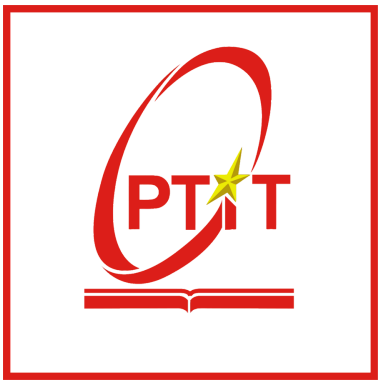
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG
KHOA CÔNG NGHỆ THÔNG TIN 1
****



BÀI TẬP LỚN IOT VÀ ỨNG DỤNG
Báo cáo giữa kỳ
Đề tài: Hệ thống tủ thông minh

Giảng viên hướng dẫn: TS. Kim Ngọc Bách

Lớp chuyên ngành: D22CNPM02 - Nhóm BTL: 9
Đề tài:

Thành viên	Mã sinh viên
Trần Mai Hương	B22DCCN424
Nguyễn Thị Khánh Vân	B22DCCN892
Nguyễn Nam Vũ	B22DCCN916
Nguyễn Thị Yến	B22DCCN928

Hà Nội, Tháng 10/2025

MỤC LỤC

A. PHÂN TÍCH YÊU CẦU

I. Giới thiệu

Trong thời đại công nghệ 4.0, nhu cầu về các thiết bị thông minh, tiện lợi và an toàn ngày càng tăng. Hệ thống tủ thông minh nhận diện khuôn mặt giúp người dùng mở khóa mà không cần chìa, đồng thời có thể điều khiển và giám sát từ xa qua Internet. Đề tài này sử dụng ESP32-CAM – một vi điều khiển có camera tích hợp và khả năng kết nối Wi-Fi, để nhận diện khuôn mặt, mở khóa tủ và gửi cảnh báo trong các tình huống bất thường.

1. Lý do chọn đề tài

- Nhu cầu bảo mật và tự động hóa trong gia đình, văn phòng, ký túc xá, phòng làm việc, trung tâm thương mại ngày càng lớn.
- Người dùng có thể giám sát từ xa hoặc được cảnh báo khi có hành vi mở khóa trái phép.
- Công nghệ nhận diện khuôn mặt và IoT hiện nay cho phép xây dựng các hệ thống an ninh nhỏ gọn, giá rẻ nhưng hiệu quả cao.
- Hệ thống loại bỏ rủi ro mất chìa khóa, bị sao chép thẻ RFID, đồng thời có khả năng cảnh báo nguy hiểm.

2. Mục tiêu đề tài

- Xây dựng hệ thống tủ thông minh có khả năng nhận diện khuôn mặt để mở khóa tự động.
- Cho phép điều khiển đóng/mở tủ từ xa
- Tích hợp tính năng cảnh báo khi phát hiện người lạ hoặc có hành vi truy cập trái phép
- Hệ thống hoạt động ổn định, chi phí thấp, dễ dàng mở rộng**.**

3. Phạm vi triển khai

- Phạm vi thiết bị
  - 1 module ESP32-CAM (camera và xử lý trung tâm).
  - 1 servo motor để điều khiển cơ cấu đóng/mở tủ.
  - 1 cảm biến chuyển động PIR.
  - Nguồn cấp: lấy từ cổng USB máy tính hoặc adapter 5V.
- Phạm vi kết nối
  - Hệ thống hoạt động qua Wi-Fi để điều khiển và gửi cảnh báo.
- Phạm vi môi trường:
  - Áp dụng cho tủ cá nhân, tủ văn phòng, ký túc xá, trung tâm thương mại hoặc không gian trong nhà.
  - Mô hình demo có thể mở rộng thành hệ thống nhiều tủ có quản lý trung tâm.

4. Tiêu chí thành công

- Độ chính xác:
  - Sai số nhận diện khuôn mặt < 10%
- Độ trễ:
  - Thời gian từ khi khi nhận diện đến khi mở tủ < 5s
  - Độ trễ khi mở khóa từ xa < 5s
- Độ tin cậy
  - Tỉ lệ truyền dữ liệu thành công > 98%.
  - Thông báo khi có mở trái phép > 95%
- Chi phí:

- Tổng chi phí phần cứng < 500.000 VNĐ.

5. Kết quả mong đợi

- Tự động hóa mở khóa tủ thông qua nhận diện khuôn mặt → nâng cao tính bảo mật và tiện lợi cho người dùng.
- Cho phép điều khiển tủ từ xa → người dùng có thể mở tủ hoặc kiểm tra trạng thái dù không có mặt trực tiếp.
- Tự động gửi cảnh báo an ninh khi phát hiện hành động xâm nhập hoặc nhận diện sai khuôn mặt.
- Tự động phát hiện chuyển động và nhận diện khuôn mặt.
- Giảm rủi ro mất cắp và nâng cao an toàn tài sản cá nhân nhờ cơ chế xác thực sinh trắc học.
- Giao diện điều khiển thân thiện, dễ mở rộng, có thể tích hợp cho nhiều tủ khác nhau trong cùng hệ thống.
- Cơ sở dữ liệu nhận diện khuôn mặt có thể mở rộng, giúp phục vụ nhiều người dùng (sinh viên, nhân viên, hộ gia đình).

II. Mô tả tổng quan

1. Môi trường

- Vị trí hoạt động: trong nhà, tại khu vực học tập, văn phòng hoặc gia đình.
- Điều kiện nhiệt độ: 20°C – 35°C, độ ẩm trung bình.
- Kết nối: Wi-Fi ổn định để truyền nhận dữ liệu.
- Nguồn điện: sử dụng nguồn 5V
- Ánh sáng: đủ sáng để camera ESP32-CAM nhận diện khuôn mặt chính xác.

2. Các ràng buộc

- Phần cứng giới hạn: ESP32-CAM có bộ nhớ nhỏ, không thể chạy mô hình AI phức tạp.
- Độ sáng ảnh hưởng: nhận diện kém trong điều kiện thiếu sáng hoặc ngược sáng.
- Phụ thuộc vào Wi-Fi: hệ thống chỉ hoạt động tốt khi có kết nối mạng ổn định.
- Cơ cấu khóa servo: hoạt động trong giới hạn điện áp định mức, cần nguồn ổn định.
- Độ trễ mạng: có thể ảnh hưởng đến tốc độ phản hồi khi mở khóa từ xa.
- Cảm biến chuyển động (PIR): Hoạt động hiệu quả trong phạm vi 3–7 mét và góc quét khoảng 110°, dễ bị nhiễu khi có nguồn nhiệt hoặc vật thể di chuyển gần (như quạt, ánh sáng mặt trời trực tiếp). Không phù hợp cho môi trường ngoài trời hoặc nơi có sự thay đổi nhiệt độ lớn.

3. Các giả định

- Khu vực lắp đặt có kết nối Wi-Fi với sóng ổn định.
- Nguồn điện cung cấp ổn định (5V, ≥1A) để đảm bảo servo và camera hoạt động.
- Hệ thống được dùng cho mục đích cá nhân hoặc học tập, không yêu cầu bảo mật cấp cao.
- Môi trường hoạt động trong nhà, tránh ánh nắng trực tiếp và bụi bẩn ảnh hưởng camera.
- Cảm biến PIR được lắp đặt ở vị trí phù hợp, không bị che khuất, và hướng về khu vực thường có người tiếp cận tủ.

