

HỌC VIỆN CÔNG NGHỆ BUỒU CHÍNH VIỄN THÔNG

KHOA CNTT1



BÁO CÁO BÀI TẬP LỚN XỬ LÝ ẢNH

ĐỀ TÀI 6: Dự án Trích xuất Thông tin từ Căn cước công dân Việt Nam

GIẢNG VIÊN HƯỚNG DẪN: THẦY PHẠM HOÀNG VIỆT

LỚP : D22HTTT- 06

NHÓM LỚP MÔN HỌC : 11

NHÓM BÀI TẬP LỚN: 03

THÀNH VIÊN NHÓM: LÊ VĂN ĐỨC - B22DCCN228

PHẠM THÀNH NAM - B22DCCN563

TÔ QUANG HUY- B22DCCN396

Hà Nội, 2025

Mục Lục

1. Bối cảnh và Mục tiêu Dự án	3
1.1. Lý do và Sứ cần thiết	3
1.2. Đối tượng nghiên cứu :	4
1.3. Mục đích nghiên cứu.....	4
2. Công nghệ sử dụng:	5
2.1 Pytorch	5
2.2 Flask API.....	5
2.3 Gradio.....	6
3. Báo cáo dự án:	6
3.1 Phân tích và thiết kế hệ thống	6
3.1.1. Sơ đồ trường hợp sử dụng tổng thể (Overall use case diagram)	6
3.1.2. Sơ đồ hoạt động (Activity diagram):	8
3.1.3: Kiểm thử test Case:	9
3.1.4: Các thử nghiệm quan trọng bao gồm:	12
3.2. Phương pháp luận (Methodology)	12
3.2.1 Phát hiện góc (Corner detection).....	13
3.2.2. Phân loại (Classification)	13
3.3 Tăng cường dữ liệu (Data Augmentation)	14
4. Tiến độ dự án :	15
5. Kết quả thực nghiệm :.....	16
5.1 Bộ dữ liệu (Dataset)	16
5.2. Hiệu suất và các chỉ số	16
5.3. Đánh giá Tổng thể	18
6. Kết Luận.....	19
6.1. Ưu điểm.....	19
6.2. Hạn chế.....	19
6.3. Cải tiến Tương lai	19
7. Tài liệu tham khảo :	20

1. Bối cảnh và Mục tiêu Dự án

1.1. Lý do và Sứ cần thiết

Trong kỷ nguyên số hiện tại, nhu cầu về tự động hóa và các hệ thống quản lý dữ liệu hiệu quả đang gia tăng trên nhiều lĩnh vực, bao gồm chính phủ, tài chính, chăm sóc sức khỏe, v.v.. Thông tin nhận dạng cá nhân đóng vai trò then chốt và thường cần được xử lý trên quy mô lớn, đặt ra những thách thức liên quan đến độ chính xác, tốc độ và bảo mật.

Tại Việt Nam, Thẻ Căn cước Công dân Quốc gia (National ID card) là hình thức nhận dạng cơ bản cho phần lớn dân số. Thẻ này chứa thông tin cá nhân quan trọng như họ tên đầy đủ, ngày sinh, giới tính, địa chỉ và một số định danh duy nhất, cần thiết cho nhiều loại dịch vụ.

Tự động hóa quá trình trích xuất dữ liệu từ các tài liệu này có thể cải thiện đáng kể hiệu quả và giảm thiểu lỗi do con người, trở thành giải pháp hấp dẫn cho nhiều tác vụ hành chính và dịch vụ. Bằng cách tự động hóa việc trích xuất thông tin cá nhân từ CCCD Việt Nam, dự án này đáp ứng nhu cầu ngày càng tăng về xử lý dữ liệu nhanh chóng, an toàn và chính xác. Hơn nữa, quá trình tự động hóa giúp hợp lý hóa quy trình làm việc trong các cơ quan chính phủ, tổ chức ngân hàng, nhà cung cấp dịch vụ chăm sóc sức khỏe và các tổ chức tư nhân khác, đảm bảo dịch vụ đáng tin cậy và dễ tiếp cận hơn cho người dùng.

Sự ra đời của các công nghệ hiện đại như Nhận dạng Ký tự Quang học (OCR) và phát hiện đối tượng (object detection) mang đến một con đường thực tế để giải quyết những thách thức này. Kỹ thuật phát hiện đối tượng có thể được sử dụng để định vị và trích xuất các khu vực quan tâm cụ thể (chẳng hạn như các trường văn bản và số định danh), trong khi OCR có thể được sử dụng để đọc và số hóa văn bản một cách chính xác. Bằng cách kết hợp hai công nghệ này, dự án nhằm mục đích tối ưu hóa quy trình trích xuất dữ liệu, cải thiện bảo mật dữ liệu và nâng cao hiệu quả dịch vụ trên nhiều lĩnh vực. Tự động hóa nhiệm vụ này sẽ giảm đáng kể thời gian dành cho việc nhập liệu thủ công, cho phép xác minh dữ liệu nhanh hơn và chính xác hơn trong các lĩnh vực như ngân hàng, chăm sóc sức khỏe và dịch vụ chính phủ.

