CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

MÔ TẢ THUYẾT MINH KỸ THUẬT

MÔ TẢ TIÊU CHÍ KỸ THUẬT CỦA PHẦN MỀM SẢN PHẨM

Hà Nội, 2025

MỤC LỤC

PHẦN 1. TIÊU CHÍ CHỨC NĂNG AI **3**

I. Chức năng/công nghệ AI *3*

1. Thuật toán AI chuyên sâu 3

2. Phân tích thời gian thực 5

3. Tích hợp đa camera, liên kết xuyên camera 6

4. Trích xuất, tổng hợp, phân loại sự kiện 7

5. Khả năng mở rộng 9

PHẦN 2. TIÊU CHÍ KỸ THUẬT **11**

I.Mô hình hệ thống đề xuất *11*

II. Tiêu chuẩn công nghệ và tiêu chuẩn tích hợp đáp ứng *11*

1. Tiêu chuẩn về công nghệ sử dụng (Techstack) 11

2. Tiêu chuẩn về hạ tầng thiết bị 12

2.1. Thiết bị Camera 12

2.2. Thiết bị xử lý tại biên AI Box 13

2.3. Hệ thống Server (máy chủ trung tâm) 14

3. Tiêu chuẩn về nền tảng triển khai và Phần mềm quản lý 14

3.1. Hạ tầng triển khai: 14

3.2. Quản lý phiên bản và phân phối phần mềm 14

3.3. Lưu trữ dữ liệu 15

3.4. Giám sát hệ thống 15

3.5. Ghi log và giám sát hoạt động 15

3.6. Web quản lý tập trung 16

4. Tiêu chuẩn về an ninh – bảo mật 16

5. Tiêu chuẩn về vận hành – bảo trì 18

5.1. Đảm bảo tính ổn định và liên tục 18

5.2. Khả năng mở rộng quy mô 18

5.3. Đảm bảo hiệu suất và độ trễ thấp 19

5.4. Cơ chế High Availability (HA) 19

5.5. Cơ chế backup và phục hồi 19

5.6. Cập nhật và quản lý phiên bản ứng dụng 20

6. Tiêu chuẩn về hiệu suất 20

7. Tiêu chuẩn về khả năng tích hợp 21

7.1. Tích hợp với các hệ thống quản trị và điều hành hiện có 21

7.2. Tích hợp với cơ sở dữ liệu quốc gia 21

7.3. Yêu cầu về kiến trúc mở 21

7.4. Cơ chế xác thực và bảo mật tích hợp 22

PHẦN 1. TIÊU CHÍ CHỨC NĂNG AI

I. Chức năng/công nghệ AI

1. Thuật toán AI chuyên sâu

Công nghệ phát hiện vật thể **(Object Detection)** là nền tảng cốt lõi trong giải pháp, giúp tự động nhận diện và xác định vị trí các đối tượng như người, phương tiện, vật thể nguy hiểm… trong khung hình video theo thời gian thực. Nhờ sự phát triển mạnh mẽ của trí tuệ nhân tạo và học sâu, nhiều phương pháp tiên tiến cần được ứng dụng, nổi bật nhất là YOLO (You Only Look Once), SSD (Single Shot Multibox Detector) và Faster R-CNN (Faster Region-based Convolutional Neural Network).

Một số công nghệ thuật toán AI chuyên sâu bao gồm:

a. Tích hợp các mô hình AI tiên tiến

- Hỗ trợ triển khai, tích hợp các mô hình học sâu (Deep Learning) hiện đại như:

- YOLO (You Only Look Once): Phát hiện nhanh đối tượng, phương tiện, biển số xe, vật thể lạ, người mang vật nguy hiểm, hành vi bất thường...
- SSD (Single Shot MultiBox Detector): Nhận diện đối tượng với tốc độ cao, phù hợp giám sát thời gian thực.
- Faster R-CNN: Nhận diện chính xác các đối tượng, hỗ trợ nhận diện khuôn mặt, biển số, vật thể nhỏ trong khung hình.
- Cho phép lựa chọn, cấu hình mô hình AI phù hợp từng kịch bản giám sát (giao thông, an ninh, khu vực trọng điểm...).

b. Nhận diện đa lớp, nhiều sự kiện đồng thời

- Khả năng phát hiện, nhận diện đa lớp (multi-class): cùng lúc phân biệt nhiều loại đối tượng (người, ô tô, xe máy, xe tải, biển số, vật thể lạ, hung khí...).

- Nhận diện đồng thời nhiều sự kiện khác nhau trên nhiều luồng camera, điểm giám sát:

- Tụ tập đông người
- Đánh nhau, gây rối, hành vi bạo lực
- Xâm nhập trái phép
- Mang hung khí (dao, súng...)
- Cháy, nổ, khói, nguy cơ hỏa hoạn
- Vi phạm giao thông, dừng đỗ sai quy định
- Nhận diện khuôn mặt đối tượng truy nã, cấm vào khu vực
- Hỗ trợ cảnh báo tức thời khi phát hiện sự kiện bất thường, đồng bộ dữ liệu về Trung tâm điều hành và Cơ sở dữ liệu dùng chung.

c. Tùy chỉnh, mở rộng và tối ưu

- Cho phép huấn luyện, cập nhật mô hình AI với dữ liệu thực tế từng khu vực, địa phương nhằm tăng độ chính xác nhận diện.

