|  |
| --- |
| HỌC VIỆN KỸ THUẬT MẬT MÃ  **KHOA CÔNG NGHỆ THÔNG TIN**  ¯¯¯¯¯¯¯¯¯¯¯¯¯¯¯¯  Logo HvKTMM |
| **BÁO CÁO MÔN HỌC**  **THỰC TẬP CƠ SỞ** |
| **Đề tài:**  **Phát triển xe tự hành dựa trên học máy**  **Sinh viên thực hiện:**  Đỗ Võ Hữu Tùng - CT040153  Nguyễn Xuân Văn - CT040154  Nhóm 8  **Giảng viên hướng dẫn:**  ThS. Phạm Văn Hưởng |
| Hà Nội, 2022 |

MỤC LỤC

[MỤC LỤC 2](#_Toc126180026)

[DANH MỤC HÌNH ẢNH i](#_Toc126180027)

[DANH MỤC BẢNG BIỂU ii](#_Toc126180028)

[LỜI MỞ ĐẦU iii](#_Toc126180029)

[CHƯƠNG 1: CƠ SỞ LÝ THUYẾT XỬ LÝ ẢNH VÀ NHẬN DẠNG 1](#_Toc126180030)

[1.1 Tổng quan về xử lý ảnh 1](#_Toc126180031)

[1.1.1 Khái niệm về xử lý ảnh 1](#_Toc126180032)

[1.1.2 Các bước cơ bản trong xử lý ảnh 1](#_Toc126180033)

[1.2 Một số khái niệm cơ bản về ảnh 2](#_Toc126180034)

[1.2.1 Ảnh và điểm ảnh 2](#_Toc126180035)

[1.2.2 Mức xám 2](#_Toc126180036)

[1.2.3 Độ phân giải của ảnh 2](#_Toc126180037)

[1.2.4 Ngưỡng 3](#_Toc126180038)

[1.3 Các bài toán thường gặp trong xử lý ảnh 4](#_Toc126180039)

[1.3.1 Khử nhiễu 4](#_Toc126180040)

[1.3.2 Trích xuất đặc trưng 4](#_Toc126180041)

[1.3.3 Nhận dạng 4](#_Toc126180042)

[CHƯƠNG 2: ỨNG DỤNG XỬ LÝ ẢNH VÀ THUẬT TOÁN SUPPORT VECTOR MACHINE TRONG BÀI TOÁN PHÁT HIỆN VÀ PHÂN LỚP BIỂN BÁO GIAO THÔNG ĐƯỜNG BỘ 6](#_Toc126180043)

[2.1 Tổng quan về hệ thống phát hiện và phân lớp biển báo giao thông 6](#_Toc126180044)

[2.2 Phát hiện biển báo bằng thuật toán xử lý ảnh 7](#_Toc126180045)

[2.2.1 Phân đoạn ảnh 7](#_Toc126180046)

[2.2.2 Phát hiện vùng ứng viên 8](#_Toc126180047)

[2.3 Trích xuất đặc trưng HOG 8](#_Toc126180048)

[2.3.1 Tổng quan về đặc trưng HOG 8](#_Toc126180049)

[2.3.2 Quy trình trích xuất đặc trung HOG trên ảnh 9](#_Toc126180050)

[2.4 Thuật toán Support Vector Machine (SVM) 11](#_Toc126180051)

[2.4.1 Siêu phẳng (hyperplane) 11](#_Toc126180052)

[2.4.2 Khoảng các từ một điểm đến siêu phẳng 12](#_Toc126180053)

[2.4.3 Siêu phẳng tối ưu 13](#_Toc126180054)

[CHƯƠNG 3: THỰC NGHIỆM 18](#_Toc126180055)

[3.1 Xử lý hình ảnh - phát hiện làn đường 18](#_Toc126180056)

[KẾT LUẬN 21](#_Toc126180057)

[TÀI LIỆU THAM KHẢO 22](#_Toc126180058)

DANH MỤC HÌNH ẢNH

[Hình 1.1 Các mục đích chính của xử lý ảnh 1](#_Toc126179410)

[Hình 1.2 Các giai đoạn cơ bản của chương trình xử lý ảnh 1](#_Toc126179411)

[Hình 1.3 Ảnh gốc (trái) và ảnh phân ngưỡng (phải) 3](#_Toc126179412)

[Hình 2.1 Quy trình tổng quát của quá trình phát hiện và phân loại biển báo 6](#_Toc126179413)

[Hình 2.2 Ảnh RGB và ảnh HSV 7](#_Toc126179414)

[Hình 2.3 Ảnh đen - trắng 7](#_Toc126179415)

[Hình 2.4 Phát hiện vùng ứng viên dựa vào tọa độ và cắt được hình ảnh của biển báo. 8](#_Toc126179416)

[Hình 2.5 Dữ liệu linearly và non linearly separable 13](#_Toc126179417)

[Hình 2.6 Hai siêu phẳng thỏa mãn rằng buộc. 14](#_Toc126179418)

[Hình 2.7 z0 là một điểm thuộc H1 16](#_Toc126179419)

DANH MỤC BẢNG BIỂU

**No table of figures entries found.**

LỜI MỞ ĐẦU

Trong những năm gần đây trên thế giới, đi cùng với sự phát triển cùa quy mô sản xuất là sự bùng nổ trong nghiên cứu và phát triển các hệ thống robot di động tự trị (autonomous mobile robot) có khả năng ứng dụng được vào những công việc khác nhau trong kho bãi, nhà xưởng. Robot di động là một máy tự động có khả năng di chuyển đến đích trong một môi trường nào đó. Đây là điều khác với các robot công nghiệp (như tay máy có khớp quay và đầu kẹp) được gắn với một không gian cố định. Khởi đầu bằng các xe vận tải tự động AGV (autonomous guided vehicles) đã được ứng dụng rất thành công trong công nghiệp, các robot di động thông minh hiện nay vẫn đang trong thời kỳ phát triển và được coi là có tiềm năng rất lớn trong tương lai gần. Các tiến bộ công nghệ mới trong việc phát triển các thiết bị cảm biến (sensor) và khả năng tính toán cùa các hệ xử lý đã thúc đẩy mức độ tự trị trong sự vận hành các robot di động. Mặt khác, ứng dụng của robot di động trong các môi trường khác như dịch vụ, giãi trí, y tế, an ninh, quân sự cũng hứa hẹn có những tiến bộ quan trọng trong quá trình thiết kế và phát triển các hệ thống này.

