

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



VƯƠNG VIỆT HÙNG

**PHÁT HIỆN VÀ NHẬN DIỆN HÀNH VI QUA VIDEO ỨNG DỤNG
TRONG GIÁM SÁT PHÒNG THỰC HÀNH**

ĐỀ ÁN TỐT NGHIỆP THẠC SĨ KỸ THUẬT

HÀ NỘI - 2025

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



VƯƠNG VIỆT HÙNG

**PHÁT HIỆN VÀ NHẬN DIỆN HÀNH VI QUA VIDEO ỨNG DỤNG
TRONG GIÁM SÁT PHÒNG THỰC HÀNH**

Chuyên ngành: Hệ thống thông tin

Mã số: 8.48.01.04

ĐỀ ÁN TỐT NGHIỆP THẠC SĨ KỸ THUẬT
(Theo định hướng ứng dụng)

NGƯỜI HƯỚNG DẪN KHOA HỌC:

PGS. TS. NGUYỄN TRỌNG KHÁNH

HÀ NỘI - 2025

LỜI CAM ĐOAN

Tôi cam đoan đây là công trình nghiên cứu của riêng tôi.

Các số liệu, kết quả nêu trong đề án là trung thực và chưa từng được công bố trong bất kỳ công trình nào khác.

Hà Nội, tháng 7 năm 2025

Tác giả đề án tốt nghiệp
(ký và ghi rõ họ tên)



Vương Việt Hùng

LỜI CẢM ƠN

Tôi xin trân trọng cảm ơn các thầy cô trong Khoa công nghệ thông tin đã tạo điều kiện cho tôi một môi trường học tập tốt, đồng thời truyền đạt cho tôi một vốn kiến thức quý báu, một tư duy khoa học để phục vụ cho quá trình học tập và công tác của tôi.

Tôi xin gửi lời cảm ơn đến các bạn trong lớp Cao học Hệ thống thông tin M23CQIS01-B khóa 2022- 2024 đã giúp đỡ tôi trong suốt thời gian học tập vừa qua.

Đặc biệt, tôi xin được bày tỏ lòng biết ơn sâu sắc đến PGS. TS. Nguyễn Trọng Khánh đã tận tình chỉ bảo cho tôi trong suốt quá trình học tập và nghiên cứu, giúp tôi có nhận thức đúng đắn về kiến thức khoa học, tác phong học tập và làm việc, tạo điều kiện thuận lợi để tôi hoàn thành đề án này.

Cuối cùng, tôi xin được gửi lời cảm ơn tới gia đình, đồng nghiệp, người thân đã động viên, giúp đỡ tôi trong quá trình hoàn thành đề án.

Hà Nội, tháng 7 năm 2025

Tác giả đề án tốt nghiệp
(ký và ghi rõ họ tên)



Vương Việt Hùng

MỤC LỤC

LỜI CAM ĐOAN	i
LỜI CẢM ƠN	ii
MỤC LỤC	iii
DANH MỤC HÌNH VẼ	viii
MỞ ĐẦU	1
CHƯƠNG 1 - CƠ SỞ LÝ THUYẾT, THỰC NGHIỆM, ĐÁNH GIÁ.....	4
1.1. Giới thiệu bài toán.....	4
1.2. Xác định hành vi bị cấm trong phòng thực hành máy tính.	4
1.3. Các phương pháp học máy, học sâu để phát hiện hành vi qua video.	4
1.3.1. Phân tích tư thế (Pose Estimation).....	5
1.3.2. Phân loại hành vi (Action Recognition).....	5
1.3.3. Phát hiện đối tượng (Object Detection)	6
1.4. Giới thiệu YOLO.....	7
1.4.1. Mô hình YOLO	7
1.4.2. Ưu điểm và nhược điểm của phát hiện đối tượng YOLO.....	9
1.4.3. Ứng dụng phù hợp:	10
1.4.4. Kiến trúc tổng thể.....	10
1.4.5. YOLOv5 Backbone: CSODarknet53	11
1.4.6. Cross Stage Partial Network (CSPNet) in YOLOv5	12
1.4.7. Kết Luận:.....	17
1.5. Kết Chương	18
CHƯƠNG 2 - PHÂN TÍCH THIẾT KẾ HỆ THỐNG	19

2.1. Tổng quan về nghiệp vụ	19
2.2. Phân tích yêu cầu chức năng:	20
2.3. Đặc tả usecase:	21
2.3.1. Chức năng đăng nhập	21
2.3.2. Chức năng đăng ký lịch thực hành	21
2.3.3. Chức năng xóa lịch thực hành đã đăng ký	22
2.3.4. Chức năng cập nhật lịch thực hành đã đăng ký	22
2.3.5. Chức năng xem lịch thực hành	23
2.3.6. Chức năng xem thông báo	23
2.3.7. Chức năng xem thông tin camera phòng thực hành	24
2.3.8. Chức năng phê duyệt lịch thực hành	26
2.3.9. Chức năng từ chối lịch thực hành	27
2.3.10. Chức năng tìm kiếm lịch thực hành	27
2.3.11. Chức năng quản lý thiết bị phòng thực hành	28
2.3.12. Chức năng quản lý phòng thực hành	29
2.3.13. Chức năng thống kê thiết bị phòng thực hành	29
2.4. Phân tích yêu cầu phi chức năng	30
2.4.1. Yêu cầu bảo mật	30
2.4.2. Yêu cầu hiệu năng	31
2.5. Thiết kế biểu đồ lớp thực thể hệ thống	31
2.6. Thiết kế cơ sở dữ liệu hệ thống	34
2.7. Kết chương	35
CHƯƠNG 3 - CÀI ĐẶT VÀ KIỂM THỬ HỆ THỐNG	37

Cấu trúc dự án	37
3.1. Mô hình triển khai hệ thống	37
3.2. Giới thiệu về MySQL:	38
3.3. Giới thiệu về Spring boot:	38
3.4. Đánh giá các chỉ số (Evaluation Metrics)	41
3.4.1. Intersection over Union (IoU)	41
3.4.2. Precision, Recall và F1 Score	41
3.4.3. Average Precision (AP) và Mean Average Precision (mAP)	42
3.4.4. Labelling	43
3.5. Thực nghiệm và đánh giá	44
3.5.1. Thực nghiệm: Quy trình thu thập và gán nhãn dữ liệu	44
3.5.2. Cấu hình tham số huấn luyện	45
3.5.3. Huấn luyện mô hình	46
3.5.4. Đánh giá	48
3.6. Giao diện hệ thống	52
3.6.1. Giao diện phía người dùng	52
3.7. Kết chương	53
KẾT LUẬN	55
DANH MỤC TÀI LIỆU THAM KHẢO	56
PHỤ LỤC 1	57
1.2 Giao diện phía quản trị viên	62

DANH MỤC CÁC THUẬT NGỮ, CHỮ VIẾT TẮT

Viết tắt	Tiếng Anh	Tiếng Việt
YOLOv5	You Only Look Once version 5	Mô hình AI nhận diện và phát hiện vật thể thời gian thực
JWT	JSON Web Token	Cơ chế xác thực và bảo mật thông tin trong giao tiếp client-server
RESTful API	Representational State Transfer Application Programming Interface	Giao diện lập trình ứng dụng theo tiêu chuẩn REST
React.js	React JavaScript	Thư viện JavaScript dùng xây dựng giao diện người dùng (UI)
Spring Boot	Spring Framework boot	Framework Java đơn giản hóa phát triển ứng dụng web
MySQL	My Structured Query Language	Hệ quản trị cơ sở dữ liệu quan hệ mã nguồn mở
WebSocket	Web Socket	Giao thức truyền thông hai chiều thời gian thực
API	Application Programming Interface	Là các phương thức/giao thức kết nối giữa ứng dụng và các thư viện dữ liệu.
UI	User Interface	Giao diện người dùng
AI	Artificial Intelligence	Trí tuệ nhân tạo được áp dụng để

		nhận diện và phát hiện hành vi
API	Application Programming Interface	Giao diện lập trình ứng dụng để frontend và backend giao tiếp
SQL	Structured Query Language	Ngôn ngữ truy vấn cơ sở dữ liệu

DANH MỤC HÌNH VẼ

Hình 1-1. Sơ đồ luồng suy luận mặc định của YOLOv5.....	11
Hình 1-2. SPPF	12
Hình 1-3. (a) DenseNet (b) CSP DenseNet	13
Hình 1-4. Cấu trúc mô-đun BottleNeckCSP:.....	14
Hình 1-5. (a) Without Feature Fusion, (b) FPN and (c) PAN.	15
Hình 1-6. YOLOv5: Neck (Within Dashed Box).....	15
Hình 1-7. Cấu trúc khối SPPF.	16
Hình 1-8. YOLOv5: Head	17
Hình 1-9. Bounding Box Regression(Hồi quy hộp giới hạn).....	17
Hình 2-1. Biểu đồ usecase tổng quan của hệ thống.....	20
Hình 2-2. Biểu đồ lớp thực thể của hệ thống.....	32
Hình 2-3. Cơ sở dữ liệu hệ thống.....	35
Hình PL 1-18. Giao diện màn hình thêm mới phần mềm.....	65
Hình PL 1-21. Giao diện chỉnh sửa phòng thực hành.	67

MỞ ĐẦU

1. Lý do chọn đề tài

Trong những năm gần đây, sự phát triển mạnh mẽ của công nghệ thông tin và trí tuệ nhân tạo (AI) đã tạo ra những bước tiến lớn trong việc tự động hóa quản lý và giám sát ở nhiều lĩnh vực. Tại các trường đại học và cơ sở giáo dục, đặc biệt là những đơn vị có phòng thực hành máy tính, việc ứng dụng công nghệ để quản lý, giám sát hiệu quả không chỉ giúp tối ưu hóa quy trình vận hành mà còn đảm bảo môi trường học tập đạt chất lượng cao.

Đảm bảo tuân thủ các quy định trong phòng thí nghiệm là điều cần thiết để duy trì một môi trường an toàn và hiệu quả. Tuy nhiên, các phương pháp giám sát truyền thống thường không đáp ứng được nhu cầu hiện đại. Việc giám sát thủ công bị giới hạn bởi sự sẵn có của nhân sự, dễ xảy ra sai sót và không đảm bảo được tính liên tục. Với sự phát triển của AI và học sâu, việc áp dụng các hệ thống tự động để phát hiện hành vi đã nổi lên như một phương pháp đầy hứa hẹn. Những hệ thống này không chỉ giảm bớt gánh nặng cho nguồn nhân lực mà còn cải thiện độ chính xác và hiệu quả của việc giám sát.

Đề tài “phát hiện và nhận diện hành vi qua video ứng dụng trong giám sát phòng thực hành ” hướng đến việc phát triển một hệ thống thông minh, tích hợp công nghệ AI (Yolov5) để quản lý và giám sát toàn diện các hoạt động tại phòng thực hành.

Một trong những điểm nổi bật của hệ thống là khả năng giám sát tự động thông qua mô hình trí tuệ nhân tạo YOLOv5 (You Only Look Once). Đây là một mô hình học sâu tiên tiến, được thiết kế để nhận diện và phân loại các vật thể trong thời gian thực. Với tốc độ xử lý nhanh và độ chính xác cao, YOLOv5 được ứng dụng để phát hiện các hành vi vi phạm nội quy trong phòng thực hành, chẳng hạn như đi dép, sử dụng điện thoại, hoặc không tuân thủ quy định về trang phục. Điều

này không chỉ đảm bảo môi trường học tập đạt chất lượng mà còn giảm thiểu sự phụ thuộc vào giám sát thủ công.

Hệ thống được phát triển dưới dạng ứng dụng web, với giao diện trực quan, dễ sử dụng và khả năng truy cập linh hoạt. Sự kết hợp giữa công nghệ web và mô hình YOLOv5 mang lại khả năng giám sát tự động hóa cao, giảm thiểu sự can thiệp thủ công và tăng hiệu quả quản lý.

Với tiềm năng lớn, đề tài này không chỉ đáp ứng nhu cầu quản lý và giám sát mà còn góp phần thúc đẩy việc ứng dụng AI vào thực tiễn, tạo ra môi trường học tập và làm việc hiện đại, chuyên nghiệp.

2. Câu hỏi và mục đích nghiên cứu

Làm thế nào để phát triển một hệ thống sử dụng AI có khả năng tự động phát hiện và giám sát các hành vi vi phạm quy tắc trong phòng thí nghiệm? Nghiên cứu tập trung vào việc đạt được các mục tiêu sau:

- Phát triển một hệ thống giám sát thông minh sử dụng mô hình học sâu YOLOv5 để phát hiện các đối tượng và hành vi.
- Phân loại chính xác các vi phạm quy tắc, như vi phạm quy định về giày dép, trang phục không phù hợp, sử dụng điện thoại, gian lận học tập và tư thế không đúng.
- Triển khai một hệ thống thời gian thực có khả năng phát cảnh báo khi phát hiện vi phạm, đồng thời lưu trữ dữ liệu để phân tích sau này. Để đạt được các mục tiêu này, điều cần thiết là phải hiểu rõ các công trình nghiên cứu hiện có liên quan đến phát hiện hành vi và phát hiện vi phạm quy tắc. Phần tiếp theo sẽ xem xét các nghiên cứu và phương pháp đã thực hiện, có liên quan đến việc phát triển các hệ thống dựa trên AI để giám sát vi phạm quy tắc trong các môi trường chuyên biệt như phòng thí nghiệm. Bằng cách phân tích các công trình liên quan, nghiên cứu này sẽ xác định những thách thức, kỹ thuật và khoảng trống trong lĩnh vực, tạo nền tảng cho hệ thống được đề xuất.

3. Phương pháp nghiên cứu

Phương pháp nghiên cứu là nghiên cứu lý thuyết kết hợp với thực nghiệm và phân tích, đánh giá kết quả.

CHƯƠNG 1 - CƠ SỞ LÝ THUYẾT, THỰC NGHIỆM, ĐÁNH GIÁ

1.1. Giới thiệu bài toán

Phòng thực hành máy tính là môi trường quan trọng trong giáo dục, nơi sinh viên thực hiện các bài tập thực hành, nghiên cứu và kiểm tra. Tuy nhiên, việc giám sát phòng thực hành là một thách thức lớn do số lượng người tham gia đông, tính chất đa dạng của các hoạt động bởi vậy nguy cơ vi phạm nội quy là không thể tránh khỏi.

Chính vì vậy sử dụng video giám sát kết hợp với công nghệ trí tuệ nhân tạo (AI), cụ thể là học máy (Machine Learning) và học sâu (Deep Learning), có thể giúp:

- Theo dõi hoạt động của sinh viên theo thời gian thực.
- Phát hiện các hành vi vi phạm nội quy, quy định trong phòng thực hành máy tính như trang phục chưa đúng quy định, tư thế chưa đúng quy định
- Tăng cường an toàn và quản lý hiệu quả hơn phòng thực hành máy tính.

1.2. Xác định hành vi bị cấm trong phòng thực hành máy tính.

Việc giám sát nhằm phát hiện và ngăn chặn các hành vi vi phạm, bao gồm:

- Gian lận trong kiểm tra, thi cử: Trao đổi bài, sử dụng tài liệu trái phép. hoặc dùng điện thoại bằng cách phân tích hành vi và vị trí của tay và mặt.
- Gây mất trật tự: Nói chuyện riêng, di chuyển nhiều, ăn uống trong phòng thực hành.
- Vi phạm về giày dép: Để đảm bảo phòng thực hành luôn được sạch sẽ.
- Vi phạm tư thế: Hệ thống giám sát video có thể phát hiện sinh viên quay sang trao đổi bài

1.3. Các phương pháp học máy, học sâu để phát hiện hành vi qua video.

Hệ thống giám sát thông minh sử dụng nhiều thuật toán để phát hiện hành vi, trong đó phổ biến nhất là:

1.3.1. Phân tích tư thế (Pose Estimation)

- Phân tích tư thế là một kỹ thuật trong thị giác máy tính (Computer Vision) nhằm xác định vị trí và hướng của các bộ phận trên cơ thể người hoặc vật trong không gian 2D hoặc 3D. Công nghệ này được ứng dụng rộng rãi trong thể thao, y tế, thực tế ảo (VR), giám sát an ninh, và điều khiển cử chỉ.

- Ứng dụng: phát hiện, gian lận trong thi cử.
- Quy trình phân tích tư thế:

- Dữ liệu đầu vào có thể đến từ: Camera RGB thông thường (webcam, điện thoại), camera chiều sâu (Depth camera) như Microsoft Kinect, Intel RealSense, cảm biến chuyển động (IMU – Inertial Measurement Unit) trong các thiết bị đeo.

- Phát hiện đối tượng (Object Detection): xác định vị trí người hoặc vật trong ảnh/video bằng thuật toán phát hiện đối tượng như: YOLO (You Only Look Once), Faster R-CNN, SSD (Single shot Detector)

- Phát hiện và xác định Keypoints: là các điểm quan trọng trên cơ thể như đầu, vai, khuỷu tay, đầu gối hay cổ tay.... Các mô hình thường được dùng: OpenPose (CMU) xác định 25 điểm trên cơ thể, PoseNet(Google) nhận diện 17 điểm chính trên cơ thể hay như là MediaPipe Pose (Google) nhận diện 33 điểm với hiệu suất cao.

- Xây dựng bộ khung cơ thể (Skeleton Reconstruction): kết nối các điểm vừa phát hiện để tạo thành mô hình khung xương 2D hoặc 3D, áp dụng mô hình dự đoán để làm mịn kết quả và loại bỏ nhiễu.

1.3.2. Phân loại hành vi (Action Recognition)

- Là một nhánh của trí tuệ nhận tạo (AI) và thị giác máy tính(Computer Vision) nó giúp xác định và phân loại hành động của con người từ dữ liệu cảm biến hoặc hình ảnh/video. Công nghệ này ứng dụng rộng rãi trong giám sát an ninh, chăm sóc sức khỏe, thể thao, hay thực tế ảo (VR)

- Quy trình phân loại hành vi:

- Thu thập dữ liệu: Dữ liệu cảm biến: gia tốc kế (Accelerometer), con quay hồi chuyển (Gyroscope), cảm biến từ trường (Magnetometer) từ điện thoại, smartwatch, cảm biến áp suất hoặc ECG dùng trong y tế. Dữ liệu hình ảnh/video: từ camera giám sát, camera điện thoại, drone ghi lại hình ảnh hoặc video hành động của con người.

- Tiền xử lý dữ liệu với dữ liệu cảm biến cần loại bỏ nhiễu và chuẩn hoá dữ liệu. Với dữ liệu hình ảnh/video cần phải trích xuất khung hình (frame extraction) xử lý hình ảnh, giảm nhiễu, căn chỉnh độ sáng.

- Trích xuất đặc trưng với dữ liệu cảm biến là tốc độ thay đổi gia tốc (Jerk) và biến thiên góc quay, độ lệch chuẩn, giá trị đỉnh, tần số phổ. Với dữ liệu hình ảnh/video phát hiện khung xương người (Pose Estimation) bằng OpenPose, MediaPipe và trích xuất đặc trưng bằng CNN (Convolutional Neural Network)

- Phân loại hành vi: với dữ liệu cảm biến sử dụng thuật toán học máy như: Hồi quy (logistic Regression), mô hình cây quyết định (Decision Tree), rừng ngẫu nhiên (Random Forest), mạng nơ-ron nhân tạo (Deep Neural Network – DNN). Với dữ liệu hình ảnh/video dùng mô hình học sâu như: CNN kết hợp RNN (LSTM, GRU) để học chuỗi hành động, 3D-CNN, I3D (Inflated 3D ConvNet) để phân tích video.

1.3.3. Phát hiện đối tượng (Object Detection)

- Là quá trình xác định, phân loại, theo dõi và trích xuất thông tin từ đối tượng trong hình ảnh hoặc video bằng trí tuệ nhân tạo (AI) và thị giác máy tính (Computer Vision)

- Dùng để xác định các vật thể như: điện thoại, sách vở, tài liệu, tai nghe.
- Ứng dụng: phát hiện sinh viên sử dụng thiết bị trái phép.
- Mô hình phổ biến: YOLO (You Only Look Once), Faster R-CNN (Single Shot MultiBox Detector).

- Quy trình phân tích đối tượng

- Thu thập dữ liệu: với dữ liệu hình ảnh/video từ camera giám sát hoặc ảnh từ các thiết bị di động. Với dữ liệu cảm biến dùng xe tự hành hoặc robot

- Tiền xử lý dữ liệu với dữ liệu là ảnh phải chuyển ảnh sang thang độ xám (grayscale), cân bằng ánh sáng, giảm nhiễu, phát hiện cạnh bằng Sobel, Canny. Với dữ liệu là video thì chia video thành từng khung hình và ổn định video để giảm rung lắc.

- Phát hiện đối tượng với các phương pháp truyền thống thì dùng HOG (Histogram of Oriented Gradients) + SVM hay Haar Cascade (dùng trong phát hiện khuôn mặt). Với các phương pháp hiện đại ta dùng (Deep Learning) như YOLO cho tốc độ nhanh và thời gian thực, Faster R-CNN có độ chính xác cao (phù hợp với y tế hay các lĩnh vực đòi hỏi độ chính xác cao), SSD (Single Shot MultiBox Detector) nó cân bằng giữa tốc độ và độ chính xác.

- Phân loại đối tượng: Sử dụng mạng nơ-ron tích chập (CNN) để gán nhãn đối tượng. Các mô hình phổ biến: ResNet, VGG, EfficientNet dùng để phân loại ảnh tĩnh, với Vision Transformer (ViT) dùng Transformer cho hình ảnh

- Nhận diện đặc điểm và trích xuất thông tin: để nhận diện màu sắc, kích thước hay hình ảnh ta dùng: OpenCV hoặc mạng Neural. Trích xuất văn bản từ đối tượng OCR (Optical Character Recognition) như Tesserract OCR.

1.4. Giới thiệu YOLO

1.4.1. Mô hình YOLO

YOLO được giới thiệu đến cộng đồng thị giác máy tính qua một bài báo vào năm 2015 bởi Joseph Redmon và các cộng sự, mang tên "You Only Look Once: Unified, Real-Time Object Detection". Bài báo này đã làm mới cách tiếp cận phát hiện đối tượng, coi nó như một bài toán hồi quy một lần duy nhất, bắt đầu từ các pixel hình ảnh và di chuyển đến các xác suất hộp bao quanh và lớp. Phương pháp được đề xuất dựa trên khái niệm "hộp nhất" cho phép dự đoán đồng thời nhiều hộp bao quanh và xác suất lớp, cải thiện cả tốc độ lẫn độ chính xác.

Kể từ khi ra đời vào năm 2016 đến năm 2020, dòng YOLO đã liên tục phát triển với tốc độ nhanh chóng. Mặc dù tác giả ban đầu (Joseph Redmon) đã dừng công việc trong lĩnh vực thị giác máy tính với YOLO-v3, nhưng tính hiệu quả và tiềm năng của khái niệm "hộp nhất" cốt lõi đã được phát triển thêm bởi nhiều tác giả khác.

Nguyên lý cốt lõi được đề xuất bởi YOLO-v1 là việc áp dụng một lưới ô có kích thước $s \times s$ lên hình ảnh. Trong trường hợp tâm của đối tượng quan tâm rơi vào một trong các ô lưới, ô lưới đó sẽ chịu trách nhiệm phát hiện đối tượng đó. Điều này cho phép các ô khác bỏ qua đối tượng đó trong trường hợp đối tượng xuất hiện nhiều lần.

