

**BỘ GIÁO DỤC
VÀ ĐÀO TẠO**

**VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM**

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



Trịnh Hiền Anh

**NGHIÊN CỨU CÁC KỸ THUẬT PHÁT HIỆN, NHẬN DIỆN
BẤT THƯỜNG THIẾT BỊ TRÊN LƯỚI ĐIỆN CAO THẾ TỪ
ẢNH THU NHẬN BẰNG UAV DỰA TRÊN HỌC SÂU**

TÓM TẮT LUẬN ÁN TIẾN SĨ MÁY TÍNH

Ngành: Hệ thống thông tin

Mã số: 9 48 01 04

Hà Nội- 2026

Công trình được hoàn thành tại: Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

Người hướng dẫn khoa học:

Người hướng dẫn 1: TS. Nguyễn Thị Thanh Tân, Khoa Công nghệ thông tin, Trường Đại học Điện lực

Người hướng dẫn 2: PGS.TS. Ngô Quốc Tạo, Viện Công nghệ thông tin, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

Phản biện 1:

Phản biện 2:.

Phản biện 3:

Luận án được bảo vệ trước Hội đồng đánh giá luận án tiến sĩ cấp Học viện họp tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam vào hồi giờ , ngày tháng năm 20

Có thể tìm hiểu luận án tại:

1. Thư viện Học viện Khoa học và Công nghệ
2. Thư viện Quốc gia Việt Nam

MỞ ĐẦU

1. Tính cấp thiết

Trong bối cảnh nhu cầu điện năng ngày càng tăng, việc đảm bảo vận hành an toàn và tin cậy của lưới truyền tải điện cao thế có ý nghĩa quan trọng đối với an ninh năng lượng quốc gia. Sự phát triển của các công nghệ giám sát như UAV, camera độ phân giải cao và ảnh nhiệt đã giúp thu thập dữ liệu kiểm tra hiệu quả hơn. Tuy nhiên, khối lượng lớn ảnh và video tạo ra vượt quá khả năng phân tích thủ công, trong khi việc phát hiện các hư hỏng nhỏ vẫn phụ thuộc nhiều vào chuyên gia, làm tăng chi phí, thời gian và nguy cơ bỏ sót bất thường.

Học sâu và thị giác máy tính đã đạt nhiều thành tựu trong phát hiện đối tượng, nhưng khi áp dụng cho dữ liệu UAV trong ngành điện vẫn gặp nhiều thách thức như: thiết bị có kích thước nhỏ và đa tỷ lệ, nền ảnh phức tạp, điều kiện ánh sáng và thời tiết thay đổi, hiện tượng che khuất, cùng với sự mất cân bằng dữ liệu do số lượng mẫu hư hỏng rất hạn chế. Ngoài ra, các mô hình còn phải đáp ứng yêu cầu xử lý nhanh và hiệu quả trên các nền tảng tính toán hạn chế như UAV hoặc thiết bị biên.

Vì vậy, luận án tập trung vào đề tài: “Nghiên cứu các kỹ thuật phát hiện, nhận diện bất thường thiết bị trên lưới điện cao thế từ ảnh thu bằng UAV dựa trên học sâu”. Nghiên cứu góp phần nâng cao hiệu quả giám sát, phát hiện sớm hư hỏng, hỗ trợ bảo trì dự báo, thúc đẩy chuyển đổi số trong ngành điện lực và nâng cao độ tin cậy vận hành của hệ thống điện quốc gia.

2. Mục tiêu nghiên cứu

Đề xuất các mô hình học sâu nhằm nâng cao hiệu quả phát hiện và nhận diện bất thường trên thiết bị lưới truyền tải điện từ ảnh UAV trong điều kiện đối tượng nhỏ, nền phức tạp và dữ liệu mất cân bằng.

Mục tiêu cụ thể: Để đạt được mục tiêu tổng quát nêu trên, luận án tập trung vào các mục tiêu cụ thể sau:

- Nghiên cứu các mô hình phát hiện đối tượng dựa trên học sâu cho ảnh UAV;
- Đề xuất mô hình phát hiện và nhận diện bất thường trên chuỗi cách điện;
- Nghiên cứu kỹ thuật xử lý mất cân bằng dữ liệu nhằm nâng cao hiệu quả huấn luyện;
- Xây dựng và đánh giá thực nghiệm hệ thống trên bộ dữ liệu UAV thực tế.

3. Nội dung nghiên cứu

Để đạt được mục tiêu đề ra, các nội dung sau được NCS quan tâm

- Nghiên cứu hiện trạng giám sát thiết bị trên lưới truyền tải điện;
- Phân tích các mô hình học sâu trong phát hiện đối tượng từ ảnh UAV;

- Xây dựng và đánh giá bộ dữ liệu thiết bị điện thu nhận bằng UAV;
- Đề xuất các giải pháp nâng cao hiệu quả phát hiện và nhận diện bất thường,

bao gồm:

- Mô hình phát hiện thiết bị trên lưới truyền tải điện;
- Mô hình phát hiện bất thường trên chuỗi cách điện;
- Kỹ thuật tăng cường và cân bằng dữ liệu.

4. Đối tượng, phạm vi nghiên cứu, phương pháp nghiên cứu

Đối tượng nghiên cứu

Với mục tiêu đã đề ra của luận án, đối tượng nghiên cứu của luận án bao gồm:

- Các mô hình học sâu trong phát hiện đối tượng từ ảnh thu nhận bằng UAV;
- Chuỗi cách điện và các dạng bất thường thường gặp trên lưới truyền tải điện;
- Các mô hình phát hiện đối tượng và phát hiện bất thường dựa trên học sâu;
- Các chiến lược xử lý mất cân bằng dữ liệu trong thị giác máy tính.

Phạm vi nghiên cứu

Luận án tập trung vào bài toán phát hiện, nhận diện bất thường thiết bị trên lưới điện cao thế từ ảnh quang học thu nhận bằng UAV trong môi trường thực tế, không mở rộng sang các bài toán dự báo theo chuỗi thời gian dài hạn hoặc mô phỏng hệ thống điện.

Phương pháp nghiên cứu

Luận án kết hợp nghiên cứu lý thuyết và thực nghiệm, dựa trên việc tổng hợp các tài liệu khoa học uy tín từ các tạp chí, hội thảo và cơ sở nghiên cứu. Các mô hình đề xuất được xây dựng, cài đặt và huấn luyện, sau đó đánh giá trên dữ liệu thực tế và so sánh với các phương pháp tiên tiến để kiểm chứng hiệu quả.

5. Đóng góp của nghiên cứu

Sau những nỗ lực nghiên cứu, luận án có những đóng góp khoa học và thực tiễn như sau:

- Xây dựng mô hình cải tiến HVE-YOLO nhằm nâng cao hiệu quả phát hiện các thiết bị trên lưới truyền tải điện từ ảnh UAV;
- Phát triển khung học sâu tích hợp phục vụ bài toán phát hiện và phân loại bất thường trên chuỗi cách điện trong điều kiện đối tượng nhỏ và nền ảnh phức tạp;
- Xây dựng chiến lược tăng cường dữ liệu dựa trên Attention-Aware Prompting nhằm khắc phục tình trạng mất cân bằng dữ liệu trong quá trình huấn luyện.

6. Cấu trúc của luận án

Luận án gồm phần mở đầu, bốn chương nội dung và phần kết luận. Phần mở đầu trình bày mục tiêu, phạm vi, phương pháp và các đóng góp của nghiên cứu. Chương 1 tổng quan bài toán giám sát lưới điện cao thế từ ảnh UAV, phân tích các phương pháp hiện có và xác định khoảng trống nghiên cứu. Chương 2 đề xuất mô hình HVE-YOLO nhằm phát hiện thiết bị trên lưới điện từ ảnh UAV. Chương 3 nghiên cứu mô hình FGS để phát hiện và phân đoạn bất thường trên chuỗi cách điện. Chương 4 đề xuất AAPID theo khuôn khổ FSCIL, nhằm nhận diện các lớp lỗi mới với số lượng mẫu hạn chế và giảm hiện tượng quên lãng thảm họa. Cuối cùng, phần kết luận tổng hợp các kết quả, đóng góp và định hướng nghiên cứu tiếp theo.

CHƯƠNG 1. GIỚI THIỆU BÀI TOÁN PHÁT HIỆN VÀ NHẬN DIỆN BẤT THƯỜNG THIẾT BỊ TRÊN LƯỚI ĐIỆN TỪ ẢNH UAV

1.1. Giới thiệu bài toán

Sử dụng UAV kết hợp với thị giác máy tính và học sâu là hướng tiếp cận tiềm năng nhằm tự động hóa giám sát lưới truyền tải điện. Tuy nhiên, ảnh UAV có nhiều thách thức như đối tượng nhỏ, góc nhìn thay đổi, nền ảnh phức tạp và thiếu dữ liệu bất thường.

Luận án tập trung giải quyết ba bài toán chính:

Phát hiện thiết bị trên lưới truyền tải từ ảnh UAV.

Nhận diện bất thường trên chuỗi cách điện trong môi trường phức tạp.

Nâng cao khả năng học khi dữ liệu bất thường bị hạn chế.

1.1.1. Phát hiện thiết bị trên lưới truyền tải điện cao thế

1.1.2. Phát hiện bất thường trên chuỗi cách điện

Chuỗi cách điện trong hệ thống truyền tải điện cao thế có nhiệm vụ cách ly dây dẫn với cột điện và chịu tải cơ học, ảnh hưởng trực tiếp đến an toàn và độ tin cậy của hệ thống

Các bất thường này thường có kích thước nhỏ, xuất hiện cục bộ và không đồng đều, khiến việc phát hiện sớm trở nên khó khăn nếu chỉ dựa vào kiểm tra thủ công.

