Tạp chí Khoa học Đại học Cần Thơ
Số chuyên đề: Khoa học và Công nghệ cho sự phát triển bền vững của vùng Đồng bằng sông Cửu Long
ISSN 1859-2333 | e-ISSN 2815-5599

DOI:10.22144/ctujos.2026.081

NGHIÊN CỨU VÀ THIẾT KẾ HỆ THỐNG GIÁM SÁT AN NINH TÍCH HỢP IOT

Võ Thị Phương Loan[1], Lê Thị Bảo Trâm[1], Nguyễn Đăng Khoa[1], Phan Bình Minh[2] và Nguyễn Đình Tứ[1*]

[1]*Khoa Kỹ thuật Cơ Khí, Trường Đại học Kỹ thuật – Công nghệ Cần Thơ, Việt Nam*
[2]*Công ty Nhiệt điện Cần Thơ, Việt Nam*
**Tác giả liên hệ (Corresponding author): ndtu@ctuet.edu.vn*

Thông tin chung (Article Information)

Nhận bài (Received): 26/01/2026
Sửa bài (Revised): 01/03/2026
Duyệt đăng (Accepted): 10/05/2026

Title: *Development of an IoT-Based Integrated Security Surveillance System*

Author(s): *Vo Thi Phuong Loan[1], Le Thi Bao Tram[1], Nguyen Dang Khoa[1], Phan Binh Minh[2] and Nguyen Dinh Tu[1*]*

Affiliation(s)*: [1]Faculty of Mechanical Engineering, Can Tho University of Technology, Viet Nam; [2]Can Tho Thermal Power Company, Viet Nam*

TÓM TẮT

Một hệ thống an ninh thông minh, tập trung vào giải quyết các hạn chế của hệ thống truyền thống, tăng cường khả năng phát hiện và cảnh báo sớm các mối đe dọa được đề xuất trong nghiên cứu. Hệ thống đề xuất bao gồm các thành phần chính như: thiết bị sử dụng mô-đun cho Internet vạn vật (IoT – Internet of Things), camera, thuật toán để nhận dạng đối tượng, công nghệ RFID kiểm soát quyền truy cập vào các khu vực khác nhau và phần mềm quản lý tập trung để hỗ trợ hoạt động bảo mật, giám sát. Kết quả cho thấy hệ thống có khả năng cải thiện tính toàn diện cho quá trình bảo mật trong thời gian thực có độ chính xác trên 80%, với các tính năng như: cải thiện độ chính xác, tối ưu hóa tốc độ xử lý và cung cấp giải pháp quản lý tập trung qua phần mềm. Qua đó, khẳng định tiềm năng của ứng dụng công nghệ IoT và trí tuệ nhân tạo (AI) trong lĩnh vực an ninh, đồng thời mở ra cơ hội phát triển các giải pháp bảo mật hiện đại hơn trong tương lai.

Từ khóa: *Công nghệ IoT, hệ thống an ninh, nhận dạng đối tượng, phần mềm quản lý, trí tuệ nhân tạo*

ABSTRACT

This study proposes an intelligent security system designed to overcome the limitations of traditional frameworks while enhancing threat detection and early warning capabilities. The proposed architecture integrates key components, including Internet of Things (IoT) modules, surveillance cameras, object recognition algorithms, RFID technology for multi-zone access control, and management software to streamline security operations and monitoring. Experimental results demonstrate that the system significantly improves the comprehensiveness of security protocols of real-time environments, offering enhanced access control efficiency and timely alerts upon incident occurrence. This confirms the transformative potential of IoT and Artificial Intelligence (AI) in the security sector, providing a foundation for the development of more progressive security solutions in the future.

Keywords*: Artificial intelligence, IoT technology, management software, object detection, security system*

1. GIỚI THIỆU

Trong bối cảnh đô thị hóa và phát triển kinh tế mạnh mẽ, vấn đề bảo đảm an ninh và an toàn tài sản đã trở thành ưu tiên hàng đầu của các cá nhân và tổ chức. Theo báo cáo của IBM (2023), chi phí trung bình cho các vụ vi phạm dữ liệu và an ninh vật lý đang có xu hướng tăng cao, khiến nhiều doanh nghiệp phải đối mặt với rủi ro tài chính nghiêm trọng. Tại Việt Nam, thị trường thiết bị an ninh thông minh cũng đang chứng kiến sự bùng nổ mạnh mẽ. Theo Statista (2024) và 6Wresearch (2020-2026), doanh thu từ thị trường này dự kiến đạt mức tăng trưởng kép hàng năm (CAGR) trên 10%, phản ánh nhu cầu cấp thiết về các giải pháp giám sát hiện đại thay thế cho các phương pháp bảo vệ truyền thống vốn bộc lộ nhiều hạn chế về tính linh hoạt và khả năng phản ứng thời gian thực (Arun et al., 2023).

Các hệ thống giám sát truyền thống chủ yếu dựa vào việc ghi hình thụ động và sự giám sát của con người, dẫn đến tình trạng chậm trễ trong việc phát hiện xâm nhập và dễ xảy ra sai sót do mệt mỏi hoặc thiếu tập trung (Abdullah & Jabber, 2018). Sự ra đời của Internet vạn vật (IoT) đã mở ra một kỷ nguyên mới cho phép kết nối các thiết bị cảm biến và camera thành một mạng lưới thông minh, hỗ trợ truyền tin tức thời và giám sát từ xa (Jyothi & Vardhan, 2016; Manu et al., 2021). Các nghiên cứu của Lulla et al. (2021) hay Sattaru et al. (2023) đã khẳng định rằng việc tích hợp IoT vào hệ thống an ninh không chỉ nâng cao khả năng quản lý mà còn giảm thiểu đáng kể chi phí vận hành cho các hộ gia đình và doanh nghiệp nhỏ.

Tuy nhiên, thách thức lớn nhất của các hệ thống camera hiện nay là khả năng phân tích nội dung hình ảnh để phân biệt giữa đối tượng gây nguy hiểm và các yếu tố môi trường (như động vật hay cây cối). Sự phát triển của trí tuệ nhân tạo AI (Artificial Intelligence), đặc biệt là các thuật toán học sâu (Deep Learning), đã giải quyết hiệu quả bài toán này. Trong đó, dòng thuật toán YOLO (You Only Look Once) nổi lên như một giải pháp tối ưu nhờ sự cân bằng giữa độ chính xác và tốc độ xử lý (Chatterjee et al., 2024). Đặc biệt, phiên bản YOLOv8 – thế hệ mới nhất của dòng YOLO – đã chứng minh được hiệu suất vượt trội trong việc nhận dạng đa đối tượng trong các điều kiện ánh sáng phức tạp (Motwani & Soumya, 2023; Swathi & Challa, 2024). Việc ứng dụng YOLOv8 vào giám sát an ninh cho phép hệ thống tự động nhận diện con người và kích hoạt cảnh báo chỉ khi có sự xâm nhập thực sự (Sudharson et al., 2023; Sanjalawe et al., 2024; Siva et al., 2025).