III. Xác định các yêu cầu chức năng

Các Actor của hệ thống gồm thành viên hệ thống và Admin, User kế thừa từ thành viên hệ thống. Các chức năng tương ứng với từng actor:

- Admin: Quản lý người dùng, quản lý tủ, điều khiển đóng mở tủ từ xa
- User: điều khiển mở tủ từ xa, nhận diện khuôn mặt và mở tủ, nhận cảnh báo khi có truy cập trái phép

Dưới đây là mô tả chi tiết các chức năng

1. Các chức năng cần có của hệ thống

a. Mở tủ:

- Hệ thống cung cấp 2 cách mở tủ: mở bằng nhận diện khuôn mặt và mở tủ từ xa
    - Mở tủ từ xa: Người dùng và admin có thể truy cập website để gửi lệnh mở/đóng tủ từ xa qua Internet. Lệnh được truyền qua Network Layer (HTTP/MQTT) đến hệ thống, kích hoạt servo motor. Hỗ trợ kiểm tra trạng thái tủ realtime (mở/đóng). Trong trường hợp của admin, chỉ thực hiện chức năng khi có yêu cầu từ người dùng.
    - Mở bằng nhận diện khuôn mặt: hệ thống sử dụng camera trên ESP32-CAM để chụp hình khuôn mặt người dùng khi tiếp cận tủ. Hình ảnh được gửi đến server AI (qua Wi-Fi/MQTT) để so sánh với cơ sở dữ liệu khuôn mặt đã đăng ký. Nếu khớp (độ chính xác >90%), servo motor sẽ tự động mở chốt tủ trong vòng <5 giây.

b. Quản lý tủ:

- Admin có thể thêm/sửa/xóa các thông tin về tủ, hỗ trợ mở rộng cho nhiều tủ

c. Gửi cảnh báo an ninh khi truy cập trái phép:

- Nếu nhận diện khuôn mặt thất bại, hệ thống sẽ gửi thông báo đẩy đến tài khoản người dùng. Bao gồm ảnh chụp người lạ và thời gian sự kiện. Độ tin cậy gửi >95%.

d. Quản lý người dùng:

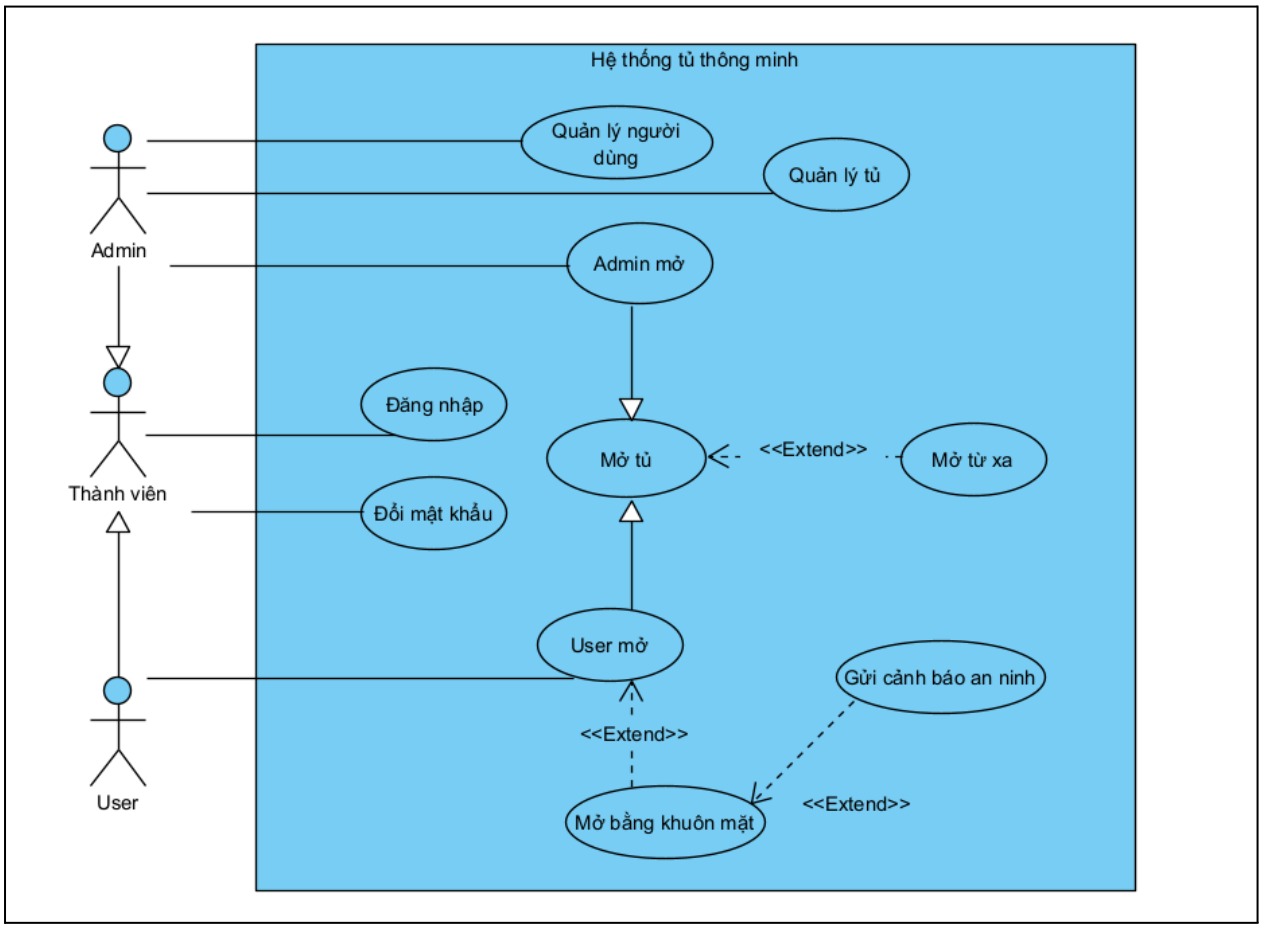
- Admin có thể thêm/xóa/sửa dữ liệu người dùng qua web. Hỗ trợ mở rộng cho nhiều user. Dữ liệu được lưu trữ an toàn trên database (MySQL/Firebase).

2. Đặc tả luồng công việc bằng sơ đồ usecase

Dưới đây là sơ đồ use case tổng quan cho hệ thống với các actor và chức năng được mô tả ở trên:

- Usecase đăng nhập và đổi mật khẩu: thành viên hệ thống được quyền đăng nhập và đổi mật khẩu của mình nếu muốn
- Usecase quản lý người dùng: admin có quyền thêm, sửa, xóa thông tin của người dùng
- Use case quản lý tủ: admin có quyền thêm tủ vào hệ thống, sửa thông tin tủ và xóa tủ ra khỏi hệ thống
- Usecase mở tủ: thành viên hệ thống có quyền mở tủ từ xa. trong đó:
    - User phải được xác thực thông tin (đăng nhập vào hệ thống hoặc khuôn mặt được nhận diện đúng) thì mới ở được tủ.
    - Admin chỉ được mở tủ từ xa thông qua quyền hệ thống khi có yêu cầu từ người dùng (lỗi, …).

- Usecase mở bằng khuôn mặt được extend từ usecase user mở tủ: hệ thống tự động nhận diện khuôn mặt của người dùng thông qua model AI và thực hiện cơ chế xác thực để mở tủ
- Usecase gửi cảnh báo an ninh được extend từ usecase mở bằng khuôn mặt: trong trường hợp xác thực khuôn mặt bị sai (không đúng user sở hữu tủ), hệ thống sẽ gửi cảnh báo an ninh về tài khoản của người dùng



IV. Xác định các yêu cầu phi chức năng

1. Hiệu năng

- Độ trễ hệ thống
  - Thời gian nhận diện khuôn mặt đến khi mở khóa tủ ≤ 5 giây.
  - Độ trễ khi gửi lệnh mở/đóng từ xa ≤ 3 giây qua Wi-Fi hoặc MQTT.
- Khả năng chịu tải:
  - Hệ thống có thể xử lý tối thiểu 10 yêu cầu truy cập/giây mà không gián đoạn.
  - Có khả năng hoạt động đồng thời với nhiều tủ thông minh (≥10 module ESP32-CAM) mà không xung đột mạng.
- Tần suất lấy mẫu:
  - Camera ESP32-CAM chụp ảnh khuôn mặt mỗi khi phát hiện chuyển động (tối đa 1 ảnh/2 giây/sự kiện).
- Tối ưu năng lượng và băng thông:
  - Ảnh khuôn mặt được nén JPEG < 50KB trước khi gửi lên server.