Để thực hiện phát hiện đối tượng, mỗi ô lưới sẽ dự đoán B hộp bao quanh cùng với kích thước và điểm tin cậy. Điểm tin cậy này chỉ ra sự có mặt hoặc không có mặt của đối tượng trong hộp bao quanh. Do đó, điểm tin cậy có thể được biểu diễn bằng công thức sau:

$$\text{Confidence Score} = P(\text{Object}) \times \text{IoU}_{\text{pred truth}}$$

Trong đó, $P(\text{Object})$ biểu thị xác suất đối tượng có mặt, với phạm vi từ 0–1, trong đó 0 chỉ ra đối tượng không có mặt, và $\text{IoU}_{\text{pred truth}}$ đại diện cho chỉ số giao nhau trên toàn bộ (intersection-over-union) giữa hộp bao quanh dự đoán và hộp bao quanh đúng.

Bảng 1.1 So sánh phiên bản YOLO

Phiên bản YOLO	Năm	Cải tiến	Tổng hợp số lớp trong mạng
YOLO v2	2017	Bao gồm bộ phân giải đơn và hộp neo	5 lớp max-pooling, 19 lớp tích chập (Darknet)
YOLO v3	2018	Cải thiện hiệu suất trên các	Tăng lên 106 lớp

		đối tượng nhỏ	
YOLO v4	2019	Tối ưu hóa tốc độ và độ chính xác	53 lớp tích chập với kích thước cụ thể
YOLO v5	2020	Giới thiệu tăng cường mosaic	80 lớp phát hiện (Chi tiết lớp không được chỉ rõ)

1.4.2. Ưu điểm và nhược điểm của phát hiện đối tượng YOLO.

a. Ưu điểm.

- Tốc độ nhanh: YOLO được thiết kế để xử lý nhanh chóng, phù hợp với các ứng dụng yêu cầu phát hiện thời gian thực như giám sát video, điều hướng xe tự lái, và phân tích bán lẻ.

- Tích hợp dự đoán duy nhất: Phương pháp dự đoán đồng thời nhiều hộp bao quanh và lớp giúp giảm thiểu thời gian tính toán so với các phương pháp khác.

- Đơn giản và dễ triển khai: YOLO sử dụng một hệ thống mạng duy nhất cho tất cả các bước dự đoán, giúp đơn giản hóa quá trình phát triển và triển khai.

- Hiệu quả với các ứng dụng yêu cầu phát hiện nhanh chóng: YOLO phù hợp cho các tình huống cần phát hiện đối tượng nhanh chóng và hiệu quả, như đếm người hoặc đối tượng trong cửa hàng.

b. Nhược điểm.

- Khó khăn với các đối tượng nhỏ: Hệ thống dự đoán dựa trên lưới có thể gặp khó khăn khi phát hiện các đối tượng nhỏ, đặc biệt khi các đối tượng này chiếm chỉ một số ô lưới nhỏ.

- Hộp bao quanh không chính xác: Các hộp bao quanh do YOLO dự đoán có thể kém chính xác hơn so với các phương pháp dựa trên khu vực như Faster R-CNN, đặc biệt là với các đối tượng có hình dạng không đều hoặc trong các cảnh đông đúc.

- Xử lý hậu kỳ tổn tài nguyên: YOLO sử dụng các kỹ thuật xử lý hậu kỳ như loại bỏ hộp bao quanh dư thừa (non-maximum suppression), điều này có thể gây ra một chút chi phí tính toán và độ chính xác không hoàn hảo.

- Độ chính xác không đủ cho các ứng dụng yêu cầu định vị chính xác: Mặc dù YOLO nổi bật về tốc độ, nhưng độ chính xác của nó có thể không đủ cho các ứng dụng đòi hỏi định vị chính xác như chẩn đoán hình ảnh y tế.

- Yêu cầu phần cứng mạnh mẽ: Để đạt được hiệu suất thời gian thực tối ưu với các đầu vào có độ phân giải cao, YOLO có thể cần phần cứng GPU mạnh mẽ hoặc phần cứng tối ưu hóa.

1.4.3. Ứng dụng phù hợp:

YOLO thích hợp với các ứng dụng ưu tiên tốc độ thay vì độ chính xác tuyệt đối. Ví dụ như giám sát video thời gian thực, điều hướng xe tự lái và phân tích bán lẻ cho việc đếm người hoặc đối tượng. Mặc dù hiệu suất có thể không tốt trong các cảnh phức tạp hoặc đông đúc, YOLO là một lựa chọn đáng tin cậy trong các tình huống cần phát hiện đối tượng nhanh chóng và hiệu quả. Tùy thuộc vào yêu cầu cụ thể của dự án – chẳng hạn như yêu cầu độ chính xác chi tiết hoặc xử lý đối tượng nhỏ – YOLO có thể cần được kết hợp với các phương pháp phát hiện khác hoặc cải thiện thông qua huấn luyện thêm.

1.4.4. Kiến trúc tổng thể

Kiến trúc tổng thể của YOLOv5 được thiết kế để phát hiện đối tượng hiệu quả và chính xác bằng cách tích hợp nhiều thành phần chính. Đầu tiên, hình ảnh đầu vào được xử lý qua lớp đầu vào và chuyển đến phần backbone, nơi các bản đồ đặc trưng (feature maps) với kích thước khác nhau được trích xuất. Các bản đồ đặc trưng này sau đó được kết hợp trong mạng feature fusion, còn được gọi là "neck", để tạo ra ba bản đồ đặc trưng khác nhau—P3, P4 và P5—with kích thước lần lượt là 80×80 , 40×40 và 20×20 . Mỗi bản đồ đặc trưng được thiết kế để phát hiện các đối tượng có kích thước khác nhau: nhỏ, vừa và lớn.

Tiếp theo, các bản đồ đặc trưng này được gửi đến prediction head, nơi các điểm tin cậy (confidence scores) và các phép hồi quy hộp bao quanh (bounding-box regressions) được tính toán cho mỗi pixel sử dụng các neo ưu tiên (prior anchors) đã được thiết lập trước. Quá trình này tạo ra một mảng đa chiều (BBboxes) chứa lớp đối tượng, điểm tin cậy, tọa độ hộp bao quanh và thông tin kích thước. Cuối cùng, các ngưỡng như điểm tin cậy (confthreshold) và sự có mặt của đối tượng (objthreshold) được áp dụng để lọc thông tin không liên quan, tiếp theo là bước non-maximum suppression (NMS) để loại bỏ các phát hiện dư thừa.

Kiến trúc đơn giản hóa này giúp YOLOv5 cung cấp kết quả phát hiện đối tượng chính xác và hiệu suất cao.

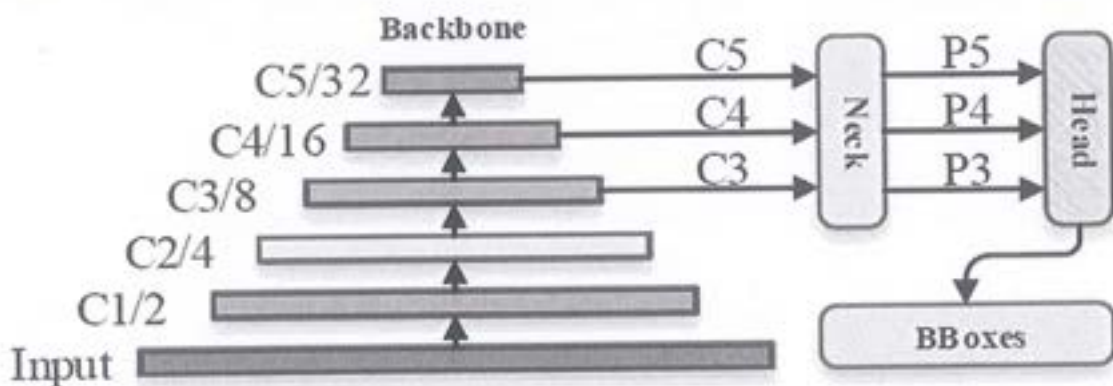


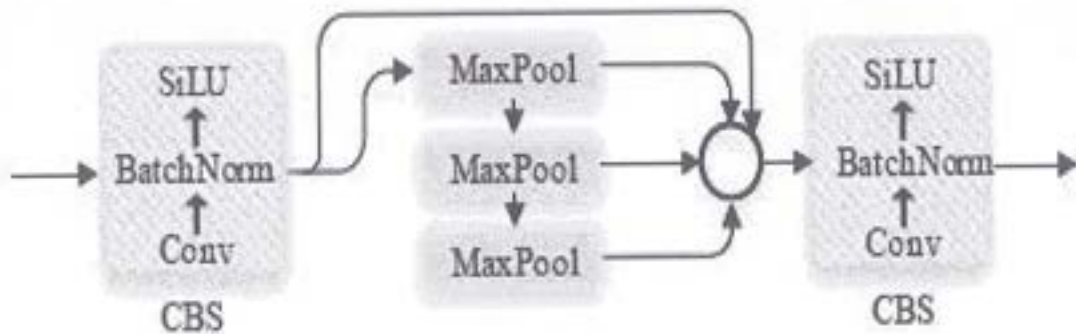
Figure 1. The default inference flowchart of YOLOv5.

Hình 1-1. Sơ đồ luồng suy luận mặc định của YOLOv5

1.4.5. YOLOv5 Backbone: CSPDarknet53

Backbone của YOLOv5 là CSPDarknet53, một phiên bản nâng cao của kiến trúc Darknet53 được giới thiệu ban đầu trong YOLOv3. Bằng cách tích hợp chiến lược Cross Stage Partial Network (CSPNet), CSPDarknet53 giải quyết hiệu quả các thách thức như thông tin gradient dư thừa trong khi vẫn duy trì việc tái sử dụng các đặc trưng một cách dày đặc. Kiến trúc của backbone chủ yếu bao gồm việc xếp chồng các mô-đun CBS (Convolution + Batch Normalization + SiLU activation) và các mô-đun C3, sau đó là mô-đun Spatial Pyramid Pooling-Fast (SPPF) ở cuối.

Các mô-đun CBS hỗ trợ các mô-đun C3 trong việc trích xuất đặc trưng bằng cách cung cấp cấu trúc cơ bản, trong khi các mô-đun C3 áp dụng nguyên lý CSPNet để phân chia và kết hợp các bản đồ đặc trưng qua các giai đoạn. Quá trình phân tích theo cấp bậc này giúp giảm dòng gradient dư thừa và cải thiện hiệu quả tính toán. Mô-đun SPPF cuối cùng nâng cao khả năng biểu diễn của các đặc trưng đã được trích xuất bằng cách áp dụng các phép max-pooling song song. Khác với phiên bản tiền nhiệm là mô-đun SPP trong SPPNet, mô-đun SPPF tránh các phép toán dư thừa bằng cách pooling trên các đặc trưng đã được pooling, dẫn đến tốc độ xử lý nhanh hơn đáng kể.



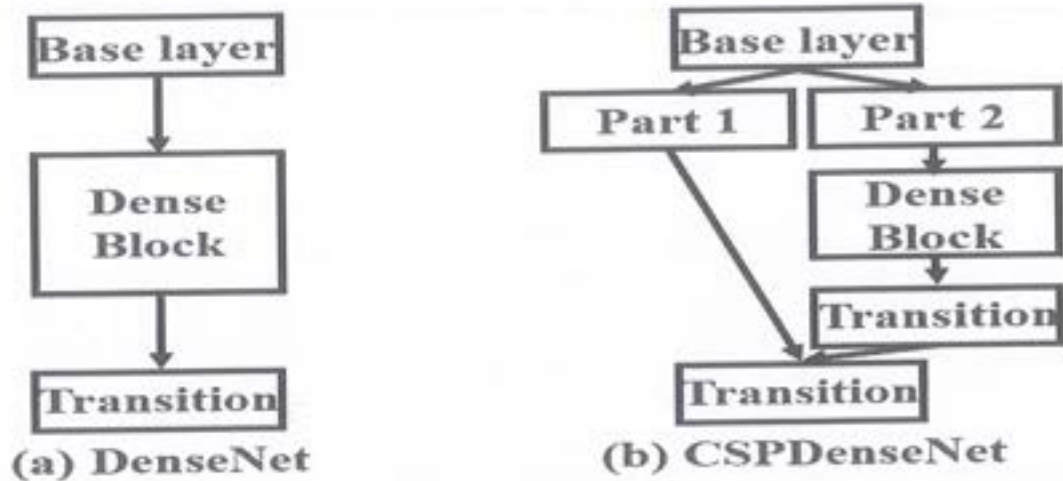
(a) SPPF

Hình 1-2. SPPF

1.4.6. Cross Stage Partial Network (CSPNet) in YOLOv5

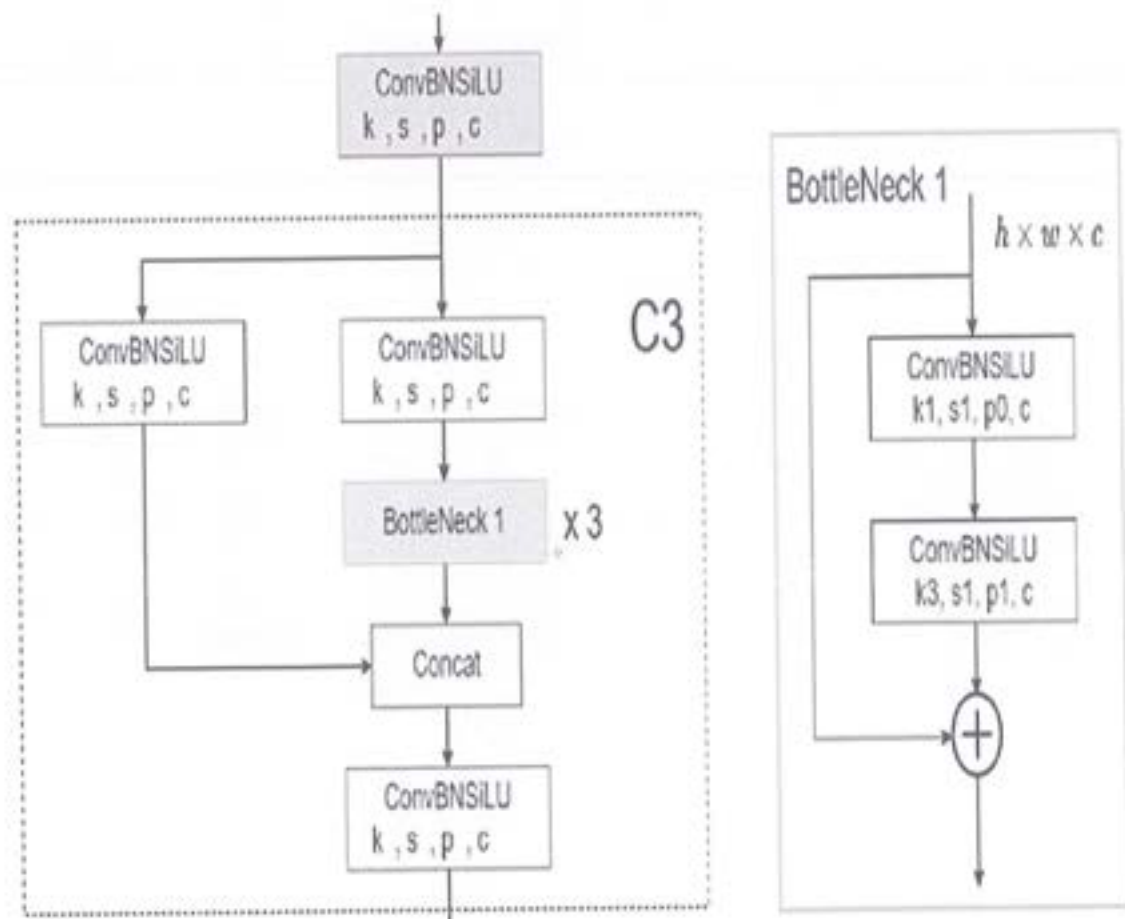
YOLOv5 tận dụng CSPNet để tối ưu hóa kiến trúc của mô hình học sâu. Các mạng sâu như YOLO thường dựa vào các khối residual và dense để tạo điều kiện cho việc truyền thông tin vào các lớp sâu hơn, từ đó giảm thiểu vấn đề hiện tượng tiêu biến gradient. Tuy nhiên, các khối này có thể dẫn đến sự dư thừa thông tin gradient quá mức. CSPNet giải quyết vấn đề này bằng cách cắt bỏ dòng gradient, điều này không chỉ bảo tồn lợi ích của việc tái sử dụng đặc trưng như trong DenseNet, mà còn giảm thiểu sự dư thừa tính toán, nhờ cách tiếp cận này, CSPNet

giúp giảm bớt độ phức tạp tính toán và cải thiện hiệu quả xử lý, đồng thời duy trì chất lượng đặc trưng đã trích xuất, mang lại sự cân bằng giữa hiệu suất và tốc độ.



Hình 1-3. (a) DenseNet (b) CSP DenseNet

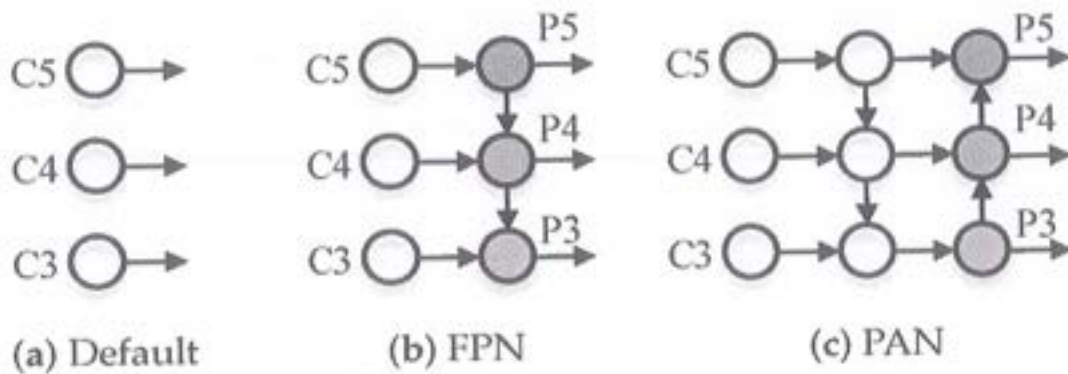
Trong CSPNet, bản đồ đặc trưng của lớp cơ sở được phân thành hai phần và kết hợp thông qua một hệ thống phân cấp xuyên lớp. Chiến lược này giúp giảm đáng kể số lượng tham số và phép toán dấu chấm động (FLOPS), từ đó tăng tốc độ suy luận một yếu tố quan trọng đối với các hệ thống phát hiện đối tượng thời gian thực. Bằng cách áp dụng phương pháp này, CSPDarknet53 giúp YOLOv5 đạt được sự cân bằng giữa độ phức tạp của mô hình và hiệu suất.



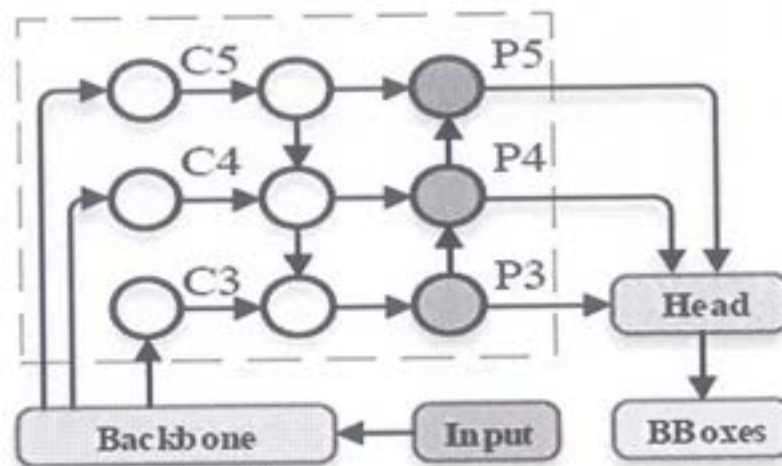
Hình 1-4. Cấu trúc mô-đun BottleNeckCSP:.

a. YOLOv5 Cổ (Neck).

Cổ của YOLOv5 đóng vai trò quan trọng trong việc tăng cường kết hợp các đặc trưng và cải thiện hiệu suất phát hiện trên nhiều tỷ lệ đối tượng. YOLOv5 kết hợp hai cải tiến chính trong thiết kế cổ của nó: một Mạng Tập hợp Đường dẫn (PANet) được sửa đổi và một mô-đun Cấp tầng Kim tự tháp Không gian Nhanh (SPPF) cập nhật. PANet, được giới thiệu lần đầu tiên trong YOLOv4, hỗ trợ dòng thông tin tốt hơn và định vị chính xác hơn bằng cách tập hợp các đặc trưng giữa các lớp. Trong YOLOv5, PANet được tinh chỉnh thêm bằng cách tích hợp chiến lược Mạng Phân đoạn Xuyên lớp (CSPNet), giúp nâng cao hiệu quả bằng cách giảm thiểu tính toán dư thừa trong khi vẫn duy trì việc tái sử dụng các đặc trưng.

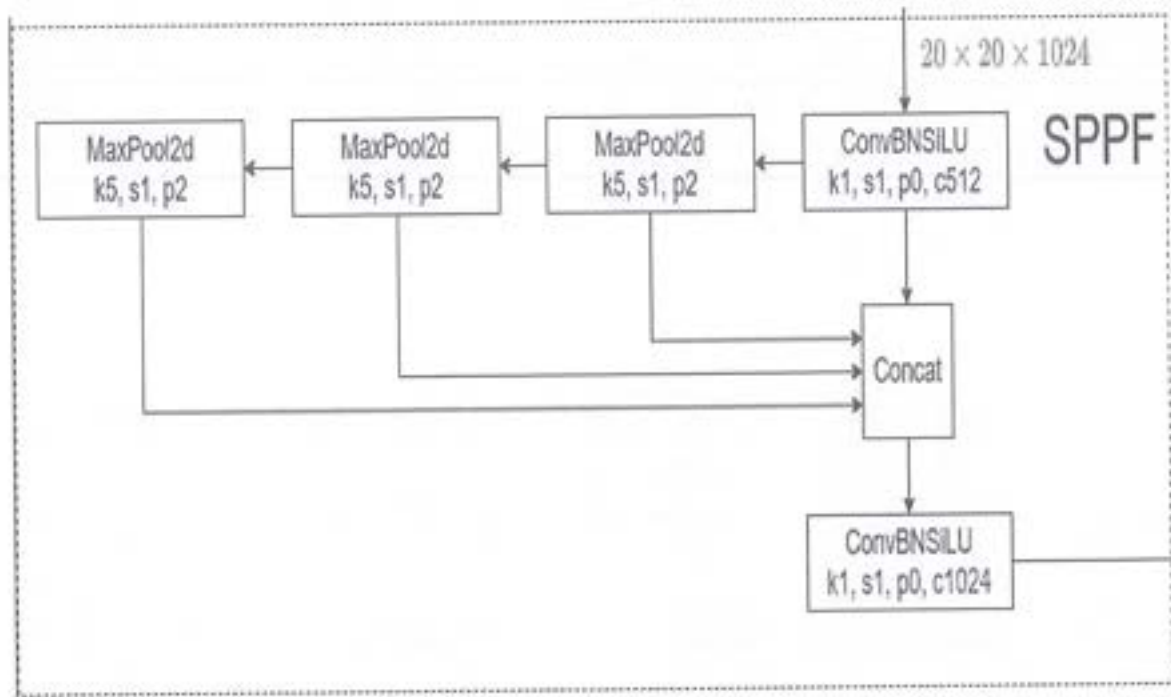


Hình 1-5. (a) Without Feature Fusion, (b) FPN and (c) PAN.



