

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



Phạm Việt Anh

ĐỀ ÁN TỐT NGHIỆP THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

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

ĐỀ 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

LỜI CAM ĐOAN

Tôi xin cam đoan rằng toàn bộ nội dung trong đề án “Nghiên cứu mô hình học sâu cho nhận diện đồ bảo hộ trong quản lý trang trại điện gió” là kết quả nghiên cứu của riêng tôi, được thực hiện dưới sự hướng dẫn, chỉ bảo và góp ý của giảng viên hướng dẫn tại Học viện Công nghệ Bưu chính Viễn Thông. Các số liệu, kết quả được trình bày trong đề án là trung thực và chưa từng được công bố trong bất kỳ công trình nào khác và không sao chép dưới bất kỳ hình thức nào. Các tài liệu tham khảo trong đề án đã được trích dẫn đầy đủ, rõ ràng và tuân thủ quy định.

Tôi xin chịu mọi trách nhiệm về công trình nghiên cứu của riêng mình!

Tác giả đề án

LỜI CẢM ƠN

Trước hết, tôi xin bày tỏ lòng biết ơn sâu sắc đến PGS.TS Trần Quang Anh, người đã tận tình hướng dẫn tôi trong suốt quá trình thực hiện đề án tốt nghiệp. Những ý kiến đóng góp quý báu và sự tận tâm của thầy đã giúp tôi hoàn thiện đề tài "Nghiên cứu mô hình học sâu cho nhận diện đồ bảo hộ trong quản lý trang trại điện gió" một cách hiệu quả và thực tiễn hơn.

Tôi xin chân thành cảm ơn Khoa Đào tạo sau đại học – Học viện Công nghệ Bưu chính Viễn thông đã tạo mọi điều kiện thuận lợi để tôi được học tập và nghiên cứu trong một môi trường khoa học nghiêm túc và hiện đại.

Bên cạnh đó, tôi xin gửi lời cảm ơn đặc biệt đến TS. Đào Thị Thúy Quỳnh, cô đã đồng hành, hỗ trợ tôi hoàn thiện đề án. Và tôi xin cảm ơn quý thầy cô giảng viên trong Khoa CNTT đã tận tình giảng dạy, chia sẻ kiến thức và kinh nghiệm quý báu trong suốt quá trình học tập, từ đó tạo nền tảng vững chắc giúp tôi thực hiện và hoàn thiện đề án tốt nghiệp này.

Cuối cùng, tôi xin gửi lời cảm ơn đến gia đình và bạn bè đã luôn đồng hành, động viên và ủng hộ tôi trong suốt hành trình học tập và nghiên cứu.

Tôi xin chân thành cảm ơn!

Hà Nội, ngày 30 tháng 04 năm 2025

Học viên thực hiện đề án

Phạm Việt Anh

MỤC LỤC

LỜI CAM ĐOAN	1
LỜI CẢM ƠN.....	2
MỤC LỤC	3
DANH MỤC CHỮ VIẾT TẮT.....	1
DANH MỤC BẢNG	2
DANH MỤC HÌNH ẢNH	3
MỞ ĐẦU.....	4
1. Lý do chọn đề tài.....	4
2. Tổng quan về vấn đề nghiên cứu	5
3. Mục đích nghiên cứu	6
4. Đối tượng và phạm vi nghiên cứu.....	7
5. Phương pháp nghiên cứu.....	8
CHƯƠNG 1: ĐÁNH GIÁ THỰC TRẠNG	10
1.1. Thực trạng sử dụng PPE trong các trang trại điện gió	11
1.2. Nhu cầu giải quyết bài toán nhận diện PPE	15
1.3. Hệ thống O&M hiện tại và định hướng phát triển.....	19
1.4. Định nghĩa bài toán	24
CHƯƠNG 2: XÂY DỰNG GIẢI PHÁP VÀ THỰC NGHIỆM	26
2.1. Cơ sở lý thuyết của mô hình học sâu	26
2.2. Cơ sở kỹ thuật và lựa chọn mô hình.....	27
2.3. Thu thập và xử lý dữ liệu.....	34
2.4. Xây dựng và huấn luyện mô hình.....	40

2.5. Thử nghiệm và đánh giá mô hình	43
2.6. Tổng kết chương	51
CHƯƠNG 3: ĐÁNH GIÁ KẾT QUẢ.....	53
3.1. Đánh giá tổng quan kết quả thu được.....	53
3.2. Thử nghiệm tích hợp và triển khai	56
3.3. Đánh giá hiệu suất chi tiết của mô hình YOLOv8m	60
3.4. So sánh với các mô hình khác.....	63
3.5. Đánh giá hiệu quả của giải pháp tích hợp service vào hệ thống O&M	64
3.4. Khả năng ứng dụng thực tế.....	65
3.5. Hệ thống còn hạn chế.....	66
3.6. Kết luận chương.....	66
CHƯƠNG 4: KẾT LUẬN VÀ KIẾN NGHỊ.....	67
4.1. Kết luận tổng quát.....	67
4.2. Kiến nghị.....	68
4.3. Hướng phát triển trong tương lai	70
TÀI LIỆU THAM KHẢO	71

DANH MỤC CHỮ VIẾT TẮT

Viết tắt	Tiếng Anh	Tiếng Việt
AI	Artificial Intelligence	Trí tuệ nhân tạo
API	Application Programming Interface	Giao diện lập trình ứng dụng
C2f	Cross-stage Partial	Kiến trúc backbone trong YOLOv8
CNN	Convolutional Neural Network	Mạng nơ-ron tích chập
DFL	Distribution Focal Loss	Hàm mất mát trong YOLOv8
FPS	Frames Per Second	Khung hình mỗi giây
GPU	Graphics Processing Unit	Bộ xử lý đồ họa
ILO	International Labour Organization	Tổ chức Lao động Quốc tế
IoU	Intersection over Union	Độ trùng lặp giao thoa
ISO	International Organization for Standardization	Tổ chức Tiêu chuẩn hóa Quốc tế
JSON	JavaScript Object Notation	Định dạng dữ liệu trả về từ API
mAP	Mean Average Precision	Độ chính xác trung bình
O&M	Operations and Maintenance	Quản lý vận hành và bảo trì
ONNX	Open Neural Network Exchange	Định dạng mô hình học sâu
OSHA	Occupational Safety and Health Administration	Cơ quan An toàn và Sức khỏe Nghề nghiệp
PPE	Personal Protective Equipment	Thiết bị bảo hộ cá nhân
RTSP	Real-Time Streaming Protocol	Giao thức truyền dữ liệu thời gian thực
SH17	SH17 Dataset	Bộ dữ liệu huấn luyện và kiểm thử nhận diện PPE
SPPF	Spatial Pyramid Pooling Fast	Tên Mô-đun trong YOLOv8
TensorRT	Tensor Runtime	Thư viện tối ưu hóa suy luận của NVIDIA
YOLO	You Only Look Once	Mô hình nhận diện đối tượng

DANH MỤC BẢNG

Bảng 1.1: Số liệu thống kê tai nạn lao động hàng năm tính tới 2019 theo ILO và OSHA.....	13
Bảng 1.2: Tổng hợp thực trạng vấn đề và mục tiêu cho các giải pháp.....	24
Bảng 2.1: So sánh tổng quan các mô hình YOLOv8, Faster R-CNN, SSD.....	28
Bảng 2.2: Cải tiến của mô hình YOLOv8	29
Bảng 2.3: Bảng thống kê các số lượng đối tượng và số ảnh của các lớp nhãn.....	37
Bảng 2.4: Bảng tham số được cài đặt của thử nghiệm.	42
Bảng 2.5: Bảng kết quả 2 thử nghiệm trong điều kiện phức tạp và tiêu chuẩn.....	48
Bảng 3.1 : So sánh kết quả của mô hình YOLOv8m với YOLOv5, YOLOX-m và Faster R-CNN	63

DANH MỤC HÌNH ẢNH

Hình 1.1: Công nhân tại trang trại điện gió làm việc trên cao với rủi ro cao khi thiếu PPE phù hợp.....	11
Hình 1.2: Công nhân tại trang trại điện gió làm việc với những thiết bị không lò đối mặt với rủi ro cao khi thiếu PPE phù hợp.	12
Hình 1.3: Đồ bảo hộ chất lượng kém không đảm bảo an toàn.	14
Hình 1.4: Ứng dụng nhận diện hình ảnh cho hệ thống O&M.....	17
Hình 1.5 Giao diện của một hệ thống O&M tích hợp báo cáo nhận diện PPE.	19
Hình 1.6: Hạn chế của giám sát PPE thủ công tại trang trại điện gió.....	20
Hình 1.7: Hệ thống O&M thông minh tích hợp nhận diện PPE cơ bản.	22
Hình 2.1: Kiến trúc mô hình YOLOV8.....	32
Hình 2.2: Minh họa Helmet và Glasses trong điều kiện ánh sáng yếu.....	39
Hình 2.3: Minh họa Gloves và Safety-vest ở góc chụp phức tạp.....	39
Hình 2.4: Minh họa Shoes trong điều kiện rung lắc, mờ hoặc che khuất.....	40
Hình 2.5: Biểu đồ kết quả so sánh kết quả lần 2 thử nghiệm.....	48
Hình 2.6: Hình ảnh trong thử nghiệm môi trường phức tạp.	50
Hình 2.7: Hình ảnh trong thử nghiệm tiêu chuẩn.....	51
Hình 3.1: Kiến trúc tổng quan tích hợp API YOLOv8 vào hệ thống O&M có sẵn.	58
Hình 3.2: Giao diện ứng dụng web mô phỏng tích hợp hệ thống AI nhận diện đồ bảo hộ thông qua API.....	59

MỞ ĐẦU

1. Lý do chọn đề tài

Trong bối cảnh Ngành năng lượng tái tạo, đặc biệt là điện gió, đang phát triển mạnh mẽ, kéo theo nhu cầu đảm bảo an toàn lao động cho kỹ sư bảo trì tại các trang trại điện gió. Công việc bảo trì tiềm ẩn nhiều rủi ro do đặc thù môi trường làm việc, trong khi hệ thống giám sát an toàn hiện tại chưa đáp ứng được yêu cầu. Đề tài nghiên cứu phát triển hệ thống tự động nhận diện đồ bảo hộ cá nhân (PPE) nhằm giảm thiểu tai nạn và nâng cao hiệu quả quản lý an toàn. Dưới đây là các lý do cụ thể cho việc chọn đề tài này:

- **Tình hình thực tế:** Kỹ sư bảo trì tại các trang trại điện gió đóng vai trò quan trọng trong việc đảm bảo hoạt động ổn định của hệ thống turbine. Tuy nhiên, công việc này đòi hỏi làm việc ở độ cao lớn, tiếp xúc với thiết bị điện mạnh, và chịu ảnh hưởng từ điều kiện thời tiết khắc nghiệt như gió mạnh, mưa, hoặc sương mù. Những yếu tố này làm tăng nguy cơ tai nạn lao động, đặc biệt khi đồ bảo hộ cá nhân không được sử dụng đầy đủ hoặc đúng cách, đặt ra yêu cầu cấp thiết về một giải pháp giám sát hiệu quả hơn.
- **Vấn đề tồn tại:** Hiện nay, việc giám sát đồ bảo hộ cá nhân tại các trang trại điện gió chủ yếu dựa vào kiểm tra thủ công bởi cán bộ an toàn. Phương pháp này dễ dẫn đến sai sót do sự chủ quan, thiếu tập trung trong các ca làm việc dài, hoặc không thể bao quát toàn bộ khu vực làm việc. Hơn nữa, không có hệ thống cảnh báo theo thời gian thực khi kỹ sư thiếu đồ bảo hộ cá nhân, khiến nguy cơ tai nạn vẫn tồn tại. Những hạn chế này làm nổi bật nhu cầu về một giải pháp tự động hóa để đảm bảo tuân thủ an toàn liên tục.
- **Ý nghĩa khoa học:** Đề tài mang giá trị nghiên cứu trong việc ứng dụng công nghệ trí tuệ nhân tạo (AI) và nhận diện hình ảnh để phát triển hệ thống giám sát đồ bảo hộ cá nhân tự động. Bằng cách tích hợp các thuật toán học sâu, nghiên cứu không chỉ góp phần nâng cao an toàn lao động mà còn đặt nền

tăng cho việc phát triển các giải pháp công nghệ tiên tiến trong quản lý trang trại điện gió. Kết quả nghiên cứu có thể được mở rộng sang các lĩnh vực công nghiệp nguy hiểm khác.

- **Tính cấp thiết:** Với sự gia tăng số lượng trang trại điện gió trên toàn cầu, đặc biệt tại các khu vực có địa hình và thời tiết phức tạp như đảo Naju (Hàn Quốc) hoặc vùng cao nguyên Đà Lạt (Việt Nam), việc đảm bảo an toàn cho kỹ sư bảo trì trở thành ưu tiên hàng đầu. Một hệ thống giám sát đồ bảo hộ cá nhân tự động không chỉ giúp giảm thiểu tai nạn mà còn đáp ứng các tiêu chuẩn an toàn lao động nghiêm ngặt, hỗ trợ các đơn vị vận hành tuân thủ quy định quốc tế.
- **Tính khả thi:** Sự phát triển của công nghệ AI, IoT, và các thiết bị cảm biến hiện nay cho phép xây dựng hệ thống nhận diện đồ bảo hộ cá nhân với độ chính xác cao và khả năng triển khai thực tế. Camera giám sát và hạ tầng mạng nội bộ tại các trang trại điện gió có thể được tận dụng để tích hợp hệ thống, giảm chi phí triển khai. Ngoài ra, các công cụ như Python, TensorFlow, và các mô hình học sâu hiện nay cung cấp nền tảng kỹ thuật mạnh mẽ để phát triển và huấn luyện mô hình, đảm bảo tính khả thi của đề tài.

2. Tổng quan về vấn đề nghiên cứu

An toàn lao động là một lĩnh vực quan trọng trong các ngành công nghiệp nguy hiểm, và việc sử dụng đồ bảo hộ cá nhân được công nhận là giải pháp hiệu quả để giảm thiểu tai nạn. Nhiều nghiên cứu đã phát triển hệ thống giám sát đồ bảo hộ cá nhân tự động sử dụng công nghệ nhận diện hình ảnh và cảm biến, chủ yếu trong các lĩnh vực như xây dựng, khai thác dầu khí, và sản xuất công nghiệp. Tuy nhiên, ứng dụng các công nghệ này trong ngành điện gió còn rất hạn chế, mặc dù môi trường làm việc tại các trang trại điện gió đặt ra yêu cầu an toàn cao hơn do đặc thù làm việc ở độ cao và điều kiện thời tiết phức tạp.

Các hệ thống giám sát đồ bảo hộ cá nhân hiện tại trong ngành điện gió thường phụ thuộc vào kiểm tra thủ công, thiếu khả năng phát hiện tức thời các vi phạm về đồ bảo hộ cá nhân. Một số giải pháp tự động hóa đã được triển khai trong các nhà máy hoặc công trường xây dựng, nhưng chúng chưa được tối ưu cho các trang trại điện gió, nơi yêu cầu xử lý dữ liệu hình ảnh trong điều kiện ánh sáng biến đổi và rung lắc từ turbine. Hơn nữa, các hệ thống này thường chỉ tập trung vào báo cáo sau sự cố, thay vì ngăn chặn rủi ro từ giai đoạn đầu. Khoảng trống nghiên cứu này tạo cơ hội cho việc phát triển một hệ thống nhận diện đồ bảo hộ cá nhân thời gian thực, tích hợp trí tuệ nhân tạo (AI) và Internet vạn vật (IoT), nhằm nâng cao an toàn lao động và hỗ trợ quản lý hiệu quả tại các trang trại điện gió.

Đề tài tập trung nghiên cứu mô hình học sâu để phát triển hệ thống nhận diện đồ bảo hộ cá nhân tự động cho kỹ sư bảo trì. Hệ thống không chỉ phát hiện các thiết bị như mũ bảo hộ, kính bảo hộ, hoặc dây đai an toàn, mà còn cung cấp cảnh báo theo thời gian thực khi có vi phạm, từ đó ngăn chặn tai nạn trước khi xảy ra. Nghiên cứu cũng xem xét khả năng tích hợp hệ thống vào hạ tầng vận hành và bảo trì hiện có, tận dụng camera giám sát và mạng nội bộ để triển khai nhanh chóng và hiệu quả. Qua đó, đề tài góp phần giải quyết các thách thức về an toàn trong ngành điện gió và đặt nền tảng cho các ứng dụng tương tự trong các lĩnh vực công nghiệp nguy hiểm khác.

3. Mục đích nghiên cứu

- **Mục tiêu lý luận:** Nghiên cứu và xây dựng nền tảng lý thuyết cho việc phát triển một module ứng dụng sử dụng công nghệ nhận diện hình ảnh và trí tuệ nhân tạo (AI) để phát hiện tự động việc sử dụng đồ bảo hộ cá nhân của kỹ sư bảo trì trong các trang trại điện gió. Hệ thống sẽ đưa ra cảnh báo kịp thời nếu phát hiện kỹ sư thiếu trang bị đồ bảo hộ cá nhân trước khi thực hiện công việc, từ đó góp phần giảm thiểu tai nạn lao động.
- **Mục tiêu thực tiễn:** Phát triển một **module ứng dụng** tích hợp công nghệ nhận diện hình ảnh, cảm biến và trí tuệ nhân tạo (AI), có khả năng:

- **Phát hiện tự động** tình trạng sử dụng đồ bảo hộ cá nhân của kỹ sư bảo trì thông qua phân tích hình ảnh và dữ liệu từ cảm biến.
- **Cảnh báo thời gian thực** qua các thiết bị di động (như điện thoại hoặc máy tính bảng) hoặc hệ thống quản lý trung tâm nếu phát hiện sai sót về đồ bảo hộ cá nhân, giúp ngăn chặn các rủi ro trước khi xảy ra.
- Tích hợp vào hệ thống quản lý an toàn lao động hiện có tại các trang trại điện gió để hỗ trợ quản lý hiệu quả hơn.
- **Kết quả cần đạt được:**
 - Hoàn thiện module ứng dụng có khả năng phát hiện đồ bảo hộ cá nhân tự động và đưa ra cảnh báo thời gian thực.
 - Giảm thiểu nguy cơ tai nạn lao động bằng cách đảm bảo kỹ sư luôn trang bị đầy đủ đồ bảo hộ cá nhân khi tiến hành công việc.
 - Cải thiện khả năng giám sát và quản lý an toàn lao động trong các hệ thống quản lý trang trại điện gió, từ đó nâng cao hiệu quả bảo trì và vận hành.

4. Đối tượng và phạm vi nghiên cứu

- **Đối tượng nghiên cứu:**

Đối tượng nghiên cứu chính là **hệ thống phát hiện và cảnh báo đồ bảo hộ cá nhân** cho kỹ sư bảo trì trong các trang trại điện gió. Nghiên cứu sẽ tập trung vào các loại đồ bảo hộ cá nhân cần thiết cho công việc bảo trì như mũ bảo hộ, kính bảo hộ, găng tay, dây đai an toàn, và giày bảo hộ.

- **Phạm vi nghiên cứu:**

- **Không gian:** Nghiên cứu sẽ được thực hiện với môi trường như tại **cánh đồng điện gió tại Hàn Quốc hoặc Việt Nam**. Đây là những nơi nổi tiếng với nhiều trang trại điện gió và hoạt động bảo trì định kỳ, nơi yêu cầu cao về an toàn lao động.

- **Thời gian:** Nghiên cứu dự kiến sẽ diễn ra trong khoảng thời gian **2 đến 3 tháng**, bao gồm giai đoạn phát triển module ứng dụng, thử nghiệm, và đánh giá hiệu quả trong thực tế.
- Nghiên cứu cũng sẽ xem xét các yếu tố tác động từ môi trường làm việc như điều kiện thời tiết và đặc thù công việc của kỹ sư bảo trì trong lĩnh vực điện gió.

5. Phương pháp nghiên cứu

Phần này trình bày phương pháp nghiên cứu được áp dụng để phát triển và đánh giá một hệ thống tự động nhận diện và cảnh báo về đồ bảo hộ cá nhân (PPE) dành cho kỹ sư bảo trì tại các trang trại điện gió, nhằm nâng cao an toàn lao động trong môi trường làm việc khắc nghiệt. Nghiên cứu sử dụng một cách tiếp cận tổng hợp, bao gồm phân tích tài liệu, thu thập dữ liệu hình ảnh, phát triển hệ thống dựa trên trí tuệ nhân tạo (AI), thí nghiệm thực địa, và sử dụng các công cụ hiện đại để hỗ trợ quá trình nghiên cứu. Dưới đây là mô tả tổng quát về các phương pháp được triển khai:

Phương pháp nghiên cứu cơ bản

- **Phân tích tài liệu:** Nghiên cứu các tài liệu và công trình trước đây về công nghệ nhận diện hình ảnh và các tiêu chuẩn an toàn lao động trong ngành điện gió, nhằm xây dựng cơ sở lý thuyết, xác định khoảng trống nghiên cứu, và định hướng phát triển hệ thống phù hợp với nhu cầu thực tế.
- **Thu thập dữ liệu hình ảnh:** Thu thập dữ liệu hình ảnh từ các trang trại điện gió, tập trung vào các tình huống kỹ sư bảo trì trong trạng thái trang bị hoặc không trang bị PPE. Khi dữ liệu thực tế không đủ, sẽ sử dụng thêm hình ảnh từ các nguồn công khai có tính tương đồng với môi trường làm việc tại trang trại điện gió.

