TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**ĐỒ ÁN 1**

**NGHIÊN CỨU XÂY DỰNG ỨNG DỤNG NHẬN BIẾT TÍN HIỆU ĐÈN GIAO THÔNG QUA CAMERA**

*Người hướng dẫn*: **TS.PHẠM VĂN HUY**

*Người thực hiện*: **NGUYỄN CÔNG MINH - 51303338**

**NGUYỄN QUANG HUY- 51303302**

Lớp **: 13050303**

Khoá  **: 17**

**THÀNH PHỐ HỒ CHÍ MINH NĂM 2017**

TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**ĐỒ ÁN 1**

**NGHIÊN CỨU XÂY DỰNG ỨNG DỤNG NHẬN BIẾT TÍN HIỆU ĐÈN GIAO THÔNG QUA CAMERA**

*Người hướng dẫn*: **TS.PHẠM VĂN HUY**

*Người thực hiện*: **NGUYỄN CÔNG MINH - 51303338**

**NGUYỄN QUANG HUY- 51303302**

Lớp **: 13050303**

Khoá  **: 17**

**THÀNH PHỐ HỒ CHÍ MINH NĂM 2017**

LỜI CẢM ƠN

Chúng em xin cảm ơn thầy Phạm Văn Huy và các bạn trường Đại học Tôn Đức Thắng khoa Công Nghệ Thông Tin đã giúp chúng em hoàn thành tốt đồ án này. Bái báo cáo có thể có sai sót mong thầy cô thông cảm.

**ĐỒ ÁN ĐƯỢC HOÀN THÀNH**

**TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

Tôi xin cam đoan đây là sản phẩm đồ án của riêng chúng tôi và được sự hướng dẫn của TS.Phạm Văn Huy. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

Ngoài ra, trong đồ án còn sử dụng một số nhận xét, đánh giá cũng như số liệu của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

**Nếu phát hiện có bất kỳ sự gian lận nào chúng tôi xin hoàn toàn chịu trách nhiệm về nội dung đồ án của mình.** Trường đại học Tôn Đức Thắng không liên quan đến những vi phạm tác quyền, bản quyền do chúng tôi gây ra trong quá trình thực hiện (nếu có).

*TP. Hồ Chí Minh, ngày tháng năm*

*Tác giả*

*(ký tên và ghi rõ họ tên)*

*Nguyễn Công Minh*

*Nguyễn Quang Huy*

PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN

**Phần xác nhận của GV hướng dẫn**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tp. Hồ Chí Minh, ngày tháng năm

(kí và ghi họ tên)

**Phần đánh giá của GV chấm bài**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tp. Hồ Chí Minh, ngày tháng năm

(kí và ghi họ tên)

MỤC LỤC

[CHƯƠNG 1. GIỚI THIỆU 3](#_Toc484168673)

[1.1. Tổng quan đề tài 3](#_Toc484168674)

[1.2. Mục tiêu đề tài – Bài toán nhận diện tín hiệu giao thông 3](#_Toc484168675)

[1.3. Phạm vi nghiên cứu 4](#_Toc484168676)

[1.4. Bố cục 4](#_Toc484168677)

[CHƯƠNG 2. CƠ SỞ LÝ THUYẾT 6](#_Toc484168678)

[2.1. Giới thiệu về openCV 6](#_Toc484168683)

[2.1.1. Cấu trúc openCV 6](#_Toc484168684)

[2.1.2. Không gian màu 8](#_Toc484168685)

[2.2. Haar – Cascade 10](#_Toc484168686)

[2.2.1. Haar – like feature 10](#_Toc484168687)

[2.2.2. Cascade Classifier 13](#_Toc484168688)

[2.2.3. Negative Sample 15](#_Toc484168689)

[2.2.4. Positive Sample 15](#_Toc484168690)

[CHƯƠNG 3. XÂY DỰNG ỨNG DỤNG NHẬN DẠNG TÍN HIỆU ĐÈN GIAO THÔNG 16](#_Toc484168691)

[3.1. Ý tưởng 16](#_Toc484168694)

[3.2. Hiện thực ứng dụng 16](#_Toc484168695)

[3.2.1. Thiết lập 16](#_Toc484168696)

[3.2.2. Quá trình huấn luyện 17](#_Toc484168697)

[3.2.3. Phân lớp và khoanh vùng đối tượng 20](#_Toc484168698)

[3.2.4. Nhận biết màu 24](#_Toc484168699)

[3.3. Kết quả chạy thực nghiệm: 25](#_Toc484168700)

[CHƯƠNG 4. KẾT LUẬN 27](#_Toc484168701)

[4.1. Ưu điểm 27](#_Toc484168703)

[4.2. Ngược điểm 27](#_Toc484168704)

[4.3. Hướng phát triển 27](#_Toc484168705)

DANH MỤC HÌNH VẼ, ĐỒ THỊ

[Hình 2.1‑1 Logo openCV 6](#_Toc483993344)

[Hình 2.1‑2 Cấu trúc openCV. 7](#_Toc483993345)

[Hình 2.1‑3 Dải màu HSV. 9](#_Toc483993346)

[Hình 2.2‑1 Quá trình huấn luyện trên console. 10](#_Toc483993347)

[Hình 2.2‑2 Các đặc trưng của Haar-like cơ bản 11](#_Toc483993348)

[Hình 2.2‑3 Tính giá trị ảnh tích phân tại điểm P có tọa độ (x, y). 12](#_Toc483993349)

[Hình 2.2‑4 Tính nhanh giá trị của vùng ảnh D. 12](#_Toc483993350)

[Hình 2.2‑5 Mô tả về phương pháp phân tầng Cascade. 13](#_Toc483993351)

[Hình 2.2‑6 Khoanh vùng đối tượng (không nhiễu). 14](#_Toc483993352)

[Hình 2.2‑7 Khoanh vùng đối tượng (nhiễu nhiều). 15](#_Toc483993353)

[Hình 3.2‑1 Các tập tin trong thư mục Haar training. 17](#_Toc483993354)

[Hình 3.2‑2 Tạo danh sách Negative. 17](#_Toc483993355)

[Hình 3.2‑3 Tạo danh sách Positive. 18](#_Toc483993356)

[Hình 3.2‑4 Danh sách Positive sau khi tạo xong. 19](#_Toc483993357)

[Hình 3.2‑5 Chỉnh sửa thông số huấn luyện. 19](#_Toc483993358)

