

# WSCA: NHẬN DẠNG RÁC TRÊN MẶT NƯỚC ỨNG DỤNG MẠNG NƠON HỌC SÂU

## WSCA: A FLOATING TRASH COLLECTOR SYSTEM APPLIED DEEP NEURAL NETWORK

Huỳnh Ngọc Thái Anh\*, Trang Thanh Trí

*Trường Công nghệ Thông tin và Truyền thông, Trường Đại học Cần Thơ, Việt Nam<sup>1</sup>*

\*Tác giả liên hệ / Corresponding author: hntanh@ctu.edu.vn

(Nhận bài / Received: 20/6/2023; Sửa bài / Revised: 18/8/2023; Chấp nhận đăng / Accepted: 08/9/2023)

**Tóm tắt** - Bài báo giới thiệu giải pháp máy thu rác trên mặt nước với chức năng nhận dạng rác tự động bằng việc huấn luyện mô hình học sâu với kiến trúc mạng YOLOv7. Đầu tiên, phương pháp Thinking Design kết hợp mô phỏng 3D được sử dụng trong quá trình thiết kế và đánh giá hiệu quả của máy thu rác trên mặt nước có tên WSCA (Water Surface Cleaning Autobot). Nhóm tác giả đề xuất phương pháp tăng cường dữ liệu để tạo thành tập FloW+ có thêm 800 ảnh so với tập data FloW (gồm 2000 ảnh với 5.271 chất thải nhựa trôi nổi). Cuối cùng là xây dựng và huấn luyện mô hình mạng học sâu để nhận dạng rác trên mặt nước với tập dữ liệu tập FloW+. Độ chính xác trên tập kiểm thử là Precision đạt 80,5%, Recall đạt 76,6%, mAP@0.5 đạt 78,8%, mAP@0.5..95 đạt 35,6% với FPS đạt trung bình là 17,6. Phương pháp này đem đến khả năng ứng dụng cao cho việc xây dựng một thiết bị thu rác trôi nổi tự động cũng như áp dụng mở rộng ở các quy mô lớn hơn.

**Từ khóa** - WSCA; mạng học sâu; nhận dạng rác trôi nổi; YOLOv7

### 1. Đặt vấn đề

Tình trạng ô nhiễm rác nhựa trong đại dương và trong các con sông đang rất đáng báo động, có hơn 5 nghìn tỷ mảnh rác nhựa đang trôi nổi trong các đại dương và 250.000 tấn trôi nổi trên mặt biển [1]. Các chất thải, đặc biệt là nhựa, từ các con sông tràn vào các đại dương đang đe dọa nghiêm trọng đến các sinh vật biển và gây ô nhiễm nghiêm trọng cho toàn bộ hệ sinh thái [24] (như minh họa trong Hình 1 (a)). Tuy nhiên, theo [2], chỉ có 15% chất thải biển trôi nổi trên bề mặt đại dương, trong khi 15% nằm trong cột nước và 70% chìm dưới đáy biển. Hiện tại đã có nhiều dự án được triển khai để thu gom rác thải trên đại dương nhưng hiệu quả hơn hết là ngăn chặn các chất thải trước khi xả vào đại dương. Các nguồn rác thải từ đất liền được chỉ ra là đầu vào chủ yếu [3]. Chính vì vậy, làm sạch rác thải trôi nổi ở các nguồn nước trên đất liền như kênh, sông, hồ và các vịnh là cách tiếp cận hiệu quả để giảm thiểu rác thải trên biển cũng như việc ô nhiễm nguồn nước.

Các phương pháp truyền thống cho vấn đề thu gom rác thải trôi nổi trong các vùng nước nội địa chủ yếu là dựa vào nhân công thu gom trực tiếp. Tuy nhiên các phương pháp này có hiệu quả thấp và tốn kém trước những khu vực có lượng lớn rác thải trôi nổi [4]. Hơn nữa, các phương pháp này có các rủi ro tiềm tàng như khí độc và các chất thải độc hại từ các hệ thống thoát nước [5], tai nạn đuối nước cùng nhiều hạn chế khác về hiệu suất làm việc của người lao động. Chính vì vậy, các phương pháp làm sạch rác thải trôi nổi trên mặt nước bằng phương pháp tự động đang thể hiện

**Abstract** - This article presents an automatic water surface garbage collector using a deep learning model with the YOLOv7 network architecture. First, the Thinking Design method was combined with 3D simulation in the design and evaluation process of the WSCA (Water Surface Cleaning Autobot). The authors then propose a data augmentation method to create the FloW+ dataset, which contains an additional 800 images compared to the FloW dataset (2000 images with 5271 floating plastic waste). Finally, the deep learning model is trained on the FloW and FloW+ datasets to identify trash on water surfaces. The accuracy on the test set in the FloW+ set is a precision of 80.5%, recall of 76.6%, mAP@0.5 of 78.8%, mAP@0.5..95 of 35.6% with an average FPS of 17.6. This method has the potential application for the construction of an autonomous floating trash-collecting robot, as well as scale it up to larger scales.

