

## An IoT-based Traffic Signal Control System Using Environmental Monitoring and Hand-Raising Behavior Recognition via Cloud-Connected Camera

Nguyen Ngan Truong<sup>1\*</sup>, Bui Nhat Thanh<sup>1</sup>, Tran Doan Hoa<sup>1</sup>, Nguyen Quoc Tinh<sup>1</sup>, Nguyen Van Truong<sup>1</sup>

<sup>1</sup>Ho Chi Minh City University of Technology and Education, Vietnam

\*Corresponding author. Email: [23162107@student.hcmute.edu.vn](mailto:23162107@student.hcmute.edu.vn)

### ARTICLE INFO

Received:

Revised:

Accepted:

Published:

### KEYWORDS

Intelligent traffic signal control;

Pedestrian behavior recognition;

Air quality monitoring;

Real-time camera processing;

Cloud-based traffic control.

### ABSTRACT

This paper presents the development of an intelligent traffic signal control system capable of responding adaptively based on pedestrian behavior and environmental conditions. The central research question addresses how to enhance safety and operational efficiency at urban intersections by detecting pedestrians who intend to cross the street and providing timely signal responses. The proposed system is designed on an Internet of Things architecture, integrating real-time image data from a camera module and air quality information from an environmental sensor. The collected data is transmitted and stored on a cloud-based platform to support decision-making in signal control. Simulation results demonstrate that the system can accurately detect hand-raising gestures indicative of a pedestrian's intention to cross, through geometric analysis of body posture, and adjust signal phases according to both pedestrian density and air pollution levels. The system exhibits fast response time and high reliability under various simulated traffic scenarios. These findings highlight the system's potential for deployment in smart city environments, contributing to improved pedestrian safety and the optimization of traffic flows at complex intersections.

## Hệ thống IoT điều chỉnh đèn tín hiệu giao thông dựa trên giám sát môi trường và nhận dạng hành vi giơ tay xin sang đường qua camera kết nối đám mây

Nguyễn Ngân Trường<sup>1\*</sup>, Bùi Nhật Thành<sup>1</sup>, Trần Doãn Hòa<sup>1</sup>, Nguyễn Quốc Tinh<sup>1</sup>, Nguyễn Văn Trường<sup>1</sup>

<sup>1</sup>Trường Đại học Sư phạm Kỹ thuật Thành phố Hồ Chí Minh, Việt Nam

\*Tác giả liên hệ. Email: [23162107@student.hcmute.edu.vn](mailto:23162107@student.hcmute.edu.vn)

### THÔNG TIN BÀI BÁO

Ngày nhận bài:

Ngày hoàn thiện:

Ngày chấp nhận đăng:

Ngày đăng:

### TỪ KHÓA

Điều khiển tín hiệu giao thông thông minh;

Nhận dạng hành vi người đi bộ;

Giám sát chất lượng không khí;

Camera thời gian thực;

Điều khiển tín hiệu giao thông qua nền tảng đám mây.

### TÓM TẮT

Bài báo này trình bày một nghiên cứu về phát triển hệ thống điều khiển tín hiệu giao thông thông minh có khả năng phản ứng linh hoạt dựa trên hành vi người đi bộ và điều kiện môi trường. Câu hỏi trung tâm đặt ra là làm thế nào để tăng mức độ an toàn và hiệu quả tại các nút giao thông đô thị, thông qua việc phát hiện người đi bộ có nhu cầu sang đường và phản hồi tín hiệu kịp thời. Hệ thống được thiết kế dựa trên nền tảng thiết bị IoT, kết hợp mô-đun xử lý hình ảnh từ camera và cảm biến chất lượng không khí, dữ liệu được truyền và lưu trữ trên nền tảng đám mây để phục vụ xử lý và ra quyết định điều khiển tín hiệu. Kết quả mô phỏng cho thấy hệ thống có khả năng nhận diện chính xác hành vi giơ tay xin sang đường thông qua phân tích tư thế cơ thể, đồng thời điều chỉnh thời lượng đèn tín hiệu một cách phù hợp với mật độ người đi bộ và mức độ ô nhiễm không khí. Hệ thống phản hồi tín hiệu nhanh, độ chính xác cao, cho thấy khả năng vận hành ổn định trong các kịch bản giao thông đa dạng. Nghiên cứu này mở ra tiềm năng ứng dụng trong các đô thị thông minh, góp phần nâng cao an toàn cho người đi bộ và tối ưu hóa lưu lượng phương tiện tại giao lộ.

Doi: <https://doi.org/10.54644/jtexxxxxx>

Copyright © JTE. This is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted use, distribution, and reproduction in any medium for non-commercial purpose, provided the original work is properly cited.