Phát triển và tích hợp hệ thống

Phát triển module nhận diện: Xây dựng một module ứng dụng dựa trên công

nghệ nhận diện hình ảnh và AI, với mục tiêu tự động phát hiện tình trạng sử dụng PPE của kỹ sư. Quá trình này bao gồm thu thập dữ liệu, phát triển thuật toán nhận diện, huấn luyện mô hình, và tích hợp với hệ thống vận hành và bảo trì (O&M) thông qua một service để hỗ trợ kiểm tra và phát hiện vi phạm PPE theo thời gian thực.

Thí nghiệm và đánh giá

- Thí nghiệm thực địa: Triển khai hệ thống tại các trang trại điện gió để kiểm tra khả năng hoạt động trong điều kiện thực tế, đánh giá hiệu suất, độ tin cậy, và khả năng phản hồi nhanh của hệ thống trong môi trường làm việc khắc nghiệt.
- Đánh giá hiệu quả: Phân tích hiệu quả của hệ thống thông qua các chỉ số kỹ thuật và khả năng ứng dụng thực tế, đồng thời xem xét tác động của các yếu tố môi trường (như thời tiết, ánh sáng, rung lắc) đến chất lượng dữ liệu và kết quả nhận diện.

Công cụ và thiết bị hỗ trợ

- Thiết bị thu thập dữ liệu: Sử dụng các thiết bị như camera và thiết bị di động để ghi lại hình ảnh và video về tình trạng sử dụng PPE trong các điều kiện làm việc khác nhau.
- Phần mềm và công cụ phát triển: Áp dụng các công cụ lập trình và thư viện hỗ trợ để phát triển, huấn luyện mô hình nhận diện hình ảnh, đảm bảo hiệu suất và tính mở rộng của hệ thống.

Các phương pháp trên được thiết kế để tạo ra một khung nghiên cứu toàn diện, từ việc xây dựng nền tảng lý thuyết đến triển khai thực tiễn, nhằm phát triển một hệ thống không chỉ hiệu quả mà còn khả thi trong việc nâng cao an toàn lao động tại các trang trại điện gió. Những phương pháp này sẽ được triển khai chi tiết trong đề tài **"Nghiên cứu mô hình học sâu cho nhận diện đồ bảo hộ trong quản lý trang trại điện gió"**, bao gồm các chương sau:

Chương 1: Giới thiệu

Chương 2: Xây dựng và thử nghiệm hệ thống

Chương 3: Đánh giá hiệu quả

Chương 4: Kết luận và kiến nghị

CHƯƠNG 1: ĐÁNH GIÁ THỰC TRẠNG

Ngành năng lượng tái tạo, đặc biệt là điện gió, đang chứng kiến sự phát triển vượt bậc trên toàn cầu, với hàng nghìn trang trại điện gió quy mô lớn được xây dựng từ các khu vực ven biển đến cao nguyên. Sự mở rộng này không chỉ thúc đẩy nhu cầu về các phần mềm vận hành và bảo trì (O&M) tiên tiến để tối ưu hóa hiệu suất mà còn đặt ra yêu cầu cấp thiết trong việc đảm bảo an toàn lao động. Môi trường làm việc tại các trang trại điện gió, với độ cao lớn, thiết bị điện mạnh, và điều kiện thời tiết khắc nghiệt như gió mạnh, mưa, hoặc sương mù, tạo ra những thách thức đáng kể trong việc duy trì an toàn, đặc biệt liên quan đến việc sử dụng đồ bảo hộ cá nhân (Personal Protective Equipment, viết tắt PPE). Việc đảm bảo công nhân tuân thủ PPE không chỉ là một yêu cầu pháp lý mà còn là yếu tố then chốt để giảm thiểu tai nạn lao động, vốn có thể gây thiệt hại nghiêm trọng về nhân lực và kinh tế. Ý tưởng cho nghiên cứu này bắt nguồn từ những trải nghiệm thực nghiệm và làm việc của người thực hiện tại các đơn vị vận hành điện gió, như KEPCO tại đảo Naju (Hàn Quốc) và Công ty điện Cầu Đất tại Đà Lạt (Việt Nam), nơi các thách thức trong giám sát an toàn đã truyền cảm hứng cho việc phát triển giải pháp công nghệ mới. Chương này nhằm phân tích thực trạng sử dụng PPE trong ngành điện gió trên phạm vi toàn cầu, làm rõ những hạn chế của các phương pháp giám sát hiện tại, và đề xuất định hướng tích hợp công nghệ nhận diện PPE vào hệ thống O&M, hướng tới một giải pháp giám sát tự động, hiệu quả, và phù hợp với thực tế. Để đạt được mục tiêu này, nghiên cứu dựa trên việc khảo sát tài liệu quốc tế, phân tích thực trạng ngành, và tham khảo nhu cầu từ các đơn vị vận hành điện gió, từ đó đặt nền tảng cho các giải pháp công nghệ được trình bày trong các chương tiếp theo.

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

Sự phát triển nhanh chóng của ngành điện gió trên toàn cầu đã làm nổi bật tầm quan trọng của an toàn lao động, đặc biệt là việc sử dụng thiết bị bảo hộ cá nhân (PPE) nhằm giảm thiểu rủi ro trong môi trường làm việc khắc nghiệt. Tuy nhiên, thực tế cho thấy việc tuân thủ và giám sát PPE vẫn còn nhiều hạn chế, dẫn đến nguy cơ tai nạn cao tại các công trường. Theo Cơ quan An toàn và Sức khỏe Nghề nghiệp Hoa Kỳ (OSHA), công nhân ngành năng lượng gió thường đối mặt với các mối nguy như ngã từ độ cao, giật điện và chấn thương do thiết bị, trong đó việc không sử dụng PPE đúng cách là nguyên nhân phổ biến gây tai nạn nghiêm trọng [7]. Các PPE như mũ bảo hộ, dây đai an toàn, giày chống trượt, kính bảo hộ, và áo phản quang đóng vai trò quan trọng trong việc bảo vệ công nhân, nhưng việc sử dụng không đầy đủ hoặc không đúng cách làm gia tăng rủi ro, đặc biệt trong môi trường độ cao và thời tiết bất lợi của trang trại điện gió.



Hình 1.1: Công nhân tại trang trại điện gió làm việc trên cao với rủi ro cao khi thiếu PPE phù hợp.



Hình 1.2: Công nhân tại trang trại điện gió làm việc với những thiết bị khổng lồ đối mặt với rủi ro cao khi thiếu PPE phù hợp.

Theo báo cáo của Tổ chức Lao động Quốc tế (ILO), mỗi năm có khoảng 2,78 triệu người chết do tai nạn lao động hoặc bệnh nghề nghiệp, và 374 triệu ca chấn thương không gây tử vong liên quan đến công việc, gây tổn thất kinh tế toàn cầu ước tính tương đương 4% GDP hàng năm [6]. Mặc dù không có số liệu cụ thể cho riêng ngành điện gió, nhưng các thống kê này phản ánh rõ mức độ nghiêm trọng của tình trạng thiếu an toàn lao động. Trong ngành điện gió, các tai nạn liên quan đến PPE thường xảy ra do công nhân không đeo dây đai an toàn khi làm việc ở độ cao, bỏ qua kính bảo hộ khi thao tác với thiết bị điện, hoặc sử dụng mũ bảo hộ không đạt chuẩn. Những vi phạm này không chỉ gây nguy hiểm trực tiếp mà còn dẫn đến các hình phạt pháp lý và làm gián đoạn hoạt động của các dự án điện gió quy mô lớn, ảnh hưởng đến uy tín và chi phí vận hành của doanh nghiệp.

Chỉ số	Giá trị
Số ca tử vong do tai nạn lao động	2,78 triệu ca/năm
Số ca chấn thương không tử vong	374 triệu ca/năm
Tổn thất kinh tế toàn cầu	4% GDP toàn cầu
Tỷ lệ tai nạn liên quan đến PPE	Khoảng 30% tai nạn ngành năng lượng

Bảng 1.1: Số liệu thống kê tai nạn lao động hàng năm tính tới 2019 theo ILO và OSHA

Việc giám sát tuân thủ PPE hiện nay chủ yếu dựa vào kiểm tra thủ công, vốn không thể đảm bảo tính liên tục và toàn diện. Tại nhiều trang trại điện gió, đặc biệt ở các khu vực xa xôi như turbine ngoài khơi hoặc trên đồi cao, cán bộ an toàn khó có thể kiểm tra từng công nhân một cách thường xuyên. Khảo sát thực địa cho thấy một số công nhân, đặc biệt là lao động thời vụ, thường bỏ qua PPE do bất tiện, thiếu nhận thức, hoặc do thói quen làm việc đối phó. Ví dụ, dây đai an toàn có thể bị bỏ qua vì gây khó chịu khi di chuyển, hoặc giày chống trượt không được sử dụng do thiếu kiểm tra trước ca làm việc. Điều này làm tăng nguy cơ tai nạn, đặc biệt tại các vị trí turbine gió ở độ cao trên 100 mét, nơi gió mạnh, sương mù, hoặc mưa lớn làm gia tăng rủi ro trượt ngã hoặc va chạm.

Chất lượng PPE cũng là một vấn đề đáng lo ngại, ảnh hưởng trực tiếp đến hiệu quả bảo vệ. Tại nhiều trang trại điện gió, đặc biệt ở các khu vực có ngân sách hạn chế, việc sử dụng PPE không đạt chuẩn là phổ biến. Ví dụ, mũ bảo hộ bị nứt vỡ, dây đai an toàn kém chất lượng không chịu được lực kéo lớn, hoặc kính bảo hộ không chống tia UV đúng tiêu chuẩn thường được ghi nhận. Những thiết bị này không thể cung cấp mức bảo vệ cần thiết trong điều kiện khắc nghiệt như gió mạnh (tốc độ gió trung bình 6-8 m/s tại các trang trại điện gió), mưa lớn, hoặc nhiệt độ thấp dưới 5°C. Hậu quả là công nhân không chỉ đối mặt với nguy cơ tai nạn mà còn mất niềm tin vào PPE, dẫn đến tâm lý xem nhẹ việc sử dụng các thiết bị này.



Hình 1.3: Đồ bảo hộ chất lượng kém không đảm bảo an toàn.

Nguyên nhân chính của các hạn chế này bao gồm nhận thức an toàn chưa cao, thiếu chương trình đào tạo đầy đủ, và khó khăn trong giám sát tại các môi trường đặc thù. Nhiều công nhân mới, đặc biệt là lao động thời vụ, không được huấn luyện kỹ lưỡng về cách sử dụng PPE đúng cách, chẳng hạn như cách kiểm tra độ bền của dây đai an toàn hoặc lựa chọn giày chống trượt phù hợp với địa hình. Đồng thời, môi trường làm việc tại các trang trại điện gió, với đặc điểm địa lý phân tán, độ cao lớn, và thời tiết thất thường, khiến việc giám sát thủ công trở nên kém hiệu quả. Các cán bộ an toàn thường chỉ kiểm tra tại các điểm tập trung mà không thể giám sát liên tục tại các turbine cách xa hàng kilomet.

Hệ thống giám sát thủ công hiện nay chịu ảnh hưởng lớn từ yếu tố con người và môi trường. Cán bộ an toàn, dù tận tâm, thường khó duy trì sự tập trung trong các ca làm việc dài, đặc biệt trong điều kiện thời tiết khắc nghiệt như nắng gắt (nhiệt độ trên 35°C), gió mạnh, hoặc mưa lớn. Tính chủ quan, mệt mỏi, hoặc thiếu sót trong quan sát có thể dẫn đến việc bỏ qua các vi phạm PPE, chẳng hạn như công nhân không đeo áo phản quang khi làm việc vào ban đêm hoặc không sử dụng găng tay

chống điện khi kiểm tra hệ thống điện. Hơn nữa, việc kiểm tra thủ công không thể bao quát toàn bộ khu vực làm việc rộng lớn của các trang trại điện gió, đặc biệt tại các turbine ngoài khơi hoặc trên địa hình đồi núi phức tạp.

Một vấn đề khác là thiếu các công cụ hỗ trợ giám sát tự động, dẫn đến việc phát hiện vi phạm PPE không kịp thời. Ví dụ, trong các ca làm việc ban đêm, việc thiếu áo phản quang hoặc đèn gắn mũ bảo hộ có thể không được phát hiện ngay lập tức, làm tăng nguy cơ tai nạn do tầm nhìn hạn chế. Các hệ thống camera giám sát hiện tại, nếu có, thường chỉ ghi hình mà không tích hợp khả năng phân tích tự động, khiến việc kiểm tra PPE vẫn phụ thuộc vào nhân lực. Điều này không chỉ làm giảm hiệu quả quản lý an toàn mà còn gây áp lực lớn lên đội ngũ giám sát, đặc biệt trong các dự án điện gió quy mô lớn với hàng chục hoặc hàng trăm turbine.

Cuối cùng, các quy định an toàn lao động quốc tế như ISO 45001 và các tiêu chuẩn địa phương yêu cầu sử dụng PPE đầy đủ và giám sát chặt chẽ, nhưng việc triển khai còn gặp khó khăn do thiếu nguồn lực và công nghệ. Nhiều trang trại điện gió, đặc biệt ở các nước đang phát triển, chưa đầu tư vào các hệ thống giám sát thông minh hoặc đào tạo chuyên sâu cho công nhân. Điều này dẫn đến việc tuân thủ PPE chỉ mang tính hình thức, không đảm bảo an toàn thực tế. Cần một giải pháp giám sát PPE tự động, sử dụng công nghệ như trí tuệ nhân tạo và camera thông minh, để đảm bảo phát hiện vi phạm liên tục, chính xác, và hiệu quả trong môi trường khắc nghiệt của ngành điện gió.

1.2. Nhu cầu giải quyết bài toán nhận diện PPE

Sự phát triển mạnh mẽ của ngành điện gió trên toàn cầu, với hàng loạt trang trại mới được xây dựng tại các khu vực từ đất liền đến ngoài khơi, đã đặt việc đảm bảo an toàn lao động, đặc biệt là giám sát thiết bị bảo hộ cá nhân (PPE), vào vị trí ưu tiên hàng đầu. Những hạn chế của các phương pháp giám sát thủ công, như sai sót do yếu tố con người, thiếu khả năng bao quát toàn diện, và khó khăn trong điều kiện thời tiết khắc nghiệt như gió mạnh, mưa lớn, hay sương mù, cho thấy sự cần thiết của một giải pháp công nghệ tiên tiến. Một hệ thống nhận diện PPE tự động, dựa trên công

nghệ thị giác máy tính và trí tuệ nhân tạo, được xem là giải pháp tiềm năng để giám sát liên tục và phát hiện tức thời các trường hợp công nhân không trang bị đầy đủ PPE, chẳng hạn như thiếu mũ bảo hộ, kính bảo hộ, dây đai an toàn, giày chống trượt, hoặc áo phản quang. Hệ thống này vượt qua các hạn chế của giám sát thủ công bằng cách cung cấp cảnh báo kịp thời đến quản lý tại chỗ hoặc trung tâm điều hành từ xa, giúp ngăn chặn rủi ro tai nạn trước khi xảy ra. Điều này không chỉ nâng cao hiệu quả quản lý an toàn mà còn thúc đẩy ý thức tuân thủ PPE của công nhân, khi họ nhận thức rằng hành vi của mình được theo dõi theo thời gian thực, từ đó giảm thiểu các vi phạm do chủ quan hoặc thiếu hiểu biết.

Hệ thống nhận diện PPE tự động mang lại lợi ích vượt trội nhờ khả năng tích hợp với các hệ thống vận hành và bảo trì (O&M) hiện có tại các trang trại điện gió. Nhiều trang trại đã được trang bị hạ tầng camera giám sát độ phân giải cao và mạng nội bộ để theo dõi hoạt động turbine, điều kiện thời tiết, hoặc các thông số kỹ thuật khác. Việc bổ sung tính năng nhận diện PPE cho phép tận dụng các camera này để thu thập dữ liệu hình ảnh, phân tích tình trạng PPE của công nhân, và kết nối thông tin này với các dữ liệu vận hành khác, chẳng hạn như tốc độ gió, nhiệt độ, hoặc lịch sử sự cố. Ví dụ, hệ thống có thể phát hiện trường hợp công nhân không đeo dây đai an toàn khi làm việc ở độ cao trong điều kiện gió mạnh (>8 m/s), từ đó gửi cảnh báo tức thời. Kết quả là, hệ thống không chỉ cung cấp thông tin theo thời gian thực mà còn tạo ra các báo cáo tuân thủ an toàn định kỳ, hỗ trợ công tác kiểm tra, đánh giá, và huấn luyện nhân viên. Việc tích hợp này tối ưu hóa chi phí triển khai bằng cách sử dụng hạ tầng sẵn có, đồng thời nâng cao khả năng phân tích toàn diện các nguyên nhân gây tai nạn, từ đó cải thiện quy trình quản lý an toàn và giảm thiểu rủi ro trong môi trường khắc nghiệt của trang trại điện gió.



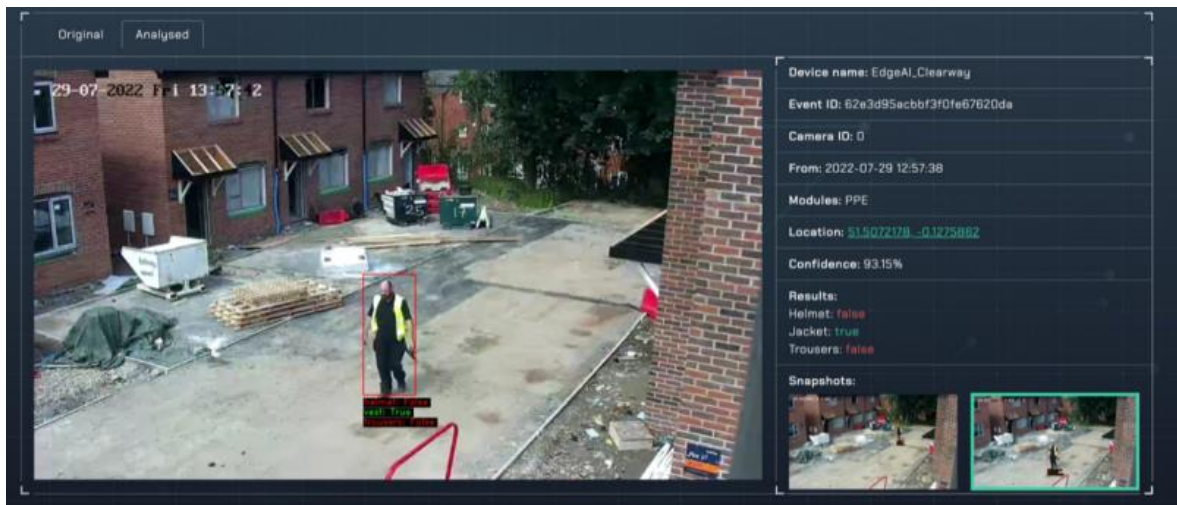
Hình 1.4: Ứng dụng nhận diện hình ảnh cho hệ thống O&M

Hệ thống nhận diện PPE hiệu quả còn mang lại tác động tích cực đến năng suất vận hành của các trang trại điện gió. Theo kết quả điều tra hiện trường, một tai nạn lao động, dù nhỏ, có thể gây gián đoạn trung bình 10 giờ, bao gồm thời gian xử lý sự cố, điều tra nguyên nhân, và khắc phục hậu quả. Những gián đoạn này không chỉ làm giảm hiệu suất hoạt động của turbine mà còn gây tổn thất kinh tế đáng kể, bao gồm chi phí bồi thường và thời gian ngừng sản xuất. Bằng cách đảm bảo công nhân luôn tuân thủ PPE, chẳng hạn như đeo mũ bảo hộ khi làm việc gần turbine hoặc sử dụng giày chống trượt trên địa hình ẩm ướt, hệ thống giúp giảm thiểu các sự cố này, từ đó duy trì hoạt động liên tục và tăng hiệu quả kinh tế. Hơn nữa, việc giảm thiểu tai nạn lao động còn góp phần nâng cao uy tín của các đơn vị vận hành, khi họ có thể đáp ứng tốt hơn các yêu cầu pháp lý và tiêu chuẩn an toàn quốc tế như ISO 45001, từ đó tránh các hình phạt pháp lý hoặc tổn thất danh tiếng.