**Key words** - WSCA; Deep learning network; Floating trash detection; Yolo v7

tính hiệu quả và cần thiết hơn bao giờ hết. Hiện tại, các thiết bị không người lái - Unmanned surface vehicles (USV) đã dần được sử dụng và chứng tỏ tiềm năng to lớn trong việc thu gom các chất thải trôi nổi [4]. Ưu điểm của các thiết bị không người lái là có thể đi vào các khu vực khó tiếp cận và hoạt động độc lập. Hiệu quả của các thiết bị không người lái được đánh giá bao gồm hệ thống lái tự động hoặc bán tự động, khả năng thu gom và tách rác trôi nổi khỏi mặt nước cùng khả năng chứa và vận chuyển các rác thải này [6].



a. Rác thải trên đại dương



b. Rác thải trên sông rạch



c. Thu rác thủ công



d. Thu rác bằng máy

**Hình 1.** Một số loại ô nhiễm rác nhựa trôi nổi và một số phương pháp thu rác phổ biến hiện nay

<sup>1</sup> The College of Information and Communication Technology, Can Tho University, Vietnam (Huỳnh Ngọc Thái Anh, Trang Thanh Trí)

So với một số thiết bị cố định, một USV tốt bao gồm tất cả các thiết bị có thể dẫn hướng (bộ điều khiển hay tự hành), các kết cấu kỹ thuật có thể loại bỏ, chứa và vận chuyển các rác trôi trên bề mặt nước [7]. Tuy nhiên, ở các vùng nước nội địa thì các loại rác như chai lọ, lon nước, ly nhựa ... có kích thước khá nhỏ và chỉ nổi một phần trên mặt nước. Điều này dẫn đến diện tích hình ảnh hiển thị trong camera cũng sẽ nhỏ, dẫn đến khó khăn trong việc nhận dạng chính xác. Đối với việc phát hiện vật thể dựa trên giải thuật học sâu thì các đối tượng có độ phân giải quá nhỏ sẽ kém chính xác hơn so với các đối tượng có kích thước vừa và lớn. Nguyên nhân là do các đối tượng này thiếu những thông tin đặc trưng cấp cao (high-level features) và chúng cũng không tách biệt rõ với nền mặt nước khi nhận dạng [8-9].

Một trong những nguyên nhân khác làm hạn chế khả năng nhận dạng các loại rác trôi nổi dựa trên camera quan trắc là vấn đề phản xạ và chói sáng của mặt nước, sự phản xạ này dẫn đến các vùng ảnh bị chói sáng làm ảnh hưởng việc nhận dạng các loại rác trôi nổi xuất hiện trong vùng ảnh này. Một nguyên nhân quan trọng nữa là các sự phức tạp của môi trường nước mặt tự nhiên ở các ao hồ và sông ngòi ở Việt Nam. Các loại rác trôi nổi sẽ dễ bị lẫn với bùn, rêu và các chất bẩn khác nếu đã trôi nổi lâu ngày, điều này trực tiếp làm giảm các đặc trưng khi nhận dạng. Các loại rác này cũng rất dễ mắc vào các loại thực vật thủy sinh như bèo tây, lục bình... trôi trên sông hoặc mọc gần bờ, lúc này phần lớn hình ảnh của các loại rác sẽ bị che đi làm giảm khả năng phát hiện. Thực tế trên khẳng định rằng việc phát hiện các loại chất thải trôi nổi dựa trên camera quan trắc vẫn còn là một bài toán chưa được giải quyết tối ưu ngay ở thời điểm hiện tại [10].

Để cải thiện khả năng nhận dạng thì một trong những giải pháp tối ưu là kết hợp camera quan trắc cùng radar thu phát sóng ở tần số 77Ghz để phát hiện các vật thể xung quanh [11]. Giải pháp này kết hợp những ưu điểm của radar là không bị ảnh hưởng bởi thời tiết, ánh sáng, khả năng phát hiện đối tượng ở phạm vi xa và cả tốc độ chuyển động của vật thể [12]. Tuy có nhiều ưu điểm nhưng vì giá thành radar phát hiện chướng ngại vật còn khá cao chính là rào cản làm cho hướng tiếp cận này chưa phổ biến khi triển khai ở Việt Nam so với việc chỉ sử dụng camera quan trắc. Hướng tiếp cận nhận dạng bằng camera chính là mục tiêu của nghiên cứu này thông qua các nội dung cụ thể sau:

- Thiết kế và đánh giá một mô hình USV là một máy thu rác trên mặt nước bằng phương pháp Thinking design.
- Nghiên cứu tập dữ liệu FloW cùng các giải pháp tăng cường dữ liệu bằng phương pháp ghép ảnh ngẫu nhiên.
- Xây dựng mô hình nhận dạng rác thải nhựa trôi nổi trên mặt nước sử dụng kiến trúc mạng YOLO v7.
- Đánh giá mô hình và đề xuất các hướng triển khai.

