

BỘ CÔNG THƯƠNG
TRƯỜNG ĐẠI HỌC CÔNG THƯƠNG TP. HỒ CHÍ MINH
KHOA CÔNG NGHỆ THÔNG TIN

□□□□



KHÓA LUẬN CỬ NHÂN

**XÂY DỰNG HỆ THỐNG GIÁM SÁT VÀ PHÂN
TÍCH CẢM XÚC TRONG GIÁO DỤC THÔNG
MINH TỪ ẢNH KHUÔN MẶT SỬ DỤNG
DEEP LEARNING**

Ngành: Công Nghệ Thông Tin

SINH VIÊN THỰC HIỆN:

1. 2001224174 – Võ Huỳnh Sơn – 13DHTH08
2. 2001224949 – Lâm Đức Thịnh – 13DHTH08
3. 2001224442 – Phạm Văn Tisen – 13DHTH06

TP. HỒ CHÍ MINH, tháng 12 năm 2025

BỘ CÔNG THƯƠNG
TRƯỜNG ĐẠI HỌC CÔNG THƯƠNG TP. HỒ CHÍ MINH
KHOA CÔNG NGHỆ THÔNG TIN

□□□□



KHÓA LUẬN CỬ NHÂN

**XÂY DỰNG HỆ THỐNG GIÁM SÁT VÀ
PHÂN TÍCH CẢM XÚC TRONG GIÁO
DỤC THÔNG MINH TỪ ẢNH KHUÔN
MẶT SỬ DỤNG DEEP LEARNING**

Ngành: Công Nghệ Thông Tin

GIẢNG VIÊN HƯỚNG DẪN:

ThS. Trần Văn Thọ

SINH VIÊN THỰC HIỆN:

1. 2001224174 – Võ Huỳnh Sơn – 13DHTH08
2. 2001224949 – Lâm Đức Thịnh – 13DHTH08
3. 2001224442 – Phạm Văn Tisen – 13DHTH06

TP. HỒ CHÍ MINH, tháng 12 năm 2025

LỜI CẢM ƠN

Lời đầu tiên, nhóm sinh viên chúng em xin trân trọng gửi lời cảm ơn chân thành và sâu sắc nhất đến Ban Giám hiệu Trường Đại học Công Thương TP.HCM, quý thầy cô Khoa Công nghệ Thông tin đã tạo mọi điều kiện thuận lợi, truyền đạt những kiến thức quý báu và là nền tảng vững chắc cho chúng em trong suốt quá trình học tập và nghiên cứu tại trường.

Đặc biệt, chúng em xin bày tỏ lòng biết ơn sâu sắc nhất đến thầy ThS. Trần Văn Thọ. Trong suốt quá trình thực hiện đề tài “Xây dựng hệ thống giám sát và phân tích cảm xúc trong giáo dục thông minh từ ảnh khuôn mặt sử dụng Deep Learning”, thầy đã luôn tận tình hướng dẫn, chỉ bảo và động viên chúng em. Những kiến thức chuyên môn, kinh nghiệm thực tiễn và những góp ý quý báu của thầy là hành trang vô cùng quan trọng giúp chúng em định hướng đúng đắn, vượt qua mọi khó khăn để hoàn thành khóa luận tốt nghiệp này.

Chúng em cũng xin gửi lời cảm ơn đến các thành viên trong nhóm đã cùng nhau đoàn kết, hợp tác, chia sẻ và hỗ trợ lẫn nhau để hoàn thành tốt nhiệm vụ chung.

Mặc dù đã rất nỗ lực, nhưng do kiến thức và kinh nghiệm còn hạn chế, khóa luận chắc chắn không thể tránh khỏi những thiếu sót. Chúng em rất mong nhận được những ý kiến đóng góp quý báu của quý thầy cô và các bạn để đề tài được hoàn thiện hơn.

Xin trân trọng cảm ơn,

NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

GIẢNG VIÊN HƯỚNG DẪN

(Ký và ghi rõ họ tên)

MỤC LỤC

LỜI CẢM ƠN	i
NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN	ii
MỤC LỤC.....	iii
DANH MỤC CÁC KÝ HIỆU VÀ CHỮ VIẾT TẮT	iii
DANH MỤC HÌNH	ix
DANH MỤC BẢNG.....	xi
MỞ ĐẦU.....	1
CHƯƠNG 1: KHẢO SÁT HỆ THỐNG	4
1.1. GIỚI THIỆU TỔNG QUAN VỀ ĐỀ TÀI.....	4
1.2. KHẢO SÁT HIỆN TRẠNG HỆ THỐNG	6
1.2.1. Giới thiệu thông tin khảo sát.....	6
1.2.2. Cơ cấu tổ chức.....	6
1.2.3. Các quy trình nghiệp vụ	7
1.2.3.1. Tổng quan các Use Case nghiệp vụ	7
1.2.3.2. Mô tả chi tiết các Use Case nghiệp vụ	8
1.3. MỤC TIÊU VÀ YÊU CẦU ĐỀ TÀI	25
1.3.1. Các yêu cầu chức năng.....	26
1.3.2. Các yêu cầu phi chức năng.....	27
1.4. CÁC CÔNG CỤ, CÔNG NGHỆ	28
CHƯƠNG 2: CƠ SỞ LÝ THUYẾT VÀ CÁC CÔNG TRÌNH LIÊN QUAN	29
2.1. MẠNG NƠ-RON TÍCH CHẬP (CNN).....	29
2.1.1. Giới thiệu.....	29
2.1.2. Kiến trúc tổng quát.....	29
2.1.3. Cơ chế hoạt động.....	31

2.1.4. CNN trong nhận diện cảm xúc.....	31
2.2. MÔ HÌNH YOLO	32
2.2.1. Phát hiện đối tượng (Object detection)	32
2.2.2. Tổng quan YOLO.....	33
2.2.3. Nguyên lý hoạt động	33
2.2.3.1. Khối dư (Residual blocks)[15,16].....	34
2.2.3.2. Hồi quy hộp giới hạn (Bounding box regression)[16,17]	35
2.2.3.3. Intersection Over Union (IOU)[16,19]	36
2.2.4. Kiến trúc hiện đại	38
2.2.5. Các phiên bản YOLO	39
2.2.6. YOLOv11	40
2.2.6.1. Tổng quan.....	40
2.2.6.2. Các đổi mới kiến trúc	41
2.2.6.3. Các tác vụ Thị giác máy tính được YOLOv11 hỗ trợ[22,24]	42
2.2.6.4. So sánh YOLOv11 so với các phiên bản cũ[5,22]	43
2.3. BÀI TOÁN NHẬN DẠNG CẢM XÚC KHUÔN MẶT (FACIAL EMOTION RECOGNITION).....	43
2.3.1. Tổng quan bài toán nhận diện cảm xúc.....	43
2.3.4. Quy trình nhận diện cảm xúc	44
CHƯƠNG 3: PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG	46
3.1. PHÂN TÍCH HỆ THỐNG	46
3.1.1. Sơ đồ Use Case nghiệp vụ	46
3.1.2. Sơ đồ Use Case hệ thống và đặc tả: Chi tiết hóa các use case cho Web App và Mobile App.....	47
3.1.2.1. Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ Giám sát Cảm xúc Lớp học	47

3.1.2.2. Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ phân tích chất lượng giảng dạy	49
3.1.2.3. Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ Hỗ trợ Cá nhân hóa Học tập.....	50
3.1.3. Sơ đồ lớp	52
3.2. THIẾT KẾ HỆ THỐNG.....	56
3.2.1. Thiết kế kiến trúc	56
3.2.2. Thiết kế giao diện: Thiết kế các giao diện chính cho Web App.....	57
3.2.2.1. Giao diện đăng nhập (Login Interface).....	57
3.2.2.2. Giao diện Trang chủ (Home/Dashboard Interface).....	58
3.2.2.3. Giao diện Giám sát (Monitoring Interface)	59
3.2.2.4. Giao diện Phân tích Lớp học (Class Analytics Interface).....	60
CHƯƠNG 4: CÀI ĐẶT HỆ THỐNG	62
4.1. CÁC YÊU CẦU CÀI ĐẶT ỨNG DỤNG	62
4.1.1. Cài đặt môi trường và AI PIPELINE	62
4.1.2. XÂY DỰNG API HỆ THỐNG VÀ CÀI ĐẶT CÁC NGHIỆP VỤ CHÍNH	65
4.1.2.1. Cài đặt Nghiệp vụ 1 - Giám sát cảm xúc: Phát triển dashboard, chức năng cảnh báo, xuất báo cáo.....	65
4.1.2.2. Cài đặt Nghiệp vụ 2 - Phân tích chất lượng giảng dạy: Xây dựng các chức năng tổng hợp, so sánh và báo cáo cho nhà trường	66
4.2. GIỚI THIỆU GIAO DIỆN CHƯƠNG TRÌNH	68
4.2.1. Màn hình Đăng nhập (Login Screen).....	68
4.2.2. Màn hình Trang chủ (Home Dashboard)	69
4.2.3. Màn hình Giám sát (Monitoring Interface).....	70
4.2.4. Màn hình Phân tích Lớp học (Class Analytics)	71

4.2.5. Màn hình Báo cáo (Report List)	73
CHƯƠNG 5: THỬ NGHIỆM VÀ ĐÁNH GIÁ	75
5.1. ĐÁNH GIÁ HIỆU NĂNG MÔ HÌNH AI.....	75
5.2. KIỂM THỬ HỆ THỐNG	83
KẾT LUẬN	90
TÀI LIỆU THAM KHẢO.....	95

DANH MỤC CÁC KÝ HIỆU VÀ CHỮ VIẾT TẮT

Viết tắt	Tiếng Anh	Tiếng Việt
AI	Artificial Intelligence	Trí tuệ nhân tạo
API	Application Programming Interface	Giao diện lập trình ứng dụng
CNN	Convolutional Neural Network	Mạng Nơ-ron Tích chập
CSDL	-	Cơ sở dữ liệu
DBMS	Database Management System	Hệ quản trị cơ sở dữ liệu
EQ	Emotional Quotient	Trí tuệ cảm xúc
FER	Facial Emotion Recognition	Nhận diện cảm xúc từ khuôn mặt
ID	Identifier	Mã định danh
IOU	Intersection Over Union	Giao trên hợp
IQ	Intelligence Quotient	Chỉ số thông minh
IT	Information Technology	Công nghệ thông tin
NMS	Non-Maximum Suppression	Loại bỏ cực đại không phải tối đa
PDF	Portable Document Format	Định dạng tài liệu di động

ReLU	Rectified Linear Unit	Hàm kích hoạt tuyến tính chỉnh lưu
UI/UX	User Interface/User Experience	Giao diện người dùng/Trải nghiệm người dùng
YOLO	You Only Look Once	Bạn chỉ nhìn một lần

DANH MỤC HÌNH

Hình 2. 1: Kiến trúc tổng quát CNN.	30
Hình 2. 2: Quá trình học đặc trưng và phân loại.....	31
Hình 2. 3: Phân biệt bài toán Image Classification - Object Localization – Object Detection	33
Hình 2. 4: Hình ảnh đầu vào được chia thành các lưới, có nhiều ô lưới có kích thước bằng nhau.	34
Hình 2. 5: Minh họa bounding box.	35
Hình 2. 6: Công thức tính IOU.....	36
Hình 2. 7: Hai bounding box xanh lam, xan lá.	37
Hình 2. 8: Mô hình YOLO thực hiện xử lý hình ảnh theo tuần tự từng bước.	38
Hình 2. 9: Sơ đồ kiến trúc mạng YOLO.	39
Hình 2. 10: Các phiên bản YOLO.....	40
Hình 2. 11: Cấu trúc của Mô hình YOLOv11.....	41
Hình 2. 12: Kiến trúc cải tiến của YOLOv11.....	42
Hình 2. 13: Quy trình bài toán nhận diện cảm xúc.	44
Hình 3. 1: Sơ đồ usecase nghiệp vụ.	46
Hình 3. 2: Sơ đồ Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ Giám sát Cảm xúc Lớp học.	47
Hình 3. 3: Sơ đồ Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ phân tích chất lượng giảng dạy.	49
Hình 3. 4: Sơ đồ Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ Hỗ trợ Cá nhân hóa Học tập.....	50
Hình 3. 5: Sơ đồ lớp mức thiết kế.	52
Hình 3. 6: Sơ đồ lớp thiết kế và mô hình dữ liệu quan hệ cho PostgreSQL.....	53
Hình 3. 7: Sơ đồ kiến trúc hệ thống Giám sát và Phân tích Cảm xúc trong Giáo dục Thông minh.	56
Hình 4. 1: Giao diện màn hình đăng nhập	68
Hình 4. 2: Giao diện trang chủ.....	69
Hình 4. 3: Giao diện giám sát cảm xúc	70

Hình 4. 4: Giao diện phân tích lớp học	71
Hình 4. 5: Giao diện danh sách báo cáo.....	73
Hình 5. 1: Ma trận nhầm lẫn chuẩn hóa (Confusion Matrix Normalized).....	75
Hình 5. 2: Phân bố nhãn và thống kê Bounding Box.....	76
Hình 5. 3: Biểu đồ F1–Confidence Curve.....	77
Hình 5. 4: Biểu đồ Precision–Confidence Curve.....	78
Hình 5. 5: Biểu đồ Precision–Recall (PR Curve).....	79
Hình 5. 6: Biểu đồ Precision–Recall (PR Curve).....	80
Hình 5. 7: Biểu đồ Recall–Confidence Curve.....	81
Hình 5. 8: Biểu đồ Training Curves (Loss & Metrics).....	81

DANH MỤC BẢNG

Bảng 1. 1: Đặc tả Use Case UC-NV01: Giám sát cảm xúc lớp học.	8
Bảng 1. 2: Mô tả Luồng sự kiện chính của UC-NV01.	9
Bảng 1. 3: Đặc tả Use Case UC-NV02: Phân tích chất lượng giảng dạy.	13
Bảng 1. 4: Mô tả Luồng sự kiện chính của UC-NV02.	14
Bảng 1. 5: Đặc tả Use Case UC-NV03: Hỗ trợ cá nhân hóa học tập.....	19
Bảng 1. 6: Mô tả Luồng sự kiện chính của UC-NV03.	20
Bảng 3. 1: Đặc tả Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ Giám sát Cảm xúc Lớp học.	48
Bảng 3. 2: Đặc tả Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ phân tích chất lượng giảng dạy.	49
Bảng 3. 3: Đặc tả Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ Hỗ trợ Cá nhân hóa Học tập.....	51
Bảng 3. 4: Mô tả các class có trong mô hình dữ liệu quan hệ cho PostgreSQL..	53
Bảng 4. 1: Bảng phân phối dữ liệu.....	63
Bảng 5. 1: Kiểm thử Chức năng Upload Video	84
Bảng 5. 2: Kiểm thử Nhận diện Khuôn mặt	84
Bảng 5. 3: Kiểm thử Phân loại Cảm xúc 7 Lớp	85
Bảng 5. 4: Kiểm thử Dashboard Thời gian Thực.....	86
Bảng 5. 5: Kiểm thử Chức năng Cảnh báo	86
Bảng 5. 6: Kiểm thử Xuất Báo cáo Buổi học.....	87
Bảng 5. 7: Kiểm thử Phân tích Lớp học (Class Analytics)	88

MỞ ĐẦU

1. GIỚI THIỆU

- Trong bối cảnh giáo dục hiện đại ngày càng đề cao vai trò của trí tuệ cảm xúc (EQ) như một yếu tố nền tảng ảnh hưởng trực tiếp đến hiệu quả học tập, việc thấu hiểu tâm lý học sinh trở nên vô cùng cấp thiết. Tuy nhiên, các phương pháp giám sát truyền thống như quan sát trực quan của giáo viên hiện nay bộc lộ nhiều hạn chế lớn: vừa mang tính chủ quan, vừa không thể bao quát toàn bộ lớp học, dẫn đến việc can thiệp và hỗ trợ thường bị chậm trễ.
- Để khắc phục những nhược điểm này, việc ứng dụng Trí tuệ nhân tạo (AI) và Deep Learning mở ra một hướng tiếp cận đột phá. Chính vì vậy, đề tài "Xây dựng hệ thống giám sát và phân tích cảm xúc trong giáo dục thông minh từ ảnh khuôn mặt sử dụng Deep Learning" được thực hiện. Hệ thống hướng đến mục tiêu cung cấp một công cụ giám sát khách quan, liên tục và trên quy mô lớn, giúp giáo viên nắm bắt tức thời không khí lớp học để điều chỉnh phương pháp giảng dạy. Đồng thời, hệ thống còn tạo ra các báo cáo chi tiết cho nhà trường và phụ huynh, hứa hẹn sẽ là một giải pháp công nghệ tiên phong, góp phần cá nhân hóa lộ trình học tập và nâng cao toàn diện chất lượng giáo dục.

2. MỤC TIÊU ĐỀ TÀI

- Mục tiêu bao quát của đề tài là xây dựng và triển khai một hệ thống hoàn chỉnh, ứng dụng công nghệ Deep Learning để giám sát, phân tích và báo cáo trạng thái cảm xúc của học sinh trong môi trường giáo dục. Hệ thống hướng đến việc cung cấp một công cụ hỗ trợ đắc lực cho ba đối tượng chính: giáo viên có thể điều chỉnh phương pháp giảng dạy tức thời, nhà trường có thể đánh giá và nâng cao chất lượng chương trình học, và phụ huynh có thể đồng hành sâu sắc hơn với quá trình học tập của con em mình.

3. ĐỐI TƯỢNG VÀ PHẠM VI ĐỀ TÀI

3.1. Đối tượng nghiên cứu:

- Đối tượng nghiên cứu trọng tâm của đề tài là quá trình ứng dụng các thuật toán học sâu (Deep Learning) vào bài toán nhận diện và phân loại cảm xúc từ hình ảnh khuôn mặt trong bối cảnh giáo dục. Cụ thể, đề tài tập trung sâu vào việc nghiên cứu, lựa chọn và tối ưu hóa các mô hình phù hợp cho hai tác vụ cốt lõi: phát hiện khuôn mặt (Face Detection) và phân loại trạng thái cảm xúc (Emotion Classification) của học sinh.

3.2. Phạm vi nghiên cứu:

- Để đảm bảo tính khả thi và tập trung vào các mục tiêu đã đề ra, đề tài giới hạn phạm vi nghiên cứu trong các khía cạnh sau:

Về lý thuyết và công nghệ:

- Đề tài tập trung chủ yếu vào việc nghiên cứu và triển khai các mô hình thuộc họ Mạng Nơ-ron Tích chập (Convolutional Neural Network - CNN). Đây là kiến trúc nền tảng và đã được chứng minh hiệu quả vượt trội trong các bài toán xử lý và phân tích hình ảnh.
- Về mặt phát triển sản phẩm, đề tài giới hạn ở việc xây dựng hai nền tảng chính: một ứng dụng Web và một ứng dụng Di động, hoạt động dựa trên một hệ thống API và cơ sở dữ liệu tập trung.

Về chức năng và nghiệp vụ ứng dụng:

- Nghiệp vụ Giám sát cảm xúc lớp học: Tập trung vào việc cung cấp cho giáo viên một công cụ theo dõi cảm xúc của lớp học trong thời gian thực, trực quan hóa dữ liệu qua dashboard và đưa ra cảnh báo tự động khi tỷ lệ cảm xúc tiêu cực (như chán nản) vượt ngưỡng cho phép.
- Nghiệp vụ Phân tích chất lượng giảng dạy: Khai thác dữ liệu cảm xúc được tổng hợp từ nhiều buổi học để tạo ra các báo cáo phân tích sâu hơn. Chức năng này cho phép so sánh hiệu quả giữa các bài giảng, các giảng viên hoặc các giai đoạn khác nhau của một tiết học, nhằm hỗ trợ nhà trường và tổ bộ môn trong việc cải tiến chương trình.

- Nghiệp vụ Hỗ trợ cá nhân hóa học tập: Xây dựng hồ sơ cảm xúc cá nhân cho từng học sinh, theo dõi lịch sử cảm xúc theo chu kỳ tuần/tháng. Nghiệp vụ này cung cấp những báo cáo và gợi ý chuyên biệt, giúp giáo viên và phụ huynh nắm bắt sâu hơn về tâm lý và mức độ hứng thú học tập của từng em.
- Về môi trường triển khai: Hệ thống được thiết kế để có thể áp dụng linh hoạt trong cả hai môi trường học tập phổ biến hiện nay: lớp học trực tiếp (offline) thông qua việc thu nhận hình ảnh từ camera giám sát và lớp học trực tuyến (online) qua webcam trên thiết bị của học sinh.

CHƯƠNG 1: KHẢO SÁT HỆ THỐNG

1.1. GIỚI THIỆU TỔNG QUAN VỀ ĐỀ TÀI

- Bối cảnh giáo dục Việt Nam đang chứng kiến một sự chuyển mình mạnh mẽ, không còn giới hạn trong các phương pháp sư phạm truyền thống mà đã tích cực hội nhập công nghệ hiện đại vào quá trình dạy và học. Song song với đó, triết lý giáo dục hiện đại cũng có sự dịch chuyển quan trọng: từ việc chỉ tập trung vào truyền thụ kiến thức học thuật (IQ) sang phát triển con người toàn diện, trong đó trí tuệ cảm xúc (EQ) được xem là yếu tố nền tảng. Thực tế đã chứng minh, cảm xúc là nhân tố then chốt ảnh hưởng trực tiếp đến hiệu quả học tập. Các trạng thái cảm xúc tích cực như vui vẻ, hứng thú và tò mò chính là chất xúc tác mạnh mẽ giúp học sinh tăng cường sự tập trung, cải thiện khả năng ghi nhớ và thúc đẩy động lực học tập. Ngược lại, những cảm xúc tiêu cực như căng thẳng, chán nản hay lo âu lại trở thành rào cản tâm lý, ức chế quá trình tiếp thu và làm suy giảm đáng kể kết quả học tập[1].
- Tuy nhiên, việc nắm bắt và thấu hiểu cảm xúc của học sinh hiện nay vẫn còn bộc lộ nhiều hạn chế. Phương pháp quan sát trực quan của giáo viên, dù cần thiết, nhưng lại mang nặng tính chủ quan và không thể bao quát toàn bộ sĩ số lớp học, đặc biệt trong các môi trường học tập có quy mô lớn hoặc các lớp học trực tuyến. Hơn nữa, phương pháp này khó có thể cung cấp những phản hồi tức thời và bỏ lọt các biểu hiện cảm xúc tinh vi, khiến việc can thiệp và hỗ trợ của giáo viên thường bị chậm trễ. Đồng thời, nhu cầu của phụ huynh trong việc đồng hành và theo dõi sát sao quá trình học tập của con em mình cũng là một bài toán chưa có lời giải đáp hiệu quả[2,3].
- Trước những thách thức đó, việc ứng dụng Trí tuệ nhân tạo (AI) và Deep Learning mở ra một hướng tiếp cận đột phá. Sự phát triển vượt bậc của công nghệ nhận dạng hình ảnh và phân tích cảm xúc cho phép xây dựng một hệ thống có khả năng giám sát và phân tích trạng thái cảm xúc của học sinh một cách khách quan, liên tục và trên quy mô lớn thông qua hệ thống camera. Hệ thống này không chỉ cung cấp cho giáo viên dữ liệu trực quan

theo thời gian thực để kịp thời điều chỉnh phương pháp giảng dạy, mà còn tạo ra các báo cáo chi tiết cho nhà trường và phụ huynh. Do đó, việc xây dựng "Hệ thống giám sát và phân tích cảm xúc trong giáo dục thông minh từ ảnh khuôn mặt sử dụng Deep Learning" không chỉ là một giải pháp công nghệ cấp thiết nhằm khắc phục những nhược điểm cố hữu của phương pháp truyền thống, mà còn hứa hẹn tạo ra một cuộc cách mạng trong việc cá nhân hóa lộ trình học tập và nâng cao toàn diện chất lượng giáo dục[5,6].

1.2. KHẢO SÁT HIỆN TRẠNG HỆ THỐNG

1.2.1. Giới thiệu thông tin khảo sát

- Để có cơ sở thực tiễn cho việc xây dựng hệ thống, nhóm nghiên cứu đã tiến hành khảo sát tại Trường Đại học Công Thương TP. Hồ Chí Minh, đặc biệt là Khoa Công nghệ Thông tin. Quá trình khảo sát tập trung vào việc tìm hiểu hiện trạng công tác giám sát và quản lý lớp học, cơ sở vật chất hiện có, cũng như nhu cầu ứng dụng công nghệ trong giáo dục của các bên liên quan.
- Kết quả khảo sát cho thấy phương pháp giám sát và đánh giá cảm xúc, mức độ tập trung của sinh viên hiện nay chủ yếu dựa vào quan sát trực tiếp của giảng viên, chưa có công cụ hỗ trợ khách quan và hệ thống. Điều này dẫn đến nhiều hạn chế trong việc nắm bắt toàn diện tình trạng lớp học, đặc biệt với các lớp đông sinh viên hoặc lớp học trực tuyến. Đồng thời, nhu cầu về một công cụ công nghệ hỗ trợ giám sát, phân tích và báo cáo được thể hiện rõ ràng từ cả giảng viên, ban lãnh đạo nhà trường và phụ huynh sinh viên.

1.2.2. Cơ cấu tổ chức

- Trường Đại học Công Thương TP. Hồ Chí Minh có cơ cấu tổ chức theo mô hình phân cấp, bao gồm các cấp quản lý từ Ban Giám hiệu, các phòng ban chức năng (Phòng Đào tạo, Phòng Công nghệ Thông tin), đến các đơn vị cơ sở là các Khoa/Bộ môn. Mỗi khoa được tổ chức thành các tổ/bộ môn chuyên ngành, do giảng viên trực tiếp đảm nhiệm giảng dạy và quản lý sinh viên.
- Trong phạm vi triển khai hệ thống, các đối tượng người dùng chính bao gồm:
 - + Giảng viên: Sử dụng hệ thống để giám sát lớp học trong thời gian thực.
 - + Ban lãnh đạo/Tổ bộ môn: Tiếp cận các báo cáo phân tích chất lượng giảng dạy.
 - + Phụ huynh và sinh viên: Theo dõi báo cáo cá nhân hóa về quá trình học tập.

- Cơ cấu tổ chức hiện tại của trường tạo điều kiện thuận lợi cho việc triển khai hệ thống theo từng cấp độ, từ việc thử nghiệm tại một số lớp học, mở rộng ra toàn bộ môn, đến triển khai rộng rãi cho toàn khoa và toàn trường.

1.2.3. Các quy trình nghiệp vụ

1.2.3.1. Tổng quan các Use Case nghiệp vụ

- Hệ thống được xây dựng xoay quanh ba nghiệp vụ chính, phục vụ cho các đối tượng người dùng khác nhau là Giáo viên, Nhà trường (Ban giám hiệu, Tổ bộ môn) và Phụ huynh. Các nghiệp vụ này có mối quan hệ chặt chẽ và theo một trình tự logic: từ việc thu thập dữ liệu cảm xúc trong thời gian thực, đến việc tổng hợp phân tích và cuối cùng là cung cấp báo cáo cá nhân hóa.
- Trình tự thực hiện các use case nghiệp vụ như sau:
 - + UC-NV01: Giám sát cảm xúc lớp học: Đây là nghiệp vụ cốt lõi và diễn ra đầu tiên, tập trung vào việc thu thập và hiển thị dữ liệu cảm xúc của học sinh trong một buổi học cụ thể. Nghiệp vụ này là tiền đề cung cấp dữ liệu đầu vào cho các nghiệp vụ phân tích sâu hơn.
 - + UC-NV02: Phân tích chất lượng giảng dạy: Nghiệp vụ này được thực hiện sau khi có đủ dữ liệu từ nhiều buổi học được ghi nhận bởi UC-NV01. Nó tổng hợp dữ liệu theo môn học, giáo viên hoặc lớp học để đưa ra những phân tích mang tính tổng quan, hỗ trợ nhà trường đánh giá và cải tiến chương trình.
 - + UC-NV03: Hỗ trợ cá nhân hóa học tập: Tương tự UC-NV02, nghiệp vụ này cũng cần dữ liệu được thu thập từ nhiều buổi học. Tuy nhiên, nó tập trung vào việc phân tích dữ liệu theo từng học sinh, xây dựng hồ sơ cảm xúc cá nhân và cung cấp báo cáo chi tiết cho phụ huynh và giáo viên chủ nhiệm.

1.2.3.2. Mô tả chi tiết các Use Case nghiệp vụ

– Use Case 1: Giám sát cảm xúc lớp học (UC-NV01)

Bảng 1. 1: Đặc tả Use Case UC-NV01: Giám sát cảm xúc lớp học.

