

NGHIÊN CỨU CÁC KỸ THUẬT THỊ GIÁC MÁY TÍNH ỨNG DỤNG THEO DÕI PHƯƠNG TIỆN, GIÁM SÁT GIAO THÔNG TRONG THỜI GIAN THỰC

RESEARCH COMPUTER VISION TECHNOLOGY FOR TRAFFIC MONITORING AND VEHICLE TRACKING
IN REAL-TIME

Trần Xuân Kiên¹, Đinh Xuân Trường²,
Phạm Tuấn Thành³, Lê Minh Dương⁴

DOI: <https://doi.org/10.57001/huih5804.2023.103>

TÓM TẮT

Bài báo trình bày về việc sử dụng YOLOv7 để xử lý ảnh trong bài toán đếm phương tiện giao thông, nhận dạng các vật cản có thể gây sự cố giao thông và dùng Haar cascade để nhận diện đối tượng và tính toán đo tốc độ phương tiện. Kết quả dựa trên nền trí tuệ nhân tạo, YOLOv7 có thể nhận dạng được chính xác trên 85 loại phương tiện (80 phương tiện và vật thể) và các vật thể các loại (5 vật thể do người dùng thêm vào). Không những nhận dạng được mà còn bám sát theo dõi hướng di chuyển, tạo vết ảnh trên màn hình để tiện theo dõi và giám sát. Kết quả đếm lưu lượng phương tiện chính xác, gồm hai tham số: là mật độ tham gia giao thông tại một nút và đếm tổng số các phương tiện trong một khoảng thời gian, đảm bảo xác định mức độ tham gia giao thông tại mỗi nút giao thông. Dựa trên thuật toán xử lý ảnh trí tuệ nhân tạo Haar Cascade, OpenCV xác định ước lượng tốc độ di chuyển của xe ô tô trên cao tốc. Giúp giám sát tốc độ của các phương tiện trực tiếp thời gian thực, lưu được ảnh xe gắn với hiển thị tốc độ hiện tại. Hỗ trợ giám sát và quản lý phương tiện trên cao tốc.

Từ khóa: Hệ thống thị giác máy tính, Open CV, YOLOv7, nhận dạng, đếm và đo tốc độ phương tiện giao thông.

ABSTRACT

This paper presents an advanced system for traffic monitoring and management, which leverages YOLOv7 for image processing. The system is capable of accurately counting traffic, identifying potential traffic hazards, and detecting objects using the Haar cascade algorithm, while also calculating vehicle speeds. By utilizing YOLOv7 the system can accurately identify over 85 types of vehicles and objects, tracking their direction of movement and creating image traces on the screen for easy monitoring and surveillance (80 types of vehicles and objects and 5 types of objects for custom data). Moreover, the traffic flow count results are highly precise, incorporating two parameters: the traffic density at a node and the total number of vehicles within a specific period, ensuring accurate determination of traffic levels at each intersection. Additionally, the system can measure the speed of cars on highways using the Haar Cascade algorithm and OpenCV, enabling real-time monitoring of vehicle speed and saving images of the vehicle with the current speed display. Overall, this system is designed to support the monitoring and management of vehicles on highways, enhancing traffic safety and efficiency.

Keywords: Computer vision system, OpenCV, YOLOv7, identification, counting and measuring vehicle speed.

¹Viện Điện tử, Viện Khoa học và Công nghệ Quân sự

²Viện Khoa học và công nghệ Giao thông vận tải

³Học viện Kỹ thuật Quân sự

⁴Lớp Cao học Tự động hóa, Khóa 33, Học viện Kỹ thuật Quân sự

*Email: txkien2003@gmail.com

Ngày nhận bài: 25/4/2023

Ngày nhận bài sửa sau phản biện: 31/5/2023

Ngày chấp nhận đăng: 15/6/2023

1. GIỚI THIỆU

Ùn tắc giao thông gây lãng phí thời gian và năng lượng dẫn đến thất thoát về tiền bạc là rất lớn. Để giảm thiểu ùn tắc giao thông cũng như trợ giúp cho các công tác điều hành, quản lý và quy hoạch giao thông cần có hệ thống