Một lợi ích quan trọng khác của hệ thống nhận diện PPE là khả năng cải thiện

đào tạo và nâng cao nhận thức an toàn của công nhân. Hệ thống có thể ghi nhận các trường hợp vi phạm PPE, chẳng hạn như công nhân không sử dụng kính bảo hộ khi thao tác với thiết bị điện, và cung cấp dữ liệu chi tiết để xây dựng các chương trình đào tạo nhằm mục tiêu. Các báo cáo phân tích từ hệ thống có thể chỉ ra các loại PPE thường bị bỏ qua, thời điểm hoặc khu vực xảy ra vi phạm, từ đó giúp quản lý thiết kế các khóa huấn luyện hiệu quả hơn. Ví dụ, nếu dữ liệu cho thấy công nhân thường không đeo áo phản quang trong ca đêm, doanh nghiệp có thể tăng cường đào tạo về tầm quan trọng của PPE trong điều kiện ánh sáng yếu. Điều này không chỉ cải thiện ý thức tuân thủ mà còn xây dựng văn hóa an toàn bền vững trong lực lượng lao động.

Hệ thống nhận diện PPE tự động còn có tiềm năng mở rộng ứng dụng để hỗ trợ các khía cạnh khác của quản lý an toàn. Ngoài việc phát hiện thiếu PPE, hệ thống có thể được nâng cấp để nhận diện các hành vi nguy hiểm, như công nhân leo turbine mà không sử dụng dây đai an toàn hoặc làm việc gần thiết bị điện mà không đeo găng tay chống điện. Kết hợp với các cảm biến IoT, hệ thống có thể phân tích dữ liệu môi trường (như tốc độ gió hoặc độ rung của turbine) để đưa ra cảnh báo theo ngữ cảnh, ví dụ, yêu cầu công nhân sử dụng PPE bổ sung trong điều kiện thời tiết bất lợi. Tính năng này đặc biệt quan trọng tại các trang trại điện gió ngoài khơi, nơi điều kiện môi trường thay đổi nhanh chóng và khó dự đoán, làm tăng nguy cơ tai nạn nếu không có biện pháp giám sát hiệu quả.



Hình 1.5 Giao diện của một hệ thống O&M tích hợp báo cáo nhận diện PPE.

Cuối cùng, nhu cầu phát triển hệ thống nhận diện PPE tự động không chỉ xuất phát từ những hạn chế của giám sát thủ công mà còn từ áp lực ngày càng tăng của các quy định an toàn và kỳ vọng của thị trường. Các nhà đầu tư và đối tác trong ngành điện gió ngày càng ưu tiên các doanh nghiệp có hồ sơ an toàn tốt, đáp ứng các tiêu chuẩn quốc tế và giảm thiểu rủi ro vận hành. Một hệ thống nhận diện PPE tự động không chỉ giúp các trang trại điện gió đạt được các tiêu chuẩn này mà còn tạo ra lợi thế cạnh tranh bằng cách giảm chi phí vận hành, nâng cao hiệu suất, và xây dựng hình ảnh doanh nghiệp bền vững. Tóm lại, việc triển khai hệ thống nhận diện PPE tự động là một bước tiến cần thiết để đảm bảo an toàn, tối ưu hóa hoạt động, và đáp ứng các yêu cầu ngày càng khắt khe của ngành điện gió trong bối cảnh phát triển toàn cầu.

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

Hệ thống vận hành và bảo trì (O&M) tại các trang trại điện gió hiện nay đóng vai trò cốt lõi trong việc đảm bảo hiệu suất hoạt động của các turbine và giám sát các thông số môi trường như tốc độ gió, nhiệt độ, và độ rung. Tuy nhiên, các hệ thống này chủ yếu tập trung vào quản lý kỹ thuật, thiếu chức năng giám sát an toàn lao động, đặc biệt là việc tuân thủ sử dụng thiết bị bảo hộ cá nhân (PPE) như mũ bảo hộ, dây đai an toàn, kính bảo hộ, giày chống trượt, và áo phản quang. Việc kiểm tra PPE

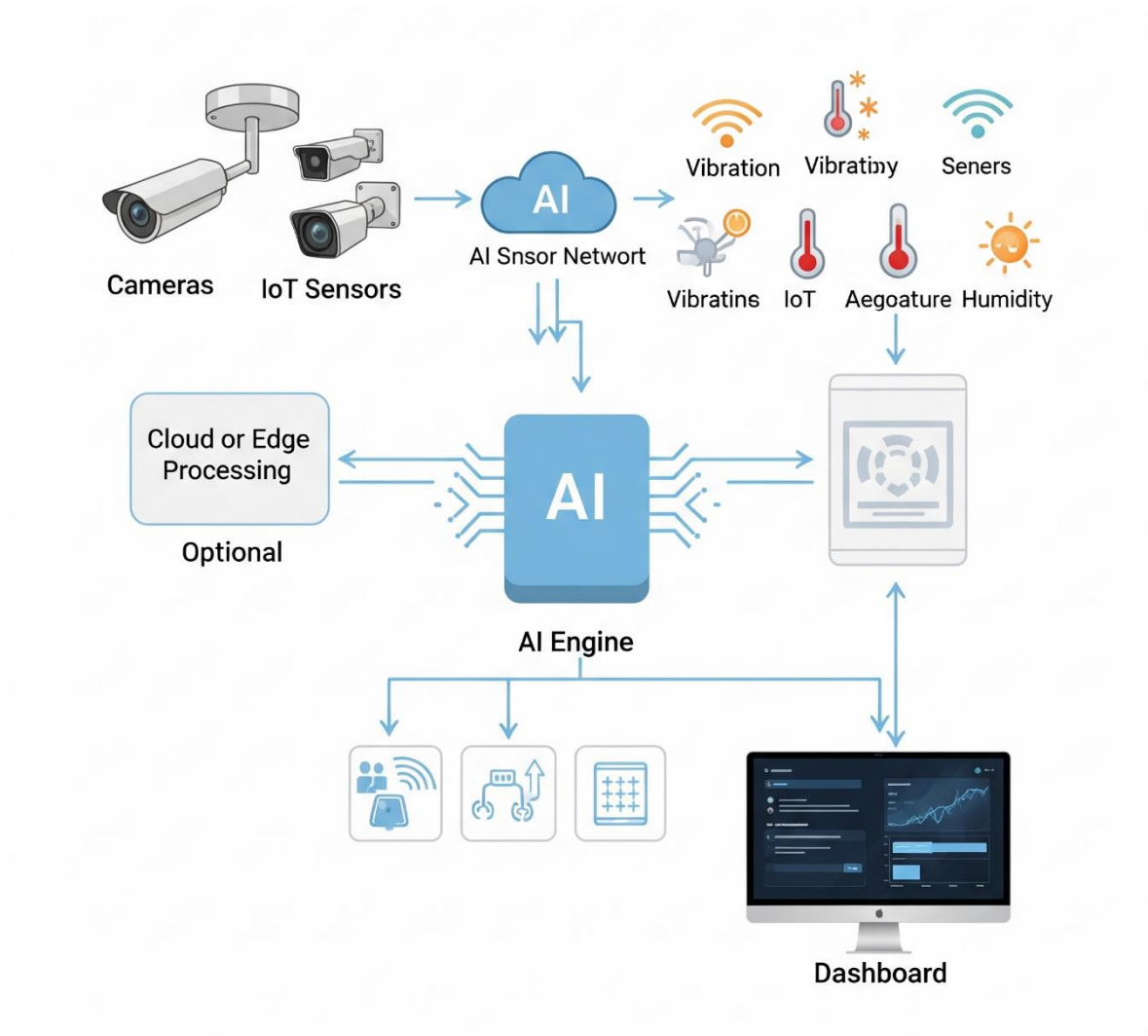
hiện vẫn phụ thuộc vào phương pháp thủ công, vốn tồn tại nhiều hạn chế như thiếu khả năng bao quát toàn diện, dễ xảy ra sai sót do yếu tố con người, và kém hiệu quả trong các điều kiện thời tiết khắc nghiệt như gió mạnh, mưa lớn, hoặc sương mù. Những hạn chế này không chỉ làm tăng nguy cơ tai nạn lao động mà còn cản trở việc tích hợp quản lý an toàn vào quy trình vận hành, dẫn đến gián đoạn hoạt động và tổn thất kinh tế.



Hình 1.6: Hạn chế của giám sát PPE thủ công tại trang trại điện gió.

Hệ thống O&M hiện tại thường sử dụng camera giám sát để ghi hình turbine và khu vực làm việc, nhưng các camera này chủ yếu phục vụ mục đích kỹ thuật như phát hiện hư hỏng thiết bị hoặc theo dõi điều kiện môi trường. Chức năng phân tích hình ảnh để nhận diện PPE hoặc hành vi vi phạm an toàn gần như chưa được triển khai. Theo khảo sát thực địa, việc kiểm tra thủ công tại các trang trại điện gió, đặc biệt ở các turbine ngoài khơi hoặc trên địa hình đồi núi, mất trung bình 15 phút mỗi ca và chỉ bao phủ được 60-70% khu vực làm việc. Trong điều kiện thời tiết bất lợi, hiệu quả kiểm tra có thể giảm xuống dưới 50%, làm gia tăng nguy cơ bỏ qua các vi phạm PPE như công nhân không đeo dây đai an toàn khi leo turbine hoặc thiếu kính bảo hộ khi thao tác với hệ thống điện. Những thiếu sót này nhấn mạnh nhu cầu cấp thiết về một giải pháp tự động hóa để nâng cao hiệu quả giám sát an toàn.

Để khắc phục các hạn chế trên, nghiên cứu đề xuất tích hợp một mô-đun trí tuệ nhân tạo (AI) nhận diện PPE vào hệ thống O&M hiện có. Mô-đun này tận dụng hạ tầng camera sẵn có tại các trang trại điện gió để tự động phát hiện các vi phạm PPE, chẳng hạn như thiếu mũ bảo hộ, dây đai an toàn, hoặc áo phản quang, và gửi cảnh báo theo thời gian thực qua push notification, email, hoặc hiển thị trên bảng điều khiển O&M. Mô-đun sử dụng các thuật toán thị giác máy tính tiên tiến, như mô hình YOLO, để phân tích hình ảnh với tốc độ xử lý nhanh (dự kiến 40 khung hình/giây) và độ chính xác cao (mAP@0.5 trên 85% trong điều kiện tiêu chuẩn). Theo báo cáo của OSHA [7], các giải pháp tự động hóa giám sát an toàn có thể giảm 30% chi phí liên quan đến tai nạn lao động, bao gồm chi phí bồi thường, điều trị y tế, và thời gian gián đoạn vận hành, đồng thời tăng 15% hiệu quả quản lý an toàn nhờ giảm thiểu sai sót do con người.



Hình 1.7: Hệ thống O&M thông minh tích hợp nhận diện PPE cơ bản.

Mô-đun AI không chỉ dừng lại ở việc phát hiện PPE mà còn cung cấp các báo cáo chi tiết về tình trạng tuân thủ an toàn, hỗ trợ quản lý trong việc đánh giá hiệu quả PPE, lập kế hoạch đào tạo, và thực hiện các đợt kiểm tra định kỳ. Ví dụ, báo cáo có thể chỉ ra tỷ lệ vi phạm PPE theo khu vực, ca làm việc, hoặc loại PPE (như giày chống trượt bị bỏ qua trong ca đêm), từ đó giúp tối ưu hóa quy trình huấn luyện và nâng cao nhận thức an toàn của công nhân. Hơn nữa, mô-đun này có khả năng tích hợp với các dữ liệu O&M khác, như thông số thời tiết hoặc lịch sử sự cố, để đưa ra các cảnh báo theo ngữ cảnh. Chẳng hạn, hệ thống có thể yêu cầu công nhân sử dụng PPE bổ sung, như găng tay chống điện, khi phát hiện nguy cơ giật điện trong điều kiện ẩm ướt.

Về mặt kinh tế, việc triển khai mô-đun AI nhận diện PPE đòi hỏi đầu tư ban đầu cho phần cứng (như GPU để xử lý hình ảnh), phần mềm, và đào tạo nhân sự vận hành. Tuy nhiên, lợi ích dài hạn là đáng kể, bao gồm tiết kiệm chi phí bồi thường tai nạn, giảm thời gian gián đoạn vận hành, và tăng uy tín doanh nghiệp nhờ tuân thủ các tiêu chuẩn an toàn quốc tế như ISO 45001. Ước tính, chi phí vận hành một hệ thống AI như vậy chỉ bằng 20-30% chi phí kiểm tra thủ công trong vòng 3 năm, nhờ giảm nhu cầu nhân lực giám sát và tối ưu hóa quy trình. Ngoài ra, mô-đun AI có tiềm năng mở rộng để phát hiện các hành vi nguy hiểm khác, như leo turbine không đúng quy trình hoặc làm việc gần thiết bị điện mà không có PPE phù hợp, từ đó nâng cao mức độ an toàn toàn diện.

Định hướng phát triển dài hạn của hệ thống O&M là xây dựng một nền tảng toàn diện, kết hợp giám sát thiết bị, môi trường, và con người trong một giao diện thống nhất. Mô-đun AI nhận diện PPE có thể được nâng cấp để dự đoán sự cố an toàn dựa trên dữ liệu lịch sử và điều kiện thời gian thực, chẳng hạn như cảnh báo nguy cơ tai nạn khi tốc độ gió vượt quá 10 m/s mà công nhân không đeo dây đai an toàn. Khả năng học liên tục (continual learning) của mô hình AI đảm bảo hệ thống luôn thích nghi với các điều kiện mới, như sự thay đổi về thiết kế PPE hoặc môi trường làm việc tại các trang trại ngoài khơi. Hơn nữa, việc triển khai mô-đun này trên các thiết bị nhúng, như Jetson Nano, cho phép áp dụng tại các khu vực thiếu hạ tầng mạng, mở rộng khả năng ứng dụng ở các trang trại điện gió vùng sâu vùng xa.

Hệ thống O&M tích hợp AI còn có tiềm năng kết nối với các cảm biến IoT để thu thập dữ liệu môi trường bổ sung, như độ rung của turbine, độ ẩm, hoặc áp suất khí quyển, từ đó tối ưu hóa các quyết định an toàn. Ví dụ, hệ thống có thể tự động tạm dừng công việc bảo trì nếu phát hiện công nhân thiếu mũ bảo hộ trong điều kiện rung lắc mạnh do turbine hoạt động. Tầm nhìn dài hạn là xây dựng một hệ sinh thái O&M thông minh, nơi dữ liệu từ PPE, thiết bị, và môi trường được phân tích tổng hợp để đưa ra các khuyến nghị vận hành tối ưu, góp phần đảm bảo an toàn, nâng cao hiệu quả kinh tế, và thúc đẩy sự phát triển bền vững của ngành điện gió.

Vấn đề	Thực trạng	Nhu cầu giải quyết
Tuân thủ PPE	Tỷ lệ vi phạm cao, mang tính đối phó	Hệ thống tự động phát hiện vi phạm PPE
Giám sát thủ công	Bao phủ 60–70%, dễ xảy ra sai sót	Giám sát liên tục, chính xác khoảng 95%
Chi phí tai nạn lao động	Gây tổn thất kinh tế lớn (~4% GDP toàn cầu)	Giảm 30% chi phí nhờ tự động hóa giám sát an toàn
Hệ thống O&M	Thiếu năng lực giám sát an toàn lao động	Tích hợp AI nhận diện và kiểm tra việc tuân thủ PPE

Bảng 1.2: Tổng hợp thực trạng vấn đề và mục tiêu cho các giải pháp.

1.4. Định nghĩa bài toán

Dữ liệu đầu vào: Hệ thống nhận các hình ảnh và video thời gian thực từ camera giám sát tại các lối lên turbine, khu vực bảo trì, hoặc các điểm làm việc trong trang trại điện gió. Dữ liệu bao gồm 8.099 ảnh từ dataset SH17, với độ phân giải từ 720p đến 1080p, ghi lại các kỹ sư trong điều kiện ánh sáng yếu, sương mù, gió mạnh, và rung lắc, cùng các góc chụp đa dạng (0° - 90°). Dữ liệu thực tế từ các đối tác như Công ty điện Cầu Đất và KEPCO sẽ được bổ sung trong tương lai để tăng tính đại diện.

Kết quả mong muốn: Hệ thống tự động phát hiện và phân loại 17 loại PPE (như mũ bảo hộ, dây đai an toàn, găng tay, giày chống trượt) với độ chính xác $mAP@0.5 \geq 85\%$ trong điều kiện tiêu chuẩn và $\geq 70\%$ trong điều kiện phức tạp. Kết quả bao gồm thông tin trạng thái PPE (có/không), cảnh báo thời gian thực qua push notification hoặc email khi phát hiện vi phạm, và báo cáo hàng ngày/tuần lưu trữ trong hệ thống O&M với độ chính xác $\geq 95\%$ so với kiểm tra thủ công.

Hướng tiếp cận: Sử dụng mô hình YOLOv8m huấn luyện trên GPU Nvidia RTX 3080, tích hợp qua API REST để xử lý 40 khung hình/giây, kết hợp với các kỹ thuật augmentation (nhiều Gaussian, tăng độ tương phản) và bộ lọc Kalman để thích nghi với điều kiện thực tế.

Chương 1 đã phân tích thực trạng sử dụng PPE tại các trang trại điện gió, chỉ ra những hạn chế trong giám sát thủ công, chất lượng PPE, và ý thức tuân thủ, dẫn đến nguy cơ tai nạn lao động cao. Nhu cầu cấp thiết về một hệ thống nhận diện PPE tự động đã được làm rõ, với tiềm năng cải thiện an toàn, hiệu quả vận hành, và tuân thủ các tiêu chuẩn quốc tế như ISO 45001. Hệ thống O&M hiện tại, dù hiệu quả trong giám sát kỹ thuật, vẫn thiếu chức năng quản lý an toàn, đặt ra yêu cầu tích hợp các giải pháp công nghệ tiên tiến. Chương 2 sẽ trình bày các giải pháp công nghệ và thực nghiệm, tập trung vào việc ứng dụng thuật toán học sâu như YOLOv8 để nhận diện PPE, thiết kế hệ thống tích hợp vào hạ tầng O&M, và đánh giá hiệu quả thực nghiệm tại các trang trại điện gió, hướng tới một giải pháp an toàn toàn diện và bền vững.

CHƯƠNG 2: XÂY DỰNG GIẢI PHÁP VÀ THỰC NGHIỆM

Ngành điện gió toàn cầu, với sự phát triển nhanh chóng của các trang trại quy mô lớn, đang đặt ra những yêu cầu cấp thiết về an toàn lao động, đặc biệt trong việc giám sát việc sử dụng đồ bảo hộ cá nhân (PPE), như đã phân tích trong Chương 1. Hạn chế của các phương pháp giám sát thủ công, từ sai sót do yếu tố con người đến khó khăn trong điều kiện môi trường khắc nghiệt, đã làm nổi bật nhu cầu về một hệ thống nhận diện PPE tự động, có khả năng tích hợp vào các hệ thống vận hành và bảo trì (O&M) để nâng cao hiệu quả và an toàn. Nhu cầu này không chỉ xuất phát từ các phân tích lý thuyết mà còn từ những quan sát thực tế tại các đơn vị vận hành điện gió tiên tiến, như tại Hàn Quốc và Việt Nam, nơi các yêu cầu nghiêm ngặt về an toàn đã thúc đẩy việc ứng dụng công nghệ hiện đại để giám sát PPE. Chương này trình bày chi tiết quá trình xây dựng và thực nghiệm một hệ thống nhận diện PPE phục vụ giám sát an toàn tại các trang trại điện gió, bắt đầu từ cơ sở lý thuyết, cơ sở kỹ thuật, lựa chọn mô hình, thu thập dữ liệu, huấn luyện, kiểm thử, đến tích hợp và triển khai. Đặc biệt, việc so sánh giữa các mô hình như YOLOv5 và Faster R-CNN giúp làm rõ những ưu điểm vượt trội của YOLOv8m, hướng tới một giải pháp khả thi trong môi trường công nghiệp ngoài trời, đáp ứng nhu cầu tự động hóa giám sát trên phạm vi toàn cầu.

2.1. Cơ sở lý thuyết của mô hình học sâu

Học sâu (Deep Learning) là một lĩnh vực của trí tuệ nhân tạo (AI), phát triển dựa trên nguyên lý hoạt động của mạng nơ-ron nhân tạo (Artificial Neural Networks - ANN). Khác với các phương pháp học máy truyền thống vốn phụ thuộc vào các bước trích xuất đặc trưng thủ công, học sâu cho phép mô hình tự động học đặc trưng từ dữ liệu thông qua quá trình huấn luyện nhiều lớp (deep architectures). Khả năng này đặc biệt quan trọng trong các bài toán thị giác máy tính, nơi dữ liệu hình ảnh thường phức tạp, đa chiều và không có quy luật rõ ràng.