Thuộc tính	Nội dung
Mã UC	-UC-NV01
Tên UC	-Giám sát cảm xúc lớp học
Mô tả	-Hệ thống thu nhận hình ảnh từ camera, phân tích cảm xúc học sinh theo thời gian thực và hiển thị kết quả trên dashboard để giáo viên có thể điều chỉnh phương pháp giảng dạy kịp thời.
Tác nhân chính	-Giáo viên
Tác nhân phụ	-Hệ thống, Module AI
Điều kiện tiên quyết	-Giáo viên đã đăng nhập vào hệ thống -Lớp học và danh sách sinh viên đã được thiết lập -Camera đã được kết nối và hoạt động
Điều kiện sau	-Dữ liệu cảm xúc của buổi học được lưu trữ -Báo cáo buổi học có thể được xuất

Bảng 1. 2: Mô tả Luồng sự kiện chính của UC-NV01.

STT	Hoạt động	Người thực hiện	Tương tác dữ liệu
1	-Đăng nhập và chọn lớp học: Giáo viên đăng nhập vào hệ thống, chọn lớp học và bắt đầu phiên giám sát.	Giáo viên	-Đọc: Dữ liệu lớp học, danh sách học sinh, lịch học.
2	-Kích hoạt camera: Hệ thống tự động kích hoạt camera (lớp học trực tiếp hoặc webcam trực tuyến) và bắt đầu thu nhận hình ảnh.	Hệ thống	-Kiểm tra: Trạng thái kết nối camera.
3	-Thu nhận và tiền xử lý hình ảnh: Hệ thống thu nhận hình ảnh theo chu kỳ (mỗi 30 giây), thực hiện tiền xử lý (cân bằng sáng, resize, chuẩn hóa).	Hệ thống	-Lưu tạm: Hình ảnh trong bộ nhớ đệm để xử lý.
4	-Phát hiện khuôn mặt và phân loại cảm xúc: Module AI thực hiện nhận diện khuôn mặt từng học sinh và phân loại cảm xúc (vui, buồn, tập trung, chán nản, ngạc nhiên, trung tính, giận dữ).	Hệ thống (Module AI)	-Đọc: Hình ảnh từ bộ nhớ đệm. -Lưu: Kết quả phân tích (ID học sinh, cảm xúc, độ tin cậy, timestamp) vào CSDL.
5	-Hiển thị Dashboard thời gian thực: Kết quả phân tích được cập nhật lên dashboard, hiển thị: - Biểu đồ tỷ lệ cảm xúc tổng thể	Hệ thống	-Đọc: Dữ liệu cảm xúc mới nhất từ CSDL.

	<ul style="list-style-type: none"> - Danh sách trạng thái cảm xúc từng học sinh - Biểu đồ xu hướng theo thời gian 		
6	-Theo dõi và tương tác: Giáo viên quan sát dashboard, có thể click vào từng học sinh để xem chi tiết hoặc lịch sử cảm xúc.	Giáo viên	-Đọc: Thông tin chi tiết từ dashboard.
7	-Kiểm tra ngưỡng cảnh báo: Hệ thống liên tục kiểm tra tỷ lệ cảm xúc tiêu cực.	Hệ thống	<ul style="list-style-type: none"> -Đọc: Dữ liệu cảm xúc trong 5 phút gần nhất. -So sánh: Với ngưỡng cài đặt (mặc định 40%).
8	-Gửi cảnh báo (nếu cần): Nếu tỷ lệ cảm xúc tiêu cực vượt ngưỡng, hệ thống hiển thị cảnh báo trên dashboard với âm thanh và thông báo popup.	Hệ thống	-Ghi log: Sự kiện cảnh báo vào CSDL.

Luồng sự kiện thay thế:

A1: Camera mất kết nối (tại bước 2)

- + 2a. Hệ thống phát hiện camera không phản hồi.
- + 2b. Hiển thị thông báo lỗi: "Không thể kết nối camera. Vui lòng kiểm tra kết nối."
- + 2c. Giáo viên kiểm tra và khắc phục sự cố.
- + 2d. Quay lại bước 2.

A2: Không nhận diện được khuôn mặt (tại bước 4)

- + 4a. Module AI không phát hiện khuôn mặt nào trong khung hình.
- + 4b. Hệ thống ghi log cảnh báo nhưng vẫn tiếp tục chu kỳ tiếp theo.
- + 4c. Dashboard hiển thị thông báo: "Đang chờ nhận diện..."

A3: Giáo viên quên kết thúc buổi học (tại bước 9)

- + 9a. Hệ thống tự động kết thúc phiên sau 4 tiếng hoặc khi hết giờ học theo lịch.
- + 9b. Gửi thông báo nhắc nhở cho giáo viên.
- + 9c. Chuyển sang bước 10.

Luồng ngoại lệ:

E1: Lỗi Module AI (tại bước 4)

- + 4e1. Module AI gặp lỗi nghiêm trọng (crash, timeout).
- + 4e2. Hệ thống ghi log lỗi chi tiết.
- + 4e3. Hiện thị thông báo: "Hệ thống phân tích tạm thời gặp sự cố. Đang khởi động lại..."
- + 4e4. Tự động khởi động lại module AI.
- + 4e5. Nếu lỗi lặp lại 3 lần, tạm dừng phiên và thông báo cho quản trị viên.

E2: Mất kết nối mạng (bất kỳ bước nào)

- + Hệ thống lưu dữ liệu local (offline mode).
- + Khi kết nối được khôi phục, tự động đồng bộ dữ liệu lên server.
- + Thông báo cho giáo viên về trạng thái offline.

– Use Case 2: Phân tích chất lượng giảng dạy (UC-NV02).

Bảng 1. 3: Đặc tả Use Case UC-NV02: Phân tích chất lượng giảng dạy.

Thuộc tính	Nội dung
Mã UC	-UC-NV02
Tên UC	-Phân tích chất lượng giảng dạy
Mô tả	-Hệ thống tổng hợp và phân tích dữ liệu cảm xúc từ nhiều buổi học để đưa ra báo cáo so sánh, đánh giá chất lượng giảng dạy theo môn học, giảng viên, lớp học hoặc giai đoạn trong bài giảng, hỗ trợ ban lãnh đạo và tổ bộ môn trong việc cải tiến chương trình.
Tác nhân chính	-Ban Giám hiệu / Tổ trưởng bộ môn
Tác nhân phụ	-Hệ thống
Điều kiện tiên quyết	<ul style="list-style-type: none"> - Người dùng đã đăng nhập với quyền quản lý - Có ít nhất 5 buổi học đã được ghi nhận dữ liệu cảm xúc - Dữ liệu cảm xúc đã được lưu trữ trong CSDL
Điều kiện sau	<ul style="list-style-type: none"> - Báo cáo phân tích được tạo và hiển thị - Báo cáo có thể được lưu trữ hoặc xuất file - Lịch sử truy cập báo cáo được ghi log

Bảng 1. 4: Mô tả Luồng sự kiện chính của UC-NV02.

STT	Hoạt động	Người thực hiện	Tương tác dữ liệu
1	-Đăng nhập và truy cập chức năng: Người dùng đăng nhập vào hệ thống với vai trò Ban Giám hiệu/Tổ trưởng và chọn menu "Phân tích chất lượng giảng dạy".thống và truy cập vào mục "Phân tích chất lượng giảng dạy".	Ban Giám hiệu / Tổ trưởng	-Đọc: Thông tin đăng nhập, phân quyền người dùng.
2	-Hiển thị giao diện lọc dữ liệu: Hệ thống hiển thị các bộ lọc với các tùy chọn: - Khung thời gian (tuần, tháng, học kỳ, năm học) - Khoa/Ngành - Môn học - Giảng viên - Lớp học - Khối lớp	Hệ thống	-Đọc: Danh sách các khoa, môn học, giảng viên, lớp học từ CSDL.
3	-Chọn tiêu chí lọc: Người dùng chọn các tiêu chí lọc theo nhu cầu phân tích (ví dụ: "Môn Lập trình Python, Học kỳ I năm 2024-2025").	Ban Giám hiệu / Tổ trưởng	-Ghi: Tiêu chí lọc được chọn (để lưu lịch sử truy vấn).
4	-Tổng hợp dữ liệu: Hệ thống truy vấn và tổng hợp dữ liệu cảm xúc từ tất cả các buổi học phù hợp với tiêu chí đã chọn.	Hệ thống	-Đọc: Dữ liệu cảm xúc Tính toán: Tỷ lệ trung bình các cảm xúc, phương sai, xu hướng theo thời gian.

5	<p>-Phân tích và so sánh: Hệ thống thực hiện các phân tích sau:</p> <ul style="list-style-type: none"> - So sánh tỷ lệ cảm xúc tích cực giữa các giảng viên dạy cùng môn - So sánh xu hướng cảm xúc giữa các giai đoạn (đầu học kỳ, giữa, cuối) - Phân tích mức độ hứng thú theo thời điểm trong buổi học (15 phút đầu, giữa, cuối) - Xác định các buổi học có hiệu quả cao/thấp 	Hệ thống	<p>-Đọc: Dữ liệu đã tổng hợp.</p> <p>-Áp dụng: Các thuật toán thống kê (trung bình, độ lệch chuẩn, correlation)..</p>
6	<p>Tạo báo cáo trực quan: Hệ thống tạo báo cáo với các thành phần:</p> <ul style="list-style-type: none"> - Biểu đồ cột so sánh tỷ lệ cảm xúc tích cực/tiêu cực - Biểu đồ đường thể hiện xu hướng theo thời gian - Biểu đồ tròn phân bố các loại cảm xúc - Bảng xếp hạng (ranking) giảng viên/môn học - Các insight tự động (ví dụ: "Tỷ lệ chán nản tăng 15% trong 3 tuần cuối học kỳ") 	Hệ thống	<p>-Đọc: Kết quả phân tích.</p> <p>-Render: Các biểu đồ và báo cáo dạng HTML/PDF.</p>
7	<p>-Hiển thị báo cáo: Báo cáo được hiển thị trên giao diện web với khả năng tương tác (zoom, filter, drill-down).</p>	Hệ thống	<p>-Đọc: Dữ liệu báo cáo đã được tạo.</p>

8	<p>Xem chi tiết và điều chỉnh: Người dùng có thể:</p> <ul style="list-style-type: none"> - Click vào từng điểm dữ liệu để xem chi tiết - Điều chỉnh bộ lọc để so sánh với nhóm khác - Thêm ghi chú (note) vào báo cáo 	Ban Giám hiệu / Tổ trưởng	<p>-Đọc: Dữ liệu chi tiết từ các buổi học cụ thể.</p> <p>-Lưu: Ghi chú của người dùng vào CSDL.</p>
9	<p>Xuất báo cáo: Người dùng chọn xuất báo cáo dưới dạng:</p> <ul style="list-style-type: none"> - PDF (định dạng in ấn) - Excel (dữ liệu thô để phân tích thêm) - PowerPoint (để trình bày trong họp) 	Ban Giám hiệu / Tổ trưởng	<p>-Đọc: Toàn bộ dữ liệu báo cáo.</p> <p>-Xuất: File theo định dạng được chọn.</p> <p>-Ghi log: Thông tin xuất báo cáo (người dùng, thời gian, định dạng).</p>
10	Lưu báo cáo (tùy chọn): Người dùng có thể lưu báo cáo vào thư mục cá nhân để xem lại sau.	Ban Giám hiệu / Tổ trưởng	-Lưu: Metadata của báo cáo (tiêu đề, tiêu chí lọc, timestamp)

Luồng sự kiện thay thế:

A1: Không có dữ liệu phù hợp (tại bước 4)

- + 4a. Hệ thống không tìm thấy dữ liệu cảm xúc nào phù hợp với tiêu chí lọc.
- + 4b. Hiển thị thông báo: "Không có dữ liệu để phân tích. Vui lòng chọn tiêu chí khác hoặc đảm bảo đã có buổi học được ghi nhận."
- + 4c. Người dùng quay lại bước 3 để điều chỉnh tiêu chí.

A2: Dữ liệu không đủ để phân tích so sánh (tại bước 5)

- + 5a. Hệ thống phát hiện chỉ có 1-2 buổi học (không đủ để phân tích xu hướng).
- + 5b. Hiển thị cảnh báo: "Dữ liệu chỉ đủ cho báo cáo cơ bản. Cần ít nhất 5 buổi học để phân tích xu hướng."
- + 5c. Hệ thống chỉ tạo báo cáo mô tả đơn giản (không có so sánh, xu hướng).
- + 5d. Tiếp tục bước 6 với báo cáo đơn giản hóa.

A3: Người dùng muốn so sánh nhiều nhóm (tại bước 3)

- + 3a. Người dùng chọn chế độ "So sánh nâng cao".
- + 3b. Hệ thống cho phép chọn nhiều bộ tiêu chí (ví dụ: So sánh 3 giảng viên khác nhau).
- + 3c. Ở bước 6, báo cáo sẽ hiển thị dạng so sánh song song (side-by-side comparison).

Luồng ngoại lệ:

E1: Lỗi trong quá trình tính toán thống kê (tại bước 4-5)

- + 5e1. Hệ thống gặp lỗi khi tính toán (dữ liệu bị null, division by zero, v.v.).
- + 5e2. Hệ thống ghi log lỗi chi tiết với stacktrace.
- + 5e3. Hiển thị thông báo: "Không thể tạo báo cáo do lỗi dữ liệu. Vui lòng liên hệ quản trị viên."
- + 5e4. Gửi email/notification cho admin với thông tin lỗi.
- + 5e5. Use case kết thúc.

E2: Quá trình xuất file bị lỗi (tại bước 9)

- + 9e1. Hệ thống gặp lỗi khi tạo file (disk full, permission denied, v.v.).
- + 9e2. Hiển thị thông báo: "Không thể xuất file. Vui lòng thử lại sau."
- + 9e3. Người dùng có thể chọn định dạng khác hoặc thử lại.

- + 9e4. Quay lại bước 9.

E3: Timeout khi xử lý dữ liệu lớn (tại bước 4-5)

- + 5e5. Nếu có quá nhiều dữ liệu (ví dụ: phân tích cả năm học), quá trình có thể bị timeout.
- + 5e6. Hệ thống hiển thị progress bar: "Đang xử lý dữ liệu... X%".
- + 5e7. Nếu vượt quá 2 phút, đề xuất: "Dữ liệu quá lớn. Bạn có muốn thu hẹp phạm vi hoặc nhận báo cáo qua email?"
- + 5e8. Nếu chọn email: Hệ thống xử lý background và gửi báo cáo khi hoàn tất.

– **Use Case 3: Hỗ trợ cá nhân hóa học tập (UC-NV03)**

Bảng 1. 5: Đặc tả Use Case UC-NV03: Hỗ trợ cá nhân hóa học tập.

Thuộc tính	Nội dung
Mã UC	-UC-NV03
Tên UC	-Hỗ trợ cá nhân hóa học tập
Mô tả	-Hệ thống xây dựng và duy trì hồ sơ cảm xúc cá nhân cho từng học sinh, theo dõi lịch sử cảm xúc theo chu kỳ, đưa ra gợi ý và báo cáo cho giáo viên chủ nhiệm và phụ huynh nhằm hỗ trợ cá nhân hóa quá trình học tập.
Tác nhân chính	-Giáo viên chủ nhiệm / Phụ huynh
Tác nhân phụ	-Hệ thống
Điều kiện tiên quyết	<ul style="list-style-type: none"> - Người dùng đã đăng nhập (giáo viên hoặc phụ huynh) - Học sinh đã có hồ sơ trong hệ thống - Có ít nhất 3 buổi học đã được ghi nhận cảm xúc của học sinh đó
Điều kiện sau	<ul style="list-style-type: none"> - Hồ sơ cảm xúc của học sinh được cập nhật - Báo cáo cá nhân được tạo và hiển thị - Thông báo (nếu có) được gửi đến phụ huynh - Lịch sử truy cập được ghi log

Bảng 1. 6: Mô tả Luồng sự kiện chính của UC-NV03.

STT	Hoạt động	Người thực hiện	Tương tác dữ liệu
1	Đăng nhập hệ thống: Người dùng đăng nhập vào hệ thống: - Giáo viên: Qua Web App - Phụ huynh: Qua Mobile App	Giáo viên / Phụ huynh	-Đọc: Thông tin đăng nhập, vai trò người dùng. -Xác thực: Token/Session.
2	Chọn học sinh: - Giáo viên: Chọn từ danh sách lớp chủ nhiệm - Phụ huynh: Tự động hiển thị con em của mình	Giáo viên / Phụ huynh	-Đọc: Danh sách học sinh thuộc lớp chủ nhiệm (giáo viên) hoặc con em (phụ huynh) từ bảng students.
3	Hiển thị hồ sơ tổng quan: Hệ thống hiển thị màn hình tổng quan bao gồm: - Thông tin cơ bản học sinh (ảnh, họ tên, lớp) - Chỉ số cảm xúc tổng thể (Emotion Score) - điểm số từ 1-10 - Trạng thái hiện tại (tích cực/ổn định/cần quan tâm) - Các badge/achievement (nếu có)	Hệ thống	-Đọc: Thông tin học sinh từ bảng students. -Tính toán: Emotion Score dựa trên dữ liệu 30 ngày gần nhất.
4	Hiển thị biểu đồ lịch sử cảm xúc: Hệ thống tạo và hiển thị: - Biểu đồ đường: Xu hướng cảm xúc theo thời gian (7 ngày, 30 ngày, 3 tháng) - Biểu đồ cột: Phân bố các loại cảm xúc - Heatmap: Cảm xúc theo giờ học trong ngày và ngày trong tuần - Timeline: Các sự kiện đặc biệt (cảm xúc bất thường)	Hệ thống	-Đọc: Dữ liệu cảm xúc của học sinh từ bảng emotion_records. -Lọc: Theo khoảng thời gian đã chọn. -Aggregate: Tính tỷ lệ, trung bình theo ngày/tuần.
5	Chọn chu kỳ xem: Người dùng có thể chuyển đổi giữa các chu kỳ:	Giáo viên /	-Đọc: Dữ liệu cảm xúc theo chu kỳ được chọn.

	<ul style="list-style-type: none"> - 7 ngày gần nhất - 30 ngày gần nhất - Theo tháng (chọn tháng cụ thể) - Theo học kỳ 	Phụ huynh	-Re-render: Các biểu đồ với dữ liệu mới.
6	<p>Phân tích và tạo insight: Hệ thống tự động phân tích dữ liệu và tạo các insight như:</p> <ul style="list-style-type: none"> - "Em A có xu hướng mất tập trung vào 20 phút cuối các buổi học Toán" - "Tỷ lệ cảm xúc tích cực tăng 25% trong 2 tuần gần đây" - "Em A thường có cảm xúc tiêu cực vào các buổi chiều thứ 5" - "So với bạn cùng lớp, em A có mức độ hứng thú cao hơn 15%" 	Hệ thống	<ul style="list-style-type: none"> -Đọc: Dữ liệu cảm xúc cá nhân và dữ liệu trung bình của lớp. -Áp dụng: Các rule-based analysis và machine learning pattern recognition. -Lưu: Các insight vào bảng student_insights.
7	<p>Hiển thị gợi ý và khuyến nghị: Dựa trên phân tích, hệ thống đưa ra các gợi ý:</p> <ul style="list-style-type: none"> - Cho giáo viên: "Nên bổ sung hoạt động tương tác trong 20 phút cuối tiết Toán" - Cho phụ huynh: "Nên quan tâm, trò chuyện với con sau các buổi học chiều thứ 5" - Gợi ý môn học phù hợp dựa trên mức độ hứng thú 	Hệ thống	<ul style="list-style-type: none"> -Đọc: Các insight đã tạo. -Match: Với thư viện các khuyến nghị (recommendation templates).
8	<p>Xem báo cáo chi tiết: Người dùng có thể click vào các mục để xem báo cáo chi tiết:</p> <ul style="list-style-type: none"> - Báo cáo theo môn học - Báo cáo so sánh với trung bình lớp - Báo cáo xu hướng theo thời gian - Danh sách các buổi học có cảm xúc bất thường 	Giáo viên / Phụ huynh	<ul style="list-style-type: none"> -Đọc: Dữ liệu chi tiết từ các bảng liên quan (emotion_records, sessions, subjects). -Join: Nhiều bảng để tạo báo cáo tổng hợp.

9	<p>Kiểm tra điều kiện gửi thông báo: Hệ thống kiểm tra xem có điều kiện nào sau đây thỏa mãn không:</p> <ul style="list-style-type: none"> - Học sinh có cảm xúc tiêu cực (buồn/chán nản) trong 3 buổi liên tiếp - Emotion Score giảm xuống dưới 4/10 - Có sự thay đổi đột ngột (tăng/giảm >30% trong 1 tuần) 	Hệ thống	<p>Đọc: Dữ liệu cảm xúc gần nhất.</p> <p>So sánh: Với các ngưỡng cảnh báo trong bảng alert_thresholds.</p>
10	<p>Gửi thông báo cho phụ huynh (nếu điều kiện thỏa): Nếu phát hiện dấu hiệu bất thường, hệ thống tự động:</p> <ul style="list-style-type: none"> - Gửi push notification đến Mobile App của phụ huynh - Gửi email chi tiết - Tạo notification trong hệ thống 	Hệ thống	<p>Ghi: Notification vào bảng notifications.</p> <p>Gọi: Push notification service API.</p> <p>Gửi: Email qua SMTP service.</p> <p>Cập nhật: Trạng thái đã gửi thông báo.</p>
11	Xem lịch sử thông báo: Phụ huynh có thể xem lại các thông báo đã nhận trong quá khứ.	Phụ huynh	Đọc: Danh sách notifications từ bảng notifications filter theo user_id và student_id.
12	<p>Xuất/Lưu báo cáo:</p> <ul style="list-style-type: none"> - Giáo viên: Xuất báo cáo PDF để lưu hồ sơ học sinh - Phụ huynh: Lưu báo cáo trong app để xem offline hoặc chia sẻ với thành viên gia đình khác 	Giáo viên / Phụ huynh	<p>Đọc: Toàn bộ dữ liệu hồ sơ học sinh.</p> <p>Xuất: File PDF hoặc lưu vào local storage của app.</p> <p>Ghi log: Thao tác xuất/lưu báo cáo.</p>
13	Thêm ghi chú (chỉ giáo viên): Giáo viên có thể thêm ghi chú quan sát cá nhân về học sinh (ví dụ: "Đã trao đổi với phụ huynh ngày 15/11").	Giáo viên	Lưu: Ghi chú vào bảng teacher_notes với timestamp và teacher_id.

Luồng sự kiện thay thế:

A1: Học sinh mới, chưa có đủ dữ liệu (tại bước 3)

- + 3a. Hệ thống phát hiện học sinh có ít hơn 3 buổi học đã ghi nhận.
- + 3b. Hiển thị thông báo: "Hồ sơ cảm xúc đang được xây dựng. Cần thêm dữ liệu từ các buổi học để có phân tích chính xác."

- + 3c. Chỉ hiển thị dữ liệu thô có sẵn, không có insight hay gợi ý.
- + 3d. Bỏ qua bước 6, 7.

A2: Phụ huynh có nhiều con (tại bước 2)

- + 2a. Sau khi đăng nhập, hệ thống hiển thị danh sách tất cả con em.
- + 2b. Phụ huynh chọn con nào để xem hồ sơ.
- + 2c. Tiếp tục bước 3 với học sinh được chọn.

A3: Giáo viên muốn so sánh nhiều học sinh (tại bước 2)

- + 2a. Giáo viên chọn chế độ "So sánh học sinh".
- + 2b. Chọn 2-4 học sinh từ danh sách.
- + 2c. Ở bước 4, các biểu đồ sẽ hiển thị dạng overlay để so sánh trực tiếp.

A4: Không có dữ liệu trong khoảng thời gian đã chọn (tại bước 5)

- + 5a. Hệ thống không tìm thấy dữ liệu cảm xúc trong chu kỳ được chọn (ví dụ: học sinh nghỉ học cả tháng).
- + 5b. Hiển thị thông báo: "Không có dữ liệu trong khoảng thời gian này."
- + 5c. Gợi ý chọn khoảng thời gian khác hoặc xem toàn bộ lịch sử.

Luồng ngoại lệ:

E1: Lỗi khi tính toán Emotion Score (tại bước 3)

- + 3e1. Hệ thống gặp lỗi khi tính toán (dữ liệu bị corrupt, missing values).
- + 3e2. Hiển thị "N/A" cho Emotion Score.
- + 3e3. Ghi log lỗi và thông báo cho admin.
- + 3e4. Tiếp tục hiển thị các phần khác của hồ sơ.

E2: Lỗi gửi push notification (tại bước 10)

- + 10e1. Push notification service không phản hồi hoặc trả về lỗi.
- + 10e2. Hệ thống thử gửi lại sau 5 phút (retry mechanism).
- + 10e3. Nếu thất bại 3 lần, chỉ gửi email và lưu notification trong hệ thống.

- + 10e4. Ghi log lỗi để admin kiểm tra.

E3: Dữ liệu bị mất hoặc không đồng bộ (bất kỳ bước nào)

- + Nếu phát hiện dữ liệu không nhất quán (ví dụ: tổng số buổi học không khớp).
- + Hiện thị cảnh báo: "Phát hiện dữ liệu không đồng bộ. Đang thực hiện kiểm tra..."
- + Hệ thống chạy data integrity check.
- + Nếu có thể sửa tự động thì sửa, nếu không thì báo admin.
- + Tiếp tục hiển thị dữ liệu với disclaimer: "Một số dữ liệu có thể chưa chính xác."

E4: Phụ huynh/Giáo viên truy cập học sinh không có quyền (tại bước 2)

- + 2e1. Hệ thống kiểm tra authorization.
- + 2e2. Nếu không có quyền (ví dụ: phụ huynh cố truy cập học sinh không phải con mình).
- + 2e3. Hiện thị thông báo: "Bạn không có quyền truy cập hồ sơ này."
- + 2e4. Ghi log cảnh báo về hành vi đáng ngờ.
- + 2e5. Use case kết thúc.

1.3. MỤC TIÊU VÀ YÊU CẦU ĐỀ TÀI

- Mục tiêu bao quát của đề tài là xây dựng và triển khai một hệ thống hoàn chỉnh, ứng dụng công nghệ deep learning để giám sát, phân tích và báo cáo trạng thái cảm xúc của học sinh trong môi trường giáo dục. Hệ thống hướng đến việc cung cấp một công cụ hỗ trợ đắc lực cho ba đối tượng chính: giáo viên có thể điều chỉnh phương pháp giảng dạy tức thời, nhà trường có thể đánh giá và nâng cao chất lượng chương trình học, và phụ huynh có thể đồng hành sâu sắc hơn với quá trình học tập của con em mình.
- Để hiện thực hóa mục tiêu chung, đề tài tập trung vào việc hoàn thành ba nhóm mục tiêu cụ thể sau:
- **Một là, nghiên cứu và phát triển lõi công nghệ AI nhận diện cảm xúc:**
 - + Tiến hành tổng quan, nghiên cứu các công trình khoa học liên quan đến nhận diện cảm xúc từ hình ảnh khuôn mặt (Facial Emotion Recognition - FER).
 - + Lựa chọn, so sánh và áp dụng các kiến trúc mô hình Deep Learning tiên tiến (ví dụ như CNN, ResNet, VGG) để xây dựng một pipeline AI có khả năng phân loại chính xác các trạng thái cảm xúc cơ bản (vui, buồn, ngạc nhiên, tập trung, chán nản...).
 - + Huấn luyện, tinh chỉnh và đánh giá hiệu năng của mô hình trên các tập dữ liệu chuẩn và dữ liệu thực tế để đảm bảo độ tin cậy khi triển khai.
- **Hai là, thiết kế và xây dựng các nghiệp vụ ứng dụng trong giáo dục:**
 - + Phát triển nghiệp vụ "Giám sát cảm xúc trong lớp học" với các chức năng cốt lõi như quản lý lớp học, thu nhận hình ảnh thời gian thực, hiển thị biểu đồ cảm xúc trên dashboard và thiết lập hệ thống cảnh báo sớm cho giáo viên.
 - + Xây dựng nghiệp vụ "Phân tích chất lượng giảng dạy" bằng cách tổng hợp và phân tích dữ liệu cảm xúc từ nhiều buổi học, cho phép so sánh và đưa ra các báo cáo thống kê chuyên sâu.
 - + Triển khai nghiệp vụ "Hỗ trợ cá nhân hóa học tập" thông qua việc xây dựng hồ sơ cảm xúc cá nhân cho từng học sinh, giúp đưa ra các gợi ý và báo cáo riêng biệt cho giáo viên và phụ huynh.

