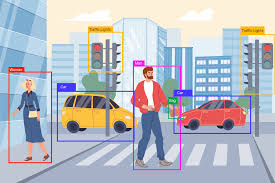
**Báo cáo AI/Machine Learning tuần 5**

1. **Bài toán object detection và image segmentation.**

* Object detection là bài toán xác định và phân loại các vật thể xuất hiện có trong một bức ảnh.
* Thuật toán object detection có đầu vào là 1 ảnh, đầu ra là các hộp bao quanh (bounding box) xung quanh mỗi đối tượng trong ảnh, đồng thời gán cho nó một nhãn (label) tương ứng với loại đối tượng đó.



* Ứng dụng của bài toán object detection:
* **Xe tự lái:** Phát hiện các phương tiện khác, người đi bộ, biển báo giao thông,.. qua đó đưa ra xử lý tình huống giao thông ngay lập tức.
* **Giám sát an ninh:** Phát hiện người lạ xâm nhập, đối tượng khả nghi từ đó đưa ra cảnh báo.
* **Tìm kiếm hình ảnh:** Tìm kiếm các hình ảnh chứa đối tượng cụ thể.
* Image Segmentation (phân đoạn vật thể trong ảnh) là một nhiệm vụ phức tạp hơn, ở đây là chia hình ảnh thành các vùng đồng nhất theo những tiêu chí nhất định như màu sắc, kết cấu, cùng đối tượng, mục tiêu là tạo ra 1 bản đồ phân đoạn, trong đó mỗi pixel được gán nhãn tương ứng với vùng mà nó thuộc về.

A person with her tongue out

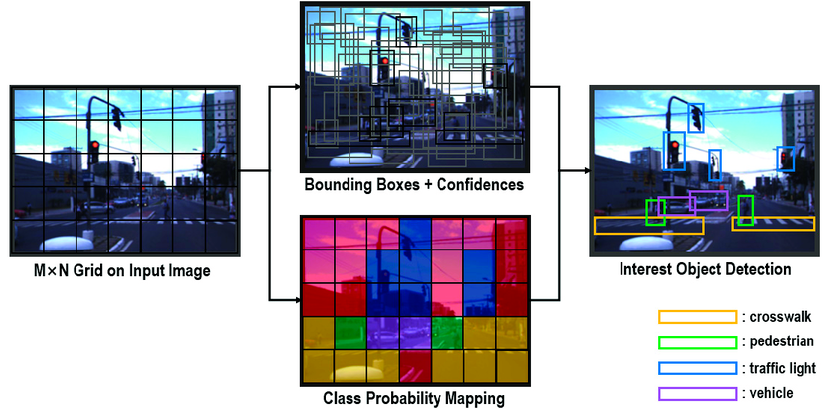
Description automatically generated

* Ứng dụng:
* Xử lý ảnh y tế: Phân đoạn các cơ quan trong cơ thể từ hình ảnh X-quang, MRI.
* Chỉnh sửa ảnh phức tạp: Tách đối tượng ra khỏi nền, lựa chọn đối tượng từ ảnh.
* So sánh Object Detection và Image Segmentation.

|  |  |
| --- | --- |
| Object Detection | Image Segmentation |
| Xác định đối tượng, phân loại vị trí của đối tượng. | Chia ảnh thành các vùng đồng nhất |
| Bounding box và labels của từng vật thể xuất hiện trong ảnh | Bản đồ phân đoạn bao quanh từng vật thể trong ảnh |
| Độ phức tạp thấp hơn | Độ phức tạp cao hơn |

1. **Model thường sử dụng và phân tích.**

* Ta có nhiều thuật toán để giải quyết 2 bài toán kể trên, tuy nhiên ta sẽ đề cập đến mô hình thông dụng có thể giải quyết được cả 2 bài toán kể trên là YOLO.
* YOLO (You Only Look Once) là một thuật toán phát hiện đối tượng nổi tiếng, được thiết kế để thực hiện phát hiện đối tượng trong thời gian thực. Tuy nhiên, YOLO ban đầu chủ yếu tập trung vào bài toán object detection, nghĩa là xác định vị trí và phân loại các đối tượng trong một hình ảnh bằng cách vẽ các bounding box xung quanh chúng.
* Cách YOLO giải quyết bài toán nhận diện đồ vật:
* Ảnh đầu vào được truyền qua 1 CNN để trích xuất các đặc trưng thuộc tính có trong ảnh.
* Các đặc trưng có được sẽ được truyền qua 1 lớp fully-connected, dự đoán thuộc tính đó thuộc về lớp đối tượng nào và tọa độ của bounding box tương ứng với đối tượng đó.
* Chia để trị: Chia ảnh ra thành lưới các ô có diện tích bằng nhau.
* Trong mỗi ô, hệ thống dự đoán xác suất mà mỗi ô đó thuộc về bounding box của đối tượng nào, từ đó cho ra xác suất mà ô thuộc đối tượng nào.
* Đầu ra của hệ thống là tập hợp các bounding boxes và xác suất thuộc về từng lớp của mỗi ô.
* Các bounding boxes sau đó được lọc thông qua thuật toán hậu xử lý non-max suppression để chọn ra box với xác suất cao nhất
* Đầu ra cuối cùng là tập hợp dự đoán bằng các bounding boxes, và nhãn các lớp cùng xác suất cho từng đối tượng xuất hiện trên ảnh.



* YOLO ban đầu không được thiết kế để thực hiện image segmentation, tức là chia một hình ảnh thành các vùng đồng nhất tương ứng với các đối tượng khác nhau. Tuy nhiên, với một số điều chỉnh và kết hợp với các kỹ thuật khác, YOLO có thể được sử dụng để giải quyết bài toán này. Phiên bản mới nhất của YOLO, YOLOv8, đã giới thiệu một mô hình chuyên biệt cho bài toán image segmentation. Nó thực hiện cả ba nhiệm vụ trong một lần dự đoán duy nhất:
* **Phát hiện đối tượng:** Xác định vị trí của các đối tượng bằng cách tạo ra các bounding box.
* **Phân loại đối tượng:** Xác định nhãn của từng đối tượng được phát hiện.
* **Phân đoạn đối tượng:** Tạo ra các mask (mặt nạ) để xác định chính xác hình dạng của từng đối tượng.
* Cách YOLO hoạt động trong bài toán image segmentation:

1. **Cách train model và các thông số đánh giá độ hiệu quả của 1 model.**
2. **Background Subtraction.**