## 2. Kết quả nghiên cứu và khảo sát

### 2.1. Phương pháp Thinking Design

Thinking design hay Design thinking còn được gọi là phương pháp “*Tư duy thiết kế*” là một phương pháp triển khai các thiết kế, sản phẩm sáng tạo bao gồm các quá trình nhận thức, xây dựng chiến lược và kiểm tra trong thực tiễn. Các khái niệm và các bước trong quy trình Thinking design

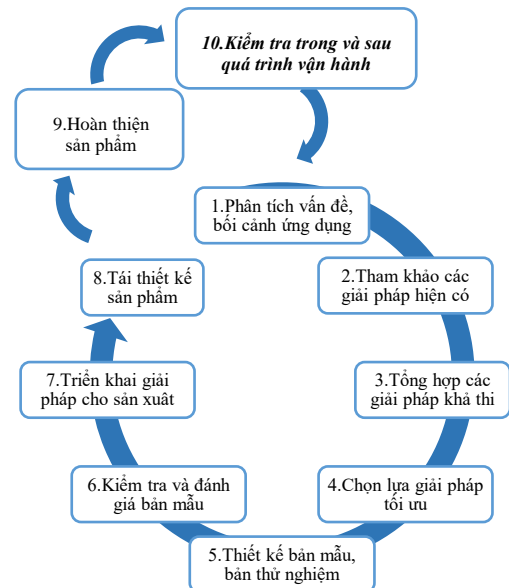
được xây dựng thông qua các nghiên cứu trong nhiều lĩnh vực khác nhau nhưng tập trung vào nhận thức của người thiết kế và hoạt động thiết kế trong bối cảnh thí nghiệm và cả trong thực tiễn [13], [14]. Tư duy thiết kế được gắn liền với quá trình sáng tạo và đổi mới các sản phẩm, dịch vụ [15], [16] cũng như thiết kế các ứng dụng khoa học kỹ thuật như robot hay các thiết bị máy móc sáng tạo [17].

Thinking là một quy trình gồm các bước sau: phân tích bối cảnh; tìm và tạo khung vấn đề; tạo ý tưởng và giải pháp; tư duy sáng tạo; phác thảo; mô hình hóa và tạo mẫu; thử nghiệm; và đánh giá [18]. Khi triển khai một bước trong quy trình gặp phải vấn đề không thể giải quyết thì sẽ quay trở về bước trước đó để thực hiện lại hoặc chọn theo một hướng giải quyết khác. Điều này được lặp lại cho đến khi vấn đề hiện tại ở một bước được giải quyết. Quy tắc quay ngược này chính là áp dụng phương pháp lý luận hồi tố (abductive reasoning).

Các đặc điểm nổi bật của tư duy thiết kế bao gồm:

- Giải quyết các vấn đề mới chưa được xác định.
- Áp dụng các chiến lược tập trung vào giải pháp.
- Sử dụng lý luận hồi tố hiệu quả.
- Ứng dụng phương tiện mô hình hóa không lời, đồ họa không gian, phác thảo và tạo mẫu [19]

Quy trình thiết kế một sản phẩm với phương pháp Thinking design được đề xuất bởi Hynes et al. [20] mô tả theo Hình 2. Trong quá trình triển khai nghiên cứu và chế tạo, nhóm tác giả có đề xuất thêm bước thứ 10 là Kiểm tra trong và sau quá trình vận hành. Thông qua bước cải tiến này, thiết kế và hiệu năng của mô hình sẽ được đánh giá lại để là cơ sở cải tiến cho các thiết kế mới về sau.



Hình 2. Quy trình thiết kế theo phương pháp Thinking design có cải tiến

### 2.2. Thiết kế máy thu rác trên mặt nước WSCA

Máy vớt rác WSCA là tên gọi viết tắt của Water Surface Cleaning Autobot – thiết bị dọn dẹp và thu gom rác tự động trên mặt nước. WSCA được nhóm tác giả thiết kế vào cuối năm 2021 với mục tiêu là hoạt động trên mặt nước sông, hồ tĩnh và mặt biển sóng nhẹ, máy có chức năng chính là thu rác nhựa trôi nổi trên bề mặt nước [10]. Ở giai đoạn đầu

của quá trình thiết kế, thiết bị được tích hợp bộ điều khiển không dây từ xa thông qua ứng dụng trên điện thoại thông minh. Quá trình thiết kế mô hình này được sự tham vấn của các chuyên gia từ UNESCO Việt Nam, thiết bị đang được vận hành thử nghiệm tại sông Hoài thuộc địa bàn phố cổ Hội An, tỉnh Quảng Nam. Quá trình thiết kế của WSCA được triển khai theo quy trình Thinking design được trình bày ở Hình 2. Mô hình bản mẫu trong bước 5 được gia công theo tỉ lệ 1:1 để đánh giá chính xác nhất về độ nổi; khả năng ổn định trước sóng; tốc độ di chuyển thuận và ngược chiều nước chảy; độ bền của cấu trúc và vật liệu ...