## 1. Giới thiệu

Trong bối cảnh đô thị hóa-hiện đại hóa, việc đảm bảo an toàn cho người đi bộ tại các điểm giao cắt đang trở thành một yêu cầu cấp bách, đặc biệt là tại những khu vực thiếu sự hiện diện của hệ thống đèn tín hiệu giao thông thông minh. Theo thống kê của tổ chức Y tế Thế giới, số lượng người đi bộ tử vong trong các vụ tai nạn giao thông tại khu vực thành thị vẫn đang ở mức cao, chiếm tỷ lệ đáng kể trong tổng số ca tử vong do tai nạn giao thông toàn cầu [1]. Bên cạnh đó, tình trạng ô nhiễm không khí tại các điểm giao thông đông đúc cũng ngày càng nghiêm trọng, ảnh hưởng trực tiếp đến sức khỏe cộng đồng và chất lượng không khí [2]. Do đó, việc phát triển một hệ thống giám sát môi trường tích hợp điều khiển giao thông thông minh dựa trên hành vi người dùng là một nhu cầu thiết yếu.

Nhiều nghiên cứu đã được thực hiện nhằm xây dựng các hệ thống hỗ trợ sang đường cho người đi bộ dựa trên phân tích hình ảnh và các công nghệ cảm biến khác nhau. Trong [3], tác giả đề xuất một mô hình nhận diện người đi bộ sử dụng camera hồng ngoại và mạng nơ-ron tích chập (CNN), tuy nhiên, hệ thống này đòi hỏi năng lực tính toán cao, chưa phù hợp với môi trường biên. Một hướng tiếp cận khác trong [4] khai thác dữ liệu từ cảm biến áp suất trên mặt đường để phát hiện người đi bộ, song thiếu khả năng nhận diện các tín hiệu phi ngôn ngữ như hành vi giơ tay – một hành vi phổ biến trong ngữ cảnh văn hóa giao thông đô thị. Ngoài ra, các nghiên cứu [5]-[7] tập trung vào tối ưu hóa luồng giao thông dựa trên dữ liệu thời gian thực nhưng chưa có sự tích hợp giữa yếu tố môi trường và hành vi người tham gia giao thông.

Từ những khoảng trống đó, nghiên cứu này đề xuất một mô hình hệ thống IoT thông minh có khả năng nhận diện hành vi “giơ tay xin sang đường” thông qua camera tích hợp ESP32-CAM với cảm biến OV2640 sử dụng mô hình YOLOv8-pose. Bên cạnh đó, hệ thống cũng đồng thời giám sát chất lượng không khí tại điểm chờ bằng cảm biến MQ135 nhằm cung cấp dữ liệu đầu vào phục vụ điều chỉnh trạng thái đèn tín hiệu một cách linh hoạt. Toàn bộ dữ liệu bao gồm hình ảnh, kết quả nhận diện, chỉ số môi trường và thời gian được đồng bộ hóa và lưu trữ trên nền tảng đám mây, hỗ trợ việc phân tích, thống kê và tối ưu hóa sau này.

Cách tiếp cận được đề xuất kết hợp công nghệ xử lý ảnh, nhận diện tư thế người, phân tích hành vi và giám sát môi trường thời gian thực, phù hợp triển khai trên thiết bị có tài nguyên hạn chế, đặc biệt thích hợp với điều kiện đô thị ở các nước đang phát triển. Các kết quả thực nghiệm ban đầu cho thấy hệ thống có khả năng phản ứng linh hoạt với các tình huống giao thông thực tế, đồng thời góp phần nâng cao mức độ an toàn và thân thiện cho người đi bộ tại các điểm giao cắt.

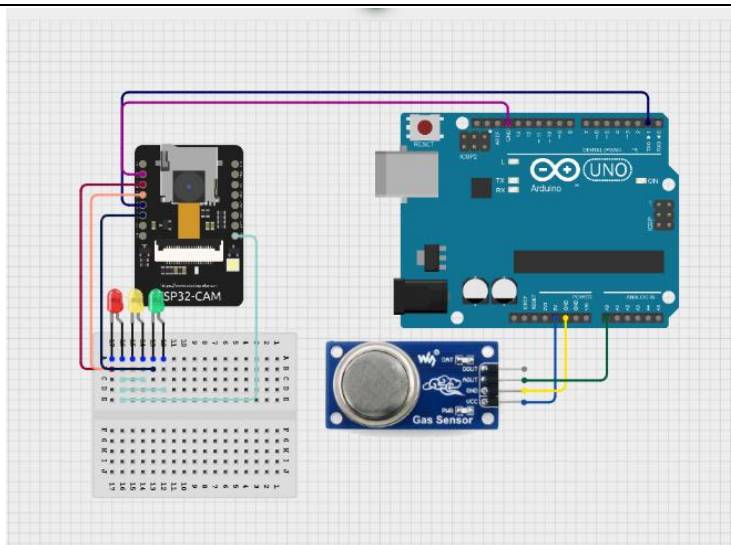
## 2. Thiết kế hệ thống

Thiết kế phần cứng và phần mềm được trình bày trong phần này. Tiếp theo đó là thảo luận quy trình hoạt động của hệ thống để đáp ứng yêu cầu làm việc.

