

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC CẦN THƠ**



**BÁO CÁO TỔNG KẾT
ĐỀ TÀI NGHIÊN CỨU KHOA HỌC CỦA SINH VIÊN**

**ỨNG DỤNG HỖ TRỢ HỌC TIẾNG ANH QUA VIỆC
NHẬN DẠNG ĐÓI TUỢNG TRONG VIDEO**

Mã số: THS2022-15

Cần Thơ, 11/2022

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC CẦN THƠ**



**BÁO CÁO TỔNG KẾT
ĐỀ TÀI NGHIÊN CỨU KHOA HỌC CỦA SINH VIÊN**

**ỨNG DỤNG HỖ TRỢ HỌC TIẾNG ANH QUA VIỆC
NHẬN DẠNG ĐÓI TUỢNG TRONG VIDEO**

Mã số: THS2022-15

Sinh viên thực hiện: Nguyễn Chí Bảo

Nam/Nữ: Nam

Dân tộc: Kinh

Lớp: DI19V7F2, Trường: CNTT-TT Năm thứ: 4 /Số năm đào tạo: 4.5

Ngành học: Công nghệ thông tin (Chất lượng cao)

Người hướng dẫn: TS. Nguyễn Thanh Hải

Cần Thơ, 11/2022

DANH SÁCH THÀNH VIÊN THAM GIA

TT	Họ và tên	Vai trò	MSSV/Đơn vị
1	Nguyễn Chí Bảo	Chủ nhiệm đề tài	MSSV: B1910619 Lớp: DI19V7F2 (CNTT Chất lượng cao) Khóa: K45
2	Lê Nguyễn Bảo Anh	Thư ký khoa học	MSSV: B1910611 Lớp: DI19V7F2 (CNTT Chất lượng cao) Khóa: K45
3	Lê Huỳnh Như	Thành viên chính	MSSV: B1910680 Lớp: DI19V7F2 (CNTT Chất lượng cao) Khóa: K45
4	Danh Bé Hai	Thành viên chính	MSSV: B1909907 Lớp: DI19V7F2 (CNTT Chất lượng cao) Khóa: K45
5	Đặng Quốc Cường	Thành viên chính	MSSV: B1910623 Lớp: DI19V7F2 (CNTT Chất lượng cao) Khóa: K45

CÁN BỘ HƯỚNG DẪN SINH VIÊN THỰC HIỆN ĐỀ TÀI

Họ và tên, MSCB	Đơn vị công tác và lĩnh vực chuyên môn	Nhiệm vụ
TS. Nguyễn Thanh Hải MSCB: 2267	Đơn vị công tác: Trường CNTT & TT Lĩnh vực chuyên môn: Phân tích hệ thống, Lập trình web, máy học, học sâu, v.v	Hướng dẫn nội dung khoa học và hướng dẫn lập dự toán kinh phí đề tài

ĐƠN VỊ PHỐI HỢP CHÍNH

- ✓ Trường Công nghệ thông tin & Truyền thông – Đại học Cần Thơ
- ✓ Phòng Quản lý Khoa học – Đại học Cần Thơ

MỤC LỤC

DANH SÁCH THÀNH VIÊN THAM GIA	I
CÁN BỘ HƯỚNG DẪN SINH VIÊN THỰC HIỆN ĐỀ TÀI.....	I
ĐƠN VỊ PHỐI HỢP CHÍNH	II
MỤC LỤC	III
DANH MỤC HÌNH.....	V
DANH MỤC BẢNG	VI
DANH MỤC CÁC TỪ VIẾT TẮT	VII
TÓM TẮT	IX
ABSTRACT	X
THÔNG TIN KẾT QUẢ NGHIÊN CỨU	XI
PHẦN 1: MỞ ĐẦU	1
I. TỔNG QUAN TÌNH HÌNH NGHIÊN CỨU TRONG VÀ NGOÀI NƯỚC.....	1
1. Trong nước	1
2. Ngoài nước	2
II. SỰ CẦN THIẾT CỦA ĐỀ TÀI.....	3
III. MỤC TIÊU CỦA ĐỀ TÀI	3
IV. NỘI DUNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU	4
1. Nội dung nghiên cứu	4
2. Phương pháp nghiên cứu	4
V. ĐỐI TƯỢNG VÀ PHẠM VI NGHIÊN CỨU	4
1. Đối tượng nghiên cứu	4
2. Phạm vi nghiên cứu	4
PHẦN 2: NỘI DUNG NGHIÊN CỨU.....	5
I. CƠ SỞ LÝ THUYẾT	5
1. Ngôn ngữ lập trình Python.....	5
2. Java	5
3. Tensorflow FrameWork.....	6
4. Tensorflow Lite FrameWork	6
5. Android Studio	6
6. Deep Learning	7
7. SQLite	9
II. Thu thập dữ liệu ảnh và phương pháp tăng dữ liệu ảnh (data augmentation)	9
1. Thu thập dữ liệu:	9
2. Phương pháp tăng dữ liệu ảnh(data augmentation):.....	14
II. PHƯƠNG PHÁP NGHIÊN CỨU	15
1. Mô hình tổng quan:	15
1.1. Cấu trúc tổng quan của model:	15
1.2. Workflow	18
2. Xử lý hình ảnh:	19
3. Object Detection:	20

4. Hệ thống lưu trữ từ điển.....	20
4.1. Lưu trữ trên SQLite.....	20
4.2. Lưu trữ trên GitHub.....	21
PHẦN 3: THỰC NGHIỆM.....	22
I. KẾT QUẢ ĐÁNH GIÁ MÔ HÌNH:.....	22
1. Phân chia training, validation and testing set:	22
2. Kết quả sau khi huấn luyện:.....	22
II. So sánh kết quả các cách tiếp cận:.....	28
1. Mô hình yolov5s và yolov5m	29
2. Mô hình sử dụng tăng cường dữ liệu và không sử dụng tăng cường dữ liệu	31
3. Tăng số lớp và số ảnh cho mô hình đã đào tạo	36
III. Ứng dụng Android:.....	40
1. Giao diện màn hình chính	40
2. Giao diện nhận dạng trực tiếp (Real-time)	42
3. Giao diện nhận dạng bằng hình ảnh chụp từ camera hoặc hình ảnh có sẵn trong máy	43
4. Giao diện từ điển	47
PHẦN 4: KẾT LUẬN VÀ KIẾN NGHỊ	49
I. KẾT LUẬN	49
II. KẾT QUẢ ĐÓNG GÓP	49
III. HƯỚNG PHÁT TRIỂN	50
TÀI LIỆU THAM KHẢO.....	51
PHẦN 5: PHỤ LỤC	52
I. Cài đặt Python trên Windows	52
II. Cài đặt Java trên Windows	55
III. Cài đặt Android Studio trên Windows.....	58
IV. Cài đặt DB Browser for SQLite trên máy Windows.....	64
V. Cài đặt yolov5 và huấn luyện mô hình	67

DANH MỤC HÌNH

Hình 1: Cơ chế hoạt động của Deep Learning.....	23
Hình 2: Hình ảnh trước và sau khi sử dụng phương pháp tăng cường dữ liệu.....	29
Hình 3: Kiến trúc YOLOv5.....	30
Hình 4: Mô-đun Focus.....	31
Hình 5: BottleneckCSP.....	32
Hình 6: Mô-đun SPP	32
Hình 7: Sơ đồ chi tiết cách thức xây dựng và đào tạo mô hình.....	33
Hình 8: Gắn nhãn hình ảnh trên Roboflow.....	35
Hình 9: Các phiên bản của YOLOv5:.....	43
Hình 10: Biểu đồ so sánh chỉ số mAP giữa 2 mô hình yolov5s và yolov5m.....	44
Hình 11: Biểu đồ cột thể hiện sự so sánh về độ chính xác giữa 2 mô hình yolov5s và yolov5m	45
Hình 12: Confusion matrix của mô hình không sử dụng tăng cường dữ liệu.....	46
Hình 13: Confusion matrix của mô hình sử dụng tăng cường dữ liệu.....	47
Hình 14: Hình ảnh nhận dạng được trước và sau khi sử dụng tăng cường dữ liệu.....	48
Hình 15: Biểu đồ cột so sánh độ chính xác của hai mô hình	48
Hình 16: Biểu đồ cột so sánh Precision và Recall của hai mô hình.....	49
Hình 17: Biểu đồ cột so sánh F1-score của hai mô hình	49
Hình 18: Biểu đồ cột so sánh mAP_0.5 và mAP_0.5:0.95 của hai mô hình	50
Hình 19: Confusion matrix của mô hình ban đầu	51
Hình 20: Confusion matrix của mô hình sau khi thêm lớp mới	52
Hình 21: Biểu đồ so sánh chỉ số mAP giữa mô hình ban đầu và mô hình sau khi thêm lớp mới	53
Hình 22: Biểu đồ so sánh kết quả đạt được sau khi mô hình được cải thiện.....	54
Hình 24: Giao diện Splash Screen	55
Hình 23: Giao diện chính của ứng dụng	55
Hình 25: Giao diện nhận dạng trực tiếp (real-time).....	56
Hình 26: Giao diện chức năng “Detect by image”	57
Hình 27: Chọn ảnh từ thư viện	58
Hình 28: Tiến hành nhận dạng.....	60
Hình 29: Đã nhận dạng thành công.....	61
Hình 30: Giao diện thư viện từ vựng	62
Hình 31: Chức năng tìm kiếm.....	63
Hình 32: Tùy chỉnh các cài đặt Python.....	67
Hình 33: Các tính năng tùy chỉnh	68
Hình 34: Chọn nơi cài đặt	68
Hình 35: Python đang được cài đặt.....	69
Hình 36: Cài đặt thành công	69
Hình 37: Trang tải JDK phiên bản mới nhất hiện tại.....	70
Hình 38: Các phiên bản JDK cho các hệ điều hành.....	70

Hình 39: Tiến hành cài đặt JDK	71
Hình 40: Chọn nơi lưu trữ.....	72
Hình 41: Đang tiến hành cài đặt	72
Hình 42: Cài đặt JDK thành công.....	73
Hình 43: Trang tải về bản cài đặt Android Studio	73
Hình 44: Cửa sổ cài đặt Android Studio	74
Hình 45: Cửa sổ lựa chọn các Components	74
Hình 46: Chọn nơi lưu trữ.....	75
Hình 47: Cửa sổ chọn Start Menu Folder	75
Hình 48: Hoàn tất việc cài đặt Android Studio.....	76
Hình 49: Cài đặt thành công	76
Hình 50: Cửa sổ thông báo Missing SDK	77
Hình 51: Cửa sổ cài đặt Components SDK	78
Hình 52: Các thông số của chương trình	78
Hình 53: Đang tiến tải và cài đặt các tệp tài nguyên	79
Hình 54: Hoàn tất việc cài đặt	79
Hình 55: Trang tải về bản cài đặt DB Browser for SQLite	80
Hình 56: Cửa sổ cài đặt DB Browser for SQLite	81
Hình 57: Cửa sổ điều khoản của DB Browser for SQLite	81
Hình 58: Lựa chọn nơi lưu trữ	82
Hình 59: Bắt đầu cài đặt.....	82
Hình 60: Cài đặt thành công	83

DANH MỤC BẢNG

Bảng 1: Mô tả dữ liệu đã thu thập.....	23
Bảng 2: Kết quả sau khi huấn luyện	35
Bảng 3: Bảng so sánh thời gian phát hiện vật thể qua các phiên bản của yolov5	42

DANH MỤC CÁC TỪ VIẾT TẮT

STT	Thuật ngữ	Điễn giải	Nghĩa Tiếng Việt
1	AI	Artificial Intelligence	Trí tuệ nhân tạo
2	API	Application Programming Interface	Giao diện lập trình ứng dụng
3	App	Application	Ứng dụng
4	ATM	Automated Teller Machine	Máy rút tiền
5	APK	Android Package Kit	Bộ gói Android
6	BN	Batch Normalization	Chuẩn hóa hàng loạt
7	CNTT	Công Nghệ Thông Tin	Công Nghệ Thông Tin
8	CT	Công thức	Công thức
9	Conv2d	Lớp tích chập 2d	Lớp tích chập 2d
10	DB	Database	Cơ sở dữ liệu
11	eKYC	Electronic Know Your Customer	Điện tử biết khách hàng của bạn
12	FPN	Feature Pyramid Networks	Mạng kim tự tháp tính năng
13	JDK	Java Development Kit	Bộ phát triển Java
15	IDE	Integrated Development Environment	Môi trường phát triển tích hợp
16	MS-DOS	Microsoft Disk Operating System	Hệ điều hành đĩa Microsoft
17	OOP	Object Oriented Programming	Lập trình hướng đối tượng
18	SPP	Spatial Pyramid Pooling	Tổng hợp kim tự tháp không gian

STT	Thuật ngữ	Điễn giải	Nghĩa Tiếng Việt
19	SDK	Software Development Kit	Bộ phát triển phần mềm
20	YOLO	You Only Look Once	Bạn Chỉ Nhìn Một Lần

TÓM TẮT

Tiếng Anh là ngôn ngữ phổ biến nhất trên thế giới, và nó ngày càng quan trọng. Tiếng Anh đã được biên soạn trong hầu hết các tài liệu, thông tin và nội dung trên internet. Tuy nhiên, với lượng từ vựng khổng lồ, việc học tiếng Anh đối với nhiều người rất khó nhọc. Chính vì vậy, nhiều công nghệ hiện đại đã ra đời để hỗ trợ học tiếng Anh như công nghệ học tiếng Anh qua trò chơi ghép chữ, tìm đồ vật, giải đố,... đã giúp trẻ hứng thú và dễ dàng tiếp cận với tiếng Anh ngay từ nhỏ. Nghiên cứu này trình bày một phương pháp hỗ trợ việc học tiếng Anh thông qua phát hiện đối tượng trong video, hình ảnh hoặc thậm chí video phát trực tiếp trong thời gian thực bằng cách sử dụng các kiến trúc học sâu chẩn đoán như You Look Only Once (YOLO) - một trong những mô hình phát hiện đối tượng vô cùng hiện đại và tốt nhất hiện nay. Phương pháp để có được mAP (Độ chính xác trung bình trung bình) là 55,6 với 17GFlops. Kết quả đạt được là một ứng dụng gồm bộ từ vựng, ý nghĩa và đặt câu với từ vừa tìm kiếm được qua khả năng nhận dạng và phát hiện vật thể và các đối tượng mà đề tài dùng để nhận dạng trong nghiên cứu lần này là các vật thể xuất hiện phổ biến ở thành phố Cần Thơ.

Để làm được điều này, phương pháp của mà đề tài chọn là chọn lọc mô hình phù hợp nhất và có độ chính xác tốt với mô hình được huấn luyện từ 7303 hình ảnh thuộc 145 lớp khác nhau với trung bình 50 ảnh cho mỗi đối tượng sử dụng ứng dụng trên thiết bị với thời gian phát hiện vật thể là phù hợp.

ABSTRACT

English is the most common language in the world, and it is increasingly important. English has been compiled in most documents, information and content on the internet. However, with a huge vocabulary, learning English for many people is difficult to remember. Therefore, many modern technologies have been born to support learning English such as technology to learn English through word games, finding objects, puzzles, etc. with English from an early age. This study presents a method to support English learning through object detection in video, images or even live streaming video in real time using deep learning architectures such as You Look Only Once (YOLO) - one of the best object groups of detection models with modern performances. The method to get mAP (Mean Mean Accuracy) is 55.6 with 17GFlops. The result is an application that includes vocabulary, meanings and sentences with the search word through object recognition and detection, and the objects that we used to identify in this study are the output, popular objects in Can Tho city.

To do this, the method chosen by the topic is to select the most suitable and accurate model with the model trained from 7303 images belonging to 145 different classes with an average of 50 images for each subject. object using the application on the device with the object detection time is appropriate.

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC CẦN THƠ**

THÔNG TIN KẾT QUẢ NGHIÊN CỨU

1. Thông tin chung:

- Tên đề tài: Ứng dụng hỗ trợ học tiếng Anh qua việc nhận dạng đối tượng trong video.
- Sinh viên thực hiện:

1. Nguyễn Chí Bảo	MSSV: B1910619
2. Lê Nguyễn Bảo Anh	MSSV: B1910611
3. Lê Huỳnh Như	MSSV: B1910680
4. Danh Bé Hai	MSSV: B1909907
5. Đặng Quốc Cường	MSSV: B1910623
- Lớp: DI19V7F2
- Chuyên ngành: Công nghệ thông tin - Chất lượng cao
- Trường Công nghệ Thông tin và Truyền thông.
- Sinh viên năm thứ/Số năm đào tạo: 4 năm/4.5 năm
- Người hướng dẫn: TS. Nguyễn Thanh Hải

2. Mục tiêu đề tài:

Mục tiêu của đề tài nghiên cứu nhằm tạo ra một chương trình ứng dụng học Tiếng Anh trên App Android. Ứng dụng hỗ trợ phát hiện, nhận dạng các vật thể trong video bằng tiếng anh, làm tiền đề giúp người sử dụng học tiếng anh một thoải mái, giải trí mà vẫn trau dồi được nhiều kiến thức từ vựng với chức năng phát hiện đối tượng, định nghĩa đối tượng, phát âm và đặt câu với đối tượng vừa được nhận dạng.

3. Kết quả nghiên cứu:

- Mô hình máy học và học sâu giúp nhận dạng đối tượng
- Ứng dụng Android phát hiện đối tượng qua camera.
- Bộ dữ liệu hình ảnh đã được xử lý
- Bộ từ vựng tiếng anh phong phú, đa dạng

4. Tác động và lợi ích mang lại của kết quả nghiên cứu:

Đối với lĩnh vực giáo dục và đào tạo:

- Phát triển phạm vi nghiên cứu về nhận dạng vật thể thông qua hình ảnh, video.
- Phục vụ trong việc học Tiếng Anh của người dùng.
- Trong quá trình nghiên cứu cũng giúp ích cho sinh viên tham gia cơ hội để rèn luyện Tiếng Anh, cách ứng dụng CNTT vào một bài toán thực tế trong việc học tập.

Giúp sinh viên nâng cao tinh thần tự học, tự nghiên cứu và sáng tạo. Góp phần hỗ trợ sinh viên nghiên cứu sinh trong lĩnh vực công nghệ máy học, phân tích và thiết kế dữ liệu.

Đối với lĩnh vực khoa học và công nghệ có liên quan

- Góp phần ứng dụng công nghệ thông tin vào giáo dục. Tăng thêm phần hứng thú và khả năng ghi nhớ khi học tiếng anh cho trẻ em.
- Góp phần hỗ trợ các nhà nghiên cứu về máy học, học sâu cải thiện và ứng dụng các giải thuật. Từ đó đưa vào các mô hình kiểm thử và thực nghiệm phục vụ cho các báo cáo khoa học hay ứng dụng thực tiễn.

Đối với phát triển kinh tế-xã hội

- Đẩy mạnh phát triển công nghệ AI trong lĩnh vực giáo dục đặc biệt là giáo dục cho trẻ nhỏ ở thời đại 4.0.
- Phát triển ứng dụng có thể mở rộng để sử dụng rộng rãi trong các lĩnh vực giáo dục khác cũng như nâng cao trình độ ngoại ngữ trong thời đại kinh tế hội nhập hiện nay.
- Giảm một phần chi phí thuê giảng dạy cũng như giúp việc học ngoại ngữ trở nên gần gũi và dễ tiếp cận hơn đối với gia đình không có điều kiện.

Đối với tổ chức chủ trì và các cơ sở ứng dụng kết quả nghiên cứu

- Cung cấp một ứng dụng học từ vựng tiếng Anh tiện lợi, gần gũi.
- Hỗ trợ học tập về nhiều mặt từ đó tiết kiệm được thời gian mà vẫn đem lại hiệu quả tốt cho người dùng.

6. Công bố khoa học của sinh viên từ kết quả nghiên cứu của đề tài, hoặc nhận xét, đánh giá của cơ sở đã áp dụng các kết quả nghiên cứu (nếu có): Không

Ngày tháng năm
Sinh viên chịu trách nhiệm chính
thực hiện đề tài
(ký, họ và tên)

Nhận xét của người hướng dẫn về những đóng góp khoa học của sinh viên thực hiện đề tài (phản này do người hướng dẫn ghi):

Xác nhận của Trường Đại học Cần Thơ
(ký tên và đóng dấu)

Ngày tháng năm
Người hướng dẫn
(ký, họ và tên)

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC CẦN THƠ

**THÔNG TIN VỀ SINH VIÊN
CHỊU TRÁCH NHIỆM CHÍNH THỰC HIỆN ĐỀ TÀI**

I. SO LƯỢC VỀ SINH VIÊN:

Họ và tên: Nguyễn Chí Bảo

Sinh ngày : 28 tháng 05 năm 2001

Nơi sinh: Cần Thơ

Lớp: DI19V7F2

Khóa: K45

Ảnh 4x6

Trường: Công nghệ Thông tin và Truyền thông

Địa chỉ liên hệ: 172/29 Trần Quang Diệu, An Thới, Bình Thủy, Cần Thơ

Điện thoại: 0353530086 Email: baob1910619@student.ctu.edu.vn

II. QUÁ TRÌNH HỌC TẬP (kê khai thành tích của sinh viên từ năm thứ 1 đến năm đang học):

* **Năm thứ 1:**

Ngành học: Công nghệ thông tin Trường: Công nghệ Thông tin và Truyền thông

Kết quả xếp loại học tập: 3.75 - Xuất sắc

Sơ lược thành tích:

Stt	Thành tích tiêu biểu	Ngày cấp	Lý do
1	Tham gia hoạt động vệ sinh môi trường Đoàn khoa	09-10-2019	Lao động HK1 2019-2020
2	Tham gia hoạt động vệ sinh môi trường Đoàn khoa	13-12-2019	Lao động HK1 2019-2020
3	Tham gia hoạt động vệ sinh môi trường Đoàn khoa	21-06-2020	Lao động HK2 2019-2020

* **Năm thứ 2:**

Ngành học: Công nghệ thông tin Trường: Công nghệ Thông tin và Truyền thông

Kết quả xếp loại học tập: 2.65 - Khá

Sơ lược thành tích:

Stt	Thành tích tiêu biểu	Ngày cấp	Lý do
1	Tham gia hoạt động văn hoá văn nghệ	12-11-2020	Văn nghệ chào đón Tân Sinh viên K46

*** Năm thứ 3:**

Ngành học: Công nghệ thông tin Trường: Công nghệ Thông tin và Truyền thông
Kết quả xếp loại học tập: 2.68 - Khá

Xác nhận của Trường Đại học Cần Thơ
(ký tên và đóng dấu)

Ngày tháng năm
Sinh viên chịu trách nhiệm chính
thực hiện đề tài
(ký, họ và tên)

PHẦN 1: MỞ ĐẦU

I. TỔNG QUAN TÌNH HÌNH NGHIÊN CỨU TRONG VÀ NGOÀI NƯỚC

1. Trong nước

Qua những tìm hiểu về các dự án liên quan đến công nghệ máy học ở Việt Nam, một số ứng dụng dựa trên lĩnh vực trí tuệ nhân tạo (AI) được nhiều nghiên cứu biết đến và sử dụng tạo nên nhiều ứng dụng có ích, giúp con người giải quyết mọi vấn đề trong cuộc sống. Trí tuệ nhân tạo được ứng dụng ở nhiều lĩnh vực trong đời sống, ví dụ như: trong y tế, vận tải, tài chính ngân hàng, giáo dục, truyền thông,... Trong năm 2021, Việt Nam đã ứng dụng AI vào lĩnh vực y tế, tham gia vào cuộc cách mạng phòng chống đại dịch COVID-19 đầu tiên phải kể đến đó là DrAid, phần mềm AI đầu tiên tại Việt Nam do Công ty VinBrain thuộc tập đoàn Vingroup phát triển từ năm 2019, giúp chẩn đoán các bệnh về phổi, tim và xương dựa trên X-quang. DrAid giúp phát hiện nhanh các dấu hiệu bất thường dựa trên ảnh X-quang ngực thẳng trong vòng chưa đầy 5 giây, kết hợp cùng xét nghiệm PCR, từ đó nâng cao độ chính xác, hỗ trợ tăng tính nhất quán và chuyên giao kiến thức của bác sĩ từ tuyến Trung ương tới cơ sở.

Nói về ứng dụng AI trong lĩnh vực ngân hàng, ông Nguyễn Việt Phương, Giám đốc phát triển phần mềm Techcombank cho biết ngân hàng này ứng dụng AI đã giúp xác thực thông tin, nhận diện khách hàng thông qua eKYC. Trong bối cảnh dịch COVID-19, nhiều khách hàng không thể đi đến ngân hàng giao dịch trực tiếp, việc ứng dụng AI giúp số lượng khách hàng mở tài khoản tăng lên nhiều. Nhiều mảng công việc, nghiệp vụ khác nhau trong lĩnh vực ngân hàng, theo ông Phương, cũng có thể ứng dụng AI như xử lý dữ liệu lớn, phân tích thông tin giao dịch để tìm ra đặc điểm, thói quen, hành vi của từng phân khúc khách hàng khác nhau để từ đó đưa ra một đề nghị dịch vụ phù hợp với khách hàng. Techcombank ứng dụng AI "phân tích" mùa nào là mùa cao điểm rút tiền từ ATM để tăng cường dòng tiền, thậm chí phân tích thông tin để phòng chống gian lận.[1]

Với những thành quả nêu trên, có thể thấy sự phát triển vượt trội của nền khoa học công nghệ thông tin của Việt Nam đang ngày càng mạnh. Nhằm góp phần tạo thêm giá trị cho nền giáo dục Việt Nam, tạo nên những định hướng phát triển tiềm năng lĩnh vực kỹ thuật và công nghệ. Nhóm nghiên cứu đề tài cũng dựa trên lĩnh vực trí tuệ nhân tạo (AI), nghiên cứu và phát triển ứng dụng học từ vựng Tiếng Anh thông qua hình ảnh, video.

Như chúng ta thấy nhu cầu sử dụng Tiếng Anh để giao tiếp quốc tế, phát triển kiến thức chuyên môn ở ngành nghề cũng như công việc của người dân Việt Nam ngày càng cao. Ngoại ngữ có vai trò rất quan trọng trong đời sống hiện nay. Có thống kê cho rằng, những

người sử dụng được ngoại ngữ rất có khả năng thành công trong nhiều lĩnh vực nghiên cứu, thăng tiến trong nghề nghiệp. Đặc biệt trong thời đại cuộc cách mạng 4.0 đang ngày càng phát triển hiện nay, các ngôn ngữ thông dụng như tiếng Nhật, tiếng Hàn, tiếng Trung,... nói chung và tiếng Anh nói riêng, lại càng đóng vai trò quan trọng trong cuộc sống chúng ta. Có thể nói rằng tiếng Anh là cầu nối giúp bạn dễ dàng có một công việc ổn định với mức lương cao hơn.

Tiếng Anh có vai trò quan trọng đến vậy, yêu cầu của xã hội về khả năng thành thạo ngôn ngữ là khá cao. Nhưng nhìn sâu vào vấn đề đó, việc học và nâng cao trình độ ngoại ngữ ở Việt Nam đặc biệt là tiếng Anh, vô cùng cần thiết đối với học sinh, sinh viên vì đó là điều kiện bắt buộc có để sinh viên ra trường. Đối với những người đã và đang đi làm ở các ngành nghề, có nhu cầu giao tiếp thành thạo tiếng Anh, thì thực trạng về năng lực ngoại ngữ của họ lại đang còn quá thấp, từ vựng bị hạn hẹp, gây cản trở, chưa đáp ứng được trong giao tiếp công việc hiện tại cũng như tương lai. Hơn thế, về phía phụ huynh luôn mong muốn con mình khởi đầu bước học tiếng Anh từ những năm bé còn 3 đến 4 tuổi, nhưng rất đau đầu vì không biết dạy con mình từ những cơ bản nào. Thực cần thiết có một phương pháp học tiếng Anh đi từ cơ bản đến nâng cao. Vì thế đề tài ứng dụng AI tạo nên sản phẩm học từ vựng Tiếng Anh thông qua hình ảnh, video dễ dàng sử dụng và hiệu quả để đóng góp phần nào về sự phát triển nền giáo dục của Việt Nam ngày càng lên cao. Phương pháp học từ vựng Tiếng Anh thông qua hình ảnh, video sẽ giúp cho mọi người tiếp thu nhanh và nhớ lâu hơn. Đặc biệt là trẻ em, ứng dụng sẽ giúp các bé không bị nhàm chán, học vẹt.. mà qua đó sẽ giúp các bé hứng thú, vui nhộn với việc học hơn. Từ đó, con người có thể tự do học và tìm hiểu từ ngữ Tiếng Anh một cách không bị quá phức tạp và gò bó. Hơn thế nữa, mọi người có thể học từ vựng mọi lúc mọi nơi mà không tốn phí, chỉ cần có chiếc điện thoại thông minh đã tải ứng dụng về và kết nối với internet.

2. Ngoài nước

Qua những tìm hiểu về các dự án, đề tài liên quan đến việc đưa công nghệ AI vào các hệ thống giáo dục nói chung và giáo dục ngoại ngữ nói riêng có thể thấy được tiềm năng của công nghệ AI trong giáo dục làm cải thiện chất lượng, nâng cao dân trí, trình độ giáo dục cho người dân. Việc ứng dụng công nghệ AI trong giáo dục tiêu biểu như phần mềm Grammarly - phần mềm kiểm tra lỗi chính tả và ngữ pháp tiếng Anh, hoặc là phần mềm AI-Robotics là công cụ tiềm năng và thiết yếu đối với giáo dục tại Mỹ. Trước hết, AI giúp tự động hóa và thực hiện các nhiệm vụ quản trị của giáo viên như chấm điểm bài thi, phân loại

giấy tờ và tiến hành các thủ tục nhập học... để giáo viên có nhiều thời gian tương tác với học sinh.[2]

Công ty Cổ phần dịch vụ và giải pháp công nghệ giáo dục PHX (PHX Smart School) đã cho ra đời Giải pháp PHX Smart School tích hợp phần mềm hỗ trợ quy trình đào tạo, giúp trường học có thể triển khai hoạt động giảng dạy và học tập trực tuyến, chia sẻ video bài giảng tiện lợi để hỗ trợ học tập hiệu quả hơn; tích hợp hệ thống thi trực tuyến, cho phép tạo các kỳ thi trắc nghiệm và chấm điểm ngay sau khi làm xong bài, hay các bài thi tự luận được giảng viên chấm tay và thông báo kết quả cho học sinh và phụ huynh qua ứng dụng. Đây còn là một phần mềm sổ liên lạc điện tử, kết nối nhà trường với phụ huynh một cách dễ dàng thông qua các app trên Android và iOs. Đặc biệt, PHX Smart School tích hợp hệ thống giám sát lộ trình di chuyển thông minh, điểm danh chính xác bằng gương mặt... tự động kết nối và cập nhật thông tin nhanh chóng tới gia đình về trạng thái của học sinh khi đến trường, đảm bảo an toàn cho học sinh.[3]

II. SỰ CẦN THIẾT CỦA ĐỀ TÀI

Sự phát triển nhanh chóng và liên tục được cải tiến không ngừng của các giải thuật học sâu, các kết quả trong việc nhận dạng vật thể có độ chính xác cao. Đề tài tập chung nghiên cứu và xây dựng một ứng dụng để học Tiếng Anh thông qua việc đánh dấu các vật thể từ hình ảnh, video giúp hỗ trợ người dùng học Tiếng Anh một cách hiệu quả. - Hiện nay, việc học Tiếng Anh là vô cùng quan trọng, trong đó việc sử dụng hình ảnh để học tiếng Anh là một cách học hiệu quả, đặc biệt đối với người học trực quan. Bằng cách nhìn thấy một thứ gì đó, ghi nhớ nó và sau đó liên kết một khái niệm với nó. Thé nê, việc ứng dụng tiềm năng của công nghệ để đánh dấu vật thể trong hình ảnh, video nhằm hỗ trợ việc học tiếng Tiếng Anh là cần thiết để giúp người dùng nâng cao hiệu quả trong việc học Tiếng Anh của mình.

III. MỤC TIÊU CỦA ĐỀ TÀI

Mục tiêu của đề tài nghiên cứu nhằm tạo ra một chương trình ứng dụng học Tiếng Anh trên App Android. Ứng dụng hỗ trợ phát hiện, nhận dạng các vật thể trong video bằng tiếng anh, làm tiền đề giúp người sử dụng học tiếng anh một thoải mái, giải trí mà vẫn trau dồi được nhiều kiến thức từ vựng với chức năng phát hiện đối tượng, phát âm, đặt câu với đối tượng vừa được nhận dạng. Đặc biệt ứng dụng còn cho phép người học xây dựng từ mới cho từ điển một cách dễ dàng, góp phần tạo nên một cộng đồng lớn mạnh cùng giúp đỡ nhau chinh phục vốn từ vựng sâu rộng và phong phú.