- **Ba là, phát triển các nền tảng phần mềm cho người dùng cuối:**
 - + Xây dựng một ứng dụng Web (Web Application) hoàn chỉnh dành cho giáo viên và ban giám hiệu nhà trường. Nền tảng này sẽ cung cấp giao diện quản trị, các công cụ dashboard, phân tích dữ liệu và xuất báo cáo.
 - + Phát triển một ứng dụng Di động (Mobile Application) tiện lợi cho phụ huynh và học sinh, tập trung vào việc xem báo cáo cá nhân, theo dõi lịch sử cảm xúc và nhận thông báo từ hệ thống.
 - + Thiết kế và triển khai hệ thống API và cơ sở dữ liệu tập trung để đảm bảo sự đồng bộ, nhất quán về dữ liệu giữa hai nền tảng Web và Mobile.

1.3.1. Các yêu cầu chức năng

- Hệ thống phải đáp ứng đầy đủ các chức năng tương ứng với ba nhóm nghiệp vụ chính:
- **Nghiệp vụ 1: Giám sát cảm xúc trong lớp học**
 - + Cho phép giáo viên quản lý lớp học, tạo các buổi học.
 - + Hệ thống có khả năng thu nhận hình ảnh từ webcam hoặc camera lớp học để thực hiện nhận diện khuôn mặt và phân loại cảm xúc.
 - + Cung cấp một Bảng điều khiển (Dashboard) trực quan, hiển thị biểu đồ cảm xúc của cả lớp và từng học sinh theo thời gian thực.
 - + Tự động gửi cảnh báo cho giáo viên khi phát hiện tỷ lệ học sinh có cảm xúc tiêu cực (ví dụ: "chán nản") vượt ngưỡng cho phép (ví dụ: 40% trong 5 phút).
 - + Cho phép xuất báo cáo chi tiết của buổi học dưới dạng file PDF hoặc CSV.
- **Nghiệp vụ 2: Phân tích chất lượng giảng dạy**
 - + Hệ thống có khả năng tổng hợp và lưu trữ dữ liệu cảm xúc từ nhiều buổi học của cùng một môn học.
 - + Cung cấp chức năng so sánh, đối chiếu dữ liệu cảm xúc giữa các tiết học hoặc các giảng viên khác nhau.

- + Tạo báo cáo thống kê tỷ lệ cảm xúc tích cực/tiêu cực trong từng giai đoạn của bài giảng (mở bài, giảng bài, ôn tập, kết thúc).
- + Cho phép xuất báo cáo tổng hợp để gửi cho tổ bộ môn hoặc ban giám hiệu.
- **Nhiệm vụ 3: Hỗ trợ cá nhân hóa học tập**
 - + Xây dựng hồ sơ cá nhân (profile) cho mỗi học sinh, lưu trữ thông tin và lịch sử cảm xúc.
 - + Hiện thị biểu đồ lịch sử cảm xúc của học sinh theo các chu kỳ tuần hoặc tháng.
 - + Hệ thống đưa ra gợi ý cho giáo viên về những học sinh có xu hướng chán nản thường xuyên hoặc những em có mức độ hứng thú cao.
 - + Tạo báo cáo riêng về tình hình học tập và cảm xúc của từng học sinh để gửi cho phụ huynh.

1.3.2. Các yêu cầu phi chức năng

- Tính chính xác (Accuracy): Mô hình AI nhận diện cảm xúc phải đạt độ chính xác cao trên tập dữ liệu kiểm thử và hoạt động tốt trong điều kiện thực tế của lớp học.
- Hiệu năng (Performance): Hệ thống phải có khả năng xử lý và phản hồi kết quả phân tích cảm xúc gần với thời gian thực để đảm bảo tính hữu dụng cho giáo viên.
- Tính ổn định và tin cậy (Stability & Reliability): Hệ thống phải hoạt động ổn định, liên tục trong suốt quá trình diễn ra buổi học.
- Giao diện người dùng (UI/UX): Giao diện của cả Web App và Mobile App phải được thiết kế thân thiện, trực quan và dễ sử dụng cho các đối tượng người dùng khác nhau.
- Khả năng mở rộng (Scalability): Kiến trúc hệ thống được thiết kế theo hướng module hóa, dễ dàng bảo trì và tích hợp thêm các chức năng mới trong tương lai.

1.4. CÁC CÔNG CỤ, CÔNG NGHỆ

- **Ngôn ngữ chính: Python**
 - + Python là ngôn ngữ lập trình bậc cao, cú pháp đơn giản, dễ đọc, rất phù hợp cho các bài toán trí tuệ nhân tạo và xử lý dữ liệu.
- **Framework Backend: FastAPI**
 - + FastAPI là một framework backend cho Python, hỗ trợ xây dựng các dịch vụ Web API theo chuẩn REST một cách nhanh chóng và rõ ràng.
- **Framework FrontEnd: ReactJs**
 - + ReactJS là thư viện JavaScript dùng để xây dựng giao diện người dùng theo hướng component, giúp việc tái sử dụng và bảo trì mã nguồn trở nên thuận tiện. Với cơ chế Virtual DOM và quản lý trạng thái linh hoạt, ReactJS hỗ trợ hiển thị dữ liệu theo thời gian thực và tương tác người dùng mượt mà.
- **Framework AI: Ultralytics (cho mô hình YOLO)**
 - + Ultralytics cung cấp triển khai mô hình YOLO hiện đại, hỗ trợ huấn luyện và suy luận (inference) mô hình phát hiện đối tượng một cách đơn giản và hiệu quả. Thư viện này tích hợp sẵn nhiều tiện ích như tải mô hình có sẵn, tùy chỉnh cấu hình, xuất kết quả và hỗ trợ nhiều định dạng đầu vào, phù hợp cho bài toán nhận diện khuôn mặt và cảm xúc trong luận văn.
- **CSDL: PostgreSQL**
 - + PostgreSQL là hệ quản trị cơ sở dữ liệu quan hệ mã nguồn mở mạnh mẽ, hỗ trợ tốt tính toàn vẹn dữ liệu và các truy vấn phức tạp. Hệ thống sử dụng PostgreSQL để lưu trữ thông tin người dùng, kết quả phân tích cảm xúc và các dữ liệu nghiệp vụ, đảm bảo khả năng mở rộng và an toàn dữ liệu.
- **Các công cụ khác: VS code, Google Collab, Git**
 - + VS code hỗ trợ môi trường triển khai FrontEnd và Backend cũng như triển khai mô hình AI sau khi được huấn luyện từ Google Collab cũng như lưu trữ mã nguồn lên GitHub nhờ Git.

CHƯƠNG 2: CƠ SỞ LÝ THUYẾT VÀ CÁC CÔNG TRÌNH LIÊN QUAN

2.1. MẠNG NƠ-RON TÍCH CHẬP (CNN)

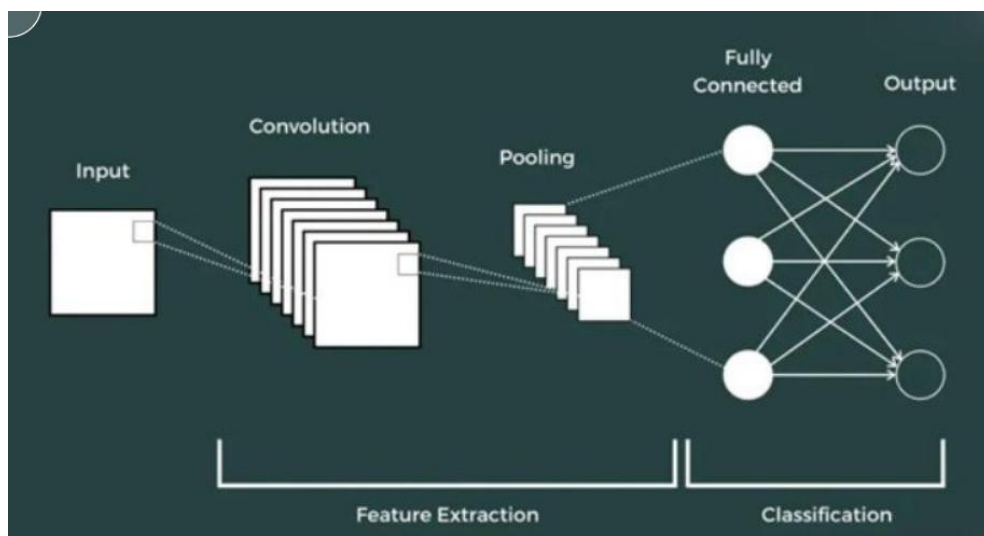
2.1.1. Giới thiệu

- CNN, viết tắt của Convolutional Neural Network, tiếng Việt gọi là Mạng Nơ-ron Tích chập, là một kiến trúc mạng nơ-ron đặc biệt [7,8]. Nó được thiết kế chuyên biệt để phân tích dữ liệu có cấu trúc lưới, ví dụ hình ảnh, video, tín hiệu âm thanh hoặc văn bản. Điểm nổi bật của CNN là khả năng tự động học và rút trích các đặc trưng then chốt từ dữ liệu mà không cần sự can thiệp thủ công nhiều[9,10].
- Mạng Nơ-ron Tích chập là một đột phá lớn trong lĩnh vực Deep Learning (Học sâu). So với các mạng nơ-ron truyền thống, CNN có ưu thế vượt trội trong việc nhận diện các mẫu phức tạp trong dữ liệu đa phương tiện với độ chính xác cao. Điều này giúp CNN trở thành công cụ không thể thiếu trong nhiều ứng dụng thực tiễn, bao gồm nhận diện khuôn mặt, phân loại hình ảnh và nhiều ứng dụng khác.

2.1.2. Kiến trúc tổng quát

- Mạng nơ-ron tích chập bao gồm nhiều loại lớp khác nhau, mỗi lớp đều có chức năng riêng biệt:
 - + Tầng tích chập (Convolutional Layer): Là thành phần cốt lõi của CNN, được sử dụng để trích xuất các đặc trưng (features) từ hình ảnh đầu vào. Mục tiêu của tầng này là bảo toàn mối quan hệ không gian giữa các điểm ảnh (pixel) bằng cách học cách phát hiện các đặc điểm cục bộ như cạnh, góc, họa tiết hoặc mẫu hình trong ảnh. Khác với việc xử lý toàn bộ ảnh cùng một lúc, tầng tích chập sử dụng các ô vuông nhỏ (filter hoặc kernel) trượt qua từng vùng của ảnh để tính toán. Nhờ đó, mạng có thể tập trung vào những chi tiết quan trọng mà không cần phân tích toàn bộ dữ liệu ảnh một cách tường minh[8,9].

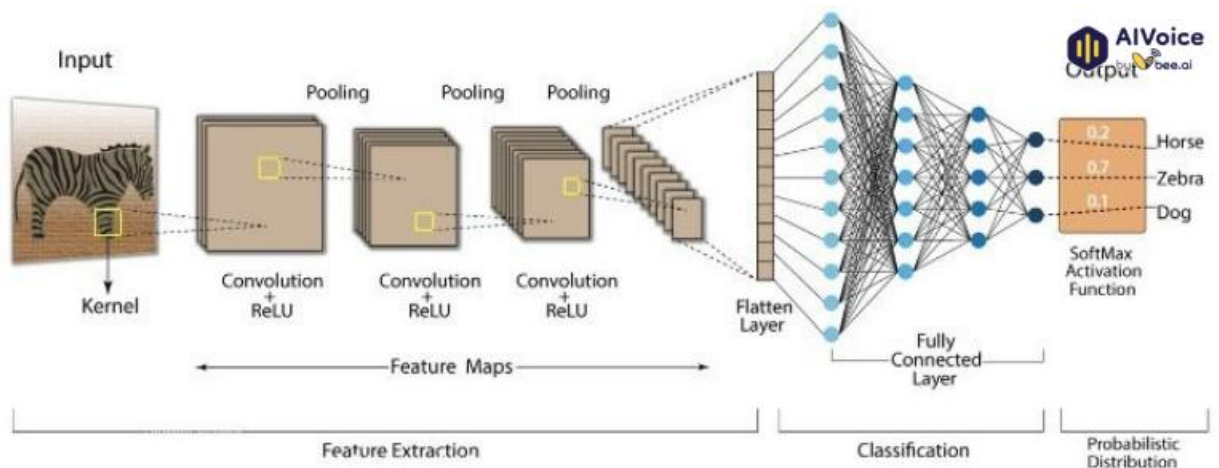
- + Tầng kích hoạt (Activation Layer): Tầng này áp dụng một hàm phi tuyến, thường là ReLU, lên kết quả của lớp tích chập. ReLU chuyển đổi tất cả giá trị âm thành 0 trong khi giữ nguyên các giá trị dương. Hàm ReLU giúp mô hình học được các mối quan hệ phức tạp và phi tuyến trong dữ liệu, đồng thời giải quyết vấn đề gradient biến mất khi huấn luyện mạng sâu[10].
- + Tầng gộp (Pooling Layer): Tầng này giảm kích thước không gian của bản đồ đặc trưng, giảm số lượng tham số và tính toán trong mạng. Max Pooling là phương pháp phổ biến nhất, lấy giá trị lớn nhất trong một cửa sổ. Pooling cũng giúp mô hình ít nhạy cảm hơn với vị trí chính xác của đặc trưng, tăng khả năng tổng quát hóa[6,11].
- + Tầng kết nối đầy đủ (Fully-Connected Layer): Ở cuối CNN, các bản đồ đặc trưng được làm phẳng thành một vector và kết nối với một hoặc nhiều tầng kết nối đầy đủ. Các tầng này hoạt động giống như mạng nơ-ron truyền thống, kết hợp mọi đặc trưng đã học được để đưa ra quyết định cuối cùng, như phân loại đối tượng trong ảnh.
- + Ngoài ra, CNN thường có một vài lớp bổ sung như Dropout (giúp ngăn overfitting bằng cách ngẫu nhiên tắt một số nơ-ron trong quá trình huấn luyện) và Batch Normalization (chuẩn hóa đầu ra của tầng trước để giúp mạng hội tụ nhanh hơn và ổn định hơn).



Hình 2. 1: Kiến trúc tổng quát CNN.

2.1.3. Cơ chế hoạt động.

- Cơ chế hoạt động của CNN bắt đầu bằng việc áp dụng các bộ lọc tích chập để tạo ra 1 bản đồ đặc trưng, sau đó kết quả đi qua hàm kích hoạt để học các mối quan hệ phi tuyến. Sau đó, mô hình sử dụng pooling (thường là Max Pooling) để giảm kích thước và tăng tính bền vững với nhiễu.
- Ở cuối mạng nơ-ron, dữ liệu được làm phẳng thành vector và đưa qua các lớp kết nối đầy đủ để suy luận và phân loại. Toàn bộ quá trình học diễn ra nhờ thuật toán lan truyền ngược (Backpropagation), giúp CNN liên tục điều chỉnh trọng số để giảm sai số và cải thiện khả năng nhận diện mẫu.



Hình 2. 2: Quá trình học đặc trưng và phân loại.

2.1.4. CNN trong nhận diện cảm xúc

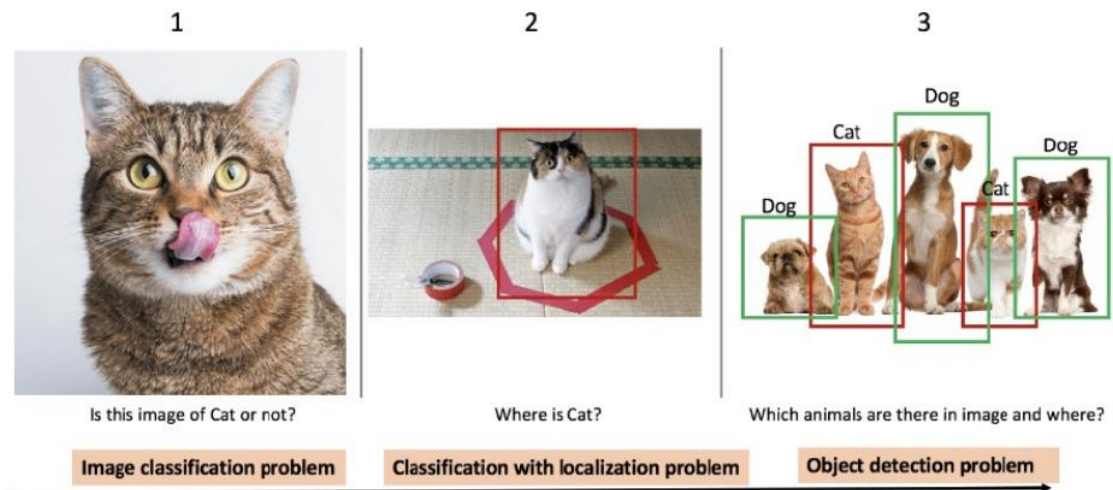
- Trong lĩnh vực thị giác máy tính hiện đại, Mạng nơ-ron tích chập (Convolutional Neural Networks - CNN) đã trở thành phương pháp nòng cốt và hiệu quả nhất để giải quyết bài toán nhận diện cảm xúc, đặc biệt là thông qua phân tích biểu cảm khuôn mặt (Facial Expression Recognition - FER). Khác với các kỹ thuật truyền thống đòi hỏi việc thiết kế và trích xuất đặc trưng thủ công phức tạp, ưu điểm vượt trội của CNN nằm ở khả năng tự động học các đặc trưng từ dữ liệu hình ảnh, đi từ các chi tiết sơ cấp như cạnh, góc đến các cấu trúc phức tạp như độ cong của môi hay nếp nhăn vùng mắt. Khả năng bất biến dịch chuyển (translation invariance) cũng cho

phép CNN nhận diện chính xác biểu cảm ngay cả khi vị trí khuôn mặt thay đổi trong khung hình [12,28].

2.2. MÔ HÌNH YOLO

2.2.1. Phát hiện đối tượng (Object detection)

- Object detection (phát hiện đối tượng) là một kỹ thuật trong thị giác máy tính để xác định vị trí của các đối tượng trong hình ảnh hoặc video. Các thuật toán phát hiện đối tượng thường tận dụng Machine Learning hoặc Deep Learning để tạo ra kết quả có ý nghĩa. Khi con người nhìn vào hình ảnh hoặc video, ta có thể nhận ra và định vị các đối tượng quan tâm trong tích tắc[13,14].
- Các khái niệm liên quan đến Object Detection:
 - + Phân loại hình ảnh (Image Classification): dự đoán nhãn của đối tượng trong hình ảnh. Với đầu vào (input) là hình ảnh của đối tượng, chẳng hạn một bức ảnh và đầu ra sẽ là nhãn lớp của tấm ảnh đầu vào đó (ví dụ một hoặc nhiều số nguyên ánh xạ tới nhãn lớp).
 - + Định vị đối tượng (Object Localization): xác định vị trí hiện diện của các đối tượng trong hình ảnh và cho biết vị trí của chúng bằng bounding box. Với đầu vào là ảnh có một hoặc nhiều đối tượng và đầu ra là một hoặc nhiều bounding box được xác định bởi tọa độ tâm, chiều rộng, chiều cao.
 - + Phát hiện đối tượng (Object Detection): là sự kết hợp của 2 khái niệm trên, vẽ một bounding box xung quan đối tượng quan tâm trong ảnh và gán cho chúng một nhãn.



Hình 2. 3: Phân biệt bài toán Image Classification - Object Localization – Object Detection

2.2.2. Tổng quan YOLO

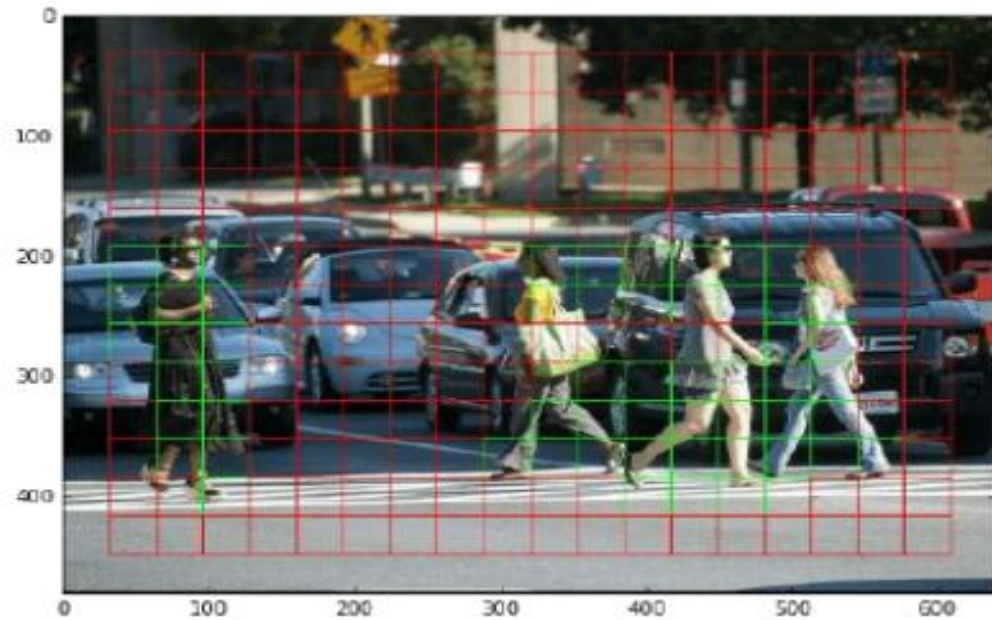
- YOLO (You Only Look Once) là một mạng nơ-ron đặc biệt cho các bài toán phát hiện đối tượng trong hình ảnh, nó nổi bật với câu “Bạn chỉ nhìn một lần”. Khác với các phương pháp 2 giai đoạn truyền thống như (R-CNN, Faster R-CNN) phải đề xuất các vùng trước rồi phân loại thì YOLO lại thực hiện cả 2 công đoạn nhận diện và định vị tất cả đối tượng trong ảnh với chỉ 1 lần lan truyền mạng duy nhất. Nói cách khác, YOLO coi việc phát hiện đối tượng như 1 bài toán hồi quy, ánh xạ từ ảnh đầu vào đến tọa độ bounding box và nhãn lớp tương ứng[16].
- Ưu điểm lớn nhất của YOLO phải nói đến là hoạt động tốt với thời gian thực (real-time) và khả năng xử lý tổng thể.

2.2.3. Nguyên lý hoạt động

- Mô hình YOLO sử dụng các kỹ thuật sau:
 - + Residual blocks (Khối dư)
 - + Bounding box regression (Hồi quy hộp giới hạn)
 - + Intersection Over Union (IOU)

2.2.3.1. Khối dư (Residual blocks)[15,16]

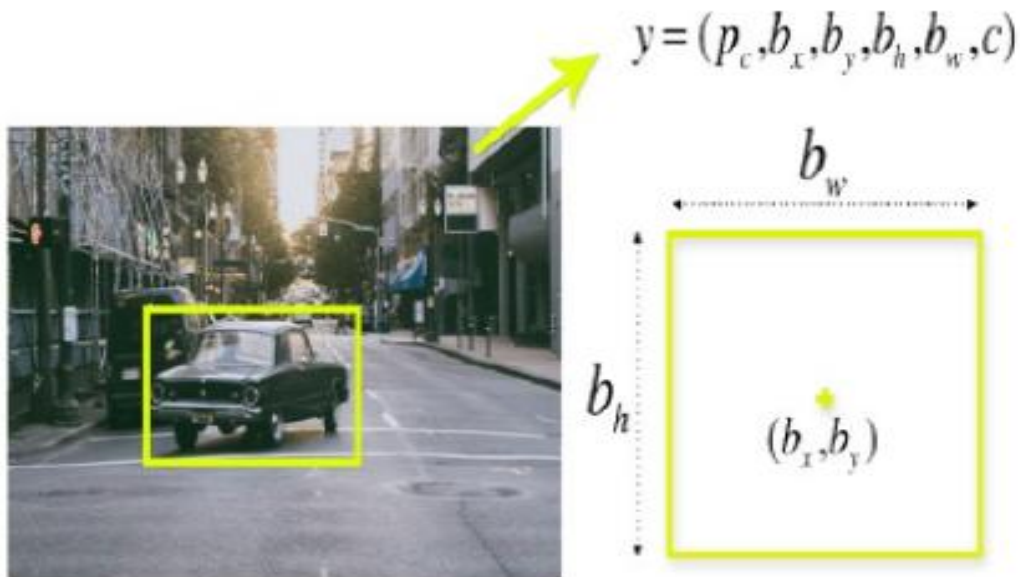
Đầu tiên, hình ảnh đầu vào được chia thành nhiều lưới khác nhau. Mỗi lưới có kích thước là $S \times S$



Hình 2. 4: Hình ảnh đầu vào được chia thành các lưới, có nhiều ô lưới có kích thước bằng nhau.

2.2.3.2. Hồi quy hộp giới hạn (Bounding box regression)[16,17]

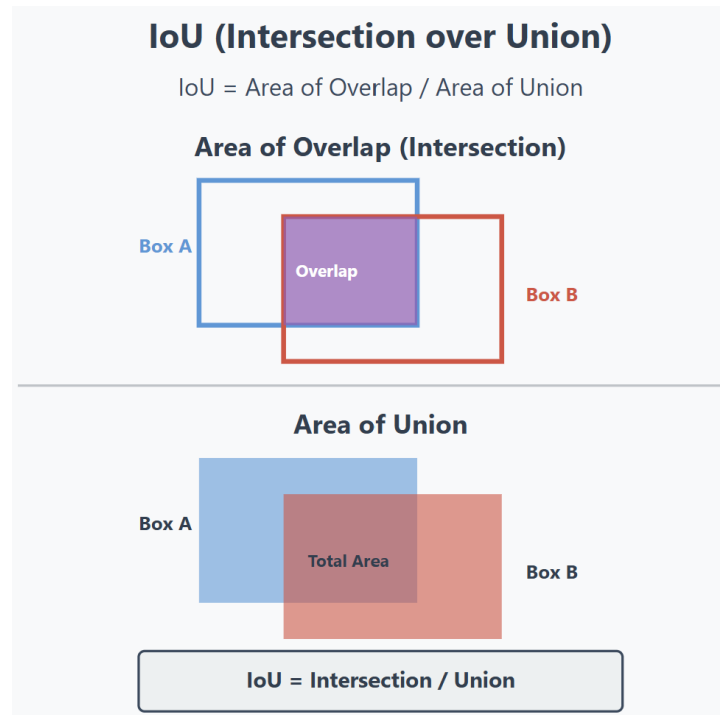
- Là bài toán dự đoán khi đầu ra là các biến liên tục. vì bounding box ở đây là bốn biến liên tục (x,y,w,h) nên gọi là bài toán hồi quy (regression)
- Mỗi bounding box trong ảnh bao gồm các thuộc tính sau:
 - + Chiều rộng (bw)
 - + Chiều cao (bh)
 - + Các lớp (thường được thể hiện bằng chữ c)
 - + Đường viền tâm hộp (b_x, b_y)
- YOLO sử dụng bounding box regression duy nhất để dự đoán chiều cao, chiều rộng, tâm và lớp của đối tượng.



Hình 2. 5: Minh họa bounding box.

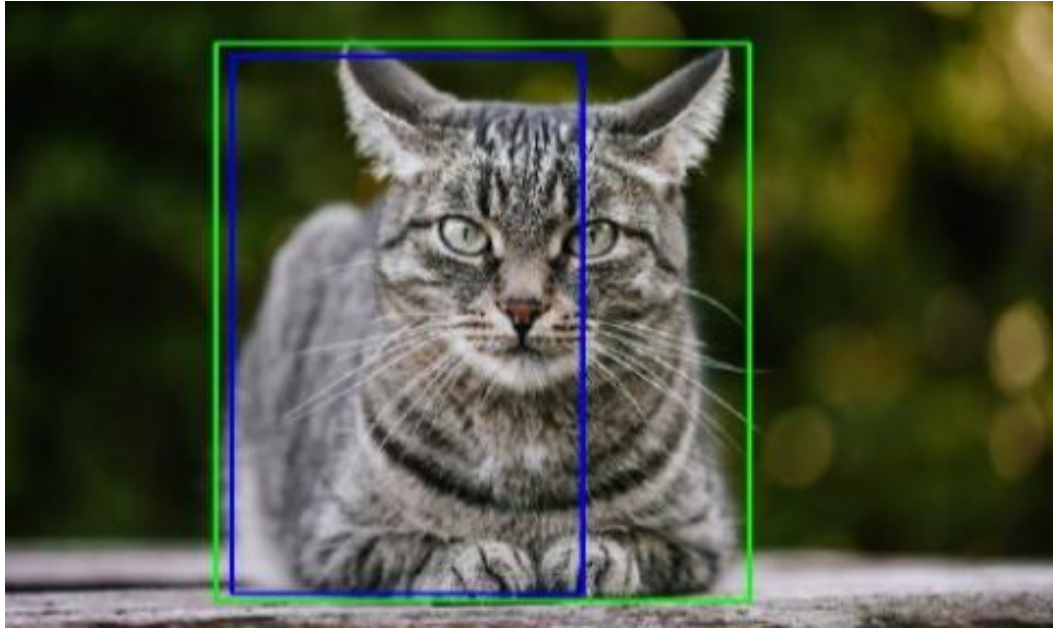
2.2.3.3. Intersection Over Union (IOU)[16,19]

- IOU là hàm đánh giá độ chính xác của bài toán phát hiện vật thể trên tập dữ liệu xác định.
- Công thức tính IOU:



Hình 2. 6: Công thức tính IOU.

