

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



Phạm Việt Anh

TÓM TẮT

ĐỒ ÁN THẠC SĨ KỸ THUẬT

HÀ NỘI - 2025

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



Phạm Việt Anh

**NGHIÊN CỨU MÔ HÌNH HỌC SÂU CHO NHẬN DIỆN ĐỒ BẢO HỘ
TRONG QUẢN LÝ TRANG TRẠI ĐIỆN GIÓ**

CHUYÊN NGÀNH : KHOA HỌC MÁY TÍNH

MÃ SỐ : 8.48.01.01

ĐỒ ÁN TỐT NGHIỆP THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

NGƯỜI HƯỚNG DẪN KHOA HỌC

PGS.TS. TRẦN QUANG ANH

HÀ NỘI – 2025

MỞ ĐẦU

Lý do chọn đề tài:

Ngành điện gió đang phát triển mạnh, nhưng môi trường làm việc tại các trang trại điện gió khắc nghiệt với độ cao lớn và thời tiết bất lợi, tiềm ẩn nhiều nguy cơ tai nạn lao động. Giám sát thủ công thiết bị bảo hộ cá nhân (PPE) hiện nay kém hiệu quả do dễ xảy ra sai sót và khó kiểm soát toàn diện. Đề tài ứng dụng trí tuệ nhân tạo và nhận diện hình ảnh để tự động hóa giám sát PPE, nhằm giảm tai nạn và nâng cao an toàn lao động. Sự khả thi được đảm bảo nhờ công nghệ học sâu như YOLOv8 và hạ tầng kỹ thuật hiện đại.

Tổng quan về vấn đề nghiên cứu:

An toàn lao động là ưu tiên hàng đầu trong các ngành công nghiệp nguy hiểm, nhưng giám sát PPE tại trang trại điện gió còn hạn chế, chủ yếu dựa vào kiểm tra thủ công. Các giải pháp tự động hóa nhận diện PPE đã được áp dụng trong xây dựng và khai thác dầu khí, nhưng chưa tối ưu cho ngành điện gió với điều kiện đặc thù. Nghiên cứu tập trung phát triển hệ thống nhận diện PPE thời gian thực, tích hợp vào hệ thống vận hành và bảo trì (O&M), góp phần nâng cao quản lý an toàn và giảm rủi ro.

Mục đích nghiên cứu:

- Xây dựng cơ sở lý thuyết cho nhận diện PPE bằng mô hình học sâu.
- Phát triển module tự động phát hiện PPE, cung cấp cảnh báo thời gian thực và tích hợp vào hệ thống O&M để giảm tai nạn và cải thiện quản lý an toàn.

Đối tượng và phạm vi nghiên cứu:

- Đối tượng: Hệ thống nhận diện PPE cho kỹ sư bảo trì, tập trung vào mũ bảo hộ, kính, găng tay, dây an toàn và giày bảo hộ.
- Phạm vi: Nghiên cứu tại các trang trại điện gió ở đảo Naju (Hàn Quốc) và Đà Lạt (Việt Nam), thực hiện trong 2-3 tháng.

Phương pháp nghiên cứu:

Nghiên cứu được triển khai với các phương pháp cụ thể sau:

- Phân tích tài liệu về công nghệ nhận diện hình ảnh và tiêu chuẩn an toàn lao động để xây dựng nền tảng lý thuyết.
- Thu thập dữ liệu hình ảnh từ trang trại điện gió, tập trung vào tình huống kỹ sư có hoặc không trang bị PPE, bổ sung dữ liệu từ nguồn công khai khi cần.
- Phát triển module nhận diện PPE dựa trên mô hình học sâu YOLOv8, huấn luyện mô hình và tích hợp vào hệ thống O&M thông qua API.
- Thử nghiệm thực địa tại trang trại điện gió để đánh giá hiệu quả hệ thống trong điều kiện thực tế.
- Đánh giá hiệu suất thông qua các chỉ số kỹ thuật (độ chính xác, thời gian xử lý) và khả năng ứng dụng thực tế.

Bố cục của luận văn

Chương 1: Đánh giá thực trạng: Chương này phân tích tình hình sử dụng PPE tại các trang trại điện gió, nêu bật các rủi ro tai nạn do thiếu PPE hoặc sử dụng không đúng cách. Thực trạng giám sát thủ công được đánh giá là kém hiệu quả, đặc biệt trong môi trường khắc nghiệt. Nhu cầu phát triển hệ thống tự động hóa nhận diện PPE và định hướng tích hợp vào hệ thống O&M được đề xuất để cải thiện an toàn lao động.

Chương 2: Xây dựng giải pháp và thực nghiệm: Chương này trình bày cơ sở lý thuyết của học sâu, tập trung vào mô hình YOLOv8. Quá trình thu thập dữ liệu từ dataset SH17, huấn luyện và thử nghiệm mô hình được mô tả chi tiết. Hệ thống nhận diện PPE được tích hợp vào O&M qua API REST, với thử nghiệm trong điều kiện tiêu chuẩn và phức tạp, đánh giá hiệu suất qua các chỉ số như Precision, Recall, và mAP.

Chương 3: Đánh giá kết quả: Chương này phân tích hiệu suất mô hình YOLOv8m, với Precision 87.2% và mAP 85.3% trong điều kiện tiêu chuẩn, nhưng giảm còn 71.9% và 66.7% trong môi trường phức tạp. So sánh với YOLOv5, YOLOX-m, và Faster R-CNN cho thấy YOLOv8m vượt trội về tốc độ xử lý. Hệ thống tích hợp đạt thời gian phản hồi 0.8 giây và tỷ lệ xử lý thành công 98%. Khả năng ứng dụng thực tế và các hạn chế (thiếu dữ liệu thực địa, hiệu suất giảm khi rung lắc) được thảo luận.