Bên cạnh việc nhận dạng hình ảnh, kiểm soát truy cập vật lý cũng là một tầng bảo mật không thể thiếu. Công nghệ nhận dạng qua tần số vô tuyến RFID (Radio Frequency Identification) đã được chứng minh là một giải pháp hiệu quả và chi phí thấp để quản lý quyền ra vào tại các khu vực trọng yếu (Farooq et al., 2014). Sự kết hợp giữa nhận dạng khuôn mặt/đối tượng bằng AI và kiểm soát thẻ từ RFID tạo nên một hệ thống bảo mật đa nhân tố, giúp ngăn chặn triệt để các hành vi xâm nhập trái phép (Jesus et al., 2019; Al-E'mari et al., 2024).

Mặc dù đã có nhiều nghiên cứu về an ninh thông minh, nhưng việc xây dựng một hệ thống tích hợp đầy đủ từ phần cứng (vi điều khiển hiệu năng cao như ESP32), phần mềm giao diện thân thiện (Qt6 Designer), đến thuật toán AI tiên tiến và khả năng cảnh báo đa kênh (Telegram, cuộc gọi, loa báo động) vẫn là một hướng đi cần nhiều sự tối ưu hóa cho điều kiện thực tế tại địa phương (Ling et al., 2017; Rao & Sudheer, 2020). Xuất phát từ nhu cầu đó, nghiên cứu này tập trung vào việc "Nghiên cứu và thiết kế hệ thống giám sát an ninh tích hợp IoT". Hệ thống đề xuất sử dụng mô hình YOLOv8n được huấn luyện trên tập dữ liệu lớn để đảm bảo nhận diện chính xác con người trong thời gian thực, kết hợp với công nghệ RFID và 4G LTE với sự hỗ trợ của mô-đun SIM A7680C để tạo ra một cơ chế phản ứng nhanh chóng và toàn diện.

Nghiên cứu này không chỉ hướng tới việc cải thiện độ chính xác của quá trình nhận dạng (như kết quả thực nghiệm cho thấy đạt Precision 80,74% và Recall 69,70%) mà còn chú trọng vào tính thực tiễn thông qua việc tối ưu hóa tốc độ xử lý trên các thiết bị nhúng và cung cấp giải pháp quản lý tập trung qua phần mềm, góp phần thúc đẩy sự ứng dụng của AI và IoT trong lĩnh vực an ninh bảo mật tại Việt Nam.

2. PHƯƠNG PHÁP NGHIÊN CỨU

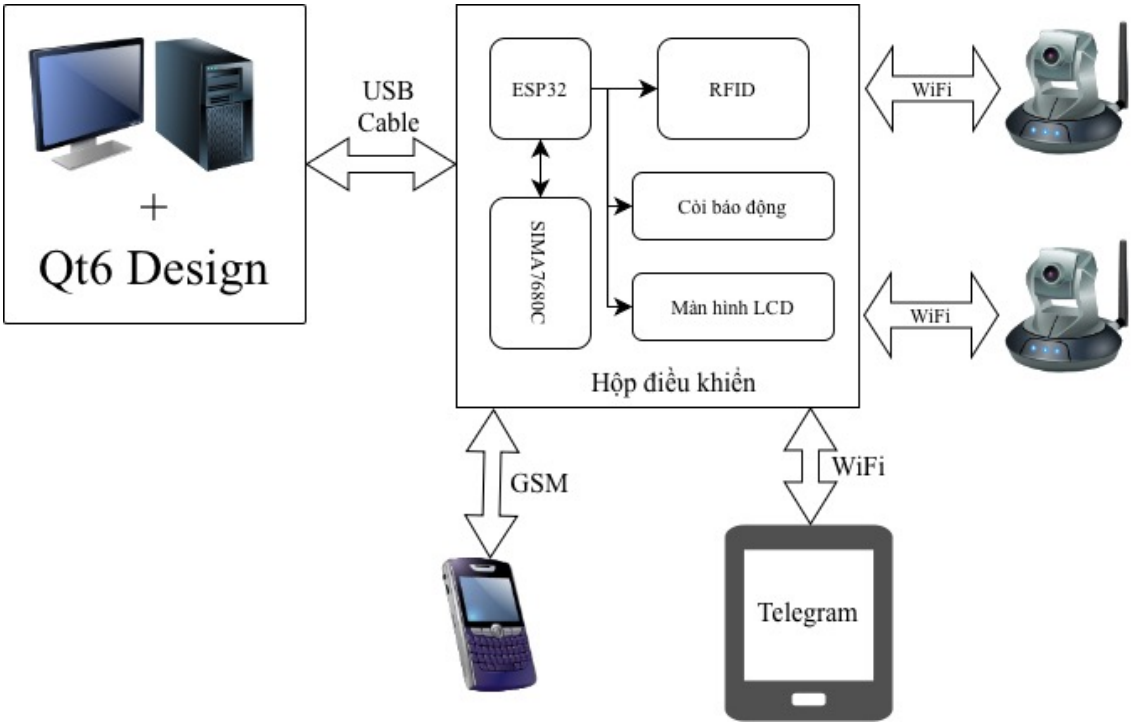
Tổng quan về hệ thống được thể hiện ở Hình 1. Trong đó, dữ liệu đầu vào từ đầu đọc RFID (để định danh) và hình ảnh từ các camera IP qua Wi-Fi. Dữ liệu này được xử lý tại bộ điều khiển (hộp điều khiển), sau đó hiển thị trạng thái lên màn hình LCD và kích hoạt còi báo động nếu có xâm nhập trái phép.

Song song đó, hệ thống duy trì ba kênh giao tiếp:

– Tại chỗ: Kết nối máy tính qua USB để quản trị bằng giao diện Qt6 (hỗ trợ lưu video trên bộ nhớ máy tính).

– Trực tuyến: Gửi thông báo và nhận lệnh điều khiển từ xa qua ứng dụng Telegram (Wi-Fi).

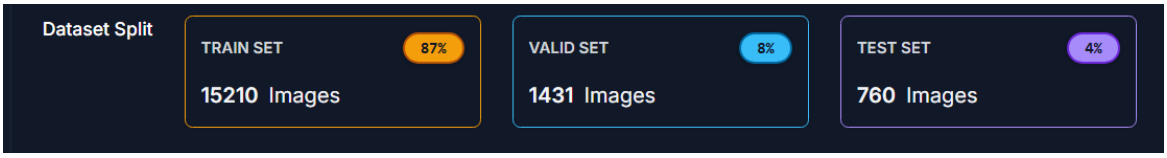
– Dự phòng: Module SIM A7680C đảm bảo gửi cảnh báo qua mạng GSM (SMS/Cuộc gọi) ngay cả khi mất internet.



Hình 1. Sơ đồ tổng quan của hệ thống đề xuất

2.1. Huấn luyện mô hình YOLOv8n và nhận dạng đối tượng

Trong nghiên cứu này, mô hình mạng nơ-ron tích chập (Convolutional Neural Network – CNN) được sử dụng để nhận dạng đối tượng, mô hình CNN được sử dụng với tập dữ liệu huấn luyện có sẵn trên trang Roboflow của tác giả Leo Ueno (2025) thể hiện ở Hình 2, tập dữ liệu gồm có 17.401 hình ảnh được chia ra 15.210 hình ảnh cho tập dữ liệu huấn luyện, 1431 hình ảnh cho tập dữ liệu kiểm chứng, 760 hình ảnh cho tập dữ liệu thử nghiệm. Các hình ảnh đã được tác giả của tập dữ liệu áp dụng các kỹ thuật tăng cường hình ảnh như xoay ngang, thu phóng 0 - 25%, áp dụng thang độ xám cho 25% hình ảnh, sắc thái ±25°, độ bão hòa ±25%, độ sáng ±25%, độ phơi sáng ±25%, làm mờ tối đa 2,5 pixel, nhiễu tối đa 1% pixel.