    - Dữ liệu điều khiển truyền qua MQTT (QoS = 1) để đảm bảo phản hồi nhanh và tiết kiệm lưu lượng.

2. Bảo mật

- Mã hóa: Dữ liệu khuôn mặt và thông tin người dùng được mã hóa bằng AES-256 khi lưu trữ,và sử dụng TLS/SSL trong quá trình truyền qua Wi-Fi.
- Xác thực và phân quyền:
    - Chỉ người dùng đã đăng ký mới có thể mở khóa hoặc quản lý tủ.
    - Admin có thể thêm/xóa/sửa dữ liệu khuôn mặt
    - Người dùng chỉ được phép mở tủ của mình.
- Bảo vệ dữ liệu cá nhân:
    - Tuân thủ Luật An ninh mạng 2018 và Nghị định 13/2023/NĐ-CP về bảo vệ dữ liệu cá nhân.
    - Người dùng có quyền yêu cầu xóa dữ liệu khuôn mặt khỏi hệ thống.
- Cảnh báo truy cập trái phép khi có người lạ định mở khóa tủ.

3. Độ tin cậy

- Tỉ lệ hoạt động ổn định: ≥ 98% uptime trong điều kiện Wi-Fi ổn định.
- Cơ chế dự phòng:
    - Hệ thống tự động retry gửi dữ liệu khi mất kết nối mạng.
    - Khi Wi-Fi yếu, có thể chuyển sang BLE để điều khiển cục bộ.
- Chống mất dữ liệu: Dữ liệu nhận diện và nhật ký mở khóa được lưu trên Firebase hoặc MySQL có cơ chế sao lưu định kỳ.
- Phục hồi lỗi: Khi mất điện, hệ thống khởi động lại và tự động đồng bộ trạng thái khóa cuối cùng.

4. Khả năng mở rộng

- Mở rộng thiết bị: Hỗ trợ thêm nhiều tủ thông minh (mỗi tủ 1 ESP32-CAM) mà không cần thay đổi kiến trúc mạng.
- Mở rộng người dùng: Cơ sở dữ liệu khuôn mặt và tài khoản có thể mở rộng đến hàng trăm người.
- Tích hợp hệ thống: Có thể kết nối với các nền tảng IoT khác (Google Firebase, Node-RED, hoặc MQTT Cloud).

5. Chi phí và năng lượng

- Chi phí phần cứng: ≤ 500.000 VNĐ/tủ, bao gồm ESP32-CAM, servo SG90, cảm biến PIR và nguồn 5V.
- Tiêu thụ điện năng:
    - Trung bình < 200 mA trong chế độ hoạt động và < 20 mA khi ở chế độ sleep.
    - Khi mở rộng nhiều tủ, có thể dùng adapter USB đa cổng để giảm chi phí cấp nguồn.
- Tiết kiệm năng lượng:
    - Kích hoạt deep sleep mode cho ESP32-CAM, chỉ bật camera khi phát hiện chuyển động.
    - Giảm tần suất gửi ảnh nếu không có thay đổi lớn.

6. Rủi ro và đối sách

- Mất kết nối Wi-Fi: Lưu lệnh cục bộ, retry định kỳ, fallback sang BLE
- Nhận diện sai khuôn mặt: Cải thiện thuật toán, yêu cầu chụp lại ảnh trong điều kiện sáng tốt.

- Mất điện: Trang bị UPS mini dự phòng 30 phút
- Quá tải bộ nhớ ESP32: Tối ưu nén ảnh, xử lý tiền xử lý trước khi gửi cloud
- Lộ dữ liệu khuôn mặt: Mã hóa dữ liệu và yêu cầu xác thực trước truy cập

V. Phân tích ràng buộc kỹ thuật và môi trường

1. Môi trường hoạt động

- Hệ thống tủ thông minh được thiết kế hoạt động chủ yếu trong môi trường trong nhà, phù hợp với các không gian như ký túc xá, văn phòng, khu vực học tập hoặc gia đình, nơi có điều kiện ổn định hơn so với môi trường ngoài trời. Các yếu tố môi trường cần xem xét bao gồm:
    - Nhiệt độ và độ ẩm: Hoạt động ở biên độ nhiệt độ 20°C – 35°C và độ ẩm trung bình 40-70%. → ESP32-CAM, cảm biến PIR và servo motor có thể chịu được dải này mà không cần vỏ bảo vệ đặc biệt, nhưng cần tránh lắp đặt gần nguồn nhiệt cao (như máy sưởi) hoặc khu vực ẩm ướt (nhà tắm) để tránh hỏng camera.
    - Nhiễu sóng: Môi trường văn phòng hoặc ký túc xá có thể có nhiều thiết bị Wi-Fi gây nhiễu 2.4GHz. → Chọn kênh Wi-Fi ít nhiễu (tự động scan qua code ESP32), hoặc fallback sang BLE nếu cần kết nối ngắn khoảng cách. Thử nghiệm thực địa để đảm bảo tín hiệu ổn định trong bán kính 10-20m.
    - Nguồn cấp: Sử dụng nguồn 5V ổn định từ adapter USB hoặc cổng sạc (≥1A). → Không có điện lưới không ổn định, nhưng cần UPS nhỏ cho demo để tránh gián đoạn khi mất điện ngắn (dự phòng 30 phút).
    - Cảm biến chuyển động PIR cần được lắp ở vị trí thoáng, không bị che khuất và hướng về khu vực người dùng thường tiếp cận tủ. Tránh đặt cảm biến gần nguồn nhiệt, ánh sáng mặt trời trực tiếp hoặc quạt gió mạnh, vì có thể gây kích hoạt sai (false trigger). → Khoảng cách phát hiện hiệu quả của PIR là 3–5 mét, góc quét khoảng 120°, phù hợp cho khu vực nhỏ như phòng hoặc hành lang.

2. Ràng buộc pháp lý

- Hệ thống liên quan đến dữ liệu sinh trắc học (khuôn mặt) và truyền dữ liệu qua mạng, nên phải tuân thủ các quy định pháp lý tại Việt Nam và quốc tế để tránh rủi ro pháp lý:
    - Quy định tần số vô tuyến: Wi-Fi hoạt động ở dải 2.4GHz và 5GHz được phép theo Quyết định 103/2016/QĐ-TTg của Bộ Thông tin và Truyền thông Việt Nam. → Cấu hình ESP32-CAM đúng dải tần, tránh can thiệp tần số không được phép để không bị xử phạt.
    - Bảo mật dữ liệu cá nhân: Dữ liệu khuôn mặt là thông tin nhạy cảm, phải tuân thủ Luật An ninh mạng 2018 và Nghị định 13/2023/NĐ-CP về bảo vệ dữ liệu cá nhân (tương tự GDPR). → Xin phép người dùng trước khi lưu trữ dữ liệu, mã hóa AES-256, và hỗ trợ quyền xóa dữ liệu (right to be forgotten). Không chia sẻ dữ liệu với bên thứ ba mà không có sự đồng ý.

