

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ



BÁO CÁO GIỮA KỲ MÔN KHOA HỌC NHẬN THỨC

Hệ thống giám sát ra/vào thông minh ứng dụng AI
(AI-Powered Smart Entrance Monitoring System)

Sinh viên thực hiện: Nguyễn Bảo Long

Mã Sinh viên: 22027537

Lớp: K67E – RE

Giảng viên: TS. Nguyễn Thé Hoàng Anh

TA. Nguyễn Minh Kiên

Hà Nội, ngày 30 tháng 10 năm 2025

I. THÔNG TIN CHUNG VỀ ĐỀ TÀI

1. Tên đề tài

Hệ thống giám sát ra/vào thông minh ứng dụng AI trong đếm người và dự đoán mật độ đám đông.

2. Tổng quan về đề tài

Đề tài này hướng đến việc phát triển một hệ thống giám sát thông minh ứng dụng Deep Learning, có khả năng:

- Phát hiện và đếm số lượng người ra vào cửa hàng theo thời gian thực.
- Ước lượng mật độ đám đông (crowd density) trong khu vực camera giám sát.
- Hiển thị dữ liệu giám sát trực quan thông qua dashboard và biểu đồ thời gian.

3. Thời gian thực hiện

Thời gian kéo dài thực hiện: 7 tuần (Từ 27/10/2025 – 8/12/2025)

4. Thông tin người thực hiện đề tài

- Họ và tên: Nguyễn Bảo Long
- Mã sinh viên: 22027537
- Lớp / Khoa: Kỹ thuật Robot / Khoa Điện tử Viễn thông – Trường Đại học Công Nghệ, ĐHQGHN

II. NỘI DUNG NGHIÊN CỨU ĐỀ TÀI

1. Giới thiệu đề tài

Trong kỷ nguyên chuyển đổi số và phát triển thành phố thông minh, việc quản lý và giám sát lưu lượng người ra vào cũng như mật độ đám đông tại các khu vực công cộng như cửa hàng, trung tâm thương mại, sân bay, tòa nhà hay khu vui chơi trở nên ngày càng quan trọng. Những dữ liệu này giúp hỗ trợ công tác điều phối nhân lực, đảm bảo an toàn, tối ưu vận hành và nâng cao trải nghiệm người dùng.

Tuy nhiên các hệ thống giám sát truyền thống hiện nay chủ yếu chỉ cung cấp dữ liệu thô từ camera hoặc phát hiện chuyển động đơn giản, chưa có khả năng phân tích nâng cao như đếm chính xác số lượng người, theo dõi hướng di chuyển, hay đánh giá mức độ đông đúc theo từng thời điểm. Điều này gây khó khăn cho việc ra quyết định và phản ứng kịp thời trong các tình huống thực tế.

Đề tài “Hệ thống giám sát ra/vào thông minh ứng dụng AI trong đếm người và dự đoán mật độ đám đông” được xây dựng nhằm khắc phục những hạn chế đó, với mục tiêu phát triển một giải pháp giám sát thông minh dựa trên công nghệ Deep Learning, có khả năng:

- Phát hiện và đếm số lượng người ra vào khu vực giám sát theo thời gian thực
- Ước lượng mật độ đám đông (crowd density) dựa trên hình ảnh từ camera.
- Trực quan hóa dữ liệu thông qua dashboard và biểu đồ thống kê theo thời gian, giúp người quản lý dễ dàng theo dõi và đưa ra quyết định nhanh chóng.

2. Mục tiêu đề tài

- Ứng dụng Deep Learning (CNN) để phát hiện và đếm người ra/vào theo thời gian thực
- Xây dựng mô hình ước lượng mật độ đám đông (Crowd Density Estimation) sử dụng mạng CNN (CSRNet).
- Phát triển dashboard trực quan hiển thị các chỉ số: số người vào, số người ra, mật độ đám đông, và cảnh báo khi vượt ngưỡng.
- Đánh giá độ chính xác của mô hình và khả năng hoạt động trong môi trường thực tế (camera cửa hàng, siêu thị, trung tâm thương mại,...)

3. Tính cấp thiết của đề tài (Lợi ích)

- Tính thực tiễn:
Việc giám sát ra/vào và dự đoán mật độ đám đông là nhu cầu thiết yếu trong quản lý an ninh, tối ưu vận hành tại các khu vực công cộng, phát hiện sớm các tình huống bất thường như tụ tập đông người, tắc nghẽn hay vượt quá giới hạn an toàn.
- Tính khoa học:
Ứng dụng các công nghệ Deep Learning tiên tiến trong thị giác máy tính, giúp tăng độ chính xác và khả năng thích ứng.
- Tính đổi mới:
Đề tài tích hợp hai hướng mạnh của AI – object detection and density estimation – trong một pipeline thống nhất, hoạt động realtime và hiển thị trực quan.

