

ƯỚC LƯỢNG KHOẢNG CÁCH AN TOÀN GIỮA CÁC ĐỐI TƯỢNG ĐANG DI CHUYỂN

ESTIMATE THE SAFE DISTANCE BETWEEN MOVING OBJECTS

HỒ THỊ HƯƠNG THOM

Khoa Công nghệ thông tin, Trường Đại học Hàng hải Việt Nam

*Email liên hệ: thomhth@vamaru.edu.vn

Tóm tắt

Giới hạn một khoảng cách là một trong những thuật ngữ đã phổ biến đối với các đối tượng như người đi bộ, xe cộ, tàu thuyền đang di chuyển/lưu thông trên các tuyến giao thông. Các đối tượng buộc phải duy trì một khoảng cách vừa đủ nhằm đảm bảo an toàn và ngăn chặn va chạm có thể xảy ra. Để có thể ước lượng khoảng cách an toàn, nghiên cứu này vận dụng thuật toán xác định khoảng cách Euclid để có thể đưa ra dự báo khoảng cách an toàn giữa các đối tượng đang di chuyển. Trước tiên sẽ sử dụng mạng nơ ron sâu YOLOv8 để phát hiện đối tượng, thư viện OpenCV để xử lý hình ảnh; sau đó khoảng cách giữa các đối tượng được ước tính với bất kỳ cặp đối tượng nào trong khung hình, sẽ được biểu thị bằng khung viền màu đỏ nếu có ghi vẩn thiếu an toàn, ngược lại sẽ là khung viền màu xanh hoặc vàng. Mô hình ước lượng này đã được thử nghiệm và đánh giá cho các đoạn video ghi lại cảnh: người đi bộ trên đường, phương tiện xe cộ trên tuyến đường hoặc tàu thuyền lưu thông trên luồng.

Từ khóa: Ước lượng khoảng cách, mạng nơ ron sâu YOLO.

Abstract

Limiting a distance is one of the popular terms for objects such as pedestrians, vehicles, and boats moving on traffic routes. Subjects must maintain a sufficient distance to ensure safety and prevent possible collisions. To be able to estimate the safe distance, this study applies the Euclidean distance determination algorithm to be able to predict the safe distance between moving objects. First, we will use the YOLOv8 deep neural network to detect objects and the OpenCV library to process images; The distance between objects is then estimated for any pair of objects in the frame, which will be indicated by a red border if unsafe, otherwise by a blue or yellow border. The proposed method has been tested and evaluated

for videos recording scenes of: pedestrians on the road, vehicles on the route or boats traveling on the channel.

Keywords: Estimate safe distance, YOLO deep neural network.

1. Giới thiệu

Giãn cách người đi lại trong đám đông tránh lây lan dịch bệnh như dịch COVID-19, ước lượng khoảng cách an toàn giữa các phương tiện giao thông đường bộ hoặc đường sông là vấn đề rất được quan tâm. Khi các đối tượng này tiếp xúc gần hoặc phương tiện giao thông va chạm có thể gây ra các rủi ro nghiêm trọng cho xã hội và cộng đồng. Chính vì vậy cần xây dựng phương án hỗ trợ giãn cách xã hội tại các nơi công cộng tự động hoặc tại các tuyến đường giao thông giúp hỗ trợ cảnh báo cho người đi bộ, người điều khiển phương tiện giao thông giúp tránh được các rủi ro có thể xảy ra.

Phương án giãn cách đối tượng chính là bài toán ước lượng khoảng cách giữa các đối tượng để có phương án đảm bảo các đối tượng đang di chuyển có khoảng cách an toàn hay không. Để làm được điều đó đầu tiên phải phát hiện được đối tượng qua hình ảnh của các camera giám sát, sau đó ước lượng khoảng cách giữa các đối tượng, đưa ra cảnh báo hoặc hướng xử lý phù hợp đối với các đối tượng đang di chuyển không đảm bảo an toàn.

Sơ đồ tổng quát của phương án xây dựng công cụ ước lượng khoảng cách đến các đối tượng được thể hiện trong Hình 1.