### 2.1. Cấu trúc phần cứng

Hệ thống đèn tín hiệu giao thông thông minh được thiết kế theo cấu trúc đơn giản nhưng hiệu quả. Mô-đun ESP32-CAM, tích hợp cảm biến hình ảnh kỹ thuật số CMOS OV2640, được sử dụng để thu nhận hình ảnh theo thời gian thực và đồng thời đóng vai trò như một bộ xử lý trung tâm của toàn hệ thống. Bên cạnh đó, hệ thống cũng sử dụng cảm biến chất lượng không khí MQ135, được kết nối và xử lý sơ bộ thông qua mạch Arduino UNO, nhằm giám sát nồng độ các chất khí tại khu vực chờ.

Các thành phần này được bố trí vị trí cao trên cột đèn giao thông để đảm bảo góc quan sát rộng và độ chính xác cao trong việc thu thập dữ liệu. Cơ sở lý thuyết về cảm biến MQ135 được trình bày trong [8]; nguyên lý hoạt động và cấu trúc của mô-đun ESP32-CAM được đề cập chi tiết tại [9]. Sơ đồ mạch được thể hiện qua *Hình 1*; mô phỏng vị trí lắp đặt các thành phần trong thực tế được thể hiện qua *Hình 2*.



**Hình 1.** Sơ đồ mạch kết nối các thành phần



**Hình 2.** Mô phỏng vị trí lắp đặt phần cứng trong thực tế

## 2.2. Thiết kế thuật toán phát hiện người đi bộ

Trong bối cảnh các hệ thống nhúng như ESP32-CAM gặp hạn chế bởi tài nguyên tính toán và bộ nhớ, việc triển khai các mô hình học sâu phức tạp để nhận diện hành vi con người trở nên khó khăn. Do đó, nghiên cứu này đề xuất một cách tiếp cận nhẹ, linh hoạt và phù hợp với thiết bị tài nguyên thấp bằng cách khai thác đặc trưng hình học từ tư thế con người, thay vì dựa hoàn toàn vào học sâu hoặc phân loại hình ảnh truyền thống. Cụ thể, phương pháp nhận diện hành vi “giơ tay xin sang đường” được xây dựng dựa trên phân tích tọa độ của các điểm đặc trưng (keypoints) từ mô hình nhận dạng tư thế người.

Công cụ chính được sử dụng trong nghiên cứu là mô hình YOLOv8-pose. Mô hình này có khả năng phát hiện chính xác các điểm như vai, cổ, đầu gối, mắt cá chân, từ đó cho phép suy luận về hành vi

người dùng thông qua các quan hệ hình học giữa các điểm trên cơ thể [10]. Thuật toán phát hiện người đi bộ giờ tay được xây dựng dựa trên hai tiêu chí chính:

- Đặc điểm 1: Giờ tay. Hành vi giờ tay được suy luận dựa trên việc so sánh vị trí tương đối giữa cổ tay và vai, Trên hệ trục tọa độ ảnh trong OpenCV bắt đầu từ góc bên trái, bao gồm trục x (từ trái sang phải) và trục y (từ trên xuống dưới) [11]. Nghĩa là vật càng ở trên thì y càng nhỏ. Do đó, nếu tọa độ y của cổ tay nhỏ hơn y của vai, ta có thể kết luận rằng tay đang giờ lên:

$$y_{cổ tay} < y_{vai}$$

Trong trường hợp đặc điểm 1, hai điều kiện này được xét riêng biệt cho tay trái và tay phải, chỉ cần một trong hai tay thỏa mãn điều kiện thì đối tượng được gán nhãn có giờ tay.

- Đặc điểm 2: Tư thế người đi bộ. Một người đi bộ thường có tư thế đứng thẳng, đồng thời mắt cá chân nằm gần sát đáy của hộp giới hạn (bounding box). Dựa trên quan sát thực nghiệm và nghiên cứu của các tác giả từ [12], mắt cá chân của người đi bộ thường có tọa độ y cách cạnh dưới bounding box không quá một ngưỡng  $\epsilon$ :

$$y_{mắt cá chân} > y_{cạnh dưới bounding box} - \epsilon, \text{ với } \epsilon = 50 \text{ px (hoặc một ngưỡng cố định phù hợp)}$$

Trong trường hợp đặc điểm 2, cả hai mắt cá chân trái và phải đều được xét; nếu một trong hai thỏa mãn điều kiện thì đối tượng được gán nhãn là đang trong tư thế đi bộ.

