

BỘ KHOA HỌC VÀ CÔNG NGHỆ
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG
—oo—



ĐỀ CƯƠNG

ĐÁNH GIÁ HIỆU QUẢ CÁC CHIẾN LƯỢC ACTIVE
LEARNING
CHO BÀI TOÁN PHÁT HIỆN RÁC THẢI SINH HOẠT
THEO VẬT LIỆU SỬ DỤNG YOLOv8

Giảng viên hướng dẫn:	Kim Ngọc Bách
Sinh viên:	Chu Tuyết Nhi
Mã sinh viên:	B23DCCE075
Lớp:	D23CQCEO6-B
Niên khóa:	2023 - 2028
Hệ đào tạo:	Đại học chính quy

Hà Nội, 2026

DANH MỤC CÁC TỪ VIẾT TẮT

Từ viết tắt	Tiếng Anh	Tiếng Việt
AL	Active Learning	Học chủ động
AP	Average Precision	Độ chính xác trung bình (một lớp)
CNN	Convolutional Neural Network	Mạng nơ-ron tích chập
COCO	Common Objects in Context	Tập dữ liệu chuẩn phát hiện đối tượng
FPN	Feature Pyramid Network	Mạng kim tự tháp đặc trưng
IoU	Intersection over Union	Tỷ lệ giao trên hợp
mAP	mean Average Precision	Độ chính xác trung bình (trên tất cả các lớp)
VOC	Visual Object Classes (PASCAL)	Tập dữ liệu chuẩn phát hiện đối tượng
YOLO	You Only Look Once	Kiến trúc phát hiện đối tượng một lần

MỤC LỤC

DANH MỤC CÁC TỪ VIẾT TẮT	i
1 GIỚI THIỆU	1
1.1 Lý do chọn đề tài	1
1.2 Mục tiêu nghiên cứu	2
1.2.1 Mục tiêu tổng quát	2
1.2.2 Mục tiêu cụ thể	2
1.2.3 Câu hỏi nghiên cứu	2
1.3 Ý nghĩa, tính ứng dụng và giá trị học thuật	2
1.4 Đối tượng và phạm vi nghiên cứu	3
1.5 Phương pháp nghiên cứu	3
2 CƠ SỞ LÝ THUYẾT VÀ CÔNG NGHỆ SỬ DỤNG	4
2.1 Phát hiện đối tượng và YOLOv8	4
2.1.1 Bài toán phát hiện đối tượng	4
2.1.2 Kiến trúc YOLOv8	5
2.2 Active Learning	5
2.2.1 Khái niệm và quy trình	5
2.2.2 Các chiến lược truy vấn	6
2.2.3 Active Learning cho phát hiện đối tượng	6
2.3 Công nghệ sử dụng	7
2.4 Tổng quan nghiên cứu liên quan	7
2.4.1 Nghiên cứu ngoài nước	7
2.4.2 Nghiên cứu trong nước	7
2.4.3 Khoảng trống nghiên cứu	8
3 PHÂN TÍCH YÊU CẦU DỰ ÁN	9
3.1 Mô tả tổng quan	9

3.2	Yêu cầu phần nghiên cứu	9
3.2.1	Bộ dữ liệu	9
3.2.2	Thiết kế thí nghiệm	9
3.2.3	Chiến lược truy vấn	10
3.2.4	Chỉ số đánh giá	10
3.3	Yêu cầu ứng dụng demo	11
4	KẾ HOẠCH THỰC HIỆN	12
4.1	Tiến độ thực hiện đề tài	12
4.2	Kết quả dự kiến	12
4.3	Bô cục dự kiến của báo cáo cuối kỳ	13

DANH SÁCH BẢNG

2.1	Công nghệ dự kiến sử dụng	7
3.1	Tham số thí nghiệm	10
3.2	Các chiến lược truy vấn	10
3.3	Yêu cầu chức năng ứng dụng demo	11
4.1	Tiến độ dự kiến thực hiện đề tài	12

CHƯƠNG 1

GIỚI THIỆU

1.1. Lý do chọn đề tài

Về mặt thực tiễn, quản lý chất thải rắn sinh hoạt là vấn đề môi trường cấp bách tại Việt Nam. Theo ông Hồ Trung Kiên, Phó Cục trưởng Cục Môi trường, tính đến ngày 30 tháng 6 năm 2025, tổng lượng chất thải rắn sinh hoạt phát sinh trên toàn quốc đạt khoảng 69,4 nghìn tấn mỗi ngày, trong đó tỷ lệ thu gom và xử lý chỉ đạt khoảng 91% [1]. Phần còn lại vẫn bị chôn lấp không phân loại hoặc xử lý không đúng quy định. Trong khi đó, công tác phân loại tại nguồn mới chỉ triển khai thí điểm ở quy mô nhỏ, hạ tầng thu gom và xử lý chưa đồng bộ trên phạm vi cả nước.