[Hình 3.2‑6 Thông số trong tập tin xml sau khi luyện. 20](#_Toc483993359)

[Hình 3.2‑7 Đèn giao thông dạng 1. 21](#_Toc483993360)

[Hình 3.2‑8 Đèn giao thông dạng 2. 21](#_Toc483993361)

[Hình 3.2‑9 Đèn giao thông dạng 3. 22](#_Toc483993362)

[Hình 3.2‑10 Khoanh vùng đèn giao thông dạng 2. 23](#_Toc483993363)

[Hình 3.2‑11 Khoanh vùng đèn giao thông dạng 3 (không nhiễu). 23](#_Toc483993364)

[Hình 3.2‑12 Khoanh vùng đèn giao thông dạng 3 (có nhiễu). 23](#_Toc483993365)

[Hình 3.3‑1 Nhận diện tín hiệu đèn đỏ 25](#_Toc483993366)

[Hình 3.3‑2 Nhận diện tín hiệu đèn xanh. 26](#_Toc483993367)

# GIỚI THIỆU

## Tổng quan đề tài

Cặp mắt là một trog những phần quan trọng nhất của con người,nó giúp ta nhìn thấy biết được màu sắc,hình dáng… của mọi sự vật.Có thể nói,cặp mắt là món quà mà con người có được từ khi mới sinh ra.

Thế nhưng những vật vô tri như máy tính,camera,robot.. thì vốn không thể “nhìn” như con người,vì thế lĩnh vực Thị giác máy tính ra đời để tìm ra các giải pháp giúp cho những vật ấy có thể “nhìn” cũng như nhận biết các sự vật trong cược sống.

Một cách tổng quát, **Thị giác máy tính** (**computer vision**) là một lĩnh vực bao gồm các phương pháp thu nhận, [xử lý ảnh kỹ thuật số](https://vi.wikipedia.org/w/index.php?title=X%E1%BB%AD_l%C3%BD_%E1%BA%A3nh_k%E1%BB%B9_thu%E1%BA%ADt_s%E1%BB%91&action=edit&redlink=1), phân tích và nhận dạng các hình ảnh, nói chung là dữ liệu đa chiều từ thế giới thực để cho ra các thông tin số hoặc biểu tượng.Việc phát triển lĩnh vực này bắt nguồn từ việc sao chép các khả năng thị giác con người bởi sự nhận diện và hiểu biết một hình ảnh mang tính điện tử. Sự nhận diện hình ảnh có thể xem là việc giải quyết vấn đề của các biểu tượng thông tin từ dữ liệu hình ảnh qua cách dùng các mô hình được xây dựng với sự giúp đỡ của các ngành [lý thuyết học](https://vi.wikipedia.org/w/index.php?title=L%C3%BD_thuy%E1%BA%BFt_h%E1%BB%8Dc&action=edit&redlink=1), [thống kê](https://vi.wikipedia.org/wiki/Th%E1%BB%91ng_k%C3%AA), [vật lý](https://vi.wikipedia.org/wiki/V%E1%BA%ADt_l%C3%BD) và [hình học](https://vi.wikipedia.org/wiki/H%C3%ACnh_h%E1%BB%8Dc).

Nhận diện tín hiệu giao thông là một nhánh của nhận diện đối tượng.

## Mục tiêu đề tài – Bài toán nhận diện tín hiệu giao thông

Nhận diện tín hiệu giao thông trong ảnh hoặc là bài toán có nhiều ứng dụng trong thực tế. Việc phát hiện tốt, chính xác những tín hiệu đèn có mặt trong ảnh hoặc sẽ tạo tiền đề tốt cho những hướng phát triển như:

* Phục vụ quá trình tham gia giao thông diễn ra đúng với luật giao thông đường bộ.
* Hướng dẫn người tham gia giao thông các hướng di chuyển, xử lý tình huống tiếp theo nếu gặp 1 tín hiệu giao thông.

Những khó khăn phải đối mặt trong đề tài là:

* Các ảnh/video thường không đạt chuẩn, backround phức tạp.
* Ảnh hưởng bất lợi về điều kiện ánh sáng.
* Sự che khuất lẫn nhau của các vật thể và tín hiệu đèn trong ảnh.
* Sự đa dạng về loại, kiểu dáng cột đèn.
* Đèn giao thông thường ở các mức tỉ lệ khác nhau.
* Một số đối tượng giống với cột đèn giao thông như **cột đèn 2 bên đường đi,…**

## Phạm vi nghiên cứu

Trong đề tài này, tôi tập trung vào việc xác định cột đèn và tín hiệu màu trong ảnh, video từ tập dữ liệu chúng tôi thu thập tại khu vực quận 7, Tp.Hồ Chí Minh. Do các điều kiện khó khăn đã nêu ở trên (mục 1.2) tôi xin đưa ra những giả định và ràng buộc sau để giảm độ phức tạp của bài toán:

* Giảm kích thước xét duyệt video, chỉ xét nửa trên của video (640x480) tức là thuật toán chỉ phân tích video (640x240). Điều này cũng phản ánh đúng thực tế vị trí của cột đèn thường nằm phía trên cao.
* Video/ảnh trong điều kiện ánh sáng bình thường.

## Bố cục

Nội dung cố đồ án này được trình bày trong 5 chương, bao gồm các nội dung về các phương pháp và vấn đề về nhận diện đối tượng. Cùng với việc áp dụng các phương pháp này vào vấn đề nhận diện tín hiệu đèn giao thông.

