

Đồ án xử lý ảnh

Hà Huy Dũng - 1510551

Lương Hoài Thiện

Nguyễn Huỳnh Đức

Trần Lê Đức Trung

Ngày 30 tháng 12 năm 2019

Mục lục

1	Tổng quan về đề tài	1
1.1	Bài toán đặt ra và mục tiêu	1
1.2	Các phương hướng giải quyết bài toán	2
2	Nhận dạng các vật thể di động	2
2.1	CNN	2
2.2	YOLO - You Only Look Once	2
3	Tính toán để đưa ra các quyết định	2
4	Kết quả	4
4.1	KTIT dataset	4
5	Kết Luận	4
5.1	Nhận xét	4
5.2	Hướng phát triển sau này	4
6	Tài liệu tham khảo	4

Danh sách hình vẽ

1	Phân vùng hình ảnh	4
2	Giải thuật ra quyết định	5

Danh sách bảng

1 Tổng quan về đề tài

1.1 Bài toán đặt ra và mục tiêu

Yêu cầu bài toán: Viết chương trình đọc các video từ camera hành trình của xe hơi, nhận dạng được các vật cản di động và tính toán để biết được xe có thể

tránh được hay cần dừng lại.

- Vì các vật thể di động trên đường có thể rất đa dạng và là một bài toán lớn nên trong phạm vi đồ án sẽ chỉ xét các vật thể di động thông dụng trên đường sau: người, xe đạp, xe hơi, xe máy, xe tải.
- Việc quyết định xe có thể tránh được hay dừng lại tùy thuộc rất nhiều vào tình huống thực tế nên trong phạm vi đồ án sẽ chỉ cố gắng đưa ra các quyết định đơn giản: rẽ trái, rẽ phải, đi thẳng và dừng để tránh dẫn tới việc va chạm.

Theo như yêu cầu của bài toán thì thông số cần tối thiểu là False Negative, tức là có nhưng không nhận dạng được.

1.2 Các phương hướng giải quyết bài toán

Về nhận dạng các vật thể, hiện nay có các hướng giải quyết chính như sau:

1. Machine Learning

- Viola-Jones object detection framework based on Haar features
- Scale-invariant feature transform
- Histogram of oriented gradients

2. Deep Learning

- Region Proposals (R-CNN, Fast R-CNN, Faster R-CNN)
- Single Shot MultiBox Detector
- You Only Look Once
- Single-Shot Refinement Neural Network for Object Detection (RefineDet)

Các phương hướng tiếp cận của Machine Learning rất phụ thuộc vào đặc tính hình của video và cần phải chỉnh sửa rất nhiều thông số mới có thể hoạt động tốt trên các video nhất định. Vì vậy, đồ án này sẽ sử dụng phương hướng tiếp cận là Deep Learning, các giải thuật này cho kết quả tốt hơn và lượng dữ liệu cho bài toán hiện nay là rất lớn.

Cụ thể hơn, giải thuật sẽ sử dụng là thuật toán YOLO và việc đưa ra các quyết định sẽ dựa trên các dữ liệu đầu ra của thuật toán này.

2 Nhận dạng các vật thể di động

2.1 CNN

2.2 YOLO - You Only Look Once

3 Tính toán để đưa ra các quyết định

Việc đưa ra quyết định của thuật toán phát triển bởi nhóm sẽ dựa vào các dữ liệu đầu vào là tâm của các vật thể đang di động mà mạng trí tuệ nhân tạo trả về sau quá trình xử lý ở phần trên. Vị trí của những tâm vật thể sẽ chứa

nhiều thông tin về vị trí tương đối của vật thể so với xe, cụ thể là camera nằm trên xe. Ta sẽ bỏ qua kích thước cụ thể của vật thể trong ảnh, vì việc tính toán chính xác kích thước vật thể sẽ làm phức tạp bài toán hơn và độ chính xác cao đến từ việc tính toán kích thước vật thể cũng không cần thiết trong dự án lần này.

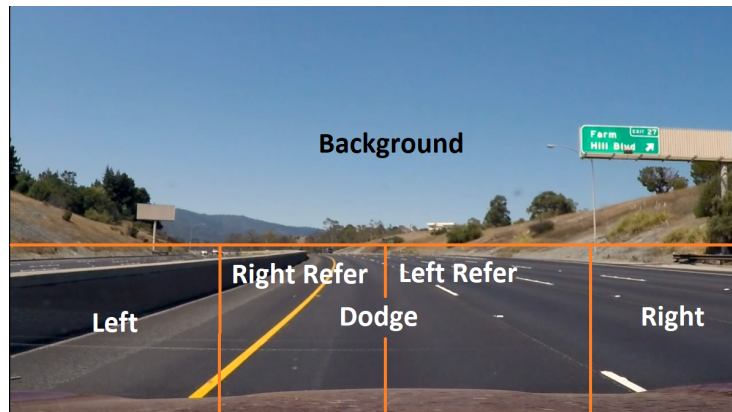
Ý tưởng của thuật toán quyết định là chia ảnh đầu vào thành các vùng khác nhau và việc tâm vật thể nằm trong vùng nào sẽ ảnh hưởng đến quyết định cuối cùng. Cụ thể, ta chia ảnh ra thành những vùng chính:

- **Background:** Trong hình ảnh, sẽ tồn tại một phần không phải là đường chạy của xe. Phần này có thể là bầu trời, nhà cao tầng,... nằm trong hình ảnh thu được bởi camera. Trong dự án này, ta sẽ bỏ qua việc làn đường bị giới hạn. Cụ thể, ta có thể xem xe có thể rẽ phải hoặc rẽ trái tùy ý, mà không xét đến lề đường, nhà cao tầng,... Mục đích cuối cùng là xác định hướng rẽ tránh được vật thể mà không va chạm vật thể khác. Vì thế, những vật thể có tâm nằm trong phần Background sẽ được bỏ qua.
- **Zone of Interest (ZOI):** Vùng không phải là Background sẽ được xem là vùng ZOI. ZOI sẽ được chia ra làm nhiều vùng nhỏ hơn, và việc tâm vật thể nằm trong vùng phụ nào của ZOI sẽ ảnh hưởng đến quyết định của thuật toán.
- **Right:** Vùng nằm bên phải của bức ảnh, và nằm trong vùng ZOI. Các tâm vật thể nằm trong vùng này cho thấy rằng bên phải không an toàn, vì có vật thể đang chắn. Việc tồn tại tâm nằm trong vùng này sẽ quyết định xe có được rẽ phải hay không.
- **Left:** Vùng nằm bên trái của bức ảnh, và nằm trong vùng ZOI. Tương tự như vùng Right, tồn tại tâm vật thể nằm trong vùng Left có nghĩa là việc rẽ trái là cần phải cân nhắc.
- **Dodge:** Nằm giữa bức ảnh, và nằm trong vùng ZOI. Vật thể nằm trong vùng này hàm ý rằng đường đi đang bị cản trở, và xe sẽ phải ra quyết định rẽ trái, rẽ phải, hoặc đi chậm lại. Việc phương tiện ưu tiên rẽ phải hoặc rẽ trái sẽ tùy thuộc vào tâm vật thể đó nằm trong phần nào của vùng Dodge. Cụ thể, ta chia vùng Dodge ra thành hai vùng Left Refer và Right Refer. Nếu vật thể nằm trong vùng Left Refer, phương tiện sẽ phải rẽ trái và tương tự, nếu vật thể nằm trong Right Refer thì phương tiện sẽ phải rẽ phải.

Việc phân chia các vùng có thể được minh họa cụ thể một cách khái quát bằng Hình.3

Sau khi nhận được tọa độ tâm của các vật thể, bằng việc xem xét vị trí tương ứng của từng tâm rơi vào vùng nào, thuật toán sẽ đưa ra quyết. Nhóm đề ra giải thuật như sau:

1. Kiểm tra có tồn tại vật thể trong vùng Dodge hay không? Nếu không có, tiếp tục đi thẳng và xét tiếp các tâm vật thể khác.
2. Nếu tồn tại ít nhất một vật thể trong vùng Dodge, xác định xem tâm đó nằm trong phần nào của vùng Dodge (Left Refer hoặc Right Refer).



Hình 1: Phân vùng hình ảnh

3. Nếu vật thể nằm trong vùng Left Refer, ta kiểm tra xem bên vùng Left có tồn tại vật thể nào hay không. Nếu có, ta cho phương tiện chạy chậm lại, nếu không, ta cho phương tiện rẽ trái.
4. Nếu vật thể nằm trong vùng Right Refer, ta kiểm tra xem vùng Right có tồn tại vật thể hay không. Nếu có, ta cho phương tiện chạy chậm lại, nếu không, ta cho phương tiện rẽ trái.

Thuật toán được minh họa bằng sơ đồ giải thuật trong Hình.3.

4 Kết quả

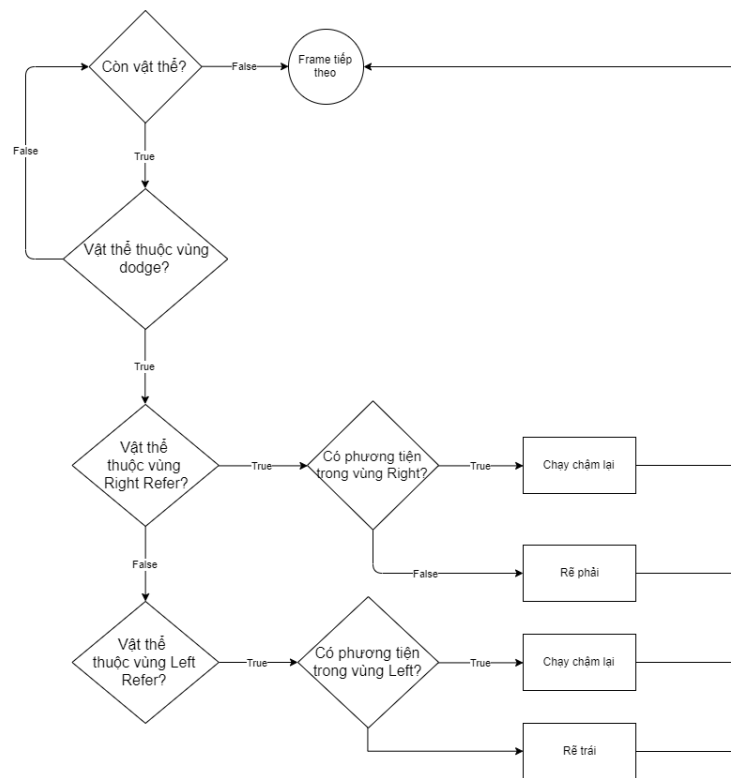
4.1 KTIT dataset

5 Kết Luận

5.1 Nhận xét

5.2 Hướng phát triển sau này

6 Tài liệu tham khảo



Hình 2: Giải thuật ra quyết định