

NHẬN DẠNG ĐỐI TƯỢNG XUẤT HIỆN VÀO VÙNG NGUY HIỂM PHÍA TRƯỚC XE Ô TÔ CỐ LẬP ĐẶT HỆ THỐNG THỊ GIÁC MÁY TÍNH

RECOGNIZING OBJECTS APPEARED IN DANGEROUS REGION IN FRONT OF VEHICLES MOUNTED COMPUTER VISION SYSTEM

Lê Mỹ Hà

Trường đại học Sư phạm Kỹ thuật TP.HCM, Việt Nam

Ngày toà soạn nhận bài 2/4/2019, ngày phân biên đánh giá 18/4/2019, ngày chấp nhận đăng 8/5/2019

TÓM TẮT

Ngày nay, sự phát triển bùng nổ của khoa học công nghệ đã đưa máy móc đến gần và hỗ trợ con người trong các ứng dụng cụ thể trong cuộc sống hằng ngày. Trong đó, đã có rất nhiều ứng dụng có kết quả khả quan, đặc biệt là lĩnh vực xe tự hành, cảnh báo tai nạn. Trong bài báo này, tác giả đề xuất phương pháp sử dụng hệ thống thị giác máy tính gắn trên xe ô tô đang di chuyển trên đường để nhận biết những đối tượng phía trước như oto, xe máy khác xuất hiện trong vùng nguy hiểm phía trước có khả năng gây va chạm. Phương pháp đề xuất dựa trên đặc trưng kênh tổng hợp (Aggregate Channel Features – ACF) để nhận dạng đối tượng xe máy và ô tô trong quá trình tham gia giao thông trên các tuyến đường khác nhau trong môi trường đô thị. Bên cạnh đó tác giả cũng kết hợp kỹ thuật xử lý ảnh để phát hiện làn đường bằng phép biến đổi mắt chim (Bird's-Eye View transformation) và ước lượng khoảng cách từ camera đến vật thể để hỗ trợ cho việc cảnh báo va chạm phía trước. Kết quả cho thấy phương pháp đề xuất hiệu quả trong việc nhận dạng với thuật toán đơn giản và tốc độ xử lý nhanh.

Từ khóa: Máy học; đặc trưng kênh tổng hợp; kỹ thuật mắt chim; nhận dạng làn đường; nhận dạng người.

ABSTRACT

Today, the rapid growth of the new techniques has brought machines close to people. There are many applications that have positive results in our life, especially in autonomous vehicle and collision warning or avoidance. In this article, the authors proposed a method using a computer vision system mounted on moving vehicle to detect the objects appeared in the dangerous region to warning for a collision. The proposed method applied Aggregate Channel Features (ACF) to identify motorbikes and cars in different urban roads. In addition, the author combined lane detection using the bird-eyes view transformation algorithm and estimated the distance from the camera to other objects to support frontal warning. The result showed that this proposed method is an efficient technique with simplicity and fast processing speeds.

Keywords: Machine learning; Aggregate Channel Features; bird-eyes image technique; lane detection; human detection.

1. GIỚI THIỆU

Một trong những ứng dụng rất quan trọng của AI đó là trong lĩnh vực giao thông – ngành công nghiệp ô tô. Nhắc tới ngành này thì ta không thể không kể đến công nghệ tự lái thông minh (hay còn gọi là xe tự lái, xe

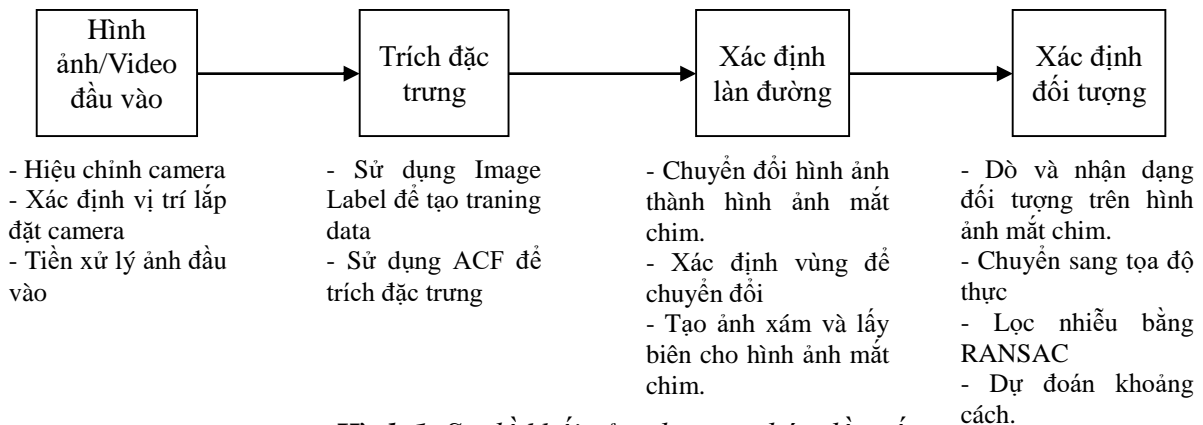
tự hành) mà không cần đến bàn tay của con người, công nghệ này gắn liền với các tên tuổi nổi bật như Ford, Testa, Daimler, Nissan, Toyota... Nhà sản xuất ô tô Nhật Bản Nissan mới đây thông báo đang có kế hoạch thử nghiệm xe tự hành trên các đường phố Nhật Bản ở gần trụ sở chính của công ty tại

