**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC NHA TRANG**

**KHOA CÔNG NGHỆ THÔNG TIN**

**🙚🕮🙘**

**HỌC MÁY**

**ĐỀ TÀI :**

**PHÂN TÍCH VÀ SỬ DỤNG CÔNG NGHỆ HỌC MÁY ĐỂ PHÁT HIỆN XE HƠI TRÊN CAO TỐC**

**GVHD:** Đinh Đồng Lưỡng

**SVTH:** Nguyễn Hưng

**MSSV:** 61132580

**Lớp:** 61.CNTT-CLC

MỤC LỤC

[Chương 1: GIỚI THIỆU SƠ LƯỢC 4](#_Toc117692437)

[1.1 Sơ lược tình trạng hiện nay 4](#_Toc117692438)

[1.2 Giới thiệu vấn đề 4](#_Toc117692439)

[1.3 Mục tiêu nghiên cứu 5](#_Toc117692440)

[1.4 Tầm quan trọng nghiên cứu 5](#_Toc117692441)

[1.5 Hạn chế 5](#_Toc117692442)

[Chương 2: CÁC CÔNG NGHỆ HỌC MÁY 6](#_Toc117692443)

[2.1 Machine Learning 6](#_Toc117692444)

[2.1.1 Học có giám sát (Supervised Learning) 6](#_Toc117692445)

[2.1.2 Học không giám sát (Unsupervised Learning) 8](#_Toc117692446)

[2.2 Support Vector Machine (SVM) 9](#_Toc117692447)

[Chương 3: PHƯƠNG PHÁP NGHIÊN CỨU 10](#_Toc117692448)

[3.1 Công cụ sử dụng 10](#_Toc117692449)

[3.1.1 Python 10](#_Toc117692450)

[3.1.2 Jupyter Notebook 10](#_Toc117692451)

[3.1.3 Dataset 10](#_Toc117692452)

[3.2 Tổng quan thực hiện 11](#_Toc117692453)

[3.2.1 Color Histogram 12](#_Toc117692454)

[3.2.2 HOG 12](#_Toc117692455)

[3.2.3 Classifiers 13](#_Toc117692456)

[3.2.4 Train Test Split 13](#_Toc117692457)

[3.2.5 Sliding window 13](#_Toc117692458)

[3.2.5 Pipeline video 13](#_Toc117692459)

[3.2.6 Mô hình tổng quan 14](#_Toc117692460)

[Chương 4: KẾT QUẢ 14](#_Toc117692461)

[4.1 Tổng quát 14](#_Toc117692462)

[4.2 Cài đặt thử nghiệm 15](#_Toc117692463)

[4.2.1 Nhập bộ dữ liệu (dataset) 15](#_Toc117692464)

[4.2.2 Trích đặc trưng sử dụng HOG 16](#_Toc117692465)

[4.2.3 Tiền xử lý dữ liệu và huấn luyện mô hình 17](#_Toc117692466)

[4.2.3 Xử lý sliding window và hoàn thiện yêu cầu 17](#_Toc117692467)

[Chương 5: TỔNG KẾT 20](#_Toc117692468)

[5.1 Sơ lược 20](#_Toc117692469)

[5.2 Khuyết điểm 20](#_Toc117692470)

[5.3 Tham khảo 21](#_Toc117692471)

Chương 1: GIỚI THIỆU SƠ LƯỢC

1.1 Sơ lược tình trạng hiện nay

Dân số và hệ thống giao thông tăng lên từng ngày, nên nhu cầu quản lý chúng cũng đồng thời tăng lên. Do đó số lượng máy móc phục vụ giao thông như các loại xe đều tăng một cách nhanh chóng. Điều đó cho ta thấy được, các chủ đề mới như giao thông, tai nạn và nhiều vấn đề khác cần được quản lý, vì vậy rất khó để quản lý chúng bằng các phương pháp cũ. Các xu hướng và công nghệ mới từ đây đã được nghiên cứu, tìm thấy và phát minh để xử lý từng giai đoạn của nhân loại. Một trong những yêu cầu này chính là giao thông trên đường cao tốc và thành phố. Nhiều phương án như đèn giao thông, biển báo, ... được đưa ra để đối phó với yêu cầu này. Nhưng như ta đã thấy, những phương án này dường như không đủ sức hoặc không quá hiệu quả khi giải quyết các vấn đề giao thông trên đường cao tốc. Các công nghệ mới như theo dõi các vật thể (bao gồm xe hơi) được phát minh ra để giám sát tự động, từ đó tạo ra các dữ liệu có thể mang lại ý nghĩa cho quá trình ra quyết định. Hiện tượng này đã được sử dụng cho các loại vấn đề khác nhau. Xu hướng mới với Intelligent Transport System (ITS) có nhiều yếu tố mà phát hiện và theo dõi đối tượng là một trong số đó . Hệ thống này được sử dụng để phát hiện các phương tiện, làn đường, biển báo giao thông. Khả năng phát hiện và phân loại các phương tiện cung cấp cho ta khả năng cải thiện giao thông và đường xá, ngăn ngừa tai nạn,…

Con người có thể dễ dàng nhận ra các loại xe trong video hoặc hình ảnh. Trong các thuật toán và chương trình máy tính, nó phụ thuộc nhiều vào các loại dữ liệu đầu vào. Một số hạn chế như thời tiết hoặc thậm chí là cường độ ánh sáng cũng góp vai trò khá quan trọng trong việc làm cho quá trình này trở nên dễ dàng hay khó khăn hơn. Đồng thời, ta còn có các loại và hình dạng khác nhau của từng phương tiện. Và hơn thế, khó khăn nhất là khi chuyển động trong video theo thời gian thực, chúng sẽ có từng kích thước khác nhau theo từng frame.