**Chương 1 – Tổng quan đề tài**: Giới thiệu tổng quan về đề tài, bài toán, mục tiêu và phạm vi đề tài.

**Chương 2 – Cơ sở lý thuyết**: Các khái niệm về *openCV*, không gian màu, tìm hiểu về phương pháp huận luyện *Haar-like* kết hợp với *Cascade Classifier.*

**Chương 3 – Xây dựng ứng dụng nhận dạng tín hiệu đèn giao thông**: Thiết lập, quá trình huấn luyện. Kết quả chạy thực nghiệm

**Chương 4 – Kết luận**: Rút ra các kết luận ưu, nhược điểm và xác định hướng phát triển tiếp theo cho bài toán.

# CƠ SỞ LÝ THUYẾT



## Giới thiệu về openCV

OpenCV[2] (Open Source Computer Vision) là một thư viện mã nguồn mở về thị giác máy với hơn 500 hàm và hơn 2500 các thuật toán đã được tối ưu về xử lý ảnh, và các vấn đề liên quan tới thị giác máy. OpenCV là thư viện được thiết kế để chạy trên nhiều nền tảng khác nhau (cross-patform) như Window, Linux, Mac, iOS …



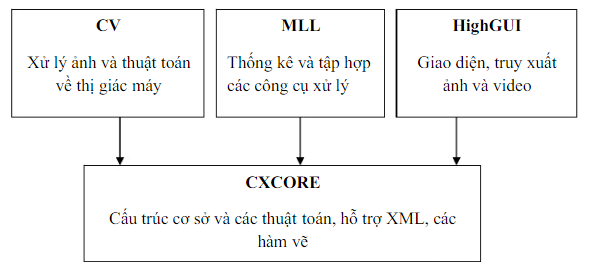
Hình ‑ Logo openCV

Dự án về OpenCV được khởi động từ những năm 1999, đến năm 2000 nó được giới thiệu trong một hội nghị của IEEE về các vấn đề trong thị giác máy và nhận dạng, tuy nhiên bản OpenCV 1.0 mãi tới tận năm 2006 mới chính thức được công bố và năm 2008 bản 1.1 pre-release) mới được ra đời. Tháng 10 năm 2009, bản OpenCV thế hệ thứ hai ra đời, phiên bản này có giao diện của C++ (khác với phiên bản trước có giao diện của C) và có khá nhiều điểm khác biệt so với phiện bản thứ nhất. Thư viện OpenCV ban đầu được sự hỗ trợ từ Intel, sau đó được hỗ trợ bởi Willow Garage, một phòng thí nghiệm chuyên nghiên cứu về công nghệ robot. Cho đến nay, OpenCV vẫn là thư viện mở, được phát triển bởi nguồn quỹ không lợi nhuận (none -profit foundation) và được sự hưởng ứng rất lớn của cộng đồng.

### Cấu trúc openCV

Thư viện OpenCV có thể được chia thành 2 phần (module) chính. Phần căn bản (basic) là mã nguồn được nhóm phát triển xây dựng và kiểm định toàn diện, gồm các thuật toán đã được thế giới công nhận và đánh giá dựa trên cơ sở lý thuyết chắc chắn. Phần mở rộng (contribution) được nhiều tổ chức khoa học khác nhau trên thế giới đóng góp, gồm nhiều thuật toán cập nhật được xây dựng dựa trên các công trình nghiên cứu, bài báo mới đăng trong thời gian gần đây. Do vậy các thuật toán trong phần mở rộng có độ ổn định và tối ưu không cao. Từ phiên bản 3.0, phần mở rộng được tách riêng không còn được gộp chung với thư viện mặc định.

OpenCV là thư viện mã nguồn mở được đóng gói thành tập tin nén. Tùy vào phiên bản dành cho các hệ điều hành khác nhau mà tập tin nén này có định dạng tương ứng. Thư viện OpenCV cung cấp cho người dùng các cấu trúc dữ liệu, đối tượng và hàm bằng cách khai báo nguyên mẫu (prototype) của chúng trong các tập tin thư viện C/C++ (\*.h, \*.hpp,…) và định nghĩa chi tiết trong các tập tin mã nguồn (\*.c, \*.cpp). Với mức độ sử dụng OpenCV, chỉ cần giải nén các tập tin đã được biên dịch sẵn rồi thực hiện các thao tác cài đặt đường dẫn cho thích hợp để hệ điều hành tìm đến đúng vị trí của các tập tin thư viện. Ở mức độ cao hơn, nếu muốn hiệu chỉnh sửa đổi thuật toán hay sử dụng phần mở rộng của OpenCV ta cần phải biên dịch mã nguồn trực tiếp trên máy trước khi cài đặt.



Hình ‑ Cấu trúc openCV.

**CXCORE:** Chứa các định nghĩa kiểu dữ liệu cơ sở. Ví dụ, các cấu trúc dữ liệu cho ảnh, điểm và hình chữ nhật được định nghĩa trong cxtypes.h. CXCORE cũng chứa đại số tuyến tính và phương pháp thống kê, chức năng duy trì và điều khiển chuỗi. Một số ít, các chức năng đồ họa để vẽ trên ảnh cũng được đặt ở đây.

**CV**: Chứa các thuật toán về xử lý ảnh và định kích cỡ camera. Các chức năng hình họa máy tính cũng được đặt ở đây.

**CVAUX**: Dùng cho nhận diện mặt và chúng được ứng dụng rộng rãi cho mục đích đó.

**HIGHGUI**: Chứa các giao diện vào ra cơ bản, nó cũng chứa các khả năng mở rộng cửa sổ và vào ra video.

**CVCAM**: Chứa các giao diện cho video truy cập qua DirectX trên nềnWindows 32 bits.

### Không gian màu

Không gian màu[1] là một mô hình toán học dùng để mô tả các màu sắc trong thực tế được biểu diễn dưới dạng số học. Trên thực tế có rất nhiều không gian màu khác nhau được mô hình để sử dụng vào những mục đích khác nhau. Trong openCV có ba không gian màu cơ bản hay được nhắc tới và ứng dụng nhiều, đó là hệ không gian màu RGB, HSV và CMYK.

#### Không gian màu RGB

RGB là không gian màu rất phổ biến được dùng trong đồ họa máy tính và nhiều thiết bị kĩ thuật số khác. Ý tưởng chính của không gian màu này là sự kết hợp của 3 màu sắc cơ bản : màu đỏ (R, Red), xanh lục (G, Green) và xanh lơ (B, Blue) để mô tả tất cả các màu sắc khác. Nếu như một ảnh số được mã hóa bằng 24bit, nghĩa là 8bit cho kênh R, 8bit cho kênh G, 8bit cho kênh B, thì mỗ kênh này màu này sẽ nhận giá trị từ 0-255. Với mỗi giá trị khác nhau của các kênh màu kết hợp với nhau ta sẽ được một màu khác nhau, như vậy ta sẽ có tổng cộng 255x255x255 = 1.66 triệu màu sắc.

