

HƯỚNG TIẾP CẬN SỬ DỤNG THƯ VIỆN HỖ TRỢ CHO BÀI TOÁN NHẬN DẠNG CHỮ TRONG HÌNH ẢNH VÀ XÂY DỰNG BỘ DỮ LIỆU VỀ NHẬN DẠNG CHỮ TRONG HÌNH ẢNH CHO NGÔN NGỮ TIẾNG VIỆT

Thái Trần Khánh Nguyên

Nguyễn Khánh Như

Lê Đoàn Thiện Nhân

Trường Đại học Công Nghệ Thông Tin

Lớp CS519.M11.KHCL

What ?

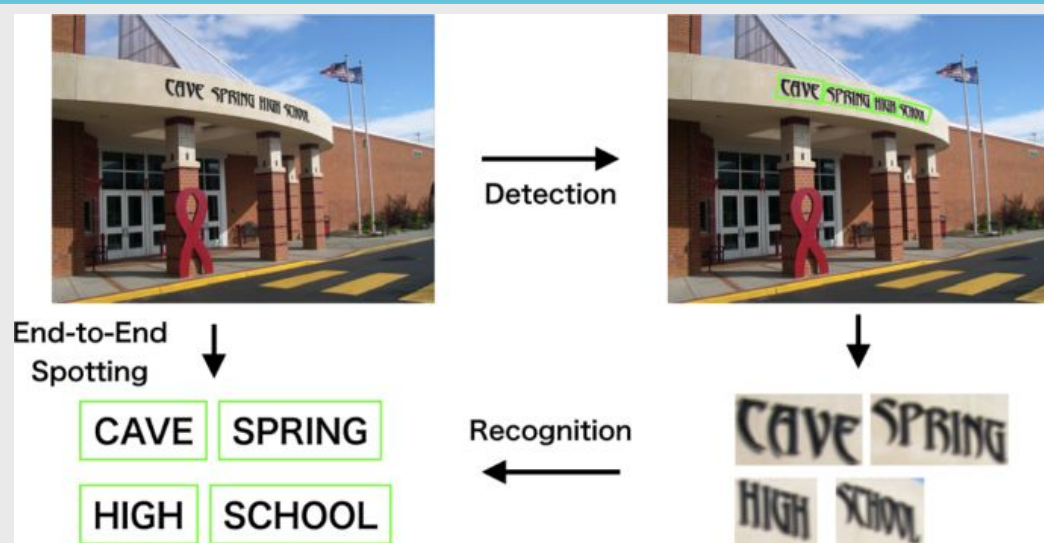
- Chúng tôi đề xuất một hướng tiếp cận mới cho việc sử dụng từ điển hỗ trợ thay vì buộc mô hình phải dự đoán đầu ra là một từ thuộc từ điển thì tích hợp từ điển vào mô hình. Từ đó, tạo ra danh sách các từ ứng cử viên phù hợp, sau đó, đưa vào module tính điểm để tìm ra từ tương thích nhất với các đặc điểm ngoại hình.
- Chúng tôi đề xuất việc xây dựng bộ dữ liệu đối với bài toán Nhận dạng văn bản trong hình ảnh cho ngôn ngữ Tiếng Việt

Why ?

Sử dụng từ điển để cải thiện hiệu quả mô hình: hướng tiếp cận hiện tại còn đơn giản, một vài bất cập khi sử dụng phương pháp. Các từ nước ngoài hoặc không có trong từ điển. Không có sự phản hồi trong quy trình suy luận khi huấn luyện. Khoảng cách dùng để xác định từ được xuất ra còn nhiều hạn chế

Dữ liệu dành cho bài toán này cho ngôn ngữ Tiếng Việt còn hạn chế, cần có một bộ dữ liệu cho cộng đồng nghiên cứu ở Việt Nam phát triển bài toán trên ngữ cảnh Tiếng Việt

Overview



Bảng phân phối dữ liệu



Mô tả chi tiết

1. Mục Tiêu Nghiên Cứu

- Đề xuất cách tiếp cận mới, thay vì buộc mô hình phải dự đoán đầu ra là một từ thuộc từ điển thì tích hợp từ điển vào mô hình. Từ đó, tạo ra danh sách các từ ứng cử viên phù hợp, sau đó, đưa vào module tính điểm để tìm ra từ tương thích nhất với các đặc điểm ngoại hình.
- So sánh kết quả của mô hình Scene Text Recognition cơ bản và phiên bản cải tiến (tích hợp từ điển vào mô hình ban đầu) bằng cách sử dụng các bộ dataset TotalText, ICDAR2013, ICDAR2015 và VinText để đánh giá
- Xây dựng bộ dataset cho bài toán Scene Text Recognition trong ngữ cảnh tiếng Việt.

3. Kết quả mong đợi

- Hướng tiếp cận mới sẽ đạt được state-of-the-art khi so sánh (h-means) với các giải pháp trước đây
- Chứng minh rằng bộ dataset cho vấn đề Scene Text Recognition đến từ team Vin (VinText) là một bộ dataset có độ tin cậy cao và đáng để được đưa vào sử dụng rộng rãi.

2. Nội dung và Phương pháp

- Nghiên cứu về các mô hình detection tìm ra các mô hình tốt nhất để sử dụng trong quá trình nghiên cứu.
- Nghiên cứu về cách thức sinh danh sách ứng cử viên trong quá trình suy luận và huấn luyện
- Nghiên cứu về các khoảng cách để có thể tìm ra cách điều chỉnh khoảng cách thích hợp với hướng tiếp cận mới này.
- Thực hiện việc thu thập dữ liệu cho bộ dữ liệu VinText thông qua các cách như thu thập trên internet và chụp hình thực tế.
- Nghiên cứu về việc sinh tập ứng cử viên (Candidate generation). Từ điển sẽ được nghiên cứu và được dùng trong việc tạo ra danh sách các từ ứng cử viên trong cả quá trình suy luận (inference) và huấn luyện (training). Trong quá trình huấn luyện, chúng tôi đặt giả thuyết rằng danh sách cách từ ứng cử viên là k từ trong bộ từ điển có khoảng cách Levenshtein nhỏ nhất đến kết quả dự đoán ban đầu. Hàm mất mát sẽ được chúng tôi nghiên cứu và chỉnh sửa sao cho phù hợp với phương pháp mà chúng tôi đề xuất.

