

## NGHIÊN CỨU THUẬT TOÁN NHẬN DIỆN VÀ PHÂN LOẠI RÁC THẢI NHỰA TỰ ĐỘNG BẰNG YOLOV8

### STUDY ON PLASTIC WASTE DETECTION OF SORTING SYSTEM USING YOLOV8

Lưu Thanh Tùng<sup>1,2</sup>, Nguyễn Trí Thanh<sup>1,2</sup>, Nguyễn Hoàng Long<sup>1,2</sup>, Tống Phước Thanh An<sup>1,2</sup>

<sup>1</sup>Khoa Cơ khí, Trường Đại học Bách Khoa, Đại học Quốc gia Thành phố Hồ Chí Minh (HCMUT)

<sup>2</sup>Đại học Quốc gia Thành phố Hồ Chí Minh (VNU-HCM)

#### TÓM TẮT

*Rác thải nhựa là một vấn đề nóng không chỉ ở Việt Nam mà còn trên toàn thế giới. Rác thải nhựa không qua xử lý và tái chế được xả thẳng ra môi trường đã làm biến đổi hệ sinh thái, giết chết các loại động vật, thực vật và gián tiếp đầu độc đến cuộc sống của con người. Vì vậy, việc phân loại và tái chế rác thải nhựa theo quy mô công nghiệp là vô cùng cần thiết để giảm thiểu mức độ ô nhiễm môi trường và tận dụng được nhựa đã qua sử dụng. Vì vậy, chúng tôi đề xuất việc sử dụng mô hình Học sâu YOLOv8, một mô hình dạng nhận diện mạnh nhất hiện nay để phân loại các loại rác thải nhựa trên dây chuyền thực tế trong nhà máy. Bằng cách huấn luyện và phân tích các chỉ số của mô hình, chúng tôi đã chứng minh sự hiệu quả của YOLOv8 trong lĩnh vực phân loại rác thải nhựa theo thời gian thực. Đây cũng là cơ sở để xây dựng nên các dây chuyền xử lý rác thải nhựa hoàn toàn tự động, đảm bảo năng suất phân loại và không gây ảnh hưởng lên sức khỏe của nhân viên làm việc.*

**Từ khóa:** Rác thải nhựa; YOLOv8; Mô hình nhận diện; Hệ thống phân loại tự động.

#### ABSTRACT

*Plastic waste is a pressing issue not only in Vietnam but also globally. Unprocessed plastic waste discharged directly into the environment has led to ecosystem disruption, the death of flora and fauna, and indirect threats to human life. Therefore, the industrial-scale sorting and recycling of plastic waste are essential to minimize environmental pollution and utilize used plastic effectively. We propose the utilization of the YOLOv8 Deep Learning model, currently one of the most robust object detection models, to classify various types of plastic waste on real production lines within factories. Through training and analyzing the model's metrics, we have demonstrated the effectiveness of YOLOv8 in real-time plastic waste classification. This forms the basis for establishing fully automated plastic waste processing lines, ensuring efficient sorting without compromising the health of factory workers.*

**Keywords:** Plastic waste; YOLOv8; Object detection models; Automation sorting system.

## 1. ĐẶT VẤN ĐỀ

Hiện nay, việc phân loại rác thải nhựa tại Việt Nam có sự tham gia của công nhân trong công đoạn phân loại một cách thủ công như ở một số doanh nghiệp Duy Tân, Coca Cola,... Bên cạnh việc gây ô nhiễm môi trường nếu không được xử lý đúng phương pháp [1], [2], rác thải nhựa gây ảnh hưởng trực tiếp đến sức khỏe của con người tham gia công việc này. Nguyên nhân xuất phát từ chất phụ gia trong sản phẩm nhựa có thể gây rối loạn nội tiết và gây ung thư thông qua con đường tiếp xúc hoặc hô hấp [3]. Do đó, cấp thiết phải có một hệ thống tự động thay thế cho nhân công thực hiện nhiệm vụ phân loại rác này, nhằm đảm bảo năng suất lao động và đồng thời bảo vệ sức khỏe của công nhân trước những hậu quả gây ra bởi phụ gia nhựa.