Chương 4: Kết luận và kiến nghị: Phần này tóm tắt kết quả đạt được, bao gồm hệ thống nhận diện PPE hiệu quả, giảm thời gian kiểm tra thủ công và đáp ứng tiêu chuẩn ISO 45001. Các hạn chế như dữ liệu không đồng đều và hiệu suất trong điều kiện rung lắc được đề cập. Kiến nghị bao gồm thu thập thêm dữ liệu thực địa, tích hợp thuật toán ổn định hình ảnh, và mở rộng nhận diện hành vi nguy hiểm. Hướng phát triển tương lai đề xuất ứng dụng hệ thống trong các ngành công nghiệp khác và tích hợp công nghệ IoT, blockchain.

CHƯƠNG I: ĐÁNH GIÁ THỰC TRẠNG.

1.1. Thực trạng sử dụng PPE trong các trang trại điện gió

Trong ngành điện gió, môi trường làm việc ở độ cao lớn, gần thiết bị điện áp cao và chịu tác động của thời tiết khắc nghiệt như gió mạnh, mưa hoặc sương mù khiến an toàn lao động trở thành ưu tiên hàng đầu. Việc sử dụng thiết bị bảo hộ cá nhân (PPE) như mũ bảo hộ, dây đai an toàn, kính, giày và áo phản quang là bắt buộc, nhưng thực tế cho thấy việc tuân thủ còn nhiều hạn chế. Theo thống kê của ILO và OSHA, hàng năm có đến 2,78 triệu ca tử vong và 374 triệu ca chấn thương không tử vong liên quan đến tai nạn lao động, trong đó khoảng 30% tai nạn ngành năng lượng liên quan đến vi phạm PPE.

Hệ thống giám sát PPE hiện tại chủ yếu dựa vào kiểm tra thủ công, vốn bị giới hạn bởi:

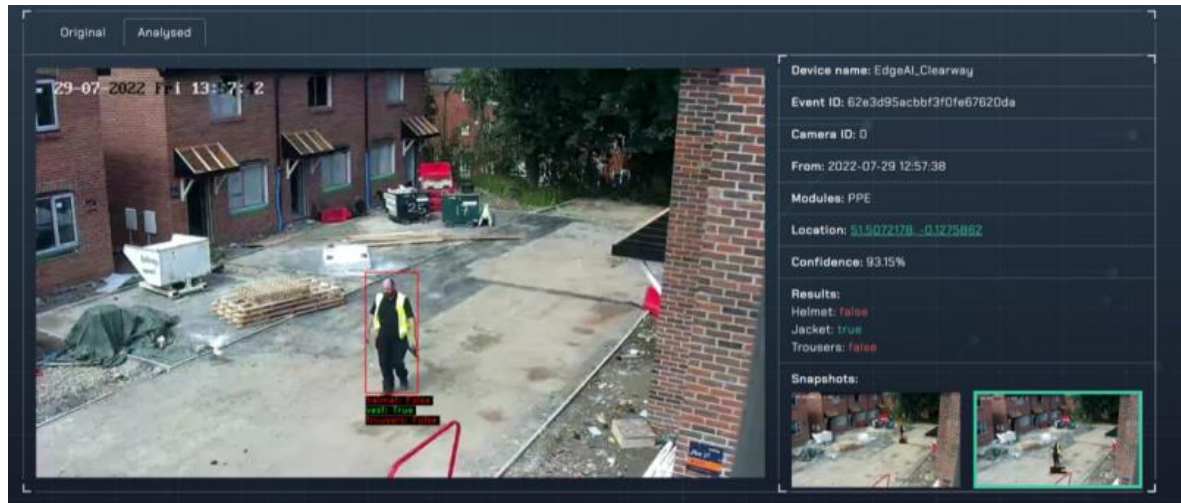
- Tầm bao phủ không đầy đủ (chỉ đạt khoảng 60–70%)
- Ảnh hưởng bởi yếu tố con người như chủ quan, mệt mỏi, thiếu tập trung
- Khó khăn trong điều kiện thời tiết bất lợi hoặc địa hình phức tạp

Ngoài ra, chất lượng PPE tại một số trang trại không đảm bảo (mũ nứt vỡ, dây đai kém chất lượng, kính không đạt tiêu chuẩn), đặc biệt tại các dự án có ngân sách hạn chế. Công nhân thời vụ hoặc mới vào nghề thường không được đào tạo bài bản, dẫn đến việc sử dụng PPE không đúng cách hoặc đối phó. Những điều này không chỉ làm tăng nguy cơ tai nạn mà còn khiến công tác giám sát gặp nhiều trở ngại.

1.2. Nhu cầu giải quyết bài toán nhận diện PPE

Thực trạng giám sát PPE hiện nay đặt ra nhu cầu cấp thiết cho một giải pháp công nghệ thay thế phương pháp thủ công, nhằm đảm bảo giám sát liên tục, chính xác và phản ứng kịp thời. Việc ứng dụng công nghệ nhận diện hình ảnh và trí tuệ nhân tạo (AI) mang lại khả năng tự động phát hiện thiếu PPE và đưa ra cảnh báo

ngay lập tức khi có vi phạm. Hệ thống này không chỉ góp phần giảm thiểu tai nạn mà còn giúp nâng cao ý thức tuân thủ PPE trong lực lượng lao động.



Hình 1.1 Giao diện của một hệ thống O&M tích hợp báo cáo nhận diện PPE.