IV. NỘI DUNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

1. Nội dung nghiên cứu

- Tìm hiểu nhu cầu sử dụng, tính cấp thiết của đề tài và đối tượng người dùng.
- Phân tích đề tài.
- Thu thập dữ liệu hình ảnh /video về một số từ thông dụng xung quanh ta (cụ thể là trong địa phận Cần Thơ).
- Xây dựng bộ từ điển tiếng anh gồm chữ và âm thanh từ các đối tượng được nhận dạng.
- Nghiên cứu khảo sát các mô hình máy học để đề xuất mô hình máy học phù hợp để nhận dạng.
- Xây dựng một ứng dụng mobile với giao diện thân thiện để tích hợp những nội dung được đề xuất.

2. Phương pháp nghiên cứu

Tổng hợp dữ liệu từ các đối tượng bên ngoài tự nhiên từ đó tạo nên mô hình nhận dạng đối tượng trong video kết hợp với dữ liệu tiếng Anh.

V. ĐỐI TƯỢNG VÀ PHẠM VI NGHIÊN CỨU

1. Đối tượng nghiên cứu

- Công nghệ AI nhận dạng đối tượng trong video
- Chuyển đổi văn bản thành giọng nói
- App Android và cấu trúc dữ liệu lưu trữ bài học

2. Phạm vi nghiên cứu

- Nghiên cứu: bộ từ điển gồm các từ có liên quan đến các đối tượng ngoài tự nhiên bắt đầu từ địa phận Cần Thơ.
- Xây dựng các tính năng: nhận dạng, lưu trữ, thêm từ mới, public, cho ví dụ với từ mới, dịch sang tiếng Việt.

PHẦN 2: NỘI DUNG NGHIÊN CỨU

I. CƠ SỞ LÝ THUYẾT

1. Ngôn ngữ lập trình Python

Python là một ngôn ngữ lập trình mã nguồn mở bậc cao do Guido van Rossum tạo ra và lần đầu ra mắt vào năm 1991. Python không chỉ dễ dàng sử dụng mà còn dễ học đối với người mới tiếp xúc với ngôn ngữ này. Hai yếu tố này đã góp phần làm cho ngôn ngữ này trở thành ngôn ngữ ưa thích cho những người mới bắt đầu phát triển phần mềm. Hơn nữa, tính linh hoạt của nó với tư cách là một ngôn ngữ lập trình đa năng khiến nó phù hợp với nhu cầu của nhiều ngành công nghiệp.

Ban đầu, Python được phát triển để chạy trên hệ điều hành Unix, theo thời gian, Python dần mở rộng sang mọi hệ điều hành từ MS-DOS đến Mac OS, OS/2, Windows, Linux và các hệ điều hành khác thuộc họ Unix.

Sử dụng:

- Python hiện nay được sử dụng rộng rãi và phổ biến, áp dụng trong mọi lĩnh vực chẳng hạn như có thể đóng vai trò như là một ngôn ngữ kịch bản cho ứng dụng web. Các bộ khung web như Django, Pylons, Tornado, Flask, Bottle..... SQLAlchemy có thể được dùng để ánh xạ dữ liệu sang một cơ sở dữ liệu quan hệ. Ngoài ra Python thường được sử dụng trong các dự án trí tuệ nhân tạo và học máy với các thư viện như TensorFlow, Keras, Pytorch và Scikit-learn hoặc sử dụng trong xử lý ngôn ngữ tự nhiên.
- Python có thể được dùng để tạo ra trò chơi thông qua các thư viện chẳng hạn như Pygame, một thư viện làm trò chơi 2D.
- Python còn được sử dụng rộng rãi trong ngành công nghiệp an toàn thông tin, bao gồm phát triển khai thác.[4]

2. Java

Java là một ngôn ngữ lập trình hướng đối tượng (OOP) và dựa trên các lớp (class). Ban đầu Java được phát triển bởi Sun Microsystems và được phát hành vào năm 1995. Được sử dụng trên nhiều nền tảng như Window, Mac hay các phiên bản của Unix và cũng như được sử dụng trong nhiều lĩnh vực.[5]

Sự phổ biến: Tính đến ngày 21 tháng 1 năm 2021, Java là ngôn ngữ lập trình phổ biến thứ nhì thế giới với tỉ lệ 11,96%. Tuy những năm gần đây độ phổ biến có xu hướng

giảm nhưng Java vẫn giữ được tỉ lệ trên 10% mặc cho sự phát triển chóng mặt của thế giới công nghệ. [6]

Sử dụng: Vì Java là một ngôn ngữ miễn phí và linh hoạt được sử dụng rộng rãi trong nhiều lĩnh vực chẳng hạn như phát triển trò chơi, điện toán đám mây, Big Data, hoặc IoT...

3. Tensorflow FrameWork

Tensorflow là một thư viện có mã nguồn mở được tạo ra và phát triển bởi đội ngũ chuyên viên của Google cụ thể là Google Brain được cấp phép hoạt động vào tháng 11 năm 2015. Được sử dụng phổ biến trong lĩnh vực machine learning. TensorFlow kết hợp một loạt các mô hình và thuật toán machine learning cùng deep learning, từ đó làm cho chúng trở nên hữu ích bằng những phép toán. TensorFlow cho phép các nhà phát triển tạo một biểu đồ để thực hiện các tính toán. Mỗi nút trong biểu đồ đại diện cho một phép toán và mỗi kết nối đại diện cho dữ liệu. Vì vậy các nhà phát triển có thể tập trung vào logic tổng thể của ứng dụng. [7]

Sử dụng: Tensorflow được sử dụng trong nhiều ứng dụng dịch vụ của Google như: Thêm và phân tích hình ảnh, nhận dạng giọng nói hay trong các dự án Magenta... Ngoài ra Tensorflow đã và đang phát triển cho việc hỗ trợ các thiết bị iOS.

4. Tensorflow Lite FrameWork

TensorFlow Lite là giải pháp gọn nhẹ của TensorFlow cho thiết bị di động và thiết bị nhúng có thể sử dụng. Nó được thiết kế để có độ trễ thấp, với các hạt nhân được tối ưu hóa cho các ứng dụng di động chẳng hạn như giúp người dùng hoạt động với phân loại hình ảnh từ camera của thiết bị trên cả Android và iOS.[8]

Ưu điểm:

- Nhanh hơn, do TensorFlow Lite cho phép thực hiện machine learning ngay trên device với độ trễ thấp.
- TensorFlow Lite tốn ít dung lượng nên
- Hỗ trợ cảm biến gia tốc của thiết bị android với Android Neural Networks API
- Tối ưu kernels cho mobile apps
- Kích hoạt trước khi hợp nhất (Refused activations)
- Lượng tử hóa kernels cho phép models nhỏ hơn và nhanh hơn [9]

5. Android Studio

Android Studio là IDE chính thức được sử dụng trong phát triển ứng dụng Android dựa trên IntelliJ IDEA. Chức năng chính của Android Studio là cung cấp các giao diện

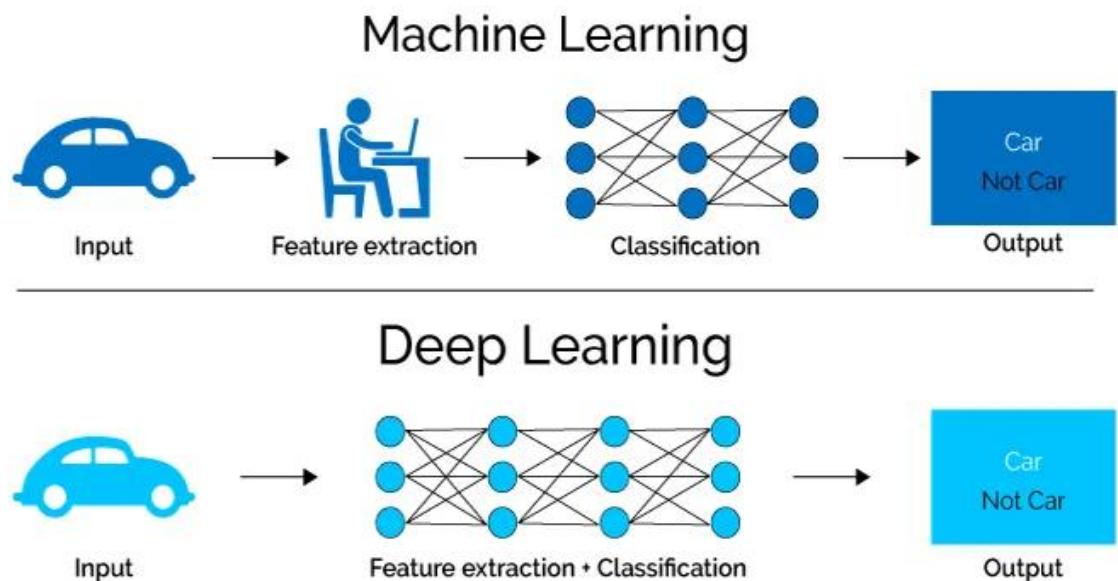
giúp người dùng có thể tạo các ứng dụng và xử lý các công cụ file phức tạp. Ngôn ngữ lập trình được sử dụng chủ yếu trong Android Studio là Java. Khi sử dụng Android Studio, người dùng chỉ cần viết, chỉnh sửa và lưu trữ file các dự án riêng các file nằm trong dự án đó. Đồng thời, Android Studio còn cung cấp quyền truy cập vào Android SDK. Hiện nay, đội ngũ phát triển đã và đang tiếp tục giúp Android Studio trở nên tiện ích và hữu ích hơn, hỗ trợ tối đa và gần gũi hơn với người dùng. [10]

6. Deep Learning

Deep Learning (học sâu) là một lĩnh vực con của Machine Learning (máy học), các máy tính sẽ học và cải thiện chính nó thông qua các thuật toán. Deep Learning được xây dựng dựa trên các khái niệm phức tạp, chủ yếu hoạt động với các mạng nơ-ron nhân tạo để bắt chước khả năng tư duy và suy nghĩ của bộ não con người.[13]

Cách thức hoạt động:

- Deep Learning là một phương pháp của Machine Learning. Mạng nơ-ron nhân tạo trong Deep Learning được xây dựng để mô phỏng khả năng tư duy của bộ não con người.
- Một mạng nơ-ron bao gồm nhiều lớp (layer) khác nhau, số lượng layer càng nhiều thì mạng sẽ càng “sâu”. Trong mỗi layer là các nút mạng (node) và được liên kết với những lớp liền kề khác. Mỗi kết nối giữa các node sẽ có một trọng số tương ứng, trọng số càng cao thì ảnh hưởng của kết nối này đến mạng nơ-ron càng lớn.
- Mỗi nơron sẽ có một hàm kích hoạt, về cơ bản thì có nhiệm vụ “chuẩn hoá” đầu ra từ nơ-ron này. Dữ liệu được người dùng đưa vào mạng nơ-ron sẽ đi qua tất cả layer và trả về kết quả ở layer cuối cùng, gọi là output layer.



Hình 1: Cơ chế hoạt động của Deep Learning [13]

- Trong quá trình huấn luyện mô hình mạng nơ-ron, các trọng số sẽ được thay đổi và nhiệm vụ của mô hình là tìm ra bộ giá trị của trọng số sao cho phán đoán là tốt nhất.
- Các hệ thống Deep Learning yêu cầu phần cứng phải rất mạnh để có thể xử lý được lượng dữ liệu lớn và thực hiện các phép tính phức tạp. Nhiều mô hình Deep Learning có thể mất nhiều tuần, thậm chí nhiều tháng để triển khai trên những phần cứng tiên tiến nhất hiện nay.

Deep Learning cho phép các nhà khoa học dữ liệu xây dựng nhiều mô hình có độ chính xác rất cao trong lĩnh vực nhận dạng ảnh, xử lý ngôn ngữ tự nhiên, xử lý giọng nói,... Một số ưu điểm vượt trội của Deep Learning gồm có:

- Kiến trúc mạng nơ-ron linh hoạt, có thể dễ dàng thay đổi để phù hợp với nhiều vấn đề khác nhau.
- Có khả năng giải quyết nhiều bài toán phức tạp với độ chính xác rất cao.
- Tính tự động hóa cao, có khả năng tự điều chỉnh và tự tối ưu.
- Có khả năng thực hiện tính toán song song, hiệu năng tốt, xử lý được lượng dữ liệu lớn.

Bên cạnh những ưu điểm, Deep Learning vẫn còn nhiều khó khăn và hạn chế như:

- Cần có khối lượng dữ liệu rất lớn để tận dụng tối đa khả năng của Deep Learning.
- Chi phí tính toán cao vì phải xử lý nhiều mô hình phức tạp.

- Chưa có nền tảng lý thuyết mạnh mẽ để lựa chọn các công cụ tối ưu cho Deep Learning.

Một số ứng dụng sử dụng deep Learning trong đời sống:

- Xe tự lái
- Phân tích cảm xúc
- Trợ lý ảo
- Mạng xã hội
- Chăm sóc sức khỏe [13]

7. SQLite

SQLite là một hệ cơ sở dữ liệu nhúng và là một ứng dụng nằm trong miền công cộng. SQLite tập hợp rất nhiều tính năng nổi bật phù hợp ứng dụng vào dự án như gọn, nhẹ, cách cài đặt, sử dụng đơn giản, không yêu cầu máy chủ riêng biệt, tốc độ và chức năng hoàn toàn tốt hơn các cơ sở dữ liệu khác đối với quy mô vừa của dự án. Ứng dụng còn tuân thủ theo nguyên tắc ACID ngay cả khi hệ thống gặp sự cố hoặc mất điện. Ngoài ra vì nằm trong miền công cộng nên SQLite được sử dụng miễn phí cho bất kỳ mục đích nào và còn được tích hợp vào tất cả các điện thoại di động nên rất phù hợp với hệ thống chạy trên ứng dụng di động của dự án này.

II. Thu thập dữ liệu ảnh và phương pháp tăng dữ liệu ảnh (data augmentation)

1. Thu thập dữ liệu:

Nhóm nghiên cứu đã sử dụng công cụ SQLite để lưu trữ và tạo ra bộ từ điển gồm 145 lớp đối tượng khác nhau phổ biến tại Cần Thơ, sau đó đề tài tiến hành thu thập dữ liệu từ nhiều nguồn ở trên mạng và từ điện thoại thông qua camera. Tùy vào độ đa dạng về hình dạng, màu sắc của đồ vật nên số lượng dữ liệu thu thập của lớp sẽ có chênh lệch, tuy nhiên đề tài đã đặt chỉ tiêu khoảng 50 ảnh cho mỗi lớp. Kết quả mà đề tài đã thu thập tổng cộng 7303 ảnh cho 144 lớp đồ vật (trung bình là 49-50 ảnh cho một lớp). Số lớp và ảnh có sự gộp mặt của lớp đã liệt kê theo bảng, tuy nhiên có một số ảnh xuất hiện nhiều lớp, vì vậy tổng số ảnh liệt kê tại bảng không phản ánh toàn bộ ảnh mà đề tài thu thập.

Bảng 1: Mô tả dữ liệu đã thu thập

Tên đồ vật	Tên lớp của đồ vật	Số lượng ảnh	Tên đồ vật	Tên lớp của đồ vật	Số lượng ảnh
Táo	000000_Apple	50	Chuồng chuồng	000076_Dragonfly	50
Mận	000001_BellFruit	50	Bọ ngựa	000077_Mantis	50

Tên đồ vật	Tên lớp của đồ vật	Số lượng ảnh	Tên đồ vật	Tên lớp của đồ vật	Số lượng ảnh
Dưa hấu	000002_WaterMelon	50	Châu chấu	000078_Grasshopper	50
Chôm chôm	000003_Rambutan	50	Dế	000079_Cricket	51
Bưởi	000004_Pomelo	50	Gián	000080_Cockroach	48
Nho	000005_Grape	50	Người	000081_Human	52
Cam	000006_Orange	50	Sâu	000082_Worm	53
Vú sữa	000007_StarApple	50	Bướm	000083_Butterfly	49
Dâu da	000008_BurmeseGrape	50	Gà con	000086_Chick	49
Bòn bon	000009_Langsat	51	Rắn	000087_Snake	50
Chuối	000010_Banana	50	Cá	000088_Fish	47
Đu đủ	000011_Papaya	50	Tủ quần áo	000094_Wardrobe	27
Măng cầu	000012_CustardApple	50	Gối chiếc	000095_Pillow	51
Khế	000013_StarFruit	50	Chăn	000096_Blanket	51
Quả na	000014_SugarApple	50	Gối ôm	000097_HugPillow	52
Măng cụt	000015_Mangosteen	49	Gấu bông	000098_TeddyBear	53
Thanh long	000016_DragonFruit	50	Gương	000099_Mirror	52
Lê	000017_Pear	50	Vợt muỗi	000100_Mosquito_Racket	50

Tên đồ vật	Tên lớp của đồ vật	Số lượng ảnh	Tên đồ vật	Tên lớp của đồ vật	Số lượng ảnh
Mít	000018_JackFruit	50	Mùng	000101_Sleeping_Nets	52
Ôi	000019_Guava	50	Nệm	000102_Mattress	62
Nhãn	000020_Longan	50	Cửa	000103_Door	53
Quýt	000021_Tangerine	51	Cửa sổ	000104_Window	51
Sầu riêng	000022_Durian	50	Bồn cầu	000106_Toilet	50
Vải	000023_Lychee	50	Bàn chải	000107_Toothbrush	92
Khóm	000024_Pineapple	50	Bồn rửa mặt	000108_Sink	49
Xoài	000025_Mango	51	Vòi sen	000109_Shower	49
Lựu	000026_Pomegranates	49	Vòi nước	000110_Faucet	50
Cóc	000027_Ambarella	50	Bồn tắm	000111_Bathtub	50
Xe máy	000028_Motobike	70	Bông tắm	000112_BathSponge	50
Xe đạp	000029_Bicycle	53	Bàn chải chà chân	000113_Brushes	50
Xe hơi	000030_Car	66	Cây lao nhà	000114_Mop	49
Xe tải	000031_Truck	60	Cây chổi	000115_Broom	50
Xe khách	000032_Bus	52	Cây chà bồn cầu	000116_ToiletBrush	12
Xe cứu Hỏa	000033_FireTruck	50	Tăm nước	000117_WaterFlosser	49

Tên đồ vật	Tên lớp của đồ vật	Số lượng ảnh	Tên đồ vật	Tên lớp của đồ vật	Số lượng ảnh
Xe cứu Thương	000034_Ambulance	50	Cạo râu	000118_Razor	50
Xe lu	000035_RoadRoller	50	Lược	000119_Comb	49
Xe cẩu	000036_Excavator	50	Máy giặt	000120_WashingMachine	50
Xe lăn	000037_Wheelchair	48	Bếp ga	000121_Gas_Stove	51
Âm trà	000041_Teapot	50	Bình ga	000122_Gas_Bottle	52
Bàn	000042_Table	50	Bếp từ	000123_Induction_Cooker	53
Bình hoa	000043_Vase	50	Nồi chiên không dầu	000124_Oil-free_fryer	52
Đèn (ống dài)	000044_LEDtubelights	50	Chảo	000125_Pan	53
Đèn trần	000045_Ceiling_lights	48	Âm nước	000126_Kettle	54
Gạt tàn	000046_Ashtray	50	Bình thủy	000127_Vase	53
Ghế	000047_Chair	50	Nồi cơm điện	000128_Electric_Cooker	50
Kệ đựng sách	000048_Bookshelf	50	Chén	000129_Bowl	51
Khung ảnh	000049_Picture_frames	50	Tô	000130_BigBowl	53
Loa	000050_Loudspeaker	50	Đũa	000131_Chopsticks	52
Ly	000051_Glass	50	Muỗng	000132_Spoon	53
Máy lạnh	000052_Air-conditioner	50	Dĩa	000133_Plate	51

Tên đồ vật	Tên lớp của đồ vật	Số lượng ảnh	Tên đồ vật	Tên lớp của đồ vật	Số lượng ảnh
Máy quạt	000053_Fan	50	Nồi	000134_Pot	52
Màn cửa sổ	000054_Window blinds	50	Dao	000135_Knife	51
Sofa	000055_Sofa	49	Nĩa	000136_Fork	50
Tivi	000056_Tivi	49	Máy giặt	000137_Washing Machine	50
Chó	000057_Dog	51	Tủ lạnh	000138_Fridge	50
Mèo	000058_Cat	50	Tăm tre	000139_Bamboo-ToothPick	50
Vịt	000059_Duck	50	Nắp nồi	000140_Lid	50
Muỗi	000060_Mosquito	50	Điện thoại	000141_SmartPhone	75
Ong	000061_Bee	50	Sạc	000142_Charger	46
Ruồi	000062_Fly	50	Ô điện	000143_Socket	50
Nhện	000063_Spider	51	Laptop	000144_Laptop	50
Tắc kè	000064_Lizard	50	Võng	000145_Hammock	50
Thằn lằn	000065_Gecko	44	Ớt chuông	000146_BellPepper	50
Chim sẻ	000066_Sparrow	50	Bầu	000147_BottleGourd	50
Bồ câu	000067_Pigeon	48	Bông cải xanh	000148_Broccoli	50
Gà	000068_Chicken	50	Cải bắp	000149_Cabbage	50

Tên đồ vật	Tên lớp của đồ vật	Số lượng ảnh	Tên đồ vật	Tên lớp của đồ vật	Số lượng ảnh
Sáo	000069_Starling	50	Củ cà rốt	000150_Carrot	50
Chuột kiểng	000070_Hamster	32	Súp lơ	0	50
Chuột	000071_Mouse	50	Dưa leo	1	52
Thỏ	000072_Rabbit	50	Khoai	2	50
Bò	000073_Cow	50	Bí ngô	3	52
Trâu	000074_Buffalo	50	Củ cải trắng	4	50
Cò	000075_Stork	50			

2. Phương pháp tăng dữ liệu ảnh(data augmentation):

Phương pháp tăng dữ liệu ảnh là một tập hợp các kỹ thuật để tăng lượng dữ liệu bằng cách giả tạo dữ liệu từ dữ liệu hiện có. Điều này giúp cải thiện chất lượng của mô hình sau khi huấn luyện, và lần này để tài quyết định lựa chọn một phương pháp tăng cường dữ liệu đó là phương pháp Lật (Flip).

Phương pháp lật (Hình) là phương pháp tăng cường dữ liệu bằng cách đảo ngược hình ảnh theo chiều ngang, bằng cách này, dữ liệu sẽ được tăng lên mà không cần phải thêm hình ảnh mới nào.



Hình gốc



Hình sau khi lật

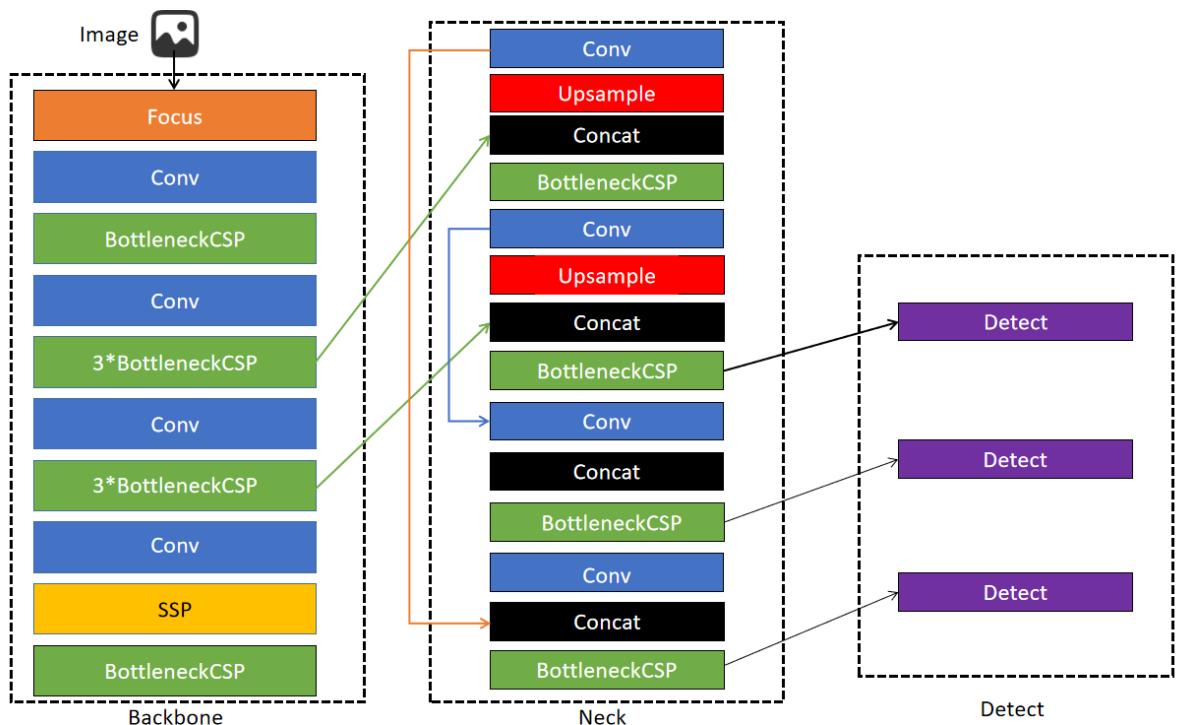
Hình 2: Hình ảnh sau khi dùng phương pháp tăng cường dữ liệu

II. PHƯƠNG PHÁP NGHIÊN CỨU

1. Mô hình tổng quan:

1.1. Cấu trúc tổng quan của model:

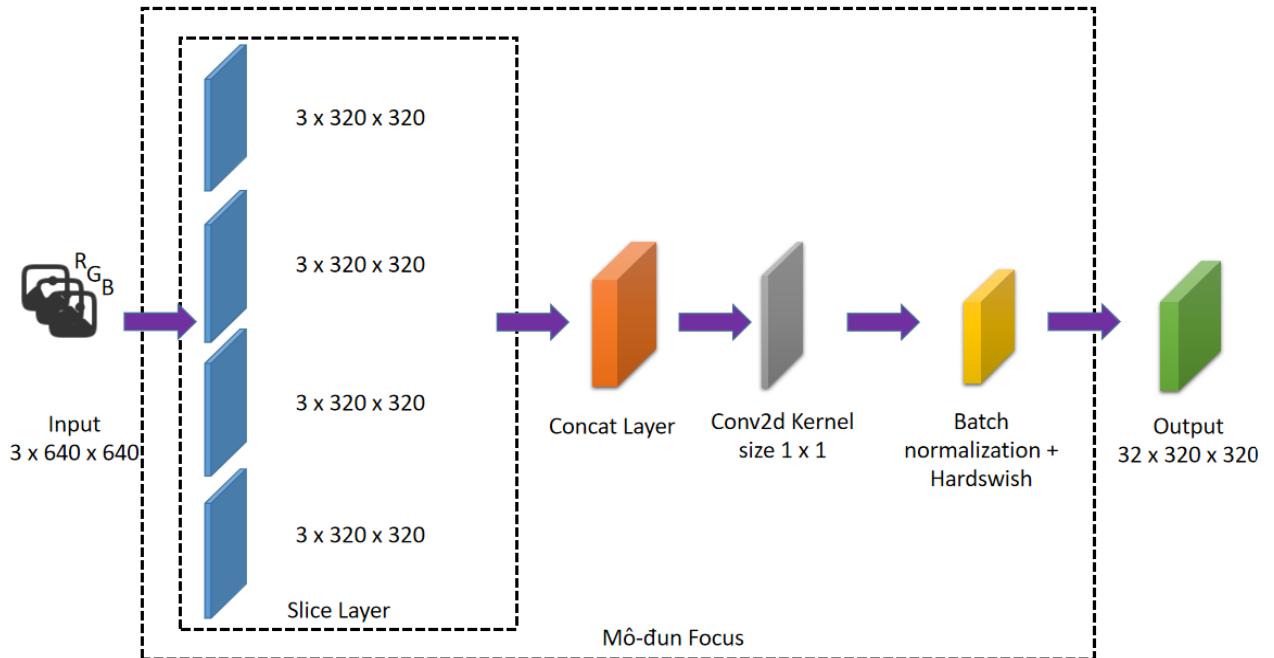
Yolov5 [14] là sản phẩm thuộc chuỗi kiến trúc Yolo. Độ chính xác phát hiện của mô hình này cao và tốc độ nhận dạng nhanh, với tốc độ phát hiện nhanh nhất lên tới 140 khung hình mỗi giây, đồng thời kích thước tệp sau khi train của mô hình Yolov5 là nhỏ, gần như nhỏ hơn 90% so với Yolov4 cho thấy độ phù hợp để triển khai cho các thiết bị nhúng để phát hiện nhận dạng vật thể ở thời gian thực. Yolov5 chứa tận 10 kiến trúc khác nhau bao gồm: YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x, YOLOv5n6, YOLOv5s6, YOLOv5m6, YOLOv5l6 và YOLOv5x6, sự khác biệt giữa chúng là số lượng các mô-đun trích xuất và kernel tích chập. Độ chính xác, thời gian nhận dạng và kích thước mô hình đã được xem xét toàn diện trong nghiên cứu với mục tiêu nhận dạng vật thể được thực hiện dựa trên kiến trúc YOLOv5 (Hình 3)



Hình 3: Kiến trúc YOLOv5 [15]

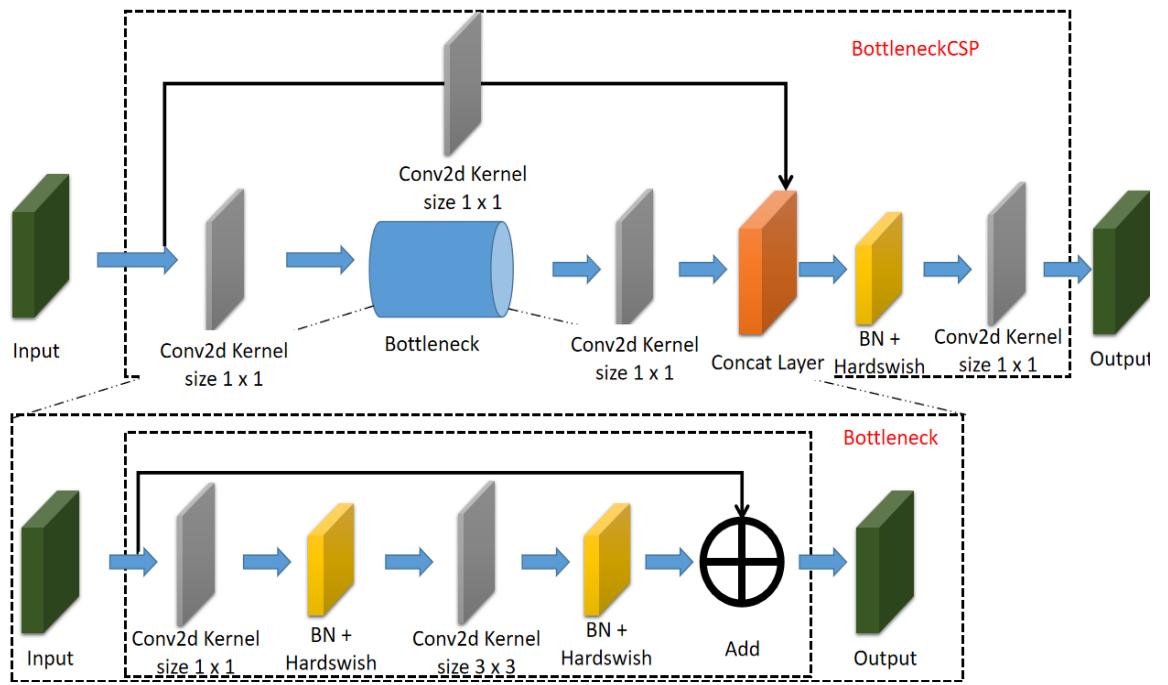
Kiến trúc Yolov5 chủ yếu gồm ba phần gồm: backbone network, neck network and detect network. Backbone network là một mạng nơ ron phức hợp tổng hợp từ các hình ảnh chi tiết khác nhau và tạo thành đặc điểm hình ảnh. Cụ thể hơn, lớp đầu tiên của phần Backbone là mô-đun Focus (Hình 4) được thiết kế để giảm tính toán của

mô hình và tăng tốc độ đào tạo. Bao gồm các chức năng sau, đầu vào là hình ảnh với 3 kênh (kích thước hình ảnh đầu vào mặc định là $3 \times 640 \times 640$) được phân đoạn thành 4 phần với kích thước $3 \times 320 \times 320$ mỗi phần, thứ hai Concat được sử dụng để kết nối vào 4 phần feature maps với kích thước của đầu ra là $12 \times 320 \times 320$, sau đó qua lớp tích chập gồm 32 kernel feature maps đầu ra có kích thước $32 \times 320 \times 320$ đã được tạo, cuối cùng thông qua lớp Batch Normalization, kết quả sẽ được xuất vào lớp kế tiếp.