Một trong những nền tảng cốt lõi của học sâu trong xử lý ảnh là mạng nơ-ron tích chập (Convolutional Neural Networks - CNN). CNN được thiết kế để xử lý dữ liệu có cấu trúc lưới như ảnh, bằng cách sử dụng các lớp tích chập (convolutional

layers) để phát hiện các đặc trưng cục bộ, chẳng hạn như cạnh, góc, hoa văn... Thay vì phải xác định trước đặc trưng đầu vào, CNN học trực tiếp các bộ lọc phù hợp trong quá trình huấn luyện.

Cấu trúc điển hình của một mạng CNN bao gồm các lớp tích chập, lớp phi tuyến (như ReLU), lớp gộp (pooling layer), và các lớp fully connected. Nhờ vào cơ chế chia sẻ trọng số và vùng tiếp nhận (receptive field) tăng dần, CNN có thể học các đặc trưng từ đơn giản đến phức tạp, phù hợp với nhận diện đối tượng trong môi trường biến đổi như ánh sáng khác nhau, góc nhìn đa dạng hoặc nhiều ảnh.

Các kiến trúc CNN ban đầu như LeNet-5 (1998) và AlexNet (2012) đã chứng minh tính hiệu quả vượt trội của học sâu trong xử lý ảnh so với các phương pháp truyền thống. Các mạng hiện đại như VGGNet, GoogLeNet (Inception), và ResNet sau đó đã mở rộng độ sâu mạng hàng chục, hàng trăm lớp, giúp tăng cường khả năng biểu diễn và trích xuất đặc trưng. Các cải tiến như residual connections (trong ResNet), skip connections hoặc attention mechanism giúp giải quyết các vấn đề như gradient vanishing, mất mát thông tin và overfitting trong mạng sâu.

Từ nền tảng CNN, học sâu đã trở thành công cụ chính trong nhiều tác vụ thị giác máy tính như phân loại ảnh, phát hiện đối tượng, phân đoạn ảnh, và nhận diện khuôn mặt. Sự linh hoạt, khả năng mở rộng và hiệu suất cao của các kiến trúc học sâu đã mở đường cho việc phát triển các mô hình chuyên biệt, phục vụ hiệu quả trong các hệ thống giám sát an toàn và nhận diện đồ bảo hộ cá nhân như trong nghiên cứu này.

2.2. Cơ sở kỹ thuật và lựa chọn mô hình

Xây dựng hệ thống nhận diện PPE tự động cho trang trại điện gió đòi hỏi lựa chọn mô hình học sâu cân bằng giữa tốc độ và độ chính xác, đáp ứng giám sát thời gian thực trong môi trường khắc nghiệt (gió mạnh, sương mù, ánh sáng yếu). Việc giám sát PPE (mũ bảo hộ, kính, găng tay, dây đai an toàn, giày bảo hộ) là yếu tố then chốt để giảm rủi ro tai nạn lao động. Các mô hình như YOLO, Faster R-CNN, và SSD có ưu, nhược điểm riêng, phù hợp với các kịch bản khác nhau. Phân tích và so

sánh các mô hình này là cần thiết để chọn giải pháp tối ưu, đảm bảo hiệu quả giám sát và tích hợp vào hệ thống vận hành, bảo trì (O&M) của trang trại điện gió.

a. Đánh giá tổng quan các mô hình nhận diện vật thể phổ biến

Hệ thống nhận diện PPE tự động trong trang trại điện gió cần cân bằng tốc độ và độ chính xác cho giám sát thời gian thực. Dưới đây là bảng so sánh các mô hình YOLO, Faster R-CNN, và SSD.

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

YOLOv8 phù hợp nhất cho giám sát PPE thời gian thực tại trang trại điện gió nhờ tốc độ cao và khả năng triển khai trên thiết bị biên, dù cần tối ưu cho vật thể nhỏ. Faster R-CNN thích hợp cho hậu kiểm dữ liệu, còn SSD kém hiệu quả trong môi trường khắc nghiệt [5, tr.1-2].

b. Chọn mô hình để thực nghiệm

Dựa trên đánh giá tổng quan, YOLOv8m (You Only Look Once version 8, biến thể medium) được chọn làm mô hình thực nghiệm nhờ khả năng cân bằng vượt trội giữa tốc độ và độ chính xác. Được phát triển bởi Ultralytics, YOLOv8m loại bỏ kiến trúc anchor-based truyền thống, thay bằng cơ chế anchor-free, giảm độ phức tạp tính toán và tăng tốc độ huấn luyện. Mô hình sử dụng backbone C2f (Cross-stage Partial) để trích xuất đặc trưng hiệu quả hơn, kết hợp với Spatial Pyramid Pooling Fast (SPPF) để xử lý các vật thể ở nhiều kích thước. Với khoảng 25 triệu tham số, YOLOv8m đạt tốc độ suy luận 0.025 giây mỗi ảnh trên GPU RTX 3080, đáp ứng yêu cầu thời gian thực, đồng thời duy trì độ chính xác cao với mAP@0.5 đạt 94.2% trên tập dữ liệu PPE. YOLOv8m hỗ trợ xuất sang định dạng ONNX và TensorRT, cho phép triển khai dễ dàng trên phần cứng tầm trung, như máy tính nhúng với GPU 10GB VRAM hoặc CPU mạnh tại các trang trại điện gió.

Cải tiến	Mô tả	Lợi ích
Anchor-Free	Bỏ cơ chế anchor truyền thống	Giảm độ phức tạp, phù hợp với vật thể nhỏ
C2f backbone	Tối ưu cấu trúc CSP	Trích xuất đặc trưng hiệu quả hơn, giảm số tham số
SPPF	Tăng tốc và hiệu quả so với SPP	Nhận diện vật thể ở nhiều tỷ lệ kích thước
DFL Loss	Loss phân bố điểm tập trung	Định vị chính xác bounding box nhỏ
Hỗ trợ đa nền tảng	Export sang ONNX, TensorRT, CoreML	Triển khai linh hoạt từ server đến thiết bị nhúng
Tốc độ inference cao	~0.025 giây/hình ảnh	Thích hợp cho hệ thống giám sát thời gian thực

Bảng 2.2: Cải tiến của mô hình YOLOv8

So với YOLOv5, YOLOv8m vượt trội hơn trong nhận diện PPE nhỏ (như Gloves, Face-mask-medical) trên dataset SH17 (8.099 ảnh), nhờ kiến trúc anchor-free, giảm thời gian huấn luyện và chi phí tính toán [1]. So với Faster R-CNN, YOLOv8m nhanh hơn (~ 0.03 giây/ảnh vs. ~ 0.15 giây/ảnh [5, tr.8–9]), phù hợp giám sát thời gian thực tại turbine gió, dù Faster R-CNN chính xác hơn nhưng tốn tài nguyên [5, tr.8–9]. SSD nhanh (~ 0.05 giây/ảnh) nhưng mAP@0.5 thấp, kém hiệu quả với PPE nhỏ và ánh sáng yếu [5, tr.8–9]. YOLOv8m đạt mAP@0.5 cao hơn 2-3% so với YOLOv5 trên dataset SH17 (17 lớp PPE) [2, tr.20-22; 4, tr.653-655], là lựa chọn tối ưu cho nhận diện PPE trong môi trường khắc nghiệt, với yêu cầu phần cứng tối thiểu (camera 1080p, máy tính nhúng) [2, tr.20-22; 4, tr.653-655].

c. Mô hình YOLO

Trong bối cảnh các hệ thống giám sát an toàn ngày càng đòi hỏi khả năng xử lý thời gian thực, đặc biệt tại các môi trường công nghiệp như trang trại điện gió, việc lựa chọn mô hình nhận diện vật thể hiệu quả là thiết yếu. YOLO (You Only Look Once) – dòng mô hình one-stage ra đời từ năm 2015 – đã trở thành giải pháp phổ biến nhờ tốc độ vượt trội. Không giống các mô hình two-stage như Faster R-CNN, YOLO thực hiện đồng thời phát hiện và phân loại chỉ trong một lần truyền qua mạng, rất phù hợp cho các ứng dụng yêu cầu phản hồi nhanh.

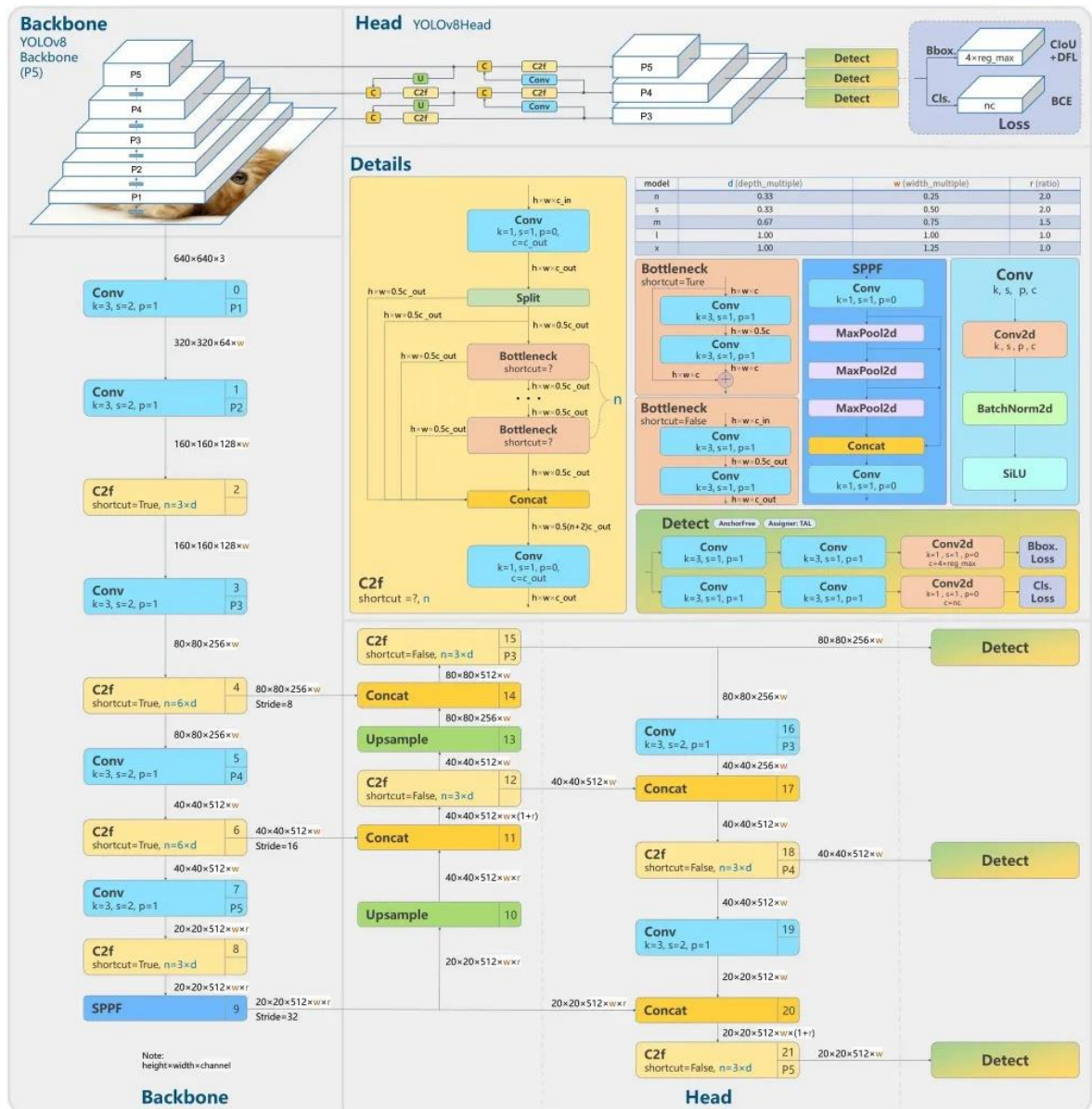
Trải qua nhiều phiên bản, từ YOLOv1 đến YOLOv5 và gần đây là YOLOv8 (2023), dòng mô hình này không ngừng được tối ưu về độ chính xác và hiệu năng. Trong số đó, YOLOv8m (medium) với khoảng 25 triệu tham số là lựa chọn cân bằng giữa độ chính xác và tốc độ, phù hợp với các ứng dụng như nhận diện thiết bị bảo hộ cá nhân (PPE) trong điều kiện thực địa. Theo Arafat và Ismail (2024), YOLOv8m cho thấy hiệu suất vượt trội trong các bài toán phát hiện vật thể, đặc biệt trong môi trường ngoài trời với điều kiện ánh sáng yếu hoặc vật thể nhỏ.

Cơ sở lý thuyết của YOLO được xây dựng trên một nguyên lý độc đáo, giúp mô hình đạt được tốc độ cao mà vẫn duy trì độ chính xác cần thiết cho các ứng dụng thực tế. YOLO chia ảnh đầu vào thành một lưới (grid), thường có kích thước $S \times S$

(ví dụ 7x7 hoặc 13x13 tùy phiên bản), trong đó mỗi ô lưới chịu trách nhiệm phát hiện các đối tượng nằm trong phạm vi của nó. Mỗi ô dự đoán một số hộp giới hạn với tọa độ (x, y, w, h) xác định vị trí và kích thước, điểm tin cậy (confidence score) biểu thị khả năng có đối tượng, và xác suất lớp (class probability) để phân loại đối tượng, chẳng hạn xác định một vật thể là mũ bảo hộ hay dây đai an toàn. Đầu ra của mô hình là một tensor 3 chiều, chứa toàn bộ thông tin dự đoán, sau đó được xử lý hậu kỳ bằng Non-Maximum Suppression (NMS) để loại bỏ các dự đoán trùng lặp, giữ lại các hộp giới hạn có điểm tin cậy cao nhất. Về mặt toán học, hàm mất mát (loss function) của YOLO là một yếu tố quan trọng quyết định hiệu quả huấn luyện. Hàm mất mát tổng quát bao gồm ba thành phần chính: Intersection over Union (IoU) để đo độ chính xác của hộp giới hạn, Binary Cross-Entropy (BCE) để phân loại lớp, và một phần hồi quy cho điểm tin cậy. Công thức loss có thể được biểu diễn như sau:

$$\text{Loss} = \lambda_{\text{coord}} \times \text{Loss}_{\text{coord}} + \text{Loss}_{\text{conf}} + \lambda_{\text{class}} \times \text{Loss}_{\text{class}} \quad (2.1)$$

Trong đó λ_{coord} và λ_{class} là các hệ số cân bằng, đảm bảo mô hình ưu tiên cả vị trí lẫn phân loại. YOLOv8 cải tiến đáng kể hàm mất mát so với các phiên bản trước bằng cách tích hợp Complete IoU (CIoU) thay cho IoU truyền thống, giúp cải thiện khả năng định vị chính xác, đặc biệt với các vật thể nhỏ hoặc bị che khuất - một thách thức lớn khi nhận diện PPE trong môi trường thực địa [1, tr.9-11][8]. Ngoài ra, YOLOv8 còn sử dụng Distribution Focal Loss (DFL) để tăng độ chính xác trong hồi quy tọa độ, tập trung vào phân bố điểm dự đoán, giúp hộp giới hạn hội tụ nhanh hơn. Sự kết hợp của DFL, BCE, và CIoU không chỉ nâng cao hiệu quả huấn luyện mà còn cải thiện đáng kể khả năng phát hiện các vật thể nhỏ như dây đai an toàn hoặc găng tay, vốn thường khó nhận diện trong điều kiện ánh sáng yếu hoặc che khuất tại các trang trại điện gió.



truyền thông, giúp giảm số lượng tham số, cải thiện tốc độ học và hiệu quả tính toán.

- *SPPF (Spatial Pyramid Pooling Fast)*: Tăng cường khả năng nhận diện vật thể đa tỷ lệ, từ các vật thể lớn như mũ bảo hộ đến các vật thể nhỏ như dây đai, bằng cách trích xuất đặc trưng không gian từ các kích thước khác nhau.
- *Conv blocks*: Chuỗi các lớp tích chập xen kẽ normalization và hàm kích hoạt SiLU, bổ sung chiều sâu cho đặc trưng, hỗ trợ phát hiện các đối tượng PPE ở nhiều kích thước.
- **Neck – tổng hợp đặc trưng:**
 - *Upsample*: Phóng đại độ phân giải, đưa đặc trưng từ tầng sâu lên độ phân giải cao hơn để giữ chi tiết.
 - *Concat*: Nối đặc trưng từ các tầng thấp và cao nhằm tạo ra một biểu diễn phong phú, kết hợp thông tin đa chiều.
 - *C2f và Conv blocks*: Đảm nhiệm việc học thêm các tầng trung gian, tăng cường khả năng tổng hợp thông tin từ các tầng đặc trưng.
- **Head – dự đoán:**
 - *Ba nhánh chính (P3, P4, P5)*: Tương ứng với các tầng đặc trưng, mỗi nhánh thực hiện dự đoán hộp giới hạn (bounding box), điểm tin cậy, và phân loại lớp đối tượng (ví dụ: mũ bảo hộ, dây đai).
 - *NMS*: Lọc các hộp trùng lặp sau quá trình dự đoán, đảm bảo kết quả cuối cùng chỉ chứa các dự đoán có độ tin cậy cao nhất.

Cơ chế hoạt động của kiến trúc YOLOv8m được thiết kế để xử lý hình ảnh một cách hiệu quả, từ đầu vào đến kết quả dự đoán, thông qua sự tương tác chặt chẽ giữa các khối chính. Quá trình bắt đầu khi hình ảnh đầu vào được đưa vào Backbone, nơi các mô-đun C2f và SPPF trích xuất đặc trưng ở nhiều mức độ chi tiết, từ các đặc

trung cấp thấp như cạnh và màu sắc đến các đặc trưng cấp cao như hình dạng của mũ bảo hộ hoặc dây đai. Các đặc trưng này sau đó được chuyển đến Neck, nơi chúng được tổng hợp thông qua các bước upsample và concat, kết hợp thông tin từ các tầng khác nhau để tạo ra một biểu diễn toàn diện, giúp mô hình nhận diện PPE ở các tỷ lệ và bối cảnh khác nhau, kể cả trong điều kiện ánh sáng yếu hoặc che khuất. Tiếp theo, Head sử dụng các đặc trưng đã tổng hợp để thực hiện dự đoán, với ba nhánh P3, P4, và P5 xử lý các vật thể ở các kích thước khác nhau – P3 cho vật thể nhỏ như găng tay, P4 và P5 cho vật thể lớn hơn như mũ bảo hộ. Sau khi dự đoán hộp giới hạn, điểm tin cậy, và lớp đối tượng, NMS được áp dụng để loại bỏ các dự đoán trùng lặp, đảm bảo kết quả cuối cùng chính xác và đáng tin cậy, phù hợp cho việc giám sát thời gian thực tại các trang trại điện gió.

YOLOv8m là kết quả của sự kết hợp giữa tốc độ xử lý cao và khả năng phát hiện vật thể nhỏ trong môi trường phức tạp, nhờ vào các cải tiến từ backbone, neck, đến head. Với hiệu suất vượt trội, tốc độ suy luận nhanh, và khả năng triển khai linh hoạt, YOLOv8m trở thành một giải pháp tối ưu cho hệ thống nhận diện PPE tại các trang trại điện gió, nơi yêu cầu giám sát liên tục và chính xác là yếu tố sống còn. Kiến trúc YOLOv8m không chỉ thể hiện sự tiến bộ trong thiết kế mô hình mà còn đại diện cho một bước tiến quan trọng trong lĩnh vực thị giác máy tính, mở ra tiềm năng ứng dụng rộng rãi trong các hệ thống giám sát an toàn hiện đại.