- Hỗ trợ đánh giá, so sánh hiệu quả các mô hình trên từng loại camera, môi trường giám sát.

- Dễ dàng tích hợp, mở rộng thêm các mô hình AI mới theo xu hướng công nghệ.

d. Một số tính năng khác

- Đảm bảo hiệu suất xử lý thời gian thực hoặc gần thời gian thực trên hệ thống đa camera, thiết bị IoT.

- Bảo mật dữ liệu, mô hình AI; phân quyền truy cập, cấu hình, cập nhật thuật toán rõ ràng.

- Giao diện cấu hình, giám sát trạng thái mô hình AI trực quan, dễ thao tác cho cán bộ vận hành.

2. Phân tích thời gian thực

Chức năng Phân tích thời gian thực giúp hệ thống phần mềm trí tuệ nhân tạo AI nâng cao khả năng phát hiện, nhận diện và xử lý các sự kiện, hành vi bất thường trên dữ liệu hình ảnh, video từ hệ thống camera giám sát an ninh. Qua đó, hỗ trợ các Trung tâm điều hành và lực lượng chức năng can thiệp kịp thời, giảm thiểu rủi ro và đảm bảo an toàn.

Một số chức năng cho tính năng Phân tích thời gian thực bao gồm:

a. Xử lý hình ảnh, video thời gian thực

- Tiếp nhận, phân tích dữ liệu hình ảnh, video từ hệ thống camera, thiết bị IoT với độ trễ xử lý nhỏ hơn 2 giây so với thời gian thực tế.

- Ứng dụng các thuật toán AI tối ưu để nhận diện đối tượng, phương tiện, hành vi bất thường, hành vi vi phạm, sự cố an ninh ngay khi phát sinh.

- Đảm bảo hệ thống có khả năng xử lý song song nhiều luồng dữ liệu từ nhiều camera/điểm giám sát khác nhau mà vẫn duy trì độ trễ nhỏ.

b. Cảnh báo và hiển thị tức thời

- Tổng quan (Dashboard) hiển thị thông báo thông tin: loại sự kiện, vị trí, thời gian phát hiện, hình ảnh/video sự kiện, mức độ cảnh báo.

- Cảnh báo có thể kích hoạt các hành động tự động (bật còi báo động, gửi lệnh đến thiết bị IoT…).

- Lưu trữ nhật ký các cảnh báo, sự kiện theo thời gian thực phục vụ tra cứu, báo cáo và đánh giá hiệu quả xử lý.

c. Hỗ trợ can thiệp nhanh

- Hệ thống cung cấp giao diện thao tác nhanh cho cán bộ vận hành như xác nhận, phân loại, chuyển tiếp cảnh báo đến đơn vị xử lý, trích xuất hình ảnh/video minh chứng.

d. Một số tính năng khác

- Đảm bảo khả năng mở rộng, nâng cấp khi số lượng camera, thiết bị tăng lên mà vẫn giữ được hiệu suất và độ trễ tối ưu.

- Đáp ứng các tiêu chuẩn bảo mật, an toàn dữ liệu trong quá trình truyền tải, xử lý và lưu trữ cảnh báo.

3. Tích hợp đa camera, liên kết xuyên camera

Chức năng tích hợp đa camera, liên kết xuyên camera giúp hệ thống phần mềm trí tuệ nhân tạo AI mở rộng khả năng giám sát an ninh, trật tự phạm vi rộng lớn, đa dạng nguồn thiết bị, hỗ trợ nhận diện, theo dõi đối tượng liên tục bất kể thay đổi vị trí hay điều kiện môi trường. Đáp ứng nhu cầu phối hợp, quản lý tập trung tại các Trung tâm điều hành và liên thông với Cơ sở dữ liệu dùng chung cho toàn thành phố hoặc địa phương.

Một số tính năng cho chức năng tích hợp đa camera, liên kết xuyên camera:

a. Hỗ trợ đa dạng hệ thống camera

- Tích hợp, kết nối với số lượng lớn camera đa dạng chủng loại (camera IP, analog, AI camera, camera PTZ, v.v.).

- Hỗ trợ nhiều chuẩn giao tiếp, kết nối (ONVIF, RTSP, HTTP, SDK riêng, v.v.), đảm bảo khả năng mở rộng, không phụ thuộc thương hiệu.

- Quản lý tập trung toàn bộ danh sách camera: vị trí, trạng thái, chủng loại, phiên bản, cấu hình, lịch sử hoạt động.

- Hỗ trợ cập nhật, mở rộng số lượng camera mà không ảnh hưởng tới hiệu suất toàn hệ thống.

b. Liên kết xuyên camera (cross-camera ReID)

- Tích hợp công nghệ ReID (Person/Vehicle Re-identification) cho phép nhận diện và theo dõi một đối tượng (người/phương tiện) xuyên suốt trên nhiều camera khác nhau trong hệ thống.

- Đối tượng có thể được phát hiện, nhận diện dù thay đổi vị trí, góc quay, điều kiện ánh sáng, trạng thái che khuất một phần hoặc thay đổi ngoại hình (quần áo, phụ kiện).

- Hỗ trợ liên kết dữ liệu nhận diện (ID, thời gian, vị trí, camera phát hiện) để tái tạo lộ trình di chuyển của đối tượng trên bản đồ số.