4. Tổng quan về vấn đề nghiên cứu cần thực hiện

Các nghiên cứu và sản phẩm hiện nay thường chỉ tập trung vào:

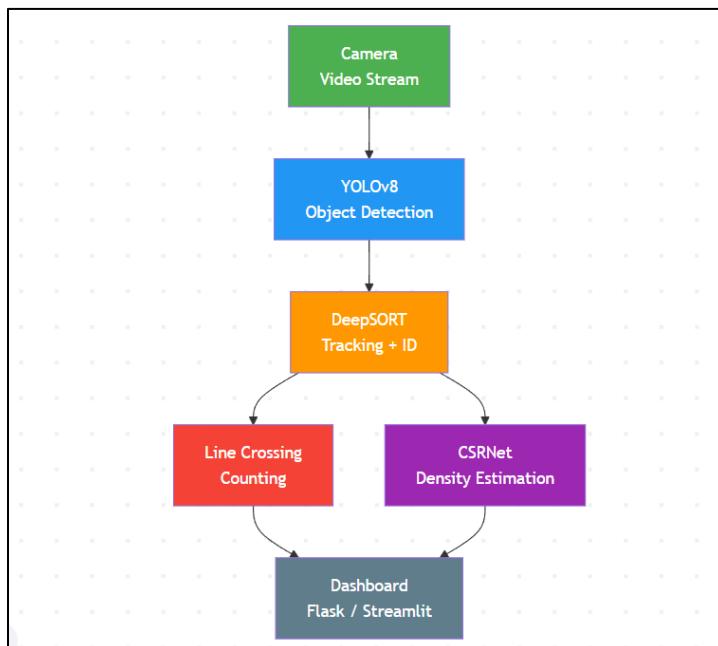
- Đếm người (People Counting) bằng cảm biến hoặc camera đơn giản.

- Ước lượng mật độ đám đông (Crowd Density) nhưng không realtime hoặc không phân tích hướng di chuyển.

Trong khi đó, đề tài này hướng đến kết hợp cả hai nhiệm vụ trong cùng một hệ thống, giúp vừa đếm chính xác người ra/vào, vừa đánh giá mật độ tổng thể trong khung hình, tạo nền tảng cho giám sát thông minh.

5. Giải pháp đề xuất (dự kiến)

- Kiến trúc tổng quan hệ thống



Hình 2.5.1. Tổng quan hệ thống AI giám sát ra/vào

- Mô tả quy trình hoạt động

1. Camera đầu vào
 - Hệ thống nhận dữ liệu hình ảnh/video theo thời gian thực từ camera giám sát đặt tại khu vực ra/vào.
 - Dữ liệu được truyền trực tiếp đến mô-đun xử lý AI để phân tích.
2. Phát hiện đối tượng (YOLOv8 Detection)
 - Mô hình YOLOv8 được sử dụng để phát hiện con người trong từng khung hình nhờ khả năng nhận dạng nhanh và chính xác
 - Kết quả đầu ra là các hộp bao (bounding boxes) chứa vị trí của từng người được phát hiện.
3. Theo dõi đối tượng (DeepSORT Tracking)
 - Thuật toán DeepSORT được sử dụng để gán ID duy nhất cho mỗi người và theo dõi chuyển động của họ qua các khung hình liên tiếp.

- Điều này giúp hệ thống phân biệt được người đi vào và người đi ra, tránh hiện tượng đếm trùng lặp.
- 4. Đếm người bằng đường ảo (Line Crossing Counting)
 - Một đường ảo (virtual line) được xác định tại vị trí cửa ra/vào.
 - Khi đối tượng (theo ID của DeepSORT) băng qua đường này, hệ thống tự động tăng bộ đếm người vào hoặc người ra tùy theo hướng di chuyển.
- 5. Ước lượng mật độ đám đông
 - Mô hình sử dụng mạng CNN sâu (Deep CNN) để tạo ra bản đồ mật độ (density map), phản ánh ước lượng mật độ đám đông trong khu vực giám sát.
- 6. Trực quan hóa dữ liệu
 - Hiển thị toàn bộ kết quả trên giao diện dashboard trực quan, giúp người quản lý dễ dàng theo dõi và giám sát hệ thống.

6. Phương pháp nghiên cứu

- Thu thập dữ liệu: sử dụng các bộ dữ liệu công khai và video thực tế tại cửa hàng, siêu thị mô phỏng
- Tiền xử lý: resize ảnh/video, gán nhãn, chuẩn hóa dữ liệu.
- Huấn luyện mô hình:
 - YOLOv8: Chỉ giữ class “person”.
 - CSRNet: Huấn luyện trên ShanghaiTech dataset (Part A & B).
- Kết hợp hệ thống: kết nối output hai mô hình để hiển thị realtime qua dashboard.
- Đánh giá:
 - People Counting: Accuracy, Precision, Recall
 - Density Estimation: RMSE

7. Dự đoán kết quả đạt được

- Độ chính xác trong phát hiện và đếm người đạt từ 90% trở lên, đảm bảo khả năng nhận diện ổn định trong nhiều điều kiện ánh sáng và góc quay khác nhau.
- Sai số trung bình trong ước lượng mật độ đám đông (MAE) được kỳ vọng dưới 10% phản ánh hiệu quả của mô hình CSRNet trong việc phân tích mật độ người trong khung hình.
- Tốc độ xử lý thời gian thực (Realtime) đạt khoảng 10-15 khung hình/giây (FPS) trên GPU tầm trung, đáp ứng yêu cầu triển khai trong các môi trường giám sát thực tế.