Yokohama vào tháng 3 năm 2018. Hệ thống giao thông thông minh – ITS là sự ứng dụng công nghệ cao điện tử, tin học và viễn thông để điều hành và quản lý hệ thống giao thông vận tải. ITS được coi là một hệ thống lớn, trong đó con người, phương tiện giao thông, mạng lưới đường giao thông là các thành phần của hệ thống, liên kết chặt chẽ với nhau. ITS được hoạch định để giảm bớt tắc nghẽn giao thông, đảm bảo an toàn, giảm nhẹ những tác động xấu tới môi trường, tăng cường năng lực vận tải hành khách. Các nghiên cứu về công nghệ trên xe tự hành chủ yếu tập trung vào 2 lĩnh vực chính, đó là phát hiện làn đường và nhận dạng đối tượng. Vấn đề phát hiện làn đường đã được nghiên cứu trong nhiều thập kỷ qua và đã được một số thành quả nhất định, được phát triển và ứng dụng trong nhiều loại xe. Đó là một trong những quá trình quan trọng dựa trên tầm nhìn trong hệ thống hỗ trợ lái xe và có thể được sử dụng để điều hướng xe, kiểm soát hướng, chống va chạm, hoặc cảnh báo khi đi trong làn đường nhỏ. Điều kiện đường xá khác nhau làm cho vấn đề này trở nên rất khó khăn bao gồm các loại khác nhau của các tuyến đường (thẳng hoặc cong), sự che khuất gây ra bởi vật cản, bóng, ánh sáng thay đổi (như thời gian ban đêm), ... đã có nhiều phương pháp tiếp cận đề xuất để giải quyết các vấn đề trên trong việc phát hiện làn đường. Bên cạnh đó, nhận dạng đối tượng là một thành phần quan trọng của hệ thống xe tự hành. Công nghệ này đã có những bước tiến bộ lớn như nhận dạng được đối tượng tĩnh như xe đạp, người đi bộ, ô tô, biển báo giao thông... và những năm gần đây, thì cũng đã có những kết quả khả quan trong việc nhận dạng đối tượng động như xe hay người đang di chuyển trên đường. Việc gia tăng mức độ tự động hóa hay khả năng tự hành có thể mang lại nhiều lợi ích, như tăng độ an toàn (có thể giảm tới hơn 90% các vụ tai nạn) và độ tin cậy, cải thiện thời gian (mẫu xe Cadillac thử nghiệm tự lái của GM có thể tự vận hành với vận tốc lên đến 70 dặm/giờ), tiết kiệm chi phí, hay giảm gánh nặng đặt lên người điều khiển trên đường cao tốc hay khi tắt đường. Trên thế giới, đã có rất nhiều nghiên cứu xoay

quanh những hướng phát triển trong công nghệ ô tô / giao thông được trình bày tóm tắt dưới đây. Thứ nhất trong nghiên cứu [1], Yongzheng Xu đã đề xuất phát hiện phương tiện giao thông từ hình ảnh vệ tinh bằng thuật toán phân lớp SVM kết hợp với đặc trưng HoG, phương pháp này có độ chính xác cao nhưng không phù hợp để áp dụng vào bài toán hướng di chuyển của xe nhiều hơn hai. Trong bài báo [2], biến đổi Hough thực hiện rất tốt trong việc theo dõi làn đường khi đã mất đổi lại một phần của xe sẽ bị mất khi xe chạy lấn sang làn đường khác. Trong bài báo [3] của Zhaojin Zhang, sử dụng Deep Neural Network (DNN) có độ chính xác cao hơn những phương pháp trước đó, nhưng tốc độ xử lý còn chậm. Nghiên cứu về nhận dạng xe dùng mạng nơ-ron tích chập của Jeffrey de Deijn [4], sử dụng CNN để nhận biết một chiếc xe có bị hư hỏng hay không, nhưng độ chính xác chưa cao vì giới hạn của bộ dữ liệu huấn luyện. Trong bài báo [5] Qingpeng Li đã sử dụng một phương pháp mới R3-Net của Neural Network (bao gồm CNN, R-RPN, R-DN kết hợp) để nhận dạng phương tiện. Tuy có thể theo dõi thêm nhiều thông tin của phương tiện như quỹ đạo nhằm nâng cao hiệu suất trong phát hiện xe đa hướng nhưng thuật toán khá phức tạp và chi phí cao. Trong công trình công bố của Daniel Neumann [6] có thể dự đoán gần đúng khoảng cách từ camera tới đối tượng được phát hiện, tuy nhiên tỉ lệ phát hiện đúng chưa cao do tập dữ liệu huấn luyện và độ phân giải của máy ảnh còn thấp. Qua các khảo sát các nghiên cứu gần đây về vấn đề nhận dạng đối tượng thì chủ yếu xoay quanh vào việc tăng độ chính xác và tốc độ xử lý bằng cách sử dụng Deep Learning (một nhánh nổi trội của AI) kết hợp các thuật toán để nhận dạng đối tượng. Với cách đó, trong bài báo này tác giả sử dụng phương pháp đặc trưng kênh tổng hợp Aggregate Channel Features (ACF) để nhận dạng khuôn mặt người hay nhận dạng chữ viết tương tự như trong [7], [8]. Xét về độ phức tạp thì thuật toán trích đặc trưng ACF đơn giản hơn so với Deep Learning, ACF còn có tốc độ nhận dạng khá nhanh do không cần xây dựng các ảnh tích phân và các kênh trong

