

Báo cáo bài tập lớn

Thành viên

- Nguyễn Hữu Vượt - 19020489
- Lê Huy Vũ - 19020056
- Kiều Thế Vinh - 19020486
- Phạm Văn Hoàn - 18020538
- Nguyễn Trung Kiên - 18020731
- Nguyễn Văn Chính - 19020232

Phương pháp sử dụng

Với ảnh có đầy đủ line:

- chuyển ảnh ban đầu sang `gray_scale` và chuyển sang binary sử dụng threshold với mức `threshold = 150` (điểm ảnh có giá trị $> \text{threshold} \Rightarrow$ set giá trị thành 255, còn lại thì set về 0)
- tách các đường line trong ảnh sử dụng phương pháp biến đổi hình thái ***open*** với kernel tương ứng:
 - phát hiện đường nằm dọc: (1,15)
 - phát hiện đường nằm ngang: (15,1)
 - Kết hợp hai kết quả ở trên thành một binary image với các line được làm dày
- sử dụng phương thức **`connectedComponentWithStats()`** để lấy ra giao điểm của các line trong image \Rightarrow tọa độ điểm trên cùng bên trái , chiều dài (width) và chiều rộng của cell (height)
- từ ảnh binary ban đầu \Rightarrow dùng tọa độ các điểm tìm được ở trên cắt ảnh binary thành các ảnh là các cell nhỏ trong ảnh, mỗi cell sẽ có background trắng và text màu đen
- với mỗi ảnh cell được cắt ra, tìm các giá trị `xmin`, `xmax`, `ymin`, `ymax` của vùng chứa text màu đen \Rightarrow detect text trong cell

Với ảnh không có line:

- Tiền xử lý ảnh:
 - áp dụng phương pháp **threshold** để chuyển ảnh sang binary, biến đổi hình thái **open**, **findContours** để xoá đi tất cả đường thẳng trong cell.
- xác định box:
 - dùng **dilation** để biến các đối tượng text đứng liền nhau thành một khối duy nhất
 - dùng **findContour** và **boundingRect** để xác định toạ độ và kích thước của box
- lưu ý: một số ảnh có các text sát liền các đường line, nên để có kết quả tốt, ta thực hiện xoá horizontal line trước, sau đó xoá vertical line hoặc ngược lại

Kịch bản training:

- B1: sử dụng phương pháp dùng với ảnh không có line để detect \Rightarrow tính toán IoU
- B2: với các ảnh có tỉ lệ chính xác $< 60\%$ \Rightarrow sử dụng phương pháp xử lý ảnh có line

Kết quả:

- thuật toán hoạt động tốt trên các ảnh có dữ liệu số, với các dữ liệu là văn bản có line rõ ràng.
- Với một số ảnh có kết quả thấp có thể thay đổi kích thước của kernel khi thực hiện phép **dilation** để cải thiện độ chính xác của thuật toán

Ưu điểm:

- phát hiện tốt trên ảnh có các cell chứa số, các cell có line rõ ràng

Nhược điểm và hạn chế:

- thuật toán hoạt động kém trên các ảnh có background tối màu hoặc nhiều màu khác nhau
- kernel hiện tại chưa thực sự phù hợp với ảnh có cell không có line, chứa văn bản nhiều dòng và khoảng cách giữa các ký tự lớn