Có khá nhiều kỹ thuật và phương pháp khác nhau để phát hiện và phân loại phương tiện. Sự đa dạng của các kỹ thuật này nằm trong các thuật toán như là Suport Vector Machine (SVM), Convolutional Neural Network (CNN), Decision Trees, Recurrent Neural Network (RNN),… Trong đồ án này, em sẽ nghiên cứu thuật toán SVM để xác định làm sao nó có thể áp dụng vào đề tài này.

1.2 Giới thiệu vấn đề

Gần đây, phát hiện vật thể đang rất được quan tâm trong các chủ đề nghiên cứu. Các nhà nghiên cứu hiện nay đang cố gắng khám phá về vấn đề này để từ đó việc phát hiện sẽ đạt đến mức độ chính xác cao hơn phù hợp được sự chấp nhận của các nhà nghiên cứu. Có nhiều kỹ thuật thực hiện công việc này nhưng để xác định mô hình tốt nhất trong số các mô hình được đề xuất, đề tài này sẽ đi sâu tìm hiểu chủ đề về cách phát hiện đối tượng (đồng thời so sánh hai mô hình gợi ý và đề xuất mô hình tốt nhất mang lại độ chính xác và hiệu suất cao nhất. )

Việc phát hiện và phân loại các đối tượng trong hình ảnh và video rất phù hợp trong học máy vì nhiệm vụ này là việc phân loại - một việc liên quan đến việc xử lý các tập dữ liệu phức tạp. Hệ thống này quan trọng đối với rất nhiều lĩnh vực, đặc biệt là đối với hệ thống giao thông và giám sát các phương tiện.

1.3 Mục tiêu nghiên cứu

Mục đích của luận văn này là tìm hiểu và phát triển thuật toán SVM để phát hiện xe ô tô trong hình ảnh và video. Nó so sánh hai thuật toán với cùng một tập dữ liệu và các phương pháp tiền xử lý.

Cụ thể ở đây, chúng ta sẽ tìm hiểu và tính toán xem phương pháp SVM sẽ cho ra kết quả với độ chính xác và hiệu suất như thế nào. Cùng với đó đề tài sẽ triển khai bộ phân loại Classifier này để từ đó dự đoán (predict) xem lớp của Vehicles và Non-Vehicles. Bộ phân loại Classifier sẽ dự đoán tọa độ của các bounding box của các chiếc xe ô tô.

1.4 Tầm quan trọng nghiên cứu

Kể từ cuộc cách mạng công nghiệp thì số lượng ô tô đã tăng lên từng ngày. Một trong những thách thức mới của xã hội hiện tại là giao thông, người dân ở các thành phố gần như sử dụng phần lớn thời gian để tham gia giao thông khi có công việc cần phải làm hay chỉ đơn giản là đi dạo quanh thành phố. Vì vậy có một hệ thống giao thông được số hóa hoạt động 24/7 sẽ làm cho các công việc trở nên dễ dàng và hiệu quả hơn, đây cũng là một điều quan trọng đối với tất cả các quốc gia trên toàn cầu. Do đó, có một hệ thống giao thông được vi tính hóa thì yêu cầu tiên quyết là phải có một hệ thống phát hiện xe ô tô. Hệ thống rõ ràng mà nói nó rất có ảnh hưởng đến với nền kinh tế, ngành công nghiệp, hay kể cả cuộc sống của người dân,…

1.5 Hạn chế

Mặc dù thực tế là đề tài đã đạt được các mục tiêu, nhưng sẽ chính xác hơn nếu dataset được cải thiện về số lượng của các loại Vehicles và Non-Vehicles, cụ thể hơn là ta cần rất rất nhiều.

Chương 2: CÁC CÔNG NGHỆ HỌC MÁY

2.1 Machine Learning

Cuộc cách mạng kỹ thuật số đã giải quyết những vấn đề mới mà con người phải đối mặt. Sự phát triển công nghệ nhanh chóng, sự tương tác của con người với các thiết bị điện tử và các công nghệ khác nhau. Từ hai thập kỷ trước, các tổ chức như trường đại học và các nhà khoa học đã cố gắng phát triển các mô hình để sử dụng những thông tin này cho nhiều mục đích khác nhau, chẳng hạn như để **detect**, **recognize**, **invest**, **distinguish**. Trên thực tế, tất cả các doanh nghiệp ngày nay đều sử dụng AI để cải thiện độ chính xác các quy trình và phương pháp làm việc của họ điển hình như nó đang được sử dụng trong các lĩnh vực quan trọng trong xã hội như y tế, kỹ thuật, tài chính, sản xuất,…