ACF được xấp xỉ bằng phương pháp ngoại suy để thu được kết quả mong muốn. Chính vì ưu điểm đó, tác giả chọn ACF để thực hiện việc nhận dạng xe máy và ô tô trong môi trường đô thị. Ngoài ra, học viên kết hợp nhận dạng làn đường và dự đoán khoảng

cách từ mono-camera tới đối tượng được nhận dạng. Những vấn đề này chính là một trong những bài toán của hệ thống lái xe tự động đang phát triển mạnh mẽ hiện nay. Phương pháp thực hiện được đề xuất theo 4 bước như trình bày trong Hình 1.



Hình 1. Sơ đồ khối của phương pháp đề xuất

2. PHÁT HIỆN LÀN ĐƯỜNG VÀ NHẬN DẠNG ĐỐI TƯỢNG

2.1. Xác định bộ thông số camera

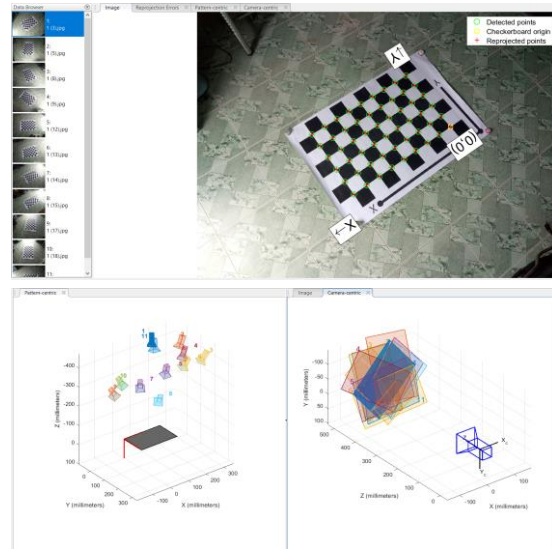
Những đối tượng trong hình ảnh được thu lại bằng camera đơn được nhận dạng chính xác hay việc dự đoán khoảng cách phụ thuộc rất nhiều vào các thông số camera sử dụng, bao gồm: thông số nội, thông số ngoại và hệ số biến dạng. Để xác định các thông số này, tác giả sử dụng Camera Calibration toolbox của Matlab. Để ước tính các thông số của camera đã sử dụng, tác giả sử dụng 11 bức ảnh chụp chess board, với khoảng cách từ camera tới vật thể trung bình là 2m. Kết quả của quá trình này sẽ xác định được các thông số của camera để sử dụng cho các bước tính toán sau này.

2.2. Phát hiện làn đường

2.2.1. Kỹ thuật mất chím

Trong bài báo này, người thực hiện sử dụng phương pháp kỹ thuật mất chím đảo tầm nhìn hình ảnh (Inverse perspective mapping – IPM) để phát hiện làn đường [9]. Kỹ thuật này đòi hỏi yêu cầu cao cho việc kiểm tra các thông số của camera đầu vào,

để đảm bảo cho việc chuyển đổi hình ảnh đạt được chất lượng cao hơn so với kỹ thuật uốn cong.

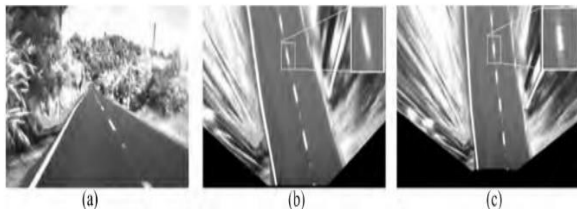


Hình 2. Kết quả hiệu chỉnh camera

Theo phương pháp toán học, IPM có thể được mô hình hóa như một phép chiếu từ một không gian 3D Euclide W , chứa các phần tử $(x, y, z) \in R^3$ thành một không gian con phẳng 2D của R^3 , được biểu diễn

bởi I, với các thành phần $(u, v) \in R^2$. Ảnh xạ từ I sang W được tính như công thức (1)