1.1.3. Đặc điểm và thách thức của dữ liệu UAV trong giám sát lưới điện

UAV như nền tảng thu nhận dữ liệu cho giám sát lưới điện

Đặc trưng của ảnh UAV trong giám sát lưới điện

Dữ liệu ảnh UAV có một số đặc trưng nổi bật: (i) độ phân giải cao và yêu cầu khai thác đặc trưng đa tỷ lệ; (ii) biến thiên lớn về góc nhìn, hình học và phối cảnh; (iii) nhiều đối tượng mục tiêu có kích thước nhỏ và dạng mảnh; (iv) nền ảnh phức tạp, đa dạng; và (v) điều kiện chiếu sáng, thời tiết thay đổi. Những đặc

điểm này làm gia tăng độ khó của bài toán phát hiện thiết bị và bất thường trên lưới điện, đồng thời đặt ra yêu cầu về các kiến trúc học sâu có khả năng bảo toàn thông tin không gian, tăng cường đặc trưng đa tỷ lệ và thích nghi tốt với môi trường thực tế.

1.2. Các nghiên cứu liên quan

Các nghiên cứu hiện nay về phát hiện thiết bị và bất thường trên lưới điện chủ yếu sử dụng học sâu và thị giác máy tính, có thể chia thành ba hướng chính:

- Ảnh thông thường: chất lượng ảnh cao, đối tượng rõ nhưng phạm vi quan sát hạn chế và chưa phản ánh đầy đủ điều kiện thực tế;
- Ảnh UAV: có phạm vi quan sát rộng và phù hợp kiểm tra thực địa, nhưng gặp khó khăn do đối tượng nhỏ, nền phức tạp, góc nhìn và ánh sáng thay đổi, mất cân bằng dữ liệu.
- Sinh dữ liệu tổng hợp: nhằm bổ sung các mẫu bất thường khan hiếm, cân bằng dữ liệu và cải thiện khả năng tổng quát hóa, đặc biệt với dữ liệu có phân phối dài đuôi.

Nhìn chung, dù các hướng trên đã đạt nhiều kết quả tích cực, vẫn còn khoảng trống trong việc kết hợp dữ liệu UAV, phát hiện đối tượng nhỏ và sinh dữ liệu bất thường. Đây là cơ sở để nghiên cứu các phương pháp học sâu và sinh dữ liệu nhằm nâng cao khả năng phát hiện, nhận diện và chẩn đoán bất thường trên chuỗi cách điện.

1.2.1. Phát hiện thiết bị và bất thường từ ảnh thông thường

Các phương pháp phát hiện đối tượng hiện nay chủ yếu được phát triển theo ba hướng: mô hình hai giai đoạn, mô hình một giai đoạn và các phương pháp dựa trên Transformer. Mô hình hai giai đoạn thực hiện đề xuất vùng trước khi phân loại và hồi quy khung bao, trong khi mô hình một giai đoạn dự đoán trực tiếp vị trí và lớp đối tượng, nhờ đó có ưu thế về tốc độ. Các phương pháp Transformer khai thác cơ chế attention để tăng khả năng mô hình hóa quan hệ giữa các đặc trưng và nâng cao hiệu quả phát hiện.

1.2.2. Phát hiện thiết bị và bất thường từ ảnh UAV

Các nghiên cứu sử dụng ảnh UAV cho giám sát lưới điện tập trung giải quyết ba thách thức chính: phát hiện đối tượng kích thước nhỏ, xử lý nền ảnh phức tạp và đảm bảo suy luận thời gian thực trên nền tảng tính toán hạn chế. Các hướng tiếp cận chủ yếu gồm mô hình hai giai đoạn, mô hình một giai đoạn và các kiến trúc Transformer hoặc mô hình lai. Tuy nhiên, sự thay đổi về góc quan sát, kích thước đối tượng, điều kiện chiếu sáng và môi trường nền vẫn gây nhiều khó khăn cho việc duy trì độ chính xác và khả năng tổng quát hóa.

1.3. Các phương pháp sinh dữ liệu tổng hợp nhằm giải quyết tình trạng khan hiếm mẫu bất thường và mất cân bằng dữ liệu

Dữ liệu giám sát lưới điện thường có sự mất cân bằng nghiêm trọng, trong đó mẫu bình thường chiếm tỷ lệ lớn còn các mẫu bất thường hiếm gặp, thậm chí dưới 1%. Sự mất cân bằng này khiến mô hình có xu hướng ưu tiên lớp đa số, dẫn đến Recall và F1-score thấp đối với các trường hợp bất thường. Các phương pháp sinh dữ liệu tổng hợp được nghiên cứu nhằm bổ sung mẫu hiếm, cân bằng phân bố dữ liệu và cải thiện khả năng phát hiện các dạng hư hỏng khó thu thập trong thực tế.

1.3.1. Khoảng trống nghiên cứu

Từ tổng quan các nghiên cứu có thể xác định ba khoảng trống chính: (i) khả năng phát hiện vật thể nhỏ trong nền ảnh UAV phức tạp còn hạn chế; (ii) chưa có giải pháp hiệu quả cho tình trạng mất cân bằng dữ liệu nghiêm trọng; và (iii) dữ liệu tổng hợp chưa được khai thác hiệu quả và còn thiếu sự tích hợp tri thức miền lưới điện.

1.3.2. Các câu hỏi nghiên cứu

Ba khoảng trống trên dẫn đến ba câu hỏi nghiên cứu chính:

- RQ1: Làm thế nào nâng cao khả năng phát hiện thiết bị nhỏ trong ảnh UAV có nền phức tạp và góc quan sát thay đổi?
- RQ2: Làm thế nào nhận diện chính xác các bất thường nhỏ, cục bộ trên chuỗi cách điện khi ranh giới giữa bình thường và bất thường không rõ ràng?
- RQ3: Làm thế nào khắc phục mất cân bằng dữ liệu và nâng cao chất lượng, tính hữu ích của dữ liệu tổng hợp bằng cách tích hợp tri thức miền lưới điện vào mô hình sinh?

1.3.3. Các nội dung nghiên cứu và kết quả chính của luận án

Để trả lời ba câu hỏi nghiên cứu, luận án tập trung vào ba nội dung chính: (i) đề xuất mô hình **HVE-YOLO** nhằm nâng cao khả năng phát hiện thiết bị kích thước nhỏ trong ảnh UAV; (ii) xây dựng khung **FGS** kết hợp phát hiện và phân đoạn để nhận diện bất thường chi tiết trên chuỗi cách điện; và (iii) đề xuất phương pháp **AAPID** nhằm học và chẩn đoán bất thường trong điều kiện dữ liệu hạn chế, đồng thời khai thác dữ liệu sinh tổng hợp và tri thức miền.

1.4. Kết luận chương

Chương 1 đã trình bày cơ sở lý thuyết, tổng quan các phương pháp và đặc điểm của bài toán giám sát lưới truyền tải điện từ ảnh UAV. Các thách thức cốt lõi gồm đối tượng kích thước nhỏ, nền ảnh phức tạp, biến thiên góc nhìn và mất cân bằng dữ liệu đã được phân tích. Trên cơ sở đó, luận án xác định ba khoảng

trồng nghiên cứu và xây dựng ba hướng giải quyết được trình bày lần lượt trong các chương tiếp theo.

CHƯƠNG 2. PHÁT HIỆN THIẾT BỊ TRÊN LƯỚI ĐIỆN CAO THỂ TỪ ẢNH UAV

Chương này tập trung vào phát hiện thiết bị lưới điện cao thể từ ảnh UAV, nhằm hỗ trợ kiểm tra, giám sát và bảo trì hệ thống điện. Bài toán gặp nhiều thách thức do đối tượng nhỏ, hình dạng mảnh, nền phức tạp và điều kiện chiếu sáng thay đổi, khiến các mô hình như Faster R-CNN và YOLO suy giảm hiệu quả.

Để giải quyết vấn đề, nghiên cứu đề xuất HVE-YOLO, một mô hình cải tiến dựa trên YOLOv11, chuyên biệt cho việc phát hiện sáu loại thiết bị cốt lõi trên lưới điện cao thể. Mô hình được đánh giá trên dữ liệu thu thập tại công trình CT6.

2.1. Các nghiên cứu liên quan về phát hiện thiết bị trên lưới điện

2.1.1. Nghiên cứu phát hiện thiết bị trên lưới điện từ ảnh UAV

Trong những năm gần đây, công nghệ thiết bị bay không người lái (UAV) đã chứng minh được hiệu quả vượt trội trong việc giám sát và thu thập dữ liệu hình ảnh các thành phần trên lưới điện cao thể [1, 2, 3].

2.1.2. Phát hiện thiết bị trên lưới điện sử dụng mô hình một giai đoạn

Với khả năng phát hiện đối tượng theo thời gian thực, độ chính xác cao và chi phí tính toán tương đối tối ưu, các phiên bản YOLO (CT1, CT4) đã được ứng dụng rộng rãi trong nhận dạng chuỗi cách điện, phụ kiện đường dây, điểm phát nhiệt bất thường cũng như các dạng hư hỏng trên thiết bị điện. Các nghiên cứu gần đây cho thấy YOLO không chỉ cải thiện hiệu quả tự động hóa công tác kiểm tra mà còn góp phần giảm thiểu rủi ro cho nhân viên vận hành và nâng cao độ tin cậy của hệ thống điện.