1.2. Đối tượng nghiên cứu :

Đối tượng nghiên cứu chính của dự án là **Thẻ Căn cước Công dân Quốc gia Việt Nam**. Thẻ này đóng vai trò là tài liệu nhận dạng chính thức và chứa các chi tiết cá nhân quan trọng, bao gồm nhưng không giới hạn ở họ tên, ngày sinh, giới tính, số ID và địa chỉ nhà của chủ thẻ.

Để trích xuất dữ liệu này một cách hiệu quả, dự án sử dụng sự kết hợp của các công nghệ hiện đại. Phát hiện đối tượng đóng vai trò quan trọng trong việc xác định và định vị các vị trí chính xác của các khôi văn bản quan trọng trên thẻ ID. Thông qua phát hiện góc (corner detection), các phần quan trọng như số ID, tên và địa chỉ được xác định, cho phép quy trình trích xuất hiệu quả. Khi các phần liên quan đã được xác định, Nhận dạng Ký tự Quang học (OCR) được sử dụng để chuyển đổi văn bản được phát hiện thành định dạng máy đọc được, đảm bảo trích xuất dữ liệu chính xác.

Nghiên cứu khám phá sự tích hợp của hai công nghệ này — phát hiện đối tượng để xác định cấu trúc của thẻ ID và OCR để nhận dạng văn bản — nhằm phát triển một giải pháp mạnh mẽ để tự động hóa việc trích xuất dữ liệu cá nhân từ CCCD Việt Nam.

1.3. Mục đích nghiên cứu

Mục đích chính của nghiên cứu này là phát triển một hệ thống tự động, toàn diện có khả năng trích xuất thông tin cá nhân từ Thẻ Căn cước Công dân Việt Nam với mức độ chính xác và hiệu quả cao.

Dự án nhằm mục đích khám phá và tích hợp các kỹ thuật phát hiện đối tượng tiên tiến — đặc biệt là phát hiện góc — để định vị các khu vực chính chứa thông tin quan trọng trên thẻ. Sau khi các khu vực này được xác định, công nghệ OCR được sử dụng để chuyển đổi thông tin văn bản thành định dạng có cấu trúc, máy đọc được.

Bằng cách kết hợp các công nghệ tiên tiến này, dự án tìm cách:

- **Nâng cao độ chính xác và hiệu quả trích xuất dữ liệu:** Phát hiện đối tượng đảm bảo các phần liên quan của thẻ ID được định vị chính xác, trong khi OCR giúp chuyển đổi văn bản thành định dạng có thể được xử lý bởi các hệ thống kỹ thuật số.

- **Giảm thiểu nhập liệu thủ công:** Hệ thống được thiết kế để tự động hóa công việc nhập dữ liệu thủ công tẻ nhạt và dễ xảy ra lỗi từ thẻ ID, giảm thiểu nguy cơ lỗi của con người và tiết kiệm thời gian.
- **Cải thiện tốc độ xử lý dữ liệu:** Bằng cách tự động hóa quá trình trích xuất, dự án nhằm mục đích tăng tốc đáng kể các hoạt động nhập dữ liệu, đặc biệt có giá trị trong các bối cảnh như dịch vụ tài chính, dịch vụ chính phủ và quản lý hồ sơ sức khỏe.

2. Công nghệ sử dụng:

Dự án đã sử dụng sự kết hợp của các công nghệ hiện đại để phát triển một hệ thống mạnh mẽ, hiệu quả và có khả năng mở rộng để trích xuất thông tin từ CCCD Việt Nam. Các công nghệ chính được sử dụng là **PyTorch**, **Flask API**, **Gradio**.

2.1 Pytorch

PyTorch là một thư viện học sâu hiệu suất cao, kiểu mệnh lệnh (imperative style). Được phát triển bởi phòng thí nghiệm AI Research của Facebook, PyTorch đã nhanh chóng trở nên phổ biến trong cả giới học thuật và công nghiệp nhờ tính linh hoạt, dễ sử dụng và hiệu suất hiệu quả. Cốt lõi của nó, PyTorch được thiết kế để cung cấp đồ thị tính toán động, cho phép các nhà phát triển xây dựng, sửa đổi và gỡ lỗi mạng nơ-ron một cách dễ dàng trong thời gian chạy.

PyTorch hỗ trợ nhiều tác vụ học máy, bao gồm thị giác máy tính, xử lý ngôn ngữ tự nhiên và học tăng cường. Khả năng tích hợp liền mạch với Python và khả năng tăng tốc GPU hiệu quả làm cho nó đặc biệt phù hợp với môi trường nghiên cứu và sản xuất.

2.2 Flask API

Flask là một khung công tác web nhẹ và linh hoạt cho Python, được thiết kế để cho phép phát triển nhanh chóng các ứng dụng web và API. Nó đặc biệt phù hợp để tạo API RESTful, làm cho nó trở thành lựa chọn lý tưởng cho phát triển back-end. Cách tiếp cận tối giản của Flask mang lại cho các nhà phát triển sự tự do để chọn các công cụ và thư viện họ cần, góp phần vào sự phổ biến của nó trong việc xây dựng các dịch vụ web có thể mở rộng và hiệu quả.

Trong dự án này, Flask được sử dụng làm khung back-end, nơi nó xử lý các yêu cầu đến, xử lý dữ liệu và trả về các phản hồi. Nó đóng vai trò là trung gian giữa giao diện người dùng front-end và logic phức tạp của các hệ thống học máy.