Hình 1. Quy trình ước lượng khoảng cách giữa các đối tượng

Chức năng chi tiết của các thành phần trong sơ đồ Hình 1. sẽ được trình bày cụ thể trong mục 2 và mục 3.2, 3.3 sau đây, mục 4 cài đặt thử nghiệm và đánh giá, phần kết luận được trình bày trong mục 5.

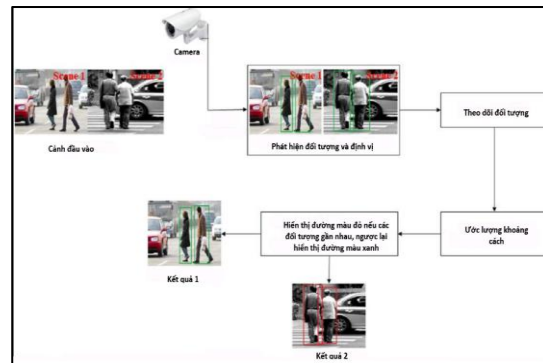
2. Phát hiện đối tượng

Đã có rất nhiều nghiên cứu về phát hiện đối tượng dựa trên hình ảnh. Trong những năm gần đây, với sự

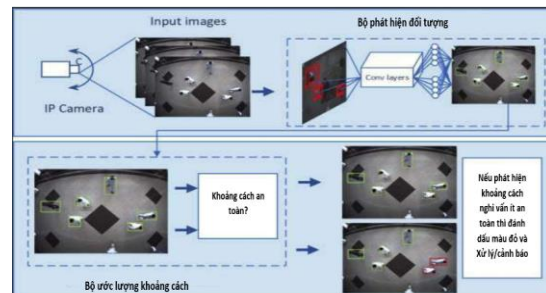
phát triển không ngừng của học sâu, nhiều nhà nghiên cứu đã đưa mạng nơ ron tích chập (CNN) vào lĩnh vực phát hiện đối tượng và đạt những kết quả được ghi nhận. Đặc biệt là mạng nơ ron YOLO trong phát hiện đối tượng. YOLO là viết tắt của cụm từ “You Only Look Once”, tức chỉ khả năng dự đoán tất cả các đối tượng có mặt trong một hình ảnh chỉ trong một lần truyền dữ liệu.

YOLO là mô hình phát hiện đối tượng theo thời gian thực tiên tiến. YOLOv1 và YOLOv2 [1] là mô hình hai thế hệ đầu tiên của YOLO. YOLOv3[2] là phiên bản thứ ba được giới thiệu vào năm 2018 như một cải tiến so với YOLOv2, nhằm tăng độ chính xác và tốc độ của thuật toán. Một trong những cải tiến chính trong YOLOv3 là việc sử dụng kiến trúc CNN mới có tên là Darknet-53. Darknet-53 là một biến thể của kiến trúc ResNet và được thiết kế dành riêng cho các nhiệm vụ phát hiện đối tượng. Sau đó là một loạt các phiên bản khác ra đời gồm YOLOv4 (năm 2020) [4] bởi Bochkovskiy và cộng sự từ một cải tiến so của YOLOv3, YOLOv5 (năm 2020) [5] bởi cùng một nhóm đã phát triển thuật toán YOLO ban đầu dưới dạng một dự án mã nguồn mở và được duy trì bởi Ultralytics, YOLOv6 (năm 2022) [6] bởi Li và cộng sự, YOLOv7 (năm 2022), có một số cải tiến so với các phiên bản trước, một trong những cải tiến chính là việc sử dụng các anchor box. YOLOv7 cũng có độ phân giải cao hơn so với các phiên bản trước. Nó xử lý hình ảnh ở độ phân giải 608 x 608 pixel, cao hơn độ phân giải 416 x 416 được sử dụng trong YOLOv3. Độ phân giải cao hơn này cho phép YOLO v7 phát hiện các đối tượng nhỏ hơn và có độ chính xác tổng thể cao hơn.

