

Nghiên cứu điều khiển hệ thống đèn giao thông thông minh dựa trên thuật toán trí tuệ nhân tạo

Research on controlling intelligent traffic light systems based on artificial intelligence algorithms

Đoàn Văn Đồng

Nhóm nghiên cứu Khoa học công nghệ cho sự phát triển bền vững (STASD), Trường Đại học Giao thông vận tải Thành phố Hồ Chí Minh

Tác giả liên hệ: dongdv@ut.edu.vn

Ngày nhận bài: 02/10/2023; Ngày chấp nhận đăng: 15/11/2023

Tóm tắt:

Hệ thống đèn giao thông thông minh đang trở thành một ứng dụng quan trọng trong việc điều phối giao thông. Trí tuệ nhân tạo với mô hình học sâu được sử dụng hiệu quả trong nhận diện các phương tiện tham gia giao thông từ hình ảnh trích xuất hoặc từ video theo thời gian thực với tốc độ nhanh và độ chính xác cao. Trong bài báo này, hệ thống đèn giao thông thông minh được đưa ra dựa trên phương pháp phân tích hình ảnh theo thời gian thực bằng thuật toán trí tuệ nhân tạo YOLOv8 để xác định mật độ phương tiện tham gia giao thông tại các nút giao nhau. Từ đó, hệ thống tự động điều khiển bộ đếm thời gian các đèn giao thông tại nhiều đường cho phù hợp với lưu lượng giao thông thực tế. Bằng phương pháp này, thời gian hoạt động các đèn được điều chỉnh một cách liên tục, đảm bảo lưu thông hiệu quả và an toàn.

Từ khóa: YOLOv8; Trí tuệ nhân tạo; Học sâu, Mạng nơ ron tích chập.

Abstract:

Intelligent traffic light systems are becoming an important application in traffic coordination. Artificial intelligence with deep learning models effectively recognizes vehicles participating in traffic from extracted images or real-time videos with high speed and accuracy. This paper presents an intelligent traffic light system based on real-time image analysis using the YOLOv8 artificial intelligence method to determine the density of vehicles participating in traffic at intersections. Consequently, the system automatically controls the timing of the traffic lights on the roads to align with the actual traffic flow. Through this method, the operation of traffic lights is continuously adjusted, ensuring efficient and safe traffic flow.

Keywords: You Only Look Once version 8; YOLOv8; Artificial intelligence; Deep learning; convolutional neural networks.

1. Giới thiệu

Theo báo cáo thống kê từ Cục Cảnh sát giao thông, trong 06 tháng đầu năm 2023, cả nước xảy ra 5.000 vụ tai nạn giao thông đường bộ, gây tử vong 2.800 người, và làm bị thương 3.500 người [1]. Tai nạn giao thông không chỉ là nỗi đau của các nạn nhân, gia đình còn là gánh nặng của xã hội và nền kinh tế. Rất nhiều giải pháp đưa ra nhằm đảm bảo lưu thông an

toàn từ thiết kế, quy hoạch, nâng cao cơ sở hạ tầng, đến các giải pháp ứng dụng công nghệ để quản lý, giám sát và điều phối các hoạt động giao thông.

Mặc dù các vụ tai nạn giao thông có xu hướng giảm về số lượng và tính chất nghiêm trọng, tuy nhiên, chưa có tính bền vững. Nguyên nhân chủ yếu là ý thức tham gia giao

thông, mật độ xe cá nhân lớn, cơ sở hạ tầng chưa đáp ứng và thời gian lưu thông kéo dài.

Tình trạng ùn tắc giao thông tại các thành phố lớn đang trở nên trầm trọng hơn khi số lượng xe cá nhân không ngừng tăng lên hàng năm, trong khi, các hệ thống đường bộ chưa được cải thiện rõ nét. Kết quả, dòng phương tiện trở nên rối loạn, và dễ dàng ách tắc khi có sự cố xảy ra.