Hình 1-6. YOLOv5: Neck (Within Dashed Box)

Mô-đun SPPF, một biến thể của khối Kim tự tháp Không gian (SPP) được sử dụng trong các phiên bản YOLO trước, được thiết kế để tổng hợp thông tin bối cảnh và tăng đáng kể trường nhận thức mà không làm giảm tốc độ mạng. Bằng cách thay thế khối SPP truyền thống bằng SPPF, YOLOv5 đạt được xử lý nhanh hơn trong khi vẫn giữ khả năng phân tách các đặc trưng bối cảnh quan trọng nhất. Cốt lõi của YOLOv5 cũng sử dụng sự kết hợp giữa Mạng Kim tự tháp Đặc trưng (FPN) và kỹ thuật PANet để tăng kích thước và kết hợp các bản đồ đặc trưng (P3, P4, và P5) được tạo ra từ xương sống (backbone). Thiết kế này đảm bảo việc kết hợp đặc trưng đa tỷ lệ hiệu quả, giúp YOLOv5 phát hiện các đối tượng có kích thước khác nhau với độ chính xác cao hơn.



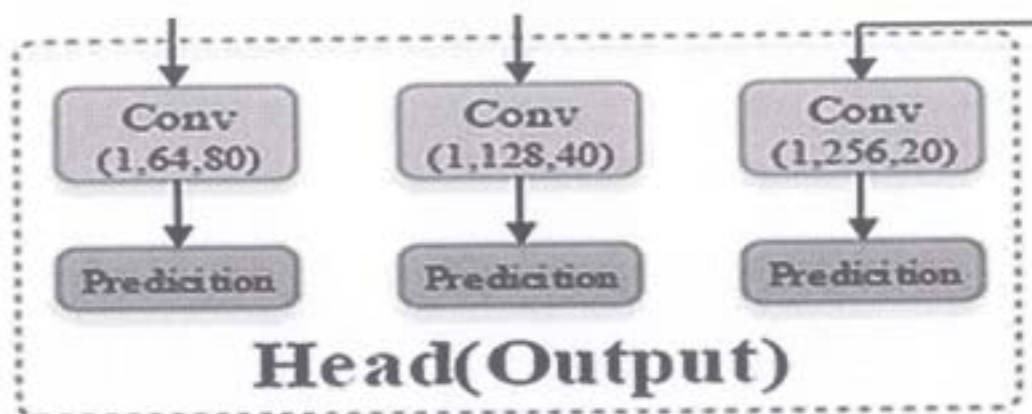
Hình 1-7. Cấu trúc khối SPPF.

b. YOLOv5 Đầu (Head).

Đầu của YOLOv5 chịu trách nhiệm cho việc hồi quy giới hạn, phân loại đối tượng và dự đoán điểm tin cậy. Nó giữ cấu trúc tương tự như các phiên bản trước, YOLOv3 và YOLOv4, bao gồm ba lớp tích chập để dự đoán tọa độ hộp giới hạn (x , y , chiều rộng, chiều cao), các lớp đối tượng và điểm tin cậy tương ứng. Tuy nhiên, YOLOv5 giới thiệu các cập nhật đối với các phương trình dùng để tính toán tọa độ mục tiêu của hộp giới hạn, làm tinh chỉnh quá trình này để cải thiện độ chính xác và hiệu suất định vị.

Quá trình hồi quy hộp giới hạn bắt đầu bằng việc đặt giá trị tọa độ của góc trên bên trái của bàn đồ đặc trưng là $(0, 0)$. Các tọa độ điểm trung tâm dự đoán chưa điều chỉnh $((r_x, r_y))$ sau đó sẽ được điều chỉnh bằng các độ dịch $((s_x, s_y))$ được tính toán bởi mô hình. Các tọa độ và kích thước đã điều chỉnh của hộp dự đoán cuối cùng $((g_x, g_y, g_w, g_h))$ được suy ra bằng cách làm tinh chỉnh thông tin neo (anchor) trước đó $((p_w, p_h))$ mà mạng cung cấp. Điều chỉnh này đảm bảo sự căn chỉnh chính xác của các hộp giới hạn với các đối tượng được phát hiện. Bằng cách

kết hợp các tinh chỉnh này, YOLOv5 đạt được độ chính xác phát hiện cao hơn và duy trì hiệu quả tính toán trong phần đầu dự đoán của nó.



Hình 1-8. YOLOv5: Head

$$g_x = 2\sigma(s_x) - 0.5 + r_x$$

$$g_y = 2\sigma(s_y) - 0.5 + r_y$$

$$g_h = p_h(2\sigma(s_h))^2$$

$$g_w = p_w(2\sigma(s_w))^2$$

Hình 1-9. Bounding Box Regression(Hồi quy hộp giới hạn)

1.4.7. Kết Luận:

Việc sử dụng YOLOv5 để phát hiện các vi phạm trong phòng thí nghiệm cho thấy kết quả hứa hẹn, đạt điểm F1 tổng thể là 0.75 với ngưỡng độ tin cậy 0.390. Trong khi các lớp như "trang phục không phù hợp" và "tư thế không phù hợp" hoạt động tốt, các lớp khác như "gian lận học thuật" lại có độ chính xác thấp do dữ liệu huấn luyện hạn chế và thường xuyên gặp phải các trường hợp dương tính giả. Những sai sót trong phân loại, như việc nền bị gán nhãn là vi phạm, làm nổi bật nhu cầu cải thiện bộ dữ liệu thông qua việc mở rộng chú thích và cân bằng phân phối lớp. Các cải tiến trong tương lai bao gồm tăng cường dữ liệu, học chuyển giao và triển khai hệ thống trong môi trường thực tế với các bộ dữ liệu lớn hơn, đa dạng hơn để xác minh tính vững chắc và khả năng ứng dụng của nó.

1.5. Kết Chương

Trong chương này, chúng ta đã giới thiệu bài toán giám sát hành vi trong phòng thực hành máy tính thông qua video. Chúng ta cũng đã xác định một số hành vi bị cấm theo nội quy, quy định của phòng thực hành máy tính. Bên cạnh đó, chương cũng trình bày tổng quan về các phương pháp học máy và học sâu phổ biến trong việc phát hiện hành vi qua video giám sát. Đây là nền tảng quan trọng để phát triển các hệ thống giám sát thông minh, góp phần nâng cao hiệu quả quản lý trong môi trường học tập.

CHƯƠNG 2 - PHÂN TÍCH THIẾT KẾ HỆ THỐNG

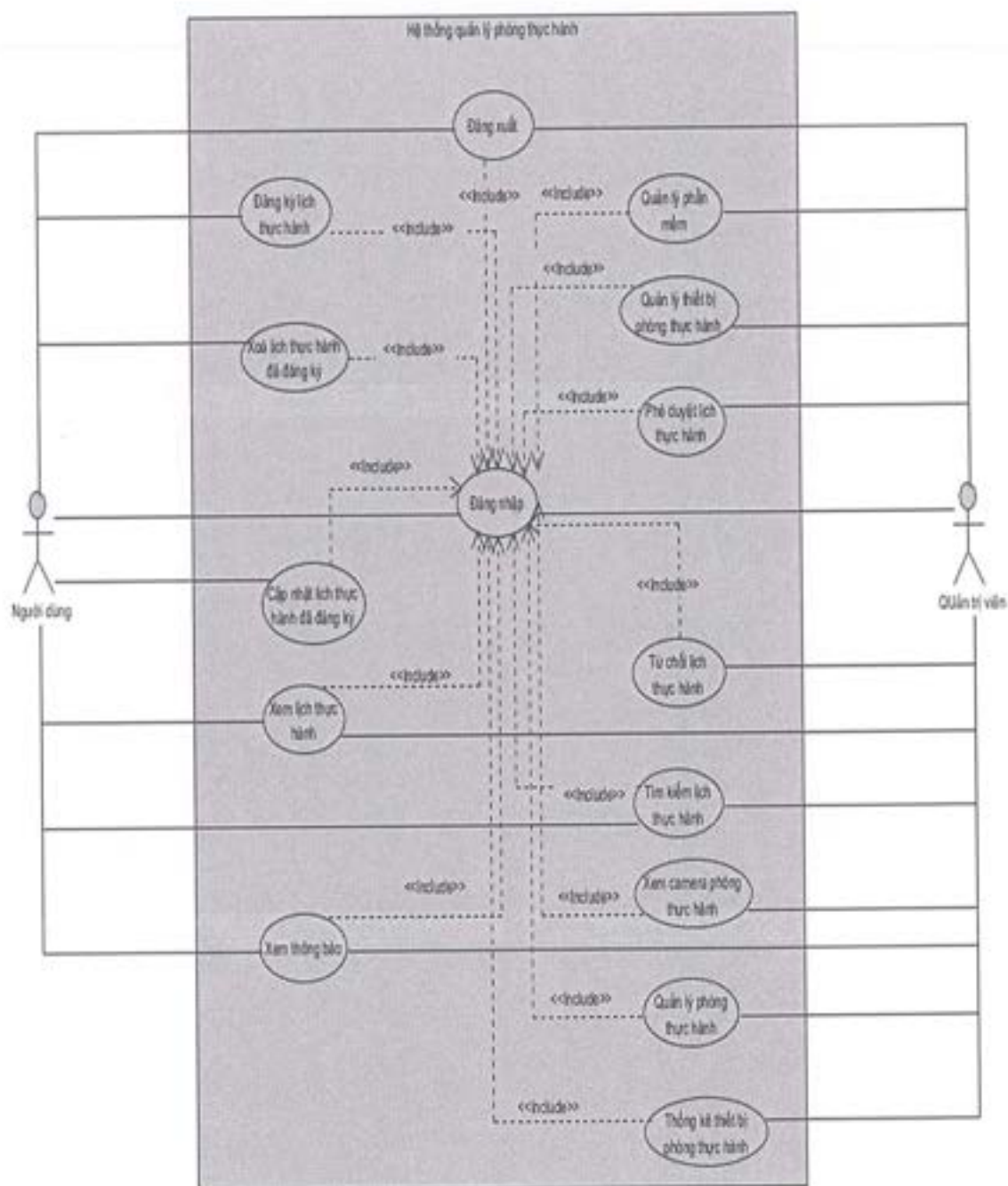
2.1. Tổng quan về nghiệp vụ

Hệ thống có 2 tác nhân chính là người dùng và quản trị viên, mỗi tác nhân đảm nhận các chức năng riêng biệt nhằm hỗ trợ quản lý và giám sát các phòng thực hành máy tính.

Tác nhân thứ 1(người dùng): Là người truy cập hệ thống bằng tài khoản được cấp. Người dùng có thể sử dụng các chức năng liên quan đến việc quản lý lịch thực hành cá nhân và giám sát phòng thực hành. Người dùng có các chức năng chính như đăng ký lịch thực hành, xem, sửa, xoá lịch thực hành đang trong trạng thái đăng ký, tìm kiếm lịch thực hành và xem camera phòng thực hành.

Tác nhân thứ 2(quản trị viên): Là người quản lý toàn bộ hệ thống, chịu trách nhiệm giám sát và đảm bảo hoạt động của phòng thực hành. Quản trị viên có các chức năng chính như quản lý thiết bị phòng thực hành(thêm, sửa xoá, tìm kiếm), quản lý lịch thực hành(phê duyệt, từ chối, tìm kiếm lịch thực hành), xem camera phòng thực hành.

2.2. Phân tích yêu cầu chức năng:



Hình 2-1. Biểu đồ usecase tổng quan của hệ thống.

(nguồn: thu thập dữ liệu tự tạo)

2.3. Đặc tả usecase:

2.3.1. Chức năng đăng nhập

+ Cho phép người dùng hoặc quản trị viên đăng nhập vào hệ thống để sử dụng các chức năng.

+ Đối tượng chính: người dùng và quản trị viên

+ Điều kiện: là người dùng đã có tài khoản được cấp bởi người quản trị

+ Luồng chính:

- Người dùng chọn chức năng đăng nhập
- Hệ thống hiển thị form đăng nhập
- Người dùng nhập email và mật khẩu
- Nhấn nút đăng nhập
- Hệ thống kiểm tra định dạng và tính đầy đủ thông tin
- Kiểm tra thông tin đăng nhập
- Nếu hợp lệ, hệ thống chuyển tới trang chủ

+ Ngoại lệ:

- Nếu sai định dạng email lúc này hệ thống hiển thị lỗi: “ Tài khoản đăng nhập không đúng định dạng email”

- Nếu sai thông tin hệ thống hiển thị lỗi: “ Tài khoản/mật khẩu không đúng”.

2.3.2. Chức năng đăng ký lịch thực hành

+ Cho phép người dùng đăng ký lịch thực hành trên hệ thống.

+ Đối tượng chính: người dùng

+ Điều kiện: người dùng đăng nhập thành công vào hệ thống

+ Luồng chính:

- Người dùng chọn 1 phòng thực hành có trên hệ thống

- Hệ thống hiển thị các lịch thực hành đã được phê duyệt bởi quản trị viên
- Người dùng chọn giờ và ngày chưa có lịch thực hành của phòng đó
- Người dùng nhập thông tin về buổi thực hành
- Người dùng bấm nút đăng ký
- Hệ thống hiển thị thông báo đăng ký thành công

+ Ngoại lệ:

- Hệ thống thông báo lỗi “ Thời gian bắt đầu phải nhỏ hơn thời gian kết thúc và không quá 4 tiếng” nếu nhập thời gian bắt đầu lớn hơn thời gian kết thúc hoặc khoảng thời gian quá 4 tiếng

2.3.3. Chức năng xóa lịch thực hành đã đăng ký

- + Cho phép người dùng xóa lịch thực hành đã đăng ký
- + Đối tượng chính: người dùng
- + Điều kiện: người dùng đăng nhập thành công vào hệ thống
- + Luồng chính:

- Người dùng chọn chức năng quản lý lịch thực hành
- Hệ thống danh sách lịch thực hành chờ phê duyệt sẽ hiện ra
- Người dùng chọn xóa 1 lịch thực hành mà người dùng đã đăng ký
- Hệ thống hiển thị popup “ Bạn có chắc muốn xóa lịch thực hành không?”
- Người dùng click vào nút xác nhận
- Hệ thống hiển thị thông báo xóa thành công và cập nhật lại danh sách

lịch thực hành chờ phê duyệt.

+ Ngoại lệ:

- Không có ngoại lệ

2.3.4. Chức năng cập nhật lịch thực hành đã đăng ký

- + Cho phép người dùng cập nhật lịch thực hành đã đăng ký của họ
- + Đối tượng chính: người dùng
- + Điều kiện: người dùng đăng nhập thành công vào hệ thống
- + Luồng chính:

- Người dùng chọn chức năng quản lý lịch thực hành

- Hệ thống danh sách lịch thực hành chờ phê duyệt
- Người dùng chọn cập nhật 1 lịch thực hành
- Hệ thống hiển thị form chưa thông tin lịch thực hành đã chọn
- Người dùng sửa thông tin lịch thực hành
- Người dùng click nút cập nhật
- Hệ thống kiểm tra thời gian đăng ký
- Hệ thống kiểm tra thời gian đăng ký có bị trùng với lịch thực hành khác không
- Hệ thống hiển thị thông báo cập nhật lịch thực hành thành công

2.3.5. Chức năng xem lịch thực hành

- + Cho phép người quản trị, người dùng xem chi tiết 1 lịch thực hành đã đăng ký
- + Đối tượng chính: người dùng, quản trị viên
- + Điều kiện: người dùng đăng nhập thành công vào hệ thống
- + Luồng chính:
 - Người dùng chọn chức năng quản lý lịch thực hành
 - Hệ thống hiển thị danh sách lịch thực hành
 - Người dùng chọn xem chi tiết 1 lịch thực hành
 - Hệ thống hiển thị chi tiết lịch thực hành cho người dùng
- + Ngoại lệ:
 - không có

2.3.6. Chức năng xem thông báo

- + Cho phép người dùng, quản trị viên xem được thông báo của họ trên hệ thống
- + Đối tượng chính: Người dùng, quản trị viên
- + Điều kiện: Người dùng đăng nhập thành công trên hệ thống
- + Luồng chính:
 - Người dùng đăng nhập vào hệ thống
 - Hệ thống lấy thông tin người dùng
 - Người dùng click vào icon thông báo ở trang chủ
 - Hệ thống hiển thị danh sách thông báo theo thông tin người dùng đã lấy

2.3.7. Chức năng xem thông tin camera phòng thực hành

+ Cho phép quản trị viên xem camera phòng thực hành

+ Đối tượng chính: quản trị viên

+ Điều kiện: quản trị viên đăng nhập thành công trên hệ thống

+ Luồng chính:

- Quản trị viên truy cập vào trang chủ hệ thống

- Quản trị viên chọn chức năng camera

- Hệ thống hiển thị camera phòng thực hành

- Chỉ cần cung cấp nguồn đầu vào cho hệ thống như là stream video hoặc một đoạn video có sẵn thì hệ thống sẽ gửi ảnh và thông báo liên tục nếu phát hiện vi phạm.

+ Chi tiết các bước thực hiện:

- Trong code backend.

- Khởi tạo Flask + SocketIO

- Tạo một web server dùng Flask.

- SocketIO được cấu hình để gửi dữ liệu thời gian thực đến client qua Websocket

- Dùng eventlet.monkey_patch() để Flask có thể chạy bất đồng bộ (async) hiệu quả

- Tải mô hình YOLOv5 đã huấn luyện

- Sử dụng torch.hub.load() để tải mô hình YOLOv5 tùy chỉnh (custom model) đã huấn luyện trước đó (để phát hiện hành vi).

- Kết nối đến luồng video RTSP từ camera

- Mở luồng video bằng cv2.VideoCapture(stream_url)

- Mỗi frame từ luồng sẽ được resize và đưa vào hàng đợi (Queue) để xử lý tuần tự

- Tiến trình buffer_frame (luồng 1)

- Đọc liên tục từng frame từ luồng video

- Resize frame (ví dụ: 640x480).

- Thêm frame vào hàng đợi frame-queue nếu chưa đầy.
- Tiến trình `process_frame` (luồng 2)
 - Liên tục lấy frame từ hàng đợi
 - Chuyển frame sang ảnh màu RGB (phù hợp với mô hình YOLOv5)
 - Đưa frame vào mô hình để dự đoán hành vi
 - Trong từng frame:
 - Mô hình trả về:
 - ✓ Tạo độ các đối tượng được phát hiện
 - ✓ Nhãn (class label)
 - ✓ Độ tin cậy (confidence)
 - Nếu phát hiện hành vi vi phạm (ví dụ: dùng điện thoại,...)
 - ✓ Vẽ bounding box + label lên frame
 - ✓ Gửi frame về client qua sự kiện `video_frame`
 - ✓ Gửi cảnh báo hành vi qua sự kiện `alert`
- Emit dữ liệu về client qua WebSocket
 - Ảnh gửi về client được encode thành base64 (để xử lý ở phía trình duyệt)
 - Sự kiện `socketio.emit(...)` được dùng để truyền ảnh (`video_frame`) và thông báo (`alert`) theo thời gian thực
- Chạy Server
 - Khi chạy `_main_`, Flask sẽ khởi chạy 2 tiến trình:
 - ✓ `Buffer)frames`: lấy video từ camera
 - ✓ `Process_frames`: xử lý AI và gửi kết quả
- Trong front-end
 - Chia sẻ màn hình (hoặc camera) từ người dùng
 - Dùng `navigator.mediaDevices.getDisplayMedia()` để lấy màn hình người dùng

- Gán vào thẻ<video> để hiển thị.
- Dùng Canvas để chụp ảnh từ video sau mỗi 1 giây.
- Gửi ảnh lên server bằng WebSocket
 - Dữ liệu ảnh được gửi dưới dạng Blob qua WebSocket
 - Server xử lý ảnh, phát hiện hành vi và trả kết quả về.
- Nhận và xử lý dữ liệu từ Server
 - Khi nhận được ảnh từ Server:
 - ✓ Upload ảnh lên Cloudinary (nếu cần)
 - ✓ Hiển thị ảnh trong danh sách ảnh nhận được.
 - ✓ Mở model nếu có hành vi được phát hiện
 - Khi nhận được thông báo (alert)
 - ✓ Hiển thị model hoặc popup cảnh báo: Tên hành vi, thời gian phát hiện, độ tin cậy

+ Ngoại lệ:

- Hệ thống hiển thị thông báo lỗi “ không lấy được luồng camera” nếu có lỗi về mạng

2.3.8. Chức năng phê duyệt lịch thực hành

+ Cho phép quản trị viên phê duyệt lịch thực hành được đăng ký bởi người dùng

+ Đối tượng chính: quản trị viên

+ Điều kiện: quản trị viên đăng nhập thành công vào hệ thống

+ Luồng chính:

- Quản trị viên đăng nhập thành công vào hệ thống
- Quản trị viên chọn chức năng quản lý lịch thực hành
- Quản trị viên chọn tab chờ phê duyệt
- Hệ thống hiển thị danh sách các lịch thực hành chờ phê duyệt
- Quản trị viên chọn một lịch thực hành để phê duyệt
- Hệ thống hiển thị popup xác nhận phê duyệt lịch thực hành
- Quản trị viên click nút phê duyệt

- Hệ thống kiểm tra xem có lịch thực hành đã được phê duyệt trùng thời gian hay không

- Hệ thống hiển thị thông báo phê duyệt lịch thực hành thành công

- + Ngoại lệ

- Hệ thống hiển thị thông báo lỗi “ Đã có lịch thực hành được phê duyệt trong khoảng thời gian này” nếu bị trùng thời gian với lịch thực hành đã phê duyệt

2.3.9. Chức năng từ chối lịch thực hành

- + Cho phép quản trị viên từ chối lịch thực hành đăng ký bởi người dùng

- + Đối tượng chính: quản trị viên

- + Điều kiện: quản trị viên đăng nhập thành công vào hệ thống và ở mục chờ phê duyệt

- + Luồng chính:

- Quản trị viên đăng nhập thành công vào hệ thống

- Quản trị viên chọn chức năng quản lý lịch thực hành

- Quản trị viên chọn tab chờ phê duyệt

- Hệ thống hiển thị danh sách các lịch thực hành chờ phê duyệt

- Quản trị viên chọn lịch thực hành để từ chối

- Hệ thống hiển thị popup xác nhận từ chối lịch thực hành

- Quản trị viên click nút xác nhận

- Hệ thống hiển thị thông báo từ chối lịch thực hành thành công

- + Ngoại lệ:

- Không có

2.3.10. Chức năng tìm kiếm lịch thực hành

- + Cho phép người dùng và quản trị viên tìm kiếm lịch thực hành trên hệ thống

- + Đối tượng chính: quản trị viên, người dùng

- + Điều kiện: quản trị viên, người dùng đăng nhập thành công vào hệ thống và ở mục quản lý lịch thực hành

- + Luồng chính:

- Quản trị viên chọn chức năng quản lý lịch thực hành

- Hệ thống hiển thị danh sách tất cả lịch thực hành
- Quản trị viên nhập tên buổi thực hành, chọn thời gian, phòng
- Quản trị viên nhấn nút tìm kiếm
- Hệ thống hiển thị danh sách các lịch thực hành phù hợp với các thông tin

tìm kiếm

+ Ngoại lệ:

- Không có

2.3.11. Chức năng quản lý thiết bị phòng thực hành

- + Cho phép người quản trị viên xem, thêm, sửa, xóa các thiết bị phòng thực hành có trên hệ thống
- + Đối tượng chính: quản trị viên
- + Điều kiện: quản trị viên đăng nhập thành công vào hệ thống và ở mục

quản lý thiết bị

+ Hệ thống cập nhật danh sách thiết bị:

- Quản trị viên chọn chức năng thêm hoặc sửa thiết bị
- Hệ thống hiển thị form thêm mới thiết bị hoặc cập nhật thiết bị
- Quản trị viên nhập đầy đủ thông tin thiết bị
- Quản trị viên nhấn nút thêm mới hoặc cập nhật thiết bị
- Hệ thống kiểm tra tính hợp lý và đầy đủ của các trường quản trị viên nhập
- Hệ thống chuyển về trang danh sách thiết bị và hiển thị thông báo thêm

hoặc sửa thiết bị thành công và cập nhật lại danh sách thiết bị

- Quản trị viên chọn chức năng xóa
- Hệ thống xóa thiết bị mà quản trị viên vừa chọn
- Quản trị viên chọn xem thiết bị
- Hệ thống hiển thị thông tin chi tiết thiết bị được chọn

+ Ngoại lệ:

- Hệ thống thông báo lỗi “ Vui lòng điền đầy đủ thông tin nếu nhập thiếu một trong các trường trên”

2.3.12. Chức năng quản lý phòng thực hành

- + Cho phép quản trị viên quản lý tất cả phòng thực hành
- + Đối tượng chính: quản trị viên
- + Điều kiện: quản trị viên đăng nhập thành công vào hệ thống và ở mục phòng thực hành
- + Luồng chính:
 - Quản trị viên chọn chức năng thêm hoặc cập nhật phòng thực hành
 - Hệ thống hiển thị form thêm mới phòng thực hành hoặc cập nhật phòng thực hành
 - Quản trị viên nhập đầy đủ thông tin phòng thực hành
 - Quản trị viên nhấn nút thêm mới hoặc cập nhật phòng thực hành
 - Hệ thống kiểm tra tính hợp lệ và đầy đủ của các trường quản trị viên nhập
 - Hệ thống chuyển về trang danh sách phòng thực hành và hiển thị thông báo thêm hoặc sửa phòng thực hành thành công và cập nhật lại danh sách phòng thực hành
- + Ngoại lệ:
 - Hệ thống thông báo lỗi “ Vui lòng điền đầy đủ thông tin” nếu nhập thiếu một trong các trường trên

2.3.13. Chức năng thống kê thiết bị phòng thực hành

- + Cho phép quản trị viên thống kê tất cả trạng thái và số lượng thiết bị của phòng thực hành
- + Đối tượng chính: quản trị viên
- + Điều kiện: quản trị viên đăng nhập thành công vào hệ thống và ở mục phòng thực hành. Quản trị viên đang ở màn hình danh sách phòng thực hành và đã chọn một phòng cụ thể

+ Luồng chính:

- Quản trị viên chọn một phòng thực hành từ danh sách phòng
- Hệ thống lấy dữ liệu thống kê thiết bị của phòng được chọn
- Hệ thống trả kết quả về giao diện người dùng
- Giao diện hiển thị danh sách thống kê thiết bị

+ Ngoại lệ:

- Không tìm thấy thiết bị nào trong phòng nếu phòng này chưa cập nhật các thiết bị

2.4. Phân tích yêu cầu phi chức năng

2.4.1. Yêu cầu bảo mật

Đồ án sử dụng Spring Boot làm backend đã được xây dựng với mục tiêu đảm bảo các yêu cầu bảo mật theo hai hình thức chính: bảo mật phân quyền và bảo mật qua việc xác thực mã hóa token.

Về bảo mật phân quyền, hệ thống phân chia quyền truy cập theo hai vai trò chính: quản trị viên (ADMIN) và người dùng (USER). Người dùng chỉ có quyền thực hiện các chức năng cơ bản phù hợp với quyền hạn, ví dụ: tìm kiếm lịch thực hành, xem và đăng ký lịch, v.v. Trong khi đó, quản trị viên có quyền thực hiện các chức năng cao cấp hơn như quản lý người dùng, phê duyệt hoặc từ chối đặt lịch, thống kê và quản lý thiết bị hệ thống. Việc này đảm bảo rằng các chức năng nhạy cảm hoặc quản trị chỉ được thực hiện bởi những người có thẩm quyền.

Về bảo mật thông tin, hệ thống sử dụng cơ chế xác thực qua JWT (JSON Web Token). Mỗi lần người dùng đăng nhập, hệ thống sẽ cấp một token được ký và mã hóa bằng thuật toán HMAC-SHA512. Token này sẽ được kiểm tra tính hợp lệ khi người dùng gửi yêu cầu đến server. Đồng thời, hệ thống cũng có cơ chế quản lý danh sách đen (token blacklist) để đảm bảo các token đã bị thu hồi hoặc hết hạn không thể sử dụng lại, ngăn chặn các lỗ hổng bảo mật.

Ngoài ra, cơ chế CORS (Cross-Origin Resource Sharing) được cấu hình để chỉ cho phép truy cập từ các nguồn đáng tin cậy, giảm thiểu nguy cơ tấn công CSRF (Cross-Site Request Forgery). Hệ thống cũng đảm bảo bảo vệ các endpoint quan trọng bằng cách bắt buộc xác thực và phân quyền chi tiết.

Với cách tiếp cận này, đồ án không chỉ đảm bảo tính an toàn, bảo mật dữ liệu mà còn hỗ trợ quản lý hiệu quả, nâng cao trải nghiệm người dùng trong hệ thống.

2.4.2. Yêu cầu hiệu năng

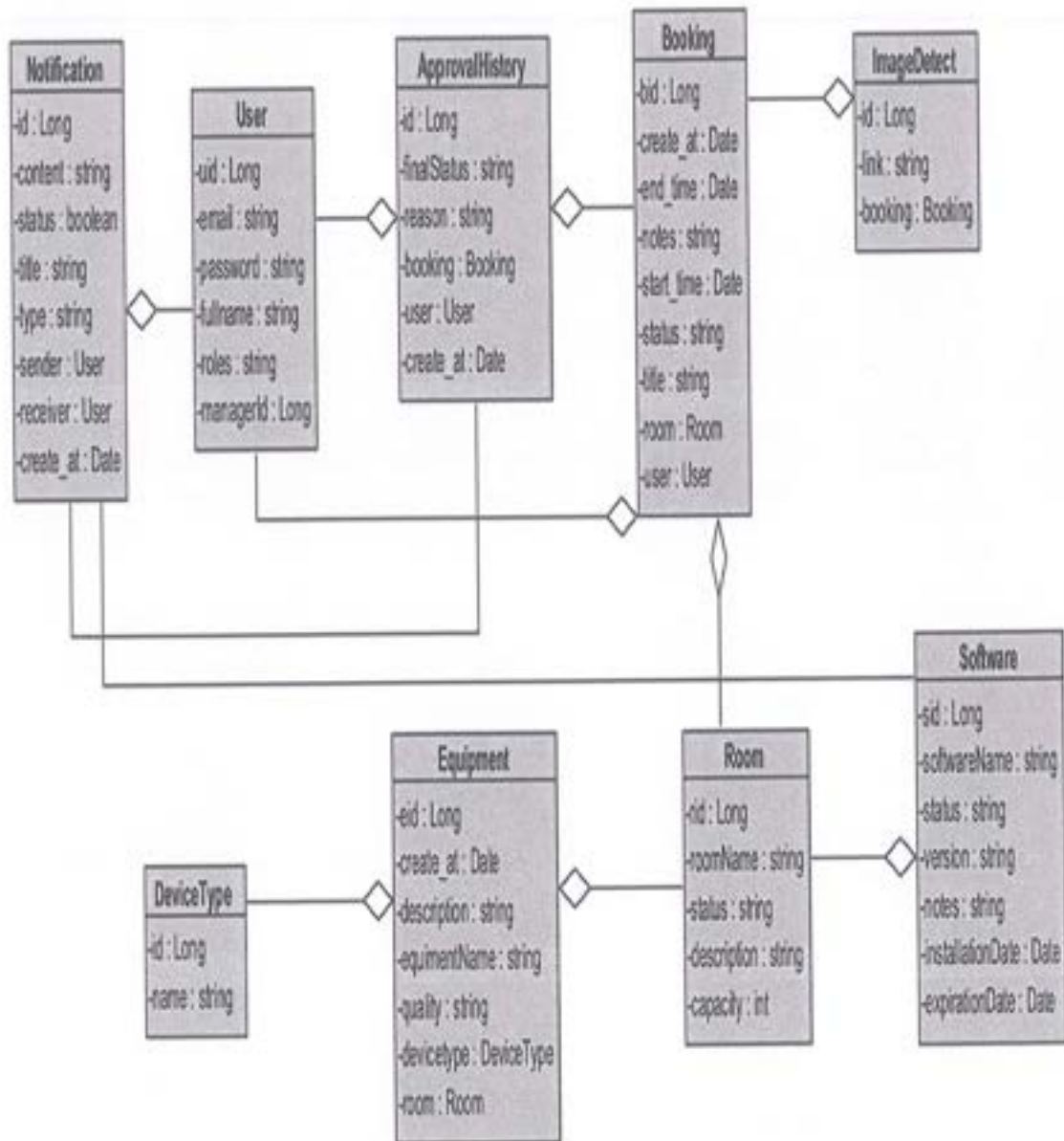
Hệ thống được xây dựng theo mô hình client-server, sử dụng ReactJS cho frontend và Spring Boot cho backend. Cách tiếp cận này tách biệt giao diện người dùng và logic nghiệp vụ, tối ưu hiệu năng, khả năng mở rộng và bảo trì. Client (ReactJS) tập trung hiển thị giao diện tương tác, mang lại trải nghiệm người dùng mượt mà. Sử dụng JavaScript, JSX và các thư viện như Ant Design, React Router để quản lý giao diện và trạng thái ứng dụng. Client sử dụng AJAX (Fetch API hoặc Axios) để giao tiếp bất đồng bộ với server, giúp cập nhật giao diện mà không cần tải lại toàn bộ trang. Áp dụng các kỹ thuật tối ưu như code splitting, lazy loading để giảm thời gian tải ban đầu. Server (Spring Boot) xử lý logic nghiệp vụ, bao gồm xử lý yêu cầu từ client, truy vấn cơ sở dữ liệu và trả về dữ liệu dạng JSON hoặc XML. Xây dựng API RESTful, sử dụng HTTP methods (GET, POST, PUT, DELETE) để tương tác với tài nguyên.

2.5. Thiết kế biểu đồ lớp thực thể hệ thống

Dựa vào biểu đồ lớp thực thể ở pha phân tích của chương 2, xây dựng biểu đồ lớp thực thể bằng cách tiến hành các bước như sau:

- Bổ sung thuộc tính id cho lớp không kế thừa từ lớp khác
- Bổ sung kiểu dữ liệu cho các thuộc tính
- Chuyển đổi sang các quan hệ association sang các dạng aggregation/ composition
- Bổ sung thuộc tính đối tượng của các lớp thực thể (nếu có)

Từ đó, ta có biểu đồ lớp thực thể ở pha thiết kế như sau:



Hình 2-2. Biểu đồ lớp thực thể của hệ thống.

(nguồn: thu tập dữ liệu tự tạo)

Sau đây là bảng chi tiết thông tin của các lớp:

Bảng 2-1. Bảng mô tả các lớp trong hệ thống.

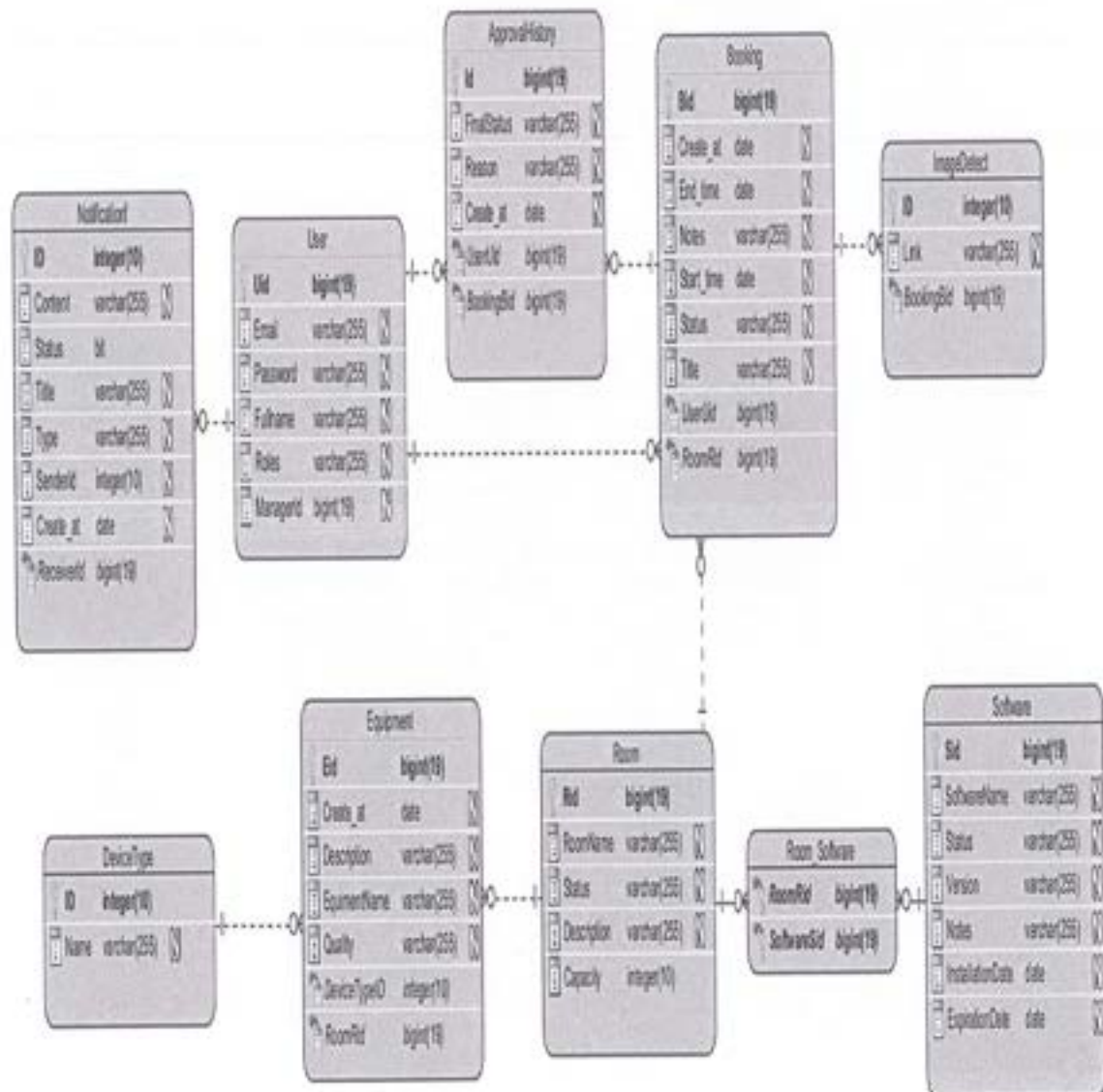
STT	Tên lớp	Mô tả
1	User	Lớp User chứa tất cả thông tin chi tiết của người dùng như tài khoản, mật khẩu, họ tên, email, vai trò và thông tin quản lý (nếu có).
2	Notification	Lớp Notification chứa thông tin về các thông báo trong hệ thống như nội dung, tiêu đề, trạng thái (đã đọc/chưa đọc), loại thông báo, người gửi và người nhận.
3	ApprovalHistory	Lớp ApprovalHistory lưu trữ lịch sử phê duyệt của các đơn đặt phòng, bao gồm trạng thái cuối cùng, lý do, thông tin người thực hiện phê duyệt, và ngày tạo.
4	Booking	Lớp Booking chứa thông tin về các đơn đặt phòng, bao gồm thời gian bắt đầu, thời gian kết thúc, trạng thái, tiêu đề, ghi chú, và liên kết với thông tin người đặt cùng phòng đặt.
5	Room	Lớp Room lưu trữ thông tin về các phòng thực hành, bao gồm tên phòng, mô tả, trạng thái (có sẵn, đang sử dụng, bảo trì), và sức chứa của phòng.
6	Equipment	Lớp Equipment chứa thông tin về các thiết bị trong hệ thống như mô tả, tên thiết bị, chất lượng, loại thiết bị (DeviceType), và liên kết với phòng nơi thiết bị được đặt.

7	DeviceType	Lớp DeviceType chứa thông tin về loại thiết bị, ví dụ như máy tính, máy in, hoặc thiết bị ngoại vi khác.
8	Software	Lớp Software lưu trữ thông tin về các phần mềm cài đặt trong phòng, bao gồm tên phần mềm, trạng thái, phiên bản, ghi chú, ngày cài đặt và ngày hết hạn sử dụng.
9	ImageDetect	Lớp ImageDetect chứa thông tin về các hình ảnh được nhận diện trong quá trình sử dụng phòng thực hành, bao gồm liên kết tới hình ảnh và thông tin liên kết với một đơn đặt phòng cụ thể.

2.6. Thiết kế cơ sở dữ liệu hệ thống

Để xây dựng thiết kế cơ sở dữ liệu cho hệ thống với đầu vào là biểu đồ lớp thực thể của cần phải thực hiện bước như sau:

- Với mỗi lớp thực thể, đề xuất một bảng dữ liệu tương ứng. Nên đặt tên bảng theo cách gọi nhớ tương ứng
- Với mỗi lớp thực thể, bỏ qua các thuộc tính kiểu đối tượng, lấy các thuộc tính kiểu cơ bản đưa sang làm thuộc tính của bảng tương ứng, và chuyển kiểu dữ liệu theo hệ quản trị cơ sở dữ liệu tương ứng
- Xét quan hệ số lượng giữa các lớp thực thể, quan hệ số lượng giữa hai lớp thực thể cũng là quan hệ số lượng giữa hai bảng tương ứng với hai lớp đấy. Ta có thiết kế cơ sở dữ liệu như sau:



Hình 2-3. Cơ sở dữ liệu hệ thống.

(nguồn: dữ liệu thu thập tự tạo)

2.7. Kết chương

Trong chương 2, các yêu cầu chức năng và phi chức năng của hệ thống đã được phân tích kỹ lưỡng thông qua biểu đồ use case và đặc tả chi tiết từng chức năng. Đồng thời, các yêu cầu bảo mật và hiệu năng cũng được xem xét để đảm bảo hệ thống hoạt động ổn định và an toàn. Bên cạnh đó là việc hoàn thành thiết kế hệ

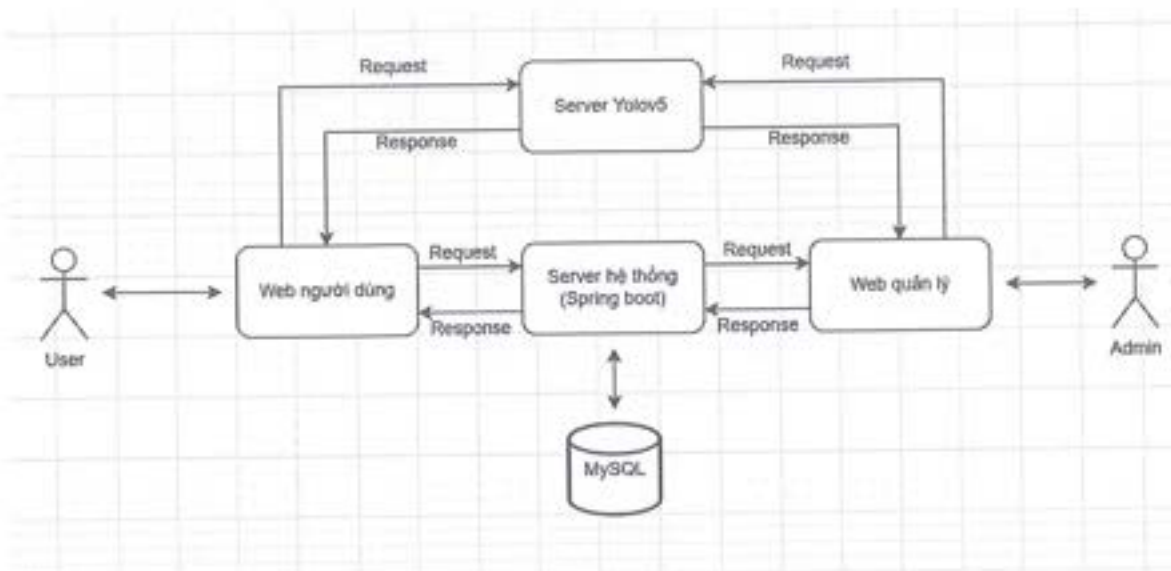
thống, bao gồm biểu đồ lớp thực thể, cơ sở dữ liệu, và biểu đồ tuần tự cho từng chức năng. Việc thiết kế chi tiết các thành phần hệ thống giúp đảm bảo quá trình cài đặt và triển khai được thực hiện một cách chính xác, tối ưu và đáp ứng đầy đủ các yêu cầu đã được phân tích trước đó.

CHƯƠNG 3 - CÀI ĐẶT VÀ KIỂM THỬ HỆ THỐNG

Cấu trúc dự án

- Tầng FE: cấu trúc giao diện chia thành các Folder tương ứng với các chức năng liên quan
- Tầng BE: chia thành các thư mục chính như: Controller, Service, Repository
- Tầng cơ sở dữ liệu: Sử dụng MySQL với các bảng như sau:

3.1. Mô hình triển khai hệ thống.



Hình 3-1. Sơ đồ kiến trúc hệ thống

Hệ thống được xây dựng theo kiến trúc Client-Server. Trong đó máy khách gửi yêu cầu HTTP request đến máy chủ. Máy chủ tiếp nhận yêu cầu và xử lý, sau đó trả kết quả về cho máy khách. Hệ thống sử dụng 2 server chính bao gồm: Server cho chức năng quản lý thông thường và Server dùng để phát hiện các hành vi xấu từ hình ảnh cho mô hình YOLOv5. Các dữ liệu được trao đổi giữa Client và Server sẽ có dạng là JSON.

Môi trường triển khai gồm có:

- Ngôn ngữ lập trình: Java, JavaScript, Python
- IDE: IntelliJ, Visual Code, PyCharm

- Cơ sở dữ liệu: MySQL
- Các công cụ khác: Ant Design để thiết kế giao diện chuyên nghiệp, mang lại trải nghiệm tương tác tốt nhất, Cloudinary dùng để lưu trữ các file hình ảnh, PyTorch hỗ trợ xử lý và chạy mô hình YOLOv5, cung cấp nền tảng mạnh mẽ cho việc phát triển AI, OpenCV để phân tích và xử lý hình ảnh/video từ camera giám sát, giúp tăng độ chính xác của hệ thống nhận diện

3.2. Giới thiệu về MySQL:

MySQL là một hệ quản trị cơ sở dữ liệu mã nguồn mở, phổ biến trên toàn cầu, được phát triển bởi Oracle Corporation. Với khả năng lưu trữ và quản lý dữ liệu mạnh mẽ, MySQL thường được sử dụng để hỗ trợ các ứng dụng web, hệ thống quản lý và phân tích dữ liệu.