YOLOv8 là phiên bản mới được công bố vào đầu năm 2023 [8], YOLOv8 đã mang lại nhiều điểm tích cực so với phiên bản tiền nhiệm, như phát hiện không dùng anchor, giới thiệu lớp tích chập C3 và tăng

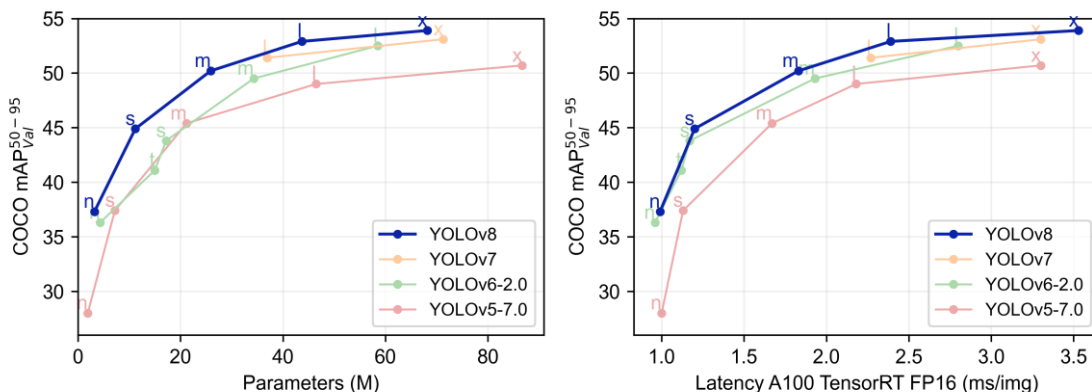


Hình 3. Chế độ xem phối cảnh thực tế



Hình 4. Chế độ xem từ trên xuống để ước lượng khoảng cách

cường mosaic. YOLOv8 có nhiều tham số hơn so với các phiên bản tiền nhiệm như YOLOv5, nhưng ít tham số hơn so với YOLOv6. Nó cung cấp khoảng 33% mAP nhiều hơn cho các mô hình kích thước n và mAP lớn hơn nói chung. YOLOv8 có thời gian suy luận nhanh hơn so với tất cả các phiên bản YOLO khác. Trong YOLOv8, ta có các kích thước mô hình khác nhau: YOLOv8 - n - nano, s - small, m - medium, l - large và x - extra large. Kích thước mô hình tương ứng tuyến tính với mAP và nghịch đảo tương ứng với thời gian suy luận. Các mô hình lớn mất nhiều thời gian suy luận để phát hiện đối tượng một cách chính xác



Hình 2. So sánh mAP của YOLOv8 với một số phiên bản trước

với mAP cao hơn, Hình 2 đưa ra so sánh giá trị trung bình mAP50-95 của YOLOv8 trên dữ liệu COCO [3] với phiên bản YOLOv6 và YOLOv7 cho thấy độ chính xác cao hơn hẳn.

Đặc biệt YOLOv8 hỗ trợ giải quyết được ba vấn đề trên một khung hình đó là: Phát hiện đối tượng (Object Detection), phân vùng đối tượng (Instance Segmentation) và phân loại đối tượng (Image Classification). Do các ưu việt của YOLOv8 nên trong nghiên cứu này sử dụng YOLOv8 để phát hiện đối tượng người đi bộ, phương tiện tham gia giao thông đường bộ và đường thủy.

3. Ước lượng khoảng cách

Ước lượng chính xác khoảng cách giữa hai đối tượng trong không gian thực (3 chiều - 3D) là bài toán không phải dễ dàng và phụ thuộc vào nhiều yếu tố. Để đơn giản hơn, ở đây chuyển đổi ảnh bối cảnh (không gian thực) sang không gian hai chiều (mặt phẳng) theo hướng nhìn từ trên xuống, khi đó mọi đối tượng đều có thể coi là đang đứng trên cùng một mặt phẳng, và có thể dễ dàng ước lượng khoảng cách giữa các đối tượng bằng công thức Euclid.

3.1. Phương pháp ước lượng

Ban đầu, mạng sử dụng nơ ron học sâu CNN được dùng để phát hiện đối tượng tàu thuyền, ở đây lựa chọn mạng học sâu CNN dựa trên thuật toán YOLOv8 [8] được sử dụng để phát hiện đối tượng trong khung hình video. Từ kết quả phát hiện, chỉ ghi nhận các đối tượng quan tâm, các lớp đối tượng khác bỏ qua trong ứng dụng này. Do đó, hộp giới hạn phù hợp nhất cho từng đối tượng được phát hiện có thể được vẽ trong các khung hình của video và các dữ liệu về đối tượng đã phát hiện này sẽ được sử dụng để ước lượng khoảng cách đảm bảo an toàn.