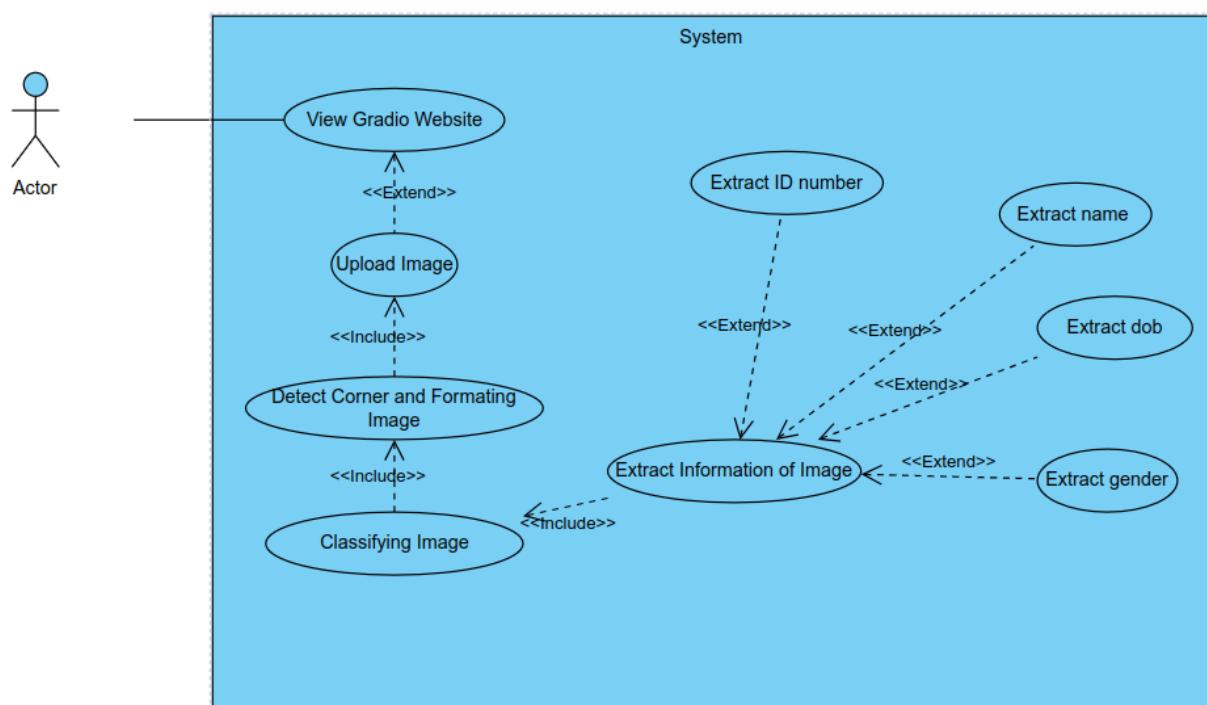
2.3 Gradio

Gradio là một thư viện Python mạnh mẽ giúp đơn giản hóa việc tạo giao diện người dùng tương tác, dựa trên web cho các mô hình học máy. Trong dự án này, Gradio được sử dụng để phát triển một giao diện trực quan và thân thiện với người dùng được thiết kế đặc biệt để tạo điều kiện trích xuất thông tin từ hình ảnh CCCD Việt Nam. Giao diện cho phép người dùng dễ dàng tải lên hình ảnh, xem dữ liệu được trích xuất và tương tác với hệ thống một cách liền mạch.