3. Tài nguyên thiết bị

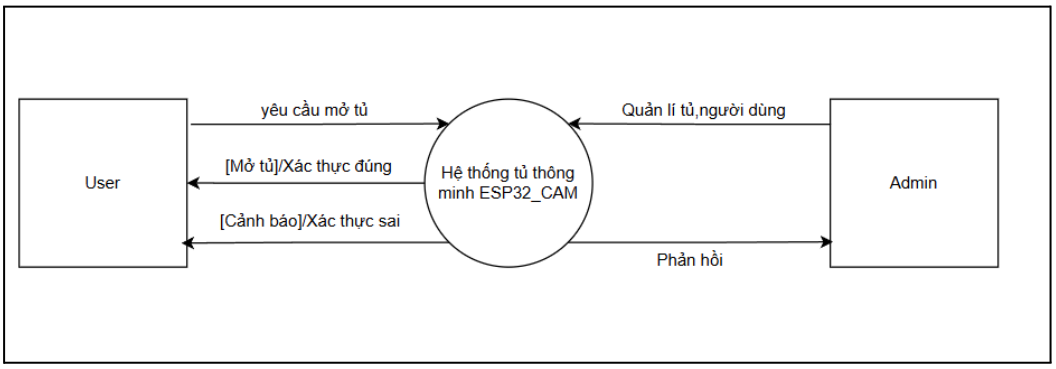
- Hệ thống sử dụng **ESP32-CAM làm vi điều khiển trung tâm**, tích hợp sẵn chip **ESP32-WROOM** (SoC Wi-Fi + Bluetooth) và camera **OV2640 2MP**, giúp xử lý hình ảnh, kết nối mạng và điều khiển thiết bị ngoại vi mà không cần thêm bộ vi điều khiển riêng biệt.

- Các tài nguyên phần cứng chính bao gồm:
  - **ESP32-CAM:**
    - CPU Dual-core Xtensa LX6 (240 MHz), RAM 520KB, Flash 4MB
    - Vi điều khiển trung tâm, tích hợp camera OV2640, chip Wi-Fi ESP32-WROOM, và các chân GPIO điều khiển servo/PIR.
  - **Camera OV2640:**
    - Nén JPEG trực tiếp trên ESP32-CAM
    - Thu nhận hình ảnh khuôn mặt, độ phân giải tối đa 1600x1200.
  - **Servo motor SG90:**
    - Điện áp 5V, góc quay 180°, dòng 500mA
    - Đóng/mở cơ cấu chốt tủ theo tín hiệu PWM.
  - **Cảm biến PIR (HC-SR501):**
    - Điện áp 5V, dòng 50µA khi standby.
    - Phát hiện chuyển động để kích hoạt camera chụp ảnh hoặc đánh thức ESP32-CAM khỏi chế độ ngủ (Deep Sleep).
  - **Nguồn cấp:**
    - Cấp nguồn cho toàn bộ hệ thống.
    - 5V – 1A (adapter hoặc cổng USB), có thể bổ sung UPS mini dự phòng 30 phút

VI. Xây dựng mô hình yêu cầu

1. Sơ đồ luồng DFD

a. DFD mức 0

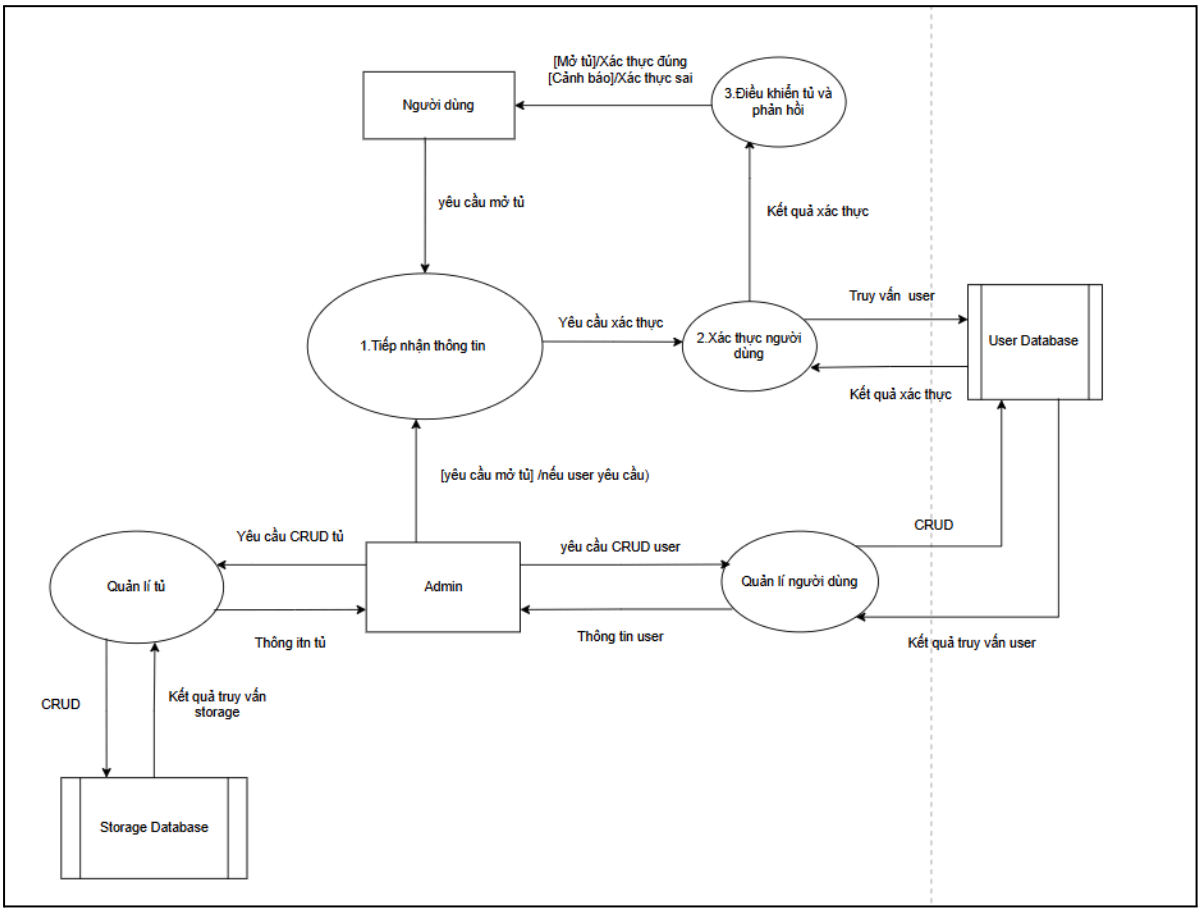


- Các thành phần trong sơ đồ

Thành phần	Vai trò
User	Tác động vật lý lên tủ thông qua nút mở. Truy cập website để gửi lệnh mở/đóng tủ từ xa qua Internet. Nhận cảnh báo khi xác thực sai.
Tủ thông minh (ESP32-CAM)	Là thiết bị IoT chính: nhận lệnh mở từ người dùng, chụp ảnh khuôn mặt, gửi ảnh lên server AI, nhận phản hồi xác thực và điều khiển server mở khóa. Nếu nhận diện khuôn mặt thất bại, hệ thống sẽ gửi thông

	báo đẩy đến tài khoản người dùng (Bao gồm ảnh chụp người lạ và thời gian sự kiện).
Admin	Truy cập website để gửi lệnh mở/đóng tủ từ xa qua Internet (chỉ khi có yêu cầu từ người dùng). Thêm/sửa/xóa các thông tin về tủ, hỗ trợ mở rộng cho nhiều tủ.