Giả sử máy ghi hình được chụp ở góc cố định làm khung hình của video và các khung hình này được coi là chế độ xem phối cảnh (Hình 3) được chuyển thành chế độ xem từ trên xuống hai chiều để ước tính đo khoảng cách chính xác hơn (Hình 4). Trong phương pháp này, giả định rằng người đi bộ trong khung hình video đang di chuyển trên cùng một mặt phẳng. Bốn điểm mặt phẳng được chọn từ khung hình và sau đó chuyển sang chế độ xem từ trên xuống. Vị trí của mỗi đối tượng có thể được ước tính dựa trên chế độ xem từ trên xuống. Khoảng cách giữa các đối tượng có thể đo và thu nhỏ. Tùy thuộc vào khoảng cách tối thiểu đặt trước, bất kỳ khoảng cách nào nhỏ hơn khoảng cách chấp nhận được giữa hai đối tượng bất kỳ sẽ được biểu thị bằng các đường màu đỏ dùng làm cảnh báo phòng ngừa.

3.2. Hiệu chỉnh chế độ xem khung hình (video)

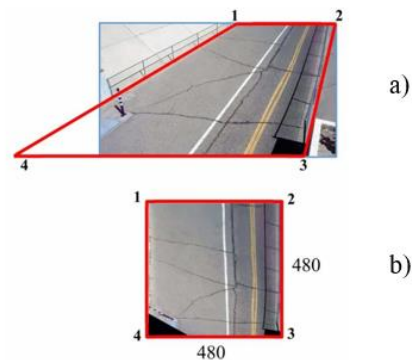
Vùng quan tâm (ROI) được chuyển sang miền 2D, sau đó biến đổi từ chế độ phối cảnh (như minh họa một bối cảnh chụp trên đường phố trong Hình 5a) thành chế độ theo hướng nhìn từ trên xuống (minh họa trong Hình 5b). Trong thư viện mã nguồn mở về thị giác máy tính (OpenCV), chuyển đổi phối cảnh là một phương pháp hiệu chỉnh hướng nhìn bao gồm việc chọn bốn điểm trong chế độ xem phối cảnh (Hình 5a) và ánh xạ chúng tới các góc của hình chữ nhật (Hình 5b). Do đó, mọi đối tượng đều được coi là đang đứng trên cùng một mặt phẳng. Khoảng cách giữa những đối tượng di chuyển tương ứng với số lượng pixel ở chế độ xem từ trên xuống có thể được ước lượng.

3.3. Ước lượng khoảng cách

Các bước của quy trình ước lượng khoảng cách như sau, vị trí của hộp giới hạn cho mỗi đối tượng là (x, y, w, h) trong chế độ xem phối cảnh và chuyển đổi thành chế độ xem theo hướng từ trên xuống. Đối với mỗi người đi bộ, vị trí trong chế độ xem từ trên xuống được ước tính dựa trên điểm chính giữa dưới cùng của hộp giới hạn. Khoảng cách giữa mỗi cặp đối tượng có thể được tính toán từ chế độ xem từ trên xuống và khoảng cách được chia tỷ lệ theo hệ số tỷ lệ ước tính từ hiệu chỉnh chế độ xem camera. Cho vị trí của đối tượng trong ảnh lần lượt là (x_1, y_1) và (x_2, y_2) , khoảng cách giữa hai đối tượng, ký hiệu d , có thể được tính theo công thức Euclid như sau:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (1)$$

Cặp đối tượng có khoảng cách dưới khoảng cách tối thiểu có thể chấp nhận được, ký hiệu là t , được đánh dấu màu đỏ và những cặp đối tượng còn lại được



Hình 5. a) Chế độ xem phối cảnh ban đầu, b) chế độ xem theo hướng từ trên xuống sau khi hiệu chỉnh

đánh dấu màu xanh lá cây. Một đường màu đỏ cũng được vẽ giữa cặp đối tượng có khoảng cách dưới ngưỡng xác định trước. Cách đánh màu cho hộp giới hạn, ký hiệu c , được định nghĩa như sau:

$$c = \begin{cases} red & d < t \\ yellow & d = t \\ green & d > t \end{cases} \quad (2)$$

Trong nghiên cứu này giá trị t (khoảng cách giữa hai đối tượng) được tính theo số pixel (ứng với tỉ lệ khoảng cách mét (m) tương ứng), ví dụ với người đi bộ thì giá trị t từ 50 pixel - 80 pixel (tương đương khoảng cách 1,524m - 1,828m).

