

	VIETTEL AI RACE	TD068
	Thu thập và tiền xử lý dữ liệu hình ảnh	Lần ban hành: 1

1. Thu thập và Tiền xử lý Dữ liệu Hình ảnh

Để một hệ thống thị giác máy tính hoạt động hiệu quả, dữ liệu hình ảnh chất lượng cao là nền tảng quan trọng nhất. Việc thu thập dữ liệu không chỉ đơn thuần là “chụp ảnh”, mà đòi hỏi một quy trình bài bản:

- Nguồn dữ liệu đa dạng: Kết hợp ảnh từ camera dây chuyền sản xuất, dữ liệu cảm biến quang học, và các bộ dữ liệu công khai như ImageNet hoặc Open Images.
- Tiêu chuẩn hóa bối cảnh: Đảm bảo nhiều góc chụp, ánh sáng khác nhau, và các điều kiện sản xuất khác nhau để mô hình có khả năng tổng quát tốt.
- Định dạng thông nhất: Giữ định dạng hình ảnh đồng nhất (JPEG/PNG) và độ phân giải tối thiểu Full HD nhằm duy trì chi tiết cần thiết cho phân tích.

Sau khi thu thập, dữ liệu phải trải qua tiền xử lý để tăng tính ổn định:

- Cân bằng dữ liệu: Dùng kỹ thuật oversampling hoặc augmentation như xoay, lật, điều chỉnh độ sáng.
- Loại bỏ nhiễu: Ứng dụng các bộ lọc như Gaussian blur hoặc median filter để giảm nhiễu cảm biến.
- Chuẩn hóa kích thước: Đưa tất cả ảnh về cùng kích thước đầu vào của mô hình, thường là 224×224 hoặc 512×512 pixel.

Quy trình này đảm bảo dữ liệu “sạch”, đa dạng và nhất quán, là nền tảng để huấn luyện các mô hình thị giác máy tính công nghiệp đạt hiệu quả cao.

	VIETTEL AI RACE	TD068
	Thu thập và tiền xử lý dữ liệu hình ảnh	Lần ban hành: 1

2. Xây dựng Kiến trúc Mô hình

Sau khi có dữ liệu chất lượng, bước tiếp theo là thiết kế và lựa chọn kiến trúc mô hình phù hợp. Việc chọn kiến trúc dựa trên mục tiêu cụ thể:

- Phân loại hình ảnh: Các kiến trúc CNN như ResNet, EfficientNet, DenseNet nổi bật nhờ khả năng trích xuất đặc trưng mạnh mẽ.
- Phát hiện vật thể: Các mô hình YOLO (phiên bản mới nhất YOLOv8) hay Faster R-CNN được sử dụng để phát hiện lỗi sản phẩm trên dây chuyền.
- Phân đoạn ảnh: U-Net, DeepLab, hay Mask R-CNN phù hợp cho việc khoanh vùng chính xác các vùng khuyết tật.
- Xu hướng mới: Vision Transformer (ViT) và các mô hình hybrid CNN–Transformer giúp mô hình hiểu mối liên hệ xa trong ảnh, cải thiện hiệu năng trong môi trường phức tạp.

Chiến lược huấn luyện:

- Pretraining + Fine-tuning: Khởi đầu với mô hình đã huấn luyện trên ImageNet, sau đó tinh chỉnh trên dữ liệu nội bộ.
- Regularization: Dùng dropout, weight decay để giảm overfitting.
- Augmentation online: Tạo biến thể ảnh trong suốt quá trình huấn luyện để tăng khả năng tổng quát.

Quá trình này giúp mô hình đạt độ chính xác cao trong thời gian huấn luyện tối ưu, đồng thời giảm yêu cầu về tài nguyên.

	VIETTEL AI RACE	TD068
	Thu thập và tiền xử lý dữ liệu hình ảnh	Lần ban hành: 1

3. Tích hợp và Triển khai Thực Tế

Sau khi mô hình đạt hiệu suất mong muốn, bước triển khai vào môi trường công nghiệp đòi hỏi sự phối hợp chặt chẽ giữa phần mềm và phần cứng:

- Giao tiếp thời gian thực: Dịch vụ inference nên được triển khai dưới dạng API (REST/gRPC) để xử lý video trực tiếp từ camera.
- Đồng bộ với PLC/SCADA: Hệ thống cần tương thích với các bộ điều khiển công nghiệp để tự động loại bỏ sản phẩm lỗi hoặc điều chỉnh dây chuyền.
- Xử lý tại biên (Edge Computing): Đưa mô hình đã nén (quantization, pruning) lên các thiết bị như NVIDIA Jetson để giảm độ trễ và tiết kiệm băng thông.

Triển khai không chỉ là chạy mô hình, mà còn liên quan đến kiến trúc hệ thống:

- Tách biệt microservices cho từng nhiệm vụ (xử lý hình ảnh, lưu trữ, giao diện).
- Thiết kế dashboard trực quan cho phép kỹ sư giám sát luồng dữ liệu, trạng thái từng camera, và kết quả kiểm tra theo thời gian thực.

4. Giám sát, Cập nhật và Bảo trì

Môi trường sản xuất thay đổi liên tục, từ ánh sáng cho tới cấu hình máy móc, dẫn đến hiện tượng drift dữ liệu – khi phân phối dữ liệu thực tế khác so với dữ liệu huấn luyện ban đầu.

	VIETTEL AI RACE	TD068
	Thu thập và tiền xử lý dữ liệu hình ảnh	Lần ban hành: 1

- Theo dõi drift dữ liệu: Sử dụng công cụ như EvidentlyAI hoặc WhyLabs để cảnh báo khi thông kê dữ liệu đầu vào thay đổi.
- Huấn luyện định kỳ: Lập lịch huấn luyện lại (ví dụ hàng tháng) với dữ liệu mới để mô hình luôn phản ánh thực tế.
- A/B Testing: Triển khai song song hai phiên bản mô hình để so sánh hiệu suất trước khi cập nhật chính thức.

Về bảo mật:

- Mã hóa luồng video (TLS/SSL).
- Phân quyền người dùng và ghi log chi tiết để truy xuất quyết định của mô hình khi cần điều tra.

Việc giám sát và bảo trì liên tục đảm bảo hệ thống hoạt động ổn định, tránh giảm hiệu năng khi môi trường thay đổi.

5. Tối ưu Chi phí và Mở Rộng Hệ Thống

Để giải pháp thị giác máy tính bền vững và có thể mở rộng:

- Điện toán đám mây lai: Kết hợp cloud và edge giúp giảm chi phí băng thông nhưng vẫn đảm bảo khả năng xử lý khối lượng lớn khi cần.
- Mô hình kinh doanh linh hoạt: “Computer Vision as a Service” – tính phí dựa trên số lượng khung hình hoặc camera đang hoạt động.

	VIETTEL AI RACE	TD068
	Thu thập và tiền xử lý dữ liệu hình ảnh	Lần ban hành: 1

- Tự động hóa quản trị: Sử dụng hạ tầng như Kubernetes để tự động scale khi số lượng camera tăng, giảm thiểu can thiệp thủ công.
- Chuẩn hóa tài liệu kỹ thuật: Xây dựng bộ quy trình vận hành chuẩn (SOP) giúp việc nhân rộng sang nhiều nhà máy diễn ra nhanh chóng.

Một hệ thống được tối ưu chi phí và linh hoạt trong mở rộng sẽ dễ dàng thích ứng khi doanh nghiệp phát triển hoặc khi yêu cầu sản xuất tăng cao.