theo dõi và giám sát, dự đoán lưu lượng giao thông sử dụng thị giác máy tính (Computer Vision). Đã có một số nghiên cứu trên thế giới về việc áp dụng thị giác máy tính trong việc quản lý giao thông cũng như trong việc điều khiển đèn tín hiệu. Trong nghiên cứu "Ứng dụng thị giác

máy tính và khai thác dữ liệu để dự đoán tình trạng tắc nghẽn giao thông trên đường bộ” của Mohsen Amoei [1], tác giả đã sử dụng thị giác máy tính để phát hiện và dự báo tình trạng tắc nghẽn trên đường. Bằng cách trích xuất thông tin từ các hình ảnh thu thập được, tác giả đã tạo ra một tập dữ liệu về số lượng xe ở mỗi địa điểm. Sau khi phân tích và đánh giá dữ liệu, các khu vực tắc nghẽn nhất đã được phát hiện ra, hành vi của luồng giao thông trong ngày, giờ cao điểm, mức độ ảnh hưởng của từng loại phương tiện vào giao thông. Cuối cùng, tác giả đã dự đoán được tổng số sự cố tắc nghẽn trong bảy ngày dựa trên dữ liệu lịch sử. Bên cạnh đó, tác giả cũng có thể dự đoán tổng số các loại phương tiện khác nhau trên đường. Để thực hiện nhiệm vụ này, tác giả đã phát triển nhiều mô hình hồi quy, học sâu và dự báo chuỗi thời gian.

Được giới thiệu trong nghiên cứu “Hệ thống phát hiện và đếm phương tiện dựa trên thị giác máy tính sử dụng kỹ thuật học sâu trên đường cao tốc”, Huansheng Song [2] nêu ra việc phát hiện và đếm phương tiện thông minh ngày càng trở nên quan trọng trong lĩnh vực quản lý đường cao tốc. Trong hệ thống phát hiện và đếm phương tiện được đề xuất, mặt đường cao tốc trong ảnh lần đầu tiên được trích xuất và phân chia thành khu vực xa và khu vực gần bằng một phương pháp phân đoạn mới được đề xuất, phương pháp này rất quan trọng để cải thiện khả năng phát hiện phương tiện. Sau đó, hai khu vực trên được đưa vào mạng YOLO để phát hiện loại và vị trí của phương tiện. Cuối cùng, quỹ đạo của xe được thu thập bằng thuật toán ORB, thuật toán này có thể được sử dụng để đánh giá hướng lái của xe và thu được số lượng xe khác nhau.

Earnest Paul Ijjina [3] trong nghiên cứu “Phát hiện tai nạn dựa trên thị giác máy tính trong giám sát giao thông” cũng đã sử dụng thị giác máy tính để phát hiện tai nạn giao thông trên đường. Nghiên cứu sử dụng thị giác máy tính để phát hiện phương tiện và theo dõi đối tượng để dự đoán xác suất về tai nạn dựa trên sự bất thường về tốc độ và quỹ đạo của một xe sau khi chồng lên phương tiện khác.

Amir-UI-Haque Bhuiyan [4] cho rằng giám sát giao thông và phân tích giao thông là rất cần thiết để đảm bảo một hệ thống giao thông hiện đại và thuận tiện. Tuy nhiên, đây là một nhiệm vụ rất khó khăn vì điều kiện giao thông thay đổi liên tục khiến việc duy trì giao thông theo cách truyền thống là hoàn toàn không thể. Việc thiết kế một hệ thống giao thông thông minh cũng là điều tất yếu đối với các thành phố lớn và sầm uất. Vì vậy, tác giả đã đề xuất một hệ thống giám sát giao thông dựa trên thị giác máy tính sẽ giúp duy trì hệ thống giao thông một cách thuận tiện. Nhóm nghiên cứu tạo ra một phân tích về giao thông trong một khoảng thời gian nhất định, điều này sẽ hữu ích để thiết kế một hệ thống giao thông thông minh và khả thi cho một thành phố sầm uất. Trong phương pháp được đề xuất, các tác giả sử dụng trình phân loại “Adaboost” dựa trên tính năng “Haar” để phát hiện xe từ video. Các tác giả cũng đếm số lượng phương tiện xuất hiện trong video bằng cách sử dụng hai đường phát hiện ảo.