3. Báo cáo dự án:

3.1 Phân tích và thiết kế hệ thống

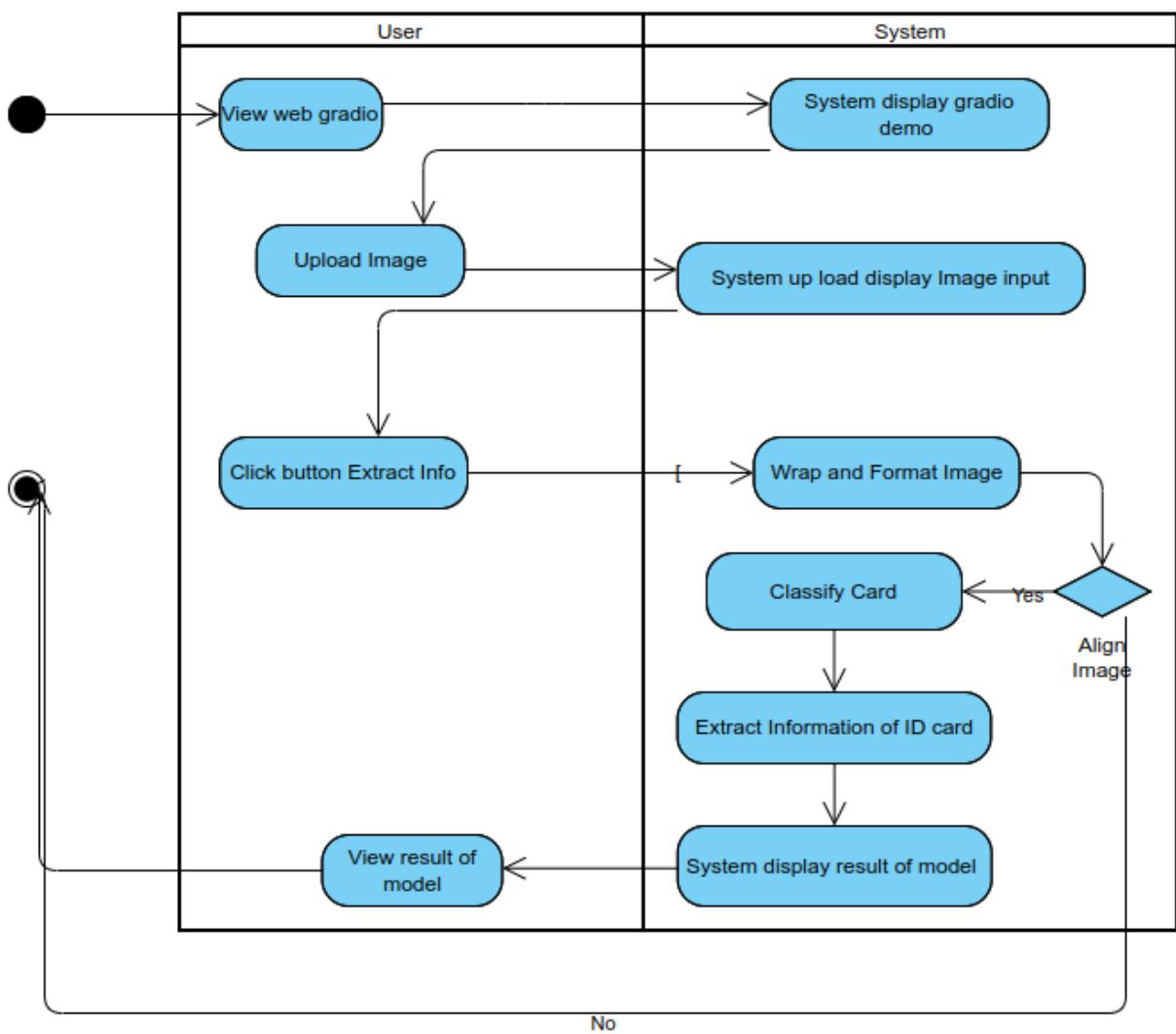
3.1.1. Sơ đồ trường hợp sử dụng tổng thể (Overall use case diagram)



Sơ đồ minh họa sự tương tác giữa người dùng (actor) và hệ thống được thiết kế để trích xuất thông tin từ thẻ ID Việt Nam thông qua trang web Gradio.

- **Actor:** Người dùng tương tác với trang web Gradilo.
- **System Boundary:** Chức năng được cung cấp bởi trang web Gradilo.
- **Use Cases** (Các trường hợp sử dụng):
 - **Xem Trang web Gradilo (View Gradilo Website).**
 - **Tải lên Hình ảnh (Upload Image):** Đây là một bước bắt buộc (<<include>>) trong quy trình làm việc.
 - **Phát hiện Góc và Định dạng Hình ảnh (Detecting Corner and Formatting Image).**
 - **Phân loại Hình ảnh (Classifying Image).**
 - **Trích xuất Số ID (Extract ID Number).**
 - **Trích xuất Tên (Extract Name).**
 - **Trích xuất Ngày sinh (Extract Date of Birth - DOB).**
 - **Trích xuất Giới tính (Extract Gender).**

3.1.2. Sơ đồ hoạt động (Activity diagram):



Sơ đồ phác thảo quy trình trích xuất thông tin từ thẻ ID, bao gồm các bước chính sau:

- Bắt đầu (Start):** Quá trình bắt đầu khi thẻ ID được trình bày để quét.
- Quét Thẻ ID (Scan ID Card).**
- Tiền xử lý Hình ảnh (Image Preprocessing):** Hình ảnh đã quét được tăng cường và căn chỉnh để rõ ràng.
- Trích xuất Thông tin (Extract Information):** Dữ liệu liên quan (tên, số ID, ngày sinh) được trích xuất bằng cách sử dụng **OCR hoặc quét mã vạch**.
- Xác thực Dữ liệu (Validate Data).**

6. Quyết định (Dữ liệu hợp lệ?): Nếu hợp lệ, quá trình tiếp tục; nếu không, một lần quét lại hoặc xử lý lỗi được kích hoạt.

7. Lưu trữ Dữ liệu (Store Data): Dữ liệu hợp lệ được lưu trữ trong hệ thống hoặc cơ sở dữ liệu.

8. Kết thúc (End).

3.1.3: Kiểm thử test Case:

D	Tính năng	Mô tả Test Case	Dữ liệu kiểm thử	Kết quả mong đợi	Người kiểm thử	Ngày	Kết quả
1	Tải ảnh từ file local	Kiểm tra tải ảnh từ local trong demo Gradio	Ảnh từ các định dạng khác nhau (png, jpg)	Tải lên và hiển thị ảnh thành công	Lê Văn Đức	16/10/2025	Đạt
2	Trích xuất thông tin từ ảnh đã xử lý	Kiểm tra kết quả sau khi trích xuất từ ảnh đã xử lý	CMND/CCC D thật để trích xuất	Thông tin sẽ được trích xuất và trả về dạng key-value	Lê Văn Đức	16/10/2025	Đạt
3	Phát hiện góc	Xác minh độ chính xác phát hiện góc trên ảnh	Ảnh CMND/CCC D chất lượng	Cả 4 góc phải được xác định	Tô Quang Huy	22/10/2025	Không đạt