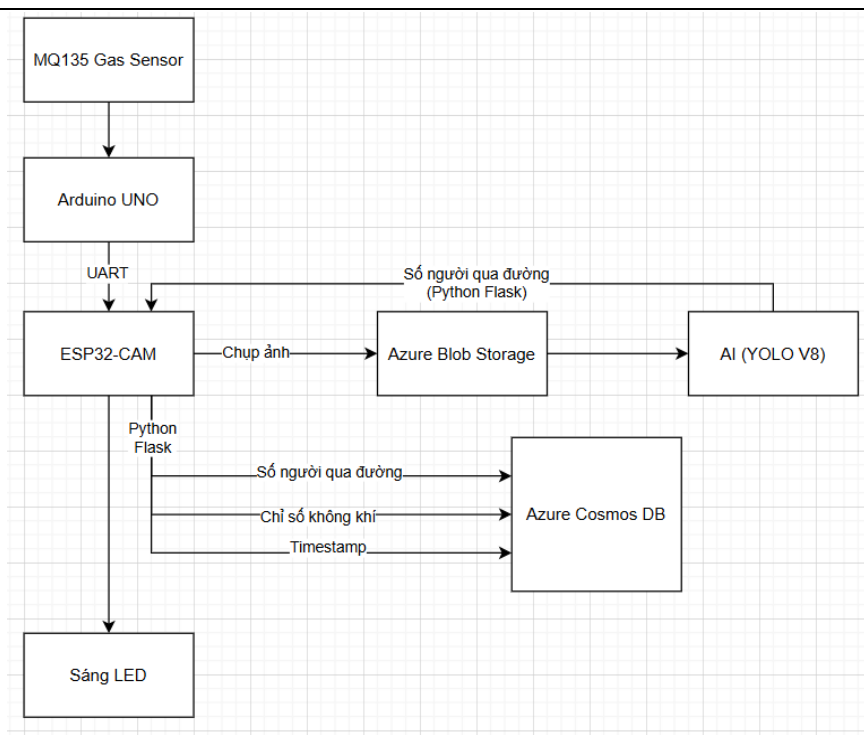
Đối tượng được xác định là “người đi bộ giơ tay xin sang đường” nếu đồng thời thỏa cả hai điều kiện trên. Phân tích này hoàn toàn không yêu cầu huấn luyện mô hình mới mà chỉ dựa vào suy luận hình học sau bước suy luận tư thế, do đó phù hợp cho hệ thống thời gian thực và thiết bị biên.

### 2.3. Quy trình hoạt động

Hệ thống được triển khai dưới dạng nguyên mẫu vật lý gồm ba thành phần chính: cảm biến chất lượng không khí MQ135, mô-đun xử lý trung tâm ESP32-CAM, và hệ thống xử lý dữ liệu trên nền tảng đám mây. MQ135 đảm nhận vai trò giám sát nồng độ các chất khí tại khu vực chờ và truyền dữ liệu về vi điều khiển Arduino UNO để xử lý sơ bộ. Dữ liệu sau đó được gửi qua giao tiếp UART đến ESP32-CAM, thiết bị trung tâm này vừa ghi nhận dữ liệu cảm biến vừa chụp ảnh thời gian thực tại khu vực vạch kẻ đường.

Các hình ảnh thu được sẽ được truyền lên dịch vụ Azure Blob Storage và được phân tích bởi mô hình học sâu YOLOv8 nhằm nhận diện và đếm số người đi bộ. Kết quả phân tích được gửi trả về cho ESP32-CAM, nơi dữ liệu được tổng hợp với chỉ số môi trường và timestamp, sau đó lưu trữ vào Azure Cosmos DB để phục vụ phân tích sau này. Hệ thống được điều khiển bằng API viết bằng Python Flask. Dựa trên thông tin thu thập, ESP32-CAM ra quyết định điều chỉnh trạng thái đèn LED tương ứng với tình huống thực tế.

Sơ đồ biểu diễn luồng dữ liệu được thể hiện qua Hình 3.



Hình 3. Sơ đồ biểu diễn luồng dữ liệu của hệ thống

### 3. Kết quả và hướng phát triển

Các kết quả thí nghiệm nhằm đánh giá hiệu quả hoạt động của hệ thống đèn tín hiệu giao thông thông minh được trình bày trong phần này. Hệ thống được kiểm chứng cả trong môi trường mô phỏng và thực tế, bao gồm khả năng nhận diện hành vi giao tay, phản hồi tín hiệu đèn và xử lý dữ liệu môi trường

#### 3.1. Kiểm chứng kết quả trong môi trường mô phỏng

##### a) Kiểm chứng khả năng kết nối và thu nhận dữ liệu thời gian thực

Các mô-đun thành phần bao gồm cảm biến MQ135 và camera ESP32-CAM đã được cấu hình kết nối thông qua mạng nội bộ Wi-Fi (2.4GHz). Dữ liệu chất lượng không khí và hình ảnh được đẩy thành công lên Azure Blob Storage thông qua giao diện Python Flask. Mô hình YOLOv8 được triển khai trên server hỗ trợ GPU đã thực hiện xử lý và trả về kết quả nhận dạng độ trễ trung bình khoảng 0.7 giây mỗi ảnh.

##### b) Đánh giá độ chính xác của hành vi “giao tay xin sang đường”

Trong bối cảnh chưa có bộ dữ liệu công khai với nhãn hành vi “giao tay xin sang đường” một cách rõ ràng, nghiên cứu sử dụng phương pháp suy luận hành vi dựa trên mô hình nhận diện tư thế người (YOLOv8-pose) kết hợp với phân tích quan hệ hình học giữa các điểm đặc trưng (keypoints). Theo đó, hành vi “giao tay” được xác định khi tạo độ theo trục tung của cổ tay nhỏ hơn vai cùng phía, tức là cổ tay nằm cao hơn vai trong không gian hình ảnh, phù hợp với hệ tọa độ ảnh trong OpenCV.

