

BÀI THU HOẠCH

CHỦ ĐỀ: VISION IN SELF-DRIVING CARS

I. Problem statement

- Input: bao gồm các dữ liệu từ các cảm biến trên xe, bao gồm camera, radar và lidar, các dữ liệu về bản đồ và môi trường xung quanh, cùng với các quy tắc và hướng dẫn về lưu lượng giao thông và an toàn.
- Output: các quyết định liên quan đến điều khiển xe, bao gồm hướng đi, tốc độ và các hành động khác để đảm bảo xe di chuyển an toàn, hiệu quả và thuận tiện cho người sử dụng

Trong hệ thống xe tự hành của Tesla:

- Input: 8 camera được đặt xung quanh xe, thu ảnh trong thời gian thực để xử lý.
- Output: vector space không gian 3D chứa các thông tin xung quanh xe cần trong quá trình lái xe (làn xe, vạch kẻ đường, biển báo, đèn giao thông, xe cộ xung quanh, người, tốc độ, ...)

II. Challenge

- Yêu cầu công nghệ tự lái phải hoạt động một cách chính xác để bảo vệ tính mạng của người dùng.
- Điều kiện đường xá sẽ thay đổi theo từng nơi. Trong một số trường hợp, có những đường cao tốc rộng rãi, có những nơi có đường sá xuống cấp, không có vạch phân lane, có ổ gà, đường hầm, tín hiệu chỉ đường không rõ ràng.
- Điều kiện thời tiết cũng có ảnh hưởng, một ô tô tự lái thì phải hoạt động được trong mọi điều kiện thời tiết.
- Có những nơi không có đèn giao thông, sẽ rất khó để điều khiển. Giả sử người qua đường xuất hiện, các ô tô tự lái phát hiện và dừng lại, nếu có nhiều người di chuyển ở nơi đó với nhiều thời gian liên tiếp nhau, thì các ô tô đó phải chờ đợi đến khi hết người, điều đó có thể dẫn đến tắc nghẽn giao thông.

III. Methods

Các khía cạnh quan trọng của xe tự lái là: phát hiện làn đường (lane detection), phát hiện phương tiện và chướng ngại vật (vehicle and obstacle detection), traffic detection và feature extraction bằng cách sử dụng các cameras được ghép nối với các phần mềm thị giác máy tính.

- **Lane detection**

- Trong xe tự lái, việc phát hiện làn đường được sử dụng để hỗ trợ giữ cho xe trong một làn đường cụ thể. Nó cũng đóng một vai trò quan trọng trong việc di chuyển xe đến một làn đường khác. Các nghiên cứu đã đề xuất các mô hình phát hiện làn đường dựa trên việc kết hợp thông tin từ thị giác máy tính và cảm biến. Các kỹ thuật như sensor fusion và path planning được sử dụng để xác định đường đi tối ưu giữa hai điểm.
- Các phương pháp phát hiện làn đường bao gồm việc lựa chọn màu sắc, chọn đường thẳng, chọn cạnh và biến đổi Hough. Tuy nhiên, các thuật toán này không hoạt động tốt trong điều kiện ánh sáng yếu, như ban đêm. Vì vậy, các nghiên cứu khác đề xuất sử dụng độ chênh lệch tốc độ giữa các xe trên đường để giúp phát hiện làn đường. Để phát hiện được các làn đường trong môi trường khác nhau như ánh sáng thay đổi, các thuật toán dựa trên nhiều đặc trưng đầu vào khác nhau được đề xuất, bao gồm gradient-based features, intensity-based features và texture-based features.

- **Obstacle detection for self-driving cars:** Phát hiện chướng ngại vật như xe cộ và người đi bộ.