- Giao diện dashboard trực quan, thân thiện, hiển thị liên tục các chỉ số như số người vào/ra, mật độ đám đông, biểu đồ theo thời gian, đồng thời tự động cập nhật dữ liệu và phát cảnh báo khi phát hiện vượt ngưỡng an toàn.

Hệ thống kỳ vọng mang lại giải pháp giám sát thông minh, hiệu quả và dễ triển khai, góp phần ứng dụng công nghệ AI vào công tác quản lý an ninh và vận hành trong đô thị thông minh.

III. THỰC HIỆN ĐỀ TÀI

1. Tiến độ thực hiện

Tuần 1: Thu thập dữ liệu và xử lý dữ liệu	<ul style="list-style-type: none"> - Nghiên cứu tổng quan, thu thập các bộ dữ liệu và xử lý video. - Tiền xử lý và chuẩn hóa dữ liệu để chuẩn bị cho quá trình huấn luyện.
Tuần 2-3: Thiết kế và huấn luyện mô hình	<ul style="list-style-type: none"> - Cài đặt môi trường, thiết kế và huấn luyện mô hình Deep Learning phù hợp với hệ thống. - Đánh giá độ chính xác của mô hình trên tập dữ liệu kiểm tra.
Tuần 4: Điều chỉnh tham số và cải tiến mô hình	<ul style="list-style-type: none"> - Xem xét và đánh giá độ chính xác của mô hình để tìm hiểu những điểm yếu của mô hình. - Cải tiến mô hình để nâng cao độ chính xác và độ tin cậy.
Tuần 5-6: Thực hiện các thử nghiệm và đánh giá hiệu quả	<ul style="list-style-type: none"> - Tiến hành các thử nghiệm với mô hình đã được cải tiến, tích hợp với nhau. - Đánh giá hiệu quả của mô hình dựa trên độ chính xác và độ tin cậy.
Tuần 7: Hoàn thiện báo cáo và trình bày kết quả	<ul style="list-style-type: none"> - Viết báo cáo về kết quả của dự án. - Trình bày kết quả của dự án cho lớp học.

Bảng 1. Tiến độ thực hiện dự án

2. Phân công nhiệm vụ

Nguyễn Bảo Long	<ul style="list-style-type: none"> - Xây dựng ý tưởng và thiết kế kiến trúc hệ thống. - Cài đặt, huấn luyện mô hình YOLOv8 & DeepSORT. - Tích hợp mô hình và phát triển hệ thống giám sát.
Nguyễn Huy Thắng	<ul style="list-style-type: none"> - Thu thập, xử lý và chuẩn hóa dữ liệu video. - Huấn luyện mô hình CSRNet và đánh giá kết quả ước lượng mật độ. - Hỗ trợ kiểm thử và tối ưu hiệu năng hệ thống.
Trương Ngọc Anh	<ul style="list-style-type: none"> - Xây dựng dashboard hiển thị dữ liệu realtime. - Thiết kế giao diện, biểu đồ và kết nối cơ sở dữ liệu.

- Tổng hợp báo cáo và chuẩn bị slide thuyết trình.

Bảng 2. Phân công nhiệm vụ

IV. KẾT LUẬN

Đề tài hướng đến xây dựng một hệ thống AI giám sát thông minh, có khả năng đếm người ra/vào và dự đoán mật độ đám đông theo thời gian thực, ứng dụng Deep Learning trong thị giác máy tính.

Hệ thống có thể triển khai trong môi trường cửa hàng, siêu thị, tòa nhà hoặc khu công cộng, mang lại giá trị thực tiễn cao trong quản lý và an ninh.

V. TÀI LIỆU THAM KHẢO

1. Zhang, Y. et al. (2018). *CSRNet: Dilated Convolutional Neural Networks for Understanding Highly Congested Scenes*, CVPR 2018.
Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
https://openaccess.thecvf.com/content_cvpr_2018/papers/Li_CSRNet_Dilated_Convolutional_CVPR_2018_paper.pdf
2. Ultralytics (2023). *YOLOv8 Documentation and Implementation*
<https://github.com/ultralytics/ultralytics>
3. Wojke, N. et al. (2017). *Simple Online and Realtime Tracking with a Deep Association Metric*.
IEEE International Conference on Image Processing (ICIP).
<https://arxiv.org/pdf/1703.07402>
4. *Multiple Object Tracking Benchmark (MOT17 Dataset Paper)*
<https://motchallenge.net/>