Không kể hoạt động của các bộ phận gắn trên robot di động, bài toán dẫn đường cho sự di chuyển của robot từ một điểm xuất phát tới đích một cách an toàn, được gọi tắt là “điều hướng cho robot di động”, là bài toán chính yếu trong các nghiên cứu về robot di động hiện nay. Không khác nhiều so với hành vi của con người trong việc điều hướng, muốn giải quyết được bài toán này thì robot phải tự xác định được vị trí cùa mình trong môi trường (positioning), xác lập được bản đồ môi trường nếu cần thiết (mapping), vạch ra được quỹ đạo đi tới đích (path planning) và xuất ra cách thức điều khiển đi trên quỹ đạo (path control) và tránh vật cản (obstacle avoidance) trên đường đi. Để làm được điều đó, một hệ thống điều hướng của robot di động thông thường có thê chia thành 4 khối [1]: hệ thống cảm biến, hệ thống nhận thức, hệ thống lập kế hoạch di chuyển và hệ thống điều khiển động cơ. Robot di động cảm nhận về môi trường bên ngoài bằng nhiều cảm biến khác nhau gắn trên nó. Dữ liệu từ cảm biến được hệ thống nhận thức xử lý, kết hợp thành các thông tin có ý nghĩa. Hệ thống lập kế hoạch di chuyển sẽ sử dụng những thông tin đó để đưa ra đường đi tối ưu đến vị trí đích. Cuối cùng, cơ cấu chấp hành là hệ thống động cơ, chịu trách nhiệm đưa chiếc xe di chuyển theo đúng kế hoạch đã lập ra..

Xe tự hành là một trường hợp đặc biệt của robot tự hành, khi mà môi trường hoạt động của nó không chỉ gói gọn ờ khu vực nhà xưởng (in-door). Phát triển xe tự hành theo hướng đã nêu có thể giúp chiếc xe hoạt động chính xác ở các môi trường đó, nơi có rất nhiều vách ngăn, tường bao để các cảm biến có thể thu thập thông tin. Tuy nhiên, trong môi trường đường phố thực tế (out-door), chiếc xe cần tuân thủ theo các chỉ dẫn làn đường và biển báo giao thông, cần nhận biết các đối tượng giao thông khác như xe cộ, người đi đường một cách chính xác ở khoảng cách an toàn. Do vậy phương pháp kết hợp cảm biến cho các công việc này trờ nên không khả thi. Đây chính là lý do khiến dữ liệu hình ảnh trở thành một nguồn thông tin vô cùng hữu ích để chiếc xe có thể thực sự quan sát về môi trường cùa mình [2] [3].

Chính vì thế chúng em quyết định chọn để tài “**Phát triển xe tự hành dựa trên học máy”** nghiên cứu về lĩnh vự xử lý ảnh, ứng dụng các phương pháp xử lý ảnh kết hợp cùng mô hình phân lớp SVM (Support Vector Machine) trong công việc phát hiện và nhận dạng biển báo giao thông đường bộ. Triển khai lập trình hệ thống thử nghiệm dựa trên nền tảng ngôn ngữ lập trình Python và các thư viện hỗ trợ liên qua.

Trong quá trình làm báo cáo, do kiến thức còn hạn chế và thiếu nhiều kinh nghiệm nên không thể tránh khỏi những sai sót, chúng em rất mong nhận được sự đóng góp của các thầy, cô và các bạn để báo cáo được hoàn thiện hơn.

**Nhóm sinh viên thực hiện**

# CƠ SỞ LÝ THUYẾT XỬ LÝ ẢNH VÀ NHẬN DẠNG

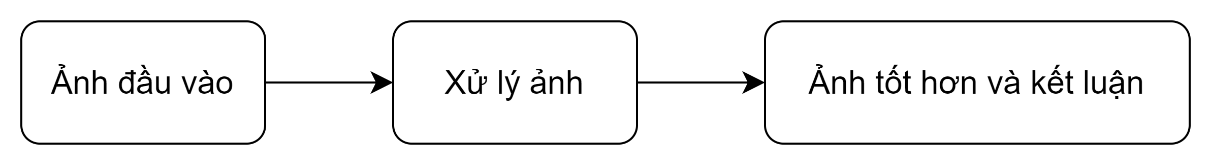
## Tổng quan về xử lý ảnh

### Khái niệm về xử lý ảnh

Xử lý ảnh là một trong những mảng nghiên cứu quan trọng nhất trong lĩnh vự thị giác máy tính (Computer Vision), là tiền đề cho nhiều ứng dụng mới thuộc lĩnh vực này. Hai nhiệm vụ cơ bản của quá trình xử lý ảnh là nâng cao chất lượng thông tin hình ảnh va xử lý số liệu cung cấp cho các quá trình khác trong đó có việc ứng dụng thị giác máy tính vào điều khiển.

Con người thu nhận thông tin qua các giác quan trong đó thị giác đóng vai trò quan trọng nhất. Sự phát triển nhanh của phần cứng máy tính, xử lý ảnh và đồ họa đã phát triển mạnh mẽ và ngày càng có nhiều ứng dụng trong cuộc sống. Xử lý ảnh đóng một vai trò quan trọng trong tương tác người - máy.

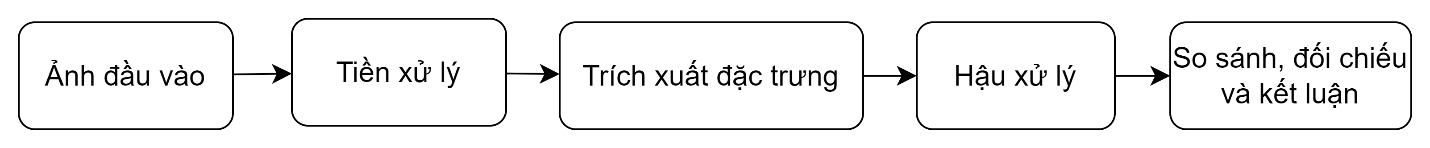
Quá trình xử lý nhận dạng ảnh là một quá trình thao tác nhằm biến đổi một ảnh đầu vào để cho ra một kết quả mong muốn kết quả đầu ra của quá trình xử lý ảnh có thể là một ảnh chất lượng tốt hơn hoặc một kết luận.



