**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**

**KHOA CÔNG NGHỆ THÔNG TIN**

--------------------------------



**ĐỒ ÁN TỐT NGHIỆP**

**NHẬN DIỆN BIỂN SỐ XE ĐỂ KIỂM TRA LƯU LƯỢNG THAM GIA GIAO THÔNG**

Sinh viên thực hiện: Trần Trung Kiên

Lớp/Khóa: CNTTK17K-KHMT

Giáo viên hướng dẫn: Ths.Đinh Khánh Linh

**Thái Nguyên – 2023**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**

**KHOA CÔNG NGHỆ THÔNG TIN**

--------------------------------



**ĐỒ ÁN TỐT NGHIỆP**

**NHẬN DIỆN BIỂN SỐ XE ĐỂ KIỂM TRA LƯU LƯỢNG THAM GIA GIAO THÔNG**

Sinh viên thực hiện: Trần Trung Kiên

Lớp/Khóa: CNTTK17K-KHMT

Giáo viên hướng dẫn: Ths.Đinh Khánh Linh

**Giáo viên hướng dẫn**  *(Ký và ghi họ tên)*

**Thái Nguyên – 2023**

**Mục Lục**

[CHƯƠNG 1: : GIỚI THIỆU 9](#_Toc128040887)

[1.1 Nội dung 9](#_Toc128040888)

[1.2 Nhiệm vụ đề tài 9](#_Toc128040889)

[1.3 Tính khoa học và cấp thiết của đề tài. 9](#_Toc128040890)

[1.3.1 Một số khái niệm cơ bản trong xử lý ảnh 10](#_Toc128040891)

[1.3.2 Tổng quan về một hệ thống xử lý ảnh 11](#_Toc128040892)

[1.3.3 Một số thao tác cơ bản trong xử lý ảnh 12](#_Toc128040893)

[1.4 Biểu diễn ảnh trong máy tính 15](#_Toc128040894)

[1.4.1 Ảnh màu 15](#_Toc128040895)

[1.4.2 Ảnh xám 17](#_Toc128040896)

[1.5 Phép tích chập trong xử lý ảnh 18](#_Toc128040897)

[ Padding 19](#_Toc128040898)

[ Ý nghĩa của phép tính convolution 19](#_Toc128040899)

[1.5.1 Phương pháp số phân vùng ảnh. 20](#_Toc128040900)

[1.5.2 Phương pháp phân loại theo cấu trúc 23](#_Toc128040901)

[CHƯƠNG 2: BÀI TOÁN NHẬN DIỆN BIỂN SỐ XE 25](#_Toc128040902)

[2.1 Tổng quan 25](#_Toc128040903)

[2.1.1 Khái niệm biển số xe 25](#_Toc128040904)

[2.1.2 Xử lý ảnh và Open CV 25](#_Toc128040905)

[2.1.3 Hướng giải quyết 26](#_Toc128040906)

[2.2 PHÁT HIỆN VỊ TRÍ VÀ TÁCH BIỂN SỐ XE 27](#_Toc128040907)

[2.2.1 Hướng giải quyết 27](#_Toc128040908)

[2.2.2 Chuyển ảnh xám 27](#_Toc128040909)

[2.2.3 Tăng độ tương phản 28](#_Toc128040910)

[2.2.4 Giảm nhiễu bằng bộ lọc Gauss 32](#_Toc128040911)

[2.2.5 Nhị phân hóa với ngưỡng động (Adaptive Threshold) 34](#_Toc128040912)

[2.2.6 Phát hiện cạnh Canny (Canny Edge Detection) 35](#_Toc128040913)

[2.2.7 Lọc biển số với contour 38](#_Toc128040914)

[2.3 PHÂN ĐOẠN KÍ TỰ 43](#_Toc128040915)

[2.3.1 Hướng giải quyết 43](#_Toc128040916)

[2.3.2 Xoay biển số 43](#_Toc128040917)

[2.3.3 Tìm vùng đối tượng 45](#_Toc128040918)

[2.3.4 Tìm và tách kí tự 46](#_Toc128040919)

[2.4 NHẬN DIỆN KÍ TỰ 47](#_Toc128040920)

[2.4.1 Tổng quan về lý thuyết 47](#_Toc128040921)

[2.4.2 Hướng giải quyết 52](#_Toc128040922)

[CHƯƠNG 3: KẾT QUẢ THỰC HIỆN 54](#_Toc128040923)

[3.1 Cách thức đo đạc, thử nghiệm 54](#_Toc128040924)

[3.2 Kết quả và giải thích 54](#_Toc128040925)

[KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 59](#_Toc128040926)

[Kết luận 59](#_Toc128040927)

[Hướng phát triển 59](#_Toc128040928)

[TÀI LIỆU THAM KHẢO 60](#_Toc128040929)

**DANH MỤC CÁC HÌNH ẢNH**

[hình 1.1: Các giai đoạn chính trong xử lý ảnh 12](#_Toc128040773)

[hình 1.2:: minh họa RGB 15](#_Toc128040774)

[hình 1.3:Ví dụ về ảnh màu 16](file:///C:\Users\The%20Kaye\Downloads\DO_AN_TRANTRUNGKIEN.docx#_Toc128040775)

[hình 1.4:biểu diễn theo tensor 3 chiều 17](file:///C:\Users\The%20Kaye\Downloads\DO_AN_TRANTRUNGKIEN.docx#_Toc128040776)

[Hình 1.5: ví dụ ảnh xám 17](file:///C:\Users\The%20Kaye\Downloads\DO_AN_TRANTRUNGKIEN.docx#_Toc128040777)

[Hình 1.6 : minh họa phép tích chập 18](file:///C:\Users\The%20Kaye\Downloads\DO_AN_TRANTRUNGKIEN.docx#_Toc128040778)

[Hình 1.7: ma trận đầu ra 19](file:///C:\Users\The%20Kaye\Downloads\DO_AN_TRANTRUNGKIEN.docx#_Toc128040779)

[Hình 1.8: Stride=2, padding= 19](#_Toc128040780)

[Hình 1.9: Một số bộ lọc Kerne1 trong xử lý ảnh 20](file:///C:\Users\The%20Kaye\Downloads\DO_AN_TRANTRUNGKIEN.docx#_Toc128040781)

[Hình 1.10: phương pháp lưới 21](#_Toc128040782)

[Hình 1.11: phương pháp cung 21](#_Toc128040783)

[Hình 1.12: Biểu diễn mẫu bằng tập kí hiệu 24](#_Toc128040784)

[Hình 2.1 26](#_Toc128040785)

[Hình 2.2 27](#_Toc128040786)

[Hình 2.3 28](#_Toc128040787)

[Hình 2.4 29](#_Toc128040788)

[Hình 2.5 29](#_Toc128040789)

[Hình 2.6 30](#_Toc128040790)

[Hình 2.7 30](#_Toc128040791)

[Hình 2.8 31](#_Toc128040792)

[Hình 2.9 31](#_Toc128040793)

[Hình 2.10 32](#_Toc128040794)

[Hình 2.11 32](#_Toc128040795)

[Hình 2.12 33](#_Toc128040796)

[Hình 2.13 34](#_Toc128040797)

[Hình 2.14 35](#_Toc128040798)

[Hình 2.15 36](#_Toc128040799)

[Hình 2.16 37](#_Toc128040800)

[Hình 2.17: Lọc ngưỡng 37](#_Toc128040801)

[Hình 2.18: Ảnh sau khi phát hiện biên Canny 38](#_Toc128040802)

[Hình 2.19 : Thuật toán Square Tracing 39](#_Toc128040803)

[Hình 2.20 : Thuật toán Square Tracing chạy đúng 39](#_Toc128040804)

[Hình 2.21: 3 Thuật toán Square Tracing chạy sai 39](#_Toc128040805)

[Hình 2.22 Thuật toán Moore - Neighbor. 40](#_Toc128040806)

[Hình 2.235 Vẽ Contour với OpenCV 41](#_Toc128040807)

[Hình 2.24: Contour chưa xấp xỉ đa giác 41](#_Toc128040808)

[Hình 2.25 : Contour đã xấp xỉ đa giác 41](#_Toc128040809)

[Hình 2.26: Các bước chính trong phân đoạn kí tự 43](#_Toc128040810)

[Hình 2.27: Ảnh biển số chưa xoay 44](#_Toc128040811)

[Hình 2.28 : Ảnh biển số đã xoay 44](#_Toc128040812)

[Hình 2.29: Ảnh nhị phân 45](#_Toc128040813)

[Hình 2.30: Tìm vùng đối tượng 45](#_Toc128040814)

[Hình 2.31 Ảnh kí tự sau khi cắt 46](#_Toc128040815)

[Hình 2.32 : Các lớp trong AI 47](#_Toc128040816)

[Hình 2.33: phân loại Machine Learning 48](#_Toc128040817)

[Hình 2.34: Ví dụ về KNN 50](#_Toc128040818)

[Hình 2.35: Xét khả năng tiêu thụ của xe 50](#_Toc128040819)

[Hình 2.36: Dữ liệu sau chuẩn hóa 51](#_Toc128040820)

[Hình 2.37:Sau khi tính khoảng cách và xếp hạng 51](#_Toc128040821)

[Hình 2.38: Tập dữ liệu huấn luyện 52](#_Toc128040822)

[Hình 2.39 : Biển số trước khi nhận diện 53](#_Toc128040823)

[Hình 2.40: Biển số sau khi nhận diện 53](#_Toc128040824)

[Hình 2.41: Biển số xe được in ra trên hình gốc 53](#_Toc128040825)

[Hình 3.1: không tìm thấy kí tự 55](#_Toc128040826)

[Hình 3.2: ảnh gốc 55](#_Toc128040827)

[Hình 3.3 :tìm tháy 5 kí tự 55](#_Toc128040828)

[Hình 3.4: tìm thấy 9 kí tự 55](#_Toc128040829)

[Hình 3.5: Lấy ngưỡng cao/rộng 1.5 56](#_Toc128040830)

[Hình 3.6 : Lấy ngưỡng cao/rộng 1.4 56](#_Toc128040831)

[Hình 3.7 : Lỗi xấp xỉ Contour 56](#_Toc128040832)

[Hình 3.8: Ảnh nhị phân bị đứt 57](#_Toc128040833)

[Hình 3.9: Đường contour bị đứt đoạn 57](#_Toc128040834)

[Hình 3.10: ảnh gốc với 3 biển số 58](#_Toc128040835)

[Hình 3.11: biển 1 58](#_Toc128040836)

[Hình 3.12: biển 2 58](#_Toc128040837)

[Hình 3.13: biển 3 58](#_Toc128040838)

[Hình 3.14:Không khoanh được vùng kí tự 58](#_Toc128040839)

**LỜI CẢM ƠN**

Để hoàn thành báo cáo đồ án tốt nghiệp này trước hết em xin chân thành cảm ơn quý thầy, cô trong khoa Công nghệ thông tin, Trường Đại Học Công nghệ thông tin và Truyền thông – Đại học Thái Nguyên đã tận tình truyền đạt kiến thức trong những năm em học tập. Với vốn kiến thức được tiếp thu trong quá trình học không chỉ là nền tảng mà còn là hành trang quý báu để em bước vào đời một cách vững chắc và tự tin.

Đặc biệt, em xin gửi đến thầy **Th.s Đinh Khánh Linh**– người đã tận tình hướng dẫn, đưa ra những lời nhận xét, những kinh nghiệm quý báu, những kinh nghiệm đó sẽ giúp em hoàn thiện bản thân hơn trong môi trường làm việc sau này.

Trong quá trình thực hoàn thành đồ án cũng như là trong quá trình làm bài báo cáo đồ án, còn nhiều sai sót, rất mong các thầy, cô bỏ qua. Đồng thời do trình độ lý luận cũng như kinh nghiệm thực tiễn còn hạn chế nên bài báo cáo không thể tránh khỏi những thiếu sót, em rất mong nhận được ý kiến đóng góp thầy, cô để em học thêm được nhiều kinh nghiệm và sẽ hoàn thành tốt hơn bài báo cáo tốt nghiệp sắp tới.

Một lần nữa em xin chân thành cảm ơn!

|  |  |
| --- | --- |
|  | Thái Nguyên, ngày tháng năm 2023  Sinh viên thực hiện |

**Trần Trung Kiên**

**LỜI CAM ĐOAN**

Tôi xin cam đoan đồ án này do chính tôi thực hiện, dưới sự hướng dẫn khoa học của **Th.s Đinh Khánh Linh**, các kết quả lý thuyết được trình bày trong đồ án là sự tổng hợp từ các kết quả đã được công bố và có trích dẫn đầy đủ, kết quả của chương trình thực nghiệm trong đồ án này được tác giả thực hiện là hoàn toàn trung thực, nếu sai tôi hoàn toàn chịu trách nhiệm.

*Thái Nguyên, tháng năm 2023*

**Học viên**

**Trần Trung Kiên**

# : GIỚI THIỆU

## **Nội dung**

* + Tìm hiểu về biển số xe và hệ thống nhận dạng biển số xe .
  + Phát biểu bài toán và hướng giải quyết .
  + Nghiên cứu một số thuật toán xử lý ảnh và nhận dạng kí tự ứng dụng trong việc nhận dạng biển số xe.
  + Áp dụng bài toán nhận diện biển số xe vào quản lý lưu lượng tham gia giao thông.

## Nhiệm vụ đề tài

Từ nội dung nêu trên, đề tài của em sẽ bao gồm các nhiệm vụ sau:

* + Tìm hiểu khái quát về xử lý ảnh và bài toán nhận dạng biển số xe.
  + Tìm hiểu thông tin về biển số xe và phân loại biển số xe của Việt Nam.
  + Tìm hiểu các công đoạn chính của bài toán nhận dạng biển số xe gồm 3 khâu chính:
    - Phát hiện vị trí và tách biển số xe.
    - Phân đoạn kí tự trong biển số xe.
    - Nhận dạng kí tự.
  + Cài đặt thử nghiệm.