Đặc điểm nổi bật của MySQL:

- Hiệu suất cao: MySQL có khả năng xử lý lượng lớn dữ liệu và truy vấn nhanh, phù hợp với các ứng dụng từ nhỏ đến quy mô lớn.
- Độ tin cậy và bảo mật: Được tích hợp các cơ chế bảo mật mạnh mẽ như mã hóa dữ liệu, kiểm soát truy cập người dùng, giúp bảo vệ thông tin trước các mối đe dọa.
- Dễ sử dụng: Giao diện thân thiện và hệ sinh thái phong phú giúp MySQL trở thành lựa chọn hàng đầu cho cả người mới học và các chuyên gia.
- Tính linh hoạt: Hỗ trợ đa nền tảng (Windows, Linux, macOS) và khả năng tích hợp với nhiều ngôn ngữ lập trình như Java, PHP, Python.
- Hỗ trợ nhiều loại lưu trữ: Với các engine lưu trữ như InnoDB, MyISAM, MySQL phù hợp với đa dạng nhu cầu lưu trữ từ giao dịch trực tuyến đến phân tích dữ liệu.

3.3. Giới thiệu về Spring boot:

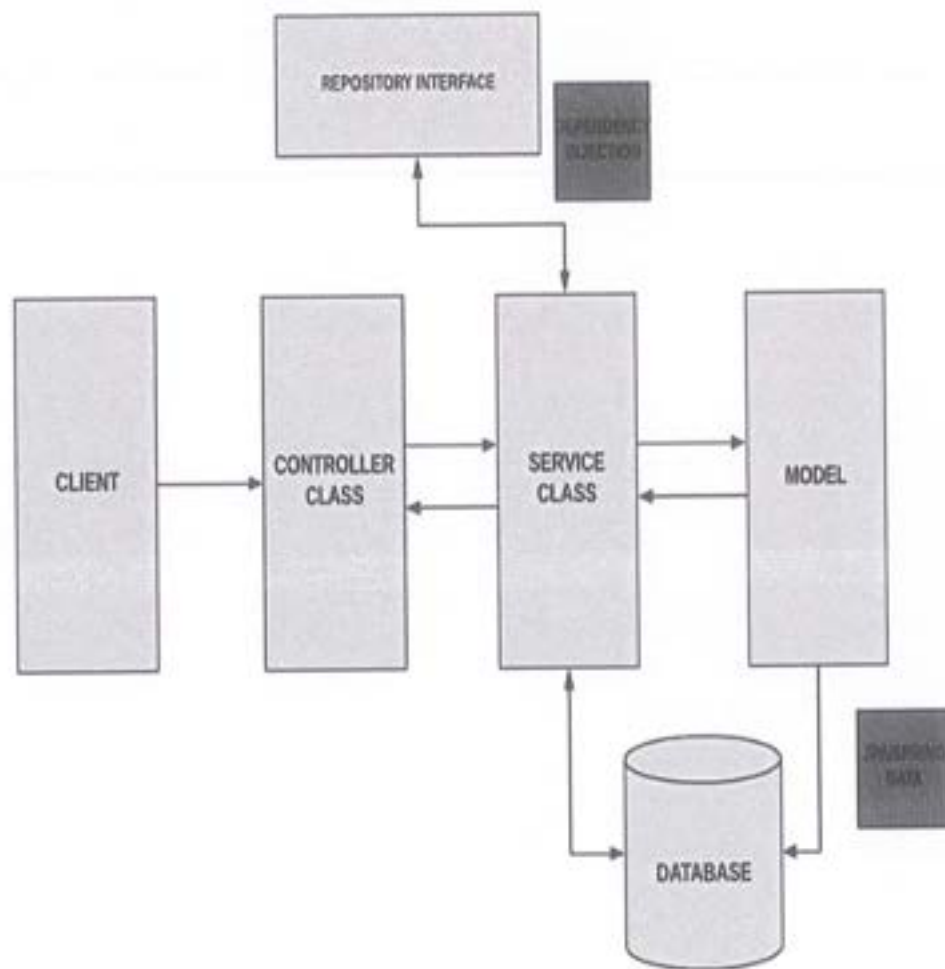
Spring Boot được xây dựng dựa trên nền của Spring Framework, phát triển bởi ngôn ngữ Java. Nó là một phiên bản có tính tự động hoá cao hơn và đơn giản

hơn của Spring. Spring Boot khiến cho việc tạo và khởi chạy ứng dụng chỉ diễn ra trong vòng vài phút.

Spring là một tập hợp các framework con, bao gồm nhiều module khác nhau như Spring MVC, Spring JDBC, Spring AOP, Spring ORM và Spring Test. Tính năng cốt lõi của Spring Framework chính là Dependency Injection (DI), cho phép các đối tượng được tiêm (injected) vào nhau một cách linh hoạt mà không cần phải tạo mới hoặc quản lý chúng trực tiếp. Điều này làm cho ứng dụng trở nên dễ dàng quản lý và mở rộng hơn.

Mặc dù Spring đã giúp cho việc phát triển các ứng dụng Java trở nên dễ dàng hơn, nhưng vẫn còn đó sự phức tạp và độ rộng lớn mà framework này mang lại. Việc học Spring Framework từ đầu có thể là một hành trình dài và đầy thách thức đối với nhiều lập trình viên. Đó cũng chính là lúc Spring Boot tỏa sáng với những cải tiến vượt trội.

Spring Boot ra đời với mục tiêu đơn giản hóa việc phát triển các ứng dụng Java bằng cách giảm thiểu cấu hình thủ công và cung cấp các thiết lập mặc định hợp lý. Thay vì phải cấu hình một cách phức tạp và chi tiết như trong Spring, Spring Boot giúp lập trình viên khởi tạo và chạy ứng dụng một cách nhanh chóng chỉ với một vài dòng mã. Spring Boot còn tích hợp sẵn các công cụ như máy chủ nhúng (embedded server), hỗ trợ cấu hình bằng YAML, và tự động hóa nhiều tác vụ mà trước đây cần phải làm thủ công.



Hình 3-2. Luồng hoạt động của Spring Boot.

Workflow của Spring Boot hoạt động như sau:

1. Client tạo một HTTP request.
2. Class Controller nhận HTTP request.
3. Controller hiểu loại request và xử lý nó.
4. Nếu cần, nó gọi class Service.
5. Class Service sẽ xử lý logic nghiệp vụ, với dữ liệu từ database
6. Nếu mọi thứ hoạt động tốt, ta sẽ trả về một trang giao diện.

3.4. Đánh giá các chỉ số (Evaluation Metrics)

3.4.1. Intersection over Union (IoU)

IoU đo lường mức độ chồng lấp giữa hộp giới hạn dự đoán và hộp giới hạn thực tế. Nó được tính toán như sau:

$$\text{IoU} = \frac{\text{Diện tích chồng lấp (Area of Overlap)}}{\text{Diện tích liên hợp (Area of Union)}}$$

- IoU rất quan trọng để xác định các kết quả đúng (true positives) ($\text{IoU} \geq$ ngưỡng) và kết quả sai (false positives) ($\text{IoU} < \text{ngưỡng}$). YOLO tối ưu hóa giá trị IoU cao để cải thiện độ chính xác định vị. Các ngưỡng IoU phổ biến bao gồm:

- $\text{IoU} \geq 0.5$: Ngưỡng chuẩn cho phát hiện đối tượng.
- $\text{IoU} \geq 0.75$: Ngưỡng nghiêm ngặt hơn cho các ứng dụng yêu cầu độ chính xác cao trong việc định vị.
- IoU trong phạm vi (ví dụ: $[0.5; 0.95]$): Được sử dụng trong các bộ dữ liệu như COCO để đánh giá Độ chính xác Trung bình (mean Average Precision, mAP) trên một loạt các ngưỡng.

3.4.2. Precision, Recall và F1 Score

Precision (độ chính xác) đo lường khả năng của mô hình trong việc dự đoán chính xác các đối tượng dương tính, được chỉ ra bởi tỷ lệ các hộp giới hạn dự đoán chính xác so với tất cả các hộp giới hạn mà mô hình dự đoán. Precision cao có

nghĩa là phần lớn các phát hiện dự đoán là chính xác. Tuy nhiên, precision cao một mình không đảm bảo mô hình phát hiện tất cả các đối tượng trong hình ảnh.

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}}$$

- True Positive (TP): Một hộp giới hạn dự đoán chính xác xác định một đối tượng với IoU vượt qua một ngưỡng nhất định (ví dụ: $\text{IoU} \geq 0.5$).
- False Positive (FP): Một hộp giới hạn dự đoán không chính xác khớp với bất kỳ đối tượng thực tế nào, do phân loại sai hoặc định vị kém.

Recall đo lường khả năng của mô hình trong việc phát hiện tất cả các đối tượng liên quan trong một hình ảnh. Nó là tỷ lệ các hộp giới hạn dự đoán chính xác so với tất cả các hộp giới hạn thực tế.

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}}$$

- False Negative (FN): Một đối tượng thực tế mà mô hình không phát hiện được.

Tuy nhiên, có một sự đánh đổi giữa precision và recall: Việc tăng ngưỡng phát hiện (điểm tin cậy) sẽ cải thiện precision nhưng có thể làm giảm recall (một số true positives bị bỏ sót), trong khi giảm ngưỡng sẽ tăng recall nhưng có thể làm giảm precision (nhiều false positives được bao gồm). Để đánh giá sự đánh đổi này, thường vẽ một Đường cong Precision-Recall, thể hiện precision dưới dạng hàm của recall.

- F1 Score là trung bình điều hòa của precision và recall, cung cấp đánh giá cân bằng về hiệu suất của mô hình bằng cách xem xét cả false positives và false negatives. Một F1 score cao cho thấy sự cân bằng giữa precision và recall.

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

3.4.3. Average Precision (AP) và Mean Average Precision (mAP)

Average Precision (AP) là một chỉ số được sử dụng để đánh giá độ chính xác của mô hình phát hiện đối tượng đối với một lớp cụ thể. AP được tính toán như diện

tích dưới đường cong precision-recall, được xây dựng bằng cách vẽ precision theo recall khi ngưỡng tin cậy của các dự đoán thay đổi.

$$AP = \int_0^1 P(r)dr$$

Mean Average Precision (mAP) là giá trị trung bình của các giá trị AP trên tất cả các lớp đối tượng. Nó cung cấp một thước đo tổng thể về khả năng của mô hình trong việc phát hiện và phân loại các đối tượng trong một tập dữ liệu. Đối với một tập dữ liệu có N lớp, mAP được tính như sau:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i$$

mAP cao chỉ ra độ chính xác tốt trong việc phát hiện và định vị đối tượng trên các lớp, trong khi mAP thấp cho thấy có vấn đề trong phân loại, định vị, hoặc cả hai. Tuy nhiên, các mô hình được huấn luyện trên các tập dữ liệu mất cân bằng có thể có AP cao đối với các lớp thường gặp nhưng AP thấp đối với các lớp hiếm gặp, làm giảm mAP. mAP phù hợp để đánh giá tổng quan hiệu suất của mô hình.

3.4.4. Labellmg

Labellmg là một công cụ chú thích hình ảnh đồ họa phổ biến được sử dụng để tạo bộ dữ liệu cho các tác vụ phát hiện đối tượng.

Trước khi bắt đầu, xác định các hành vi hoặc vi phạm cụ thể và các đối tượng cần phát hiện. Tạo một sơ đồ gán nhãn phân tách các hành động và đối tượng thành các danh mục riêng biệt. Trong tình huống này, có các vi phạm: vi phạm giày dép, trang phục không phù hợp, vi phạm điện thoại, gian lận học thuật, và tư thế không phù hợp.

Đối với các hành động này, tập trung vào việc gán nhãn các yếu tố chính, ví dụ, cá nhân thực hiện hành động hoặc các phần cụ thể của cá nhân liên quan đến hành động; và các đối tượng trong bối cảnh, ví dụ, các yếu tố nền giúp cung cấp bối cảnh cho hành động. Chi tiết sẽ được thể hiện như dưới đây.

a. Vi phạm giày dép

Để bắt xác định việc vi phạm quy tắc giày dép hay trang phục. Vẽ các hộp giới hạn xung quanh giày, dép hoặc bất kỳ loại giày dép nào không tuân thủ hướng dẫn yêu cầu. Ngoài ra, cũng bao gồm các đối tượng xung quanh bối cảnh vi phạm, chẳng hạn như sàn nhà, quần tây, hoặc các khu vực gần đó làm nổi bật giày dép không phù hợp.

b. Trang phục không phù hợp

Tập trung vào các món đồ, trang phục vi phạm quy định. Vẽ các hộp giới hạn xung quanh các món đồ như thiếu cà vạt, áo sơ mi không phù hợp, hoặc các vi phạm về trang phục khác. Trong trường hợp trang phục không phù hợp bao gồm nhiều yếu tố hơn chỉ trang phục (ví dụ, đồng phục không khớp), hãy xem xét toàn bộ ngoại hình.

c. Gian lan thi cử.

Nổi bật các hành động không phù hợp như sao chép từ màn hình của một sinh viên khác hoặc sử dụng tài liệu. Vẽ các hộp giới hạn xung quanh những hành vi này để nhận diện vi phạm rõ ràng. Bao gồm các đối tượng bối cảnh như màn hình, sách vở trong cảnh để cung cấp một bức tranh đầy đủ về hành động gian lận.

d. Tư thế không phù hợp.

Tập trung vào các bộ phận cơ thể của cá nhân thể hiện vấn đề tư thế. Ví dụ, gán nhãn vị trí chân không đúng. Để cung cấp bối cảnh cho tư thế, bao gồm các không gian làm việc gần đó, chẳng hạn như bàn làm việc hoặc ghế, vì chúng giúp làm nổi bật tư thế không phù hợp trong cảnh.

3.5. Thực nghiệm và đánh giá

3.5.1. Thực nghiệm: Quy trình thu thập và gán nhãn dữ liệu

a. Trích xuất hình ảnh từ video:

Các khung hình video được trích xuất sử dụng OpenCV trong Python với tần suất một khung hình mỗi giây. Điều này giúp cân bằng kích thước tập dữ liệu và giảm thiểu sự dư thừa.

b. Gán nhãn dữ liệu:

- Công cụ gán nhãn: Sử dụng LabelImg để gán nhãn các đối tượng.
- Lớp nhãn (classes): Danh sách tên lớp được định nghĩa trước để đảm bảo tính nhất quán.
- Quy trình:
 - Mở từng hình ảnh trong LabelImg sau đó vẽ khung bao (bounding boxes) xác định vùng chứa hành vi cần nhận diện.
 - Gán tên lớp tương ứng: Đảm bảo gán đúng tên lớp cho mỗi khung bao.
 - Lưu file nhãn: Gán nhãn và lưu dưới định dạng TXT theo chuẩn YOLO

c. Tổ chức tập dữ liệu:

- Tạo cấu trúc thư mục hợp lý:
 - images/train: chứa hình ảnh dùng để huấn luyện.
 - images/val: chứa hình ảnh dùng để kiểm tra.
 - labels/train: chứa file nhãn tương ứng cho tập huấn luyện.
 - labels/val: chứa file nhãn tương ứng cho tập kiểm tra.

- Tỷ lệ chia: 80% huấn luyện – 20% xác thực. Đảm bảo phân bố lớp cân bằng.

3.5.2. Cấu hình tham số huấn luyện.

- Các tham số quan trọng trong quá trình huấn luyện YOLOv5 gồm:

- + Kích thước ảnh: 640x640 pixel – đủ chi tiết mà vẫn đảm bảo tốc độ xử lý.
- + Batch size: 16 – 32 phụ thuộc vào dung lượng của GPU.
- + Số epoch: 100 – 200 đủ để mô hình học nhưng tránh overfitting.

- Cấu hình dữ liệu được định nghĩa trong file *data.yaml*, bao gồm:

- + Đường dẫn đến tập dữ liệu.
- + số lượng lớp (nc).
- + Tên các lớp (names).

3.5.3. Huấn luyện mô hình

Việc triển khai mô hình YOLOv5 được thực hiện trên Google Colab, tận dụng tài nguyên GPU miễn phí. Quá trình bắt đầu với việc thiết lập môi trường, trong đó Google Colab được khởi tạo với hỗ trợ GPU và các thư viện cần thiết được cài đặt. Kho lưu trữ YOLOv5 được clone từ GitHub, và các phụ thuộc cần thiết được cài đặt bằng lệnh:

- + Cài đặt môi trường

```
bash
```

```
copyEdit
```

```
!git clone https://github.com/ultralytics/yolov5
```

```
%cd yolov5
```

```
!pip install -r requirements.txt
```

- + Gắn Google Driver

Để lưu trữ và truy cập dữ liệu một cách thuận tiện, Google Drive được gắn vào Colab bằng lệnh sau:

```
python
```

```
copyEdit
```

```
from google.colab import drive
```

```
drive.mount('/content/drive')
```

Sau khi môi trường được thiết lập, dữ liệu huấn luyện được tải lên một thư mục cụ thể trong Google Drive, nơi nó được tổ chức thành các thư mục hình ảnh và nhãn. Cấu hình bộ dữ liệu được định nghĩa thông qua tệp `data.yaml`, chỉ rõ các đường dẫn tới bộ dữ liệu, số lượng lớp (`nc`), và danh sách tên lớp (`names`). Việc huấn luyện được bắt đầu bằng cách sử dụng script huấn luyện YOLOv5 tích hợp sẵn với các siêu tham số được thiết lập ở các giá trị tối ưu, chẳng hạn như kích thước hình ảnh (`--img 640`), kích thước batch (`--batch 16`), và số lượng epoch (`--epochs 150`):

+ Khởi chạy huấn luyện

bash

copyEdit

```
!python train.py --img 640 --batch 16 --epochs 150 --data <data.yaml> --weights <weights> --name <model_name>
```

Khi mô hình đã được huấn luyện, kết quả, bao gồm các tệp mô hình đã huấn luyện và nhật ký hiệu suất, được lưu trữ trong một thư mục được chỉ định trên Google Drive để dễ dàng truy cập và đánh giá thêm. Mô hình được lưu bằng script Python sau:

bash

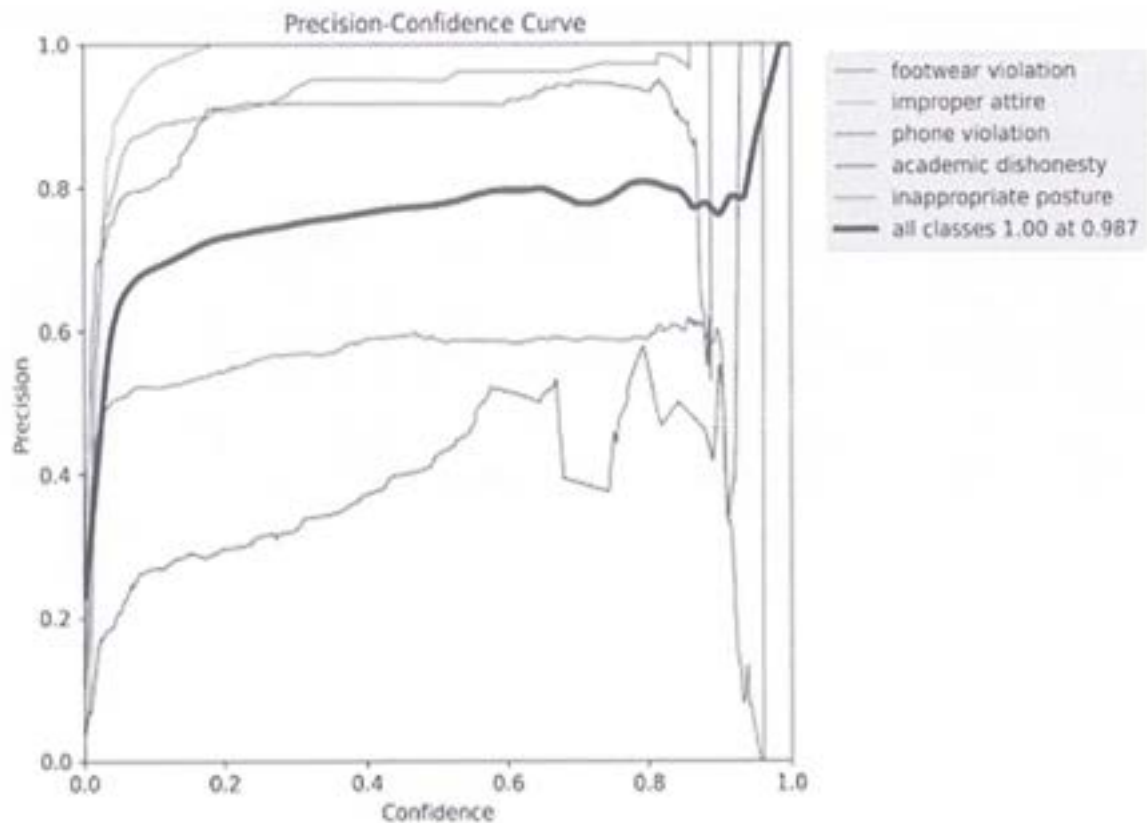
copyEdit

```
!cp -r /content/yolov5/runs/train/<model_folder> /content/driver/MyDriver/<path>
```

Thư mục mô hình chứa mô hình đã huấn luyện (`best.pt`), đồ thị hiệu suất và nhật ký. Sau khi lưu, mô hình có thể được tải hoặc kiểm tra trực tiếp trong Colab hoặc triển khai vào các ứng dụng thực tế. Cài đặt này đảm bảo một quy trình huấn luyện hiệu quả và an toàn, với kết quả sẵn có cho việc sử dụng và phát triển trong tương lai.