Hình 1.1 Các mục đích chính của xử lý ảnh

### Các bước cơ bản trong xử lý ảnh

Quá trình xử lý một ảnh đầu vào nhằm thu được một ảnh đầu ra mong muốn thường phải trải qua rất nhiều bước khác nhau. Các bước cơ bản của một quá trình xử lý ảnh được thể hiện thông qua hình sau:



Hình 1.2 Các giai đoạn cơ bản của chương trình xử lý ảnh

Dựa vào sơ đồ trên ta có thể thấy quá trình xử lý ảnh thông qua 5 bước chính:

* Thu thập ảnh đầu vào
* Tiền xử lý
* Trích xuất đặc trưng
* Hậu xử lý
* So sánh, đối chiếu và kết luận

## Một số khái niệm cơ bản về ảnh

### Ảnh và điểm ảnh

Điểm ảnh (pixel) được xem như là dấu hiệu hay cường độ sáng tại một tạo đột rong không gian của đối tượng.

Ảnh đuọc xem như là một tập hợp các điểm ảnh, có thể coi ảnh là mảng số thực 2 chiều Im\*n, có kích thước m\*n bao gồm m\*n điểm ảnh, trong đó mỗi điểm ảnh có một giá trị xác định biều thị mức xám tại vị trí m\*n tương ứng.

### Mức xám

Một điểm ảnh có hai đặc trung cơ bản là vị trí (x, y) của điểm ảnh và độ xám của nó. Dưới đây chúng ta xem xét một số khái niệm và thuật ngữ thường dung xử lý ảnh.

Mức xám của điểm ảnh là cường độ sáng của nó được gán bắng giá trị số tại điểm đó. Các thang giá trị mức xám thông thường: 16, 32, 64, 128, 256 (mức 256 là mức phổ dụng. Lý do là kỹ thuật máy tính dùng 1 byte (8 bit) để biểu diễn mức xám. Mức xám dùng 1 byte biểu diễn: 28 = 256 mức, tức là từ 0-255).

### Độ phân giải của ảnh

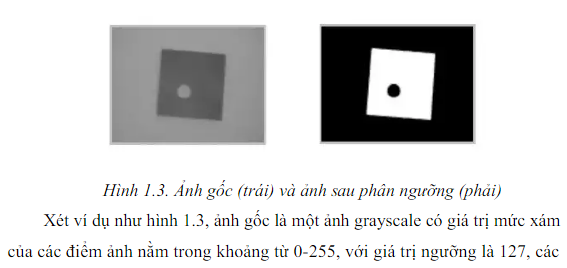
Độ phân giải của ảnh là mật độ điểm ảnh được ấn định trên ảnh được hiển thị. Khoảng cách giữa các điểm ảnh phải được chọn sao cho mắt người vẫn thấy được sự liên tục của ảnh. Việc lựa chọn khoảng cách thích hợp tạo nên một mật độ phân bố, đó chính là độ phân giải và được phân bố theo trục x và y trong không gian hai chiều.

### Ngưỡng

Ngưỡng là một khái niệm quen thuộc trong xử lý ảnh cũng như rất nhiều giải thuật khác. Ngưỡng được dùng khá hiệu quả trọng việc nhận dạng và tìm biên ảnh. Nó dùng để chỉ một giá trị mà người ta dựa vào để phân hoạch một tập hợp thành các miền phân biệt. Trong xử lý ảnh các vùng có giá trị thỏa mãn ngưỡng thường được quy về một giá trị hay nói cách khác là giá trị mức xám tại điểm xét thỏa mãn sẽ được phân hoạch vào vùng còn lại.

Giá trị ngưỡng thường được xác định dựa vào những điểm đặc biệt hoặc dựa vào kinh nhiệm khỏa sát. Nếu dựa vào số lượng ngưỡng áp dụng cho cùng một tập dữ liệu người ta sẽ phân ra các phương pháp ứng dụng ngưỡng đơn, ngưỡng kép, hay đa ngưỡng.

Nếu dựa vào sự biến thiên của giá trị ngưỡng, trong cùng phạm vi ứng dụng người ta sẽ phân ra các phương pháp dùng ngưỡng cố định và không cố định. Ngưỡng không cố định nghĩa là giá trị của nó sẽ thay đổi tùy theo sự biến thiên của tập dữ liệu theo không gian và thời gian. Thông thường giá trị này được xác định thông qua khảo sát tập dữ liệu bằng phương pháp thống kê.



Hình 1.3 Ảnh gốc (trái) và ảnh phân ngưỡng (phải)

Xét ví dụ như hình 1.3, ảnh gốc là một ảnh grayscale có giá trị mức xám của các điểm ảnh nằm trong khoảng từ 0-255, với giá trị ngưỡng là 127, các điểm ảnh nào có giá trị lớn hơn ngưỡng ta cho nó thành màu đen (0), ngược lại là trắng (1) và cho ảnh kết quả.

## Các bài toán thường gặp trong xử lý ảnh

### Khử nhiễu

Có 2 loại nhiều cơ bản trong quá trình thu nhận ảnh:

* Nhiễu hệ thống: Nhiễn có quy luật và có thể khử bằng các phép biến đổi.
* Nhiễu ngẫu nhiên: nhiễu không rõ nguyên nhân, điều này được khắc phục bằng các phép lọc.

### Trích xuất đặc trưng

Các đặc điểm của đối tượng được trích xuất tùy theo mục đích nhận dạng trong quá trình xử lý ảnh. Có thể nêu ra một số đặc điểm của ảnh sau đây:

* Đặc điểm không gian: Phân bổ mức xám, phân bổ xác suất, phân bổ biên độ.
* Đặc điểm biến đổi: Các đặc điểm loại này được trích xuất bằng việc thực hiện lọc vùng (zonal filtering). Các bộ vùng được gọi là mặt nạ đặc trưng (feature mask) thường là các khe hẹp với hình dạng khác nhau (chữ nhật, tam giác, cung tròn v.v..).
* Đặc điểm biên và đường biên: đặc trung cho đường biên của đối tượng và do vậy rất hữu ích trong việc trích xuất các thuộc tính bất biến được dùng khi nhận dạng đối tượng. Các đặc điểm này có thể được trích xuất nhờ toán từ gradient, toán tử la bàn, toán tử Laplace, toán tử “chéo không” (zero crossing),…
* Việc trích xuất chọn hiệu quả các đặc điểm giúp chọn nhận dạng các đối tượng ảnh chính xác, với tốc độ tính toán cao và giảm dung lượng bộ nhớ lưu trữ.