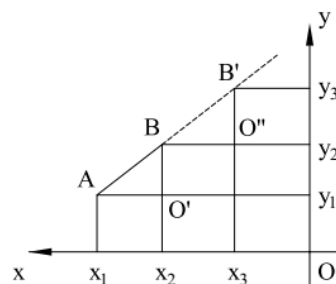
Việc áp dụng công nghệ AI vào giám sát giao thông ở nước ta đã xuất hiện trên nhiều tuyến cao tốc và đường đô thị. Tại các trung tâm điều hành của các tuyến đường cao tốc như Pháp Vân - Cầu Giẽ, Hà Nội - Hải Phòng và một số tuyến đường khác đều có các thiết bị giám sát để có thể phát hiện kịp thời các sự cố trên đường. Khi có hành vi vi phạm sẽ được camera giám sát ghi lại và truyền về cho máy nhận diện ở điểm cuối cao tốc để lực lượng chức năng xử phạt.

Để làm chủ công nghệ, phát triển hệ thống cảnh báo, giám sát, đo tốc độ, nhận dạng phương tiện và vật cản có thể gây sự cố giao thông trong thời gian thực. Bài báo trình bày kết quả của nghiên cứu và theo cấu trúc như sau: Mục 2 trình bày thuật toán đo tốc độ và nhận dạng vật thể; Mục 3 trình bày về thử nghiệm và đánh giá; Mục 4 là kết luận.

2. CÁC THUẬT TOÁN TÍNH TOÁN TỐC ĐỘ VÀ NHẬN DẠNG VẬT THỂ

2.1. Tính toán tốc độ chuyển động của ô tô

Bài toán tính toán tốc độ của ô tô trên đường cao tốc.



$$AB = \sqrt{AO^2 + BO^2} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

$$x_2 : \text{current} \begin{vmatrix} A \\ O \end{vmatrix} \quad x_1 : \text{start} \begin{vmatrix} B \\ 1 \end{vmatrix} \quad y_2 : \text{current} \begin{vmatrix} A \\ O \end{vmatrix} \quad y_1 : \text{start} \begin{vmatrix} B \\ 1 \end{vmatrix}$$

Hình 1. Xác định quãng đường của xe đi

Haar Cascade là một phương pháp phát hiện đối tượng được sử dụng để định vị đối tượng quan tâm trong hình ảnh. Trong bài báo sử dụng YOLOv7 cho việc nhận dạng phương tiện như ô tô, xe tải, xe buýt, xe máy, còn sử dụng Haar Cascade để cho bài toán nhận dạng phương tiện và đo tốc độ phương tiện. Trong bài toán đo tốc độ di chuyển của phương tiện giao thông, nhóm tác giả sử dụng phương pháp nhận dạng phát hiện phương tiện giao thông dựa trên Haar Cascade. Thuật toán được đào tạo trên một số lượng lớn các mẫu âm và dương, trong đó các mẫu dương là hình ảnh có chứa đối tượng quan tâm. Mẫu âm bản là hình ảnh có thể chứa bất kỳ thứ gì ngoại trừ đối tượng mong muốn. Sau khi được đào tạo, bộ phân loại sau đó có thể định vị đối tượng quan tâm trong bất kỳ hình ảnh mới nào.