## Tính khoa học và cấp thiết của đề tài.

Ứng dụng của công nghệ phân loại hiện nay đang phát triển rất mạnh ở rất nhiều lĩnh vực như: học thuật, kinh doanh, bảo mật, y tế... và các ở các đối tượng như: nhà nghiên cứu xã hội, chính phủ và các tổ chức phi lợi nhuận khác. Vì các tổ chức này sở hữu một lượng lớn dữ liệu không có cấu trúc và việc xử lý dữ liệu sẽ trở nên dễ dàng hơn rất nhiều nếu như các dữ liệu này được chuẩn hóa bởi các chủ đề/nhãn. Nền tảng công nghệ để thực hiện bài toán phân loại chính là trí tuệ nhân tạo (Artificial Intelligence – AI) và học sâu (Deep Learning - DL).

Trong ngành Thị giác máy tính (Computer Vision - CV), nhờ những thành tựu của lĩnh vực học sau mà trong những năm gần đây, ta đã chứng kiến được nhiều thành tựu vượt bậc. Các hệ thống xử lý ảnh lớn như Facebook, Google hay Amazon đã đưa vào sản phẩm của mình những chức năng thông minh như nhận diện khuôn mặt người dùng, phát triển xe hơi tự lái hay drone giao hàng tự động.

Việc sử dụng thị giác máy tính vào quản lý giao thông là một ứng dụng mới nhưng rất cấp thiết tại nước ta hiện nay. Việc quản lý này có thể hạn chế tối đa việc trộm cắp phương tiện giao thông , gây tai nạn giao thông xong trốn chạy hay những hành vi tượng tự và vô cùng cấp thiết.

Được sự gợi ý của thầy giáo hướng dẫn tôi đã chọn đề tài: “Nhận diện biển số xe và quản lý lưu lượng tham gia giao thông” làm đồ án tốt nghiệp của mình. Mục tiêu chính của đồ án là tìm hiểu về bài toán phân nhận diện biển số xe.

* Nhưng bài toán nhận diện biến số xe là một hướng phát triển của xử lý ảnh nên ta vần phải hiểu về xử lý ảnh trước.

### Một số khái niệm cơ bản trong xử lý ảnh

* ***Ảnh số***

Ảnh số thực tế là biểu diễn số học của hình ảnh trong máy tính, thường là biểu diễn nhị phân. Có thể phân ảnh số thành 2 loại: ảnh xám và ảnh màu.

Ảnh xám thực chất là một hàm hai chiều của cường độ sáng f(x,y), trong đó x và y là các toạ độ không gian và giá trị của hàm f tại một điểm (x,y) tỷ lệ với cường độ sáng của ảnh tại điểm đó. Nếu chúng ta có một ảnh mầu thì f là một vector mà mỗi thành phần của vector đó chỉ ra cường độ sáng của ảnh tại điểm (x,y) đó tương ứng với dải mầu [[2]](#_bookmark122) .

Mỗi thành phần của mảng (x,y) được gọi là một điểm ảnh (pixel: picture element) và là phần tử nhỏ nhất cấu tạo nên ảnh. Điểm ảnh được hiểu như 1 dấu hiệu hay cường độ sáng tại một tọa độ xác định trong không gian. Hình ảnh được xem như là 1 tập hợp các điểm. Với cùng kích thước nếu sử dụng càng nhiều điểm ảnh thì bức ảnh càng đẹp, càng mịn và càng thể hiện rõ hơn chi tiết của ảnh người ta gọi đặc điểm này là độ phân giải.

#### Cường độ sáng của một ảnh tại một ví trí điểm ảnh

Mỗi điểm ảnh của một ảnh tương ứng với một phần của một đối tượng vật lý tồn tại trong thế giới thực. Đối tượng vật lý này được chiếu sáng bởi một vài tia sáng mà tia sáng này bị phản xạ một phần hay hấp thụ một phần khi chiếu lên đối tượng vật lý đó. Phần ánh sáng phản xạ lại đi tới các bộ cảm biến được sử dụng để tạo ảnh cảm nhận và tạo ra các giá trị ghi nhận được đối tượng đối với từng điểm ảnh. Giá trị thu nhận được phụ thuộc vào phổ ánh sáng phản xạ. Giá trị cường độ sáng của các điểm ảnh khác nhau chỉ có ý nghĩa tương đối mà không có ý nghĩa trong các toán hạng tuyệt đối [[2]](#_bookmark122) .

#### Số bits cần thiết để lưu trữ một ảnh

Ở đây chúng ta chỉ quan tâm tới ảnh xám, nếu ảnh được lưu trữ dưới dạng một mảng hai chiều với kích thước NxN và có 2m mức xám thì số bits cần thiết để lưu trữ ảnh là:

*b*  *N x N x* 2*m*

(1.1)

#### Độ phân giải ảnh

Độ phân giải ảnh biểu diễn mức độ chi tiết của ảnh mà chúng ta có thể nhìn rõ đối tượng. Khi thay đổi các giá trị m và N trong phương trình thì sẽ có các hiện tượng thay đổi khác nhau. Xong thực nghiệm cho thấy khi giữ nguyên kích thước ảnh N và tăng số mức xám m lên thì sẽ thể hiện rõ hơn mức độ chi tiết trong ảnh.

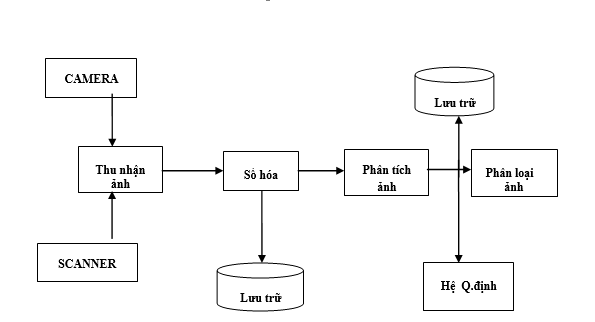
### Tổng quan về một hệ thống xử lý ảnh

Xử lý ảnh là đối tượng nghiên cứu của lĩnh vực thị giác máy, là quá trình biến đổi từ một ảnh ban đầu sang một ảnh mới với các đặc tính và tuân theo ý muốn của người sử dụng. Xử lý ảnh có thể gồm quá trình phân tích, phân lớp các đối tượng, làm tăng chất lượng, phân đoạn và tách cạnh, gán nhãn cho vùng hay quá trình biên dịch các thông tin hình ảnh của ảnh [[2]](#_bookmark122) .

Cũng như xử lý dữ liệu bằng đồ hoạ, xử lý ảnh số là một lĩnh vực của tin học ứng dụng. Xử lý dữ liệu bằng đồ hoạ đề cập đến những ảnh nhân tạo, các ảnh này được xem xét như là một cấu trúc dữ liệu và được tạo ra bởi các chương trình. Xử lý ảnh số bao gồm các phương pháp và kĩ thuật để biến đổi, để truyền tải hoặc mã hoá các ảnh tự nhiên. Mục đích của xử lý ảnh gồm:

* + - * Biến đổi ảnh, làm tăng chất lượng ảnh.
      * Tự động nhận dạng, đoán nhận, đánh giá các nội dung của ảnh.

Các bước cần thiết trong xử lý ảnh được mô tả chi tiết bao gồm các bước sau:



hình 1.1: Các giai đoạn chính trong xử lý ảnh

|  |
| --- |
| Đầu tiên là quá trình thu nhận ảnh. Ảnh có thể thu nhận được qua camera. Thường khi thu nhận ảnh qua camera là tín hiệu tương tự (loại camera ống kiểu CCIR), nhưng cũng có thể là tín hiệu số hóa (loại CCD- Charge Coupled Device). Ảnh cũng có thể thu nhận từ vệ tinh qua các bộ cảm ứng (sensor), hay ảnh tranh được quét trên scanner. Tiếp theo là quá trình số hóa (Digitalizer) để biến đổi tín hiệu tương tự sang tín hiệu rời rạc (lấy mẫu) và số hóa bằng lượng hóa, trước khi chuyển sang giai đoạn xử lý, phân tích hay lưu trữ lại. Trước hết là công việc tăng cường ảnh để nâng cao chất lượng ảnh. Do những nguyên nhân khác nhau: có thể do chất lượng thiết bị thu nhận ảnh, do nguồn sáng hay do nhiễu, ảnh có thể bị suy biến do vậy cần phải tăng cường và khôi phục lại ảnh để làm nổi bật một số đặc tính chính của ảnh, hay làm cho ảnh gần giống nhất với trạng thái gốc – trạng thái trước khi bị biến dạng. Giai đoạn tiếp theo là phát hiện các đặc tính như biên, phân vùng ảnh, trích chọn các đặc tính...v.v...  Cuối cùng tùy theo mục đích của ứng dụng, sẽ là giai đoạn nhận dạng, phân loại hay các quyết định khác. |

### Một số thao tác cơ bản trong xử lý ảnh

* ***Biểu diễn ảnh***

Trong biểu diễn ảnh, người ta thường dùng các phần tử đặc trưng của ảnh là pixel. Nhìn chung có thể một hàm hai biến chứa các thông tin như biểu diễn của một

ảnh. Các mô hình biểu diễn cho ta một mô tả logic hay định lượng các tính chất của hàm này. Trong biểu diễn ảnh cần chú ý đến tính trung thực hoặc các tiêu chuẩn “thông minh” để đo chất lượng ảnh hoặc tính hiệu quả của các kĩ thuật xử lý.

Một số mô hình thường được dùng trong biểu diễn ảnh: mô hình bài toán, mô hình thống kê. Trong mô hình bài toán, ảnh hai chiều được biểu diễn nhờ các hàm hai biến trực giao gọi là các hàm cơ sở. Còn mô hình thống kê, một ảnh được coi như một phần tử của một tập hợp đặc trưng bởi các đại lượng như: kỳ vọng toán học, hiệp biến, phương sai, moment.

#### Biến đổi ảnh (Image Transform)

Thuật ngữ biến đổi ảnh thường dùng để nói tới một lớp các ma trận đơn vị và các kĩ thuật dùng để biến đổi ảnh.

Biến đổi ảnh nhằm làm giảm các nguyên nhân của ảnh để việc xử lý hiệu quả hơn. Như làm rõ hơn các thông tin mà người dùng quan tâm nhưng người dùng phải chấp nhận mất đi một số thông tin cần thiết.

#### Phân tích ảnh

Phân tích ảnh liên quan đến việc xác định các độ đo định lượng của 1 ảnh để đưa ra một mô tả đầy đủ về ảnh.

Quá trình phân tích ảnh thực chất bao gồm nhiều công đoạn nhỏ. Trước hết là công việc tăng cường ảnh để nâng cao chất lượng ảnh, giai đoạn tiếp theo là phát hiện các đặc tính như phát hiện biên, phân vùng ảnh, trích chọn các đặc tính..v.v..

#### Tăng cường ảnh – khôi phục ảnh

Tăng cường ảnh là một bước quan trọng, tạo tiền đề cho xử lý ảnh. Nó gồm các kỹ thuật như: lọc độ tương phản, khử nhiễu, nổi màu…

Khôi phục ảnh là nhằm loại bỏ các suy giảm trong ảnh.

#### Xử lý biên ảnh

Biên là vấn đề chủ yếu trong phân tích ảnh vì các điểm trích chọn trong quá trình phân tích ảnh đều dựa vào biên. Mỗi điểm ảnh có thể là biên nếu ở đó có sự thay

đổi đột ngột về mức xám. Tập hợp các điểm biên tạo thành biên hay đường bao quanh của ảnh.

#### Phân vùng ảnh

Phân vùng là bước then chốt trong xử lý ảnh. Giai đoạn này nhằm phân tích ảnh thành những thành phần có tính chất nào đó dựa theo biên hay các vùng liên thông. Tiêu chuẩn để xác định các vùng liên thông có thể là mức xám, cùng màu hay độ tương phản.

#### Nhận dạng ảnh

Nhận dạng ảnh là quá trình liên quan đến các mô tả đối tượng mà người ta muốn đặc tả nó. Quá trình nhận dạng thường đi sau quá trình trích chọn các đặc tính chủ yếu của đối tượng. Có hai kiểu mô tả đối tượng:

Mô tả tham số (nhận dạng theo tham số).

Mô tả theo cấu trúc (nhận dạng theo cấu trúc).

Trên thực tế người ta đã áp dụng kỹ thuật nhận dạng khá thành công với nhiều đối tượng khác nhau như: nhận dạng ảnh vân tay, nhận dạng chữ viết.

#### Nén ảnh

Dữ liệu ảnh cũng như các dữ liệu khác cần phải lưu trữ hay truyền đi trên mạng mà lượng thông tin để biểu diễn cho một ảnh là rất lớn. Do đó làm giảm lượng thông tin hay nén dữ liệu là một nhu cầu cần thiết.

Nén dữ liệu là quá trình làm giảm lượng thông tin “dư thừa” trong dữ liệu gốc và do vậy lượng thông tin thu được sau khi nén thường nhỏ hơn dữ liệu gốc rất nhiều.

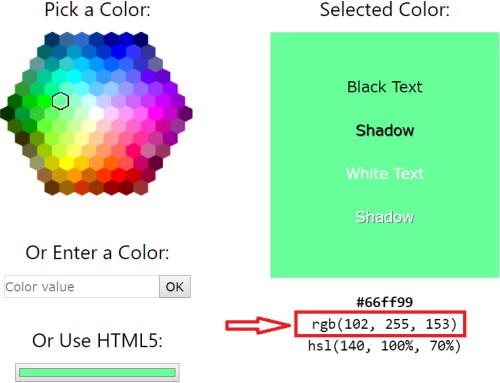
## Biểu diễn ảnh trong máy tính

### 

### Ảnh màu

* ***Hệ màu RGB***

RGB viết tắt của red (đỏ), green (xanh lục), blue (xanh lam), là ba màu chính của ánh sáng khi tách ra từ lăng kính. Khi trộn ba màu trên theo tỉ lệ nhất định có thể tạo thành các màu khác nhau.



hình 1.2:: minh họa RGB

Hình 2.2 minh họa việc chọn màu thường thấy trong các chương trình máy tính. Khi ta chọn một màu thì sẽ ra một bộ ba số tương ứng (r,g,b) màu được chọn. Ở đây là rgb(102, 255, 153), nghĩa là r=102, g=255, b=153.