$$\begin{aligned} x(u, v) &= h \cdot \cot \left\{ \left(\theta_0 - \alpha_u \right) + u \frac{2\alpha_u}{m-1} \right\} \cdot \cos \left\{ \left(\gamma_0 - \alpha_v \right) + v \frac{2\alpha_v}{n-1} + l \right\} \\ y(u, v) &= 0 \\ z(u, v) &= h \cdot \cot \left\{ \left(\theta_0 - \alpha_u \right) + u \frac{2\alpha_u}{m-1} \right\} \cdot \cos \left\{ \left(\gamma_0 - \alpha_v \right) + v \frac{2\alpha_v}{n-1} + d \right\} \end{aligned} \quad (1)$$



Hình 3. Hình ảnh từ kỹ thuật mắt chim sử dụng WPM và IPM [10]

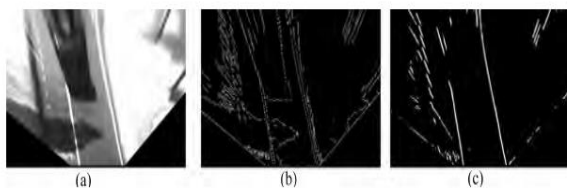
2.2.2. Phát hiện biên bằng kỹ thuật mắt chim

Bằng việc so sánh điểm ảnh hiện tại với điểm ảnh kế bên để phát hiện ra cạnh của một hình ảnh.

$$\begin{aligned} B_{+m}(x, y) &= b(x, y) - b(x+m, y) \\ B_{-m}(x, y) &= b(x, y) - b(x-m, y) \end{aligned} \quad (2)$$

Theo công thức (3), nếu hiệu của 2 điểm ảnh này lớn hơn một mức ngưỡng thì có thể phát hiện ra cạnh của làn đường.

$$r(x, y) = \begin{cases} 1, & \text{if } B_{+m} \geq 0, B_{-m} \geq 0, \text{ and } B_{+m} + B_{-m} \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$



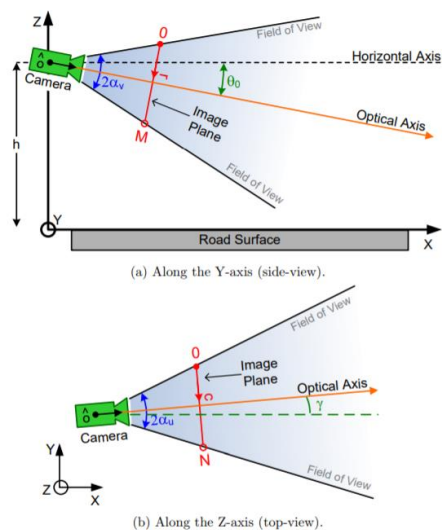
Hình 4. Phát hiện biên [9]

2.2.3. Dự đoán khoảng cách từ camera tới hình ảnh mắt chim

Nhắc lại một số thông số khi chuyển đổi hình ảnh từ tầm nhìn sang hình ảnh mắt chim.

- h là chiều cao của máy ảnh từ mặt đất,
- θ_0 : góc được hình thành bởi trục quang học và trục ngang.
- Góc quan sát của máy ảnh là $2\alpha_u$ theo hàng và $2\alpha_v$ theo cột.
- γ_0 : góc được hình thành bởi phép chiếu của trục quang trên mặt phẳng xz

- n x m là độ phân giải của máy ảnh
- f là tiêu cự camera

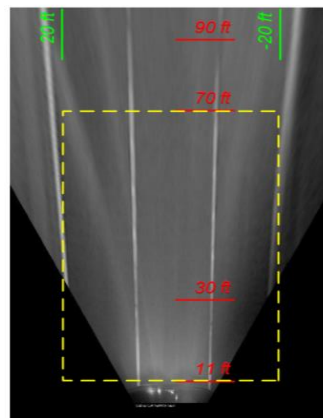


Hình 5. Cấu hình camera

Ta có:

$$\theta_0 = \tan^{-1} \left[\left[1 - 2 \left(\frac{r_H - 1}{m-1} \right) \right] \tan \alpha_v \right] \quad (4)$$

Từ công thức (4) có thể ước tính được vị trí của vật thể đang di chuyển trên đường với camera đơn gắn cố định với góc nghiêng xác định trước, được kí hiệu là r_H .



Hình 6. Vị trí (r_H) của vật thể theo chiều ngang trong hình ảnh camera

Các vị trí của các khoảng cách khác nhau phía trước xe trong hình ảnh mắt chim. Khu vực được sử dụng để xử lý tiếp theo được thể hiện bằng hình chữ nhật màu vàng chấm dứt nét.