Kiến trúc backbone Yolo11

Hạn chế mô hình

Mặc dù YOLOv11 đạt được những cải tiến đáng kể về độ chính xác và tốc độ suy luận trong các bài toán phát hiện đối tượng thời gian thực, mô hình vẫn bộc lộ một số hạn chế khi áp dụng cho bài toán phát hiện thiết bị trên lưới truyền tải điện cao thể. Một trong những thách thức quan trọng là khả năng nhận diện các đối tượng có kích thước nhỏ như chốt khóa, spacer, bộ giảm chấn hoặc các vết nứt nhỏ trên chuỗi cách điện.

Một hạn chế đáng chú ý khác của YOLOv11 là sự phụ thuộc lớn vào tập dữ liệu huấn luyện có chất lượng cao và được gán nhãn đầy đủ.

2.2. Mô hình HVE-YOLO (High Voltage Equipment Detection)

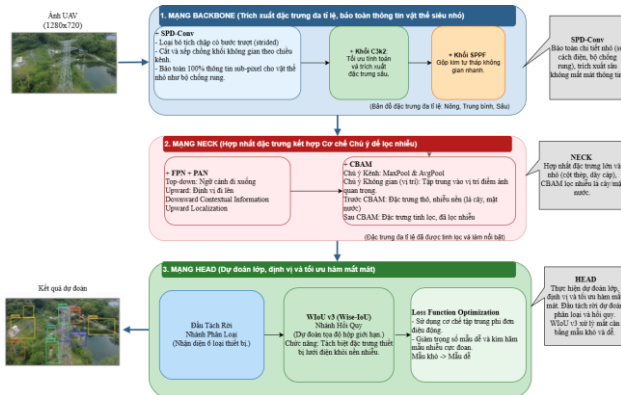
Nghiên cứu sinh đề xuất HVE-YOLO nhằm nâng cao khả năng phát hiện thiết bị trên lưới truyền tải điện cao thế từ ảnh UAV, kế thừa kiến trúc Backbone–Neck–Head và cơ chế dự đoán không anchor của YOLO11. Mô hình được cải tiến với ba thành phần chính: (i) SPD-Conv thay thế downsampling bằng stride để bảo toàn thông tin không gian và tăng khả năng phát hiện đối tượng nhỏ; (ii) CBAM sau hợp nhất đặc trưng đa tỷ lệ nhằm giảm ảnh hưởng của nhiễu nền và tăng cường đặc trưng quan trọng; (iii) Wise-IoU v3 nhằm nâng cao chất lượng hồi quy hộp bao, đặc biệt đối với các mẫu khó.

2.2.1. Mô hình HVE-YOLO

Mô hình HVE-YOLO [CT6] được xây dựng nhằm nâng cao hiệu quả phát hiện và nhận dạng thiết bị trên lưới truyền tải điện cao thế từ ảnh UAV. Mô hình có khả năng phát hiện sáu loại thiết bị cơ bản trên lưới điện, bao gồm: *cột đơn thân, cột thép hình, dây dẫn điện, tạ chống rung, chuỗi cách điện thủy tinh và chuỗi cách điện silicon*. Đây đều là các thành phần quan trọng trong công tác giám sát vận hành và bảo trì lưới điện truyền tải.

2.2.2. Kiến trúc tổng thể của mô hình HVE-YOLO

Kiến trúc tổng thể của mô hình được xây dựng theo hướng tối ưu hóa khả năng phát hiện đối tượng nhỏ và nâng cao độ bền vững của mô hình trong các điều kiện môi trường phức tạp. Kiến trúc mô hình HVE-YOLO được mô tả trong Hình 2.1 với ba thành phần chính Backbone-Neck-Head.



Hình 2.1: Kiến trúc mô hình HVE-YOLO

Ảnh đầu vào thu được từ UAV trước tiên được đưa vào Backbone để trích xuất các đặc trưng phân cấp. Các đặc trưng này sau đó được tổng hợp thông qua Neck bằng cơ chế Feature Pyramid Network (FPN) và Path Aggregation

Network (PAN), đồng thời được tinh chỉnh bằng Attention CBAM. Cuối cùng, Detection Head thực hiện đồng thời ba nhiệm vụ: phân loại đối tượng, hồi quy hộp bao và ước lượng xác suất tồn tại của đối tượng.

2.2.3. Các cải tiến trong HVE-YOLO

Cải tiến backbone nhằm tăng khả năng phát hiện vật thể nhỏ

Backbone HVE-YOLO được cải tiến với SPD-Conv, C3k2 và SPPF nhằm bảo toàn thông tin không gian, tăng khả năng biểu diễn đặc trưng và khai thác ngữ cảnh đa tỷ lệ. SPD-Conv hạn chế mất mát thông tin khi giảm mẫu, đặc biệt hiệu quả với các đối tượng nhỏ và mảnh trong ảnh UAV. C3k2 tăng cường khả năng học đặc trưng và phân biệt đối tượng với nền phức tạp, trong khi SPPF mở rộng trường cảm thụ và tổng hợp thông tin ở nhiều tỷ lệ với chi phí tính toán thấp.

Sự kết hợp ba thành phần giúp mô hình nâng cao khả năng phát hiện đối tượng nhỏ, bị che khuất hoặc có độ tương phản thấp, đồng thời duy trì độ phức tạp hợp lý và khả năng suy luận thời gian thực, phù hợp với GPU nhúng và nền tảng UAV Edge AI.

CBAM Attention

HVE-YOLO sử dụng FPN và PAN để hợp nhất đặc trưng đa tỷ lệ, kết hợp thông tin ngữ nghĩa với đặc trưng không gian nhằm nâng cao khả năng phát hiện các thiết bị có kích thước đa dạng, đặc biệt là đối tượng nhỏ. Mô hình tiếp tục tích hợp CBAM gồm Channel Attention và Spatial Attention để tập trung vào các đặc trưng và vị trí quan trọng, đồng thời giảm ảnh hưởng của nền cảnh phức tạp.

Sự kết hợp FPN, PAN và CBAM giúp HVE-YOLO tăng khả năng phát hiện đối tượng nhỏ, bị che khuất hoặc có độ tương phản thấp, trong khi vẫn duy trì kiến trúc nhẹ và chi phí tính toán phù hợp cho suy luận thời gian thực.

Cải tiến hàm mất mát định vị

Để khắc phục hạn chế cho nhận dạng thiết bị điện có kích thước nhỏ, HVE-YOLO sử dụng Wise-IoU phiên bản WIoU-v3 làm hàm mất mát định vị. WIoU-v3 áp dụng cơ chế tập trung động phi đơn điệu (*dynamic non-monotonic focusing*), trong đó mức độ đóng góp của mỗi mẫu được điều chỉnh thông qua hệ số khuếch đại gradient. Cơ chế này giúp giảm ảnh hưởng của các mẫu quá dễ hoặc quá khó, đồng thời tập trung tối ưu vào các mẫu có độ khó phù hợp và có khả năng mang lại thông tin học hữu ích.

Tổng hàm mất mát của HVE-YOLO được xác định bởi

$$\mathcal{L} = \lambda_1 \mathcal{L}_{WIoU-v3} + \lambda_2 \mathcal{L}_{cls} + \lambda_3 \mathcal{L}_{obj} \quad (2.1)$$

trong đó $\mathcal{L}_{WIoU-v3}$ là mất mát định vị, $\mathcal{L}_{cls}$ là mất mát phân loại, $\mathcal{L}_{obj}$ là

mất mát xác suất tồn tại đối tượng, còn $\lambda_1, \lambda_2, \lambda_3$ là các hệ số cân bằng.

$$\text{với hệ số tập trung } r = \frac{\beta}{\delta \cdot \alpha(\beta - \delta)} \quad \beta = \frac{\mathcal{L}_{IoU}}{\mathcal{L}_{IoU}} \quad (2.3)$$

trong đó phản ánh mức độ ngoại lai của mẫu, $\overline{\mathcal{L}_{IoU}}$ là giá trị trung bình động của mất mát IoU, còn α và δ là các siêu tham số điều khiển cơ chế tập trung động.

Khác với cơ chế tăng trọng số đơn điệu theo độ khó của mẫu, WIoU-v3 sử dụng tập trung phi đơn điệu để hạn chế ảnh hưởng của các mẫu nhiễu hoặc quá khó, đồng thời ưu tiên các mẫu có chất lượng học phù hợp. Nhờ đó, HVE-YOLO cải thiện khả năng hồi quy khung bao đối với các thiết bị nhỏ, bị che khuất hoặc quan sát dưới góc phức tạp, góp phần nâng cao độ chính xác và tính ổn định của mô hình trong môi trường giám sát lưới điện bằng UAV.

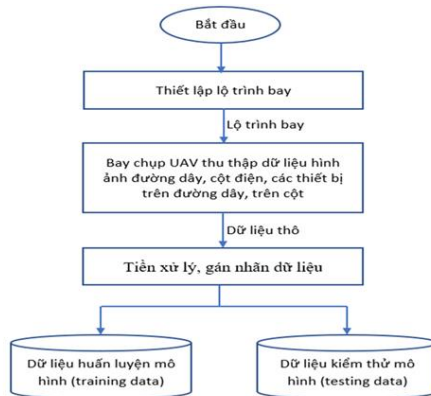
2.2.4. Quy trình huấn luyện mô hình HVE-YOLO

Quy trình xây dựng và huấn luyện mô hình HVE-YOLO được thực hiện qua ba giai đoạn chính: (i) chuẩn bị dữ liệu, (ii) xây dựng và huấn luyện mô hình, và (iii) đánh giá mô hình. Các giai đoạn được tổ chức liên kết nhằm bảo đảm mô hình đạt độ chính xác cao và có khả năng tổng quát hóa khi triển khai trong môi trường thực tế.