Một số giải pháp tạm thời đã được áp dụng như điều chỉnh giờ làm việc, giờ học, khoanh vùng cấm đỗ, đường cấm lưu thông trong thời gian cao điểm, cũng như tăng cường xử phạt hành chính đối với hành vi vi phạm. Tuy nhiên, cần có giải pháp tổng thể và toàn diện để giải quyết tình trạng ùn tắc giao thông trong điều kiện cơ sở hạ tầng chưa đáp ứng như hiện nay. Hệ thống giao thông thông minh (Intelligent Transport System-ITS) là một trong những giải pháp đầy hứa hẹn, nhằm áp dụng công nghệ và viễn thông để việc di chuyển trở nên an toàn và hiệu quả hơn [2], [3]. Cho đến nay, ITS đã khẳng định được vai trò trong việc mang lại các lợi ích thực tế [4]:

- Giảm tai nạn, ùn tắc và ô nhiễm môi trường;
- Tiết kiệm thời gian, nhiên liệu, và tăng cường khả năng lưu thông hàng hóa;
- Điều tiết lưu thông của các phương tiện giao thông;
- Hiệu quả quá trình giám sát, quản lý, và điều hành các hoạt động giao thông;
- Nâng cao trải nghiệm cho người tham gia giao thông.

ITS được áp dụng vào thực tế, làm chuyển biến nhận thức của một bộ phận lớn người tham gia giao thông. Tuy nhiên, mật độ tham gia giao thông quá lớn như hiện nay, cần thiết phải có nhiều hơn nữa các giải pháp hỗ trợ trong từng tình hình cụ thể. Trong bài báo [5], hệ thống đèn giao thông thông minh được thực hiện dựa theo độ dài hàng đợi. Độ dài hàng đợi là chiều dài con đường chứa lưu lượng xe chờ đèn đỏ tại nút giao thông. Thế nhưng, yếu tố

độ dài hàng đợi chỉ mang tính tương đối, chưa hoàn toàn phản ánh lượng phương tiện thực tế tham gia giao thông. Hình ảnh thời gian thực đưa vào phân tích để xác định mật độ và cài đặt thời gian cho đèn giao thông được đề xuất tại nghiên cứu [6]. Tuy nhiên, nhận dạng sẽ thiếu chính xác khi thực hiện phương pháp truyền thống bằng việc tính toán cố định các yếu tố (mật độ phương tiện giao thông, loại phương tiện, tình trạng tắc nghẽn,...) để định trước thời gian của các đèn giao thông. Đặc biệt, các phương pháp này gặp nhiều khó khăn khi thực hiện trong điều kiện các phương tiện giao thông không đồng nhất như tại Việt Nam.

Với việc sử dụng công nghệ trí tuệ nhân tạo để xử lý ảnh có khả năng phát hiện và nhận dạng các phương tiện di chuyển trên đường dễ dàng. YOLOv8 là lựa chọn tối ưu, có thể phát hiện và nhận dạng các phương tiện di chuyển trong thời gian thực. Ứng dụng YOLOv8 rất hiệu quả khi phát hiện xe tự hành trong môi trường hỗn hợp nhiều phương tiện tham gia giao thông [7], và để phát hiện được các hành động trộm cắp xe tức thời nhằm giảm thiểu thời gian tra xét [8].

Trong bài báo này, tác giả ứng dụng thuật toán trí tuệ nhân tạo YOLOv8 trong việc nhận diện và phân loại phương tiện giao thông tại các nút giao (ngã tư). Từ đó, xác định số lượng phương tiện và tính toán mật độ lưu lượng phương tiện tham gia giao thông. Cuối cùng, thiết lập tỉ lệ thời gian lưu thông đèn xanh của các hướng phù hợp với nhu cầu thực tế. Kết quả thực nghiệm của bài báo được so sánh với các phương pháp nghiên cứu khác để đánh giá hiệu quả của thuật toán cũng như khả năng ứng dụng thực tế.

Phần còn lại của bài báo được cấu trúc như sau. Phần 2 trình bày cơ sở lý thuyết về YOLOv8. Mô hình hệ thống giao thông thông minh được đề xuất trong phần 3. Phần 4 là đánh giá và phân tích các kết quả đạt được. Cuối cùng, các kết luận được trình bày tại phần 5.