**Hình 3.** Bản mẫu 3D và bản thử nghiệm của WSCA vận hành thử nghiệm tại chợ nổi Cái Răng, thành phố Cần Thơ

Hiện tại, hệ thống nhận dạng rác tự động sử dụng kiến trúc YOLO V7 vừa được nhóm tác giả xây dựng thành công sẽ được tích hợp vào mô hình ở giai đoạn 2 của dự án nghiên cứu. Mục tiêu của giai đoạn này là phát triển WSCA trở thành một thiết bị USV cơ bản từ bán tự động trở thành tự động hoàn toàn. WSCA sẽ được vận hành trong các vùng nước chảy chậm hoặc tĩnh trong đất liền như sông và các hồ nước để liên tục thu mới các dữ liệu quan trắc bao gồm ảnh và video. Quy cách triển khai thiết kế mô hình máy thu rác WSCA được trình bày trong Bảng 1 và mô hình thiết kế, ảnh vận hành được giới thiệu trong Hình 4 và Hình 5.

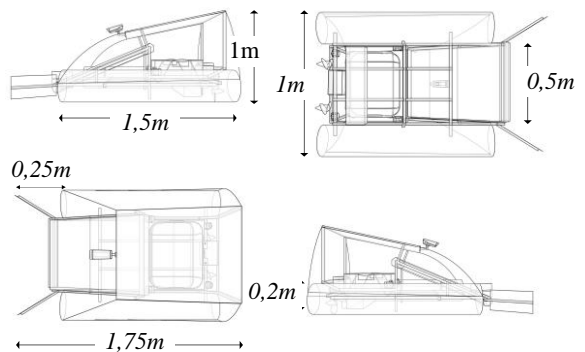
**Bảng 1.** Triển khai Thinking design trong chế tạo WSCA

STT	Quy trình Design Thinking	Quy cách triển khai
1	Phân tích vấn đề và bối cảnh ứng dụng	Tìm hiểu yêu cầu thiết kế, mục tiêu thiết kế, nơi vận hành, thời điểm vận hành, đối tượng sử dụng, tính chất vùng mặt nước vận hành v.v...
2	Tìm kiếm và tham khảo các giải pháp hiện có	Tìm kiếm trên các bài báo khoa học, tạp chí khoa học kỹ thuật cùng nhiều nguồn thông tin trong và ngoài nước
3	Tổng hợp các giải pháp khả thi	Liệt kê các giải pháp thu rác trên mặt nước
4	Đánh giá và chọn lựa giải pháp tối ưu để triển khai	Lập bản so sánh và đánh giá các giải pháp, so sánh đối chiếu với điều kiện thực tế tại Việt Nam và nơi áp dụng
5	Thiết kế bản mẫu, bản thử nghiệm	Thiết kế trên mô hình 3D để đánh giá cấu trúc, kích thước cùng mô phỏng vận hành. Thiết kế mô hình thật với cùng kích thước và trọng lượng để đánh giá vận hành thực tế
6	Kiểm tra và đánh giá bản mẫu	Đưa mô hình vận hành thử nghiệm với điều kiện thực tế
7	Triển khai giải pháp cho bên sản xuất	Chuyển giao bản thiết kế cho bên đơn vị gia công

8	Tái thiết kế sản phẩm	Tái thiết kế sản phẩm với bản vẽ chính thức, kiểm tra chất lượng gia công
9	Hoàn thiện sản phẩm	Xây dựng bản hướng dẫn vận hành và hướng dẫn kỹ thuật
10	Kiểm tra trong và sau quá trình vận hành	Xây dựng bản ghi thông tin các sự cố kỹ thuật và lỗi vận hành cùng nguyên nhân tạo lỗi



**Hình 4.** Mô hình thiết kế 3D và thực tế vận hành của WSCA



**Hình 2.** Kích thước kỹ thuật của WSCA

### 2.3. Tập dữ liệu Flow

Dữ liệu của tập Flow được tác giả Yuwei Cheng cùng các cộng sự tại Đại học Thanh Hoa, Đại học Bách khoa Tây Bắc, Trung Quốc phối hợp cùng Viện kỹ thuật AI Quebec, Canada xây dựng bằng cách thu thập các dữ liệu rác thải trôi nổi khi vận hành thiết bị USV mang tên SMURF20 do nhóm chế tạo [10]. Môi trường vận hành là các vùng nước tĩnh hoặc chảy nhẹ, loại rác nhận dạng chính là các loại rác thải nhựa. Một camera quan trắc có dải nhạy sáng rộng (chip AR0230 và độ phân giải cảm biến là 1280x720) được

đặt ở phía trước thiết bị sẽ chụp và ghi hình lại toàn bộ quá trình vận hành. Điều kiện vận hành và thiết kế của vị trí camera của SMURF20 khá tương đồng với WSCA, điều này đảm bảo cho hiệu quả áp dụng của tập dữ liệu FloW với WSCA.