Chuẩn bị dữ liệu

Dữ liệu được thu thập bằng thiết bị bay không người lái (UAV) trong quá trình khảo sát thực tế trên lưới điện truyền tải cao thế theo quy trình trong Hình 2.2.

Hình 2.2: Quy trình thu thập dữ liệu trên lưới điện cao thế bằng thiết bị bay không người lái



Dữ liệu được chia thành tập huấn luyện và tập kiểm chứng theo tỷ lệ 80%/20% bằng phương pháp phân tầng, bảo đảm sự phân bố tương đối đồng đều giữa các lớp thiết bị. Đồng thời, các ảnh được thu nhận liên tiếp tại cùng vị trí được kiểm tra để hạn chế rò rỉ dữ liệu giữa hai tập. Quy trình này góp phần bảo đảm kết quả đánh giá phản ánh khách quan khả năng tổng quát hóa của HVE-YOLO trên dữ liệu chưa được sử dụng trong quá trình huấn luyện.

Xây dựng và huấn luyện mô hình

HVE-YOLO được huấn luyện từ tham số tiền huấn luyện nhằm rút ngắn thời gian hội tụ và nâng cao khả năng phát hiện thiết bị nhỏ trong nền ảnh phức tạp.

Quá trình huấn luyện được mô tả trong thuật toán 1.

Thuật toán 1: Quy trình huấn luyện và suy luận HVE-YOLO

Đầu vào: Tập huấn luyện $T = \{(I_k, G_k)\}_{k=1}^M$, tốc độ học η , số epoch E

Đầu ra: Bộ tham số tối ưu Θ^* và kết quả phát hiện $\mathcal{D}_k$.

1: 1. Khởi tạo tham số mô hình Θ .

2. For epoch = 1 to E:

2.1. For mỗi lô dữ liệu $\mathcal{B}$:

$F_b \leftarrow \text{Backbone}(I_k)$.

$F_n \leftarrow \text{Neck}(F_b)$.

$(P_b, P_c, P_o) \leftarrow \text{Head}(F_n)$.

$\mathcal{L}_{\text{box}} \leftarrow \mathcal{L}_{\text{WIoU-v3}}(P_b, G_k)$.

$\mathcal{L}_{\text{cls}} \leftarrow \mathcal{L}_{\text{cls}}(P_c, G_k)$.

$\mathcal{L}_{\text{obj}} \leftarrow \mathcal{L}_{\text{obj}}(P_o, G_k)$.

$\mathcal{L}_{\mathcal{B}} \leftarrow \lambda_1 \mathcal{L}_{\text{box}} + \lambda_2 \mathcal{L}_{\text{cls}} + \lambda_3 \mathcal{L}_{\text{obj}}$.

$\Theta \leftarrow \Theta - \eta \nabla_{\Theta} \mathcal{L}_{\mathcal{B}}$.

3. $\Theta^* \leftarrow \Theta$.

4. Giai đoạn suy luận:

5. For mỗi ảnh thử nghiệm I_k :

$F_b \leftarrow \text{Backbone}(I_k)$.

$F_n \leftarrow \text{Neck}(F_b)$.

$(P_b, P_c, P_o) \leftarrow \text{Head}(F_n)$.

$S \leftarrow P_c \odot P_o$.

$\mathcal{D}_k \leftarrow \text{NMS}(P_b, S)$.

6. Return $\{\Theta^*, \mathcal{D}_k\}$.

Đánh giá mô hình

HVE-YOLO được đánh giá trên tập kiểm chứng bằng các chỉ số Precision, Recall, F1-score, AP, mAP và FPS, kết hợp phân tích trực quan khả năng phát hiện đối tượng nhỏ, bị che khuất và trong nền phức tạp. Kết quả đạt $mAP@0.5 = 0,972$, cho thấy hiệu quả phát hiện cao trong điều kiện dữ liệu nghiên cứu.

Tuy nhiên, kết quả chưa thể trực tiếp khẳng định khả năng tổng quát trên các tuyến 220 kV và 500 kV, do khoảng cách UAV lớn hơn có thể làm giảm kích thước đối tượng trên ảnh và tăng độ khó phát hiện. Khả năng duy trì hiệu năng phụ thuộc vào chất lượng ảnh, độ phân giải, độ ổn định của UAV và khả năng bảo toàn đặc trưng chi tiết.

Trong tương lai, cần mở rộng dữ liệu theo nhiều cấp điện áp và khoảng cách quan sát, đồng thời nghiên cứu siêu phân giải, phát hiện vật thể siêu nhỏ và hợp nhất dữ liệu đa cảm biến để nâng cao độ tin cậy khi giám sát từ xa.

2.3. Thử nghiệm

Nghiên cứu xây dựng môi trường thử nghiệm nhằm đánh giá khả năng triển khai mô hình trên UAV, với mô hình được huấn luyện bằng Python 3 và PyTorch, sau đó triển khai trên NVIDIA Jetson Orin Nano và tối ưu bằng TensorRT. Mô hình được chuyển đổi sang FP16 và INT8 kết hợp PTQ để đánh giá độ chính xác, tốc độ suy luận, mức sử dụng tài nguyên và khả năng triển khai Edge AI.

2.3.1. Dữ liệu thử nghiệm

Bộ dữ liệu được thu thập bằng UAV DJI Mavic 2 Pro trên lưới điện 110 kV, trong nhiều điều kiện môi trường. Bộ dữ liệu có 69.955 đối tượng thuộc 6 lớp, gồm dây dẫn, chuỗi cách điện silicon, chuỗi cách điện thủy tinh, cột thép, cột đơn và tạ chống rung. Trong đó, 53.784 đối tượng dùng để huấn luyện và 16.171 đối tượng dùng để kiểm tra, với nhãn được lưu theo định dạng YOLO.

2.3.2. Kịch bản thử nghiệm

HVE-YOLO được đánh giá thông qua 5 nhóm thử nghiệm nhằm kiểm chứng toàn diện khả năng của mô hình:

(1) Độ chính xác tổng thể: Huấn luyện với tỷ lệ 80/20%, ảnh đầu vào 640×640 , optimizer AdamW, learning rate 10-3, batch size 16 và 300 epoch; đánh giá bằng Precision, Recall và $mAP@0.5 : 0.95$.

(2) Phát hiện đối tượng nhỏ: Tập trung vào dây dẫn, chuỗi cách điện và tạ chống rung, đặc biệt các trường hợp có kích thước dưới 5% diện tích ảnh, bị che khuất hoặc góc quan sát phức tạp.

(3) Độ bền vững môi trường: Đánh giá mô hình dưới các điều kiện khác nhau về ánh sáng, thời tiết, nền cảnh, góc nhìn và chất lượng ảnh thông qua mức suy giảm mAP, Recall và F1-score.

(4) Hiệu năng thời gian thực: Triển khai trên NVIDIA Jetson Orin Nano, thực hiện Intensive Test, Parallel Processing Test và Stress Test; đánh giá FPS, thời gian suy luận, tài nguyên, nhiệt độ và công suất.

(5) So sánh mô hình: Đối chiếu HVE-YOLO với HVE-Mask R-CNN và các phiên bản YOLO theo các tiêu chí độ chính xác, phát hiện vật thể nhỏ, tốc độ, tài nguyên và khả năng triển khai thời gian thực.

2.3.3. Kết quả thử nghiệm

Kết quả thực nghiệm cho thấy HVE-YOLO đạt hiệu năng cao hơn HVE-Mask R-CNN trên phần lớn các tiêu chí đánh giá. Cụ thể, HVE-YOLO đạt $mAP@0.5 = 0,972$, F1-score cực đại 0,950 và Recall xem bảng 2.1

Bảng 2.1. So sánh hiệu năng giữa HVE-YOLO và HVE-Mask R-CNN

Chỉ số	HVE-YOLO	HVE-Mask R-CNN
Kiến trúc	One-stage	Two-stage
$mAP@0.5$	0,972	0,855
F1-score cực đại	0,950	0,851
Recall tổng thể	0,990	0,906
FPS – Intensive Test	162,9	75,8
FPS – Stress Test	165,5	82,3
FPS – Parallel Processing	95,6	98,6

2.4. Kết luận chương 2

Chương 2 phát triển mô hình HVE-YOLO để phát hiện sáu loại thiết bị trên lưới truyền tải 110 kV từ ảnh UAV. Mô hình được cải tiến bằng đặc trưng đa tỷ lệ, lớp P2 và cơ chế attention, giúp nâng cao khả năng phát hiện đối tượng nhỏ trong nền ảnh phức tạp.

Thực nghiệm cho thấy HVE-YOLO đạt $mAP@0.5 = 0,972$ và tốc độ tối đa 165,5 FPS, đáp ứng yêu cầu phát hiện thời gian thực và duy trì độ ổn định trước các thay đổi về góc quan sát, ánh sáng, khoảng cách và nền ảnh. Kết quả định vị chuỗi cách điện trên 96% tạo cơ sở cho việc phân tích trạng thái thiết bị. Đây là cơ sở để Chương 3 tiếp tục nghiên cứu phát hiện và nhận diện bất thường trên chuỗi cách điện từ ảnh UAV.

