Chương 2. Độ chính xác của các phương pháp trích xuất văn bản từ hình ảnh

**1.Độ chính xác (Accuracy metric)**

Khi xây dựng bất kì thuật toán nào, các lập trình viên luôn mong muốn đạt tỉ lệ chính xác là 100%. Tuy nhiên, không phải lúc nào thuật toán cũng có thể được xây dựng để luôn đạt được kết quả đúng.

Ở 2 phương thức phân tách văn bản từ hình ảnh được đề cặp đến trong bài báo cáo này, ta có thể tính độ chính xác của thuật toán theo Accuracy metric . Công thức:

Accuracy =

Trong đó :

TP: True positive rate ,các văn bản xuất hiện trong ảnh được nhận diện đúng.

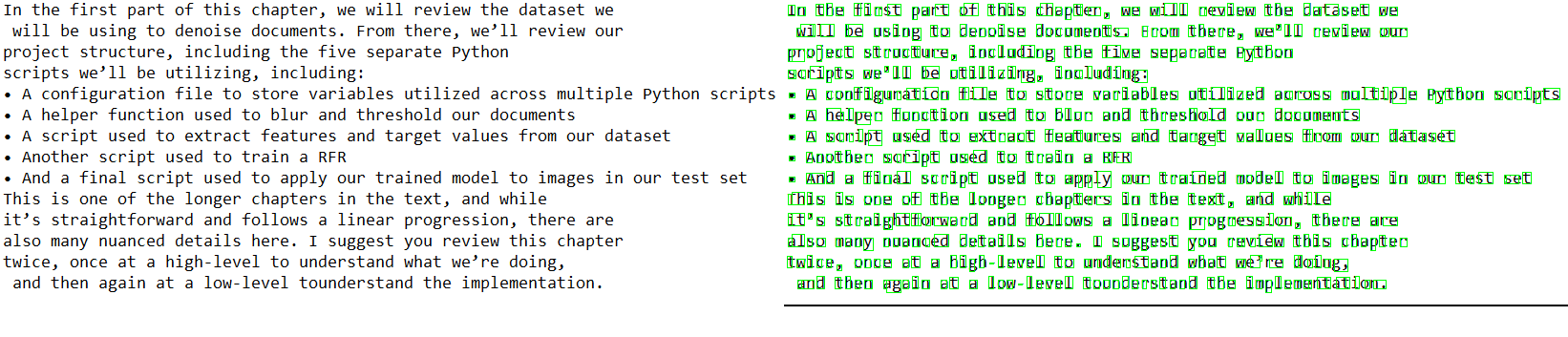
TN: True negative rate, các văn bản không xuất hiện trong hình được nhận diện đúng

FP :False positive rate, các văn bản bị nhận diện sai trong trong hình ảnh

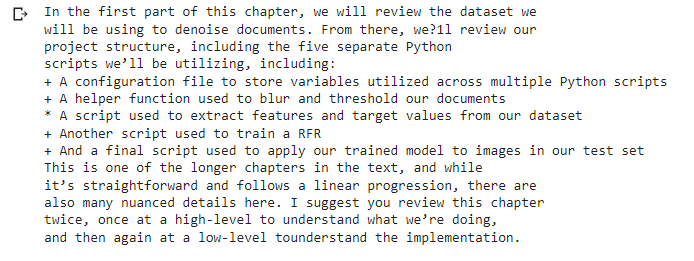
FN:False negative rate, các văn bản bị bỏ xót trong hình ảnh.