Sau quá trình thu thập và làm sạch dữ liệu, thì tập dữ liệu FloW-Img đã được xây dựng có 2.000 hình ảnh với 5.271 chất thải nhựa trôi nổi được gắn nhãn. Tập dữ liệu này được chia theo tỉ lệ 6:4 cho tập huấn luyện (tương ứng 1.200 ảnh) và tập kiểm tra (800 ảnh) [10].

Các loại rác trôi nổi được phát hiện có diện tích hiển thị nằm trong 3 mức, nhỏ hơn 32x32 pixel, trong khoảng 32x32 pixel đến 96x96 pixel và lớn hơn 96x96 pixel.

#### 2.4. Giới thiệu kiến trúc YOLO v7

YOLO (You Only Look Once) là một mô hình mạng nơ-ron tích chập (CNN) được phát triển cho mục tiêu phát hiện, nhận dạng và phân loại đối tượng. YOLO được tạo ra từ việc kết hợp giữa các lớp tích chập có nhiệm vụ trích xuất các đặc trưng ảnh (backbone) và lớp fully-connected (head output) để tính toán xác định tọa độ của đối tượng, nhận dạng đối tượng và dự đoán xác suất có tồn tại đối tượng hay không.

Kiến trúc YOLO v7 được 2 tác giả là WongKinYiu và AlexeyAB giới thiệu trong bài báo YOLO v7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors vào ngày 6 tháng 7 năm 2022 [7]. Đây là phiên bản mới nhất trong tất cả các phiên bản YOLO từng ra mắt tính đến thời điểm hiện tại. Các ưu điểm nổi bật nhất của phiên bản này là có tốc độ suy luận tính toán nhanh hơn và độ chính xác cao hơn so với các phiên bản trước theo công bố của tác giả.

Bước đầu trong quy trình nhận dạng là các khung ảnh sẽ được trích xuất đặc trưng thông qua backbone với các mức tỷ lệ giảm dần độ phân giải. Các đặc trưng sau khi được trích xuất sẽ được tổng hợp ở phần kiến trúc FPN (Feature Pyramid Networks). Tiếp theo là bộ lọc đặc trưng áp dụng nguyên lý kim tự tháp đảm bảo cân bằng giữa độ chính xác và tốc độ xử lý. Các đặc trưng kết hợp này sẽ chuyển đến phần Head của kiến trúc, là nơi YOLO dự đoán các bounding boxes cùng với vị trí, nhận dạng và phân lớp. Giai đoạn cuối là xử lý hậu kỳ qua NMS (Non Maximum Suppression) để đưa ra dự đoán cuối cùng.

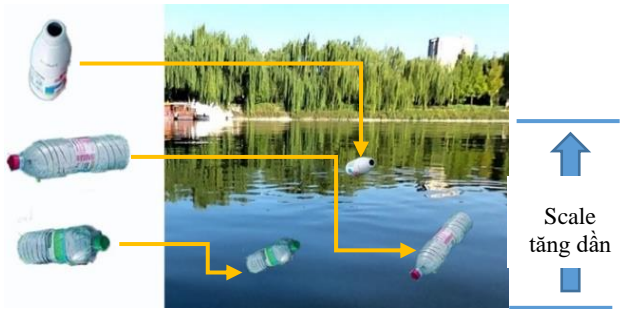
#### 2.5. Giải pháp tăng cường dữ liệu

Để mở rộng kích thước của tập huấn luyện thì giải pháp của nhóm tác giả là sử dụng 150 ảnh rác thải nhựa đã được tách nền và ghép ngẫu nhiên vào các vị trí trên mặt nước ở các ảnh trong tập FloW\_img. Việc ghép ngẫu nhiên này bao gồm 3 bước xử lý:

1. Các ảnh tách nền này sẽ được điều chỉnh tỷ lệ ngẫu nhiên trong khoảng 16x16 pixel đến 124x124 pixel.
2. Áp dụng ngẫu nhiên các giải thuật xoay, lật cho các ảnh tách nền này.
3. Ghép các ảnh vừa xử lý vào các vị trí ngẫu nhiên trên ảnh mặt nước với độ scale tăng dần theo khoảng cách từ gần đến xa.

