

# NHẬN DẠNG CHỮ TIẾNG VIỆT Ở NGOẠI CẢNH

Nguyễn Ích Tài

Nguyễn Quế Phong

Đường Minh Khang

<sup>1</sup> Trường Đại học Công nghệ Thông tin  
TPHCM

<sup>2</sup> Đại học Quốc gia  
Thành phố Hồ Chí Minh

<sup>3</sup> {18521362, 21520398, 21520957}  
@gm.uit.edu.vn

## Mục tiêu

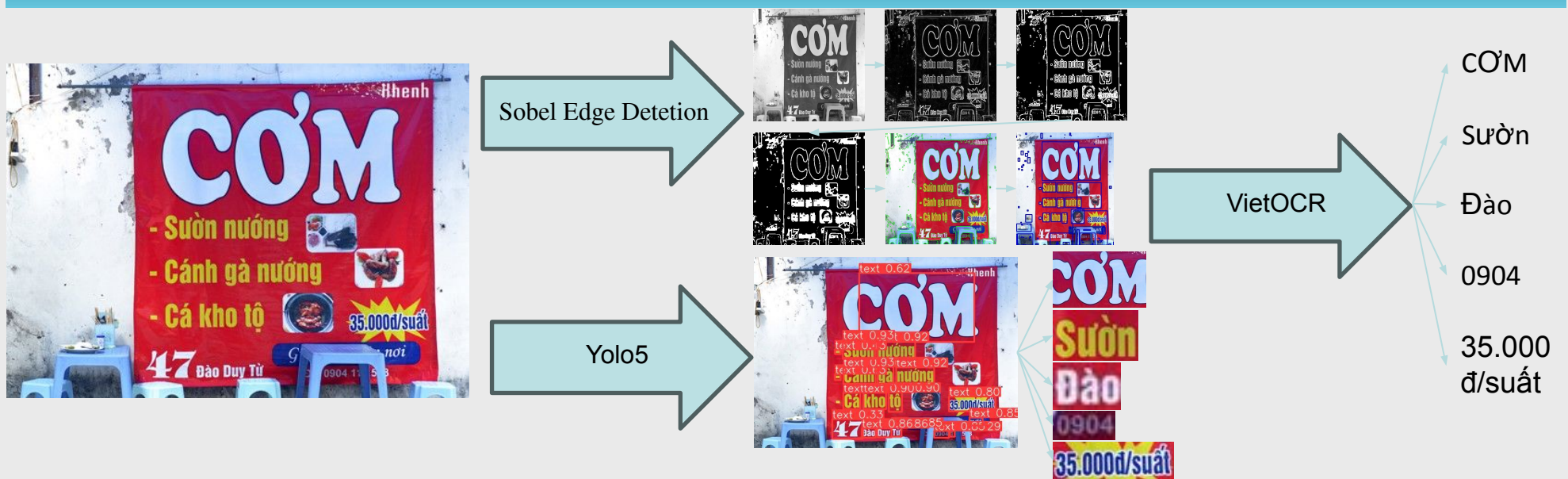
Hai phương pháp nhận diện chữ tiếng Việt trong ngoại cảnh kết hợp giữa Edge Detection(hoặc Yolo5) và VietOCR được nghiên cứu nhằm mục đích:

- Xây dựng một mô hình tổng quan, dễ cài đặt với độ chính xác cao.
- Áp dụng được mô hình với nhiều vấn đề thực tiễn.
- Nghiên cứu, thử nghiệm và đánh giá mô hình để đưa ra định hướng phát triển và hoàn thiện các vấn đề hạn chế.

## Lý do chọn đề tài

- Trước sự bùng nổ của dữ liệu hình ảnh trên Internet và sự phổ biến của các thiết bị di động với khả năng chụp ảnh ngày càng cao, nhu cầu tự động hóa việc nhận diện và phân tích chữ viết trên ảnh đã trở nên ngày càng quan trọng.
- Việc nghiên cứu bài toán nhận dạng chữ tiếng Việt trong ngoại cảnh mang lại nhiều ứng dụng thực tế như: Tự động nhận diện biển số xe, chuyển đổi thông tin từ các tài liệu giấy... Hơn hết thảy, đây là đề tài có thể được ứng dụng, cải thiện và phát triển cho người Việt Nam chúng ta.