3.5.4. Đánh giá

a. Đường cong Precision



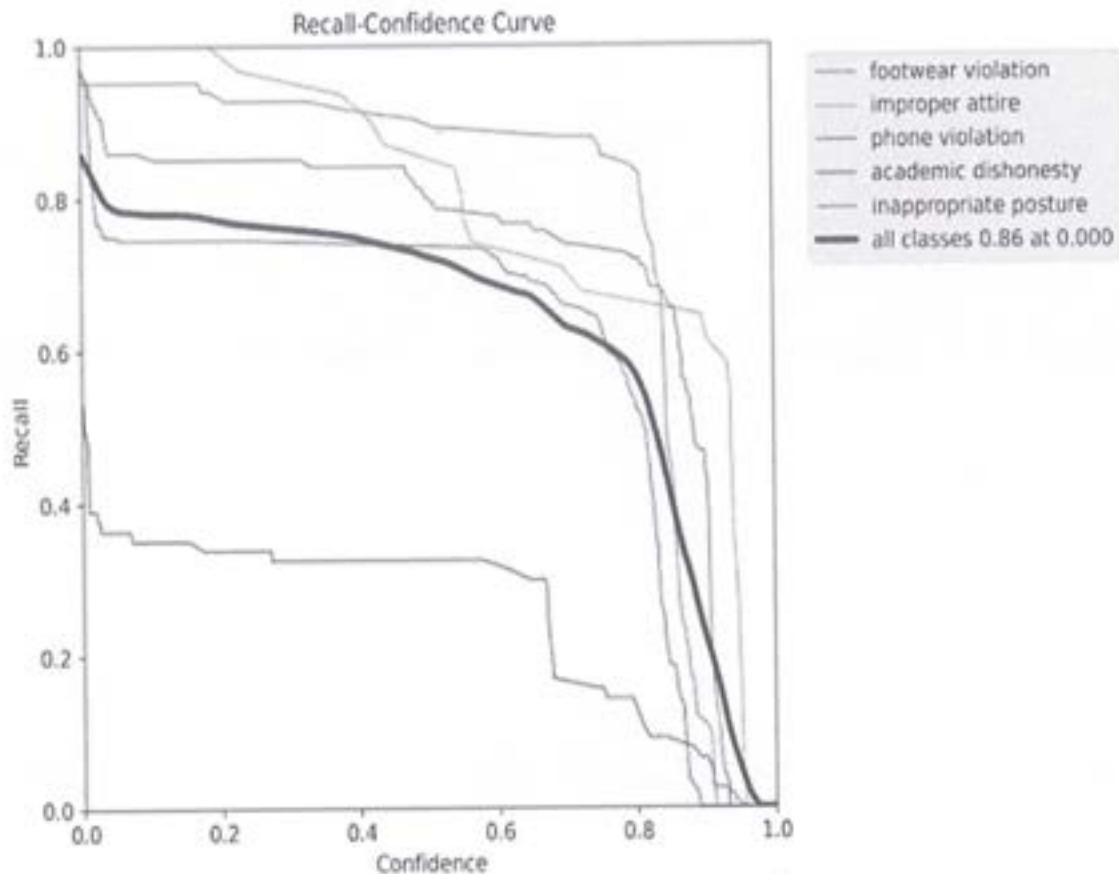
Hình 3-3. Đường cong Precision

(Nguồn: dữ liệu thu thập tự tạo)

Các lớp "improper attire" và "inappropriate posture" có giá trị precision cao trong hầu hết các ngưỡng độ tin cậy, cho thấy mô hình rất chính xác trong việc nhận diện các lớp này khi đưa ra dự đoán với độ tin cậy cao. Tuy nhiên, lớp "academic dishonesty" có precision thấp rõ rệt ở các ngưỡng độ tin cậy, chỉ ra rằng mô hình gặp khó khăn trong việc dự đoán chính xác lớp này. Khi độ tin cậy tăng (gần 1.0), precision của tất cả các lớp thường tăng vì mô hình trở nên chọn lọc hơn, chỉ đưa ra dự đoán khi nó rất chắc chắn.

Đường kẻ xanh đậm tổng hợp hiệu suất của tất cả các lớp. Đường này cho thấy sự gia tăng đều đặn của precision khi độ tin cậy tăng, đạt đỉnh gần như chính xác tuyệt đối (1.0) khi ngưỡng độ tin cậy rất cao (ví dụ, 0.987).

b. Đường cong Recall



Hình 3-4. Đường cong Recall

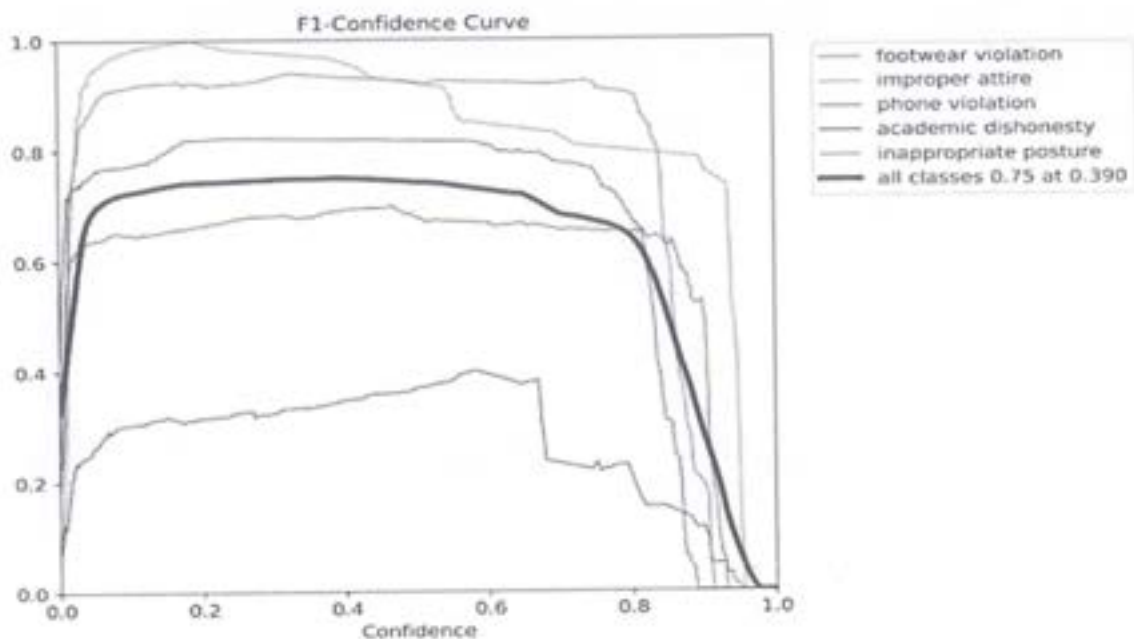
(Nguồn: dữ liệu thu thập tự tạo)

Các lớp "improper attire" và "inappropriate posture" duy trì recall cao trong một phạm vi lớn các ngưỡng độ tin cậy, cho thấy mô hình rất giỏi trong việc phát hiện hầu hết các trường hợp của chúng, ngay cả với các ngưỡng nghiêm ngặt. Tuy nhiên, lớp "academic dishonesty" giảm mạnh recall khi độ tin cậy tăng, cho thấy mô hình gặp khó khăn trong việc phát hiện lớp này một cách đáng tin cậy.

Recall tổng hợp cho tất cả các lớp bắt đầu ở mức cao (~ 0.86 tại ngưỡng độ tin cậy 0.0), nhưng giảm dần khi ngưỡng độ tin cậy tăng. Điều này phản ánh sự đánh đổi giữa việc duy trì recall cao và tăng precision, vì các ngưỡng nghiêm ngặt hơn dẫn đến việc bỏ sót các phát hiện.

Tại các ngưỡng độ tin cậy rất cao (~ 0.9), recall giảm mạnh ở hầu hết các lớp vì chỉ giữ lại những dự đoán có độ tin cậy cao nhất. Mặc dù các dự đoán này có khả năng chính xác cao (dẫn đến precision cao như đã thấy trong Đường cong Precision-Confidence), nhưng nhiều true positives bị bỏ sót, khiến recall giảm mạnh.

c. Đường cong F1 Score



Hình 3-5. Đường cong F1

(Nguồn: dữ liệu thu thập tự tạo)

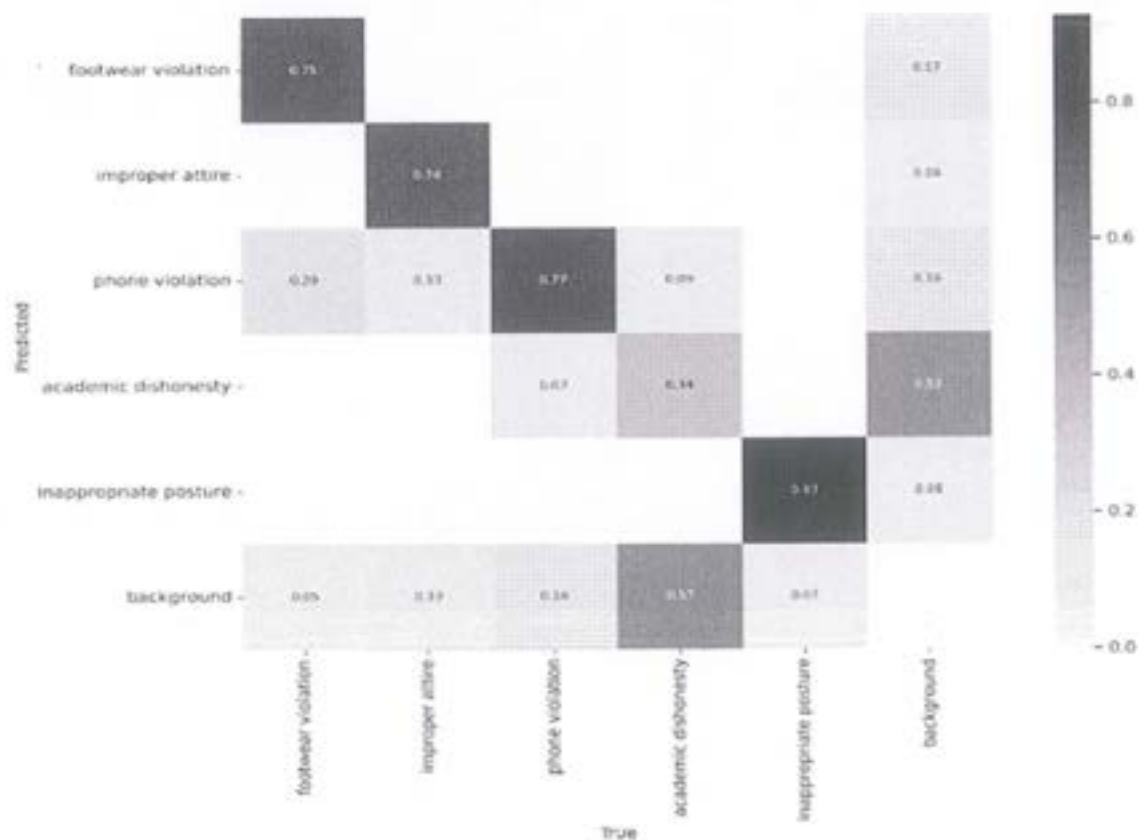
Mỗi đường cong màu (ngoại trừ đường xanh dày) tương ứng với một lớp cụ thể. Điểm F1 cao hơn cho thấy sự cân bằng tốt hơn giữa precision và recall cho lớp đó tại ngưỡng độ tin cậy nhất định.

Đường cong F1 tổng hợp, là đường xanh dày, đại diện cho hiệu suất kết hợp của tất cả các lớp. Giá trị "0.75 tại 0.390" cho thấy điểm F1 cao nhất đạt được ở tất cả các lớp ($F1 = 0.75$) tại ngưỡng độ tin cậy 0.390.

Từ đồ thị, lớp "improper attire" đạt điểm F1 cao nhất trên một phạm vi rộng các ngưỡng độ tin cậy, cho thấy hiệu suất tốt hơn của mô hình đối với lớp này. Trong khi đó, lớp "academic dishonesty" có điểm F1 thấp nhất, cho thấy mô hình gặp khó khăn trong việc phát hiện lớp này một cách chính xác. Do đó, lớp này cần thêm dữ liệu huấn luyện và cải thiện chú thích để nâng cao hiệu quả.

Đường cong F1 tổng hợp đạt đỉnh tại 0.390, có nghĩa là ngưỡng này cung cấp sự đánh đổi tối ưu giữa precision và recall cho tất cả các lớp.

d. Ma trận nhầm lẫn



Hình 3-6. Ma trận nhầm lẫn

(Nguồn: dữ liệu thu thập tự tạo)

Các giá trị trên đường chéo đại diện cho mức độ chính xác mà mô hình của bạn dự đoán đúng mỗi lớp. Do đó, lớp "tư thế không phù hợp" được phát hiện rất tốt (0.93), trong khi lớp "gian lận học thuật" (0.34) cần cải thiện.

Mặt khác, các giá trị cao ở các ô ngoài đường chéo chỉ ra sự nhầm lẫn giữa các lớp hoặc nền. Việc nền bị phân loại sai thành "gian lận học thuật" (0.57) cho thấy có thể xảy ra nhiều trường hợp dương tính giả cho lớp này; "vi phạm giấy dép" đôi khi bị nhầm lẫn với "vi phạm điện thoại" (0.20).

Dựa trên kết quả đã cho, bộ dữ liệu cần được cải thiện để bao quát sự đa dạng của các lớp, đặc biệt là các lớp có giá trị đường chéo thấp, trong trường hợp này là "gian lận học thuật".

3.6. Giao diện hệ thống

3.6.1. Giao diện phía người dùng

Khi người dùng thực hiện đăng nhập thành công vào hệ thống, màn hình trang chủ sẽ hiển thị như sau:



Hình 3-7. Giao diện trang chủ.

Trong giao diện trang chủ có 1 số chức năng như: Đăng ký lịch thực hành, người dùng và thống kê hành vi.

- ✓ Đăng ký lịch thực hành: Để người dùng sau khi đăng nhập có thể đăng ký lịch thực hành trên hệ thống
- ✓ Người dùng: đây là chức năng cho phép người dùng có thể biết được lịch đăng ký đang ở chế độ chờ, hoặc người dùng biết được thông tin khi lịch thực hành được phê duyệt.
- ✓ Thông kê hành vi: Chức năng này sẽ lưu lại các hành vi vi phạm mà hệ thống phát hiện.

- Các giao diện còn lại được trình bày dưới phụ lục.

3.7. Kết chương

Chương 3 đã trình bày chi tiết quá trình cài đặt và kiểm thử hệ thống, từ cấu trúc dự án, các thành phần frontend và backend, cho đến giao diện người dùng. Kết quả kiểm thử cho thấy hệ thống hoạt động ổn định, đáp ứng được các yêu cầu nghiệp vụ đã đặt ra.

Quá trình huấn luyện mô hình được thực hiện trên nền tảng Google Colab sử dụng GPU, giúp tối ưu hóa tốc độ xử lý và hiệu suất học của mô hình. Việc gán nhãn dữ liệu được thực hiện cẩn thận bằng công cụ LabelImg với cấu trúc tổ chức tập dữ liệu hợp lý, đảm bảo phân bố lớp cân bằng và tính nhất quán trong gán nhãn. Các siêu tham số huấn luyện được lựa chọn tối ưu như kích thước ảnh, số lượng epoch và batch size, phù hợp với mục tiêu phát hiện hành vi chính xác và hiệu quả.

Kết quả thực nghiệm cho thấy mô hình đạt hiệu suất tốt ở nhiều lớp, đặc biệt là các hành vi như "trang phục không phù hợp" và "tư thế không chuẩn" với các chỉ số Precision, Recall và F1 Score ở mức cao. Tuy nhiên, lớp "gian lận học thuật" vẫn còn gặp nhiều khó khăn do đặc thù phức tạp và sự chồng chéo về ngữ cảnh trong hình ảnh. Ma trận nhầm lẫn và các biểu đồ phân tích đã chỉ ra những điểm mạnh cũng như hạn chế của mô hình, từ đó cung cấp cơ sở để cải tiến bộ dữ liệu và thuật toán nhận diện trong tương lai.

Ngoài ra, hệ thống còn xây dựng giao diện trực quan và thân thiện với người dùng, hỗ trợ các chức năng đăng ký lịch thực hành, theo dõi vi phạm và thống kê hành vi. Việc áp dụng MySQL trong lưu trữ dữ liệu cùng Spring Boot trong xử lý backend giúp hệ thống hoạt động ổn định, dễ bảo trì và mở rộng.

Tổng thể, hệ thống không chỉ cung cấp giải pháp kỹ thuật khả thi trong giám sát hành vi vi phạm mà còn đặt nền móng cho các nghiên cứu mở rộng trong lĩnh vực ứng dụng trí tuệ nhân tạo vào giáo dục và quản lý đào tạo. Trong tương lai, hệ thống có thể được nâng cấp để tích hợp thêm các mô hình thị giác máy tính tiên tiến, cải thiện độ chính xác và khả năng thích ứng với các tình huống phức tạp hơn trong môi trường thực tế.

KẾT LUẬN

Đề án “ phát hiện và nhận diện hành vi qua video ứng dụng trong giám sát phòng thực hành “ đã hoàn thành toàn bộ và bám sát nội dung đề cương được phê duyệt, các kết quả chính được thực hiện trong đề án gồm:

1. Phát triển một hệ thống giám sát thông minh sử dụng mô hình học sâu YOLOv5 để phát hiện các đối tượng và hành vi.
2. Phân loại chính xác các vi phạm quy tắc, như quy định vi phạm về giày dép, trang phục không phù hợp, dùng điện thoại hay gian lận trong thi cử.
3. Triển khai một hệ thống thời gian thực có khả năng phát cảnh báo khi phát hiện vi phạm, đồng thời lưu trữ dữ liệu để phân tích sau này.

Đề án được nghiên cứu và thực hiện trong thời gian ngắn cùng với kiến thức có hạn của tác giả nên không tránh khỏi thiếu sót, kính mong nhận được ý kiến đóng góp của quý thầy cô và bạn đọc.

DANH MỤC TÀI LIỆU THAM KHẢO

[1] C. -Y. Wang, H. -Y. Mark Liao, Y. -H. Wu, P. -Y. Chen, J. -W. Hsieh and I. -H. Yeh, "CSPNet: A New Backbone that can Enhance Learning Capability of CNN," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2020, pp. 1571-1580, doi: 10.1109/CVPRW50498.2020.00203.

[2] HE, Kaiming, ZHANG, Xiangyu, REN, Shaoqing, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, 2015, vol. 37, no 9, p. 1904-1916.

[3] WANG, Chien-Yao, LIAO, Hong-Yuan Mark, WU, Yuch-Hua, et al. CSPNet: A new backbone that can enhance learning capability of CNN. In : *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. 2020. p. 390-391.

[4] Ultralytics YOLOv5 [Online].Available:

<https://docs.ultralytics.com/vi/models/yolov5/>

[5] Reactjs Tutorial [Online].Available:

<https://react.dev/learn>

[6] MySQL Tutorial [Online].Available:

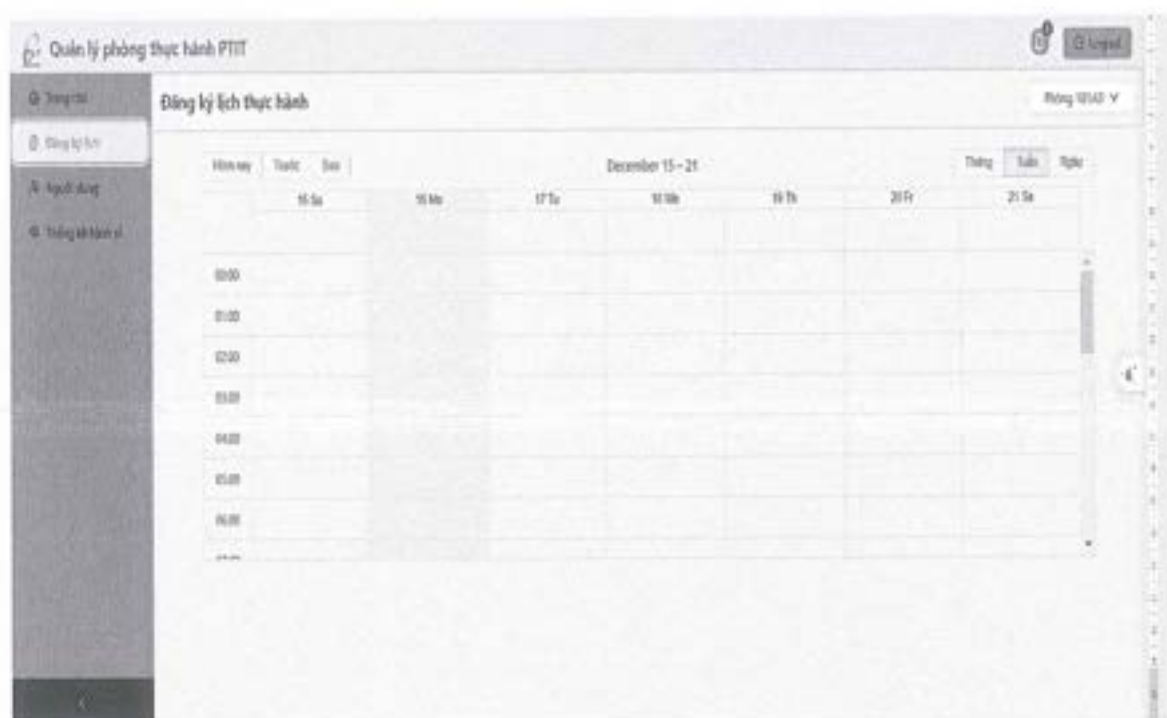
<https://dev.mysql.com/doc/>

[7] Spring Boot Tutorial [Online].Available:

<https://spring.io/projects/spring-boot>

PHỤ LỤC 1

Khi click vào màn đăng ký lịch giao diện sẽ được hiển thị như sau:



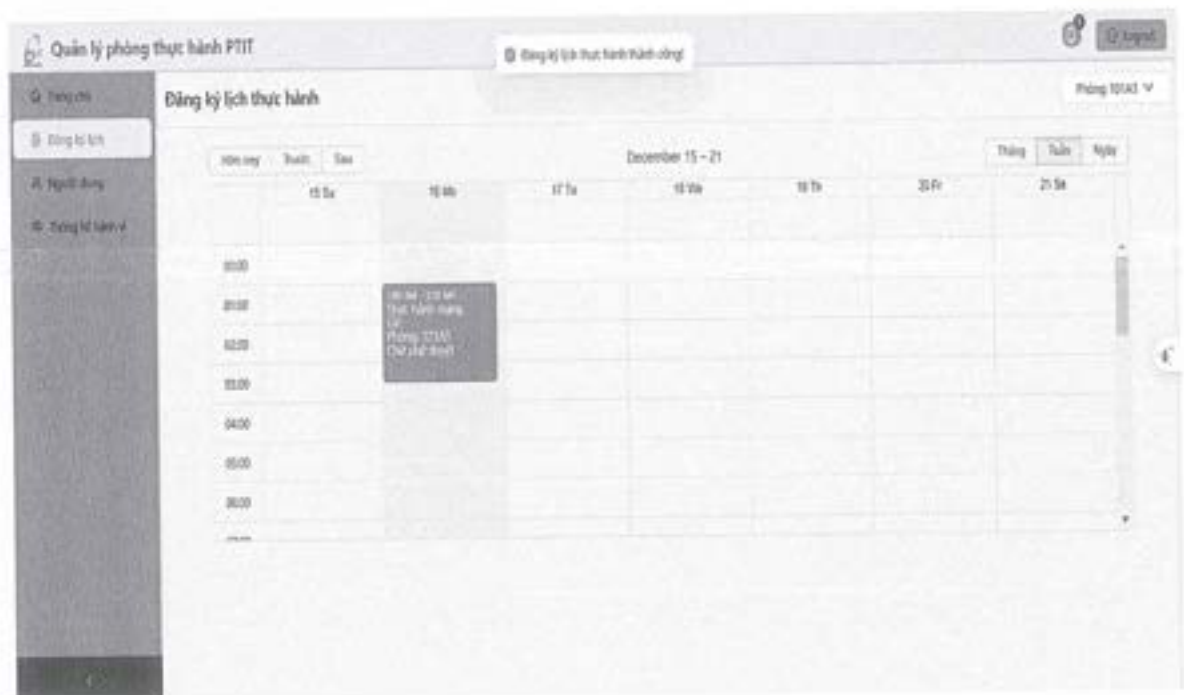
Hình PL 1-1. Giao diện đăng ký lịch thực hành.

