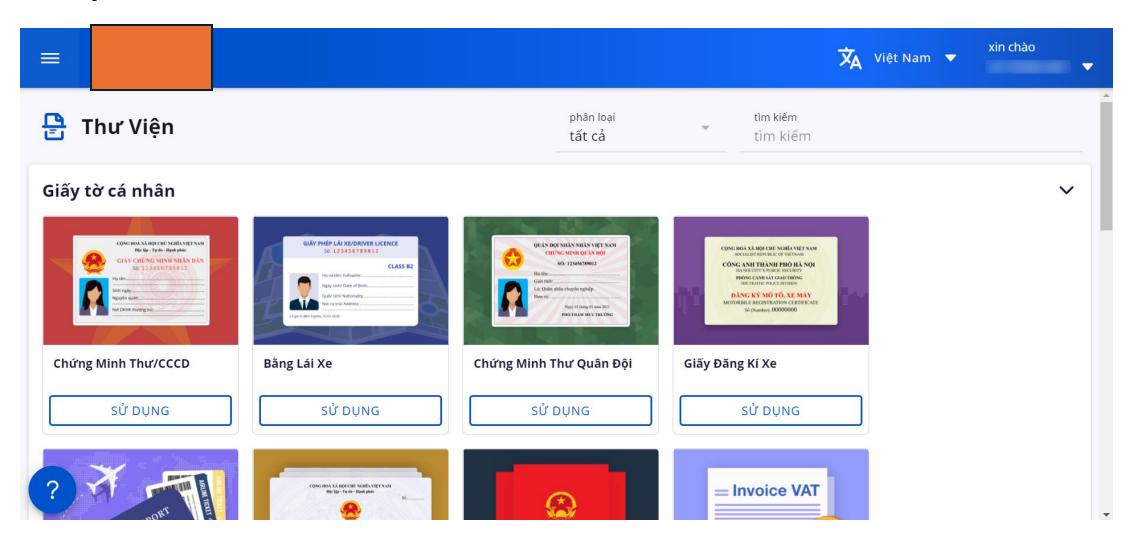
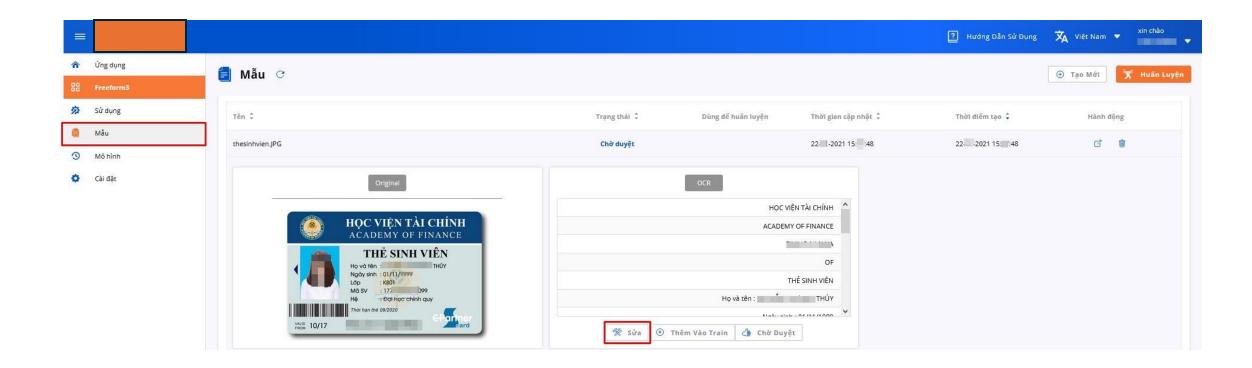