Thuật ngữ Machine Learning (ML) được Arthur Samuel sử dụng vào năm 1959 và được tạo ra và hoàn thiện bởi các nhà nghiên cứu khác nhau cho đến nay. ML là một phần của Trí tuệ nhân tạo (AI) sử dụng các chiến lược thống kê để cung cấp khả năng cho các ứng dụng máy tính và cho phép chúng thu được dữ liệu một cách linh hoạt mà không cần hướng dẫn của con người trước đó. Đối với các mục đích này, ML sử dụng các thuật toán máy tính, phần lớn trong số đó là các mô hình được làm việc để thu được từ dữ liệu, đưa ra dự đoán về một loại thông tin tương tự đã được chuẩn bị. Các ứng dụng truyền thống đang hoạt động trong các nguyên tắc hạn chế và ràng buộc, được đặc trưng bởi các kỹ sư phần mềm. Các tính chất tiến bộ và tự học của thuật toán ảnh hưởng đến chúng để chinh phục các ứng dụng thông thường vì việc xây dựng các ứng dụng định hướng thông tin như tầm nhìn máy tính hoặc sàng lọc email, v.v. gần như không khả thi với các chiến lược lập trình truyền thống. Các thuật toán này giúp chúng tôi đưa ra quyết định tốt hơn và mang lại sự đáng tin cậy. ML có 3 loại cơ bản chúng như sau

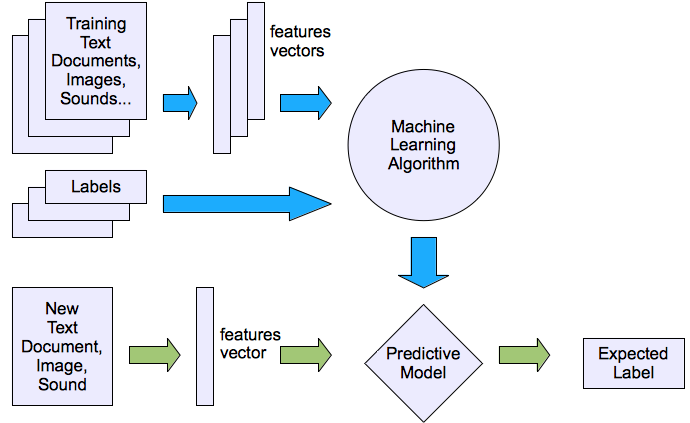
2.1.1 Học có giám sát (Supervised Learning)

Trong hầu hết các trường hợp, học có giám sát thường được sử dụng cho các vấn đề liên quan đến trí tuệ nhân tạo (AI). Có hai nhân tố chủ yếu trong học có giám sát, một nhân tố đầu vào của thông tin để huấn luyện mô hình và một yếu tố đầu ra kết quả tỉ lệ chính xác của mô hình.

Các thuật toán cố gắng tìm ra luật để thành thạo việc ánh xạ giữa hai nhân tố này thông qua một mapping work. Các thuật toán được chuẩn bị với các nguồn thông tin cần thiết sẽ tính toán kết quả để từ đó cho đầu ra có thể chấp nhận được. Trong giai đoạn huấn luyện các mô hình, việc quan trọng nhất cần phải để ý và thực hiện đó chính là kiểm tra việc học (learning) hay nói cách khác là kiểm tra sự đúng đắn của mô hình sau khi nó được huấn luyện.

Tại thời điểm khi các thuật toán dự đoán để đưa ra các đầu ra, nó sẽ kiểm tra xem những kết quả phản hồi phù hợp là đúng hay là sai, kết quả đó như thế nào so với đầu ra, gần hoặc sai khác một khoảng so với đầu ra mong muốn. Các quy trình này sẽ giúp các thuật toán học hỏi và cải thiện hiệu suất của bản thân chúng. Quy trình huấn luyện sẽ kết thúc khi mà nó đạt đến mức độ chính xác mà chúng ta thiết lập. Tại thời điểm khi dữ liệu mới đến, việc nó cố gắng dự đoán kết quả đầu ra phụ thuộc vào quá trình huấn luyện trong quá khứ và sự ánh xạ giữa các đầu vào và đầu ra của thông tin.

Các thuật toán học có giám sát được chia thành hai loại chính gồm hồi quy (Regression) và phân loại (Classification). Có một số vấn đề cho rằng các thuật toán này sẽ hoạt động như thế nào khi các dữ liệu được gắn nhãn, vì dữ liệu cũng giống như các tài nguyên hữu hạn có trên Trái Đất hiện này, tất cả dữ liệu không phải là miễn phí và việc sử dụng thông tin dữ liệu như vậy để huấn luyện mô hình cho các thuật toán sẽ rất khó khăn và tốn kém. Thông thường mà nói, rất nhiều thuật toán học có giám sát đã được các nhà nghiên cứu tiến hành thử nghiệm, và cho thấy rằng mỗi loại trong số chúng đều có những ưu và nhược điểm của riêng từng loại. Từ đó ta có thể thấy được rằng không có một thuật toán học có giám sát nào là xử lý tốt nhất cho một công việc cụ thể nào cả, mỗi loại đều mạnh cũng như yếu hơn ở một khía cạnh nhất định nào đó . Do đó, việc chọn thuật toán phù hợp là một vấn đề quan trọng trước khi thực hiện một dự án cụ thể liên quan đến học máy, đây là một số thuật toán được sử dụng nhiều nhất:

* Support Vector Machines
* Logistic regression
* Linear regression
* Decision trees
* K-Nearest Neighbor
* Neural Networks

Hình 2.1 - Mô tả mô hình Học có giám sát

2.1.2 Học không giám sát (Unsupervised Learning)

(Việc các thuật toán học không giám sát hoàn toàn không giống với các thuật toán học giám sát không có câu trả lời đúng). Như tên gọi của nó, thuật toán học không giám sát không có giá trị biến đầu ra (các nhãn) và một người có thể hướng dẫn nó giải quyết việc học mà nó phải tự mình tìm ra các mẫu nằm ẩn trong các dữ liệu đầu vào. Sẽ không có nhãn nào được gắn vào các dữ liệu để các thuật toán này sử dụng để học hỏi và nâng cao sự chính xác của mô hình mà nó huấn luyện ra. Các thuật toán không giám sát cũng được chia thành hai loại chính bao gồm các bài toán phân cụm (Clustering) và kết hợp (Association).

