng tôi là ở giai đoạn nhận dạng
- Chúng tôi cũng sử dụng từ điển trong giai đoạn đào tạo
- Chúng tôi sử dụng kết hợp phương pháp của chúng tôi với các mô-đun tiên tiến nhất hiện nay
- Quy trình đào tạo



Kết quả dự kiến

- Khi kết quả thành công (kết quả thực nghiệm cao hơn các mô-đun tiên tiến nhất hiện tại, cụ thể là H-mean score) sẽ mở ra một hướng tiếp cận mới cho toán Scene Text Recognition
- Bộ dữ liệu VinText sẽ là bộ dữ liệu tiêu chuẩn cho bài toán Scene Text Recognition ở Việt Nam
- Triển khai vào ứng dụng thực tế để cải thiện hiệu suất, chất lượng tự động ghi lại thông tin chi tiết của container xuất nhập khẩu (mã, loại, kích cỡ,...) từ tàu vào bờ và từ bờ đến.

Tài liệu tham khảo

- [1]. V. I. Levenshtein. Binary codes capable of correcting insertions and reversals. 1966
- [2]Y. Liu, Hao Chen, Chunhua Shen, Tong He, Lian-Wen Jin, and L. Wang. ABCNet: Real-time scene text spotting with adaptive bezier-curve network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020.
- [3] Minghui Liao, Guan Pang, J. Huang, Tal Hassner, and X. Bai. Mask textspotter v3: Segmentation proposal network for robust scene text spotting. ArXiv, 2020
- [4] Y. Liu, Hao Chen, Chunhua Shen, Tong He, Lian-Wen Jin, and L. Wang. ABCNet: Real-time scene text spotting with adaptive bezier-curve network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020