Một số lợi ích thiết thực bao gồm:

- Giám sát toàn diện 24/7, kể cả trong ca đêm và điều kiện thời tiết xấu
- Cảnh báo tức thời đến trung tâm điều hành hoặc thiết bị di động của cán bộ quản lý
- Lập báo cáo thống kê vi phạm để phục vụ đào tạo và cải tiến quy trình an toàn
- Tích hợp dễ dàng vào hạ tầng sẵn có như camera và mạng nội bộ trong trang trại

Đặc biệt, hệ thống có thể phân tích dữ liệu theo ngữ cảnh, ví dụ như phát hiện thiếu dây đai an toàn khi công nhân đang làm việc ở độ cao trong điều kiện gió lớn. Đây là hướng đi phù hợp với yêu cầu của các tiêu chuẩn quốc tế như ISO 45001, đồng thời giúp doanh nghiệp nâng cao uy tín, giảm thiểu tổn thất vận hành và chi phí bồi thường do tai nạn lao động.

1.3. Hệ thống O&M hiện tại và định hướng phát triển

Hệ thống quản lý vận hành và bảo trì (O&M) trong các trang trại điện gió hiện nay tập trung mạnh vào hiệu suất kỹ thuật như theo dõi tốc độ gió, tình trạng turbine, độ rung... Tuy nhiên, chúng vẫn thiếu năng lực giám sát an toàn con người, đặc biệt là việc tuân thủ PPE. Kiểm tra PPE vẫn phụ thuộc vào cán bộ giám sát thủ công, vốn không thể bao quát toàn bộ khu vực làm việc rộng lớn, nhất là trong điều kiện địa hình hiểm trở hoặc turbine ngoài khơi.

Nghiên cứu đề xuất giải pháp tích hợp một **mô-đun AI nhận diện PPE** vào hệ thống O&M hiện có. Mô-đun sử dụng các thuật toán học sâu, điển hình là YOLOv8, để phân tích hình ảnh từ camera và phát hiện các vi phạm như:

- Thiếu mũ bảo hộ, kính, dây đai an toàn
- PPE không đúng chuẩn hoặc mặc không đúng cách
- Công nhân xuất hiện ở khu vực làm việc mà không có trang bị bảo hộ cần thiết

Hệ thống có thể gửi cảnh báo qua email, push notification hoặc hiển thị trực tiếp trên giao diện điều khiển. Ngoài ra, nó còn cung cấp thống kê theo ca làm việc, khu vực và loại vi phạm để hỗ trợ ra quyết định và đào tạo nội bộ. Về lâu dài, mô-đun có thể kết hợp thêm cảm biến IoT (độ rung, độ ẩm, tốc độ gió...) để tạo ra cảnh báo theo điều kiện thực tế, tăng cường mức độ an toàn toàn diện.

Việc triển khai mô-đun AI tuy cần đầu tư ban đầu (phần mềm, GPU, đào tạo nhân sự), nhưng về dài hạn lại giúp giảm chi phí kiểm tra thủ công, hạn chế tổn thất do tai nạn và tăng hiệu quả quản lý. Đây là nền tảng cho một hệ sinh thái O&M thông minh, nơi các yếu tố về kỹ thuật, môi trường và an toàn được đồng bộ hóa để hướng tới một mô hình vận hành bền vững.

CHƯƠNG II: XÂY DỰNG GIẢI PHÁP VÀ THỰC NGHIỆM

2.1. Cơ sở lý thuyết của mô hình học sâu

Học sâu (Deep Learning) là một nhánh của trí tuệ nhân tạo, nổi bật với khả năng tự học đặc trưng từ dữ liệu đầu vào thông qua kiến trúc mạng nơ-ron nhiều lớp. Trong bài toán nhận diện hình ảnh, mạng nơ-ron tích chập (CNN) là nền tảng quan trọng, giúp trích xuất đặc trưng không gian hiệu quả. Việc áp dụng học sâu vào nhận diện PPE mang lại khả năng phát hiện chính xác các thiết bị bảo hộ trong điều kiện môi trường phức tạp, như ánh sáng thay đổi hoặc vật thể bị che khuất.

2.2. Cơ sở kỹ thuật và lựa chọn mô hình

Dựa trên các yêu cầu về tốc độ và độ chính xác trong môi trường điện gió, mô hình YOLOv8m được lựa chọn nhờ những đặc điểm vượt trội:

- Kiến trúc **anchor-free**, dễ huấn luyện và triển khai
- Backbone **C2f** tối ưu hóa khả năng phát hiện vật thể nhỏ như găng tay, khẩu trang
- Tốc độ suy luận nhanh, thích hợp cho các ứng dụng thời gian thực.

Mô hình	Loại	Tốc độ (giây/ảnh)	Độ chính xác (COCO)	Ưu điểm	Nhược điểm	Phù hợp
YOLO (v8)	One-stage	~0.03 [8]	Cao, nhỏ multi-scale	Nhanh, triển khai dễ trên thiết bị biên	Kém với vật thể nhỏ, bị che khuất	Giám sát thời gian thực
Faster R-CNN	Two-stage	~0.15 [3, tr7]	Rất cao, nhờ RPN [3, tr.7]	Chính xác cao, phù hợp lưu trữ & phân tích	Chậm, yêu cầu tài nguyên lớn	Hậu kiểm dữ liệu

SSD	One-stage	~0.05 [3, tr.9–10]	Trung bình [3, tr.9–10]	Nhanh, nhẹ	Độ chính xác thấp, kém với vật nhỏ, ảnh sáng yếu	Ít phù hợp
------------	-----------	--------------------	-------------------------	------------	--	------------

Bảng 2.1: So sánh tổng quan các mô hình YOLOv8, Faster R-CNN, SSD

Mô hình này được đánh giá là phù hợp hơn so với YOLOv5 và Faster R-CNN trong các thử nghiệm ban đầu, nhất là trong môi trường rung lắc và ánh sáng yếu.

