



# Xác định vị trí đối tượng (Object Localization)

# Phát biểu bài toán

- Input:

1. Tập dữ liệu gồm có  $N$  phần tử đã được gán nhãn:  $\mathcal{D} = \{(\mathbf{x}_i, y_i, \mathbf{b}_i)\}_{i=1}^N$ .

- Tập các nhãn của  $\mathcal{D}$  là:  $\mathcal{L} = \bigcup_{i=1}^N y_i$
- Trong mỗi phần tử của  $\mathcal{D}$  :
  - $\mathbf{x}_i$  là ảnh số  $\mathcal{I}$
  - $y_i$  là giá trị nhãn (label) được gán cho  $\mathbf{x}_i$
  - $\mathbf{b}_i$  là hộp giới hạn bao quanh đối tượng (bounding box) có nhãn là  $y_i$  trong  $\mathbf{x}_i$

2. Ảnh  $\mathbf{x} \in \mathcal{I}$

- Output:

- $\hat{y} \in \mathcal{L}$ : giá trị nhãn dự đoán của ảnh  $\mathbf{x}$
- $\hat{\mathbf{b}}$  : hộp giới hạn bao quanh đối tượng  $\hat{y}$  trong ảnh  $\mathbf{x}$



# Giới thiệu

---

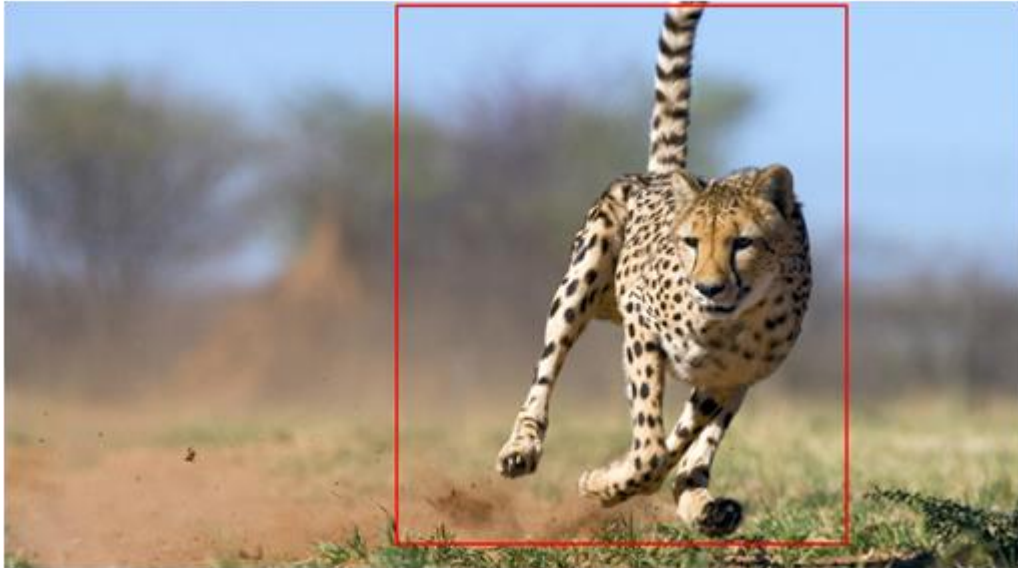
- Định nghĩa Object Localization
- Phân biệt Object Detection và Object Localization
- Tầm quan trọng và ứng dụng của Object Localization



# Định nghĩa Object Localization

- **Object Localization** là một bài toán trong thị giác máy tính, trong đó mục tiêu là **xác định vị trí của một đối tượng** trong ảnh hoặc video và đánh dấu nó bằng cách sử dụng một **hộp giới hạn (bounding box)** bao quanh đối tượng đó.
- Nhiệm vụ của Object Localization không chỉ dừng lại ở việc **phân loại đối tượng** (ví dụ như xác định đối tượng là một con mèo hay một chiếc ô tô) mà còn yêu cầu **xác định tọa độ vị trí của đối tượng** đó trong ảnh

# Định nghĩa Object Localization





# Định nghĩa Object Localization

---

## Các yếu tố chính của Object Localization:

- **Phân loại đối tượng:** Xác định đối tượng đó thuộc loại nào (ví dụ: người, xe, động vật).
- **Xác định vị trí:** Xác định vị trí chính xác của đối tượng bằng cách sử dụng các tọa độ của hộp giới hạn (bounding box) xung quanh đối tượng.