### Nhận dạng

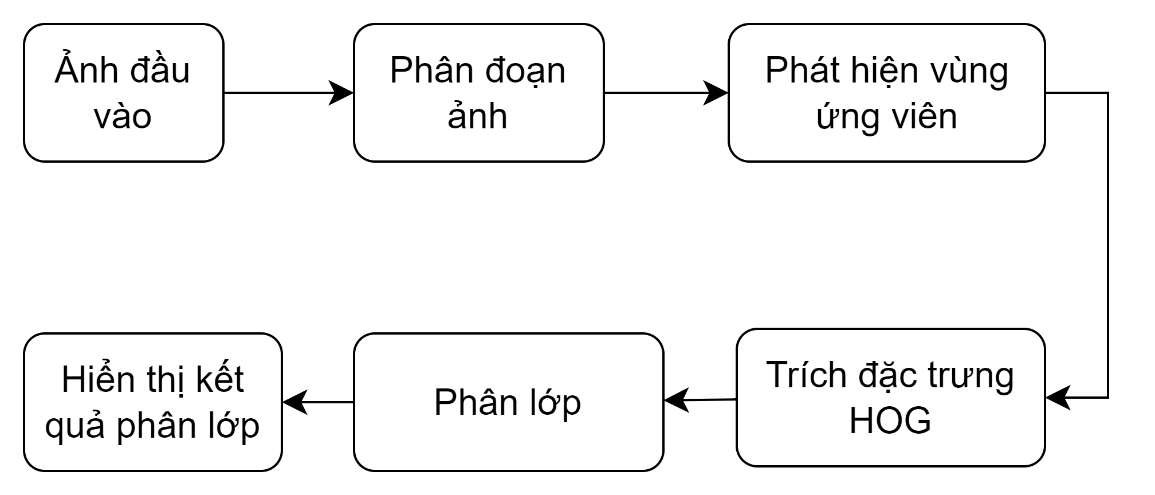
Nhận dạng tự động mô tả đối tượng, phân loại và phân nhóm các mẫu là nhưng vẫn đề quan trọng trong thị giác máy tính, được ứng dụng trong nhiều ngành khoa học khác nhau. Quá trình nhận dạng thường đi kèm với mẫu. Mẫu có thể là ảnh của vân tay, ảnh của một vật nào đó được chụp, một chữ viết, khuôn mặt người hoặc một ký đồ tín hiệu tiếng nói. Khi biết một mẫu nào đó, để nhận dạng hoặc phân loại mẫu đó, ta có thể làm theo hai hướng:

* Phân loại có mẫu: chẳng hạn phân tích phân biệt trong đó mẫu đầu vào được định danh như một thành phần của một lớp đã xác định, khi đó ta chỉ cần so sánh đối tượng đầu vào với mẫu để đưa ra kết luận.
* Phân loại không có mẫu: trong đó các mẫu được gán vào các lớp khác nhau dựa trên một tiêu chuẩn đồng dạng nào đó. Các lớp này cho đến thời điểm phân loại vẫn chưa biết hay chưa được định danh. Việc giải quyết bài toán nhận dạng trong những ứng dụng mới, nảy sinh trong cuộc sống không chỉ tạo ra những thách thức về thuật giải, mà còn đặt ra những yêu cầu về tốc độ tính toán. Đặc điểm chung của tất cả những ứng dụng đó là cần nhiều những đặc trưng cần thiết, không thể do chuyên gia đề xuất mà phải được trích chọn dựa trên các thủ tục phân tích dữ liệu.

# ỨNG DỤNG XỬ LÝ ẢNH VÀ THUẬT TOÁN SUPPORT VECTOR MACHINE TRONG BÀI TOÁN PHÁT HIỆN VÀ PHÂN LỚP BIỂN BÁO GIAO THÔNG ĐƯỜNG BỘ

## Tổng quan về hệ thống phát hiện và phân lớp biển báo giao thông

Quy trình xử lý tổng quá của phương pháp được trình bày như trong hình 2.1.



Hình 2.1 Quy trình tổng quát của quá trình phát hiện và phân loại biển báo

Đầu tiên, ảnh đầu vào thu được từ camera sẽ được áp dụng các phương pháp phân đoạn để loại bỏ nền, làm nổi bật các đối tượng có thể là biển báo. Tiếp theo, biên của các đối tượng này sẽ đượng xử lý lọc kích thước và phân tích hình dáng để chọn ra các vùng ứng viên có thể là biển báo giao thông.

Sau đó, các vùng ứng viên sẽ được trích đặc trưng HOG, các đặc trưng này sẽ đượng đưa ra qua mô hình SVM đã được huấn luyện trước đó để nhận dạng, đưa ra kết quả cuối cùng về phân lớp của biển báo này.

Nhìn chung, hệ thống có thể chia làm hai quá trình lớn:

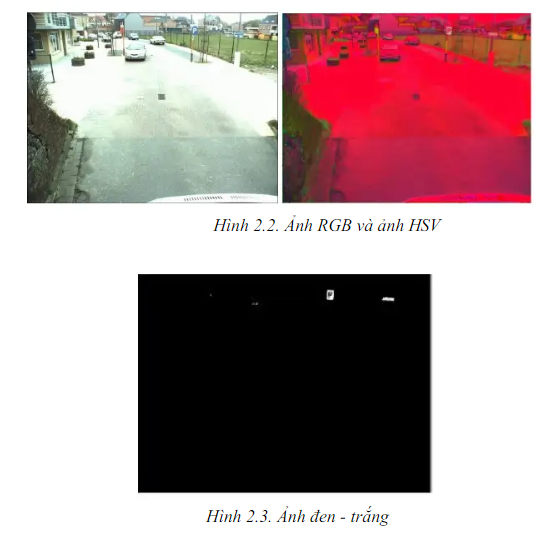
* Phát hiện biển báo trong khung hình sử dụng các thuật toán xử lý ảnh.
* Phân lớp biển báo phát hiện được bằng mô hình SVM.

## Phát hiện biển báo bằng thuật toán xử lý ảnh

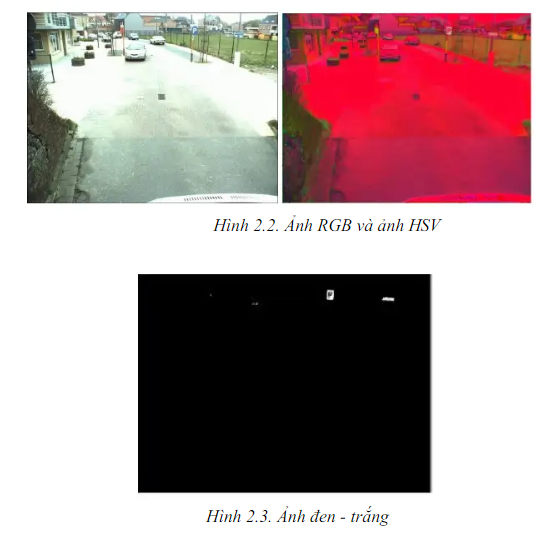
### Phân đoạn ảnh

Phân đoạn ảnh là một trong những bước chính yếu của quá trình phát hiện và phân loại biển báo giao thông, với rất nhiều tham số khác nhau cần tinh chỉnh và chọn lựa.