2. Cơ sở lý thuyết về YOLOv8

2.1. Giới thiệu YOLOv8

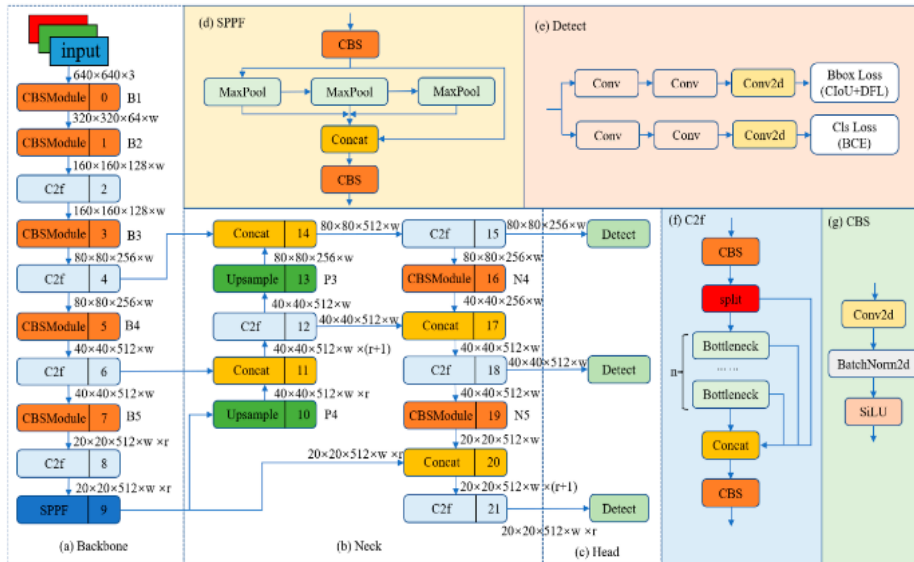
YOLOv8 là viết tắt của “You Only Look Once Version 8”, với ý nghĩa là chỉ cần nhìn một lần để nhận diện đối tượng. Điều này có nghĩa là mô hình có khả năng nhận diện đối tượng trong ảnh chỉ bằng một lần chạy qua mạng nơ ron, không cần chia ảnh thành các vùng nhỏ và áp dụng nhiều lần như các phương pháp truyền thống khác.

Mô hình YOLOv8 sử dụng kiến trúc mạng nơ ron tích chập (CNN) để học các đặc trưng của đối tượng trong ảnh. Quá trình huấn luyện

mô hình bao gồm việc sử dụng dữ liệu huấn luyện có nhãn để điều chỉnh các trọng số của mạng nơ ron, từ đó, tăng cường khả năng nhận diện của mô hình. Một trong những ưu điểm nổi bật của YOLOv8 là khả năng xử lý nhanh chóng và hiệu quả. Nhờ sự tối ưu hóa và cải tiến của mô hình, YOLOv8 có thể xử lý ảnh với tốc độ gần như thời gian thực, giúp tiết kiệm thời gian và tăng cường hiệu suất trong các ứng dụng thực tế.

2.2. Kiến trúc YOLOv8

Kiến trúc mạng YOLOv8 chủ yếu bao gồm phần cốt lõi (backbone, neck, và head), như được hiển thị trong Hình 1.



Hình 1. Kiến trúc YOLOv8.

Backbone: Đóng vai trò là một mạng nơ ron tích chập (CNN), giúp trích xuất đặc trưng đầu vào. Trong YOLOv8, Backbone được xây dựng dựa trên kiến trúc Darknet-53, với mục tiêu tăng cường khả năng nhận diện và giảm thiểu lỗi phân loại mạng. Backbone trong YOLOv8 sử dụng các kích thước khác nhau của các lớp tích chập. Điều này giúp mô hình có khả năng nhận diện vật thể ở các kích thước khác nhau trong ảnh đầu vào.

Neck: Có nhiệm vụ kết hợp các đặc trưng từ các lớp trước đó và tạo ra các đặc trưng tập trung hơn để đưa ra dự đoán chính xác về vị trí và đối tượng trong ảnh. Neck thường được thiết kế dưới dạng các lớp convolution và

pooling để giảm kích thước của đầu vào, tạo ra các đặc trưng tập trung. Các lớp này được xếp chồng lên nhau để tăng cường khả năng học và trích xuất thông tin của mô hình. Một số phương pháp phổ biến được dùng trong Neck bao gồm Spatial Pyramid Pooling Fast (SPPF) và Path Aggregation Network (PAN).

Head: Chức năng chính của phần này là tạo ra các dự đoán về vị trí, kích thước và loại đối tượng trong ảnh. Để thực hiện được điều này, phần Head sử dụng một mạng nơ ron network với nhiều lớp Fully Connected và Convolutional để học các đặc trưng của các đối tượng trong ảnh. Cấu trúc đầu ra tách rời sử dụng hai nhánh riêng biệt cho phân loại đối

tượng và dự báo hộp giới hạn, các hàm mất mát khác nhau được sử dụng cho hai loại nhiệm vụ này. Đối với nhiệm vụ phân loại, sử dụng hàm mất mát nhị phân chéo - Entropy (BCE Loss). Đối với nhiệm vụ dự báo hộp giới hạn, sử dụng hàm mất mát Focal phân phối (DFL) và CIoU.

