

BÁO CÁO PHÂN TÍCH CHUYÊN SÂU HỆ THỐNG NHẬN DIỆN BIỂN BÁO GIAO THÔNG

1. Giới thiệu bài toán

Trong bối cảnh giao thông ngày càng phức tạp, việc đảm bảo an toàn cho người tham gia giao thông là một vấn đề cấp thiết. Một trong những yếu tố quan trọng hỗ trợ quá trình lái xe an toàn là khả năng nhận biết nhanh chóng và chính xác các biển báo giao thông. Tuy nhiên, trong thực tế, việc quan sát bằng mắt thường của con người thường bị hạn chế bởi tốc độ di chuyển, điều kiện thời tiết, ánh sáng hoặc sự tập trung của tài xế. Do đó, việc ứng dụng các phương pháp xử lý ảnh và trí tuệ nhân tạo để tự động tìm và phân loại biển báo giao thông đang ngày càng được quan tâm nghiên cứu.

Bài toán Tìm và phân loại biển báo giao thông (Traffic Sign Detection and Classification) bao gồm hai bước chính:

1. Tìm (Detection):

2. Xác định vị trí của biển báo trong ảnh hoặc trong khung hình video. Đây là bước quan trọng nhằm khoanh vùng đúng khu vực có chứa biển báo, phân biệt với các yếu tố nền như cây cối, xe cộ hay nhà cửa.

3. Phân loại (Classification):

4. Sau khi tìm được biển báo, hệ thống sẽ nhận diện loại biển báo tương ứng (ví dụ: biển cấm, biển chỉ dẫn, biển nguy hiểm, biển hiệu lệnh...). Việc phân loại chính xác giúp hệ thống đưa ra cảnh báo hoặc hỗ trợ lái xe kịp thời.

Bài toán này có nhiều ứng dụng thực tiễn như:

- **Hỗ trợ lái xe an toàn (ADAS):** cung cấp cảnh báo biển báo cho tài xế trong thời gian thực.
- **Xe tự hành (Autonomous Vehicles):** là thành phần quan trọng để xe có thể tự động điều chỉnh hành vi lái.
- **Giám sát và quản lý giao thông:** hỗ trợ hệ thống camera giao thông tự động ghi nhận, thống kê và cảnh báo vi phạm.

Hiện nay, các kỹ thuật xử lý ảnh truyền thống kết hợp với học sâu (Deep Learning) đã đạt được nhiều thành tựu trong việc giải quyết bài toán này. Các mô hình mạng nơ-ron tích chập (CNN), cũng như các kiến trúc hiện đại như YOLO, Faster R-CNN hay EfficientDet, được ứng dụng để vừa phát hiện vừa phân loại biển báo với độ chính xác và tốc độ cao.

2. Phương pháp thực hiện

Quy trình được xây dựng theo ba giai đoạn chính:

(a) Tiền xử lý và xác định đối tượng (Localize)

- Dữ liệu đầu vào là ảnh màu định dạng .png cùng annotation .xml theo chuẩn Pascal VOC.
- Từ file XML, hệ thống trích xuất tọa độ bounding box (xmin, ymin, xmax, ymax).
- Ảnh biển báo được cắt (crop) ra từ ảnh gốc dựa vào bounding box.
- Trên ảnh crop, tiến hành chuyển đổi không gian màu RGB → HSV để dễ dàng phân tách các vùng màu đặc trưng (đỏ, xanh).

- Áp dụng Thresholding theo kênh màu và phép toán hình thái học (morphological operations) để loại bỏ nhiễu.
- Cuối cùng, Contour Detection được sử dụng để xác định chính xác vùng chứa biển báo, kết quả được lưu lại thành ảnh con.

(b) Trích xuất đặc trưng (Feature Extraction)

- Hai loại đặc trưng được khai thác:
 1. HOG (Histogram of Oriented Gradients): mô tả hình dạng và biên của biển báo.
 2. Histogram màu: mô tả phân bố màu sắc, hỗ trợ phân biệt các loại biển báo có màu đặc trưng.
- Ngoài ra, một phiên bản đơn giản là ảnh xám vector hoá cũng được dùng để so sánh.