#### Không gian màu CMYK

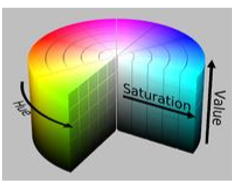
CMYK là không gian màu được sử dụng phổ biến trong ngành công nghiệp in ấn. Ý tưởng cơ bản của hệ không gian này là dùng 4 màu sắc cơ bản để phục vụ cho việc pha trộn mực in. Trên thực tế, 3 màu là C=Cyan: xanh lơ, M=Magenta: hồng xẫm, và Y=Yellow: vàng để biểu diễn các màu sắc khác nhau. Nếu lấy màu hồng xẫm cộng với vàng sẽ ra màu đỏ, màu xẫm kết hợp với xanh lơ sẽ cho xanh lam ... Sự kết hợp của 3 màu trên sẽ cho ra màu đen, tuy nhiên màu đen ở đây khôn phải là đen tuyệt đối và thường có độ tương phản lớn, nên trong ngành in, để tiết kiệm mực in người ta thêm vào màu đen để in những chi tiết có màu đen thay vì phải kết hợp 3 màu sắc trên.

Nguyên lý làm việc của hệ màu này như sau : Trên một nền giấy trắng, khi mỗi màu này được in lên sẽ loại bỏ dần đi thành phần màu trắng. 3 màu C, M, Y khác nhau in theo những tỉ lệ khác nhau sẽ loại bỏ đi thành phần đó một cách khác nhau và cuối cùng cho ta màu sắc cần in. Khi cần in màu đen, thay vì phải in cả 3 màu người ta dùng màu đen để in lên. Nguyên lý này khác với nguyên lý làm việc của hệ RGB ở chỗ hệ RGB là sự kết hợp của các thành phần màu, còn hệ CMYK là sự loại bỏ lẫn nhau của các thành phần màu.

#### Không gian màu HSV

HSV và cũng gần tương tự như HSL là không gian màu được dùng nhiều trong việc chỉnh sửa ảnh, phân tích ảnh và một phần của lĩnh vực thị giác máy tính. Hệ không gian này dựa vào 3 thông số sau để mô tả màu sắc H = Hue: màu sắc, S = Saturation: độ đậm đặc, sự bảo hòa, V = value: giá trị cường độ sáng.

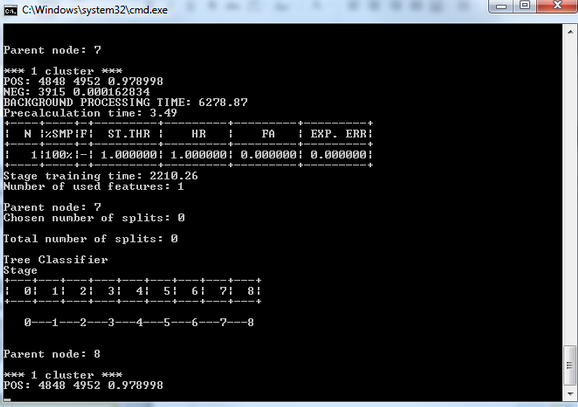
Không gian màu này thường được biểu diễn dưới dạng hình trụ hoặc hình nón. Theo đó, đi theo vòng tròn từ 0 -360 độ là trường biểu diễn màu sắc(Hue). Trường này bắt đầu từ màu đỏ đầu tiên (red primary) tới màu xanh lục đầu tiên (green primary) nằm trong khoảng 0-120 độ, từ 120 - 240 độ là màu xanh lục tới xanh lơ (green primary - blue primary). Từ 240 - 360 là từ màu đen tới lại màu đỏ.



Hình ‑ Dải màu HSV.

## Haar – Cascade

Haar – Cascade là giải thuật kết hợp giữa phương thức Haar-like feature và phương pháp Cascade, được sử dụng trong xử lý ảnh về nhận dạng đối tượng (Object detection) bằng cách huấn luyện trên hai tập dữ liệu hình ảnh có sẵn là Positive và Negative thành file .xml hoặc .yml. Tổng số hình phải luyện là trên 1000 hình để chương trình có thể nhận diện đúng, càng nhiều hình thì kết quả nhận diện càng hiệu quả. Quá trình huấn luyện được gọi là Haar Training (Cascade Training), sau khi huấn luyện xong Cascade Classifier sẽ kiểm tra kết quả qua một mẫu thử (hình ảnh, video hoặc trực tiếp qua thiết bị ghi hình).

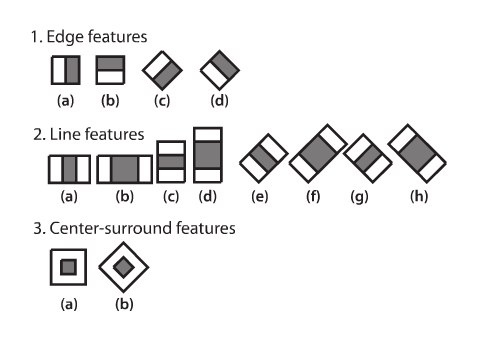


Hình ‑ Quá trình huấn luyện trên console.

### Haar – like feature

Haar – like feature [11] là một phương thức dùng để phát hiện đối tượng trong hình ảnh qua các cấu trúc cụ thể trong hình.

Hình được chia thành nhiều vùng cụ thể và giải thuật sẽ tra cứu trong từng điểm ảnh (pixel) tìm ra sự khác biệt. Những sự khác biệt đó sẽ được mã hóa thành các thông số và lưu trữ lại thông qua quá trình luyện.



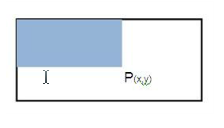
Hình ‑ Các đặc trưng của Haar-like cơ bản

Để nhận diện được tín hiệu đèn giao thông ta sẽ áp dụng đặc trưng xung quanh tâm

 Lợi ích của các đặc trưng Haar-like là diễn đạt được tri thức về các đối tượng trong ảnh (bởi vì nó biểu diễn mối liên hệ giữa các bộ phận của đối tượng), điều mà bản thân từng điểm ảnh không diễn đạt được. Để tính giá trị các đặc trưng haar-like, ta tính sự chênh lệch giữa tổng của các pixel của các vùng đen và các vùng trắng như trong công thức sau: f(x) = Tổng vùng đen(pixel) - Tổng vùng trắng(pixel) Như vậy ta có thể thấy rằng, để tính các giá trị của đặc trưng Haar-like, ta phải tính tổng của các vùng pixel trên ảnh. Nhưng để tính tốn các giá trị của các đặc trưng Haar-like cho tất cả các vị trí trên ảnh địi hỏi chi phí tính tốn khá lớn, không đáp ứng được cho các ứng dụng địi hỏi tính run-time. Do đó Viola và Jones đưa ra một khái niệm gọi là Integral Image để tính tốn nhanh cho các đặc trưng cơ bản.