2.3. Hoạt động của YOLOv8

Nguyên lý hoạt động của mạng YOLOv8 dựa trên việc chia ảnh đầu vào thành một lưới ô vuông và dự đoán các hộp giới hạn (Bounding Boxes) chứa các đối tượng trong từng ô vuông. Mỗi ô vuông có một số lượng hộp giới hạn được dự đoán, và mỗi hộp giới hạn có một điểm số (confidence score) thể hiện mức độ chắc chắn của việc dự đoán.

Mạng YOLOv8 sử dụng kiến trúc Darknet-53 làm bộ trích xuất đặc trưng cho ảnh đầu vào. Darknet-53 là một mạng nơ ron tích chập sâu (Deep Convolutional Neural Network) với 53 lớp, được xây dựng để rút trích các đặc trưng quan trọng từ ảnh. Các mạng nơ ron tích chập sâu như Darknet-53 đã được chứng minh là rất hiệu quả trong việc nhận dạng và phân loại hình ảnh.

Sau khi ảnh được rút trích đặc trưng bởi Darknet-53, mạng YOLOv8 tiếp tục với các lớp tích chập để kết hợp thông tin từ nhiều ô vuông khác nhau và dự đoán các hộp giới hạn. Mỗi hộp giới hạn được xác định bởi tọa độ của góc trên bên trái và góc dưới bên phải, cũng như điểm số thể hiện mức độ chắc chắn của việc dự đoán.

Để tăng cường độ chính xác, YOLOv8 sử dụng nhiều kỹ thuật tiên tiến để giúp cho việc học hiệu quả hơn như khối dư “residual block”. Đây là kiến trúc dùng kết nối tắt đồng nhất để sử dụng xuyên qua một hay nhiều lớp nhằm giảm thiểu hiện tượng độ dốc biến mất “vanishing gradient” nguyên nhân dẫn đến quá trình học tập “learning” không tốt. Ngoài ra, thuật toán cũng sử dụng các kỹ thuật tăng cường dữ liệu “Data Augmentation” bằng việc

biến đổi ảnh đầu vào để tăng cường khả năng tổng quát hóa của mô hình.

Mạng YOLOv8 còn sử dụng các kỹ thuật như “Non-Maximum Suppression” để loại bỏ những hộp giới hạn trùng lặp và giữ lại những hộp giới hạn có điểm số cao nhất. Điều này giúp tăng độ chính xác của mô hình và giảm số lượng hộp giới hạn không cần thiết.

YOLOv8 có các kích thước khác nhau: YOLOv8n (Nano), YOLOv8s (Small), YOLOv8m (Medium), YOLOv8l (Large) và YOLOv8x (Extra Large). Khả năng phân loại mạnh mẽ của nó xuất phát từ việc được huấn luyện trước trên tập dữ liệu lớn “ImageNet”, chứa hàng triệu hình ảnh.

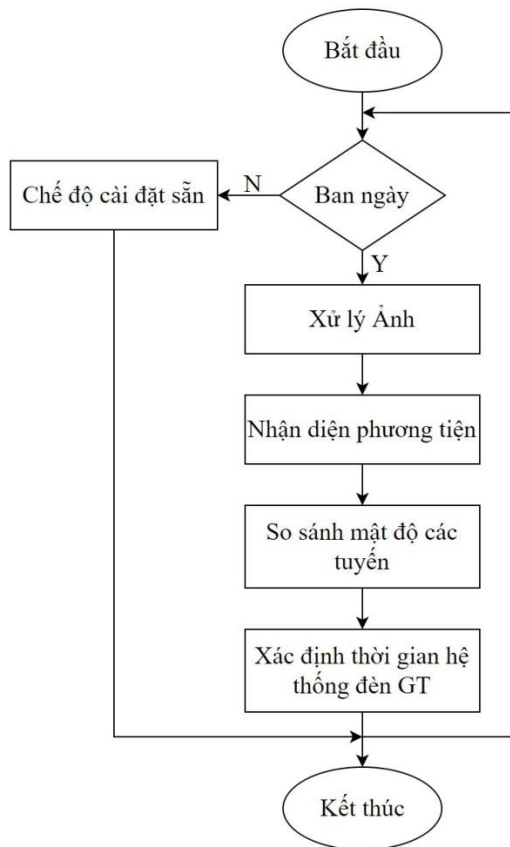
3. Hệ thống đèn giao thông thông minh

Hệ thống đèn giao thông được áp dụng cho các nút giao thông xuyên xảy ra ùn tắc giao thông. Cụ thể, nút giao giữa đường Bạch Đằng và đường Đinh Bộ Lĩnh được thực hiện và kiểm chứng trong bài báo này (Hình 2).