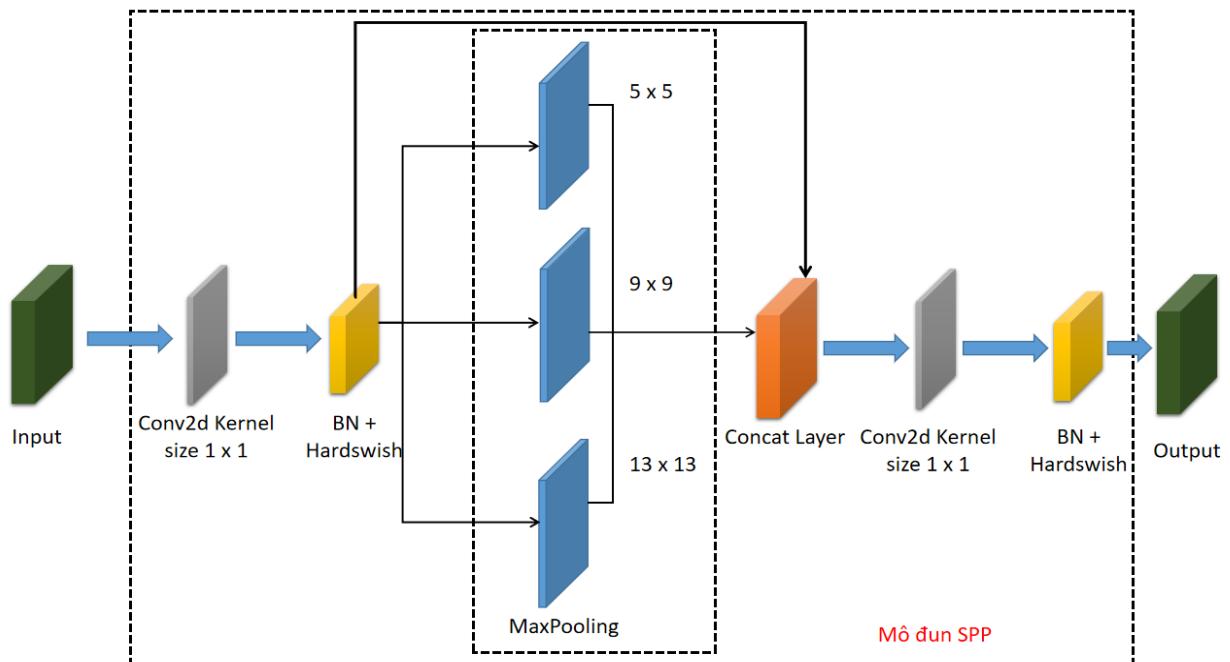
Hình 4: Mô-đun Focus [15]

Lớp thứ ba trong phần Backbone là lớp BottleneckCSP (hình 5) được thiết kế để trích xuất tốt hơn các đặc điểm của hình ảnh. Đầu vào ban đầu được chia ra thành hai nhánh với số lượng kênh của feature maps được giảm đi phân nửa thông qua thao tác tích chập ở hai nhánh .Lớp BottleneckCSP bao gồm một mô đun Bottleneck là (a residual network architecture that connects a convolutional layer) có kích thước kernel tích chập là 1×1 , và thêm một lớp tích chập có kích thước là 3×3 . Đầu ra cuối cùng của mô-đun bottleneck là phần đầu ra bổ sung cùng với phần dư từ đầu vào ban đầu. Từ đầu ra của Bottleneck ở nhánh một và lớp Conv2d ở nhánh hai, đầu ra feature maps của hai nhánh được kết nối sâu với thao tác concat. Cuối cùng, đầu ra feature maps của mô-đun thu được sau khi lần lượt đi qua lớp BN (batch normalization) và lớp Conv2d, kích thước của đầu ra feature maps lớp BottleneckCSP tương tự với kích thước đầu vào.



Hình 5: BottleneckCSP [15]

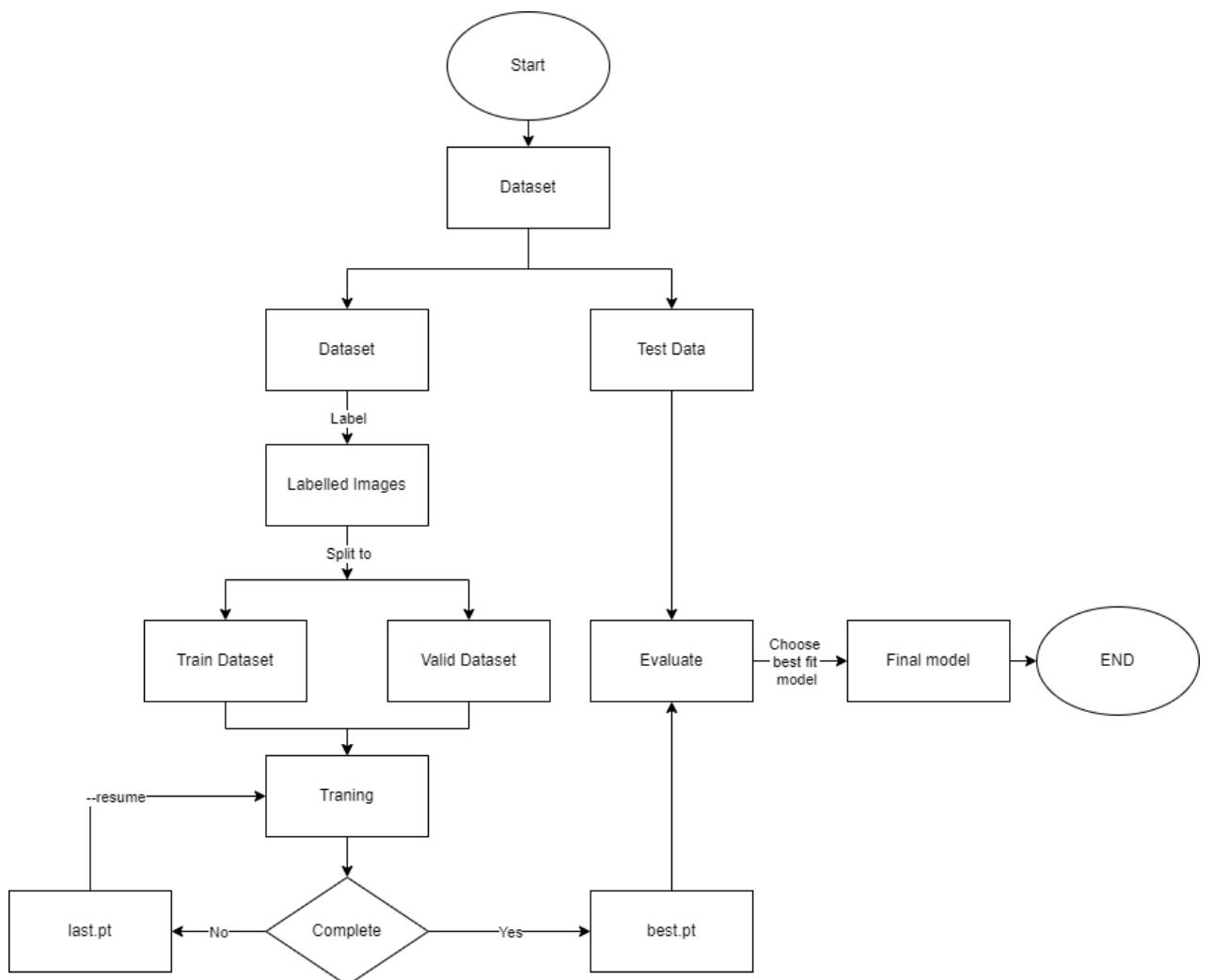
Lớp thứ chín trong phần Backbone là mô-đun SPP (spatial pyramid pooling) (Hình 5) được thiết kế để cải thiện trường tiếp nhận của mô hình bằng cách chuyển đổi các kích thước của các feature maps thành một vector đặc trưng có kích thước cố định. Đầu tiên feature maps sẽ đi qua lớp tích chập có kích thước là 1×1 . Sau đó, sẽ chia ra làm hai nhánh, một nhánh sẽ được lấy mẫu phụ thông qua ba lớp Maxpooling song song rồi gộp lại với nhánh còn lại thu được feature maps cuối cùng sau khi đi qua được lớp tích chập.



Hình 6: Mô đun SPP [15]

Phần Neck là tổng hợp của một loạt các “feature aggregation layers of mix” và “combined image feature”, được sử dụng để tạo FPN (feature pyramid networks) và sau đó đầu ra feature map được truyền đến prediction network. Vì trình trích xuất tính năng của mạng này áp dụng cấu trúc FPN nên việc truyền các “low-level features” được tăng cường giúp cải thiện khả năng phát hiện các đối tượng có kích thước và tỉ lệ khác nhau. Phần Detect chủ yếu sử dụng cho phần phát hiện cuối cùng của mô hình.

1.2. Workflow



Hình 7: Sơ đồ chi tiết cách thức xây dựng và đào tạo mô hình

Quá trình xây dựng mô hình được chia làm 3 giai đoạn:

- Giai đoạn xử lý ảnh: Đầu tiên ta chuẩn bị bộ dataset gồm 145 lớp đối tượng với trung bình 50 ảnh mỗi lớp để đào tạo và 15 ảnh mỗi lớp không trùng với các ảnh ban đầu để làm tập kiểm thử(test data). Gắn nhãn bộ dataset dùng để đào tạo mô hình bằng công cụ roboflow và chia chúng ra hai tập train và tập validation với tỉ lệ 4:1.

- Giai đoạn đào tạo: Sau khi đã chuẩn bị xong ta tiến hành đào tạo mô hình sử dụng thuật toán yolov5. Để tối ưu hóa mô hình, nhóm đã cho máy học đến tám trăm lần (800 epochs) với số ảnh mỗi lần đọc là bốn (batch = 4) và phải đọc hết một nghìn bốn trăm năm mươi ba lần mới được một lần học và cứ lặp lại 800 lần học ta được mô hình hoàn chỉnh. Sau mỗi lần đào tạo, ta sẽ được 2 mô hình là file best.pt và last.pt, file best chính là mô hình tốt nhất ta dùng để nhận dạng và phát hiện đối tượng còn last.pt ta sử dụng để tiếp tục training nếu quá trình học tập của máy học bị gián đoạn hoặc khi ta muốn tăng dữ liệu cho mô hình.
- Giai đoạn kiểm thử: Sau khi có mô hình hoàn chỉnh, ta tiếp tục đánh giá, so sánh các mô hình với nhau để chọn ra mô hình phù hợp nhất cho bài toán nhận dạng đối tượng này.

2. Xử lý hình ảnh:

Ảnh vật thể sau khi thu thập xong sẽ thông qua bước gán nhãn, việc gán nhãn sẽ được thực hiện trên trang web Roboflow. Roboflow là một sản phẩm phần mềm dưới dạng dịch vụ giúp người dùng quản lý tệp hình ảnh, chú thích, nhãn, tiền xử lý, phương pháp tăng cường dữ liệu, định dạng tệp và huấn luyện mô hình với thao tác đơn giản, giúp phục vụ công việc thị giác máy tính dễ dàng hơn.



Hình 8: Gắn nhãn hình ảnh trên Roboflow

Sau khi gán nhãn xong, sẽ dùng các công cụ từ roboflow tạo ra phiên bản dataset từ các dữ liệu mà đã thu thập và gán nhãn, phiên bản đó sẽ cài đặt lại kích thước ảnh thành 640 x 640.

3. Object Detection:

Object Detection là một nhiệm vụ vô cùng quan trọng của thị giác máy tính (computer vision) liên quan tới đến khả năng của hệ thống máy tính và phần mềm để xác định, định vị các đối tượng trong một hình ảnh. Object Detection đã được sử dụng rộng rãi để phát hiện khuôn mặt, phát hiện xe, đếm số người đi bộ, hệ thống bảo mật và xe không người lái. Có nhiều cách để nhận diện đối tượng có thể được sử dụng cũng như trong nhiều lĩnh vực thực hành và trong nghiên cứu này Object Detection được sử dụng để phát hiện các đối tượng nhằm xây dựng chức năng hỗ trợ cho việc học Tiếng Anh thông qua video.

4. Hệ thống lưu trữ từ điển

Hệ thống lưu trữ từ điển trong dự án bao gồm hai phần: các cấu trúc bảng, ràng buộc, nội dung chữ, đường dẫn liên kết được lưu trữ trong hệ quản trị cơ sở dữ liệu SQLite; và phần còn lại là các tập tin âm thanh được lưu trữ trên GitHub.

4.1. Lưu trữ trên SQLite

Để sử dụng cơ sở dữ liệu SQLite, chúng tôi đã cài đặt một ứng dụng hỗ trợ có tên là “DB Browser for SQLite” và phần mềm quản lý cơ sở dữ liệu “DBeaver”.

Cấu trúc của cơ sở dữ liệu sẽ bao gồm hai bảng:

- Bảng “Doituong” dùng để lưu trữ tên của đối tượng đã được thu thập và đường dẫn liên kết tới GitHub để lấy tập tin âm thanh chứa các định nghĩa của đối tượng. Bảng sẽ gồm 145 đối tượng được đánh số ID tương ứng và tất cả đều nhập vào bảng song ngữ cả tiếng Anh và tiếng Việt.

```
CREATE TABLE "Doituong" (
    "Iddt"      INTEGER,
    "TendtEN"    TEXT,
    "TendtVN"    TEXT,
    "DinhnghiaEN"  BLOB,
    "DinhnghiaVn"  BLOB,
    PRIMARY KEY("Iddt")
);
```

- Bảng “Datcau” dùng để lưu trữ các câu có liên quan đến đối tượng nhằm đưa ra ví dụ, tình huống sử dụng đối tượng trong đời sống. Bảng chứa hơn 700 câu song ngữ với trung bình 5 câu cho một đối tượng. Câu được thu thập chủ yếu bằng các trang web Collins Dictionary, YourDictionary, Ichacha và Google Translate.

```
CREATE TABLE "Datcau" (
    "Iddc"      INTEGER,
    "CauEN"    TEXT,
    "CauVN"    TEXT,
    "Iddt"      INTEGER,
    FOREIGN KEY("Iddt") REFERENCES "Doituong" ("Iddt"),
    PRIMARY KEY("Iddc")
);
```

4.2. Lưu trữ trên GitHub

GitHub là một hệ thống quản lý dự án và phiên bản code, hoạt động gần giống như một trang mạng xã hội. Github cho phép lưu trữ các tập tin âm thanh cũng như chuyển kho lưu trữ thành trang lưu trữ thông qua GitHub Pages, vì vậy có thể dễ dàng lấy đường liên kết các tập tin định nghĩa đối tượng bằng âm thanh đem lưu trữ vào SQLite. Định nghĩa các đối tượng chủ yếu là hướng đến những đặc điểm về ngoại hình, màu sắc, tính chất, cân nặng,... của đối tượng được thu thập chủ yếu từ trang web Wikipedia. Các định nghĩa sau khi thu thập được chuyển thành tập tin âm thanh thông qua trang web TTSFree.com trước khi được đưa lên GitHub.

PHẦN 3: THỰC NGHIỆM

I. KẾT QUẢ ĐÁNH GIÁ MÔ HÌNH:

1. Phân chia training, validation and testing set:

Phân chia tập dataset thành ba loại: training validation và testing. Đề tài chia theo tỉ lệ 8/2/0 vì tập test sẽ làm riêng sau khi huấn luyện mô hình (trung bình mỗi lớp 15 ảnh test) ảnh thành hai tập: tập training và tập validation, số lượng ảnh tại tập training là 5809 ảnh, còn tập validation là khoảng 1494 ảnh. Sau khi áp dụng phương pháp tăng dữ liệu ảnh, số lượng ảnh ở tập training là 11683 ảnh, còn ở tập validation vẫn là 1494.

2. Kết quả sau khi huấn luyện:

Kết quả trình bày trong bảng 2 dưới đây là kết quả được train bằng YOLOv5m, dùng dataset bình thường không sử dụng phương pháp tăng cường dữ liệu.

Các chỉ số trình bày theo bảng 2 dưới đây bao gồm :

- **Precision:** biểu thị tỷ lệ giữa số sample được tính là True Positive với tổng số sample được phân loại là Positive.

Công thức tính Precision:

$$\text{Precision} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalsePositive}}$$

- **Recall:** biểu thị tỷ lệ phần trăm các positive sample được phân loại chính xác so với tất cả các positive sample.

Công thức tính Recall:

$$\text{Recall} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative}}$$

- **mAP50:** là chỉ số thể hiện độ chính xác của mô hình, được đánh giá dựa trên tập validation.

Bảng 2: Kết quả sau khi huấn luyện

<u>Class</u>	<u>Tên đối tượng</u>	<u>Precision</u>	<u>Recall</u>	<u>mAP50</u>
0	Súp lơ	0.760	0.941	0.917
1	Dưa leo	0.628	0.950	0.916
2	Khoai	0.891	0.966	0.990
3	Bí ngô	0.678	0.944	0.859

<u>Class</u>	<u>Tên đối tượng</u>	<u>Precision</u>	<u>Recall</u>	<u>mAP50</u>
4	Củ cải Trắng	0.832	0.870	0.902
000000_Apple	Táo	0.802	0.793	0.917
000001_BellFruit	Mận	0.750	0.931	0.890
000002_WaterMelon	Dưa hấu	0.985	1.000	0.995
000003_Rambutan	Chôm chôm	0.956	0.870	0.954
000004_Pomelo	Bưởi	0.956	0.717	0.884
000005_Grape	Nho	0.866	0.698	0.774
000006_Orange	Cam	0.812	0.914	0.925
000007_StarApple	Vú sữa	0.730	0.953	0.927
000008_BurmeseGrape	Dâu da	0.586	0.462	0.504
000009_Langsat	Bòn bon	0.937	1.000	0.991
000010_Banana	Chuối	0.942	0.824	0.991
000011_Papaya	Đu đủ	0.650	0.884	0.921
000012_CustardApple	Măng cụt	0.925	1.000	0.995
000013_StarFruit	Khê	0.861	0.852	0.908
000014_SugarApple	Quả na	0.894	1.000	0.966
000015_Mangosteen	Măng cụt	0.870	0.879	0.918
000016_DragonFruit	Thanh long	0.912	0.075	0.840
000017_Pear	Lê	0.798	0.542	0.654
000018_JackFruit	Mít	0.842	0.839	0.888
000019_Guava	Ổi	0.759	0.913	0.912
000020_Longan	Nhãn	0.882	0.577	0.629
000021_Tangerine	Quýt	0.772	0.891	0.917
000022_Durian	Sầu riêng	0.896	0.938	0.956
000023_Lychee	Vải	0.823	0.970	0.973

<u>Class</u>	<u>Tên đối tượng</u>	<u>Precision</u>	<u>Recall</u>	<u>mAP50</u>
000024_Pineapple	Khóm	0.666	0.760	0.793
000025_Mango	Xoài	0.612	0.910	0.864
000026_Pomegranates	Lựu	0.795	0.960	0.938
000027_Ambarella	Cóc	0.724	0.982	0.960
000028_Motobike	Xe máy	0.621	0.857	0.728
000029_Bicycle	Xe đạp	0.766	1.000	0.962
000030_Car	Xe hơi	0.880	0.457	0.480
000031_Truck	Xe tải	0.783	0.714	0.855
000032_Bus	Xe khách	0.803	1.000	0.989
000033_FireTruck	Xe cứu Hỏa	0.845	0.917	0.915
000034_Ambulance	Xe cứu Thương	0.944	1.000	0.995
000035_RoadRoller	Xe lu	0.959	1.000	0.995
000036_Excavator	Xe cẩu	0.846	0.900	0.942
000037_Wheelchair	Xe lăn	1.000	0.761	0.972
000041_Teapot	Âm trà	0.975	1.000	0.995
000042_Table	Bàn	0.546	0.967	0.906
000043_Vase	Bình hoa	0.778	1.000	0.986
000044_LEDtubelights	Đèn (óng dài)	0.837	1.000	0.973
000045_Ceiling lights	Đèn trần	0.843	0.895	0.927
000046_Ashtray	Gạt tàn	0.854	0.778	0.769
000047_Chair	Ghế	0.917	1.000	0.995
000048_ Bookshelf	Kệ đựng sách	0.661	0.789	0.862
000049_Picture frames	Khung ảnh	0.813	0.889	0.908
000050_Loudspeaker	Loa	0.847	0.975	0.962
000051_Glass	Ly	0.849	1.000	0.995

<u>Class</u>	<u>Tên đối tượng</u>	<u>Precision</u>	<u>Recall</u>	<u>mAP50</u>
000052_Air-conditioner	Máy lạnh	0.856	0.812	0.881
000053_Fan	Máy quạt	0.950	1.000	0.995
000054_Window blinds	Màn cửa sổ	0.761	0.911	0.942
000055_Sofa	Sofa	0.944	1.000	0.995
000056_Tivi	Tivi	0.866	0.923	0.951
000057_Dog	Chó	0.773	0.570	0.667
000058_Cat	Mèo	0.866	0.649	0.936
000059_Duck	Vịt	1.000	0.860	0.975
000060_Mosquito	Muỗi	0.881	0.909	0.968
000061_Bee	Ong	0.871	0.412	0.588
000062_Fly	Ruồi	0.929	1.000	0.995
000063_Spider	Nhện	0.860	0.400	0.666
000064_Lizard	Tắc kè	0.729	0.541	0.860
000065_Gecko	Thằn lằn	0.670	0.625	0.621
000066_Sparrow	Chim sẻ	0.876	0.818	0.893
000067_Pigeon	Bồ câu	0.831	0.273	0.424
000068_Chicken	Gà	0.947	0.833	0.940
000069_Starling	Sáo	0.891	1.000	0.995
000070_Hamster	Chuột kiêng (hamster)	0.399	0.444	0.473
000071_Mouse	Chuột	0.658	0.578	0.564
000072_Rabbit	Thỏ	0.605	0.769	0.847
000073_Cow	Bò	0.857	0.737	0.841
000074_Buffalo	Trâu	0.821	0.909	0.940
000075_Stork	Cò	0.945	0.917	0.923

<u>Class</u>	<u>Tên đối tượng</u>	<u>Precision</u>	<u>Recall</u>	<u>mAP50</u>
000076_Dragonfly	Chuồng chuồng	0.958	0.900	0.954
000077_Mantis	Bọ ngựa	0.916	0.700	0.912
000078_Grasshopper	Châu chấu	0.688	0.900	0.901
000079_Cricket	Dé	0.624	0.498	0.732
000080_Cockroach	Gián	0.909	0.625	0.619
000081_Human	Người	0.755	0.411	0.651
000082_Worm	Sâu	0.997	0.818	0.892
000083_Butterfly	Bướm	0.902	0.923	0.977
000086_Chick	Gà con	0.819	0.857	0.893
000087_Snake	Rắn	0.602	0.456	0.618
000088_Fish	Cá	0.662	0.327	0.453
000094_Wardrobe	Tủ quần áo	0.870	0.609	0.854
000095_Pillow	Gối chiếc	0.326	0.571	0.423
000096_Blanket	Chăn	0.696	0.688	0.745
000097_HugPillow	Gối ôm	0.295	0.167	0.200
000098_TeddyBear	Gấu bông	0.879	0.882	0.872
000099_Mirror	Gương	0.638	0.643	0.664
000100_Mosquito_Racket	Vợt muỗi	0.864	1.000	0.995
000101_Sleeping_Nets	Mùng	0.895	0.855	0.977
000102_Mattress	Nệm	0.627	0.750	0.803
000103_Door	Cửa	0.965	0.846	0.900
000104_Window	Cửa sổ	0.687	0.909	0.874
000106_Toilet	Bồn cầu	0.868	1.000	0.995
000107_Toothbrush	Bàn chải	0.609	0.459	0.539
000108_Sink	Bồn rửa mặt	0.728	0.833	0.942

<u>Class</u>	<u>Tên đối tượng</u>	<u>Precision</u>	<u>Recall</u>	<u>mAP50</u>
000109_Shower	Vòi sen	0.862	0.893	0.917
000110_Faucet	Vòi nước	0.903	0.727	0.877
000111_Bathtub	Bồn tắm	0.925	1.000	0.995
000112_BathSponge	Bông tắm	0.935	0.907	0.981
000113_Brushes	Bàn chải chà chân	0.625	0.750	0.795
000114_Mop	Cây lao nhà	0.959	0.846	0.911
000115_Broom	Cây chổi	1.000	0.830	0.921
000117_WaterFlosser	Tăm nước	0.962	1.000	0.995
000118_Razor	Cạo râu	0.746	0.769	0.644
000119_Comb	Lược	0.710	0.933	0.838
000120_WashingMachine	Máy giặt	0.827	0.917	0.975
000121_Gas_Stove	Bếp ga	0.686	1.000	0.914
000122_Gas_Bottle	Bình ga	0.699	0.820	0.784
000123_Induction_Cooker	Bếp từ	0.932	0.667	0.721
000124_Oil-free_fryer	Nồi chiên không dầu	0.908	0.909	0.959
000125_Pan	Chảo	0.873	0.647	0.647
000126_Kettle	Âm nước	0.817	0.917	0.882
000127_Vase	Bình thủy	0.920	0.720	0.844
000128_Electric_Cooker	Nồi cơm điện	0.684	0.909	0.729
000129_Bowl	Chén	0.497	0.263	0.305
000130_BigBowl	Tô	0.421	0.656	0.569
000131_Chopsticks	Đũa	0.790	0.631	0.722
000132_Spoon	Muỗng	0.678	0.765	0.710
000133_Plate	Dĩa	0.635	0.350	0.447

<u>Class</u>	<u>Tên đối tượng</u>	<u>Precision</u>	<u>Recall</u>	<u>mAP50</u>
000134_Pot	Nồi	0.897	0.917	0.940
000135_Knife	Dao	0.854	0.390	0.543
000136_Fork	Nĩa	0.567	0.656	0.612
000138_Fridge	Tủ lạnh	0.800	0.895	0.835
000139_Bamboo-ToothPick	Tăm tre	0.964	0.895	0.930
000140_Lid	Nắp nồi	1.000	0.546	0.908
000141_SmartPhone	Điện thoại	0.915	0.862	0.880
000142_Charger	Sạc	0.715	0.818	0.856
000143_Socket	Ổ điện	0.854	0.909	0.906
000144_Laptop	Laptop	0.982	0.667	0.838
000145_Hammock	Võng	0.833	0.794	0.547
000146_BellPepper	Ớt chuông	0.884	0.958	0.968
000147_BottleGourd	Bầu	0.844	1.000	0.959
000148_Broccoli	Bông cải xanh	0.818	0.875	0.928
000149_Cabbage	Cải bắp	0.779	0.941	0.894
000150_Carrot	Cà rốt	0.768	0.894	0.905

II. So sánh kết quả các cách tiếp cận:

Trong nghiên cứu này, đề tài đã sử dụng các chỉ số sau để thực hiện so sánh giữa các cách tiếp cận để lựa chọn mô hình tối ưu nhất.

- **Accuracy(acc):** Độ chính xác tổng thể của mô hình, được đánh giá dựa trên tập test.

$$CT: Acc = n/N$$

Trong đó:

n: Số sample dự đoán đúng

N: Tổng số sample

- **Mean Average Precision (mAP_0.5):** Độ chính xác của mô hình, được đánh giá dựa trên tập validation.

- **Mean Average Precision (mAP_0.5:0.95):** Độ chính xác trung bình (với IoU>0.5), được đánh giá dựa trên tập validation.

Trong đó: IoU(Intersection over Unit) là độ đo overlap giữa ground truth bounding box với predicted bounding box mà mô hình dự đoán.

$$\text{IoU} \text{ được xác định bằng CT: } \text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

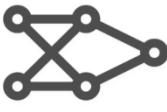
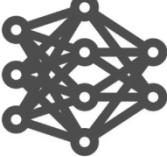
- **F1-score:** Là số dung hòa Recall và Precision giúp ta có căn cứ để lựa chọn model. F1 càng cao càng tốt.

$$\text{CT: } F1 = 2 \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

1. Mô hình yolov5s và yolov5m

Theo như những so sánh về các phiên bản của yolov5[16]:

- Về thời gian xử lý: 5s (xử lý nhanh nhất) < 5m < 5l < 5x (xử lý chậm nhất).
- Về độ chính xác: 5s (độ chính xác thấp nhất) < 5m < 5l < 5x (độ chính xác cao nhất)

			
Small YOLOv5s	Medium YOLOv5m	Large YOLOv5l	XLarge YOLOv5x
14 MB _{FP16} 2.2 ms _{V100} 36.8 mAP _{COCO}	41 MB _{FP16} 2.9 ms _{V100} 44.5 mAP _{COCO}	90 MB _{FP16} 3.8 ms _{V100} 48.1 mAP _{COCO}	168 MB _{FP16} 6.0 ms _{V100} 50.1 mAP _{COCO}

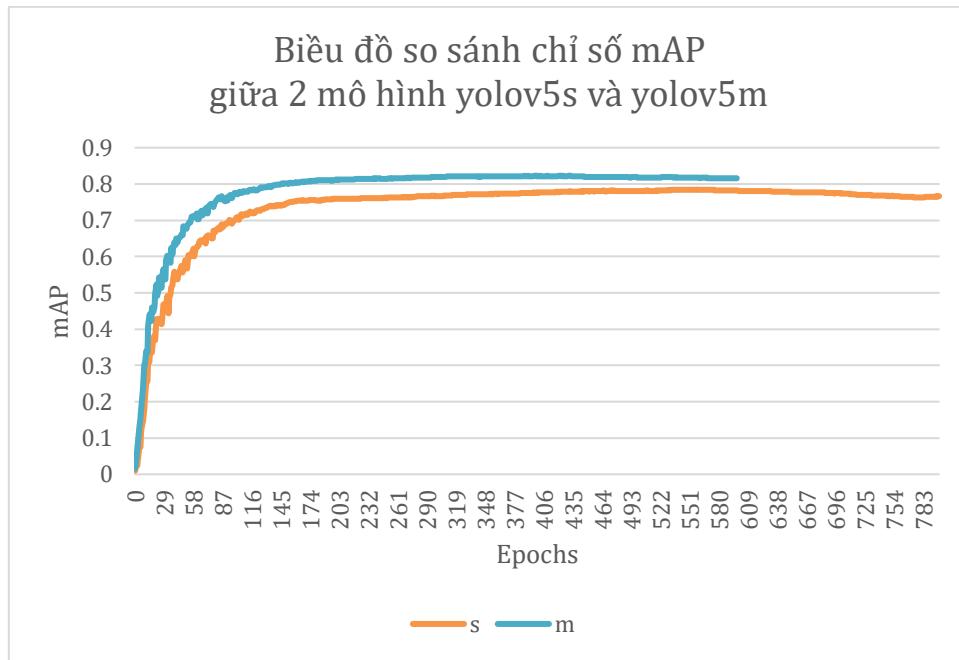
Hình 9: Các phiên bản của YOLOv5 [16]

Để chọn phiên bản phù hợp nhất cho thiết bị so sánh thời gian thực (real-time) mà mô hình có thể hoạt động. Sau khi đã thử nghiệm trên các phiên bản thì đây là real-time mà đè tài có được:

Bảng 3: Bảng so sánh thời gian phát hiện vật thể qua các phiên bản của yolov5

Phiên bản	YOLOv5s	YOLOv5m	YOLOv5l	YOLOv5x
Real-time (s)	10	10-15	20	25-30

Qua so sánh trên, thì nhóm đã chọn 2 phiên bản s và m phù hợp với thời gian thực nhận dạng được trên thiết bị. Sau đó tiến hành so sánh chúng ta tiến hành so sánh 2 phiên bản:

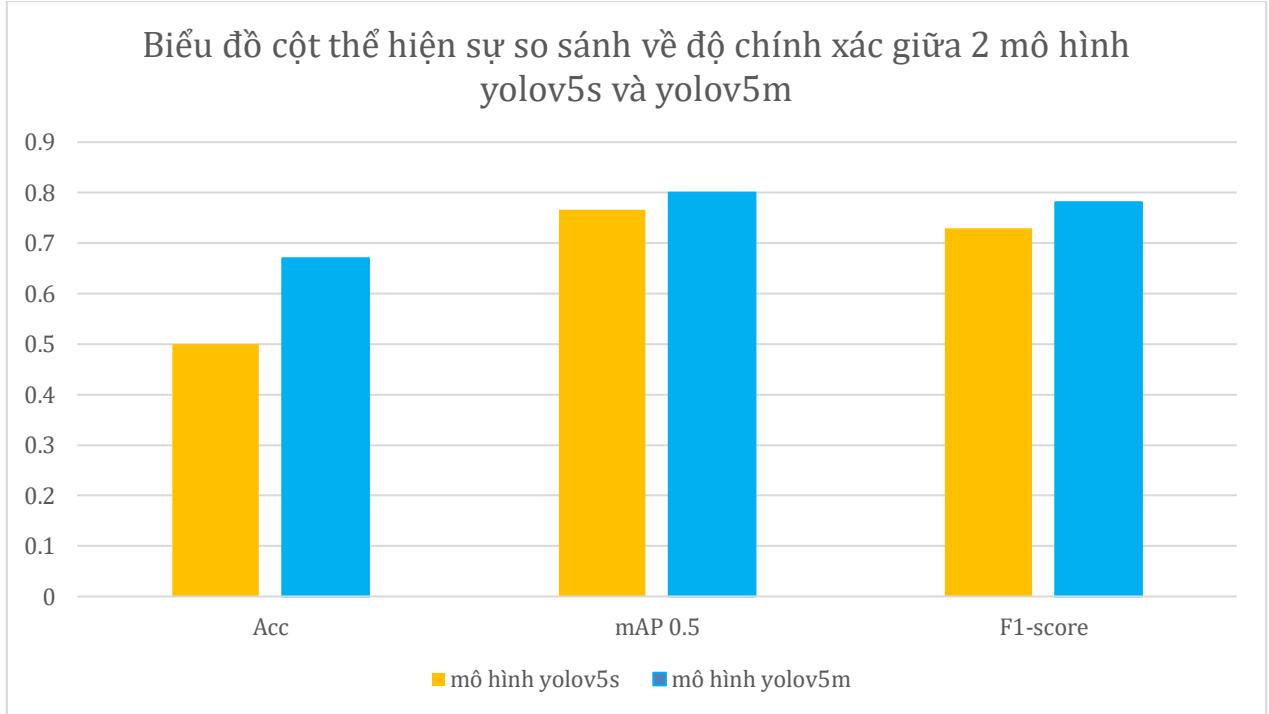


Hình 10: Biểu đồ so sánh chỉ số mAP giữa 2 mô hình yolov5s và yolov5m

Đường màu xanh đại diện cho phiên bản yolov5m và đường màu cam đại diện cho phiên bản yolov5s.