CHƯƠNG 3. PHÁT HIỆN BẤT THƯỜNG TRÊN CHUỖI CÁCH ĐIỆN TỪ ẢNH UAV

Chương 2 đề xuất HVE-YOLO cho phát hiện thiết bị trên lưới điện cao thế từ ảnh UAV, đạt $mAP@0.5 = 97\%$ và 165,5 FPS, đáp ứng yêu cầu giám sát thời gian thực. Tuy nhiên, phát hiện thiết bị mới là bước đầu; cần tiếp tục phân tích trạng thái và phát hiện bất thường, đặc biệt trên chuỗi cách điện do hư hỏng thường nhỏ, đa dạng và thiếu dữ liệu.

Vì vậy, Chương 3 tập trung phát triển mô hình FGS (FPN-GELAN-Diffusion Segmentation), kết hợp đặc trưng đa tỷ lệ và cơ chế khuếch tán nhằm nâng cao khả năng phát hiện, phân đoạn bất thường trên chuỗi cách điện trong điều kiện dữ liệu hạn chế và môi trường phức tạp.

3.1. Các nghiên cứu liên quan về phát hiện bất thường trên chuỗi cách điện

3.1.1. Các dạng bất thường trên chuỗi cách điện

3.1.2. Nghiên cứu phát hiện bất thường trên chuỗi cách điện từ ảnh thu nhận bằng UAV

Trong ảnh UAV, bất thường trên chuỗi cách điện thường có kích thước nhỏ, hình thái đa dạng và dễ bị che khuất bởi nền phức tạp, gây khó khăn cho việc trích xuất đặc trưng và phát hiện. Các mô hình YOLO được sử dụng phổ biến nhờ tốc độ cao và khả năng phát hiện đối tượng nhỏ, với các phiên bản khai thác FPN, PAN và gần đây kết hợp DETR hoặc RT-DETR để tăng cường biểu diễn đặc trưng. Tuy nhiên, DETR đòi hỏi dữ liệu huấn luyện lớn, trong khi YOLOv9 cải thiện khả năng biểu diễn thông qua kiến trúc GELAN, góp phần nâng cao hiệu quả phát hiện bất thường trên ảnh UAV.

Backbone YOLOv9

Ưu điểm và hạn chế của YOLOv9 trong bài toán kiểm tra chuỗi cách điện từ ảnh UAV

Ưu điểm: Tốc độ xử lý cao, nâng cao khả năng phát hiện các đối tượng nhỏ, có khả năng học và thích nghi cao với thay đổi về góc quan sát, tỷ lệ đối tượng và điều kiện ánh sáng.

Hạn chế: Yêu cầu lượng dữ liệu lớn, hạn chế với các lỗi mới

3.2. Mô hình đề xuất FGS (FPN-GELAN-Diffusion Segmentation)

3.2.1. Mô hình FGS (FPN-GELAN-Diffusion Segmentation)

Trong ảnh UAV, việc phát hiện bất thường trên chuỗi cách điện gặp nhiều khó khăn do kích thước lỗi nhỏ, hình thái đa dạng, nền ảnh phức tạp và điều kiện thu nhận biến động. Các yếu tố như rung UAV, góc và khoảng cách quan sát,

thời tiết, ánh sáng và sự khác biệt quang học giữa các loại cách điện làm suy giảm tính ổn định của đặc trưng ảnh. Đồng thời, quá trình downsampling có thể làm mất các chi tiết bất thường nhỏ như nứt, bong tróc hoặc nhiễm bẩn cục bộ.

Để giải quyết các hạn chế này, luận án đề xuất FGS (FPN–GELAN–Diffusion Segmentation), kết hợp hợp nhất đặc trưng đa tỷ lệ, GELAN và mô hình Diffusion nhằm bảo toàn thông tin chi tiết, tăng khả năng phát hiện bất thường nhỏ và thực hiện phân đoạn chính xác trong điều kiện ảnh UAV phức tạp.

3.2.2. Kiến trúc tổng quan

Mô hình FGS bao gồm ba thành phần chính là mạng đặc trưng đa tỷ lệ FPN, backbone GELAN và mô hình khuếch tán (Diffusion Model). Đồng thời, hệ thống còn tích hợp nhánh phân đoạn mức điểm ảnh nhằm nâng cao độ chính xác trong việc xác định hình dạng và phạm vi của vùng hư hỏng.



Hình 3.1: Kiến trúc tổng thể mô hình FGS

Hình 3.1 minh họa cấu trúc tổng thể của hệ thống FGS. Kiến trúc được xây dựng theo hướng tích hợp thống nhất giữa các giai đoạn trích xuất đặc trưng, đề xuất vùng, tăng cường biểu diễn và phân đoạn bất thường.

3.2.3. Các cải tiến trong mô hình FGS Backbone GELAN

Backbone GELAN được tích hợp trong FGS nhằm giải quyết trực tiếp vấn đề trên thông qua cơ chế tổng hợp đặc trưng hiệu quả giữa các tầng mạng. Khác với các backbone CNN truyền thống chỉ truyền thông tin theo chiều sâu tuần tự, GELAN sử dụng cơ chế kết nối đa nhánh và tái sử dụng đặc trưng nhằm bảo toàn tối đa thông tin gradient trong quá trình huấn luyện.

Cải tiến này mang lại ba lợi ích quan trọng.

Thứ nhất: tăng cường khả năng phát hiện các bất thường nhỏ có hình dạng phức tạp.

Thứ hai: hỗ trợ kết hợp nhiều nhánh xử lý như FGS.

Thứ ba, backbone GELAN tối ưu hóa hiệu quả tính

Ngoài ra, cơ chế chú ý SE (Squeeze-and-Excitation) được tích hợp trong backbone cho phép mô hình học trọng số thích nghi theo từng kênh đặc trưng. Nhờ đó, mạng có xu hướng tập trung mạnh hơn vào các vùng chứa thông tin bất thường thay vì các thành phần nền không liên quan.

Mô hình khuếch tán (Diffusion Model)

Một trong những điểm khác biệt quan trọng của FGS so với các mô hình phát hiện bất thường truyền thống là việc tích hợp mô hình khuếch tán (Diffusion Model) vào giai đoạn tăng cường đặc trưng.

Về bản chất, quá trình khuếch tán giúp mạng học cách phân biệt giữa nhiễu ngẫu nhiên và cấu trúc đặc trưng thực sự của cách điện. Điều này đặc biệt quan trọng trong môi trường ảnh được thu nhận bằng UAV.

Vùng quan tâm và các nhánh dự đoán

Trong mô hình đề xuất, lớp RoI Align đóng vai trò trung gian quan trọng trong việc kết hợp thông tin từ nhiều nguồn đặc trưng khác nhau. Lớp này tiếp nhận ba đầu vào chính gồm: đặc trưng từ RPN, đặc trưng từ backbone GELAN và đặc trưng tăng cường từ mô hình khuếch tán. Việc kết hợp đa nguồn đặc trưng giúp mô hình xây dựng được biểu diễn RoI có độ phân biệt cao hơn.

Giai đoạn hậu xử lý (Post-processing)

Sau khi các nhánh RCNN Head, Mask Head và MaskIoU Head hoàn thành quá trình suy luận, hệ thống thu được tập các dự đoán bao gồm vị trí đối tượng, nhãn lớp và mặt nạ phân đoạn tương ứng. Tuy nhiên, các kết quả đầu ra ban đầu vẫn có thể tồn tại các dự đoán trùng lặp, các đối tượng có độ tin cậy thấp hoặc các mặt nạ phân đoạn chưa đạt chất lượng mong muốn. Vì vậy, mô hình FGS tích hợp một giai đoạn hậu xử lý nhằm nâng cao độ chính xác và độ ổn định của kết quả trước khi đưa ra quyết định cuối cùng.

3.2.4. Quy trình huấn luyện mô hình FGS

Quy trình huấn luyện mô hình FGS được xây dựng theo hướng end-to-end với các giai đoạn (i) chuẩn bị dữ liệu; (ii) cấu hình và quá trình huấn luyện và (iii) đánh giá mô hình.

Chuẩn bị dữ liệu

Cấu hình và quá trình huấn luyện

Quá trình huấn luyện mô hình được thực hiện với thuật toán huấn luyện mô hình FGS được thực hiện theo một quy trình gồm năm bước chính. Đầu tiên, các thành phần của mô hình gồm GELAN-SE-DWConv, Diffusion/DiT, FPN, RPN, RoIAlign, RCNN Head, Mask Head và MaskIoU Head được khởi tạo và huấn luyện từ đầu.

Đánh giá mô hình

3.3. Thử nghiệm

3.3.1. Thiết lập thử nghiệm

Toàn bộ các thử nghiệm được triển khai trên nền tảng Python 3.10 sử dụng

bộ xử lý Intel Core i9-13900K, bộ nhớ RAM 64 GB và card đồ họa NVIDIA RTX 4090 với 24 GB bộ nhớ đồ họa với ảnh đầu 640×640 trong 50 epoch với bộ tối ưu AdamX và tốc độ học khởi tạo bằng 5×10^{-4} .

Trong năm epoch đầu tiên, cơ chế linear warm-up được áp dụng; tiếp đến tốc độ được giảm dần theo cosine decay. Để hạn chế sự phụ thuộc quá mức vào dữ liệu tăng cường, toàn bộ kỹ thuật augmentation được tắt trong 10 epoch cuối.