Khi bấm vào chọn vào một ngày trong lịch sẽ hiển thị modal nhập thông tin đăng ký lịch:



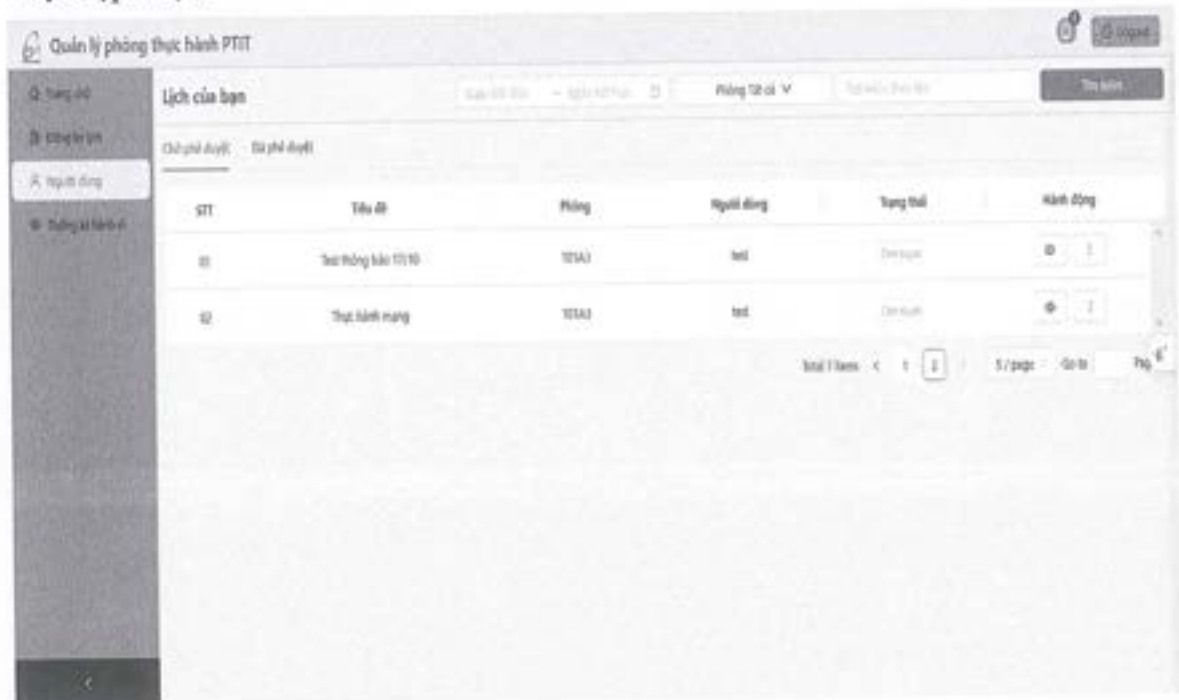
Hình PL 1-2. Model đăng ký lịch thực hành.

Giao diện khi đăng ký thành công:



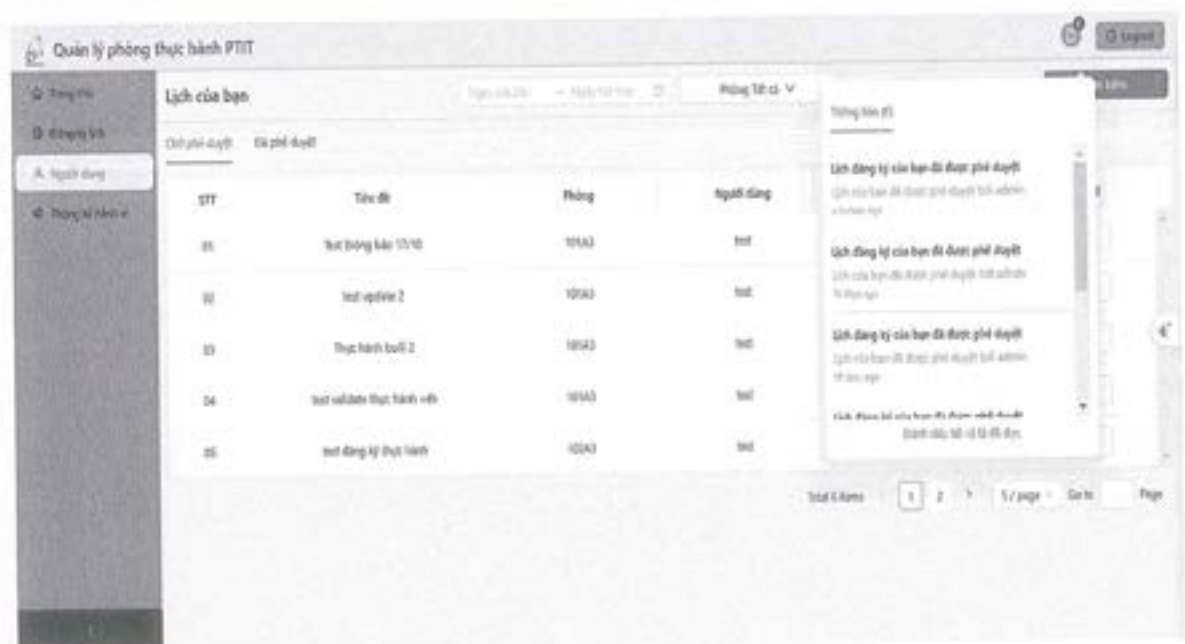
Hình PL 1-3. Giao diện sau khi đăng ký lịch thực hành thành công .

Lịch thực hành sau khi đăng ký sẽ ở trạng thái chờ phê duyệt và có thể xóa hoặc cập nhật:



Hình PL 1-4. Giao diện tab chờ phê duyệt của màn người dùng .

Khi lịch thực hành đăng ký được phê duyệt bởi quản trị viên thì sẽ nhận được thông báo



Hình PL 1-5. Thông báo khi lịch thực hành đã được phê duyệt.



Hình PL 1-6. Giao diện đăng ký lịch khi đã được phê duyệt



Hình PL 1-7. Giao diện xem chi tiết lịch thực hành ở tab đã phê duyệt

Quản lý phòng thực hành PTIT

Thống kê hành vi

STT	Tên đề	Phòng	Số học	Hành động
01	Thực hành phòng 101	01A01	01 học	
02	Thực hành phòng 101	01A01	01 học	
03	Thực hành phòng 101	01A01	01 học	
04	Thực hành phòng 101	01A01	01 học	
05	Thực hành phòng 101	01A01	01 học	

Tổng 11 học

Hình PL 1-8. Giao diện thống kê hành vi lịch thực hành

Quản lý phòng thực hành PTIT

Thống kê hành vi

STT	Link hình ảnh	Phòng	Ngày học
01		01A01	10/01/2024 09:10 PM
02		01A01	10/01/2024 09:10 PM
03		01A01	10/01/2024 09:10 PM

Hình PL 1-9 Model thống kê hình ảnh hành vi



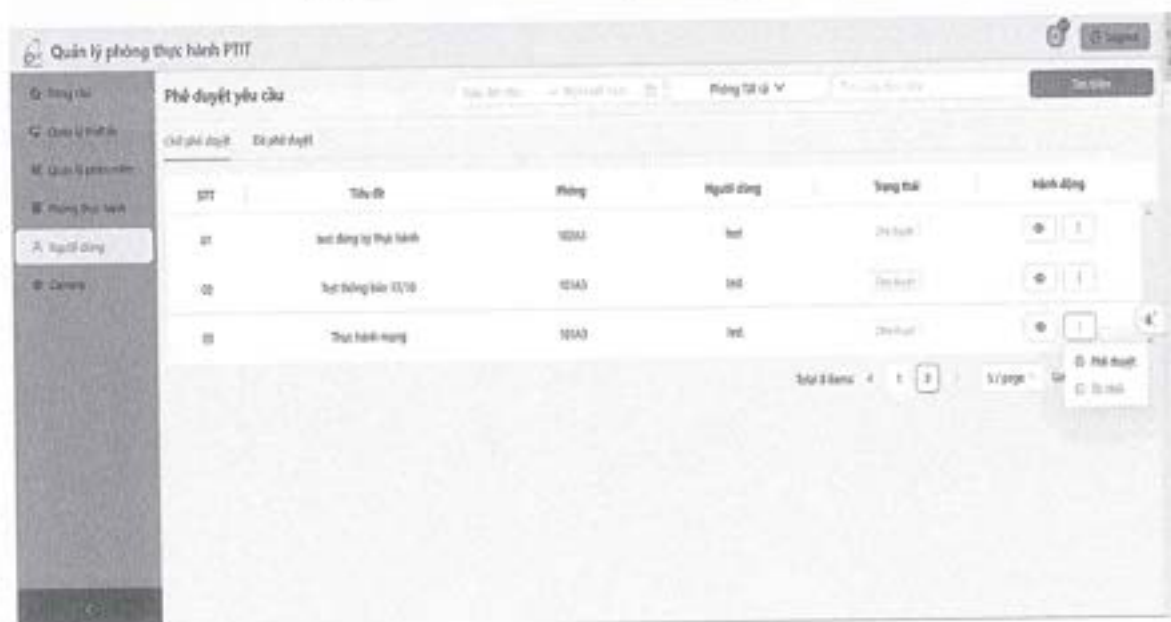
Hình PL 1-10 Model xem chi tiết hình ảnh hành vi

1.2 Giao diện phía quản trị viên

Khi có người dùng đăng ký lịch thực hành phía quản trị viên sẽ nhận được thông báo và khi bấm vào thông báo sẽ điều hướng tới tab chờ phê duyệt của màn người dùng của quản trị viên.

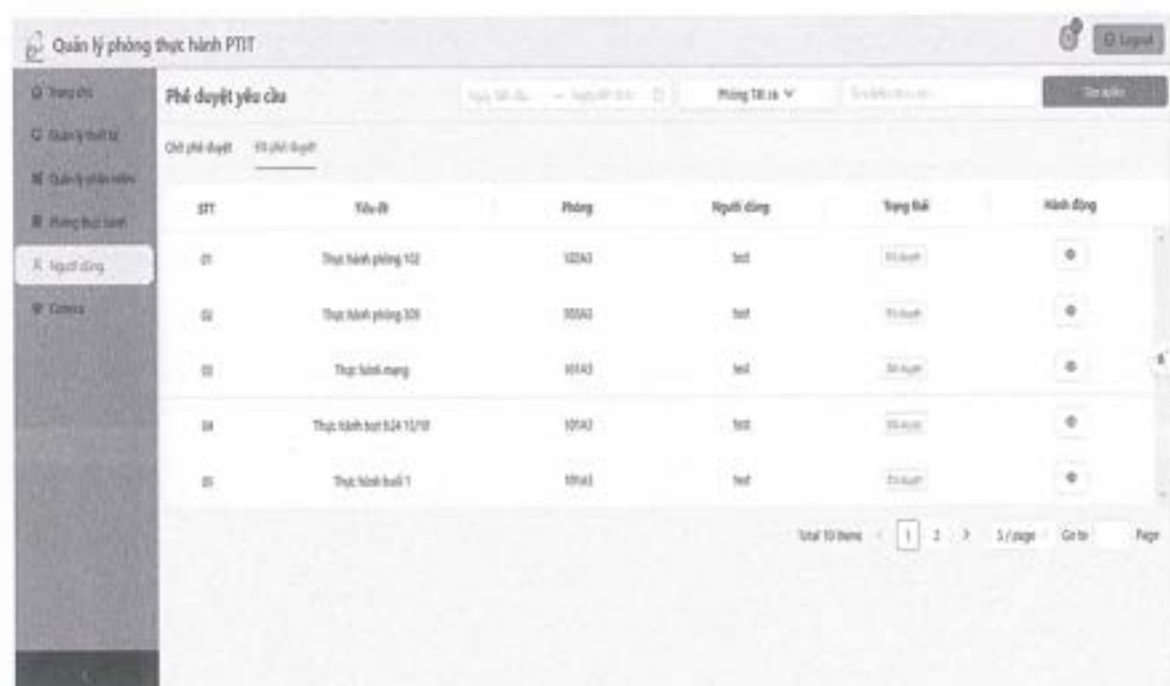


Hình PL 1-11. Giao diện của quản trị viên.



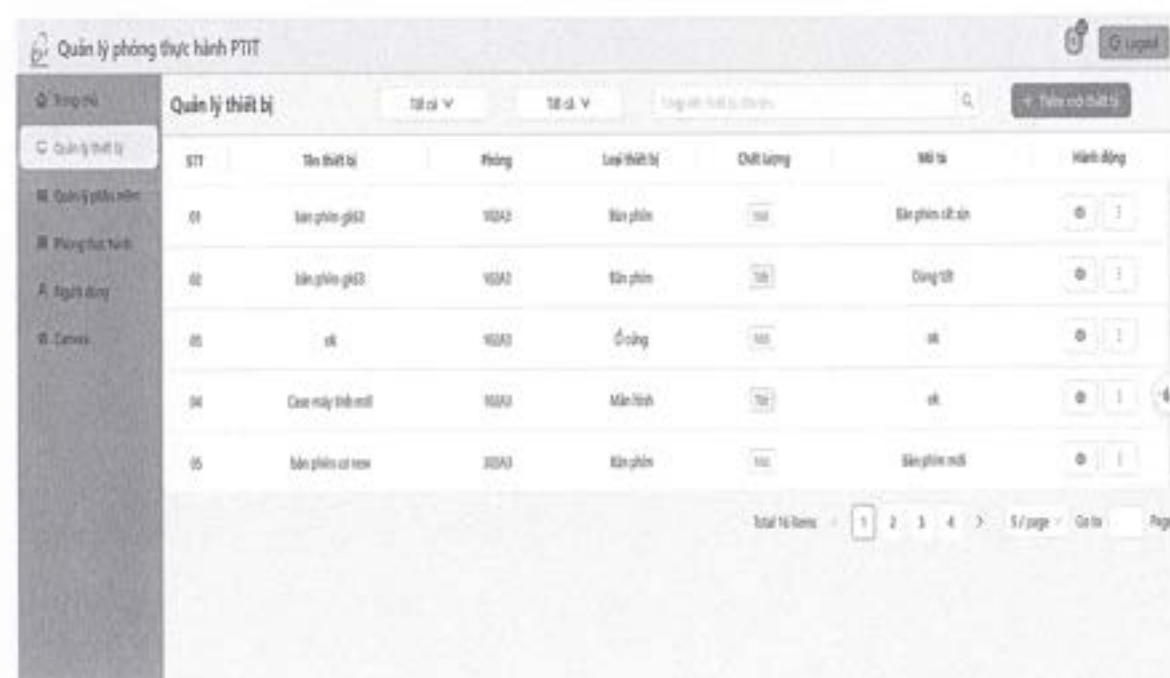
Hình PL 1-12. Giao diện phê duyệt hoặc từ chối lịch thực hành.

Khi quản trị viên phê duyệt thì lịch thực hành sẽ chuyển trạng thái là đã phê duyệt và gửi thông báo cho người dùng đăng ký



Hình PL 1-13. Giao diện tab đã phê duyệt của quản trị viên.

Giao diện màn hình quản lý thiết bị



Hình PL 1-14. Giao diện màn hình quản lý thiết bị.

Quản lý phòng thực hành PTTT

Quản lý thiết bị

STT	Tên thiết bị	Phòng	Loại thiết bị	Chất lượng
01	Sàn phòng giặt	100A1	Sàn phòng	OK
02	Sàn phòng giặt	100A2	Sàn phòng	OK
03	OK	100A3	Ổ cứng	OK
04	Quạt máy treo tường	100A4	Sàn phòng	OK
05	Sàn phòng giặt	100A5	Sàn phòng	OK

Thêm mới thiết bị

Tên thiết bị

Tên phòng

Loại thiết bị

Chất lượng

Mô tả

X Hủy Lưu

Hình PL 1-15. Giao diện màn hình thêm mới thiết bị.

Quản lý phòng thực hành PTTT

Quản lý thiết bị

STT	Tên thiết bị	Phòng	Loại thiết bị	Chất lượng	Mô tả	Hành động
01	Sàn phòng giặt	100A1	Sàn phòng	OK	Sàn phòng giặt	
02	Sàn phòng giặt	100A2	Sàn phòng	OK	Sàn phòng giặt	
03	OK	100A3	Ổ cứng	OK	Ổ cứng	
04	Quạt máy treo tường	100A4	Máy móc	OK	Máy móc	
05	Sàn phòng giặt	100A5	Sàn phòng	OK	Sàn phòng giặt	

Kiểm tra nhận xét

Bạn có chắc muốn xóa sàn phòng giặt?

OK Hủy

total 16 items 1 2 3 4 5 1 page Go to Page

Hình PL 1-16. Giao diện màn hình xóa thiết bị.

Giao diện màn hình quản lý phần mềm

Quản lý phòng thực hành CNTT

Quản lý phần mềm

Tìm kiếm phần mềm

Thêm mới phần mềm

STT	Tên phần mềm	Phiên bản	Ngày cài đặt	Ngày hết hạn	Tạng thái	Hành động
1	Eclipse	1.8.8	2024-03-08	2025-03-21	Hoạt động	✎ ✕
2	Eclipse	1.8.2	2024-03-08	2024-03-10	Ngừng hoạt động	✎ ✕
3	Eclipse	1.8.8	2024-03-08	2024-03-10	Hoạt động	✎ ✕
4	Eclipse 1.8.8	1.8.8	2024-03-08	2024-03-10	Ngừng hoạt động	✎ ✕
5	Eclipse	1.8.8	2024-03-08	2024-03-10	Hoạt động	✎ ✕

Total 11 items: 1 2 3 > 1 / page Go to Page

Hình PL 1-17. Giao diện màn hình quản lý phần mềm.

Quản lý phòng thực hành CNTT

Quản lý phần mềm

Tìm kiếm phần mềm

Thêm mới phần mềm

STT	Tên phần mềm	Phiên bản	Ngày cài đặt	Ngày hết hạn
1	Eclipse	1.8.8	2024-03-08	2025-03-21
2	Eclipse	1.8.2	2024-03-08	2024-03-10
3	Eclipse	1.8.8	2024-03-08	2024-03-10
4	Eclipse 1.8.8	1.8.8	2024-03-08	2024-03-10
5	Eclipse	1.8.8	2024-03-08	2024-03-10

Thêm mới phần mềm

Phiên bản

Tạng thái

Tên phần mềm

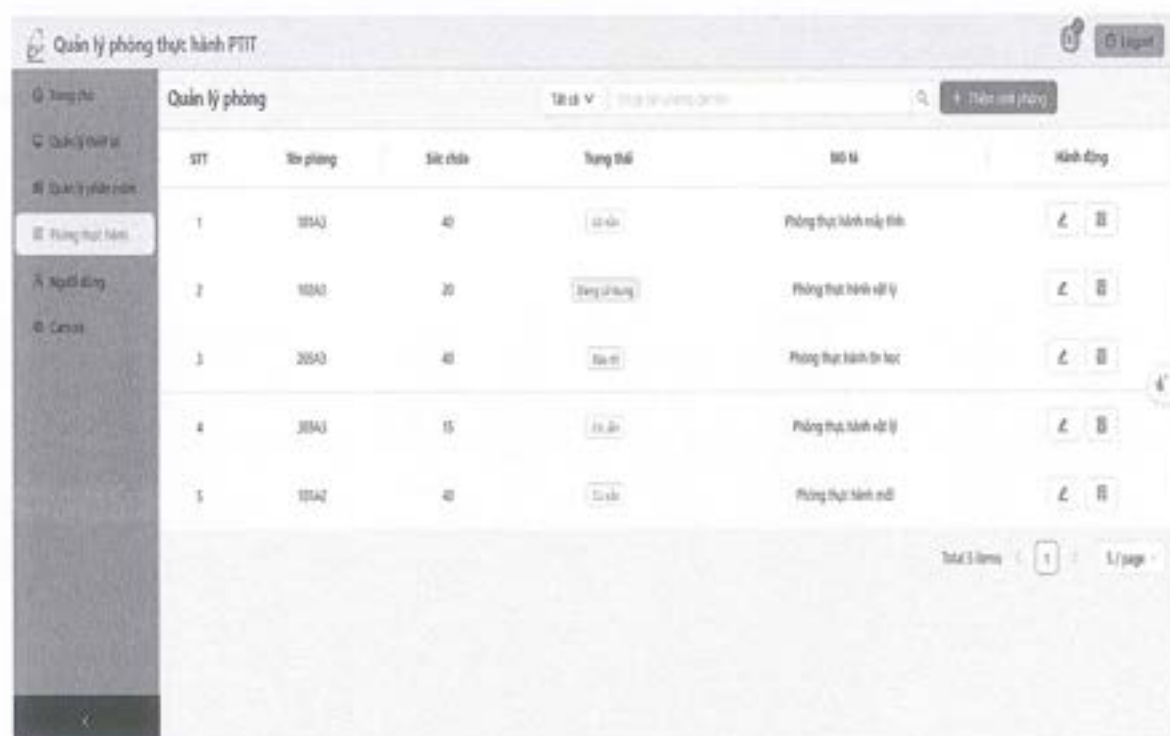
Ngày hết hạn

Giới chú

X Hủy [Lưu](#)

Hình PL 1-181. Giao diện màn hình thêm mới phần mềm.

Giao diện quản lý phòng thực hành



Quản lý phòng thực hành PTIT

Quản lý phòng

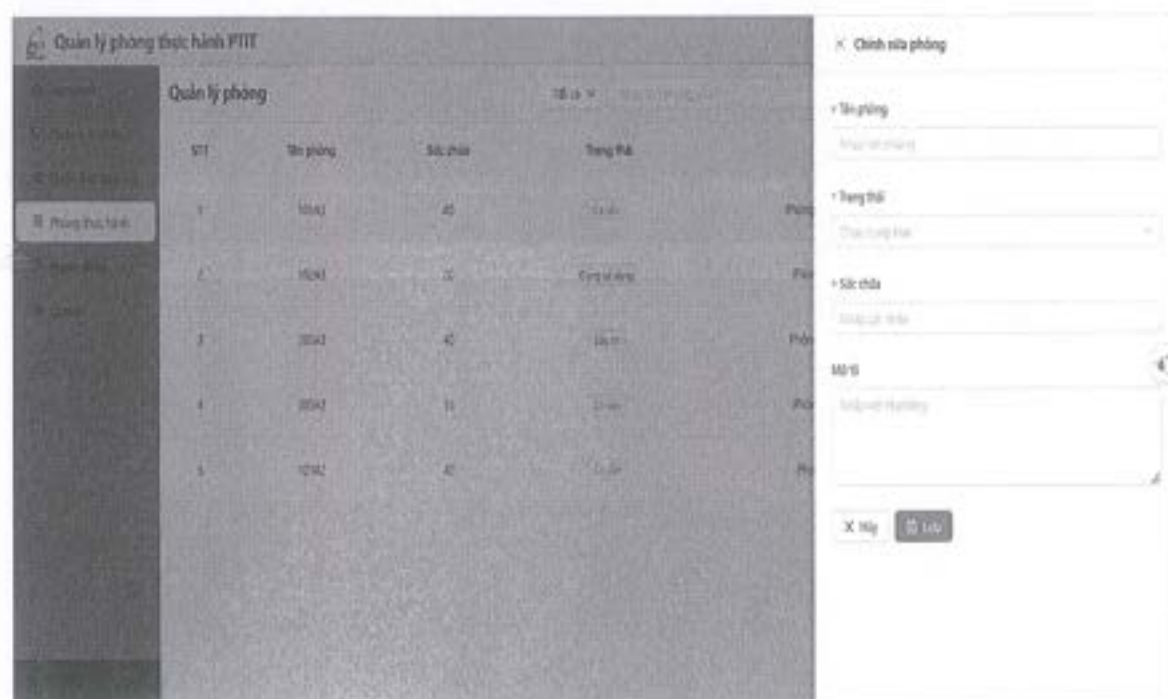
Tìm kiếm

+ Thêm mới phòng

STT	Tên phòng	Sức chứa	Trạng thái	Mô tả	Hành động
1	101A1	40	Đã sẵn	Phòng thực hành máy tính	✎ ✕
2	101A2	20	Đang chờ dùng	Phòng thực hành vật lý	✎ ✕
3	201A3	40	Đã sẵn	Phòng thực hành tin học	✎ ✕
4	301A3	15	Đã sẵn	Phòng thực hành vật lý	✎ ✕
5	101A2	40	Đã sẵn	Phòng thực hành mới	✎ ✕

Total 5 items 1 1 / page

Hình PL 1-19. Giao diện quản lý phòng thực hành.