(c) Phân loại (Classify)

- Thực hiện hai mô hình phân loại:
 1. Softmax Regression: đóng vai trò baseline, mô hình tuyến tính đa lớp.
 2. SVM (Support Vector Machine): mô hình nâng cao, sử dụng hàm kernel để phân tách dữ liệu phi tuyến.
- Đầu vào classifier: vector đặc trưng HOG hoặc histogram màu.
- Đầu ra classifier: nhãn lớp của biển báo (speedlimit, no_parking, turn_left, ...).

3.1 Thí nghiệm

- Tập dữ liệu: gồm ảnh giao thông thật kèm annotation Pascal VOC. Mỗi ảnh chứa ít nhất một biển báo.

- **Chia dữ liệu:** 70% training, 20% validation, 10% test.
 - **Thực hiện:**
 - Giai đoạn 1: Crop ảnh biển báo từ bounding box, tiền xử lý HSV + thresholding + contour.
 - Giai đoạn 2: Trích xuất đặc trưng HOG và histogram màu.
 - Giai đoạn 3: Huấn luyện Softmax và SVM trên tập huấn luyện.
 - **Cấu hình:**
 - Ảnh đầu vào resize về kích thước chuẩn (ví dụ 64×64).
 - Thư viện sử dụng: OpenCV (cv2), scikit-learn (svm, logistic regression), scikit-image (HOG).
-

3.2 Kết quả

- **Softmax Regression:**
 - Training accuracy ~90–95%.
 - Validation accuracy thấp (~15–20%).
 - Hiện tượng overfitting, mô hình tuyênlính không đủ mạnh để khái quát.
- **SVM:**
 - Training accuracy cao (>95%).
 - Validation accuracy cải thiện đáng kể (~40–50%).
 - SVM với đặc trưng HOG cho kết quả ổn định hơn histogram màu.

4. Phân tích Mô hình hiện tại

4.1. Phân tích Ưu điểm

- **Hiệu quả tính toán và tốc độ xử lý:** Mô hình HOG + Softmax/SVM có chi phí tính toán thấp. Việc trích xuất đặc trưng HOG và huấn luyện mô hình phân loại là các tác vụ không quá phức tạp, cho phép hệ thống chạy mượt mà ngay cả trên các thiết bị có tài nguyên hạn chế như Raspberry Pi hoặc các hệ thống nhúng (embedded systems). Điều này làm cho nó trở thành một lựa chọn lý tưởng cho các ứng dụng yêu cầu tốc độ xử lý theo thời gian thực (real-time).
- **Tính trực quan và dễ hiểu:** Mỗi thành phần trong hệ thống đều có một vai trò rõ ràng và dễ hiểu. HOG tập trung vào các đặc trưng hình dạng, trong khi Softmax/SVM thực hiện việc phân loại dựa trên các đặc trưng đó. Điều này giúp các nhà phát triển dễ dàng gỡ lỗi (debug) và tối ưu hóa từng bước một cách độc lập.
- **Tính ổn định đối với biến thể hình học:** Đặc trưng HOG được thiết kế để nắm bắt các đặc trưng hình dạng như đường viền và góc cạnh của đối tượng. Nhờ đó, mô hình có khả năng nhận diện biến báo khá tốt ngay cả khi chúng bị xoay, nghiêng hoặc thay đổi kích thước.

4.2. Phân tích Nhược điểm và Hạn chế

- **Độ bền vững (Robustness) thấp trước các điều kiện môi trường:** Đây là hạn chế lớn nhất của hệ thống này. HOG nhạy cảm với các yếu tố như:
 - Ánh sáng: Biến báo quá sáng hoặc quá tối có thể làm mất đi các gradient cần thiết, dẫn đến việc trích xuất đặc trưng không chính xác.
 - Che khuất: Nếu biến báo bị che khuất một phần (ví dụ: bởi cành cây, tuyết), HOG có thể không thu được đặc trưng toàn diện, gây ra sai lệch trong phân loại.