Hình ảnh kết quả của quá trình cân bằng histogram sẽ được chuyển đổi từ không gian màu RGB sang không gian màu HSV. Sau khi chuyển ảnh sang không gian màu HSV, chúng em xác định các giới hạn trên và dưới của các mặt nạ (mask) trong không gian màu này để lọc ra các vùng ảnh có xuất hiện màu đỏ và màu xanh. Thông số chi tiết của các mặt nạ được thể hiện dưới dạng véc-tơ [H, S, V] được biểu diễn như hình dưới.



Hình 2.2 Ảnh RGB và ảnh HSV



Hình 2.3 Ảnh đen - trắng

### Phát hiện vùng ứng viên

Hình ảnh kết quả từ quá trình xử lý trên được cung cấp cho hàm findContours() của thư viện OpenCV để dò tìm các đường viện (contour) xuất hiện trong ảnh. Mặc dù kích thước của vùng ứng viên biển báo giao thông thưc sự có tỷ lệ w/h (chiều rộng/chiều cao) xấp xỉ 1 nhưng khi truy xuất các khung hình từ camera của xe, tỉ lệ này sẽ thay đổi tùy theo khoảng cách, thời gian và góc nhìn. Tùy theo góc nhìn mà có thể chiều cao sẽ lớn hơn chiều rộng và ngược lại. Thực nghiệm cho thấy tỷ lệ phù hợp cho các vùng ứng viên với các góc nhìn khác nhau thỏa mãn w/h < 1/3 và h/w < 1/7. Vì vậy, sau khi trích được các vùng ứng viên, chúng em sử dụng ràng buộc tỷ lệ chiều rộng w và chiều cao h thỏa w/h < 1/3 và h/w < 1/7 để chọn các vùng ứng viên thực sự.



Hình 2.4 Phát hiện vùng ứng viên dựa vào tọa độ và cắt được hình ảnh của biển báo.

## Trích xuất đặc trưng HOG

### Tổng quan về đặc trưng HOG

HOG viết tắt của Histogram of Oriented Gradient là một phương pháp giúp trừu tượng hóa đối tượng bằng cách trích xuất ra những đặc trưng của đối tượng đó và bỏ đi những thông tin không hữu ích. Vì vậy, HOG được sử dụng chủ yếu để mô tả hình dạng và sự xuất hiện của một đối tượng trong ảnh.

Bản chất của phương phát HOG là sử dụng thông tin về sự phân bổ của các cường độ gradient (intensity gradient) hoặc của hướng biên (edge directions) để mô tả các đối tượng cục bộ trong ảnh. Các toán tử HOG được cài đặt bằng cách chia nhỏ một bức ảnh thành các vùng được gọi là cells và với mỗi cell ta sẽ tính toán một histogram về các hướng của gradients cho các điểm nằm trong cell. Ghép các histogram lại với nhau ta sẽ có một biểu diễn cho bức ảnh ban đầu. Để tăng cường hiệu quả nhận dạng, các histogram cục bộ có thể được chuẩn hóa về độ tương phản bằng cách tính một ngưỡng cường độ trong một vùng lớn hơn cell, gọi là các khối (blocks) và sử dụng giá trị ngưỡng đó để chuẩn hóa tất cả các cell trong khối. Kết quả sau bước chuẩn hóa sẽ là một vector đặc trưng có tính bất biến cao hơn đối với các thay đổi về điều kiện ánh sáng.

Đặc trưng HOG cho phép miêu tả tốt các loại biển báo giao thông có hình dạng khác nhau. Vì vậy chúng em sử dụng đặc trưng HOG để trích xuất đặc trung của biển báo trên vùng ứng viên đã được tìm ra ở phần trước để phục vụ cho thao tác nhận dạng.

### Quy trình trích xuất đặc trung HOG trên ảnh

**Bước 1:** Tính cường độ và hướng biến thiên tại mỗi pixel theo công thức (2.4), (2.5).

Cường độ:  (2.4)

Hướng:  (2.5)

Trong đó,  và  là giá trị pixel (x, y).

**Bước 2:** Phân chia ảnh

Chia ảnh đầu ra ở bước trên thành nhiều khối (block), mối khối có số cell bằng nhau, mỗi cell có số pixels bằng nhau. Số khối được tính bằng công thức (2.6). Trong đó: Wimage, Himage, Wblock, Hblock, Wcell, Hcell lần lượt là chiều rộng, chiều cao của ảnh, khối và ô.

 (2.6)

**Bước 3:** Tính vector đặc trung cho từng khối

Sau khi xác định số block và kích thước mỗi block, cell, để tính toán vector đặc trưng cho từng cell, chúng ta cần:

* Chia không gian hướng thành p bin (số chiều vector đặc trưng của ô).
* Rời rạc hóa góc hướng nghiêng tại mỗi điểm ảnh vào trong các bin với độ lớn .

Trường hợp rời rạc hóa unsigned-HOG với p=9:

 (2.7)

Trường hợp rời rạc hóa signed-HOG với p=18:

 (2.8)

Giá trị bin được định lượng bởi tổng cường độ biến thiên của các pixels thuộc về bin đó. Sau khi tính toán đặc trưng ô, ta sẽ nối các vector đặc trưng ô để thu được vector đặc trưng khối. Số chiều vector đặc trưng khối tính theo công thức:

 (2.9)

Trong đó: n là số ô trong khối và size feature/cell là số chiều vector đặc trưng của ô bằng 9 (unsigned-HOG) hoặc 18 (signed-HOG).

**Bước 4:** Chuẩn hóa khối

Để tăng cường hiêu năng nhận dạng, các histogram cục bộ sẽ được chuẩn hóa về độ tương phản bằng cách tính một ngưỡng cường độ trong một khối và sử dụng giá trị đó để chuẩn hóa tất cả các ô trong khối. Kết quả sau bước chuẩn hóa sẽ là một vector đặc trưng có tính bất biến cao hơn đối với các thay đổi về điều kiện ánh sáng.