- a. Sử dụng một hệ thống học sâu được gọi là Faster Region based CNN (R-CNN):

Phát hiện và phân loại đối tượng bằng cách sử dụng mạng nơ-ron tích chập dựa trên khu vực để phát hiện và phân loại các đối tượng như xe cộ, người đi bộ và động vật. Phát hiện xe cộ là phức tạp do sự khác biệt về kích thước xe, điều kiện ánh sáng và môi trường xung quanh. Phương pháp này sử dụng một

biến thể của mạng nơ-ron tích chập được gọi là Faster Region based CNN (R-CNN) để phát hiện và phân loại đối tượng. R-CNN trước tiên nhận vào một hình ảnh đầu vào thô và trích xuất các đặc trưng. Sau đó, nó sử dụng phân mẫu trên các tầng khác nhau của mạng nơ-ron và cũng áp dụng các hàm kích hoạt để phân loại các đối tượng. Nhiều thương hiệu ô tô và công ty xe tự hành đang đầu tư hàng tỉ đô la vào công nghệ tự lái. Các ví dụ bao gồm những tên tuổi công nghệ như Tesla.

- b. Sử dụng một thuật toán xử lý ảnh gọi là feature matching (khớp đặc trưng):

Phát hiện vật cản trên đường như xe cộ hay người đi bộ. Thuật toán này đầu tiên trích xuất các đặc trưng từ ảnh và phân loại chúng thành các điểm chính để xác định các đối tượng khác nhau trong ảnh. Sau đó, quá trình khớp đặc trưng so sánh các đặc trưng của khung hình hiện tại với các đặc trưng của các đối tượng đã biết. Khi các đặc trưng khớp nhau, một đối tượng có thể được phát hiện.

Để kiểm soát nhiễu, thuật toán sử dụng một bộ lọc gọi là "David Lowe's ratio test". Sau quá trình xử lý, một hình ảnh được lọc có thể được vẽ với một đa giác bao quanh để đại diện cho khu vực quan tâm. Khoảng cách giữa các đối tượng được tính toán bằng cách sử dụng công thức khoảng cách, được định nghĩa:

$$D = 2H / 0.026A$$

Đây là công thức tính khoảng cách D của một đối tượng từ camera, dựa trên chiều cao H của đối tượng và chiều cao A được tính toán từ camera.

Đặt một ngưỡng giá trị cho D, dưới mức đó các hành động của xe cộ phải được thực hiện. Bằng cách này, thuật toán sử dụng trích xuất đặc trưng trong xử lý ảnh để phát hiện các đối tượng và khoảng cách giữa xe cộ và các đối tượng đã phát hiện cung cấp thông tin để điều khiển xe cộ trong các kịch bản khác nhau.

- **Computer vision cho self-driving cars:**

Các vấn đề cần được giải quyết khi sử dụng computer vision trong xe tự hành:

- Occlusion: khi một phần đối tượng bị che khuất bởi 1 đối tượng khác.
- Inter-class variability: sự khác biệt giữa các lớp đối tượng.
- Pose variability: sự thay đổi màu sắc và kết cấu của đối tượng khi nó được xoay trên mặt phẳng 2 chiều.

Để giải quyết occlusion, YOLO là một trong những phương pháp CNN real time được đề xuất để phát hiện đối tượng trong ảnh.

- **Hệ thống đề xuất cho việc nhận diện làn đường trên ô tô tự lái:**

Sử dụng các kỹ thuật image processing: Đầu tiên, ảnh được chuyển đổi sang ảnh xám và được làm mịn để giảm nhiễu bằng cách sử dụng một hàm Gaussian Blur. Tiếp theo, áp dụng hàm Canny để phát hiện cạnh. Sau khi áp dụng hàm Canny, các cạnh trong ảnh được lấy bằng cách đo độ dốc giữa các pixel liền kề. Sau đó, một khu vực quan tâm được tạo ra tương ứng với nửa dưới của ảnh vì đó là nơi các làn đường được tìm thấy. Tiếp theo, đường làn đường được tạo ra bằng cách sử dụng phép biến đổi Hough. Cuối cùng, hình ảnh được kết hợp bằng cách chồng các đường làn với ảnh gốc.