Từ ngày 01 tháng 01 năm 2025, Luật Bảo vệ Môi trường 2020 quy định phân loại chất thải rắn sinh hoạt tại nguồn là nghĩa vụ bắt buộc của hộ gia đình và cá nhân theo Điều 75 và Điều 79 [2]. Tuy nhiên, người dân vẫn thiếu công cụ hỗ trợ nhận biết và phân loại nhanh chóng theo vật liệu. Một ứng dụng có khả năng tự động phát hiện và phân loại rác thải từ ảnh chụp sẽ giúp người dùng tra cứu nhanh loại vật liệu, từ đó hỗ trợ phân loại đúng quy định.

Về mặt lý luận, YOLOv8 đã được kiểm chứng đạt độ chính xác cao trong phát hiện và phân loại rác thải, tuy nhiên các nghiên cứu hiện nay chủ yếu được huấn luyện và đánh giá trên các tập dữ liệu đã gán nhãn đầy đủ [3, 4]. Trong khi đó, bài toán phát hiện đối tượng (*object detection*) nói chung đòi hỏi gán nhãn bounding box, vốn tốn kém và phức tạp hơn đáng kể so với gán nhãn phân loại ảnh (*classification*), làm gia tăng chi phí xây dựng dữ liệu huấn luyện. Điều này đặt ra nhu cầu nghiên cứu các chiến lược active learning (AL) nhằm tối ưu hóa quá trình lựa chọn mẫu và giảm chi phí gán nhãn [5].

Mặc dù hiệu quả của AL đã được xác nhận trên các tập dữ liệu tổng quát như COCO và VOC [6, 7], hiện tại vẫn còn rất ít tài liệu kiểm chứng một cách hệ thống phương pháp này trên các dataset rác thải sinh hoạt đa dạng. Đây chính là khoảng trống nghiên cứu mà đề tài hướng tới.

Xuất phát từ nhu cầu thực tiễn về công cụ hỗ trợ phân loại rác thải và từ câu hỏi lý luận về hiệu quả của active learning trong bối cảnh dữ liệu rác thải sinh hoạt, đề tài đặt mục tiêu đánh giá và so sánh các chiến lược active learning cho bài toán phát hiện rác theo vật liệu sử dụng YOLOv8, đồng thời xây dựng ứng dụng web demo nhằm minh họa khả năng triển khai trong thực tế.

1.2. Mục tiêu nghiên cứu

1.2.1. Mục tiêu tổng quát

Đánh giá và so sánh hiệu quả các chiến lược active learning cho bài toán phát hiện rác thải sinh hoạt theo vật liệu sử dụng YOLOv8 trong điều kiện ngân sách gán nhãn bị giới hạn, đồng thời xây dựng ứng dụng web demo minh họa kết quả phát hiện và phân loại.

1.2.2. Mục tiêu cụ thể

1. Xây dựng bộ dữ liệu phát hiện rác thải sinh hoạt với nhãn bounding box theo năm lớp vật liệu chính (nhựa, giấy, kim loại, thủy tinh, thực phẩm) nhằm hỗ trợ phân loại theo ba nhóm quy định tại Luật Bảo vệ Môi trường 2020.
2. So sánh bốn chiến lược truy vấn: random sampling, uncertainty sampling, diversity sampling (CoreSet), và hybrid.
3. Xác định tỷ lệ dữ liệu gán nhãn tối thiểu để đạt $mAP@50 \geq 85\%$ so với mô hình oracle baseline.
4. Phân tích ảnh hưởng của kích thước mô hình khi chuyển từ YOLOv8n sang YOLOv8s.
5. Xây dựng ứng dụng web demo bằng Gradio, cho phép người dùng tải ảnh lên và xem kết quả phát hiện, phân loại rác thải trực tiếp trên trình duyệt.

1.2.3. Câu hỏi nghiên cứu

1. Chiến lược truy vấn nào đạt $mAP@50$ cao nhất khi ngân sách gán nhãn bị giới hạn trên miền dữ liệu rác thải sinh hoạt theo vật liệu?
2. Cần gán nhãn bao nhiêu % tập dữ liệu để đạt ngưỡng $mAP@50$ chấp nhận được?
3. Hiệu quả tương đối của các chiến lược có thay đổi khi chuyển từ YOLOv8n sang YOLOv8s không?
4. Ứng dụng web demo có thể trình diễn kết quả phát hiện và phân loại rác thải với thời gian phản hồi chấp nhận được không?