Về thời gian huấn luyện, ta thấy đường màu cam dài hơn đường màu xanh, đường màu xanh dừng lại vì mô hình yolov5m đã đạt được kết quả sớm hơn mô hình yolov5s với 600 lần học (epochs) trong khi phiên bản yolov5s phải tới tận max epoch mà vẫn chưa đạt được kết quả tốt nhất.

Về độ chính xác ta thấy đường màu cam thấp hơn đường màu xanh, điều đó thể hiện chỉ số mAP của phiên bản yolov5s thấp hơn phiên bản yolov5m trong khi thời gian đào tạo của phiên bản yolov5s là dài hơn.



Hình 11: Biểu đồ cột thể hiện sự so sánh về độ chính xác giữa 2 mô hình yolov5s và yolov5m

Ngoài chỉ số mAP, theo biểu đồ cột ở trên ta còn thấy được cột màu cam thấp hơn cột xanh cả về độ chính xác được xác định trên tập kiểm thử (Accuracy) và F1-score trên tập xác thực.

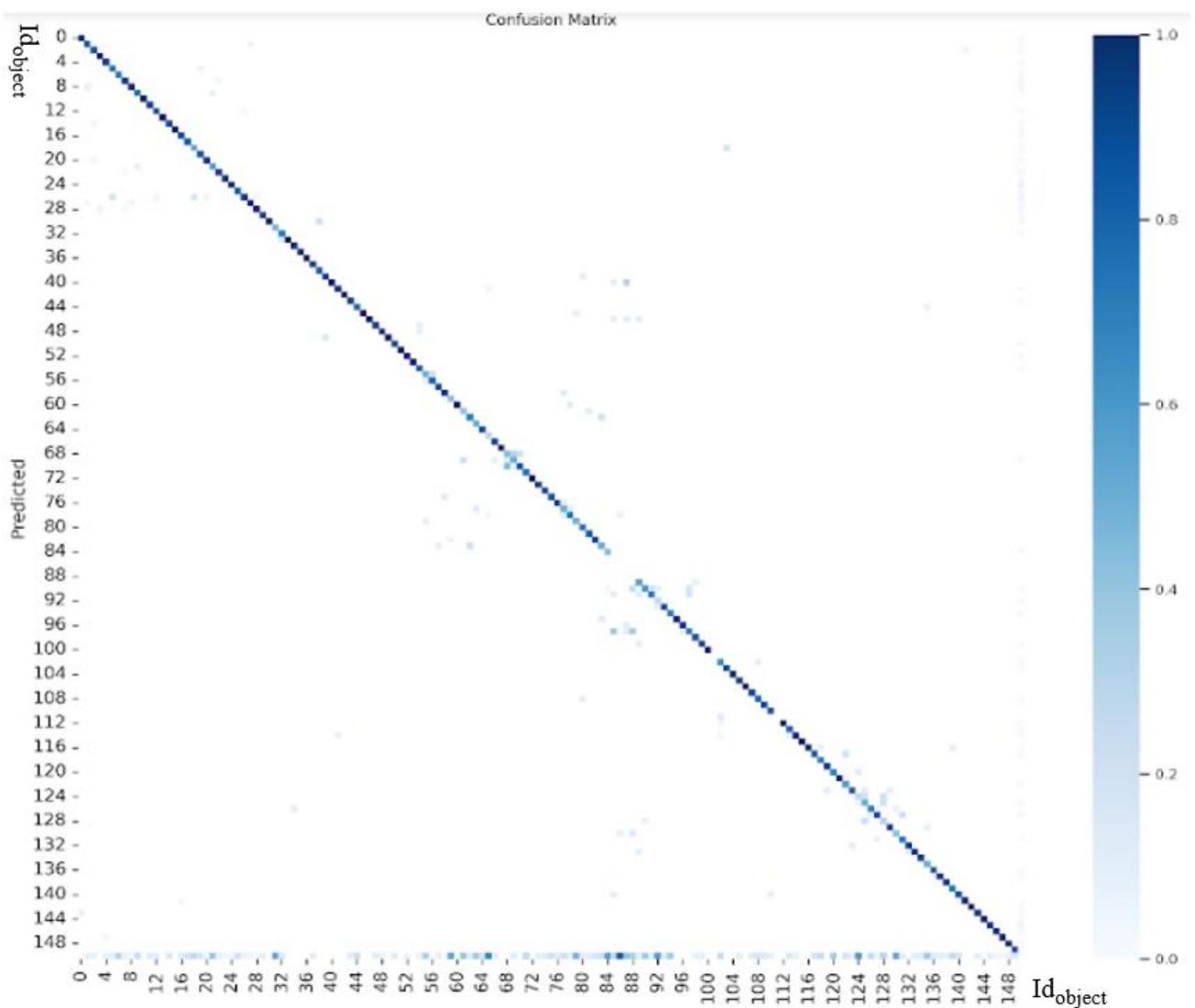
Qua các so sánh trên nhóm quyết định chọn mô hình yolov5m vì độ chính xác, thời gian huấn luyện của nó chênh lệch quá lớn so với phiên bản s mà thời gian phát hiện vật thể của nó còn gần bằng so với phiên bản s.

2. Mô hình sử dụng tăng cường dữ liệu và không sử dụng tăng cường dữ liệu

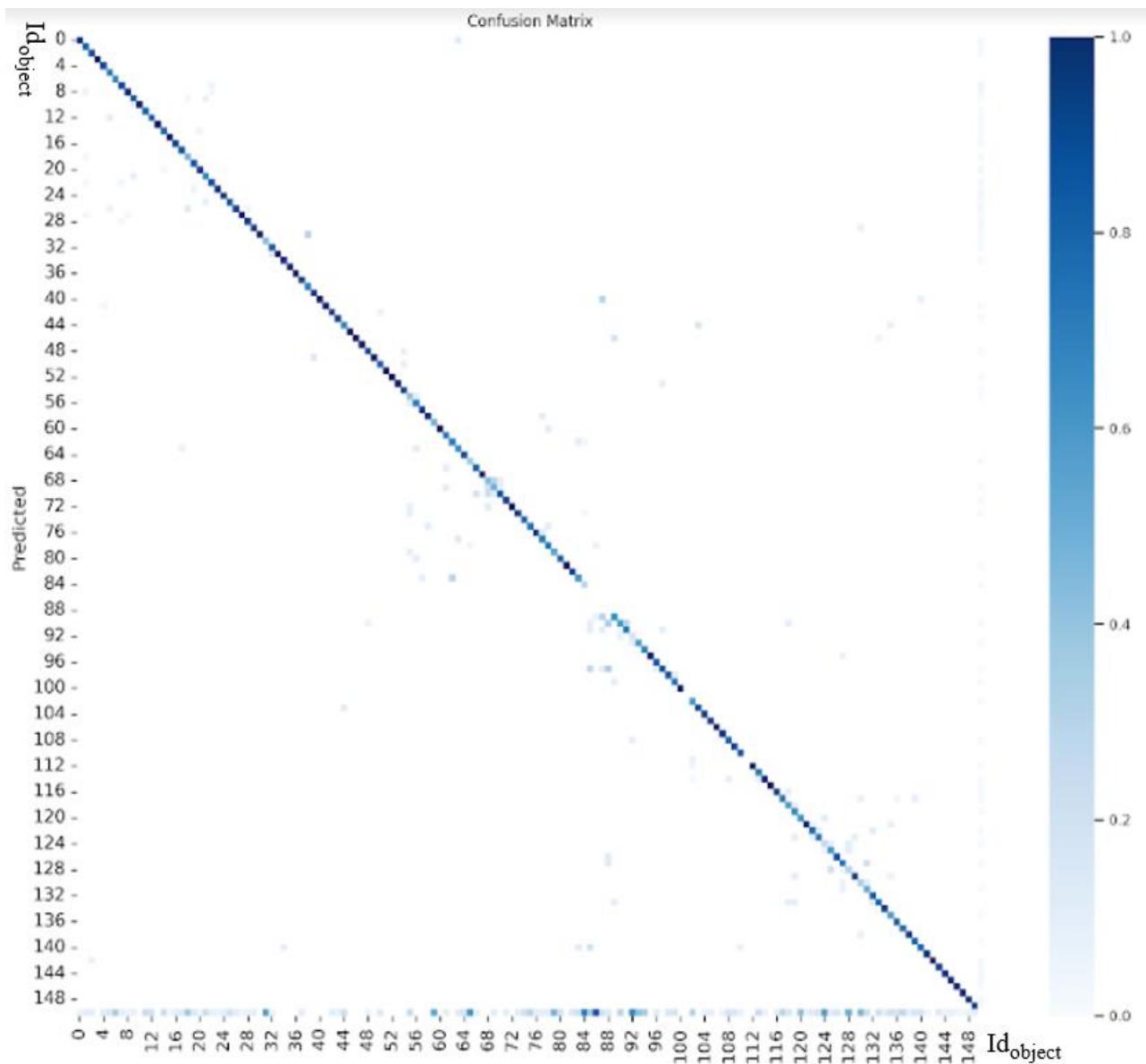
Sau khi train của mô hình Yolov5m của hai loại dataset khác nhau, một loại bình thường và loại đã trải qua tăng cường dữ liệu dựa vào dataset gốc. Kết quả của nó sẽ được đem ra so sánh nhằm có thể tối ưu khả năng của ứng dụng trong việc nhận dạng vật thể. Chỉ số dùng để so sánh hai mô hình được đề cập trong đầu phần 2. Và sau đây là kết quả so sánh được của hai mô hình:

- Accuracy:

Chỉ số về độ chính xác của hai mô hình sẽ được trình bày cụ thể hơn thông qua bảng (*) sau khi trải qua bước kiểm tra với 2175 ảnh của 145 class khác nhau như chúng ta có thể thấy độ chính xác của hai mô hình đối với từng class đa số là khác nhau, tuy nhiên chỉ số của chúng gần như là không có chênh lệch gì quá cao, so sánh về confusion matrix của hai mô hình tại hình 12 và hình 13.

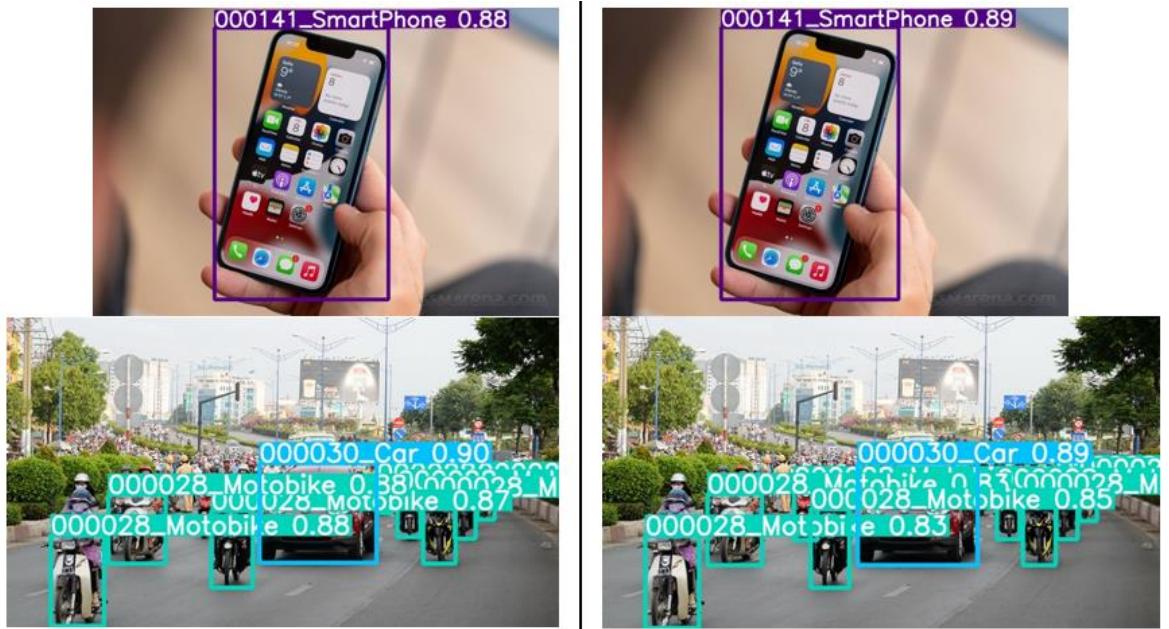


Hình 12: Confusion matrix của mô hình không sử dụng tăng cường dữ liệu



Hình 13: Confusion matrix của mô hình sử dụng tăng cường dữ liệu

Khả năng dò được nhiều đối tượng trong một hình cũng được đề tài lưu ý, khi kiểm tra khả năng trên của hai mô hình, nhận thấy rằng, khả năng dò được nhiều class trong một tấm ảnh của mô hình không sử dụng tăng cường dữ liệu không được tốt so với mô hình có sử dụng tăng cường dữ liệu, độ chính xác mô hình không sử dụng phương pháp tăng cường dữ liệu khi dò các hình xuất hiện nhiều vật thể khác nhau là không tốt nếu so với mô hình còn lại, tuy nhiên khi dò các tấm hình có vật thể đơn độc thì mô hình không sử dụng phương pháp tăng cường dữ liệu lại tốt hơn, dù chênh lệch không cao. Ví dụ ở hình 14



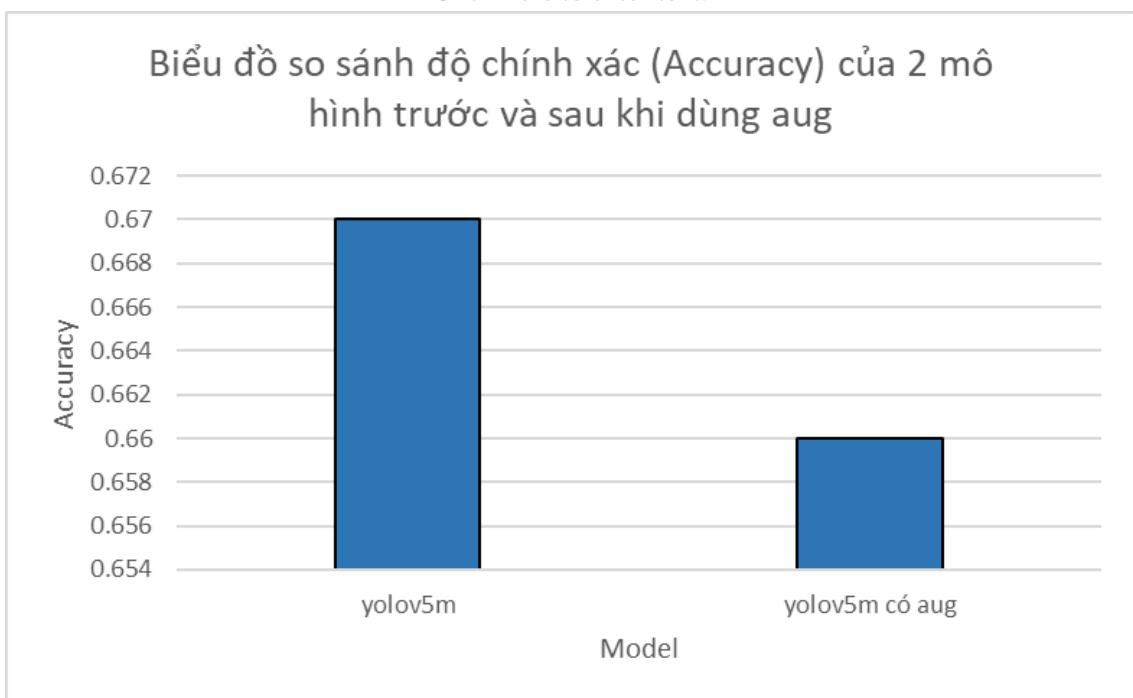
Yolov5m data augmentation

Yolov5m

Hình 14: Hình ảnh nhận dạng được trước và sau khi sử dụng tăng cường dữ liệu

So sánh về độ chính xác(Acc) của hai mô hình được trình bày ở hình 15

Click here to enter text.

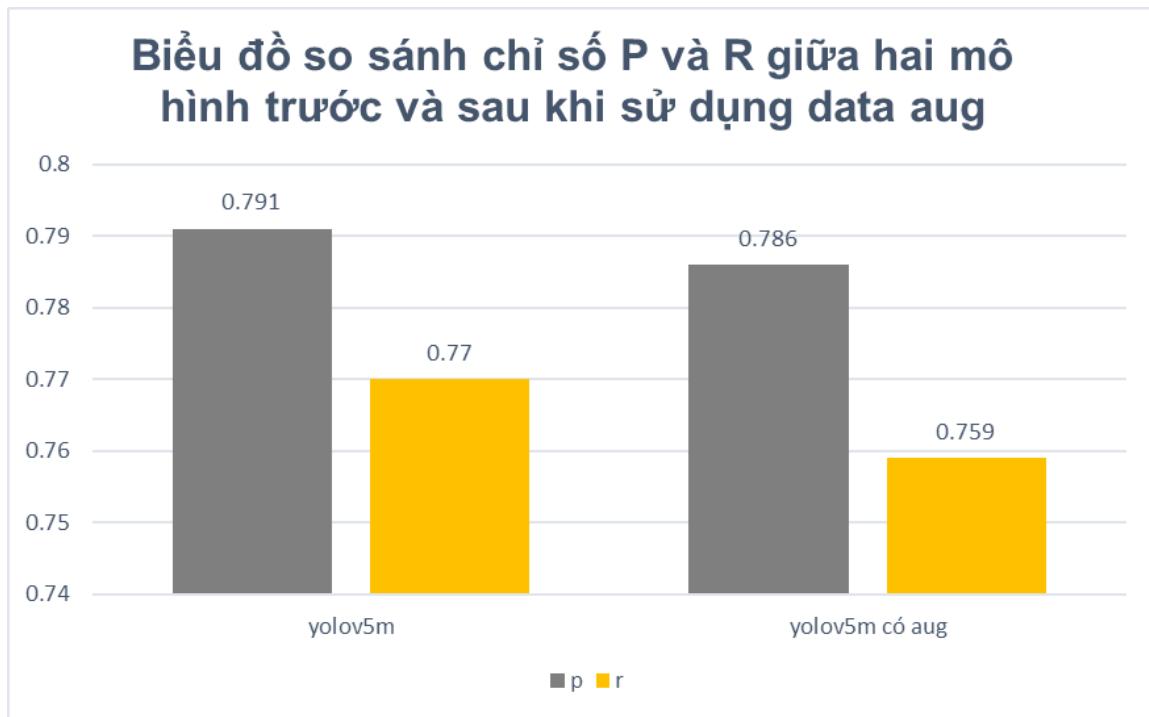


Hình 15: Biểu đồ cột so sánh độ chính xác của hai mô hình

- F1-score

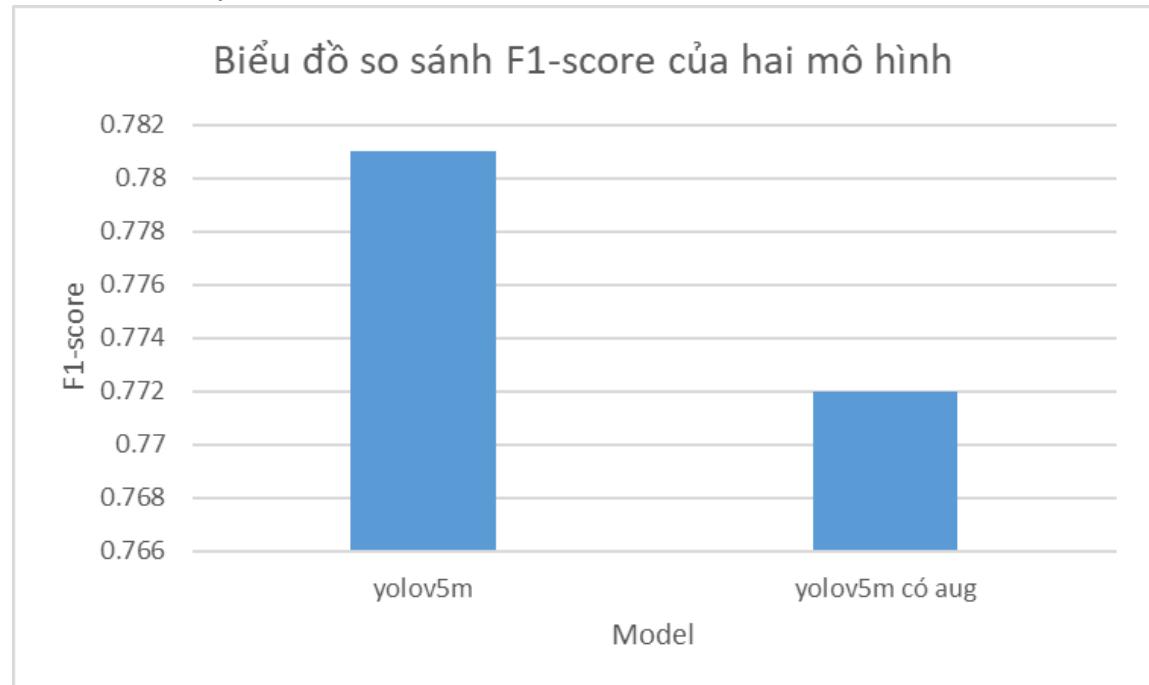
Chỉ số về F1-score của hai mô hình sau khi train sẽ được trình bày rõ ràng sau khi ta trình bày chi tiết hơn về số precision và recall của hai mô hình

So sánh chỉ số precision và recall so sánh của hai model được trình bày ở hình 16 [Click here to enter text](#).



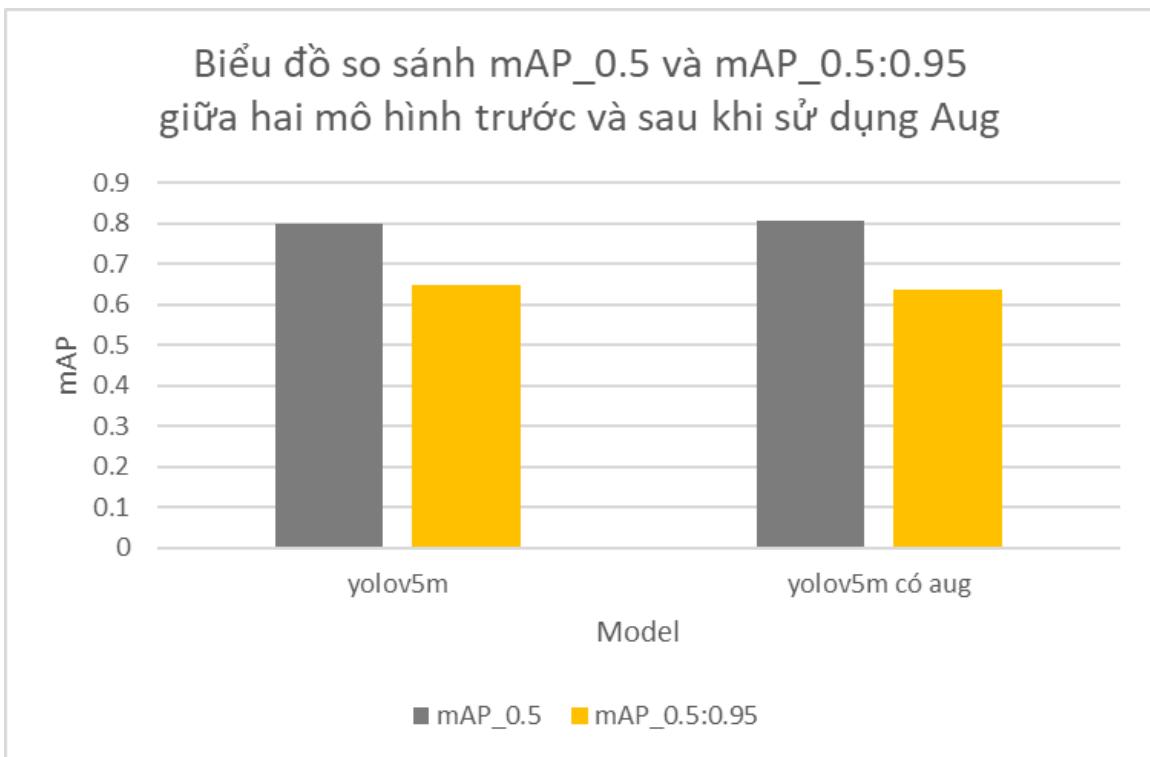
Hình 16: Biểu đồ cột so sánh Precision và Recall của hai mô hình

Từ số liệu Precision và Recall từ hình trên ta có thể tính được F1-score. So sánh F1-score được trình bày ở hình



Hình 17: Biểu đồ cột so sánh F1-score của hai mô hình

- mAP
Chỉ số về mAP của hai mô hình thì sẽ được trình bày ở cả hai loại là chỉ số mAP 0.5 và chỉ số mAP 0.5-0.95, so sánh trình bày ở hình



Hình 18: Biểu đồ cột so sánh mAP_0.5 và mAP_0.5:0.95 của hai mô hình

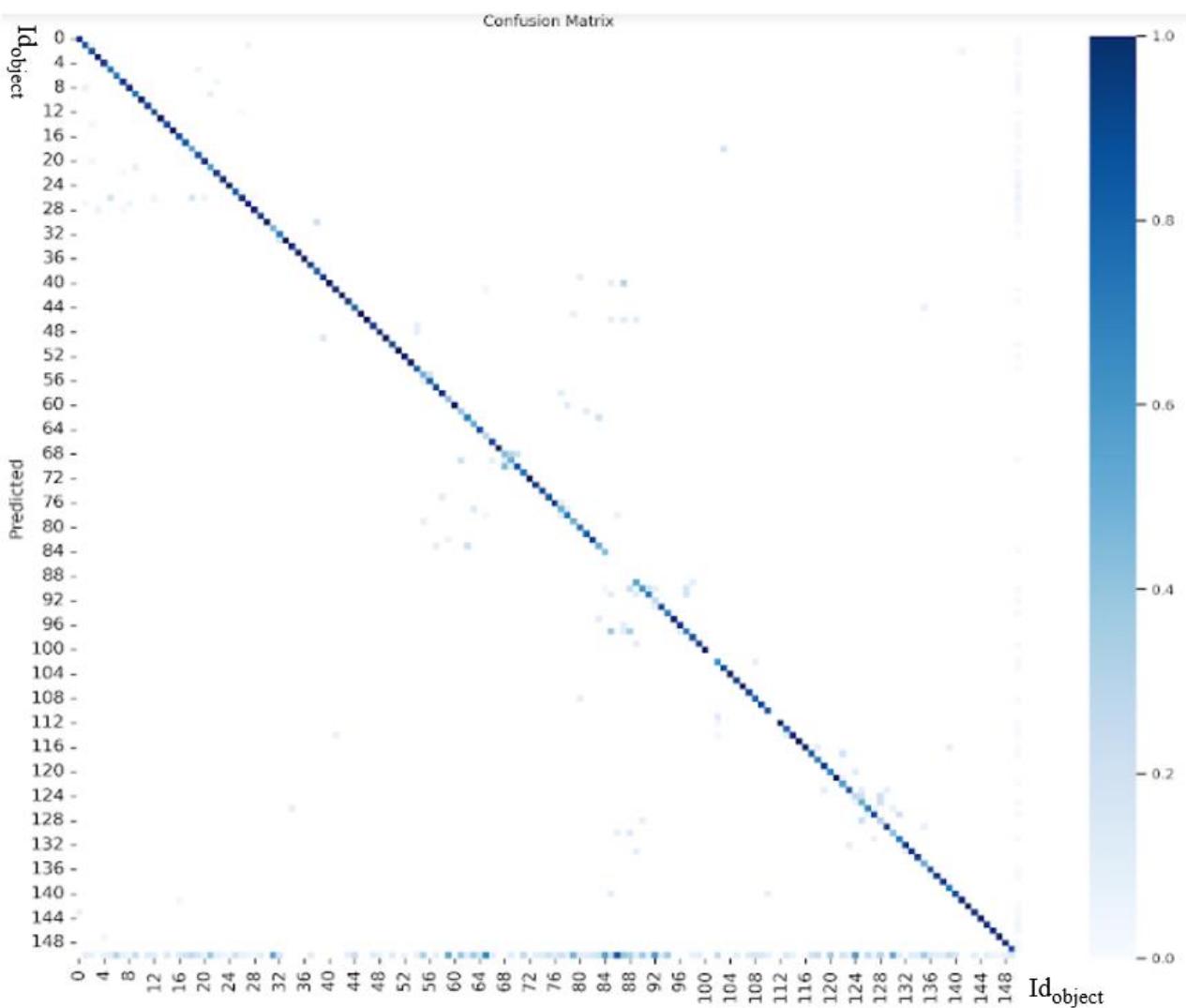
3. Tăng số lớp và số ảnh cho mô hình đã đào tạo

Sau khi mô hình hoàn thành, nhóm tiếp tục thêm 2 lớp khác vào mô hình được chọn là tốt nhất sau khi đã đánh giá ở 2 phần trước đó (mô hình yolov5m không sử dụng tăng cường dữ liệu)