b. DFD mức 1



Thành phần	Vai trò
User	Tác động vật lý lên tủ thông qua nút mở. Truy cập website để gửi lệnh mở/đóng tủ từ xa qua Internet. Nhận cảnh báo khi xác thực sai.
Hệ thống tủ thông minh (ESP32-CAM)	Được phân rã thành 3 tiến trình: 1.Tiếp nhận thông tin: Nhận yêu cầu từ người dùng hoặc admin. 2.Xác thực người dùng: Kiểm tra thông tin trong *User Database*. 3.Điều khiển tủ và phản hồi: Gửi lệnh mở/đóng tủ, phản hồi kết quả cho người dùng

Admin	Truy cập website để gửi lệnh mở/đóng tủ từ xa qua Internet (chỉ khi có yêu cầu từ người dùng). Thêm/sửa/xóa các thông tin về tủ, hỗ trợ mở rộng cho nhiều tủ. Thêm/sửa/xóa thông tin của người dùng
Quản lý người dùng	Thực hiện chức năng thêm/sửa/xóa người dùng
User Database	Cơ sở dữ liệu chứa các thông tin người dùng
Quản lý tủ	Thực hiện chức năng thêm/sửa/xóa tủ
Storage Database	Cơ sở dữ liệu chứa các thông tin tủ

2. Sơ đồ kiến trúc IoT 3 lớp

- Hệ thống Tủ thông minh nhận diện khuôn mặt được thiết kế dựa trên mô hình kiến trúc 3 lớp IoT (Perception – Network – Application).
- Cách tiếp cận này giúp phân tách rõ vai trò phần cứng, truyền thông và xử lý dữ liệu, đảm bảo khả năng mở rộng và bảo trì dễ dàng.

a. Perception Layer (Tầng Cảm nhận – Phần cứng):

- Chức năng: Thu thập dữ liệu khuôn mặt, trạng thái vật lý của tủ, phát hiện chuyển động và điều khiển cơ cấu đóng/mở tủ..
- Thành phần chính:
    - ESP32-CAM: Vi điều khiển có camera OV2640 (2MP), chụp hình khuôn mặt và xử lý sơ bộ (nén JPEG).
    - Servo motor SG90: Thực hiện đóng/mở chốt tủ theo lệnh điều khiển từ hệ thống sau khi xác thực khuôn mặt thành công..
    - Cảm biến PIR (HC-SR501): Phát hiện chuyển động hoặc người tiếp cận tủ, kích hoạt ESP32-CAM chụp ảnh và gửi dữ liệu.
    - Nguồn 5V DC: Cung cấp năng lượng cho toàn bộ hệ thống, có thể lấy từ adapter hoặc pin sạc dự phòng.
- Yêu cầu kỹ thuật:
    - Thu thập ảnh realtime (2–5 giây/sự kiện).
    - Nén ảnh trước khi gửi để giảm tải đường truyền.
    - Đảm bảo hoạt động ổn định trong điều kiện nhiệt độ phòng (20–35°C).

b. Network Layer (Tầng mạng – Truyền thông)

- Chức năng: Truyền dữ liệu giữa phần cứng (ESP32-CAM) và máy chủ (server AI & web).
- Thành phần và công nghệ:
    - Wi-Fi (IEEE 802.11 b/g/n): Kết nối ESP32-CAM với mạng nội bộ (bán kính 10–20m).
    - MQTT (Message Queuing Telemetry Transport): Giao thức lightweight truyền dữ liệu thời gian thực (publish/subscribe) giữa thiết bị và server Node.js → Dùng cho các sự kiện mở khóa, cảnh báo, log truy cập.

- HTTP/HTTPS: Gửi ảnh khuôn mặt (HTTP POST) và nhận phản hồi xác thực từ server AI. → Hỗ trợ mã hóa TLS/SSL để đảm bảo an toàn dữ liệu sinh trắc học.
- Bảo mật: Mã hóa TLS/SSL, sử dụng QoS = 1 cho các gói tin điều khiển realtime (mở khóa, cảnh báo).
- Dự phòng: Có thể mở rộng sử dụng Bluetooth Low Energy (BLE) khi Wi-Fi yếu.

- Yêu cầu:
  - Đảm bảo truyền dữ liệu ổn định, hỗ trợ cơ chế retry khi mất kết nối.
  - Độ trễ truyền ảnh < 3 giây; độ tin cậy truyền tin > 98%.

c. Application Layer (Tầng ứng dụng – Phân tích và điều khiển)

- Chức năng: Phân tích dữ liệu khuôn mặt, xác thực người dùng, quản lý tủ và hiển thị thông tin qua giao diện web/app.
- Thành phần:
  - Server AI (Node.js + mô hình AIt): xử lý bằng mô hình nhận diện khuôn mặt.
  - Cơ sở dữ liệu (MySQL / Firebase): Lưu trữ thông tin người dùng, lịch sử truy cập và nhật ký cảnh báo.
  - Website (React.js): Hiển thị dashboard realtime: trạng thái tủ (mở/đóng), log truy cập, cảnh báo, và hỗ trợ điều khiển mở tủ từ xa.
  - Firebase Cloud Messaging (FCM): Gửi thông báo realtime khi có người lạ hoặc hành vi truy cập trái phép.
- Yêu cầu:
  - Giao diện trực quan, hiển thị trạng thái realtime.
  - Hỗ trợ phân quyền: chỉ người dùng hợp lệ mới mở khóa hoặc cấp quyền cho người khác.
  - Cho phép mở rộng quản lý **nhiều tủ – nhiều người dùng** trên cùng hệ thống.

B. CƠ SỞ LÝ THUYẾT & CÔNG NGHỆ ÁP DỤNG

I. Kiến thức nền tảng liên quan

Các nội dung trình bày sau là các khái niệm cơ bản trong IoT, AI và an ninh mạng. Những kiến thức này là nền tảng để nhóm có thể phát triển được hệ thống tủ thông minh.

1. Internet of Things (IoT) và kiến trúc 3 lớp

IoT là mạng lưới các thiết bị vật lý kết nối Internet để thu thập, trao đổi dữ liệu mà không cần can thiệp con người, nhằm tự động hóa và tối ưu hóa quy trình (theo định nghĩa từ IEEE). Kiến trúc IoT 3 lớp bao gồm Perception Layer với chức năng thu thập dữ liệu từ cảm biến, Network Layer truyền thông qua Wi-Fi/MQTT, và Application Layer để xử lý dữ liệu trên cloud/server. Trong đề tài, mô hình này được áp dụng để ESP32-CAM thu thập hình ảnh khuôn mặt, truyền qua Wi-Fi, và xử lý nhận diện trên server AI, giúp hệ thống hoạt động ổn định với độ trễ và chi phí phù hợp.

2. Nhận diện khuôn mặt (Face Recognition)

Nhận diện khuôn mặt là kỹ thuật sử dụng AI để xác định danh tính cá nhân dựa trên đặc trưng khuôn mặt (như khoảng cách mắt, mũi), với độ chính xác >90% trong điều kiện ánh sáng tốt (sử dụng mô hình như TensorFlow Lite hoặc Firebase ML Kit). Đây là một phần của sinh trắc học, loại bỏ rủi ro từ chìa khóa vật lý hoặc RFID, và phù hợp với hệ thống an ninh gia đình/văn phòng. Trong đề tài, chức năng này được offload lên cloud để vượt qua hạn chế bộ nhớ của ESP32-CAM, đảm bảo sai số <10% và hỗ trợ mở rộng cho nhiều người dùng.