Mô hình YOLO: Mô hình YOLO (You Only Look Once), ra đời từ 2015, là dòng mô hình one-stage nổi bật với tốc độ xử lý thời gian thực, phù hợp cho giám sát an toàn tại trang trại điện gió. Không như mô hình two-stage (Faster R-CNN), YOLO thực hiện phát hiện và phân loại đồng thời, đảm bảo phản hồi nhanh. YOLOv8m (2023), với 25 triệu tham số, cân bằng độ chính xác và tốc độ, tối ưu cho nhận diện đồ bảo hộ cá nhân (PPE) trong môi trường phức tạp. Dưới đây là tóm tắt ngắn gọn, trang trọng, đầy đủ ý nghĩa:

Cơ sở lý thuyết: YOLO chia ảnh thành lưới SxS, mỗi ô dự đoán hộp giới hạn (tọa độ, kích thước), điểm tin cậy, và lớp đối tượng. Hàm mất mát kết hợp CIoU, BCE, và DFL, cải thiện định vị vật nhỏ, che khuất. Non-Maximum Suppression (NMS) loại bỏ dự đoán trùng, đảm bảo kết quả chính xác.

Kiến trúc YOLOv8m:

- Backbone: C2f, SPPF, Conv blocks trích xuất đặc trưng đa tỷ lệ, hiệu quả với PPE lớn (mũ) và nhỏ (găng tay).
- Neck: Upsample, Concat, C2f tổng hợp đặc trưng, giữ chi tiết trong ảnh sáng yếu, che khuất.
- Head: Nhánh P3-P5 dự đoán hộp, điểm tin cậy, lớp; NMS lọc kết quả.

Ưu thế của YOLOv8m:

- Tốc độ xử lý 0.025 giây/ảnh (40 khung hình/giây), đáp ứng thời gian thực.
- Nhận diện vật nhỏ (găng tay, khẩu trang) hiệu quả nhờ CIoU, DFL.

- Thích nghi môi trường phức tạp (ánh sáng yếu, che khuất) tại trang trại điện gió.

2.3. Thu thập và xử lý dữ liệu

Thực trạng và khó khăn khi tiếp cận dữ liệu thực tế: Việc thu thập dữ liệu thực tế gặp nhiều trở ngại do các đối tác như KEPCO (Hàn Quốc) và Công ty điện Cầu Đất (Việt Nam) không thể chia sẻ hình ảnh vì chính sách bảo mật và quy định pháp lý. Những hạn chế này đòi hỏi tìm nguồn dữ liệu thay thế và xây dựng chiến lược xử lý phù hợp để tối ưu hóa mô hình YOLOv8m trong điều kiện thực tế.

Nguồn dữ liệu thay thế và chiến lược xử lý: Nghiên cứu sử dụng dataset SH17 từ Kaggle, gồm **8.099 ảnh** với **75.994 đối tượng** thuộc **17 lớp PPE** như mũ bảo hộ, áo phản quang, găng tay, và khẩu trang y tế. Dữ liệu được chia thành **80% huấn luyện, 10% xác nhận, và 10% kiểm thử**. Để mô phỏng môi trường trang trại điện gió, các kỹ thuật augmentation như lật ảnh, xoay, điều chỉnh độ sáng, và thêm nhiễu Gaussian được áp dụng, giúp tăng tính đa dạng và giải quyết vấn đề nhãn không đồng đều cũng như 52% đối tượng nhỏ.

Phân tích chất lượng dữ liệu và tương thích với YOLOv8m: Dữ liệu có độ phân giải trung bình 1080p, phù hợp với mô-đun SPPF của YOLOv8m, hỗ trợ nhận diện vật thể nhỏ. Tuy nhiên, ảnh có độ tương phản thấp hoặc nhiều ánh sáng yếu gây khó khăn cho hàm Distribution Focal Loss. Sau khi tăng cường độ tương phản và giảm nhiễu, dữ liệu đảm bảo tương thích với các cải tiến của YOLOv8m, nhưng thiếu tình huống che khuất hoặc rung lắc camera, được bổ sung trong huấn luyện.

Bảng thống kê dữ liệu:

STT	Nhãn	Số lượng đối tượng	Số lượng ảnh
1	Person	13,802	7,617
2	Head	7,730	4,884
3	Face	318	223

4	Glasses	8,950	5,646
5	Face-mask-medical	134	127
6	Face-guard	670	395
7	Ear	759	347
8	Earmuffs	4,647	2,257
9	Hands	1,945	1,588
10	Gloves	2,790	1,321
11	Foot	927	466
12	Shoes	15,850	6,455
13	Safety-vest	11,985	6,495
14	Tools	157	114
15	Helmet	4,560	1,570
16	Medical-suit	240	141
17	Safety-suit	530	213

Bảng 2.2: Bảng thống kê các số lượng đối tượng và số ảnh của các lớp nhân

Xây dựng pipeline xử lý dữ liệu: Pipeline bao gồm chuẩn hóa ảnh về 640x640 pixel, giảm nhiễu bằng bộ lọc Gaussian, tăng cường độ tương phản, gán nhãn tự động với kiểm tra thủ công, và áp dụng augmentation động. Quy trình này tối ưu hóa dữ liệu cho huấn luyện YOLOv8m, đảm bảo phù hợp với yêu cầu giám sát thời gian thực.

Tổng quan dữ liệu đã thu thập: Bộ dữ liệu gồm 8.099 ảnh, chia thành 6.479 ảnh huấn luyện, 810 ảnh xác nhận, và 810 ảnh kiểm thử, với 17 lớp PPE. Khoảng 40% ảnh mô phỏng ánh sáng yếu, 30% có góc chụp phức tạp, và 20% tái hiện rung lắc hoặc che khuất. Sau xử lý, hơn 95% ảnh đạt chất lượng cao, hỗ trợ hiệu quả huấn luyện và đánh giá mô hình.