Ảnh tăng cường được miêu tả ở Hình 6 và kết quả áp dụng phương pháp tăng cường dữ liệu tập dữ liệu gốc 2000

ảnh đã tăng thành 2800 ảnh (gọi là tập FloW+). Phương pháp tăng cường dữ liệu này được áp dụng trước giai đoạn huấn luyện. Nhóm tác giả cũng áp dụng các phương pháp tăng cường dữ liệu trong giai đoạn huấn luyện bao gồm: tăng cường ngẫu nhiên cường độ màu HSV, chỉnh sửa kích thước, kéo dẫn ảnh, lật ảnh, ghép ảnh.



Hình 6. Giải pháp tăng cường dữ liệu

#### 2.6. Kết quả thực nghiệm

##### 2.6.1. Cài đặt chương trình

Nhóm tác giả đã tiến hành cài đặt hệ thống chương trình nhận dạng rác trôi nổi tự động bằng ngôn ngữ lập trình Python với kiến trúc Yolo v7. Tất cả kết quả nghiên cứu đều được thực nghiệm trên cùng hệ thống máy tính có cấu hình như sau: CPU: AMD Ryzen 9 5900X; GPU: 2x Nvidia GTX 1080Ti, 11GB GDDR5; RAM: 64GB DDR4; Operating system: Windows 10; Python 3.8.8.

##### 2.6.2. Mô tả dữ liệu huấn luyện

Tập dữ liệu là FloW+ có 2.000 ảnh của tập FloW được chia theo tỉ lệ 6:4. Tập huấn luyện có 1200 bổ sung 800 ảnh tăng cường đạt tổng 2000 ảnh, tập kiểm thử có 800 ảnh.

Tỷ lệ giao nhau (IoU)

$$IoU = \frac{area(B_p \cap B_{gt})}{area(B_p \cup B_{gt})}$$

$B_p$  là vùng ảnh mô hình dự đoán đối tượng.

$B_{gt}$  là vùng ảnh chứa đối tượng (theo nhãn).

##### 2.6.3. Độ đo được sử dụng trong mô hình

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

Hình 7. Minh họa IoU

Độ chính xác (Precision và Recall)

$$Precision = \frac{TP}{TP + FP} \quad Recall = \frac{TP}{TP + FN}$$

True positive (TP): Nhận dạng đúng lớp đối tượng.

False positive (FP): Không nhận dạng đúng đối tượng.

False Negative (FN): Nhận dạng sai với đối tượng đúng.

Độ chính xác trung bình (AP)

$$p_{interp}(r) = \max_{r' \geq r} p(r')$$

$$AP = \sum_{i=1}^{n-1} (r_{i+1} - r_i) p_{interp}(r_i + 1)$$

( $r_1, r_2, r_3, \dots, r_n$ ) là các cấp độ recall được xếp theo thứ tự tăng dần (i.e., 0.0, 0.1, 0.2, ..., 1.0).

#### Độ đo COCO

mAP IoU=.50..05..95: Trung bình mAP trên các ngưỡng IoU khác nhau (i.e., 0.50, 0.55, 0.60, ..., 0.95).

#### 2.6.4. Kết quả mô hình

Kết quả thu được từ mô hình huấn luyện được trình bày trong Bảng 2 và kết quả so sánh hiệu năng với các phương pháp khác được trình bày trong Bảng 3

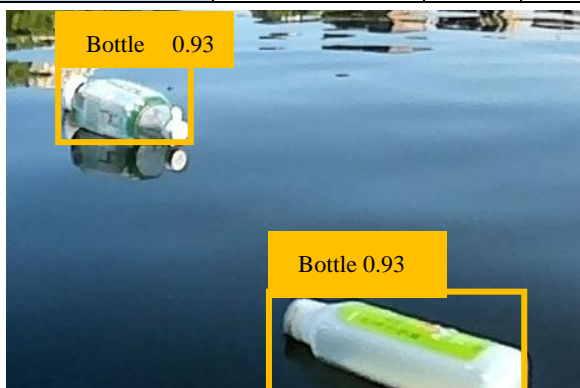
**Bảng 2.** Kết quả độ chính xác của mô hình

Precision	Recall	mAP@.5:	mAP@.5..95	FPS:
<b>0,805</b>	<b>0,766</b>	<b>0,788</b>	<b>0,356</b>	<b>17,6</b>

Trong các mô hình học sâu như DSSD, RetinaNet, YOLO v3, Faster R-CNN, FPN và Cascade R-CNN có thể thấy, mô hình DSSD đạt được tốc độ dự đoán nhanh nhất là 28,6fps, trong khi đó thấp nhất là Cascade R-CNN 3,9fps. Mô hình Cascade R-CNN đạt được độ chính xác cao nhất là 0,434 trên độ đo mAP nhưng không đáp ứng được yêu cầu tính toán theo thời gian thực. Phương pháp nhóm tác giả đề xuất đã cải thiện được độ chính xác mAP@.5..95 là 0,334 trên tập FloW và 0,356 trên tập FloW+. Kết quả huấn luyện trên tập dữ liệu tăng cường FloW+ có độ chính xác mAP@.5 là 0,788 cao hơn mức 0,764 trên tập FloW. Kết quả này tốt hơn so với các phương pháp FPN, RetinaNet, YOLO v3, Faster R-CNN và DSSD, đồng thời đạt được tốc độ tính toán là 17.6fps nhanh hơn so với Cascade R-CNN. Mô hình nhận dạng này phù hợp với hệ thống camera quan trắc thời gian thực được tích hợp với máy thu rác WSCA khi vận hành. Kết quả nhận dạng được giới thiệu ở Bảng 3 và Hình 8.