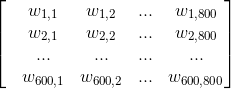
#### Biểu diễn ảnh màu

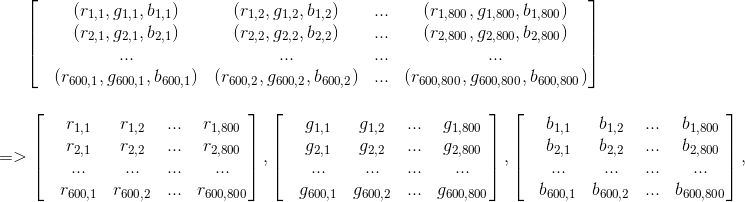


hình 1.3:Ví dụ về ảnh màu

Hình 2.3 là một ma trận các pixel mà mỗi pixel biểu diễn một điểm màu. Mỗi điểm màu được biểu diễn bằng bộ 3 số (r,g,b). Để tiện cho việc xử lý ảnh thì sẽ tách ma trận pixel ra 3 channel red, green, blue.

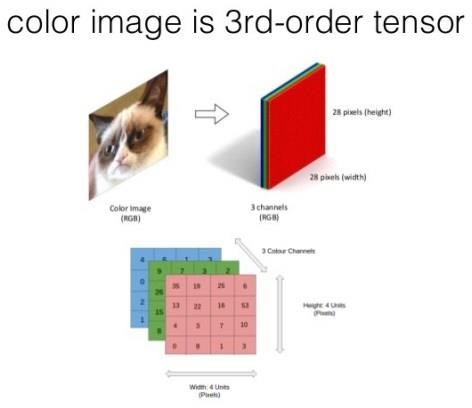
Bức ảnh trên [Hình 1.3](#_bookmark18) có kích thước 800 pixel \* 600 pixel, bức ảnh này có thể biểu diễn dưới dạng một ma trận kích thước 600 \* 800

(1.2)

Trong đó mỗi phần tử w*ij* là một pixel. Tuy nhiên để biểu diễn 1 màu ta cần 3 thông số (r,g,b) nên gọi w*ij*  (r*ij* , g*ij* , b*ij* ) ta có thể để biểu diễn dưới dạng ma trận như sau:

Mỗi ma trận được tách ra (r, g, b) được gọi là 1 channel nên ảnh màu được gọi là 3 channel: channel red, channel green, channel blue.

Ảnh màu trên máy tính sẽ được biểu diễn dưới dạng tensor 3 chiều chồng lên nhau. [Hình 1.4](#_bookmark20) mô tả biểu diễn một ảnh màu kích thước 28\*28 trên máy tính. Trong đó, ảnh được biểu diễn dưới dạng tensor 3 chiều kích thước 28\*28\*3 do có 3 ma trận (channel) màu red, green, blue kích thước 28\*28 chồng lên nhau.



hình 1.4:biểu diễn theo tensor 3 chiều

### Ảnh xám



Hình 1.5: ví dụ ảnh xám

Hình 2.5 mô tả ảnh xám của bắc ảnh màu trong [Hình 1.3](#_bookmark18). Tương tự ảnh màu, ảnh xám cũng có kích thước 800 pixel \* 600 pixel, có thể biểu diễn dưới dạng một ma trận kích thước 600 \* 800 như [(1.2)](#_bookmark19).

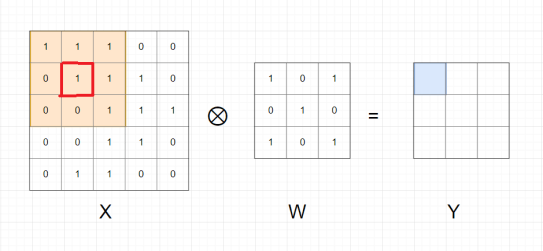
Tuy nhiên mỗi pixel trong ảnh xám chỉ cần biểu diễn bằng một giá trị nguyên trong khoảng từ [0,255] thay vì (r,g,b) như trong ảnh màu. Giá trị 0 là màu đen, 255 là màu trắng và giá trị pixel càng gần 0 thì càng tối và càng gần 255 thì càng sáng. Do đó khi biểu diễn ảnh xám trong máy tính chỉ cần một ma trận là đủ.

## Phép tích chập trong xử lý ảnh

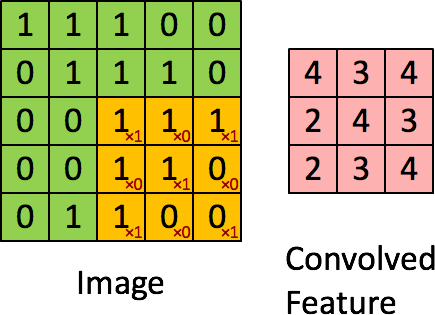
Phép tích chập (Convolution) là kỹ thuật quan trọng trong xử lý ảnh, được sử dụng chính yếu trong các phép toán trên ảnh như: đạo hàm ảnh, làm trơn ảnh, trích xuất biên cạnh trong ảnh…Kí hiệu phép tính convolution là : *Y*  *X* *W*

Theo toán học, tích chập là phép toán tuyến tính, cho ra kết quả là một hàm bằng việc tính toán dựa trên hai hàm đã có (X và W). Để cho dễ hình dung mình sẽ lấy ví dụ trên ảnh xám, tức là ảnh được biểu diễn dưới dạng ma trận X kích thước m\*n [[2]](#_bookmark122) .

Công thức tích chập giữa hàm ảnh X(x, y) và bộ lọc W(x, y) (kích thước mxn)

kernel thường lẻ) gọi là ma trận A. Sau đó tính tổng các phần tử của phép tính element-wise của ma trận A và ma trận W, rồi viết vào ma trận kết quả Y.

Hình 1.6 : minh họa phép tích chập

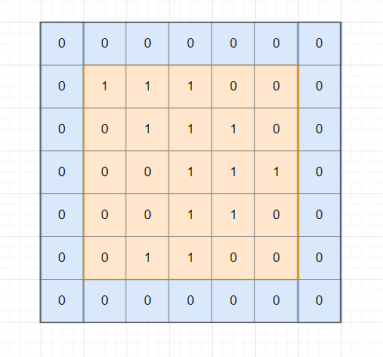


Hình 1.7: ma trận đầu ra

Làm tương tự với các phần tử còn lại trong ma trận ta thu được kết quả là ma trận Y đầu ra ([Hình 1.7](#_bookmark26)). Ma trận Y này có kích thước nhỏ hơn ma trận X. Kích thước của ma trận Y là (m-k+1) \* (n-k+1).

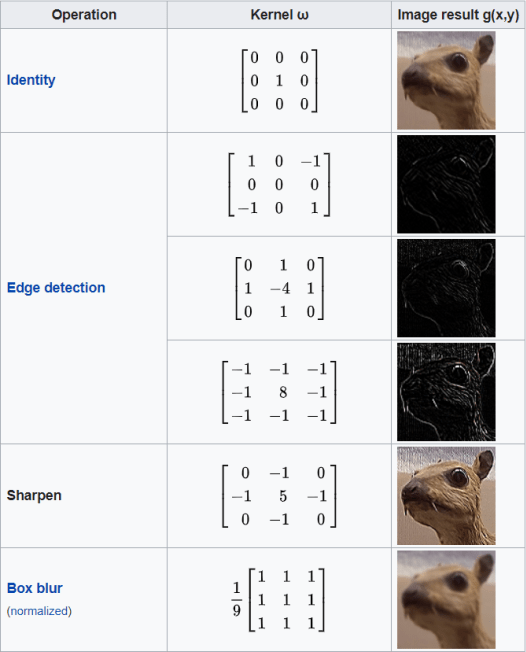
### Padding

Như đã phân tích ở trên, mỗi lần thực hiện phép tính convolution xong, kích thước ma trận Y đều nhỏ hơn X. Nếu muốn ma trận Y thu được có kích thước bằng ma trận X ta phải thêm giá trị 0 ở viền ngoài ma trận X. Phép tính này gọi là convolution với padding=1. Padding=k nghĩa là thêm k vector 0 vào mỗi phía của ma trận.



Hình 1.8: Stride=2, padding=

### Ý nghĩa của phép tính convolution



Hình 1.9: Một số bộ lọc Kerne1 trong xử lý ảnh

Mục đích của phép tính convolution trên ảnh là làm mở, làm nét ảnh; xác định các đường;… Mỗi kernel khác nhau thì sẽ phép tính convolution sẽ có ý nghĩa khác nhau. Hình 2.9 minh họa một số bộ lọc Kernel được xử dụng trong các thao tác xử lý ảnh như làm nét ảnh, xác định biên ảnh, làm mờ ảnh.

### Phương pháp số phân vùng ảnh.

Phương pháp số phân vùng ảnh biểu diễn các mẫu dưới dạng các giá trị số và quá trình phân lớp tập mẫu chính là quá trình thực hiện việc sắp xếp các giá trị số này thành từng lớp riêng biệt.

* + - * Trích chọn đặc trưng mẫu:

Nhiệm vụ đặt ra cho bước trích chọn đặc trưng mẫu là phải rút ra được các đặc trưng riêng của từng đối tượng ảnh trong tập mẫu. Sau đó mỗi đặc trưng của đối tượng được mô tả bằng các giá trị số và các giá trị này sẽ tạo thành vectơ mô tả tập mẫu. Để tìm ra các đặc trưng riêng của đối tượng ta có thể xét đến các đặc trưng đơn giản như đặc trưng về hình học, topo...Ngoài ra có thể dùng một số phương pháp đặc biệt để có thể phát hiện được các đặc trưng phức tạp hơn của đối tượng đặc biệt là với các trường hợp mẫu có hình dạng phức tạp. Có thể kể ra đây một số phương pháp như sau:

* + - * Phương pháp lưới:

3 điểm cắt

3 ô trùm

5 ô trùm

2 điểm cắt

Hình 1.10: phương pháp lưới

Một lưới vuông chuẩn được chùm lên đối tượng. Số lượng điểm cắt của mỗi nan lưới ngang và dọc với đối tượng sẽ là đặc trưng của đối tượng. Đồng thời số lượng mắt lưới vuông có trùm lên đối tượng theo từng chiều dọc và ngang cũng được sử dụng làm đặc trưng của đối tượng. Mỗi đường dọc và ngang của lưới sẽ được gán cho một trọng số nhất định. Với phương pháp này, việc tiêu chuẩn hoá đối tượng rất quan trọng vì nó sẽ giúp cho việc xác định kích thước của lưới chuẩn được sử dụng

* + - * Phương pháp cung:

3 điểm cắt

*4* dải quạt phủ

*(vành khăn 2)*

Hình 1.11: phương pháp cung

Có thể loại bỏ ảnh hưởng của hướng đối tượng trong phương pháp lưới bằng cách thay thế lưới vuông chuẩn bằng lưới hình vành khăn, tức là các nan lưới là các đường tròn đồng tâm (có thể coi đây là phương pháp lưới dùng trong hệ toạ độ cực). Điểm tâm của các vòng tròn này chính là trọng tâm của đối tượng và ta cần xác định điểm này trước tiên. Từ điểm này ta sẽ kẻ các đường bán kính chuẩn chia đều các đường tròn thành các cung. Số lượng các điểm cắt với đối tượng dọc theo một bán kính sẽ là đặc trưng của đối tượng. Tương tự như phương pháp lưới, số lượng cung của mỗi vòng tròn phủ lên đối tượng xét cũng coi như là đặc trưng của đối tượng. Mỗi bán kính và vành khăn sẽ được gán một trọng số.

* + - * Kỹ thuật phân lớp mẫu:

Các đặc trưng của đối tượng được biểu diễn bởi các giá trị số và các giá trị này được xem là các thành phần của các vectơ biểu diễn mẫu. Khi ta đưa vào hệ thống một tập các mẫu chuẩn thì quá trình trích chọn đặc trưng sẽ tạo nên các vectơ mẫu chuẩn phân bố trong không gian mẫu và với mỗi vectơ mẫu chuẩn này thì ta biết được ánh xạ từ nó sang không gian diễn dịch, tức là biết tên của nó. Như vậy các vectơ mẫu chuẩn đã được phân thành các lớp mà mỗi lớp ứng với một tên. Những lớp này ta gọi là lớp chuẩn.

Khi đưa các mẫu chưa xác định (mẫu cần phân loại) vào hệ thống thì việc nhận dạng mẫu chính là tìm ra một quy tắc để sắp xếp vectơ biểu diễn mẫu đó vào một lớp chuẩn nào đó. Để có thể đạt được mục đích này thì trước hết cần phải tạo được một sự phân định rõ ràng giữa các lớp chuẩn, tức là trong không gian mẫu phải có một sự phân hoạch rõ ràng. Trong thực tế thì không gian mẫu không phải lúc nào cũng đạt được đến sự phân tách hoàn toàn sau quá trình phân lớp mẫu. Nguyên nhân là do chúng ta chưa chọn được bộ đặc trưng tối ưu để phân tách đối tượng. Bởi vậy mà có thể xảy ra trường hợp một vectơ mẫu nào đó sẽ rơi vào vùng chồng lên nhau của 2 hay nhiều lớp mẫu chuẩn. Trong trường hợp này ta phải chọn lớp có xác suất cao hơn hoặc phải đánh dấu để chi ra rằng mẫu đó không phân lớp được. Quá trình xây dựng các lớp mẫu chuẩn như thế gọi là quá trình học. Việc xây dựng một thư viện mẫu chuẩn có vai trò rất quan trọng cho khả năng nhận dạng của hệ thống.