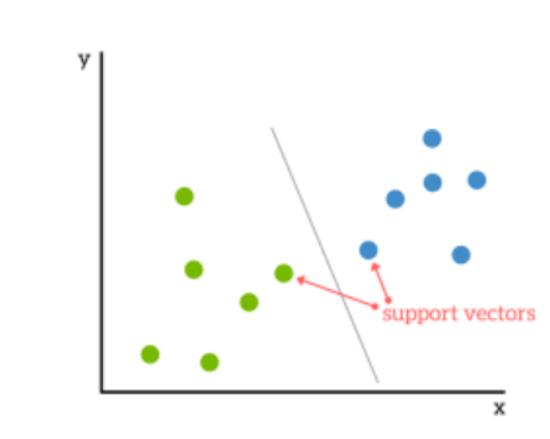
* Phân cụm: các thông tin dữ liệu sẽ được chia thành các nhóm (hay cụm - cluster).
* Kết hợp: các thuật toán sẽ cố gắng hiểu các tiêu chuẩn tương tự của các dữ liệu và từ đó sẽ học được các phần dữ liệu mở rộng (liên quan).

Diagram

Description automatically generated

Hình 2.2 – Mô tả mô hình Học không giám sát

2.2 Support Vector Machine (SVM)

Đây là một loại thuật toán học có giám sát có thể được sử dụng cho các tác vụ phân loại và hồi quy. Cách thức chung trong SVM chính là tìm ra một siêu mặt phẳng từ đó có thể tách tập dữ liệu thành hai lớp tương ứng như Hình 2.3.

Hình 2.3 - Mô tả SVM Splittion

Chương 3: PHƯƠNG PHÁP NGHIÊN CỨU

3.1 Công cụ sử dụng

Giống như các nghiên cứu khác, đề tài này này sử dụng một số các apps và công cụ để sử dụng cho việc huấn luyện mô hình và thực nghiệm trên các mô hình đó. Đối với những nghiên cứu như vậy, nhiều công cụ rất cần được sử dụng để hỗ trợ ta nhanh chóng xử lý những tác vụ khó nhằn mà không mất quá nhiều thời gian, ví dụ như đề tài này đã sử dụng python làm ngôn ngữ chính để phát triển các mô hình hoặc một tập dataset với rất nhiều hình ảnh thuộc lớp Vehicles và Non-Vehicles được sử dụng hoặc cũng như nhiều thư viện python cần thiết,…

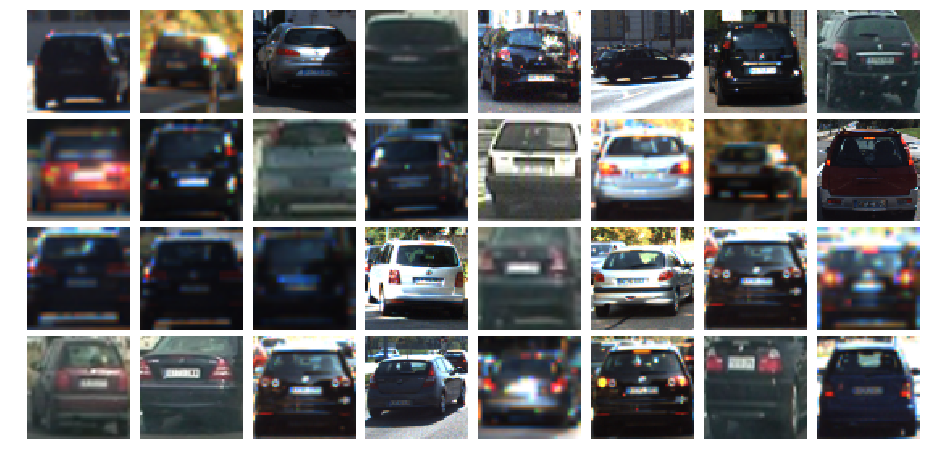
3.1.1 Python

Python là một ngôn ngữ lập trình được sử dụng rộng rãi trong các ứng dụng web, phát triển phần mềm, khoa học dữ liệu và máy học (ML). Mục đích chính của việc tạo python là làm cho việc lập trình trở nên dễ dàng hơn để từ đó mọi người trên thế giới đều có thể tiếp cận và học coding một cách dễ dàng hơn. Hay nói cách khác, Python phổ biến trên toàn thế giới vì sự đơn giản dễ tiếp cận của nó. Mặc dù là python được tạo ra dành cho trẻ em, nhưng hiện tại nó đã được sử dụng để thực thi công việc của hầu hết ngôn ngữ lập trình hết (không phải tất cả) nhiều lĩnh vực. Python có thể làm bất cứ điều gì mà các ngôn ngữ lập trình khác có thể làm. Từ web đến thuật toán đến các window apps,… Trong vài năm qua, python cũng đã dần trở nên phổ biến trong các ứng dụng trí tuệ nhân tạo và học máy do có cung cấp rất nhiều thư viện hiệu quả và tiện dụng, giúp công việc liên quan đến trí tuệ nhân tạo hay học máy dễ dàng và nhanh chóng hơn rất nhiều.