Hình 2. Tập dữ liệu huấn luyện, kiểm chứng và thử nghiệm

Mô hình đề xuất sử dụng YOLOv8n – một biến thể của YOLOv8 - cho quá trình huấn luyện và suy diễn của hệ thống. Việc lựa chọn YOLOv8 dựa trên những cải tiến vượt trội của phiên bản này so với các thế hệ trước đó, trong bối cảnh học sâu và thị giác máy tính ngày càng phát triển mạnh mẽ. YOLOv8 hiện được xem là một trong những phiên bản tối ưu nhất, nhờ vào hiệu suất vượt trội về tốc độ xử lý, độ chính xác, tính linh hoạt và hiệu quả tổng thể. Phiên bản này hiện nay đã hỗ trợ đa tác vụ về thị giác máy tính, bao gồm nhận dạng đối tượng (detection), phân đoạn đối tượng (segmentation), ước lượng tư thế đối tượng (pose estimation), theo dõi (tracking) và phân loại (classification). Sự đa dạng này cho phép người dùng tận dụng hết mọi khả năng của YOLOv8 trong nhiều ứng dụng và lĩnh vực khác nhau, điều đó giúp

YOLOv8 ngày càng trở nên phổ biến và được sử dụng rộng rãi, đánh dấu một bước tiến quan trọng trong thế giới của thị giác máy tính cũng như các ứng dụng liên quan.

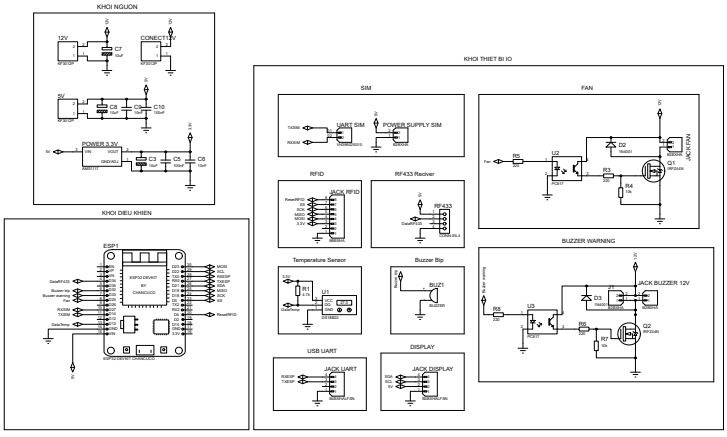
2.2. Thiết phần mềm giao diện

Giao diện phần mềm trong nghiên cứu này được thiết kế trên máy tính (laptop) bằng công cụ Qt6 Designer. Cách tiếp cận này mang đến sự tích hợp tốt với hệ sinh thái Qt6 bao gồm dễ dàng chuyển đổi sang Python (PySlide6), hỗ trợ cho quá trình đồng bộ với phần lập trình xử lý ảnh và áp dụng học sâu trên nền tảng Python cho toàn bộ hệ thống đề xuất. Qt6 Designer dựa trên cách tiếp cận kéo/thả các widget. Qt6 Designer là một phần mềm tạo giao diện đồ họa (WYSIWYG - What You See Is What You Get) cho phép bạn kéo và thả các widget (như nút ấn, hộp văn bản, nhãn,...) lên một form để thiết kế giao diện của ứng dụng. Các widget này có thể được tùy chỉnh về kích thước, font chữ, màu sắc và các thuộc tính khác. Qt là một framework đa nền tảng và là bộ dụng cụ hỗ trợ tạo giao diện người dùng và đồ họa (Fitzpatrick, 2021).

2.3. Thiết kế hộp báo động

Phần mềm SOLIDWORKS được sử dụng để thiết kế phần vỏ hộp cho hệ thống, bao gồm các chi tiết để lắp đặt với phần cứng như: Jack nguồn, kết nối USB to TTL, còi báo động, các giao tiếp ăng-ten và phần hiển thiện qua LCD (Hình 3). Ngoài ra, sơ đồ nguyên lý phần cứng của hệ thống được thể hiện ở Hình 3 với thiết bị ESP32 là bộ điều khiển chính, thiết bị Module SIM A7680C để thực hiện tính năng gọi điện đến các số điện thoại được nhập từ phần mềm giao diện, thiết bị loa báo động SFM 27 để phát ra tiếng cảnh báo đến người dùng, thiết bị thu sóng RF433 WL218T để nhận các tín hiệu điều khiển kích hoạt tính năng trên hộp báo động từ thiết bị remote, thiết bị đọc/ghi RFID RC522 để thêm/xóa và nhận thẻ RFID được thêm, thiết bị cảm biến nhiệt độ DS18B20 để đo nhiệt độ của hộp báo động, thiết bị quạt tản nhiệt để làm mát cho linh kiện bên trong hộp báo động, thiết bị màn hình OLED 128 × 32 để hiển thị rõ nét thông số nhiệt độ, tính năng RFID, tín hiệu của SIM, tín hiệu mạng của hộp báo động. Bên cạnh đó, có các linh kiện bán dẫn như opto PC817 và MOSFET IRFZ44N để kích quạt tản nhiệt và loa báo động hoạt động với điện áp cao hơn điện áp điều khiển, các linh kiện giảm áp LM2596 tạo mức điện áp 5 V và AMS1117 tạo mức điện áp 3,3 V, các linh kiện thụ động như điện trở để hạn dòng và các tụ điện để lọc nhiễu. Nguồn cấp cho hộp báo động được sử dụng với nguồn điện một chiều 12 V – 5 A (Hình 4).