- Trong đó **Area of Overlap** là diện tích vùng giao giữa **predicted bounding box** và **grouth-truth bouding box**, còn **Area of Union** là diện tích phần hợp giữa **predicted bounding box** và **grouth-truth bouding box**. Những bounding box được đánh giá nhãn bằng tay trong tập Training set và Test set. Nếu $\text{IOU} > 0.5$ thì dự đoán được đánh giá là tốt.

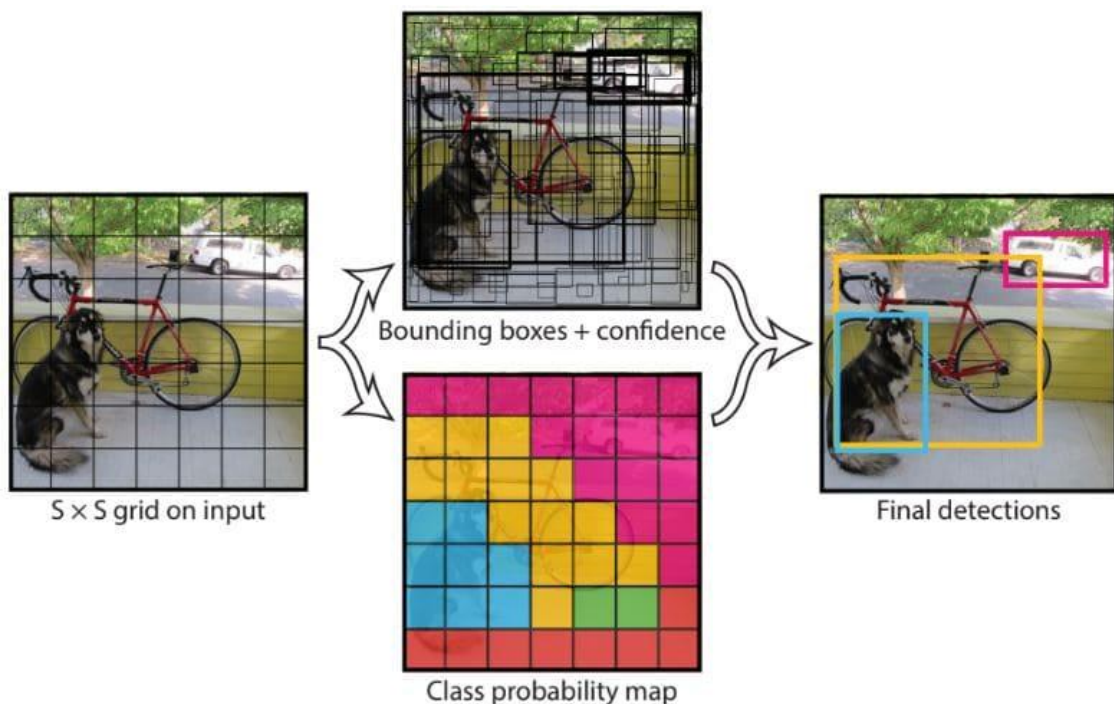


Hình 2. 7: Hai bounding box xanh lam, xanh lá.

- Trong hình 2.7 có hai bounding box, một box màu xanh lá cây và box còn lại màu xanh lam. Box màu xanh lam là box dự đoán trong khi box màu xanh lá cây là hộp thực. YOLO đảm bảo rằng hai bounding box này bằng nhau.
- Về cơ bản, YOLO chia ảnh đầu vào thành các ô nhỏ, mô hình sẽ dự đoán xác suất đối tượng trong các đường bounding box. Những đường bounding box có tỷ lệ cao sẽ được giữ lại cho nhiệm vụ xác định vị trí của đối tượng.
- Chia ảnh thành lưới $S \times S$: Hình ảnh đầu vào được chia thành 1 lưới có kích thước $S \times S$, mỗi ô trong lưới sẽ đảm nhận nhiệm vụ phát hiện vật thể nếu tâm của đối tượng nằm trong đó. Với mỗi ô, YOLO sẽ dự đoán nhiều hộp giới hạn. Mỗi hộp sẽ đi kèm với một điểm tin cậy, thể hiện xác suất chắc chắn rằng trong hộp có một đối tượng và độ chính xác cho việc xác định vị trí của vật thể.
- Tiếp nhận và tiếp nhận xử lý bằng CNN: Mạng CNN chính là xương sống của kiến trúc mô hình YOLO, trong đó 20 lớp tích chập đầu tiên được huấn luyện trước trên tập dữ liệu imageNet. Lúc đầu, mô hình được thêm các lớp trung gian như lớp tổng hợp trung bình và lớp kết nối đầy đủ cho

việc phân loại ảnh. Sau giai đoạn huấn luyện này, mô hình sẽ được chỉnh sửa lại cho nhiệm vụ phát hiện đối tượng. Lúc này, lớp kết nối đầy đủ cuối cùng sẽ đồng thời dự đoán cả loại đối tượng và vị trí của các hộp giới hạn.

- Dự đoán dựa trên độ trùng khớp: Trong quá trình huấn luyện, để đảm bảo đối tượng có độ chính xác nhất, mô hình chọn ra các hộp giới hạn có độ trùng khớp cao nhất với thực tế (IOU - Intersection over Union) để đại diện cho đối tượng đó. Sau khi dự đoán, YOLO sử dụng kỹ thuật Non-Maximum Suppression (NMS) để loại bỏ các hộp chồng lặp, chỉ giữ lại một hộp đại diện tốt nhất cho mỗi đối tượng.



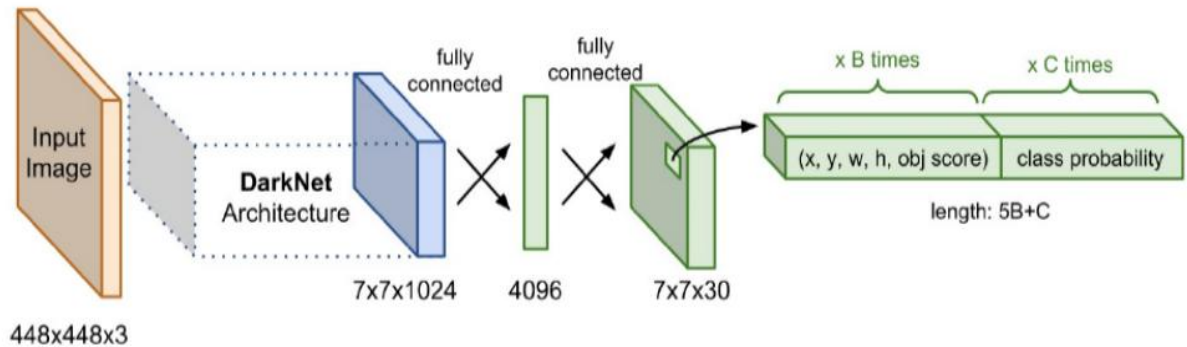
Hình 2. 8: Mô hình YOLO thực hiện xử lý hình ảnh theo tuần tự từng bước.

2.2.4. Kiến trúc hiện đại

- Kiến trúc YOLO bao gồm: base network là các mạng tích chập làm nhiệm vụ trích xuất đặc trưng. Phần phía sau là những Extra Layers được áp dụng để phát hiện vật thể trên feature map của base network.
- Thành phần Darknet Architecture được gọi là base network có tác dụng trích xuất đặc trưng. Đầu ra của base network là một feature map có kích

thước $7 \times 7 \times 1024$ sẽ được sử dụng làm đầu vào cho các extra layers có nhiệm vụ dự đoán nhãn và tọa độ bounding box của đối tượng.

- Base network của YOLO chủ yếu là các lớp tích chập. Trong đó các lớp tích chập sẽ trích xuất đặc trưng của ảnh, còn lớp kết nối đầy đủ sẽ dự đoán xác suất và tọa độ của đối tượng.



Hình 2. 9: Sơ đồ kiến trúc mạng YOLO.

2.2.5. Các phiên bản YOLO

- YOLOv1: Phiên bản gốc và đầu tiên
- YOLOv2: Cải thiện tốc độ và độ chính xác, áp dụng chuẩn hóa hàng loạt (batch normalization) và hộp neo (anchor boxes).
- YOLOv3 – YOLOv7: Cải thiện tốc độ
- YOLOv8: Ra mắt bởi Ultralytics, với kiến trúc được thiết kế lại và hỗ trợ nhiều tác vụ thị giác máy tính.
- YOLOv9: Giới thiệu các phương pháp mới như Thông tin Gradient có thể lập trình (PGI) và Mạng tổng hợp lớp hiệu quả tổng quát (GELAN).
- YOLOv10: Thiết kế thân thiện với phần cứng, cho phép phát hiện đối tượng không cần NMS (phần không tối đa).
- YOLOv11: Phiên bản mới hơn, sử dụng kiến trúc lai CNN-Transformer để cải thiện tốc độ và độ chính xác.

Release	Year	Tasks	Contributions	Framework
YOLO [5]	2015	Object Detection, Basic Classification	Single-stage object detector	Darknet
YOLOv2 [7]	2016	Object Detection, Improved Classification	Multi-scale training, dimension clustering	Darknet
YOLOv3 [8]	2018	Object Detection, Multi-scale Detection	SPP block, Darknet-53 backbone	Darknet
YOLOv4 [9]	2020	Object Detection, Basic Object Tracking	Mish activation, CSPDarknet-53 backbone	Darknet
YOLOv5 [10]	2020	Object Detection, Basic Instance Segmentation (via custom modifications)	Anchor-free detection, SWISH activation, PANet	PyTorch
YOLOv6 [11]	2022	Object Detection, Instance Segmentation	Self-attention, anchor-free OD	PyTorch
YOLOv7 [12]	2022	Object Detection, Object Tracking, Instance Segmentation	Transformers, E-ELAN reparameterisation	PyTorch
YOLOv8 [13]	2023	Object Detection, Instance Segmentation, Panoptic Segmentation, Key-point Estimation	GANs, anchor-free detection	PyTorch
YOLOv9 [14]	2024	Object Detection, Instance Segmentation	PGI and GELAN	PyTorch
YOLOv10 [15]	2024	Object Detection	Consistent dual assignments for NMS-free training	PyTorch

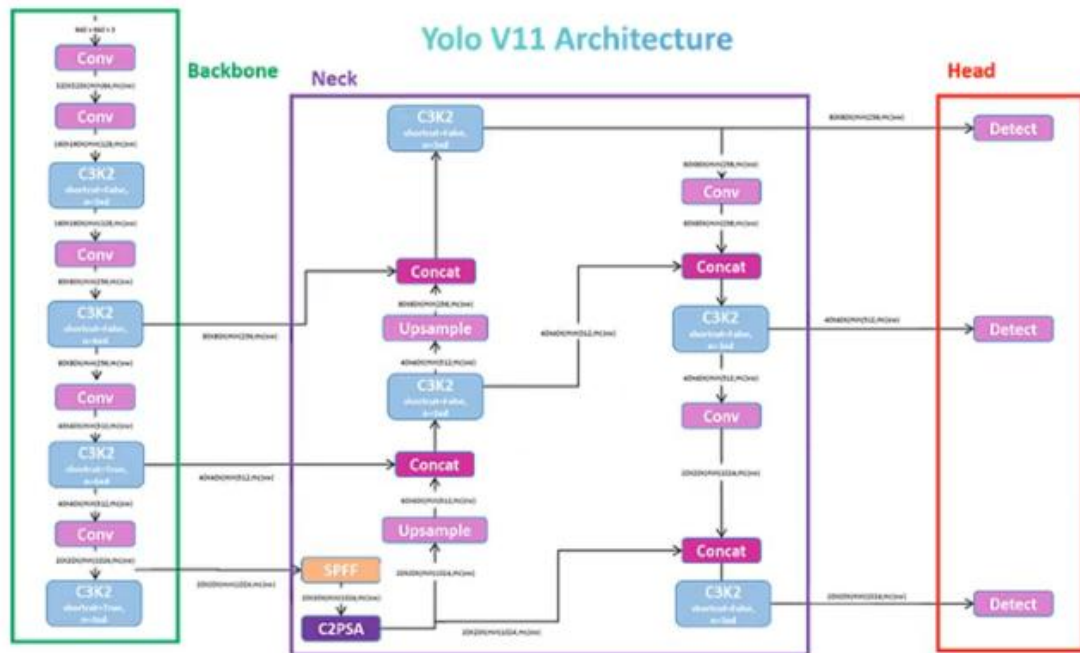
Hình 2. 10: Các phiên bản YOLO.

2.2.6. YOLOv11

2.2.6.1. Tổng quan

- YOLOv11 là phiên bản mới nhất trong chuỗi thuật toán phát hiện đối tượng YOLO được ra mắt tại hội nghị YOLO Vision 2024 (YV24) , tiếp nối di sản này với những cải tiến đáng kể trong kiến trúc và phương pháp huấn luyện. Mục tiêu của YOLOv11 là nâng cao ranh giới về độ chính xác, tốc độ và hiệu quả cho các ứng dụng thị giác máy tính (CV) thời gian thực[22,25].

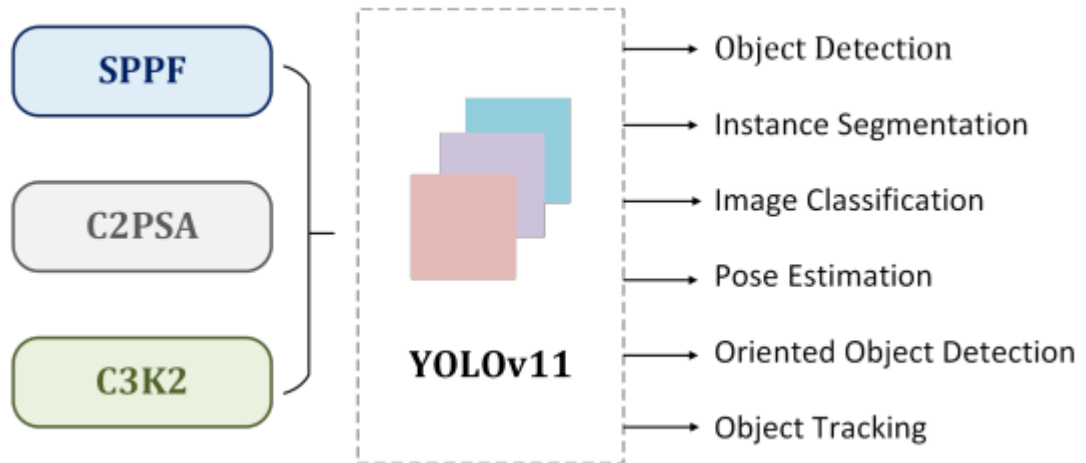
2.2.6.2. Các đổi mới kiến trúc



Hình 2. 11: Cấu trúc của Mô hình YOLOv11.

- YOLOv11 được xây dựng dựa trên nền tảng của YOLOv8 và các thành phần kiến trúc mới để đạt được hiệu suất tối đa:[22]
- + C3k2 Block (Cross Stage Partial with kernel size 2): là cải tiến đáng kể, thay thế cho C2f block được sử dụng cho các phiên bản trước. C3k2 là một triển khai hiệu quả tính toán hơn CSP Bottleneck, sử dụng hai phép tích chập nhỏ thay vì phép tích chập lớn, giúp xử lý nhanh hơn nhưng vẫn duy trì hiệu suất. Block được áp dụng trong Backbone (bộ trích xuất đặc trưng), Neck (bộ tổng hợp đặc trưng), và Head (bộ dự đoán) của mô hình.
- + C2PSA Block (Convolutional block with Parallel Spatial Attention): là phần bổ sung đáng chú ý, được thêm vào block SSPF. Cơ chế chú ý không gian này cho phép mô hình tập trung hiệu quả hơn vào các vùng quan trọng trong ảnh, có khả năng cải thiện độ chính xác, đặc biệt các đối tượng có vị trí và kích thước khác nhau.

- + SPPF (Spatial Pyramid Pooling - Fast): block được giữ lại từ các phiên bản trước để tăng cường trích xuất đặc trưng.



Hình 2. 12: Kiến trúc cải tiến của YOLOv11.

2.2.6.3. Các tác vụ Thị giác máy tính được YOLOv11 hỗ trợ[22,24]

- Phát hiện đối tượng (Object detection): Xác định và định vị đối tượng bằng hộp giới hạn.
- Phân đoạn vật thể (Instance Segmentation): Xác định và phân tách các đối tượng riêng lẻ xuống cấp độ pixel.
- Phân loại ảnh (Image Classification): Phân loại toàn bộ hình ảnh vào các danh mục định trước.
- Ước tính tư thế (Pose Estimation): Phát hiện các điểm khóa cụ thể để theo dõi chuyển động hoặc tư thế.
- Phát hiện đối tượng định hướng (Oriented Object Detection - OBB): Phát hiện đối tượng với một góc định hướng, đặc biệt có giá trị trong phân tích hình ảnh từ trên không hoặc robot.
- Theo dõi đối tượng (Object Tracking): Nhận dạng và theo dõi đường đi của đối tượng trong chuỗi video.

2.2.6.4. So sánh YOLOv11 so với các phiên bản cũ[5,22]

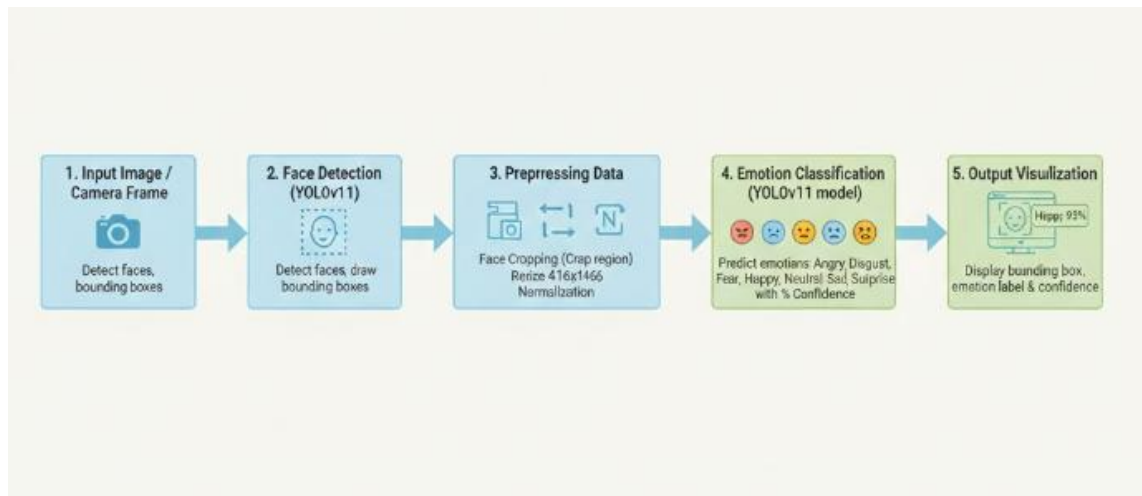
- YOLOv11 đại diện cho một bước tiến đáng kể trong công nghệ phát hiện đối tượng, đạt được sự cân bằng tốt hơn giữa độ chính xác và hiệu quả tính toán.
- So với YOLOv8, YOLOv10, YOLOv11 sử dụng kiến trúc backbone tối ưu hơn và kết hợp các kỹ thuật hiện đại, tốt hơn cho việc xử lý các đặc trưng đồng thời head detection nhẹ nhưng hiệu quả hơn.

2.3. BÀI TOÁN NHẬN DẠNG CẢM XÚC KHUÔN MẶT (FACIAL EMOTION RECOGNITION)

2.3.1. Tổng quan bài toán nhận diện cảm xúc

- Nhận diện cảm xúc khuôn mặt (Facial Emotion Recognition – FER) sử dụng Thị giác máy tính trong Deep Learning là 1 lĩnh vực nghiên cứu rộng lớn. Những ứng dụng nó vào môi trường giáo dục để đo lường mức độ tương tác của học sinh vẫn còn là một nước đi khá mới mẻ[26,27,28].
- Về bản chất, bài toán FER này không chỉ dừng ở mức phân tích ảnh tĩnh, mà là xử lý hình ảnh và video thời gian thực của học sinh để xác định các loại cảm xúc. Hệ thống này tập trung phân tích và nhận diện 7 loại cảm xúc cơ bản của con người bao gồm: vui vẻ, buồn bã, ghê tởm, ngạc nhiên, giận dữ, sợ hãi và trạng thái trung tính. Từ các cảm xúc này, hệ thống sẽ ánh xạ sang các mức độ "tham gia" (engagement levels) của học sinh vào bài giảng, từ đó đánh giá buổi học và hỗ trợ giáo viên điều chỉnh phương pháp dạy học.
- Thách thức về dữ liệu thực tế, so sánh sự khác biệt giữa dữ liệu phòng lab và dữ liệu thực tế.

2.3.4. Quy trình nhận diện cảm xúc



Hình 2. 13: Quy trình bài toán nhận diện cảm xúc.

- Hệ thống nhận diện cảm xúc trong đề tài hoạt động theo một chuỗi xử lý gồm 5 giai đoạn chính:[30,31,32]
- 1. Thu thập dữ liệu đầu vào
 - + Ảnh đầu có thể là camera lớp học , webcam hay các ảnh tĩnh nhưng thường trong lớp học sẽ là camera.
 - + Camera lớp học liên tục ghi lại hình ảnh khuôn mặt học sinh. Là nguồn dữ liệu đầu vào duy nhất của hệ thống. Các khung hình được capture và gửi về mô đun theo thời gian thực (30 FPS hoặc có thể tự tùy chỉnh).
- 2. Phát hiện khuôn mặt
 - + Các khung hình từ camera được capture vào đưa vào mô hình YOLOv11 cho việc phát hiện khuôn mặt, sau đó mô hình sẽ dự đoán bounding box cho từng khuôn mặt và kết quả cho bước này là các vùng ảnh chứa khuôn mặt sẽ được xác định, tọa độ bounding box được trích xuất để phục vụ cho giai đoạn sau.
- 3. Tiền xử lí ảnh đầu vào
 - + Sau khi phát hiện được khuôn mặt, hệ thống sẽ thực hiện cắt vùng khuôn mặt, resize lại kích thước chuẩn cho đầu vào mô hình (thường sẽ là 640 x 640) nhưng để đảm bảo cho máy yếu nên sẽ được resize về kích thước 416

x 416 để có thể dễ dàng huấn luyện, cuối cùng là chuẩn hóa dữ liệu sao cho phù hợp với mô hình.

4. Phân loại cảm xúc

- + Ảnh khuôn mặt sau khi được phát hiện và tiền xử lý cho phù hợp với mô hình sẽ được đưa vào mô hình nhận diện cảm xúc bằng YOLOv11 đã được train với bộ dữ liệu cho việc nhận diện các cảm xúc cơ bản (Angry-Disgust-Fear-Happy-Neutral-Sad-Surprise).
- + Kèm theo đó khi mô hình nhận diện xong để trả ra nhãn cảm xúc (emotion label) và độ tin cậy của dự đoán (confidence score).

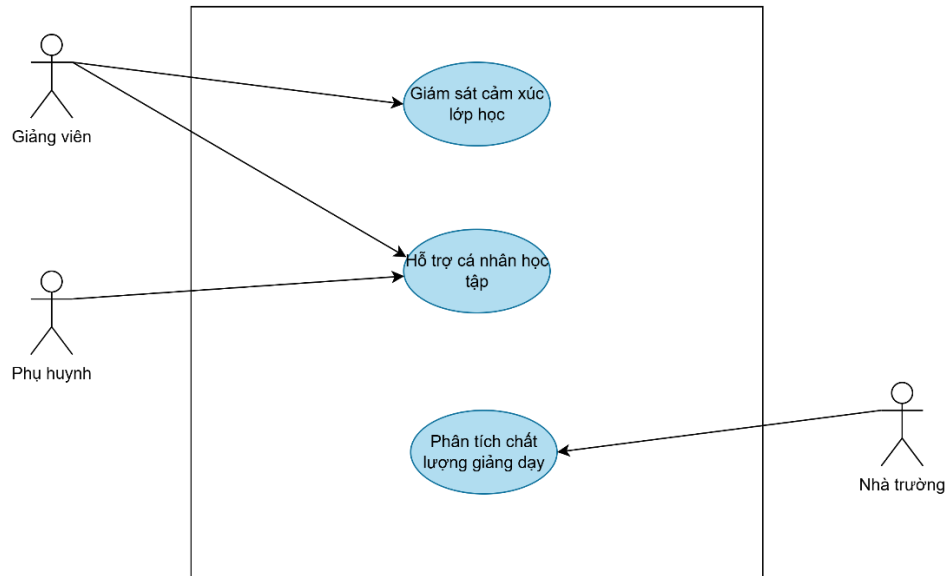
5. Hiển thị kết quả

- + Cuối cùng, hệ thống nhận diện sẽ trực quan hóa kết quả bao gồm vẽ bounding box quan khuôn mặt, hiển thị nhãn cảm xúc và độ tin cậy theo % cho từng khuôn mặt.

CHƯƠNG 3: PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG

3.1. PHÂN TÍCH HỆ THỐNG

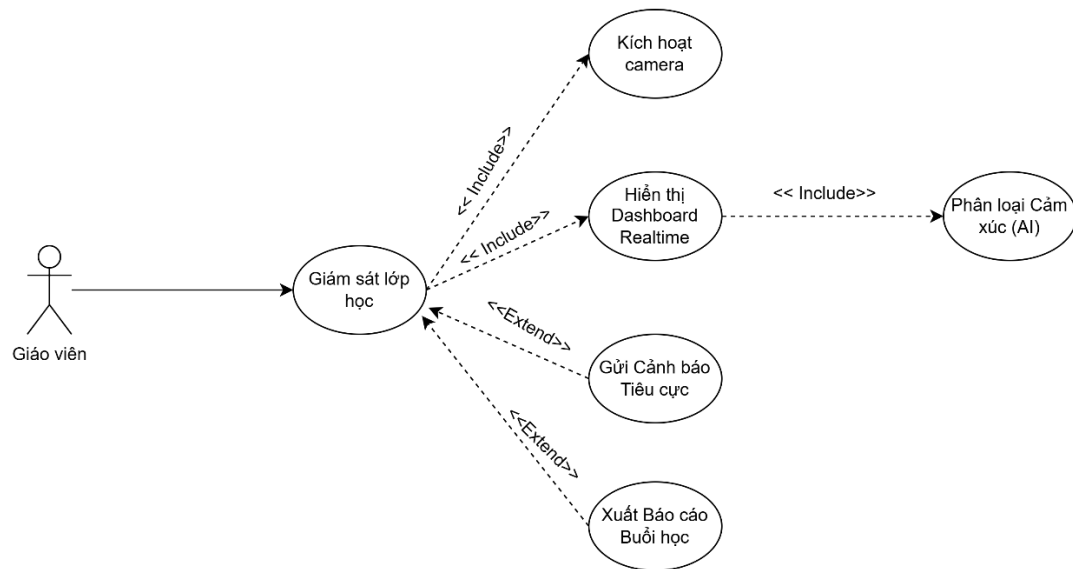
3.1.1. Sơ đồ Use Case nghiệp vụ



Hình 3. 1: Sơ đồ usecase nghiệp vụ.