2.3. Đặc trưng kênh tổng hợp (Aggregate Channel Features – ACF)

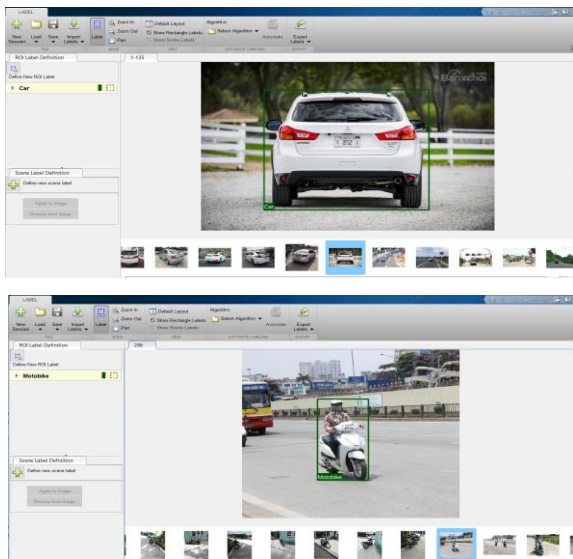
Đặc trưng kênh tổng hợp (Aggregate Channel Features – ACF) được đề xuất bởi Dollar [10] sử dụng việc tra cứu điểm ảnh trong các kênh tổng hợp để giảm thời gian trích đặc trưng mà không cần xây dựng các ảnh tích phân. Hiệu quả phối hợp của các kênh đặc trưng và việc chiến lược lựa chọn đặc trưng bằng cách xếp chồng các bộ phân loại làm các khung cửa sổ ACF hoạt động tốt hơn hầu hết các phương pháp khác về cả độ chính xác, tốc độ xử lý và có cách thực hiện đơn giản hơn.

2.3.1. Đặc điểm

a. Tập dữ liệu huấn luyện

Các dữ liệu huấn luyện (training data) được tạo từ Image Label Toolbox do Matlab hỗ trợ. Trong bài báo này, tập dữ liệu bao gồm:

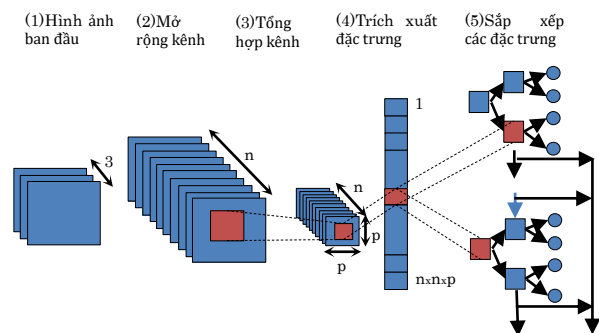
- Tập dữ liệu huấn luyện của xe máy: 207 bức ảnh trong đó có một số bức ảnh có nhiều hơn một chiếc xe máy.
- Tập dữ liệu huấn luyện của xe hơi: 331 bức ảnh trong đó có một số bức ảnh có nhiều hơn một chiếc xe hơi.



Hình 7. Tracking đối tượng cho dữ liệu huấn luyện xe máy và xe hơi

b. Mô hình thực hiện

Với hình I ngõ vào được tính toán với với một vài kênh $C = \Omega(I)$, tổng hợp mỗi khối pixel trong C, làm phẳng các kênh có độ phân giải thấp. Các đặc trưng được tra cứu trong các điểm ảnh đơn lẻ trong các kênh tổng hợp. Tầng cường được sử dụng để huấn luyện và kết hợp trên các tính năng này theo dạng cây để phân biệt đối tượng từ nền và sử dụng phương pháp tiếp cận đa cửa sổ trượt. Với sự lựa chọn thích hợp của các kênh và cấu trúc thiết kế đáng tin cậy, ACF đạt được hiệu quả tốt trong phát hiện người (hình 8).



Hình 8. Mô hình ACF trong nhận dạng đối tượng

Kênh (Channels)

ACF sử dụng 10 kênh như sau:

- Biên độ dốc (gradient magnitude): 1 kênh, độ lớn: $M(x, y) = \sqrt{I_x^2 + I_y^2}$. Với I_x và I_y là giá trị cường độ màu theo trục x và y tại vị trí thứ (x, y).
- Hướng dốc (gradient histogram): 6 kênh, từ $0^\circ \sim 30^\circ$; $31^\circ \sim 60^\circ$; $61^\circ \sim 90^\circ$; $91^\circ \sim 120^\circ$; $121^\circ \sim 150^\circ$; $151^\circ \sim 180^\circ$).
- Kênh màu LUV (3 kênh): L – độ sáng, U và V – giá trị của màu trong ảnh. Kênh màu LUV ít bị ảnh hưởng từ các điều kiện sáng khác nhau.

Trước khi tính toán 10 kênh, hình ảnh ngõ vào I được làm phẳng với một bộ lọc có kích thước $[1 \ 2 \ 1]/4$.