Hình 3. Mô phỏng của thiết kế vỏ hộp cảnh báo trên SOLIDWORKS

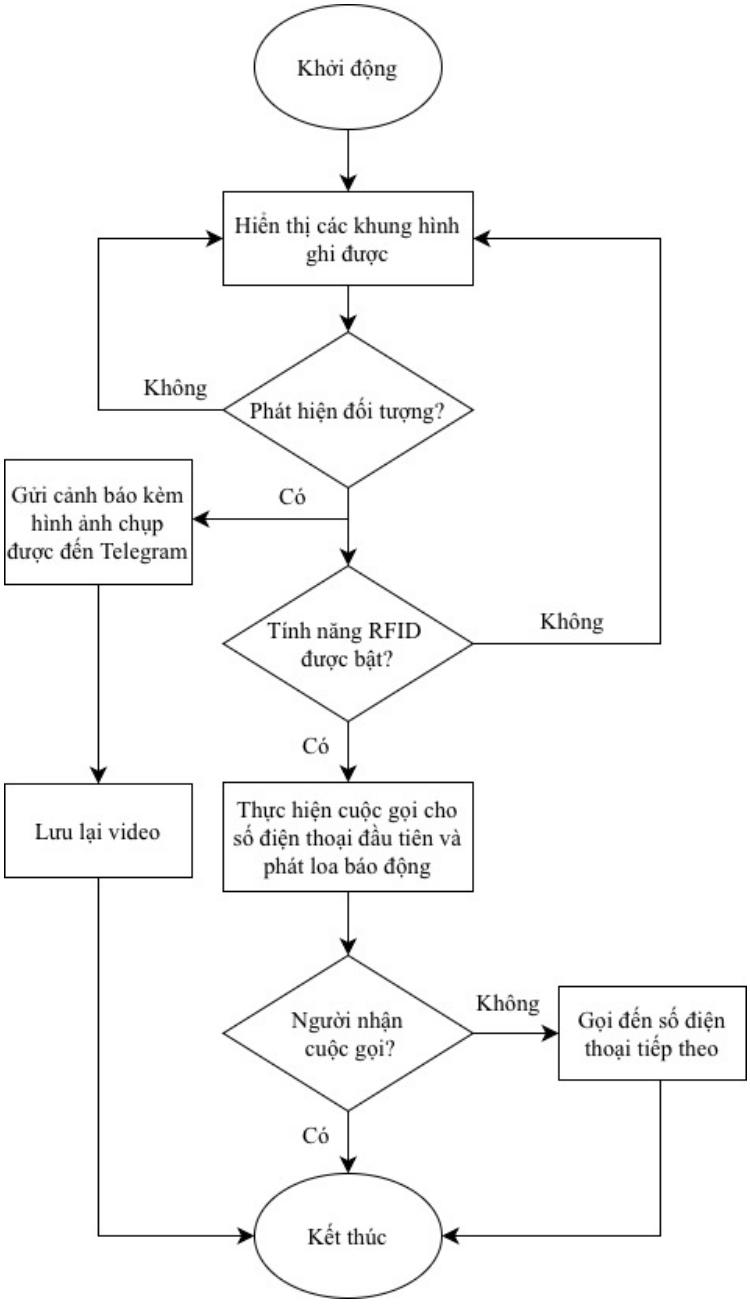


Hình 4. Sơ đồ nguyên lý của hệ thống

Ghi chú: (a) Khối điều khiển và khối nguồn; (b) khối thiết bị I/O

2.4. Lưu đồ thuật toán và nguyên lý hoạt động

Kết quả được trình bày ở Hình 5 cho thấy lưu đồ giải thuật cho quá trình hoạt động của hệ thống đề xuất. Khi hệ thống được khởi động, các camera IP được kết nối ghi lại các khung hình tại khu vực cần giám sát. Khi có đối tượng bên trong khung hình mô hình AI nhận dạng và gửi cảnh báo kèm hình ảnh chụp lại đối tượng đến ứng dụng Telegram và thực hiện lưu lại video, đồng thời gửi tín hiệu đến hộp báo động được kết nối với phần mềm thông qua phương thức truyền UART. Trước khi thực hiện báo động thực hiện kiểm tra xem chức năng RFID có được bật hay không, nếu có thì không thực hiện báo động, ngược lại hộp báo động bắt đầu thực hiện cuộc gọi theo trình tự các số điện thoại đã được nhập trên phần mềm, đồng thời phát loa báo động. Khi người nhận bắt máy, cuộc gọi sẽ kết thúc và loa báo động sẽ ngưng.

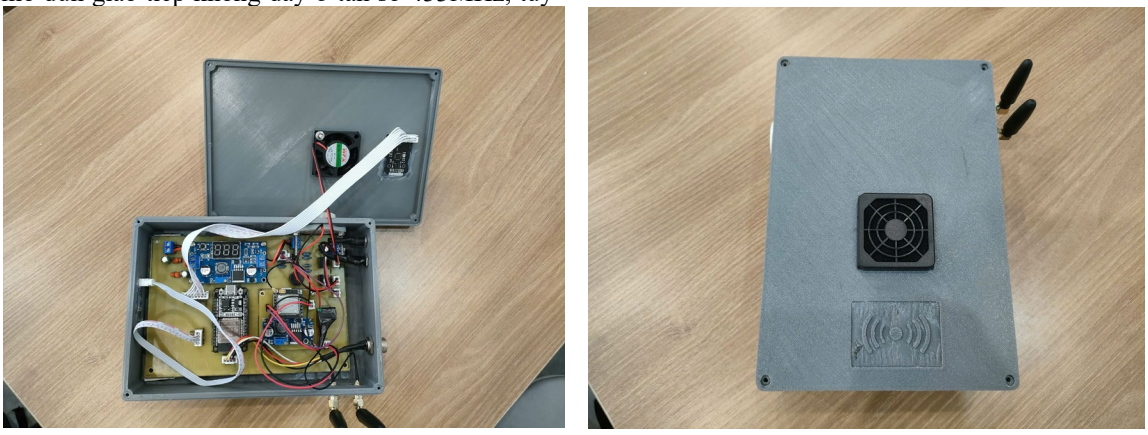


Hình 5. Lưu đồ hoạt động của hệ thống

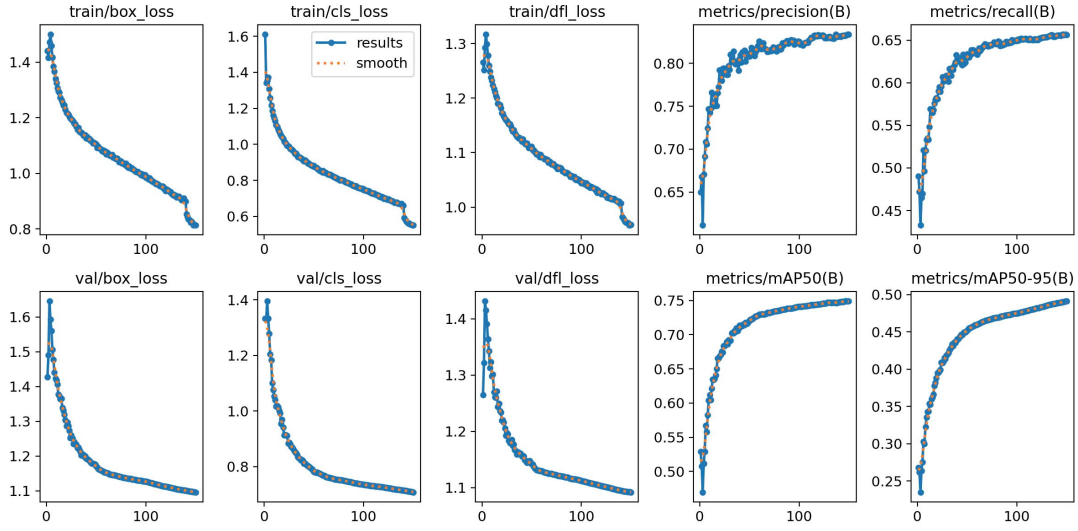
3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Kết quả thực nghiệm của hệ thống

Giao diện được thiết kế hoàn thiện để kết nối với hộp cảnh báo (Hình 6). Giao diện đề xuất bao gồm hai tùy chọn chính là "Quan sát" và "Tùy chỉnh". Chức năng của tuỳ chọn "Quan sát" là thực hiện hiển thị các khung hình ghi được từ camera và bật/tắt loa báo động. Trang "Thiết bị" thực hiện các tính năng tùy chỉnh các số điện thoại được nhập vào, có thể thêm/xóa thẻ RFID, thêm/xóa các thiết bị thông qua mô đun giao tiếp không dây ở tần số 433MHz, tùy chỉnh khoảng thời gian chạy mô hình YOLOv8 của camera lắp đặt. Tuỳ chỉnh "Cài đặt" thực hiện các tính năng thêm các đường dẫn RTSP của camera IP để kết nối đến phần mềm, thêm các đường dẫn trên máy để lưu trữ video, thêm token bot và chat id để kết nối đến ứng dụng Telegram; Trang "Kết nối" thực hiện các tính năng kết nối WiFi ở hộp báo động, kết nối giữa máy tính với hộp báo động thông qua giao thức truyền thông UART. Bên cạnh đó, phần mềm có tính năng bảo mật trước khi vào trang tùy chỉnh để tránh người khác vào thay đổi thông tin.