4. Cài đặt và thử nghiệm

Mô hình được cài đặt bằng ngôn ngữ lập trình Python, các thư viện hỗ trợ như OpenCV, Numpy, Pillow và Ultralytics cho YOLOv8.

Thử nghiệm gồm hai giai đoạn: Giai đoạn đầu thu thập dữ liệu và huấn luyện mô hình bằng YOLOv8 để phát hiện đối tượng; giai đoạn sau sẽ thử nghiệm ước lượng khoảng cách giữa các đối tượng phát hiện.

Bộ dữ liệu dùng để huấn luyện mô hình phát hiện

đối tượng người đi bộ gồm 747 ảnh (với 697 ảnh được tải về từ nguồn [11] và 50 ảnh tự thu thập và gán nhãn), tách ra thành ba tập ảnh: Tập train là 523 ảnh (70%), tập valid là 149 ảnh (20%) và tập test là 75 ảnh (10%). Sử dụng YOLOv8n (YOLOv8n (Nano model) để huấn luyện trên Colab Google với chu kỳ là 100 epochs, trong thời gian 0,454 giờ (27 phút 52 giây) được mô hình với các thông số tin cậy P (Precision), R (Recall) và mAP (Mean Average Precision), mAP50-95 như trong Bảng 1. Với bộ dữ liệu này mAP50-95 của mô hình là 94,3% cao hơn mAP50-95 của bộ dữ liệu của COCO [3] (cho 80 đối tượng) là 57,3% [8].

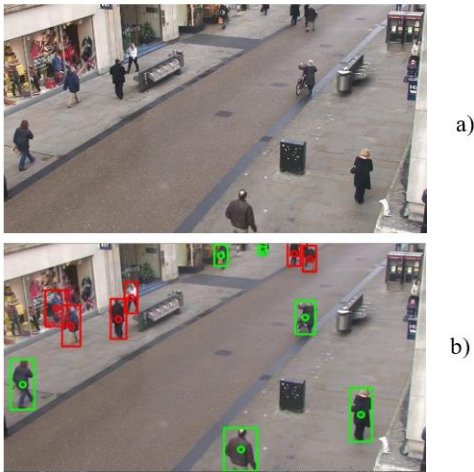
Bộ dữ liệu dùng để huấn luyện mô hình cho phát hiện đối tượng tàu thuyền gồm 899 ảnh với 849 ảnh được tải về từ nguồn [9] và 50 ảnh tự thu thập và gán nhãn, tách bộ dữ liệu thành ba tập: Tập train là 629 ảnh (70%), tập valid là 180 ảnh (20%) và tập test là 90 ảnh (10%). Sử dụng YOLOv7n và YOLOv8n để huấn luyện trên Colab Google với chu kỳ là 100 epochs, được mô hình với các thông số P, R, mAP và mAP⁵⁰⁻⁹⁵ như trong Bảng 1. Với chu kỳ huấn luyện từ 100 đến 140 epochs cho kết quả tương tự Bảng 1.

Bảng 1. Kết quả huấn luyện mạng YOLOv7n và YOLOv8n phát hiện đối tượng người đi bộ và tàu thuyền

TT	Mạng	Tập dữ liệu	Số ảnh	P (%)	R (%)	mAP (%)	mAP ⁵⁰⁻⁹⁵ (%)
1	YOLOv7	Người đi bộ	747	88.2	90.7	95.1	91.1
2		Tàu thuyền	899	85.7	89.8	91.5	62.7
3	YOLOv8	Người đi bộ	747	98.2	97.7	99.1	95.1
4		Tàu thuyền	899	89.7	91.8	94.5	65.7
4	YOLOv8	COCO[8]	330K				57.3