## Tổng quan



## Chi tiết

### 1. Dataset

- Dự kiến sử dụng bộ dataset: Vintext được sử dụng trong cuộc thi AICChallenge 2021

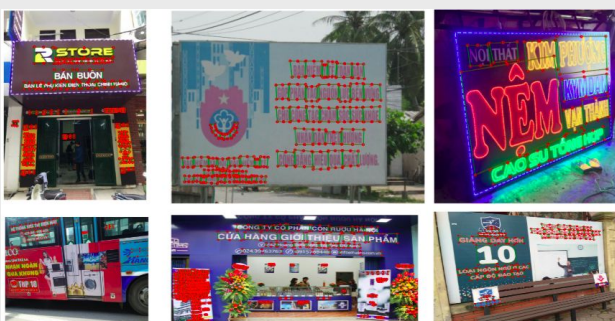


Figure 1. Vintext dataset của AICChallenge 2021.

### 2. Phương pháp

#### • Edge Detection

- Từ ảnh input đầu vào, sử dụng thuật toán Sobel Edge Detection và các hàm. Có sẵn trong thư viện OpenCV để tìm vị trí chứa văn bản.
- Cắt ảnh các vị trí đã tìm được.
- Sử dụng phép Morphology Closing để lấp những khoảng trống với kích thước kernel là 5x5.
- Dùng hàm contour để tìm các đường bao quanh các biên cạnh tìm thấy được sau khi sử dụng Closing.
- Cắt các vùng được tìm thấy bởi contour và đưa ảnh đã cắt vào mô hình VietOCR để dự đoán.

#### • YOLOV5

- Từ ảnh input đầu vào, đưa ảnh qua mô hình YOLOv5 để tìm xác định vị trí của vùng chứa văn bản.
- Tiến hành cắt ảnh từ các vị trí tìm thấy qua mô hình YOLOv5.



Figure 2. YOLOv5 + VietOCR.

#### • VietOCR

- Để cho ra được kết quả cao, nhóm tiến hành pre-trained VietOCR trên bộ dữ liệu Vintext bản Original với định dạng vị trí 4 đỉnh và nhãn. Tiến hành cắt ảnh từ các vị trí tìm thấy qua mô hình YOLOv5.
- Tiến hành cắt ảnh với vị trí 4 đỉnh tương ứng với nhãn của chúng.
- Đưa dữ liệu đã được xử lý và huấn luyện trên mô hình pre-trained VGG19 - Transformer (vgg\_transformer) với 5000 iters.

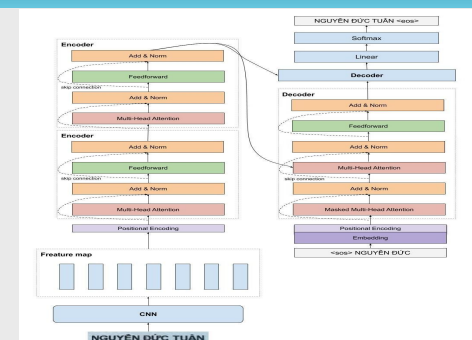


Figure 3. VietOCR.

### 3. Kết quả dự kiến

- Phương pháp YOLOV5 + VietOCR sẽ mang lại kết quả khả quan hơn trên tập dữ liệu được sử dụng.
- Tạo ra một mô hình nhận diện chữ cái tiếng Việt có thể sử dụng được trên nhiều bức ảnh thực tế với độ sai số không đáng kể.
- Cải thiện được những mặt hạn chế, nghiên cứu thêm nhiều phương pháp khác nhằm vận dụng và hoàn thiện mô hình nhưng vẫn giữ được mục tiêu đặt ra ban đầu.

### 4. Kế hoạch thực hiện