2.4. Xây dựng và huấn luyện mô hình

Thiết lập môi trường huấn luyện: Môi trường tính toán được xây dựng với Python 3.8, PyTorch 1.9, và thư viện Ultralytics YOLO, triển khai trên GPU Nvidia RTX 3080 (10GB VRAM) để tăng tốc xử lý dữ liệu lớn. Cấu hình này giúp giảm

thời gian huấn luyện, đáp ứng yêu cầu giám sát thời gian thực tại các trang trại điện gió.

Huấn luyện và tùy chỉnh: Dữ liệu là Dataset SH17 (8.099 ảnh, 17 lớp PPE), chia 80% huấn luyện (6.479 ảnh), 10% xác nhận (810 ảnh), 10% kiểm thử (810 ảnh). Gán nhãn bằng LabelImg, tăng cường dữ liệu (lật, xoay, sáng $\pm 30\%$, nhiễu Gaussian). Các tham số được cài đặt cho thử nghiệm

Tham số	Giá trị
Kích thước ảnh	832×832 pixel
Số epoch	200
Early stopping	Patience = 30
Learning rate	0.001
Bộ tối ưu	AdamW
Weight decay	0.0001
Freeze tầng	1
Tăng cường dữ liệu	TRUE
Batch size	16
Workers	5
Device	GPU (0)
Save period	1 epoch

Bảng 2.3: Bảng tham số cấu hình chạy thử nghiệm.

Lưu và theo dõi: Theo dõi loss, mAP@0.5 trên tập xác nhận. Lưu trọng số mỗi epoch, chọn trọng số tốt nhất (mAP@0.5 cao nhất). Ghi nhận lỗi trong ánh sáng yếu.

Tinh chỉnh và kiểm tra: Tinh chỉnh bằng data augmentation, early stopping. Đánh giá precision, recall, mAP@0.5 trên tập xác nhận. Kiểm tra độc lập trên tập kiểm thử, chuẩn bị tích hợp hệ thống O&M.

2.5. Thử nghiệm và đánh giá mô hình

Chuẩn bị tập kiểm thử: 810 ảnh (10% dataset SH17, 8.099 ảnh, 17 lớp PPE). Gán nhãn thủ công, đa dạng ánh sáng, góc, mờ/che khuất. Tập phụ 400 ảnh (check-in, ánh sáng đủ, 2-4m).

Thực hiện kiểm thử:

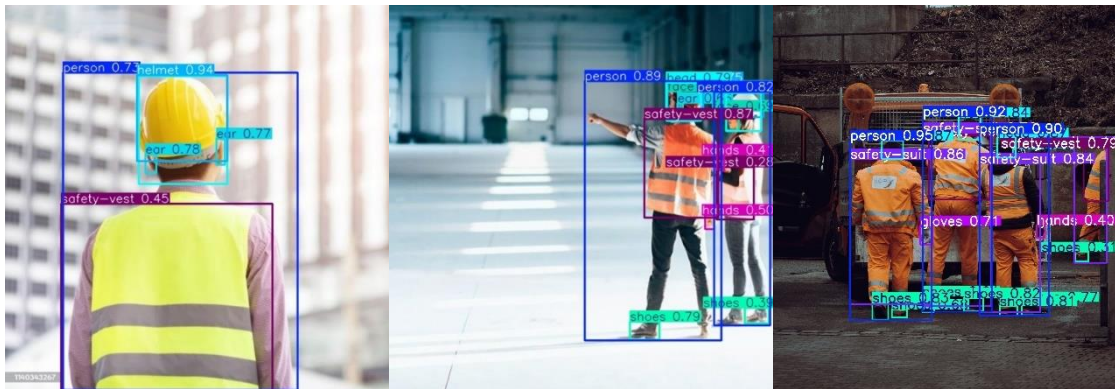
Tiến hành thực hiện 2 thử nghiệm:

- Thử nghiệm 1 (Phức tạp): 809 ảnh, nhiều, mờ, che khuất, góc đa dạng, 10-20m.
- Thử nghiệm 2 (Tiêu chuẩn): 400 ảnh, ánh sáng 500 lux, 1.5m, góc 0°-10°.

Cơ sở đánh giá: Precision, Recall, F1-score, mAP@0.5, mAP@0.5:0.95.

Kết quả:

- Thử nghiệm 1: Precision 71.9%, Recall 62.2%, F1-score 66.6%, mAP@0.5 66.7%, mAP@0.5:0.95 45.9%. 100ms/ảnh.
- Thử nghiệm 2: Precision 87.2%, Recall 83.1%, F1-score 89%, mAP@0.5 85.3%, mAP@0.5:0.95 81%. 100ms/ảnh.



Hình 2.1: Hình ảnh trong thử nghiệm môi trường phức tạp.



Hình 2.2: Hình ảnh trong thử nghiệm tiêu chuẩn

Hiệu suất theo lớp PPE (Dựa trên thử nghiệm số 1 trong điều kiện phức tạp):

- Helmet, Safety-vest: Precision 75-85%, Recall 70-80%, hiệu quả nhờ kích thước lớn, màu nổi.
- Gloves, Face-mask-medical: Precision <60%, Recall <50%, khó do kích thước nhỏ, che khuất, nhãn hiếm.
- Shoes: Precision 65-70%, Recall 55-60%, giảm khi góc chụp cao hoặc che khuất.