Có nhiều phương pháp có thể được dùng để chuẩn hóa khối. v là vector đặc trưng ban đầu của khối, vk  là k-norm của v (k = 1, 2), e là hằng số nhỏ. Khi đó, các giá trị chuẩn hóa có thể tính bằng một trong các công thức sau:

L2-norm:  (2.10)

L1-norm:  (2.11)

L1-sqrt:  (2.12)

Ghép các vector đặc trưng khối sẽ thu được vector đặc trưng HOG cho ảnh. Số chiều vector đặc trưng ảnh tính theo công thức với n là số khối của ảnh, sizeblock là số chiều của vector đặc trưng khối:

 (2.13)

## Thuật toán Support Vector Machine (SVM)

Support Vector Machine (SVM) là phương pháp phân tích lớp dựa trên lý thuyết thống kê của Vapnik và Alexei Chervonenkis xây dựng vào năm 1960 và dược sử dụng nhiều trong các ứng dụng nhận dạng chữ viết tay, nhân dạng khuôn mặt, phân loại tài liệu, tin sinh học… So với các phương pháp phân loại khác, khả năng phân loại của SVM là tương đương hoặc tốt hơn đáng kể [5]. Trong dự án này, nhóm chúng em sẽ sử dụng phương pháp phân loại SVM để phân loại biển báo giao thông.

### Siêu phẳng (hyperplane)

Chúng ta có thể hiểu siêu phẳng là gì thông qua việc nghiên cứu khái niệm đường thẳng.

Trong không gian 2 chiều, phương trình đường thẳng có dạng:

Nếu ta thay *y = x2*và *x = x1*, phương trình trở thành:

Tương đương với:

Giả sử vector **x** = (*x1, x2*) và **w** = (*a, -1*), tích vô hướng **w** và **x** là

Thay vào phương trình ta được:

Phương trình trình trên là phương trình của một siêu phẳng. Từ phương trình này, chúng ta có thể có một cái nhìn sâu sắc khác về siêu phẳng là gì: nó là tập hợp các điểm thỏa mãn .

Chúng ta có thể suy ra phương trình siêu phẳng từ phương trình của một đường thẳng, đó là bởi vì một đường thẳng cũng là một siêu phẳng. Ta sẽ nhận thấy rằng, trên thực tế, một đường thẳng là một không gian hai chiều được bao quanh bởi một mặt phẳng có ba chiều. Tương tự, điểm và mặt phẳng cũng là siêu phẳng. Lợi ích của siêu phẳng là nó giúp ta làm việc dễ dàng hơn khi không gian có lớn hơn 2 chiều. Vector **w**  sẽ luôn là **pháp tuyến** của siêu phẳng

### Khoảng các từ một điểm đến siêu phẳng

Trong không gian 2 chiều, ta biết rằng khoảng cách từ một điểm có toạ độ (*x0, y0*) tới đường thẳng có phương trình được xác định bởi:

Trong không gian ba chiều, khoảng cách từ một điểm có toạ độ (*x0, y0, z0*) tới một mặt phẳng có phương trình được xác định bởi:

Việc này có thể được tổng quát lên không gian nhiều chiều: Khoảng cách từ một điểm (vector) có toạ độ x0 tời siêu phẳng có phương trình được xác định bởi:

Với với d là số chiều của không gian.

### Siêu phẳng tối ưu

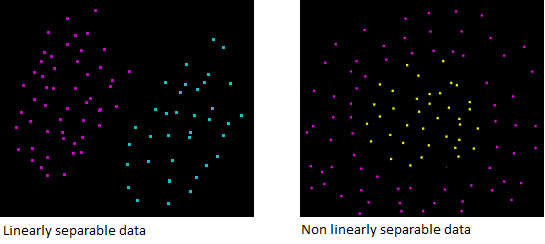
Để phân chia hai lớp dữ liệu, rõ ràng là có rất nhiều siêu phẳng có thể làm được điều này. Mặc dù vậy, mục tiêu của chúng ta là tìm ra siêu phẳng có lề rộng nhất tức là có khoảng cách tới các điểm của hai lớp là lớn nhất.

Điều này có thể được thức hiện thông qua các bước sau:

**Bước 1:** Ta có một tập dữ liệu cần phải phân loại có dạng

Với mang giá trị 1 hoặc -1, xác định lớp của điểm **xi**. Mỗi **xi** làmột vector thực p-chiều. Ta cần tìm siêu phẳng có lề lớn nhất chia các điểm yi = 1 và các điểm yi = -1.

**Bước 2:**  Chọn 2 siêu phẳng phân chia tập dữ liệu sao cho không có điểm nào nằm giữa chúng. Ta chỉ có thể làm điều này nếu tập dữ liệu là linearly separable.



Hình 2.5 Dữ liệu linearly và non linearly separable

Vì vậy, ta giả sử rằng tập dữ liệu là linearly separable.

Cho một siêu phẳng tách tập dữ liệu và thỏa mãn:

Ta chọn 2 siêu phẳng và khác có các phương trình sau:

Và

Ta thấy cách đều và . Vì hằng số không quá quan trọng nên ta đặt để cho đơn giản.

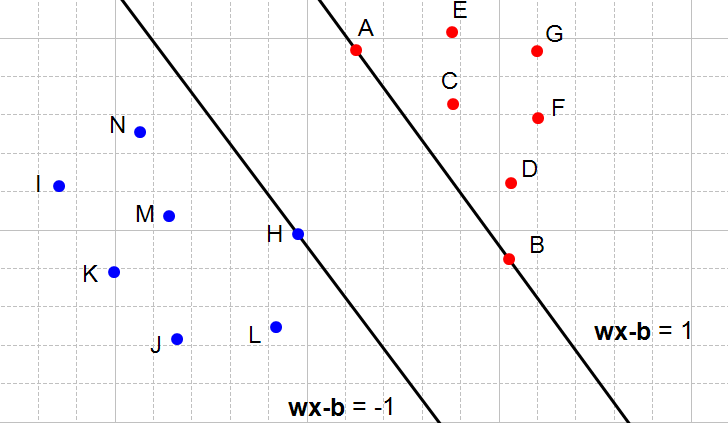
Và

Bây giờ để đảm bảo không có điểm nào nằm giữa chúng, ta sẽ chỉ lựa chọn cách siêu phẳng thỏa mãn điều kiện:

(2.14)

hoặc

(2.15)



Hình 2.6 Hai siêu phẳng thỏa mãn rằng buộc.

Bằng cách xác định các ràng buộc này, ta có thể tìm ra 2 siêu phẳng không có điểm nào nằm giữa chúng. Và điều này có thể được áp dụng với p-chiều.