1.3. Ý nghĩa, tính ứng dụng và giá trị học thuật

Đề tài góp phần giảm chi phí xây dựng tập dữ liệu cho các bài toán thị giác máy tính liên quan đến rác thải, thông qua việc xác định chiến lược active learning phù hợp. Ứng dụng web demo

giúp người dùng tải ảnh và xem kết quả phân loại trực tiếp trên trình duyệt, minh họa khả năng ứng dụng thực tế của mô hình. Về mặt học thuật, đề tài cung cấp đánh giá định lượng về hiệu quả các chiến lược active learning trên miền rác thải sinh hoạt với YOLOv8. Đường cong học tập (*learning curve*) và phân tích ngân sách gán nhãn theo từng chiến lược sẽ là tài liệu tham khảo cho các nghiên cứu tiếp theo.

1.4. Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu gồm các chiến lược active learning pool-based, mô hình YOLOv8n và YOLOv8s, cùng ảnh rác thải sinh hoạt theo năm lớp vật liệu chính (nhựa, giấy, kim loại, thủy tinh, thực phẩm) nhằm hỗ trợ phân loại theo ba nhóm quy định pháp luật trong tương lai.

Phạm vi nội dung tập trung vào năm lớp vật liệu trên, không bao gồm rác công nghiệp hoặc y tế. Phạm vi không gian là ảnh từ các nền tảng dữ liệu mở uy tín (Roboflow Universe, Kaggle, GitHub) và tự chụp tại Thành phố Hà Nội. Phạm vi thời gian từ tháng 02 năm 2026 đến tháng 07 năm 2026.

1.5. Phương pháp nghiên cứu

Đề tài sử dụng phương pháp thực nghiệm, gồm các bước chính:

1. Khảo sát tài liệu về active learning cho phát hiện đối tượng.
2. Thu thập, chuẩn hóa và gán nhãn bộ dữ liệu rác thải sinh hoạt theo 5 lớp vật liệu.
3. Cài đặt vòng lặp active learning pool-based với 4 chiến lược truy vấn.
4. So sánh định lượng hiệu quả các chiến lược theo mAP@50 qua từng vòng lặp.
5. Phân tích ảnh hưởng của kích thước mô hình (YOLOv8n, YOLOv8s).
6. Phát triển ứng dụng web demo bằng Gradio.

CHƯƠNG 2

CƠ SỞ LÝ THUYẾT VÀ CÔNG NGHỆ SỬ DỤNG

2.1. Phát hiện đối tượng và YOLOv8

2.1.1. Bài toán phát hiện đối tượng

Phát hiện đối tượng là bài toán yêu cầu xác định đồng thời vị trí và lớp của từng đối tượng trong ảnh, với đầu ra là danh sách các hộp giới hạn (*bounding box*) kèm nhãn lớp và điểm tin cậy. Mạng nơ-ron tích chập (*Convolutional Neural Network*, CNN) là nền tảng chung cho các kiến trúc phát hiện đối tượng hiện đại [8].

Hiệu năng mô hình được đánh giá qua các chỉ số precision, recall, IoU và mAP [9]. Precision đo tỷ lệ dự đoán đúng trên tổng dự đoán:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2.1)$$

Recall đo tỷ lệ đối tượng thực được phát hiện đúng:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2.2)$$

IoU (*Intersection over Union*) đánh giá độ chồng lấp giữa hộp dự đoán và hộp thực:

$$\text{IoU} = \frac{\text{Area}(B_p \cap B_{gt})}{\text{Area}(B_p \cup B_{gt})} \quad (2.3)$$

Average Precision (AP) là diện tích dưới đường cong precision-recall cho một lớp. Mean Average Precision (mAP) là trung bình AP trên tất cả các lớp:

$$\text{mAP} = \frac{1}{C} \sum_{c=1}^C \text{AP}_c \quad (2.4)$$

Đề tài sử dụng mAP@50 ($\text{IoU} \geq 0,5$) làm chỉ số chính vì đây là metric phổ biến trong các công trình về YOLO và phát hiện đối tượng thực tế, cho phép so sánh trực tiếp với các nghiên cứu trước đồng thời cân bằng giữa độ chính xác vị trí và khả năng phát hiện.

2.1.2. Kiến trúc YOLOv8

YOLO được Redmon et al. giới thiệu năm 2016 với cách tiếp cận hồi quy thực hiện trong một lần truyền xuôi duy nhất, cho phép phát hiện đối tượng theo thời gian thực [10]. Các phiên bản sau liên tục cải tiến backbone, neck và head để nâng cao độ chính xác và tốc độ [11].