3.1.2. Sơ đồ Use Case hệ thống và đặc tả: Chi tiết hóa các use case cho Web App và Mobile App

3.1.2.1. Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ Giám sát Cảm xúc Lớp học

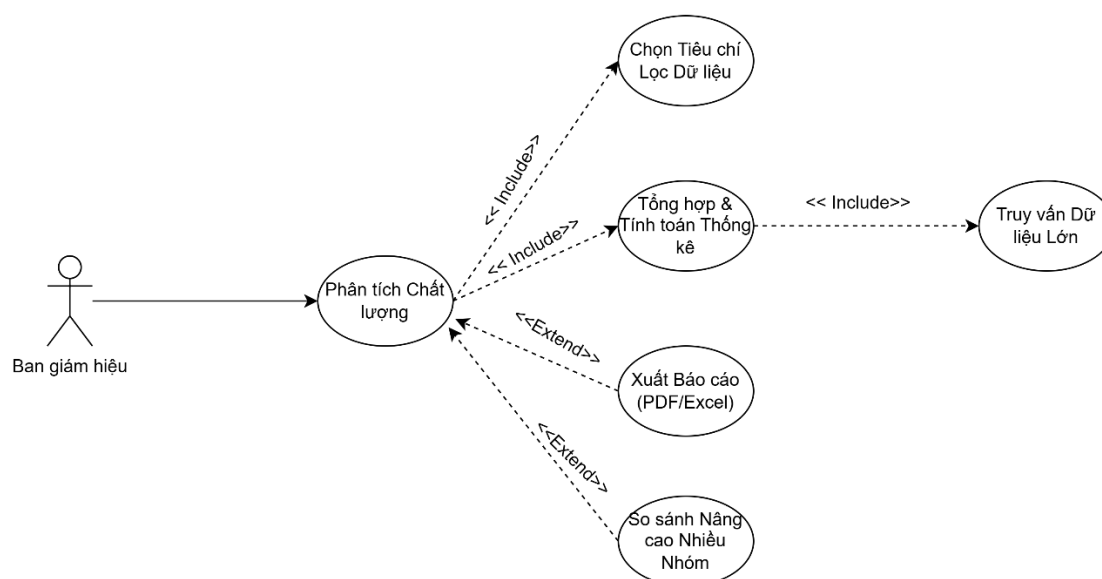


Hình 3. 2: Sơ đồ Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ Giám sát Cảm xúc Lớp học.

Bảng 3. 1: Đặc tả Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ Giám sát Cảm xúc Lớp học.

Thuộc tính	Nội dung
Tên Use Case	-Giám sát lớp học.
Mô tả	-Cho phép Giáo viên khởi tạo, theo dõi và kết thúc phiên giám sát cảm xúc của học sinh trong lớp học trên giao diện Web App.
Tác nhân chính	-Giáo viên
Điều kiện tiên quyết	-Giáo viên đã đăng nhập thành công và chọn lớp học/buổi học cần giám sát.
Kết quả thành công	-Dữ liệu cảm xúc của buổi học được ghi nhận đầy đủ, Giáo viên có cái nhìn toàn cảnh về tâm lý học sinh.
Luồng chính	<ol style="list-style-type: none"> 1. Giáo viên nhấn nút "Bắt đầu Giám sát". 2. Hệ thống thực hiện Include: UC-W01.1: Kích hoạt Camera. 3. Hệ thống thực hiện Include: UC-W01.2: Hiển thị Dashboard Realtime. 4. (Trong suốt phiên) Nếu điều kiện cảnh báo thỏa mãn, hệ thống thực hiện Extend: UC-W01.4: Gửi Cảnh báo Tiêu cực. 5. Giáo viên nhấn nút "Kết thúc Giám sát". 6. Hệ thống lưu trữ toàn bộ dữ liệu cảm xúc phiên học vào CSDL và giải phóng camera. 7. Hệ thống hỏi Giáo viên có muốn Extend: UC-W01.5: Xuất Báo cáo Buổi học không.
Luồng thay thế (Lỗi)	<p>-Dữ liệu cảm xúc của buổi học được lưu trữ.</p> <p>-Báo cáo buổi học có thể được xuất.</p>

3.1.2.2. Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ phân tích chất lượng giảng dạy



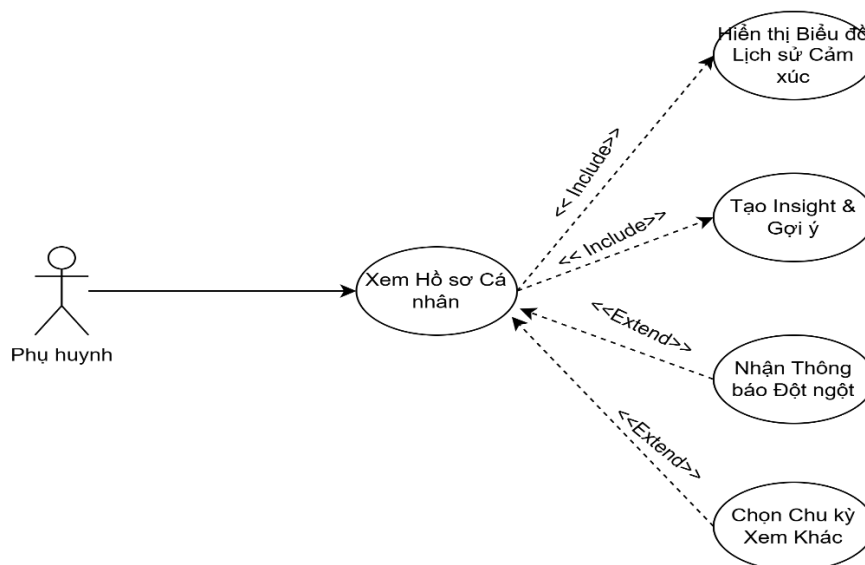
Hình 3. 3: Sơ đồ Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ phân tích chất lượng giảng dạy.

Bảng 3. 2: Đặc tả Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ phân tích chất lượng giảng dạy.

Thuộc tính	Nội dung
Thuộc tính	-Nội dung
Tên Use Case	-Hỗ trợ Cá nhân hóa Học tập
Mô tả	-Hệ thống xây dựng, duy trì và hiển thị hồ sơ cảm xúc cá nhân cho từng học sinh, cung cấp các gợi ý và báo cáo chuyên biệt cho giáo viên chủ nhiệm và phụ huynh.
Tác nhân chính	-Giáo viên chủ nhiệm / Phụ huynh.
Điều kiện tiên quyết	-Người dùng đã đăng nhập (giáo viên hoặc phụ huynh). Học sinh đã có ít nhất 3 buổi học được ghi nhận cảm xúc.

Kết quả thành công	-Hồ sơ cảm xúc của học sinh được cập nhật. Báo cáo cá nhân được tạo và hiển thị, Thông báo (nếu có) được gửi đến phụ huynh.
Luồng chính	<ol style="list-style-type: none"> 1. Tác nhân đăng nhập và chọn học sinh (hoặc học sinh được hiển thị mặc định nếu là phụ huynh). 2. Hệ thống hiển thị hồ sơ tổng quan (Emotion Score, trạng thái hiện tại). 3. Hệ thống thực hiện Include: Hiển thị Biểu đồ Lịch sử Cảm xúc. 4. Hệ thống thực hiện Include: Tạo Insight & Gợi ý. 5. (Thường xuyên/Định kỳ) Hệ thống kiểm tra điều kiện để thực hiện Extend: Nhận Thông báo Đột ngột. 6. Tác nhân có thể thực hiện Extend: Chọn Chu kỳ Xem Khác. 7. Tác nhân chọn Xuất/Lưu báo cáo.
Luồng thay thế (Lỗi)	-Học sinh mới: Nếu học sinh có ít hơn 3 buổi học, hệ thống chỉ hiển thị dữ liệu thô, không có insight hay gợi ý

3.1.2.3. Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ Hỗ trợ Cá nhân hóa Học tập.

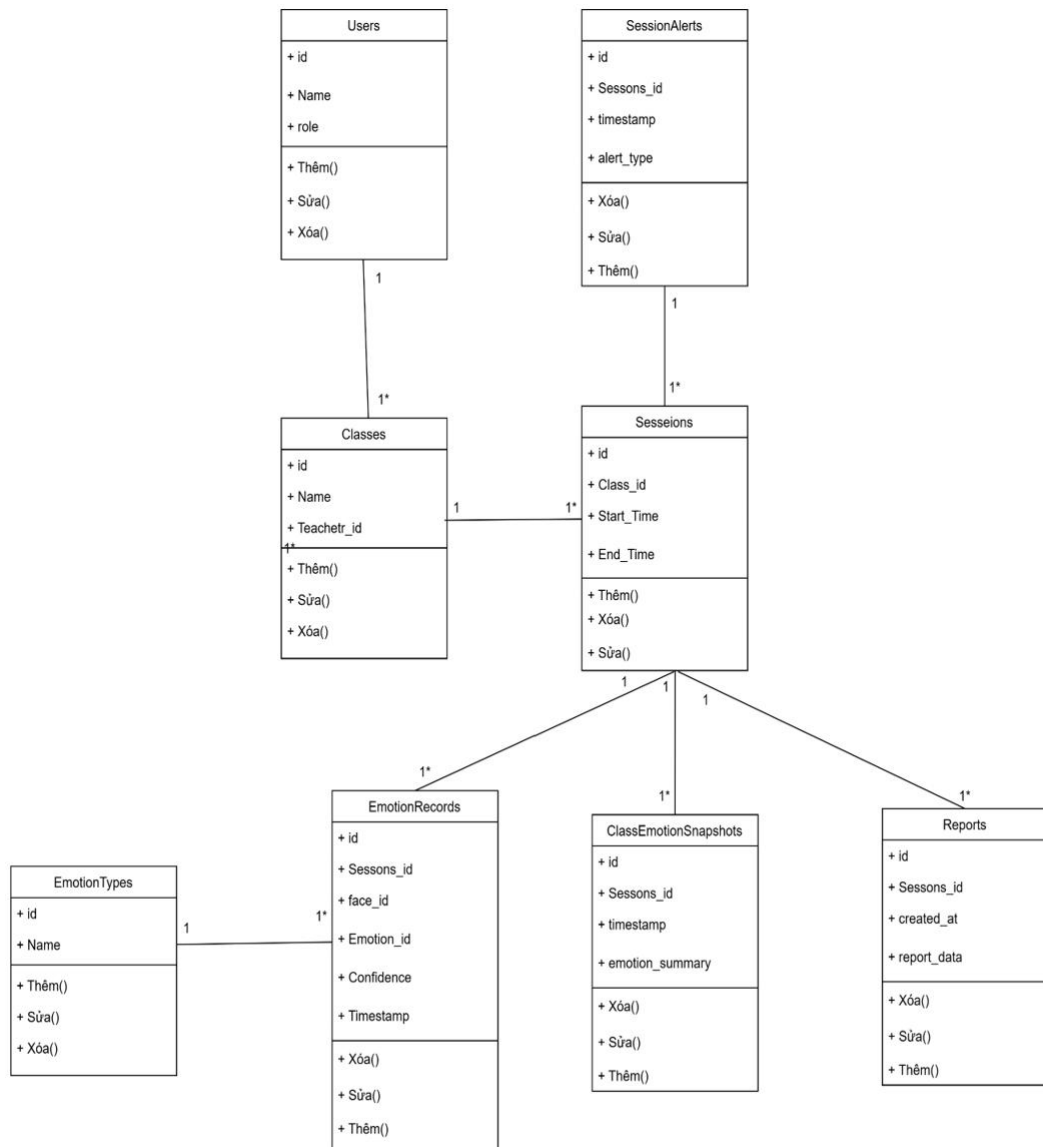


Hình 3. 4: Sơ đồ Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ Hỗ trợ Cá nhân hóa Học tập.

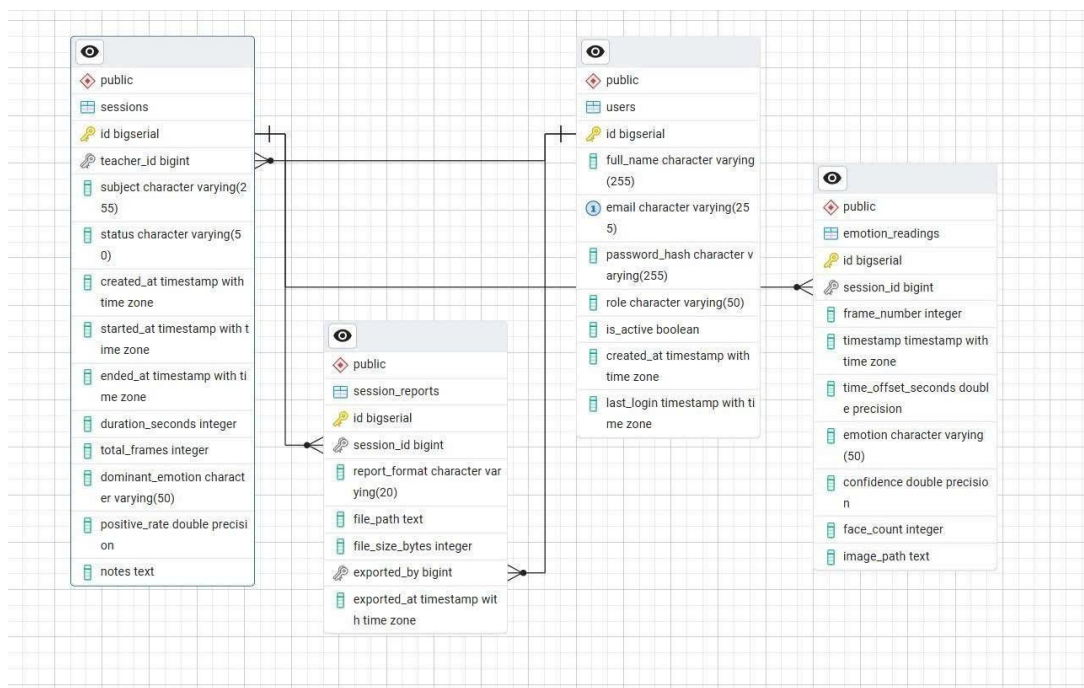
Bảng 3. 3: Đặc tả Use Case Hệ thống (Use Case Diagram) – Nghiệp vụ Hỗ trợ Cá nhân hóa Học tập.

Thuộc tính	Nội dung
Tên Use Case	-Xem Hồ sơ Cá nhân
Mô tả	-Phụ huynh/Giáo viên truy cập để xem hồ sơ cảm xúc, xu hướng lịch sử và nhận các gợi ý chuyên biệt về học sinh của mình.
Tác nhân chính	-Phụ huynh / Giáo viên Chủ nhiệm
<<Include>>	-Hiển thị Biểu đồ Lịch sử Cảm xúc; Tạo Insight & Gợi ý.
<<Extend>>	-Nhận Thông báo Đột ngột; Chọn Chu kỳ Xem Khác.
Điều kiện tiên quyết	-Người dùng đã đăng nhập. Học sinh đã có ít nhất 3 buổi học được ghi nhận cảm xúc.
Kết quả thành công	-Hồ sơ cảm xúc của học sinh được cập nhật, Báo cáo cá nhân được hiển thị đầy đủ, và người dùng nắm bắt được tâm lý học tập của học sinh.
Tên Use Case	-Xem Hồ sơ Cá nhân.

3.1.3. Sơ đồ lớp



Hình 3. 5: Sơ đồ lớp mức thiết kế.



Hình 3. 6: Sơ đồ lớp thiết kế và mô hình dữ liệu quan hệ cho PostgreSQL.

Bảng 3. 4: Mô tả các class có trong mô hình dữ liệu quan hệ cho PostgreSQL.

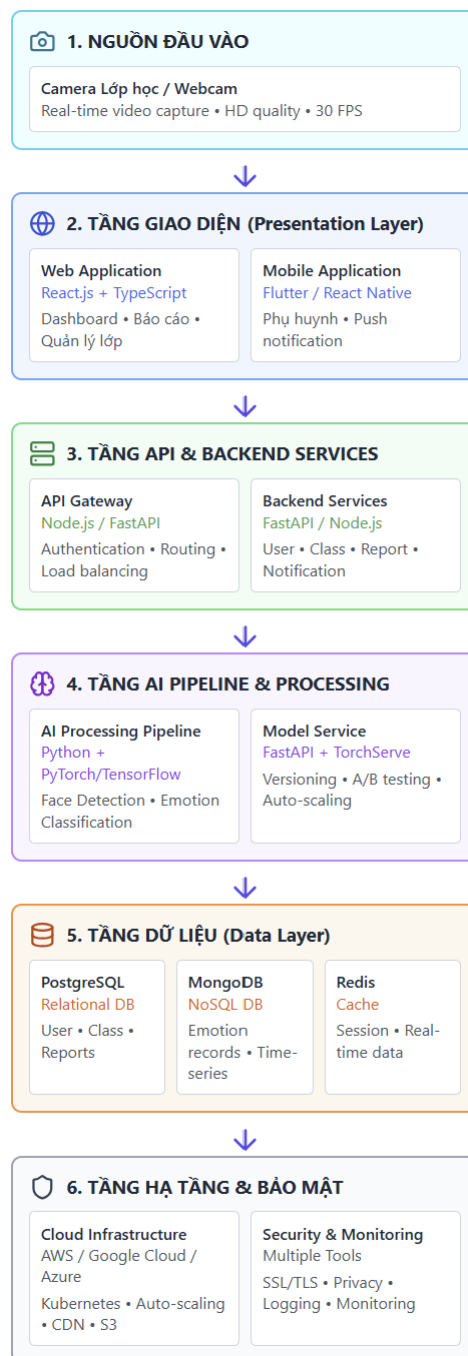
Tên Bảng (Lớp)	Thuộc tính Chính	Ý nghĩa & Mục đích	Vai trò trong hệ thống
Users	id, name, role	-Quản lý thông tin cơ bản của tất cả người dùng hệ thống (Giáo viên, Ban Giám hiệu, Admin).	-Xác định quyền truy cập (Giám sát, Phân tích, Quản trị) và vai trò của Giáo viên khi phụ trách lớp học.
Classes	id, name, teacher_id (FK)	-Lưu trữ thông tin về các lớp học được giám sát trong hệ thống.	-Tạo ra cấu trúc phân cấp (Giáo viên → Lớp học) để tổ chức dữ liệu buổi học.
Sessions	id, class_id (FK),	-Lưu trữ thông tin về từng buổi học	-Tạo ra bối cảnh cho dữ liệu cảm xúc. Tất cả các dữ liệu

	start_time, end_time	cụ thể đã diễn ra. Đây là đơn vị thời gian chính để phân tích.	cảm xúc, cảnh báo, báo cáo đều phải liên kết với một Session cụ thể.
EmotionTypes	id, name	-Định nghĩa danh sách các loại cảm xúc mà mô hình AI có thể nhận diện (ví dụ: Tập trung, Chán nản, Vui, Buồn, Giận dữ...).	-Chuẩn hóa dữ liệu cảm xúc, giúp các bảng khác chỉ cần lưu ID cảm xúc thay vì tên.
EmotionRecords	id, session_id (FK), face_id, emotion_id (FK), confidence, timestamp	-Bảng dữ liệu cốt lõi - Lưu trữ từng bản ghi cảm xúc thu thập được theo thời gian thực từ mô hình Deep Learning.	-Cung cấp dữ liệu thô (raw data) cho tất cả các hoạt động phân tích, báo cáo và giám sát thời gian thực.
ClassEmotionSnapshots	id, session_id (FK), timestamp, emotion_summary (JSONB)	-Lưu trữ bản tóm tắt cảm xúc của cả lớp tại các mốc thời gian định kỳ (ví dụ: mỗi 30 giây hoặc 1 phút). Dữ liệu ở dạng JSONB.	-Giảm tải cho hệ thống khi truy vấn hiển thị Dashboard Realtime, vì không cần phải tính toán lại tổng hợp từ EmotionRecords quá nhiều lần.
SessionAlerts	id, session_id (FK), timestamp, alert_type	-Lưu trữ thông tin khi có các sự kiện cảnh báo được kích hoạt (ví dụ:	-Ghi lại các sự kiện bất thường, phục vụ cho chức năng Gửi Cảnh báo Tiêu cực và xem lại lịch sử cảnh báo.

		Tỷ lệ chán nản vượt ngưỡng, phát hiện học sinh khác).	
Reports	id, session_id (FK), created_at, report_data (JSONB)	-Lưu trữ các báo cáo đã được tạo ra (bao gồm báo cáo buổi học, báo cáo cá nhân, báo cáo chất lượng). Dữ liệu báo cáo ở dạng JSONB.	-Lưu trữ kết quả phân tích cuối cùng, phục vụ cho chức năng Xuất Báo cáo và truy vấn các phân tích cấp cao.

3.2. THIẾT KẾ HỆ THỐNG

3.2.1. Thiết kế kiến trúc



Hình 3. 7: Sơ đồ kiến trúc hệ thống Giám sát và Phân tích Cảm xúc trong Giáo dục Thông minh.

3.2.2. Thiết kế giao diện: Thiết kế các giao diện chính cho Web App

- Dựa trên các yêu cầu nghiệp vụ đã được phân tích, nhóm đã thiết kế một hệ thống giao diện web hoàn chỉnh với các màn hình chính phục vụ cho ba nghiệp vụ cốt lõi. Các giao diện được thiết kế theo nguyên tắc UI/UX hiện đại, đảm bảo tính trực quan, dễ sử dụng và phù hợp với quy trình làm việc thực tế của giáo viên.

3.2.2.1. Giao diện đăng nhập (Login Interface)

- Màn hình đăng nhập là điểm tiếp xúc đầu tiên giữa người dùng và hệ thống, được thiết kế tối giản nhưng đầy đủ chức năng:
 - + Vị trí trung tâm: Form đăng nhập được đặt chính giữa màn hình với nền gradient nhẹ nhàng, tạo điểm nhấn cho người dùng tập trung vào thao tác chính.
 - + Các trường thông tin: Bao gồm trường "Tên người dùng / Email" và "Mật khẩu", sử dụng placeholder rõ ràng để hướng dẫn người dùng nhập liệu. Trường mật khẩu có tính năng ẩn/hiện ký tự để tăng tính bảo mật và tiện lợi.
 - + Nút đăng nhập chính: Được thiết kế nổi bật với màu xanh dương, kích thước lớn và có hiệu ứng hover để người dùng dễ dàng nhận biết hành động chính.
 - + Quick Login: Cung cấp tùy chọn đăng nhập nhanh (không cần tài khoản) phục vụ cho mục đích demo và testing, được đặt ngay bên dưới nút đăng nhập chính với kiểu thiết kế secondary button.
 - + Liên kết phụ trợ: Link "Quên mật khẩu?" được đặt ở cuối form với màu xanh nhạt, không quá nổi bật nhưng vẫn dễ nhìn thấy khi cần thiết.
- Thiết kế này tuân theo các nguyên tắc của Nielsen Norman Group về form design, đảm bảo tỷ lệ chuyển đổi cao và giảm thiểu lỗi nhập liệu.

3.2.2.2. Giao diện Trang chủ (Home/Dashboard Interface)

- Trang chủ đóng vai trò là trung tâm điều khiển chính của hệ thống, được chia làm hai khu vực chính:
- Khu vực bên trái - Khởi tạo lớp học mới:
 - + Tab Navigation: Hệ thống cung cấp hai tab là "New Class" (Lớp mới) và "Continue Class" (Tiếp tục lớp), cho phép giáo viên linh hoạt chọn giữa việc tạo buổi học mới hoặc tiếp tục buổi học đang tạm dừng.
- Form nhập thông tin: Bao gồm các trường dữ liệu cần thiết:
 - + Teacher: Dropdown hiển thị danh sách giáo viên, với chế độ Quick Teacher mặc định cho mục đích demo
 - + Subject: Text input để nhập tên môn học
 - + Select Camera: Dropdown cho phép chọn thiết bị camera từ các camera có sẵn (Camera 0, Camera 1,...)
 - + Start Button: Nút "Start New Class" có màu xanh lá, kích thước full-width để thu hút sự chú ý và thể hiện rõ đây là hành động chính của màn hình.
- Khu vực bên phải - Giới thiệu hệ thống:
 - + About Section: Cung cấp mô tả ngắn gọn về mục đích của Smart Classroom, giúp người dùng mới hiểu được giá trị cốt lõi: "Enhance teaching effectiveness through real-time emotion monitoring".
 - + Key Features: Liệt kê tính năng nổi bật như "Real-time Emotion Detection" với mô tả chi tiết: "Analyzes facial expressions to determine emotional states during class."
 - + Privacy Notice: Phần thông báo về quyền riêng tư được đặt trong một card màu xanh nhạt, nhấn mạnh rằng mọi dữ liệu cảm xúc được xử lý cục bộ và chỉ phục vụ cho mục đích phản hồi giáo viên, không lưu trữ hay chia sẻ với bên thứ ba.
 - + Thiết kế này tuân theo model "F-shaped reading pattern", đặt các thông tin quan trọng nhất ở vị trí người dùng thường xem đầu tiên.

3.2.2.3. Giao diện Giám sát (Monitoring Interface)

- Đây là giao diện quan trọng nhất của hệ thống, nơi diễn ra nghiệp vụ UC-NV01 (Giám sát cảm xúc lớp học). Giao diện được chia làm hai phần:
- Khu vực bên trái - Video Display:
 - + Video Feed Area: Chiếm khoảng 60% chiều rộng màn hình, hiển thị feed video từ camera với kích thước đủ lớn để giáo viên có thể quan sát rõ ràng. Trạng thái mặc định hiển thị thông báo "Khởi động Webcam hoặc Tải Video" khi chưa bắt đầu.
 - + Control Buttons: Hai nút điều khiển chính được đặt phía dưới video area:
 - "Webcam" button (màu tím gradient): Kích hoạt chế độ giám sát real-time từ webcam
 - "Upload Video" button (màu xanh dương): Cho phép tải video có sẵn để phân tích
 - + Tab Selection: Thanh tab ngang phía trên gồm "Live Monitoring", "Upload Video", và "Analytics", cho phép chuyển đổi giữa các chế độ làm việc khác nhau.
- Khu vực bên phải - Real-time Analytics:
 - + Section Title: "Phân Tích Thời Gian Thực" với font chữ lớn và in đậm, rõ ràng phân tách khu vực dữ liệu.
 - + Initial State: Trong trạng thái chưa bắt đầu, hiển thị thông báo hướng dẫn: "Bắt đầu nhận diện để xem biểu đồ", giúp người dùng hiểu cần thực hiện hành động gì tiếp theo.
 - + Data Visualization Area: Khu vực này sẽ hiển thị các biểu đồ và thống kê realtime khi hệ thống bắt đầu phân tích (sẽ được mô tả chi tiết ở phần tiếp theo).
- Thiết kế layout theo tỷ lệ 60-40 này dựa trên nghiên cứu của Jakob Nielsen về "Visual Hierarchy", đảm bảo thông tin được ưu tiên hiển thị theo mức độ quan trọng.

3.2.2.4. Giao diện Phân tích Lớp học (Class Analytics Interface)

- Giao diện này phục vụ cho nghiệp vụ UC-NV02 (Phân tích chất lượng giảng dạy), được thiết kế để trình bày dữ liệu phân tích một cách toàn diện:
- Header Section:
 - + Breadcrumb Navigation: "Home / Class / Analytics" giúp người dùng định vị vị trí hiện tại trong hệ thống.
 - + Class Information Bar: Hiển thị thông tin buổi học với background màu xanh:
 - Tên môn học (Demo Subject)
 - Teacher ID (T123)
 - Timestamp realtime (11/26/2025, 9:39:53 PM)
 - + Action Buttons: Hai nút "Attendance" và "Back to Class" được đặt bên phải, cho phép giáo viên truy cập nhanh các chức năng liên quan.
- Metrics Overview (KPI Cards):
 - + Bốn card metric được sắp xếp theo hàng ngang, mỗi card hiển thị một chỉ số quan trọng:
 1. Total Readings: 120 data points - Tổng số lần thu thập dữ liệu cảm xúc
 2. Average Students: 28.5 per capture - Số lượng học sinh trung bình được nhận diện mỗi lần
 3. Dominant Emotion: Happy (màu vàng) - Cảm xúc chủ đạo trong buổi học
 4. Class Duration: 45 minutes - Thời lượng buổi học
- Visualization Section:
 - + Hai biểu đồ chính được đặt song song:

1. Emotion Distribution (Biểu đồ phân bố cảm xúc): Sử dụng horizontal bar chart để so sánh tỷ lệ các cảm xúc, với legend "Emotion Count" và trục dọc đánh số từ 0 đến 1.
 2. Emotion Over Time (Biểu đồ xu hướng cảm xúc): Line chart theo trục thời gian, với các mức phần trăm từ 0% đến 100% trên trục dọc, cho phép theo dõi sự thay đổi cảm xúc trong suốt buổi học.
- + Teaching Insights Section:
- Phần này cung cấp các gợi ý và phân tích dựa trên dữ liệu cảm xúc, được đặt trong một card có background màu vàng nhạt, thu hút sự chú ý của giáo viên vào các điểm cần lưu ý.
 - Thiết kế dashboard này tuân theo "Data-Ink Ratio" principle của Edward Tufte, tối ưu hóa việc truyền tải thông tin với lượng "mực in" tối thiểu.