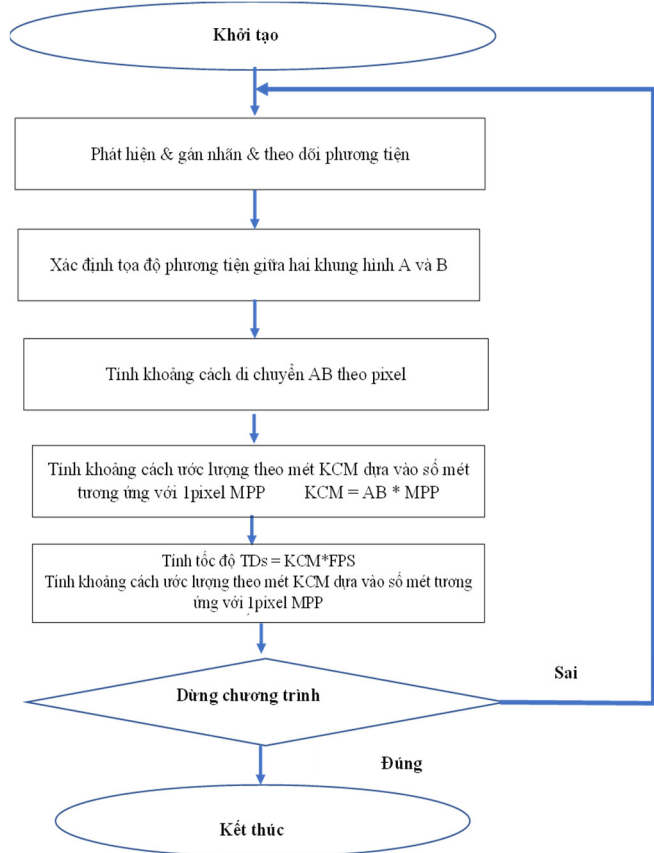
Khoảng cách theo mét là KCM và tính theo công thức:

$$KCM = AB * MPP$$

MPP (số mét tương ứng với 1 pixel, MPP - pixels_per_meter). Chúng ta phải dựa vào đo đạc thực tế và dựa vào hình ảnh trên video để xác định xem bao nhiêu pixel trên ảnh thì ứng với 1 mét ngoài thực địa.

FPS là số khung hình xử lý trong 1 giây (FPS - Frame per second).