YOLOv8 do Ultralytics phát triển năm 2023 là phiên bản cải tiến với ba thành phần chính [11]:

- **Backbone** (CSPDarknet với mô-đun C2f): trích xuất đặc trưng đa tầng hiệu quả.
- **Neck** (PANet kết hợp FPN): tổng hợp đặc trưng ở nhiều độ phân giải.
- **Head** (anchor-free, decoupled): tách riêng nhánh phân loại và hồi quy vị trí, giảm số lượng tham số.

Để tài chọn hai biến thể YOLOv8n và YOLOv8s để khảo sát ảnh hưởng của kích thước mô hình đến hiệu quả active learning [11]. Cả hai biến thể đều có tốc độ huấn luyện nhanh, phù hợp với yêu cầu chạy nhiều vòng lặp trên tài nguyên hạn chế, trong khi khoảng cách hiệu năng giữa hai biến thể đủ để quan sát ảnh hưởng của kích thước mô hình lên kết quả active learning.

2.2. Active Learning

2.2.1. Khái niệm và quy trình

Active learning là phương pháp học máy trong đó mô hình chủ động chọn các mẫu cần gán nhãn từ tập chưa gán nhãn, thay vì gán nhãn toàn bộ dữ liệu [5]. Phương pháp này đặc biệt phù hợp khi chi phí gán nhãn cao, vì cho phép mô hình đạt hiệu năng tương đương với lượng nhãn thấp hơn đáng kể. Trong kịch bản pool-based, quy trình mỗi vòng lặp gồm năm bước:

1. Huấn luyện mô hình trên tập đã gán nhãn \mathcal{L} .
2. Đánh giá và chấm điểm từng mẫu trong tập chưa gán nhãn U theo một hàm truy vấn cụ thể.
3. Chọn batch b mẫu có điểm cao nhất theo chiến lược truy vấn.
4. Gán nhãn batch và chuyển từ U sang \mathcal{L} .
5. Lặp lại đến khi hết ngân sách hoặc hiệu năng bão hòa.

2.2.2. Các chiến lược truy vấn

Uncertainty sampling ưu tiên mẫu mà mô hình kém tự tin nhất [5]. Hai cách đo phổ biến là least confidence:

$$x_{LC}^* = \arg \min_{x \in \mathcal{U}} \max_y P(y | x; \theta) \quad (2.5)$$

và entropy:

$$x_H^* = \arg \max_{x \in \mathcal{U}} \left[- \sum_y P(y | x; \theta) \log P(y | x; \theta) \right] \quad (2.6)$$

Ưu điểm của uncertainty sampling là đơn giản, dễ cài đặt và chi phí tính toán thấp; nhưng điểm là dễ chọn outlier khi mô hình chưa ổn định [5].

Diversity sampling (CoreSet) tối đa hóa độ bao phủ của tập được chọn trong không gian đặc trưng học được của backbone [12]:

$$B^* = \arg \min_{B \subset \mathcal{U}, |B|=b} \max_{x \in \mathcal{U} \setminus B} \min_{x' \in \mathcal{L} \cup B} d(f(x), f(x')) \quad (2.7)$$

trong đó $f(x)$ là vector đặc trưng và $d(\cdot, \cdot)$ là khoảng cách Euclidean. Ưu điểm là giảm redundancy trong tập huấn luyện; nhưng điểm là chi phí tính toán cao khi tập \mathcal{U} lớn [12].

Hybrid kết hợp uncertainty và diversity nhằm tận dụng ưu điểm của cả hai. Theo khảo sát của Li et al. [13], nhiều chiến lược hybrid kết hợp uncertainty và diversity có thể đạt hiệu năng cạnh tranh chỉ với một phần nhỏ lượng gán nhãn so với toàn bộ tập dữ liệu. Phương án kết hợp cụ thể của đề tài sẽ được xác định sau giai đoạn khảo sát ban đầu.

2.2.3. Active Learning cho phát hiện đối tượng

Áp dụng active learning cho phát hiện đối tượng phức tạp hơn phân loại ảnh vì mỗi ảnh chứa nhiều detection với điểm uncertainty riêng biệt [14]. Để tổng hợp uncertainty từ mức detection lên mức ảnh, đề tài áp dụng max aggregation:

$$U_{\text{image}} = \max_{i=1}^N U_{\text{det},i} \quad (2.8)$$

Max aggregation được chọn vì ưu tiên ảnh chứa ít nhất một detection khó, giúp mô hình tập trung vào mẫu mang nhiều thông tin nhất thay vì trung bình hóa. Choi et al [14] xác nhận chiến lược này hiệu quả trong các bài toán phát hiện đối tượng.