3.1.2 Jupyter Notebook

Jupyter Notebook là một ứng dụng dựa trên web giúp chúng ta có thể tạo và sửa đổi các dòng mã lệnh trực tiếp với các hình ảnh rất trực quan. Đây là một notebook mã nguồn mở hỗ trợ cho rất nhiều ngôn ngữ lập trình. Nó được sử dụng cho các mục đích khác nhau như học máy, mô phỏng số liệu, trực quan hóa các thông tin,… các nhà nghiên cứu đã quyết định sử dụng notebook này để viết các mã lệnh có thể đọc được một cách dễ dàng và thực hiện các thuật toán học máy.

3.1.3 Dataset



Hình 3.1 - Dataset của Vehicles

A picture containing different, window, image, same

Description automatically generated

Hình 3.2 - Dataset của Non-Vehicles

3.2 Tổng quan thực hiện

Phát hiện và theo dõi các phương tiện là tiền đề hết sức quan trọng trong việc xây dựng công nghệ tự lái của các xe ô tô hiện nay. Ở mục tiêu đề tài này, ta sẽ xây dựng một pipeline cho phần mềm để phát hiện các phương tiện trong video.

Để thực hiện việc xây dựng một pipeline, ta cần phải thực hiện các bước sau đây:

* Thực hiện trích chọn đặc trưng Histogram of Oriented Gradients (HOG) trên một tập hợp hình ảnh có gắn nhãn(dataset) được sử dụng để huấn luyện, cùng với đó là huấn luyện một bộ phân loại Linear SVM Classifier dựa trên tập hợp các hình ảnh đó.
* Thực hiện kỹ thuật window sliding và sử dụng bộ phân loại được huấn luyện ở bên trên để từ đó tìm kiếm các phương tiện trong hình ảnh đầu vào.
* Tiến hành chạy pipeline trên một video và tạo một heat mapping của các phương tiện đã được phát hiện lần lượt có trong từng khung hình (frame) của video để từ đó loại bỏ các trường hợp ngoại lệ không cần thiết khi phát hiện và theo dõi các phương tiện đó.
* Ước tính một bounding box cho các phương tiện đã được phát hiện trong các khung hình.

Tập hình ảnh (dataset) chứa hai lớp Vehicles và Non-Vehicles được chuyển đổi sang kiểu dữ liệu dạng mảng và được truyền vào hai danh sách riêng biệt bằng cách sử dụng các function có sẵn của thư viện Numpy. Có nhiều kỹ thuật trích đặc trưng khác nhau đã được sử dụng để huấn luyện bộ phân loại phát hiện các phương tiện một cách hiệu quả.

Mặc dù có thể rất phức tạp với việc ba kênh màu của các hình ảnh đều có độ phân giải đầy đủ, nhưng ta có thể thực hiện phân vùng không gian (spatial binning) trên hình ảnh mà vẫn có thể giữ được đầy đủ thông tin trên các hình ảnh nằm tìm kiếm và phát hiện các phương tiện.

3.2.1 Color Histogram

Đối với bên chuyên nhiếp ảnh, các histogram chỉ đơn giản là một biểu diễn đồ họa của số lượng các pixel trong ảnh được nằm trong một phạm vi, với độ sáng hoặc màu sắc nhất định. Ví dụ: đối với biểu đồ biểu thị độ sáng bình thường, biểu đồ sẽ hiển thị số pixel cho mỗi độ chói (luminance) hay còn gọi là mức độ sáng từ đen sang trắng. Các mức độ trong một khoảng chứa (bins) càng cao thì càng có nhiều pixel ở mức độ chói đó. Color Histogram với nguyên tắc cũng gần giống như thế nhưng thay vì là mức độ sáng đen hoặc trắng, bây giờ ta sẽ có mỗi màu của pixel sẽ nằm trong ba màu (RGB).

Color Histogram đơn giản chỉ là histogram hiển thị mức màu cho từng kênh màu RGB riêng lẻ.

Chúng ta có thể phân biệt hai hình ảnh khác nhau chỉ dựa trên sự khác biệt trong từng histogram của chúng. Như mong đợi, hình ảnh của chiếc ô tô màu đỏ có cường độ giá trị nằm trong khoảng chứa lớn hơn nhiều trong R histogram (Red Channel) so với R histogram của ô tô màu xanh lam, và ngược lại.

3.2.2 HOG

Bộ mô tả đặc trưng có thể nói một cách khác chính là đại diện của hình ảnh giúp ta đơn giản hóa hình ảnh bằng cách trích xuất các thông tin hữu ích và loại bỏ thông tin không liên quan có trong ảnh đang xử lý. Kỹ thuật này đếm số lần xuất hiện của hướng gradient trong các phần được gộp chung lại của hình ảnh. Trong bộ mô tả đặc trưng HOG, sự phân bố (histogram) của các hướng gradient (oriented gradients) được coi như là các đặc trưng được phân bố trong một vùng của ảnh đang xử lý. Các gradients của một ảnh rất có ích khi xác định biên của vật thể nằm trong ảnh đó vì độ lớn của một gradient lớn hơn hẳn các gradients xung quanh (ở các vùng có cường độ thay đổi đột ngột) và chúng ta cũng biết rằng các cạnh và góc chứa nhiều thông tin hơn về biên của đối tượng hơn các vùng phẳng. Tiếp theo ta sẽ chọn các tham số phù hợp để huấn luyện bộ phân loại.