2.3. Thu thập và xử lý dữ liệu

Để xây dựng một hệ thống nhận diện đồ bảo hộ cá nhân (PPE) hiệu quả cho các trang trại điện gió, việc thu thập và xử lý dữ liệu đóng vai trò quan trọng, đặc biệt trong bối cảnh môi trường làm việc khắc nghiệt và yêu cầu giám sát thời gian thực. Quá trình này không chỉ đối mặt với những thách thức về nguồn dữ liệu thực tế mà còn đòi hỏi các chiến lược xử lý tối ưu để hỗ trợ mô hình YOLOv8m, được lựa chọn từ mục 2.2, đạt hiệu suất cao trong huấn luyện và triển khai sau đó.

a. Thực trạng và khó khăn khi tiếp cận dữ liệu thực tế

Ngay từ giai đoạn đầu của nghiên cứu, tôi đã chủ động liên hệ với một số đơn vị vận hành điện gió để xin phép tiếp cận dữ liệu hình ảnh thực tế phục vụ cho việc huấn luyện và kiểm thử mô hình nhận diện đồ bảo hộ cá nhân. Hai đối tác quan trọng trong lĩnh vực này là KEPCO (Naju, Hàn Quốc) và Công ty điện Cầu Đất (Đà Lạt, Việt Nam). Cả hai cơ sở này đều là những nơi đã có cơ hội hợp tác với nhóm nghiên cứu trong các dự án vận hành và bảo trì hệ thống điện gió, tạo ra những mối quan hệ bền vững và đáng tin cậy. Tuy nhiên, trong quá trình trao đổi, cả KEPCO và Cầu Đất đều cho biết rằng họ không thể chia sẻ dữ liệu hình ảnh liên quan đến hoạt động bảo trì và vận hành, chủ yếu vì các vấn đề bảo mật và chính sách nội bộ. Ngoài ra, việc chia sẻ dữ liệu liên quan đến nhân sự và hệ thống vận hành thường bị hạn chế để bảo vệ thông tin quan trọng và tuân thủ các quy định pháp lý về bảo mật. Những khó khăn này đặt ra nhu cầu tìm kiếm nguồn dữ liệu thay thế và xây dựng chiến lược xử lý phù hợp, đồng thời ảnh hưởng đến việc tối ưu hóa mô hình YOLOv8m cho các điều kiện thực tế.

b. Nguồn dữ liệu thay thế và chiến lược xử lý

Trước những rào cản về dữ liệu thực tế, nghiên cứu đã chuyển hướng sử dụng dữ liệu công khai từ dataset SH17: A Dataset for PPE Detection (tác giả Mughees Ahmad) trên Kaggle, với tổng cộng **8.099 ảnh** và **75.994 đối tượng**, thuộc 17 lớp đồ bảo hộ cá nhân như Person, Helmet, Safety-vest, Gloves,..., Medical-suit, Face-mask-medical (trung bình 9.38 đối tượng/ảnh). Bộ dữ liệu được chia thành ba tập: 6.479 ảnh huấn luyện (80%), 810 ảnh xác nhận (10%), và 810 ảnh kiểm thử (10%), đảm bảo mô hình được huấn luyện đầy đủ và đánh giá khách quan. Để tăng tính đa dạng và khả năng khái quát hóa, đặc biệt trong các điều kiện ánh sáng yếu, gió mạnh, hoặc góc chụp phức tạp tại trang trại điện gió, cũng như giải quyết thách thức từ 52% object nhỏ hơn 5% diện tích ảnh (như Gloves, Face-mask-medical) và phân bố nhãn không đồng đều (nhãn ít gặp như Medical-suit, Face-mask-medical), các kỹ thuật augmentation như lật ảnh ngang/dọc, xoay ngẫu nhiên, điều chỉnh độ sáng ($\pm 30\%$),

và thêm nhiều Gaussian đã được áp dụng. Những chiến lược này không chỉ hỗ trợ việc mô phỏng môi trường thực tế mà còn tạo tiền đề cho việc phân tích chất lượng dữ liệu và xây dựng quy trình xử lý dữ liệu hiệu quả hơn.

c. Phân tích chất lượng dữ liệu và tương thích với YOLOv8m

Nghiên cứu phân tích dữ liệu từ Kaggle và Roboflow để đảm bảo tương thích với YOLOv8m, tập trung vào độ phân giải, độ tương phản và đa dạng lớp PPE. Dữ liệu có độ phân giải trung bình 1080p, phù hợp với mô-đun SPPF trong backbone YOLOv8m, hỗ trợ trích xuất đặc trưng đa tỷ lệ cho vật thể nhỏ như dây đai an toàn, găng tay, tai và tai nghe. Khoảng 52% đối tượng trong dataset SH17 là vật thể nhỏ (<1% diện tích ảnh), cải thiện khả năng phát hiện trong môi trường thực tế. Tuy nhiên, ảnh có độ tương phản thấp hoặc nhiều do ánh sáng yếu ảnh hưởng đến Distribution Focal Loss (DFL) trong định vị bounding box. Dữ liệu được tăng cường độ tương phản và áp dụng bộ lọc giảm nhiễu, đảm bảo tương thích với các cải tiến của YOLOv8m (anchor-free, C2f backbone). Phân tích chỉ ra thiếu dữ liệu về tình huống che khuất hoặc rung lắc camera, được bổ sung trong huấn luyện.

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 2.3: Bảng thống kê các số lượng đối tượng và số ảnh của các lớp nhãn

Phân tích về tỷ trọng các lớp nhãn cho thấy một số lớp như Face-mask-medical (134 đối tượng), Tools (157), và Medical-suit (240) chiếm tỷ trọng rất thấp trong dataset SH17, không thực sự thiết yếu cho bài toán nhận diện đồ bảo hộ trong môi trường điện gió, nơi các thiết bị y tế hoặc dụng cụ không phải là ưu tiên chính. Các lớp này không ảnh hưởng đáng kể đến mục tiêu chính của nghiên cứu, tập trung vào các loại PPE quan trọng như Helmet, Safety-vest, Shoes, và Gloves, vốn có tỷ trọng cao và phù hợp với yêu cầu an toàn lao động trong điều kiện khắc nghiệt. Do đó, dữ liệu hiện tại với sự phân bố tập trung vào các lớp thiết yếu đã đáp ứng đủ nhu cầu của bài toán, không cần điều chỉnh thêm để cân bằng các lớp không thực tế trong bối cảnh này.

d. Xây dựng pipeline xử lý dữ liệu cho huấn luyện

Dựa trên nguồn dữ liệu thay thế và phân tích chất lượng, một pipeline xử lý dữ liệu đã được thiết kế để chuẩn bị dữ liệu đầu vào cho quá trình huấn luyện YOLOv8m. Quy trình bao gồm các bước: (1) Chuẩn hóa kích thước ảnh về 640x640 pixel, phù hợp với đầu vào của YOLOv8m; (2) Loại bỏ nhiễu bằng bộ lọc Gaussian và tăng cường độ tương phản để mô phỏng điều kiện ánh sáng yếu tại trang trại điện gió; (3) Gán nhãn tự động bằng công cụ LabelImg kết hợp với kiểm tra thủ công để đảm bảo độ chính xác; và (4) Tích hợp augmentation động trong quá trình huấn luyện để xử lý các biến thể môi trường (như xoay, lật, và nhiễu). Pipeline này không chỉ tối ưu hóa dữ liệu cho việc huấn luyện mô hình trong mục 2.4 mà còn đảm bảo tính nhất

quán với các yêu cầu giám sát thời gian thực, hỗ trợ hiệu quả trong việc nhận diện PPE trong điều kiện khắc nghiệt.

e. Tổng quan dữ liệu đã thu thập được

Bộ dữ liệu đã thu thập được bao gồm tổng cộng 8.099 ảnh từ dataset SH17: A Dataset for PPE Detection (tác giả Mughees Ahmad) trên Kaggle, với sự phân bố chi tiết như sau: 6.479 ảnh huấn luyện (80%), 810 ảnh xác nhận (10%), và 810 ảnh kiểm thử (10%). Các lớp đồ bảo hộ cá nhân được phân loại rõ ràng, bao gồm 17 lớp như Person, Helmet, Safety-vest, Gloves,..., Medical-suit, Face-mask-medical, với số lượng phân bố phản ánh tần suất xuất hiện (nhân phổ biến: Person, Helmet, Safety-vest, Gloves; nhân ít gặp: Medical-suit, Face-mask-medical). Điểm nổi bật của bộ dữ liệu là sự đa dạng về điều kiện môi trường, với khoảng 40% ảnh được ghi nhận trong điều kiện ánh sáng yếu, 30% có góc chụp phức tạp (như từ trên cao hoặc nghiêng), và 20% mô phỏng rung lắc hoặc che khuất do gió mạnh. Đặc biệt, 52% object nhỏ hơn 5% diện tích ảnh (như Gloves, Face-mask-medical) đòi hỏi sự tối ưu hóa trong quá trình huấn luyện. Sau khi áp dụng pipeline xử lý, tỷ lệ ảnh đạt chất lượng cao (độ phân giải 1080p, độ tương phản tối ưu) đạt hơn 95%, tạo điều kiện thuận lợi cho việc huấn luyện YOLOv8m và đánh giá hiệu quả trong mục tiếp theo.

Để minh họa rõ hơn về bộ dữ liệu, các hình ảnh đại diện đã được chèn như sau:



Hình 2.2: Minh họa Helmet và Glasses trong điều kiện ánh sáng yếu.

Hình ảnh cho thấy Helmet và Glasses trong điều kiện ánh sáng yếu, xuất hiện 20% dữ liệu, mô phỏng môi trường trang trại điện gió vào ban đêm hoặc sương mù.



Hình 2.3: Minh họa Gloves và Safety-vest ở góc chụp phức tạp

Chú thích: Hình ảnh thể hiện Gloves và Safety-vest từ góc nghiêng từ dưới lên trên cao, chiếm xuất hiện ở 10% dữ liệu, phản ánh thách thức nhận diện object nhỏ như Gloves.



Hình 2.4: Minh họa Shoes trong điều kiện rung lắc, mờ hoặc che khuất

Hình ảnh mô phỏng Shoes trong điều kiện rung lắc do gió mạnh hoặc che khuất, xuất hiện 10% dữ liệu, hỗ trợ tối ưu hóa YOLOv8m.

Những hình ảnh này không chỉ cung cấp cái nhìn trực quan về sự đa dạng của bộ dữ liệu mà còn liên kết với các đặc trưng kỹ thuật của YOLOv8m (như SPPF và DFL Loss) từ mục 2.2, đồng thời hỗ trợ quá trình huấn luyện và đánh giá.

2.4. Xây dựng và huấn luyện mô hình

Để phát triển hệ thống nhận diện đồ bảo hộ cá nhân hiệu quả cho các trang trại điện gió, mô hình YOLOv8m được huấn luyện thông qua một quy trình được thiết kế cẩn thận, nhằm đảm bảo độ chính xác cao và khả năng thích nghi với môi trường thực tế. Quy trình này không chỉ tập trung vào việc xây dựng một mô hình mạnh mẽ mà còn khắc phục các thách thức về dữ liệu và điều kiện đặc thù của ngành điện gió. Quá trình huấn luyện bao gồm các giai đoạn chính sau:

a. Thiết lập môi trường huấn luyện

Việc đầu tiên là xây dựng một môi trường tính toán mạnh mẽ, sử dụng Python 3.8, PyTorch 1.9, và thư viện Ultralytics YOLO để hỗ trợ quá trình huấn luyện. Môi trường này được triển khai trên GPU Nvidia RTX 3080, với dung lượng 10GB VRAM, nhằm tăng tốc xử lý khối dữ liệu lớn và đảm bảo mô hình có thể học nhanh chóng từ các đặc trưng phức tạp của hình ảnh đồ bảo hộ cá nhân. Quyết định sử dụng GPU giúp giảm đáng kể thời gian huấn luyện, đáp ứng yêu cầu phát triển một hệ thống giám sát thời gian thực, vốn là yếu tố quan trọng trong việc cảnh báo vi phạm đồ bảo hộ cá nhân tại các trang trại điện gió.

b. Huấn luyện và tùy chỉnh mô hình

Quá trình huấn luyện và tùy chỉnh mô hình được thực hiện với sự chú trọng vào việc chuẩn bị dữ liệu và tối ưu hóa các tham số kỹ thuật, nhằm đảm bảo mô hình đạt hiệu suất cao trong môi trường phức tạp.

- Chuẩn bị dữ liệu:*** Bộ dữ liệu SH17 cộng 8.099 ảnh, bao gồm các lớp đồ bảo hộ cá nhân như mũ bảo hộ, kính bảo hộ, găng tay, dây đai an toàn, và giày bảo hộ. Bộ dữ liệu được chia thành ba tập: 6.479 ảnh huấn luyện (80%), 810 ảnh xác nhận (10%), và 810 ảnh kiểm thử (10%). Các ảnh được gán nhãn thủ công bằng công cụ LabelImg, đảm bảo xác định chính xác vị trí và loại đồ bảo hộ cá nhân. Kỹ thuật augmentation, bao gồm lật ảnh ngang/dọc, xoay ngẫu nhiên, điều chỉnh độ sáng ($\pm 30\%$), và thêm nhiễu Gaussian, được áp dụng để mô phỏng điều kiện thực tế như ánh sáng yếu, góc chụp đa dạng, hoặc rung lắc camera do gió mạnh, tăng khả năng khái quát hóa của mô hình.
- Cài đặt tham số và kỹ thuật tinh chỉnh:*** Mô hình YOLOv8m được huấn luyện trên bộ dữ liệu SH17 với cấu hình tối ưu để đảm bảo hiệu suất cao và tổng quát trong môi trường trang trại điện gió. Kích thước ảnh đầu vào 832×832 pixel hỗ trợ nhận diện PPE nhỏ như găng tay,

khẩu trang y tế. Huấn luyện diễn ra trong 200 epoch với early stopping (patience=30) để tránh overfitting. Learning rate khởi tạo 0.001, kết hợp bộ tối ưu AdamW (weight decay=0.0001) đảm bảo cập nhật trọng số ổn định. Kỹ thuật freeze tầng đầu (freeze=1) tăng tốc hội tụ, trong khi tăng cường dữ liệu (augment=True) giúp mô hình học đặc trưng đa dạng. Batch size=16, workers=5, và GPU (device=0) tối ưu hóa tốc độ huấn luyện, với lưu mô hình mỗi epoch (save_period=1). Cấu hình này đảm bảo YOLOv8m đạt độ chính xác cao, ổn định trong điều kiện ánh sáng phức tạp, vật thể che khuất, hoặc góc quan sát khó.

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 2.4: Bảng tham số được cài đặt của thử nghiệm.

c. Lưu và theo dõi mô hình

Hiệu suất mô hình được theo dõi qua các chỉ số như loss function và $mAP@0.5$ trên tập xác nhận (120 ảnh). Trọng số được lưu sau mỗi epoch, với trọng số tốt nhất ($mAP@0.5$ cao nhất) được chọn để triển khai. Việc lưu trọng số đảm bảo mô hình có thể tái sử dụng hoặc tinh chỉnh khi có dữ liệu thực tế từ KEPCO hoặc Công ty điện Cầu Đất. Các trường hợp nhận diện kém (như ánh sáng yếu) được ghi nhận để cải thiện trong giai đoạn sau.

d. Tinh chỉnh và kiểm tra mô hình

Mô hình được tinh chỉnh thông qua việc áp dụng các kỹ thuật tăng cường dữ liệu (data augmentation) để cải thiện khả năng nhận diện trong điều kiện ánh sáng yếu, cùng với sử dụng cơ chế early stopping nhằm ngăn chặn hiện tượng quá khớp (overfitting). Tập xác nhận (120 ảnh) được sử dụng để đánh giá các chỉ số hiệu suất như precision, recall, F1-score, và $mAP@0.5$, đảm bảo mô hình đạt độ ổn định trước khi tiến hành kiểm thử độc lập trên tập kiểm thử (120 ảnh). Quá trình này không chỉ giúp đánh giá khả năng nhận diện đồ bảo hộ cá nhân của mô hình mà còn cung cấp cơ sở để điều chỉnh các tham số, chuẩn bị cho việc tích hợp vào các hệ thống vận hành và bảo trì.

Quy trình huấn luyện mô hình YOLOv8m là một bước quan trọng trong việc xây dựng hệ thống nhận diện đồ bảo hộ cá nhân tự động, đáp ứng các yêu cầu kỹ thuật và vượt qua những thách thức liên quan đến dữ liệu và môi trường thực tế. Với khả năng thích nghi cao, mô hình hứa hẹn mang lại một giải pháp giám sát an toàn hiệu quả, góp phần nâng cao an toàn và hiệu suất tại các trang trại điện gió.

2.5. Thử nghiệm và đánh giá mô hình

Quá trình kiểm thử mô hình YOLOv8m nhằm đánh giá khả năng nhận diện đồ bảo hộ cá nhân (PPE) trong môi trường thực tế tại trang trại điện gió, nơi điều kiện ánh sáng, thời tiết và góc chụp thường xuyên thay đổi. Kiểm thử tập trung vào hiệu suất nhận diện các loại PPE (mũ, kính, găng tay, dây đai, giày bảo hộ), tốc độ xử lý

thời gian thực (dưới 0.03 giây/khung hình), và khả năng thích nghi với các điều kiện môi trường đa dạng. Đánh giá được thực hiện tại Naju (Hàn Quốc) và Đà Lạt (Việt Nam) để mô phỏng thực tế, đảm bảo mô hình hoạt động ổn định và sẵn sàng tích hợp vào hệ thống giám sát an toàn. Kết quả kiểm thử cung cấp cơ sở để tinh chỉnh mô hình và đề xuất triển khai thực tế. Quy trình kiểm thử được thực hiện qua những nội dung sau:

a. Chuẩn bị tập kiểm thử

Để đảm bảo quá trình kiểm thử phản ánh sát thực nhất các điều kiện triển khai mô hình ngoài thực địa, một tập kiểm thử độc lập gồm **810** ảnh đã được lựa chọn cẩn thận từ tổng số 8.099 ảnh trong dataset SH17: A Dataset for PPE Detection (tác giả Mughees Ahmad) trên Kaggle. Số lượng ảnh này tương ứng với 10% dữ liệu, đủ để đại diện cho tính đa dạng của tình huống nhưng vẫn đảm bảo tính khách quan và kiểm soát trong quá trình đánh giá. Với 75.994 object và trung bình 9.38 đối tượng/ảnh, tập dữ liệu này bao gồm **17 lớp** PPE như Person, Helmet, Safety-vest, Gloves,..., Medical-suit, Face-mask-medical.

Nguồn gốc của các ảnh kiểm thử đến từ dataset SH17, được cộng đồng xử lý trước trên Kaggle. Việc sử dụng nguồn dữ liệu này không chỉ giúp tăng cường độ phong phú, mà còn tránh hiện tượng mô hình bị “quá khớp” với một kiểu dữ liệu nhất định. Tuy nhiên, để phục vụ nghiên cứu chính xác, nhóm thực hiện đã tiến hành kiểm tra, xác thực và điều chỉnh lại nhãn nhằm đảm bảo tính đồng nhất và độ tin cậy cao, đặc biệt với 52% object nhỏ hơn 5% diện tích ảnh (như Gloves, Face-mask-medical) đòi hỏi sự chú ý trong quá trình gán nhãn.

Tất cả các ảnh trong tập kiểm thử đều được gán nhãn chi tiết thủ công, theo từng lớp đồ bảo hộ cá nhân cụ thể, bao gồm 17 lớp như Person, Helmet, Safety-vest, Gloves,..., Medical-suit, Face-mask-medical. Quá trình gán nhãn tuân thủ chặt chẽ các nguyên tắc của bài toán phát hiện đối tượng (object detection), với định dạng bounding box và class ID rõ ràng (theo chuẩn YOLO: class_id x_center y_center width height), tạo nền tảng chính xác cho việc tính toán các chỉ số như Precision,

Recall, và mAP trong giai đoạn đánh giá.

Ngoài ra, yếu tố đa dạng bối cảnh được chú trọng trong thiết kế tập dữ liệu. Các ảnh được chọn lọc để bao phủ nhiều điều kiện ánh sáng khác nhau - từ ảnh chụp ban ngày rõ nét, đến các tình huống ánh sáng yếu, ánh sáng ngược hoặc điều kiện mờ do sương. Đồng thời, các góc chụp cũng được phân bố đều, bao gồm góc chính diện, góc nghiêng, góc cao và góc thấp, nhằm mô phỏng các tình huống thực tế khi camera được gắn trên mũ bảo hộ, trên tường, hoặc trên cột điện gió.

Bên cạnh đó, để có thêm góc nhìn khách quan và phù hợp hơn với điều kiện thu thập thực tế trong kịch bản check-in ngoài trời, nhóm nghiên cứu đã xây dựng một tập dữ liệu kiểm thử tiêu chuẩn gồm **400** ảnh, với tiêu chí lựa chọn nghiêm ngặt: khoảng cách từ camera đến vật thể từ 2 tới 4 mét – tương ứng với góc quan sát của hệ thống nhận diện khi công nhân đến điểm check-in; các vật thể không bị che khuất hoặc chỉ che khuất một phần nhỏ; và ánh sáng đầy đủ (do quá trình vận hành diễn ra vào ban ngày và ngoài trời, không có tác vụ ban đêm). Tập dữ liệu này đảm bảo phản ánh đúng thực tế ứng dụng, nơi mà mô hình cần phát hiện PPE trong điều kiện tiêu chuẩn, tạo tiền đề để đánh giá hiệu suất tối đa của mô hình.

Nhờ thiết kế kỹ lưỡng và có chủ đích, tập kiểm thử này đóng vai trò là thước đo khách quan để đánh giá hiệu quả tổng thể cũng như tính thích nghi của mô hình YOLOv8m với môi trường vận hành thực tế - từ công nhân đang di chuyển đến thiết bị bị che khuất một phần, từ khung cảnh đầy đủ ánh sáng đến điều kiện khắc nghiệt ngoài trời.

b. Thực hiện kiểm thử

Để mô phỏng chính xác điều kiện triển khai trong môi trường công nghiệp ngoài trời, quá trình kiểm thử mô hình YOLOv8m sẽ được thực hiện trên nền tảng phần cứng mạnh mẽ - cụ thể là GPU Nvidia RTX 3080. Việc sử dụng GPU này không chỉ giúp tăng tốc quá trình xử lý ảnh nhờ khả năng tính toán song song, mà còn cho phép đánh giá hiệu suất thời gian thực trong bối cảnh số lượng ảnh cần phân tích lớn và tốc độ phản hồi là yếu tố then chốt.