Trong toán học, mọi người thích mọi thứ được diễn đạt ngắn gọn.

Các phương trình (2.14) và (2.15) có thể được kết hợp thành một ràng buộc duy nhất:

Bắt đầu với phương trình (2.15) khi ta nhân cả 2 vế với *yi*(trong trường hợp này *yi* = -1 vì các phần tử thuộc lớp -1). Phương trình (2.15) được viết thành:

(2.16)

Tương tự với phương trình (2.14), nhưng lúc này *yi* = 1 nên nó không làm thay đổi dấu của bất phương trình.

(2.17)

Kết hợp phương trình (2.16) và (2.17) ta được:

(2.18)

**Bước 3:** Cực đại hóa khoảng cách giữa 2 siêu phẳng

Trước khi cố gắng tối đa hóa khoảng cách giữa hai siêu phẳng, trước tiên chúng ta sẽ tìm cách nào để tính khoảng cách đó.

Giả sử:

* là siêu phẳng có phương trình:
* là siêu phẳng có phương trình:
* là một điểm thuộc siêu phẳng

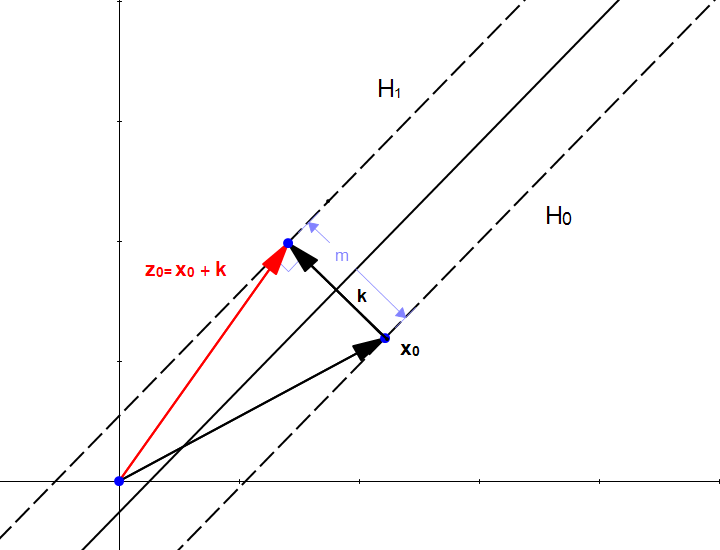
Gọi là khoảng cách vuông góc từ đến siêu phẳng .Theo định nghĩa, chính là margin (lề). Vì thuộc nên chính là khoảng cách giữa 2 siêu phẳng và **.**

Vector đơn vị của vector **w** là: . Nếu ta nhân **u** với  ta sẻ được vector **k** = với:

* (vì **)**
* **k** vuông góc với (vì **k** cùng hướng với **w** - vector pháp tuyến của siêu phẳng )

(2.19)

Ta có: là một điểm thuộc siêu phẳng



Hình 2.7 z0 là một điểm thuộc H1

Vì z0 thuộc  nên:

(2.20)

Thay :

(2.21)

Thay **k**  bằng phương trình (2.19):

(2.22)

(2.23)

Tích vô hướng của một vectơ với chính nó là bình phương của độ dài nên:

(2.24)

(2.25)

(2.26)

Vì **x**0 thuộc  nên:

(2.27)

(2.28)

(2.29)

Bây giờ chúng ta có một công thức để tính toán lợi nhuận:

Biến duy nhất chúng ta có thể thay đổi trong công thức này là độ dài của **w**. Ta có thể thấy, độ dài càng lớn thì margin càng nhỏ. Như vậy cực đại hóa margin cũng chính là cực tiểu độ dài của **w.** Như vậy trong số tất cả các siêu phẳng có thể đáp ứng các ràng buộc, chúng ta sẽ chọn siêu phẳng có nhỏ nhất bởi vì khi đó ta sẽ thu được margin là lớn nhất.

# THỰC NGHIỆM

## Xử lý hình ảnh - phát hiện làn đường

Bước 1: Thresholing

Vì đường cho xe chạy sử dụng giấy trắng A4 để làm nên chúng em chỉ cần sử dụng tính năng phát hiện màu để tìm đường dẫn.

Ý tưởng ở đây chỉ đơn giản là chuyển đổi hình ảnh sang không gian HSV và sau đó áp dụng một dải màu để tìm ra đường ( chọn mọi thứ trừ màu trắng)

Bước 2: Warping Lane

Vì chúng em không muốn xử lý toàn bộ hình ảnh mà camera thu thập được, vỉ em chỉ muốn RasPi biết có khúc cua trên con đường ngay tại thời điểm đó chứ không phải là vài giây phía trước nên em cần cắt hình ảnh mà camera thu thập được. Nhưng hơn thế nữa, em còn muốn PasPi nhìn con đường như thể đang nhìn từ trên xuống. Đây gọi là “bird eye view” và nó rất quan trọng vì nó sẽ cho phép RasPi dễ dàng tìm thấy đường cong hơn.

Để làm điều này trước tiên chúng em xác định các điểm ban đầu. Những điểm này thường được xác định thủ công vì thế em có thể sử dụng các trackbar để thử nghiệm các giá trị khác nhau với ý tưởng là để có được một hình chữ nhật khi đường thẳng.

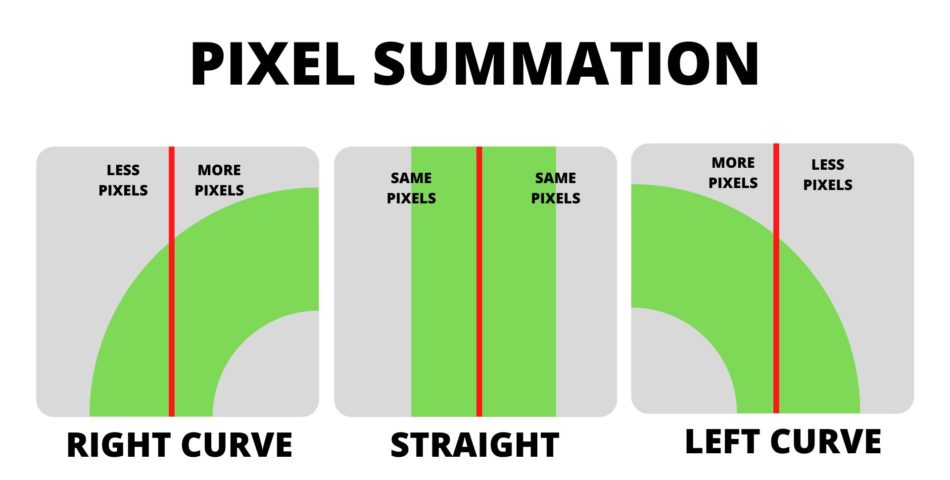
Bước 3: Histogram

Đây là phần quan trọng nhất để tìm ra khúc cua trên đường đi. Ý tưởng là sử dụng tổng các pixel mà các ô là màu trắng theo chiều y