3.2.3 Classifiers

Support Vector Machine (SVM) là một tập hợp các phương pháp học có giám sát được sử dụng cho Classification, Regression và phát hiện ngoại lệ (outlier detection). Đề tài này sẽ sử dụng kernel LinearSVC làm bộ phân loại.

3.2.4 Train Test Split

Hàm StandardScaler() sẽ giả định dữ liệu được thể hiện trong mỗi đối tượng và sẽ chia tỷ lệ chúng sao cho các giá trị dữ liệu hiện tại sẽ tồn tại giữa 0 và 1 ([0, 1]), với độ lệch chuẩn là 1. Các giá trị ‘X’ sẽ được biến đổi bằng cách sử dụng hàm này và nhận được kết quả đầu ra là giá trị của biến scaled\_X.

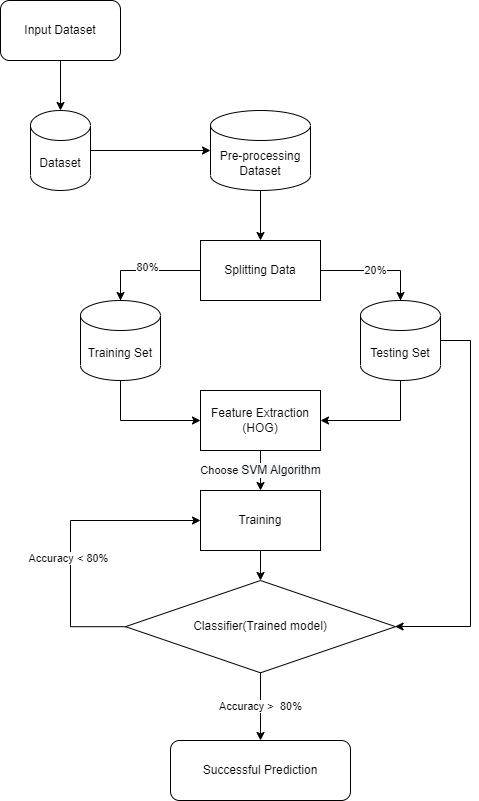
Hàm train\_test\_split() từ thư viện Scikit-Learn sẽ hỗ trợ ta việc phân chia tập dữ liệu (dataset) thành hai tập – training dùng để huấn luyện bộ phân loại (mô hình) và testing dùng để kiểm tra độ chính xác của bộ phân loại, theo tỉ lệ nhất định tùy vào tham số mà người phát triển tùy biến.

3.2.5 Sliding window

Theo như tên gọi của nó, cửa sổ trượt (Sliding window) là một vùng hình chữ nhật có chiều rộng và chiều cao cố định “trượt” qua một hình ảnh. Đối với mỗi window, chúng ta thường lấy vùng window và áp dụng bộ phân loại hình ảnh ở trên để xác định xem trong window đang xét có đối tượng mà chúng ta cần phải phát hiện hay không.

3.2.5 Pipeline video

Cuối cùng tạo video pipeline bằng cách xử lý từng khung hình của hình ảnh với các kỹ thuật trên và tạo video từ các khung đã xử lý. Hàm find\_cars () trích tất cả các bounding box được phát hiện cho các chiếc xe trong ảnh.

3.2.6 Mô hình tổng quan

Hình 3.3 – Sơ lược mô hình phát triển đề tài

Chương 4: KẾT QUẢ

4.1 Tổng quát

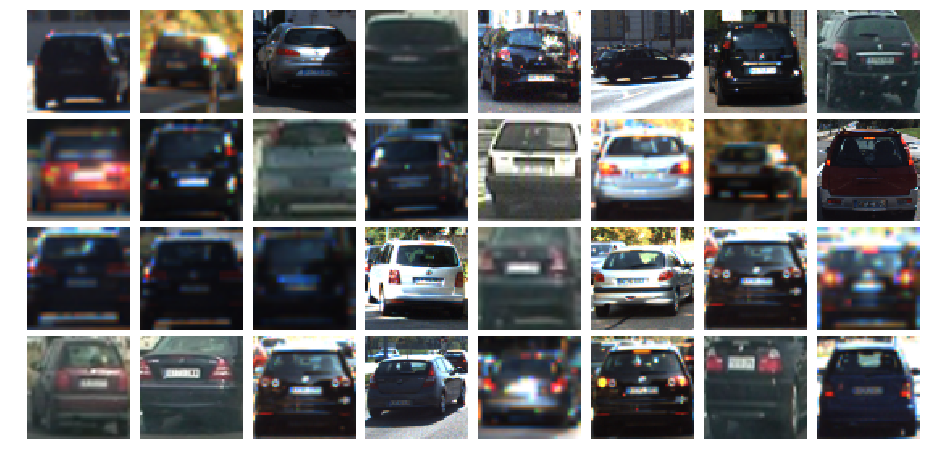
Trong quá trình phát triển các thuật toán học máy, có rất nhiều ngôn ngữ mà ta có thể lựa chọn, chẳng hạn như MATLAB, Python và R. Vì mỗi lựa chọn đều có ưu nhược điểm nhất định, nên các nhà nghiên cứu đã quyết định chọn python do tính dễ dàng và thư viện phong phú có sẵn cho việc xử lý các công việc. Nhiều thư viện đã được sử dụng để thực hiện các công việc khác nhau như Numpy được sử dụng để nhập và xử lý dữ liệu, matplotlib được sử dụng để hiển thị các đặc trưng màu được trích xuất, Scikit-Learn được dùng để tìm hiểu cách tách dữ liệu thành các phần để training và testing, các chức năng classify để tạo ra các mô hình và huấn luyện chúng với dữ liệu có sẵn trong dataset.