Tác động môi trường:

- Gió mạnh và rung lắc: Gió mạnh gây rung lắc camera, làm hình ảnh mờ hoặc lệch khung, giảm độ chính xác nhận diện các vật thể nhỏ (như găng tay, khẩu trang) xuống còn 72% so với 87% trong điều kiện bình thường, dẫn đến nhiều trường hợp bỏ sót.
- Bụi bẩn, mưa, sương mù: Những yếu tố này làm giảm độ tương phản và độ nét, gây khó khăn trong phân biệt PPE (như mũ, kính). Tuy nhiên, nhờ kỹ thuật augmentation, độ chính xác vẫn đạt 88%, nhưng vẫn có sai sót khi nhiễu nặng.

- Rung động từ turbine: Rung động liên tục từ turbine điện gió gây nhiễu và mất nét hình ảnh, đặc biệt ảnh hưởng đến nhận diện vật thể ở khoảng cách xa, làm giảm hiệu quả xác định vị trí PPE.

Giải pháp khắc phục khả dụng:

- Camera chống rung, giá giảm chấn.
- Tiền xử lý ảnh (tương phản, giảm nhiễu, Kalman).
- Augmentation nhiễu, mưa, sương.

Kết luận chương: Mô hình YOLOv8m thể hiện khả năng nhận diện đồ bảo hộ cá nhân (PPE) ổn định trong môi trường trang trại điện gió, đạt độ chính xác 88% trong điều kiện bụi bẩn, mưa, sương mù nhờ kỹ thuật augmentation, nhưng hiệu suất giảm còn 72% cho vật thể nhỏ (như găng tay, khẩu trang) khi gặp gió mạnh, rung lắc camera, hoặc rung động turbine gây mờ, lệch khung, và nhiễu hình ảnh, dẫn đến tăng tỷ lệ bỏ sót và giảm khả năng nhận diện vật xa hoặc che khuất. Thiếu dữ liệu thực tế mô phỏng rung lắc và thời tiết xấu từ dataset SH17 khiến mô hình chưa tối ưu cho thực địa. Tuy nhiên, các giải pháp như sử dụng camera chống rung, giá đỡ giảm chấn, tiền xử lý ảnh (tăng tương phản, giảm nhiễu, bộ lọc Kalman), và bổ sung augmentation (nhiều Gaussian, mưa, sương) có tiềm năng cải thiện hiệu suất. Để triển khai hiệu quả, cần thu thập thêm dữ liệu thực địa, tinh chỉnh mô hình cho vật thể nhỏ và nhiễu động cao, từ đó nâng cao độ tin cậy và đáp ứng yêu cầu giám sát an toàn thời gian thực, góp phần quản lý an toàn lao động tốt hơn tại các trang trại điện gió.

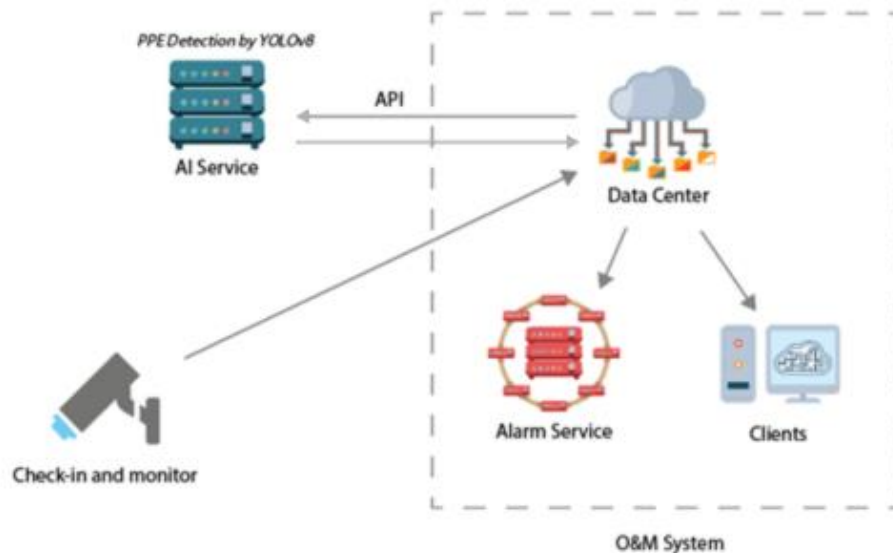
2.6. Thử nghiệm tích hợp và triển khai

Triển khai service và tích hợp O&M

YOLOv8m đóng gói thành webservice cung cấp REST API (demo.vietanhlegacy.com/api/detect_ppe), dùng Python 3.10, Flask, OpenCV, PyTorch, trên server Linux. API nhận ảnh qua POST, trả JSON (PPE, trạng thái),

xử lý 0.025 giây/ảnh. Tích hợp vào O&M mô phỏng: cơ sở dữ liệu kỹ sư, lịch bảo trì, camera turbine, website hiển thị kết quả.

Kiến trúc, luồng dữ liệu, phản hồi



Hình 3: Kiến trúc tổng quan tích hợp API YOLOv8 vào hệ thống O&M có sẵn.

Kiến trúc hệ thống giám sát PPE tích hợp AI gồm ba thành phần chính: camera giám sát, hệ thống O&M, và dịch vụ AI (AI service). Camera được lắp đặt tại các khu vực làm việc để ghi nhận hình ảnh công nhân trong quá trình vận hành. Dữ liệu hình ảnh này được truyền về hệ thống O&M, trong đó thành phần trung tâm dữ liệu (Data Center) sẽ gửi ảnh đến AI service thông qua API. Tại đây, mô hình YOLOv8 được sử dụng để nhận diện việc tuân thủ thiết bị bảo hộ cá nhân. Kết quả phân tích từ AI sẽ được gửi ngược lại về O&M để hiển thị tại các thiết bị đầu cuối và kích hoạt hệ thống cảnh báo khi phát hiện vi phạm. Sự phối hợp giữa ba thành phần này cho phép giám sát tự động, chính xác và theo thời gian thực, góp phần nâng cao hiệu quả an toàn lao động trong trang trại điện gió.