Bảng 2. Kết quả huấn luyện mạng YOLOv8n cho phát hiện đối tượng phương tiện giao thông đường bộ

TT	Đối tượng	P (%)	R (%)	mAP (%)	mAP 50-95 (%)
1	Big bus	81.3	39.7	74.1	55.8
2	Big truck	83.4	45	64.1	42.4
3	Bus-l	5	87.5	4.9	2.5
4	Bus Simulator	24	83.3	25.8	20.7
5	Car	84.8	74.2	82.5	52.2
6	Mid truck	74.9	41.2	42.7	32.9
7	Small bus	22.2	24.5	14.2	8.7
8	Truck Simulator	70.3	55.8	61.5	40.4
9	Truck-l	46.3	68.5	46.6	36.4
10	Truck-m	41.7	67.4	37.6	29.7
11	Truck-s	29.8	55.2	23.8	16.7
12	Truck-xl	49.1	77.7	58.1	45.9
	Trung bình	51.1	60.0	44.7	32.0



Hình 6. Ước lượng khoảng cách giữa an toàn giữa những người đi bộ. a) là ảnh đầu vào, b) là ảnh kết quả ước lượng

Từ Bảng 1 có thể thấy giá trị tin cậy P, R, mAP sử dụng mạng huấn luyện của YOLOv8 và YOLOv7 đều khá cao và của YOLOv8 cao hơn YOLOv7, đặc biệt giá trị mAP⁵⁰⁻⁹⁵ của YOLOv8 còn cao hơn khi sử dụng bộ dữ liệu huấn luyện COCO[8]. Do đó có thể tin cậy lựa chọn YOLOv8 trong việc phát hiện đối tượng để giải quyết bài toán của nghiên cứu này. Vì vậy sẽ sử dụng YOLOv8 trong huấn luyện phát hiện đối tượng xe cơ giới đường bộ và tàu thuyền.

Bộ dữ liệu dùng để huấn luyện mô hình cho phát hiện đối tượng phương tiện giao thông đường bộ (12 loại xe cơ giới - được liệt kê trong Bảng 2) gồm 4058 ảnh được thu thập từ nguồn [10] (trong đó tập train là 2634 ảnh (65%), tập valid là 966 ảnh (24%) và tập test là 458 ảnh (11%)). Sử dụng YOLOv8 để huấn luyện với chu kỳ là 80 epochs, trong thời gian 2,024 giờ được các tỉ lệ tin cậy trung bình cho 12 phương tiện như sau: P là 51,1%, R là 60% và mAP = 44,7%, mAP₅₀₋₉₅ là 32% (mAP₅₀₋₉₅ trên bộ dữ liệu COCO là 53,7% [8]). Cụ thể của từng loại phương tiện theo Bảng 2.

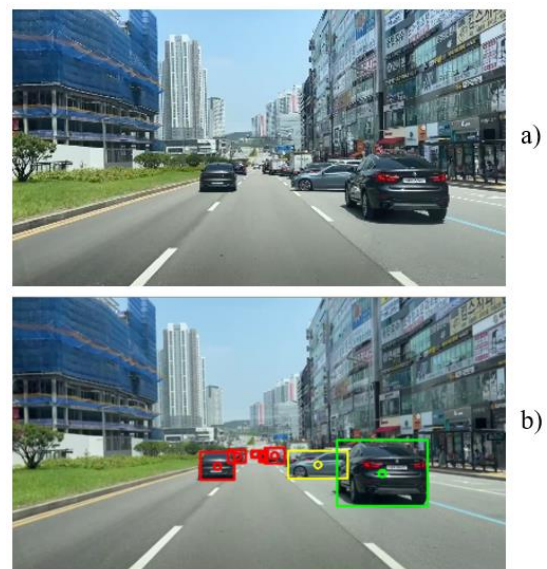
Do sử dụng phát hiện cho 12 đối tượng xe cộ, nhiều đối tượng có hình dạng gần như tương tự nhau chỉ khác nhau về kích thước, dẫn đến phát hiện nhầm lẫn có thể xảy ra, theo Bảng 2 dễ dàng thấy độ tin cậy của một số loại phương tiện loại cỡ nhỏ khá thấp như: Bus-l, Bus - s, Small bus,... Để cải thiện độ tin cậy, có thể nhóm các loại phương tiện vào ba nhóm: Bus, Car, Truck. Gán lại nhãn và huấn luyện lại mô hình phát hiện phương tiện xe cộ, được bảng kết quả độ tin cậy sau khi huấn luyện theo Bảng 3 với mAP₅₀₋₉₅ trung bình cho ba đối tượng là 66,3, cải thiện hơn so với trường hợp phát hiện cho 12 đối tượng.