Tốc độ tính bằng công thức: $TDs = KCM * FPS$

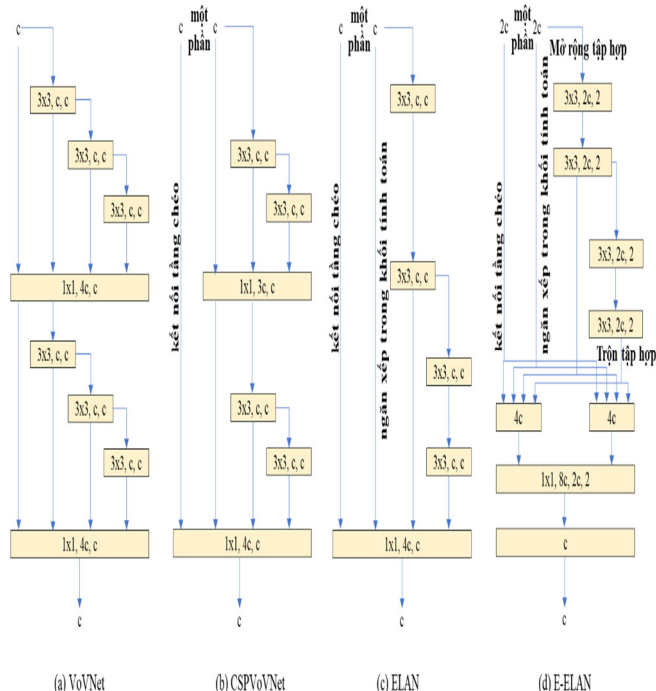


Hình 2. Lưu đồ thuật toán tính vận tốc xe

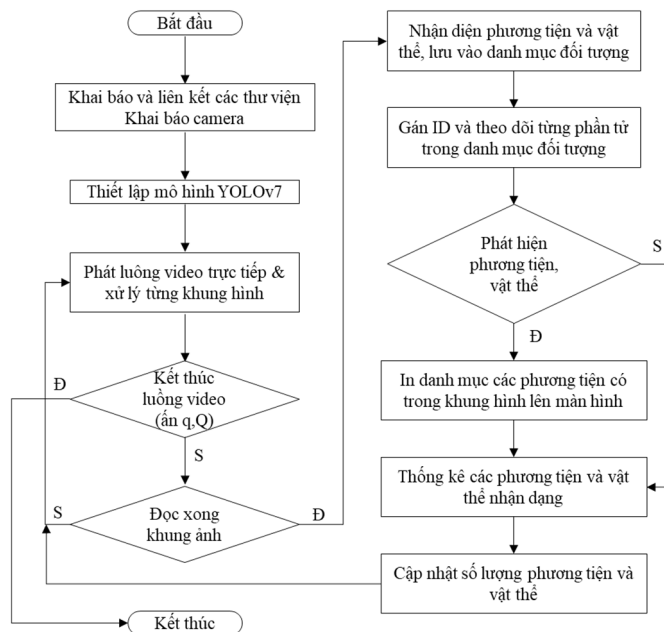
2.2. Nhận dạng vật thể dựa trên YOLOv7

YOLO v7, phiên bản mới nhất của YOLO (đến 2/2023), có một số cải tiến so với các phiên bản trước. Một trong những cải tiến chính là việc sử dụng các hộp neo (anchor box). Các hộp neo là một tập hợp các hộp được xác định trước với các tỷ lệ khung hình khác nhau được sử dụng để phát hiện các đối tượng có hình dạng khác nhau. YOLOv7 sử dụng chín hộp neo, cho phép YOLO phát hiện phạm vi hình dạng và kích thước đối tượng rộng hơn so với các phiên bản trước, do đó giúp giảm số lượng xác định sai. Một cải tiến quan trọng trong YOLOv7 là việc sử dụng một hàm lỗi (loss function) mới gọi là "lỗi trọng tâm" (focal loss). Các phiên bản trước của YOLO đã sử dụng hàm lỗi entropy chéo tiêu chuẩn, được biết là kém hiệu quả hơn trong việc phát hiện các đối tượng nhỏ. Lỗi tập trung giải quyết vấn đề này bằng cách giảm trọng số mất mát cho các ví dụ được phân loại tốt và tập trung vào các đối tượng khó phát hiện. 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 pixel được sử dụng trong YOLOv3. Độ phân giải cao hơn này cho phép YOLOv7 phát hiện các đối tượng nhỏ hơn và có độ chính xác tổng thể cao hơn.

Một trong những ưu điểm chính của YOLOv7 là tốc độ. Nó có thể xử lý hình ảnh với tốc độ 155 khung hình mỗi giây, nhanh hơn nhiều so với các thuật toán phát hiện đối tượng hiện đại khác. Điều này làm cho nó phù hợp với các ứng dụng thời gian thực nhạy cảm như giám sát giao thông và ô tô tự lái, trong đó tốc độ xử lý cao hơn là rất quan trọng. Về độ chính xác, YOLOv7 thể hiện tốt so với các thuật toán phát hiện đối tượng khác. Nó đạt được độ chính xác trung bình là 37,2% ở ngưỡng IoU (giao điểm trên hợp nhất) là 0,5 trên bộ dữ liệu COCO phổ biến, có thể so sánh với các thuật toán phát hiện đối tượng hiện đại khác.



Hình 3. Thay đổi sơ đồ tổng hợp lớp của YOLOv7



Hình 4. Thuật toán nhận dạng đối tượng trên YOLOv7

Là một thuật toán phát hiện đối tượng mạnh mẽ và hiệu quả nhưng YOLOv7 cũng có một số hạn chế như:

- Giống như nhiều thuật toán phát hiện đối tượng, gặp khó khăn trong việc phát hiện các đối tượng nhỏ.

- Nó có thể không phát hiện chính xác các đối tượng trong các cảnh đông đúc hoặc khi các đối tượng ở xa máy ảnh.

- YOLOv7 cũng không hoàn hảo trong việc phát hiện các đối tượng ở các tỷ lệ khác nhau. Điều này có thể gây khó khăn cho việc phát hiện các đối tượng rất lớn hoặc rất nhỏ so với các đối tượng khác trong cảnh.

- YOLOv7 có thể nhạy cảm với những thay đổi về ánh sáng hoặc các điều kiện môi trường khác, vì vậy có thể bất tiện khi sử dụng trong các ứng dụng thực, nơi điều kiện ánh sáng có thể thay đổi.

- YOLO v7 có thể đòi hỏi nhiều tính toán, điều này gây khó khăn khi chạy trong thời gian thực trên các thiết bị hạn chế về tài nguyên.

3. THỬ NGHIỆM VÀ ĐÁNH GIÁ

3.1. Thử nghiệm đo tốc độ

Thực hiện lắp đặt hệ thống và thử nghiệm đo tốc độ tại nút giao thông Láng - Cầu Giấy. Xe chạy với các vận tốc và kết quả đo thể hiện trên bảng 1 và thể hiện trên hình 5.