**Bảng 3.** So sánh độ chính xác và tốc độ xử lý

	mmAP	FPS	
<b>DSSD [22]</b>	0,275	<b>28,6</b>	
<b>RetinaNet [23]</b>	0,249	<b>7,6</b>	
<b>YOLO-v3 [24]</b>	0,128	<b>23,2</b>	
<b>Faster R-CNN [25]</b>	0,184	<b>9,3</b>	
<b>FPN [26]</b>	0,334	<b>7,4</b>	
<b>Cascade</b>	<b>R-CNN [27]</b>	0,434	<b>3,9</b>
<b>YOLO-v7 (FloW)</b>	mAP@.5	0,764	<b>17,6</b>
	mAP@.5..95	0,334	
<b>YOLO-v7 (FloW+)</b>	mAP@.5	<b>0,788</b>	<b>17,6</b>
	mAP@.5..95	<b>0,356</b>	



**Hình 8.** Kết quả nhận dạng thực nghiệm

### 3. Kết luận

Trong nghiên cứu này, nhóm tác giả đã trình bày một quy trình thiết kế kỹ thuật cho máy thu rác trên mặt nước WSCA ứng dụng phương pháp Thinking design. Đánh giá tính hiệu quả của WSCA là nhiều kiểm tra thực nghiệm trong quá trình vận hành. Các kết quả cho thấy, việc ứng dụng phương pháp Thinking design giúp đẩy nhanh được quá trình phân tích và thiết kế mô hình. Bên cạnh đó, việc thiết kế và giả lập mô hình trong không gian 3 chiều giúp đánh giá các thông số thiết kế trước khi gia công giúp giảm chi phí nguyên vật liệu cho khâu thử nghiệm hoặc do lỗi thiết kế. Lỗi vận hành sẽ được phân tích để làm cơ sở phát triển các phiên bản tiếp theo của WSCA.

Để giải quyết vấn đề nhận dạng rác thải trôi nổi, nhóm tác giả đã nghiên cứu mô hình học sâu ứng dụng kiến trúc YOLO v7 mới nhất hiện nay. Tập dữ liệu được sử dụng là FloW có chứa 2000 ảnh được chụp trong môi trường vận hành tương tự với WSCA, với 5.271 chất thải nhựa trôi nổi được gắn nhãn. Kết hợp với phương pháp tăng cường dữ liệu, 800 ảnh tăng cường đã được thêm vào tập huấn luyện. Dựa trên kết quả thực nghiệm, đánh giá YOLO v7 đem lại độ chính xác cao và có tốc độ nhận dạng nhanh hơn so với các phiên bản trước đó. Độ chính xác Precision đạt 80,5%, Recall đạt 76,6%, mAP@0.5 đạt 78,8%, mAP@0.5..95 đạt 35,6% với FPS đạt trung bình là 17,6

Nghiên cứu này mở ra nhiều hướng phát triển đầy hứa hẹn. Một trong số đó là hướng đến việc thu thập các hình ảnh rác nhựa trong quá trình vận hành của WSCA để tạo ra một tập dữ liệu mới, đồng thời giúp nâng cao khả năng nhận dạng và phân loại rác nhựa trên mặt nước. Bên cạnh đó, nghiên cứu đã đề xuất được phương pháp phù hợp cho bài toán thu thập rác trôi nổi trên mặt nước để hướng tới mục tiêu giảm thiểu lượng rác nhựa không lồ trong môi trường hiện nay.

**Lời cảm ơn:** Nghiên cứu này được tài trợ chi phí chế tạo máy WSCA.2.0 bởi UNESCO Việt Nam thông qua Chương trình Tìm kiếm Ý tưởng sáng tạo vì một Đại dương không Nhựa