# OCR sổ đỏ bằng Vision Language Model

# Inspiration



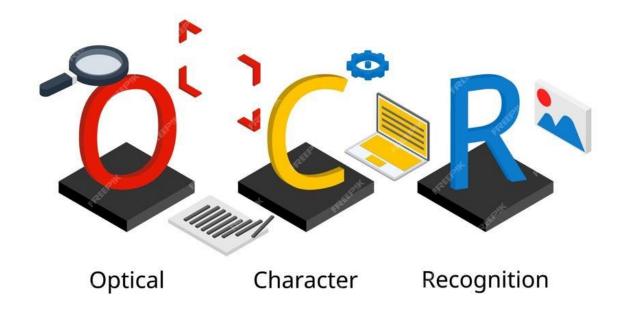
# Inspiration



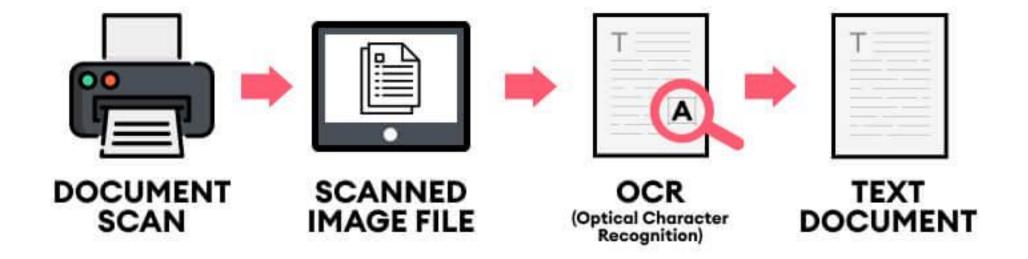
## Inspiration



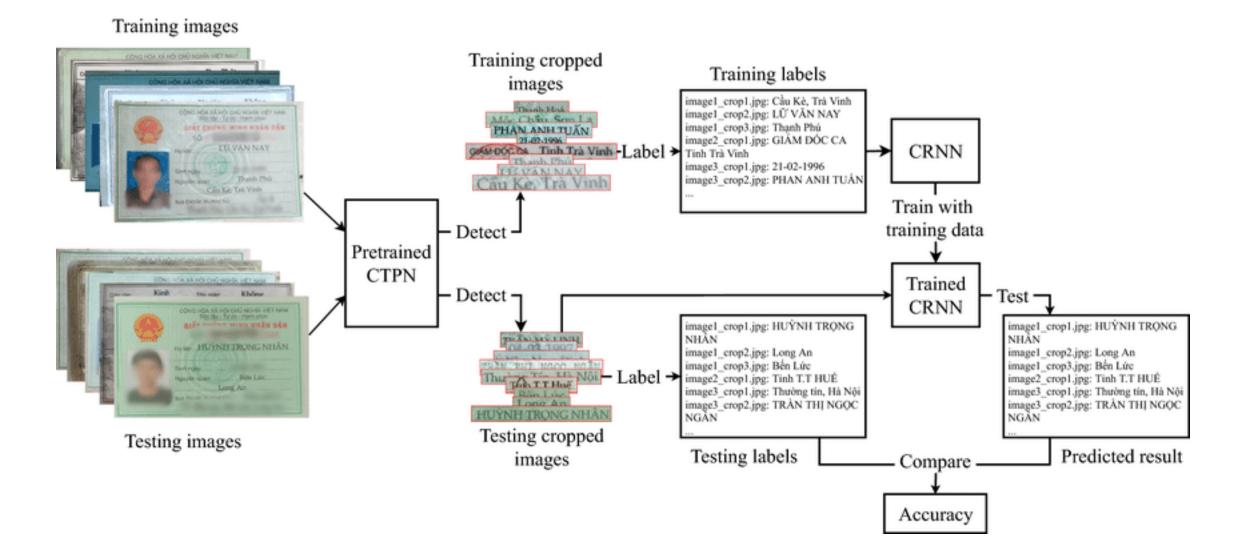
## What is OCR?