Hình 2. Nút giao giữa đường Đinh Bộ Lĩnh và đường Bạch Đằng.

Các camera được đặt tại các cột đèn giao thông để thu nhận hình ảnh và được kết nối với khối xử lý chính theo sơ đồ hoạt động như Hình 3. Đầu tiên, với 142 hình ảnh về các loại phương tiện phổ biến tại Việt Nam như xe máy, xe ô tô, xe buýt,... được thu thập. Sau đó, dữ liệu đào tạo dần được tăng lên đến 4.700 hình ảnh.



Hình 3. Sơ đồ nguyên lý hoạt động.

Tiếp theo, khi các loại phương tiện có thể nhận dạng nhanh chóng từ dữ liệu thời gian thực, mật độ các phương tiện giao thông có thể được xác định tương ứng. Với vị trí phù hợp từ camera hợp lý khi xét cùng một thời điểm thì sự chênh lệch về mật độ các phương tiện tham gia giao thông là yếu tố quyết định thời gian đèn xanh lưu thông. Cuối cùng, bên cạnh đèn giao thông tự động cài đặt thời gian đèn xanh, vàng, đỏ theo mật độ tham gia giao thông thì hệ thống có thể điều chỉnh các trường hợp “khẩn cấp” từ xa. Trường hợp này áp dụng đối với các sự cố hay tình huống khẩn cấp xảy ra ngay tại nút giao thông.

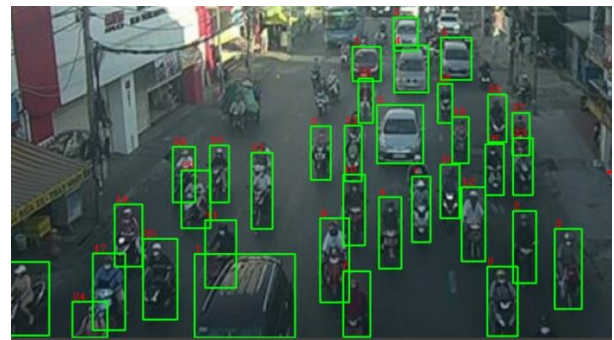
4. Kết quả mô phỏng

Mô hình đã nhận diện và đếm phương tiện với tỷ lệ chính xác cao lên đến 98% ở mỗi tuyến đường (Bảng 1). Trong đó, “Precision” là thông số biểu thị độ chính xác trong dự đoán, “Recall” là thông số biểu thị khả năng phát

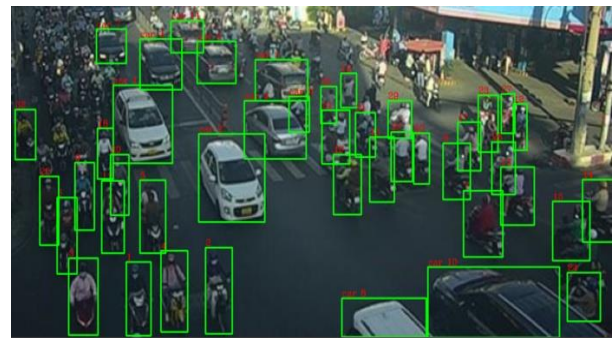
hiện đối tượng trong dữ liệu đầu vào (độ nhạy), và mAP50 là độ chính xác trung bình nhận dạng được. Với các hình ảnh đưa vào đào tạo, độ chính xác trung bình có thể dự đoán đến 98% đối với ô tô và 95,4% đối với xe máy. Tỷ lệ này có thể tăng lên nếu như tăng đầu vào hình ảnh các loại phương tiện.

Bảng 1. Thông số cài đặt.