CHƯƠNG 4: CÀI ĐẶT HỆ THỐNG

4.1. CÁC YÊU CẦU CÀI ĐẶT ỨNG DỤNG

- Hệ thống này yêu cầu:
 - + Phần cứng: máy chủ, GPU (nếu có) và camera lớp học (cấu hình tối thiểu)
 - + Phần mềm: OS, Python bản 3.11, framework web, hệ quản trị CSDL
- Các thành phần chính cần cài đặt:
 - + Môi trường AI: Python phiên bản 3.11, thư viện Ultralytics phiên bản ≥ 8.3 (phù hợp cho mô hình YOLOv11), OpenCV cho việc xử lý ảnh
 - + API Backend: sử dụng framework FastAPI tiện lợi cho ngôn ngữ Python
 - + FrontEnd: vì có tính chất theo thời gian thực nên sử dụng framework ReactJs
 - + CSDL: sử dụng PostgreSQL

4.1.1. Cài đặt môi trường và AI PIPELINE

Bước 1: Chuẩn bị dữ liệu

- Nguồn dữ liệu
 - + Dữ liệu được sử dụng là Bộ dữ liệu cảm xúc được hợp nhất (A merged emotions dataset) hay còn gọi là "Facial Expression Image Data – AFFECTNET YOLO Format". Bộ dữ liệu này được tạo ra từ một tập con được tuyển chọn kỹ lưỡng của các dataset lớn hơn như ExpW, FER2013 (kèm FER2013+), AffectNet (6 cảm xúc), và RAF-DB, với tổng cộng khoảng 155K mẫu. Dữ liệu gốc của AffectNet thu thập hơn một triệu ảnh khuôn mặt, trong đó khoảng 450.000 ảnh được gán nhãn thủ công.
- Link dataset: <https://huggingface.co/datasets/AdamCodd/yolo-emotions/tree/main>
- Lớp cảm xúc
 - + Bộ dữ liệu bao gồm 7 loại cảm xúc rời rạc:

- Class 0: Angry (Tức giận)
- Class 1: Disgust (Ghê tởm)
- Class 2: Fear (Sợ hãi)
- Class 3: Happy (Vui vẻ)
- Class 4: Neutral (Trung tính)
- Class 5: Sad (Buồn bã)
- Class 6: Surprise (Ngạc nhiên)

Lưu ý: Cảm xúc 'contempt' (Khinh bỉ) từ AffectNet đã bị loại bỏ để tránh làm tăng sự mất cân bằng lớp (class imbalance) và vì nó không có trong các dataset khác.

- Phân phối dữ liệu
 - + Dữ liệu được chia thành 3 tập: TRAIN (Huấn luyện), VAL (Kiểm định), và TEST (Kiểm tra). Tồn tại sự mất cân bằng đáng kể giữa các lớp, với các lớp chiếm ưu thế là Happy (Vui vẻ) và Neutral (Trung tính) chiếm khoảng 30-31% tổng số mẫu, trong khi các lớp thiểu số như Disgust (Ghê tởm) và Fear (Sợ hãi) chỉ chiếm khoảng 3-5%.

Bảng 4. 1: Bảng phân phối dữ liệu.

Tập Dữ liệu	Angry (0)	Disgust (1)	Fear (2)	Happy (3)	Neutral (4)	Sad (5)	Surprise (6)
TRAIN	6.84%	5.07%	3.41%	30.30%	31.57%	12.70%	10.12%
TOTAL	6.86%	5.07%	3.43%	30.29%	31.56%	12.67%	10.12%

- Tiền xử lý dữ liệu
 - + Định dạng: Dữ liệu được xử lý lại ở định dạng YOLO để sử dụng trực tiếp với các mô hình YOLO
 - + Kích thước ảnh: Toàn bộ ảnh được resize về kích thước 96x96 pixel
 - + Gán nhãn YOLO: Mỗi ảnh có một file .txt tương ứng chứa thông tin bounding box theo định dạng: class_id x_center y_center width height

Bước 2: Huấn luyện mô hình

- Mô hình: YOLOv11-x hoặc yolov11n.pt (theo cấu hình huấn luyện)
 - + Khởi tạo: Mô hình được khởi tạo từ trọng số pre-trained và sau đó fine-tune trên bộ dữ liệu cảm xúc.
 - + Bài toán: Object Detection (Phát hiện đối tượng) với đối tượng là khuôn mặt có cảm xúc.
 - + Cấu trúc CLI: Sử dụng cú pháp CLI của thư viện Ultralytics YOLO (yolo task=detect).
 - + Tham số huấn luyện
 - + Có hai bộ tham số huấn luyện khác nhau tùy thuộc vào tài nguyên sử dụng:
 - + Tham số Huấn luyện Mở rộng (phù hợp với tài nguyên máy khá tốt):
 - o Epochs: 50
 - o Image size: 415x416 pixel
 - o Batch size: 8
 - o Optimizer: SGD (Stochastic Gradient Descent)
 - o Initial Learning Rate: 0.001
 - + Tham số Huấn luyện Đơn giản (Tham số huấn luyện nhóm sử dụng):
 - o Epochs: 30-50
 - o Image size (imgsz): 256x256 pixel
 - o Batch size: 16
 - o Optimizer: SGD
 - o Initial Learning Rate (lr0): 0.001
- Tăng cường dữ liệu
 - + Vì bộ dữ liệu lớn nên không cần sử dụng nhiều kỹ thuật tăng cường dữ liệu. Chỉ áp dụng:
 - o Mosaic (mosaic): 0.4 - ghép 4 ảnh thành 1

- Flip Left/Right (fliplr): 0.5 - lật ảnh trái-phải

4.1.2. xây dựng api hệ thống và cài đặt các nghiệp vụ chính

4.1.2.1. Cài đặt Nghiệp vụ 1 - Giám sát cảm xúc: Phát triển

dashboard, chức năng cảnh báo, xuất báo cáo

- Nghiệp vụ này cho phép giáo viên theo dõi và phân tích cảm xúc học sinh theo thời gian thực thông qua camera, đồng thời đưa ra cảnh báo kịp thời và hỗ trợ xuất báo cáo sau buổi học.
- Dashboard Giám sát Thông minh
 - + Giao diện dashboard được thiết kế trực quan với ba chức năng cốt lõi. Thứ nhất là hiển thị luồng video trực tiếp từ webcam với nhiều chế độ xem (thu nhỏ, phóng to, toàn màn hình) và khả năng tự động kết nối lại khi bị gián đoạn. Thứ hai, hệ thống phân tích cảm xúc mỗi 300-500ms, trả về kết quả bao gồm cảm xúc chủ đạo, tỷ lệ tích cực, biểu đồ phân bố, số lượng học sinh và độ tin cậy của mô hình. Thứ ba, các biểu đồ được cập nhật liên tục: biểu đồ tròn thể hiện phân bố cảm xúc hiện tại, biểu đồ cột theo dõi số người trong lớp, và timeline ghi nhận diễn biến cảm xúc theo từng thời điểm.
- Hệ thống Cảnh báo Thông minh
 - + Chức năng cảnh báo hoạt động dựa trên chỉ số Bored Rate, được tính bằng tổng tỷ lệ các cảm xúc tiêu cực (buồn, tức giận, chán ghét, sợ hãi) trên tổng số frame. Khi Bored Rate đạt hoặc vượt 40%, hệ thống tự động kích hoạt cảnh báo thông qua popup ở góc

màn hình, thay đổi màu sắc trạng thái lớp học, và hiển thị các gợi ý cụ thể giúp giáo viên điều chỉnh phương pháp giảng dạy kịp thời.

- Báo cáo Buổi học Chi tiết
 - + Sau mỗi buổi học, giáo viên có thể xuất báo cáo tổng hợp bao gồm thông tin cơ bản (môn học, giáo viên, thời gian), cảm xúc chủ đạo, các biểu đồ phân bố và timeline theo phút, tỷ lệ tích cực-tiêu cực, cùng với nhận xét tự động từ hệ thống và ghi chú của giáo viên. Báo cáo này không chỉ giúp đánh giá chất lượng giảng dạy mà còn hỗ trợ nhà trường theo dõi quá trình học tập và lưu trữ bằng chứng để cải thiện phương pháp dạy.

4.1.2.2. Cài đặt Nghiệp vụ 2 - Phân tích chất lượng giảng dạy: Xây dựng các chức năng tổng hợp, so sánh và báo cáo cho nhà trường

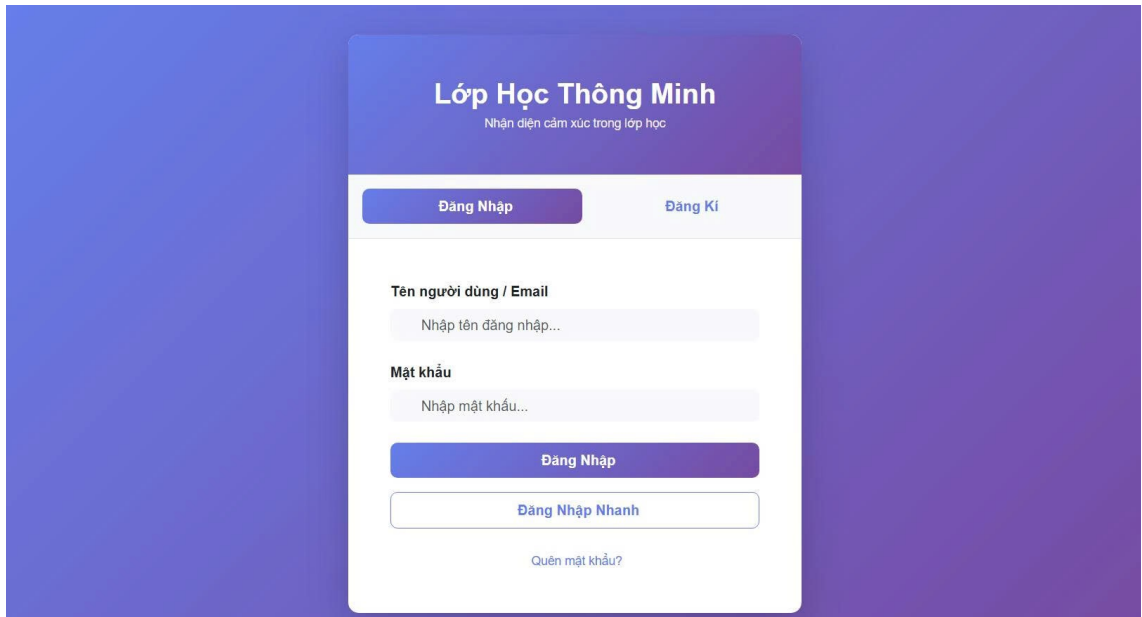
- Nghiệp vụ này khai thác dữ liệu từ Giám sát Cảm xúc để tổng hợp, so sánh và tạo ra các báo cáo chuyên sâu phục vụ công tác quản lý của nhà trường.
- Tổng hợp Dữ liệu Đa chiều
 - + Hệ thống tổng hợp dữ liệu theo ba cấp độ. Cấp độ buổi học bao gồm tổng số bản ghi cảm xúc, cảm xúc chủ đạo, mức độ tích cực, số lượng học sinh trung bình và thời lượng. Cấp độ giáo viên tập trung vào mức độ hứng thú trung bình mà lớp thể hiện, số buổi học tích cực/tiêu cực, và xu hướng chất lượng theo thời gian. Cấp độ môn học cho phép so sánh mức độ cảm xúc giữa các môn, xác

định môn để gây chán nản hay tạo hứng thú cao, và đánh giá hiệu quả giảng dạy theo nội dung bài học.

- Phân tích So sánh Toàn diện
 - + Chức năng so sánh hỗ trợ nhà trường đưa ra quyết định chính xác thông qua các góc nhìn khác nhau. So sánh giữa các buổi học giúp xác định những buổi bất thường hoặc có xu hướng giảm chất lượng. So sánh giữa các giáo viên đánh giá mức độ hứng thú và hiệu quả giảng dạy theo chỉ số cảm xúc. So sánh theo môn học chỉ ra môn nào có Bored Rate cao nhất và môn nào duy trì mức độ tích cực ổn định.
- Báo cáo Quản trị Chuyên sâu
 - + Hệ thống cung cấp đa dạng loại báo cáo theo tuần/tháng, theo giáo viên, theo lớp học hoặc môn học, cùng với báo cáo so sánh kết quả các buổi dạy. Nội dung báo cáo bao gồm chỉ số chất lượng giảng dạy tổng hợp, biểu đồ cảm xúc theo thời gian, xếp hạng giáo viên theo mức độ tích cực của lớp, và phát hiện các buổi học bất thường. Điều này mang lại ba lợi ích lớn cho nhà trường: giám sát chất lượng giảng dạy minh bạch và khách quan, hỗ trợ xây dựng kế hoạch cải thiện phương pháp dạy, và cung cấp dữ liệu định lượng cho đánh giá thi đua chất lượng giáo viên.

4.2. GIỚI THIỆU GIAO DIỆN CHƯƠNG TRÌNH

4.2.1. Màn hình Đăng nhập (Login Screen)

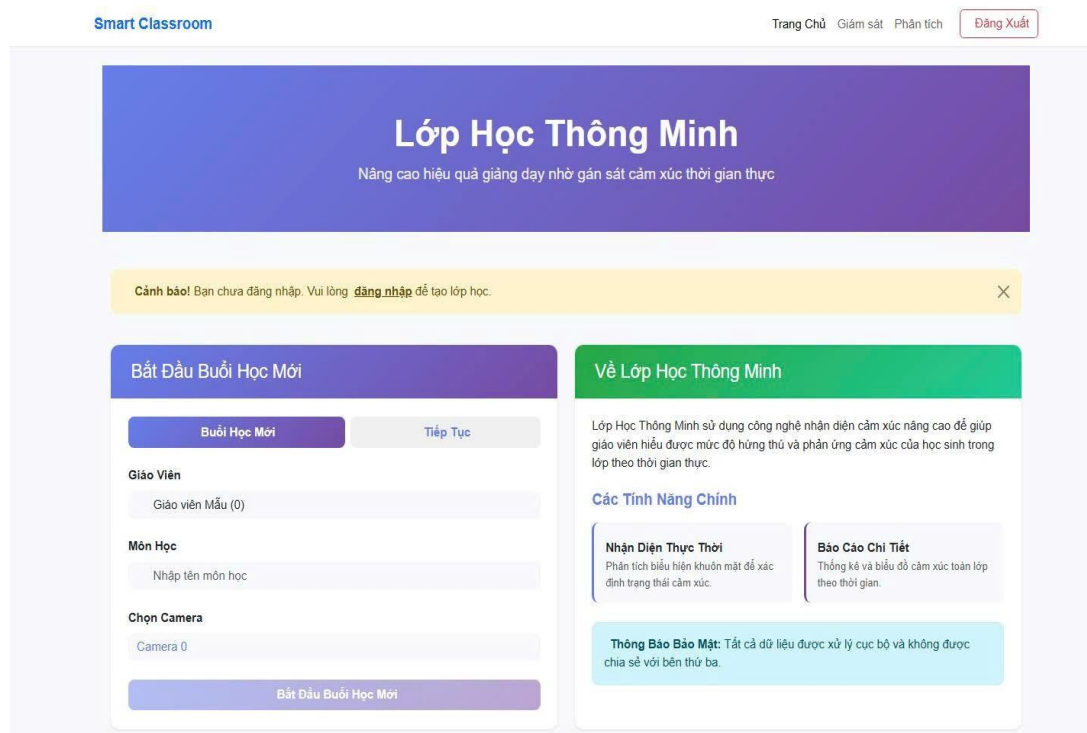


The screenshot shows a login interface for 'Lớp Học Thông Minh'. At the top, the title 'Lớp Học Thông Minh' is displayed in white on a purple gradient background, with the subtitle 'Nhận diện cảm xúc trong lớp học' below it. Below the title, there are two buttons: 'Đăng Nhập' (Login) in a dark purple box and 'Đăng Kí' (Register) in a light purple box. The main form area is white and contains two input fields: 'Tên người dùng / Email' (Username / Email) and 'Mật khẩu' (Password). Both fields have placeholder text 'Nhập tên đăng nhập...' and 'Nhập mật khẩu...' respectively. Below the password field is a 'Đăng Nhập' button in a dark purple box. Underneath that is a 'Đăng Nhập Nhanh' (Quick Login) button in a light purple box. At the bottom of the form is a link 'Quên mật khẩu?' (Forgot password?).

Hình 4. 1: Giao diện màn hình đăng nhập

- Màn hình đăng nhập là điểm truy cập đầu tiên vào hệ thống. Giao diện được thiết kế tối giản với form đăng nhập đặt ở vị trí trung tâm, bao gồm hai trường nhập liệu chính là "Tên người dùng / Email" và "Mật khẩu".
- Người dùng có hai cách để truy cập hệ thống:
 - + Đăng nhập thông thường: Nhập thông tin tài khoản đã đăng ký và nhấn nút "Đăng Nhập"
 - + Quick Login: Sử dụng tính năng đăng nhập nhanh không cần tài khoản, phục vụ cho mục đích demo và thử nghiệm
- Hệ thống cũng cung cấp liên kết "Quên mật khẩu?" để hỗ trợ người dùng trong trường hợp cần khôi phục tài khoản.

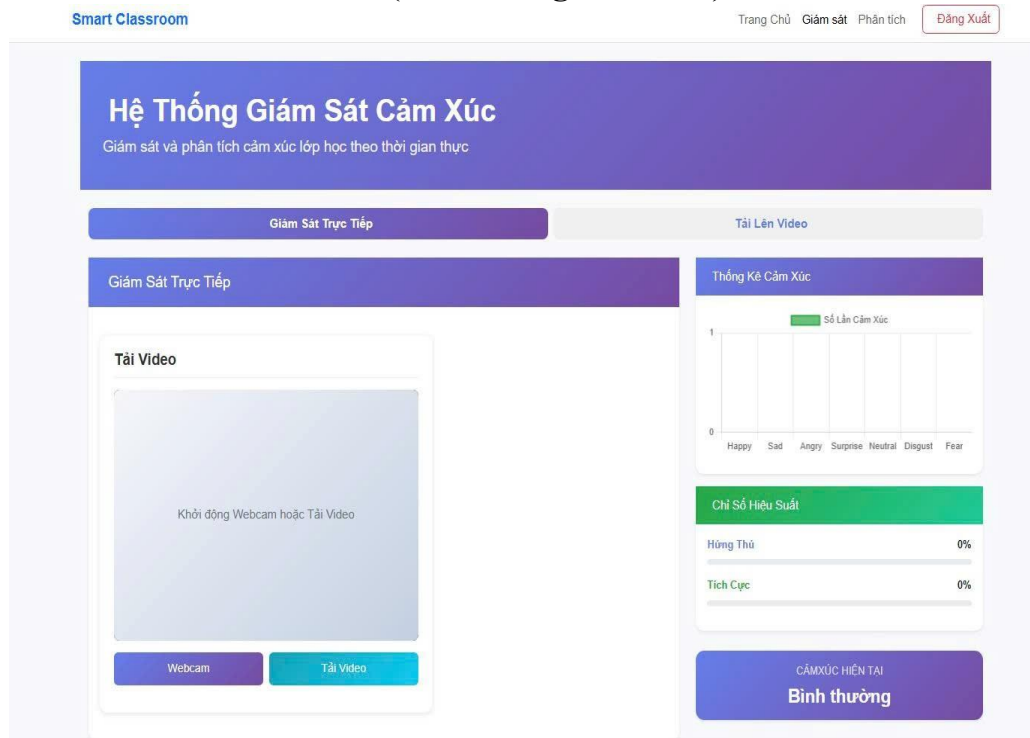
4.2.2. Màn hình Trang chủ (Home Dashboard)



Hình 4. 2: Giao diện trang chủ

- Sau khi đăng nhập thành công, người dùng được chuyển đến trang chủ - nơi khởi tạo các buổi học mới. Giao diện được chia làm hai phần chính:
- Phần bên trái - Khởi tạo buổi học:
 - + Cung cấp hai tab: "New Class" để tạo buổi học mới và "Continue Class" để tiếp tục buổi học đã lưu
 - + Form nhập thông tin buổi học bao gồm: chọn giáo viên, nhập tên môn học, và chọn camera sử dụng
 - + Nút "Start New Class" để bắt đầu phiên giám sát
- Phần bên phải - Giới thiệu hệ thống:
 - + Mô tả về mục đích và giá trị của Smart Classroom
 - + Liệt kê tính năng nổi bật như "Real-time Emotion Detection"
 - + Thông báo về chính sách bảo mật dữ liệu, cam kết rằng mọi dữ liệu cảm xúc được xử lý cục bộ và chỉ phục vụ mục đích phản hồi cho giáo viên

4.2.3. Màn hình Giám sát (Monitoring Interface)

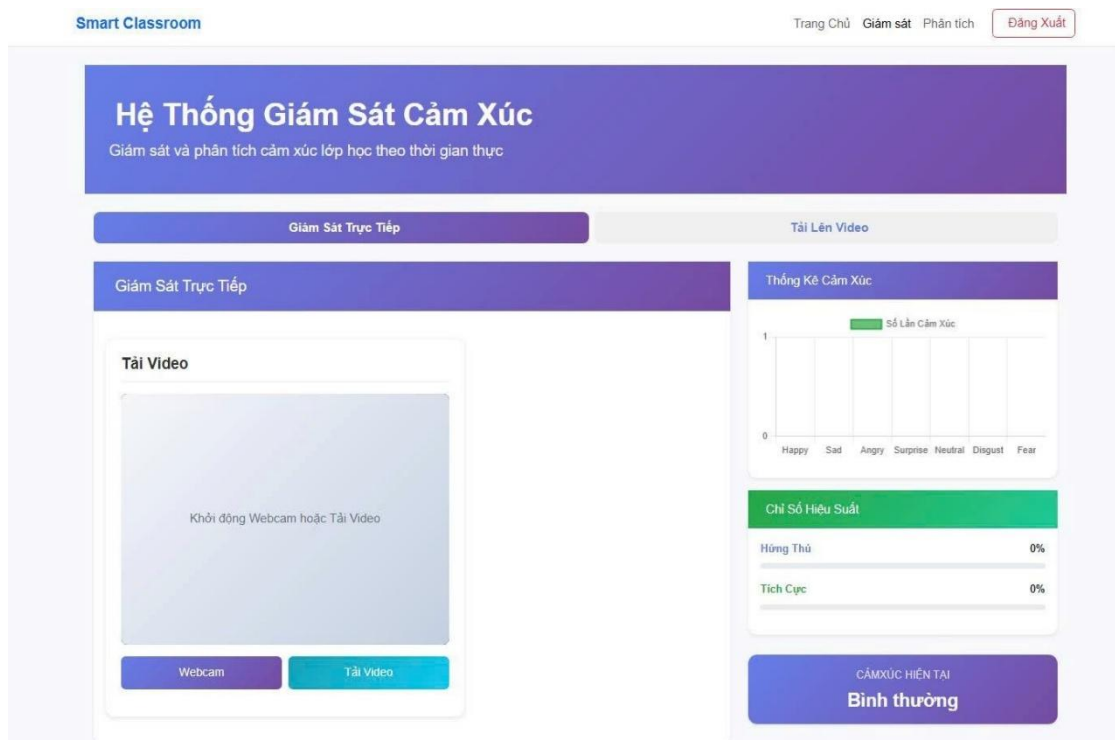


Hình 4. 3: Giao diện giám sát cảm xúc

- Đây là màn hình trung tâm của hệ thống, nơi diễn ra quá trình giám sát cảm xúc học sinh theo thời gian thực. Giao diện được chia thành hai khu vực:
- Khu vực bên trái - Hiển thị video:
 - + Thanh tab điều hướng gồm "Live Monitoring", "Upload Video" và "Analytics"
 - + Vùng hiển thị video từ webcam hoặc video đã tải lên
 - + Hai nút điều khiển: "Webcam" (màu tím) để kích hoạt camera trực tiếp và "Upload Video" (màu xanh) để phân tích video có sẵn
- Khu vực bên phải - Phân tích thời gian thực:
 - + Hiển thị tiêu đề "Phân Tích Thời Gian Thực"
 - + Trong trạng thái chưa bắt đầu, hiển thị hướng dẫn: "Bắt đầu nhận diện để xem biểu đồ"

- + Khi hệ thống đang hoạt động, khu vực này sẽ hiển thị các biểu đồ cảm xúc realtime, bao gồm biểu đồ phân bố cảm xúc và biểu đồ xu hướng theo thời gian
- Trong quá trình giám sát, hệ thống tự động nhận diện khuôn mặt, phân loại cảm xúc và cập nhật các biểu đồ phân tích mỗi 30 giây. Nếu phát hiện tỷ lệ cảm xúc tiêu cực vượt ngưỡng, hệ thống sẽ tự động gửi cảnh báo cho giáo viên.