2.4. Thuật toán RANSAC

Từ tập dữ liệu ban đầu, ta sẽ có hai loại dữ liệu nội tuyến và ngoại tuyến (outliers và

inliers), vì thế ta phải đi tính toán để tìm ra mô hình tốt nhất cho tập dữ liệu. Việc tính toán và chọn ra mô hình tốt nhất sẽ được lặp đi lặp lại k lần, với giá trị được chọn sao cho đủ lớn để đảm bảo xác suất p (thường rơi vào giá trị 0.99) của tập dữ liệu mẫu ngẫu nhiên không chứa dữ liệu nhiễu. Nếu gọi u là ước lượng dữ liệu không nhiễu thì $v = 1 - u$ là ước lượng dữ liệu nhiễu và m là số lượng dữ liệu đầu vào cần xây dựng mô hình. Khi đó:

$$1 - p = (1 - u^m)^k \quad (5)$$

k sẽ được tính theo công thức:

$$k = \frac{\log(1 - p)}{\log(1 - (1 - v)^m)} \quad (6)$$

Kết quả thu được sẽ là mô hình cần xây dựng phù hợp nhất với dữ liệu đầu vào, tập các dữ liệu nhiễu và tập các dữ liệu không nhiễu.

3. KẾT QUẢ THỰC NGHIỆM

Để thực hiện việc phát hiện làn đường và các đối tượng trên đường, tác giả sử dụng những video trên đường thực tế trong đô thị, lưu lại bằng camera điện thoại với cảm biến chính có độ phân giải 13MP, khẩu độ f/2.2 và kích thước cảm biến lớn 1/3inch, xử lý offline bằng phần mềm Matlab 2018a, với cấu hình máy tính như sau:

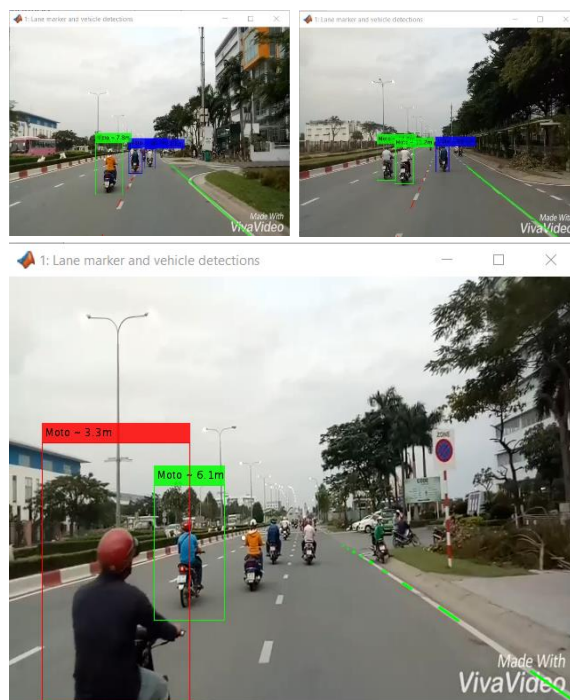
- System Information: Intel(R) Core(TM) i5-5200U CPU @ 2.20GHz (4 CPUs), ~2.2GHz. Memory: 4096MB RAM.
- Display Devices: Card name: NVIDIA GeForce 940M. Memory: 3994 MB

Dữ liệu 1: Trên đường Phạm Văn Đồng, được quay lúc 12h trưa, với vị trí đặt của camera khoảng 1m7 tính từ mặt đất, góc nghiêng khoảng 10° .



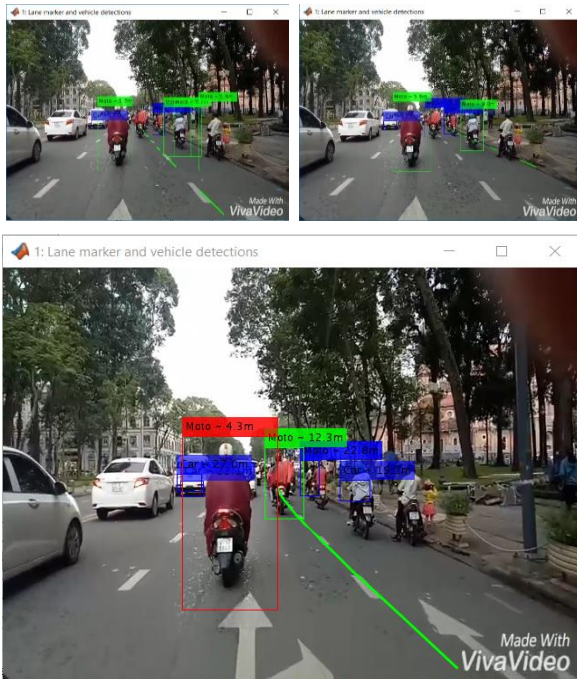
Hình 9. Kết quả nhận dạng từ dữ liệu 1

Dữ liệu 2: Trên đường Khu Công Nghệ cao quận 9, được quay lúc 16h, với vị trí đặt của camera khoảng 2m tính từ mặt đất, góc nghiêng khoảng 12° .