		CMND/CCC D rõ nét	cao với các góc rõ ràng	chính xác với độ tin cậy >95 %			
4	Phát hiện góc	Kiểm tra phát hiện góc trên ảnh CMND/CCC D xoay	Ảnh CMND/CCC D được xoay ở các góc độ khác nhau (45°, 90°, 180°)	Cả 4 góc phải được xác định chính xác bất kể góc xoay	Tô Quang Huy	22/10/202 5	Khôn g đạt
3	Phát hiện góc	Xác minh độ chính xác phát hiện góc trên ảnh CMND/CCC D rõ nét	Ảnh CMND/CCC D chất lượng cao với các góc rõ ràng	Cả 4 góc phải được xác định chính xác với độ tin cậy >95 %	Tô Quang Huy	24/10/202 5	Đạt
4	Phát hiện góc	Kiểm tra phát hiện góc trên ảnh CMND/CCC D xoay	Ảnh CMND/CCC D được xoay ở các góc độ khác nhau (45°, 90°, 180°)	Cả 4 góc phải được xác định chính xác bất kể góc xoay	Lê Văn Đức	24/10/202 5	Đạt

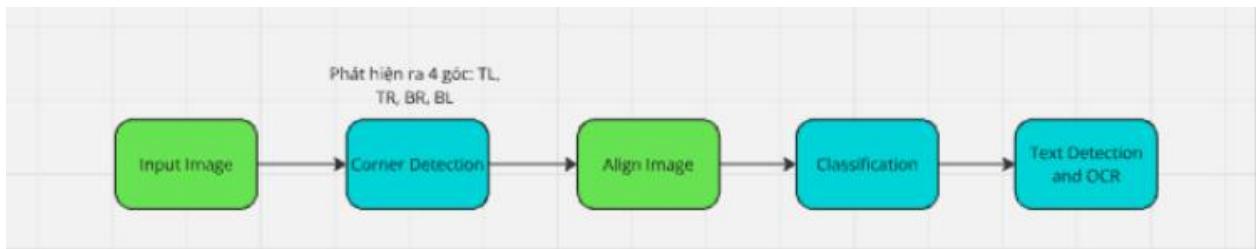
5	Phát hiện góc	Kiểm tra phát hiện góc trên CMND/CCC D bị che khuất một phần	CMND/CCC D có một góc bị che khuất một phần	Hệ thống phải phát hiện các góc nhìn thấy được và ước lượng vị trí góc bị che khuất	Tô Quang Huy	2/11/2025	Đạt
6	Phân loại	Kiểm tra độ chính xác phân loại CCCD gắn chip	Nhiều CCCD gắn chip	Hệ thống xác định chính xác là "CCCD gắn chip" với độ tin cậy >95 %	Phạm Thành Nam	17/10/2025	Đạt
7	Phân loại	Kiểm tra độ chính xác phân loại CCCD không gắn chip	Nhiều CCCD không gắn chip	Hệ thống xác định chính xác là "CCCD" với độ tin cậy >98 %	Nguyễn Hữu Sơn	17/03/2025	Đạt

3.1.4: Các thử nghiệm quan trọng bao gồm:

- **Tải lên hình ảnh:** Kiểm tra tải lên hình ảnh từ máy tính cục bộ trong bản demo Gradio (Đạt).
- **Trích xuất thông tin:** Kiểm tra kết quả sau khi trích xuất từ hình ảnh đã xử lý (Đạt).
- **Phát hiện góc:**
 - Kiểm tra độ chính xác trên hình ảnh rõ nét (Đã đạt >95% độ tin cậy).
 - Kiểm tra trên hình ảnh xoay (45° , 90° , 180° - Đạt, tất cả bốn góc phải được xác định chính xác bất kể độ xoay).
 - Kiểm tra trên thẻ ID bị che khuất một phần (Hệ thống nên phát hiện các góc có thể nhìn thấy và ước tính vị trí góc bị che khuất - Đạt).
- **Phân loại:** Kiểm tra độ chính xác phân loại cho CCCD có chip, không chip và thẻ ID kiểu cũ (Tất cả đều Đạt với độ tin cậy từ 90% trở lên).
- **OCR:**
 - Kiểm tra trích xuất số ID (Đạt 100% độ chính xác ký tự).
 - Kiểm tra trích xuất tên tiếng Việt có dấu phụ (diacritics - Đạt, trích xuất đầy đủ tên với tất cả các dấu phụ được giữ nguyên).
- **Quy trình đầu cuối (End-to-end):** Kiểm tra toàn bộ quy trình từ tải lên đến trích xuất thông tin trong vòng 5 giây (Đạt).