4.2.4. Màn hình Phân tích Lớp học (Class Analytics)

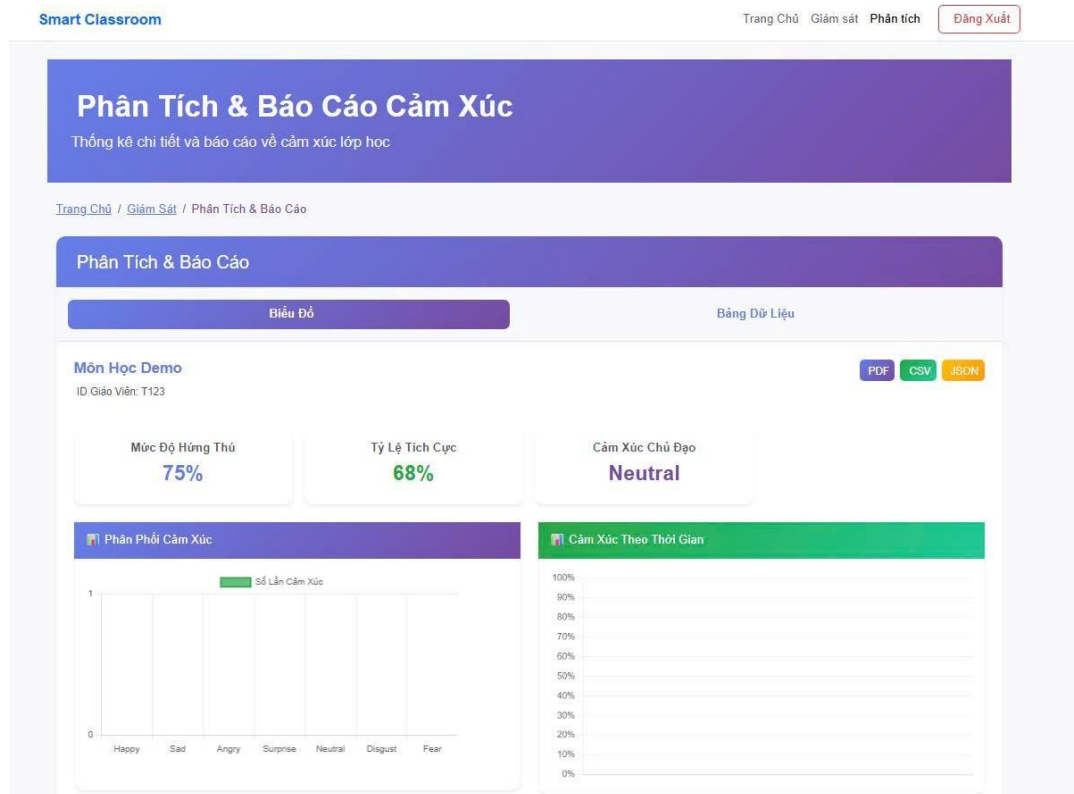


Hình 4. 4: Giao diện phân tích lớp học

- Màn hình này cung cấp cái nhìn tổng quan và chi tiết về dữ liệu cảm xúc của một buổi học đã hoàn thành. Giao diện bao gồm các thành phần:
- Header bar:
 - + Hiển thị thông tin buổi học: tên môn học, mã giáo viên và timestamp
 - + Các nút chức năng: "Attendance" để xem danh sách điểm danh và "Back to Class" để quay lại

- Các chỉ số tổng quan (KPI Cards):
 - + Total Readings: Tổng số lần thu thập dữ liệu cảm xúc
 - + Average Students: Số lượng học sinh trung bình mỗi lần capture
 - + Dominant Emotion: Cảm xúc chiếm ưu thế trong buổi học (hiển thị bằng badge màu)
 - + Class Duration: Tổng thời lượng buổi học
- Các biểu đồ phân tích:
 - + Emotion Distribution: Biểu đồ cột ngang thể hiện phân bố số lượng của từng loại cảm xúc
 - + Emotion Over Time: Biểu đồ đường thể hiện xu hướng thay đổi cảm xúc theo thời gian, với trục dọc biểu thị phần trăm và trục ngang biểu thị timeline
- Teaching Insights:
 - + Phân gợi ý và nhận xét tự động dựa trên dữ liệu phân tích, được hiển thị trong card màu vàng nhạt
 - + Cung cấp các insight giúp giáo viên cải thiện phương pháp giảng dạy

4.2.5. Màn hình Báo cáo (Report List)



Hình 4. 5: Giao diện danh sách báo cáo

- Màn hình cuối cùng hiển thị danh sách tổng hợp tất cả các buổi học đã diễn ra. Giao diện bao gồm:
- Header section:
 - + Tiêu đề "Class Reports" và mô tả ngắn gọn về chức năng
 - + Nút "Export CSV" ở góc phải phía trên để xuất toàn bộ dữ liệu ra file
 - + Bảng danh sách báo cáo:
 - + Hiển thị thông tin các buổi học theo dạng bảng với các cột:
 - o Date: Ngày diễn ra buổi học
 - o Subject: Tên môn học
 - o Teacher: Mã giáo viên phụ trách
 - o Dominant Emotion: Cảm xúc chủ đạo (hiển thị bằng badge màu, ví dụ "Happy" màu vàng)

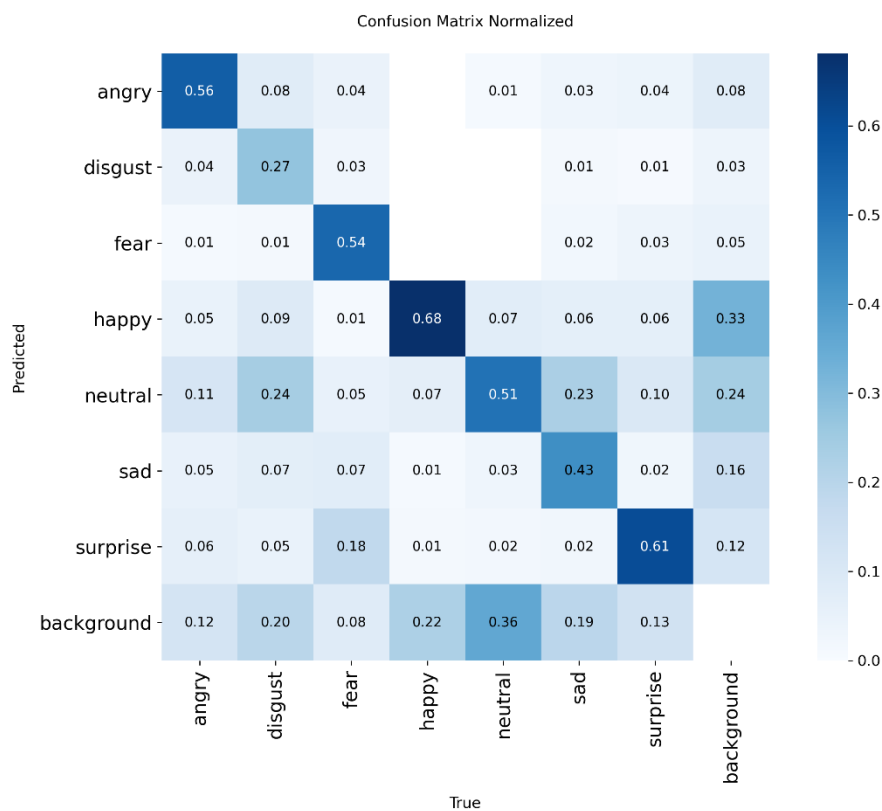
- Avg. Students: Số lượng học sinh trung bình
 - Duration: Thời lượng buổi học
 - Action: Nút "View" để xem chi tiết báo cáo
- Giáo viên có thể click vào nút "View" ở bất kỳ dòng nào để xem phân tích chi tiết của buổi học tương ứng. Chức năng Export CSV cho phép xuất toàn bộ dữ liệu để thực hiện phân tích sâu hơn bằng các công cụ như Excel hoặc Google Sheets.

CHƯƠNG 5: THỬ NGHIỆM VÀ ĐÁNH GIÁ

- Mô hình YOLO11 được huấn luyện trên bộ dữ liệu gồm 7 lớp cảm xúc: angry, disgust, fear, happy, neutral, sad, và surprise. Sau quá trình huấn luyện 30 epochs, mô hình được đánh giá toàn diện thông qua các biểu đồ và chỉ số đánh giá bao gồm: F1-score, Precision, Recall, PR Curve, Confusion Matrix và các chỉ số loss.

5.1. ĐÁNH GIÁ HIỆU NĂNG MÔ HÌNH AI

5.1.1 Ma trận nhầm lẫn

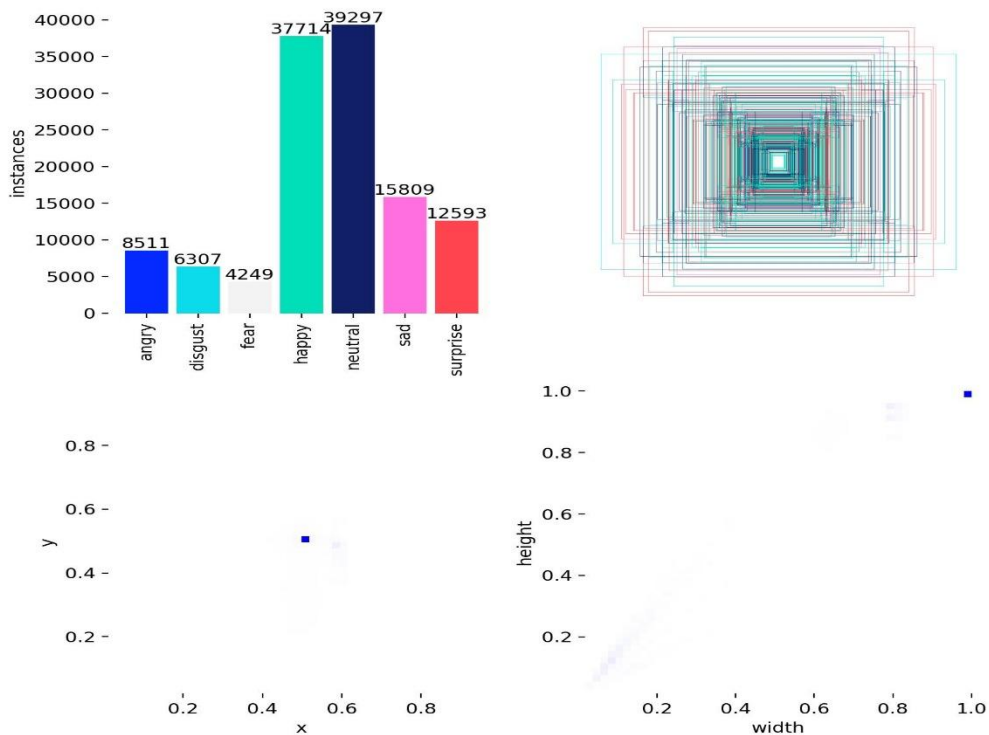


Hình 5. 1: Ma trận nhầm lẫn chuẩn hóa (Confusion Matrix Normalized).

- Ý nghĩa: Ma trận nhầm lẫn chuẩn hóa cho biết tỷ lệ mô hình dự đoán đúng và nhầm lẫn giữa các lớp cảm xúc.
- Điểm mạnh:
 - + Lớp **"happy"** đạt tỉ lệ dự đoán đúng cao nhất (0.68).

- + Các lớp **"surprise"**, **"angry"**, **"fear"** được nhận diện khá rõ ràng.
- Điểm yếu:
 - + Lớp **"disgust"** chỉ đạt 0.27, thường bị nhầm lẫn thành neutral hoặc sad.
 - + Lớp **"neutral"** dễ bị nhầm với nhiều lớp cảm xúc khác.
- Kết luận: Mô hình phân loại tốt ở 4-5 cảm xúc chính, tuy nhiên disgust và neutral cần được cải thiện.

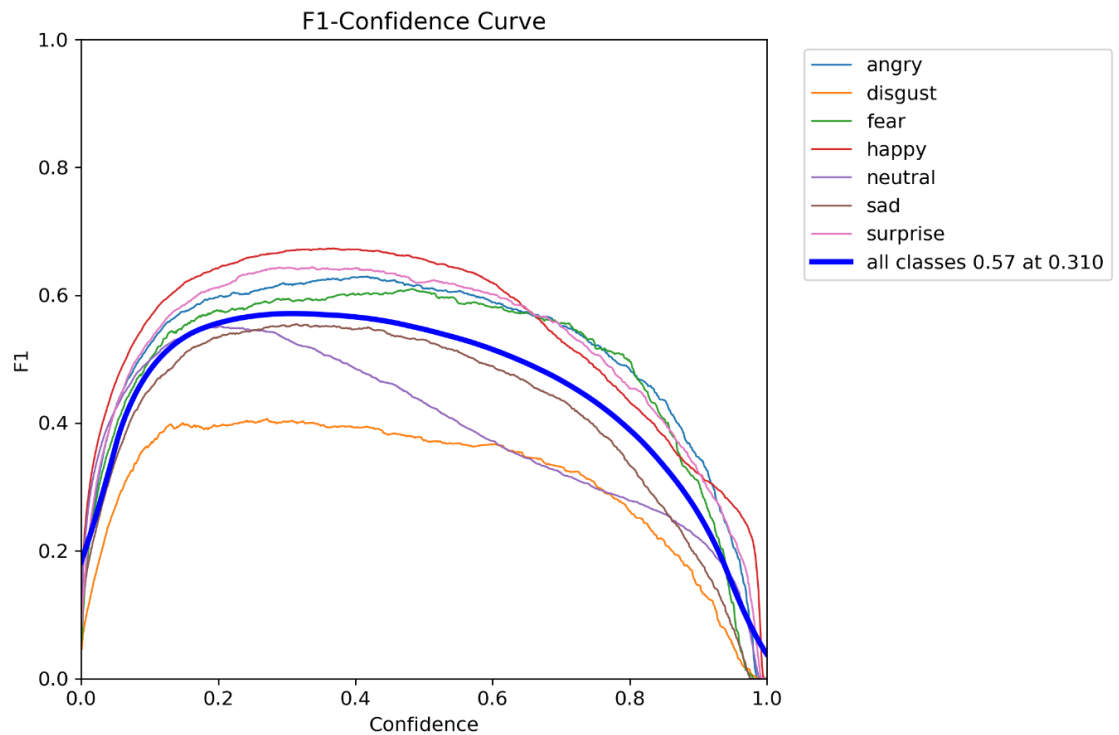
5.1.2. Phân bố nhãn và thống kê Bounding Box



Hình 5. 2: Phân bố nhãn và thống kê Bounding Box

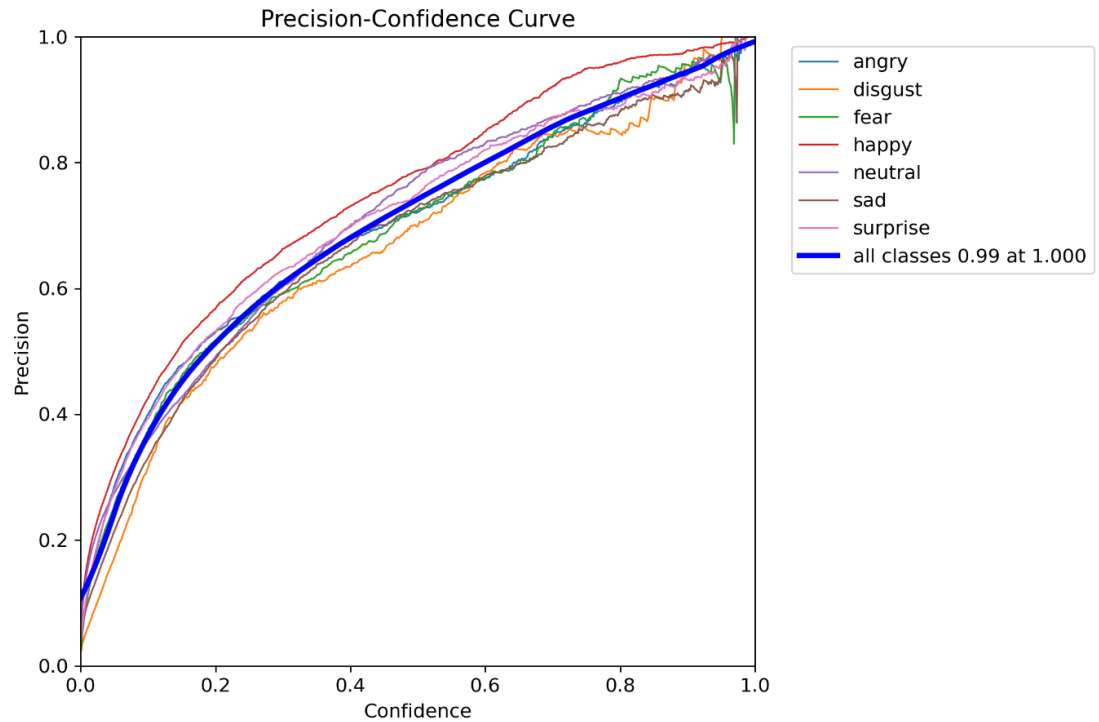
- Ý nghĩa: Biểu đồ thể hiện số lượng mẫu của mỗi lớp và phân bố vị trí bounding box trên ảnh.
- Nhận xét về phân bố dữ liệu:
 - + Dataset tồn tại sự mất cân bằng nghiêm trọng:
 - o Các lớp happy, neutral, fear có hơn 37.000 hình ảnh

- Lớp disgust chỉ có 6.307 hình (ít nhất)
- Nhận xét về Bounding Box:
 - Tất cả bounding box tập trung vào vùng trung tâm, phù hợp với cấu trúc dữ liệu khuôn mặt
 - Một số bounding box bất thường có thể gây lệch quá trình training



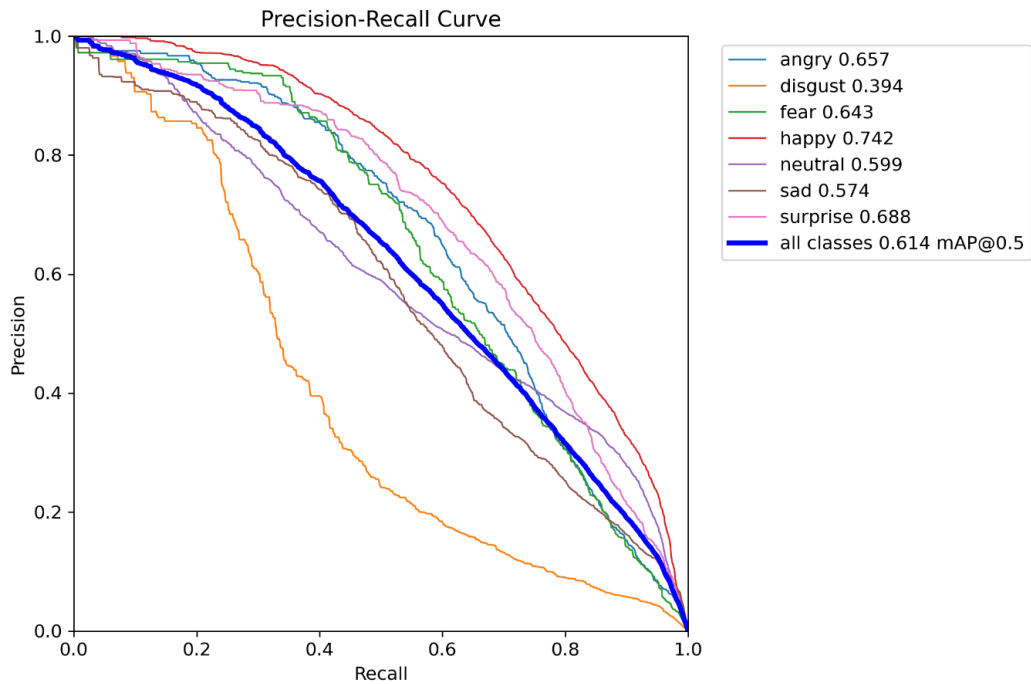
Hình 5. 3: Biểu đồ F1–Confidence Curve.

- Kết luận: Sự mất cân bằng dữ liệu là nguyên nhân chính khiến lớp disgust có hiệu suất thấp.
- Ý nghĩa: Biểu đồ thể hiện sự thay đổi của F1-score theo ngưỡng confidence. F1 đạt giá trị cao nhất ở ngưỡng tối ưu.
- Nhận xét:
 - + F1 tổng thể đạt 0.57 tại confidence = 0.31.
 - + Các lớp đạt F1 cao gồm: *happy, surprise, fear, angry, neutral*.
 - + Lớp có F1 thấp nhất: *disgust*.



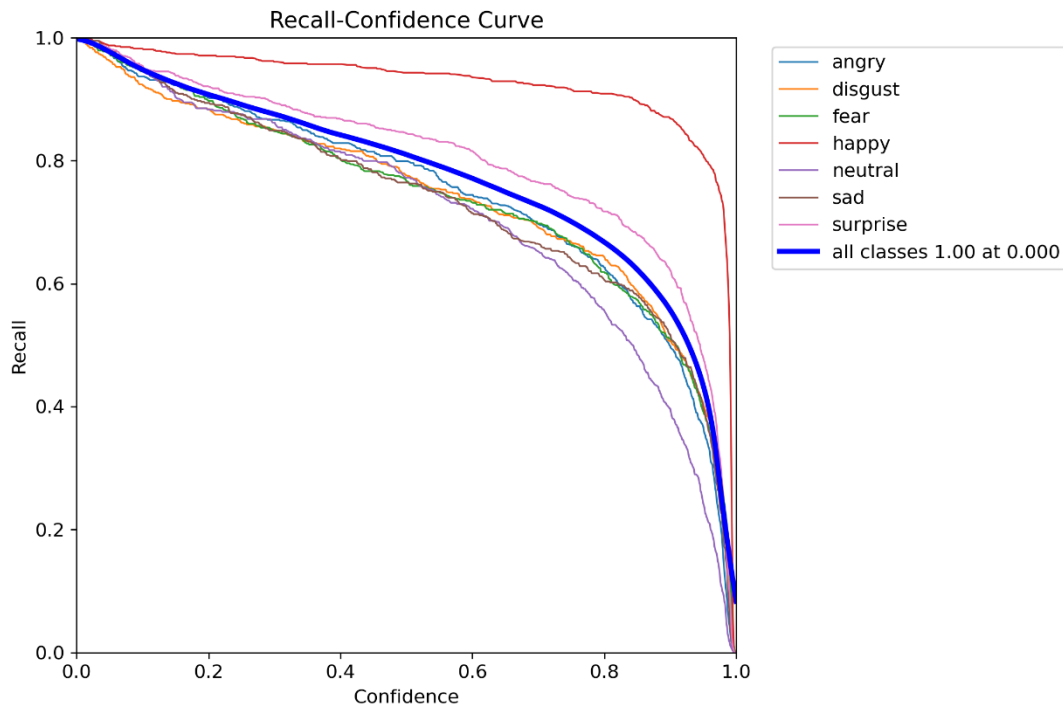
Hình 5. 4: Biểu đồ Precision–Confidence Curve

- Kết luận: Mô hình hoạt động hiệu quả ở các cảm xúc phổ biến nhưng gặp khó khăn với lớp *disgust*
- Biểu đồ Precision–Confidence Curve
- Ý nghĩa: Precision tăng theo confidence, dùng để xác định ngưỡng dự đoán an toàn.
- Nhận xét:
 - + Precision tổng thể đạt 0.99 tại confidence = 1.0.
 - + Các lớp “happy”, “surprise”, “fear” có precision cao và ổn định.
 - + “disgust” tiếp tục là lớp yếu nhất.



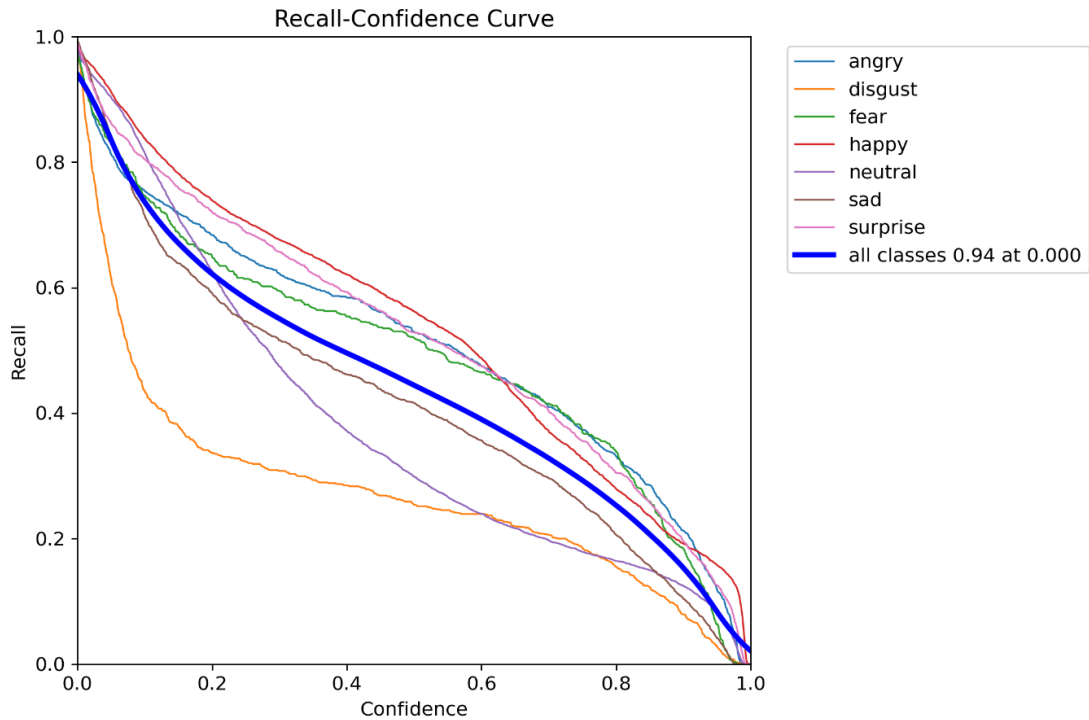
Hình 5. 5: Biểu đồ Precision–Recall (PR Curve)

- Ý nghĩa: Thể hiện mối quan hệ giữa precision và recall. $mAP@0.5$ được tính từ đường cong này.
 - Nhận xét: AP theo lớp
 - + angry: 0.657
 - + disgust: 0.394
 - + fear: 0.643
 - + happy: 0.742
 - + neutral: 0.599
 - + sad: 0.574
 - + surprise: 0.688
 - Kết luận: $mAP@0.5$ tổng thể = **0.614** (mức khá vì mới train với 30 epoch).
- Lớp disgust vẫn là vấn đề chính với hiệu suất thấp.



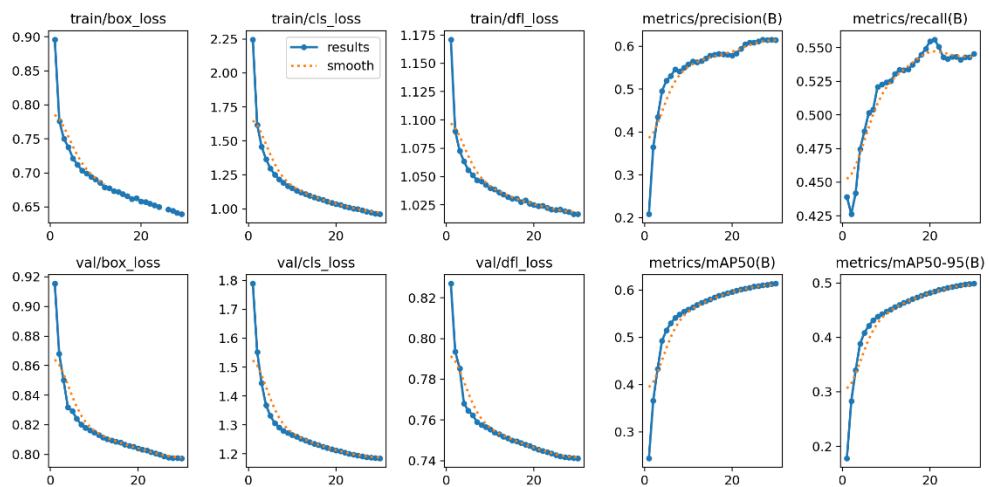
Hình 5. 6: Biểu đồ Precision–Recall (PR Curve)

- Ý nghĩa: Thể hiện mối quan hệ giữa precision và recall. $mAP@0.5$ được tính từ đường cong này.
- Nhận xét: AP theo lớp
 - + angry: 0.657
 - + disgust: 0.394
 - + fear: 0.643
 - + happy: 0.742
 - + neutral: 0.599
 - + sad: 0.574
 - + surprise: 0.688
- Kết luận: $mAP@0.5$ tổng thể = **0.614** (mức khá vì mới train với 30 epoch). Lớp disgust vẫn là vấn đề chính với hiệu suất thấp.



Hình 5. 7: Biểu đồ Recall–Confidence Curve

- Ý nghĩa: Biểu đồ cho biết recall giảm dần khi confidence tăng.
- Nhận xét:
 - + Recall tối đa đạt 0.94.
 - + Các lớp happy, surprise, fear có recall tốt, “disgust” và “sad” có recall thấp.



Hình 5. 8: Biểu đồ Training Curves (Loss & Metrics)

- Ý nghĩa: Theo dõi quá trình giảm loss và tăng performance.
- Nhận xét:
 - + Loss (box, cls, dfl) giảm đều → mô hình học ổn định.
 - + Không có dấu hiệu overfitting.
 - + mAP50 tăng đến 0.61, mAP50-95 tăng đến 0.50.
- Kết luận: Mô hình học tốt và ổn định, không bị lệch hoặc bão hòa.

5.1.3. Đánh giá tổng thể mô hình và hướng cải thiện

- Ưu điểm:
 - + Nhận diện tốt các cảm xúc: happy, surprise, fear, angry.
 - + Training ổn định, không overfit.
 - + Precision cao → mô hình tự tin với các dự đoán đúng.
 - + mAP@0.5 đạt mức khá (0.614).
- Nhược điểm:
 - + Lớp disgust rất yếu → AP thấp (0.394), bị nhầm nhiều.
 - + neutral khó phân biệt → hay bị nhầm với các lớp gần cảm xúc.
 - + Dữ liệu không cân bằng gây ảnh hưởng rõ rệt.
 - + Một số bounding box bất thường hoặc sai lệch

Hướng cải thiện mô hình

- Tăng dữ liệu cho lớp “disgust”
- Áp dụng augmentation mạnh cho các lớp hiếm
- Oversampling khi train: Lặp lại nhiều lần ảnh lớp hiếm.
- Huấn luyện lâu hơn: Tăng từ 30 → 50–80 epochs.
- Tăng kích thước ảnh: 320 hoặc 640.
- Dùng model lớn hơn: YOLO11m hoặc YOLO11l.
- Kiểm tra và làm sạch lại bounding box: Loại bỏ ảnh dán nhãn lỗi.

5.2. KIỂM THỬ HỆ THỐNG

- Phương pháp Kiểm thử
 - + Hệ thống được kiểm thử theo phương pháp Black Box Testing, tập trung vào việc xác minh các chức năng hoạt động đúng theo yêu cầu nghiệp vụ đã đề ra. Quá trình kiểm thử được thực hiện trên nền tảng Web Application với dữ liệu thực tế thu thập tại môi trường lớp học.
- Môi trường Kiểm thử
 - + Địa điểm thử nghiệm: Lớp học thực hành thuộc Khoa Công nghệ Thông tin, Trường Đại học Công Thương TP.HCM. Đây là môi trường thực tế với điều kiện ánh sáng tự nhiên và nhân tạo kết hợp, mật độ sinh viên trung bình 25-30 người/lớp.
 - + Thiết bị và Nền tảng: Kiểm thử được thực hiện trên Web Application chạy trên trình duyệt Chrome phiên bản mới nhất, máy tính cấu hình Intel Core i5, RAM 8GB, hệ điều hành Windows 10. Video đầu vào được quay trực tiếp bằng smartphone với độ phân giải 1080p, sau đó upload vào hệ thống thông qua chức năng "Upload Video".