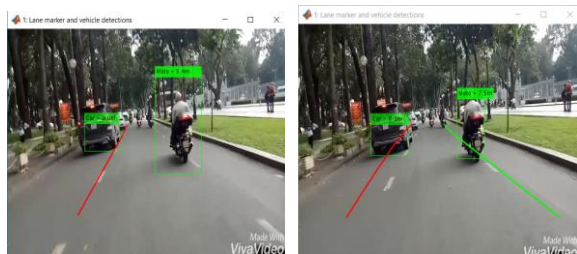
Hình 10. Kết quả nhận dạng từ dữ liệu 2

Dữ liệu 3: Trên đường Lê Duẩn, trong điều kiện hơi âm u (có mưa), được quay lúc 16h, với vị trí đặt của camera khoảng 2m tính từ mặt đất, góc nghiêng khoảng 4° .



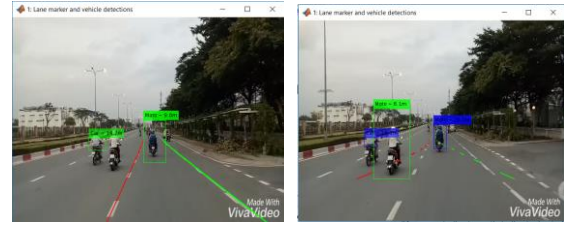
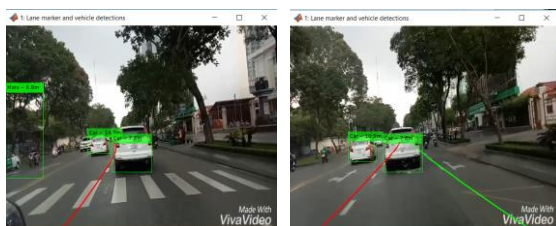
Hình 11. Kết quả nhận dạng từ dữ liệu 3

Dữ liệu 4: trên đường Nam Kỳ Khởi Nghĩa, trong điều kiện hơi âm u (có mưa), được quay lúc 16h, với vị trí đặt của camera khoảng 2m tính từ mặt đất, góc nghiêng khoảng 4° .



Hình 12. Kết quả nhận dạng từ dữ liệu 4

Trường hợp nhận dạng chưa chính xác: Làn đường có thể bị phát hiện chưa đúng do phụ thuộc vào cảm biến máy ảnh trong việc tạo hình ảnh mắt chim; nhận dạng sai giữa phương tiện là xe hơi hay xe máy hay không do dữ liệu huấn luyện không nhiều.



Hình 13. Một số kết quả nhận dạng chưa chính xác

Đánh giá độ chính xác

Để đánh giá độ chính xác của kết quả nhận dạng, tác giả sử dụng thông số ROC (Receiver Operating Characteristic), để xác định là phần nhận được có tín hiệu hay chỉ là do nhiễu.

- *True positive (TP)*: Nếu kết quả từ dự đoán là có xe máy và giá trị thực tế cũng là có xe máy – Phân loại chính xác.
- *True negatives (TN)*: Xảy ra khi cả kết quả dự đoán và giá trị thực tế là không có xe máy.
- *False positives (FP)*: Nếu giá trị thực tế là không có xe máy thì nó được cho có xe máy – phân loại không chính xác
- *False negatives (FN)*: Khi kết quả dự đoán là không có xe máy trong khi giá trị thực tế là có xe máy.

Đánh giá độ chính xác theo công thức (7):

$$accuracy = \frac{TP + TN}{P + N} \quad (7)$$

Bảng 1. Thống kê độ chính xác của 4 video clip trong vấn đề nhận dạng xe hơi, xe máy và làn đường

	TP	TN	FP	FN
Video 1	15	20	0	5
	ACC = 87.5 %			
Video 2	18	10	10	2
	ACC = 70 %			
Video 3	15	13	7	5
	ACC = 70 %			
Video 4	16	16	4	4
	ACC = 80 %			

Thời gian xử lý: Trong nghiên cứu này, tác giả đã thực nghiệm nhận dạng đối tượng offline từ các đoạn dữ liệu trên môi trường Matlab với thiết bị phần cứng là máy tính cá nhân với cấu hình trình bày bên trên. Các frame ảnh từ video đầu vào được điều chỉnh kích thước lại [480x320]. Thời gian nhận dạng các đối tượng trong ảnh xấp xỉ 0,03 giây/ 1 frame.

Bảng 2. So sánh độ chính xác lý của phương pháp sử dụng đặc trưng ACF và phương pháp dùng Adaboost kết hợp đặc trưng Haar liked

Data	Phương pháp dùng ACF	Phương pháp dùng Adaboost kết hợp Haar liked
Video 1	87.5 %	85.8%
Video 2	70 %	69%
Video 3	70 %	73.4%
Video 4	80 %	78%

Khi so sánh độ chính xác của phương pháp dùng đặc trưng ACF và phương pháp dùng bộ phân loại Adaboost kết hợp đặc trưng Haar like trên một số tập dữ liệu thực được khảo sát dễ dàng nhận thấy rằng đối với những đối tượng có cấu trúc đối xứng và hình khối độ chính xác nhận dạng dùng đặc trưng Haar like chiếm ưu thế, trong những trường hợp còn lại, kết quả nhận dạng khi dùng đặc trưng ACF cho kết quả tốt hơn.