Loại	Precision	Recall	mAP50	Map50-95
Ô tô	0.935	0.938	0.954	0.717
Xe máy	0.967	0.937	0.98	0.645

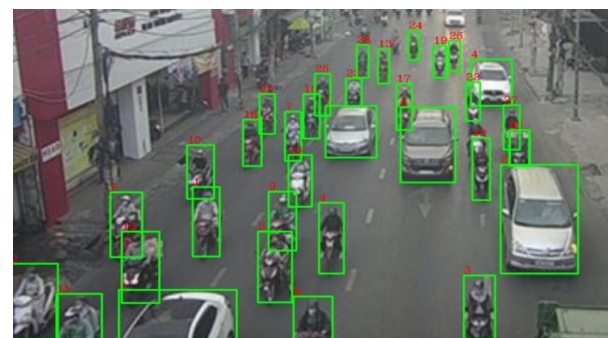


(a) Nút giao hướng đường Đinh Bộ Lĩnh.

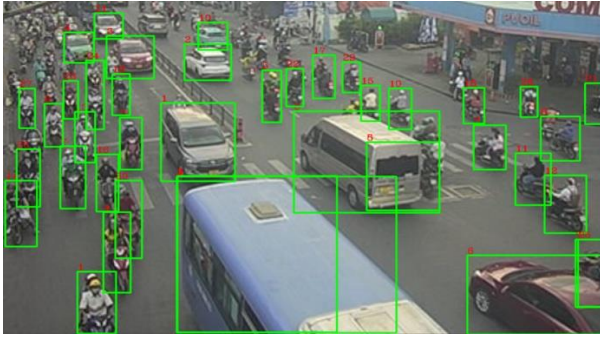


(b) Nút giao hướng đường Bạch Đằng.

Hình 5. Nhận diện phương tiện trong giờ cao điểm (7 giờ 30 phút).

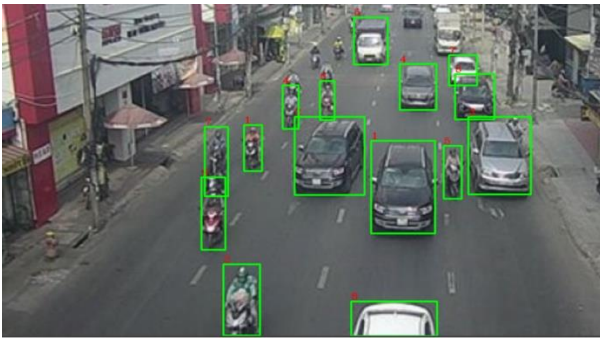


(a) Nút giao hướng đường Đinh Bộ Lĩnh.

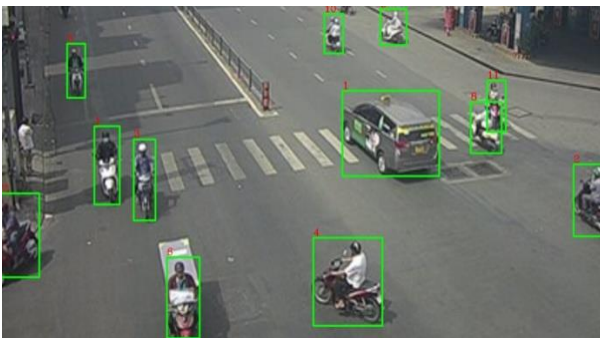


(b) Nút giao hướng đường Bạch Đằng.

Hình 6. Nhận diện phương tiện trong giờ cao điểm (16 giờ 30 phút).



(a) Nút giao hướng đường Đinh Bộ Lĩnh.



(b) Nút giao hướng đường Bạch Đằng.

Hình 7. Nhận diện phương tiện trong giờ cao điểm (11 giờ 30 phút).

Sau khi nhận diện và đếm chính xác các phương tiện, khối xử lý trung tâm xác định tỷ lệ mật độ các tuyến giao nhau để điều khiển bộ đếm thời gian cho đèn giao thông tương ứng. Cụ thể, mật độ phương tiện tại giờ cao điểm tuyến Đinh Bộ Lĩnh và tuyến Bạch Đằng lúc 7 giờ 30 là 0,714 lần và lúc 16 giờ 30 là 0,77 lần. Trong khi đó, mật độ phương tiện tại một thời điểm bình thường giữa hai tuyến này lúc 11 giờ 30 là 1,63 lần. Từ đó, thời gian đèn xanh của các tuyến đường này được điều chỉnh tuyến tính với tỷ lệ mật độ các phương tiện. Đèn vàng và đèn đỏ được điều chỉnh tương ứng với thời gian đèn xanh của mỗi tuyến.