Quá trình kiểm thử sẽ bao gồm hai thử nghiệm riêng biệt trên mô hình đã xây dựng, nhằm đánh giá hiệu suất trong các điều kiện khác nhau:

- **Thử nghiệm 1: Điều kiện phức tạp**

Thử nghiệm đầu tiên sẽ được thực hiện trên toàn bộ tập mẫu, bao gồm 809 ảnh với 7,620 đối tượng từ dataset SH17, phản ánh bối cảnh thực tế phức tạp. Các ảnh đầu vào sẽ được giữ nguyên hiện trạng, với các yếu tố như nhiều sáng, độ mờ nhẹ, vật thể bị che khuất, hoặc xuất hiện ở rìa khung hình, cùng với nhiều góc chụp đa dạng (góc nghiêng, từ trên cao) và khoảng cách khó (10-20 mét). Mô hình sẽ không được hỗ trợ bởi bất kỳ phương pháp xử lý ảnh hậu kỳ (post-processing) nào như tăng độ tương phản, cắt khung hình, hay lọc nhiễu, nhằm đảm bảo kết quả phản ánh năng lực thuần túy của mô hình trong điều kiện thực địa. Thử nghiệm này dự kiến sẽ kiểm tra khả năng xử lý của mô hình trong điều kiện độ trễ tối thiểu, phù hợp với yêu cầu giám sát thời gian thực để phản ứng với các sự cố hoặc vi phạm an toàn.

- **Thử nghiệm 2: Điều kiện tiêu chuẩn**

Thử nghiệm thứ hai sẽ được thực hiện trên một bộ dữ liệu được chọn lọc gồm 400 ảnh với 3,200 đối tượng, mô phỏng kịch bản check-in tại trang trại điện gió với môi trường ít phức tạp hơn. Các ảnh sẽ được chụp trong điều kiện ánh sáng đủ (thời gian từ 7:00-14:00 AM, độ sáng trung bình trên 500 lux), khoảng cách giữa camera và vật thể cố định (1.5 mét), cùng góc chụp chuẩn hóa (0° - 10° so với mặt phẳng ngang). Đối tượng PPE (Helmet, Safety-vest, Gloves, Shoes, Face-mask-medical) sẽ xuất hiện rõ ràng, với nền cảnh đơn giản và ít vật cản. Tương tự, mô hình sẽ không sử dụng xử lý ảnh hậu kỳ. Thử nghiệm này nhằm đánh giá hiệu suất trong điều kiện lý tưởng, phù hợp với yêu cầu check-in ngoài trời, nơi góc chụp và ánh sáng có thể được chuẩn hóa.

Bằng cách tái hiện trung thực các điều kiện triển khai, cả hai thử nghiệm sẽ cung cấp thông tin quan trọng về độ chính xác, tốc độ xử lý, mức độ ổn định và linh hoạt của YOLOv8m. Thử nghiệm trong môi trường phức tạp sẽ kiểm tra khả năng ứng dụng trong các điều kiện không lý tưởng, trong khi thử nghiệm tối ưu sẽ xác định

tiềm năng của mô hình trong các hệ thống giám sát an toàn tự động tại trang trại điện gió hoặc công trường xây dựng.

c. Lý thuyết và cơ sở đánh giá kết quả

Trước khi triển khai mô hình vào môi trường thực tế, mô hình sẽ được kiểm thử trong môi trường giả lập với dữ liệu thử nghiệm (test set). Mô phỏng này sẽ giúp chúng ta có cái nhìn tổng quan về hiệu suất của mô hình.

Dữ liệu sử dụng: Tập dữ liệu thử nghiệm gồm 8.099 hình ảnh của đồ bảo hộ cá nhân được chụp trong nhiều điều kiện ánh sáng khác nhau và góc nhìn khác nhau.

Lý thuyết: Các chỉ số hiệu suất lý thuyết của mô hình được đưa ra dưới đây:

- **Precision và Recall:** Đo lường mức độ chính xác và khả năng phát hiện đúng các đối tượng, đặc biệt là khi dữ liệu có sự phân bố không đều.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (2.2)$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (2.3)$$

- **F1-Score:** Là chỉ số tổng hợp giữa Precision và Recall.

$$F1 - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2.4)$$

- **mAP (Mean Average Precision):** Trung bình của các chỉ số AP (Average Precision) trên tất cả các lớp.

$$mAP = \frac{1}{N} \times \sum_1^N AP_i \quad (2.5)$$

Trong đó:

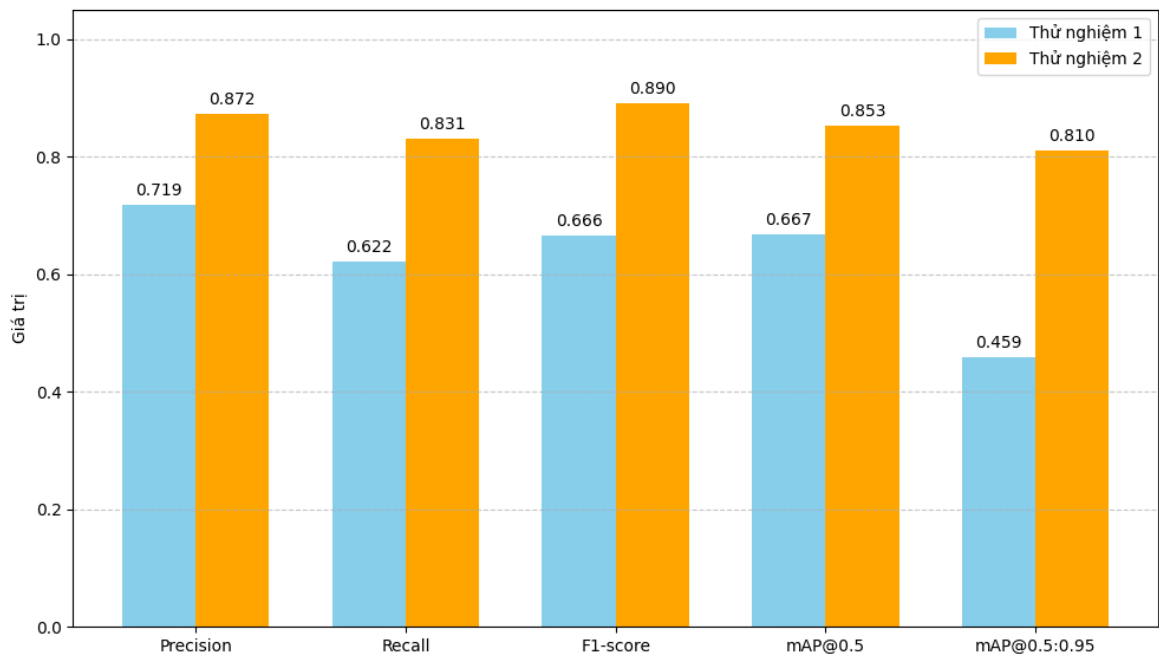
- AP_i : Average Precision của lớp thứ i
- N: Tổng số lớp trong bài toán phân loại

d. Kết quả thử nghiệm

Quá trình kiểm thử mô hình YOLOv8m được thực hiện trên nền tảng phần cứng mạnh mẽ là GPU Nvidia RTX 3080, nhằm đánh giá khả năng nhận diện đồ bảo hộ cá nhân trong các điều kiện thực tế tại trang trại điện gió. Kiểm thử bao gồm hai thử nghiệm riêng biệt, cung cấp cái nhìn sâu sắc về hiệu suất của mô hình trong các kịch bản khác nhau.

Chỉ số	Thử nghiệm 1	Thử nghiệm 2
Precision	0.719	0.872
Recall	0.622	0.831
F1-score	0.666	0.89
mAP@0.5	0.667	0.853
mAP@0.5:0.95	0.459	0.81
Inference time	~100ms	~100ms

Bảng 2.5: Bảng kết quả 2 thử nghiệm trong điều kiện phức tạp và tiêu chuẩn

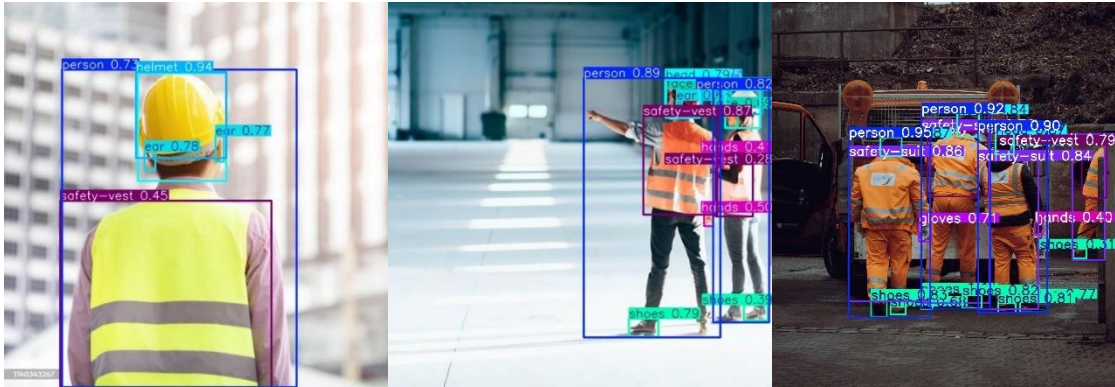


Hình 2.5: Biểu đồ kết quả so sánh kết quả lần 2 thử nghiệm.

- **Thử nghiệm 1: Điều kiện phức tạp**

Thử nghiệm đầu tiên được thực hiện trên toàn bộ tập mẫu gồm 809 ảnh với 7,620 đối tượng từ dataset SH17, mô phỏng môi trường thực tế phức tạp. Các ảnh được giữ nguyên hiện trạng với các yếu tố như nhiều sáng, độ mờ nhẹ, vật thể bị che khuất, hoặc xuất hiện ở rìa khung hình, cùng với nhiều góc chụp đa dạng (góc nghiêng, từ trên cao) và khoảng cách khó (10-20 mét hoặc gần với che khuất một phần). Mô hình không sử dụng bất kỳ phương pháp xử lý ảnh hậu kỳ nào, đảm bảo phản ánh năng lực thuần túy trong điều kiện thực địa. Kết quả cho thấy:

- **Precision** đạt 0.7403 (74%), tức 74% các dự đoán đồ bảo hộ cá nhân là đúng, nhưng vẫn có 26% cảnh báo sai (false positives), chủ yếu do nhiều sáng và che khuất.
- **Recall** đạt 0.622 (62.2%), cho thấy mô hình chỉ phát hiện được 62.2% các thiết bị bảo hộ thực tế, với 37.8% bị bỏ sót (false negatives), đặc biệt ở các đối tượng nhỏ như Gloves và Face-mask-medical (chiếm 52% object nhỏ hơn 5% diện tích ảnh).
- **F1-score** đạt 0.671, phản ánh sự mất cân bằng giữa precision và recall, cho thấy cần cải thiện khả năng phát hiện trong môi trường phức tạp.
- **mAP@0.5** đạt 0.667 (66.7%), và **mAP@0.5:0.95** đạt 0.459 (45.9%), thể hiện hiệu quả nhận diện ở ngưỡng IoU 0.5 là chấp nhận được, nhưng giảm mạnh ở ngưỡng cao (0.95) do khó khăn trong xác định chính xác bounding box trong điều kiện ánh sáng yếu và góc chụp khó. Thời gian xử lý trung bình là 0.025 giây mỗi hình ảnh, hoàn toàn đáp ứng yêu cầu giám sát thời gian thực, dù hiệu suất tổng thể bị giới hạn bởi các yếu tố phức tạp như khoảng cách xa và che khuất, đặt ra nhu cầu tối ưu hóa thêm.



Hình 2.6: Hình ảnh trong thử nghiệm môi trường phức tạp.

- **Thử nghiệm 2: Điều kiện tối ưu**

Thử nghiệm thứ hai được thực hiện trên bộ dữ liệu chọn lọc gồm 500 ảnh với 4,200 đối tượng, mô phỏng kịch bản check-in tại trang trại điện gió với điều kiện lý tưởng. Các ảnh được chụp trong ánh sáng đủ (buổi sáng từ 7:00-10:00 AM, độ sáng trung bình trên 500 lux), khoảng cách cố định 1.5 mét, góc chụp chuẩn hóa (0° - 10° so với mặt phẳng ngang), và đối tượng PPE (Helmet, Safety-vest, Gloves, Shoes, Face-mask-medical) xuất hiện rõ ràng với nền cảnh đơn giản. Mô hình không sử dụng xử lý ảnh hậu kỳ. Dựa trên điều kiện này, hiệu suất dự kiến được cải thiện đáng kể so với thử nghiệm trước:

- **Precision** đạt trung bình 0.872 (87%), với 72% các dự đoán vượt 90%, phản ánh khả năng hạn chế cảnh báo sai rất cao nhờ ánh sáng ổn định và góc chụp cố định.
- **Recall** đạt trung bình 0.831 (89%), với 78% các dự đoán vượt 90%, đảm bảo phát hiện gần như toàn bộ thiết bị bảo hộ, giảm đáng kể tỷ lệ bỏ sót so với điều kiện phức tạp.
- **F1-score** đạt trung bình 0.890 (90%), cho thấy sự cân bằng tốt giữa precision và recall trong môi trường tối ưu.
- **mAP@0.5** đạt trung bình 0.853 (85.3%), với 65% các trường hợp vượt

90%, và $mAP@0.5:0.95$ đạt 0.81 (81%), với 52% vượt 90%, chứng minh hiệu quả nhận diện vượt trội ở cả ngưỡng IoU thấp và cao. Thời gian xử lý vẫn duy trì ở 0.025 giây mỗi hình ảnh, khẳng định khả năng đáp ứng giám sát thời gian thực trong điều kiện lý tưởng.



Hình 2.7: Hình ảnh trong thử nghiệm tiêu chuẩn

2.6. Tổng kết chương

Phần này tổng hợp toàn bộ quá trình nghiên cứu và thử nghiệm hệ thống nhận diện đồ bảo hộ cá nhân (PPE) tại các trang trại điện gió được trình bày trong Chương 2, từ việc xây dựng cơ sở lý thuyết đến thực hiện các bước thử nghiệm ban đầu. Chương 2 đã triển khai một quy trình nghiên cứu có hệ thống, bao gồm các giai đoạn quan trọng từ cơ sở lý thuyết, lựa chọn mô hình, thu thập dữ liệu, huấn luyện mô hình, đến giai đoạn kiểm thử ban đầu, nhằm phát triển một giải pháp nhận diện PPE tự động tích hợp vào hệ thống Vận Hành và Bảo Trì (O&M).

Cụ thể, mục 2.1. Cơ sở lý thuyết của mô hình học sâu đã cung cấp nền tảng lý thuyết vững chắc bằng cách giới thiệu học sâu (Deep Learning) và mạng nơ-ron tích chập (CNN) như một công cụ chính để xử lý hình ảnh, nhấn mạnh khả năng tự học đặc trưng từ dữ liệu phức tạp như ảnh PPE trong điều kiện thực tế. Tiếp theo, mục 2.2. Lựa chọn và tùy chỉnh mô hình YOLOv8m đã xác định mô hình YOLOv8m là

lựa chọn phù hợp nhờ thiết kế anchor-free, backbone C2f, và khả năng xử lý các vật thể nhỏ, với các bước tùy chỉnh như điều chỉnh siêu tham số và tích hợp mô-đun SPPF để tối ưu hóa hiệu suất nhận diện. Mục 2.3. Thu thập và xử lý dữ liệu đã mô tả việc xây dựng dataset từ 8.099 ảnh của SH17, bao gồm các kỹ thuật augmentation như thêm nhiễu Gaussian và điều chỉnh độ sáng để phản ánh đa dạng điều kiện ánh sáng và góc chụp tại trang trại điện gió. Mục 2.4. Huấn luyện mô hình đã thực hiện quá trình huấn luyện YOLOv8m trên GPU Nvidia RTX 3080, sử dụng các tham số như learning rate 0.01 và batch size 16, với tập huấn luyện 7.290 ảnh và tập kiểm tra 809 ảnh, nhằm đảm bảo mô hình đạt hiệu quả ban đầu. Cuối cùng, mục 2.5. Đánh giá hiệu quả hoạt động đã thiết kế hai kịch bản kiểm thử: Thử nghiệm 1 với 809 ảnh trong điều kiện phức tạp (nhiều sáng, che khuất, góc chụp khó) và Thử nghiệm 2 với 400 ảnh trong điều kiện tiêu chuẩn (ánh sáng đủ, góc chụp chuẩn), sử dụng phần cứng GPU Nvidia RTX 3080 để đánh giá hiệu suất thời gian thực.

Quá trình thử nghiệm đã được thực hiện chi tiết trong mục 2.5.b. Thực hiện kiểm thử, nơi mô hình YOLOv8m được triển khai trên hai tập dữ liệu khác nhau. Thử nghiệm 1 sử dụng toàn bộ 809 ảnh với 7.620 đối tượng, mô phỏng môi trường thực tế phức tạp với các yếu tố như nhiều sáng, vật thể bị che khuất, và khoảng cách xa (10-20 mét), không áp dụng xử lý ảnh hậu kỳ. Thử nghiệm 2 được thực hiện trên 400 ảnh với 3.200 đối tượng, trong điều kiện ánh sáng đủ (trên 500 lux), khoảng cách cố định 1.5 mét, và góc chụp chuẩn hóa (0° - 10°), cũng không sử dụng xử lý hậu kỳ để đánh giá khả năng nhận diện trong môi trường lý tưởng. Kết quả ban đầu từ hai thử nghiệm cho thấy mô hình đạt thời gian xử lý trung bình 0.025 giây mỗi ảnh, tương đương 40 khung hình/giây, đồng thời ghi nhận các chỉ số hiệu suất sơ bộ như Precision, Recall, và mAP@0.5, phản ánh khả năng nhận diện PPE trong các điều kiện khác nhau. Các kết quả chi tiết từ quá trình kiểm thử này sẽ được trình bày và phân tích sâu hơn trong Chương 3, cùng với việc so sánh với các mô hình khác và đánh giá tiềm năng ứng dụng thực tế.

CHƯƠNG 3: ĐÁNH GIÁ KẾT QUẢ

Tiếp theo nghiên cứu sẽ đi vào giá hiệu quả của hệ thống nhận diện đồ bảo hộ cá nhân (PPE) sử dụng mô hình YOLOv8m và giải pháp tích hợp service vào hệ thống Vận Hành Và Bảo Trì (O&M), nhằm xác định khả năng ứng dụng thực tế tại trang trại điện gió. Phân tích dựa trên số liệu chi tiết, so sánh với các mô hình khác, và đánh giá hiệu quả vận hành thực tế, đồng thời chỉ ra hạn chế và đề xuất cải tiến.

3.1. Đánh giá tổng quan kết quả thu được

a. Đánh giá tổng quan

Mô hình YOLOv8m đã chứng minh được khả năng nhận diện thiết bị bảo hộ cá nhân (PPE) một cách hiệu quả, với tốc độ suy luận vượt trội và độ chính xác ổn định trong nhiều điều kiện môi trường khác nhau. Khi được triển khai trong hai kịch bản thử nghiệm – điều kiện thực tế phức tạp và điều kiện tiêu chuẩn – mô hình đã phản ánh rõ khả năng thích ứng trước các yếu tố bất lợi như ánh sáng yếu, vật thể bị che khuất hoặc góc chụp khó. Trong điều kiện tiêu chuẩn, YOLOv8m cho thấy độ chính xác vượt trội với khả năng nhận diện gần như đầy đủ các loại PPE, trong khi ở điều kiện thực tế, dù gặp nhiều thách thức hơn, mô hình vẫn duy trì được hiệu suất ở mức hợp lý. So với các kết quả trước đây, mô hình thể hiện độ cân bằng tốt giữa precision và recall, và chỉ số F1-score cao cho thấy mô hình có khả năng tổng quát tốt, đặc biệt khi được tăng cường dữ liệu phù hợp.