Hình 6. Thiết kế của hộp cảnh báo sau khi hoàn thiện



Hình 7. Các chỉ số huấn luyện và kiểm chứng qua 150 epochs của mô hình YOLOv8s, bao gồm hiệu suất trên cả tập dữ liệu huấn luyện (train) và kiểm chứng (val)

Kết quả được thể hiện tại Hình 7 cho thấy quá trình huấn luyện của mô hình, kết quả đã được biểu diễn rõ ràng qua các epochs, với việc giảm lỗi và tăng hiệu suất chỉ số. Khi không có dấu hiệu overfitting, điều này cho thấy mô hình có khả năng hoạt động tốt trong quá trình huấn luyện và đánh giá dữ liệu, giúp giảm thiểu các cảnh báo sai trong hệ thống giám sát. Nhờ vào việc triển khai mô hình YOLOv8 được huấn luyện trên tập dữ liệu đề xuất, hệ thống đã đạt được độ chính xác và độ nhạy tối ưu. Kết quả này cho phép phương pháp đề xuất không chỉ nhận dạng đối tượng một cách tin cậy mà

còn duy trì khả năng hoạt động ổn định trong các điều kiện môi trường thay đổi phức tạp, đặc biệt là khả năng giám sát xuyên suốt cả ngày và đêm.

Mô hình thể hiện khả năng tổng quát hóa tốt và ổn định trong các điều kiện ánh sáng khác nhau, từ đó đảm bảo độ tin cậy của hệ thống giám sát. Train/cls_loss và val/cls_loss đều giảm dần, mặc dù độ lỗi phân loại ở tập đánh giá vẫn hơi cao hơn so với tập huấn luyện. Sự chênh lệch này có thể do sự đa dạng trong dữ liệu đánh giá, yêu cầu mô hình tiếp tục tinh chỉnh để đáp ứng với các trường hợp phức tạp hơn. Tuy nhiên, sự giảm lỗi đều đặn vẫn là minh chứng cho việc cải thiện khả năng phân loại chính xác theo thời gian. Các chỉ số đánh giá như *m*AP và độ chính xác cho thấy mô hình có thể phát hiện chính xác các đối tượng ở nhiều ngưỡng IoU khác nhau, giảm thiểu tỷ lệ báo động giả. Đồng thời, khả năng phân loại chính xác các lớp đối tượng giúp hệ thống đưa ra các quyết định phù hợp. Các chỉ số hiệu suất Precision và Recall đều tăng trưởng, với precision đạt trên 0,8 và recall duy trì ở khoảng 0,65. Mặc dù độ nhạy có thể được cải thiện hơn nữa, mức này đã cho thấy mô hình có khả năng phát hiện phần lớn các đối tượng, giảm thiểu tình trạng bỏ sót các hoạt động bất thường. Độ chính xác cao đồng nghĩa với ít cảnh báo sai hơn, giúp hệ thống giám sát hoạt động hiệu quả hơn và giảm thiểu các phản hồi không cần thiết. Trong việc đánh giá hiệu quả của mô hình sử dụng YOLOv8, nghiên cứu dựa trên độ chính xác trung bình (*m*AP), thể hiện ở (1) hoặc độ chính xác trung bình (AP),

$$m\mathrm{AP} = \frac{1}{N}\sum_{i=1}^{N} AP_i \quad (1)$$

Đánh giá dựa trên *m*AP: *m*AP50 và *m*AP50-95 cho thấy mô hình có khả năng phát hiện chính xác đối tượng ở nhiều ngưỡng khác nhau, từ 50% đến 95%. Chỉ số *m*AP50-95 đạt khoảng 0,5 cho thấy mô hình duy trì hiệu suất ổn định trong nhiều điều kiện và độ khó khác nhau. Điều này rất quan trọng trong giám sát an ninh thực tế, nơi mô hình cần đáp ứng với các mức độ cảnh báo và tình huống khác nhau từ các hành động bất thường nhẹ đến các tình huống nguy hiểm cao.

Trong nghiên cứu, nhóm tác giả thực hiện việc xây dựng một giao diện người dùng bằng công cụ Qt6 Designer. Việc cho phép tạo giao diện bằng phương pháp kéo thả widget giúp cho việc xây dựng giao diện dễ dàng hơn. Giao diện sử dụng widget QstackedWidget để tạo các trang chính, widget QTab Widget để tạo các trang con, widget Qlabel để tạo các tiêu đề và các trạng thái cập nhật, widget QPushButton để tạo các nút nhấn chức năng, widget QLineEdit để tạo các ô nhập thông tin, widget QTimeEdit để tạo các cài đặt thời gian, widget QComboBox để tạo các danh sách lựa chọn. Bên cạnh đó, các icon cũng được thêm vào để tăng tính sinh động cho giao diện (Hình 8).

Dựa vào các thông số trong ma trận nhầm lẫn thể hiện ở Hình 9 có thể tính được các giá trị Precision, Recall và F1-Score cho Bảng 1 thông qua các công thức (1), (2) và (3) bên dưới.

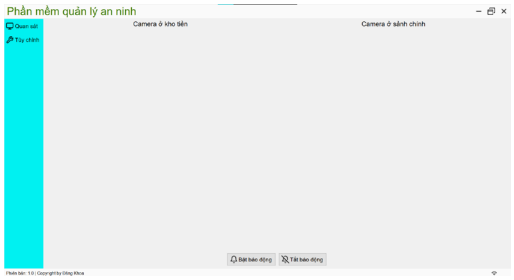
$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \times 100 = \frac{7.430}{7.430+1.772} \times 100 \approx 80,74\% \quad (1)$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \times 100 = \frac{7.430}{7.430+3.230} \times 100 \approx 69,70\% \quad (2)$$

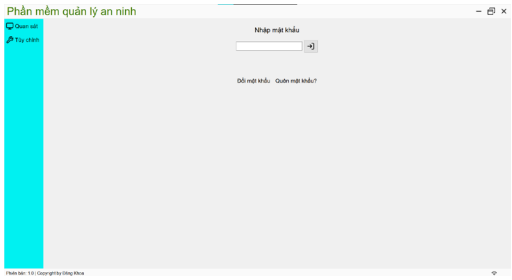
$$\mathrm{F1} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} = 2 \times \frac{80,74\% \times 69,70\%}{80,74\% + 69,70\%} = 74.81\% \quad (3)$$

Dựa vào quá trình tính toán ở công thức (3), ta thấy trị số F1 Score nằm ở mức khá tốt để nhận diện đối tượng. Vậy đối với vấn đề nhận diện đối tượng là con người trong thời gian thực với tần suất và số lượng thấp thì mô hình này có thể dùng được trong hệ thống được xây dựng.