- **Advanced lane detection using computer vision:**

Để cải thiện việc nhận dạng làn đường, ta có thể sử dụng các kỹ thuật trượt cửa sổ (sliding window), ngưỡng khác nhau cho không gian màu sắc và độ dốc, biến đổi góc nhìn để phù hợp với đường lái xe. Để đạt được sự cải thiện này, các bước sau cần được thực hiện:

- Tính toán ma trận hiệu chỉnh máy ảnh và hệ số méo hình học
- Áp dụng hiệu chỉnh méo hình ảnh
- Tạo một hình ảnh nhị phân ngưỡng bằng cách sử dụng chuyển đổi màu sắc và độ dốc
- Tạo ra một góc nhìn chim của hình ảnh bằng cách sử dụng một biến đổi góc nhìn

- Phát hiện các điểm ảnh và ranh giới của làn đường
- Phát hiện vị trí của xe và độ cong của làn đường liên quan đến trung tâm
- Hiện thị ước tính số liệu về độ cong của làn đường và vị trí của xe và lớp phủ các đường biên làn đã phát hiện lên hình ảnh gốc.

- **Transfromer networks:**

Spatial Transformer Network (STN) - một kỹ thuật được sử dụng để loại bỏ tính không đổi không gian từ hình ảnh và tăng độ chính xác của bộ phân loại trong mạng neuron tích chập. STN giúp giảm các vấn đề gây ra bởi biến đổi đầu vào, bao gồm tỷ lệ, góc nhìn và nền nhiễu. STN được thiết kế để hoạt động theo các giai đoạn, bao gồm định nghĩa ma trận biến đổi tuyến tính, tạo lưới chồng và sử dụng mạng neuron để học và tạo ra theta, một hàm kích hoạt được chuyển đến bilinear sampler để thay đổi đầu vào của mạng neuron tiếp theo. STN có thể được chèn vào bất kỳ phần nào của mô hình và được đào tạo bằng một thuật toán back propagation.

- **Traffic sign detection in self-driving cars:**

Việc phát hiện và diễn giải chính xác các biển báo giao thông là rất quan trọng đối với hoạt động của một xe tự hành. Một mạng nơ-ron tích chập sử dụng Adam Optimizer được sử dụng để phát hiện các biển báo giao thông. Thuật toán tối ưu hóa là rất quan trọng vì nó cải thiện thời gian xử lý từ vài phút đến vài giờ, thậm chí là vài ngày.

Adam Optimizer là sự mở rộng của thuật toán stochastic gradient descent và được sử dụng rộng rãi trong các giải pháp cho các vấn đề liên quan đến thị giác máy tính. Adam Optimizer có nhiều lợi ích, bao gồm đơn giản hóa quá trình triển khai, hiệu suất tính toán tốt, yêu cầu bộ nhớ nhỏ, khả năng phát hiện đối tượng không tĩnh, phù hợp với các bài toán phức tạp và lượng dữ liệu lớn. Adam Optimizer kết hợp ưu điểm của Adaptive Gradient Algorithm và Root Mean Square Propagation technique. Adam Optimizer sử dụng trung bình của hai moment đầu tiên và thứ hai của gradient để tính toán learning rate và giảm thiểu sai số.

IV. Conclusion:

Điều khiển tự động là một trong những xu hướng công nghệ phát triển mạnh mẽ nhất của thế kỷ 21. Xe tự lái sử dụng các công nghệ mới như máy học, trí tuệ nhân tạo, cảm biến, hệ thống giám sát và dẫn đường để tự động hoá việc lái xe. Trong tương lai, nó sẽ là một phần không thể thiếu của giao thông đô thị và môi trường đô thị thông minh.

Các ứng dụng của xe tự lái rất đa dạng, từ vận chuyển khách hàng và hàng hóa đến giải trí và du lịch. Nó cũng có thể được sử dụng để giảm thiểu tai nạn giao thông và tăng cường an toàn cho người lái và hành khách. Hơn nữa, nó có thể giúp giảm thiểu tắc nghẽn giao thông và ùn tắc, làm giảm tiêu thụ năng lượng và khí thải.

Tuy nhiên, việc triển khai công nghệ này còn đối mặt với nhiều thách thức. Các hệ thống phải được thiết kế để đảm bảo tính an toàn và đáng tin cậy trong mọi tình huống, bao gồm cả những tình huống khẩn cấp. Hơn nữa, việc phát triển phần mềm và thiết bị phải được kiểm soát chặt chẽ để đảm bảo tính tương thích và ổn định.