3. Giao thức truyền thông trong IoT

Giao thức như MQTT (Message Queuing Telemetry Transport) và HTTP được sử dụng để truyền dữ liệu nhẹ, realtime giữa thiết bị và server, với QoS=1 để đảm bảo độ tin cậy >90%. Wi-Fi hỗ trợ kết nối không dây bán kính 10-20m, kết hợp TLS/SSL để mã hóa dữ liệu. Trong hệ thống, MQTT được áp dụng để gửi lệnh mở tủ từ xa hoặc cảnh báo an ninh, giảm độ trễ mạng và fallback.

II. Công nghệ phần cứng

Phần cứng là nền tảng của hệ thống IoT, bao gồm các module thu thập dữ liệu, xử lý và thực thi. Trong đề tài **"Tủ thông minh nhận diện khuôn mặt"**, hệ thống được xây dựng dựa trên các thành phần chính sau:

1. **ESP32-CAM - Vi điều khiển trung tâm:**

- ESP32-CAM là một module camera phát triển dựa trên vi điều khiển ESP32 của Espressif, tích hợp khả năng xử lý hình ảnh và truyền dữ liệu qua Wi-Fi, rất phù hợp cho các ứng dụng IoT(Internet of Things) như giám sát an ninh, nhận diện khuôn mặt, và các dự án camera thông minh. Với kích thước nhỏ gọn và giá thành

thấp, ESP32-CAM đã trở thành một lựa chọn phổ biến trong cộng đồng lập trình viên và các nhà phát triển hệ thống nhúng

Hình: Module ESP32 - CAM - MB

- ESP32-CAM được tích hợp đầy đủ chức năng xử lý hình ảnh, kết nối Wi-Fi, Bluetooth và các cổng giao tiếp mở rộng.
    - Camera OV2640 (2MP, hỗ trợ độ phân giải lên đến 1600x1200)
    - Wi-Fi 802.11 b/g/n và Bluetooth 4.2.
- **Thông số kỹ thuật chính:**

Thuộc tính	Giá trị
Vi xử lý	Dual-core Xtensa LX6, xung nhịp 240 MHz
Bộ nhớ	520 KB SRAM, 4MB Flash
Camera tích hợp	OV2640 – 2MP (hỗ trợ độ phân giải 1600×1200)
Kết nối không dây	Wi-Fi 802.11 b/g/n, Bluetooth 4.2 BLE
Giao tiếp hỗ trợ	UART, SPI, I2C, PWM, GPIO
Nguồn hoạt động	5V DC
Mức tiêu thụ điện	< 200 mA khi hoạt động
Kích thước	Khoảng 40 × 27 mm

- **Ưu điểm nổi bật:**
    - Tích hợp camera, không cần module rời.
    - Giá thành rẻ, khoảng dưới 200.000 VNĐ, dễ mua và thay thế.

- Lập trình đơn giản qua Arduino IDE, hỗ trợ cập nhật OTA (Over-the-Air).
- Xử lý ảnh sơ bộ và nén JPEG trực tiếp, giảm tải cho server.

- **Vai trò trong hệ thống:**
  - ESP32-CAM là bộ điều khiển trung tâm, thực hiện:
  - Chụp ảnh khuôn mặt khi phát hiện chuyển động.
  - Gửi ảnh qua Wi-Fi đến server AI để nhận dạng.
  - Nhận phản hồi xác thực và điều khiển servo mở khóa.
  - Cảnh báo người lạ.

2. Servo Motor SG90 – Bộ truyền động cơ khóa

- Servo SG90 được sử dụng làm bộ chấp hành (actuator) để điều khiển cơ cấu khóa của tủ. Khi hệ thống xác thực khuôn mặt hợp lệ, ESP32-CAM gửi tín hiệu PWM đến servo để mở chốt khóa, sau đó tự động đóng lại sau một khoảng thời gian định sẵn

Hình: Servo SG90

- **Thông số kỹ thuật:**

Thuộc tính	Giá trị
Điện áp hoạt động	5V DC
Dòng tiêu thụ	Khoảng 500 mA
Góc quay	0° – 180°
Điều khiển	Tín hiệu PWM từ ESP32-CAM
Trọng lượng	9g
Tốc độ phản hồi	0.1 s/60°

- **Ưu điểm:**
    - Kích thước nhỏ, dễ lắp đặt trong khoang tủ.
    - Hoạt động ổn định và chính xác, phù hợp với nguồn USB 5V.
    - Phản hồi nhanh, hỗ trợ điều khiển mượt qua PWM.

3. Cảm biến PIR (HC-SR501)

- Cảm biến PIR giúp phát hiện chuyển động của con người để kích hoạt camera chụp ảnh hoặc đánh thức ESP32-CAM khỏi chế độ deep sleep, giúp tiết kiệm năng lượng.

Hình: Cảm biến PIR (HC-SR501)

- **Thông số kỹ thuật:**

Thuộc tính	Giá trị
Điện áp hoạt động	5V DC
Dòng tiêu thụ	<50 µA (standby), ~100 µA (hoạt động)
Khoảng cách phát hiện	3 – 7 mét
Góc phát hiện	~120°
Thời gian trễ	0.5 – 200 giây (có thể điều chỉnh)

- **Ưu điểm:**
    - Tiêu thụ năng lượng cực thấp, phù hợp IoT.
    - Giúp giảm tải CPU khi không có người đến gần.
    - Nâng cao tính thông minh của hệ thống: camera chỉ kích hoạt khi có người.

4. Tổng quan hệ thống phần cứng

- Tất cả các module được cấp nguồn 5V DC (qua adapter USB hoặc pin sạc).

- Hệ thống được thiết kế hoạt động độc lập, không cần máy tính trung gian.
- Nguồn cấp:
    - **Nguồn chính:** Adapter 5V – 1A, đảm bảo đủ dòng cho ESP32-CAM và servo.
    - **Nguồn dự phòng:** Pin LiPo 3.7V (kèm mạch boost 5V) hoặc UPS mini 1000 mAh, hoạt động 30–60 phút khi mất điện.
- Kết nối giữa các thành phần:
    - PIR sensor → GPIO Input ESP32-CAM (phát hiện chuyển động).
    - Servo motor → GPIO PWM ESP32-CAM (điều khiển chốt).
    - ESP32-CAM ↔ Wi-Fi ↔ Server Node.js (gửi ảnh, nhận lệnh mở khóa).
    - Server ↔ Web/App (React/Firebase) để hiển thị và điều khiển realtime

III. Các giao thức truyền thông

Giao thức truyền thông là cầu nối giữa các lớp trong kiến trúc IoT, đảm bảo dữ liệu được truyền tải an toàn, ổn định và theo thời gian thực. Trong hệ thống Tủ thông minh nhận diện khuôn mặt, nhóm sử dụng Wi-Fi làm nền tảng kết nối chính, kết hợp MQTT cho truyền thông hai chiều realtime và HTTP/HTTPS cho các dịch vụ web và API cloud.

1. Wi-Fi (IEEE 802.11)

- Wi-Fi là giao thức truyền thông không dây phổ biến, hoạt động ở dải tần 2.4GHz, cho tốc độ truyền dữ liệu lên tới 150Mbps và phạm vi hoạt động từ 10–50 mét trong môi trường trong nhà.
- Module ESP32-CAM được tích hợp Wi-Fi và có thể hoạt động ở hai chế độ:
    - Station (Client Mode): Kết nối vào mạng Wi-Fi sẵn có để gửi dữ liệu lên server hoặc cloud.
    - Access Point (AP Mode): Tạo mạng nội bộ để người dùng điều khiển và cấu hình hệ thống trực tiếp.
- Ưu điểm:
    - Triển khai dễ dàng, không cần gateway trung gian.
    - Băng thông lớn, đủ để truyền ảnh hoặc video.
    - Hỗ trợ giao tiếp song song với nhiều thiết bị khác trong cùng mạng.
- Hạn chế:
    - Tiêu thụ điện năng cao hơn so với BLE hoặc ZigBee.
    - Phụ thuộc vào độ ổn định mạng Wi-Fi khi truyền dữ liệu liên tục.