- Ghi nhận và lưu trữ lịch sử di chuyển, tần suất xuất hiện, các điểm ra/vào của từng đối tượng phục vụ điều tra, giám sát chuyên sâu.

c. Tích hợp, hiển thị và cảnh báo tập trung

- Cho phép hiển thị trạng thái, hình ảnh đối tượng được nhận diện xuyên camera trên giao diện quản lý tập trung tại Trung tâm điều hành.

- Phát hiện sự kiện bất thường, cảnh báo tự động khi đối tượng xuất hiện tại vị trí nhạy cảm hoặc di chuyển theo lộ trình bất thường.

- Tích hợp tính năng tra cứu, trích xuất nhanh hành trình di chuyển, thời điểm xuất hiện và các thông tin liên quan của đối tượng trên toàn hệ thống camera.

d. Một số tính năng khác

- Đảm bảo khả năng mở rộng, chịu tải lớn khi tích hợp hàng nghìn camera và xử lý đồng thời nhiều luồng dữ liệu.

- Đáp ứng các tiêu chuẩn bảo mật, phân quyền truy cập, đảm bảo an toàn thông tin xuyên suốt quá trình liên kết và chia sẻ dữ liệu.

- Giao diện trực quan, hỗ trợ thao tác lọc, tìm kiếm, theo dõi đối tượng linh hoạt.

4. Trích xuất, tổng hợp, phân loại sự kiện

Chức năng Trích xuất, tổng hợp, phân loại sự kiện giúp hệ thống trí tuệ nhân tạo AI chủ động ghi nhận, phân tích, phân loại các sự kiện phát sinh từ nhiều luồng camera giám sát an ninh. Qua đó, hỗ trợ Trung tâm điều hành và lực lượng

chức năng tổng hợp, truy xuất, khai thác dữ liệu sự kiện phục vụ điều tra, xử lý, báo cáo và ra quyết định điều hành kịp thời.

Một số tính năng cho chức năng Trích xuất, tổng hợp, phân loại sự kiện bao gồm:

a. Tự động ghi nhận, phân tích và phân loại sự kiện

- Hệ thống AI tự động thu nhận dữ liệu hình ảnh, video từ nhiều luồng camera/thiết bị IoT, nhận diện và ghi nhận các sự kiện phát sinh trên toàn hệ thống.

- Phân tích, xác định loại sự kiện: an ninh (đánh nhau, xâm nhập trái phép, mang hung khí...), trật tự (tụ tập đông người, gây rối...), giao thông (vi phạm, ùn tắc, dừng đỗ sai quy định...), cháy nổ, sự cố môi trường, v.v.

- Tự động phân loại sự kiện theo các tiêu chí: loại sự kiện, mức độ nguy cơ (nghiêm trọng, cảnh báo, thông báo...), khu vực, thời gian, nguồn phát hiện (camera, IoT).

b. Tổng hợp, thống kê và xuất dữ liệu sự kiện

- Tổng hợp, thống kê số lượng, tần suất, phân bố sự kiện theo khu vực địa lý, loại thiết bị, khung thời gian, mức độ nguy cơ…

- Hỗ trợ sắp xếp, lọc, truy vấn sự kiện linh hoạt theo nhiều tiêu chí phục vụ nghiệp vụ điều tra, xử lý.

- Cho phép xuất dữ liệu sự kiện dưới các định dạng phổ biến (Excel, PDF, CSV, hình ảnh…) hoặc chuyển tiếp qua API về các hệ thống liên quan.

- Lưu trữ đầy đủ nhật ký, hình ảnh/video minh chứng, thông tin chi tiết từng sự kiện (thời điểm, vị trí, camera ghi nhận, trạng thái xử lý…).

c. Hỗ trợ điều tra, xử lý và báo cáo

- Kết nối dữ liệu sự kiện với các module hỗ trợ điều tra (truy vết đối tượng, hành trình di chuyển, kiểm tra lịch sử sự kiện…).

- Cung cấp giao diện tổng hợp sự kiện trên Tổng quan: hiển thị bản đồ, biểu đồ, báo cáo phân tích, cảnh báo nhanh cho cán bộ vận hành.

- Hỗ trợ đánh dấu, phân loại trạng thái các sự kiện (chưa xử lý, đang xử lý, đã xử lý, xác minh sai…) và lưu vết các thao tác liên quan.

d. Một số tính năng khác

- Đảm bảo hiệu suất xử lý dữ liệu lớn, nhiều luồng camera, nhiều loại sự kiện cùng lúc.

- Đáp ứng tiêu chuẩn bảo mật, phân quyền truy cập, khai thác dữ liệu sự kiện.

5. Khả năng mở rộng

Chức năng Khả năng mở rộng giúp hệ thống phần mềm trí tuệ nhân tạo AI đáp ứng nhu cầu phát triển lâu dài, dễ dàng bổ sung số lượng lớn camera, mở rộng vùng giám sát, tăng số lượng người dùng và các phân hệ chức năng mà không ảnh hưởng đến hiệu suất, độ ổn định và bảo mật của toàn hệ thống. Đảm bảo hệ thống luôn sẵn sàng thích ứng với quy mô đô thị hiện đại, các dự án mở rộng trong tương lai và các Trung tâm điều hành quy mô lớn.