# Định nghĩa Object Localization

---

Object Localization là bước cơ bản trong các bài toán phức tạp hơn như **Object Detection** (phát hiện nhiều đối tượng trong ảnh) hoặc **Instance Segmentation** (phân vùng từng đối tượng cụ thể).

# Phân biệt Object Detection và Object Localization



## Object Localization

- **Mục tiêu:** Xác định vị trí của một đối tượng duy nhất trong ảnh, kết hợp với phân loại đối tượng.
- **Đầu ra:** Một **bounding box** (hộp giới hạn) quanh đối tượng cùng với nhãn phân loại.
- **Giới hạn:** Chỉ xử lý một đối tượng duy nhất trong ảnh. Nếu có nhiều đối tượng, Object Localization sẽ chỉ xác định và phân loại đối tượng quan trọng hoặc lớn nhất trong ảnh.



# Phân biệt Object Detection và Object Localization



## Object Detection

- **Mục tiêu:** Xác định và phân loại **nhiều đối tượng** trong cùng một ảnh.
- **Đầu ra:** Nhiều bounding box và nhãn phân loại cho từng đối tượng.
- **Ứng dụng:** Các ứng dụng thực tiễn như giám sát an ninh, xe tự hành và phân tích video yêu cầu Object Detection do khả năng phát hiện và phân loại nhiều đối tượng cùng lúc.

# Phân biệt Object Detection và Object Localization



Yếu tố	Object Localization	Object Detection
Số lượng đối tượng	Một đối tượng	Nhiều đối tượng
Đầu ra	Bounding box và nhãn phân loại cho 1 đối tượng	Bounding box và nhãn phân loại cho nhiều đối tượng
Độ phức tạp	Thấp hơn	Cao hơn, yêu cầu thuật toán phức tạp hơn
Ứng dụng	Định vị đối tượng đơn lẻ, phân loại ảnh	Phát hiện đối tượng trong video, giám sát, xe tự hành



## Localization

---

- **Tầm quan trọng của Object Localization** nằm ở khả năng xác định vị trí và phân loại các đối tượng trong ảnh, cung cấp thông tin không chỉ về loại đối tượng mà còn cả vị trí không gian của chúng. Điều này giúp các hệ thống thị giác máy tính hiểu rõ hơn về bố cục và ngữ cảnh trong một hình ảnh hay video. Object Localization là nền tảng cho các bài toán phức tạp hơn như Object Detection và Instance Segmentation.



# Ứng dụng của Object Localization

---

- **Xe tự hành (Autonomous Vehicles):**
  - Giúp nhận diện vị trí của các phương tiện khác, người đi bộ, tín hiệu giao thông, và các chướng ngại vật.
  - Nâng cao an toàn và khả năng xử lý tình huống trong môi trường phức tạp.



# Ứng dụng của Object Localization

---

- **Giám sát an ninh (Security Surveillance):**

- Định vị và nhận diện các đối tượng nghi vấn trong các hệ thống giám sát video.
- Ứng dụng trong phát hiện hành vi khả nghi hoặc đám đông tại các địa điểm công cộng.



# Ứng dụng của Object Localization

---

- **Y tế và chẩn đoán hình ảnh (Medical Imaging):**
  - Phát hiện và xác định vị trí các khối u hoặc tổn thương trong ảnh chụp X-quang, MRI, CT scan.
  - Giúp bác sĩ phát hiện sớm các bệnh lý và có phương án điều trị kịp thời.