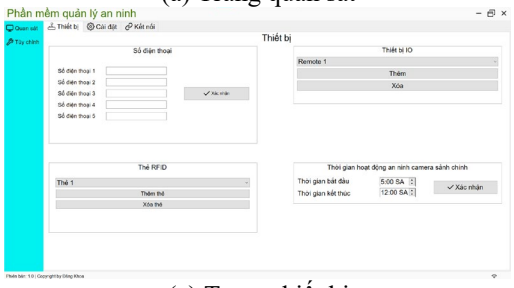
Kết quả được thể hiện tại Bảng 1 cho thấy kết quả sau khi huấn luyện, mô hình đạt độ chính xác (Precision) 80,74%, cho phép giảm thiểu các cảnh báo sai, đảm bảo hệ thống chỉ phát ra cảnh báo khi thực sự cần thiết. Độ nhạy (Recall) đạt 69,70%, phản ánh khả năng phát hiện con người trong khu vực giám sát với một số trường hợp bị bỏ sót. Chỉ số F1-Score, kết hợp giữa Precision và Recall, đạt 74,69%, cho thấy hiệu suất tổng thể phù hợp cho các ứng dụng trong thời gian thực.



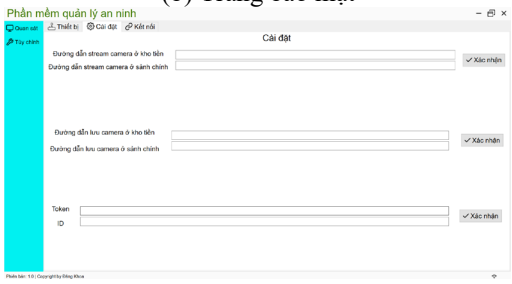
(a) Trang quan sát



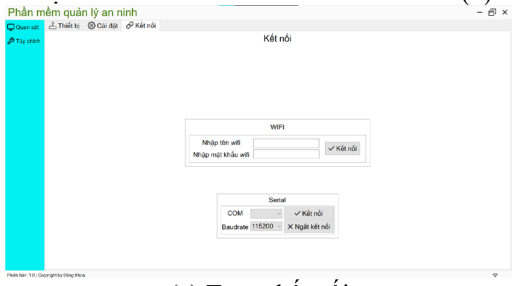
(b) Trang bảo mật



(c) Trang thiết bị

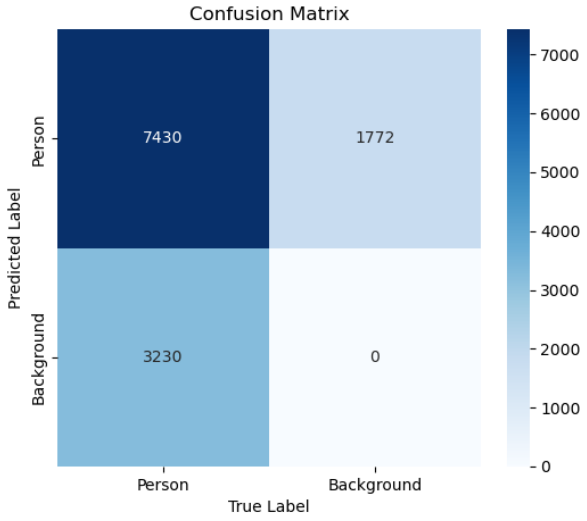


(d) Trang cài đặt



(e) Trang kết nối

Hình 7. Giao diện phần mềm quản lý



Hình 9. Ma trận nhầm lẫn

Bảng 1. Kết quả sau khi huấn luyện tập dữ liệu

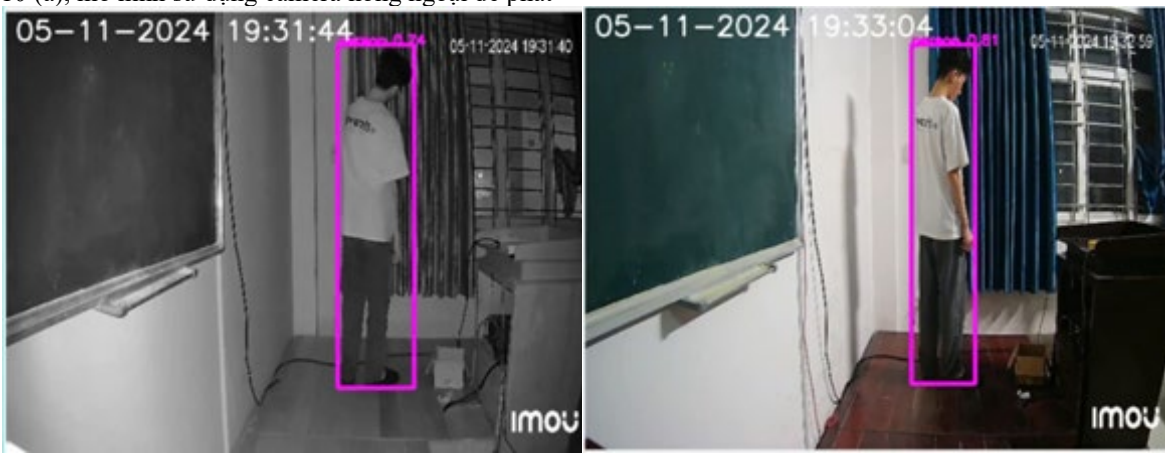
Precision	Recall	F1-Score	Tốc độ nhận dạng
80,74%	69,70%	74,69%	0,127 Sec/image

Tốc độ xử lý của mô hình là 0,127 giây trên mỗi ảnh, tương đương với khoảng 7,87 fps, đáp ứng yêu cầu giám sát liên tục trong thời gian thực. Những kết quả này chứng minh mô hình có tiềm năng lớn trong việc ứng dụng vào hệ thống cảnh báo an ninh trong thực tế, góp phần nâng cao hiệu quả bảo vệ tài sản và đảm bảo an toàn cho khách hàng cũng như nhân viên. Tuy nhiên, để tăng cường hiệu suất trong các môi trường thực tế phức tạp, như điều kiện ánh sáng kém hoặc khu vực đông đúc, các bước cải tiến tiếp theo tập trung vào việc mở rộng dữ liệu huấn luyện và tối ưu hóa cấu trúc mô hình. Những điều chỉnh này hứa hẹn giúp hệ thống trở nên đáng tin cậy hơn, đáp ứng các tiêu chuẩn an ninh nghiêm ngặt theo xu hướng hiện nay.

3.2. Đánh giá quá trình hoạt động của hệ thống

Trong điều kiện ánh sáng kém thể hiện ở Hình 10 (a), mô hình sử dụng camera hồng ngoại để phát hiện đối tượng với độ tự tin (confidence) là 0,74. Kết quả này minh chứng rằng mô hình đã được tối ưu hóa để hoạt động trong môi trường tối, hỗ trợ giám sát hiệu quả trong khoảng thời gian ban đêm, một yêu cầu quan trọng cho các hệ thống an ninh hoạt động 24/7. Khả năng này giúp hệ thống phát hiện các đối tượng khả nghi mà không bị phụ thuộc vào điều kiện ánh sáng.