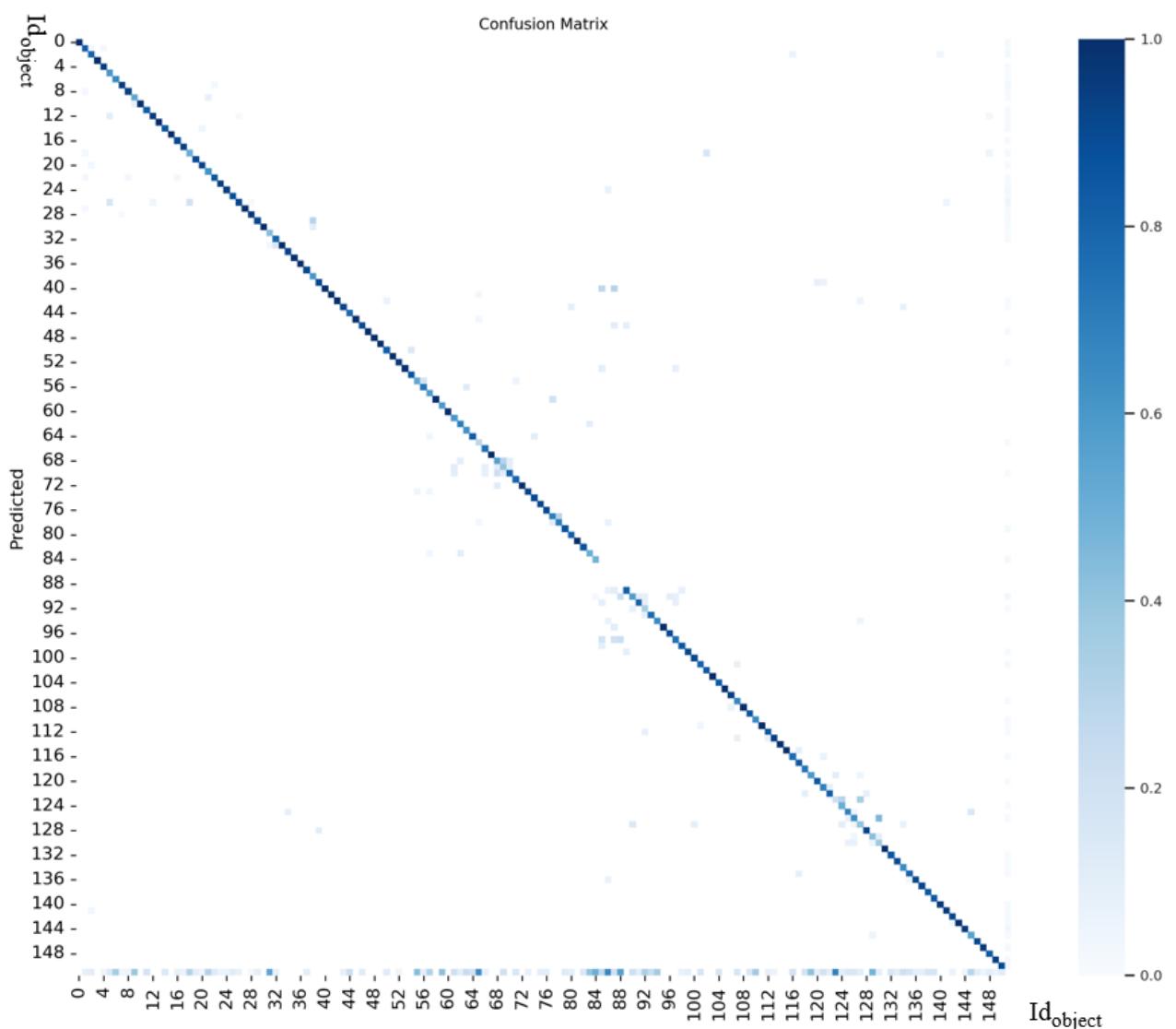
Hai lớp được thêm là:

- 000151_Pen :cây bút
- 000152_LipStick :son môi

Và mô hình dường như không hề bị ảnh hưởng bởi các lớp mới, có thể nhận thấy điều đó qua confusion matrix của mô hình ban đầu (hình 19) và confusion matrix của mô hình sau khi đã thêm lớp mới (hình 20)

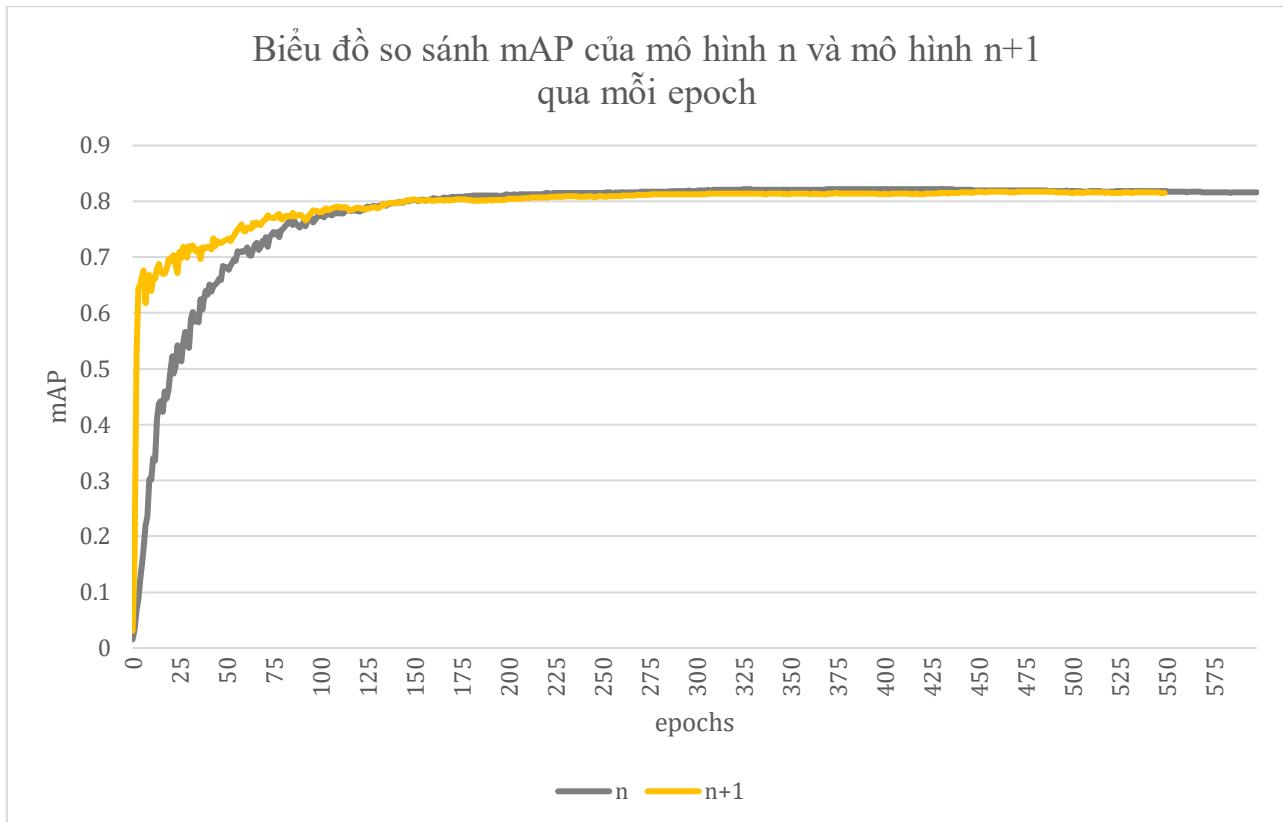


Hình 19: Confusion matrix của mô hình ban đầu



Hình 20: Confusion matrix của mô hình sau khi thêm lớp mới

Ngoài ra, mô hình còn hoàn thành huấn luyện và đạt kết quả tối đa trong thời gian ngắn hơn với số 550 epochs so với việc huấn luyện mô hình ban đầu cần tới tận 600 epochs để đạt được tối đa (hình 21)



Hình 21: Biểu đồ so sánh chỉ số mAP giữa mô hình ban đầu và mô hình sau khi thêm lớp mới

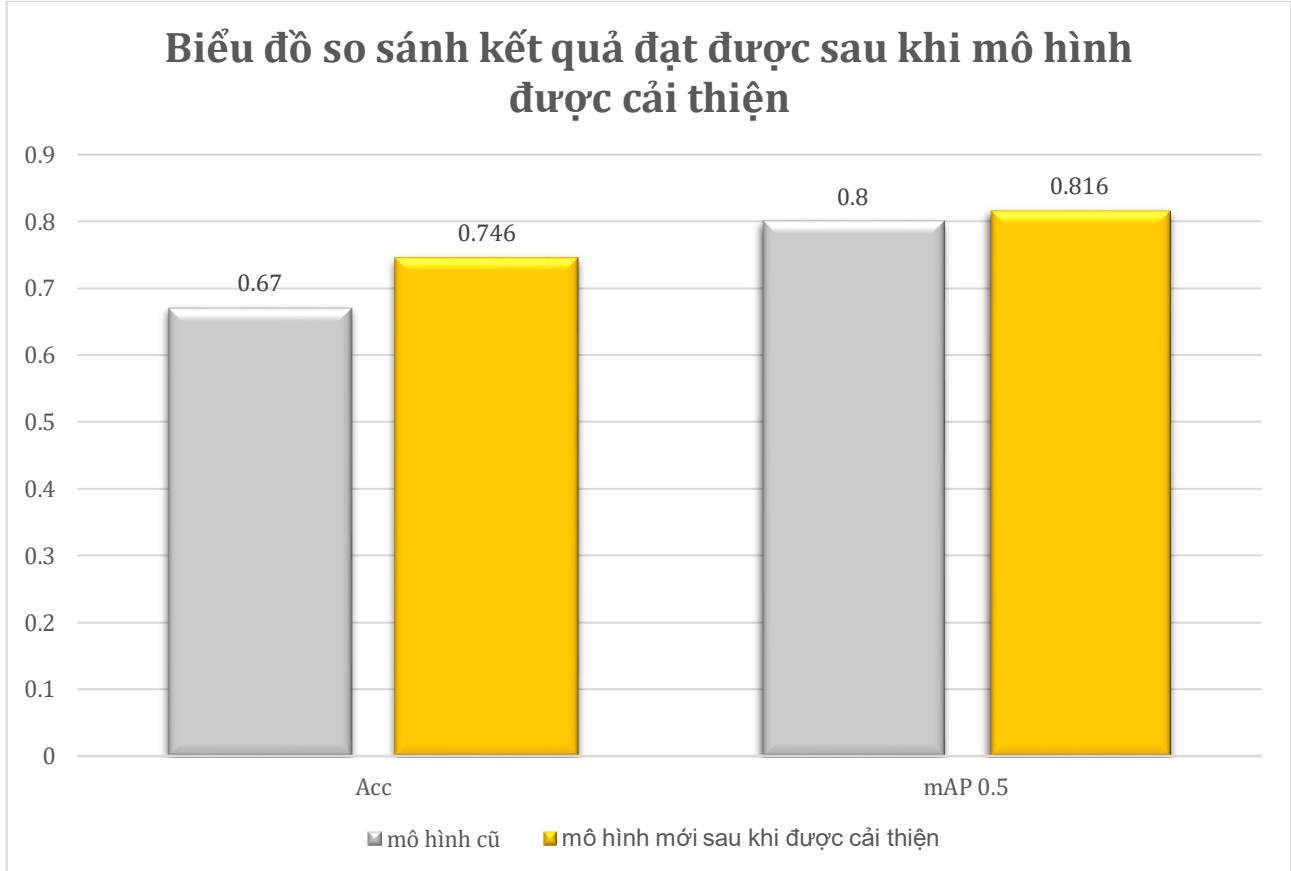
Trong đó:

n: là mô hình ban đầu

n+1: là mô hình sau khi tăng thêm lớp mới

Qua biểu đồ trên, ngoài đạt được kết quả tốt nhất nhanh hơn mô hình ban đầu (cố 50 epochs), mô hình n+1 còn hội tụ sớm hơn mô hình n ở những lần học đầu tiên của mô hình.

Và trong khi huấn luyện mô hình mới nhóm đề tài cũng đã cải thiện mô hình cũ bằng cách tăng số ảnh cho các lớp thấp và sửa một số lỗi trong khi gắn nhãn và kết quả đạt được là:



Hình 22: Biểu đồ so sánh kết quả đạt được sau khi mô hình được cải thiện

Qua biểu đồ trên, ta thấy sự thay đổi tích cực rõ rệt cả về độ chính xác (accuracy) trên tập test và chỉ số mAP được đánh giá trên tập validation.

Điều này chứng tỏ mô hình không chỉ có thể tăng thêm đối tượng còn có thể cải thiện độ chính xác để đạt được mô hình tốt nhất.

III. Ứng dụng Android:

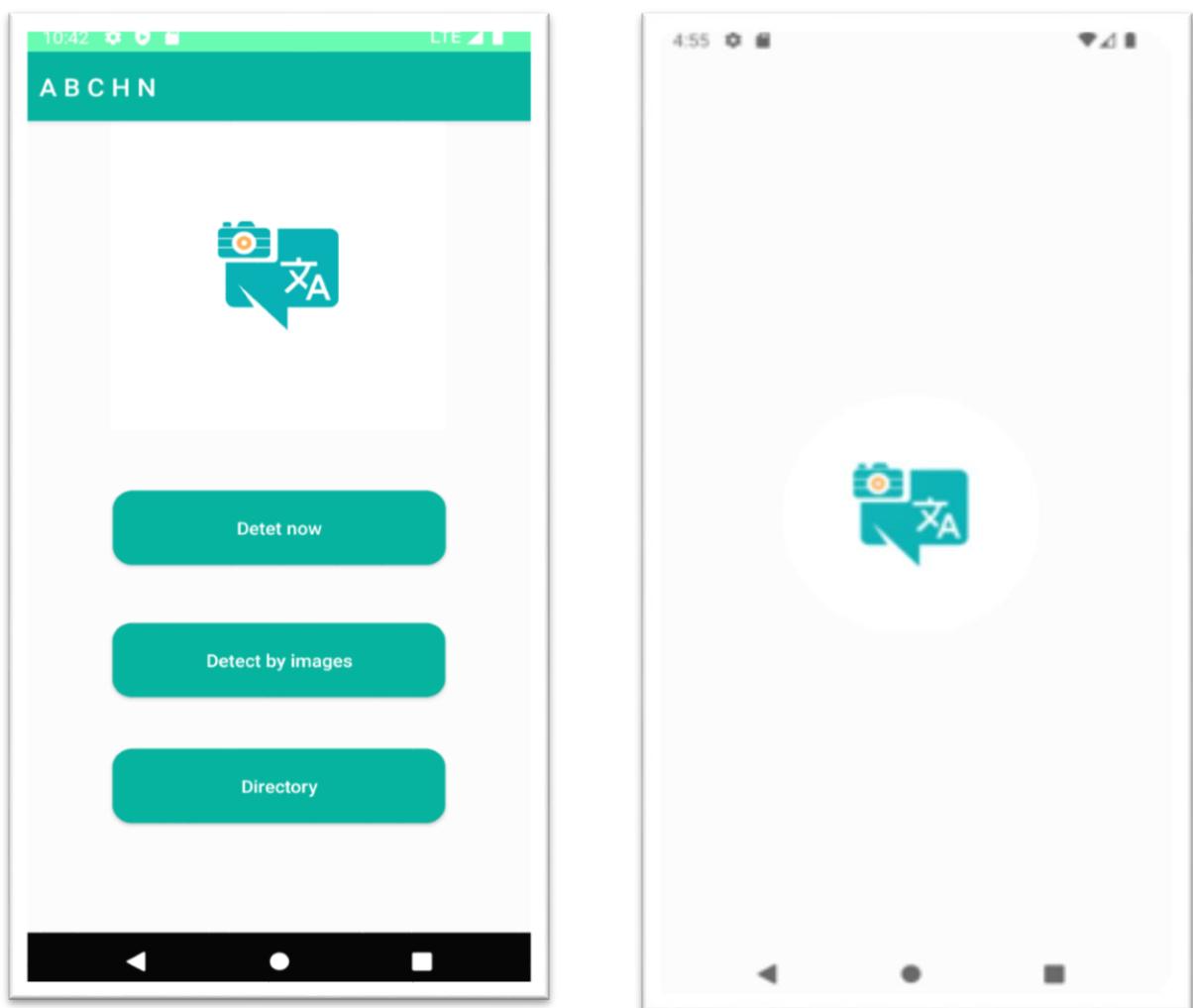
1. Giao diện màn hình chính

Khi khởi động ứng dụng, giao diện Splash Screen sẽ hiển thị trong thoáng chốc và giao diện chính của ứng dụng sẽ xuất hiện. Ở giao diện chính sẽ hiển thị 3 button tương ứng với 3 chức năng sau:

Button “Detect now”: chức năng nhận dạng trực tiếp đối tượng có trong video stream (real-time).

Button “Detect by images”: chức năng nhận dạng hình ảnh chụp từ camera hoặc hình ảnh có sẵn trong máy.

Button “Directory” : chức năng từ điển.



Hình 24: Giao diện chính (trái) và Splash screen (phải)

2. Giao diện nhận dạng trực tiếp (Real-time)

Để tiến hành nhận dạng đồ vật một cách trực tiếp (video steam), người dùng chọn vào chức năng Detect now. App sẽ tiến hành nhận dạng trực tiếp đồ vật xuất hiện trong camera như hình bên dưới:

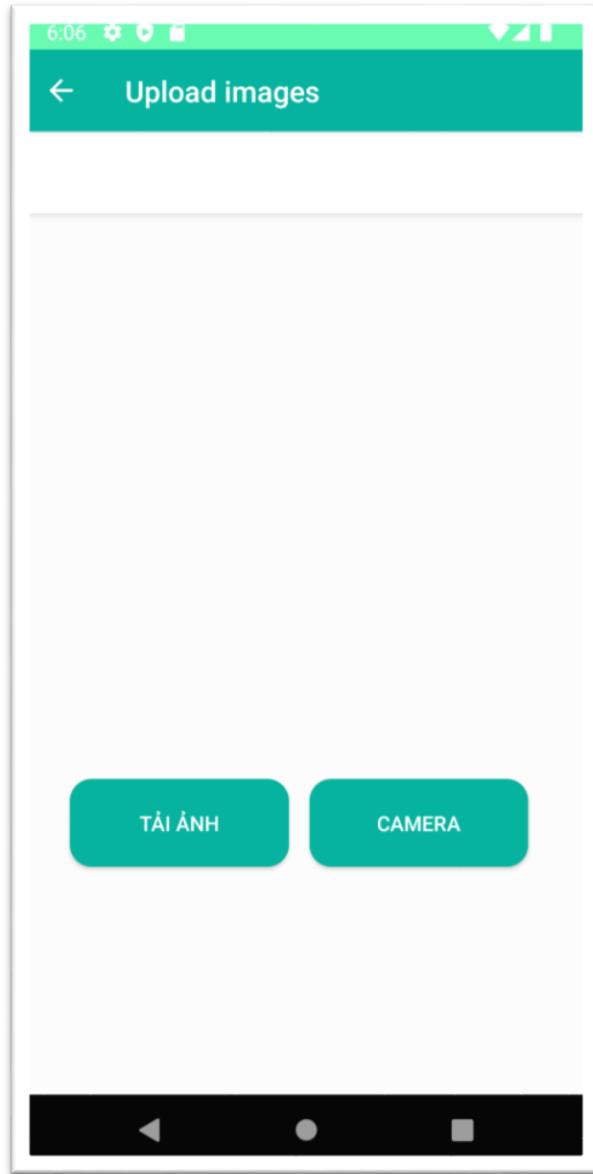


Hình 25: Giao diện nhận dạng trực tiếp (real-time)

Đối tượng xuất hiện trên màn hình sẽ được nhận dạng và trên màn hình sẽ xuất hiện một khung bao xung quanh đối tượng đó, trên khung sẽ có tên tiếng Anh của đồ vật được nhận dạng. Ở hình trên, đối tượng đã được nhận dạng là chiếc máy quạt và tên tiếng Anh của đối tượng là Fan.

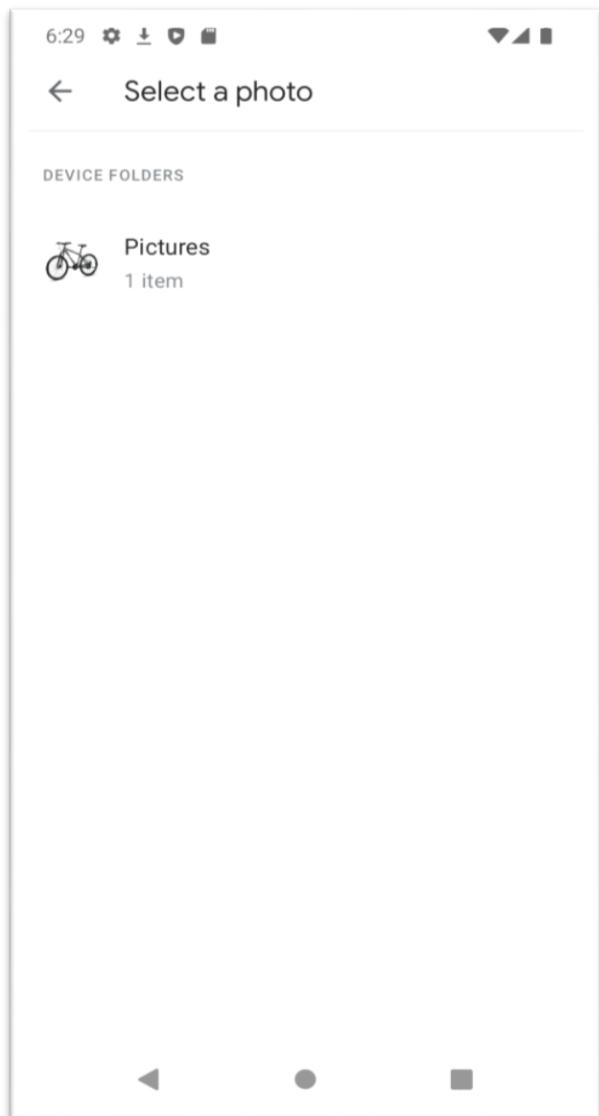
3. Giao diện nhận dạng bằng hình ảnh chụp từ camera hoặc hình ảnh có sẵn trong máy.

Để thực hiện nhận dạng bằng hình ảnh chụp từ camera, người dùng chọn chức năng “ Detect by image”, giao diện nhận dạng bằng hình ảnh sẽ xuất hiện:



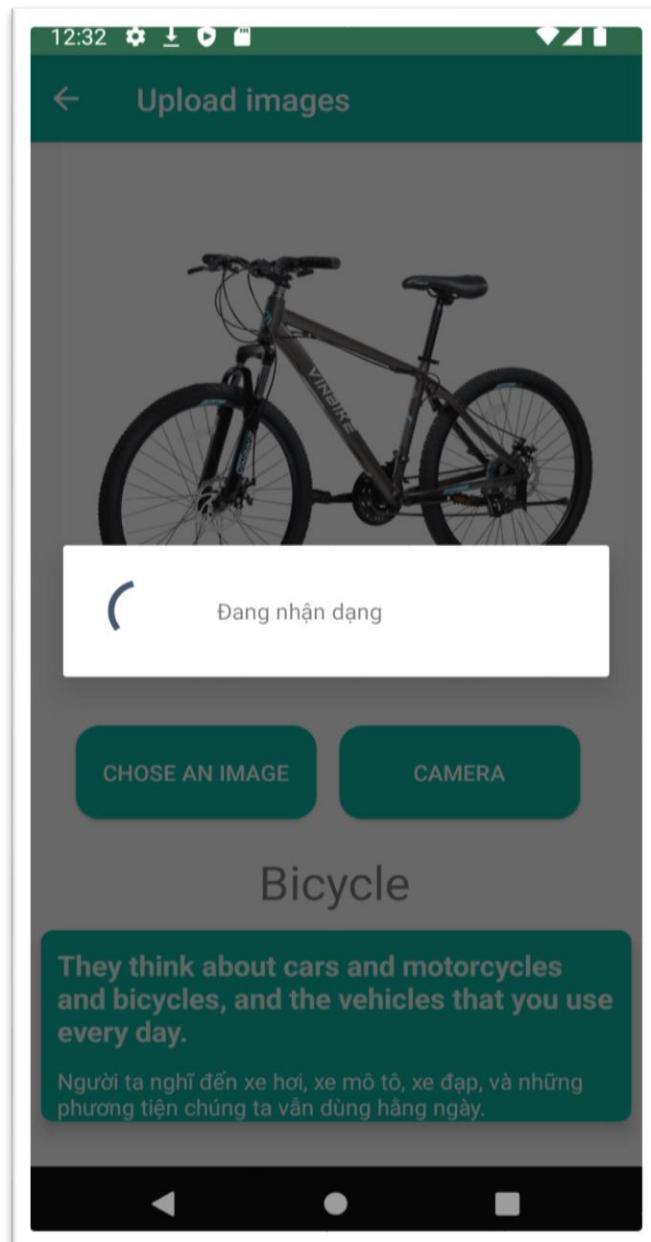
Hình 26: Giao diện chức năng “Detect by image”

Người dùng có thể chọn nhận dạng bằng hình ảnh từ thư viện ảnh trong thiết bị bằng cách nhấn vào nút “Tải ảnh”:



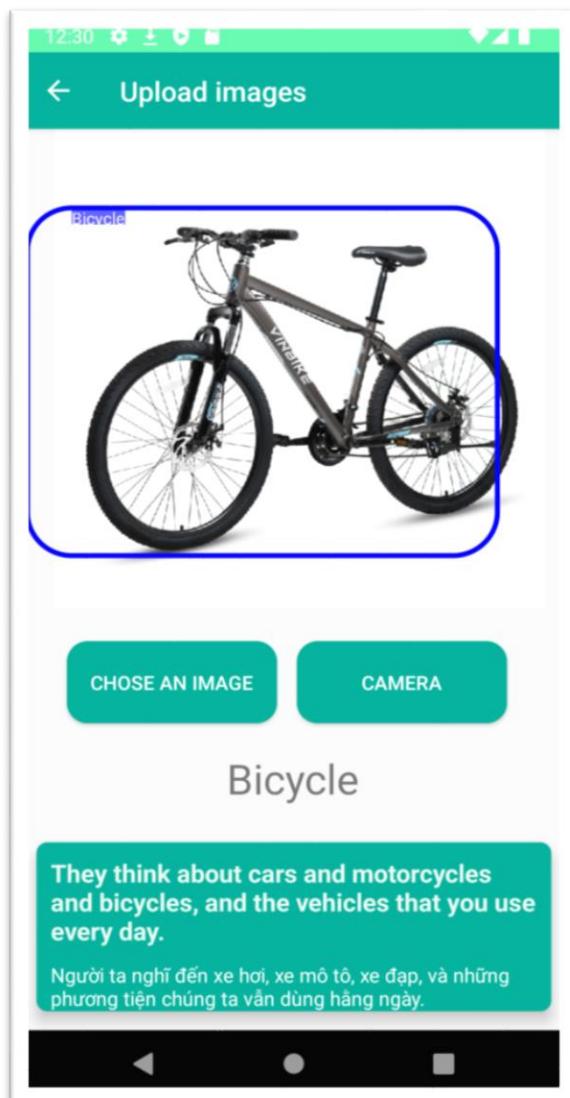
Hình 27: Chọn ảnh từ thư viện

Sau khi người dùng chọn hình để nhận dạng, ứng dụng tiến hành nhận dạng:



Hình 28: Tiến hành nhận dạng

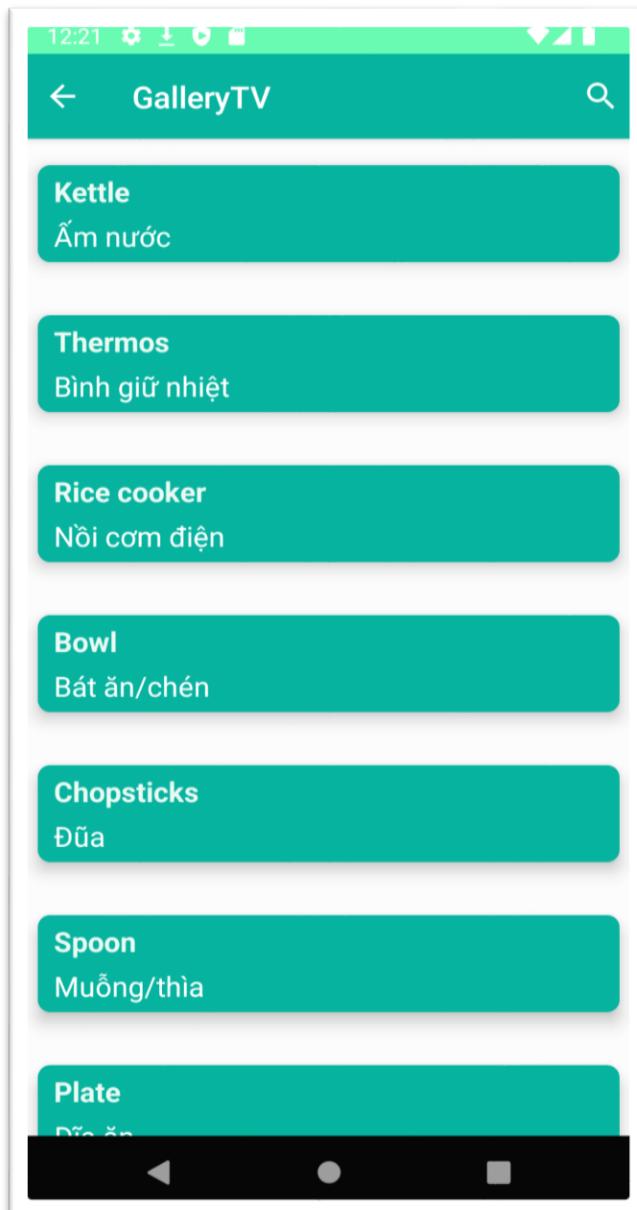
Sau khi nhận dạng thành công, sẽ xuất hiện một khung bao quanh đối tượng và tên tiếng Anh của đối tượng sẽ được hiện ở trên khung và ở bên dưới hình, kèm theo là một số câu ví dụ bằng tiếng Anh kèm theo nghĩa tiếng Việt ở bên dưới của đối tượng đó:



Hình 29: Đã nhận dạng thành công

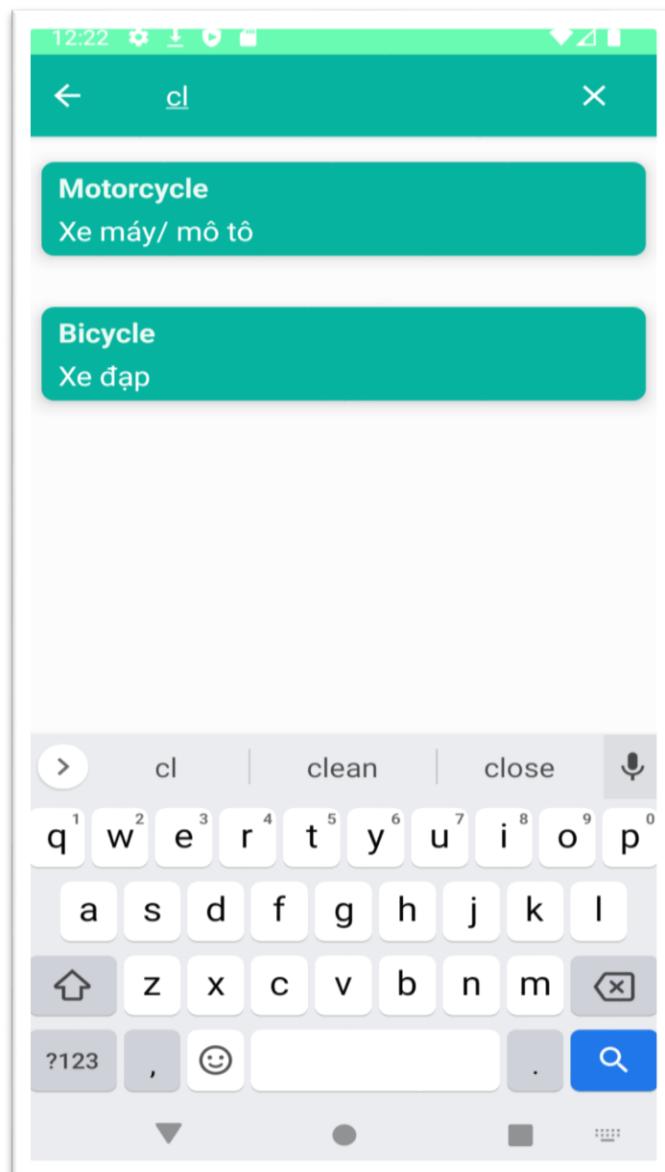
4. Giao diện từ điển

Người dùng có nhu cầu tra cứu từ vựng có thể vào thư viện để tra cứu, để sử dụng chức này, người dùng chọn chức năng “Directory” ở màn hình chính. Giao diện của thư viện sẽ xuất hiện:



Hình 30: Giao diện thư viện từ vựng

Để tra thông tin một đối tượng nào đó, người dùng có thể tìm kiếm từ vựng muốn tra cứu trên thanh tìm kiếm, ngoài ra khi người dùng bấm vào từ vựng, một đoạn audio giải thích về từ vựng đó sẽ được phát ra:



Hình 31: Chức năng tìm kiếm

Khi bấm vào “Bicycle” một đoạn audio định nghĩa giải thích về từ vựng đó bằng tiếng Anh sẽ phát ra hoặc người dùng muốn định nghĩa bằng tiếng Việt, chỉ cần bấm vào “Xe đạp”, đoạn audio định nghĩa bằng tiếng Việt sẽ tự động phát.