3.3.2. Chuẩn bị dữ liệu

Nghiên cứu sử dụng hai bộ dữ liệu VNElectric và Insulator Defect Detection để đánh giá hiệu quả và khả năng tổng quát hóa của phương pháp đề xuất. VNElectric gồm 1.596 ảnh UAV thu nhận trên lưới điện 110 kV tại miền Bắc Việt Nam, với đa dạng góc quan sát, điều kiện môi trường và nền ảnh phức tạp; dữ liệu được chia theo tỷ lệ 3:1:1 cho huấn luyện, xác thực và kiểm tra, kết hợp các kỹ thuật tăng cường dữ liệu nhằm nâng cao khả năng tổng quát hóa của mô hình.

3.3.3. Kết quả thử nghiệm

Kết quả thực nghiệm trên bộ dữ liệu VNElectric cho thấy mô hình FGS đạt Precision 92,3%, Recall 91,4% và mAP@50 93,7%, lần lượt cải thiện 1,7, 1,1 và 0,9 điểm phần trăm so với YOLOv9-E. Kết quả này cho thấy sự kết hợp giữa backbone GELAN, cơ chế hợp nhất đặc trưng đa tỉ lệ FPN và Diffusion Model góp phần nâng cao khả năng phát hiện các bất thường có kích thước nhỏ trong nền ảnh UAV phức tạp. Phân tích ablation khẳng định vai trò hỗ trợ của các thành phần, trong đó FPN có ảnh hưởng lớn nhất khi loại bỏ làm mAP@50 giảm 5,6 điểm phần trăm, tiếp theo là GELAN và Diffusion Model với mức suy giảm lần lượt 3,9 và 1,7 điểm phần trăm. Kết quả trực quan cũng cho thấy FGS duy trì khả năng phát hiện tương đối ổn định trong các điều kiện nền phức tạp, đối tượng nhỏ, che khuất và ánh sáng không đồng đều, qua đó chứng minh tính hiệu quả của phương pháp đối với dữ liệu ảnh UAV thực tế.

3.4. Kết luận chương 3

Chương 3 đề xuất mô hình FGS (FPN-GELAN-Diffusion Segmentation) để phát hiện và phân đoạn bất thường trên chuỗi cách điện từ ảnh UAV. Việc kết hợp GELAN, FPN và Diffusion Transformer giúp tăng cường trích xuất và hợp nhất đặc trưng đa tỷ lệ, nâng cao khả năng nhận diện các bất thường nhỏ, tương phản thấp trong nền ảnh phức tạp.

Trên bộ dữ liệu VNElectric, FGS đạt Precision 92,3% và mAP@50 = 93,7%, vượt YOLOv9-E và được chứng minh qua các thí nghiệm ablation. Tuy nhiên,

mô hình còn chi phí tính toán và số lượng tham số cao, cần tiếp tục tối ưu để triển khai trên nền tảng Edge AI và mở rộng khả năng tổng quát hóa.

CHƯƠNG 4. PHƯƠNG PHÁP HỌC GIA TĂNG VỚI DỮ LIỆU HẠN CHẾ CHO CHẨN ĐOÁN BẤT THƯỜNG CHUỖI CÁCH ĐIỆN

Chương 4 tập trung giải quyết vấn đề thiếu dữ liệu bất thường trong giám sát lưới điện bằng UAV, đặc biệt đối với các dạng lỗi hiếm hoặc mới xuất hiện. Luận án đề xuất AAPID theo khuôn khổ Few-Shot Class-Incremental Learning (FSCIL), kết hợp prompt bất biến, prompt thích nghi, Information Bottleneck Loss và Anchor Loss để học các lớp lỗi mới từ số lượng mẫu hạn chế mà vẫn duy trì tri thức đã học.

Thực nghiệm trên CIFAR100 và VPI-Fault-2500 cho thấy AAPID đạt hiệu quả tốt, đồng thời có tiềm năng nhận diện và thích nghi với các dạng bất thường mới trên chuỗi cách điện.

4.1. Các nghiên cứu liên quan về sinh dữ liệu trong huấn luyện các mô hình học sâu

4.1.1. Các phương pháp sinh dữ liệu truyền thống

Hiệu quả của học sâu phụ thuộc lớn vào chất lượng và số lượng dữ liệu, trong khi dữ liệu lỗi thiết bị điện thường khan hiếm và mất cân bằng. Các phương pháp tăng cường truyền thống còn hạn chế vì chủ yếu biến đổi dữ liệu có sẵn. Do đó, GAN, VAE và Diffusion Model được sử dụng để tạo dữ liệu đa dạng, chân thực, bổ sung các mẫu lỗi hiếm và cải thiện khả năng tổng quát hóa của mô hình khi dữ liệu thực tế hạn chế.

4.1.2. Các phương pháp sinh dữ liệu và học với dữ liệu hạn chế trong phát hiện bất thường

Trong điều kiện dữ liệu bất thường khan hiếm, Prompt Learning kết hợp Few-Shot và Incremental Learning giúp tận dụng tri thức tiền huấn luyện, học lớp mới với ít mẫu và hạn chế quên lãng. Các phương pháp như L2P, DualPrompt, CodaPrompt và ASP đã khai thác prompt để thích nghi với lớp mới.

Tuy nhiên, các phương pháp hiện tại chưa giải quyết tốt đặc thù của bất thường chuỗi cách điện như đối tượng nhỏ, tương phản thấp, nền phức tạp và lỗi mới xuất hiện. Vì vậy, luận án đề xuất AAPID, hướng đến duy trì tri thức cũ, học hiệu quả lớp lỗi mới với ít mẫu và tăng khả năng phân biệt đặc trưng trong bài toán FSCIL.

4.2. AAPID (Attention-aware Adaptive Prompting for Insulator Diagnosis) – Mô hình Prompt thích nghi dựa trên cơ chế Attention cho chẩn đoán lỗi chuỗi cách điện

Các mô hình học sâu truyền thống có hiệu quả cao trong phát hiện và nhận diện bất thường chuỗi cách điện nhưng vẫn phụ thuộc vào dữ liệu lỗi khan hiếm, mất cân bằng và các lớp lỗi mới. Để khắc phục, luận án áp dụng Few-Shot Class-Incremental Learning (FSCIL) và đề xuất AAPID, nhằm học các lớp lỗi mới với số lượng mẫu hạn chế, duy trì tri thức đã học và hạn chế quên lãng thảm họa, qua đó nâng cao khả năng thích nghi của hệ thống.

4.2.1. Tổng quan kiến trúc

AAPID được đề xuất trong khuôn khổ FSCIL, dựa trên Vision Transformer (ViT), nhằm học các lớp bất thường mới với ít mẫu đồng thời duy trì tri thức cũ và hạn chế quên lãng thảm họa.

Mô hình đóng băng bộ mã hóa ViT tiền huấn luyện, chỉ cập nhật các tham số prompt và bộ sinh prompt thích nghi, giúp giảm quá khớp và chi phí huấn luyện. Kiến trúc gồm bốn thành phần chính: ViT cố định, Task-Invariant Prompt (TIP), Prompt Encoder và hàm mất mát kết hợp Information Bottleneck Loss với Anchor Loss, qua đó tăng khả năng thích nghi và khả năng phân biệt giữa các lớp lỗi mới.

4.2.2. Cơ chế Prompt thích nghi dựa trên Attention Bộ mã hóa Vision Transformer cố định

AAPID sử dụng Vision Transformer (ViT) tiền huấn luyện làm backbone và đóng băng toàn bộ tham số trong quá trình FSCIL. Ảnh được chia thành các patch, tạo patch embedding, bổ sung thông tin vị trí và đưa qua các khối Transformer để trích xuất đặc trưng.

Việc đóng băng backbone giúp giảm quá khớp, bảo toàn tri thức đã học, giảm số tham số cần tối ưu và tăng tính ổn định khi huấn luyện few-shot. Tuy nhiên, backbone không thể trực tiếp thích nghi với lớp mới, nên cơ chế prompting đóng vai trò chính trong việc thích nghi, tạo cơ sở cho hệ thống prompt thích nghi của AAPID.

Prompt bất biến theo nhiệm vụ

Task-Invariant Prompt (TIP) trong AAPID là tập prompt có thể học được, nhằm bổ sung tri thức đặc trưng chung cho mô hình mà không cần thay đổi backbone. TIP học các đặc điểm phổ quát của chuỗi cách điện như hình dạng, màu sắc, cấu trúc và đặc điểm bề mặt.

TIP gồm token được khởi tạo cùng giá trị, sau đó ghép với các patch token của ảnh để cùng tham gia self-attention trong Transformer. Nhờ đó, TIP cung cấp tri thức nền bất biến theo nhiệm vụ, giúp mô hình thích nghi tốt hơn với các lớp bất thường mới trong bài toán FSCIL.

Bộ sinh Prompt thích nghi

Task-Specific Prompt (TSP) được AAPID bổ sung nhằm khắc phục hạn chế của TIP trong việc phân biệt các loại lỗi khác nhau. Thay vì sử dụng prompt cố định theo từng lớp, AAPID sử dụng Prompt Encoder để sinh prompt thích nghi trực tiếp từ đặc trưng của ảnh đầu vào.

TSP giúp mã hóa các đặc trưng đặc thù của từng dạng bất thường như nứt bề mặt, sứt vảnh, vỡ thân và bám bẩn, qua đó tăng khả năng phân biệt các loại lỗi. Nhờ cơ chế sinh prompt theo từng ảnh, AAPID có thể học và nhận diện các lớp bất thường mới hiệu quả ngay cả khi dữ liệu huấn luyện hạn chế.