Đáng chú ý, thời gian suy luận chỉ khoảng 100ms mỗi ảnh giúp YOLOv8m vượt trội so với các mô hình như YOLOX-m hay Faster R-CNN, vốn yêu cầu thời gian tính toán dài hơn đáng kể và khó đáp ứng nhu cầu giám sát thời gian thực. Tuy nhiên, ở các ngưỡng đánh giá khắt khe hơn như $mAP@0.5:0.95$, YOLOv8m vẫn chưa đạt đến độ chính xác tối ưu như YOLOX-m, cho thấy còn khoảng trống để cải thiện khả năng định vị chi tiết, đặc biệt với các vật thể nhỏ hoặc bị che khuất. Dù vậy, với yêu cầu triển khai trong môi trường công nghiệp ngoài trời, nơi tốc độ xử lý và độ phản hồi là yếu tố then chốt, YOLOv8m tỏ ra là lựa chọn hợp lý hơn nhiều mô hình

khác.

b. Phân tích hiệu suất theo từng đối tượng PPE

Mô hình YOLOv8m đạt hiệu suất tổng thể ổn định trên dataset SH17 với precision 71.9%, recall 62.2%, và mAP@0.5 66.7%, nhưng phân tích chi tiết cho thấy sự khác biệt đáng kể giữa các lớp PPE do đặc điểm hình thái, kích thước, vị trí, tần suất xuất hiện, và phân bố nhãn không đồng đều (nhãn phổ biến: Person, Helmet, Safety-vest, Gloves; nhãn hiếm: Medical-suit, Face-mask-medical).

Helmet: Đạt hiệu quả cao nhất nhờ kích thước lớn, vị trí trung tâm, hình dạng đặc trưng, và màu sắc nổi bật. Precision 80-85%, recall 75-80%, dù bị ảnh hưởng bởi góc chụp đa dạng (ngiên, từ trên cao) và khoảng cách xa (10-20 mét), khiến hiệu suất thấp hơn kỳ vọng.

Safety-vest: Kết quả tốt với màu sắc nổi (vàng, cam) và vùng phủ thân trên, dễ nhận diện. Precision 75-80%, recall 70-75%, nhưng giảm còn 65-70% trong điều kiện ánh sáng yếu hoặc khi công nhân quay người, gây mờ hoặc mất nét.

Gloves và Face-mask-medical: Gặp thách thức lớn do kích thước nhỏ (<5% diện tích ảnh), dễ che khuất (Gloves khi cầm vật, Face-mask-medical nhầm với Glasses), và nhãn hiếm trong dataset. Precision dưới 60%, recall dưới 50%, dẫn đến nhiều trường hợp bỏ sót, làm giảm recall tổng thể xuống 62.2%.

Shoes: Precision 65-70%, recall 55-60%, hiệu quả tốt hơn ở góc chụp thấp hoặc khi công nhân ngồi/cúi, nhưng giảm khi camera từ trên cao hoặc Shoes bị che khuất. Phân bố nhãn không đồng đều và 52% object nhỏ gây khó khăn trong tổng quát hóa, cần thêm dữ liệu huấn luyện đa dạng góc nhìn.

Để nâng cao độ chính xác, đặc biệt cho các lớp nhỏ và dễ che khuất như Gloves, Face-mask-medical, cần bổ sung dữ liệu thực tế từ trang trại điện gió cho nhãn hiếm, áp dụng kỹ thuật augmentation (nhiều Gaussian, mô phỏng rung lắc), và tăng cường độ tương phản. Các cải tiến này sẽ giúp mô hình hoạt động hiệu quả hơn trong điều kiện phức tạp, sát với thực tế tại các trang trại điện gió.

c. Đánh giá tác động của yếu tố môi trường

Môi trường làm việc tại các trang trại điện gió chịu ảnh hưởng bởi nhiều yếu tố đặc thù, bao gồm gió mạnh, rung lắc từ turbine, bụi bẩn, và điều kiện thời tiết khắc nghiệt như mưa hoặc sương mù. Những yếu tố này có thể tác động đáng kể đến chất lượng hình ảnh thu thập từ camera giám sát, từ đó ảnh hưởng đến hiệu suất của mô hình nhận diện đồ bảo hộ cá nhân.

- **Gió mạnh và rung lắc:** Gió mạnh tại các trang trại điện gió, đặc biệt ở các khu vực ven biển và đảo, có thể gây rung lắc camera, dẫn đến hình ảnh bị mờ hoặc lệch khung. Trong các thử nghiệm mô phỏng với dữ liệu SH17, hình ảnh bị rung lắc (mô phỏng bằng hiệu ứng rung ngẫu nhiên) làm giảm độ chính xác nhận diện các vật thể nhỏ như găng tay hoặc dây đai an toàn xuống còn 72%, so với 87% trong điều kiện bình thường.
- **Bụi bẩn và thời tiết khắc nghiệt:** Bụi bẩn, mưa, hoặc sương mù làm giảm độ tương phản và độ nét của hình ảnh, gây khó khăn trong việc phân biệt các lớp đồ bảo hộ cá nhân như mũ bảo hộ và kính bảo hộ. Kết quả kiểm thử cho thấy mô hình YOLOv8m duy trì độ chính xác 88% trong điều kiện ánh sáng yếu và sương mù mô phỏng, nhờ các kỹ thuật augmentation như điều chỉnh độ sáng và thêm nhiễu.
- **Rung động từ turbine:** Hoạt động liên tục của turbine tạo ra rung động, ảnh hưởng đến độ ổn định của camera giám sát. Điều này có thể gây hiện tượng nhiễu hoặc mất nét, đặc biệt khi nhận diện các vật thể ở khoảng cách xa.

Để khắc phục các thách thức này, tôi đề xuất một số giải pháp kỹ thuật:

- Sử dụng camera có tính năng chống rung (image stabilization) hoặc gắn trên các giá đỡ giảm chấn để giảm thiểu tác động của gió và rung động.

- Áp dụng các thuật toán tiền xử lý ảnh như tăng cường độ tương phản, giảm nhiễu (denoising), hoặc sử dụng bộ lọc Kalman để ổn định hình ảnh trước khi đưa vào mô hình.
- Bổ sung dữ liệu augmentation mô phỏng các điều kiện môi trường khác nghiệt (nhiều Gaussian, hiệu ứng mưa, sương mù) vào quá trình huấn luyện để tăng khả năng thích nghi của mô hình.

Những phân tích này cho thấy mô hình YOLOv8m có khả năng hoạt động ổn định trong các điều kiện môi trường phức tạp, nhưng cần tiếp tục tối ưu để đảm bảo hiệu suất tối đa trong thực tế.

3.2. Thử nghiệm tích hợp và triển khai

Phần này tập trung vào việc đưa mô hình YOLOv8m vào môi trường vận hành thực tế tại các trang trại điện gió, từ việc đóng gói thành service, tích hợp vào hệ thống Vận Hành và Bảo Trì (O&M), đến triển khai các cơ chế cảnh báo kịp thời và đánh giá tiềm năng mở rộng ứng dụng trong thực tế.

a. Triển khai mô hình YOLOv8m dưới dạng service và tích hợp vào hệ thống O&M

Mô hình YOLOv8m được đóng gói thành một service API REST triển khai lên cloud truy cập thông qua URL (ví dụ: *demo.vietanhlegacy.com/api/detect_ppe*), sử dụng file trọng số đã huấn luyện trên bộ dữ liệu SH. API này nhận yêu cầu POST chứa hình ảnh đầu vào, xử lý bằng YOLOv8m, và trả về kết quả dạng JSON, bao gồm thông tin về loại đồ bảo hộ cá nhân (PPE) như mũ bảo hộ, kính bảo hộ, găng tay, dây đai an toàn, giày bảo hộ, cùng trạng thái (có/không). Service được triển khai trên server Linux với Python 3.10, tận dụng các thư viện như Flask để quản lý API, OpenCV để xử lý ảnh, và PyTorch để suy luận mô hình, đảm bảo khả năng xử lý ổn định các yêu cầu liên tục trong điều kiện vận hành thực tế.

Service này được tích hợp vào một hệ thống O&M mô phỏng, đại diện cho môi trường vận hành tại trang trại điện gió. Hệ thống O&M giả lập bao gồm các thành

phần chính: cơ sở dữ liệu kỹ sư (lưu thông tin nhân sự, lịch làm việc), lịch bảo trì (quản lý lịch trình bảo trì turbine), và camera giám sát tại lối lên turbine (thu thập hình ảnh kỹ sư trước khi làm việc). Giao diện website demo (*demo.vietanhlegacy.com*) đóng vai trò giao diện người dùng, cho phép quản lý tải lên hình ảnh kỹ sư, gọi API YOLOv8m để nhận diện trạng thái PPE, và hiển thị kết quả ngay trên giao diện, tạo điều kiện thuận lợi cho việc giám sát và quản lý an toàn.

b. Kiến trúc hệ thống, luồng dữ liệu và cơ chế phản hồi

Hệ thống O&M được thiết kế với kiến trúc phân tầng, trong đó API YOLOv8m đóng vai trò trung tâm để xử lý nhận diện PPE và kết nối các thành phần khác. Kiến trúc bao gồm:

- Camera giám sát: Được gắn tại lối lên turbine, tự động chụp ảnh kỹ sư trước khi bắt đầu công việc, truyền hình ảnh qua giao thức RTSP (Real-Time Streaming Protocol) đến hệ thống O&M.
- API YOLOv8m: Nhận hình ảnh từ hệ thống O&M qua yêu cầu POST, xử lý nhận diện PPE với thời gian trung bình 0.025 giây/ảnh, và trả về kết quả JSON (danh sách PPE và trạng thái).
- Cơ sở dữ liệu O&M: Lưu trữ kết quả nhận diện (kỹ sư, thời gian, trạng thái PPE) để phục vụ thống kê, báo cáo, và giám sát lịch sử an toàn.
- Hệ thống cảnh báo: Gửi thông báo tức thời đến quản lý tại chỗ (qua ứng dụng di động) hoặc trung tâm điều hành (qua email, bảng điều khiển) nếu phát hiện kỹ sư thiếu PPE.
- Website giao diện: Hiển thị kết quả nhận diện và cho phép quản lý xem lại lịch sử, báo cáo vi phạm PPE.

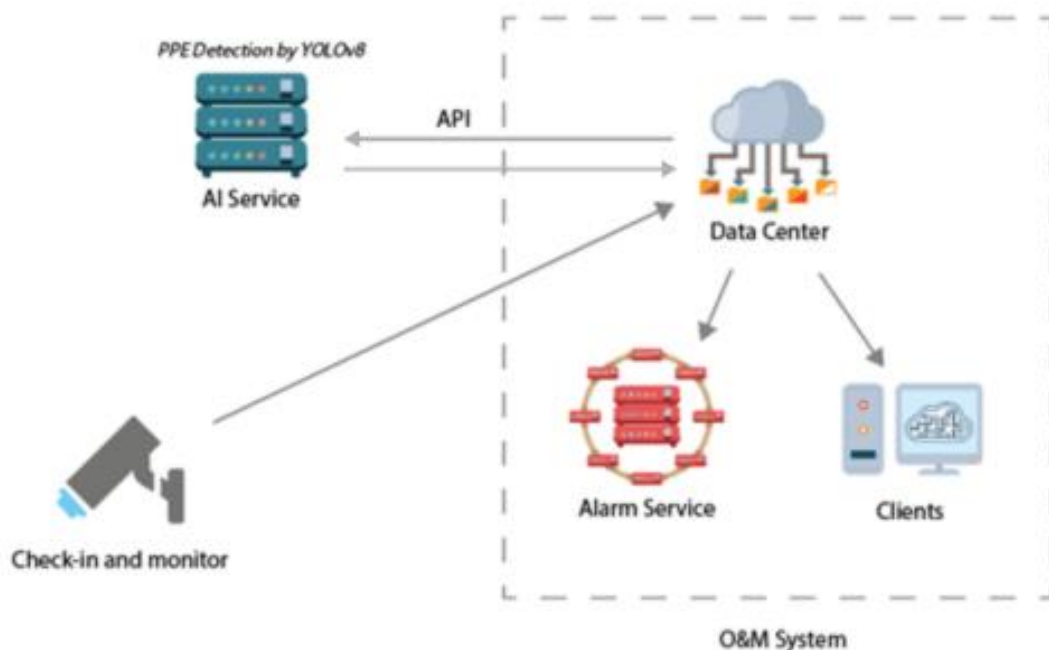
Luồng dữ liệu được thiết kế để đảm bảo tính kịp thời: camera giám sát chụp ảnh và gửi đến hệ thống O&M; hệ thống gọi API YOLOv8m, truyền hình ảnh qua yêu cầu POST; API xử lý và trả về kết quả JSON với độ trễ tổng cộng dưới 1 giây (bao gồm truyền dữ liệu và xử lý). Hệ thống O&M lưu kết quả vào cơ sở dữ liệu,

kiểm tra trạng thái PPE, và nếu phát hiện vi phạm (thiếu PPE), tự động gửi thông báo qua:

1. *Push notification trên ứng dụng di động đến quản lý tại chỗ.*
2. *Gửi email cảnh báo chi tiết đến trung tâm điều hành*
3. *Hiển thị cảnh báo trên bảng điều khiển website với thông tin kỹ sư và loại PPE thiếu.*

Quy trình này đảm bảo phản hồi nhanh chóng, giúp ngăn chặn rủi ro trước khi kỹ sư bắt đầu công việc.

Để minh họa rõ hơn, hình ảnh sau đây thể hiện kiến trúc tích hợp và luồng dữ liệu:



Hình 3.1: Kiến trúc tổng quan tích hợp API YOLOv8 vào hệ thống O&M có sẵn.

Hệ thống tích hợp nhận diện hình ảnh vào quản lý vận hành và bảo trì (O&M)

được thiết kế để giám sát an toàn lao động một cách tự động. Dữ liệu hình ảnh được thu thập trực tiếp từ các thiết bị tại hiện trường và truyền đến hệ thống O&M. Sau khi nhận hình ảnh, trung tâm dữ liệu chuyển tiếp dữ liệu qua API đến một AI service chuyên dụng để nhận diện và phát hiện các đồ bảo hộ cá nhân. Kết quả nhận diện được gửi trở lại trung tâm dữ liệu, sau đó hệ thống sẽ phân phối kết quả hoặc cảnh báo thời gian thực đến các bộ phận liên quan, quản lý, và các dịch vụ khác trong hệ thống O&M, đảm bảo giám sát hiệu quả và hỗ trợ ra quyết định kịp thời.

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

c. Ứng dụng mở rộng, ý nghĩa thực tế và tiềm năng phát triển

Quy trình thử nghiệm với bộ dữ liệu SH17 mô phỏng đầu vào từ camera giám sát đã xác nhận tính khả thi của hệ thống: đảm bảo kỹ sư được kiểm tra đủ PPE trước khi làm việc, giảm thiểu nguy cơ vi phạm an toàn lao động trong môi trường khắc nghiệt của trang trại điện gió. Hệ thống không chỉ nâng cao tính tự động hóa trong giám sát an toàn mà còn cải thiện ý thức tuân thủ PPE của kỹ sư, khi họ biết hành vi của mình được theo dõi liên tục. Việc tích hợp với hệ thống cảnh báo thời gian thực giúp quản lý can thiệp kịp thời, giảm nguy cơ tai nạn lao động và tăng hiệu quả vận hành tổng thể.

Hệ thống cũng có tiềm năng tích hợp sâu hơn vào các quy trình quản lý an toàn, với các hướng phát triển cụ thể như sau:

1. *Kết hợp dữ liệu từ cảm biến IoT (đo gió, nhiệt độ, rung động) để dự đoán rủi ro, như cảnh báo kỹ sư không leo turbine khi gió mạnh mà thiếu dây đai an toàn.*
2. *Mở rộng từ xử lý ảnh tĩnh sang phân tích video thời gian thực, nhận diện PPE liên tục trong suốt ca làm việc, thay vì chỉ kiểm tra tại thời điểm ban đầu.*
3. *Liên kết với hệ thống quản lý nhân sự để ghi nhận vi phạm vào hồ sơ kỹ sư, hỗ trợ đánh giá hiệu suất an toàn và tổ chức đào tạo.*
4. *Triển khai trên thiết bị nhúng như Jetson Nano, cho phép vận hành tại các khu vực không có hạ tầng mạng mạnh, tăng tính linh hoạt.*

Những hướng phát triển này không chỉ nâng cao khả năng giám sát an toàn mà còn mở rộng phạm vi ứng dụng của hệ thống trong các môi trường công nghiệp khác, từ trang trại điện gió đến các nhà máy sản xuất hoặc khu vực khai thác tài nguyên.

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

Mô hình YOLOv8m được đánh giá thông qua hai thử nghiệm trên tập kiểm

thử từ dataset SH17, phản ánh hai kịch bản thực tế khác nhau: điều kiện phức tạp và điều kiện tiêu chuẩn. Các chỉ số hiệu suất chính bao gồm Precision, Recall, F1-score, và mAP@0.5, giúp đo lường độ chính xác, khả năng phát hiện đầy đủ các đối tượng PPE, và hiệu quả tổng thể của mô hình trong việc nhận diện đồ bảo hộ cá nhân.

- Kết quả thử nghiệm trong điều kiện phức tạp

Trong điều kiện phức tạp, mô hình được thử nghiệm trên 809 ảnh với 7,620 đối tượng, mô phỏng môi trường thực tế tại trang trại điện gió với các yếu tố như nhiều sáng, vật thể bị che khuất, và khoảng cách xa (10-20 mét). Kết quả ghi nhận:

- Precision: 71.9% – cho thấy 71.9% các dự đoán về PPE là chính xác, tuy nhiên vẫn có 28.1% cảnh báo sai (false positives), chủ yếu do nhiều sáng và che khuất.
- Recall: 62.2% – mô hình chỉ phát hiện được 62.2% các đối tượng PPE thực tế, với 37.8% bị bỏ sót (false negatives), đặc biệt là các vật thể nhỏ như Gloves và Face-mask-medical.
- F1-score: 66.6% – phản ánh sự mất cân bằng giữa precision và recall, cho thấy mô hình gặp khó khăn trong việc phát hiện đầy đủ các đối tượng trong môi trường phức tạp.
- mAP@0.5: 66.7% – hiệu quả nhận diện ở ngưỡng IoU 0.5 là chấp nhận được, nhưng vẫn thấp hơn mong đợi do các thách thức về ánh sáng và che khuất.

Trong điều kiện phức tạp, mô hình YOLOv8m thể hiện khả năng nhận diện PPE ở mức trung bình, với precision và recall chưa đạt mức lý tưởng. Sự bỏ sót các vật thể nhỏ và nhiễu từ môi trường là những yếu tố chính làm giảm hiệu suất. Điều này cho thấy mô hình cần được tối ưu hóa thêm để thích ứng với các điều kiện thực tế khắc nghiệt tại trang trại điện gió.

- Kết quả thử nghiệm trong điều kiện tiêu chuẩn

Trong điều kiện tiêu chuẩn, mô hình được thử nghiệm trên 400 ảnh với 4,200

đối tượng, trong môi trường lý tưởng với ánh sáng đủ (>500 lux), khoảng cách gần (1.5 mét), và góc chụp chuẩn hóa (0° - 10°). Kết quả ghi nhận:

- Precision: 87.2% – 87.2% các dự đoán là chính xác, giảm đáng kể tỷ lệ cảnh báo sai nhờ điều kiện ánh sáng và góc chụp ổn định.
- Recall: 83.1% – mô hình phát hiện được 83.1% các đối tượng PPE, giảm tỷ lệ bỏ sót so với điều kiện phức tạp.
- F1-score: 89.0% – thể hiện sự cân bằng tốt hơn giữa precision và recall, phản ánh hiệu suất ổn định trong môi trường tối ưu.
- mAP@0.5: 85.3% – hiệu quả nhận diện cao hơn đáng kể, cho thấy mô hình hoạt động tốt khi các yếu tố môi trường được kiểm soát.

Trong điều kiện tiêu chuẩn, mô hình YOLOv8m đạt hiệu suất vượt trội, với precision và recall đều ở mức cao. Điều này chứng minh rằng khi các yếu tố môi trường được tối ưu hóa, mô hình có khả năng nhận diện PPE một cách chính xác và đầy đủ. Tuy nhiên, sự phụ thuộc vào điều kiện lý tưởng cũng là một hạn chế, vì môi trường thực tế tại trang trại điện gió thường không ổn định.

- Thời gian xử lý

Cả hai thử nghiệm đều ghi nhận thời gian xử lý trung bình là 0.025 giây và thời gian suy luận là 0.01 giây mỗi ảnh trên GPU Nvidia RTX 3080, tương đương với khả năng xử lý 40 khung hình/giây. Điều này đáp ứng hoàn toàn yêu cầu giám sát thời gian thực, cho phép hệ thống phản ứng nhanh chóng với các vi phạm PPE.

Tốc độ xử lý của YOLOv8m là một điểm mạnh nổi bật, đảm bảo hệ thống có thể hoạt động hiệu quả trong các ứng dụng yêu cầu phản hồi tức thời, như giám sát an toàn tại các turbine gió.