Có hai phương pháp ứng dụng công nghệ thông tin phổ biến trong việc phân loại rác thải nhựa. Thứ nhất, là trường phái ứng dụng các thuật toán Machine Learning như Linear Discriminant Analysis (LDA), Partial Least-Squares Discriminant Analysis (PLS-DA), Spectral Angle Mapper (SAM) và Support Vector Machine (SVM) [4] để phân loại. Một trường phái nghiên cứu khác là sử dụng mạng nơ-ron tích chập sâu CNN. Trong nước, tác giả Đỗ Thị Nhung và cộng sự đã sử dụng mạng nơ-ron tích chập CNN kết hợp với dữ liệu chụp được từ UAV phân loại hình ảnh chụp được từ thiết bị này có hay không rác thải nhựa trôi dạt trên biển. Độ chính xác trong phân loại chỉ đạt 0.87 [5]. Huỳnh Ngọc Thái Anh và Trang Thanh Trí ứng dụng mạng nơ-ron YOLOv7 để nhận dạng rác trên mặt nước. Độ chính xác của mô hình là 80.5%, với tốc độ chuyển khung hình FPS 17,6 [6]. Bên cạnh tình hình trong nước, Bobulski và cộng sự đề xuất mạng CNN 15 lớp, giảm số lớp so với mạng AlexNet gồm

có 23 lớp để phù hợp với ảnh chụp công nghiệp hơn. Tuy nhiên, độ chính xác của mô hình phân loại chỉ đạt 74% [7].

Ở nghiên cứu này, chúng tôi thử nghiệm xây dựng mô hình YOLOv8 để phân loại rác thải nhựa.

## 2. PHƯƠNG PHÁP

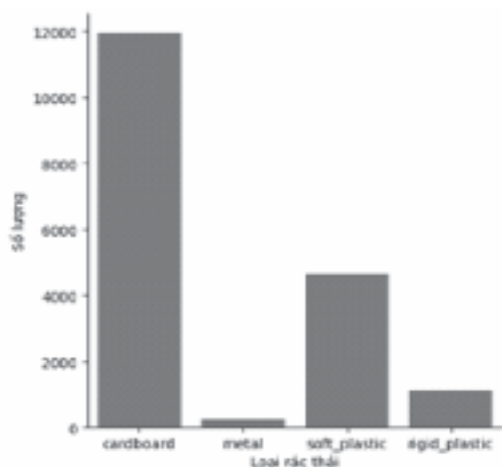
### 2.1. Bộ dữ liệu

Đối với rác thải nhựa, tùy vào mục tiêu ứng dụng của nghiên cứu, cách thu thập và xây dựng bộ dữ liệu cũng khác nhau. Trong nghiên cứu này, với mục đích thử nghiệm YOLOv8 cho dây chuyền thực tế trong các nhà máy nhựa tái chế, bộ dữ liệu được sử dụng cần phải đáp ứng được tiêu chí sát với thực tế, có liên quan đến dây chuyền tự động và có phân bố rác thải trên băng tải một cách ngẫu nhiên với số lượng rác thải trong mỗi khung hình ở số lượng lớn. Để có thể thu thập được dữ liệu có yêu cầu khó khăn như vậy, cần phải trực tiếp vào các nhà máy nhựa tái chế tại Việt Nam để ghi lại và cần một kinh phí, thời gian và công sức rất lớn để gán nhãn cho dữ liệu thu thập được. Với điều kiện hạn chế về mặt kinh phí và nhân lực, nhóm nghiên cứu quyết định sử dụng một bộ dữ liệu đã được công bố trước đây có liên quan đến rác thải nhựa và phù hợp với các yêu cầu đã được đề cập, bộ dữ liệu rác thải nhựa công nghiệp đầu tiên trên thế giới ZeroWaste [8]. Bộ dữ liệu được thu thập trong một nhà máy tại Mỹ, với các lớp chính trong bộ dữ liệu là Cardboard, Rigid plastic, Soft plastic và Metal. Ngoài ra, bộ dữ liệu cũng có các bảng thống kê độ chính xác của các mô hình khác đối với bộ dữ liệu này như Mask R-CNN [9], RetinaNet [10] và TridentNet [11]. Với những gì mà bộ dữ liệu ZeroWaste có được, đây chính là giải pháp để đánh giá mô hình YOLOv8 trong tác vụ xử lý

phân loại rác thải nhựa với điều kiện tài chính và thời gian không nhiều.