Sau này, Lienhart kế thừa Integral Image (SAT) và đưa thêm khái niệm Rotated Summed Area Table (RSAT) dùng để tính tốn nhanh cho các đặc trưng xoay 1 góc 45 độ. Integral Image là một mảng 2 chiều với kích thước bằng với kích của ảnh cần tính các đặc trưng Haar-like, với mỗi phần tử của mảng này được tính bằng cách tính tổng của điểm ảnh phía trên (dạng-1) và bên trái (cột-1) của nĩ. Bắt đầu từ vị trí trên, bên trái đến vị trí dưới, phải của ảnh, việc tính tốn này đơn thuần chỉ đựa trên phép cộng số nguyên đơn giản, do đó tốc độ thực hiện rất nhanh.

Dưới đây là mô tả cách tính ảnh tích phân:

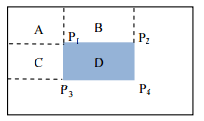


Hình ‑ Tính giá trị ảnh tích phân tại điểm P có tọa độ (x, y).

Giá trị của ảnh tích phân tại điểm P có tọa độ (x, y) được tính như sau:



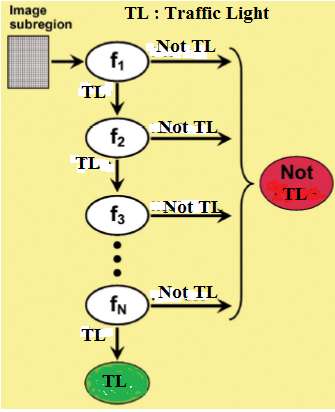
Sau khi đã tính được ảnh tích phân, việc tính tổng các giá trị mức xám của một vùng ảnh bất kỳ nào đó trên ảnh thực hiện theo cách sau, ví dụ tính giá trị của vùng D trong hình 5 như sau: **D=A+B+C+D-(A+B)-(A+C)+A.**



Hình ‑ Tính nhanh giá trị của vùng ảnh D.

### Cascade Classifier

Ta thấy quá trình huấn luyện, bộ phân loại phải duyệt qua tất cả các đặc trưng của các mẫu trong tập training. Việc này tốn rất nhiều thời gian. Tuy nhiên, trong các mẫu đưa vào, không phải mẫu nào cũng thuộc loại khĩ nhận dạng, có những mẫu background rất dễ nhận ra (ta gọi đây là những mẫu background đơn giản). Đối với những mẫu này, ta chỉ cần xét một hay vài đặc trưng đơn giản là có thể nhận diện được chứ không cần xét tất cả các đặc trưng. Nhưng đối với các bộ phân loại thơng thường thì cho dù mẫu cần nhận dạng là dễ hay khĩ thì nĩ vẫn sẽ xét tất cả các đặc trưng mà nĩ rút ra được trong quá trình học. Do đó, chúng tốn thời gian xử lý một cách không cần thiết.



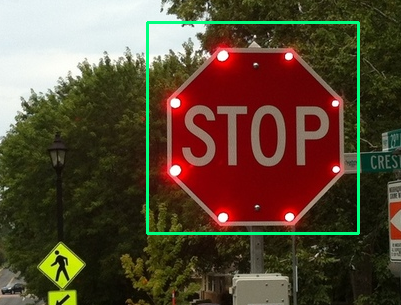
Hình ‑ Mô tả về phương pháp phân tầng Cascade.

 Cascade of Classifiers [11] được xây dựng chính là nhằm rút ngắn thời gian xử lý, giảm thiểu false alarm cho bộ phân loại. Cascade tree gồm nhiều stage (hay cịn gọi là layer), mỗi stage của cây sẽ là một stage classifier. Một mẫu để được phân loại là đối tượng thì nĩ cần phải đi qua hết tất cả các stages của cây. Các stage classifiers ở stage sau được huấn luyện bằng những mẫu negative mà stage classifier trước nĩ nhận dạng sai, tức là nĩ sẽ tập trung học từ các mẫu background khĩ hơn, do đó sự kết hợp các stage classifiers này lại sẽ giúp bộ phân loại có false alarm thấp.

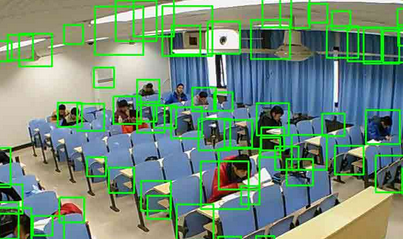
Với cấu trúc này, những mẫu background dễ nhận diện sẽ bị loại ngay từ những stages đầu tiên, giúp đáp ứng tốt nhất đối với độ phức tạp gia tăng của các mẫu đưa vào, đồng thời giúp rút ngắn thời gian xử lý.

Thuật toán Cascade training:

Sau khi huấn luyện xong, chương trình sẽ gọi kết quả để kiểm tra (file .xml hoặc .yml). OpenCV có hỗ trợ lớp “CascadeClassifier” và hiện thực bằng cách gọi phương thức “detectMultiScale(param 1, param 2, …)”, phương thức trên sẽ phân loại qua các pixel và khoanh vùng các đối tượng bằng một vector<Shape>. Dựa vào vector<Shape> chương trình sẽ hiển thị trên hình ảnh cho người lập trình biết vùng nhận diện đối tượng bằng các hình chữ nhật hay hình tròn hoặc có thể là hình đa giác. Qua đó, người lập trình có thể biết được hiệu quả của tập huấn luyện hiện tại và có thể đưa ra cách chỉnh sửa tập huấn luyện đó để có kết quả chính xác hơn.



Hình ‑ Khoanh vùng đối tượng (không nhiễu).



Hình ‑ Khoanh vùng đối tượng (nhiễu nhiều).

### Negative Sample

Negative sample mang ý nghĩa là một mẫu tiêu cực, đây là các hình ảnh không chứa các đối tượng muốn nhận diện và có thể gọi là hình ảnh gây nhiễu. Các hình ảnh này thường là các phông nền, các đối tượng ở gần đối tượng muốn nhận diện,...Các hình ảnh negative này luyện càng nhiều thì mức độ khử nhiễu càng cao.