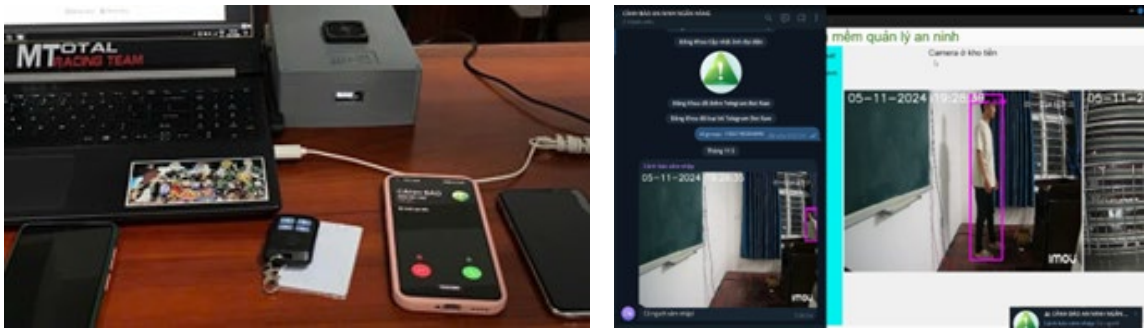
Trong điều kiện ánh sáng tốt thể hiện ở Hình 10 (b), mô hình đạt độ tự tin 0,81 cao hơn so với môi trường thiếu sáng. Điều này cho thấy rằng mô hình có khả năng phát hiện và nhận dạng đối tượng rõ ràng hơn trong điều kiện ánh sáng thuận lợi. Khả năng nhận dạng ổn định bất kể điều kiện ánh sáng giúp hệ thống duy trì hiệu quả giám sát liên tục, giảm thiểu khả năng bỏ sót trong cả ngày lẫn đêm.



(a) (b)

Hình 10. Kết quả nhận dạng của mô hình trong (a) điều kiện ánh sáng kém; (b) điều kiện ánh sáng tốt

Khi nhận dạng được đối tượng trong khung hình hệ thống thực hiện cảnh báo gọi điện thể hiện ở Hình 11 (a) đồng thời gửi cảnh báo đến nhóm trên ứng dụng Telegram thể hiện ở Hình 11 (b). Kết quả thử nghiệm cho thấy tốc độ gửi cảnh báo đến Telegram nhanh hơn tốc độ gọi điện đến số điện thoại.



(a) (b)

Hình 11. Thực hiện cảnh báo khi camera nhận dạng đối tượng qua phương thức (a) gọi qua số điện thoại; (b) gửi cảnh báo qua Telegram

Bảng 2. Kết quả thử nghiệm của hệ thống

Điều kiện môi trường	Số lần thử nghiệm	Số lần phát hiện đối tượng	Số lần không phát hiện đối tượng	Số lần phát hiện và cảnh báo	Số lần phát hiện nhưng không cảnh báo
Ánh sáng đầy đủ	20	19	1	18	1
Thiếu sáng	20	17	3	14	3

Dựa trên kết quả thử nghiệm thể hiện ở Bảng 2 trong hai điều kiện ánh sáng, mô hình phát hiện đối tượng thể hiện hiệu suất đáng chú ý trong việc phát hiện đối tượng và cảnh báo an ninh. Trong điều kiện ánh sáng đầy đủ, với tổng cộng 20 lần thử nghiệm, mô hình phát hiện được 19 trường hợp có đối tượng, trong đó có 18 lần mô hình phát hiện và cảnh báo kịp thời. Tuy nhiên, có một lần mô hình phát hiện con người nhưng không đưa ra cảnh báo. Điều này cho thấy mô hình hoạt động khá chính xác và đáng tin cậy trong điều kiện ánh sáng đầy đủ, với tỷ lệ cảnh báo chính xác cao.

Trong điều kiện thiếu sáng, mặc dù mô hình vẫn duy trì khả năng phát hiện đối tượng tốt, nhưng hiệu suất có sự giảm sút. Cụ thể, trong 20 lần thử nghiệm, mô hình phát hiện được 17 trường hợp có đối tượng, trong đó có 14 lần phát hiện và cảnh báo đúng, trong đó, có 03 lần không phát hiện đối tượng chính xác, dẫn đến không cảnh báo. Điều này cho thấy rằng mô hình gặp khó khăn hơn trong việc phát hiện đối tượng dưới điều kiện thiếu sáng, dẫn đến sự giảm hiệu quả của hệ thống cảnh báo.

Bảng 3. So sánh tính hiệu quả của phương pháp đề xuất với các nghiên cứu khác

Nghiên cứu	Phương pháp	Ưu điểm/phạm vi	Hiệu suất/Hạn chế
Jyothi và ctv. (2016)	Sử dụng phương pháp xử lý ảnh truyền thống	Sử dụng phương pháp xử lý ảnh truyền thống	Không ứng dụng được đối với ảnh hồng ngoại ban đêm
Sodhro và ctv. (2025)	Phát hiện xâm nhập trái phép sử dụng YOLOv5 và YOLOv8	Hiệu quả cho các ứng dụng giám sát có độ trễ thấp (5 giây) và sử dụng thuật toán trên máy tính nhúng Jetson nano	Chưa thử nghiệm trong điều kiện ban đêm với ảnh hồng ngoại
Siva và ctv. (2025)	YOLOv8	Phát hiện thời gian thực, chi phí thấp và cải thiện quyền riêng tư	Chỉ sử dụng duy nhất một camera. Chưa áp dụng được các môi trường khác nhau (ngày và đêm)
Phương pháp đề xuất	YOLOv8	Phát hiện thời gian thực cho nhiều camera, có giao diện và đảm bảo tính bảo mật quyền truy cập. Giám sát thời gian thực ngày và đêm	Đòi hỏi tập dữ liệu lớn để cải thiện độ chính xác

Kết quả được thể hiện ở Bảng 3 cho thấy sự so sánh giữa phương pháp đề xuất và các nghiên cứu liên quan. So với nghiên cứu của Jyothi et al. (2016) vốn bị giới hạn bởi các kỹ thuật xử lý ảnh truyền thống và không thể hoạt động trong môi trường ánh sáng yếu (ban đêm) hay nghiên cứu của Sodhro et al. (2025) sử dụng YOLOv5 và YOLOv8 gặp khó khăn khi chưa thử nghiệm dưới điều kiện ban đêm, phương pháp đề xuất đã giải quyết được bài toán môi trường bằng khả năng giám sát ổn định cả ngày và đêm.

Mặc dù Siva et al. (2025) đã triển khai YOLOv8 để tối ưu chi phí và quyền riêng tư nhưng nghiên cứu này vẫn bị hạn chế khi chỉ hoạt động trên một camera duy nhất chưa thích ứng với đa dạng môi trường. Phương pháp đề xuất đã mở rộng quy mô lên hệ thống nhiều camera (multi-camera), đồng thời bổ sung giao diện người dùng và cơ chế bảo mật quyền truy cập chặt chẽ hơn.

Nhìn chung, hệ thống có thể hoạt động tốt trong điều kiện ánh sáng đầy đủ, nhưng cần cải thiện độ chính xác để xử lý hiệu quả hơn trong môi trường thiếu sáng (đòi hỏi tập dữ liệu lớn hơn), đảm bảo độ chính xác và tính kịp thời của cảnh báo.

4. KẾT LUẬN

Mô hình hệ thống an ninh đề xuất, tích hợp IoT, camera không dây, RFID và phần mềm quản lý, đã chứng minh hiệu quả vượt trội trong việc giám sát an ninh. Hệ thống không chỉ tối ưu hóa khả năng nhận dạng, phân quyền linh hoạt mà còn tăng cường tốc độ xử lý tình huống khẩn cấp, giảm thiểu rủi ro tài sản. Kết quả bước đầu khẳng định ứng dụng AI và IoT là giải pháp tối ưu để xây dựng hệ thống bảo mật đa lớp, thông minh và đáng tin cậy.