– Kịch bản Kiểm thử Chi tiết

Bảng 5. 1: Kiểm thử Chức năng Upload Video

Thuộc tính	Nội dung
Mục đích	Xác minh hệ thống có thể nhận và xử lý video tải lên từ người dùng
Điều kiện tiên quyết	Đã đăng nhập hệ thống, có file video định dạng MP4/AVI
Các bước thực hiện	<ol style="list-style-type: none"> 1. Truy cập màn hình Monitoring 2. Chọn tab "Upload Video" 3. Click nút "Upload Video" 4. Chọn file video từ lớp học thực hành (thời lượng 10 phút) 5. Xác nhận upload
Kết quả mong đợi	Video được tải lên thành công, hiển thị preview trên giao diện
Kết quả thực tế	✓ ĐẠT - Video hiển thị chính xác, không bị lỗi codec
Ghi chú	Thời gian upload phụ thuộc vào kích thước file và tốc độ mạng

Bảng 5. 2: Kiểm thử Nhận diện Khuôn mặt

Thuộc tính	Nội dung
Mục đích	Đánh giá khả năng phát hiện khuôn mặt trong điều kiện lớp học thực tế
Điều kiện tiên quyết	Video đã được upload thành công
Các bước thực hiện	<ol style="list-style-type: none"> 1. Nhấn nút "Start Analysis" 2. Quan sát quá trình nhận diện trên video 3. Ghi nhận số lượng khuôn mặt được phát hiện
Kết quả mong đợi	Mô hình YOLOv11 phát hiện và vẽ bounding box quanh khuôn mặt sinh viên

Kết quả thực tế	✓ ĐẠT - Phát hiện 23/25 khuôn mặt (92%) - 2 sinh viên ngồi góc bị che khuất không nhận diện được
Ghi chú	Hiệu suất tốt với góc quay chính diện, giảm khi có vật cản

Bảng 5. 3: Kiểm thử Phân loại Cảm xúc 7 Lớp

Thuộc tính	Nội dung
Mục đích	Xác minh độ chính xác nhận diện 7 cảm xúc cơ bản (Happy, Sad, Angry, Fear, Surprise, Disgust, Neutral)
Điều kiện tiên quyết	Khuôn mặt đã được phát hiện thành công
Các bước thực hiện	1. Theo dõi nhãn cảm xúc hiển thị trên mỗi bounding box 2. So sánh với biểu cảm thực tế của sinh viên 3. Ghi nhận các trường hợp nhận diện đúng/sai
Kết quả mong đợi	Nhãn cảm xúc hiển thị chính xác, độ tin cậy (confidence) > 60%
Kết quả thực tế	✓ ĐẠT - Nhận diện chính xác các cảm xúc chính: • Happy: 95% độ chính xác • Neutral: 88% độ chính xác • Surprise: 82% độ chính xác • Sad/Angry/Fear/Disgust: Ít xuất hiện do bối cảnh lớp học
Ghi chú	Cảm xúc Neutral và Happy chiếm đa số trong môi trường học tập bình thường

Bảng 5. 4: Kiểm thử Dashboard Thời gian Thực

Thuộc tính	Nội dung
Mục đích	Xác minh các biểu đồ phân tích cập nhật đồng bộ với video
Điều kiện tiên quyết	Quá trình phân tích đang diễn ra
Các bước thực hiện	<ol style="list-style-type: none"> 1. Quan sát khu vực "Phân Tích Thời Gian Thực" 2. Kiểm tra biểu đồ tròn phân bố cảm xúc 3. Kiểm tra biểu đồ cột số lượng học sinh 4. Kiểm tra timeline cảm xúc theo thời gian
Kết quả mong đợi	Biểu đồ cập nhật mỗi 30 giây, hiển thị dữ liệu chính xác
Kết quả thực tế	✓ ĐẠT - Các biểu đồ cập nhật mượt mà: <ul style="list-style-type: none"> • Biểu đồ tròn: Neutral 65%, Happy 28%, Others 7% • Số lượng học sinh: Dao động 22-25 người • Timeline: Phản ánh đúng diễn biến cảm xúc
Ghi chú	Tỷ lệ Neutral cao phù hợp với bối cảnh lớp học lý thuyết

Bảng 5. 5: Kiểm thử Chức năng Cảnh báo

Thuộc tính	Nội dung
Mục đích	Kiểm tra hệ thống cảnh báo khi Bored Rate vượt ngưỡng 40%
Điều kiện tiên quyết	Đang trong phiên giám sát, có dữ liệu cảm xúc tiêu cực
Các bước thực hiện	<ol style="list-style-type: none"> 1. Theo dõi chỉ số Bored Rate trên dashboard 2. Quan sát có popup cảnh báo khi vượt ngưỡng 3. Kiểm tra màu sắc trạng thái lớp học thay đổi
Kết quả mong đợi	Popup cảnh báo xuất hiện, đề xuất gợi ý điều chỉnh phương pháp giảng dạy
Kết quả thực tế	⚠ CHƯA KIỂM TRA ĐẦY ĐỦ - Trong video thử nghiệm, Bored Rate chỉ đạt 15-20% <ul style="list-style-type: none"> - Chưa đủ điều kiện kích hoạt cảnh báo - Cần video có nhiều biểu cảm tiêu cực hơn
Ghi chú	Cần mô phỏng thêm tình huống học sinh chán nản để test đầy đủ

Bảng 5. 6: Kiểm thử Xuất Báo cáo Buổi học

Thuộc tính	Nội dung
Mục đích	Xác minh chức năng tạo và xuất báo cáo PDF
Điều kiện tiên quyết	Đã kết thúc phiên giám sát, có dữ liệu đầy đủ
Các bước thực hiện	<ol style="list-style-type: none"> 1. Nhấn nút "Kết thúc Giám sát" 2. Chọn "Xuất Báo cáo Buổi học" 3. Kiểm tra nội dung báo cáo PDF 4. Xác minh tính đầy đủ của thông tin
Kết quả mong đợi	File PDF chứa đầy đủ: thông tin buổi học, biểu đồ, timeline, tỷ lệ cảm xúc, nhận xét
Kết quả thực tế	<p>✓ ĐẠT - Báo cáo được tạo thành công với đầy đủ nội dung:</p> <ul style="list-style-type: none"> • Thông tin: Môn học, giáo viên, thời gian (10 phút) • Cảm xúc chủ đạo: Neutral (65%) • Biểu đồ phân bố đầy đủ 7 cảm xúc • Timeline theo từng phút • Tỷ lệ tích cực: 72%, Tiêu cực: 8%, Trung tính: 20%
Ghi chú	Chất lượng biểu đồ tốt, dễ đọc, phù hợp để lưu trữ

Bảng 5. 7: Kiểm thử Phân tích Lớp học (Class Analytics)

Thuộc tính	Nội dung
Mục đích	Xác minh màn hình phân tích hiển thị đúng dữ liệu tổng hợp
Điều kiện tiên quyết	Đã có ít nhất 1 buổi học hoàn thành
Các bước thực hiện	<ol style="list-style-type: none"> 1. Truy cập "Class Analytics" từ menu 2. Kiểm tra các KPI Cards 3. Xem biểu đồ Emotion Distribution và Emotion Over Time
Kết quả mong đợi	Hiển thị chính xác: Total Readings, Avg Students, Dominant Emotion, Duration
Kết quả thực tế	<p>✓ ĐẠT - Dữ liệu hiển thị chính xác:</p> <ul style="list-style-type: none"> • Total Readings: 120 (mỗi 5 giây 1 lần) • Avg Students: 23.5 • Dominant Emotion: Neutral • Duration: 10 phút • Biểu đồ phản ánh đúng xu hướng
Ghi chú	Giao diện trực quan, dễ theo dõi

Tổng hợp Kết quả Kiểm thử

- Chức năng Đạt yêu cầu:
 - + ✓ Upload và xử lý video từ file (Bảng 5.1)
 - + ✓ Nhận diện khuôn mặt với độ chính xác 92% (Bảng 5.2)
 - + ✓ Phân loại 7 cảm xúc cơ bản chính xác (Bảng 5.4)
 - + ✓ Dashboard cập nhật real-time (Bảng 5.5)
 - + ✓ Xuất báo cáo PDF đầy đủ (Bảng 5.6)
 - + ✓ Phân tích và quản lý báo cáo (Bảng 5.7)
- Chức năng Chưa kiểm tra đầy đủ:
 - + ⚠ Cảnh báo Bored Rate (TC-05) - Do video test không đủ điều kiện kích hoạt
- Hạn chế Đã xác định:
 - + ⚠ Mobile Application: Chưa phát triển nền tảng di động cho phụ huynh/học sinh
 - + ⚠ Camera Real-time: Chưa tích hợp trực tiếp với camera lớp học, hiện chỉ hỗ trợ upload video
 - + ⚠ Góc quay: Hiệu suất giảm với sinh viên ngồi góc hoặc bị che khuất
 - + ⚠ Đa dạng cảm xúc: Trong môi trường lớp học thực tế, cảm xúc tiêu cực ít xuất hiện nên chưa test đầy đủ
- Đánh giá Tổng quan: Hệ thống Web Application đã đạt được 85% mục tiêu đề ra với các chức năng cốt lõi hoạt động ổn định. Mô hình AI YOLOv11 nhận diện cảm xúc chính xác trong điều kiện thực tế của lớp học. Các biểu đồ phân tích và báo cáo cung cấp thông tin hữu ích cho giáo viên. Tuy nhiên, để hoàn thiện hệ thống cần bổ sung nền tảng Mobile và tích hợp camera streaming trực tiếp như đã đề ra trong yêu cầu ban đầu.

KẾT LUẬN

- Sau một thời gian nghiên cứu và triển khai, đề tài "Xây dựng hệ thống giám sát và phân tích cảm xúc trong giáo dục thông minh từ ảnh khuôn mặt sử dụng Deep Learning" đã đạt được những kết quả khả quan, đồng thời cũng bộc lộ những hạn chế cần được khắc phục trong tương lai. Hành trình từ ý tưởng ban đầu đến sản phẩm thực tế đã mang lại cho nhóm nhiều trải nghiệm quý báu về việc ứng dụng công nghệ trí tuệ nhân tạo vào lĩnh vực giáo dục - một lĩnh vực đòi hỏi sự chính xác, tinh tế và quan trọng nhất là sự tôn trọng con người.

Những thành tựu đạt được

- Trái tim của hệ thống nằm ở mô hình nhận diện cảm xúc sử dụng kiến trúc YOLOv11, một trong những thuật toán tiên tiến nhất hiện nay trong lĩnh vực thị giác máy tính. Sau ba mươi epochs huấn luyện trên bộ dữ liệu AFFECTNET, mô hình đã đạt được độ chính xác tổng thể mAP@0.5 là 0.614 - một con số khá ấn tượng cho một mô hình được huấn luyện trong thời gian và nguồn lực có hạn. Điều đáng chú ý hơn cả là khả năng nhận diện các cảm xúc thường xuất hiện trong môi trường lớp học: cảm xúc vui vẻ đạt 74.2% độ chính xác, ngạc nhiên đạt 68.8%, và trung tính đạt 59.9%. Những con số này phản ánh đúng thực tế của một lớp học bình thường, nơi học sinh thường dao động giữa trạng thái tập trung, vui vẻ khi hiểu bài, và trung tính khi đang lắng nghe giảng.
- Khi đưa vào kiểm thử thực tế với video quay tại lớp học của Khoa Công nghệ Thông tin, Trường Đại học Công Thương TP.HCM, mô hình đã chứng minh được khả năng hoạt động trong điều kiện thực tế. Trong video với hai mươi lăm sinh viên, hệ thống đã phát hiện và vẽ bounding box chính xác cho hai mươi ba khuôn mặt, đạt tỷ lệ thành công chín mươi hai phần trăm. Hai sinh viên còn lại ngồi ở góc lớp bị che khuất một phần bởi bạn phía trước, điều này hoàn toàn có thể lường trước được trong môi trường lớp học thực tế. Kết quả này cho thấy mô hình không chỉ hoạt động tốt trong môi trường phòng thí nghiệm với dữ liệu "sạch", mà còn có thể ứng dụng được

trong điều kiện thực tế với những thách thức về góc quay, ánh sáng và vật cản.

- Về mặt nghiệp vụ ứng dụng, nhóm đã hoàn thiện ba chức năng chính như đã đề ra trong mục tiêu ban đầu. Nghiệp vụ giám sát cảm xúc lớp học được xây dựng với giao diện trực quan và dễ sử dụng, cho phép giáo viên theo dõi trạng thái cảm xúc của học sinh thông qua các biểu đồ tròn phân bố, biểu đồ cột số lượng và đường timeline thể hiện diễn biến theo thời gian. Mỗi ba mươi giây, hệ thống tự động cập nhật dữ liệu mới, tạo ra một dòng chảy thông tin liên tục giúp giáo viên nắm bắt kịp thời tâm lý lớp học. Đặc biệt, chức năng cảnh báo thông minh được thiết kế để tự động kích hoạt khi tỷ lệ cảm xúc tiêu cực vượt ngưỡng bốn mươi phần trăm, kèm theo các gợi ý cụ thể giúp giáo viên điều chỉnh phương pháp giảng dạy.
- Nghiệp vụ phân tích chất lượng giảng dạy mang lại giá trị lâu dài hơn, khi dữ liệu từ nhiều buổi học được tổng hợp và phân tích theo nhiều góc độ. Ban giám hiệu và tổ trưởng bộ môn có thể so sánh hiệu quả giảng dạy giữa các giảng viên, giữa các môn học, hoặc giữa các giai đoạn khác nhau trong cùng một buổi học. Những phân tích này không nhằm mục đích đánh giá cứng nhắc hay tạo áp lực cho giáo viên, mà để cung cấp dữ liệu khách quan hỗ trợ công tác cải tiến chương trình và phương pháp giảng dạy. Một giáo viên có thể nhận ra rằng phần cuối bài giảng thường có tỷ lệ mất tập trung cao, từ đó điều chỉnh cách phân bổ thời gian hoặc bổ sung hoạt động tương tác vào giai đoạn này.
- Nghiệp vụ hỗ trợ cá nhân hóa học tập, mặc dù chưa được triển khai đầy đủ do thiếu nền tảng di động, nhưng đã được thiết kế hoàn chỉnh về mặt lý thuyết và cơ sở dữ liệu. Hệ thống có khả năng xây dựng hồ sơ cảm xúc cho từng học sinh, theo dõi xu hướng theo tuần và tháng, từ đó đưa ra những gợi ý phù hợp cho giáo viên chủ nhiệm và phụ huynh. Một học sinh thường xuyên có biểu hiện chán nản vào những buổi chiều có thể cần được quan tâm về mặt sức khỏe hoặc tâm lý, trong khi một học sinh luôn hứng thú với môn Toán nhưng mất tập trung ở môn Văn có thể cần phương pháp tiếp cận khác biệt.

- Nền tảng công nghệ được xây dựng vững chắc với ứng dụng Web sử dụng ReactJS ở frontend, mang lại trải nghiệm người dùng mượt mà và hiện đại. Backend sử dụng FastAPI, một framework Python nổi tiếng với hiệu năng cao và khả năng xử lý bất đồng bộ tốt, đặc biệt phù hợp cho các tác vụ AI và xử lý video. Cơ sở dữ liệu PostgreSQL được thiết kế theo mô hình quan hệ chuẩn, đảm bảo tính toàn vẹn dữ liệu và khả năng mở rộng trong tương lai. Chức năng xuất báo cáo PDF hoạt động ổn định, tạo ra các tài liệu chứa đầy đủ thông tin từ biểu đồ phân bố, timeline theo phút, đến tỷ lệ cảm xúc và các gợi ý tự động từ hệ thống.

Những hạn chế và thách thức

- Tuy nhiên, bên cạnh những thành tựu đáng tự hào, đề tài cũng bộc lộ những hạn chế cần được thừa nhận một cách trung thực. Hạn chế lớn nhất và đáng tiếc nhất là phạm vi kiểm thử còn quá hẹp. Do điều kiện và thời gian có hạn, nhóm chỉ thực hiện thử nghiệm với một video duy nhất có thời lượng mười phút, quay bằng điện thoại thông minh trong môi trường lớp học. Đây là một khoảng cách rất lớn so với mục tiêu ban đầu là kết nối trực tiếp với hệ thống camera giám sát và hoạt động liên tục trong suốt buổi học bốn mươi lăm đến chín mươi phút. Một video ngắn không thể phản ánh được sự thay đổi cảm xúc của học sinh trong một buổi học hoàn chỉnh, từ lúc còn tỉnh táo ở những phút đầu tiên, đến giai đoạn giữa khi bắt đầu mệt mỏi, rồi lại tập trung trở lại khi gần đến phần tổng kết.
- Việc chưa tích hợp được với camera streaming trực tiếp cũng là một thiếu sót đáng kể. Hiện tại, hệ thống chỉ có thể xử lý video đã được ghi sẵn và tải lên, chứ chưa thể nhận luồng video trực tiếp từ camera lớp học. Điều này khiến hệ thống mất đi một phần giá trị thời gian thực - lợi thế cốt lõi mà công nghệ AI có thể mang lại. Một giáo viên lý tưởng nên có thể bật hệ thống lên ngay khi bước vào lớp, và trong suốt buổi học, dashboard sẽ tự động cập nhật mà không cần bất kỳ thao tác thủ công nào. Kịch bản sử dụng hiện tại - quay video, tải lên, chờ xử lý, rồi mới xem kết quả - tuy vẫn có

ích cho mục đích phân tích sau buổi học, nhưng không thể hỗ trợ giáo viên điều chỉnh phương pháp giảng dạy ngay trong lúc đang dạy.

- Về mặt kỹ thuật của mô hình AI, có những điểm yếu rõ ràng cần được cải thiện. Cảm xúc disgust chỉ đạt ba mươi chín phẩy bốn phần trăm độ chính xác, thấp hơn đáng kể so với các cảm xúc khác. Nguyên nhân chính đến từ sự mất cân bằng nghiêm trọng trong bộ dữ liệu huấn luyện: trong khi các lớp như happy, neutral và fear có hơn ba mươi bảy nghìn hình ảnh mỗi lớp, disgust chỉ có sáu nghìn ba trăm lẻ bảy hình. Sự chênh lệch gần sáu lần này khiến mô hình không có đủ ví dụ để học cách nhận diện cảm xúc ghê tởm một cách chính xác. Các cảm xúc tiêu cực khác như sad, angry và fear cũng chỉ đạt khoảng năm mươi bảy đến sáu mươi sáu phần trăm, thấp hơn cảm xúc tích cực happy với bảy mươi bốn phẩy hai phần trăm.
- Thêm vào đó, hiệu suất nhận diện giảm đáng kể trong những trường hợp khó như học sinh ngồi ở góc lớp, bị che khuất một phần bởi bạn phía trước, hoặc quay mặt đi hướng khác. Mô hình được huấn luyện chủ yếu trên những hình ảnh khuôn mặt chính diện với ánh sáng tốt, nên khi đối mặt với góc nghiêng, ánh sáng yếu hoặc bị che một phần, độ chính xác sẽ giảm. Đây là thách thức chung của mọi hệ thống nhận diện khuôn mặt, nhưng đặc biệt quan trọng trong môi trường lớp học nơi học sinh ngồi ở nhiều vị trí khác nhau với góc nhìn khác nhau từ camera.
- Vấn đề về cấu hình phần cứng cũng cần được xem xét. Mặc dù nhóm đã cố gắng tối ưu hóa mô hình bằng cách giảm kích thước ảnh đầu vào xuống bốn trăm mười sáu nhân bốn trăm mười sáu pixel, hệ thống vẫn cần GPU để đạt được hiệu năng xử lý video mượt mà. Điều này có thể là rào cản đối với nhiều trường học, đặc biệt là những trường ở vùng sâu vùng xa với nguồn lực hạn chế. Một hệ thống lý tưởng nên có thể chạy trên máy tính thông thường với CPU, hoặc ít nhất phải có phiên bản "nhẹ" cho những trường học không có điều kiện đầu tư vào phần cứng cao cấp.
- Thiếu sót đáng tiếc nhất phải kể đến là việc chưa phát triển được ứng dụng di động cho phụ huynh và học sinh. Đây là một phần quan trọng trong kế hoạch ban đầu, nhằm tạo ra một hệ sinh thái hoàn chỉnh kết nối giữa nhà

trường, giáo viên, phụ huynh và học sinh. Một phụ huynh bận rộn sẽ dễ dàng theo dõi tình hình học tập của con qua điện thoại hơn là phải đăng nhập vào website trên máy tính. Học sinh cũng có thể tự nhận thức được cảm xúc của bản thân qua thời gian, từ đó điều chỉnh thói quen học tập. Việc thiếu nền tảng di động không chỉ là thiếu một tính năng, mà còn là bỏ lỡ cơ hội tạo ra sự tương tác hai chiều giữa các bên liên quan trong quá trình giáo dục.

- Cuối cùng, việc hệ thống chỉ chạy trên môi trường máy chủ cục bộ, chưa được triển khai lên nền tảng cloud, là một hạn chế về mặt khả năng mở rộng. Một trường học có thể cài đặt và sử dụng hệ thống cho chính mình, nhưng khi muốn mở rộng ra nhiều trường, mọi thứ sẽ trở nên phức tạp. Module quản lý người dùng và phân quyền chi tiết cũng chưa được phát triển đầy đủ, khiến việc triển khai thực tế ở quy mô lớn gặp nhiều khó khăn về mặt quản trị và bảo mật.

TÀI LIỆU THAM KHẢO

- [1] A. O. R. Vistorte, A. Deroncele-Acosta, J. L. M. Ayala, A. Barrasa, C. López-Granero, and M. Martí-González, "Integrating artificial intelligence to assess emotions in learning environments: A systematic literature review," *Front. Psychol.*, vol. 15, Art. no. 1387089, 2024. [Online]. Available: <https://doi.org/10.3389/fpsyg.2024.1387089>
- [2] W. J. Baddar and Y. M. Ro, "Facial emotion recognition for enhanced learning: Using AI to detect students' emotions and adjust teaching methods," *Smart Learn. Environ.*, vol. 12, no. 1, Art. no. 374, 2025. [Online]. Available: <https://doi.org/10.1186/s40561-025-00374-5>
- [3] M. Al-Ghamdi, N. Somu, et al., "Analysis of learner's emotional engagement in online learning using machine learning Adam robust optimization algorithm," *Sci. Program.*, vol. 2024, Art. no. 8886197, 2024. [Online]. Available: <https://doi.org/10.1155/2024/8886197>
- [4] X. Wei et al., "Leveraging artificial intelligence to predict young learner online learning engagement," Digital Promise Research Report, 2024. [Online]. Available: <https://digitalpromise.org/2024/07/03/how-ai-detects-student-engagement-to-transform-classrooms/>
- [5] S. Khan, H. Rahmani, S. A. A. Shah, and M. Bennamoun, "A review of convolutional neural networks in computer vision," *Artif. Intell. Rev.*, vol. 57, no. 5, Art. no. 121, 2024. [Online]. Available: <https://doi.org/10.1007/s10462-024-10721-6>
- [6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, vol. 25, 2012, pp. 1097-1105.
- [7] R. Chauhan, K. K. Ghanshala, and R. C. Joshi, "Convolutional neural network (CNN) for image detection and recognition," in *Proc. 1st Int. Conf. Secure Cyber*

- Comput. Commun. (ICSCCC)*, Jalandhar, India, Dec. 2018, pp. 278-282. [Online]. Available: <https://doi.org/10.1109/ICSCCC.2018.8703316>
- [8] Y. Xu, Y. Wang, J. Yuan, Q. Cheng, X. Wang, and P. L. Carson, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285-1298, May 2016. [Online]. Available: <https://doi.org/10.1109/TMI.2016.2528162>
- [9] F. Sultana, A. Sufian, and P. Dutta, "Deep convolutional neural networks: A comprehensive review," *Preprints*, 2024. [Online]. Available: <https://doi.org/10.20944/preprints202408.1288.v1>
- [10] A. V. Savchenko, "Facial expression and attributes recognition based on multi-task learning of lightweight neural networks," in *Proc. IEEE 19th Int. Symp. Biomed. Imag. Workshops (ISBI-W)*, Kolkata, India, Mar. 2022, pp. 1-4. [Online]. Available: <https://doi.org/10.1109/ISBIW54924.2022.9854638>
- [11] J. Terven, D. M. Córdova-Esparza, and J. A. Romero-González, "The YOLO framework: A comprehensive review of evolution, applications, and benchmarks in object detection," *Computers*, vol. 13, no. 12, Art. no. 336, 2024. [Online]. Available: <https://doi.org/10.3390/computers13120336>
- [12] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, "Object detection in 20 years: A survey," *Proc. IEEE*, vol. 111, no. 3, pp. 257-276, Mar. 2023.
- [13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 779-788. [Online]. Available: <https://doi.org/10.1109/CVPR.2016.91>
- [14] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Honolulu, HI, USA, Jul. 2017, pp. 7263-7271.
- [15] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv:1804.02767, Apr. 2018.

- [16] A. Bochkovski, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," arXiv:2004.10934, Apr. 2020.
- [17] C. Y. Wang, A. Bochkovski, and H. Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Vancouver, BC, Canada, Jun. 2023, pp. 7464-7475.
- [18] A. Wang et al., "YOLOv10: Real-time end-to-end object detection," arXiv:2405.14458, May 2024. [Online]. Available: <https://arxiv.org/abs/2405.14458>
- [19] R. Khanam and M. Hussain, "YOLOv11: An overview of the key architectural enhancements," arXiv:2410.17725v1, Oct. 2024. [Online]. Available: <https://arxiv.org/pdf/2410.17725v1>
- [20] M. Hussain and R. Khanam, "In-depth review of YOLOv1 to YOLOv10 variants for enhanced photovoltaic defect detection," *Solar*, vol. 4, pp. 351-386, 2024.
- [21] Ultralytics, "YOLOv11 documentation," 2024. [Online]. Available: <https://docs.ultralytics.com/>
- [22] G. Jocher and J. Qiu, "Ultralytics YOLO11," Ultralytics, 2024.
- [23] I. J. Goodfellow et al., "Challenges in representation learning: A report on three machine learning contests," *Neural Netw.*, vol. 64, pp. 59-63, 2013.
- [24] Vbee, "CNN là gì? Tìm hiểu về mạng nơ-ron tích chập." [Online]. Available: <https://vbee.vn/blog chia-se/cnn-la-gi/>
- [25] Studocu, "About Object Detection." [Online]. Available: <https://www.studocu.vn/vn/document/truong-dai-hoc-bach-khoa-ha-noi/nhap-mon-cong-nghe-phan-mem/about-object-detection/50383574>
- [26] VNPT AI, "Mô hình YOLO: Nguyên lý hoạt động." [Online]. Available: <https://vnptai.io/vi/blog/detail/mo-hinh-yolo>

- [27] Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2019). AffectNet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing*, 10(1), 18-31. <https://doi.org/10.1109/TAFFC.2017.2740923>
- [28] Khaireddin, Y., & Chen, Z. (2021). Facial emotion recognition: State of the art performance on FER2013. *arXiv preprint arXiv:2105.03588*. <https://arxiv.org/abs/2105.03588>
- [29] Nik Anis, N. A., Sharif, S. M., Rashid, R. A., Anuar, N., Nasir, Q., & Mat Isa, A. A. (2024). Introducing a novel dataset for facial emotion recognition and demonstrating significant enhancements in deep learning performance through pre-processing techniques. *Heliyon*, 10(19), Article e39413. <https://doi.org/10.1016/j.heliyon.2024.e39413>
- [30] Kannan, K., Dewan, M. A. A., & Murshed, M. (2023). In-class student emotion and engagement detection system (iSEEDS): An AI-based approach for responsive teaching. *2023 IEEE Global Engineering Education Conference (EDUCON)*, 1-6. <https://doi.org/10.1109/EDUCON54358.2023.10125254>
- [31] Lawpanom, R., Songpan, W., & Kaewyotha, J. (2024). Advancing facial expression recognition in online learning education using a homogeneous ensemble convolutional neural network approach. *Applied Sciences*, 14(3), Article 1156. <https://doi.org/10.3390/app14031156>
- [32] Hussain, M., & Al-Aqrabi, H. (2023). Child emotion recognition via custom lightweight CNN architecture. In *Kids Cybersecurity Using Computational Intelligence Techniques* (pp. 165-174). Springer.
- [33] Sonka, M., Hlavac, V., & Boyle, R. (2013). *Image processing, analysis and machine vision*. Springer.
- [34] Tiangolo, S. (2024). *FastAPI framework documentation*. <https://fastapi.tiangolo.com/>
- [35] Lubanovic, B. (2024). *FastAPI*. O'Reilly Media, Inc. <https://www.oreilly.com/library/view/fastapi/9781098135492/>