## 2.2. Trích xuất thông tin từ bộ dữ liệu

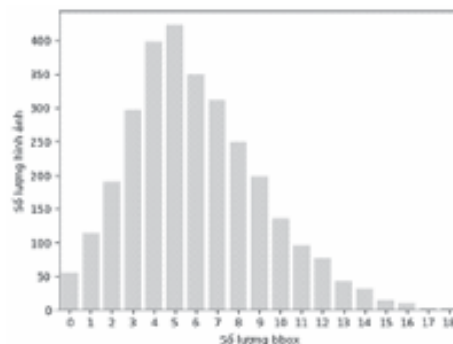
Để hiểu hơn về bộ dữ liệu ZeroWaste nhằm tiến hành tinh chỉnh mô hình YOLOv8 phù hợp, nhóm nghiên cứu đã tiến hành trích xuất các thông tin quan trọng từ bộ dữ liệu ZeroWaste thông qua các phương pháp thống kê. Đầu tiên là thống kê về số lượng các vật thể theo mỗi lớp trong bộ dữ liệu ZeroWaste. Theo đó, Cardboard là phân lớp chiếm số lượng nhiều nhất với gần 12000 vật thể, nhiều hơn gấp đôi so với lớp thứ hai là Soft plastic (khoảng 5000 vật thể). Bên cạnh đó, bộ dữ liệu cũng có rất ít vật thể thuộc lớp Metal và Rigid Plastic. Điều này cho thấy bộ dữ liệu ZeroWaste có dấu hiệu bị mất cân bằng trong dữ liệu (data imbalance) và có thể gây ảnh hưởng xấu đến mô hình nếu không được xử lý đúng cách.



Hình 1. Số lượng vật thể của các loại rác thải trong bộ dữ liệu ZeroWaste

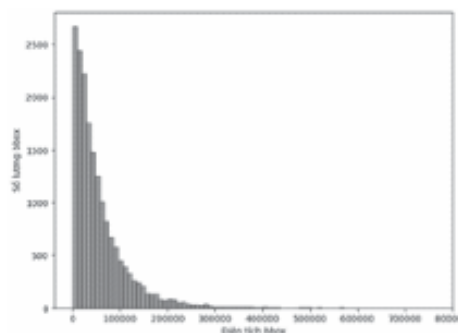
Tiếp theo, nhóm nghiên cứu tiến hành kiểm tra số lượng vật thể được gán nhãn (ground-truth bounding box) trong mỗi hình ảnh của bộ dữ liệu thu được. Kết quả thu được cho thấy, số lượng vật thể được gán nhãn trong

một ảnh phổ biến từ 5-6 vật thể và phần lớn ảnh có chứa ít hơn 12 vật thể được gán nhãn. Ngoài ra, số vật thể được gán nhãn lớn nhất trong một ảnh là 18 vật thể. Từ đây có thể thấy rằng, số lượng vật thể được gán nhãn trong mỗi ảnh là rất lớn, sát với thực tế trên các băng tải trong nhà máy công nghiệp hiện nay.



Hình 2. Số lượng vật thể được gán nhãn trong mỗi ảnh của bộ dữ liệu ZeroWaste

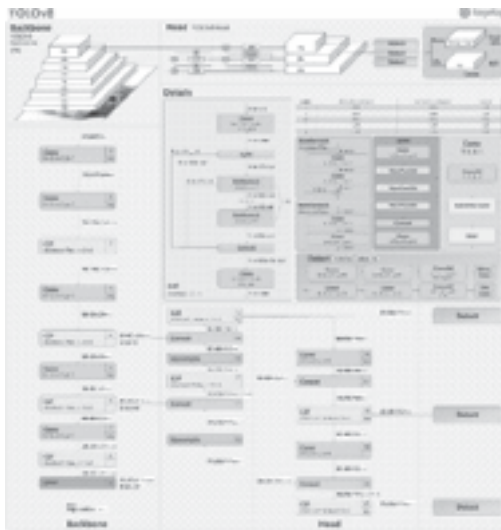
Cuối cùng, nhóm nghiên cứu tiến hành khảo sát diện tích của các vật thể được gán nhãn trong mỗi ảnh. Kết quả thu được cho thấy sự phân bố tương đối rõ ràng, khi vật thể nhỏ ( $< 32^2$  pixels) chiếm số lượng nhiều nhất, tiếp đến là vật thể vừa ( $> 32^2$  và  $< 96^2$  pixels) và cuối cùng là vật thể lớn. Cả ba loại vật thể đều có tổng số lượng khá tương đồng với nhau, cho thấy sự phân bố ngẫu nhiên và không có hình dạng cụ thể xác định của rác thải nhựa trong nhà máy thực tế.