### TÀI LIỆU THAM KHẢO

- [1] M. Eriksen *et al.* "Plastic pollution in the world's oceans: more than 5 trillion plastic pieces weighing over 250,000 tons afloat at sea", *PLoS one*, Vol. 9, No. 12, e111913, 2014.
- [2] W. C. Li, H. F. Tse, and L. Fok "Plastic waste in the marine environment: A review of sources, occurrence and effects", *Science of the Total Environment*, Vol. 566, pp. 333-349, 2016.
- [3] L.C. M. Lebreton, J. Zwet, J.-W. Damsteeg, B. Slat, A. Andrady, and J. Reisser. "River plastic emissions to the world's oceans", *Nature communications*, Vol. 8, No. 1, pp. 1-10, 2017.
- [4] N. Ruangpayoongsak, J. Sumroengrit, and M. Leanglum. "A floating waste scooper robot on water surface", *In: 2017 17th International Conference on Control, Automation and Systems (ICCAS)*, IEEE, 17430768, 2017, pp. 1543-1548. DOI: 10.23919/ICCAS.2017.8204234
- [5] B. Y. Sanda and I. Ibrahim. "Causes, categories and control of water pollution", *International Journal of Scientific Engineering and Science*, Vol. 4, No. 9, pp. 84-90, 2020.
- [6] O. Helinski, C. Poor and J. Wolfand. "Ridding our rivers of plastic: A framework for plastic pollution capture device selection", *Marine pollution bulletin*, Vol. 165, pp. 112095, 2021.

- [7] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023, pp. 7464-7475.
- [8] Y. Bai; Y. Zhang; M. Ding; and B. Ghanem. "Sod-mtgan: Small object detection via multitask generative adversarial network", In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 206–221.
- [9] J. Li, X. Liang, Y. Wei, T. Xu, J. Feng, and S. Yan. "Perceptual generative adversarial networks for small object detection", In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 2017, pp. 1222–1230. <https://doi.org/10.48550/arXiv.1706.05274>
- [10] Y. Cheng *et al.*, "FloW: A Dataset and Benchmark for Floating Waste Detection in Inland Waters", In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 10953-10962.
- [11] S. M. Patole; M. Torlak; D. Wang; and M. Ali "Automotive radarrs: A review of signal processing techniques", *IEEE Signal Processing Magazine*. Vol. 34, No. 2, pp. 22 - 35, 2017.
- [12] A. Ouaknine, A. Newson, J. Rebut, F. Tupin, and P. Pérez, "Carradar Dataset: Camera and Automotive Radarr with Range-Angle-Doppler Annotations". arXiv preprint arXiv:2005.01456, 2021, DOI:10.1109/ICPR48806.2021.9413181.
- [13] W. Visser, *"The Cognitive Artifacts of Designing"*. Lawrence Erlbaum Associates, 2006.
- [14] N. Cross, "Design Cognition: Results from Protocol and Other Empirical Studies of Design Activity". In: C. Eastman, M. McCracken, and W. Newstatter (eds.) *Design Knowing and Learning: Cognition in Design Education*, Elsevier, Oxford, 2001, pp. 79-103. ISBN 0 08 043868 7.
- [15] T. Brown, "Design Thinking". Harvard Business Review, June 2008.
- [16] K. Dorst, "Frame Innovation: Create New Thinking by Design". Cambridge, MA: MIT Press, 2012. ISBN 978-0-262-32431-1.
- [17] C. McGinn, E. Bourke, A. Murtagh, C. Donovan, P. Lynch, M. F. Cullinan, and K. Kelly, "Meet Stevie: A Socially Assistive Robot Developed Through Application of a 'Design-Thinking' Approach". *Journal of Intelligent & Robotic Systems*, Vol. 98, No. 1, pp. 39-58, 2020.
- [18] N. Cross, "Design Thinking: Understanding How Designers Think and Work". Berg, 2011. ISBN 9781847886361.
- [19] N. Cross, "The Nature and Nurture of Design Ability". *Design Studies*, Vol. 11, pp. 127–140, 1990.
- [20] M. Hynes *et al.*, "Infusing Engineering Design into High School STEM Courses", *National Center for Engineering and Technology Education*, vol. 165, pp. 1-7, 2011.
- [21] X. Long *et al.*, "PP-YOLO: An Effective and Efficient Implementation of Object Detector". arXiv preprint arXiv:2007.12099.
- [22] C.-Y. Fu, W. Liu, A. Ranga, A. Tyagi, and A. C. Berg, "DSSD: Deconvolutional Single Shot Detector". arXiv preprint arXiv:1701.06659, 2017. <https://doi.org/10.48550/arXiv.1701.06659>
- [23] T.-Y. Lin; P. Goyal; R. Girshick; K. He; and P. Dollár, "Focal Loss for Dense Object Detection". In *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2980–2988. DOI: 10.1109/ICCV.2017.324
- [24] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement". arXiv preprint arXiv:1804.02767, 2018. <https://doi.org/10.48550/arXiv.1804.02767>
- [25] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks". In *Advances in Neural Information Processing Systems*, 2015, pp. 91–99. <https://doi.org/10.48550/arXiv.1506.01497>
- [26] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature Pyramid Networks for Object Detection". In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2117–2125.
- [27] Z. Cai and N. Vasconcelos, "Cascade R-CNN: Delving into High-Quality Object Detection", In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6154–6162. <https://doi.org/10.48550/arXiv.1712.00726>