2. MQTT (Message Queuing Telemetry Transport)

- MQTT là giao thức truyền thông nhẹ (lightweight protocol) được thiết kế tối ưu cho các hệ thống IoT có tài nguyên hạn chế.
- Giao thức này hoạt động theo mô hình Publish / Subscribe, giúp các thiết bị gửi và nhận dữ liệu qua broker trung gian mà không cần kết nối trực tiếp.
- Đặc điểm nổi bật:
    - Gói tin nhỏ (header chỉ 2 bytes) → tiết kiệm băng thông.
    - Hỗ trợ QoS (Quality of Service) – đảm bảo tin nhắn được truyền đến nơi nhận:
        - QoS 0: gửi một lần, không xác nhận.
        - QoS 1: gửi ít nhất một lần (dùng cho lệnh mở khóa).
        - QoS 2: gửi đúng một lần (cho dữ liệu quan trọng).
    - Hỗ trợ truyền dữ liệu hai chiều (bi-directional) giữa thiết bị và server theo thời gian thực.
- Vai trò trong hệ thống:

- Truyền thông báo realtime từ ESP32-CAM đến server khi phát hiện khuôn mặt hoặc xâm nhập.
- Nhận lệnh điều khiển mở/đóng khóa từ ứng dụng web hoặc mobile.
- Giúp hệ thống hoạt động ổn định, phản hồi nhanh và tiết kiệm năng lượng hơn so với HTTP polling.

3. HTTP / HTTPS (HyperText Transfer Protocol Secure)

- HTTP/HTTPS được sử dụng cho các chức năng trao đổi dữ liệu với web server và API cloud.
- Cụ thể, ESP32-CAM gửi hình ảnh khuôn mặt đến server AI hoặc Firebase bằng phương thức HTTP POST, sau đó nhận phản hồi xác thực từ server để điều khiển servo mở khóa.
- Tính năng chính:
  - Mã hóa bảo mật: HTTPS kết hợp TLS/SSL để bảo vệ dữ liệu khuôn mặt và thông tin người dùng.
  - Tương thích cao: Dễ dàng tích hợp với các nền tảng web, cloud và database.
  - Phù hợp với dashboard web: hỗ trợ cảnh báo qua trình duyệt.

IV. Công nghệ phần mềm

Phần mềm đóng vai trò kết nối, xử lý và hiển thị thông tin, giúp hệ thống Tủ thông minh nhận diện khuôn mặt hoạt động trơn tru từ tầng thiết bị đến người dùng cuối. Các thành phần phần mềm được chia thành 3 lớp chính: lập trình nhúng (firmware), xử lý backend và giao diện frontend.

1. Lập trình nhúng (Firmware cho ESP32-CAM)

- Công cụ phát triển:
  - Arduino IDE, hỗ trợ ngôn ngữ C/C++, thuận tiện cho lập trình vi điều khiển.
  - Hỗ trợ cập nhật firmware từ xa (OTA – Over-the-Air) giúp dễ bảo trì và khắc phục lỗi.
- Chức năng chính:
  - Quản lý kết nối Wi-Fi và MQTT client để truyền/nhận dữ liệu.
  - Chụp ảnh khuôn mặt bằng camera OV2640, nén ảnh JPEG trước khi gửi.
  - Gửi ảnh đến server qua HTTP POST hoặc MQTT topic.
  - Nhận phản hồi từ server để điều khiển servo SG90 mở/đóng khóa.
  - Gửi log truy cập và cảnh báo bất thường lên server
- Thư viện sử dụng:
  - ESP32Cam.h, WiFi.h, PubSubClient.h, ESPAsyncWebServer.h.
  - Hỗ trợ streaming video, gửi ảnh dạng Base64 hoặc JPEG.
  - Hỗ trợ deep sleep mode để tiết kiệm năng lượng.

2. Backend – Xử lý dữ liệu và API

- Ngôn ngữ & Framework:
  - Node.js (JavaScript runtime) kết hợp Express.js cho RESTful API.
  - MQTT Broker: sử dụng Mosquitto hoặc thư viện mqtt.js để xử lý giao tiếp realtime giữa ESP32 và server.
- Chức năng chính:

- Nhận và xử lý hình ảnh khuôn mặt từ ESP32-CAM qua HTTP POST hoặc MQTT topic.
- Xác thực và lưu trữ dữ liệu (thông tin người dùng, lịch sử mở khóa, cảnh báo) vào Firebase hoặc MySQL.
- Gửi lệnh điều khiển (mở/khóa tủ) và cảnh báo realtime đến frontend qua Socket.IO.
- Cung cấp API quản lý người dùng: thêm, sửa, xóa, hoặc cập nhật khuôn mặt.

3. Frontend – Giao diện Web và Ứng dụng người dùng

- Công nghệ sử dụng:
  - React.js: thư viện JavaScript xây dựng giao diện người dùng linh hoạt.
  - Socket.IO Client: cập nhật realtime trạng thái thiết bị.
  - Axios: giao tiếp API với backend Node.js.
  - Bootstrap / Tailwind CSS: thiết kế giao diện trực quan, responsive.
  - Firebase Cloud Messaging (FCM): gửi thông báo khi phát hiện xâm nhập hoặc khuôn mặt lạ.
- Chức năng chính:
  - Hiển thị Dashboard realtime:
    - Trạng thái tủ (mở/đóng).
    - Cảnh báo an ninh.
  - Điều khiển từ xa:
    - Gửi lệnh mở tủ qua API hoặc MQTT.
    - Hỗ trợ chia sẻ quyền mở tủ (OTP hoặc token tạm thời).
  - Quản lý người dùng:
    - Admin thêm/sửa thông tin, upload ảnh khuôn mặt mới.
    - Theo dõi danh sách thiết bị, hoạt động theo thời gian thực.

V. Công nghệ AI

1. Lý thuyết về nhận diện khuôn mặt

- Hệ thống nhận diện khuôn mặt hoạt động dựa trên ba giai đoạn chính:
  - Phát hiện khuôn mặt (Face Detection): ESP32-CAM chụp ảnh người đứng trước tủ và phát hiện vùng khuôn mặt trong khung hình (sử dụng mô hình nhẹ như Haar Cascade hoặc MTCNN).
  - Trích xuất đặc trưng (Feature Extraction): Ảnh khuôn mặt được gửi lên server AI hoặc Firebase ML Kit, nơi mô hình học sâu (như FaceNet hoặc MobileFaceNet) biến ảnh thành vector đặc trưng (embedding).
  - So sánh và nhận diện (Face Matching): Vector này được so sánh với dữ liệu trong cơ sở dữ liệu người dùng.
    - Nếu độ tương đồng vượt ngưỡng (ví dụ 90%) → hệ thống gửi lệnh mở khóa.
    - Nếu không → gửi cảnh báo về web/app của người dùng.