### Phương pháp phân loại theo cấu trúc

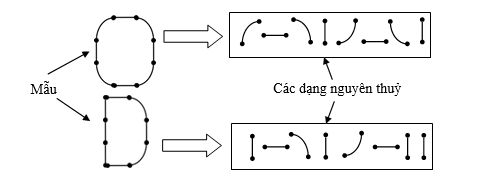
Bên cạnh phương pháp số, phương pháp cấu trúc là một trong những phương pháp truyền thống để nhận dạng mẫu. Trong khi ở phương pháp số, người ta thực hiện việc gán ý nghĩa cho các mẫu riêng biệt thì trong phương pháp cấu trúc lại xem xét các đối tượng như là các cấu trúc phức tạp tổ hợp từ các dạng nguyên thuỷ đơn giản và mối quan hệ giữa các dạng nguyên thuỷ này. Việc mô hình hoá các quá trình như vậy là tương đối khó khăn, bởi vậy mà phương pháp nhận dạng theo cấu trúc vẫn chưa được dùng phổ biến như phương pháp số.

* + - * Xây dựng bộ ký hiệu

Trước hết, ta cần xây dựng một tập các dạng nguyên thuỷ và các mẫu sẽ được biểu diễn bởi các dạng nguyên thuỷ này và mối quan hệ giữa các dạng nguyên thuỷ đó. Các dạng nguyên thuỷ phải được chọn sao cho khi dạng nguyên thuỷ này được sắp xếp theo một trật tự nào đó đối với nhau thì ta sẽ tạo ra được tất cả các dạng cấu trúc từ đơn giản đến phức tạp của tập các đối tượng cần nhận dạng. Các dạng nguyên thuỷ có thể được chọn ví dụ như đoạn thẳng, cung, điểm ngoặt, điểm kết thúc..

Để biểu diễn dạng nguyên thủy và quan hệ giữa chúng một cách thuận tiện ta dùng một bộ ký hiệu. Mỗi ký hiệu sẽ được đặc trưng cho một dạng nguyên thuỷ và một mẫu như vậy sẽ được biểu diễn bằng một chuỗi ký hiệu. Với việc biểu diễn dạng nguyên thuỷ bởi một bộ ký hiệu thì việc xử lý các mẫu sẽ đơn giản hơn rất nhiều.

* + - * Trích chọn đặc trưng cấu trúc:



Hình 1.12: Biểu diễn mẫu bằng tập kí hiệu

Quá trình trích chọn đặc trưng mẫu có thể hiểu là quá trình chuyển đổi tập mẫu sang các chuỗi ký hiệu. Tất cả các ký hiệu của một chuỗi phải được định nghĩa từ trước. Kết quả của quá trình trích chọn đặc trưng mẫu sẽ phân tập mẫu ra thành từng các phân lớp dựa theo cấu trúc của các chuỗi ký hiệu. Mỗi phân lớp sẽ được đặc trưng bằng một nguyên mẫu đại diện.

* + - * Kỹ thuật phân lớp mẫu:

Quá trình phân lớp mẫu là quá trình đánh giá sự tương tự của các mẫu với các nguyên mẫu đại diện cho từng phân lớp tức là ta phải so sánh chuỗi kí hiệu của mẫu với chuỗi kí hiệu của các nguyên mẫu đại diện. Căn cứ vào kết quả thu được ta sẽ phân chia được các mẫu mới vào từng phân lớp chuẩn.

# BÀI TOÁN NHẬN DIỆN BIỂN SỐ XE

## Tổng quan

### Khái niệm biển số xe

Ở [Việt Nam](https://vi.wikipedia.org/wiki/Vi%E1%BB%87t_Nam), biển kiểm soát xe cơ giới (hay còn gọi tắt là biển kiểm soát, biển số xe) là [tấm biển gắn trên mỗi xe cơ giới](https://vi.wikipedia.org/wiki/Bi%E1%BB%83n_%C4%91%C4%83ng_k%C3%BD_xe), được cơ quan [công an](https://vi.wikipedia.org/wiki/C%C3%B4ng_an_nh%C3%A2n_d%C3%A2n_Vi%E1%BB%87t_Nam) cấp (đối với xe quân sự do Bộ Quốc phòng cấp) khi mua xe mới hoặc chuyển nhượng xe. Biển số xe được làm bằng hợp kim [nhôm](https://vi.wikipedia.org/wiki/Nh%C3%B4m) [sắt](https://vi.wikipedia.org/wiki/S%E1%BA%AFt), có dạng [hình chữ nhật](https://vi.wikipedia.org/wiki/H%C3%ACnh_ch%E1%BB%AF_nh%E1%BA%ADt) hoặc hơi [vuông](https://vi.wikipedia.org/wiki/H%C3%ACnh_vu%C3%B4ng), trên đó có in [số](https://vi.wikipedia.org/wiki/S%E1%BB%91) và [chữ](https://vi.wikipedia.org/wiki/Ch%E1%BB%AF_vi%E1%BA%BFt) (biển xe dân sự không dùng các chữ cái I, J, O, Q, W. Chữ R chỉ dùng cho xe rơ-moóc, sơ-mi rơ-moóc) cho biết: Vùng và địa phương quản lý, các con số cụ thể khi tra trên máy tính còn cho biết danh tính người chủ hay đơn vị đã mua nó, thời gian mua nó phục vụ cho công tác an ninh, đặc biệt trên đó còn có hình [Quốc huy Việt Nam](https://vi.wikipedia.org/wiki/Qu%E1%BB%91c_huy_Vi%E1%BB%87t_Nam) dập nổi.

**Tiêu chuẩn về kích thước: Ở mỗi nước thường có tiêu chuẩn về** kích thước nhất định, còn riêng Việt Nam tỉ lệ kích thước giữa các biển số là gần như giống nhau. Biển số xe có 2 loại, kích thước như sau: Loại biển số dài có chiều cao 110 mm, chiều dài 470 mm; loại biển số ngắn có chiều cao 200 mm, chiều dài 280 mm nên ta sẽ giới hạn tỉ lệ cao/rộng là   
3.5 ≤ cao/rộng ≤ 6.5 (biển một hàng) và 0.8 ≤ cao/rộng ≤ 1.5 (biển hai hàng).

**Số lượng kí tự trong biển số xe nằm trong khoảng [7,9]. Chiều cao của chữ và số: 80mm, chiều rộng của chữ và số: 40mm.**

**Từ những đặc điểm trên ta có thể thiết lập nhưng thông số, điều khiển để lọc chọn những đối tương phù hợp mà ta cần.**

### Xử lý ảnh và Open CV

Xử lý ảnh là một phân ngành trong xử lý số tín hiệu với tín hiệu xử lý là ảnh. Đây là một phân ngành khoa học mới rất phát triển trong những năm gần đây. Xử lý ảnh gồm 4 lĩnh vực chính: xử lý nâng cao chất lượng ảnh, nhận dạng ảnh, nén ảnh và truy vấn ảnh. Sự phát triển của xử lý ảnh đem lại rất nhiều lợi ích cho cuộc sống của con người. Ngày nay xử lý ảnh đã được áp dụng rất rộng rãi trong đời sống như: photoshop, nén ảnh, nén video, nhận dạng biển số xe, nhận dạng khuôn mặt, nhận dạng chữ viết, xử lý ảnh thiên văn, ảnh y tế,....

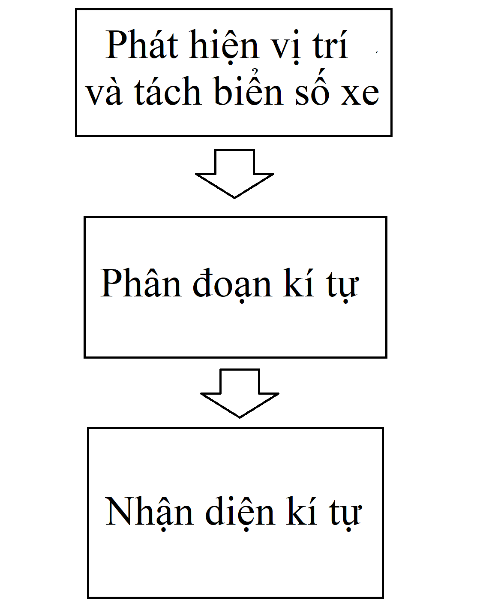
OpenCV (Open Computer Vision) là một thư viện mã nguồn mở hàng đầu cho xử lý về thị giác máy tính, machine learning, xử lý ảnh. OpenCV đươc viết bằng C/C++, vì vậy có tốc độ tính toán rất nhanh, có thể sử dụng với các ứng dụng liên quan đến thời gian thực. Opencv có các interface cho C/C++, Python Java vì vậy hỗ trợ được cho Window, Linux, MacOs lẫn Android, iOS OpenCV có cộng đồng hơn 47 nghìn người dùng và số lượng download vượt quá 6 triệu lần. Opencv có rất nhiều ứng dụng như:

* Nhận dạng ảnh
* Xử lý hình ảnh
* Phục hồi hình ảnh/video
* Thực tế ảo
* Các ứng dụng khác

### Hướng giải quyết

Hiện nay trên thế giới đã có rất nhiều cách tiếp cận khác nhau với việc nhận dạng biển số xe, tuy nhiên trong phạm vi đồ án này em sẽ giải quyết vấn đề theo 3 bước chính:

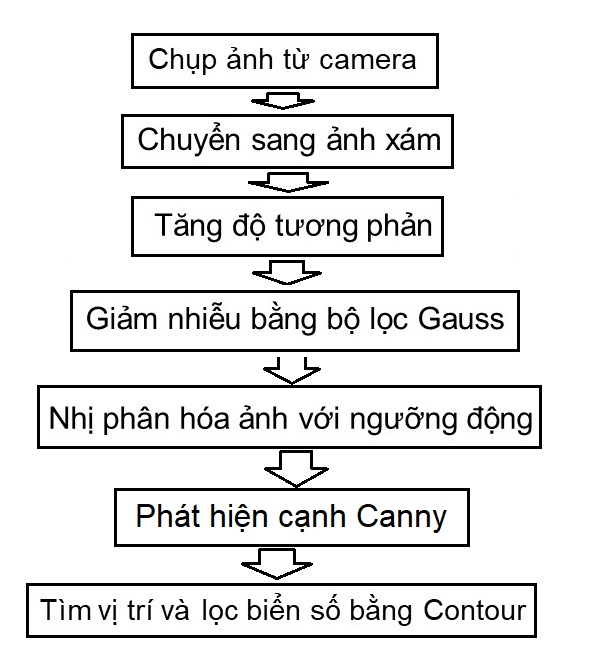
* 1. Phát hiện vị trí và tách biển số xe từ một hình ảnh có sẵn từ đầu vào là camera
  2. Phân đoạn các kí tự có trong biển số xe
  3. Nhận diện các kí tự đó rồi đưa về mã ASCII



Hình 2.1

## PHÁT HIỆN VỊ TRÍ VÀ TÁCH BIỂN SỐ XE

### Hướng giải quyết



Hình 2.2

Đầu tiên từ clip ta sẽ cắt từng frame ảnh ra từ clip đầu vào để xử lý, tách biển số. Ở phạm vi đồ án này, ý tưởng chủ yếu là nhận diện được biển số từ sự thay đổi đột ngột về cường độ ánh sáng giữa biển số và môi trường xung quanh nên ta sẽ loại bỏ các dữ liệu màu sắc RGB bằng cách chuyển sang ảnh xám. Tiếp theo ta tăng độ tương phản với hai phép toán hình thái học Top Hat và Black Hat để làm nổi bật thêm biển số giữa phông nền, hỗ trợ cho việc xử lý nhị phân sau này. Sau đó, ta giảm nhiễu bằng bộ lọc Gauss để loại bỏ những chi tiết nhiễu có thể gây ảnh hưởng đến quá trình nhận diện, đồng thời làm tăng tốc độ xử lý.

Việc lấy ngưỡng sẽ giúp ta tách được thông tin biển số và thông tin nền, ở đây em chọn lấy ngưỡng động (Adaptive Threshold). Tiếp đó ta sử dụng thuật toán phát hiện cạnh Canny để trích xuất những chi tiết cạnh của biển số. Trong quá trình xử lý máy tính có thể nhầm lẫn biển số với những chi tiết nhiễu, việc lọc lần cuối bằng các tỉ lệ cao/rộng hay diện tích của biển số sẽ giúp xác định được đúng biển số. Cuối cùng, ta sẽ xác định vị trí của biển số trong ảnh bằng cách vẽ Contour bao quanh.

### Chuyển ảnh xám

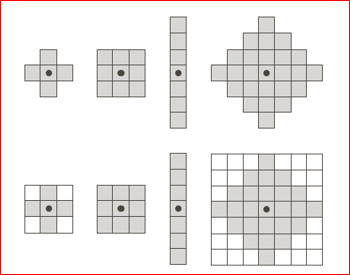
Ảnh xám (Gray Scale) đơn giản là một hình ảnh trong đó các màu là các sắc thái của màu xám với 256 cấp độ xám biến thiên từ màu đen đến màu trắng, nằm trong giải giá trị từ 0 đến 255, nghĩa là cần 8 bits hay 1 byte để biểu diễn mỗi điểm ảnh này. Lý do cần phải phân biệt giữa ảnh xám và các ảnh khác nằm ở việc ảnh xám cung cấp ít thông tin hơn cho mỗi pixel. Với ảnh thông thường thì mỗi pixel thường được cung cấp 3 trường thông tin trong khi với ảnh xám chỉ có 1 trường thông tin, việc giảm khối lượng thông tin giúp tăng tốc độ xử lý, đơn giản hóa giải thuật nhưng vẫn đảm bảo các tác vụ cần thiết

Ở bài này em sẽ chuyển ảnh xám từ hệ màu HSV thay vì RGB vì với không gian màu HSV ta có ba giá trị chính là: Vùng màu (Hue), độ bão hòa (Saturation), cường độ sáng (Value). Vì lý do đó không gian màu HSV thích nghi tốt hơn đối với sự thay đổi ánh sáng từ môi trường ngoài. Khi chuyển đổi, ảnh xám ta cần là ma trận các giá trị cường độ sáng tách ra từ hệ màu HSV.