4.2.3. Học biểu diễn Prompt bằng Information Bottleneck Ràng buộc Information Bottleneck

Để nâng cao khả năng tổng quát hóa của Prompt Encoder, AAPID áp dụng nguyên lý *Information Bottleneck* (IB), nhằm giữ lại thông tin liên quan đến nhãn lớp và loại bỏ các thông tin dư thừa từ dữ liệu đầu vào. Hàm mục tiêu IB được biểu diễn bởi min $I(P; X) - \gamma I(P; y)$

trong đó $I(P; X)$ và $I(P; y)$ lần lượt biểu thị lượng thông tin giữa prompt với dữ liệu đầu vào và nhãn lớp, còn γ là hệ số cân bằng.

Do việc tính trực tiếp thông tin tương hỗ khó thực hiện, AAPID sử dụng một xấp xỉ dựa trên cận dưới biến phân: $L_{IB} = E[-\log P(y|x)] + \beta KLD_{KL}(P(p|x)|r(p))$ trong đó thành phần thứ nhất là hàm mất mát cross-entropy, đảm bảo prompt duy trì thông tin cần thiết cho quá trình phân loại; thành phần thứ hai là phân kỳ Kullback–Leibler (KL), đóng vai trò điều chuẩn bằng cách hạn chế lượng thông tin riêng biệt mà prompt mã hóa từ từng mẫu. Với $r(p)$ thường được giả định là một phân phối Gaussian, cơ chế này giúp prompt tập trung vào các đặc trưng có tính phân biệt và giảm sự phụ thuộc vào các đặc trưng nhiễu hoặc đặc thù của từng mẫu. Nhờ đó, Prompt Encoder có thể sinh ra các prompt mang tính tổng quát cao, phù hợp với bài toán few-shot learning và khả năng thích nghi với các lớp mới.

4.2.4. Anchor Loss và tối ưu không gian đặc trưng

Mặc dù Information Bottleneck giúp prompt tập trung vào các thông tin giàu ngữ nghĩa, không gian đặc trưng vẫn có thể xuất hiện hiện tượng chồng lấn giữa các lớp. Để tăng tính phân biệt của biểu diễn, AAPID sử dụng thêm *Anchor Loss*. Với mỗi lớp k , một mẫu đại diện được lựa chọn làm anchor và biểu diễn đặc trưng tương ứng được xác định bởi $c^*k = f * \theta, p(xk)$.

Anchor Loss được xây dựng dựa trên độ tương đồng cosine giữa đặc trưng của mẫu xk và anchor:

$$L_{anchor} = 1 - c^*kT f * \theta, p(xk) \\ |c^*k|^2 |f * \theta, p(xk)|^2$$

Cơ chế này khuyến khích các mẫu cùng lớp có biểu diễn gần anchor tương ứng, từ đó hình thành các cụm đặc trưng compact và tăng khả năng phân tách giữa các lớp. Điều này đặc biệt hữu ích trong bối cảnh few-shot, khi số lượng mẫu huấn luyện cho mỗi lớp rất hạn chế.

4.2.5. Hàm mục tiêu và phân tích mô hình Hàm mục tiêu tổng thể

AAPID tối ưu đồng thời khả năng biểu diễn ngữ nghĩa của prompt và tính phân biệt của không gian đặc trưng thông qua hàm mục tiêu:

$$L = L * IB + \lambda L * anchor, (4.2)$$

trong đó λ là hệ số cân bằng giữa Information Bottleneck Loss và Anchor Loss. Sự kết hợp này tạo thành cơ chế tối ưu kép: $L * IB$ giúp loại bỏ thông tin dư thừa và duy trì các đặc trưng liên quan đến nhãn lớp, trong khi $L * anchor$ tăng tính compact và khả năng phân tách của các biểu diễn đặc trưng. Nhờ đó, mô hình duy trì hiệu quả phân loại và khả năng thích nghi tốt hơn khi số lượng mẫu của các lớp mới rất hạn chế.

Đánh giá kiến trúc

AAPID kết hợp ba hướng tiếp cận gồm Vision Transformer tiền huấn luyện, Prompt Learning và Few-Shot Class-Incremental Learning. Điểm nổi bật của kiến trúc là Prompt Encoder dựa trên Information Bottleneck, cho phép sinh prompt thích nghi và chuyển giao tri thức từ các lớp đã học sang lớp mới với số lượng mẫu hạn chế. Bên cạnh đó, Anchor Loss tăng tính phân biệt của không gian đặc trưng, góp phần hạn chế hiện tượng quên thảm họa trong quá trình học gia tăng. Nhờ sự kết hợp này, AAPID phù hợp với bài toán nhận dạng lỗi thiết bị điện, trong đó các dạng lỗi mới thường xuất hiện liên tục nhưng có rất ít dữ liệu huấn luyện.

4.3. Thử nghiệm

Các thử nghiệm được thực hiện nhằm đánh giá toàn diện hiệu quả của AAPID trong bài toán Few-Shot Class-Incremental Learning (FSCIL) cho nhận diện lỗi chuỗi cách điện từ ảnh UAV. Ba khía cạnh chính được xem xét gồm: khả năng học các lớp mới với số lượng mẫu hạn chế, khả năng duy trì tri thức của các lớp đã học và khả năng hình thành không gian đặc trưng có tính phân biệt cao. Thử nghiệm được tiến hành trên cả bộ dữ liệu chuẩn CIFAR100 và bộ dữ liệu chuyên ngành VPI-Fault-2500 nhằm đánh giá đồng thời tính hiệu quả về phương pháp luận và khả năng ứng dụng thực tế.

4.3.1. Môi trường thử nghiệm

AAPID được triển khai trên PyTorch 2.0 và CUDA, sử dụng ViT-B/16 tiền huấn luyện trên ImageNet-1K làm backbone. Ảnh được chuẩn hóa theo

ImageNet và đưa về kích thước 224×224 pixel, kết hợp các phép tăng cường dữ liệu cơ bản.

Giai đoạn học cơ sở được huấn luyện 20 epoch với SGD, learning rate 0,01 và batch size 48. Độ dài prompt lần lượt là 3 token trên CIFAR100 và 5 token trên VPI-Fault-2500. Hệ số được lựa chọn trên tập xác thực. Ở các phiên học tăng dần, mô hình chỉ sử dụng 5 mẫu cho mỗi lớp mới, phù hợp với thiết lập few-shot.

4.3.2. Dữ liệu thử nghiệm

Thực nghiệm được đánh giá trên **hai bộ dữ liệu** gồm CIFAR100 và VPI-Fault-2500. **CIFAR100** được sử dụng làm benchmark FSCIL chuẩn với 60 lớp cơ sở và 8 phiên tăng dần theo thiết lập 5-way 5-shot, ảnh được resize về 224×224 cho ViT-B/16. **VPI-Fault-2500** gồm 2.500 ảnh chuỗi cách điện thuộc 5 lớp, trong đó 3 lớp *Normal*, *SC* và *RC* là lớp cơ sở, còn *BB* và *SSC* được đưa vào phiên tăng dần theo thiết lập **2-way 5-shot**. AAPID chỉ sử dụng **5 ảnh cho mỗi lớp lỗi mới** và không sử dụng lại dữ liệu lớp cơ sở, qua đó mô phỏng điều kiện nhận diện các dạng lỗi mới với dữ liệu hạn chế.

4.3.3. Kết quả thử nghiệm

Thực nghiệm được thiết kế để trả lời ba câu hỏi nghiên cứu:

- RQ1: AAPID có cải thiện hiệu năng so với các phương pháp CIL và FSCIL hiện có hay không?
- RQ2: Mức độ đóng góp của từng thành phần trong AAPID đối với hiệu năng tổng thể là như thế nào?
- RQ3: AAPID có hình thành được không gian đặc trưng có tính phân biệt cao giữa các lớp cơ sở và lớp mới hay không?

Chỉ số đánh giá

Hiệu quả của mô hình được đánh giá bằng ba chỉ số chính:

- Average Accuracy (A_{avg}): độ chính xác trung bình trên toàn bộ các lớp và các phiên học, phản ánh hiệu năng tổng thể.
- Performance Dropping (PD): mức suy giảm độ chính xác của các lớp cơ sở sau quá trình học tăng dần: $PD = A_0 - A_T$ trong đó A_0 và A_T lần lượt là độ chính xác trên các lớp cơ sở ở phiên đầu và phiên cuối.
- Harmonic Accuracy (HAcc): đánh giá sự cân bằng giữa khả năng duy trì lớp cơ sở và học lớp mới: $H_{Acc} = \frac{2A_0A_n}{A_0+A_n}$ trong đó A_0, A_n là độ chính xác trên các lớp cơ sở và là độ chính xác trung bình trên các lớp mới.

Kết quả đánh giá tổng thể

Kết quả trên Bảng 4.1 và Bảng 4.2 cho thấy AAPID đạt hiệu năng cao nhất trên cả hai bộ dữ liệu.