1. Độ chính xác của phương pháp Pytesseract OCR:

2.1.Thực hiện một quá trình nhận diện các kí tự văn bản trên 1 ảnh



Hình 10. Ảnh input đầu vào (trái) và ảnh sau khi detect text area từng contours (phải)

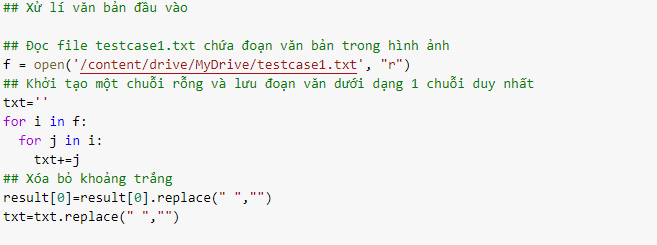


Hình 11. Kết quả thu được từ hình ảnh đầu vào thông qua phương pháp Pytesseract OCR:

2.2 Thực hiện tính toán độ chính xác của phương pháp:

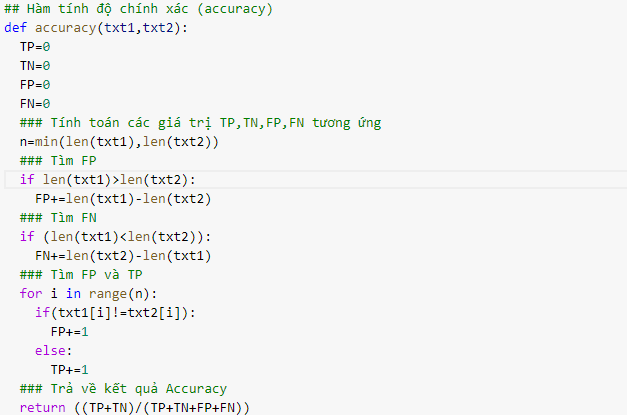
2.2.1 Tính độ chính xác theo kí tự

Xử lí dữ liệu đầu vào



Hình 12. Đoạn mã xử lí dữ liệu đầu vào để tính độ chính xác (Accuracy)

Hàm tính độ chính xác (Accuracy)



Hình 13.Hàm tính độ chính xác(Accuracy)

Kết quả:

Kết quả thu được khi thực hiện tính độ chính xác của phương pháp Pytesseract OCR



Hình 14. Kết quả độ chính xác (Accuracy)

2.2.2 Tính độ chính xác theo từ

Xử lí dữ liệu đầu vào



Hình 15 Đoạn mã xử lí dữ liệu đầu vào để tính độ chính xác (Accuracy)

Kết quả:



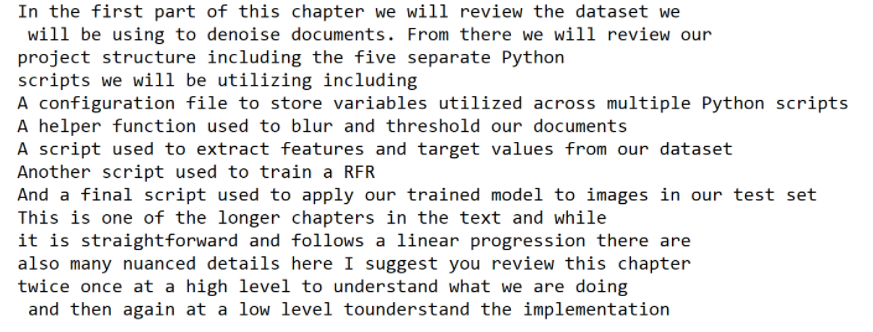
2.3 Đánh giá kết quả

Đối với Testcase trên, phương pháp Pytesseract OCR tuy không đạt được độ chính xác tuyệt đối những vẫn ở một mức tương đối cao (xấp xĩ 99%). Trong thực tế, một vài yếu tố từ hình ảnh đầu vào sẽ ảnh hưởng đến độ chính xác của thuật toán (chất lượng hình ảnh, kích thước hình ảnh,độ nghiêng của các văn bản trong ảnh,…), điều này sẽ ảnh hưởng khá lớn đến độ chính xác của các phương pháp trích xuất văn bản từ hình ảnh.