2.3. Công nghệ sử dụng

Bảng 2.1: Công nghệ dự kiến sử dụng

Công nghệ	Phiên bản	Vai trò
Python	3.10+	Ngôn ngữ lập trình chính
Ultralytics YOLOv8	8.3+	Huấn luyện và suy luận mô hình
PyTorch [15]	2.0+	Framework deep learning nền tảng
Label Studio [16]	–	Công cụ gán nhãn bounding box
Gradio [17]	4.0+	Xây dựng giao diện web demo

2.4. Tổng quan nghiên cứu liên quan

2.4.1. Nghiên cứu ngoài nước

Rastari et al. đề xuất mô hình phát hiện và phân loại rác tái chế sử dụng YOLOv8 cho quản lý rác thời gian thực, đạt precision 95,3% trên bài toán phân loại vật liệu [3]. Nghiên cứu khẳng định khả năng ứng dụng YOLOv8 trong miền rác thải nhưng chưa xem xét chi phí gán nhãn và hiệu quả khi dữ liệu bị hạn chế.

Zambri et al. xây dựng hệ thống nhận dạng rác tái chế bằng YOLOv8 phục vụ giáo dục môi trường, đạt accuracy trên 94% và mAP@50 bằng 94,9% [4]. Hạn chế của nghiên cứu là sử dụng tập dữ liệu đã gán nhãn đầy đủ, chưa đề cập đến các chiến lược giảm chi phí gán nhãn.

Kao et al. giới thiệu chiến lược localization-aware active learning cho phát hiện đối tượng trên tập VOC và COCO, giảm được 25% lượng nhãn cần gán so với random sampling [6]. Đây là một trong những nghiên cứu nền tảng về active learning cho phát hiện đối tượng nhưng chưa được kiểm chứng trên miền rác thải.

Wu et al. kết hợp entropy với ràng buộc đa dạng tiến triển trong mô hình Faster R-CNN, vượt trội so với baseline tại mức 40% dữ liệu gán nhãn trên COCO [7]. Kết quả này là căn cứ lý thuyết để đề tài xây dựng chiến lược hybrid.

Qin et al. áp dụng active learning vào kiến trúc DETR cho phát hiện rác nhà bếp, chứng minh hiệu quả của active learning trong bài toán rác thải nhưng giới hạn ở DETR, một kiến trúc nặng hơn và khó triển khai thực tế hơn YOLOv8 [18].

2.4.2. Nghiên cứu trong nước

Vo et al. xây dựng tập dữ liệu VN-trash gồm 5.904 ảnh rác thải Việt Nam và áp dụng deep transfer learning để phân loại theo ba nhóm lớn, đạt accuracy 98% trên tập dữ liệu này [19].

Đây là công trình nền tảng quan trọng nhất về rác thải trong nước, nhưng chỉ dừng ở mức phân loại ảnh, chưa sử dụng bounding box và chưa kết hợp active learning.

2.4.3. Khoảng trống nghiên cứu

Qua khảo sát tài liệu, chưa có nghiên cứu nào thực hiện đánh giá so sánh hệ thống các chiến lược active learning trên mô hình YOLOv8 cho bài toán phát hiện rác thải sinh hoạt theo vật liệu. Đặc biệt, rác thải sinh hoạt có những đặc thù rõ ràng so với các tập dữ liệu chuẩn COCO hay VOC: hình dạng đối tượng biến đổi nhiều do rác thường bị vò nát hoặc biến dạng, bề mặt phức tạp với nhựa trong suốt hay kim loại phản chiếu, và tỷ lệ che khuất (*occlusion*) cao khi rác chồng chất. Những đặc thù này đặt ra câu hỏi liệu các chiến lược AL đã được kiểm chứng trên dữ liệu tổng quát có còn hiệu quả khi áp dụng vào miền này hay không. Ngoài ra, các nghiên cứu hiện có hoặc ứng dụng YOLOv8 trên dữ liệu gán nhãn đầy đủ mà chưa xem xét chi phí gán nhãn, hoặc kết hợp AL với kiến trúc khác (DETR) mà khó triển khai thực tế hơn. Khoảng trống này là cơ sở trực tiếp để thực hiện đề tài.

CHƯƠNG 3

PHÂN TÍCH YÊU CẦU DỰ ÁN

3.1. Mô tả tổng quan

Đề tài gồm hai phần chính. Phần thứ nhất là nghiên cứu thực nghiệm so sánh các chiến lược active learning cho bài toán phát hiện rác thải sinh hoạt theo vật liệu với YOLOv8. Phần thứ hai là ứng dụng web demo sử dụng mô hình đạt kết quả tốt nhất từ thực nghiệm, minh họa khả năng triển khai thực tế của hệ thống.