### Tăng độ tương phản

#### Phép toán hình thái học

Hình thái học toán học là một lý thuyết và kỹ thuật để phân tích và xử lý cấu trúc hình học, hình ảnh đầu ra được xác định chủ yếu dựa vào phần từ cấu trúc (structuring elements/kernel)



Hình 2.3

Hình thái học toán học đã được phát triển cho hình ảnh nhị phân, và sau đó được mở rộng cho ảnh xám,... Đây là một trong những kỹ thuật được áp dụng trong giai đoạn tiền xử lý. Hai phép toán thường dùng là phép giãn nở (Dilation) và phép co (Erosion) . Từ hai phép toán cơ bản này người ta phát triển thành một số phép toán như phép đóng (Closing) và phép mở (Opening) và phép Top Hat, Black Hat.

1. Phép co

Phép toán co có ứng dụng trong việc giảm kích thước của đối tượng, tách rời các đối tượng gần nhau, làm mảnh và tìm xương của đối tượng.

|  |  |
| --- | --- |
|  |  |
| Ảnh gốc | Ảnh sau khi dùng phép co |

Hình 2.4

1. Phép giãn nở

Phép toán này có tác dụng làm cho đối tượng ban đầu trong ảnh tăng lên về kích thước (giãn nở ra). Ứng dụng của phép giãn nở là làm cho đối tượng trong ảnh được tăng lên về kích thước, các lỗ nhỏ trong ảnh được lấp đầy, nối liền đường biên ảnh đối với những đoạn rời nhỏ.

|  |  |
| --- | --- |
|  |  |
| Ảnh gốc | Ảnh sau khi dùng phép giãn nở |

Hình 2.5

1. Phép mở

Là thực hiện phép co trước sau đó mới thực hiện phép giãn nở. Phép toán mở được ứng dụng trong việc loại bỏ các phần lồi lõm và làm cho đường bao các đối tượng trong ảnh trở nên mượt mà hơn.

|  |  |
| --- | --- |
|  |  |
| Ảnh gốc | Ảnh sau khi dùng phép mở |

Hình 2.6

1. Phép đóng

Thực hiện phép giãn nở trước sau đó mới thực hiện phép co. Phép toán đóng được dùng trong ứng dụng làm trơn đường bao các đối tượng, lấp đầy các khoảng trống biên và loại bỏ những hố nhỏ.

|  |  |
| --- | --- |
|  |  |
| Ảnh gốc | Ảnh sau khi dùng phép đóng |

Hình 2.7

1. Phép Top Hat

Phép Top Hat là kết quả của phép trừ ảnh của ảnh ban đầu với ảnh sau khi thực hiện phép mở, dùng để làm nổi bật nhưng chi tiết trắng trong nền tối

|  |  |
| --- | --- |
|  |  |
| Ảnh gốc | Ảnh sau khi dùng phép Top Hat |

Hình 2.8

1. Phép Black Hat

Phép Black Hat là kết quả của phép trừ ảnh của ảnh sau khi thực hiện phép đóng với ảnh ban đầu. Dùng làm nổi bật chi tiết tối trong nền trắng.

|  |  |
| --- | --- |
|  |  |
| Ảnh gốc | Ảnh sau khi dùng phép Black Hat |

Hình 2.9

#### Tăng độ tương phản

Để làm tăng độ tương phản của biển số, em sử dụng chủ yếu hai phép Top Hat và Black Hat. Ý tưởng chung là ảnh đầu ra sẽ là ảnh gốc cộng thêm ảnh qua phép Top Hat và trừ đi ảnh qua phép Black Hat. Những chi tiết đã sáng sẽ sáng hơn và những chi tiết tối lại càng tối hơn, từ đó sẽ làm tăng độ tương phản cho biển số.

|  |  |
| --- | --- |
|  |  |
| Ảnh gốc | Ảnh sau khi tăng độ tương phản |

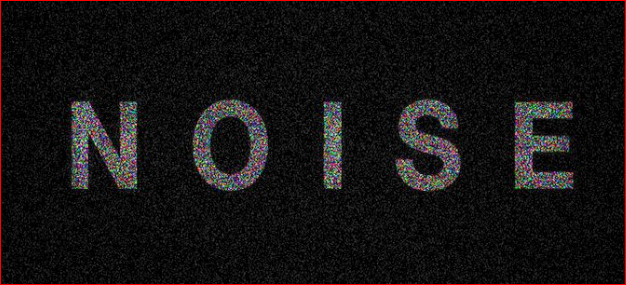
Hình 2.10

### Giảm nhiễu bằng bộ lọc Gauss

#### Nhiễu

Noise (nhiễu) được hiểu cơ bản là các dạng chấm hạt nhỏ phân bố trên hình ảnh. Noise có thể làm biến dạng các chi tiết trong ảnh khiến cho chất lượng ảnh thấp.

Trên thực tế có nhiều loại nhiễu, nhưng người ta thường chia làm ba loại: nhiễu cộng, nhiễu nhân và nhiễu xung. Bản chất của nhiễu thường tương ứng với tần số cao và cơ sở lý thuyết của bộ lọc là chỉ cho những tín hiệu có tần số nhất định đi qua, nên người ta thường sử dụng bộ lọc thông thấp hay trung bình.

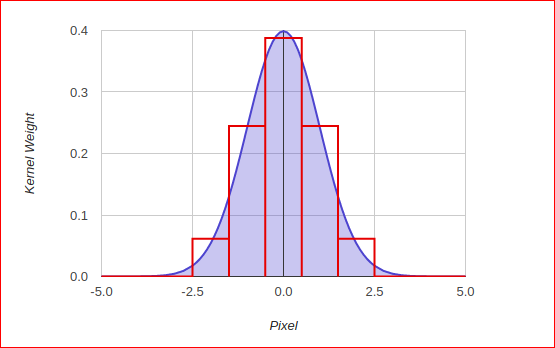


Hình 2.11

#### Bộ lọc Gauss (Gauss filter)

Bộ lọc Gauss được cho là bộ lọc hữu ích nhất, được thực hiện bằng cách nhân chập ảnh đầu vào với một ma trận lọc Gauss sau đó cộng chúng lại để tạo thành ảnh đầu ra.

Ý tưởng chung là giá trị mỗi điểm ảnh sẽ phụ thuộc nhiều vào các điểm ảnh ở gần hơn là các điểm ảnh ở xa. Trọng số của sự phụ thuộc được lấy theo hàm Gauss (cũng được sử dụng trong quy luật phân phối chuẩn).



Hình 2.12

Giả sử ảnh là một chiều. Điểm ảnh ở trung tâm sẽ có trọng số lớn nhất. Các điểm ảnh ở càng xa trung tâm sẽ có trọng số giảm dần khi khoảng cách từ chúng tới điểm trung tâm tăng lên. Như vậy điểm càng gần trung tâm sẽ càng đóng góp nhiều hơn vào giá trị điểm trung tâm.

|  |  |
| --- | --- |
|  |  |
| Ảnh gốc | Ảnh sau khi làm mờ, giảm nhiễu |

Hình 2.13

### Nhị phân hóa với ngưỡng động (Adaptive Threshold)

#### Ảnh nhị phân

Là ảnh mà giá trị của các điểm ảnh chỉ được biểu diễn bằng hai giá trị là 0 (Đen) và 255 (Trắng)

#### Nhị phân hóa

Là quá trình biến đổi một ảnh xám thành ảnh nhị phân.

* + Gọi giá trị cường độ sáng tại một điểm ảnh là I(x,y) .
  + INP(x,y) là cường độ sáng của điểm ảnh trên ảnh nhị phân .
  + (Với 0 < x < image.width) và (0 < y < image.height).

Để biến đổi ảnh xám thành ảnh nhị  phân. Ta so sánh giá trị cường độ sáng của điểm ảnh với một ngưỡng nhị phân **T**.

* + Nếu I(x,y) > **T** thì INP(x, y) = 0.
  + Nếu I(x,y) > **T** thì INP(x, y) = 255.

#### Nhị phân hóa với ngưỡng động

Việc nhị phân hóa ảnh với ngưỡng toàn cục như thông thường sẽ rất khó khăn khi phải tự tính toán và chọn mức ngưỡng phù hợp cho từng ảnh khác nhau. Nhị phân hóa ảnh ngưỡng động sẽ giúp tính toán ngưỡng cho phù hợp với từng ảnh, lợi thế thứ hai chính là nó rất phù hợp khi ảnh có vùng bị quá chói hoặc quá tối dẫn đến có thể mất luôn hình ảnh tại vùng đó nếu sử dụng ngưỡng toàn cục.

Về ý tưởng chính sẽ theo 3 bước sau:

1. Chia tấm ảnh thành nhiều khu vực, cửa sổ (Region) khác nhau.
2. Dùng một thuật toán để tìm một giá trị **T** phù hợp với từng cửa sổ.
3. Áp dụng phương pháp nhị phân hóa cho từng khu vực, cửa sổ với ngưỡng **T** phù hợp.

Có rất nhiều phương pháp để tìm **T**, ở đây em sử dụng một kiểu thuật toán mà thư viện OpenCV hỗ trợ là ADAPTIVE\_THRESH\_GAUSSIAN\_C tức lấy trung bình các giá trị xung quanh điểm ngưỡng động đang xét **T**(x,y) với phân phối Gauss rồi trừ đi hằng số C.



Hình 2.14

### Phát hiện cạnh Canny (Canny Edge Detection)

Trong hình ảnh, thường tồn tại các thành phần như: vùng trơn, góc/cạnh và nhiễu. Cạnh trong ảnh mang đặc trưng quan trọng, thường là thuộc đối tượng trong ảnh. Do đó, để phát hiện cạnh trong ảnh, có nhiều giải thuật khác nhau như toán tử Sobel, toán tử Prewitt, Zero crossing .... nhưng ở đây em chọn giải thuật Canny vì hương pháp này hơn hẳn các phương pháp khác do ít bị tác động của nhiễu và cho khả năng phát hiện các biên yếu. Phương pháp này đi theo 4 bước chính:

1. Giảm nhiễu (Noise reduction)
2. Tính toán Gradient (Gradient calculation)
3. Loại bỏ những điểm không phải là cực đại (Non-maximum suppression)
4. Lọc ngưỡng (Double threshold)
5. Giảm nhiễu

Làm mờ ảnh, giảm nhiễu dùng bộ lọc Gauss kích thước 5x5. Kích thước 5x5 thường hoạt động tốt cho giải thuật Canny

1. Tính toán Gradient

Ta dùng 2 bộ lọc Sobel X và Sobel Y (3x3) để tính đạo hàm Gx và Gy

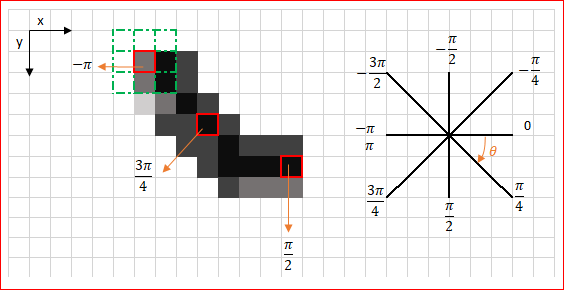
|  |  |
| --- | --- |
|  |  |

Tìm gradient và hướng được làm tròn về 4 hướng: hướng ngang (0 độ), hướng chéo bên phải (45 độ), hướng dọc (90 độ) và hướng chéo trái (135 độ).

|  |  |
| --- | --- |
|  |  |

1. Loại bỏ những điểm không phải là cực đại

Ở bước này, dùng một filter 3x3 lần lượt chạy qua các pixel trên ảnh gradient. Trong quá trình lọc, xem xét xem độ lớn gradient của pixel trung tâm có phải là cực đại so với các gradient ở các pixel xung quanh. Nếu là cực đại, ta sẽ ghi nhận sẽ giữ pixel đó lại. Còn nếu pixel tại đó không phải là cực đại lân cận, ta sẽ set độ lớn gradient của nó về zero. Ta chỉ so sánh pixel trung tâm với 2 pixel lân cận theo **hướng gradient.** Ví dụ: nếu hướng gradient đang là 0 độ, ta sẽ so pixel trung tâm với pixel liền trái và liền phải nó. Trường hợp khác nếu hướng gradient là 45 độ, ta sẽ so sánh với 2 pixel hàng xóm là góc trên bên phải và góc dưới bên trái của pixel trung tâm

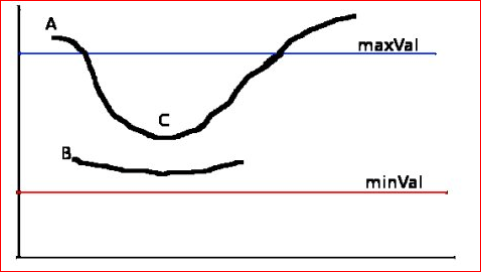


Hình 2.15



Hình 2.16

**Lọc ngưỡng: ta sẽ xét các pixel dương trên mặt nạ nhị phân kết quả của bước trước. Nếu giá trị gradient vượt ngưỡng max\_val thì pixel đó chắc chắn là cạnh. Các pixel có độ lớn gradient nhỏ hơn ngưỡng min\_val sẽ bị loại bỏ. Còn các pixel nằm trong khoảng 2 ngưỡng trên sẽ được xem xét rằng nó có nằm liền kề với những pixel được cho là "chắc chắn là cạnh" hay không. Nếu liền kề thì ta giữ, còn không liền kề bất cứ pixel cạnh nào thì ta loại. Sau bước này ta có thể áp dụng thêm bước hậu xử lý loại bỏ nhiễu (tức những pixel cạnh rời rạc hay cạnh ngắn) nếu muốn**



Hình 2.17: Lọc ngưỡng

1. Kết quả

Sau khi sử dụng phát hiện biên canny, dù đã trích xuất được những chi tiết cạnh của biển số, tuy nhiên vẫn còn quá nhiều chi tiết thừa trong hình ảnh, từ đây chúng ta sẽ vẽ contour, áp dụng nhưng đặc điểm của biển số để lọc lấy ra biển số chính xác.