Bảng 4.1: So sánh hiệu năng trên tập dữ liệu CIFAR100 (Thiết lập: 60+8×5-way 5-shot)

Phương pháp	<i>Aavg</i> (%)↑	PD(%)↓	HAcc(%)↑
iCaRL	78.4	27.2	57.5
Foster	73.9	35.9	11.0
CEC	81.0	19.0	64.1
FACT	78.6	21.7	55.5
TEEN	87.3	8.8	81.2
L2P	71.1	37.0	0.0
DualP	70.6	36.9	0.1
CodaP	72.0	37.4	0.0
DualP+	81.1	8.5	75.3
AAPID (đề xuất)	89.5	4.8	86.1

Bảng 4.2: So sánh hiệu năng trên tập dữ liệu VPI-Fault-2500 (Thiết lập: 3+2-way 5-shot)

Phương pháp	<i>Aavg</i> (%)↑	PD(%)↓	HAcc(%)↑
iCaRL	85.1	15.4	78.2
Foster	82.6	20.1	65.7
TEEN	91.5	6.2	88.3
DualP+	88.9	8.0	85.1
AAPID (đề xuất)	94.2	3.1	92.5

Nghiên cứu loại bỏ – Ablation Study (RQ2)

Để đánh giá đóng góp của từng thành phần, các mô-đun trong AAPID được lần lượt loại bỏ. Kết quả trong Bảng 4.3 cho thấy mọi biến thể loại bỏ đều làm giảm *Aavg* trên cả CIFAR100 và VPI-Fault-2500, qua đó khẳng định tính hỗ trợ giữa các thành phần của mô hình

Phương pháp	CIFAR100	VPI-Fault-2500
AAPID (full)	89.5	94.2
Without Task-Invariant Prompts (TIP)	88.9	93.6
Without Task-Specific Prompts (TSP) (TSP)	87.8	92.1
Without Anchor Loss	88.7	93.4
Without IB Loss	88.2	92.9

Không gian đặc trưng (RQ3)

Để đánh giá khả năng hình thành không gian đặc trưng có tính phân biệt, các đặc trưng trên Query Set của VPI-Fault-2500 được trực quan hóa bằng t-SNE như minh họa trong Hình 4.2



Hình 4.2: Trực quan hóa t-SNE kết quả nghiên cứu loại bỏ trên bộ dữ liệu VPI-Fault-2500.

4.4. Kết luận chương 4

Chương 4 đề xuất AAPID nhằm nhận diện lỗi chuỗi cách điện trong điều kiện dữ liệu lớp mới khan hiếm theo khuôn khổ Few-Shot Class-Incremental Learning (FSCIL). Phương pháp sử dụng Vision Transformer tiền huấn luyện kết hợp prompting thích nghi, gồm Task-Invariant Prompts, Task-result Prompts, Information Bottleneck Loss và Anchor Loss.

KẾT LUẬN

Những kết quả chính của luận án

Luận án đã nghiên cứu và phát triển các kỹ thuật học sâu phục vụ giám sát lưới truyền tải điện từ ảnh UAV, tập trung vào ba bài toán chính: phát hiện thiết bị, nhận diện và phân đoạn bất thường trên chuỗi cách điện, và học tăng dần ít mẫu đối với các dạng lỗi mới.

Thứ nhất, mô hình HVE-YOLO được đề xuất cho phát hiện sáu nhóm thiết bị trên lưới điện đạt $mAP@0.5:0.95 = 88,5\%$, đồng thời đạt 45,2 FPS, 11,4 GFLOPs và 9,8 W trên NVIDIA Jetson Orin Nano, cho thấy khả năng triển khai trên thiết bị biên.

Thứ hai, mô hình FGS trong phát hiện bất thường trên chuỗi cách điện, đạt Precision = 92,3%, Recall = 91,4% và $mAP@0.5 = 93,7\%$ trên bộ dữ liệu VNelectric. Các thí nghiệm ablation khẳng định hiệu quả của GELAN, RPN và Diffusion được xây dựng cho phát hiện, phân loại và phân đoạn bất thường trên chuỗi cách điện..

Thứ ba, phương pháp AAPID được đề xuất cho bài toán Few-Shot Class-Incremental Learning (FSCIL), cho phép học các lớp lỗi mới với số lượng mẫu rất hạn chế đồng thời duy trì tri thức đã học. Mô hình đạt Aavg = 89,5%, PD = 4,8%, HAcc = 86,1% trên CIFAR100 và 94,2%, 3,1%, 92,5% trên VPI-Fault-2500 với chỉ 5 mẫu/lớp mới.

Về thực tiễn, luận án xây dựng các bộ dữ liệu VNelectric (69.955 đối tượng) và VPI-Fault-2500 (2.500 ảnh), tạo cơ sở cho quy trình giám sát nhiều mức từ phát hiện thiết bị → nhận diện, phân đoạn bất thường → thích nghi với lỗi mới, góp phần thúc đẩy giám sát tự động và bảo trì thông minh lưới điện.

Hạn chế của luận án

Tuy nhiên, luận án còn một số hạn chế: (i) hiệu quả mô hình phụ thuộc vào chất lượng ảnh và điều kiện môi trường; (ii) một số mô hình, đặc biệt FGS, có chi phí tính toán cao; (iii) số lượng dạng lỗi mới được đánh giá trong thiết lập FSCIL còn hạn chế; và (iv) AAPID vẫn phụ thuộc vào khả năng biểu diễn của mô hình ViT tiền huấn luyện và chưa được tối ưu đầy đủ cho phần cứng tài nguyên thấp.

Hướng nghiên cứu tiếp theo

Các hướng nghiên cứu tiếp theo của luận án tập trung vào năm định hướng:

- **Thứ nhất, tích hợp dữ liệu đa phương thức:** kết hợp ảnh quang học với ảnh nhiệt, LiDAR và dữ liệu cảm biến nhằm nâng cao khả năng phát hiện các bất thường khó quan sát từ ảnh thông thường.

- **Thứ hai, tối ưu mô hình cho triển khai thời gian thực:** nghiên cứu pruning, lượng tử hóa, chưng cất tri thức, low-rank adaptation và các kiến trúc nhẹ nhằm giảm chi phí tính toán, đặc biệt đối với các mô hình có quy mô lớn như FGS.

- **Thứ ba, dự báo sớm sự cố:** kết hợp dữ liệu ảnh theo thời gian, ảnh nhiệt và dữ liệu cảm biến với các mô hình LSTM, Transformer hoặc học sâu theo chuỗi thời gian để theo dõi xu hướng suy giảm và dự báo nguy cơ sự cố.

- **Thứ tư, nâng cao khả năng tổng quát hóa:** mở rộng dữ liệu và thiết bị, đồng thời nghiên cứu transfer learning, domain adaptation và continual learning để mô hình thích nghi với các khu vực, điều kiện và hệ thống điện khác nhau.

- **Thứ năm, xây dựng hệ thống hỗ trợ ra quyết định thông minh:** tích hợp các mô hình phát hiện thiết bị, nhận diện bất thường và học tăng dần với dữ liệu UAV, ảnh nhiệt, LiDAR, SCADA và tri thức chuyên gia nhằm đánh giá rủi ro và hỗ trợ kiểm tra, bảo trì.

Định hướng dài hạn: xây dựng hệ thống giám sát lưới điện thông minh dựa trên UAV và trí tuệ nhân tạo, kết hợp dữ liệu đa phương thức, học sâu và học liên tục nhằm nâng cao mức độ tự động hóa, khả năng thích nghi và độ tin cậy trong kiểm tra, vận hành và bảo trì lưới điện

DANH MỤC CÔNG TRÌNH CÔNG BỐ LIÊN QUAN ĐẾN LUẬN ÁN

- CT1.** **Trịnh Hiền Anh**, Nguyễn Thị Thanh Tân, Bùi Xuân Hùng, Đoàn Ngọc Duy, “*Phát hiện thiết bị trên lưới truyền tải điện cao thế 110kV sử dụng kỹ thuật học sâu*”. Hội thảo Quốc gia lần thứ XXV: Một số vấn đề chọn lọc của công nghệ thông tin và truyền thông, Hà Nội, 8–9/12/2022, pp.174–179.
- CT2.** **Anh Trinh Hien**, A. D. Tran, D. C. Viet, Q. D. T. Thuy and Q. N. Huu, “Insulator defect detection based on feature pyramid network and diffusion model”. *Signal, Image and Video Processing*, 2025, Vol. 19, Issue 5, p.372 (SCIE-Q3).
- CT3.** **Anh Trinh Hien**, Tao Ngo Quoc, Thanh-Tan Nguyen Thi, “AttentionAware Prompting for Learning New Faults on Insulators with Limited Data”. *Journal of Information Hiding and Multimedia Signal Processing*, Vol. 16, No. 4, December 2025, pp.1199–1208 (SCOPUS-Q4).
- CT4.** **Trịnh Hiền Anh**, Nguyễn Thị Thanh Tân, “Một phương pháp hiệu quả để phát hiện đường dây, cột điện và các thiết bị trên lưới truyền tải điện cao thế 110kV bằng kỹ thuật học sâu”. *Tạp chí Khoa học và Công nghệ Năng lượng – Trường Đại học Điện lực*, Số 32, 2023, ISSN: 1859-4557, pp.1–11.
- CT5.** **Trịnh Hiền Anh**, Nguyễn Thị Thanh Tân, “Một kỹ thuật tăng cường chất lượng ảnh UAV thu nhận trên hệ thống truyền tải điện cao thế 110kV”. Hội thảo Quốc gia lần thứ XXVI: Một số vấn đề chọn lọc của công nghệ thông tin và truyền thông, Bắc Ninh, 5–6/10/2023, pp.206–209.
- CT6.** Thi-Thanh-Tan Nguyen, **Anh Trinh Hien**, Tao Ngo Quoc, “HVE-YOLO: A multi-scale neural network integrated for high-voltage power grid equipment recognition”, *Journal of Information Hiding and Multimedia Signal Processing*, Accepted