PHẦN 4: KẾT LUẬN VÀ KIẾN NGHỊ

I. KẾT LUẬN

Với sự phát triển không ngừng của xã hội, trí tuệ nhân tạo ngày càng được nhiều người quan tâm đến và dần trở nên phổ biến, thấy được tình hình đó đề tài đã tìm hiểu và cho ra đời sản phẩm “Ứng dụng hỗ trợ học tiếng anh qua video” kết hợp với hỗ trợ học tiếng anh cho mọi đối tượng sản phẩm càng chứng minh được tính hữu ích của mình khi được đưa vào đời sống để mọi người cùng sử dụng.

Trong nghiên cứu lần này, nhóm được tìm hiểu kiến thức về Deep Learning, các cách xử lý ảnh, các thuật toán giúp phát hiện đối tượng và nhận dạng vật thể, cách xây dựng mô hình thuật toán YOLOv5 và ứng dụng hoàn chỉnh trên nền tảng Android.

Sau khi chọn lọc, đề tài đưa ra mô hình phù hợp nhất là mô hình YOLOv5 phiên bản m với khả năng nhận dạng cao và thời gian nhận dạng vừa phải và nhóm cũng đã chứng minh được rằng mô hình hoàn toàn có thể phát triển thêm khi tăng được số đối tượng mới và có thể cải thiện độ chính xác cho từng đối tượng bằng cách tăng số ảnh vào mô hình.

Kết hợp với hợp với bộ từ điển gồm 147 vật thể khác nhau, ứng dụng còn có thể dịch sang ngữ nghĩa tiếng Việt và đặt câu với từ vừa kiểm được

II. KẾT QUẢ ĐÓNG GÓP

Học tiếng anh qua việc nhận đối tượng trong video cung cấp cho chúng ta một ứng dụng học tập kiểu mới, kích thích sự tìm tòi học hỏi đối với những đối tượng thân thuộc xung quanh bản thân chúng ta. Nhờ sự tiến bộ của máy học kết hợp với mật độ điện thoại di động được sử dụng rộng rãi như ngày nay thì việc học tập từ vựng tiếng anh lại càng trở nên dễ dàng hơn bao giờ hết.

Trong nghiên cứu này đã sử dụng YOLOv5 để huấn luyện mô hình các đối tượng quen thuộc gần gũi với đời sống từ các đồ vật như là bàn, ghế cho tới các động vật như chó, mèo, gà, vịt. Tất cả đều được nhận dạng bằng từ vựng tiếng anh, nhưng bên cạnh đó đề tài đã hỗ trợ thêm các nghĩa tiếng việt bổ sung, cũng như là có thêm ví dụ, câu minh họa, định nghĩa cho từng đối tượng cụ thể. Điều đó là cơ sở để thúc đẩy việc tìm ra nhiều phương pháp mới để hỗ trợ học tiếng anh hơn nữa mà không làm các học viên cảm thấy chán nản, không đem lại nhiều kết quả tốt như cách học truyền thống.

Trong quá trình chuẩn bị cho mô hình máy học, nhóm đề tài đã thu thập được bộ dữ liệu gồm 7467 ảnh và đã được xử lí, gắn khung nhãn với 147 đối tượng khác nhau xuất hiện phổ biến ở thành phố Cần Thơ, đóng góp vào phần mô hình giúp ích cho người sử dụng khác.

Ngoài ra khi nghiên cứu phần mô hình, đề tài còn phát hiện được khả năng học hỏi thêm của mô hình sau khi thêm hai lớp mới và tiến hành huấn luyện tiếp tục. Kết quả cho thấy được huấn luyện thêm lần này tốn ít thời gian hơn huấn luyện mô hình ban đầu. Và kết quả cho thấy không những bổ sung thêm khả năng nhận dạng được lớp mới mà còn cải thiện được nhận dạng lớp cũ khi tăng thêm số ảnh mới vào các lớp đã huấn luyện trước đó.

Ứng dụng trên điện thoại di động được thiết kế thân thiện với người dùng, dễ sử dụng, thời gian phát hiện và nhận dạng đối tượng ngắn từ 10 đến 15 giây, bên cạnh đó còn có thư viện đầy đủ cho 145 lớp, do đó phù hợp với nhu cầu người dùng.

III. HƯỚNG PHÁT TRIỂN

Với khả năng phát triển của mô hình, việc gia tăng số lượng lớp mới trong tương lai là khả thi nhằm tạo ra bộ từ điển đa dạng hơn kết hợp với một số chức năng hỗ trợ học tiếng anh như bổ sung về ngữ pháp, đặt câu cũng được tăng lên làm cho vốn từ vựng trong ứng dụng trở nên phong phú đa dạng hơn, giúp người dùng có thêm nhiều phương pháp tiếp cận để chinh phục được vốn từ vựng sâu rộng trong ngữ pháp. Ngoài tiếng anh ra thì có thể cài tiến mô hình hỗ trợ nhiều loại ngôn ngữ khác nhau như tiếng nhật, tiếng hàn, tiếng trung, ... các loại ngôn ngữ khác nhau cho người dùng tha hồ mà lựa chọn loại ngôn ngữ phù hợp mà bản thân cần trong tương lai.

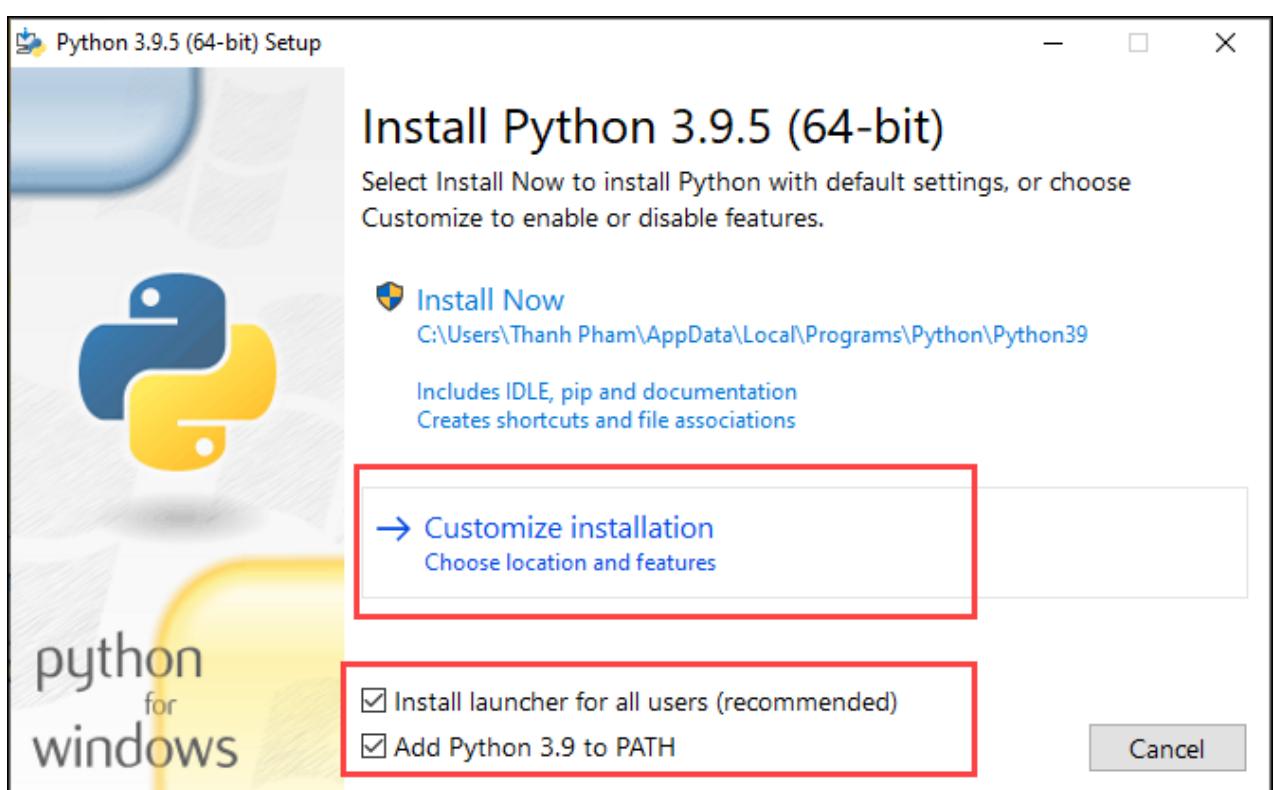
TÀI LIỆU THAM KHẢO

- [1] “Ứng dụng AI trong một số lĩnh vực tại Việt Nam và đề xuất”.
<https://ictvietnam.vn/ung-dung-ai-trong-trong-mot-so-linh-vuc-tai-viet-nam-va-de-xuat-20211106091652131.htm>
- [2] “Emerging Technology Research & Trends - Think with Google APAC”, *Think with Google*. <https://www.thinkwithgoogle.com/intl/en-apac/future-of-marketing/emerging-technology/>
- [3] “Trang chủ”, *PHX Smart School*. <https://www.phx-smartschool.com/>
- [4] “Python (ngôn ngữ lập trình)”
- [5] “Ngôn ngữ lập trình Java: Khái niệm, Đặc điểm và Ứng dụng”.
<https://magenest.com/vi/ngon-ngu-lap-trinh-java>
- [6] “Java (ngôn ngữ lập trình)”
- [7] “TensorFlow là gì? Tìm hiểu chi tiết về TensorFlow”, 26 Tháng Bảy 2021.
<https://wiki.tino.org/tensorflow-la-gi>
- [8] “Học máy di động bằng cách sử dụng TensorFlow Lite [Video]”.
<https://helpex.vn/article/hoc-may-di-dong-bang-cach-su-dung-tensorflow-lite-video>
- [9] “Android TensorFlow Lite Machine Learning Example”, 8 Tháng Năm 2018.
<https://viblo.asia/p/android-tensorflow-lite-machine-learning-example-924IJxXmKPM>
- [10] “Android Studio là gì? Hướng dẫn sử dụng Android Studio”.
<https://itnavi.com.vn/blog/android-studio-la-gi>
- [11] “Android Studio là gì? Chạy nó lại khiến 1 chiếc laptop bị đơ”.
<https://ghiencongnghe.info/android-studio-la-gi.html>
- [12] “Android Studio là gì? Cách sử dụng ra sao? - Daotaotester.com”, *Khóa học Tester*, 28 Tháng Sáu 2021. <https://daotaotester.com/android-studio-la-gi/>
- [13] “Deep Learning là gì? Tìm hiểu về Deep Learning từ A-Z”. <https://vietnix.vn/deep-learning-la-gi/>
- [14] “ultralytics/yolov5”. Ultralytics, 21 Tháng Mười-Một 2022. Truy cập: 21 Tháng Mười-Một 2022. [Online]. Available at:
<https://github.com/ultralytics/yolov5/blob/72cad39854a7d9ebbd4d58994cefa966b0da8fc1/README.md>
- [15] “<https://www.mdpi.com/2072-4292/13/9/1619>”.
- [16] “Series YOLO: #5 YOLOv5: Fast, compact, lightweight and easy to use. - DevAI”.
<https://devai.info/2021/09/01/series-yolo-5-yolov5-nhanh-gon-nhe-va-de-su-dung/>

PHẦN 5: PHỤ LỤC

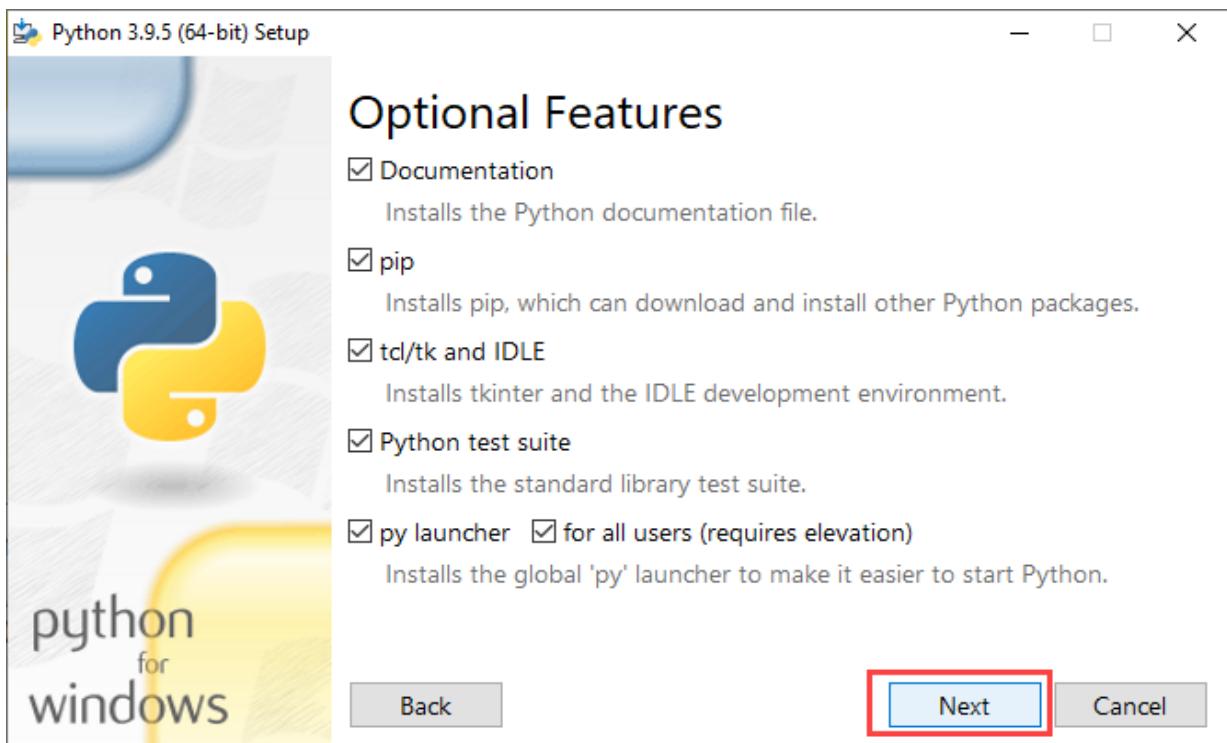
I. Cài đặt Python trên Windows

1. Truy cập vào trang web [python.org](https://www.python.org) để tải xuống trình cài đặt Python
2. Click đúp chuột vào file vừa tải xuống để tiến hành cài đặt Python. Chú ý: Nếu bạn gặp sự cố trong khi cài đặt, hãy tạm thời vô hiệu hóa phần mềm chống vi-rút của bạn trong khi cài đặt, sau đó bật lại sau khi quá trình cài đặt kết thúc. Nếu bạn đã cài đặt cho tất cả người dùng, hãy gỡ cài đặt Python và cài đặt lại nó chỉ cho người dùng của bạn và thử lại.
3. Chọn **Customize installation:**



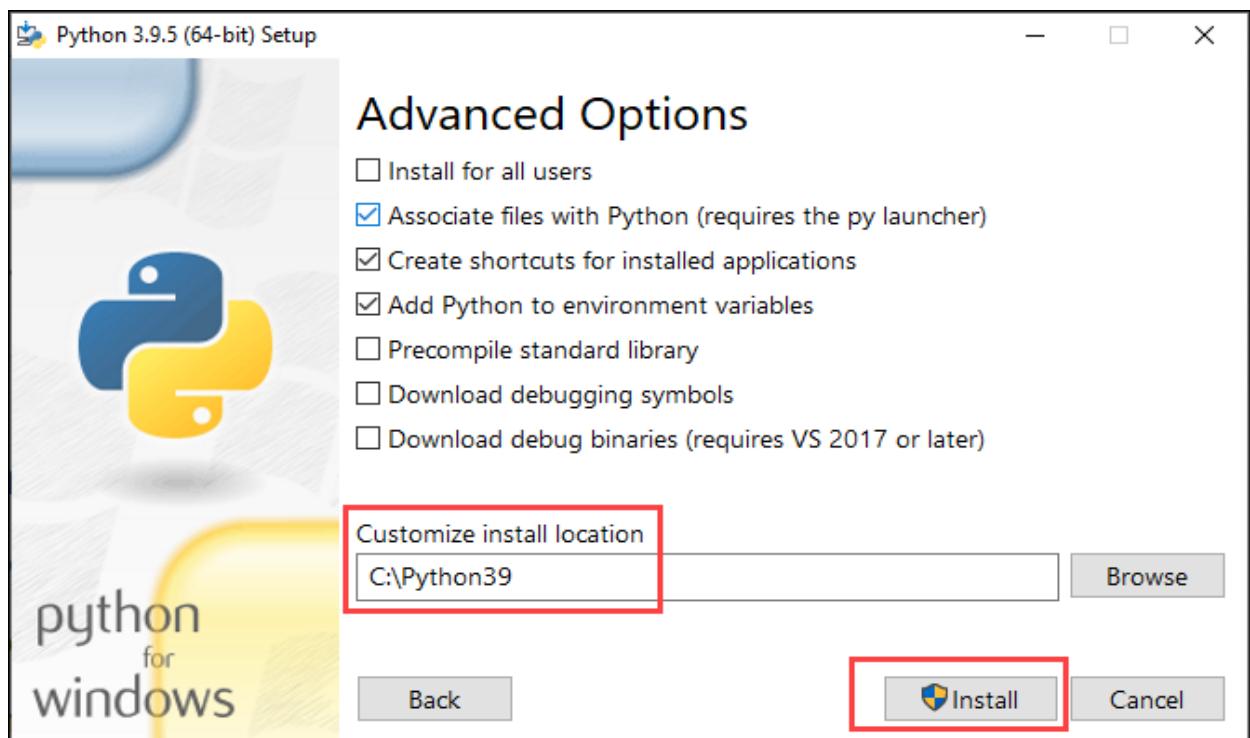
Hình 32: Tùy chỉnh các cài đặt Python

4. Chọn tất cả các tính năng tùy chọn, click Next:



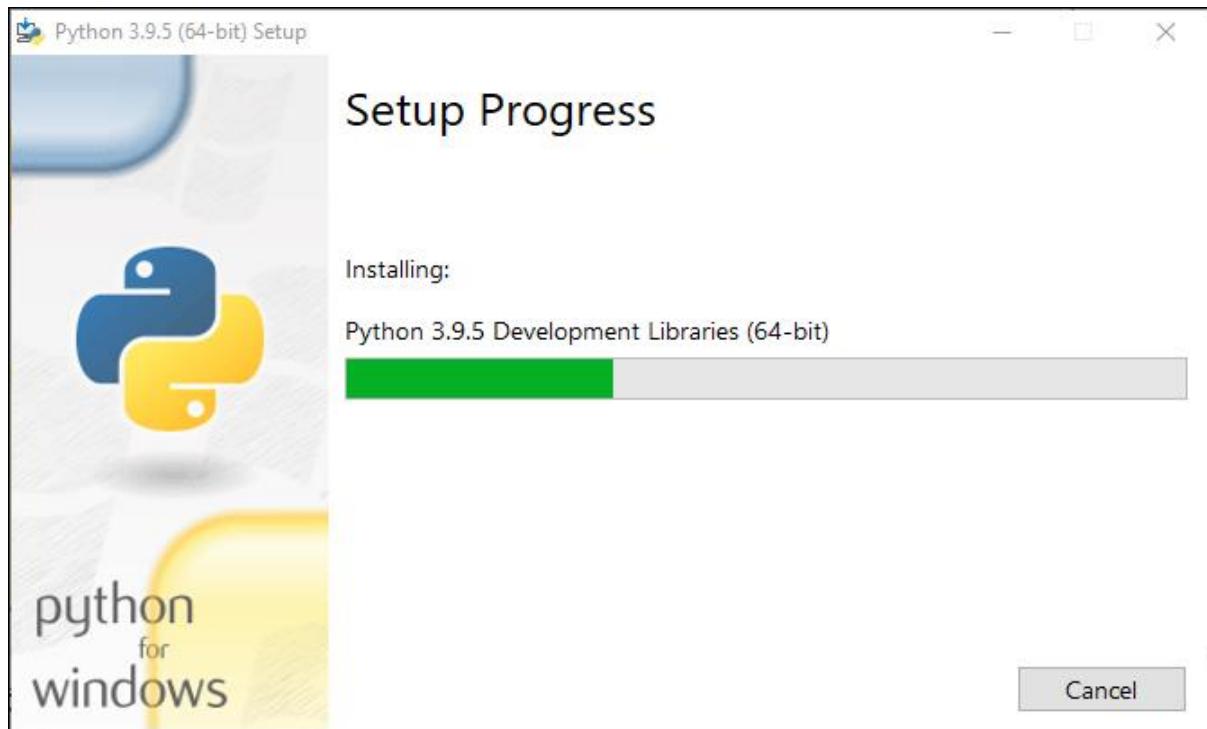
Hình 33: Các tính năng tùy chỉnh

5. Chọn các tính năng tùy chọn, click Install:



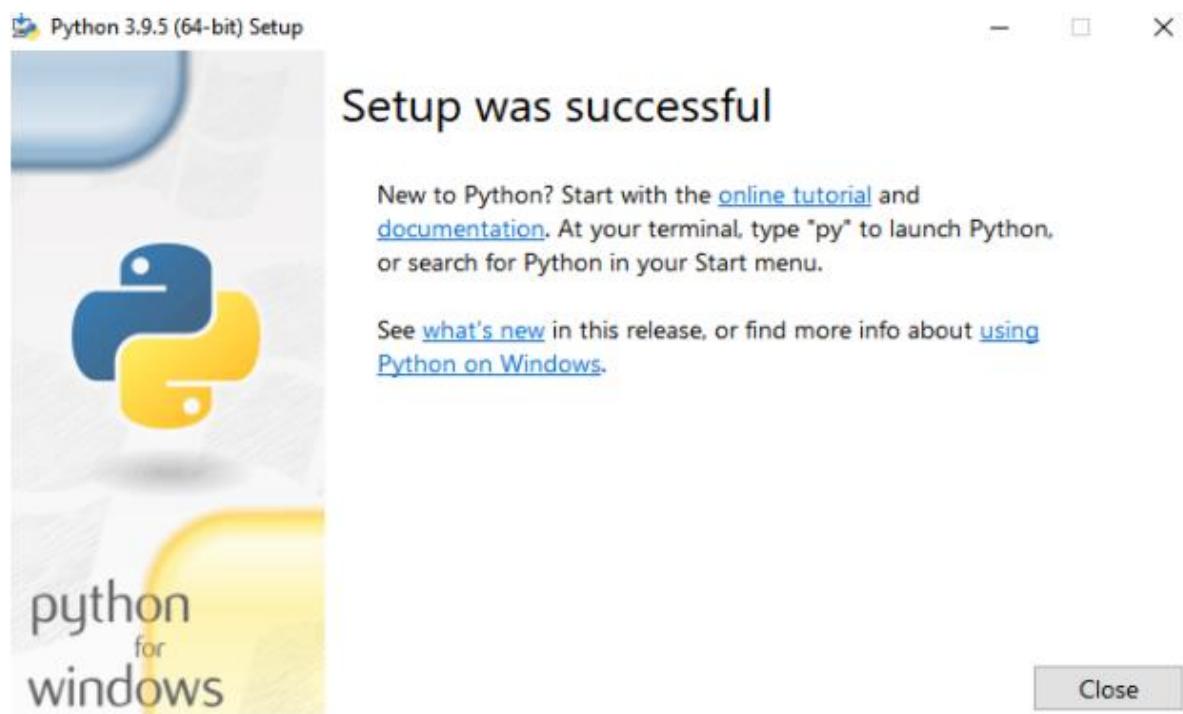
Hình 34: Chọn nơi cài đặt

6. Python 3.9.5 đang được cài đặt:



Hình 35: Python đang được cài đặt

7. Hoàn tất quá trình cài đặt



Hình 36: Cài đặt thành công

II. Cài đặt Java trên Windows

- Tải file cài đặt Java phiên bản mới nhất từ trang chủ

<https://www.oracle.com/java/technologies/downloads/#java19> của Oracle.

The screenshot shows the Oracle Java Downloads page. At the top, there's a navigation bar with tabs for 'Java downloads', 'Tools and resources', and 'Java archive'. Below the navigation bar, a banner says 'Looking for other Java downloads?' with buttons for 'OpenJDK Early Access Builds' and 'JRE for Consumers'. A section titled 'Java 19 and Java 17 available now' is displayed, stating that Java 17 LTS is the latest long-term support release. It includes links for 'Learn about Java SE Subscription' and 'Java 19' and 'Java 17'. A section for 'Java SE Development Kit 19.0.1 downloads' follows, with a note about the Java™ Platform, Standard Edition Development Kit (JDK™). It lists download options for 'Linux', 'macOS', and 'Windows'. The 'Windows' option is underlined, indicating it is selected.

Hình 37: Trang tải JDK phiên bản mới nhất hiện tại

- Chọn phiên bản phù hợp với thiết bị (Khuyến khích nên chọn phiên bản Windows x64 Installer) và tiến hành tải về

The screenshot shows the Oracle Java Downloads page for Java 19. The navigation bar and Java 19/17 availability section are identical to the previous screenshot. The 'Java SE Development Kit 19.0.1 downloads' section is also present. The 'Windows' tab is selected, showing download links for three different installer types: 'x64 Compressed Archive', 'x64 Installer', and 'x64 MSI Installer'. Each item has its file size and a download link. The URL in the browser address bar is https://www.oracle.com/java/technologies/downloads/#jdk19-windows.

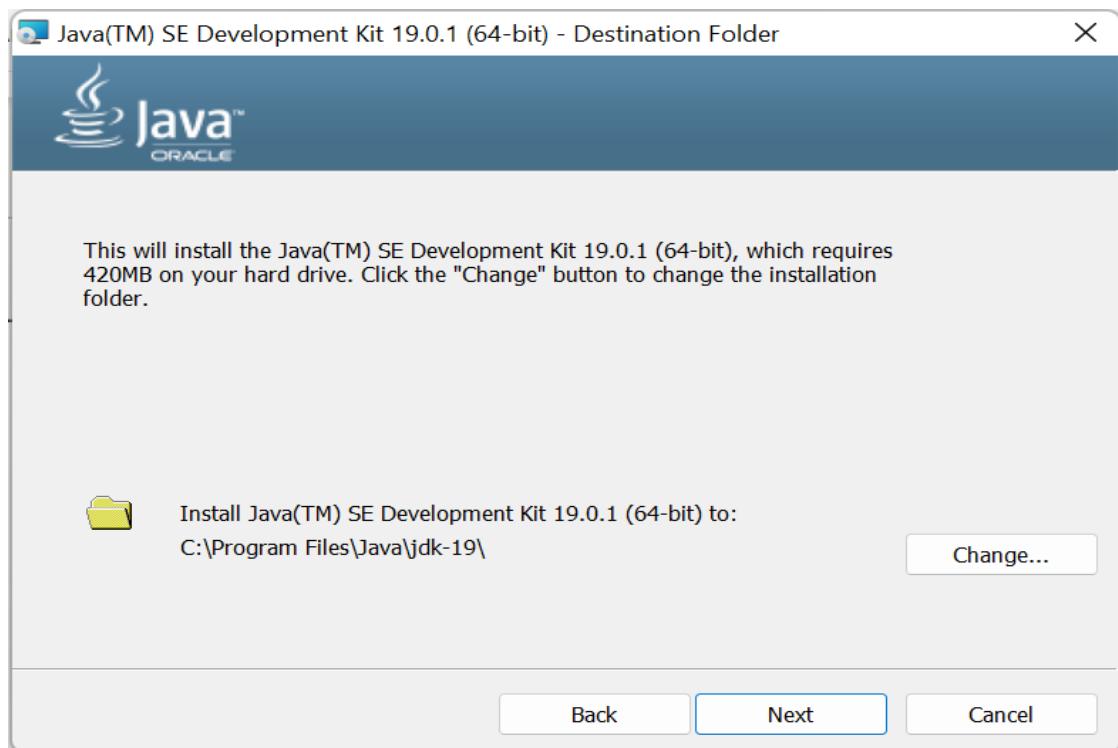
Hình 38: Các phiên bản JDK cho các hệ điều hành

- Nhấn đúp vào tệp vừa tải về và tiến hành cài đặt, hộp thoại cài đặt JDK sẽ xuất hiện và click Next.



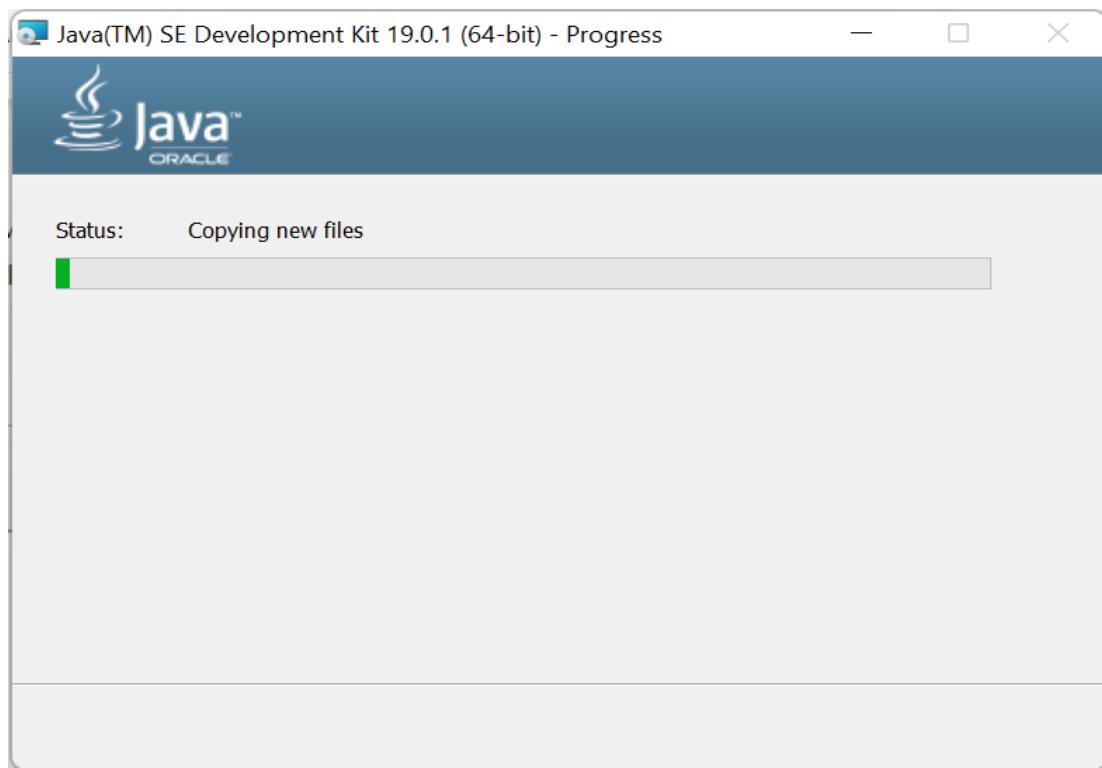
Hình 39: Tiến hành cài đặt JDK

- Chọn nơi lưu trữ file JDK và click Next để tiến hành cài đặt.



Hình 40: Chọn nơi lưu trữ

5. Tiến hành cài đặt JDK cho máy.



Hình 41: Đang tiến hành cài đặt

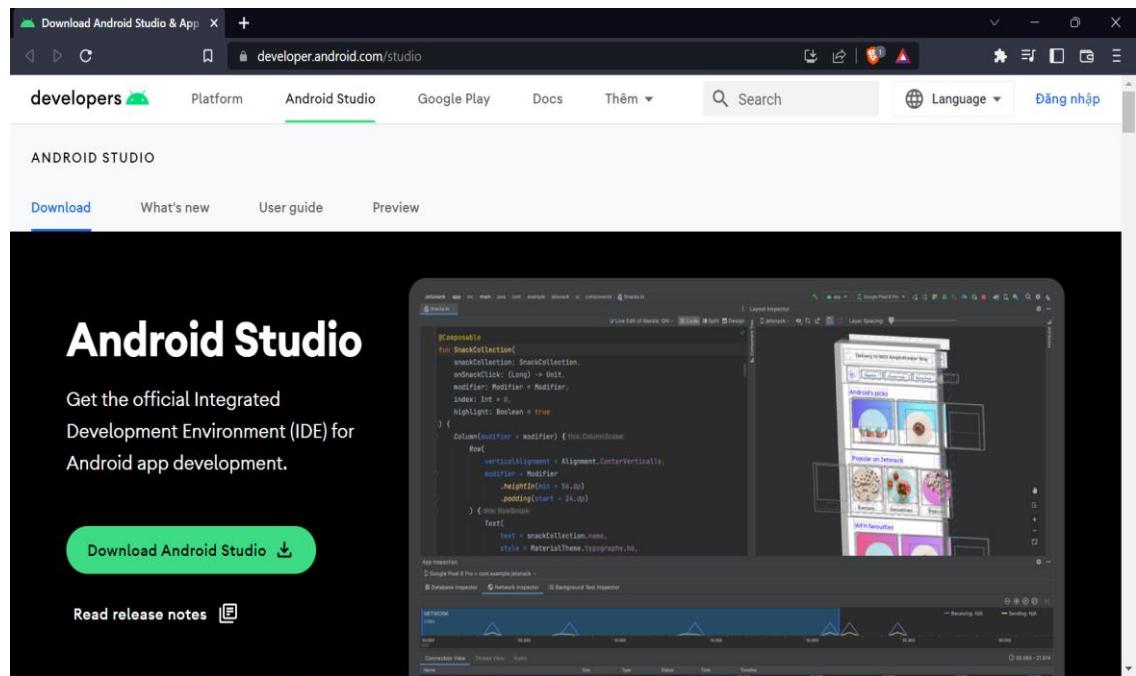
6. Hoàn tất cài đặt JDK cho máy.