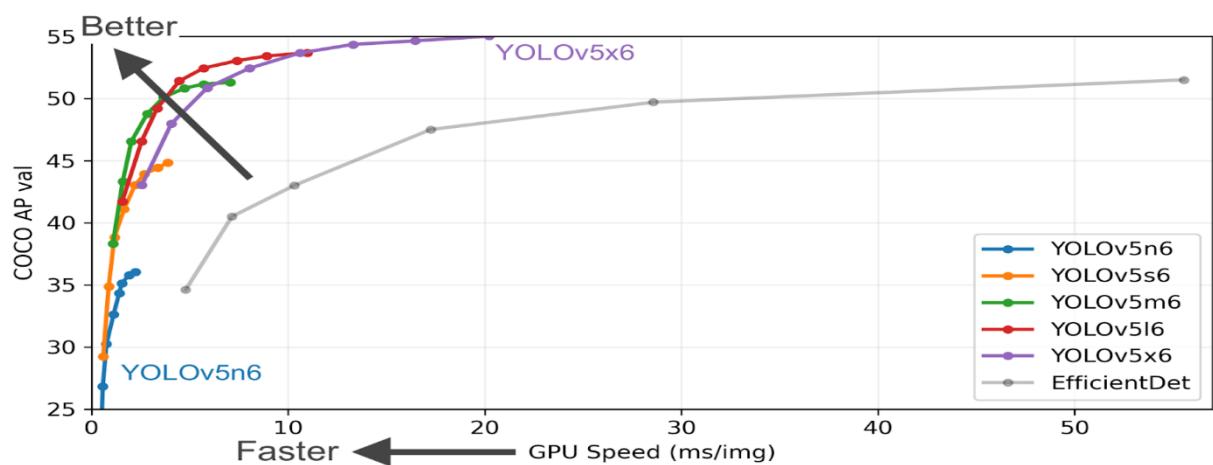
3.2. Phương pháp luận (Methodology)

Phương pháp đề xuất có ba thành phần chính: Phát hiện góc, phân loại thẻ ID và phát hiện văn bản OCR.



3.2.1 Phát hiện góc (Corner detection)

- Mô hình được chọn: YOLOv5.



- Ưu điểm của YOLOv5:

- Độ chính xác cao: YOLOv5 thể hiện hiệu suất mạnh mẽ về Độ chính xác Trung bình (AP) trên tập dữ liệu COCO, cho thấy hiệu quả của nó trong việc phát hiện chính xác các đối tượng, ngay cả trong các kịch bản khó khăn.
- Tốc độ xử lý thời gian thực: Đạt được thời gian suy luận (inference times) nhanh hơn đáng kể so với các mô hình khác, lý tưởng cho các ứng dụng yêu cầu kết quả ngay lập tức và trích xuất các tính năng chính xác như góc và khu vực văn bản trên thẻ ID.
- Mục tiêu: Xác định chính xác và trích xuất các tính năng như góc của thẻ, khu vực văn bản và các khu vực quan tâm khác giúp trích xuất thông tin.

3.2.2. Phân loại (Classification)

- Mô hình được chọn: MobileNetV2.
- Ưu điểm của MobileNetV2:

- **Hiệu quả về tốc độ và độ chính xác:** Nổi tiếng với kiến trúc nhẹ và hiệu quả cao về cả tính toán và độ chính xác.
- Sử dụng **phép chập sâu phân tách (depthwise separable convolutions)**, giúp giảm số lượng tham số và phép tính so với các lớp chập truyền thống, dẫn đến thời gian suy luận nhanh hơn.
- **Nhẹ và Phù hợp cho Thiết bị Di động:** Được thiết kế tối ưu cho hiệu suất với kích thước mô hình nhỏ hơn trong khi vẫn duy trì độ chính xác cao, làm cho nó lý tưởng để triển khai trong các môi trường hạn chế tài nguyên.

c. OCR

- **Mô hình được chọn: vietOCR.**

- **Ưu điểm của vietOCR:**

- **Độ chính xác cao cho văn bản tiếng Việt:** Được điều chỉnh chuyên biệt để xử lý sự phức tạp của ngôn ngữ Việt Nam.
- Nó thành thạo trong việc nhận dạng **dấu phụ tiếng Việt (diacritics)**, điều này rất quan trọng để giải thích từ ngữ chính xác, đảm bảo độ chính xác cao cho văn bản CCCD.
- **Khả năng thích ứng với nhiều phông chữ khác nhau:** Được xây dựng để thích ứng với các phông chữ khác nhau trên CCCD, đảm bảo văn bản được trích xuất đáng tin cậy.
- **Công cụ tiền xử lý và hậu xử lý:** Đi kèm với các công cụ tích hợp sẵn cho tiền xử lý (như nhị phân hóa và loại bỏ nhiễu) và hậu xử lý (như sửa lỗi chính tả), giúp cải thiện chất lượng trích xuất văn bản.

3.3 Tăng cường dữ liệu (Data Augmentation)

Phần này mô tả các kỹ thuật được sử dụng để tăng cường dữ liệu đào tạo nhằm cải thiện tính tổng quát hóa của mô hình, bao gồm:

- **Biến đổi hướng hình ảnh (Image Orientation Transformation):** Bao gồm lật ngang (Horizontal Flip), lật dọc (Vertical Flip), và xoay (Rotate).
- **Ưu điểm:** Giúp tăng kích thước dữ liệu đào tạo, cải thiện tính tổng quát hóa của mô hình và tính mạnh mẽ đối với các kịch bản thực tế.