Hình 3. Số lượng của các loại vật thể theo kích thước trong bộ dữ liệu ZeroWaste

Từ các đồ thị thống kê đã được trình bày, có thể thấy rằng bộ dữ liệu ZeroWaste là một bộ dữ liệu có sự mất cân bằng giữa các lớp trong bộ dữ liệu. Ngoài ra, số lượng các vật thể được gán nhãn trong mỗi ảnh cũng là tương đối lớn, kèm theo sự phân bố ngẫu nhiên về mặt kích thước và hình dáng của các vật thể. Nhóm nghiên cứu có đủ cơ sở để kết luận rằng, việc sử dụng bộ dữ liệu ZeroWaste là hợp lý với mục tiêu ứng dụng cho nhà máy thực tế.

### 2.3. Mô hình YOLOv8



Hình 4. Sơ đồ nguyên lý mạng YOLOv8 [12]

Xương sống (backbone) của YOLOv8 là bản sửa đổi của kiến trúc CSPDarknet53 [13] gồm 53 lớp tích chập và sử dụng các kết nối một phần xuyên giai đoạn để tạo điều kiện thuận lợi cho luồng thông tin giữa các lớp khác nhau. Xương sống này chịu trách nhiệm trích xuất các đặc điểm từ hình ảnh đầu vào, trong khi đó ở đầu (head) mạng, quá trình phát hiện đối tượng diễn ra sử dụng các đặc trưng được trích xuất để dự đoán xác suất của các lớp.

Bounding box được xác định tọa độ ở mỗi ô lưới trong ảnh đầu vào. Thêm vào đó, đầu mạng cũng có một cơ chế self-attention,

cho phép mô hình tập trung vào những đặc trưng phù hợp nhất với phát hiện đối tượng.

Khác với các phiên bản YOLO trước đây, YOLOv8 là mô hình anchor-free trực tiếp dự đoán tâm của đối tượng thay vì độ lệch của đối tượng so với anchor box đã biết. Cơ chế này cải thiện đáng kể các phiên bản trước của YOLO, vì nó loại bỏ sự cần thiết của các anchor box, vốn khó có thể bắt những vật có hình dạng hay kích thước bất thường.

YOLOv8 sử dụng phương pháp huấn luyện thích nghi để tối ưu tốc độ học và cân bằng hàm mất mát, cuối cùng dẫn đến một mô hình có hiệu suất tốt hơn. Mô hình YOLOv8 cũng sử dụng kỹ thuật tăng cường dữ liệu nâng cao như MixUP [14], và CutMix [15] để tăng tính mạnh mẽ và tổng quát của mô hình.

### 3. THÍ NGHIỆM

Nhóm nghiên cứu đã tiến hành huấn luyện mô hình YOLOv8 với sự hỗ trợ của Google Colaboratory. Đầu tiên, nhóm tác giả đã chuyển đổi file dữ liệu từ dạng COCO format sang thành dạng YOLO để có thể tiến hành training cho mô hình YOLOv8. Nhóm nghiên cứu huấn luyện mô hình YOLOv8 trong 50 epochs, batch size là 21 (tùy theo GPU được sử dụng của Google Colaboratory) và kích thước ảnh đầu vào được đặt là 640. Ngoài ra, các thông số của mô hình được thể hiện trong Hình 5.