### Positive Sample

Ngược lại với Negative sample, Positive sample là mẫu tích cực gồm cac hình ảnh của đối tượng muốn nhận diện. Các hình ảnh này phải được luyện ở các góc độ khác nhau, các thời điểm khác nhau để lấy được nhiều thông tin về hình dạng của đối tượng nhận diện. Các hình ảnh positive này luyện càng nhiều thì độ chính xác càng cao.

# XÂY DỰNG ỨNG DỤNG NHẬN DẠNG TÍN HIỆU ĐÈN GIAO THÔNG



## Ý tưởng

Ứng dụng sẽ sử dụng song song hai chức năng là khoanh vùng đèn giao thông và xác định màu của đèn. Đầu tiên, dựa vào tập huấn luyện xác định vị trí của đèn, hiển thị lên cho người dùng. Sau đó, duyệt trên các vùng đã xác định, kiểm tra xem có tồn tại một trong ba màu (xanh, đỏ, vàng) hay không. Nếu có thì sẽ hiển thị thông báo bằng cách vẽ một hình tròn có màu tương ứng dưới góc phải màn hình. Nếu ba màu trên không tồn tại trong vùng xác định thì sẽ không có thông báo nào cả để tránh hiện tượng nhận sai các vùng bị nhiễu.

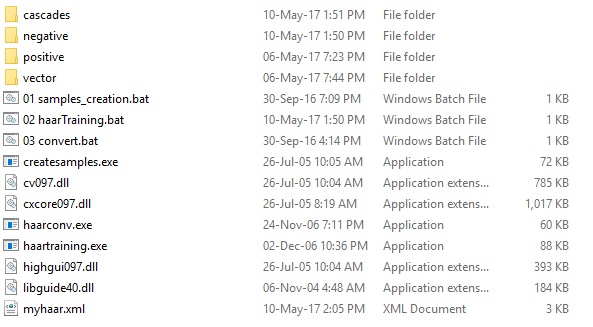
## Hiện thực ứng dụng

### Thiết lập

* OpenCV 3: Các phiên bản openCV 3.0, 3.1, 3.2, …
* Hệ điều hành: Window 10.
* Ngôn ngữ: Ứng dụng được viết bằng ngôn ngữ C++.
* Sử dụng phần mềm Visual Studio 2015 để lập trình.
* Hai tập dữ liệu hình ảnh Positive và Negative (mỗi tập khoảng hơn 500 hình):
  + Để có được tập dữ liệu hình ảnh ta phải tải từng ảnh trên internet như vậy sẽ rất mất thời gian nên chúng tôi đã sử dụng phần mềm **Total Video Converter. Tải tại:** [**http://download.cnet.com/Total-Video-Audio-Converter/3000-2194\_4-76208009.html**](http://download.cnet.com/Total-Video-Audio-Converter/3000-2194_4-76208009.html)
  + Add File (Video cần cắt ảnh) -> Option -> Chọn Frame rate 5000Hz (5 frame hình / giây) -> OK -> Nhấn Convert.

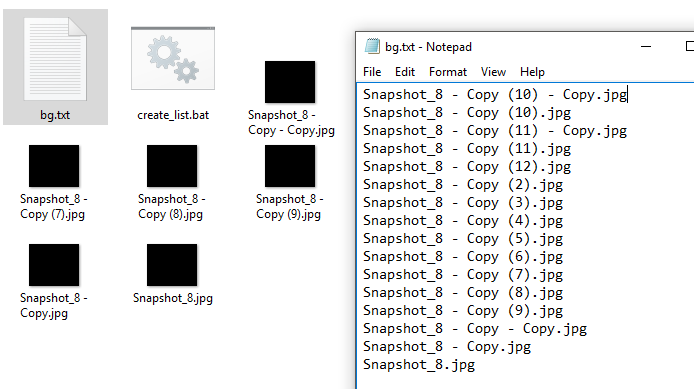
### Quá trình huấn luyện

Với hệ điều hành Window, người lập trình có thể tải xuống bộ công cụ luyện Haar-Training của Mahdi Rezaei về mảng khoa học máy tính tại Đại học Auckland, New Zealand (<https://www.cs.auckland.ac.nz/~m.rezaei/Tutorials/Haar-Training.zip>).



Hình ‑ Các tập tin trong thư mục Haar training.

* Bước 1: Tạo danh sách Negative.
* Chuyển tất cả dữ liệu hình Negative vào thư mục **rawdata** trong thư mục **negative**.
* Chạy file **create\_list.bat** để tạo file **bg.txt** chứa tên của các file ảnh

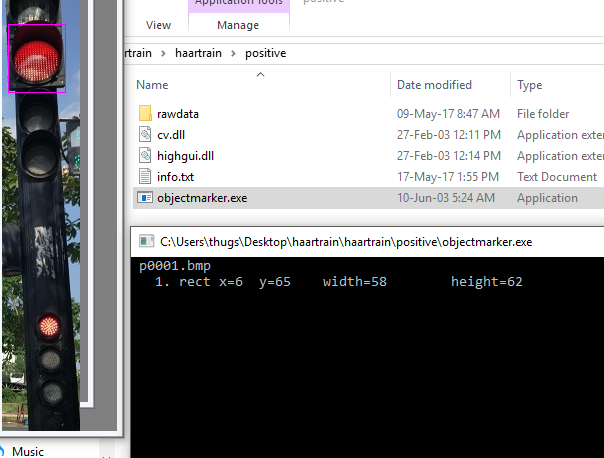


Hình ‑ Tạo danh sách Negative.

* Bước 2: Tạo danh sách Positive.

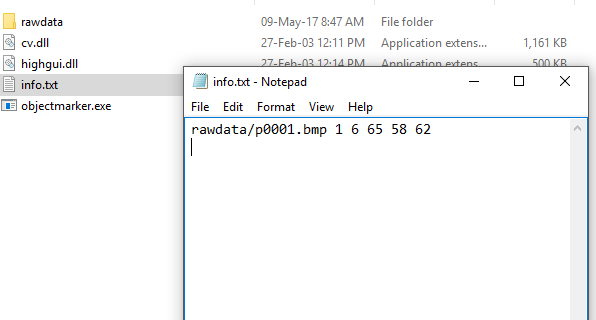
- Chuyển tất cả dữ liệu hình Positive vào thư mục **rawdata** trong thư mục **positive**.