Hình 42: Cài đặt JDK thành công

III. Cài đặt Android Studio trên Windows

- Truy cập vào trang <https://developer.android.com/studio> để tiến hành tải tệp cài đặt.



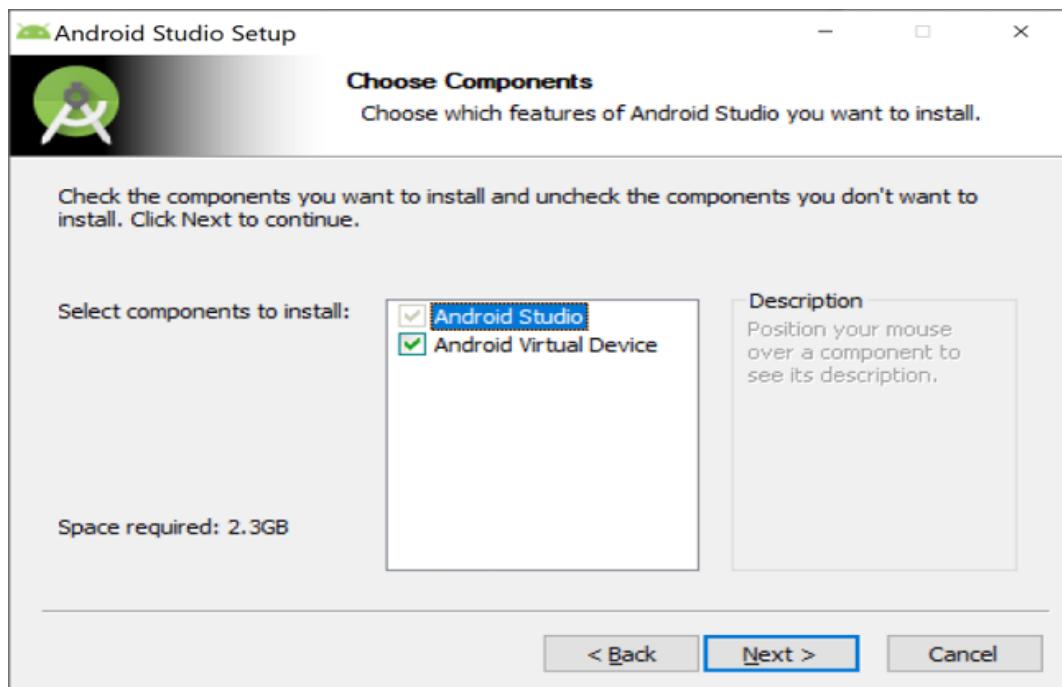
Hình 43: Trang tải về bản cài đặt Android Studio

- Nháy đúp vào tệp vừa tải về và click Next để tiến hành cài đặt.



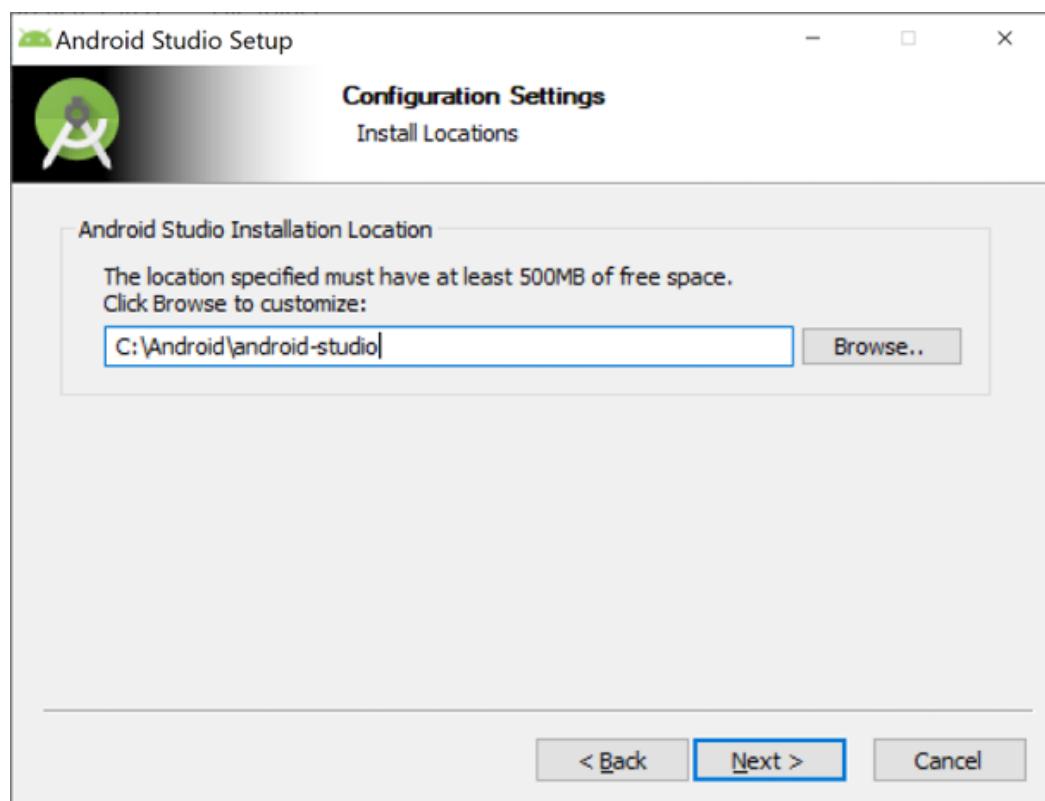
Hình 44: Cửa sổ cài đặt Android Studio

- Chọn các components (Khuyến khích nên để mặc định) và click Next.



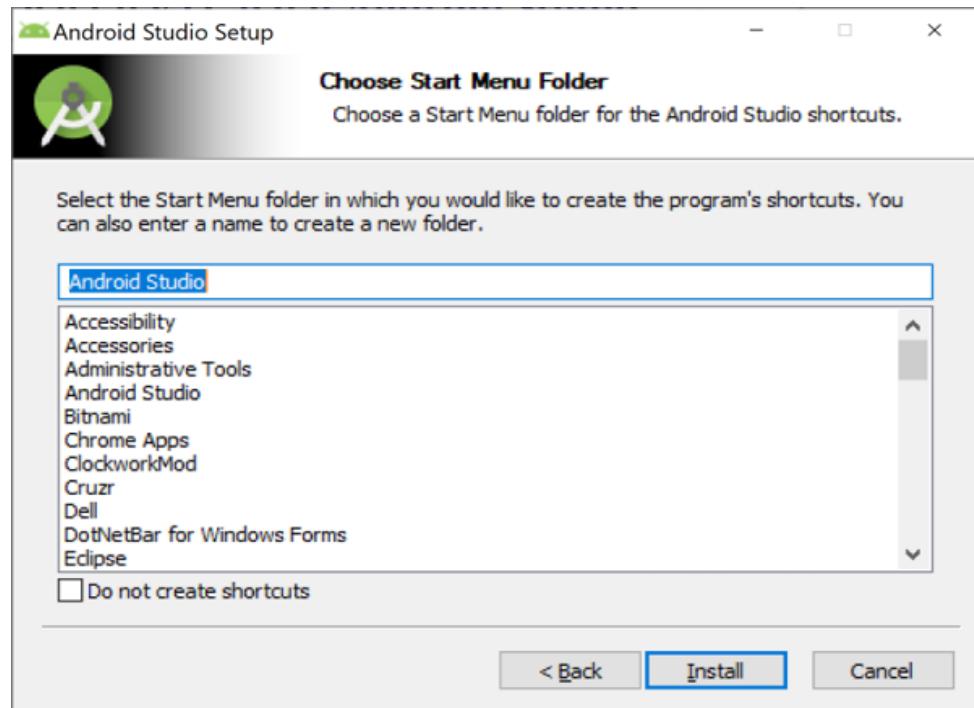
Hình 45: Cửa sổ lựa chọn các Components

- Chọn nơi lưu trữ tài nguyên chương trình và click Next.



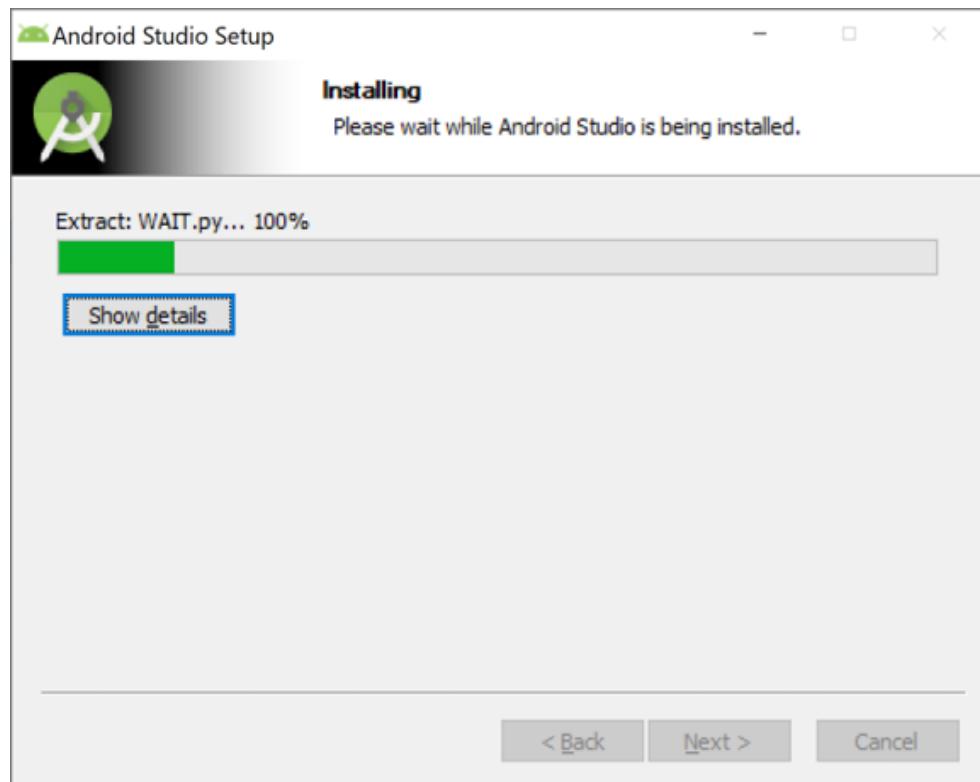
Hình 46: Chọn nơi lưu trữ

5. Chọn Start Menu Folder (Android Studio) và click Install.

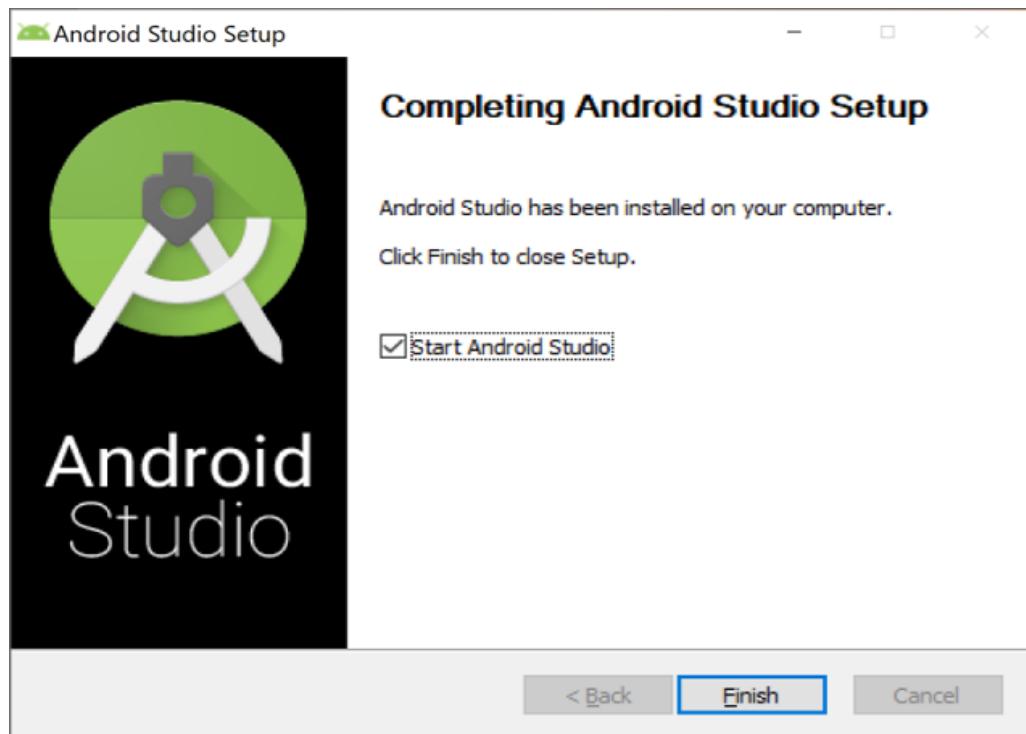


Hình 47: Cửa sổ chọn Start Menu Folder

6. Chờ chương trình tiến hành cài đặt Android Studio trên máy.

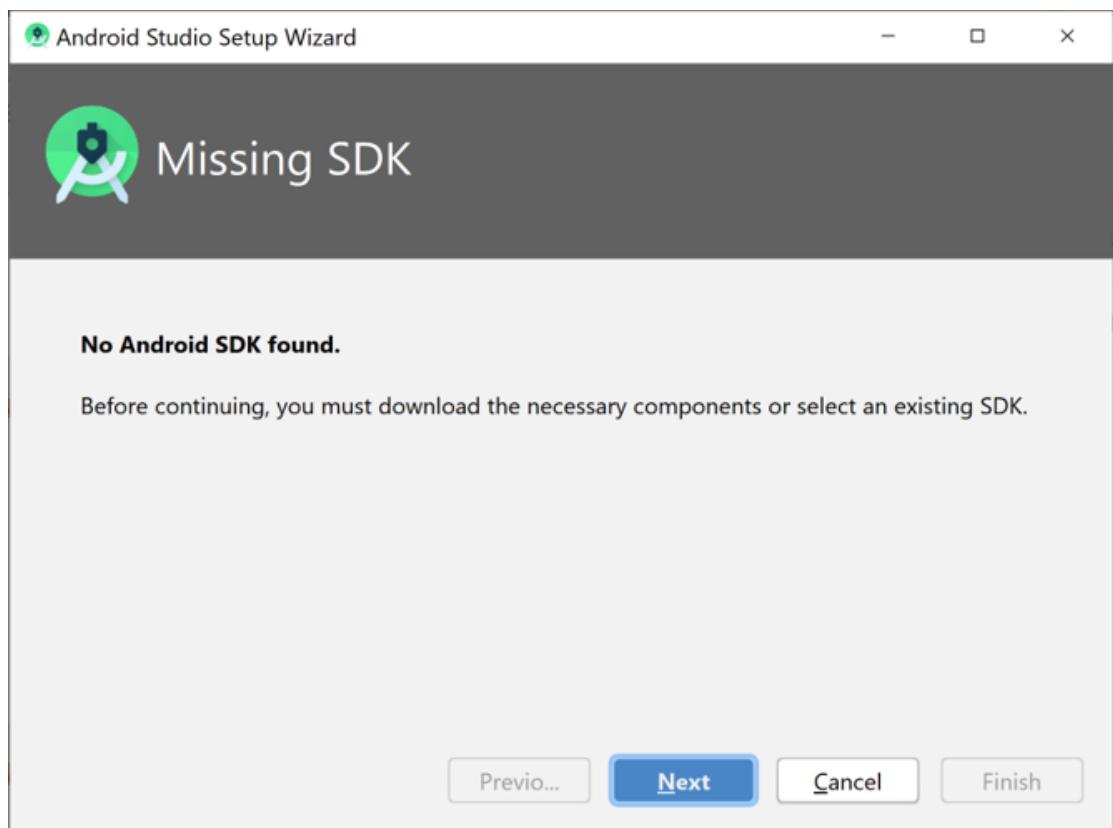


Hình 48: Hoàn tất việc cài đặt Android Studio



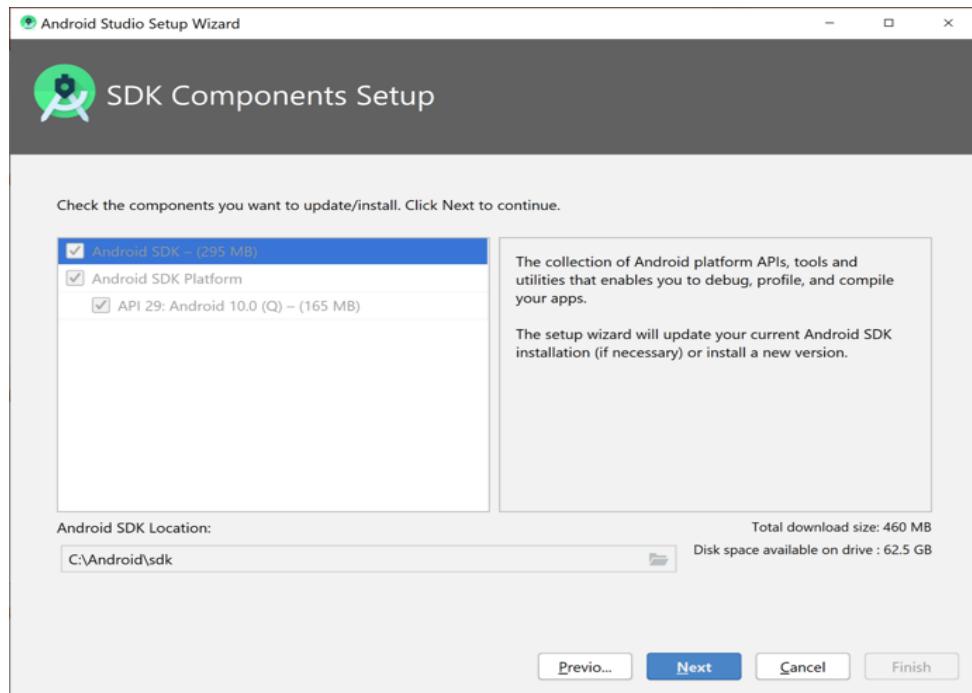
Hình 49: Cài đặt thành công

7. Tiến hành cài đặt SDK cho máy, cửa sổ cài đặt SDK xuất hiện và click Next.



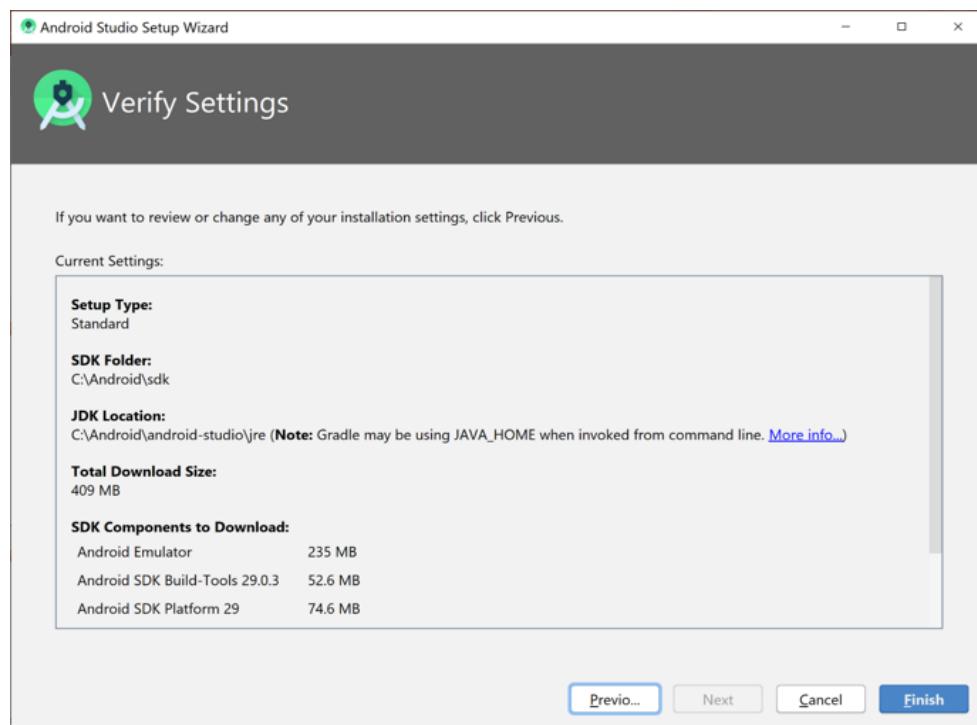
Hình 50: Cửa sổ thông báo Missing SDK

8. Cài đặt các Components và nơi lưu trữ file SDK (lưu ý Android SDK Location ta chọn đúng nơi mà ta đã tạo thư mục trước đó)



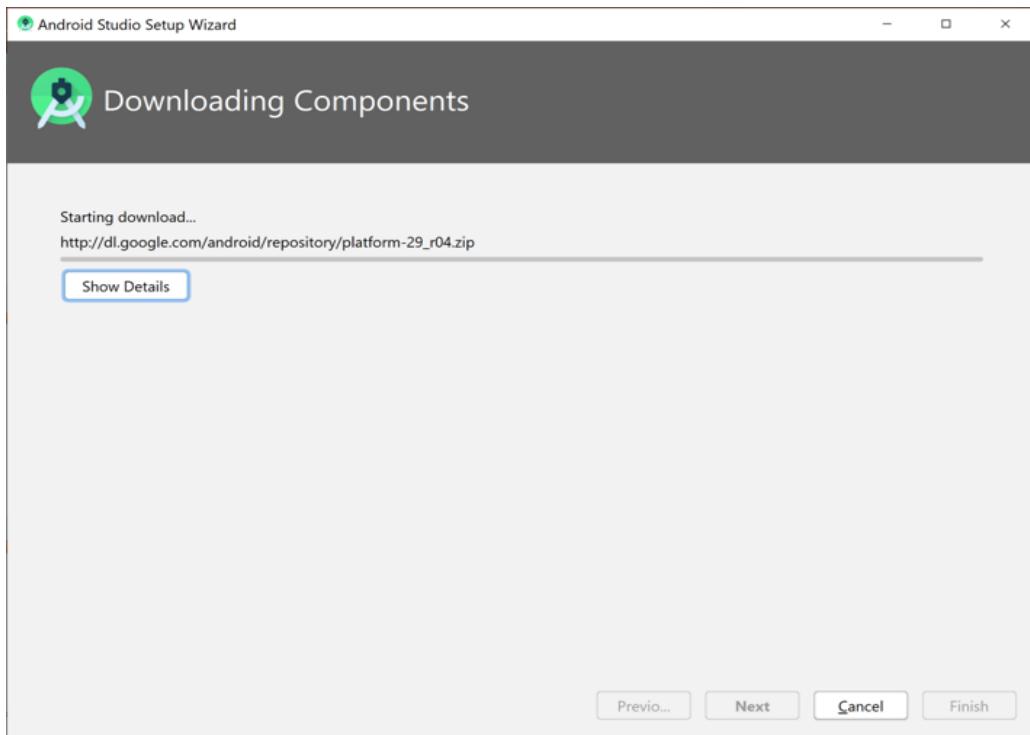
Hình 51: Cửa sổ cài đặt Components SDK

9. Cửa sổ các thông số của SDK xuất hiện, click Finish.



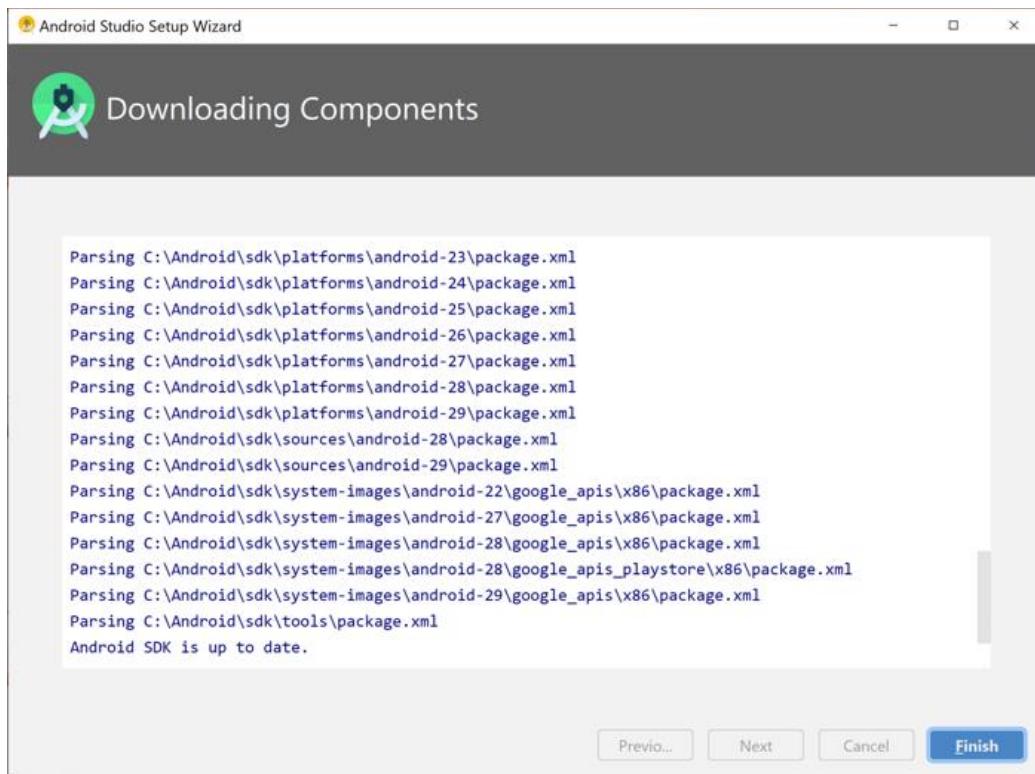
Hình 52: Các thông số của chương trình

10. Chương trình đang tiến hành cài đặt



Hình 53: Đang tiến tải và cài đặt các tệp tài nguyên

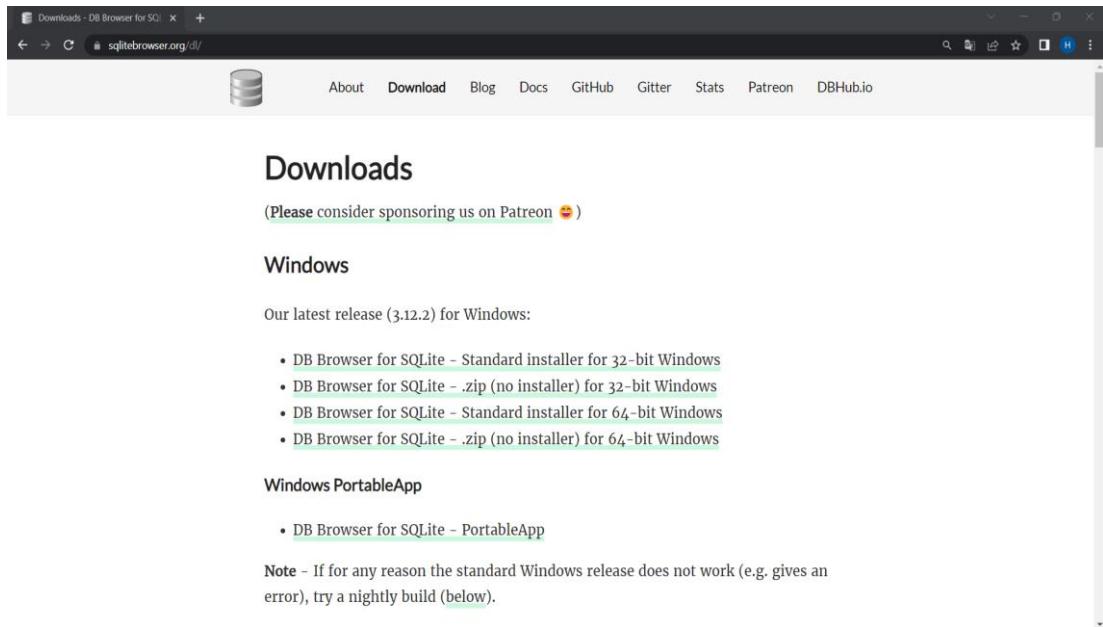
11. Hoàn tất việc cài đặt, click Finish để bắt đầu.