## What is OCR?



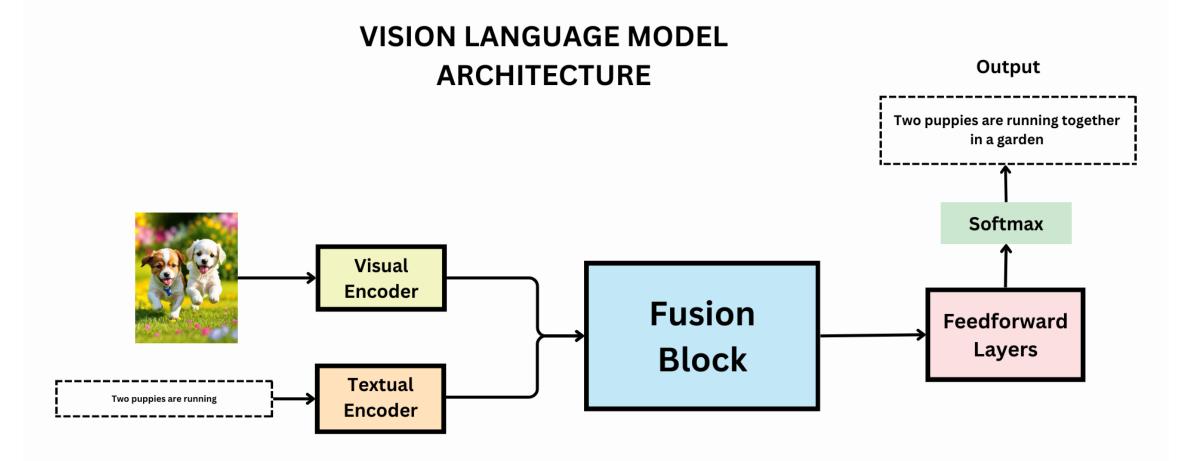
## OCR in the past



## OCR in the past



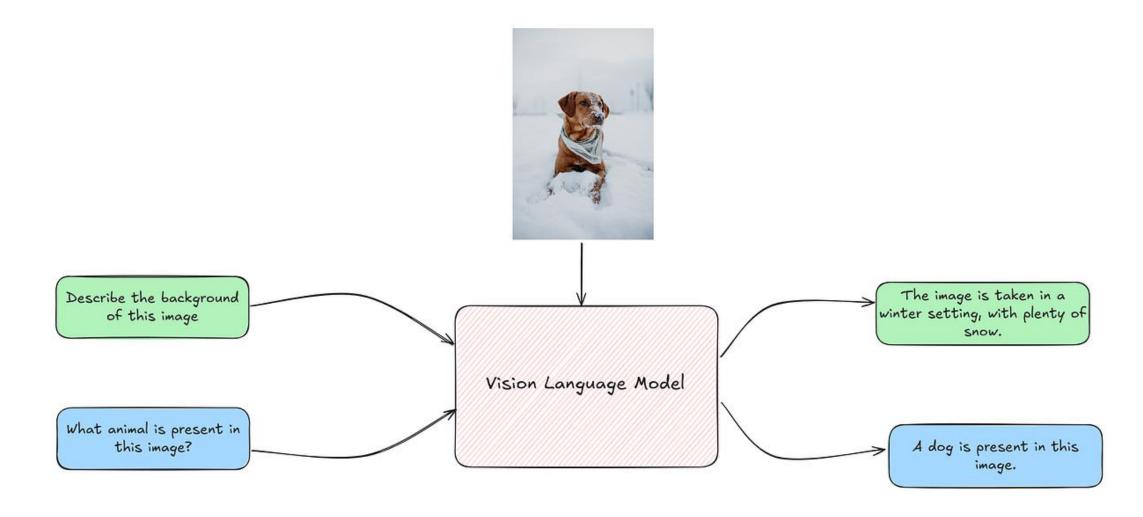
## VLM has arrived!



### VLM has arrived!

# Vision Language Model A typical architecture Image input Text input Image Tokenizer encoder Embedding projector Concatenated Tokens Text Decoder Text output

# Example



# Compare to traditional OCR?

Open question, please comment your opinion below

- Cost
- Speed
- Accurracy
- Hallucination

### Handson

- Phần cứng: VPS có GPU A4000 thuê tại Thuegpu.vn
- Phần mềm:
  - vLLM for serving
  - Model is: <u>5CD-AI/Vintern-1B-</u>v3 5
- Bài toán:
  - OCR sổ đỏ :D
  - Define template by user



### CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM Độc lập - Tự do - Hạnh phúc



### GIẤY CHỨNG NHẬN

### QUYỀN SỬ DỤNG ĐẤT, QUYỀN SỞ HỮU TÀI SẢN GẦN LIỀN VỚI ĐẤT

1. Người sử dụng đất và chủ sở hữu tài sản gắn liền với đất:

Ông: Nguyễn Văn A, số CCCD: 123456789012

Bà: Nguyễn Thị B, số CCCD: 123456789012

- 2. Thông tin thửa đất:
- a. Thừa đất số: 0000; tờ bản đồ số: 01; Diện tích: 10.012 m2 (Mười nghìn không trăm mười hai mét vuông)
- b. Loại đất: Đất ở tại nông thôn 400 m2; Đất trồng cây lâu năm 9.112 m2; Đất nuôi trồng thủy sản 500 m2.
- c. Thời hạn sử dụng đất: Đất ở tại nông thôn: lâu dài; Đất trồng cây lâu năm: đến ngày 15/10/2043; Đất nuôi trồng thủy sản: đến ngày 15/10/2043.
- d. Hình thức sử dụng đất: Sử dụng riêng.
- đ. Nguồn gốc sử dụng đất: Đất ở tại nông thôn: Nhà nước công nhận quyền sử dụng đất như giao đất có thu tiền sử dụng đất; Đất trồng cây lâu năm, đất nuôi trồng thủy sản: Nhà nước công nhận quyền sử dụng đất như giao đất không thu tiền sử dụng đất.
- e. Địa chỉ: Thôn Hà Xá, xã Đại Hưng, huyện Mỹ Đức, Thành phố Hà Nội
- 3. Thông tin tài sản gắn liền với đất:

Địa chỉ: Thôn Hà Xá, xã Đại Hưng, huyện Mỹ Đức, Thành phố Hà Nội

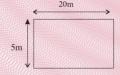
Tên tài sản/Hạng mục công trình	Diện tích xây dựng (m2)	Diện tích sàn hoặc công suất	Kết cấu chủ yếu	Cấp công trình	Số tằng	Năm HT xây dựng	Thời hạn sở hữu
Nhà ở	150		Bê tông, cốt thếp	4	2	2023	////
			ا فرام روام ا				
			A A		E VI	0.5 (3) [3]	(Y/45)

#### 4. Ghi chú:

Thừa đất có 150 m2 đất trồng cây lâu năm thuộc quy hoạch mở rộng Tinh lộ 419

Mỹ Đức, ngày tháng năm 2025 VĂN PHÔNG ĐĂNG KÝ ĐÁT ĐAI GIÁM ĐỐC

#### 5. Sơ đồ thừa đất:



AA 123456

Người được cấp Giấy chứng nhận không được sửa chữa, tấy xóa hoặc bồ xung bất kỳ nội dung nào trong Giấy chứng nhận; khi bị mất hoặc hư hóng Giấy chứng nhận phải khai báo ngay với cơ quan cấp Giấy