1. Độ chính xác của phương pháp Keras OCR:

3.1.Thực hiện một quá trình nhận diện các kí tự văn bản trên 1 ảnh

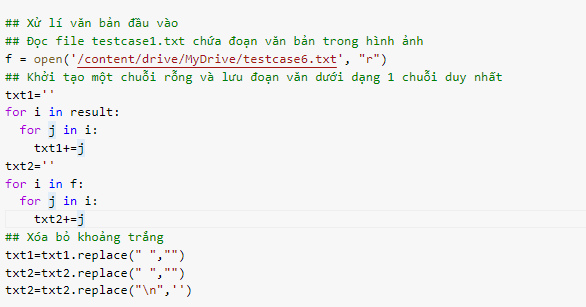
Hình ảnh đầu vào thực hiện đánh giá độ chính xác

Hình 16. Ảnh input đầu vào

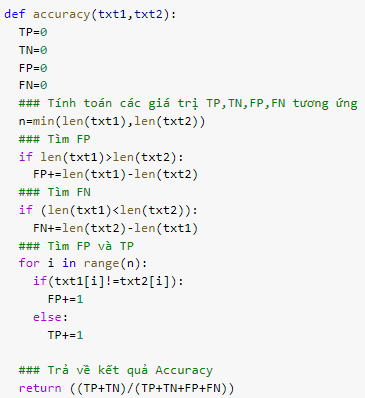
3.2 Thực hiện tính toán độ chính xác của phương pháp:

3.2.1 Thực hiện tính độ chính xác theo kí tự:

Xử lí dữ liệu đầu vào:

Hình 17 Đoạn mã xử lí dữ liệu đầu vào để tính độ chính xác (Accuracy)

Hàm tính độ chính xác



Hình 18.Hàm tính độ chính xác(Accuracy)

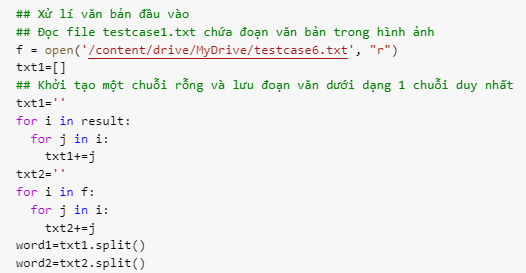
Kết quả:



Hình 19. Kết quả accuracy

3.2.2 Thực hiện tính độ chính xác theo từ:

Xử lí dữ liệu đầu vào:



Hình 20 Đoạn mã xử lí dữ liệu đầu vào để tính accuracy

Kết quả:



3.3 Đánh giá kết quả:

Đối với Testcase trên, phương pháp Pytesseract OCR tuy không đạt được độ chính xác tuyệt đối những vẫn ở một mức tương đối cao (xấp xĩ 98% đối với phương pháp tính theo kí tự và 89% theo từ). Tương tự với phương pháp Pytesseract OCR, một vài yếu tố từ hình ảnh đầu vào sẽ ảnh hưởng đến độ chính xác của thuật toán (chất lượng hình ảnh, kích thước hình ảnh,độ nghiêng của các văn bản trong ảnh,…), điều này sẽ ảnh hưởng khá lớn đến độ chính xác của các phương pháp trích xuất văn bản từ hình ảnh.

1. Một số phương pháp làm tăng độ chính xác khi xử lí ảnh:
   1. Nâng cấp chất lượng hình ảnh

Chất lượng hình ảnh đóng vai trò rất quan trọng trong các vấn đề xử lí ảnh. Một bức ảnh có độ phân giải, độ sắc nét và độ tương phản cao sẽ giúp cải thiện rất lớn độ chính xác của các phương pháp trích xuất văn bản từ hình ảnh.

4.2 Loại bỏ các nhiễu trong hình ảnh

Các nhiễu trong ảnh cũng là một trong những yếu tố đáng quan tâm đến, cần phải giảm thiểu các nhiễu này đến mức tối thiểu trước khi thực hiện quá trình trích xuất văn bản để đạt được kết quả với độ chính xác cao.

* 1. Một số yếu tố khác ảnh hưởng đến chất lượng của quá trình trích xuất dữ liệu văn bản từ hình ảnh

Để tăng độ chính xác của các phương phác phân tách, ta còn cần chú ý đến một vài vấn đề khác .Các văn bản trong hình ảnh có thể xảy ra các trường hợp ảnh hưởng đến khả năng phân tách của các thuật toán như : độ cong, độ nghiêng, độ bóp méo, loại font chữ của các kí tự.