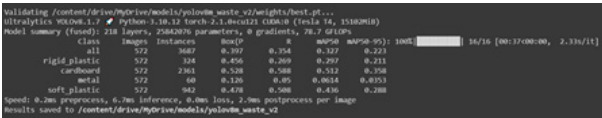
The screenshot shows the output of a YOLOv8 model training process. It displays various parameters and statistics, including the number of parameters, the number of layers, and the training progress. The output is presented in a structured format with columns for different metrics.

Hình 5. Các thông số của mô hình YOLOv8

4. KẾT QUẢ VÀ THẢO LUẬN

Kết quả của quá trình huấn luyện mô hình YOLOv8 qua 50 epochs là khá hứa hẹn. Mô hình đạt được độ chính xác (mAP) là 22.3, độ chính xác tại mAP@50 là 32.7, không cách xa độ chính xác mà các mô hình khác đạt được. Tuy nhiên, YOLOv8 lại cho thấy sự nhẹ hơn trong yêu cầu về phần cứng và xử lý khi thấp hơn gần 3 lần so với Mask R-CNN (79.1 GFLOPs so với 206 GFLOPs). Điều này cho thấy YOLOv8 có tốc độ xử lý nhanh hơn rất nhiều so với Mask R-CNN trong khi độ chính xác không chênh lệch quá nhiều.

Bên cạnh đó, kết quả mAP của model sau khi train xong, có thể thấy là đối với lớp Metal (có ít vật thể nhất) thì tỷ lệ dự đoán được rất rất thấp, kéo theo mAP của tổng thể thấp xuống. Lớp chiếm tỉ lệ nhiều nhất là Cardboard cũng chỉ có mAP@50 là 0.512, nghĩa là model chỉ đoán đúng một nửa.

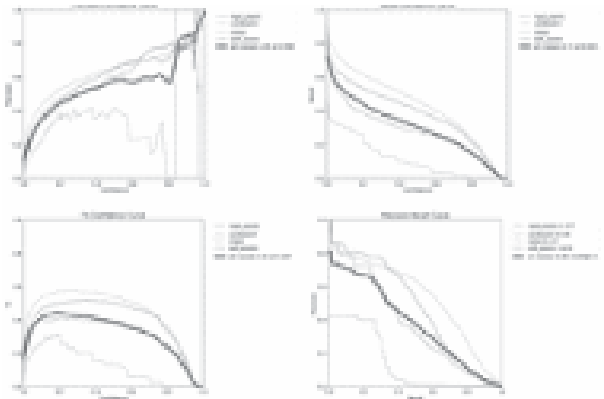


Hình 6. Kết quả huấn luyện mô hình YOLOv8 sau 50 epochs

Bảng 1. So sánh độ chính xác và gánh nặng tính toán giữa YOLOv8 và các mô hình khác

| Model      | Epochs | GFLOPs | Params | AP   | AP50 |
|------------|--------|--------|--------|------|------|
| RetinaNet  | 107    | 206    | 38M    | 21.0 | 33.5 |
| MaskRCNN   | 107    | -      | 44M    | 22.8 | 34.9 |
| TridentNet | 107    | 822    | 34M    | 24.2 | 36.3 |
| YOLOv8     | 50     | 79.1   | 26M    | 22.3 | 32.7 |

Ngoài ra, ở các đồ thị thống kê về Precision, Recall, F1-Score và Precision-Recall của mô hình, YOLOv8 vẫn chưa cho thấy toàn bộ khả năng của mình, khi các đồ thị vẫn còn tương đối xấu và có hiện tượng gãy tại threshold nhất định và điểm tương đối thấp.



Hình 7. Đồ thị thống kê cho mô hình YOLOv8

5. KẾT LUẬN

Trong nghiên cứu này, nhóm nghiên cứu giới thiệu phương pháp sử dụng YOLOv8 để phân loại rác thải nhựa trên dây chuyền thực tế tại nhà máy. Kết quả chỉ ra rằng, YOLOv8 phù hợp với các hệ thống có khả năng tính toán và tốc độ xử lý thấp vì khối lượng tính toán của YOLOv8 thấp hơn các mô hình thông thường khác ít nhất 3 lần. Ngoài ra, YOLOv8 cũng đảm bảo độ chính xác ở mức chấp nhận được và hoàn toàn thích hợp để ứng dụng vào các tác vụ phân loại khác trong công nghiệp.