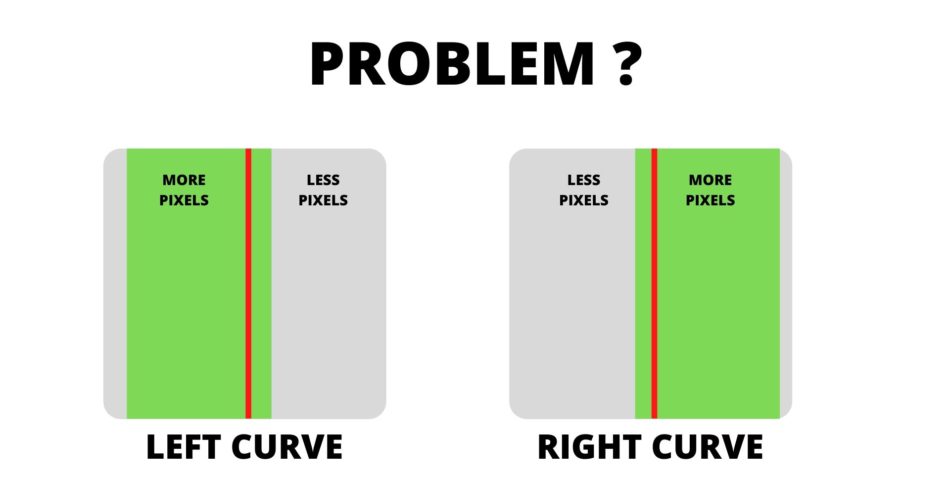
Chart

Description automatically generated

Hình trên hiển thị tất cả các pixel màu trắng có giá trị 255 và tất cả các pixel màu đen có giá trị 0. Bây giờ nếu tính tổng các pixel trong cột đầu tiên thì kết quả là 1275. Cứ như thế ta tính ra được tổng pixel ở từng cột. Vì hình ảnh có kích thước 480 x 240 px nên ta có 480 giá trị. Sau khi tính tổng, ta có thể xem có bao nhiêu giá trị vượt quá ngưỡng nhất định ( do ta chọn) ví dụ như 1000. Như vậy, trong ví dụ trên ta có 8 cột bên trái và 3 cột bên phải có tổng số pixel vượt ngưỡng 1000 suy ra đường cong đang hướng về bên trái.



Tuy nhiên phương pháp này có một số vấn đề:



Ở đây, mặc dù không có đường cong nhưng vì có nhiều pixel hơn một bên, thuật toán sẽ kết luận là đường cong về bên trái hoặc phải. Vì thế để khắc phục chúng em phải điều chỉnh đường trung tâm sau đó mới so sánh các pixel ở cả hai bên.

Bước 4: Tính trung bình

Khi có giá trị đường cong, chúng em sẽ thêm nó vào danh sách có thể tính giá trị trung bình. Tính trung bình sẽ cho phép chuyển động trơn tru hơn.

Bước 5: Hiển thị

KẾT LUẬN

TÀI LIỆU THAM KHẢO

1. Takacs, Arpad & Rudas, Imre & Bosl, Dominik & Haidegger, Tamas. (2018). Highly Automated Vehicles and Self-Driving Cars [Industry Tutorial]. IEEE Robotics & Automation Magazine. 25. 106-112. 10.1109/MRA.2018.2874301.
2. Treml, M., et al.:Speeding up Semantic Segmentaion for Autonomous Driving. Proceedings of the Conference on Neural Information Processing Systems (NIPS), Barcelona, Spain, 1-7 (2016).
3. Yuan, Chang & Chen, Hui & Liu, Ju & Zhu, Di & Xu, Yanyan, Robust Lane Detection for Complicated Road Environment Based on Normal Map. IEEE Access. PP. 1-1. 10.1109/ACCESS.2018.2868976, 2018.
4. Nguyễn Văn Long, “Tìm hiểu và đề xuất phương pháp nhận dạng và phân loại biển báo giao thông ở Việt Nam”, Luận văn Thạc sĩ, Đại học Duy Tân, 2016, tr33-83.
5. Hà Thị Kim Duyên, Lê Mạnh Long, Nguyễn Đức Duy, Phan Sỹ Thuần, Nguyễn Ngọc Hải, Nguyễn Thị Tú Uyên, “NGHIÊN CỨU VÀ PHÁT TRIỂN HỆ THỐNG XE TỰ HÀNH ỨNG DỤNG TRÍ TUỆ NHÂN TẠO”, hội nghị HAUI tổ chức vào tháng 6/2020.
6. Hiep Do Quang, Tien Ngo Manh, Cuong Nguyen Manh, Dung Pham Tien, Manhj Tran Van, Kiem Nguyen Tien, Duy Nguyen Duc, “An Approach to Design Navigation System for Omnidirectional Mobile Robot Based on ROS”, 2020 3rd International Conference on Control, Robotics and Informations (ICCRI 2020).
7. Van Nguyen Thi Thanh, Tien Ngo Manh, Cuong Nguyen Manh, Dung Pham Tien, Manh Tran Van, Duyen Ha Thi Kim, Duy Nguyen Duc, “Autonomous Navigation for Omnidirectional Robot Based on Deep Reinforcement Learing”, 2020 3rd International Conference on Control, Robotics and Informations (ICCRI 2020).
8. Maitre, Henri. (2013). Image Processing: Overview and Perspectives. Intelligent Video Surveillance Systems. 1-12. 10.1002/9781118577851.ch1.
9. Mishra, Manish & Srivastava, Monika. (2014). A view of Artificial Neural Network. 1-3. 10.1109/ICAETR.2014.7012785.
10. O’Shea, Keiron & Nash, Ryan. (2015). An Introduction to Convolutional Neural Networks. ArXiv e-prints.
11. Saadna, Yassmina & Behloul. Ali. (2017). An overview of traffic sign detection and classification methods. International Journal of Multimedia Information Retrieval. 6. 1-18. 10.1007/s13735-017-0129-8.