Bảng 3. Kết quả huấn luyện mạng YOLOv8n cho phát hiện đối tượng phương tiện giao thông đường bộ

TT	Đối tượng	P (%)	R (%)	mAP (%)	mAP ₅₀₋₉₅ (%)
1	Bus	86.3	78.7	74.1	62.8
2	Truck	83.4	77.7	82.1	61.9
3	Car	85.8	82.2	82.5	63.2
Trung bình		85.17	79.53	79.57	62.63

Tiếp theo thử nghiệm trên video ước lượng khoảng cách an toàn giữa những người đi bộ trên đường phố với ngưỡng khoảng cách ước lượng pixel là từ 50 pixel đến 80 pixel ứng với 5 feet - 6 feet (tương đương khoảng cách 1,524m - 1,828m). Hình 6 là kết quả ước lượng khoảng cách giữa người đi bộ trên một khung hình của video. Những trường hợp đối tượng người được bao hình chữ nhật màu đỏ là trường hợp ước lượng cảnh báo khoảng cách không an toàn, ngược lại những đối tượng người có hình chữ nhật màu xanh là an toàn.

Thử nghiệm cho video ước lượng khoảng cách an toàn giữa các phương tiện xe ô tô trên đường phố (vận tốc dưới 60km/h - theo Luật số: 23/2008/QH12 - Luật giao thông đường bộ [12]) với ngưỡng khoảng cách ước lượng là trên 35m), các trường hợp khác khi ứng dụng thực tế có thể phải hiệu chỉnh cho phù hợp. Hình 7 là kết quả ước lượng khoảng cách giữa các phương tiện ô tô trên một khung hình của video.



Hình 7. Ước lượng khoảng cách giữa an toàn giữa những phương tiện ô tô. a) là ảnh đầu vào, b) là ảnh kết quả ước lượng

Khoảng cách an toàn của các tàu thuyền khi tham gia lưu thông trên luồng tuyến đường sông chưa được chỉ rõ cụ thể bao nhiêu trong Thông tư số 19/2013/TT-BGTVT ngày 06/8/2013 của Bộ trưởng Bộ Giao thông vận tải [13]. Ở đây chỉ thử nghiệm ước lượng khoảng cách an toàn giữa các phương tiện tàu thuyền lưu thông trên luồng tuyến với ngưỡng khoảng cách ước lượng là từ 35m - 55m, khoảng cách này có thể hiệu chỉnh cho phù hợp với hoàn cảnh cụ thể và kích thước dài ngắn khác nhau của các loại tàu thuyền. Hình 8 là kết quả ước lượng khoảng cách giữa các tàu thuyền di chuyển trên sông trên một khung hình của video.

Trong mô hình ước lượng khoảng cách thì khâu phát hiện đối tượng là quan trọng nhất, vì khi phát hiện được đối tượng thì mới có thể ước lượng được khoảng cách giữa các đối tượng. Từ Bảng 1 và 3 có thể thấy mô hình phát hiện đối tượng với YOLOv8 là khá tốt, đặc biệt đối với đối tượng là người đi bộ và đối tượng tàu thuyền có mAP50-95 lần lượt là 95,1% và 65,7% cao hơn mAP50-95 của mô hình YOLOv8 với dữ liệu COCO. Độ tin cậy cao giúp phát hiện đối tượng chính xác hơn, điều này làm cho việc ước lượng giữa các đối tượng gặp nhiều thuận lợi.