◦ *Nhược điểm*: Không phù hợp với nội dung nhạy cảm với hướng (chẳng hạn như văn bản trên thẻ ID, vì văn bản được nhân đôi sẽ trở nên không đọc được).

• **Kỹ thuật làm mịn, làm mờ và nhiễu (Smoothing, Blurring, and Noise Techniques):**

Bao gồm Làm mờ (Blurring), Làm mịn (Smoothing), và Thêm nhiễu (Noise Addition - như nhiễu muối tiêu, nhiễu Gaussian).

◦ *Ưu điểm*: Giảm nhiễu tàn số cao, cải thiện tiền xử lý cho OCR/Phát hiện, và tăng tính mạnh mẽ của mô hình đối với các khiếm khuyết hình ảnh trong thế giới thực (ví dụ: hình ảnh di động, ánh sáng yếu).

◦ *Nhược điểm*: Lạm dụng có thể làm mất các chi tiết quan trọng (văn bản nhỏ, số ID), và hình ảnh bị mờ hoặc làm mịn nặng có thể dẫn đến độ chính xác OCR kém.

• **Biến đổi màu sắc và cường độ (Color and Intensity Transformations):** Bao gồm điều chỉnh độ sáng (Brightness Adjustment), độ tương phản (Contrast Adjustment), và độ bão hòa/sắc độ (Saturation and Hue Jitter).

◦ *Ưu điểm*: Phai bớt mô hình với các kịch bản thiếu/thừa sáng, cải thiện khả năng nhận dạng trong các điều kiện chiếu sáng đa dạng. Điều chỉnh độ tương phản giúp các nét văn bản nổi bật.

◦ *Nhược điểm*: Thay đổi độ sáng quá mức có thể làm văn bản bị trôi màu hoặc không đọc được. Tăng độ tương phản mạnh có thể khuếch đại nhiễu hoặc kết cấu nền.

4. Tiết độ dự án :

Đánh giá tổng thể dự án đạt **90%** hoàn thành.

Hầu hết các nhiệm vụ đều đã hoàn thành 100%:

- Thu thập và chuẩn bị dữ liệu.
- Chú thích dữ liệu (Data annotation).
- Nghiên cứu về mô hình và ứng dụng.
- Huấn luyện mô hình phát hiện đối tượng và cải thiện độ chính xác.
- Áp dụng mô hình OCR để trích xuất thông tin.
- Xây dựng API để tương tác với mô hình.

- Xây dựng trang web demo.

- Chuẩn bị báo cáo cuối kỳ.

5. Kết quả thực nghiệm :

5.1 Bộ dữ liệu (Dataset)

- **Phát hiện góc (Corner Detection):**

- Mục tiêu: Gán nhãn bốn góc (Trên-trái, Trên-phải, Dưới-trái, Dưới-phải) để cho phép hệ thống **chuẩn hóa phối cảnh** và kích thước của tài liệu (document rectification).

- Kích thước: 1380 mẫu huấn luyện (Train set), 346 mẫu xác thực (Val set) – **Tổng cộng 1726.**

- **Phân loại (Classification):**

- Mục tiêu: Cho phép hệ thống xử lý các loại tài liệu khác nhau một cách thích hợp và cải thiện độ chính xác của OCR. Các loại bao gồm: Thẻ ID Quốc gia kiểu cũ (CMND), CCCD có chip, Giấy phép lái xe, Hóa đơn tiền điện.

- Kích thước: 469 mẫu huấn luyện, 118 mẫu xác thực – **Tổng cộng 587.**

5.2. Hiệu suất và các chỉ số

- a. Kết quả Phát hiện Đối tượng sử dụng YOLOv5

- Mục tiêu: Phát hiện vùng chứa thẻ ID trong hình ảnh đầu vào.

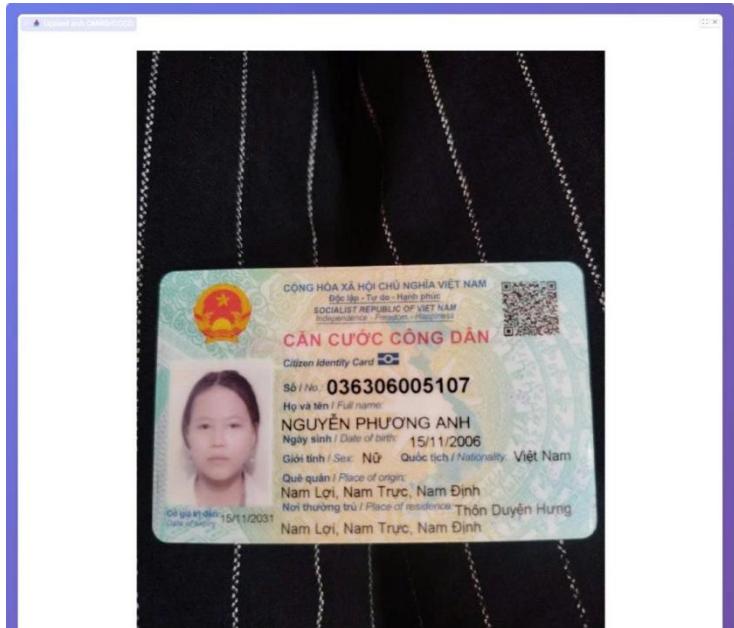
- Chỉ số hiệu suất:

- Độ chính xác (Precision): Khoảng 0.92 – 0.95.