Một số tính năng chức năng Khả năng mở rộng bao gồm:

a. Thiết kế module, phân lớp rõ ràng

- Hệ thống được xây dựng trên kiến trúc module, mỗi thành phần (quản lý camera, AI nhận diện, lưu trữ dữ liệu, dashboard, quản lý người dùng, báo cáo…) hoạt động độc lập, dễ dàng thêm mới hoặc mở rộng mà không ảnh hưởng tới các module khác.

- Hỗ trợ tích hợp module mới bằng giao diện lập trình ứng dụng mở (API), cho phép kết nối các dịch vụ, thiết bị, giải pháp AI của bên thứ ba một cách linh hoạt.

b. Mở rộng số lượng camera, vùng giám sát

- Cho phép bổ sung số lượng lớn camera giám sát, thiết bị IoT ở nhiều khu vực khác nhau mà không cần thay đổi hay dừng hệ thống đang vận hành.

- Hệ thống tự động phân phối tải, cân bằng xử lý giữa các server, cụm AI, đảm bảo không xảy ra hiện tượng nghẽn cổ chai khi quy mô tăng nhanh.

- Hỗ trợ quản lý, cấu hình, giám sát trạng thái camera mới từ giao diện tập trung mà không cần thao tác kỹ thuật phức tạp.

c. Mở rộng số lượng người dùng, nhóm vận hành

- Tăng số lượng người dùng, nhóm người dùng, phân quyền chi tiết mà không ảnh hưởng hiệu suất hoặc bảo mật hệ thống.

- Hỗ trợ đồng thời nhiều người dùng truy cập, thao tác, khai thác dữ liệu trên toàn hệ thống mà không xảy ra xung đột hay giảm hiệu suất.

- Cho phép tích hợp xác thực tập trung (Single Sign-On, LDAP/AD…) để mở rộng quy mô người dùng dễ dàng.

d. Tối ưu hiệu suất, đảm bảo ổn định khi mở rộng

- Ứng dụng các giải pháp công nghệ hiện đại như ảo hóa, container hóa (Docker, Kubernetes...), phân tán dữ liệu, load balancing để đảm bảo hệ thống hoạt động ổn định khi mở rộng quy mô.

- Hỗ trợ giám sát hiệu suất, cảnh báo sớm khi có dấu hiệu quá tải, tự động chuyển đổi tài nguyên phù hợp với nhu cầu thực tế.

e. Mở rộng tích hợp với các hệ thống, cơ sở dữ liệu ngoài

- Cho phép kết nối, chia sẻ, đồng bộ dữ liệu với các hệ thống, cơ sở dữ liệu dùng chung, các Trung tâm điều hành, hệ thống phân tích, báo cáo của các đơn vị khác nhau mà không giới hạn về quy mô, số lượng kết nối.

PHẦN 2. TIÊU CHÍ KỸ THUẬT

I.Mô hình hệ thống đề xuất

Mô hình hệ thống được thiết kế theo kiến trúc **module – phân tán – tập trung**, bảo đảm linh hoạt khi triển khai theo yêu cầu thiết kế riêng của từng khách hàng, từ đó tạo ra một nền tảng giám sát – điều hành toàn diện, liền mạch và hiệu quả.

II. Tiêu chuẩn công nghệ và tiêu chuẩn tích hợp đáp ứng

1. Tiêu chuẩn về công nghệ sử dụng (Techstack)

Hệ thống sử dụng kiến trúc hiện đại, kết hợp các công nghệ mới nhất trong lĩnh vực trí tuệ nhân tạo (AI), xử lý dữ liệu lớn (Big Data), giám sát tập trung, và triển khai phân tán, nhằm đảm bảo khả năng xử lý mạnh mẽ tại biên, đồng thời quản trị linh hoạt từ trung tâm.

Thành phần	Công nghệ
Xử lý trí tuệ nhân tạo (AI)	Python, C++, TensorRT – phục vụ suy luận mô hình AI với hiệu suất cao tại Thiết bị xử lý tại biên AI Box
Hệ thống cổng quản trị (Web Portal)	ASP.NET Core (.NET 8), Angular – đảm bảo hiệu năng và giao diện hiện đại, dễ sử dụng
Cơ sở dữ liệu cục bộ tại Thiết bị xử lý tại biên AI Box	SQLite – gọn nhẹ, phù hợp cho thiết bị nhúng và xử lý ngoại tuyến
Cơ sở dữ liệu trung tâm	MySQL hoặc NoSQL – linh hoạt trong xử lý nghiệp vụ truyền thống và dữ liệu phi cấu trúc
Hệ thống lưu trữ tệp	MinIO – triển khai dạng phân tán, tương thích chuẩn S3, hỗ trợ mở rộng linh hoạt
Hàng đợi xử lý bất đồng bộ	Kafka – đảm bảo truyền dữ liệu sự kiện thời gian thực và khả năng mở rộng cao