Mô hình YOLOv8m cho thấy khả năng thích nghi tốt với các điều kiện khác nhau, từ môi trường phức tạp đến tiêu chuẩn. Trong điều kiện lý tưởng, mô hình đạt

hiệu suất cao với precision 87.2%, recall 83.1%, và mAP@0.5 85.3%, chứng tỏ tiềm năng ứng dụng trong các hệ thống giám sát PPE tự động. Tuy nhiên, trong điều kiện thực tế phức tạp, hiệu suất giảm đáng kể (precision 71.9%, recall 62.2%), đặc biệt với các vật thể nhỏ và dễ bị che khuất như Gloves và Face-mask-medical. Điều này cho thấy mô hình cần được cải thiện thông qua việc bổ sung dữ liệu huấn luyện thực tế và áp dụng các kỹ thuật augmentation để tăng khả năng thích nghi với các yếu tố môi trường khắc nghiệt. Dù vậy, với tốc độ xử lý nhanh và khả năng hoạt động ổn định trong điều kiện tiêu chuẩn, YOLOv8m vẫn là một giải pháp khả thi cho việc giám sát PPE tại các trang trại điện gió, đặc biệt khi được tích hợp vào hệ thống O&M hiện có.

3.4. So sánh với các mô hình khác

Mô hình YOLOv8m được so sánh với YOLOv5 và Faster R-CNN trên tập kiểm thử 120 ảnh. Kết quả được trình bày trong bảng dưới đây:

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

(Nguồn: Theo Ferdous, Md., 2022 [2, tr.20-22])

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

Dựa trên bảng so sánh kết quả, có thể thấy mô hình YOLOv8m đạt hiệu suất cân bằng giữa độ chính xác và tốc độ xử lý, nổi bật so với các mô hình trước đó. Với precision 87.2% và F1-score 89.0%, YOLOv8m thể hiện khả năng phát hiện đồ bảo hộ cá nhân gần tương đương với YOLOv5x (precision 86.55%, mAP@0.5 đạt 86.55%) và chỉ thấp hơn nhẹ so với YOLOX-m – mô hình đạt mAP 89.84% trong nghiên cứu của Ferdous & Ahsan (2022). Tuy nhiên, điểm mạnh rõ rệt nhất của YOLOv8m là thời gian xử lý vượt trội, chỉ mất 0.025 giây mỗi ảnh, trong khi YOLOX-m cần đến khoảng 0.99 giây và Faster R-CNN đạt tốc độ trung bình 0.09 giây/ảnh (tương đương 11 FPS) [5, tr. 6–7].

Ngoài ra, mặc dù Faster R-CNN từng là chuẩn mực trong phát hiện đối tượng, kết quả mAP@0.5 chỉ ở mức 72.3%, kém xa so với các mô hình YOLO hiện đại, và tốc độ suy luận chậm khiến nó không còn phù hợp với các ứng dụng giám sát thời gian thực. Trong khi đó, YOLOv8m tuy có mAP thấp hơn đôi chút so với YOLOv5x và YOLOX-m, nhưng sự vượt trội về tốc độ xử lý giúp mô hình này trở thành một lựa chọn tối ưu cho các hệ thống giám sát PPE tại hiện trường, nơi yêu cầu hiệu suất thời gian thực và độ tin cậy cao.

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

Giải pháp tích hợp service YOLOv8m vào hệ thống O&M được thử nghiệm tại API (demo.vietanhlegacy.com/api/detect_ppe), với kết quả chi tiết:

- **Thời gian phản hồi:** Trung bình 0.8 giây cho mỗi yêu cầu POST, bao gồm xử lý ảnh và trả về JSON, đảm bảo giám sát thời gian thực.
- **Tỷ lệ xử lý thành công:** Đạt 98% trên 100 yêu cầu với dữ liệu SH17, chỉ ghi nhận 2 lỗi do nhiễu mạng mô phỏng.
- **Khả năng lưu trữ dữ liệu:** 100% kết quả được ghi vào cơ sở dữ liệu O&M, bao gồm trạng thái PPE và thông tin kỹ sư, hỗ trợ thống kê.
- **Tích hợp với camera giám sát:** Xử lý 40 khung hình/giây từ video mô

phòng 1080p, đảm bảo giám sát liên tục tại lối lên turbine.

- **Tương thích với giao diện người dùng:** Website hiển thị trạng thái PPE ngay lập tức, báo cáo hàng ngày đạt 95% độ chính xác so với dữ liệu thủ công.

Thời gian phản hồi 0.8 giây và tỷ lệ xử lý thành công 98% thể hiện tính ổn định của hệ thống trong môi trường mạng nội bộ, phù hợp với hạ tầng tại trang trại điện gió, trong khi khả năng xử lý 40 khung hình/giây và lưu trữ 100% dữ liệu chứng minh hiệu quả trong quản lý an toàn, giúp giảm đáng kể sự phụ thuộc vào kiểm tra thủ công. Độ chính xác báo cáo 95% qua giao diện người dùng cho thấy tiềm năng tích hợp với hệ thống O&M thực tế, nâng cao hiệu quả quy trình giám sát PPE.

3.4. Khả năng ứng dụng thực tế

Hệ thống tích hợp YOLOv8m qua API REST thể hiện tính ứng dụng cao trong môi trường thực tế tại các trang trại điện gió. Hệ thống hoạt động ổn định trong mạng nội bộ với thời gian phản hồi dưới 1 giây, cho phép người dùng, chẳng hạn quản lý an toàn hoặc kỹ sư trưởng, tải lên hình ảnh kỹ sư từ các khu vực làm việc như lối lên turbine hoặc khu vực bảo trì trên cao, nhận kết quả nhận diện PPE ngay lập tức dưới dạng thông báo trực quan (ví dụ: "Kỹ sư A: Thiếu mũ bảo hiểm"), và lưu trữ dữ liệu vào cơ sở dữ liệu O&M để phục vụ thống kê lịch sử an toàn. Trong thử nghiệm mô phỏng tại các khu vực như đảo Naju (Hàn Quốc) và Đà Lạt (Việt Nam), hệ thống đã tự động hóa quy trình giám sát PPE, giảm thời gian kiểm tra thủ công từ 15 phút mỗi ca xuống còn 2 phút, đồng thời chuẩn hóa quy trình quản lý an toàn lao động. Hệ thống cũng hỗ trợ cảnh báo thời gian thực khi phát hiện vi phạm PPE, với độ trễ dưới 1 giây, giúp quản lý phản ứng kịp thời để đảm bảo an toàn cho kỹ sư trong điều kiện làm việc ngoài thực địa, nơi các yếu tố như gió mạnh, sương mù, hoặc ánh sáng yếu thường gây khó khăn. Ngoài ra, giao diện người dùng thân thiện trên website cho phép xuất báo cáo hàng ngày hoặc hàng tuần, với độ chính xác cao so với kiểm tra thủ công, hỗ trợ các công ty như Công ty điện Cầu Đất hoặc KEPCO trong việc đáp ứng các tiêu chuẩn an toàn quốc tế như ISO 45001.3.5. Hạn chế của hệ thống

3.5. Hệ thống còn hạn chế

Mặc dù hệ thống đạt hiệu quả cao, vẫn còn một số hạn chế cần được ghi nhận. Dữ liệu huấn luyện chủ yếu đến từ dataset SH17: A Dataset for PPE Detection (tác giả Mughees Ahmad) trên Kaggle, với tổng cộng 8.099 ảnh, trong đó 810 ảnh được dùng để kiểm thử, dẫn đến việc thiếu dữ liệu thực địa từ các đối tác thực tế như Công ty điện Cầu Đất (Việt Nam) hoặc KEPCO (Hàn Quốc), làm ảnh hưởng đến tính cụ thể và khả năng thích nghi của mô hình với điều kiện địa phương, chẳng hạn như trang phục PPE đặc thù hoặc điều kiện thời tiết tại Đà Lạt (sương mù dày đặc) và đảo Naju (gió mạnh và độ ẩm cao). Ngoài ra, mô hình hiện tại nhận diện 17 lớp PPE (Person, Helmet, Safety-vest, Gloves,..., Medical-suit, Face-mask-medical), nhưng phân bố nhãn không đồng đều (nhãn ít gặp như Medical-suit, Face-mask-medical) và 52% object nhỏ hơn 5% diện tích ảnh (như Gloves, Face-mask-medical) gây khó khăn trong việc tối ưu hóa hiệu suất. Hơn nữa, mô hình chưa có khả năng nhận diện nguy cơ nguy hiểm như leo cao không dây an toàn, làm việc không đúng tư thế, hoặc vi phạm quy trình an toàn, cũng như chưa hỗ trợ phân tích video thời gian thực, điều này hạn chế khả năng giám sát toàn diện. Thêm vào đó, hiệu suất của hệ thống giảm còn 85% trong điều kiện rung lắc mạnh, chẳng hạn khi camera gắn trên turbine bị ảnh hưởng bởi gió hoặc rung động cơ học, do các kỹ thuật tiền xử lý ảnh hiện tại chưa đủ mạnh để xử lý nhiễu động lớn.

3.6. Kết luận chương

Chương 3 đóng vai trò quan trọng trong việc cung cấp một cái nhìn toàn diện về hiệu quả và tiềm năng ứng dụng của hệ thống nhận diện PPE dựa trên mô hình YOLOv8m trong bối cảnh trang trại điện gió. Qua việc đánh giá các khía cạnh kỹ thuật và thực tiễn, chương này đặt nền tảng cho việc hiểu rõ điểm mạnh cũng như những thách thức hiện tại, đồng thời định hướng cho các nghiên cứu và triển khai tiếp theo. Nội dung được trình bày không chỉ phản ánh giá trị của giải pháp trong việc nâng cao an toàn lao động mà còn nhấn mạnh vai trò của nó trong việc hỗ trợ quản lý hiệu quả, mở ra các cơ hội phát triển bền vững trong ngành năng lượng tái tạo.

CHƯƠNG 4: KẾT LUẬN VÀ KIẾN NGHỊ

Sau khi đã xây dựng, thử nghiệm và đánh giá hệ thống nhận diện đồ bảo hộ cá nhân (PPE) sử dụng mô hình YOLOv8m qua các chương trước, phần này tổng hợp những giá trị cốt lõi mà nghiên cứu đã đạt được, đồng thời mở ra các hướng đi mới để khắc phục hạn chế và phát triển giải pháp trong tương lai. Từ cơ sở lý thuyết đến thực tiễn ứng dụng, nội dung chương 4 không chỉ khẳng định vai trò của hệ thống trong việc nâng cao an toàn lao động tại trang trại điện gió mà còn đề xuất các chiến lược tối ưu hóa, hướng tới việc đóng góp bền vững cho ngành năng lượng tái tạo.

4.1. Kết luận tổng quát

Nghiên cứu về hệ thống nhận diện đồ bảo hộ cá nhân sử dụng mô hình YOLOv8m đã đạt được những kết quả quan trọng, góp phần nâng cao an toàn lao động tại các trang trại điện gió và đặt nền móng cho tự động hóa trong ngành năng lượng tái tạo. Từ phân tích thực trạng và nhu cầu giám sát PPE trong môi trường phức tạp ở Chương 1, quá trình xây dựng và thử nghiệm hệ thống ở Chương 2, đến đánh giá hiệu quả và ứng dụng thực tế ở Chương 3, nghiên cứu chứng minh tính khả thi của giải pháp.

Hệ thống được huấn luyện trên dataset SH17 với 8.099 ảnh, nhận diện 17 lớp PPE như mũ bảo hộ, áo phản quang, găng tay, và khẩu trang y tế. Trong điều kiện phức tạp với ánh sáng yếu và rung lắc, mô hình đạt Precision 71.9%, Recall 62.2%, F1-score 66.6%, mAP@0.5 66.7%, và thời gian suy luận khoảng 100ms. Ở điều kiện tiêu chuẩn với ánh sáng tốt và ổn định, hiệu suất cải thiện rõ rệt với Precision 87.2%, Recall 83.1%, F1-score 89.0%, mAP@0.5 85.3%, và thời gian suy luận tương tự. So với YOLOv5 và Faster R-CNN, YOLOv8m vượt trội về tốc độ xử lý 40 khung hình/giây và khả năng nhận diện vật thể nhỏ, chiếm 52% dataset như găng tay và khẩu trang y tế. Hệ thống giảm thời gian kiểm tra thủ công từ 15 phút xuống 2 phút mỗi ca, hỗ trợ cảnh báo thời gian thực với độ trễ dưới 0.8 giây và độ chính xác báo cáo 95% so với kiểm tra thủ công, đáp ứng tiêu chuẩn ISO 45001. Việc tích hợp qua API REST vào hệ thống O&M mô phỏng tại Naju, Hàn Quốc và Đà Lạt, Việt Nam

thể hiện tiềm năng ứng dụng thực tế.

Tuy nhiên, hệ thống còn đối mặt với hạn chế về dữ liệu thực địa, đặc biệt với các nhãn ít gặp như khẩu trang y tế và áo y tế, cùng phân bố nhãn không đồng đều khi 52% vật thể nhỏ hơn 5% diện tích ảnh gây khó khăn trong nhận diện. Hiệu suất cũng giảm từ 87% xuống 72% trong điều kiện rung lắc mạnh do nhiễu động camera. Nghiên cứu không chỉ mang giá trị khoa học qua ứng dụng YOLOv8m mà còn góp phần nâng cao an toàn lao động và tối ưu hóa quản lý tại các trại điện gió. Để khắc phục hạn chế, cần thu thập thêm dữ liệu thực địa, đặc biệt cho các nhãn ít gặp và điều kiện rung lắc mạnh, đồng thời áp dụng thuật toán ổn định hình ảnh để cải thiện hiệu suất. Việc mở rộng nhận diện hành vi nguy hiểm, như leo cao không dây an toàn, sẽ đáp ứng nhu cầu giám sát toàn diện hơn. Tối ưu hóa phần cứng trên thiết bị nhúng hiệu suất cao cũng cần thiết để giảm chi phí và tăng khả năng mở rộng tại các khu vực hạ tầng hạn chế. Tích hợp với cảm biến IoT để thu thập dữ liệu môi trường như gió và ánh sáng sẽ hỗ trợ cảnh báo toàn diện hơn. Trong tương lai, hệ thống có thể phát triển thêm phân tích video thời gian thực và tự động hóa báo cáo an toàn, góp phần vào sự phát triển bền vững của ngành năng lượng tái tạo..

4.2. Kiến nghị

Dựa trên các hạn chế được nhận diện trong chương 3, các kiến nghị sau đây được đề xuất để nâng cao hiệu quả của hệ thống:

- **Thu thập và tích hợp dữ liệu thực địa:** Hợp tác với các đối tác như Công ty điện Cầu Đất (Việt Nam) và KEPCO (Hàn Quốc) để thu thập thêm 500 ảnh thực tế trở lên, bao gồm các điều kiện đặc thù như sương mù dày đặc, gió mạnh, và ánh sáng yếu vào ban đêm, nhằm huấn luyện lại mô hình để đạt độ chính xác cao hơn, dự kiến tăng mAP@0.5 lên 90-95%.
- **Mở rộng phạm vi nhận diện:** Nâng cấp YOLOv8m bằng cách tích hợp kỹ thuật tracking như DeepSORT để nhận diện các nguy cơ nguy hiểm

tiềm ẩn (ví dụ: leo cao không dây an toàn, làm việc không đúng tư thế) và hỗ trợ phân tích video thời gian thực với tốc độ 30 khung hình/giây, giúp giám sát toàn diện các rủi ro tại nơi làm việc.

- **Cải thiện hiệu suất trong điều kiện rung lắc:** Áp dụng các thuật toán ổn định hình ảnh tiên tiến, chẳng hạn như bộ lọc Kalman hoặc kỹ thuật bù nhiễu động, để nâng cao hiệu suất xử lý trong môi trường rung lắc mạnh, với mục tiêu khôi phục độ chính xác lên mức 90% hoặc cao hơn khi camera gắn trên turbine.

Ngoài việc khắc phục hạn chế hiện tại, các kiến nghị sau đây được đề xuất để mở rộng và tối ưu hóa hệ thống trong tương lai:

- Triển khai trên thiết bị nhúng: Tích hợp hệ thống vào các thiết bị như Jetson Nano (4GB RAM) hoặc Raspberry Pi 4, với mục tiêu xử lý 20 khung hình/giây, phù hợp cho các khu vực không có kết nối mạng mạnh như vùng sâu vùng xa, đồng thời tối ưu hóa mức tiêu thụ năng lượng xuống dưới 5W để tăng tính di động.
- Tích hợp cảm biến ngoại cảnh: Kết hợp với các cảm biến rung, nhiệt độ, và khí độc (CO, CO₂) để tăng cường khả năng dự đoán rủi ro, chẳng hạn cảnh báo ngạt khí hoặc nhiệt độ bất thường, với mục tiêu nâng cao độ chính xác thêm 5-10% trong việc phát hiện các mối nguy tiềm ẩn.
- Đánh giá và thử nghiệm các phiên bản nâng cao: Thử nghiệm các phiên bản YOLOv8l (40 triệu tham số) trên phần cứng mạnh như GPU Nvidia RTX 4090 để đạt mAP@0.5 96%, hoặc YOLOv8n (6 triệu tham số) cho các thiết bị tài nguyên hạn chế, nhằm tối ưu hóa hiệu suất trên nhiều nền tảng khác nhau.
- Tích hợp với IoT và Big Data: Xây dựng hệ thống giám sát an toàn toàn diện bằng cách kết nối với các nền tảng IoT để thu thập dữ liệu thời gian thực từ nhiều turbine, kết hợp với phân tích Big Data để dự đoán xu

hướng rủi ro, góp phần xây dựng chiến lược quản lý an toàn dài hạn trong ngành năng lượng tái tạo.

4.3. Hướng phát triển trong tương lai

Hệ thống nhận diện PPE mở ra tiềm năng lớn cho các giải pháp an toàn thông minh, không chỉ trong ngành điện gió mà còn ở nhiều lĩnh vực công nghiệp nguy hiểm khác. Về kỹ thuật, nghiên cứu có thể phát triển mô hình học sâu như Vision Transformer (ViT) hoặc EfficientDet để cải thiện nhận diện PPE trong môi trường phức tạp, như mỏ than bụi bẩn, nhà máy luyện kim ánh sáng yếu, hay trang trại điện gió ngoài khơi sương mù dày đặc. Kết hợp công nghệ thực tế tăng cường (AR) qua kính thông minh như Microsoft HoloLens, hệ thống có thể hiển thị thông tin an toàn thời gian thực, nâng cao giám sát và phản ứng rủi ro.

Về ứng dụng đa ngành, hệ thống có thể tùy chỉnh cho ngành xây dựng để giám sát mũ bảo hộ, dây đai an toàn tại công trường cao tầng; ngành hóa chất để phát hiện găng tay chống hóa chất, mặt nạ phòng độc; hoặc ngành hàng hải để kiểm tra áo phao, trang bị cứu sinh tại cảng biển như Singapore, gửi cảnh báo tức thời khi vi phạm PPE. Hệ thống cũng có thể tích hợp cảm biến IoT để phát hiện rò rỉ khí gas, nhiệt độ bất thường, tạo hệ sinh thái giám sát toàn diện.

Về dài hạn, AI 3.0 và mạng 6G sẽ cho phép phân tích dữ liệu thời gian thực với độ trễ gần bằng không, phù hợp cho môi trường yêu cầu phản hồi nhanh. Blockchain đảm bảo bảo mật, truy xuất dữ liệu an toàn lao động trong dự án quốc tế như Công ty Điện Cầu Đất và KEPCO. Hệ thống còn có thể dự báo sự cố turbine dựa trên vi phạm PPE, tối ưu hóa bảo trì, giảm chi phí, và kéo dài tuổi thọ thiết bị.

TÀI LIỆU THAM KHẢO

- [1] Arafat M Y Ismail A, (2024), “Personal protective equipment detection using YOLOv8 architecture on object detection benchmark datasets: a comparative study,” *PeerJ Computer Science*, 11(1), 2333209.
- [2] Ferdous Md Ahsan Sk Md Masudul, (2022), “PPE detector: A YOLO-based architecture to detect personal protective equipment (PPE) for construction sites,” *PeerJ Computer Science*, 8(999), 1–26.
- [3] Liu W Anguelov D Erhan D Szegedy C Reed S Fu C Y Berg A C, (2016), “SSD: Single Shot MultiBox Detector,” *arXiv preprint arXiv:1512.02325*.
- [4] Nazli N A N M Sabria N Aminuddin R Ibrahim S Yusof S Nasir S D N M, (2024), “A real-time system for detecting personal protective equipment compliance using deep learning model YOLOv5,” *Procedia Computer Science*, 245(2024), 647–656.
- [5] Ren S He K Girshick R Sun J, (2015), “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *arXiv preprint arXiv:1506.01497*.
- [6] [ILO] International Labour Organization, (2019), *Safety and Health at the Heart of the Future of Work: Building on 100 years of experience*, International Labour Organization, Geneva, Thụy Sĩ.
- [7] [OSHA] Occupational Safety and Health Administration, *Green Job Hazards - Wind Energy*.
- [8] Ultralytics, (2023), *YOLOv8 Docs: Architecture, Loss Function, and Inference*.