Bảng 1. Bảng so sánh tốc độ phương tiện

STT	1	2	3	4	5	6	7	8	9	10
Tốc độ xe thực tế (km/h)	48	30	19	27	26	55	52	38	5.3	25
Tốc độ xe đo được (km/h)	48,4	30,18	18,55	27,2	25,71	54,64	52,75	38,77	5,36	25,15
Sai số (%)	0,83	0,60	-2,37	0,74	-1,12	-0,65	1,44	2,03	1,13	0,60

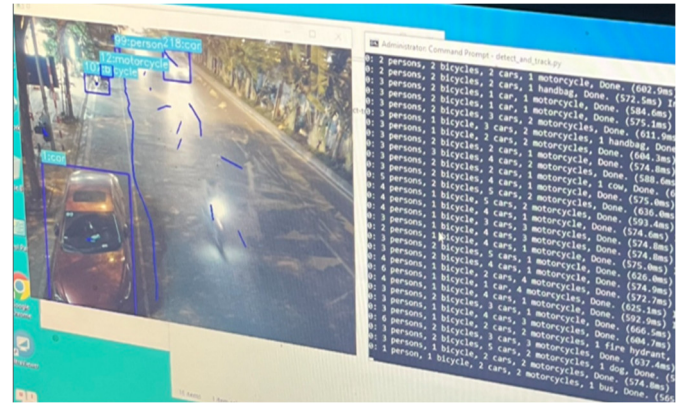


Hình 5. Kết quả đo tốc độ xe, phía trên là hình ảnh thực, phía dưới là màn hình camera

Tốc độ xe thực tế xác định bằng cách cho xe mẫu chạy với tốc độ xác định như theo hàng 1 của bảng 1.

3.2. Thử nghiệm nhận dạng phương tiện

Nhận dạng các phương tiện ban đêm: Hình 6 thể hiện hình ảnh camera và bảng thống kê giám sát đếm phương tiện. Tại thời điểm buổi tối máy tính nhận dạng được các đối tượng sau: 1 người, 1 xe đạp, 2 ô tô, 2 xe máy, 1 xe buýt.



Hình 6. Nhận dạng phương tiện ban đêm

Vào ban ngày và chỉ nhận dạng, đếm ô tô có 5 xe trong phạm vi giám sát, có 4 xe đỗ và xe thứ 39 đang chạy qua trong tổng số đếm được 39 xe đã chạy qua. Kết quả được thể hiện trong bảng 2 và hình 7.

Bảng 2. So sánh kết quả đếm phương tiện

STT	1	2	3	4	5	6
Loại xe	Ô tô	Xe máy	Xe đạp	Người	Xe buýt	Xe tải
Số đếm thực (chiếc/người)	84	197	1	127	1	3
Số đếm trên máy (chiếc/người)	84	197	1	127	1	3
Sai số (%)	0	0	0	0	0	0



Hình 7. Nhận dạng phương tiện chỉ riêng ô tô ban ngày

Các vấn đề gặp khi thực tế là điều kiện thời tiết, các loại xe đi quá gần nhau, nên gây khó khăn rất nhiều khi nhận dạng, đặc biệt là khi xảy ra ùn tắc thì không thể đếm và nhận dạng phương tiện. Thêm nữa là vấn đề là lắp đặt camera phải lắp trên trụ sắt được đặt trực tiếp trên nền đất, nếu đặt trên cầu, thành tường các phương tiện chạy qua sẽ gây rung, sẽ làm ảnh hưởng đến việc xử lý và nhận dạng, đếm phương tiện.

3.3. Thử nghiệm nhận dạng các vật thể gây cản trở giao thông

Trong nghiên cứu nhóm nhận thấy các tai nạn giao thông phần do chủ quan của tài xế như vượt quá tốc độ, sức khỏe không đảm bảo, ngủ gật, mất tập trung (dùng di động, làm việc khác khi lái xe), lái xe có nồng độ cồn, dùng chất kích thích, phần khác do nguyên nhân khách quan như tránh vật cản như túi rác, cây đổ, lốp xe, tránh các phương tiện dừng đỗ sai quy định,... Chính vì vậy nhóm nghiên cứu xây dựng phần mềm nhận dạng và phát hiện các vật thể gây sự cố giao thông. Kết quả thử nghiệm nhận dạng như hình 8.