Hình 2.18: Ảnh sau khi phát hiện biên Canny

### Lọc biển số với contour

#### Một số phương pháp tìm contour

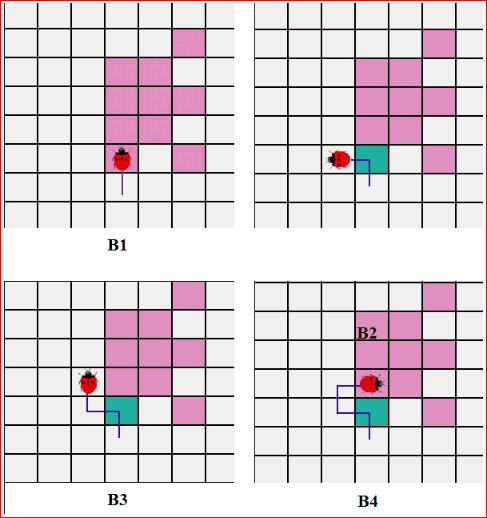
Có thể hiểu Contour là tập hợp các điểm tạo thành đường cong kín bao quanh một đối tượng nào đó. Thường dùng để xác định vị trí, đặc điểm của đối tượng. Có 4 thuật toán Contour Tracing chung nhất. Hai trong số đó có tên là: Square Tracing algorithm và Moore – Neighbor Tracing là dễ để thực hiện và thường xuyên được dùng để dò tìm contour của một mẫu. Với thư viện OpenCV người ta áp dụng thuật toán Suzuki’s Contour tracing. Dưới đây em sẽ trình bày kĩ hơn về 3 phương pháp trên:

1. Thuật toán Square Tracing

Duyệt từ pixel ngoài cùng bên trái phía dưới, đi lên cho tới khi gặp pixel có giá trị bằng 255 (pixel này sẽ được gọi là pixel start) thì bắt đầu di chuyển theo quy tắc sau:

* Nếu gặp Pixel có giá trị bằng 255 thì rẽ trái.
* Nếu gặp Pixel có giá trị bằng 0 thì rẽ phải.
* Di chuyển cho tới khi quay lại pixel start thì dừng lại.

Hình sau mô tả cách hoạt động của thuật toán:



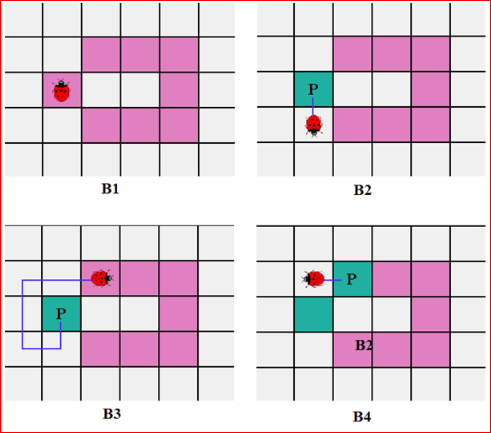
Hình 2.19 : Thuật toán Square Tracing

Thuật toán sẽ kết thúc đúng khi di chuyển vào pixel start lần thứ 2 sau khi đi qua n pixel khác và theo đúng hướng đi vào pixel start lần đầu tiên. Và sai khi di chuyển vào pixel start mà không đúng hướng ban đầu. Vậy thuật toán này chỉ chạy đúng trên đối tượng   
4 - connected.

|  |  |
| --- | --- |
| Hình 2.20 : Thuật toán Square Tracing  chạy đúng | Hình 2.21: 3 Thuật toán Square Tracing  chạy sai |

1. Thuật toán Moore – Neighbor

Thuật toán này có chút khác biệt so với thuật toán Square Tracking, cụ thể: khi gặp pixel có giá trị bằng 255 đầu tiên (pixel start) thì ta sẽ quay lại pixel trước đó, sau đó đi vòng qua các pixel thuộc 8-connected theo chiều kim đồng hồ cho tới khi gặp pixel khác có giá trị bằng 255. Và điệu kiện kết thúc cũng giống như thuật toán Square Tracking



Hình 2.22 Thuật toán Moore - Neighbor.

1. Thuật toán Suzuki’s Tracing

Đây là thuật toán được thư viện OpenCV sử dụng, ngoài khả năng xác định được biên của vật thể như hai phương pháp trên. Phương pháp Suzuki’s Tracing còn có khả năng phân biệt được đường biên ngoài (Outer) và đường biên trong (Hole) của vật thể.

Hàm trong OpenCV được biểu diễn như sau:

**findContours**(InputOutputArray image, OutputArrayOfArrays contours, OutputArray hierarchy, int mode, int method, Point offset=Point())

Các tham số:   
**image** : hình ảnh cần tìm biên, là ảnh nhị phân.  
**contours** : lưu trữ các đường biên tìm được, mỗi đường biên được lưu trữ dưới dạng một vector của các điểm.  
**hierarchy** :  chứa thông tin về hình ảnh như số đường viền, xếp hạng các đường viền theo kích thước, trong ngoài, ..  
**mode** :  
CV\_RETR\_EXTERNAL : khi sử dựng cờ này nó chỉ lấy ra những đường biên bên ngoài, nhưng biên bên trong của vật thể bị loại bỏ.  
CV\_RETR\_LIST : Khi sử dụng cờ này nó lấy ra tất cả các đường viền tìm được.  
CV\_RETR\_CCOMP : khi sử dụng cờ này nó lấy tất cả những đường biên và chia nó làm 2 level, những đường biên bên ngoài đối tượng, và những đường biên bên trong đối tượng.  
CV\_RETR\_TREE : khi sử dụng cờ này nó lấy tất cả các đường biên và tạo ra một hệ thống phân cấp đầy đủ của những đường lồng nhau.  
**method** :  
CV\_CHAIN\_APPROX\_NONE : sử dụng cờ này sẽ lưu trữ tất cả các điểm của đường viền .  
CV\_CHAIN\_APPROX\_SIMPLE : Ví dụ : một hình chữ nhật sẽ được mã hoá bằng toạ độ của 4 đỉnh.  
CV\_CHAIN\_APPROX\_TC89\_L1 or CV\_CHAIN\_APPROX\_TC89\_KCOS :  Áp dụng thuật toán xấp xỉ Tech-Chin.



Hình 2.235 Vẽ Contour với OpenCV

Trong ảnh, những đường màu hồng là đường contour bao quanh vật thể, nhưng vì có quá nhiều đường bao quanh các vật thể không phải biển số nên chúng ta sẽ áp dụng những đặc trưng riêng về tỉ lệ cao/rộng, diện tích trong khung hình cố định để lọc ra đúng biển số cần tìm.

#### Lọc biển số

Đầu tiên ta làm xấp xỉ contour thành một hình đa giác và chỉ lấy những đa giác nào chỉ có 4 cạnh. Nghĩa là lúc xấp xỉ contour bộ nhớ chỉ ghi nhớ vị trí các đỉnh của đa giác đó thành một mảng. Số cạnh của đa giác sẽ bằng số đỉnh và bằng chiều dài của mảng đó.

|  |  |
| --- | --- |
| Hình 2.24: Contour chưa xấp xỉ đa giác | Hình 2.25 : Contour đã xấp xỉ đa giác |

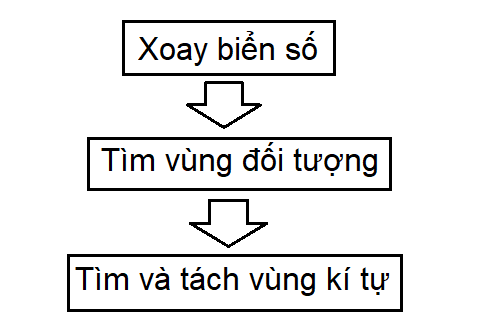
Tiếp theo ta tính toán tỉ lệ cao/rộng và diện tích của biển số phù hợp, sau đó ta lưu tất cả những biển số có trong hình dưới dạng tọa độ các đỉnh

Từ đây, ta cắt hình ảnh biển số từ các tọa độ vị trí đã biết để phục vụ cho mục đích tiếp theo “Tách các kí tự trong biển số”. Lưu ý ở đây ta cắt từ ảnh nhị phân luôn để máy tính xử lý nhanh hơn, tốn ít thời gian hơn.

## PHÂN ĐOẠN KÍ TỰ

### Hướng giải quyết

Ở giai đoạn này có những bước chính sau: Xoay biển số để tăng khả năng nhận diện, Tìm tất cả các vùng kín cho là kí tự và lọc ra những kí tự đúng. Tách hình ảnh nhưng kí tự đó ra và đưa vào bộ nhận diện



Hình 2.26: Các bước chính trong phân đoạn kí tự

### Xoay biển số

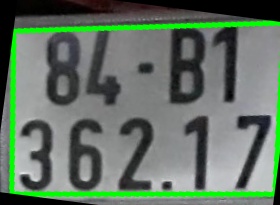
Khi chụp ảnh đầu vào, không phải lúc nào biển số cũng ở chính diện, có thể bị méo sang trái, sang phải, nghiêng góc dẫn đến nếu cứ sử dụng ảnh biển số đã cắt mà không điều chỉnh góc độ dẫn đến ảnh kí tự được cắt ra đưa vào bộ nhận diện rất dễ bị sai. Ví dụ giữa số 1 và số 7, số 2 và chữ Z, chữ B và số 8,...



Hình 2.27: Ảnh biển số chưa xoay

Phương pháp xoay ảnh em sử dụng ở đây là:

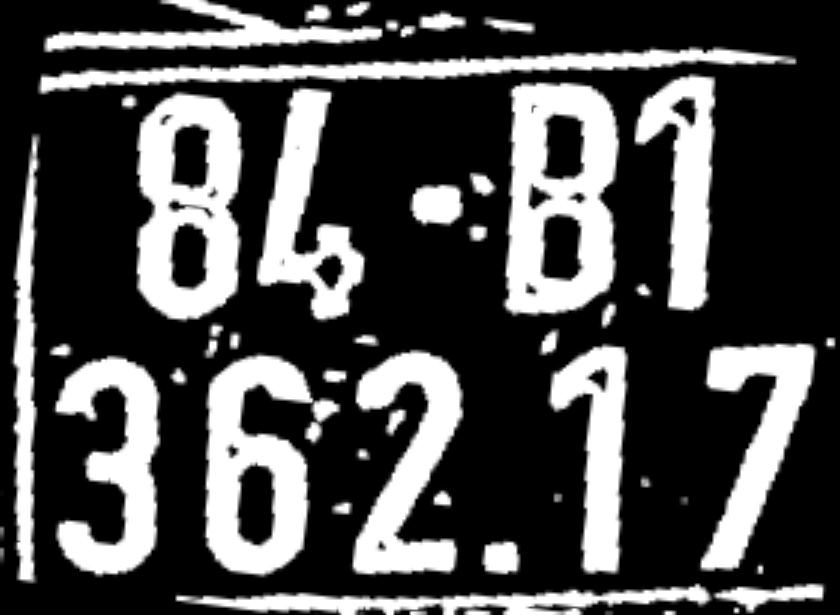
1. Lọc ra tọa độ 2 đỉnh A,B nằm dưới cùng của biển số
2. Từ 2 đỉnh có tọa độ lần lượt là A(x1, y1) và B(x2,y2) ta có thể tính được cạnh đối và cạnh kề của tam giác ABC
3. Ta tính được góc quay
4. Xoay ảnh theo góc quay đã tính. Nếu ngược lại điểm A nằm cao hơn điểm B ta cho góc quay âm



Hình 2.28 : Ảnh biển số đã xoay

### Tìm vùng đối tượng

Từ ảnh nhị phân, ta lại tìm contour cho các kí tự (phần màu trắng). Sau đó vẽ những hình chữ nhật bao quanh các kí tự đó. Tuy nhiên việc tìm contour này cũng bị nhiễu dẫn đến việc máy xử lý sai mà tìm ra những hình ảnh không phải kí tự. Ta sẽ áp dụng các đặc điểm về tỉ lệ chiều cao/rộng của kí tự, diện tích của kí tự so với biển số



Hình 2.29: Ảnh nhị phân



Hình 2.30: Tìm vùng đối tượng

Trong hình 5.5 những đường màu vàng là đường contour và nếu so sánh với ảnh nhị phân hình 5.4 thì có rất nhiều đường nhiễu như đường viền biển số, dấu gạch, dấu chấm... Sau khi đã áp dụng các điều kiện thì sẽ vẽ ra những hình chữ nhật màu xanh bao quanh các kí tự.

### Tìm và tách kí tự

Sau khi đã nhận dạng từng kí tự bằng hình chữ nhật và cũng đã có tọa độ vị trí 4 đỉnh của hình đó, ta lúc này có thể cắt hình ảnh kí tự đó ra phục vụ cho giai đoạn sau “Nhận diện kí tự”. Lưu ý ở đây ta cắt ảnh nhị phân chứ không cắt từ ảnh gốc.