Ghi log hệ thống	Serilog → Kafka → Elasticsearch – ghi nhận và phân tích log tập trung
Giao diện phân tích dữ liệu log	Kibana – trực quan hóa log, truy vết và phân tích lỗi hiệu quả
Giám sát thiết bị và tài nguyên	Node Exporter, PushGateway – theo dõi tình trạng Thiết bị xử lý tại biên AI Box và thiết bị hệ thống
Cảnh báo và giám sát tài nguyên trực quan	Grafana – thiết lập biểu đồ, cảnh báo tự động khi có bất thường
Quản lý phiên bản phần mềm	Docker Image kết hợp giao diện Web quản trị – kiểm soát và triển khai phiên bản từ xa
Hạ tầng triển khai	Docker, Kubernetes – triển khai dịch vụ dạng container, hỗ trợ High Availability
Cập nhật phần mềm Thiết bị xử lý tại biên AI Box từ xa	Web Server gửi lệnh cập nhật qua Kafka, có khả năng rollback khi cần thiết
Kết nối SSH từ Portal tới Thiết bị xử lý tại biên AI Box	Cơ chế reverse SSH – không cần IP tĩnh tại Thiết bị xử lý tại biên AI Box, đảm bảo bảo mật và kết nối thuận tiện

2. Tiêu chuẩn về hạ tầng thiết bị

2.1. Thiết bị Camera

Các thiết bị Camera sử dụng trong hệ thống đáp ứng các tiêu chí sau:

- Hỗ trợ ONVIF Profile S/G/T, RJ45 PoE.
- Bảo mật nâng cao: HTTPS/TLS, secure boot, bảo vệ firmware.
- Tương thích với nhiều VMS: Milestone, NX Witness, i-PRO Active Guard, Qognify, Luxriot…
- Hỗ trợ ghi hình edge recording, metadata AI.
- Lưu trữ dữ liệu:

- Lưu trên NVR/Server tại trung tâm điều hành, tối thiểu 30–60 ngày.
- Sao lưu định kỳ về trung tâm dữ liệu dự phòng hoặc Cloud an toàn.
- Dữ liệu được mã hóa, phân quyền chặt chẽ; chỉ cán bộ được ủy quyền mới có thể truy cập.
- Sự kiện quan trọng có thể đẩy lên hệ thống cấp trên phục vụ điều hành – điều tra.

2.2. Thiết bị xử lý tại biên AI Box

Thiết bị xử lý tại biên AI Box đóng vai trò trung tâm trong việc xử lý dữ liệu hình ảnh/video tại biên, trước khi truyền tải về hệ thống trung tâm, nhằm giảm tải đường truyền và phản hồi sự kiện nhanh chóng.

- Khả năng xử lý AI tại biên: Tích hợp mô hình AI, cho phép nhận diện, phân tích hành vi, cảnh báo tại chỗ
- Khả năng hoạt động ổn định: Vận hành 24/7 tại các điểm lắp đặt
- Giao tiếp mạng: Ethernet 10/100/1000 Mbps, hỗ trợ WiFi tùy chọn
- Kết nối thiết bị ngoại vi: GPIO, USB, RS485/RS232 nếu cần
- Nguồn tiêu thụ: sử dụng điện thông thường không cần đầu tư hệ thống lưới điện riêng đảm bảo các tiêu chuẩn tiết kiệm điện năng giảm hiệu ứng nhà kính, khí thải C02
- Hệ điều hành: Ubuntu 18.04 hoặc 20.04 64-bit, hoặc JetPack (với thiết bị sử dụng NVIDIA)
- Hỗ trợ mô hình AI: Tương thích TensorRT, OpenVINO, ONNX
- Triển khai phần mềm: Qua Docker, cho phép cập nhật tự động hoặc từ xa
- Khả năng hoạt động offline: Lưu dữ liệu tạm thời, đồng bộ lại khi có kết nối
- Bảo trì thuận tiện: Hỗ trợ reverse SSH từ hệ thống trung tâm mà không cần IP tĩnh
- Dễ lắp ráp, thay thế: Phù hợp triển khai hàng loạt với chi phí tối ưu

2.3. Hệ thống Server (máy chủ trung tâm)

- Đảm bảo tính ổn định và khả năng thu thập dữ liệu từ Thiết bị xử lý tại biên AI Box trên toàn quốc
- Lưu trữ video ngắn (video ngắn làm bằng chứng vi phạm), ảnh, log, dữ liệu AI
- Cung cấp giao diện quản trị, cảnh báo, báo cáo về thiết bị trên toàn hệ thống
- Hỗ trợ mở rộng theo chiều ngang (horizontal scaling)

3. Tiêu chuẩn về nền tảng triển khai và Phần mềm quản lý

3.1. Hạ tầng triển khai:

a. Server:

- Toàn bộ các thành phần dịch vụ trung tâm (API, Kafka, MinIO, Elasticsearch, Prometheus, Grafana...) được triển khai dưới dạng container và quản lý bằng nền tảng Docker hoặc Kubernetes (K8s).
- Hỗ trợ triển khai trên các hạ tầng Docker/K8s.
- Các ứng dụng backend viết bằng .NET 8 (ASP.NET Core), ứng dụng AI viết bằng Python/C++, giao tiếp qua API nội bộ hoặc Kafka.

b. Thiết bị xử lý tại biên Thiết bị xử lý tại biên AI Box:

- Hệ điều hành Ubuntu hoặc JetPack.
- Cài đặt ứng dụng bằng Docker.
- Các thành phần chạy trên Thiết bị xử lý tại biên AI Box gồm: AI module, API nội bộ (.NET), Kafka consumer, service đồng bộ dữ liệu.
- Quản lý cập nhật phần mềm qua Docker Image được đồng bộ từ Server.