- Chạy tập tin **objectmarker.exe** để lưu lại vị trí của đối tượng trong hình.



Hình ‑ Tạo danh sách Positive.

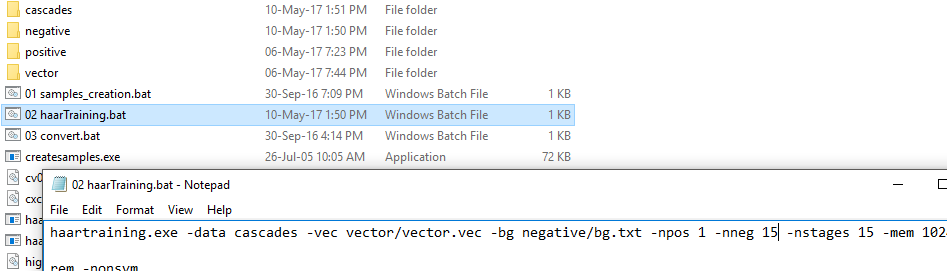
- Tập dữ liệu positive sẽ được lưu lại thành một danh sách trong tập tin **info.txt**.



Hình ‑ Danh sách Positive sau khi tạo xong.

* Bước 3: Chỉnh sửa thông số huấn luyện.

- Nhấp chuột phải vào tập tin **haarTraining.bat** chọn **edit** để chỉnh sửa thông số huấn luyện.



Hình ‑ Chỉnh sửa thông số huấn luyện.

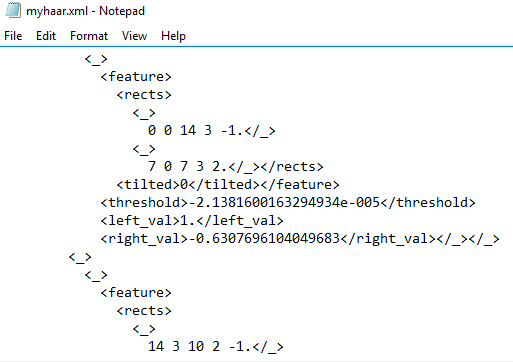
-npos <số lượng ảnh positive>.

-nneg <số lượng ảnh negative>.

* Bước 4: Huấn luyện.

- Chạy lần lượt 3 tập tin theo thứ tự samples\_creation.bat, haarTraining.bat và convert.bat.

- Kết quả cuối cùng sẽ tạo ra một tập tin .xml chứa thông tin dữ liệu hình ảnh đã được mã hóa.



Hình ‑ Thông số trong tập tin xml sau khi luyện.

### Phân lớp và khoanh vùng đối tượng

#### Đối tượng nhận dạng

Đối tượng nhận dạng sẽ là các cột đèn giao thông trên đường Nguyễn Hữu Thọ và đường Nguyễn Thị Thập, Quận 7, thành phố Hồ Chí Minh.

Sau khi khảo sát, nhóm đã xác định được hai đoạn đường trên có các loại đèn giao thông như sau:



Hình ‑ Đèn giao thông dạng 1.



Hình ‑ Đèn giao thông dạng 2.



Hình ‑ Đèn giao thông dạng 3.

#### Cài đặt

Để có một kết quả tốt thì tối thiểu phải luyện là 1000 dữ liệu hình ảnh (500 positive và 500 negative). Sau khi luyện xong, tập tin .xml sẽ dược gọi để sử dụng bằng lớp **CascadeClassifier**.

- Source code:

CascadeClassifier classifier;

classifier.load(CASCADE\_FILE\_NAME);

Sau khi tải tập tin .xml, cần chuyển hình ảnh (image) hoặc khung hình (frame) về mức xám và cân bằng Histogram (biểu đồ phân bố cường độ điểm ảnh có giá trị).

- Source code:

cvtColor(mFrame, mGray, COLOR\_BGR2GRAY);

equalizeHist(mGray, mGray);

Trong tập huấn luyện, các hình ảnh positive của đối tượng nhận diện có kích thước không tương đồng với nhau so với đối tượng trong mẫu thử (hình ảnh hoặc các khung hình trong video hay trực tiếp). Vì thế, việc nhận diện sẽ khó khăn nên openCV có hỗ trợ phương thức detectMultiScale (nhận diện với nhiều tỉ lệ). Ở trong ứng dụng nhóm quy định nhận diện đối tượng khi đối tượng có kích thước tối thiểu là 30x30 pixel.

- Source code:

classifier.detectMultiScale(mGray,vFound,1.1,2,0|CASCADE\_SCALE\_IMAGE, Size(30, 30));

#### Kiểm tra khoanh vùng

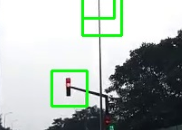
Khi chạy thực nghiệm, ứng dụng có thể nhận diện được các đèn giao thông nhưng vẫn còn nhiễu bởi các đối tượng có cấu trúc tương tự: cột đèn, cột điện,…



Hình ‑ Khoanh vùng đèn giao thông dạng 2.



Hình ‑ Khoanh vùng đèn giao thông dạng 3 (không nhiễu).



Hình ‑ Khoanh vùng đèn giao thông dạng 3 (có nhiễu).

### Nhận biết màu

#### Dải màu hsv của đèn giao thông

Sử dụng hàm **inRange** trong openCV để lấy các vùng màu cần thiết trong hình hoặc khung hình (lưu ý: phải chuyển sang dải màu HSV trước khi dùng hàm **inRange**). Các vùng màu cần lấy lần lượt là:

- Màu đỏ:

<1>: Scalar(0, 100, 100) -> Scalar(10, 255, 255)

<2>: Scalar(160, 100, 100) -> Scalar(179, 255, 255)

- Màu vàng:

<1>: Scalar(0, 100, 100) -> Scalar(10, 255, 255)

<2>: Scalar(5, 50, 50) -> Scalar(15, 255, 255)

- Màu xanh lá:

<1>: Scalar(30, 59, 50) -> Scalar(79, 158, 171)

<2>: Scalar(0, 255, 0) -> Scalar(229, 255, 204)

Mỗi màu đều có 2 vùng màu, sử dụng hàm **addWeighted** để hai vùng màu của mỗi màu thành một.

Sau khi thực hiện các bước trên, kết quả sẽ thu lại được 3 threshold của 3 màu đỏ, vàng và xanh lá.