Hình 2.31 Ảnh kí tự sau khi cắt

## NHẬN DIỆN KÍ TỰ

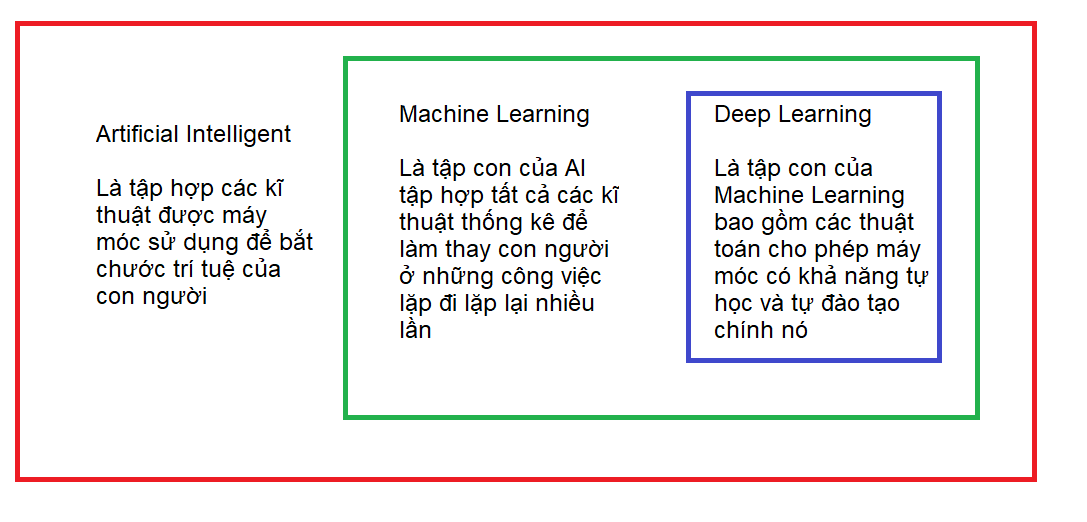
Thực chất quá trình nhận diện kí tự chính là quá trình chuyển đổi từ một hình ảnh là ma trận giá trị các điểm ảnh về một dạng thông tin khác như trong đề tài này là mã ASCII để có thể giao tiếp với người dùng. Để hiểu hơn về nhận diện, chúng ta cần đi ngược lại về ngành khoa học trí tuệ nhân tạo (Artificial Intelligent) hay còn gọi là AI

### Tổng quan về lý thuyết

#### AI (Artificial Intelligent)

Trí tuệ nhân tạo là dạng trí tuệ được biểu hiện ở máy móc, không giống như những dạng trí tuệ tự nhiên ở con người hay động vật. Thông thường cụm từ “trí tuệ nhân tạo” sẽ được hiểu như máy móc có khả năng bắt chước hành vi và nhận thức của con người như việc học tập hay xử lý vấn đề. Trong vài năm trở lại đây, AI thật sự bùng nổ, đặc biệt là từ năm 2015. Phần lớn trong số đó liên quan đến những tiện ích sẵn có của GPU khiến cho việc xử lý song song nhanh, rẻ và mạnh mẽ hơn.

Bên trong AI bao gồm hai khối nhỏ là Máy học (Machine learning) và Học sâu (Deep learning)



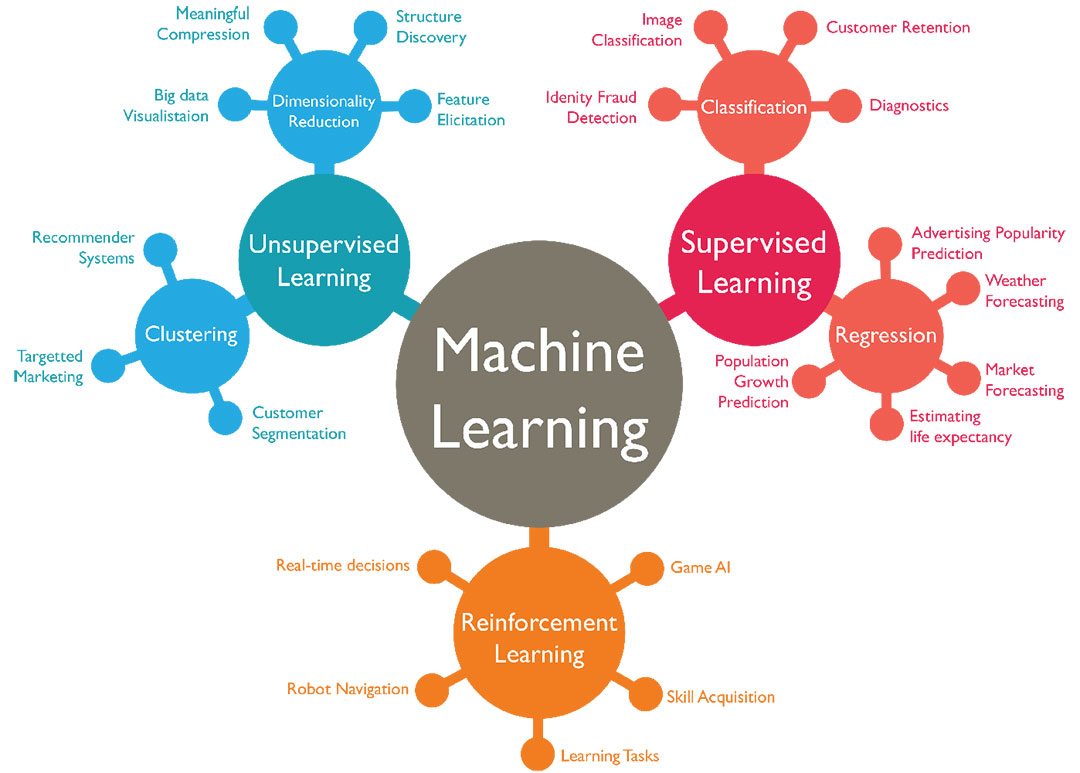
Hình 2.32 : Các lớp trong AI

#### Machine Learning

Machine learning là một lĩnh vực sử dụng các thuật toán cho phép máy tính có thể học từ dữ liệu để thực hiện các công việc thay vì được lập trình một cách rõ ràng. Một số ứng dụng có thể kể đến như xử lý hình ảnh (gắn thẻ, nhận diện kí tự, ô tô tự lái), phân tích văn bản (lọc mail spam, phân tích ngữ nghĩa và khai thác thông tin),...

Về cách hoạt động của Machine learning nói nôm na ta cần một tập dữ liệu huấn luyện bao gồm nhiều mẫu huấn luyện. Mỗi mẫu huấn luyện sẽ là một thể hiện của bài toán (có đầu vào và lời giải). Machine learning sẽ học từ các thể hiện đó để tìm ra lời giải phù hợp với từng đầu vào mới. Nó giống như khi dạy một đứa trẻ cách tâng cầu, ta sẽ tâng cầu vài lần cho đứa trẻ quan sát. Sau đó đứa trẻ sẽ bắt đầu học để tự tâng cầu. Trong đề tài đồ án này, em sẽ đưa vào tập dữ liệu mẫu là hình ảnh các chữ số, kí tự của biển số Việt Nam đã được gắn nhãn là một mã ASCII tương ứng để máy tự học.

Machine Learning có nhiều thuật toán khác nhau nhưng nổi bật nhất là hai thuật toán **Học có giám sát** (Supervised learning) và **Học không giám sát** (Unsupervised learning)



Hình 2.33: phân loại Machine Learning

1. Học có giám sát.

Là thuật toán dự đoán đầu ra (outcome) của một dữ liệu mới (new input) dựa trên các cặp (input, outcome) đã biết từ trước. Cặp dữ liệu này còn được gọi là (data, label), tức (dữ liệu, nhãn). Ví dụ như thuật toán dò các khuôn mặt trong một bức ảnh đã được phát triển từ rất lâu. Thời gian đầu, Facebook sử dụng thuật toán này để chỉ ra các khuôn mặt trong một bức ảnh và yêu cầu người dùng tag friends- tức gán nhãn cho mỗi khuôn mặt. Số lượng cặp dữ liệu (khuôn mặt, tên người) càng lớn, độ chính xác ở những lần tự động tag tiếp theo sẽ càng lớn.

Thuật toán học có giám sát lại tiếp tục được chia ra hai phần nhỏ gồm bài toán phân loại (Classification) khi các nhãn của dữ liệu đầu vào chia thành hữu hạn các nhóm và bài toán hổi quy (Regression) khi nhãn không được chia thành các nhóm mà là một giá trị thực cụ thể. Ví dụ: một cô gái x tuổi, cao y m, thu nhập hàng tháng z triệu đồng thì sẽ kết hôn khi nào?

1. Học không giám sát

Trong thuật toán này, chúng ta không biết được đầu ra hay nhãn mà chỉ có dữ liệu đầu vào. Thuật toán học không giám sát sẽ dựa vào cấu trúc của dữ liệu để thực hiện một công việc nào đó, ví dụ như phân nhóm (clustering) hoặc giảm số chiều của dữ liệu (dimension reduction) để thuận tiện trong việc lưu trữ và tính toán.

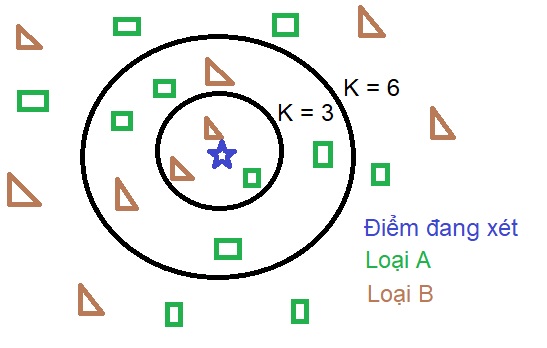
Bao gồm hai loại chính: Phân nhóm (clustering) toàn bộ dữ liệu đầu vào thành các nhóm nhỏ dựa trên sự liên quan giữa các dữ liệu trong mỗi nhóm ví dự như chia nhóm khách hàng dựa trên hành vi mua hàng, độ tuổi, giới tính của họ. Còn lại là bài toán kết hợp (Association) được dùng khi ta muốn khám phá ra một quy luật nào đó dựa trên tập dữ liệu cho trước như những khách hàng mua laptop thường có xu hướng mua thêm con chuột hay chó thường trung thành với chủ hơn mèo.

#### Thuật toán KNN (K - Nearest Neighbor)

KNN là một trong những thuật toán học có giám sát đơn giản nhất trong Machine Learning, có thể sử dụng cho cả bài toán phân loại và hồi quy. Về ý tưởng là gán kết quả với dữ liệu training gần giống với mẫu nhất. Ví dụ như khi đi câu cá, ta không biết cá câu lên là cá rô hay cá chép, nhưng khi so sánh những đặc điểm về mắt, mang, vây,... từ những con cá rô, cá chép đã thấy thì cuối cùng có thể quyết định xem con cá mình câu được thuộc nhóm cá nào.

KNN hoạt động theo quy trình gồm 4 bước chính:

1. Xác định tham số K (số láng giềng gần nhất).
2. Tính khoảng cách từ điểm đang xét đến tất cả các điểm trong tập dữ liệu cho trước
3. Sắp xếp các khoảng cách đó theo thứ tự tăng dần
4. Xét trong tập K điểm gần nhất với điểm đang xét, nếu số lượng điểm của loại nào cao hơn thì coi như điểm đang xét thuộc loại đó



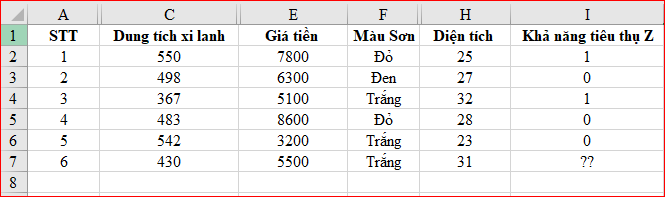
Hình 2.34: Ví dụ về KNN

Việc điểm đang xét thuộc loại nào còn phụ thuộc vào hệ số K hay trọng số khoảng cách... mà người dùng đặt sao cho phù hợp với bài toán đang xét. Chẳng hạn ở hình trên nếu ta xét K = 3 thì điểm đang xét sẽ thuộc loại B, ngược lại nếu K = 6 thì nó thuộc loại A. Ngoài ra người ta có thể để trọng số cao hơn cho những điểm gần hơn hay ít khi sử dụng K = 1 để đảm bảo kết quả đầu ra được tối ưu.

Thông thường việc tính khoảng cách đến các điểm sẽ theo công thức Euclid:

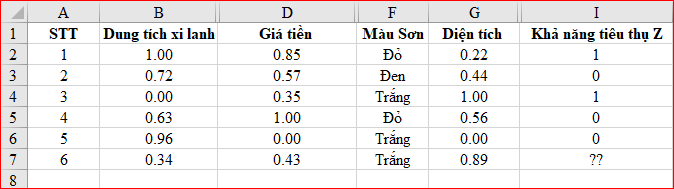
Khi thực hiện so sánh có thể bỏ qua dấu căn bậc 2. Ngoài ra nếu như khoảng cách giữa các biến quá lớn như biến x lớn hơn xấp xỉ 1000000 lần thì ta cũng cần chuẩn hóa lại dữ liệu theo công thức:

Để dễ hiểu hơn em xin lấy một ví dụ về một công ty muốn đưa dòng xe ô tô mới của mình vào thị trường và muốn biết xe sẽ bán chạy hay ế. Họ tổng hợp các dữ liệu về dung tích xi lanh, giá tiền, màu sơn, diện tích xe của các dòng xe đã xuất ra thị trường. Đặt Z là khả năng tiêu thụ cụ các dòng xe. Z = 0 nghĩa là xe ế, Z = 1 nghĩa là xe sẽ bán chạy



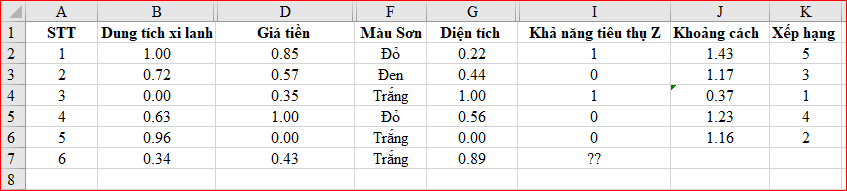
Hình 2.35: Xét khả năng tiêu thụ của xe

Sau đó ta chuẩn hóa dữ liệu:



Hình 2.36: Dữ liệu sau chuẩn hóa

Sau đó ta tính khoảng cách giữa điểm đang xét đến các dữ liệu cho trước. Lưu ý ở các biến có giá trị định tính nếu khác nhau khoảng cách sẽ bằng 1, giống nhau sẽ bằng 0. Ví dụ ta tính khoảng các giữa dòng xe 6 và 1:



Hình 2.37:Sau khi tính khoảng cách và xếp hạng

Nhìn hình trên nếu chọn K=1 thì kết quả Z=1, nếu K=3 thì Z=0, nếu K=5 thì Z=0.