3.2. Quản lý phiên bản và phân phối phần mềm

- Server tích hợp Private Docker Registry:
    - Là nơi lưu trữ và phân phối các Docker Image ứng dụng cho các Thiết bị xử lý tại biên AI Box.
    - Cho phép cập nhật phiên bản ứng dụng, rollback khi cần thiết.
- Giao diện Web hỗ trợ:

- Quản lý phiên bản phần mềm từng Thiết bị xử lý tại biên AI Box.
- Triển khai (install), cập nhật (update), và quay lại phiên bản trước (rollback) cho từng ứng dụng theo Thiết bị xử lý tại biên AI Box.
- Kiểm soát và theo dõi trạng thái triển khai phiên bản.

3.3. Lưu trữ dữ liệu

- Sử dụng MinIO để lưu trữ:
  - Ảnh đầu vào từ camera.
  - Video ngắn hành vi bất thường.
  - Kết quả AI (ảnh crop khuôn mặt, ảnh phương tiện vi phạm…).
  - MinIO triển khai dạng Distributed mode đảm bảo tính sẵn sàng cao (HA) và khả năng mở rộng.
  - Tương thích S3 API, hỗ trợ tích hợp với các hệ thống phân tích dữ liệu, AI training, hoặc backup cloud.

3.4. Giám sát hệ thống

- Sử dụng Prometheus để thu thập và lưu trữ các metrics từ:
  - Thiết bị Thiết bị xử lý tại biên AI Box (thông qua PushGateway hoặc Node Exporter).
  - Container (sử dụng cAdvisor hoặc Exporter riêng).
  - Dịch vụ Server (API, Kafka, Elasticsearch, MinIO...).
- Sử dụng Grafana để:
  - Trực quan hóa dữ liệu hệ thống.
  - Thiết lập cảnh báo (Alert) khi có sự cố (CPU cao, nhiệt độ cao, Box mất kết nối, lỗi AI, Disk đầy…).
  - Phân nhóm theo Thiết bị xử lý tại biên AI Box, theo vị trí triển khai.

3.5. Ghi log và giám sát hoạt động

- Hệ thống ghi log tập trung sử dụng:
  - Serilog (trên .NET) ghi log theo định dạng JSON.

    - Kafka làm message bus trung gian để gom log từ các Thiết bị xử lý tại biên AI Box.
    - Consumer phía Server xử lý và đẩy log vào Elasticsearch.
- Kibana dùng để:
    - Phân tích log.
    - Lọc log theo thời gian, theo Thiết bị xử lý tại biên AI Box, theo ứng dụng.
    - Cảnh báo khi có lỗi hoặc hành vi bất thường.

3.6. Web quản lý tập trung

- Hệ thống phải cung cấp một giao diện Web trực quan cho phép:
    - Quản lý danh sách các Thiết bị xử lý tại biên AI Box, camera kết nối.
    - Theo dõi phiên bản phần mềm trên mỗi Thiết bị xử lý tại biên AI Box.
    - Theo dõi trạng thái hoạt động, log lỗi, cảnh báo realtime.
    - Quản lý thông tin cảnh báo AI, ảnh sự kiện, các báo cáo định kỳ.
    - Quản lý danh sách model AI, phiên bản AI đang chạy ở các Thiết bị xử lý tại biên AI Box.

4. Tiêu chuẩn về an ninh – bảo mật

Các tiêu chuẩn an ninh – bảo mật bao gồm:

- Tuân thủ quy định của Bộ Công an và pháp luật liên quan:
    - Hệ thống phải phù hợp tiêu chuẩn kỹ thuật về an ninh vật lý, an ninh mạng và các yêu cầu về bảo vệ dữ liệu.
    - Cam kết tuân thủ đầy đủ các hướng dẫn, quy phạm trong triển khai hệ thống giám sát và thiết bị an ninh.
- Chuẩn an ninh, bảo mật thiết bị:
    - Camera, Thiết bị xử lý tại biên AI Box, Edge Server và thiết bị mạng phải có nguồn gốc, xuất xứ rõ ràng, được chứng nhận CE, FCC hoặc tương đương.