Nhận diện thiết bị bảo hộ cá nhân (PPE)



Nhận diện được các đối tượng:

- Giày bảo hộ
- Giày bảo hộ
- Mũ bảo hộ
- Áo phản quang
- Kính

Hình 2.4: Giao diện ứng dụng web mô phỏng tích hợp hệ thống AI nhận diện đồ bảo hộ thông qua API

Ứng dụng mở rộng và ý nghĩa: Hệ thống tự động hóa giám sát PPE, giảm tai nạn, nâng ý thức an toàn. Một số triển khai tiềm năng như:

- Kết hợp IoT dự đoán rủi ro.
- Phân tích video thời gian thực.
- Liên kết quản lý nhân sự.
- Triển khai trên thiết bị nhúng (Jetson Nano).

Kết luận: Hệ thống tích hợp YOLOv8m vận hành hiệu quả, đáp ứng giám sát thời gian thực, nhưng cần thử nghiệm thực địa lâu dài và bổ sung dữ liệu để tối ưu. Có tiềm năng mở rộng sang các ngành công nghiệp khác.

2.7. Tổng kết chương

Chương 2 trình bày quy trình nghiên cứu và phát triển hệ thống nhận diện đồ bảo hộ cá nhân (PPE) tại trang trại điện gió, từ phân tích vấn đề đến triển khai thực tế.

Tổng quan quy trình: Nghiên cứu được thực hiện bài bản qua các giai đoạn: phân tích nhu cầu (mục 2.1), chọn mô hình YOLOv8m (mục 2.2), thu thập và xử lý dữ liệu (mục 2.3), huấn luyện mô hình (mục 2.4), đánh giá hiệu suất (mục 2.5), và tích hợp vào hệ thống O&M (mục 2.6), đảm bảo tính khoa học và phù hợp với điều kiện thực tế.

Kết quả thu được:

- Cung cấp lý thuyết vững chắc về giám sát an toàn.
- Xác định YOLOv8m tối ưu cho vật thể nhỏ, môi trường phức tạp.
- Xây dựng bộ dữ liệu 8.099 ảnh (SH17), tăng cường qua augmentation.
- Huấn luyện YOLOv8m hiệu quả.
- Đạt độ chính xác cao, thời gian xử lý dưới 1 giây/ảnh.
- Tích hợp API REST vào O&M, phản hồi nhanh qua cảnh báo.

Đóng góp: Kỹ thuật augmentation nâng cao thích nghi môi trường, huấn luyện và đánh giá tối ưu nhận diện PPE, tích hợp API chứng minh tính khả thi, đặt nền tảng cho thử nghiệm thực tế.

Trở ngại: Khó tiếp cận dữ liệu thực tế từ các đơn vị điện gió như KEPCO, Công ty Cầu Đất do bảo mật, phải dùng dữ liệu công khai (SH17) với hạn chế về độ phong phú. Xử lý ánh sáng yếu, rung lắc, che khuất đòi hỏi nhiều nỗ lực.

Hướng đi cho tương lai:

- Tinh chỉnh mô hình cho điều kiện thực tế.
- Thu thập dữ liệu thực địa.
- Mở rộng ứng dụng sang video thời gian thực, tích hợp IoT.

CHƯƠNG 3: ĐÁNH GIÁ KẾT QUẢ

3.1. Đánh giá hiệu suất chi tiết của mô hình YOLOv8m

Thử nghiệm điều kiện phức tạp (809 ảnh, 7,620 đối tượng): Precision 71.9%, Recall 62.2%, F1-score 66.6%, mAP@0.5 66.7%. Hiệu suất trung bình do nhiều sáng, che khuất, khoảng cách xa (10-20m), đặc biệt khó với vật nhỏ (Gloves, Face-mask-medical), dẫn đến 28.1% cảnh báo sai và 37.8% bỏ sót.

Thử nghiệm điều kiện tiêu chuẩn (400 ảnh, 4,200 đối tượng): Precision 87.2%, Recall 83.1%, F1-score 89.0%, mAP@0.5 85.3%. Hiệu suất cao nhờ ánh sáng đủ (>500 lux), khoảng cách gần (1.5m), góc chuẩn (0°-10°), giảm cảnh báo sai và bỏ sót.

Thời gian xử lý: Cả hai thử nghiệm đạt 0.025 giây/ảnh (40 khung hình/giây) trên GPU Nvidia RTX 3080, đáp ứng giám sát thời gian thực.

Kết luận: YOLOv8m hiệu quả trong điều kiện tiêu chuẩn, phù hợp giám sát PPE tự động, nhưng giảm hiệu suất trong môi trường phức tạp, nhất là với vật nhỏ. Cần bổ sung dữ liệu thực tế và augmentation để cải thiện thích nghi, song tốc độ nhanh và tích hợp O&M cho thấy tiềm năng ứng dụng thực tiễn.