Để đánh giá độ chính xác của thuật toán suy luận hành vi, nhóm nghiên cứu đã xây dựng một tập dữ liệu mô phỏng gồm 200 ảnh đầu vào được trích chọn từ các video thực nghiệm, trong đó, người đi bộ thực hiện các tư thế khác nhau. Mỗi ảnh chứa từ 1-3 đối tượng người, và hành vi “giao tay” được gán nhãn tự động theo tiêu chí hình học nêu trên.

Kết quả đánh giá được trình bày trong Bảng 1 cho thấy thuật toán đạt độ chính xác tương đối cao trong việc phát hiện hành vi giao tay. Tổng cộng có 110 trường hợp “giao tay” được phát hiện trong bộ dữ liệu, trong đó có 97 trường hợp được nhận diện đúng, tương ứng với độ chính xác (Precision) và độ thu hồi (Recall) đều đạt 88.2%. Các chỉ số này cho thấy tính khả thi của thuật toán trong môi trường thử nghiệm, dù chưa được kiểm chứng trên bộ dữ liệu chuẩn hóa.



**Bảng 1.** Hiệu suất nhận diện hành vi “giơ tay xin sang đường” trên tập ảnh mô phỏng

ID	Số ảnh đầu vào	Số người giơ tay	Nhận diện đúng	Nhận diện sai	Precision (%)	Recall (%)	F1-Score (%)
Subset_1	100	52	47	5	90.4	90.4	90.4
Subset_2	100	58	50	8	86.2	86.2	86.2
Tổng	200	110	97	13	88.2	88.2	88.2

*c) Phản hồi tín hiệu đèn theo nhận diện hành vi người đi bộ*

Trong quá trình mô phỏng, hệ thống nhận diện hành vi giơ tay của người đi bộ thông qua YOLOv8-pose và thực hiện điều khiển thời gian pha đèn đỏ dựa trên số lượng người được phát hiện. Cụ thể:

- Phát hiện 1-2 người: đèn đỏ kéo dài 5 giây.
- Phát hiện 3-5 người: đèn đỏ kéo dài 8 giây.
- Trên 5 người: đèn đỏ kéo dài 10 giây.

Trong tất cả 12 kịch bản mô phỏng khác nhau được thiết lập, hệ thống đều thực hiện điều chỉnh đèn đúng logic và đưa ra phản hồi trong thời gian không tới 2 giây kể từ khi hành vi giơ tay được phát hiện. Điều này cho thấy độ ổn định và khả năng đáp ứng thời gian thực của hệ thống trong môi trường mô phỏng.

*d) Ảnh hưởng của chất lượng không khí đến quyết định của đèn*

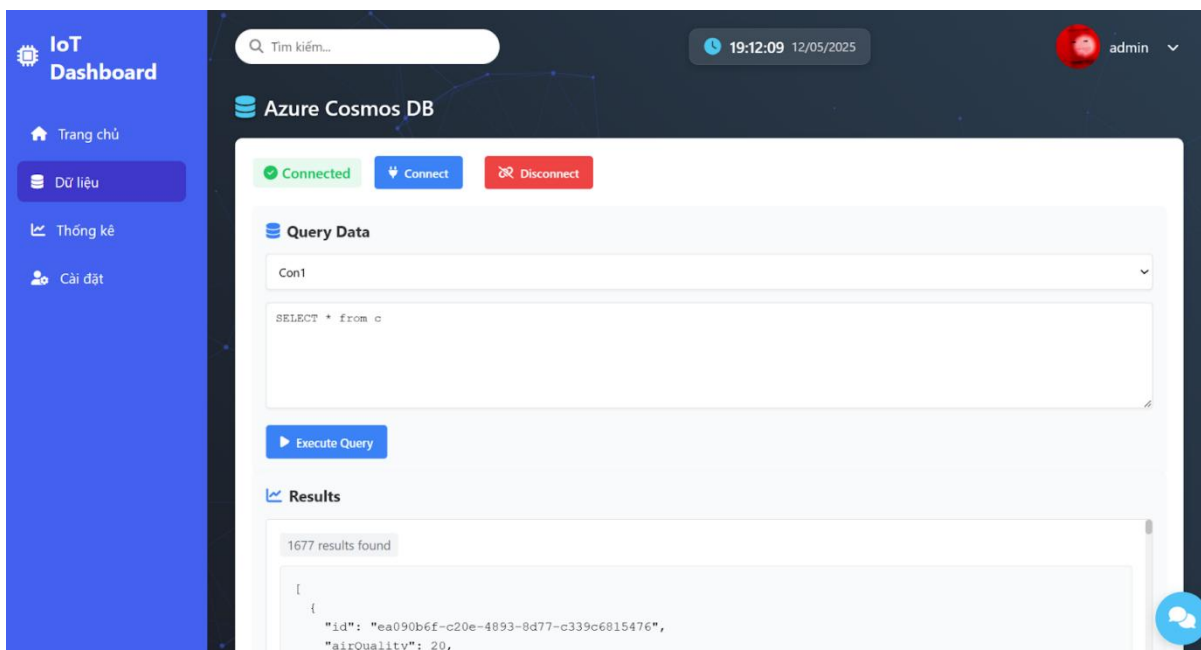
Cảm biến MQ135 được tích hợp trong mô phỏng với dữ liệu biến thiên chất lượng không khí (AQI) trong khoảng 20-60 đơn vị. Trong kịch bản mô phỏng, khi giá trị AQI vượt ngưỡng 30, tương ứng với mức chất lượng không khí kèm theo chuẩn đánh giá nội bộ, hệ thống tự động thực hiện:

- Tăng thời lượng đèn xanh thêm 20% so với mức bình thường để giảm thời gian phơi nhiễm của người đi bộ.
- Kích hoạt tín hiệu cảnh báo LED nhấp nháy với tần suất 2Hz.