Để hoàn thiện hệ thống hướng tới sự toàn diện và bền bỉ, các nghiên cứu tiếp theo tập trung vào các nhóm giải pháp sau:

– Trang bị hệ thống nguồn dự phòng (UPS) và tích hợp cảm biến hồng ngoại (IR) cùng với hệ camera có sẵn chuyên dụng để đảm bảo hệ thống vận hành liên tục 24/7 trong mọi điều kiện ánh sáng.

– Ứng dụng các giao thức bảo mật tiên tiến cho phần mềm quản lý, thiết lập quy trình cập nhật và kiểm tra lỗ hổng định kỳ để đối phó với các mối đe dọa không gian mạng.

– Nghiên cứu tích hợp các mô hình Deep Learning tiên tiến nhằm nâng cao độ chính xác trong phân tích hành vi và nhận dạng đối tượng phức tạp.

– Tích hợp máy tính nhúng (vi xử lý) vào trong hộp cảnh báo trung tâm sử dụng để chạy mô hình AI giúp cho thiết bị không cần phải kết nối đến máy tính, tối ưu cho việc lắp đặt như nghiên cứu của Sodhro et al. (2025).

LỜI CẢM TẠ

Nhóm tác giả xin chân thành cảm ơn Trường Đại học Kỹ thuật – Công nghệ Cần Thơ đã tài trợ kinh phí cho nghiên cứu này thông qua dự án mã số DTCS2025-11.

TÀI LIỆU THAM KHẢO

Abdullah, H. S., Jabber, S. A. (2018). “Intelligent monitoring to detect and recognized the unauthorized persons”. *Journal of College of Education for Pure Science*, *8*(2), 48-62.

Sanjalawe, Y., Alqudah, H. (2024). Integrating enhanced security protocols with moving object detection: A YOLO-based approach for real-time surveillance. *2024 2nd International Conference on Cyber Resilience (ICCR)*, pp. 1-6. https://doi.org/10.1109/ICCR61006.2024.10532863

Arun, G., Ajay, B., & Valarmathi, R. (2023). "IoT Based Anti-Theft Detection System," *Intelligent computing and control for engineering and business systems (ICCEBS)*, Chennai, India. IEEE, pp. 1-4. https://doi.org/10.1109/ICCEBS58601.2023.10448636

Chatterjee, N., Singh, A. V., & Agarwal, R. (2024). "You Only Look Once (YOLOv8) Based Intrusion Detection System for Physical Security and Surveillance,". *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Noida, India, pp. 1-5. https://doi.org/10.1109/ICRITO61523.2024.10522139

Farooq, U., ul Hasan, M., Amar, M. H., & Asad, M. U. (2014). "RFID based security and access control system," *International Journal of Engineering and Technology, Vol. 6*, No. 4. https://doi.org/10.7763/IJET.2014.V6.718

Fitzpatrick, M. (2021). *Create GUI applications with Python & Qt6*

IBM. (2023). IBM report: *Half of breached organizations unwilling to increase security spend despite soaring breach costs.* IBM Newsroom. https://newsroom.ibm.com/2023-07-24-IBM-Report-Half-of-Breached-Organizations-Unwilling-to-Increase-Security-Spend-Despite-Soaring-Breach-Costs

Jesus, O., Andres, M., & Oscar, G. (2019). Implementation of a banking system security in embedded systems using artificial intelligence. *Advances in Natural and Applied Sciences, Vol. 10*(17); pp: 95-101.

Jyothi, S. N., Vardhan, K. V. (2016). "Design and implementation of real time security surveillance system using IoT," *2016 International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2016, pp. 1-5. https://doi.org/10.1109/CESYS.2016.7890003

Leo, U. (2025). "*Roboflow Universe, Roboflow*". https://universe.roboflow.com/leo-ueno/people-detection-o4rdr

Ling, Z., Liu, K., Xu, Y., Jin, Y., & Fu, X. (2017). An end-to-end view of IoT security and privacy. *GLOBECOM 2017-2017 IEEE Global Communications Conference* (pp. 1-7). https://doi.org/10.1109/GLOCOM.2017.8254011

Lulla, G., Kumar A., Pole, G., & Deshmukh, G. (2021). "IoT based Smart Security and Surveillance System," *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India, pp. 385-390. https://doi.org/10.1109/ESCI50559.2021.9396843

Manu, Y. M., Shashikala, S. V., Darshan, N., Dheeraj, A. P., Hemanth, N. J., Nishanth Gowda, B. (2024). Smart Surveillance Camera Using AI. *Second International Conference on Advances in Information Technology (ICAIT)* (pp. 1-6). https://doi.org/10.1109/ICAIT61638.2024.10690287

Motwani, N. P., Soumya, S. (2023). "Human Activities Detection using DeepLearning Technique- YOLOv8," *International Conference on Data Science and Advanced Computing (ICDSAC 2023)*, Coimbatore, India, vol. 56, pp. 1-8. https://doi.org/10.1051/itmconf/20235603003

Pacal, I., Karaboga, D., Basturk, A., Akay, B., & Nalbantoglu, U. (2020). A comprehensive review of deep learning in colon cancer. *Computers in Biology and Medicine*, 126, 104003. https://doi.org/10.1016/j.compbiomed.2020.104003

Rao, B. N., Sudheer, R. (2020). Surveillance camera using IOT and Raspberry Pi. *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 1172-1176). https://doi.org/10.1109/ICIRCA48905.2020.9182983

Sattaru, P. K., Burugula, K. V., Channagiri, R., & Kavitha, S. (2023). Smart home security system using IoT and ESP8266. *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 469-474. https://doi.org/10.1109/ICSSIT55814.2023.10061059

Siva, P., Pujitha, G. B., Krishna, G. S., Hemanth, G., Teja, B. M. S. (2025). Smart Surveillance Systems Using YOLOv8: A Scalable Approach for Crowd and Threat Detection. *International Journal of Recent Advances in Engineering and Technology*, *14*(1), 51-62. https://doi.org/10.65521/ijacect.v14i1.171

Sodhro, A., Kannam, S., Jensen, M. (2025). Real-time efficiency of YOLOv5 and YOLOv8 in human intrusion detection across diverse environments and recommendation. *Internet of Things*, 101707. https://doi.org/10.1016/j.iot.2025.101707

Statista. (2024). *Smart home market outlook in Vietnam*.

Sudharson, D., Srinithi, J., Akshara, S., Abhirami, K., Sriharshitha, P., Priyanka, K. J. P. C. S. (2023). Proactive headcount and suspicious activity detection using YOLOv8. *Procedia Computer Science*, 230, 61-69. https://doi.org/10.1016/j.procs.2023.12.061

Swathi, Y., & Challa, M. (2024). YOLOv8: Advancements and innovations in object detection. *International Conference on Smart Computing and Communication*, pp. 1-13. https://doi.org/10.1007/978-981-97-1323-3_1

Weber, T. L. (2016). *Alarm systems and theft prevention*. Elsevier.

Wresearch. (2020-2026). *Vietnam Video Surveillance Market Report*. https://www.6wresearch.com/industry-report/vietnam-video-surveillance-market-2020-2026.