Thuật toán SVM được sử dụng hầu hết cho các công việc phân loại. Các mô hình này hoạt động dựa trên việc khám phá ra một siêu phẳng chia dữ liệu thành các lớp

* SVM hoạt động một cách hoàn hảo khi xử lý các dữ liệu chưa xác định hay các dữ liệu bán cấu trúc hoặc phi cấu trúc
* Trong SVM thông thường sẽ không bị overfitting

4.2 Cài đặt thử nghiệm

4.2.1 Nhập bộ dữ liệu (dataset)

Giống như đã đề cập trước đó, có rất nhiều thư viện có thể giúp chúng ta triển khai các mô hình một cách dễ dàng. Đầu tiên, ta phải import các thư viện này vào hệ thống, sau đó ta sẽ đọc và tải các dữ liệu vào các arrays đã được chuẩn bị sẵn trong hệ thống.

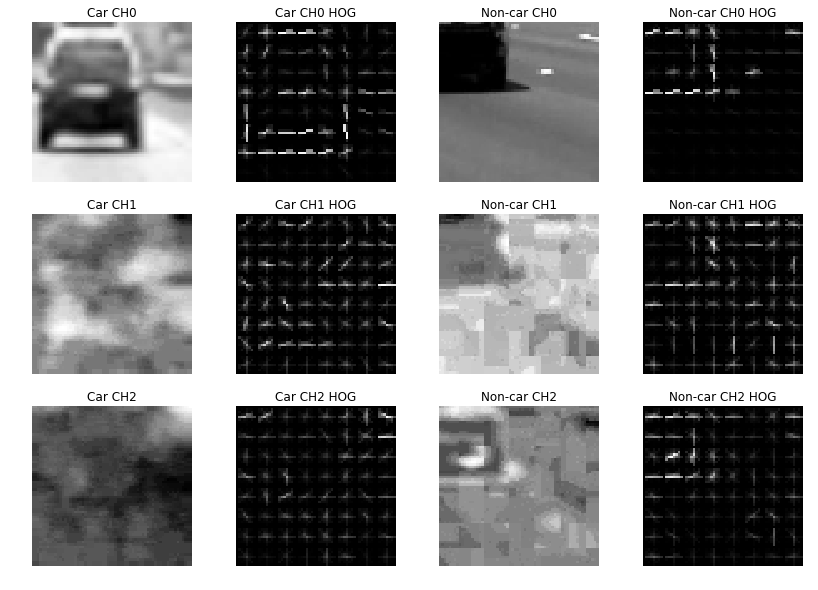
Hình 4.1 - Dataset của Vehicles



Hình 4.2 - Dataset của Non-Vehicles

4.2.2 Trích đặc trưng sử dụng HOG

Một Histogram là một mô tả chính xác về việc phổ biến các thông tin dưới dạng số. Nó chính là thước đo khả năng lặp lại của một giá trị biến liên tục. Nó thay đổi so với biểu đồ bar, vì trong biểu đồ bar có liên quan đến hai yếu tố, nhưng histogram chỉ liên quan đến một yếu tố. Để phát triển một histogram, bước đầu tiên là đặt các bins — tức là, phân chia toàn bộ phạm vi giá trị thành một chuỗi các khoảng thời gian — và sau đó kiểm tra xem số lượng giá trị rơi vào mỗi bin. Các bin (phạm vi) phải liền kề và thường xuyên (không bắt buộc) có kích thước tương tự.

Sau khi tìm ra được các histogram từ việc xét các phạm vi giá trị, ta sẽ trích xuất ra được các HOG (Histogram of Oriented Gradients). HOG là một bộ mô tả đặc trưng được sử dụng trong xử lý ảnh và thị giác máy tính (computer vision) nhằm mục đích phát hiện đối tượng. Kỹ thuật này đếm số lần xuất hiện của hướng gradient (vector) trong các phần được thu hẹp lại của hình ảnh. Phương pháp này tương tự như phương pháp của edge oriented histograms, bộ mô tả biến đổi đặc trưng với tỉ lệ bất biến, nhưng khác ở chỗ nó được tính toán trên một lưới dày đặc các ô có khoảng cách đều nhau và sử dụng chuẩn hóa tương phản cục bộ chồng chéo để cải thiện độ chính xác.

Hình 4.3 - Ví dụ HOG trên vài ảnh mẫu

4.2.3 Tiền xử lý dữ liệu và huấn luyện mô hình

Ở bước tiếp theo, ta sẽ thực hiện tiền xử lý dữ liệu (data pre-processing) – một bước hết sức quan trọng trước khi ta training bất kì một thuật toán hay các thao tác chia tỉ lệ,… Vì tập dữ liệu đầu vào (dataset) là một tập các hình ảnh, do đó ta cần phải thực hiện chia nhỏ dữ liệu.

A picture containing scatter chart

Description automatically generatedDữ liệu được chia thành hai phần training và testing nhờ sử dụng hàm train\_test\_split() từ thư viện Scikit-Learn. Tỷ lệ phần trăm cho cả hai ta đặt mặc định là 80% cho tập data training và 20% cho testing(các dữ liệu cần phải được chuẩn hóa). Trong tiến trình xử lý của trình phân loại SVC, hàm StandardScaler() được sử dụng để chuẩn hóa dữ liệu.