Độ chênh lệch mật độ các phương tiện qua nhiều thời điểm khác nhau cho thấy việc điều chỉnh thời gian đèn giao thông tương ứng là rất hiệu quả. Tuy nhiên, việc xác định tỷ lệ mật độ các tuyến phụ thuộc vào các yếu tố nhận dạng hình ảnh như chất lượng ảnh đầu vào, ánh sáng, điều kiện thời tiết,... Do đó, mô hình đèn giao thông thông minh chỉ tự động điều chỉnh theo thời gian thực trong khung giờ có đủ ánh sáng (ban ngày). Còn đối với ban đêm, hệ thống đèn trở về chế độ tự động cài đặt sẵn.

5. Kết luận

Hệ thống đèn giao thông thông minh được thiết lập dựa trên giải thuật trí tuệ nhân tạo để điều chỉnh thời gian các đèn theo thời gian thực. Với kết quả đạt được, luồng giao thông tại các nút giao điều tiết một cách hiệu quả và thường xuyên dựa theo lưu lượng thực tế phương tiện tham gia giao thông.

Trong tương lai, việc điều khiển bộ đếm thời gian sẽ được thực hiện khi cân nhắc thêm nhiều yếu tố khác như tốc độ phương tiện, độ dài hàng đợi, mức độ ưu tiên các tuyến khác nhau.

Tài liệu tham khảo

- [1] Cục Cảnh sát giao thông. “Điểm báo ngày 11/7/2023.” csgt.vn. Truy cập: 12/07/2023. [Online) Available: <https://www.csgt.vn/m/tintuc/17501/Diem-bao-ngay-11/7/2023.html>.
- [2] P. D. Macheret, R. R. Savchuk and G. I. Shkuratov, “Intelligent Transport Systems: Analysis of the Current State and Prospects of Development,” in *Proc. 2021 Int. Conf. Qual. Manage., Transp. Inf. Secur., Inf. Technol.*, Yaroslavl, Russian Federation, Sep. 6-10, 2021, doi: 10.1109/ITQMIS53292.2021.9642830.
- [3] N. Morri, S. Hadouaj, and L. B. Said, “Towards an Intelligent control system for public transport traffic efficiency KPIs optimization,” in *Proc. 2020 Glob. Congr. Elect. Eng.*, Valencia, Spain, Sep. 4-6, 2020, doi: 10.23919/GC-ElecEng48342.2020.9286268.

- [4] F. -Y. Wang *et al.*, “Transportation 5.0: The DAO to Safe, Secure, and Sustainable Intelligent Transportation Systems,” *IEEE Trans. on Intell. Transp. Syst.*, vol. 24, no. 10, pp. 10262-10278, Oct. 2023, doi: 10.1109/TITS.2023.3305380.
- [5] L. Chai, G. Shen and W. Ye, “The Traffic Flow Model for Single Intersection and its Traffic Light Intelligent Control Strategy,” in *Proc. 6th W. Congr. Intell. Contr. Automat.*, Dalian, Jun. 21-23, 2006, doi: 10.1109/WCICA.2006.1713650.
- [6] A. Dubey, Akshdeep and S. Rane, “Implementation of an intelligent traffic control system and real time traffic statistics broadcasting,” in *Proc. 2017 Int. Conf. Elect., Commun. Aerosp. Technol.*, Coimbatore, India, Apr. 20-22, 2017, doi: 10.1109/ICECA.2017.8212827.
- [7] Afdhal, K. Saddami, S. Sugiarto, Z. Fuadi and N. Nasaruddin, “Real-Time Object Detection Performance of YOLOv8 Models for Self-Driving Cars in a Mixed Traffic Environment,” in *Proc. 2nd Int. Conf. Comput. Syst. Inf. Technol. Elect. Eng.*, Banda Aceh, Indonesia, August. 02-03, 2023, doi: 10.1109/COSITE60233.2023.10249521.
- [8] S. Shafi, T. P. S. K. Reddy, R. Silla and M. Yasmeen, “Deep Learning based Real-time Stolen Vehicle Detection Model with Improved Precision and Reduced Look Up Time,” in *Proc. 3rd Int. Conf. Intell. Technol.*, Hubli, India, Jun. 23-25, 2023, doi: 10.1109/CONIT59222.2023.10205684.