Bọn em đang thực hiện dự án “Scaner (Extract Document from Image)”, chức năng chính của dự án này là:

* **Quét hình ảnh từ nhiều nguồn***:* Hệ thống có khả năng quét hình ảnh từ nhiều nguồn khác nhau như máy quét, ảnh chụp từ thiết bị di động, hoặc các tệp hình ảnh đã tồn tại.
* **Nhận dạng và trích xuất văn bản*:*** Sử dụng công nghệ OCR, hệ thống nhận dạng và trích xuất văn bản từ hình ảnh, bao gồm cả ký tự in và viết tay.
* **Tiên đoán văn bản:** Phục vụ cho mục đích phục hồi tài liệu hay thao tác với chúng.
* **Xử lý ảnh và cải thiện chất lượng:** Hệ thống có thể thực hiện các bước xử lý ảnh như điều chỉnh độ sáng, loại bỏ nhiễu, cải thiện độ tương phản để tăng khả năng nhận dạng và chính xác của OCR.
* **Tự động phân loại và định dạng tài liệu:** Hệ thống có thể phân loại các loại tài liệu khác nhau và áp dụng các định dạng văn bản phù hợp (ví dụ: PDF, văn bản thuần túy, văn bản được soạn thảo, vv.).
* **Tương tác người dùng:** Hệ thống cung cấp giao diện người dùng thân thiện để người dùng có thể xem trước hình ảnh quét, chỉnh sửa văn bản nếu cần thiết, và tương tác với dữ liệu đã trích xuất.

Bọn em đã làm được hết phần chức năng 1 và khi đang làm được một nửa ở chức năng 2 và chức năng 4 thì bọn em gặp một số khó khăn như là:

* **Xóa background:** Bọn em đã xoá được gần như là background của các ảnh nhưng mà chỉ cần background bị ánh sáng chiếu vào tạo nên một vùng sáng thì hệ thống tự nhận diện lấy luôn viền phần đó mặc dù bọn em đã chỉnh toạ độ ABCD rồi ạ. Cái khác thì lấy được hết khung rồi nhưng chữ lại có màu liền với nền tài liệu thì hệ thống lại không nhận ra được mặt chữ. Nó lại dẫn đến vấn đề thứ 2 ạ.
* **Làm rõ chữ:** Bọn em đã làm được tới bước này rồi nhưng mà như vấn đề em nói trên khi mà màu chữ bị loang ra với màu nền tài liệu thì nó tự động bỏ luôn chữ và hệ thống cũng chưa nhận dạng được mặt chữ và không lấy được chữ nếu các chữ quá xích nhau và bị mờ, bọn em cũng đã thử cách tăng độ phân giải lên 2000 2400 …, nhưng mà độ phân giải càng cao thì hệ thống trải qua các bước càng nặng dẫn đến lag hệ thống và đơ máy+code luôn ạ.
* **Khử nhiễu:** Có một số ảnh tài liệu phần background bên ngoài nó bị nhiều tác nhân gây cản trở cho việc nhận dạng viền tài liệu, ví dụ như giấy tờ đó được chụp trên khăn trải bàn có rất nhiều màu sắc và các hình dạng như hoa lá, hình vuông, hình tròn,… tạo nên nhiễu gây cho hệ thống nhận dạng luôn các phần đó mà không tập trung vào các chữ cũng như là viền ạ.
* **Mô hình AI và training:** Phần này thì bọn em đã đọc nhiều tài liệu cũng như xem video về lý thuyết này nhưng mà bọn em vẫn không hiểu được cách nó hoạt động cũng như triển khai như thế nào ạ.
* **Giao diện cho người dùng:** Nhóm em muốn hỏi là khi mình hoàn thành xong hệ thống rồi thì làm sao để có thể đưa hệ thống này lên giao diện web được ạ?

Em cảm ơn thầy đã dành thời gian đọc câu hỏi của nhóm bọn em ạ. Em xin cảm ơn.