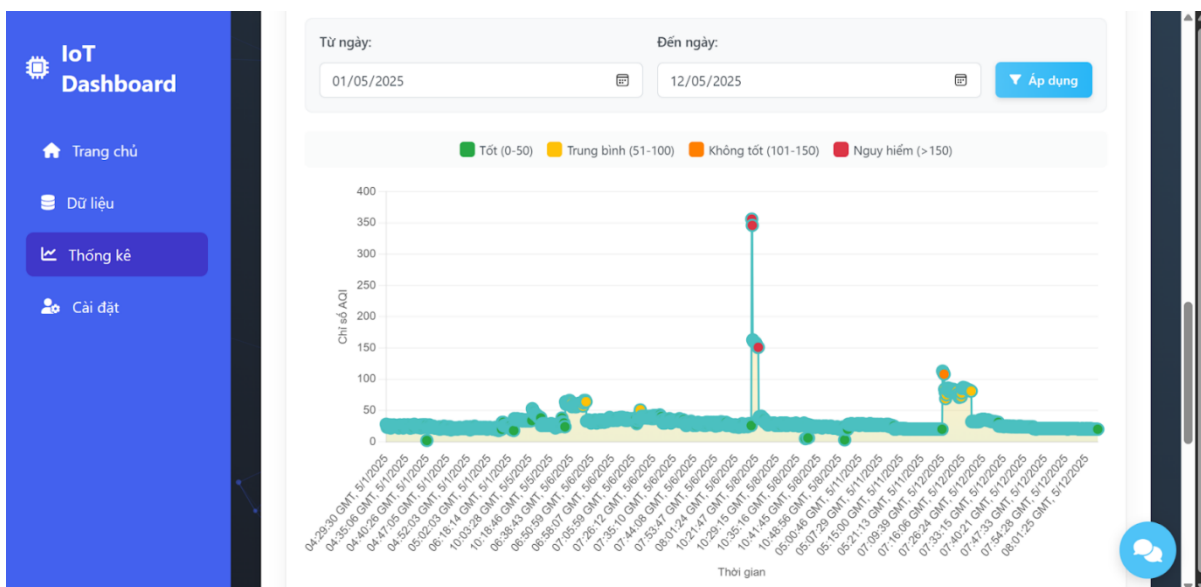
Kết quả mô phỏng cho thấy hệ thống phản hồi chính xác trong toàn bộ 8/8 kịch bản có xuất hiện AQI vượt ngưỡng, đảm bảo logic điều khiển được thực thi đầy đủ và nhanh chóng.

*e) Kiểm chứng khả năng truy xuất dữ liệu thông qua giao diện người dùng*

Thông qua giao diện web được xây dựng, ta có thể truy xuất cơ sở dữ liệu bằng tài khoản admin (Hình 4 và 5).



Hình 4. Giao diện Front-end truy xuất cơ sở dữ liệu



Hình 5. Kết quả thử nghiệm truy xuất cơ sở dữ liệu

### 3.2. Hướng phát triển

Một định hướng phát triển trọng tâm của hệ thống đèn giao thông thông minh được đề xuất là việc tăng cường các biện pháp đảm bảo an toàn và ưu tiên cho đối tượng trẻ em. Nghiên cứu này sẽ tập trung vào việc tích hợp các module camera chuyên dụng, đồng thời phát triển và thẩm định các thuật toán thị giác máy tính tiên tiến, với khả năng nhận diện trẻ em dựa trên các tham số nhân trắc học như chiều cao ước tính và các đặc điểm hình thái khác. Dữ liệu thu thập không chỉ phục vụ mục đích phát hiện tức thời mà còn được phân tích theo trục thời gian và không gian để hình thành một cơ sở dữ liệu động, định lượng hóa "các khu vực có mật độ trẻ em cao" (ví dụ, gần trường học, khu vui chơi). Dựa trên cơ sở dữ liệu này, một thuật toán điều khiển thích ứng sẽ được phát triển, cho phép hệ thống tự động điều chỉnh

chu kỳ tín hiệu, cụ thể là kéo dài pha đèn xanh cho người đi bộ khi phát hiện trẻ em, đồng thời kích hoạt các cơ chế cảnh báo phụ trợ (trực quan, âm thanh) nhằm nâng cao nhận thức của người điều khiển phương tiện. Mục tiêu của hướng nghiên cứu này là định lượng hóa sự cải thiện về an toàn giao thông cho trẻ em tại các nút giao được trang bị hệ thống.