2. Hệ thống mô hình AI

a. Model AI phân biệt giả mạo - CNN Classification

- Một vấn đề rất lớn trong việc nhận diện khuôn mặt là làm sao để phát hiện được đâu là thật và đâu là giả, bởi người dùng có thể sử dụng hình ảnh, video có sẵn trên điện thoại để thực hiện các hành vi giả mạo. Để giải quyết vấn đề này, mô hình **Convolutional Neural Network (CNN**) được sử dụng để phân loại ảnh thành hai nhãn: thật và giả, dựa trên việc phân tích các đặc trưng ảnh chi tiết. Hình ảnh khuôn mặt thật và giả thường có các khác biệt đặc trưng, mà CNN có thể nhận diện thông qua việc học tự động từ dữ liệu. Một số đặc trưng tiêu biểu giúp phân loại bao gồm:
    - **Kết cấu da (Skin Texture)**: Khuôn mặt thật có kết cấu da phức tạp với các chi tiết nhỏ như nếp nhăn, lỗ chân lông, và ánh sáng phản chiếu tự nhiên. Trong khi đó, khuôn mặt giả từ ảnh in hoặc video phát lại thường thiếu kết cấu chi tiết này, và có thể bị nhòe (blur) hoặc xuất hiện các mẫu lặp (patterns).
    - **Ánh sáng và bóng đổ (Lighting and Shadows):** Ảnh thật thể hiện ánh sáng và bóng đổ tự nhiên theo hướng nguồn sáng, trong khi ảnh giả thường có ánh sáng không đồng nhất hoặc thiếu bóng đổ do sự cố định của hình ảnh.
    - **Tần số không gian (Spatial Frequency):** Hình ảnh giả thường thiếu các tần số cao, đặc trưng cho các chi tiết sắc nét trong ảnh thật. Điều này xảy ra do quá trình in ấn hoặc nén video làm giảm chất lượng hình ảnh.
    - **Hiệu ứng động (Dynamic Cues):** Trong video, khuôn mặt thật thường thể hiện các chuyển động tự nhiên của mắt, miệng, hoặc biểu cảm. Trong khi đó, các video phát lại hoặc hình ảnh tĩnh không có những hiệu ứng động này

- **CNN** phân loại hình ảnh thông qua một chuỗi các bước:
    - **Trích xuất đặc trưng (Feature Extraction):** Các tầng tích chập (Convolutional Layers) tự động trích xuất các đặc trưng như kết cấu, ánh sáng, và tần số từ dữ liệu đầu vào.
    - **Giảm kích thước (Pooling):** Các tầng pooling giảm kích thước của các đặc trưng, giúp loại bỏ nhiễu và tăng hiệu quả tính toán.
    - **Phân loại (Classification):** Sau khi trích xuất và xử lý, các đặc trưng được đưa qua các tầng kết nối đầy đủ (Fully Connected Layers) để đưa ra dự đoán khuôn mặt là thật hay giả.

- Hệ thống CNN được huấn luyện trên tập dữ liệu bao gồm cả hình ảnh thật và giả, từ đó học cách nhận diện các đặc điểm độc nhất của từng loại. Nhờ khả năng tự động trích xuất và phân tích đặc trưng, CNN mang lại độ chính xác cao trong việc phát hiện các hành vi giả mạo như sử dụng ảnh in, video phát lại, hoặc mặt nạ 3D.

b. Model AI phát hiện khuôn mặt - Face Detection (MT-CNN)

- **MT-CNN (Multi-Task Convolutional Neural Network)** là một mô hình học sâu mạnh mẽ được thiết kế đặc biệt cho nhiệm vụ phát hiện khuôn mặt từ hình ảnh đầu vào. Trước khi muốn nhận diện chi tiết khuôn mặt, mô hình phải phát hiện được đâu là khuôn mặt để tránh học những dữ liệu nhiễu từ nền (background) hoặc từ những bộ phận khác của cơ thể con người.
- MT-CNN được cấu trúc thành ba mạng con:
    - **P-Net (Proposal Network):** Được sử dụng để xác định các khu vực tiềm năng chứa khuôn mặt và trả về bounding box ban đầu.
    - **R-Net (Refine Network):** Tinh chỉnh bounding box từ P-Net và loại bỏ các vùng nhiễu không chứa khuôn mặt
    - **O-Net (Output Network):** Trích xuất đặc trưng chi tiết và định vị chính xác các đặc điểm khuôn mặt như mắt, mũi, và miệng.
- Nhờ vào khả năng đa nhiệm và kết hợp của ba mạng, MT-CNN đảm bảo hiệu suất cao trong việc phát hiện khuôn mặt, kể cả trong các tình huống khó khăn như góc nghiêng, ánh sáng yếu, hoặc khuôn mặt nhỏ trong ảnh.
- MT-CNN đóng vai trò tiền xử lý quan trọng trong pipeline phân loại khuôn mặt thật giả, vì nó giúp chuẩn hóa và tập trung vào khu vực khuôn mặt trước khi chuyển dữ liệu vào các mô hình nhận diện hoặc phân loại tiếp theo.

c. Model AI phân biệt khuôn mặt - Face Verification (FaceNet)

- **FaceNet** là một mô hình học sâu tiên tiến được phát triển bởi Google, nhằm giải quyết bài toán nhận diện và so sánh khuôn mặt với độ chính xác cao. Thay vì trực tiếp phân loại khuôn mặt, FaceNet học một không gian nhúng (embedding space), nơi mỗi khuôn mặt được biểu diễn bằng một vector đặc trưng cố định**.**
- Cách hoạt động của FaceNet:
    - **Học không gian nhúng (Embedding Space)**: FaceNet ánh xạ mỗi khuôn mặt vào một vector có kích thước cố định (thường là 128 chiều), sao cho khoảng cách giữa các vector của cùng một khuôn mặt nhỏ hơn khoảng cách với các vector của khuôn mặt khác.
    - **Triplet Loss Function:** Để tối ưu hóa không gian nhúng, FaceNet sử dụng hàm mất mát Triplet Loss, với ba vector:
      Anchor: Vector nhúng của khuôn mặt hiện tại.
      Positive: Vector nhúng của cùng một người.
      Negative: Vector nhúng của người khác.
      Hàm mất mát tối thiểu hóa khoảng cách giữa Anchor và Positive, đồng thời tối đa hóa khoảng cách với Negative.
- Ứng dụng của FaceNet trong hệ thống: FaceNet được huấn luyện trên hàng triệu hình ảnh khuôn mặt, giúp nó học được các đặc trưng tinh vi và tổng quát. Sau khi hoàn thành huấn luyện, mô hình có thể được áp dụng trên các bài toán nhận diện khuôn mặt cụ thể bằng cách:
    - Trích xuất vector nhúng từ khuôn mặt cần nhận diện.
    - So sánh vector này với vector của các khuôn mặt đã biết để tìm ra sự tương đồng

        - Kết hợp với các thuật toán phân loại truyền thống như SVM (Support Vector Machine) để phân loại khuôn mặt thật và giả dựa trên vector nhúng.
    - **Ưu điểm:**
        - Giảm khối lượng tính toán: Nhờ vào không gian nhúng tối ưu, việc xử lý chỉ cần thực hiện trên một số khuôn mặt cụ thể
        - Tính linh hoạt: FaceNet có thể áp dụng cho nhiều bài toán nhận diện khác nhau mà không cần thiết kế lại mô hình.
    - FaceNet, khi kết hợp với MT-CNN, tạo nên một pipeline hoàn chỉnh: MT-CNN thực hiện bước phát hiện và chuẩn hóa khuôn mặt, trong khi FaceNet đảm nhận việc nhận diện hoặc phân loại dựa trên đặc trưng vector nhúng

3. Vai trò của AI trong hệ thống tủ thông minh

- Xác thực người dùng: Phân biệt khuôn mặt hợp lệ với khuôn mặt lạ.
- Ra quyết định mở/khóa tủ: Tự động gửi lệnh điều khiển servo nếu nhận dạng chính xác.
- Cảnh báo an ninh: Gửi thông báo khi phát hiện khuôn mặt không trùng khớp.
- Học và mở rộng dữ liệu: Cho phép thêm người dùng mới, cải thiện độ chính xác qua thời gian.