Hình 4.4 - Kết quả training SVC Classifier

4.2.3 Xử lý sliding window và hoàn thiện yêu cầu

Khi mô hình được train với phần training của dataset, sliding window là công việc tiếp theo mà ta cần phải tiến hành. Sau khi chạy sliding, hệ thống cần tìm các cửa sổ mà chúng ta sẽ chạy trình phân loại. Hàm này sẽ trả về các cửa sổ đã lọc trong đó bộ phân loại dự đoán kết quả đầu ra là một chiếc xe ô tô. Sau đó, chức năng vẽ cửa sổ chính xung quanh những chiếc xe được thực hiện như là một điều kiện cần thiết để xác định các cửa sổ. Sliding window và chiếc xe được xác định trong cửa sổ được vẽ xung quanh chiếc xe đó trong hình 4.5 và 4.6.

A screenshot of a video game

Description automatically generated with medium confidenceGraphical user interface, application, PowerPoint

Description automatically generatedA road with trees on the side

Description automatically generated with low confidence Quá trình vẽ các ô xung quanh những chiếc xe càng chính xác hơn. Sau đó, tất cả các heat mapping được áp dụng để tăng pixel bên trong mỗi ô. Sau đó, chúng ta cần áp dụng một giá trị ngưỡng cho hình ảnh để lọc ra các ô pixel thấp và tìm các pixel với mỗi số xe và vẽ các ô giới hạn cuối cùng. Hình ảnh bản đồ nhiệt thử nghiệm được thể hiện trong hình 4.7.

Hình 4.5 - Window Coverage

Hình 4.6 - Window Refined

Hình 4.7 - Heat Map trên ảnh thử nghiệm

Graphical user interface, application, PowerPoint

Description automatically generatedSau bước này, mô hình được huấn luyện đã phát hiện được các chiếc ô tô và vẽ các bounding box cho chúng, có thể nói mục tiêu phát hiện các chiếc ô tô đã được hoàn thành. Mặc dù có thể nói rằng hệ thống này có thể phát hiện những chiếc ô tô nhưng không thể nào theo dõi chúng trong một đoạn video. Để có thể làm được điều đó, ta cần phải tạo ra một pipeline với video truyền vào được sử dụng bởi thư viện VideoFileClip(). Hình 4.8 biểu thị ví dụ của pipeline trên từng ảnh thử nghiệm, hay đó là các ảnh nằm trong từng frame của video mà chúng ta sẽ thử nghiệm sau. Trong khi những chiếc ô tô chạy qua trong video, chương trình sẽ phát hiện và theo dõi những chiếc xe bằng các bounding box được bao phủ xung quanh chiếc ô tô đó.

Hình 4.8 - Ví dụ Pipeline trên ảnh thử nghiệm

Chương 5: TỔNG KẾT

5.1 Sơ lược

Sự phát triển nhanh chóng của các ngành công nghiệp ô tô và giao thông cũng như sự gia tăng dân số trên toàn thế giới đã kéo theo nhu cầu về các công nghệ cần thiết, đặc biệt trong đó là các giải pháp công nghệ để quản lý giao thông trong các khu vực giao thông thành phố và khu vực đông dân cư. Trong đó, phát hiện vật thể là một công nghệ thiết yếu để sử dụng cho các vấn đề giao thông hiện nay, nó cũng có thể được sử dụng ở rất nhiều lĩnh vực khác nhau để góp phần thêm sự tiện nghi, thuận tiện cho con người.

Cụ thể phát hiện vật thể được sử dụng trong các ngành công nghiệp hiện đại, đô thị thông minh, nghiên cứu, môi trường,... Ta có thể thấy rằng tầm quan trọng của công nghệ này ngày càng lớn hơn khi mà thế giới đang trải qua Cách mạng Công Nghiệp 4.0. Do đó có thể nói rằng nghiên cứu này góp phần không nhỏ trong việc phát triển nền kinh tế và xã hội cụ thể là về giao thông – một trong số các vấn đề thiết yếu khi phát triển đất nước.

Đề tài này em đã sử dụng và phát triển thuật toán phân loại học có giám sát Support Vector Machine (SVM) để huấn luyện và tính toán độ chính xác của mô hình đó dựa trên những kết quả đã thực hiện ở chương trước đó. Với độ chính xác khá cao lên tới 98%, mô hình này gần như đã hoàn thành mục tiêu phát hiện các phương tiện ở trong một hình ảnh. Nhưng nó vẫn thiếu sót những điều kiện nhất định để trở thành một mô hình hoàn chỉnh trong việc sử dụng phát hiện các phương tiện.

5.2 Khuyết điểm

* Bộ dữ liệu (dataset): do dataset còn rất hạn chế (chỉ khoảng 16000 file ảnh) nên việc tính toán độ chính xác đối với các phương tiện nếu không nằm trong dataset này dường như là không thể.
* Thuật toán SVM: thời gian huấn luyện cho mô hình của thuật toán này sẽ lớn hơn khi dataset ngày càng lớn; thuật toán này sẽ giảm hiệu suất đáng kể khi một số dữ liệu đầu vào bị nhiễu

5.3 Tham khảo

1. [Support Vector Machines](https://scikit-learn.org/stable/modules/svm.html)
2. [Dataset](https://www.kaggle.com/datasets/brsdincer/vehicle-detection-image-set)
3. [Source Code](https://github.com/charleswongzx/Vehicle-Detection)