Quản lý phòng thực hành PTIT

Quản lý phòng

Tìm kiếm

+ Thêm mới phòng

✕ Đóng cửa phòng

+ Tên phòng

Phòng thực hành

+ Trạng thái

Phòng thực hành

+ Sức chứa

Phòng thực hành

Mô tả

Phòng thực hành mới

X Hủy Lưu

Hình PL 1-20. Giao diện thêm mới phòng thực hành.

Quản lý phòng thực hành PTIT

X. Chỉnh sửa phòng

Tên phòng: 101A3

Tầng trệt: Sân sáng

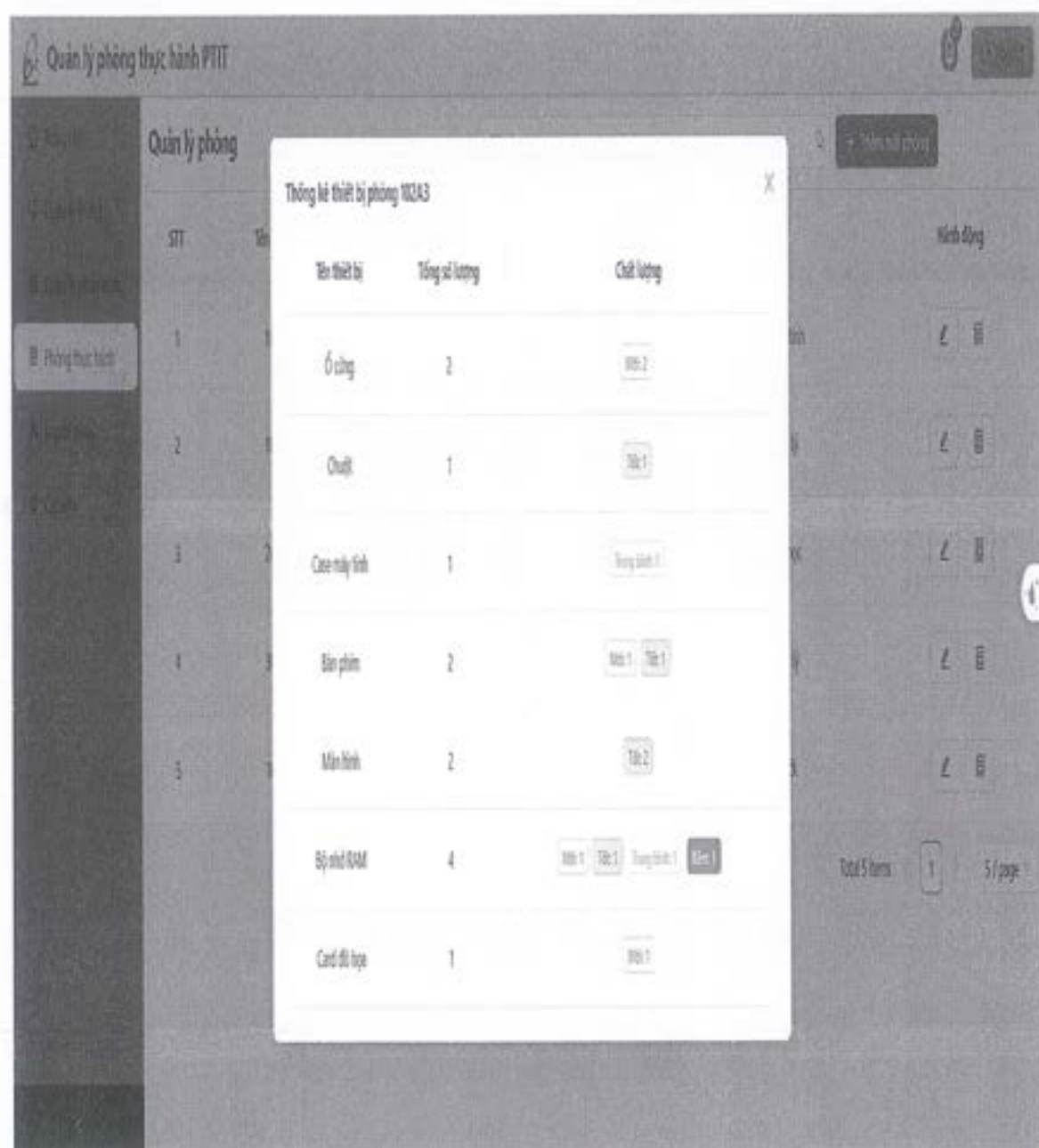
Sức chứa: 40

Mô tả: Phòng thực hành máy tính

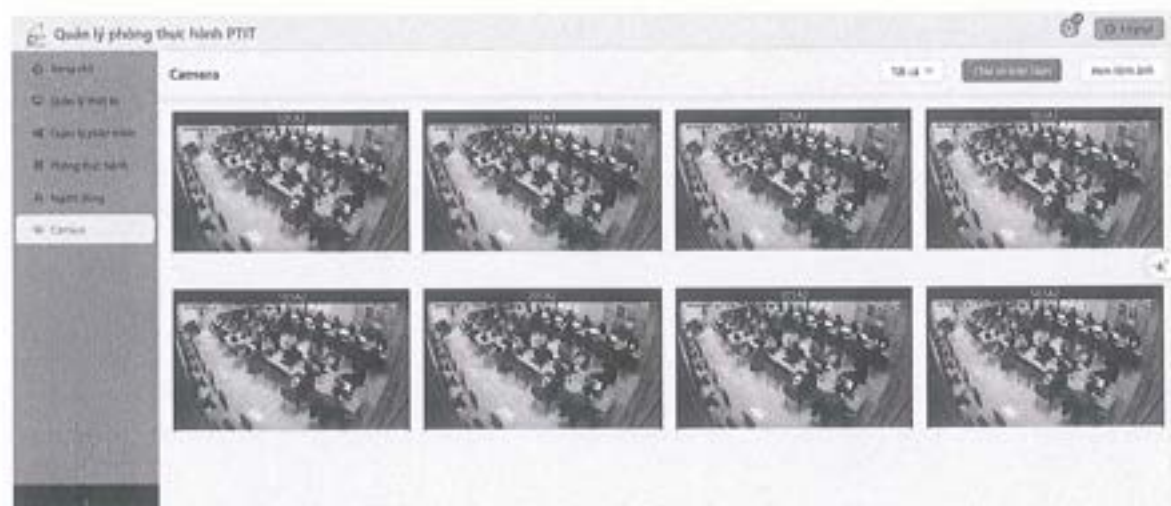
X Hủy Lưu

STT	Tên phòng	Sức chứa	Tầng trệt	Phòng
1	101A3	40	Sân sáng	Phòng
2	101A5	30	Phòng thí nghiệm	Phòng
3	101A5	40	Phòng	Phòng
4	101A1	15	Phòng	Phòng
5	101A2	40	Sân sáng	Phòng

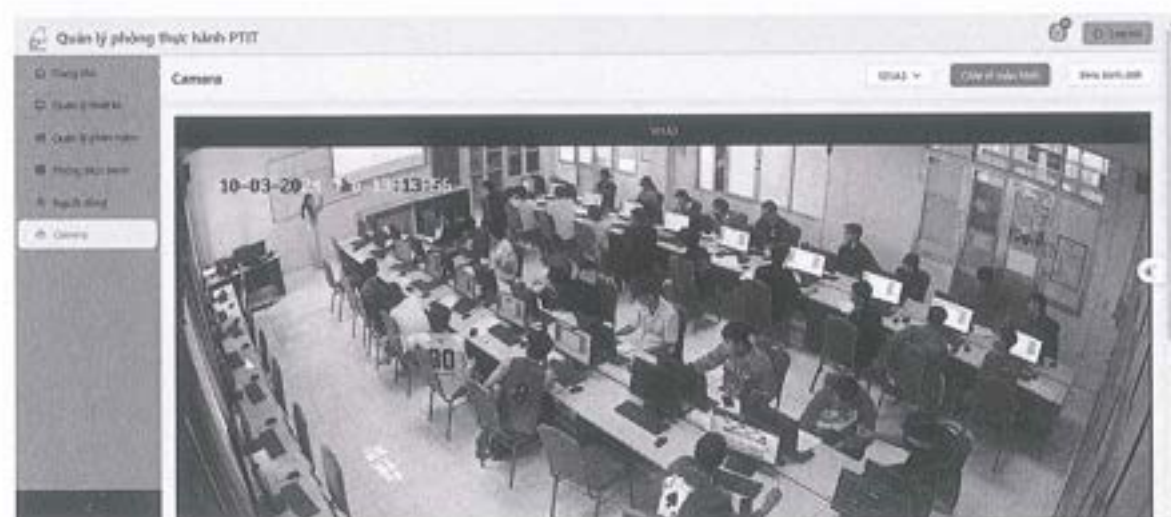
Hình PL 1-212. Giao diện chỉnh sửa phòng thực hành.



Hình PL 1-22. Giao diện thống kê thiết bị phòng thực hành.



Hình PL 1-23 Giao diện camera phòng thực hành



Hình PL 1.24 Giao diện camera phòng thực hành



BÁO CÁO KIỂM TRA TRÙNG LẬP

Thông tin tài liệu

Tên tài liệu:	B23CHIS009-Vương Việt Hưng-M23CQIS01-B
Tác giả:	HUNG VIET
Điểm trùng lặp:	8
Thời gian tải lên:	13:26 29/07/2025
Thời gian sinh báo cáo:	13:27 29/07/2025
Các trang kiểm tra:	78/78 trang



Kết quả kiểm tra trùng lặp



Có 8% nội dung trùng lặp



Có 92% nội dung không trùng lặp



Có 0% nội dung người dùng loại trừ



Có 0% nội dung hệ thống bỏ qua

Nguồn trùng lặp tiêu biểu

123docz.net tailieu.vn arxiv.org

Học viên
(Ký và ghi rõ họ tên)

Vương Việt Hưng

Người hướng dẫn
(Ký và ghi rõ họ tên)

Nguyễn Trọng Khánh

BÁO CÁO GIẢI TRÌNH
SỬA CHỮA, HOÀN THIỆN ĐỀ ÁN TỐT NGHIỆP

Họ và tên học viên: Vương Việt Hùng

Chuyên ngành: HTTT

Khóa: 2023 đợt 1

Tên đề tài: Phát hiện và nhận diện hành vi qua video ứng dụng trong giám sát phòng thực hành

Người hướng dẫn khoa học: PGS. TS. Nguyễn Trọng Khánh

Ngày bảo vệ: 19/07/2025

Các nội dung học viên đã sửa chữa, bổ sung trong đề án tốt nghiệp theo ý kiến đóng góp của Hội đồng chấm đề án tốt nghiệp:

TT	Ý kiến hội đồng	Sửa chữa của học viên
1	Chỉnh sửa lại theo nhận xét của phản biện	Tiếp thu góp ý của Hội đồng, tác giả đã chỉnh sửa theo yêu cầu phản biện
2	Làm rõ mục tiêu đề án, làm rõ nhãn dữ liệu	Tiếp thu góp ý của Hội đồng, tác giả đã bổ sung làm rõ mục tiêu đề án tại mục 1.2 trong chương 1 và làm rõ nhãn dữ liệu tại mục 3.4.4 trong chương 3
3	Mô tả chi tiết mô hình, tham số, các bước huấn luyện, độ đo.	Tiếp thu góp ý của Hội đồng, tác giả đã mô tả chi tiết mô hình, tham số, các bước huấn luyện tại mục 3.5 trong chương 3
4	Làm rõ việc tích hợp mô hình và hệ thống theo dõi đang triển khai	Tiếp thu góp ý của Hội đồng, tác giả đã làm rõ việc tích hợp mô hình và hệ thống theo dõi đang triển khai tại mục 2.3.7 trong chương 2.
5	Chỉnh sửa lại bố cục đề án	Tiếp thu góp ý của Hội đồng, tác giả đã chỉnh sửa lại bố cục đề án.

Hà Nội, ngày 01 tháng 8 năm 2025

Ký xác nhận của

CHỦ TỊCH HỘI ĐỒNG
CHẤM ĐỀ ÁN



PGS. TS Lê Hữu Lập

THƯ KÝ HỘI ĐỒNG



TS. Đinh Trường Duy

NGƯỜI HƯỚNG DẪN
KHOA HỌC



PGS. TS Nguyễn Trọng Khánh

HỌC VIÊN



Vương Việt Hùng

**BIÊN BẢN
HỌP HỘI ĐỒNG CHẤM ĐỀ ÁN TỐT NGHIỆP THẠC SĨ**

Căn cứ quyết định số Quyết định số 1098/QĐ-HV ngày 26 tháng 06 năm 2025 của Giám đốc Học viện Công nghệ Bưu chính Viễn thông về việc thành lập Hội đồng chấm đề án tốt nghiệp thạc sĩ. Hội đồng đã họp vào hồi...9...giờ...10...phút, ngày 19 tháng 07 năm 2025 tại Học viện Công nghệ Bưu chính Viễn thông để chấm đề án tốt nghiệp thạc sĩ cho:

Học viên: **Vương Việt Hùng**

Tên đề án tốt nghiệp: **Phát hiện và nhận diện hành vi qua video ứng dụng trong giám sát phòng thực hành**

Chuyên ngành: **Hệ thống thông tin**

Mã số: **8480104**

Các thành viên của Hội đồng chấm đề án tốt nghiệp có mặt: ...5.../ 05

TT	HỌ VÀ TÊN	TRÁCH NHIỆM TRONG HD	GHI CHÚ
1	PGS.TS. Lê Hữu Lập	Chủ tịch	
2	TS. Đinh Trường Duy	Thư ký	
3	PGS.TS. Nguyễn Long Giang	Phản biện 1	
4	PGS.TS. Phạm Văn Cường	Phản biện 2	
5	TS. Nguyễn Chí Thành	Ủy viên	

Các nội dung thực hiện:

1. Chủ tịch Hội đồng điều khiển buổi họp. Công bố quyết định của Giám đốc Học viện Công nghệ Bưu chính Viễn thông về việc thành lập Hội đồng chấm đề án tốt nghiệp thạc sĩ.
2. Người hướng dẫn khoa học hoặc thư ký đọc lý lịch khoa học và các điều kiện bảo vệ đề án tốt nghiệp của học viên. (có bản lý lịch khoa học và kết quả các môn học cao học của học viên kèm theo).
3. Học viên trình bày tóm tắt đề án tốt nghiệp.
4. Phản biện 1 đọc nhận xét (có văn bản kèm theo)
5. Phản biện 2 đọc nhận xét (có văn bản kèm theo)

6. Các câu hỏi của thành viên Hội đồng:

Giải thích quá trình tái huấn luyện và định chỉnh các tham số cho mô hình YOLO
làm rõ các bài toán cần giải quyết và hệ dữ liệu sử dụng cho mỗi bài toán

7. Trả lời của học viên:

Học viên thảo luận và trả lời câu hỏi của Hội đồng
Học viên tiếp thu ý kiến và tự chỉnh sửa

8. Thư ký đọc nhận xét về quá trình thực hiện đề án tốt nghiệp của học viên (có văn bản kèm theo).

9. Hội đồng họp riêng:

- Bầu Ban kiểm phiếu:

1. Trưởng Ban kiểm phiếu: PGS.TS. Phạm Văn Uông
2. Ủy viên Ban kiểm phiếu: TS. Đinh Trường Dục
3. Ủy viên Ban kiểm phiếu: PGS.TS. Nguyễn Long Giang

- Hội đồng chấm đề án tốt nghiệp bằng bỏ phiếu kín.

- Ban kiểm phiếu làm việc:

- Trưởng Ban kiểm phiếu báo cáo kết quả kiểm phiếu (có Biên bản họp Ban kiểm phiếu kèm theo)

- Điểm trung bình của đề án tốt nghiệp: 6,4

Kết luận:

1. Các nội dung cần chỉnh sửa, hoàn thiện sau bảo vệ đề án tốt nghiệp:

- Cần chỉnh sửa lại theo nhận xét của phản biện
- làm rõ mục tiêu của đề tài, làm rõ phạm vi nghiên cứu
- Mô tả chi tiết về hình thức, phạm vi, các bước nghiên cứu, đồ số
- làm rõ phạm vi kết hợp về hình thức và nội dung
- theo dõi đang triển khai
- Viết lại phần tích hợp kỹ thuật
- Chính xác các tài liệu của đề án

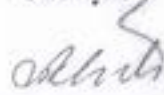
2. Đề nghị Học viện công nhận (hoặc không) và cấp bằng (hoặc không) thạc sĩ cho học viên:

Vương Việt Hùng

3. Đề án tốt nghiệp có thể phát triển thành đề tài nghiên cứu cho NCS.....

Buổi làm việc kết thúc vào 9.4.55... cùng ngày.

Chủ tịch



PGS.TS. Lê Hữu Lập

Thư ký



TS. Đinh Trường Dục

NHẬN XÉT CỦA NGƯỜI PHẢN BIỆN

Tên đề tài: Phát hiện và nhận diện hành vi qua video ứng dụng trong giám sát phòng thực hành

Chuyên ngành: Hệ thống Thông tin

Mã chuyên ngành: 8.48.01.04

Tên học viên: Vương Việt Hùng

Họ và tên người nhận xét: Phạm Văn Cường.....

Học hàm, học vị: PGS.TS.....

Chuyên ngành: Khoa học Máy tính

Nơi công tác: Học viện Công nghệ Bưu chính Viễn thông

NỘI DUNG NHẬN XÉT

I/ Cơ sở khoa học và thực tiễn, tính cấp thiết của đề tài:

Có ý nghĩa thực tế trong các ứng dụng giám sát thực hành, thi cử, bảo đảm chất lượng đào tạo.

II/ Về nội dung, chất lượng của luận văn, kết quả đạt được:

- Luận văn được chia thành 4 chương với 69 trang. Các chương được cấu trúc phù hợp với đề cương; Chương 1 trình bày tổng quan về bài toán xác định hành vi; Chương 2 trình bày về mô hình YOLO và một số độ đo đánh giá; chương 3 trình bày về phân tích và thiết kế hệ thống; chương 4 trình bày về cài đặt và thử nghiệm hệ thống.

III/ Những điểm chưa nhất trí:

Luận văn được trình bày còn lan man mà chưa làm nổi bật trọng tâm. Chương 1 & 2 nên gộp lại với nhau và nên bổ sung các nghiên cứu liên quan về nhận diện hành vi. Chương 3 cần bổ sung thiết kế kiến trúc và thiết kế giao diện của hệ thống. Chương 4 phân tích các kết quả thử nghiệm còn sơ sài, chưa đánh giá được mô hình nhận diện hành vi. Nhìn chung các kết quả còn nhiều hạn chế. có nhiều lỗi chính tả, gõ máy;

Câu hỏi:

1. Em trình bày lại quá trình tái huấn luyện và tinh chỉnh các tham số cho mô hình YOLO?

III/ Kết luận:

Đồng ý cho phép học viên được bảo vệ luận văn tốt nghiệp.

Hà Nội, ngày 15 tháng 7 năm 2025

Người nhận xét



Phạm Văn Cường

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

BẢN NHẬN XÉT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ
(Dùng cho người phản biện)

Tên đề án: Phát hiện và nhận diện hành vi qua Video ứng dụng trong giám sát phòng thực hành

Chuyên ngành: Hệ thống thông tin

Mã số: 8.48.01.04

Họ và tên học viên: Vương Việt Hùng

Họ và tên người nhận xét: PGS.TS. Nguyễn Long Giang

Cơ quan công tác: Viện Công nghệ thông tin, Viện Hàn Lâm KH&CN Việt Nam

Số điện thoại: 0904739189

Email: nlgiang@ioit.ac.vn

NỘI DUNG NHẬN XÉT

1. Cơ sở khoa học và thực tiễn, sự cần thiết lựa chọn đề tài

Đề án giải quyết bài toán phát hiện và nhận diện hành vi từ dữ liệu video bằng mô hình học sâu YOLOv5, trên cơ sở đó ứng dụng trong giám sát phòng thực hành. Chủ đề nghiên cứu của đề án có ý nghĩa khoa học và thực tiễn.

2. Nội dung của đề án, các kết quả đã đạt được

Đề án đã đạt được kết quả chính như sau:

- 1) Ứng dụng mô hình YOLOv5 phát hiện đối tượng (khuôn mặt) và nhận dạng hành vi từ dữ liệu VIDEO
- 2) Xây dựng, thử nghiệm, đánh giá hệ thống quản lý, giám sát phòng thực hành có ứng dụng mô hình phát hiện, nhận dạng hành vi

Tuy nhiên, các đóng góp và các kết quả của đề án chưa rõ ràng cả về Phương pháp giải quyết bài toán phát hiện đối tượng, nhận dạng hành vi bằng YOLO cũng như hệ thống quản lý phòng thực hành ứng dụng mô hình được phát triển. Chi tiết được mô tả ở phần sau.

3. Những vấn đề học viên cần chỉnh sửa, bổ sung

- Rà soát, chỉnh sửa đề án về trình bày. Nhìn chung trình bày các đề mục khá sơ sài, chưa có trích dẫn tài liệu tham khảo. Các hình vẽ chụp hình ảnh quá to và bị vỡ hình (Hình 0-1, 0-2, 0-3...)

- Cấu trúc lại phần trình bày chương mục cho hợp lý, coi phần mô hình là quan trọng và phần xây dựng hệ thống là mục nhỏ

- Phần trình bày mô hình giải quyết bài toán ứng dụng YOLO, cần viết lại cho rõ ràng để làm rõ: bài toán cần giải quyết (phát hiện khuôn mặt/nhận dạng hành vi gì? Hay phát hiện bình thường/bất thường), với mỗi bài toán làm rõ đầu vào/đầu ra, đặc tả rõ ràng dữ liệu và bộ dữ liệu được sử dụng (thu thập hay bộ dữ liệu có sẵn), quy trình gắn nhãn, làm rõ mô hình với mỗi bài toán, đặc tả các tham số thử nghiệm

- Phần mô tả hệ thống cần chỉnh sửa lại: mô tả cô đọng các chức năng/không để phụ lục, cần đặc tả rõ chức năng nào phát hiện/nhận diện hành vi ?

4. Kết luận

- Đề nghị học viên chỉnh sửa, bổ sung luận văn theo góp ý.

Câu hỏi: 1) Làm rõ các bài toán cần giải quyết và bộ dữ liệu sử dụng cho mỗi bài toán ?

Hà Nội, ngày 14 tháng 07 năm 2025

XÁC NHẬN CỦA CƠ QUAN

NGƯỜI NHẬN XÉT



Nguyễn Long Giang