- **Nhiễu:** Các vùng nhiễu trên ảnh (noise) hoặc vật thể tương tự biển báo có thể làm sai lệch đặc trưng HOG.
 - **Phụ thuộc vào chất lượng trích xuất đặc trưng:** Toàn bộ hiệu suất của hệ thống phụ thuộc vào độ chính xác của bước HOG. Nếu HOG không trích xuất được đặc trưng tốt, mô hình phân loại (Softmax/SVM) dù có mạnh đến đâu cũng không thể đưa ra kết quả đúng. Hệ thống không thể tự động học các đặc trưng mới hoặc phức tạp hơn từ dữ liệu.
 - **Khả năng phân biệt thấp giữa các lớp tương đồng:** Đối với các biển báo có hình dạng rất giống nhau (ví dụ: cấm đi thẳng và cấm rẽ phải), chỉ dựa vào đặc trưng hình dạng HOG có thể không đủ để phân biệt.
-

4.3. So sánh với các mô hình Học sâu (Deep Learning)

Để có cái nhìn toàn diện hơn, cần so sánh hệ thống này với các phương pháp học sâu, đặc biệt là **Mạng nơ-ron tích chập** (Convolutional Neural Network - CNN).

- **Khả năng tự động học đặc trưng:** CNNs không yêu cầu trích xuất đặc trưng thủ công. Thay vào đó, chúng tự động học và trích xuất các đặc trưng đa cấp (từ các đường nét cơ bản đến các đặc trưng phức tạp như cấu trúc toàn bộ biển báo) qua các lớp tích chập. Điều này giúp CNNs bền vững hơn nhiều trước các điều kiện môi trường bất lợi.
- **Độ chính xác:** CNNs thường đạt được độ chính xác cao hơn đáng kể, vượt qua ngưỡng của các mô hình học máy truyền thống trên các bộ dữ liệu phức tạp.
- **Yêu cầu tài nguyên:** Nhược điểm chính của CNNs là chúng đòi hỏi phần cứng mạnh (GPU) và lượng dữ liệu lớn để huấn luyện hiệu quả. Quá trình huấn luyện cũng tốn nhiều thời gian hơn.

5. Kết luận và Hướng phát triển

Hệ thống nhận diện biển báo dựa trên HOG + Softmax/SVM là một công cụ học thuật tuyệt vời để hiểu các nguyên lý cơ bản của xử lý ảnh và học máy. Nó cung cấp một giải pháp nhanh chóng và hiệu quả cho các bài toán đơn giản hoặc các ứng dụng có giới hạn về tài nguyên.

Tuy nhiên, để tạo ra một hệ thống có khả năng hoạt động ổn định và chính xác trong các tình huống giao thông thực tế, cần phải vượt qua các hạn chế của phương pháp cổ điển. Hướng phát triển chiến lược được đề xuất là chuyển sang sử dụng các mô hình học sâu.

5.1. Hướng Phát triển Ngắn hạn

- **Kết hợp đặc trưng:** Cải thiện mô hình hiện tại bằng cách kết hợp đặc trưng HOG với các đặc trưng bổ sung khác như Histogram màu (Color Histogram). Màu sắc là một thông tin quan trọng để phân biệt biển báo giao thông và có thể giúp tăng độ chính xác đáng kể.
- **Tăng cường dữ liệu (Data Augmentation):** Tạo ra các biến thể mới của dữ liệu huấn luyện bằng cách áp dụng các phép biến đổi như xoay, lật, thay đổi độ sáng, độ tương phản và thêm nhiễu. Điều này sẽ giúp mô hình hiện tại trở nên bền vững hơn.

5.2. Hướng Phát triển Dài hạn

- **Triển khai CNN cho bài toán Phân loại:** Thay thế toàn bộ pipeline HOG + Softmax/SVM bằng một mạng CNN đơn giản như LeNet hoặc VGG. Điều này sẽ loại bỏ bước trích

xuất đặc trưng thủ công và cho phép mô hình tự học các đặc trưng hiệu quả hơn.

- **Chuyển sang bài toán Phát hiện Đối tượng (Object Detection):** Nâng cấp hệ thống từ việc chỉ phân loại một biển báo đã được crop sẵn thành khả năng tự động tìm kiếm và phân loại tất cả các biển báo trong một khung hình. Các mô hình như YOLO (You Only Look Once), Faster R-CNN, hoặc SSD là các lựa chọn hàng đầu cho tác vụ này. Một mô hình YOLO sẽ không chỉ cho bạn biết biển báo là gì mà còn cung cấp tọa độ chính xác của nó trên ảnh.