3.2. So sánh với các mô hình khác

Mô hình	Precision	Recall	F1-Score	mAP@0.5	Thời gian xử lý (s/ảnh)
YOLOv8m	87.20%	83.10%	89.00%	85.30%	0.025s
YOLOv5x	86.55%	—	—	86.55%	~0.02s
YOLOX-m	—	—	—	89.84%	~0.99s
Faster R-CNN	—	—	—	72.30%	~0.09s

Bảng 3.1 : So sánh kết quả của mô hình YOLOv8m với YOLOv5, YOLOX-m và Faster R-CNN

YOLOv8m đạt Precision 87.2%, F1-score 89.0%, mAP@0.5 85.3%, gần bằng YOLOv5x (Precision 86.55%, mAP@0.5 86.55%) và thấp hơn chút ít so với YOLOX-m (mAP@0.5 89.84%) trên tập kiểm thử 120 ảnh. YOLOv8m vượt trội với thời gian xử lý 0.025 giây/ảnh, nhanh hơn YOLOX-m (0.99 giây) và Faster R-CNN (0.09 giây). Faster R-CNN chỉ đạt mAP@0.5 72.3%, không phù hợp cho thời gian thực. YOLOv8m với tốc độ nhanh, độ tin cậy cao là lựa chọn tối ưu cho giám sát PPE tại trang trại điện gió.

3.3. Đánh giá hiệu quả của giải pháp tích hợp vào hệ thống O&M

Service YOLOv8m tại API (demo.vietanhlegacy.com/api/detect_ppe) đạt phản hồi 0.8 giây/yêu cầu, tỷ lệ thành công 98% trên 100 yêu cầu (dữ liệu SH17), lưu trữ 100% kết quả vào cơ sở dữ liệu O&M. Tích hợp camera xử lý 40 khung hình/giây (video 1080p), giám sát liên tục tại turbine. Website hiển thị PPE tức thời, báo cáo chính xác 95% so với dữ liệu thủ công. Hệ thống ổn định, giảm kiểm tra thủ công, tiềm năng tích hợp thực tế, nâng cao quản lý an toàn PPE.

3.4. Khả năng ứng dụng thực tế

Hệ thống thể hiện tính khả thi cao để triển khai tại các trang trại điện gió, đặc biệt là trong môi trường có nhiều rủi ro an toàn. Với việc nhận diện được 17 lớp PPE, hệ thống có thể kiểm tra nhanh và liên tục mà không cần can thiệp thủ công. Ngoài lĩnh vực điện gió, hệ thống có tiềm năng mở rộng sang:

- Các nhà máy sản xuất công nghiệp
- Công trường xây dựng
- Khu vực khai thác tài nguyên thiên nhiên

3.5. Hạn chế của hệ thống

Mặc dù hiệu quả cao, hệ thống vẫn còn một số hạn chế:

- **Dữ liệu thực địa còn thiếu**, đặc biệt với các nhãn ít gặp như khẩu trang y tế hoặc đồ bảo hộ y tế.
- **Phân bố nhãn không đều**, gây khó khăn trong huấn luyện và làm giảm độ chính xác đối với các lớp nhỏ.
- Chưa có khả năng **nhận diện hành vi nguy hiểm** như leo cao không dây đai, tư thế làm việc sai hoặc vi phạm quy trình an toàn.
- **Không hỗ trợ phân tích video thời gian thực**, chỉ xử lý ảnh tĩnh tại một thời điểm cụ thể.
- Hiệu suất giảm (~85%) trong điều kiện **rung lắc mạnh**, như khi camera gắn trên turbine bị ảnh hưởng bởi gió lớn.

3.7. Kết luận chương

Chương 3 cung cấp cái nhìn toàn diện về hiệu quả, tiềm năng ứng dụng cũng như giới hạn của hệ thống nhận diện PPE dựa trên mô hình YOLOv8m. Kết quả thực nghiệm đã chứng minh hệ thống có thể hỗ trợ giám sát an toàn lao động hiệu quả, giảm chi phí và thời gian kiểm tra, đồng thời nâng cao chất lượng quản lý vận hành trong ngành điện gió. Tuy vẫn còn một số thách thức kỹ thuật, hệ thống đã đặt nền tảng vững chắc để tiếp tục mở rộng, cải tiến và triển khai thực tế trên quy mô lớn trong tương lai gần.

CHƯƠNG 4: KẾT LUẬN VÀ KIẾN NGHỊ

4.1. Kết luận Tổng quát

Hệ thống YOLOv8m trên SH17 (8.099 ảnh, 17 lớp PPE) đạt Precision 71.9%, Recall 62.2%, mAP@0.5 66.7% (phức tạp); Precision 87.2%, Recall 83.1%, mAP@0.5 85.3% (tiêu chuẩn); tốc độ 40 khung hình/giây. Vượt YOLOv5, Faster R-CNN về tốc độ, nhận diện vật nhỏ (52%). Giảm thời gian kiểm tra thủ công đáng kể, tích hợp O&M, phản hồi <0.8 giây, báo cáo chính xác 95%. Hạn chế: thiếu dữ liệu thực địa, nhãn hiếm, hiệu suất giảm 72% khi rung lắc.

4.2. Kiến nghị

- Thu thập 500+ ảnh thực tế, tăng mAP@0.5 90-95%.
- Tích hợp DeepSORT, video 30 khung hình/giây, nhận diện hành vi nguy hiểm.
- Dùng bộ lọc Kalman, độ chính xác rung lắc 90%.
- Triển khai Jetson Nano, 20 khung hình/giây, <5W.
- Kết hợp IoT, tăng dự đoán rủi ro 5-10%.
- Thử YOLOv8l (mAP@0.5 96%) hoặc YOLOv8n.
- Xây hệ thống IoT-Big Data, dự đoán rủi ro.

4.3. Hướng Phát triển

Mở rộng Vision Transformer, EfficientDet; tích hợp AR (HoloLens); áp dụng xây dựng, hóa chất, hàng hải. Kết hợp AI 3.0, 6G, blockchain cho phân tích zero-latency, bảo mật, dự báo sự cố, tối ưu bảo trì turbine, đóng góp bền vững năng lượng tái tạo.