    - Phần mềm quản lý phải hỗ trợ mã hóa dữ liệu (HTTPS/TLS), xác thực người dùng nhiều lớp, cơ chế chống tấn công mạng, và bảo vệ firmware theo yêu cầu của ngành Công an.
- Bảo mật dữ liệu theo quy định về bí mật nhà nước:
    - Dữ liệu nhận diện cá nhân, dữ liệu hình ảnh, video, nhật ký truy cập… phải được bảo mật nghiêm ngặt theo Thông tư 24/2020/TT-BCA và các văn bản quy định về quản lý bí mật nhà nước.
    - Việc lưu trữ, khai thác và chia sẻ dữ liệu phải đảm bảo đúng thẩm quyền, đúng quy trình và có ghi nhật ký truy xuất.
- Tiêu chí đánh giá "An toàn về ANTT" theo Thông tư 124/2021/TT-BCA:
    - Trụ sở phải có hệ thống camera giám sát đầy đủ các khu vực trọng yếu, hệ thống cảnh báo – báo động và kiểm soát ra vào đáp ứng tiêu chuẩn kỹ thuật.
    - Giải pháp đầu tư đáp ứng cả tiêu chí kỹ thuật, quy chế vận hành, và phân công trách nhiệm giữa Công an và chính quyền sở tại.
    - Hệ thống hoạt động ổn định 24/7, đảm bảo khả năng ghi hình và hỗ trợ xử lý tình huống theo yêu cầu công tác.
- Tiêu chuẩn tích hợp và tương thích hệ thống:
    - Thiết bị phải hỗ trợ chuẩn ONVIF Profile S/G/T, API mở, đảm bảo tương thích với hệ thống phần mềm nghiệp vụ, hệ thống nhận diện quốc gia và các dự án trọng điểm như Đề án 06.
    - Mọi dữ liệu trao đổi giữa hệ thống phường/xã và cấp huyện/tỉnh đều được thực hiện qua mạng truyền số liệu chuyên dùng của Bộ Công an, bảo đảm độ an toàn cao.
- Bảo vệ hệ thống và bảo đảm vận hành liên tục:
    - Máy chủ, thiết bị mạng, UPS và tủ rack phải đáp ứng tiêu chuẩn an toàn điện, an toàn cháy nổ và đảm bảo tính liên tục trong trường hợp mất điện hoặc sự cố bất thường.

- Hệ thống phải có biện pháp dự phòng (backup), RAID, sao lưu định kỳ, mã hóa khi lưu và khi truyền tải.

5. Tiêu chuẩn về vận hành – bảo trì

5.1. Đảm bảo tính ổn định và liên tục

- Hạ tầng mạng LAN an ninh riêng biệt, sử dụng VLAN và QoS để bảo đảm lưu lượng Camera luôn ổn định.
- Sử dụng UPS và nguồn dự phòng để duy trì hoạt động 1–2 giờ khi mất điện; hỗ trợ máy phát điện cho những sự cố kéo dài.
- Hệ thống phải có cảnh báo tự động khi thiết bị gặp lỗi: Camera ngoại tuyến, hết dung lượng lưu trữ, lỗi AI, mất nguồn PoE.
- Quy trình vận hành đầu ca được chuẩn hóa: kiểm tra đèn Camera, khóa cửa, trạng thái thiết bị, chất lượng hình ảnh…
- Lực lượng trực ban được đào tạo sử dụng hệ thống để thao tác báo động, phân tích tình huống và yêu cầu ứng cứu kịp thời.

5.2. Khả năng mở rộng quy mô

- Có khả năng bổ sung camera, Thiết bị xử lý tại biên AI Box, đầu đọc ra/vào, thiết bị IoT mà không ảnh hưởng hệ thống hiện hữu.
- Edge Server và hệ thống lưu trữ hỗ trợ tăng dung lượng, nâng cấp CPU/RAM, hoặc mở rộng theo dạng module.
- Hỗ trợ chuẩn ONVIF S/G/T và API mở để dễ dàng tích hợp với các hệ thống khác như:
    - CSDL quốc gia về dân cư
    - Hệ thống kiểm soát ra vào
    - Đề án 06
    - Trung tâm chỉ huy Công an cấp huyện/tỉnh
- Quy trình nâng cấp hàng năm: đánh giá hiệu quả hệ thống, xác định điểm nghẽn và đầu tư bổ sung (thay camera cũ, nâng AI, mở rộng lưu trữ…).

5.3. Đảm bảo hiệu suất và độ trễ thấp

- Mạng Gigabit Ethernet, switch PoE chuyên dụng, đảm bảo luồng camera 4MP–4K (4–10 Mbps/camera) không gây nghẽn mạng.
- Thiết bị xử lý tại biên AI Box xử lý ở tầng biên (edge AI), giảm độ trễ và giảm tải cho máy chủ trung tâm.
- Edge Server phải đủ cấu hình để:
    - Xử lý phân tích AI
    - Lưu trữ cục bộ
    - Truy xuất video
    - Xuất báo cáo mà không ảnh hưởng tốc độ xem trực tiếp
- Kết nối giữa camera – Thiết bị xử lý tại biên AI Box – Edge Server – Storage phải duy trì độ trễ thấp và ổn định.

5.4. Cơ chế High Availability (HA)

- Các dịch vụ quan trọng (AI phân tích, quản lý hệ thống, NVR/VMS, điều khiển báo động) phải hỗ trợ cơ chế dự phòng nóng/lạnh.
- Edge Server có thể triển khai cluster HA hoặc sử dụng máy chủ dự phòng.
- Khi luồng camera lỗi, hệ thống tự động chuyển sang đường truyền dự phòng hoặc node dự phòng.
- Cơ chế nguồn kép: UPS + nguồn điện dự phòng giúp duy trì tính liên tục.

5.5. Cơ chế backup và phục hồi

- Lưu trữ video và metadata AI trên RAID 5/6, tránh mất dữ liệu khi hỏng ổ cứng.
- Sao lưu định kỳ (ngày/tuần/tháng) lên:
    - Cloud Storage (lưu trữ tập trung)
    - Trung tâm dữ liệu cấp tỉnh qua mạng chuyên dùng Công an
- Hỗ trợ phục hồi nhanh cấu hình, video, nhật ký truy cập khi có sự cố.
- Ghi log đầy đủ mọi hoạt động:
    - Xem video
    - Mở cửa

- Chỉnh sửa cấu hình
- Thêm/xóa tài khoản nhằm phục vụ kiểm tra và điều tra nghiệp vụ.