#### Tìm vị trí màu trên ảnh

Dùng hàm **findContours** của openCV để nhận biết viền trong mỗi threshold của từng màu. Sau đó, sử dụng lớp đối tượng **Moments** của openCV để xác định vị trí và độ lớn của các cụm màu trong threshold.

Trước khi lưu lại các vị trí, cần phải khử nhiễu bằng cách kiểm tra kích thước của các cụm màu. Nếu cụm màu có kích thước nhỏ hơn 20x20 pixel thì đó là nhiễu cần loại bỏ. Nếu cụm màu lớn hơn hoặc bằng 2/3 màn hình hiển thị thì đó là do vật ở quá gần hoặc đó phông nền (background).

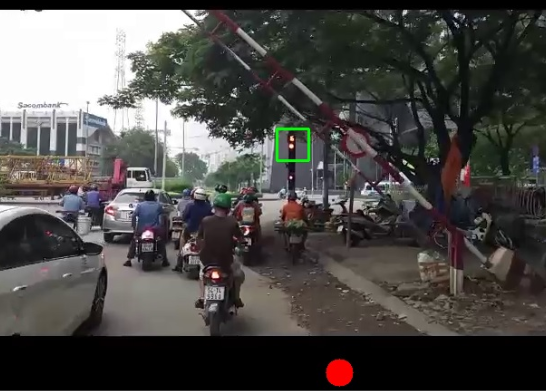
Kết quả thu về dược một danh sách các vị trí màu đã được khử nhiễu.

#### Kết hợp nhận diện màu với khoanh vùng

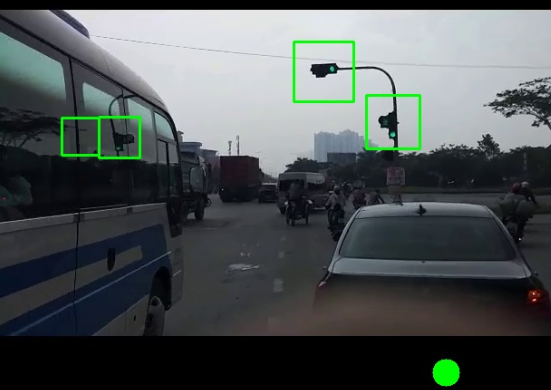
Sử dụng vòng lặp duyệt tuần tự, cứ mỗi khoanh vùng đối tượng thì kiểm tra bên trong có cụm màu nào không. Nếu có thì trả về màu nhận biết được (khi kiểm tra màu nên sử dụng lệnh **else if** để lấy duy nhất một màu đèn, tránh hiện tượng 1 lúc hiển thị 3 màu đèn).

## Kết quả chạy thực nghiệm:

Ứng dụng chạy thực nghiệm trên một video với độ dài 7 phút được ghi hình trên tuyến đường Nguyễn Hữu Thọ và Nguyễn Thị Thập, Quận 7, Thành phố Hồ Chí Minh. Kết quả sẽ hiển thị như sau:



Hình ‑ Nhận diện tín hiệu đèn đỏ



Hình ‑ Nhận diện tín hiệu đèn xanh.

**Nhận xét:**

- Tỉ lệ nhận dạng trên 90%. Hầu hết đều nhận và khoanh vùng được đèn, nhưng trong lúc chạy ứng dụng đã không nhận dạng được đèn trong một vài khung hình và lý do chính không nhận diện được là khi những đèn bị che khuất bởi cây hoặc các vật khác.

- Dù đã khử nhiễu rất nhiều nhưng vẫn chưa triệt để như hình hình 3.3-2 và 3.2-12, nguyên nhân là do tập dữ liệu hình còn ít nên ứng dụng vẫn chưa đủ một số dữ kiện cho các hình dạng khó như: cây cối, cột đèn đường,… và các cấu trúc vật thể khác có hình dạng giống như đèn giao thông.

# KẾT LUẬN



## Ưu điểm

* Nhận biết được hình dạng đèn giao thông và phân biệt được trạng thái của màu đèn.
* Khử nhiễu tốt.
* Có thể nhận dạng trên 3 phương diện: Hình ảnh, video, real time (thiết bị ghi hình).

## Nhược điểm

* Thuật toán khá nặng dẫn đến việc xử lý lâu khi áp dụng trên video và real time.
* Vẫn còn nhiễu tại một số điểm nhỏ dẫn đến sai sót đôi chút.
* Tập huấn luyện hình ảnh vẫn khá ít.

## Hướng phát triển

* Tối ưu và cải thiện thuật toán để quá trình xử lý tốt hơn.
* Tăng số lượng tập huấn luyện (khoảng trên 2000 hình positive, và trên 2000 hình negative) để giảm nhiễu và nhận diện chính xác hơn.
* Cập nhật thêm tính năng nhận biết thời gian chờ của đèn giao thông.
* Cập nhật thêm tính năng thông báo dành cho người đi bộ của các đèn giao thông hiện đại.

TÀI LIỆU THAM KHẢO

1. M.Amutha Bharathi. “Color Detection in OpenCV”.
2. OpenCV - <http://opencv.org/> .
3. Jesse Levinson, Jake Askeland, Jennifer Dolson, Sebastian Thrun. “Trafﬁc Light Mapping, Localization, and State Detection for Autonomous Vehicles”.
4. Nguyễn Văn Long. “Ứng dụng xử lý ảnh trong thực thế với thư viện OpenCV C/C++”.
5. Sander Soo. “Object detection using Haar-cascade Classifier”.
6. N. Dalal and B. Triggs, “Histograms of oriented gradients for human dectection” in IEEE conference on computer vision and Pattern Recognition, 2005.
7. Moises Diaz-Cabrera and Pietro Cerri. “Traﬃc Light Recognition During the Night Based on Fuzzy Logic Clustering”.
8. Mahmoud Abdallah and Daniel Eiland. “Traffic Signal Detection and Tracking”, 2011.
9. Subhransu Maji et al. Classification using Intersection Kernel Support Vector Machines is Effcient. IEEE Computer Vision and Pattern Recognition 2008.
10. Bài toán nhận dạng đối tượng chuyển động trong video - <http://doan.edu.vn/do-an/bai-toan-nhan-dang-doi-tuong-chuyen-dong-trong-video-29567/>
11. Bài toán phát hiện mặt người trong ảnh (trang 24-28) - <http://luanvan.co/luan-van/bai-toan-phat-hien-mat-nguoi-trong-anh-44827/>