3.2. Yêu cầu phần nghiên cứu

3.2.1. Bộ dữ liệu

Tập dữ liệu được xây dựng từ hai nguồn: ảnh thu thập từ các nền tảng dữ liệu mở uy tín (Roboflow Universe, Kaggle, GitHub) và ảnh tự thu thập tại Thành phố Hà Nội. Dữ liệu được chia theo tỷ lệ 80% huấn luyện, 10% kiểm định và 10% kiểm tra. Tập kiểm tra được tách cố định ngay từ đầu và không tham gia vào vòng lặp active learning, nhằm đảm bảo tính khách quan trong đánh giá.

Nhãn được gán theo năm lớp vật liệu chính (nhựa, giấy, kim loại, thủy tinh, thực phẩm) nhằm hỗ trợ phân loại theo ba nhóm quy định tại Luật Bảo vệ Môi trường 2020. Toàn bộ nhãn bounding box được chuẩn hóa sang định dạng YOLO bằng Label Studio [16]. Trong trường hợp phân phối lớp mắt cá bằng nghiêm trọng, kỹ thuật tăng cường dữ liệu sẽ được áp dụng trong quá trình huấn luyện.

3.2.2. Thiết kế thí nghiệm

Thí nghiệm sử dụng quy trình pool-based active learning với các tham số chính được trình bày trong Bảng 3.1.

Bảng 3.1: Tham số thí nghiệm

Tham số	Giá trị	Ghi chú
Tập khởi tạo	~5% tập huấn luyện	Chọn phân tầng theo lớp
Kích thước truy vấn	~5% tập huấn luyện	Số ảnh chọn thêm mỗi vòng
Số vòng lặp	8–10	Dừng khi hết ngân sách hoặc bão hòa
Tổng ngân sách	50% tập huấn luyện	Giới hạn tối đa ảnh được gán nhãn

Quy trình gồm các bước: tách tập kiểm tra cố định, chọn tập khởi tạo phân tầng, lặp vòng active learning, ghi nhận mAP@50 sau mỗi vòng. Oracle baseline (mô hình tròn) là mô hình được huấn luyện trên 100% tập dữ liệu huấn luyện đã được gán nhãn, đóng vai trò ngưỡng hiệu năng tham chiếu tối đa. Mỗi cấu hình được chạy lặp lại 3 lần và sai khác giữa các chiến lược được kiểm định bằng paired t-test ($\alpha = 0,05$) để đảm bảo tính thống kê.

3.2.3. Chiến lược truy vấn

Bốn chiến lược truy vấn được triển khai và so sánh, chi tiết tại Bảng 3.2.

Bảng 3.2: Các chiến lược truy vấn

ID	Chiến lược	Mô tả
S0	Random sampling	Chọn ngẫu nhiên, đóng vai trò baseline so sánh
S1	Uncertainty sampling	Ưu tiên ảnh có điểm uncertainty cao nhất sau max aggregation
S2	Diversity sampling	Ưu tiên tập ảnh tối đa hóa độ bao phủ và mang tính đại diện cao nhất trong không gian đặc trưng.
S3	Hybrid	Kết hợp uncertainty và diversity theo cơ chế được xác định sau khảo sát ban đầu

3.2.4. Chỉ số đánh giá

Chỉ số chính là mAP@50, phản ánh hiệu năng phát hiện đối tượng tổng thể theo tiêu chuẩn IoU $\geq 0,5$. Ngoài ra, đề tài thu thập đường cong học tập theo số vòng lặp, tỷ lệ ngân sách gán nhãn cần thiết để đạt ngưỡng mục tiêu, và AP riêng từng lớp vật liệu để phân tích hành vi từng chiến lược trên các lớp có đặc thù khác nhau.

3.3. Yêu cầu ứng dụng demo

Các yêu cầu chức năng của ứng dụng web demo được liệt kê trong Bảng 3.3.

Bảng 3.3: Yêu cầu chức năng ứng dụng demo

ID	Chức năng	Mô tả
FR-01	Tải ảnh	Người dùng tải lên ảnh rác thải từ thiết bị
FR-02	Phát hiện và phân loại	Xác định vị trí và loại vật liệu của từng đối tượng
FR-03	Hiển thị kết quả	Hiển thị ảnh kèm bounding box, nhãn lớp và điểm tin cậy
FR-04	Chọn mô hình	Chuyển đổi giữa YOLOv8n và YOLOv8s

Ứng dụng được xây dựng bằng Gradio [17], cho phép truy cập trực tiếp qua trình duyệt web mà không cần cài đặt phần mềm phía người dùng.