5.6. Cập nhật và quản lý phiên bản ứng dụng

- Camera, Thiết bị xử lý tại biên AI Box và Edge Server phải hỗ trợ cập nhật firmware có chữ ký số, chống giả mạo (secure boot, firmware verification).
- Việc cập nhật được lập lịch trong khung thời gian bảo trì, có phê duyệt của quản trị viên.
- Hệ thống quản lý phiên bản tập trung giúp:
  - Theo dõi phiên bản đang chạy
  - Nhận cảnh báo khi phiên bản lỗi thời
  - Nâng cấp từng thiết bị hoặc nâng cấp đồng loạt
- Theo Thông tư 35/2025/TT-BCA, việc bàn giao – nghiệm thu – bảo hành – bảo trì phải được ghi chép, công khai, minh bạch.
- Bảo trì định kỳ hàng quý: vệ sinh camera, kiểm tra kết nối, cập nhật firmware, backup dữ liệu.

6. Tiêu chuẩn về hiệu suất

- Hệ thống phải đảm bảo hiệu suất xử lý gần thời gian thực (near real-time) để phát hiện và cảnh báo các hành vi bất thường, sự cố, vi phạm một cách kịp thời và chính xác.
- Xử lý AI tại biên: giảm tải băng thông, chỉ gửi dữ liệu cần thiết.
- Buffer nội bộ: xử lý tình huống mất mạng tạm thời.
- Tối ưu GPU sử dụng TensorRT/OpenVINO để tăng tốc độ suy luận.
- Theo dõi hiệu suất Thiết bị xử lý tại biên AI Box theo thời gian thực, cảnh báo nếu CPU/GPU/RAM vượt ngưỡng.

7. Tiêu chuẩn về khả năng tích hợp

7.1. Tích hợp với các hệ thống quản trị và điều hành hiện có

- Tích hợp với hệ thống điều hành, phần mềm quản lý camera, hệ thống kiểm soát ra/vào, hệ thống báo động, Cổng dịch vụ công và các giải pháp nghiệp vụ hiện có.
- Cho phép chia sẻ dữ liệu video, nhật ký sự kiện, cảnh báo AI về Trung tâm điều hành của huyện/tỉnh thông qua mạng truyền số liệu chuyên dùng Công an.
- Tương thích với các giải pháp VMS phổ biến (Milestone, NX Witness, i-PRO, Qognify…), hỗ trợ chuẩn ONVIF để đảm bảo khả năng liên thông thiết bị.
- Cho phép API kết nối đến các hệ thống giám sát khác tại địa bàn như camera đô thị, camera giao thông, hệ thống kiểm soát bãi xe…

7.2. Tích hợp với cơ sở dữ liệu quốc gia

- Hỗ trợ kết nối an toàn với CSDL Quốc gia về dân cư, giải pháp định danh điện tử, và các thành phần của Đề án 06/CP.
- Cho phép đối soát dữ liệu nhận diện khuôn mặt, dữ liệu biển số phương tiện với kho dữ liệu tập trung cấp tỉnh/thành khi có sự vụ cần xác minh.
- Hỗ trợ chia sẻ sự kiện quan trọng lên Trung tâm chỉ huy 113 hoặc Trung tâm điều hành cấp tỉnh theo quy trình nghiệp vụ.
- Mọi hoạt động tích hợp phải tuân thủ quy định về bảo mật nhà nước, tiêu chuẩn mã hóa và kiểm soát truy cập của ngành Công an.

7.3. Yêu cầu về kiến trúc mở

- Thiết kế theo kiến trúc mở (Open Architecture), hỗ trợ ONVIF, RESTful API, Webhook, MQTT hoặc các giao thức tích hợp mở rộng.
- Cho phép bổ sung camera, Thiết bị xử lý tại biên AI Box, Edge Server, Storage, thiết bị IoT hoặc tích hợp thêm các thuật toán AI mới mà không cần thay đổi kiến trúc tổng thể.
- Hỗ trợ khả năng mở rộng theo chiều ngang (scale-out) và chiều dọc (scale-up) tại cả cấp phường/xã và trung tâm tỉnh/thành.

- Tương thích với hệ thống hạ tầng ảo hóa, container hóa hoặc microservices nếu triển khai về sau.

7.4. Cơ chế xác thực và bảo mật tích hợp

- Xác thực người dùng bằng mật khẩu mạnh, xác thực hai lớp (2FA), phân quyền theo vai trò (RBAC).
- Mọi giao tiếp giữa hệ thống phường/xã và trung tâm phải được mã hóa đầu cuối (TLS 1.2/1.3, HTTPS).
- Hỗ trợ chữ ký số, token xác thực (JWT/OAuth2), secure API key, chống giả mạo truy cập.
- Tất cả hoạt động xem video, tải dữ liệu, cấu hình thiết bị… phải được ghi log đầy đủ, đảm bảo tính minh bạch và truy vết khi cần điều tra.
- Thực hiện theo đúng quy định về bảo mật của Bộ Công an: Thông tư 24/2020/TT-BCA (bảo mật dữ liệu), Thông tư 35/2025/TT-BCA (quản lý thiết bị – minh bạch mua sắm), các quy định về an ninh mạng và an toàn hệ thống.