Tuy nhiên phương pháp này có một số ưu và nhược điểm như sau:

* Ưu điểm:
  + Dễ sử dụng và cài đặt
  + Độ phức tạp tính toán nhỏ
  + Việc dự đoán kết quả rất đơn giản
* Nhược điểm:
  + Với K nhỏ, khi gặp nhiễu dễ đưa ra kết quả không chính xác
  + Cần nhiều thời gian lưu trainning set và nếu test tăng lên sẽ tốn rất nhiều thời gian

### Hướng giải quyết

Ở giai đoạn cuối này được thực hiện theo những bước sau:

1. Tạo tập dữ liệu để huấn luyện
2. Huấn luyện mô hình KNN
3. Đưa hình ảnh từ bước “Phân đoạn kí tự” vào mô hình KNN đã tạo để đưa ra kết quả
4. In ra kết quả biển số

Bước 1 và 2 ta sẽ tạo ra mô hình KNN riêng biệt với code chính. Để khi cần nhận diện kí tự ta không cần phải làm lại các bước từ đầu. Đầu tiên em tạo tập dữ liệu (tập hình ảnh của các chữ số và kí tự) để train từ phần mềm paint. Trong phần mềm Paint ta viết các chữ số và kí tự (trừ kí tự O, I, J) với phông chữ “Biển số xe Việt Nam”, có thể xoay các kí tự này lần lượt với các góc . Kết quả có dạng như sau:



Hình 2.38: Tập dữ liệu huấn luyện

Tiếp theo ta lấy ngưỡng, vẽ contour và cắt từng kí tự. Vì mỗi kí tự có kích thước khác nhau xử lý phức tạp nên cần chuẩn hóa hình ảnh lại với kích thước cao:rộng là 30:20 pixels. Thay vì mỗi kí tự đưa vào mô hình để máy nhận diện thì những kí tự này sẽ được ta gắn nhãn bằng những phím bấm trên máy tính. Sau khi gắn nhãn hết các kí tự ta sẽ lưu hai file .txt là classifications.txt và flattened\_images.txt. File classifications.txt có nhiệm vụ lưu các mã ASCII của các kí tự đó và file flattened\_images.txt sẽ lưu giá trị các điểm ảnh có trong hình ảnh kí tự (hình 20x30 pixel có tổng cộng 600 điểm ảnh có giá trị 0 hoặc 255)

Bước 3 và 4. Ta thực hiện đưa ảnh đang xét vào và tính khoảng cách đến tất cả các điểm trong mẫu, kết quả sẽ là mã ASCII đại điện cho hình ảnh đó. Cuối cùng ta in biển số xe ra hình. Tuy nhiên ở Việt Nam có hai loại biển số là biển một hàng và biển hai hàng. Về ý tưởng chung để phân biệt hai hàng này ta dựa vào vị trí của hình ảnh kí tự, nếu vị trí nằm thấp 1/3 chiều cao của biển số thì kí tự sẽ được xếp vào hàng một. Ngược lại sẽ được xếp vào hàng hai.

|  |  |
| --- | --- |
| Hình 2.39 : Biển số trước khi nhận diện | Hình 2.40: Biển số sau khi nhận diện |



Hình 2.41: Biển số xe được in ra trên hình gốc

# KẾT QUẢ THỰC HIỆN

## Cách thức đo đạc, thử nghiệm

Clip đầu vào được quay bởi camera từ điện thoại Samsung J7 Prime có các thông số:

* Kích thước clip: FHD 1920x1080
* Camera: 9.6 MP
* 8 ảnh/s

Sau đó clip sẽ được chương trình xử lý cắt từng frame ảnh ra để nhận diện và in ra biển số. Chương trình xử lý được viết trên giao diện ứng dụng Visual Studio 2017 bằng ngôn ngữ Python (3.7), bao gồm các thư viện OpenCV, Numpy, Math.

Em sẽ tạo hai clip test dành riêng cho biển số một hàng và hai hàng được quay với đủ các góc độ nghiêng trái, phải, trên dưới, xa gần, và các môi trường có phông nền, ánh sáng khác nhau. Với mỗi clip test có độ dài X frame ảnh, mỗi frame ảnh có chứa Y biển số, ta sẽ được:

Tổng biển số =. (Y = 3 với biển 2 hàng, Y = 1 với biển một hàng).

Tỉ lệ tìm thấy biển số xe = 100 (%)

Ta sẽ tìm tỉ lệ biển sai n kí tự trên tổng số biển bắt được trong clip. Lưu ý sai ở đây nghĩa là kí tự bị nhận diện sai, không khoanh được vùng hoặc khoanh vùng kí tự sai vị trí.

Tỉ lệ nhận diện = 100 (%)

## Kết quả và giải thích

|  |  |  |  |
| --- | --- | --- | --- |
| Loại biển | Tổng biển số | Số biển  tìm thấy | Tỉ lệ tìm thấy biển số xe (%) |
| Biển 1 hàng | 370 | 182 | 49,2% |
| Biển 2 hàng | 2349 | 924 | 39,3% |

Ở biển số 1 hàng có tỉ lệ nhận diện cao hơn vì tập mẫu còn ít và mỗi frame ảnh chỉ có 1 biển, dẫn đến biển số thường nằm ở vị trí trung tâm nên dễ nhận diện hơn. Ngược lại, với biển 2 hàng có tập mẫu lớn, và số biển số trong 1 frame nhiều hơn. Chưa kể, biển số ở đây được tính là vùng hình bình hành có tối thiểu 7 kí tự mới được tính là biển số, dẫn đến có nhiều biển số đã cắt đúng những vẫn không được tính vào. Tuy nhiên cách làm này lại giúp loại bỏ đáng kể những chi tiết, biển số sai lệch từ môi trường ngoài.

|  |  |
| --- | --- |
| Hình 3.1: không tìm thấy kí tự | Hình 3.2: ảnh gốc |
| Hình 3.3 :tìm tháy 5 kí tự |
| Hình 3.4: tìm thấy 9 kí tự |

Ở hình 7.2 ta thấy chương trình xử lý tìm ra 3 vùng được cho là biển số, tuy nhiên số kí tự tìm ra ở hình 7.1 và 7.3 không đủ nên sẽ bị loại dẫn đến trong ảnh chỉ còn lại duy nhất biển số hình 7.4

Khi ta quay theo nhiều góc độ, nhiều vị trí dẫn đến khi tính toán diện tích, tỉ lệ cao/rộng của biển số không còn thỏa điều kiện đặt ra nên đã bị loại. Biển số có thể bị ảnh hưởng bởi những chi tiết ngoài nên khi xấp xỉ contour không ra hình tứ giác, dẫn đến cũng gây mất biển số. Lỗi này đặc biệt xảy ra ở những xe ô tô vì ô tô thường có nền xung quanh biển số là những vật liệu phản chiếu ánh sáng mạnh, gây ảnh hưởng lớn đến quá trình xác định vùng biển số.

|  |  |
| --- | --- |
| Hình 3.5: Lấy ngưỡng cao/rộng 1.5 | Hình 3.6 : Lấy ngưỡng cao/rộng 1.4 |



Hình 3.7 : Lỗi xấp xỉ Contour

Ở hình 7.7 mặc dù đường contour màu hồng có bao quanh biển số, tuy nhiên sau khi xấp xỉ chỉ còn lại hình 2, 3 cạnh, để khắc phục được các lỗi tỉ lệ, xấp xỉ contour ta cần điều chỉnh mức ngưỡng cho tối ưu.

Trong quá trình xử lý, việc xử lý nhị phân cũng đóng vai trò quan trọng, ở hình 6.2 - 9 ta thấy ảnh bị nhiễu và bản thân biển số bị tối, dính nhiều bụi dẫn đến khi xử lý nhị phân sẽ bị đứt đoạn và vẻ contour bị sai, để khắc phục cần sử dụng những phép toán hình thái học như phép nở, phép đóng để làm liền những đường màu trắng trong ảnh nhị phân.

|  |  |
| --- | --- |
| Hình 3.8: Ảnh nhị phân bị đứt | Hình 3.9: Đường contour bị đứt đoạn |

Dưới đây ta xét khả năng khoanh vùng và nhận diện kí tự tương ứng với giai đoạn “Phân đoạn kí tự” và “Nhận diện kí tự” đã đặt ra ở đầu bài toán.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Loại biển | Số biển tìm thấy | Không sai | Sai 1 kí tự | Sai 2 kí tự | Sai 3 kí tự  trở lên |
| Biển 1 hàng | 182 | 61 | 88 | 19 | 14 |
| Tỉ lệ (%) | | 33,5 | 48,4 | 10,4 | 7,7 |

Tỉ lệ nhận diện sai kí tự ở biển 1 hàng

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Loại biển | Số biển tìm thấy | Không sai | Sai 1 kí tự | Sai 2 kí tự | Sai 3 kí tự  trở lên |
| Biển 2 hàng | 924 | 286 | 273 | 175 | 190 |
| Tỉ lệ nhận diện (%) | | 31 | 29,5 | 18,9 | 20,6 |

Tỉ lệ nhận diện sai kí tự ở biển 2 hàng

Nhìn chung mô hình nhận diện KNN cũng khá tốt, có những kí tự dù bị mờ, bị nghiêng vẫn nhận diện đúng. Điều này một phần nhờ vào chương trình đã xoay biển số lại cho để tăng khả năng nhận diện, cho dù nghiêng thì kí tự cũng chỉ nghiêng từ đến .Tuy nhiên vẫn còn nhầm lẫn nhiều giữa các kí tự như số 1 với số 7. Chữ G, chữ D, số 6 với số 0. Chữ B với số 8...



Hình 3.10: ảnh gốc với 3 biển số

|  |  |  |
| --- | --- | --- |
| Hình 3.11: biển 1 | Hình 3.12: biển 2 | Hình 3.13: biển 3 |



Hình 3.14:Không khoanh được vùng kí tự

Ở ảnh 7.14 chỗ số 0 bị dính con ốc dẫn đến không thể khoanh vùng được, Tuy nhiên ở chữ F mặc dù vẫn dính con ốc trong ảnh cắt ra để nhận diện, mô hình KNN vẫn cho ra đáp án đúng.

## Bài toán nhận quán lý lưu lương tham gia giao thông.

-Kiểm tra và kiểm soát được số lượng xa ra và vào của 1 khu vực (các trạm BOT , hoặc các trạm gác nơi xe cộ đi lại với vận tốc chậm và theo lượt)

-Nhận diện được từng xe kèm theo biển số.

-Áp dụng vào ứng dụng để quản lý bãi đỗ xe do chưa thể tiếp cận với những trạm BOT và áp dụng chương trình

-Có thể kiểm tra được xe đi vào, xe đi ra và tự động tính tiền gửi xe của mỗi xe dựa trên số giờ trong bãi.

### Cách hoạt động của bài toán.

-Mỗi khi xe đi qua trạm thì camera sẽ tự nhận diện biển số xe và nhập biển số vào cơ sở dữ liệu, kèm theo đó là thời điểm vào bãi của xe.

-Sau khi xe rời bãi đỗ xe, camera sẽ nhận diện lại biển đõ xe thêm 1 lần nữa và dựa vào dữ liệu đã có và thời điểm ra để tính được số tiền cần trả,

# KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

## Kết luận

Bằng xử lý ảnh và thuật toán KNN có những ưu và khuyết điểm sau:

* Ưu điểm:
  + Dễ cài đặt và sử dụng.
  + Khá nhẹ nên máy tính với cấu hình yếu cũng có thể xử lý mượt mà so với các thuật toán khác như CNN, SVM.
  + Phù hợp cho đối tượng sinh viên muốn tìm hiểu căn bản về xử lý ảnh hay trí tuệ nhân tạo.
* Khuyết điểm:
  + Khả năng nhận diện của KNN còn thấp, khi tập dữ liệu quá nhiều sẽ tăng thời gian xử lý vì phải quét hết tập dữ liệu train.
  + Nhận diện kém với sự phản chiếu của biển số, sự di ảnh, chói sáng từ môi trường ngoài, những biển có phần chữ số không rõ ràng, với biển số xe ô tô

Vì vậy tốt nhất cần đặt camera cố định, với môi trường ánh sáng xung quanh được cài đặt trước. Phông nền cần hạn chế tối da nhưng chi tiết lóe sáng gây nhiễu. Phương pháp này vẫn còn cần sự giám sát của con người nhiều chứ chưa thể hoàn toàn tự động.

## Hướng phát triển

Cần thay đổi thuật toán nhận diện KNN sang những thuật toán khác tinh vi và phức tạp hơn như CNN, SVM hoặc có thể sử dụng những bộ thư viện đã có sẵn trên thế giới như YOLO, YOLOv3 ....

Sử dụng camera chuyên dụng cho việc nhận diện biển số xe vì có khả năng chống chịu với sương mù, đêm tối, chói sáng,...

Sử dụng các thuật toán xử lý ảnh khác để xác định vị trí biển số tốt hơn như phương pháp biến đổi Hough để nhận diện đường thẳng, xác định bằng màu sắc, những thuật toán làm hạn chế sự di ảnh khi xe đang di chuyển.

Kết hợp với những chương trình khác để quản lý kho bãi, quản lý các phương tiện đang tham gia giao thông, tìm xe thất lạc, theo dấu,...

# TÀI LIỆU THAM KHẢO

1. Nhận dạng biển số xe với opencv step by step. <https://thorpham.github.io/blog/2018/04/11/regconite-plate-car/>
2. Find and draw Contours - OpenCV 3.4 with python 3. <https://www.youtube.com/watch?v=_aTC-Rc4Io0>
3. Contour Features. <https://docs.opencv.org/trunk/dd/d49/tutorial_py_contour_features.html>
4. Tìm hiểu về Contour, moments trong xử lý ảnh. <https://congdongopencv.blogspot.com/2017/11/tim-hieu-ve-contour-moments-trong-xu-ly.html>
5. K-nearest neighbor trong opencv2.   
   <https://viblo.asia/p/k-nearest-neighbour-trong-opencv2-V3m5W2wWlO7>
6. Son Nguyen (2017), Machine learning với OpenCV python - Bài 10 THUẬT TOÁN K-Nearest Neighbour láng giềng gần nhất. <https://www.youtube.com/watch?v=ETOqRZIrLY8>