Định hướng phát triển thứ hai tập trung vào việc nâng cao năng lực nhận diện và phân tích đa chiều đối với phương tiện cơ giới. Hệ thống được đề xuất mở rộng để không chỉ thực hiện chức năng đếm lưu lượng mà còn tiến hành phân tích sâu các thuộc tính của phương tiện, bao gồm vận tốc tức thời, phân loại phương tiện, mật độ lưu thông và các đặc trưng hành vi. Quá trình này đòi hỏi việc tích hợp các cảm biến tiên tiến và phát triển các mô hình học máy có khả năng xử lý và diễn giải "dữ liệu vận tốc và hành trình của từng phương tiện" với độ chi tiết cao. Kết quả phân tích sẽ là cơ sở để thiết lập "hệ thống cảnh báo sớm cho người đi bộ" về các nguy cơ tiềm ẩn từ luồng phương tiện, cũng như dự báo "các khoảng thời gian và vị trí có nguy cơ ùn tắc hoặc tai nạn cao". Hơn nữa, nghiên cứu sẽ khảo sát tính khả thi và hiệu quả của việc tích hợp một "hệ thống giám sát và xử lý vi phạm tự động" (ví dụ: phạt nguội) dựa trên dữ liệu phân tích, nhằm mục tiêu cải thiện ý thức tuân thủ luật giao thông và giảm thiểu các hành vi lái xe nguy hiểm.

Định hướng phát triển thứ ba, mang tính tổng hợp và then chốt, là việc xây dựng một kiến trúc trí tuệ nhân tạo (AI) trung tâm, có khả năng tích hợp và đồng xử lý thông tin từ nhiều nguồn dữ liệu đa dạng nhằm cung cấp các cảnh báo và hướng dẫn giao thông thông minh. Kiến trúc AI này sẽ được thiết kế để xử lý đồng thời dữ liệu về trẻ em và phương tiện (như đã mô tả ở trên), cùng với "dữ liệu phân tích chất lượng không khí từ các cảm biến môi trường", cho phép đưa ra "các cảnh báo toàn diện" về mức độ ô nhiễm. Thêm vào đó, AI sẽ thực hiện "phân tích sự xuất hiện và tần suất của người đi bộ" nói chung để tối ưu hóa các quyết định điều khiển tín hiệu và cung cấp các cảnh báo phù hợp (bao gồm cả các cảnh báo liên quan đến điều kiện khí tượng cục bộ nếu hệ thống được mở rộng module cảm biến tương ứng). Chức năng cốt lõi của AI trung tâm là cung cấp "hướng dẫn di chuyển tối ưu cho người đi bộ qua đường", dựa trên một mô hình quyết định đa yếu tố, cân nhắc đồng thời các khía cạnh an toàn, hiệu quả lưu thông và phơi nhiễm môi trường, nhằm nâng cao chất lượng trải nghiệm giao thông đô thị.

#### **4. Kết luận**

Nghiên cứu đã hiện thực hóa thành công một hệ thống điều khiển đèn tín hiệu giao thông thông minh, có khả năng xử lý đồng thời dữ liệu thị giác và môi trường trong thời gian thực. Hệ thống sử dụng mô-đun ESP32-CAM để thu thập hình ảnh khu vực giao thông, kết hợp với cảm biến MQ135 để đo chất lượng không khí, tạo nên nguồn dữ liệu đa chiều được truyền tải và lưu trữ trên nền tảng đám mây Azure Cosmos DB. Trên cơ sở đó, hệ thống thực hiện điều khiển linh hoạt pha đèn giao thông theo mật độ người đi bộ và điều kiện môi trường xung quanh.

Kết quả mô phỏng cho thấy hệ thống có khả năng phát hiện hiệu quả hành vi "giơ tay xin sang đường" – một chỉ báo trực quan thường gặp tại các nút giao thông không có công an điều tiết, đồng thời phản hồi kịp thời bằng cách điều chỉnh thời lượng đèn tín hiệu. Ngoài ra, việc tích hợp thông tin chất lượng không khí góp phần tăng cường yếu tố an toàn, đặc biệt trong các khung giờ mật độ cao hoặc khu vực ô nhiễm.

Đóng góp cốt lõi của nghiên cứu nằm ở cách tiếp cận điều khiển tín hiệu dựa trên sự phối hợp của nhiều nguồn dữ liệu khác nhau, vượt qua các giới hạn của các hệ thống truyền thống chỉ dựa trên đồng hồ hẹn giờ hoặc cảm biến đơn lẻ. Với kiến trúc IoT linh hoạt và khả năng kết nối đám mây, hệ thống có tiềm năng triển khai mở rộng trong các đô thị thông minh, đóng vai trò như một thành phần trong hạ tầng giao thông bền vững.