- Hồi tưởng (Recall): Khoảng 0.90 – 0.93.

- mAP (mean Average Precision)@0.5: Trên 0.91.

- Nhận xét: Mô hình YOLOv5 hoạt động tốt trong việc phát hiện các vùng chứa thẻ ID, đảm bảo các mô-đun tiếp theo hoạt động trên các vùng quan tâm chất lượng cao.



Hình 1 : Ảnh đưa vào



Hình 2: Ảnh được căn chỉnh

Thông tin cơ bản Địa chỉ & Quốc tịch Thời hạn

SỐ CMND/CCCD:
ID number: 036306005107
036306005107

HỌ VÀ TÊN:
Name: NGUYỄN PHƯƠNG ANH
NGUYỄN PHƯƠNG ANH

NGÀY SINH:
Dob: 15/11/2006
15/11/2006

GIỚI TÍNH:
Gender: Nữ
Nữ

Hình 3: Trích xuất thông tin 1

Thông tin cơ bản Địa chỉ & Quốc tịch Thời hạn

QUỐC TỊCH:
Nationality: Việt Nam
Việt Nam

QUỐC TỊCH:
Place of origin: NAM LỢI, NAM TRỰC, NAM ĐỊNH
Nam Lợi, Nam Trực, Nam Định

NƠI THƯỜNG TRÚ:
Place of residence: NAM LỢI, NAM TRỰC, NAM ĐỊNH
NAM LỢI, NAM TRỰC, NAM ĐỊNH

Hình 4 : Trích xuất thông tin 2

Thông tin cơ bản Địa chỉ & Quốc tịch Thời hạn

NGÀY HẾT HẠN:
Expiry date: 15/11/2031
15/11/2031

Hình 5 : Trích xuất thông tin 3

5.3. Đánh giá Tổng thể

- **Tích hợp Hệ thống:** Hệ thống tích hợp bao gồm ba bước xử lý (phát hiện, phân loại và OCR) thể hiện tính khả thi và hiệu quả cho các ứng dụng thực tế.
- **Hiệu suất:**
 - Tốc độ xử lý trung bình trên mỗi hình ảnh là chấp nhận được cho các ứng dụng thời gian thực.

- Sự kết hợp giữa YOLOv5 và MobileNetV3 cung cấp sự cân bằng giữa tốc độ và độ chính xác.

- **Khả năng mở rộng:** Hệ thống có thể dễ dàng mở rộng để xử lý các loại tài liệu khác hoặc cập nhật với các bộ dữ liệu mới.

6. Kết Luận

6.1. Ưu điểm

- Mô hình AI thể hiện hiệu suất mạnh mẽ với độ chính xác cao trong việc trích xuất thông tin từ thẻ ID.

- Ứng dụng thân thiện với người dùng và dễ điều hướng, làm cho nó dễ tiếp cận đối với nhiều người dùng khác nhau.

6.2. Hạn chế

- Cần có không gian để cải thiện độ chính xác của mô hình phát hiện góc và phân loại.

- Việc tinh chỉnh VietOCR có thể nâng cao hơn nữa hiệu suất, nhưng dữ liệu hạn chế hiện có đặt ra một thách thức.

6.3. Cải tiến Tương lai

- Tích hợp **chatbot** có thể hợp lý hóa việc truy xuất dữ liệu và tăng cường tương tác người dùng.

- Mở rộng ứng dụng sang nền tảng di động có thể cung cấp khả năng tiếp cận và tiện lợi cao hơn cho người dùng.

7. Tài liệu tham khảo :

- [1] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., & Chintala, S. (2019). PyTorch: An imperative style, high-performance deep learning library. Retrieved from <https://arxiv.org/abs/1912.01703>
- [2] Grinberg, M. (2018). Flask web development: Developing web applications with Python. O'Reilly Media, Inc.
- [3] Abid, A., & Khuwaileh, A. (2019). Gradio: Hassle-free sharing and testing of ML models in the wild. Retrieved from <https://arxiv.org/abs/1906.02569>
- [4] Merkel, D. (2014). Docker: Lightweight Linux containers for consistent development and deployment. *Linux Journal*, 2014(239), 2.
- [5] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- [6] Jocher, G. (2020). YOLOv5 by Ultralytics (Version 7.0). Zenodo. <https://doi.org/10.5281/zenodo.3908559>
- [7] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). SSD: Single shot multi-box detector. *Proceedings of the European Conference on Computer Vision (ECCV)*, 21–37. https://doi.org/10.1007/978-3-319-46448-0_2
- [8] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778.
- [9] Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv:1409.1556*.
- [10] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4510–4520.
- [11] Zacharias, E., Teuchler, M., & Bernier, B. (2018). Image Processing Based Scene-Text Detection and Recognition with Tesseract. *Proceedings of the 2018 International Conference on Image Processing and Pattern Recognition (IPPR)*, 45–52.

[12] Fu, X., Ch'ng, E., Aickelin, U., & See, S. (2019). CRNN: A Joint Neural Network for Redundancy Detection. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 458–466.

[13] Quoc, P. B. C. (2023). VietOCR: A Python Framework for Vietnamese OCR. Retrieved from <https://github.com/pbcquoc/vietocr>.

Thành viên	