Hình 54: Hoàn tất việc cài đặt

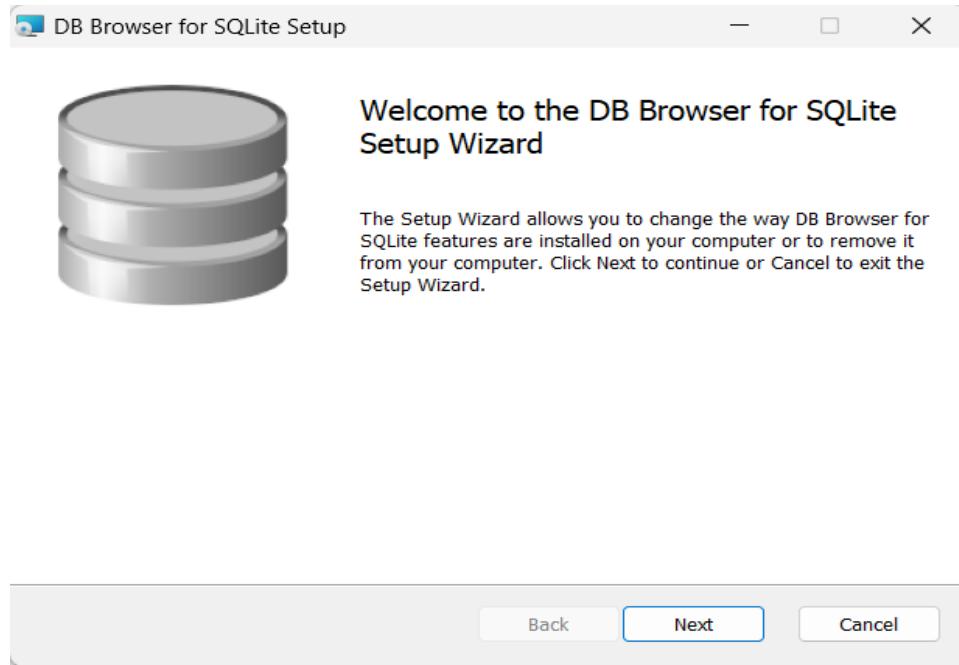
IV. Cài đặt DB Browser for SQLite trên máy Windows

- Truy cập vào trang <https://sqlitebrowser.org/dl/> để tiến hành tải tệp cài đặt. Tìm kiếm mục cài đặt cho Windows tải tệp DB Browser for SQLite - .zip (no installer) for 64-bit Windows



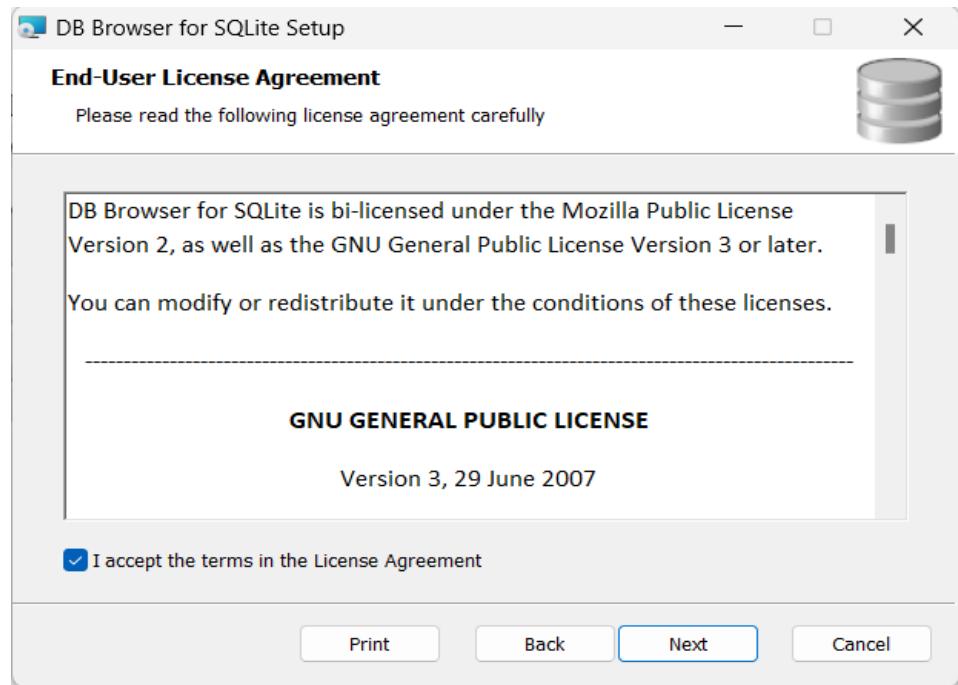
Hình 55: Trang tải về bản cài đặt DB Browser for SQLite

- Nháy đúp vào tệp vừa tải về và nhấn Next để tiếp tục cài đặt



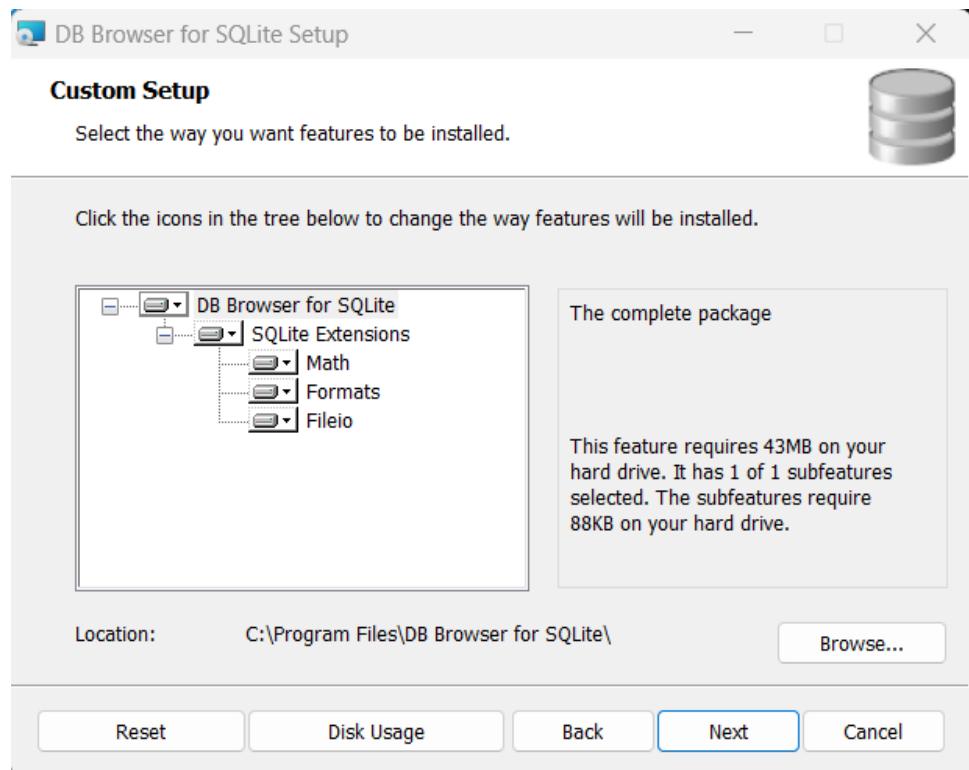
Hình 56: Cửa sổ cài đặt DB Browser for SQLite

3. Tích vào ô I accept the term in the License Agreement sau đó nhấn Next



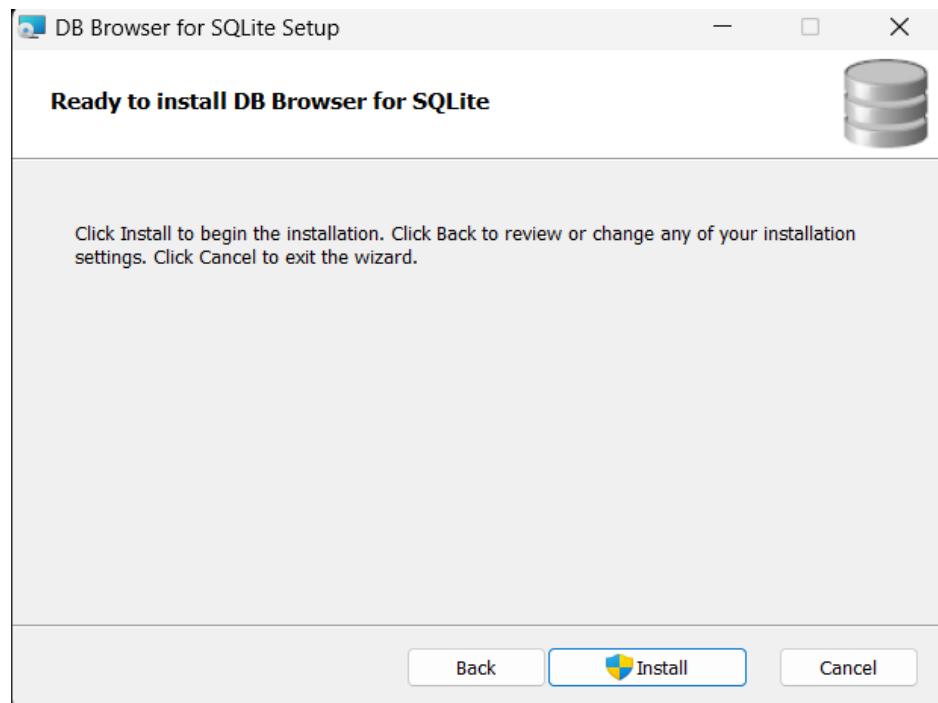
Hình 57: Cửa sổ điều khoản của DB Browser for SQLite

4. Chọn nơi lưu trữ ứng dụng và nhấn Next



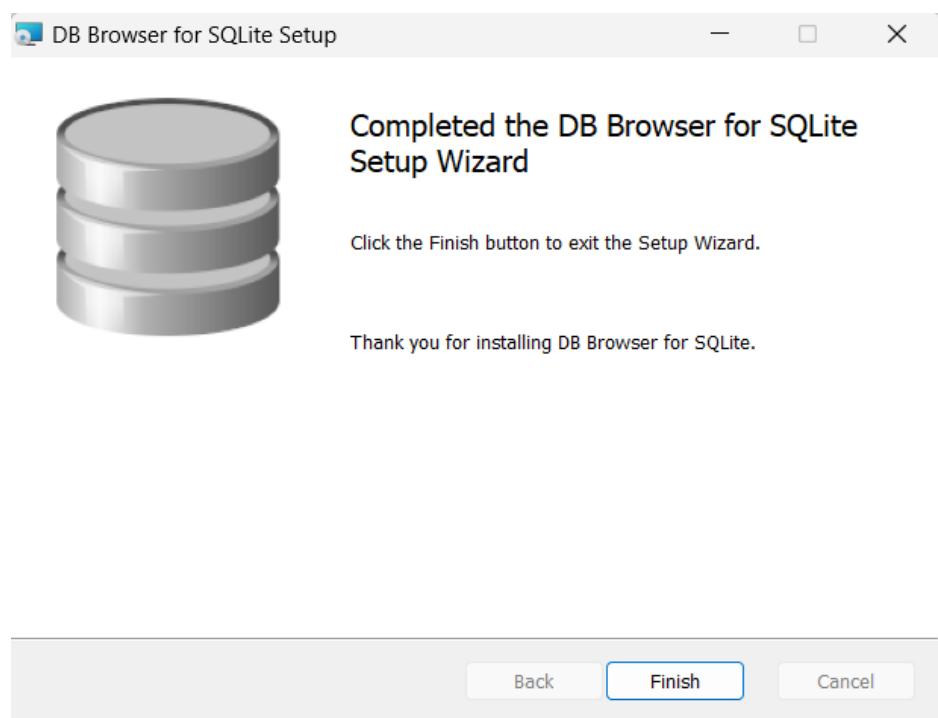
Hình 58: Lựa chọn nơi lưu trữ

5. Nhấn vào nút Install để tiến hành cài đặt



Hình 59: Bắt đầu cài đặt

6. Sau thời gian cài đặt thành công nhấn nút Finish để kết thúc cài đặt DB Browser for SQLite trên máy

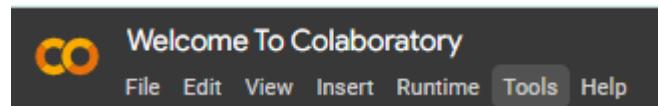


Hình 60: Cài đặt thành công

V. Cài đặt yolov5 và huấn luyện mô hình

1. Môi trường cài đặt

Để tài sử dụng google colaboratory pro để làm môi trường để đào tạo mô hình deep learning.



Mount to drive để lưu trữ mô hình đào tạo

```
▶ from google.colab import drive  
drive.mount('/content/drive')
```

2. Huấn luyện mô hình

Tải yolov5 từ github về và cài đặt thêm một số thư viện cần thiết

```
[ ] # clone YOLOv5 repository  
!git clone https://github.com/ultralytics/yolov5  
# vào thư mục yolov5 mới được cloned về  
%cd yolov5  
# (Không bắt buộc) dùng 1 commit nhất định  
# !git reset --hard 886f1c03d83957afecb059accf74296fad395b6  
# Cài đặt các thư viện cần thiết  
!pip install -qr requirements.txt # install dependencies  
  
[ ] import torch  
  
# Kiểm tra notebook đã dùng GPU (thời gian huấn luyện nhanh hơn)  
print('Setup complete. Using torch %s %s' % (torch.__version__, torch.cuda.get_device_properties(0)  
if torch.cuda.is_available() else 'CPU'))
```

GPU giúp giảm thời gian huấn luyện mô hình

3. Huấn luyện

Sau khi đã chuẩn bị xong, tiến hành đào tạo mô hình với tập train.py các tham số:

- img: kích thước ảnh (độ phân giải)
- batch: số ảnh dùng để huấn luyện trong mỗi lượt
- epochs: số lượt huấn luyện cho tất cả các ảnh trong tập dữ liệu train
- data: đường dẫn đến file cấu hình của tập dữ liệu
- weights: đường dẫn đến file weight chứa độ liên kết giữa các neuron (để "là để huấn luyện từ đầu")

- cache: dùng bộ nhớ đệm để huấn luyện nhanh hơn

```
!python train.py --img 640 --batch 4 --epochs 800 --data 'yolov5/model/data.yaml' --weights yolov5m.pt --cache
```

Sau khi việc huấn luyện kết thúc, có 2 file weights mô hình được lưu lại:

- Mô hình tốt nhất: best.pt (dựa trên chỉ số Average Precision trên tập kiểm định)
- Mô hình cuối cùng: last.pt

THUYẾT MINH ĐỀ TÀI
NGHIÊN CỨU KHOA HỌC CẤP CƠ SỞ
(do sinh viên thực hiện)

1. TÊN ĐỀ TÀI Úng dụng hỗ trợ học tiếng Anh qua việc nhận dạng đối tượng trong video.	2. MÃ SỐ THS2022-15																		
Lĩnh vực ưu tiên <ul style="list-style-type: none"> <input type="checkbox"/> Lĩnh vực 1. Ứng dụng công nghệ cao trong nông nghiệp, thủy sản và môi trường <input type="checkbox"/> Lĩnh vực 2. Quản lý và sử dụng bền vững tài nguyên thiên nhiên <input checked="" type="checkbox"/> Lĩnh vực 3. Kỹ thuật công nghệ và công nghệ thông tin – truyền thông <input checked="" type="checkbox"/> Lĩnh vực 4. Khoa học Giáo dục, Luật và Xã hội Nhân văn <input type="checkbox"/> Lĩnh vực 5. Phát triển kinh tế, thị trường <input type="checkbox"/> Không thuộc 05 Lĩnh vực ưu tiên. 																			
3. LĨNH VỰC NGHIÊN CỨU <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;">Khoa học Tự nhiên</td> <td style="width: 50%; text-align: center;"><input type="checkbox"/></td> <td style="width: 50%;">Khoa học Kỹ thuật và Công nghệ</td> <td style="width: 50%; text-align: center;"><input checked="" type="checkbox"/></td> </tr> <tr> <td style="text-align: center;">Khoa học Y, dược</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;">Khoa học Nông nghiệp</td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> <tr> <td style="text-align: center;">Khoa học Xã hội</td> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;">Khoa học Nhân văn</td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> </table>	Khoa học Tự nhiên	<input type="checkbox"/>	Khoa học Kỹ thuật và Công nghệ	<input checked="" type="checkbox"/>	Khoa học Y, dược	<input type="checkbox"/>	Khoa học Nông nghiệp	<input type="checkbox"/>	Khoa học Xã hội	<input type="checkbox"/>	Khoa học Nhân văn	<input type="checkbox"/>	4. LOẠI HÌNH NGHIÊN CỨU <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 33%;">Cơ bản</td> <td style="width: 33%;">Ứng dụng</td> <td style="width: 33%;">Triển khai</td> </tr> <tr> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input checked="" type="checkbox"/></td> <td style="text-align: center;"><input type="checkbox"/></td> </tr> </table>	Cơ bản	Ứng dụng	Triển khai	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Khoa học Tự nhiên	<input type="checkbox"/>	Khoa học Kỹ thuật và Công nghệ	<input checked="" type="checkbox"/>																
Khoa học Y, dược	<input type="checkbox"/>	Khoa học Nông nghiệp	<input type="checkbox"/>																
Khoa học Xã hội	<input type="checkbox"/>	Khoa học Nhân văn	<input type="checkbox"/>																
Cơ bản	Ứng dụng	Triển khai																	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>																	
5. THỜI GIAN THỰC HIỆN 6 tháng Từ tháng 6 năm 2022 đến tháng 11 năm 2022																			

6. ĐƠN VỊ CỦA CHỦ NHIỆM ĐỀ TÀI

Tên đơn vị: Khoa Công nghệ thông tin và Truyền thông

Điện thoại: 0292 3734713

E-mail: office@cit.ctu.edu.vn

Địa chỉ: Khu 2, đường 3/2, Phường Xuân Khánh, Q. Ninh Kiều, TP. Cần Thơ, Việt Nam

Họ và tên thủ trưởng đơn vị: TS. Nguyễn Hữu Hoà

7. CHỦ NHIỆM ĐỀ TÀI

Họ và tên: Nguyễn Chí Bảo

MSSV: B1910619

Ngày tháng năm sinh: 28/05/2001

Lớp: DI19V7F2 (CNTT Chất lượng cao F2)

Điện thoại di động: 0353530086

Khóa: K45

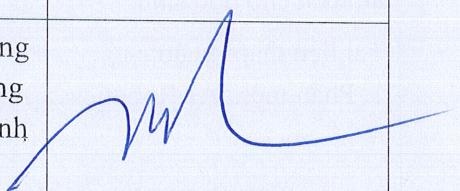
E-mail: baob1910619@student.ctu.edu.vn

8. NHỮNG THÀNH VIÊN THAM GIA NGHIÊN CỨU ĐỀ TÀI

TT	Họ và tên	MSSV, Lớp, Khóa	Nội dung nghiên cứu cụ thể được giao	Chữ ký
1	Nguyễn Chí Bảo (chủ nhiệm đề tài)	MSSV: B1910619 Lớp: DI19V7F2 (CNTT Chất lượng cao F2) Khóa: K45	Lên ý tưởng Lập kế hoạch Xây dựng bộ từ điển Thu thập dữ liệu Nghiên cứu mô hình nhận diện Triển khai mô hình nhận diện Train nhận dạng Kiểm thử Viết báo cáo	<u>bao</u>
2	Lê Nguyễn Bảo Anh (thư ký khoa học)	MSSV: B1910611 Lớp: DI19V7F2 (CNTT Chất lượng cao F2) Khóa: K45	Phân tích và thiết kế ứng dụng Nghiên cứu mô hình nhận diện Triển khai mô hình nhận diện Train nhận dạng Xây dựng chức năng Lập trình và cài đặt Kiểm thử Viết báo cáo	<u>anh</u>

3	Lê Huỳnh Như (thành viên chính)	MSSV: B1910680 Lớp: DI19V7F2 (CNTT Chất lượng cao F2) Khóa: K45	Xây dựng bộ từ điển Xây dựng bộ dữ liệu đặt câu với từ vừa nhận dạng được Lập trình và cài đặt Kiểm thử Viết báo cáo	<u>Nhu</u>
4	Danh Bé Hai (thành viên chính)	MSSV: B1909907 Lớp: DI19V7F2 (CNTT Chất lượng cao F2) Khóa: K45	Thu thập dữ liệu Phân tích và thiết kế ứng dụng Xây dựng chức năng Lập trình và cài đặt Thiết kế giao diện Kiểm thử Làm slide báo cáo Viết báo cáo	<u>hai</u>
5	Đặng Quốc Cường (thành viên chính)	MSSV: B1910623 Lớp: DI19V7F2 (CNTT Chất lượng cao F2) Khóa: K45	Thu thập dữ liệu Phân tích và thiết kế ứng dụng Xây dựng chức năng Lập trình và cài đặt Thiết kế giao diện Kiểm thử Viết báo cáo	<u>Cuong</u>

Cán bộ hướng dẫn sinh viên thực hiện đề tài

Họ và tên, MScB	Đơn vị công tác và lĩnh vực chuyên môn	Nhiệm vụ	Chữ ký
TS. Nguyễn Thanh Hải MScB: 2267	Đơn vị công tác: Khoa CNTT & TT Lĩnh vực chuyên môn: CNTT	Hướng dẫn nội dung khoa học và Hướng dẫn lập dự toán kinh phí đề tài	

9. ĐƠN VỊ PHỐI HỢP CHÍNH

Tên đơn vị trong và ngoài nước	Nội dung phối hợp nghiên cứu	Họ và tên người đại diện đơn vị
Không	Không	Không

10. TỔNG QUAN TÌNH HÌNH NGHIÊN CỨU THUỘC LĨNH VỰC CỦA ĐỀ TÀI Ở TRONG VÀ NGOÀI NƯỚC

10.1. Trong nước

Trong những năm trở lại cho ta thấy được sự phát triển đáng kinh ngạc của ngành khoa học, trong đó có vô số những lĩnh vực nghiên cứu khác nhau. Trí tuệ nhân tạo (AI) đã dẫn lối và tạo nên định hướng phát triển tiềm năng lĩnh vực kỹ thuật và công nghệ. Những gia đình người Việt, đặc biệt là những bậc cha mẹ luôn muôn tìm và định hướng cho các con của họ được học Tiếng Anh một cách dễ nhất và tiếp thu nhanh nhất. Đặc biệt là trẻ nhỏ hay những người cần tìm những ứng dụng để học những kiến thức cơ bản trước. Vì thế, việc sử dụng AI để hỗ trợ con người học những từ vựng Tiếng Anh thông qua những hình ảnh, video, hay tất tần tật những gì xung quanh mình trên chiếc smartphone để scan hình ảnh, video mà mình cần tìm hiểu từ ngữ Tiếng Anh đó là gì. Từ đó, con người có thể tự do học và tìm hiểu từ ngữ Tiếng Anh một cách không bị quá phức tạp và gò bó thông qua ứng dụng mobile học tiếng anh qua đối tượng được nhận dạng trong hình ảnh và video.

10.2. Ngoài nước

- Qua những tìm hiểu về các dự án, đề tài liên quan đến việc đưa công nghệ AI vào các hệ thống giáo dục nói chung và giáo dục ngoại ngữ nói riêng có thể thấy được tiềm năng của công nghệ AI trong giáo dục làm cải thiện chất lượng, nâng cao dân trí, trình độ giáo dục cho người dân. Việc ứng dụng công nghệ AI trong giáo dục tiêu biểu như phần mềm Grammarly - phần mềm kiểm tra lỗi chính tả và ngữ pháp tiếng Anh, hoặc là phần mềm AI-Robotics là công cụ tiềm năng và thiết yếu đối với giáo dục tại Mỹ. Trước hết, AI giúp tự động hóa và thực hiện các nhiệm vụ quản trị của giáo viên như chấm điểm bài thi, phân loại giấy tờ và tiến hành các thủ tục nhập học... để giáo viên có nhiều thời gian tương tác với học sinh.

- Công ty Cổ phần dịch vụ và giải pháp công nghệ giáo dục PHX (PHX Smart School) đã cho ra đời Giải pháp PHX Smart School tích hợp phần mềm hỗ trợ quy trình đào tạo, giúp trường học có thể triển khai hoạt động giảng dạy và học tập trực tuyến, chia sẻ video bài giảng tiện lợi để hỗ trợ học tập hiệu quả hơn; tích hợp hệ thống thi trực tuyến, cho phép tạo các kỳ thi trắc nghiệm và chấm điểm ngay sau khi làm xong bài, hay các bài thi tự luận được giảng viên chấm tay và thông báo kết quả cho học sinh và phụ huynh qua ứng dụng. Đây còn là một phần mềm sổ liên lạc điện tử, kết nối nhà trường với phụ huynh một cách dễ dàng thông qua các app trên Android và iOs. Đặc biệt, PHX Smart School tích hợp hệ thống giám sát lộ trình di chuyển thông minh, điểm danh chính xác bằng gương mặt... tự động kết nối và cập nhật thông tin nhanh chóng tới gia đình về trạng thái của học sinh khi đến trường, đảm bảo an toàn cho học sinh.

Tài liệu tham khảo:

1. Phần mềm AI-Robotics: Emerging Technology Research & Trends – Think with Google APAC

2. PHX Smart School: Trang chủ – PHX Smart School, phx-smartschool.com

10.3. Danh mục các công trình đã công bố thuộc lĩnh vực của đề tài của chủ nhiệm và những thành viên tham gia nghiên cứu

a) Của chủ nhiệm đề tài: Không

b) Của các thành viên tham gia nghiên cứu: Không

11. TÍNH CẤP THIẾT CỦA ĐỀ TÀI

- Sự phát triển nhanh chóng và liên tục được cải tiến không ngừng của các giải thuật học sâu, các kết quả trong việc nhận dạng vật thể có độ chính xác cao. Đề tài tập chung nghiên cứu và xây dựng một ứng dụng để học Tiếng Anh thông qua việc đánh dấu các vật thể từ hình ảnh, video giúp hỗ trợ người dùng học Tiếng Anh một cách hiệu quả.
- Hiện nay, việc học Tiếng Anh là vô cùng quan trọng, trong đó việc sử dụng hình ảnh để học tiếng Anh là một cách học hiệu quả, đặc biệt đối với người học trực quan. Bằng cách nhìn thấy một thứ gì đó, ghi nhớ nó và sau đó liên kết một khái niệm với nó. Thế nên, việc ứng dụng tiềm năng của công nghệ để đánh dấu vật thể trong hình ảnh, video nhằm hỗ trợ việc học tiếng Tiếng Anh là cần thiết để giúp người dùng nâng cao hiệu quả trong việc học Tiếng Anh của mình.

12. MỤC TIÊU ĐỀ TÀI

- Mục tiêu của đề tài nghiên cứu nhằm tạo ra một chương trình ứng dụng học Tiếng Anh trên App Android. Ứng dụng hỗ trợ phát hiện, nhận dạng các vật thể trong video bằng tiếng anh, làm tiền đề giúp người sử dụng học tiếng anh một thoải mái, giải trí mà vẫn trau dồi được nhiều kiến thức từ vựng với chức năng phát hiện đối tượng, phát âm, đặt câu với đối tượng vừa được nhận dạng. Đặc biệt ứng dụng còn cho phép người học xây dựng từ mới cho từ điển một cách dễ dàng, góp phần tạo nên một cộng đồng lớn mạnh cùng đỡ nhau chinh phục vốn từ vựng sâu rộng và phong phú.

13. ĐỐI TƯỢNG, PHẠM VI NGHIÊN CỨU

13.1. Đối tượng nghiên cứu

- Công nghệ AI nhận dạng đối tượng trong video
- Chuyển đổi văn bản thành giọng nói
- App Android và cấu trúc dữ liệu lưu trữ bài học

13.2. Phạm vi nghiên cứu

- Nghiên cứu: bộ từ điển gồm các từ có liên quan đến các đối tượng ngoài tự nhiên bắt đầu từ địa phận Cần Thơ.
- Xây dựng các tính năng: nhận dạng, lưu trữ, thêm từ mới, public, cho ví dụ với từ mới, dịch sang tiếng việt.

14. CÁCH TIẾP CẬN, PHƯƠNG PHÁP NGHIÊN CỨU

14.1. Cách tiếp cận

Nghiên cứu lý thuyết – Thu thập dữ liệu - Thiết kế và lập trình ứng dụng - Thủ nghiệm - Đánh giá - Ứng dụng

14.2. Phương pháp nghiên cứu:

- Tổng hợp dữ liệu từ các đối tượng bên ngoài tự nhiên từ đó tạo nên mô hình nhận dạng đối tượng trong video kết hợp với dữ liệu tiếng Anh

15. NỘI DUNG NGHIÊN CỨU VÀ TIẾN ĐỘ THỰC HIỆN

15.1. Nội dung nghiên cứu (*Mô tả chi tiết những nội dung nghiên cứu của đề tài*)

- Tìm hiểu nhu cầu sử dụng, tính cấp thiết của đề tài và đối tượng người dùng.
- Phân tích đề tài.
- Thu thập dữ liệu hình ảnh /video về một số từ thông dụng xung quanh ta (cụ thể là trong địa phận Cần Thơ).
- Xây dựng bộ từ điển tiếng anh gồm chữ và âm thanh từ các đối tượng được nhận dạng.
- Nghiên cứu khảo sát các mô hình máy học để đề xuất mô hình máy học phù hợp để nhận dạng.
- Xây dựng một ứng dụng mobile với giao diện thân thiện để tích hợp những nội dung được đề xuất.

15.2. Tiến độ thực hiện

STT	Các nội dung, công việc thực hiện	Sản phẩm	Thời gian (bắt đầu-kết thúc)	Người thực hiện và số ngày thực hiện
1.	-Tìm hiểu, phân tích đề tài -Thu thập dữ liệu	-Bản thuyết minh -Dataset	Từ 06/2022 Đến 07/2022	Nguyễn Chí Bảo (16 ngày) Lê Nguyễn Bảo Anh (16 ngày) Lê Huỳnh Như (16 ngày) Đặng Quốc Cường (2 ngày) Danh Bé Hai (2 ngày)
2.	Lập trình, cài đặt theo mô hình thiết kế - Lập trình - Cài đặt	Ứng dụng nhận dạng ngôn ngữ Tiếng Anh qua đối tượng được nhận dạng trong hình ảnh và video.	Từ 07/2022 Đến 08/2022	Nguyễn Chí Bảo (18 ngày) Lê Nguyễn Bảo Anh (20 ngày) Lê Huỳnh Như (10 ngày) Đặng Quốc Cường (18 ngày) Danh Bé Hai (18 ngày)
3.	Kiểm thử và sửa lỗi	Hoàn thành ứng dụng	Từ 08/2022 Đến 09/2022	Nguyễn Chí Bảo (10 ngày) Lê Nguyễn Bảo Anh (10 ngày) Lê Huỳnh Như (10 ngày) Danh Bé Hai (10 ngày) Đặng Quốc Cường (10 ngày)

4.	Viết báo cáo tổng kết đề tài	Báo cáo tổng kết	Từ 09/2022 Đến 11/2022	Nguyễn Chí Bảo (12 ngày) Lê Nguyễn Bảo Anh (12 ngày) Lê Huỳnh Như (10 ngày) Danh Bé Hai (12 ngày) Đặng Quốc Cường (10 ngày)
----	------------------------------	------------------	---------------------------	---

16. SẢN PHẨM

Số thứ tự	Tên sản phẩm	Số lượng	Mô tả chi tiết chất lượng sản phẩm đạt được như nội dung, hình thức, các chỉ tiêu, thông số kỹ thuật,...)
I	Sản phẩm khoa học (Các công trình khoa học sẽ được công bố: sách, bài báo khoa học...) Không.		
II	Sản phẩm đào tạo (Luận văn tốt nghiệp đại học) Không.		
III	Sản phẩm ứng dụng: Ứng dụng giáo dục sử dụng đa nền tảng với giao diện thân thiện và phù hợp với người dùng.		

17. PHƯƠNG THỨC CHUYỂN GIAO KẾT QUẢ NGHIÊN CỨU VÀ ĐỊA CHỈ ỨNG DỤNG

17.1. Phương thức chuyển giao:

Chuyển giao cho Khoa Công Nghệ Thông Tin và Truyền Thông, Trường Đại học Cần Thơ như ví dụ minh họa cho các ứng dụng của trí tuệ nhân tạo.

17.2. Địa chỉ ứng dụng:

Khoa CNTT & TT trường Đại Học Cần Thơ

Địa chỉ: Khu 2, đường 3/2, Phường Xuân Khánh, Q. Ninh Kiều, TP. Cần Thơ.

18. TÁC ĐỘNG VÀ LỢI ÍCH MANG LẠI CỦA KẾT QUẢ NGHIÊN CỨU

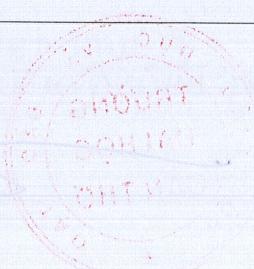
18.1. Đối với lĩnh vực giáo dục và đào tạo

- Phát triển phạm vi nghiên cứu về nhận dạng vật thể thông qua hình ảnh, video. Phục vụ trong việc học Tiếng Anh của người dùng.

- Trong quá trình nghiên cứu cũng giúp ích cho sinh viên tham gia cơ hội để rèn luyện Tiếng Anh, cách ứng dụng CNTT vào một bài toán thực tế trong việc học tập. Giúp sinh viên nâng cao tinh thần tự học, tự nghiên cứu và sáng tạo. Góp phần hỗ trợ sinh viên nghiên cứu sinh trong lĩnh vực công nghệ máy học, phân tích và thiết kế dữ liệu

18.2. Đối với lĩnh vực khoa học và công nghệ có liên quan

- Góp phần ứng dụng công nghệ thông tin vào giáo dục. Tăng thêm phần hứng thú và khả năng ghi nhớ khi học tiếng anh cho trẻ em.



- Góp phần hỗ trợ các nhà nghiên cứu về máy học, học sâu cải thiện và ứng dụng các giải thuật. Từ đó đưa vào các mô hình kiểm thử và thực nghiệm phục vụ cho các báo cáo khoa học hay ứng dụng thực tiễn.

18.3. Đối với phát triển kinh tế-xã hội

- Đẩy mạnh phát triển công nghệ AI trong lĩnh vực giáo dục đặc biệt là giáo dục cho trẻ nhỏ ở thời đại 4.0.
- Phát triển ứng dụng có thể mở rộng để sử dụng rộng rãi trong các lĩnh vực giáo dục khác cũng như nâng cao trình độ ngoại ngữ trong thời đại kinh tế hội nhập hiện nay.
- Giảm một phần chi phí thuê giảng dạy cũng như giúp việc học ngoại ngữ trở nên gần gũi và dễ tiếp cận hơn đối với gia đình không có điều kiện.

18.4. Đối với tổ chức chủ trì và các cơ sở ứng dụng kết quả nghiên cứu

- Cung cấp một ứng dụng học từ vựng tiếng Anh tiện lợi, gần gũi.
- Hỗ trợ học tập về nhiều mặt từ đó tiết kiệm được thời gian mà vẫn đem lại hiệu quả tốt cho người dùng

19. KINH PHÍ THỰC HIỆN ĐỀ TÀI VÀ NGUỒN KINH PHÍ

Kinh phí thực hiện đề tài: 15.000.000 đồng.

Trong đó:

Kinh phí Trường cấp: 15.000.000 đồng.

Các nguồn khác: 0 đồng.

Số thứ tự	Khoản chi, nội dung chi	Tổng kinh phí	Đơn vị tính: đồng	
			Kinh phí Trường cấp	Các nguồn khác
1	Chi mua vật tư, nguyên, nhiên, vật liệu	0	0	0
2	Chi tiền công lao động trực tiếp	12.100.000	12.100.000	0
3	Chi văn phòng, phẩm, thông tin liên lạc, in ấn	175.000	175.000	0
4	Chi họp hội đồng đánh giá, nghiệm thu	2.725.000	2.725.000	0
Tổng cộng		15.000.000	15.000.000	0

KHOA CNTT & TT

CÁN BỘ HƯỚNG DẪN

Ngày 01 tháng 06 năm 2022
CHỦ NHIỆM ĐỀ TÀI

Huỳnh Xuân Hiệp

Nguyễn Thanh Hải

Nguyễn Chí Bảo



Lê Nguyễn Đoan Khôi