Trong tương lai, nghiên cứu sẽ tập trung vào ba định hướng: (1) phát triển mô-đun nhận diện và ưu tiên đối tượng trẻ em, (2) mở rộng năng lực phân tích thuộc tính phương tiện giao thông và hành vi lái xe, (3) tích hợp kiến trúc AI trung tâm đồng xử lý dữ liệu hành vi – môi trường – phương tiện để đưa ra các quyết định điều khiển tối ưu. Các hướng này hứa hẹn nâng cao hơn nữa tính chủ động, độ chính xác và mức độ thích ứng của hệ thống trong các tình huống giao thông phức tạp.

#### **Lời cảm ơn**



Nhóm nghiên cứu xin trân trọng cảm ơn thầy Nguyễn Xuân Sâm, giảng viên Trường Đại học Sư phạm Kỹ thuật Thành phố Hồ Chí Minh, đã tận tình hướng dẫn và đồng hành trong suốt quá trình thực hiện đề tài. Những góp ý chuyên môn và định hướng nghiên cứu của Thầy là nền tảng quan trọng giúp công trình đạt được kết quả như mong đợi.

### Xung đột lợi ích

Các tác giả tuyên bố không có xung đột lợi ích tổng bài báo này.

### TÀI LIỆU THAM KHẢO

- [1] World Health Organization, "Global status report on road safety 2018," WHO Press, Geneva, 2018. [Online]. Available: <https://www.who.int/publications/i/item/9789241565684>
- [2] A. M. Fiore, et al., "Multimodel estimates of intercontinental source-receptor relationships for ozone pollution," *J. Geophys. Res. Atmos.*, vol. 114, no. D4, pp. 1–20, 2009, doi: 10.1029/2008JD010816
- [3] H. Zhang, Y. Liu, and L. Wang, "Pedestrian detection using deep learning neural networks," in *Proc. 14th Int. Conf. Computer Vision Theory and Applications (VISAPP)*, Prague, Czech Republic, Feb. 2019, pp. 180–187. doi: 10.5220/0007560301800187
- [4] K. S. Raghuvanshi and S. Patil, "Sensor-based smart pedestrian crossing system for smart cities," *Int. J. Recent Technol. Eng. (IJRTE)*, vol. 8, no. 4, pp. 4847–4851, Nov. 2019. doi: 10.35940/ijrte.D7560.118419
- [5] D. Lee, J. Kim, and H. Myung, "Real-time pedestrian intention prediction using a convolutional neural network," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 11, pp. 3962–3974, Nov. 2019. doi: 10.1109/TITS.2018.2878897.
- [6] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, "It's not all about size: On the role of data properties in pedestrian detection," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Munich, Germany, Sept. 2018, pp. 85–101. doi: 10.1007/978-3-030-01237-3\_6.
- [7] M. Kuutti, R. Bowden, Y. Jin, P. Barber, and S. Fallah, "A survey of deep learning applications to autonomous vehicle control," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 2, pp. 712–733, Feb. 2021. doi: 10.1109/TITS.2019.2962338.
- [8] Hanwei Electronics Co., Ltd., *MQ135 Gas Sensor Technical Data*, Zhengzhou, China, 2013. [Online]. Available: <https://www.winsen-sensor.com/d/files/PDF/MQ135.pdf>
- [9] Ai-Thinker, *ESP32-CAM Development Board Module Datasheet*, Ai-Thinker Technology Co., Ltd., Shenzhen, China, 2020. [Online]. Available: <https://docs.ai-thinker.com/en/esp32-cam..>
- [10] Ultralytics, "Pose Estimation with YOLOv8," *Ultralytics Documentation*. [Online]. Available: <https://docs.ultralytics.com/tasks/pose/#val>.
- [11] X. Zhou, V. Koltun, and P. Krähenbühl, "Tracking Objects as Points," in *Proc. of the European Conf. on Computer Vision (ECCV)*, Glasgow, UK, 2020, pp. 474–490.
- [12] G. Bradski and A. Kaehler, *Learning OpenCV 4: Computer Vision with Python and C++*, 1st ed. Sebastopol, CA, USA: O'Reilly Media, 2019.



**Nguyen Ngan Truong**, majoring in Information Security at Ho Chi Minh City University of Technology and Education, Vietnam.  
Email: [23162107@student.hcmute.edu.vn](mailto:23162107@student.hcmute.edu.vn).



**Nguyen Van Truong**, majoring in Information Security at Ho Chi Minh City University of Technology and Education, Vietnam.  
Email: [23162108@student.hcmute.edu.vn](mailto:23162108@student.hcmute.edu.vn).



**Tran Doan Hoa**, majoring in Information Security at Ho Chi Minh City University of Technology and Education, Vietnam.  
Email: [23162026@student.hcmute.edu.vn](mailto:23162026@student.hcmute.edu.vn).



**Bui Nhat Thanh**, majoring in Information Security at Ho Chi Minh City University of Technology and Education, Vietnam.  
Email: [23162091@student.hcmute.edu.vn](mailto:23162091@student.hcmute.edu.vn).



**Nguyen Quoc Tinh**, majoring in Information Security at Ho Chi Minh City University of Technology and Education, Vietnam.  
Email: [23162101@student.hcmute.edu.vn](mailto:23162101@student.hcmute.edu.vn).