Hình 8. Ước lượng khoảng cách giữa an toàn giữa những tàu thuyền. a) là ảnh đầu vào, b) là ảnh kết quả ước lượng

5. Kết luận

Nghiên cứu này đã đưa ra mô hình ước lượng khoảng cách giữa các đối tượng: Xe ô tô, người đi bộ, tàu thuyền lưu hành trên sông nước. Từ kết quả thử nghiệm cho thấy đã giải quyết phần nào bài toán ước lượng khoảng cách đảm bảo an toàn cho các đối tượng

người đi bộ, xe cộ trên đường, tàu thuyền trên sông/biển. Độ tin cậy phát hiện đối tượng sử dụng YOLOv8 được chỉ rõ trong các Bảng 1, 2 và 3 với mAP50-95 cao hơn với trường hợp sử dụng bộ dữ liệu COCO phát hiện cho 80 đối tượng.

Hướng nghiên cứu tiếp theo sẽ tập trung nghiên cứu cải thiện kết quả phát hiện đối tượng với YOLOv9 vừa được công bố vào tháng 2/2024 [14] và phát hiện thêm các đối tượng khác cho phương tiện tham gia giao thông đường bộ và các vật thể lạ (như cano, phao cứu sinh, đá ngầm, chướng ngại vật lạ,...) nổi trên biển sử dụng YOLOv8 với bộ dữ liệu thu thập thực tế hoặc giả lập trong phòng thử nghiệm mô phỏng. Việc nghiên cứu ước lượng khoảng cách đến các vật thể lạ giúp tàu thuyền tránh đâm va đảm bảo an toàn khi lưu thông trên luồng tuyến.

Ngoài ra sẽ nghiên cứu thêm phương pháp đo tốc độ vận chuyển của các đối tượng cùng di chuyển để ước lượng khoảng cách phù hợp đảm bảo tốt hơn cho giao thông đường thủy.

Lời cảm ơn

Nghiên cứu này được tài trợ bởi Trường Đại học Hàng hải Việt Nam trong đề tài mã số: **DT23-24.62**.

TÀI LIỆU THAM KHẢO

- [1] Redmon, J. (2016); Farhadi, A. *YOLO9000: Better, faster, stronger*. arXiv 2016, arXiv:1612.08242.
- [2] Redmon, J.; Farhadi (2018), A. *YOLOv3: An Incremental Improvement*. arXiv 2018, arXiv:1804.02767. [Google Scholar]
- [3] Coco dataset: <https://www.cocodataset.org>
- [4] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao, *YOLOv4: Optimal Speed and Accuracy of Object Detection*, <https://doi.org/10.48550/arXiv.2004.10934>.
- [5] Glenn Jocher et al (2021). yolo5: <https://github.com/ultralytics/yolov5>.
- [6] Chuyi Li et al, YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications, <https://doi.org/10.48550/arXiv.2209.02976>
- [7] Chien-Yao Wang et al, YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, <https://doi.org/10.48550/arXiv.2207.02696>
- [8] Om M. Khare, Shubham Gandhi, Aditya M. Rahalkar, Sunil Mane (2023), YOLOv8-Based Visual Detection of Road Hazards: Potholes, Sewer Covers, and Manholes, <https://doi.org/10.48550/arXiv.2311.00073>.

- [9] Boat dataset: <https://universe.roboflow.com/yolo-project/fishing-boat/dataset/2>,
<https://universe.roboflow.com/yamen-gm7rm/boat-xore6/dataset/1>
- [10] Vehicle dataset:
<https://universe.roboflow.com/roboflow-100/vehicles-q0x2v/dataset/2>
- [11] Person dataset:
<https://universe.roboflow.com/tank-detect/person-dataset-kzsop/dataset/3>
- [12] <https://thuvienphapluat.vn/van-ban/Giao-thong-Van-tai/Luat-giao-thong-duong-bo-2008-23-2008-QH12-82203.aspx>
- [13] <https://cangvuhanghaidanang.gov.vn/vi/thong-tu-so-192013tt-bgtvt-ngay-0682013-cua-bo-truong-bo-giao-thong-van-tai-uy-dinh-viec-ap-dung-quy>.
- [14] Chien-Yao Wang, I-Hau Yeh, and Hong-Yuan Mark Liao (2024), *YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information*, arXiv:2402.13616v2.

Ngày nhận bài:	17/01/2024
Ngày nhận bản sửa lần 01:	28/02/2024
Ngày nhận bản sửa lần 02:	15/03/2024
Ngày duyệt đăng:	23/03/2024