CHƯƠNG 4

KẾ HOẠCH THỰC HIỆN

4.1. Tiết độ thực hiện đề tài

Bảng 4.1: Tiết độ dự kiến thực hiện đề tài

Nội dung thực hiện	Tháng (năm 2026)					
	2	3	4	5	6	7
Hoàn thiện đề cương, chuẩn bị dữ liệu ban đầu						
Khảo sát tài liệu active learning cho phát hiện đối tượng						
Thu thập, gán nhãn dữ liệu						
Huấn luyện oracle baseline						
Cài đặt các chiến lược active learning						
Chạy thí nghiệm active learning						
Phân tích kết quả						
Xây dựng ứng dụng web demo						
Viết báo cáo nghiên cứu						
Hoàn thiện đề tài						

4.2. Kết quả dự kiến

Sau khi hoàn thành, đề tài dự kiến đạt được các sản phẩm sau:

1. Bộ dữ liệu phát hiện rác thải sinh hoạt với nhãn bounding box theo năm lớp vật liệu chính.
2. Bộ mô hình YOLOv8 đã huấn luyện (định dạng .pt), gồm oracle baseline và các mô hình từ từng chiến lược active learning.
3. Đường cong học tập so sánh bốn chiến lược trên hai biến thể YOLOv8n và YOLOv8s.
4. Phân tích tỷ lệ ngân sách gán nhãn cho từng chiến lược kèm kiểm định thống kê paired t-test.
5. Nhận xét và kết luận về chiến lược active learning phù hợp cho miền rác thải sinh hoạt.

6. Ứng dụng web demo nhận diện rác thải thông qua Gradio.
7. Mã nguồn thí nghiệm hoàn chỉnh.
8. Báo cáo nghiên cứu đầy đủ.

4.3. Bộ cục dự kiến của báo cáo cuối kỳ

Chương 1. GIỚI THIỆU CHUNG

- 1.1 Lý do chọn đề tài
- 1.2 Mục tiêu nghiên cứu
- 1.3 Ý nghĩa, tính ứng dụng và giá trị học thuật
- 1.4 Đối tượng và phạm vi nghiên cứu
- 1.5 Phương pháp nghiên cứu

Chương 2. CƠ SỞ LÝ THUYẾT VÀ CÔNG NGHỆ SỬ DỤNG

- 2.1 Phát hiện đối tượng và YOLOv8
- 2.2 Active Learning
- 2.3 Công nghệ sử dụng
- 2.4 Tổng quan nghiên cứu liên quan

Chương 3. PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG

- 3.1 Bộ dữ liệu
- 3.2 Thiết kế thí nghiệm active learning
- 3.3 Các chiến lược truy vấn
- 3.4 Chỉ số đánh giá
- 3.5 Thiết kế ứng dụng demo