Hình 8. Nhận dạng các vật thể có thể gây sự cố giao thông: 125- túi rác, 129- lốp xe ô tô, 131- bình nước, 128- ghế tựa

Trong hình các số 125, 131, 128 là số lần chương trình nhận dạng được các đối tượng tương ứng.

4. KẾT LUẬN

Bài báo trình bày về việc sử dụng trí tuệ nhân tạo với việc dùng YOLOv7 để xử lý ảnh trong bài toán đếm phương tiện giao thông (80 loại vật thể và phương tiện giao thông), nhận dạng các vật cản có thể gây sự cố giao thông (05 loại). Không những nhận dạng được mà còn bám sát theo dõi hướng di chuyển, tạo vết ảnh trên màn hình để tiện theo dõi và giám sát. Kết quả thử nghiệm tại một nút giao xác định đã đếm lưu lượng phương tiện chính xác tham gia giao thông tại một nút, với số liệu đó, sử dụng cho bài toán dự đoán được khả năng xảy ra tắc nghẽn giao thông tại nút đó. Bài báo cũng trình bày nội dung về thuật toán xử lý ảnh trí tuệ nhân tạo Haar Cascade để ước lượng tốc độ di chuyển của xe ô tô trên cao tốc. Giúp giám sát tốc độ của các phương tiện trực tiếp thời gian thực, lưu được ảnh xe gắn với hiển thị tốc độ hiện tại. Hỗ trợ giám sát và quản lý phương tiện trên cao tốc.

LỜI CẢM ƠN

Nghiên cứu này được tài trợ bởi Bộ Giao thông vận tải trong đề tài mã số AT22302.

TÀI LIỆU THAM KHẢO

- [1]. Mohsen Amoei, 2020. *Utilizing Computer Vision and Data Mining for Predicting Road Traffic Congestion*. Master Thesis, Concordia University, Canada.
- [2]. Huansheng Song, 2019. *Vision-based vehicle detection and counting system using deep learning in highway scenes*. European Transport Research Review volume 11, Article number: 51.
- [3]. Earnest Paul Ijjina, 2019. *Computer Vision-based Accident Detection in Traffic Surveillance*. 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT).
- [4]. B.T. Amir-Ul-Haque Bhuiyan, M. Das, M. Shamim Reza Sajib, T. Amir-Ul-Haque Bhuiyan, M. Das, M. Shamim Reza Sajib, 2019. *Computer Vision based Traffic Monitoring and Analyzing from On-Road Videos*. Global Journal of Computer Science and Technology: G Interdisciplinary Volume 19 Issue 2.
- [5]. <https://miai.vn/2020/03/09/ket-hop-object-detection-va-object-tracking-chuong-3-thu-lam-he-thong-do-toc-do-xe-hoi-tren-cao-toc/>
- [6]. <https://medium.com/augmented-startups/yolov7-training-on-custom-data-b86d23e6623>
- [7]. <https://medium.com/augmented-startups/roadmap-for-computer-vision-engineer-45167b94518c>
- [8]. <https://medium.com/augmented-startups/yolor-or-yolov5-which-one-is-better-2f844d35e1a1>
- [9]. <https://medium.com/augmented-startups/train-yolor-on-custom-data-f129391bd3d6>
- [10]. <https://medium.com/augmented-startups/develop-an-analytics-dashboard-using-streamlit-e6282fa5e0f>.

AUTHORS INFORMATION

**Tran Xuan Kien¹, Dinh Xuan Truong²,
Pham Tuan Thanh³, Le Minh Duong⁴**

¹Electronic Academy, Academy of Military Science and Technology, Hanoi, Vietnam

²Institute of Transport Science and Technology, Hanoi, Vietnam

³Department of Control Engineering, Military Technical Academy, Hanoi, Vietnam

⁴Automation Master's Class, Course 33, Military Technical Academy, Hanoi, Vietnam