Lời cảm ơn:

Nghiên cứu này được tài trợ bởi Trường Đại học Bách khoa, Đại học Quốc gia Thành



phố Hồ Chí Minh trong khuôn khổ đề tài mã số SVKSTN-2023-CK-15. Chúng tôi xin cảm ơn Trường Đại học Bách khoa, ĐHQG-HCM đã hỗ trợ cho nghiên cứu này. ❖

Ngày nhận bài: **18/02/2024**

Ngày phản biện: **27/3/2024**

### Tài liệu tham khảo:

- [1]. Prata, J. C., da Costa, J. P., Lopes, I., Duarte, A. C., & Rocha-Santos, T., "Environmental exposure to microplastics: An overview on possible human health effects". *Science of the total environment*, vol. 702, p. 134455, 2020.
- [2]. Agyeman, S., Obeng-Ahenkora, N. K., Assiamah, S., & Twumasi, G., "Exploiting recycled plastic waste as an alternative binder for paving blocks production". *Case Studies in Construction Materials*, vol. 11, p. e00246, 2019.
- [3]. Adeniran, A. A., & Shakantu, W., "The health and environmental impact of plastic waste disposal in South African Townships: A review". *International Journal of Environmental Research and Public Health*, vol. 19, no. 2, p. 779, 2022.
- [4]. Duan, Q., & Li, J., "Classification of common household plastic wastes combining multiple methods based on near-infrared spectroscopy". *Acs Es&T Engineering*, vol. 1, no. 7, pp. 1065-1073, 2021.
- [5]. Nhung, Đ. T., My, N. T. D., Mạnh, P. V., Đông, P. V., Thành, B. Q., Tuấn, N. V., & Hải, P. M., "Nghiên cứu mô hình phát hiện rác thải nhựa ven biển sử dụng ảnh máy bay không người lái và mạng nơ-ron tích chập sâu", *Tạp chí Khoa học Đo đạc và Bản đồ*, no. 49, pp. 21-29, 2021.
- [6]. Anh, H. N. T., & Trí, T. T., "WSCA: Nhận dạng rác trên mặt nước ứng dụng mạng nơ-ron học sâu", *Tạp chí Khoa học và Công nghệ, Đại học Đà Nẵng*, pp. 47-52, 2023.
- [7]. Bobulski, J., & Kubanek, M., "Deep learning for plastic waste classification system". *Applied Computational Intelligence and Soft Computing*, vol. 2021, pp. 1-7, 2021.
- [8]. Bashkistrova, D., Abdelfattah, M., Zhu, Z., Akl, J., Alladkani, F., Hu, P., ... & Saenko, K., "ZeroWaste dataset: towards deformable object segmentation in cluttered scenes". In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.
- [9]. He, K., Gkioxari, G., Dollár, P., & Girshick, R., "Mask R-CNN". In *Proceedings of the IEEE International conference on computer vision*, 2017.
- [10]. Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P., "Focal loss for dense object detection". In *Proceedings of the IEEE International conference on computer vision*, 2017.
- [11]. Li, Y., Chen, Y., Wang, N., & Zhang, Z., "Scale-aware trident networks for object detection". In *Proceedings of the IEEE/CVF International conference on computer vision*, 2019.
- [12]. Jacob Solawetz, Francesco, "Roboflow Blog". Roboflow, 11 January 2023. [Online]. Available: <https://blog.roboflow.com/whats-new-in-yolov8/>. [Accessed 5 March 2024].
- [13]. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M., "YOLOv4: Optimal speed and accuracy of object detection." *arXiv preprint arXiv:2004.10934*, 2020.
- [14]. Zhang, H., Cisse, M., Dauphin, Y. N., & Lopez-Paz, D., "Mixup: Beyond empirical risk minimization". *arXiv preprint arXiv:1710.09412*, 2017.
- [15]. Yun, S., Han, D., Oh, S. J., Chun, S., Choe, J., & Yoo, Y., "Cutmix: Regularization strategy to train strong classifiers with localizable features". In *Proceedings of the IEEE/CVF International conference on computer vision*, 2019.