Chương 4. KẾT QUẢ CÀI ĐẶT VÀ THỬ NGHIỆM

- 4.1 Oracle baseline
- 4.2 So sánh các chiến lược truy vấn

4.3 Phân tích ngân sách gán nhãn

4.4 Phân tích theo lớp vật liệu

4.5 Ứng dụng demo

4.6 Thảo luận

Chương 5. KẾT LUẬN

5.1 Kết luận

5.2 Hướng phát triển

TÀI LIỆU THAM KHẢO

PHỤ LỤC

TÀI LIỆU THAM KHẢO

- [1] T. Dương. (2025, Sep. 5). “Mỗi ngày cả nước phát sinh gần 70 nghìn tấn chất thải rắn sinh hoạt” [Online]. Từ: <https://vneconomy.vn/moi-ngay-ca-nuoc-phat-sinh-gan-70-nghin-tan-chat-thai-ran-sinh-hoat.htm>
- [2] Quốc hội nước CHXHCN Việt Nam. “Luật Bảo vệ Môi trường, Luật số 72/2020/QH14,” ngày 17 tháng 11 năm 2020. [Online]. Từ: <https://thuvienphapluat.vn/van-ban/Tai-nguyen-Moi-truong/Luat-so-72-2020-QH14-Bao-ve-moi-truong-2020-431147.aspx>
- [3] M. A. M. Rastari, R. Roslan, R. Hamzah, N. H. I. Teo, F. E. Shahbudin, and K. A. F. A. Samah, “Recycle Waste Detection and Classification Model Using YOLO-V8 for Real-time Waste Management,” in *Proc. IEEE Int. Conf. Artif. Intell. Eng. Technol. (IICAIET)*, Kota Kinabalu, Malaysia, Aug. 2024, pp. 372–377, doi: 10.1109/IICAIET62352.2024.10730703. [Online]. Từ: <https://ieeexplore.ieee.org/document/10730703/>
- [4] A. F. Zambri, S. Abdul-Rahman, N. M. Sabri, and S. Mutalib, “A Recyclable Waste Image Recognition System with YOLOv8 for Children’s Environmental Education,” *Eng. Technol. Appl. Sci. Res.*, vol. 14, no. 5, pp. 16492–16498, Oct. 2024, doi: 10.48084/etasr.7879. [Online]. Từ: <https://etasr.com/index.php/ETASR/article/view/7879>
- [5] Burr Settles, “Active Learning Literature Survey,” Computer Sciences Technical Report 1648, University of Wisconsin–Madison, Jan. 2009. [Online]. Từ: <https://burrsettles.com/pub/settles.activelearning.pdf>
- [6] C.C. Kao, T.Y. Lee, P. Sen, and M.Y. Liu, “Localization-Aware Active Learning for Object Detection,” in *Proc. Asian Conf. Comput. Vis. (ACCV)*, Perth, Australia, 2018, pp. 506–522. [Online]. Từ: <https://arxiv.org/abs/1801.05124>
- [7] J. Wu, J. Chen, and D. Huang, “Entropy-based Active Learning for Object Detection with Progressive Diversity Constraint,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New Orleans, LA, USA, 2022, pp. 9387–9396, doi: 10.1109/CVPR52688.2022.00918. [Online]. Từ: <https://arxiv.org/abs/2204.07965>
- [8] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Từ: <https://www.deeplearningbook.org/>

- [9] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “The PASCAL Visual Object Classes (VOC) Challenge,” *Int. J. Comput. Vis.*, vol. 88, no. 2, pp. 303–338, Jun. 2010, doi: 10.1007/s11263-009-0275-4. [Online]. Từ: <https://link.springer.com/content/pdf/10.1007/s11263-009-0275-4.pdf>
- [10] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91. [Online]. Từ: <https://arxiv.org/abs/1506.02640>
- [11] Ultralytics. (2023). *Ultralytics YOLOv8* [Software]. [Online]. Từ: <https://github.com/ultralytics/ultralytics>
- [12] O. Sener and S. Savarese, “Active Learning for Convolutional Neural Networks: A Core-Set Approach,” in *Proc. Int. Conf. Learn. Representations (ICLR)*, Vancouver, BC, Canada, 2018. [Online]. Từ: <https://openreview.net/forum?id=H1aIuk-RW>
- [13] X. Li, S. Liu, K. Li, and B. Han, “A Survey on Deep Active Learning: Recent Advances and New Frontiers,” *arXiv:2405.00334* [cs.LG], May 2024. [Online]. Available: <https://arxiv.org/abs/2405.00334>
- [14] J. Choi, I. Elezi, H.-J. Lee, C. Farabet, and J. M. Alvarez, “Active Learning for Deep Object Detection via Probabilistic Modeling,” in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Montreal, QC, Canada, 2021, pp. 10305–10314, doi: 10.1109/ICCV48922.2021.01014. [Online]. Từ: <https://arxiv.org/abs/2103.16130>
- [15] A. Paszke et al., “PyTorch: An Imperative Style, High-Performance Deep Learning Library,” in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 32, 2019. [Online]. Từ: <https://arxiv.org/abs/1912.01703>
- [16] HumanSignal, Inc. (2024). *Label Studio: Open Source Data Labeling Tool* [Software]. [Online]. Từ: <https://labelstud.io/>
- [17] A. Abid, A. Abdalla, A. Abid, D. Khan, A. Alfozan, and J. Zou, “Gradio: Hassle-Free Sharing and Testing of ML Models in the Wild,” *arXiv:1906.02569* [cs.LG], Jun. 2019. [Online]. Từ: <https://arxiv.org/abs/1906.02569>
- [18] H. Qin, L. Shu, L. Zhou, and Y. Wang, “Active Learning-DETR: Cost-Effective Object Detection for Kitchen Waste,” *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–15, 2024, doi: 10.1109/TIM.2024.3368494. [Online]. Từ: <https://ieeexplore.ieee.org/document/10443710>
- [19] A. H. Vo, L. H. Son, M. T. Vo, and T. Le, “A Novel Framework for Trash Classification Using Deep Transfer Learning,” *IEEE Access*, vol. 7, pp. 178631–178639, Dec. 2019,

doi: 10.1109/ACCESS.2019.2959033. [Online]. Từ: <https://www.researchgate.net/publication/344650744>