# Ứng dụng của Object Localization

---

- **Thương mại điện tử và quảng cáo (E-commerce and Advertising):**
  - Tự động nhận diện sản phẩm trong ảnh hoặc video để phục vụ cho các hệ thống gợi ý sản phẩm.
  - Phân tích hành vi của người dùng bằng cách xác định đối tượng người dùng chú ý đến trong các quảng cáo.



# Ứng dụng của Object Localization

---

- **Phân tích ảnh vệ tinh và quan sát địa lý (Satellite Imaging and Geographic Observation):**
  - Phát hiện và định vị các công trình, rừng, sông suối và các vùng ảnh hưởng do thiên tai.
  - Hỗ trợ các nhiệm vụ như theo dõi sự thay đổi địa lý và quản lý tài nguyên thiên nhiên.





# Ứng dụng của Object Localization

---

- **Thực tế tăng cường (Augmented Reality - AR):**
- Định vị các đối tượng trong môi trường thực để bổ sung các yếu tố ảo, giúp người dùng tương tác tốt hơn với không gian xung quanh.
- Ứng dụng trong ngành bán lẻ và giáo dục để nâng cao trải nghiệm người dùng.



# Nguyên lý cơ bản

---

- Cách Object Localization hoạt động
- Tổng quan về bounding box và cách xác định tọa độ
- Vai trò của các phép đo lường (IoU, Recall, Precision) trong Localization



# Cách Object Localization hoạt động

---

- Object Localization hoạt động bằng cách kết hợp hai nhiệm vụ: **phân loại đối tượng** và **xác định vị trí** của nó trong một hình ảnh.
- Dưới đây là các bước cơ bản về cách hoạt động của Object Localization:



# Cách Object Localization hoạt động

---

## 1. Tiền xử lý ảnh đầu vào

- Ảnh đầu vào có thể có kích thước và định dạng khác nhau, nên cần được tiền xử lý để đồng nhất kích thước, chuẩn hóa giá trị pixel và có thể áp dụng các kỹ thuật tăng cường dữ liệu (Data Augmentation) để làm phong phú dữ liệu huấn luyện.



# Cách Object Localization hoạt động

---

## 2. Trích xuất đặc trưng (Feature Extraction)

- Ảnh được đưa vào một mô hình mạng nơ-ron tích chập (CNN) để trích xuất các đặc trưng. CNN hoạt động bằng cách nhận diện các đặc trưng không gian trong ảnh, ví dụ như cạnh, đường viền, và các cấu trúc đặc trưng của đối tượng.
- Mỗi lớp convolutional trích xuất một tập hợp các đặc trưng, từ những chi tiết đơn giản (như cạnh và đường) đến các đặc trưng phức tạp hơn ở các lớp sâu hơn (như hình dáng tổng thể của đối tượng).



# Cách Object Localization hoạt động

## 3. Xác định Bounding Box

- **Bounding box** là một hình chữ nhật bao quanh đối tượng cần xác định. Để xác định bounding box, mạng nơ-ron sẽ dự đoán các tọa độ:  $(x,y)$   $(x, y)$   $(x,y)$  cho góc trên bên trái, chiều rộng ( $www$ ) và chiều cao ( $hhh$ ) của hộp.
- Mạng có thể sử dụng một lớp fully connected ở đầu ra để dự đoán các giá trị tọa độ bounding box cùng với nhãn phân loại.



# Cách Object Localization hoạt động

## 4. Phân loại đối tượng (Object Classification)

- Ngoài tọa độ bounding box, mạng cũng phải phân loại đối tượng trong bounding box đó thành các nhãn lớp cụ thể (ví dụ: chó, mèo, xe hơi).
- Thường thì mạng sẽ dự đoán một lớp duy nhất cho bounding box hoặc xác suất thuộc về các lớp khác nhau. Nếu đối tượng trong ảnh thuộc nhiều lớp (multi-label), mạng có thể dự đoán một vector nhãn cho từng bounding box.