4. KẾT LUẬN

Nghiên cứu này đã đề xuất phương pháp và thực nghiệm nhận dạng các đối tượng trong môi trường ngoài trời phức tạp. Tác giả đã áp dụng được kỹ thuật ACF trong nhận dạng một số đối tượng tham gia giao thông phổ biến trên các tập dữ liệu đa dạng với điều kiện khác nhau. Đề tài cũng đã sử dụng kỹ thuật mắt chim đảo tầm nhìn của hình ảnh – IPM để phát hiện làn đường. Đề tài cũng dự đoán được khoảng cách từ camera đơn đến đối tượng phía trước. Nghiên cứu cũng đã trình bày cách sử dụng Camera Calibration Toolbox trong việc hiệu chỉnh camera và Image Label Toolbox để tạo các tập dữ liệu huấn luyện. Ngoài ra, đề tài cũng sử dụng thuật toán RANSAC trong việc loại bỏ nhiễu và sử dụng thông số ROC để đánh giá độ chính xác.

Tuy nhiên, nghiên cứu vẫn còn một số hạn chế cần cải tiến trong thời gian sắp tới. Nghiên cứu chưa thực hiện trên các tập dữ liệu ban đêm hay đường đi dưới bóng cây (trong điều kiện thiếu ánh sáng). Dữ liệu huấn luyện còn hạn chế ảnh hưởng đến độ chính xác trong quá trình nhận dạng. Phần dự đoán khoảng cách chỉ sử dụng camera đơn nên cần sử dụng stereo camera để có kết quả chính xác hơn. Dựa vào những hạn chế này, đề tài có thể định hướng được hướng cải thiện và phát triển cho những nghiên cứu tiếp theo áp dụng trong lĩnh cảnh báo tai nạn, giao thông thông minh.

TÀI LIỆU THAM KHẢO

- [1] Yongzheng Xu, Guizhen Yu, Yunpeng Wang, Xinkai Wu, and Yalong Ma, “A Hybrid Vehicle Detection Method Based on Viola-Jones and HOG + SVM from UAV Images”, Sensors, 2016.
- [2] Nur Shazwani Aminuddin, Masrullizam Mat Ibrahim, Nursabillilah Mohd Ali, Syafeeza Ahmad Radzi, Wira Hidayat Mohd Saad & Abdul Majid Darsono, “A new approach to highway lane detection by using HOUGH transform technique”, Faculty of Electronic and Computer Engineering - Universiti Teknikal Malaysia Melaka, Malaysia, 2017.
- [3] Zhaojin Zhang, Cunlu Xu and Wei Feng, “Road Vehicle Detection and Classification based on Deep Neural Network”, IEEE, 2016.
- [4] Jeffrey de Deijn, “Automatic Car Damage Recognition using Convolutional Neural Networks”, Vrije Universiteit Amsterdam, 2018.

- [5] Qingpeng Li, Lichao Mou, Qizhi Xu, Yun Zhang and Xiao Xiang Zhu, “*R³-Net: A Deep Network for Multi-oriented Vehicle Detection in Aerial Images and Videos*”, IEEE, 2018.
- [6] Daniel Neumann, Tobias Langner, Fritz Ulbrich, Dorothee Spitta and Daniel Goehring, “*Online Vehicle Detection using Haar-like, LBP and HOG Feature based Image Classifiers with Stereo Vision Preselection*”
- [7] Bin Yang, Junjie Yan, Zhen Lei, Stan Z. Li, “*Aggregate Channel Features for Multi-view Face Detection*”, Center for Biometrics and Security Research & National Laboratory of Pattern Recognition Institute of Automation, Chinese Academy of Sciences, China.
- [8] Piotr Doll, Ron Appel, Serge Belongie, and Pietro Perona. “*Fast Feature Pyramids for Object Detection*, Submission to IEEE transactions on Pattern Analysis and Machine Intelligence.
- [9] Ruyi Jiang, Mutsuhiro Terauchi, Reinhard Klette, Shigang Wang, and Tobi Vaudrey, “*Low-level Image Processing for Lane Detection and Tracking*”, Shanghai Jiao Tong University, Shanghai, China, Hiroshima International University, Japan, The University of Auckland, Auckland, New Zealand, pp. 3-7, 2009.
- [10] Piotr Dollar, Ron Appel, Serge Belongie, and Pietro Perona. “*Fast Feature Pyramids for Object Detection*”, IEEE, pp. 9-10, 2014.

Tác giả chịu trách nhiệm bài viết:

Lê Mỹ Hà

Trường Đại học Sư phạm Kỹ thuật TP.HCM

Email: halm@hmute.edu.vn