# Cách Object Localization hoạt động

## 5. Tính toán hàm mất mát (Loss Calculation)

- Mô hình cần được tối ưu hóa để đưa ra dự đoán chính xác nhất. Đối với Object Localization, hàm mất mát sẽ bao gồm hai phần:
  - **Hàm mất mát phân loại (Classification Loss):** Đánh giá độ chính xác của nhãn đối tượng.
  - **Hàm mất mát định vị (Localization Loss):** Đo độ chênh lệch giữa bounding box dự đoán và bounding box thực.
- **Loss tổng hợp** = Classification Loss + Localization Loss, thường sử dụng IoU (Intersection over Union) để đánh giá mức độ chính xác của bounding box.





# Cách Object Localization hoạt động

---

## 6. Đánh giá và dự đoán đầu ra

- Sau khi hoàn thành huấn luyện, mô hình sẽ dự đoán bounding box và nhãn phân loại cho mỗi ảnh đầu vào.
- Trong quá trình kiểm tra hoặc sử dụng, mô hình sẽ tạo ra bounding box cùng với nhãn phân loại tương ứng cho đối tượng duy nhất trong ảnh.

# Tổng quan về bounding box và cách xác định tọa độ



- **Bounding box** là một hình chữ nhật được vẽ bao quanh đối tượng trong một hình ảnh, nhằm đánh dấu và xác định vị trí của đối tượng đó.
- Đây là thành phần quan trọng trong các bài toán **Object Localization** và **Object Detection**, vì nó cho phép chúng ta không chỉ phân loại đối tượng mà còn xác định chính xác vị trí không gian của nó.



# Cấu trúc của Bounding Box

---



# Vai trò của các phép đo lường (IoU, Recall, Precision) trong Localization

---

- Trong **Object Localization**, các phép đo lường như **IoU (Intersection over Union)**, **Recall**, và **Precision** đóng vai trò quan trọng trong việc đánh giá độ chính xác và hiệu quả của mô hình. Chúng giúp xác định xem mô hình có định vị đối tượng đúng và phân loại chính xác hay không.

# Các kỹ thuật Object Localization truyền thống

---



- Kỹ thuật Sliding Windows
- Selective Search và các phương pháp dựa trên vùng
- Hạn chế của các phương pháp truyền thống



# Mạng CNN và Object Localization

---

- Giới thiệu CNN trong nhận dạng và định vị vật thể
- Sử dụng lớp convolutional để trích xuất đặc trưng
- Các kiến trúc cơ bản: YOLO, SSD, và R-CNN (chỉ giới thiệu tổng quan)



# Kiến trúc hiện đại cho Object Localization

---

- Chi tiết về R-CNN và các phiên bản cải tiến (Fast R-CNN, Faster R-CNN)
- Mô hình YOLO: cách thức làm việc và ứng dụng
- SSD (Single Shot MultiBox Detector): điểm mạnh và điểm yếu



# Huấn luyện mô hình Object Localization

---

- Yêu cầu dữ liệu và gán nhãn cho Object Localization
- Kỹ thuật Data Augmentation cho dữ liệu ảnh
- Các kỹ thuật tối ưu hóa cho mô hình Localization





# Đánh giá mô hình và chỉ số hiệu quả

---

- Đo lường IoU (Intersection over Union)
- Mean Average Precision (mAP) trong đánh giá mô hình
- Confusion Matrix và các chỉ số bổ sung



# Ứng dụng thực tiễn và ví dụ

---

- Object Localization trong xe tự hành
- Ứng dụng trong y tế và giám sát
- Phân tích ảnh vệ tinh và các ví dụ khác



# Kết luận và thảo luận

---

- Tóm tắt các khái niệm chính
- Xu hướng và tương lai của Object Localization
- Câu hỏi thảo luận và phần hỏi đáp