

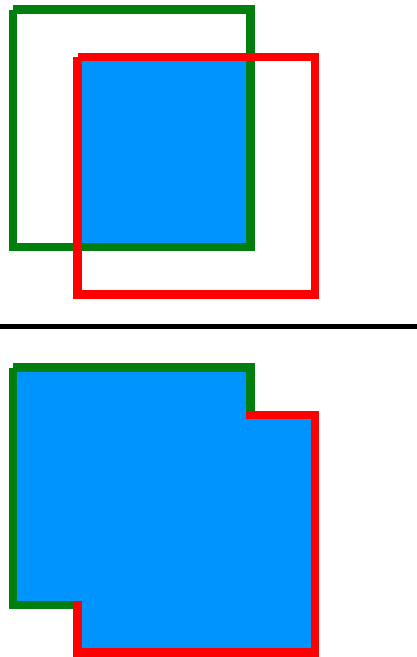
# Metrics trong bài toán object detection

## 1. Một số khái niệm quan trọng

### 1.1. Intersection Over Union (IoU)

Intersection Over Union (gọi tắt là IoU) là chỉ số được tính toán dựa trên Jaccard Index nhằm đánh giá độ trùng nhau giữa 2 bbox trong bài toán object detection (bbox groundtruth và bbox prediction).

IoU được tính bằng thương của diện tích giao của hai bbox và diện tích hợp của hai bbox.

$$IOU = \frac{\text{area of overlap}}{\text{area of union}} = \frac{\text{area of intersection}}{\text{area of union}}$$


### 1.2. TP, FP, FN và TN

Từ khái niệm về IoU, ta có các khái niệm:

- True Positive (TP): Bbox prediction được gọi là TP nếu IoU giữa bbox groundtruth và bbox prediction  $\geq$  ngưỡng IoU (IoU threshold).
- False Positive (FP): Bbox prediction được gọi là FP nếu IoU giữa bbox groundtruth và bbox prediction  $<$  ngưỡng IoU (IoU threshold).
- False Negative (FN): Là các bbox groundtruth không được dự đoán
- True Negative (TN): Không được sử dụng trong bài toán object detection

Đối với các bài toán object detection có nhiều class hơn 2 (class object và class background), bbox prediction được gọi là TP không những phải thoả mãn yêu cầu về IoU mà còn phải thoả mãn yêu cầu về class dự đoán phải đúng.



Trong quá trình đánh giá mô hình object detection, ta thường sử dụng ba ngưỡng là 0.5, 0.75 và 0.9.

### 1.3. Precision và Recall

Tương tự như việc đánh giá mô hình classification, precision và recall cũng được sử dụng để đánh giá mô hình object detection.

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

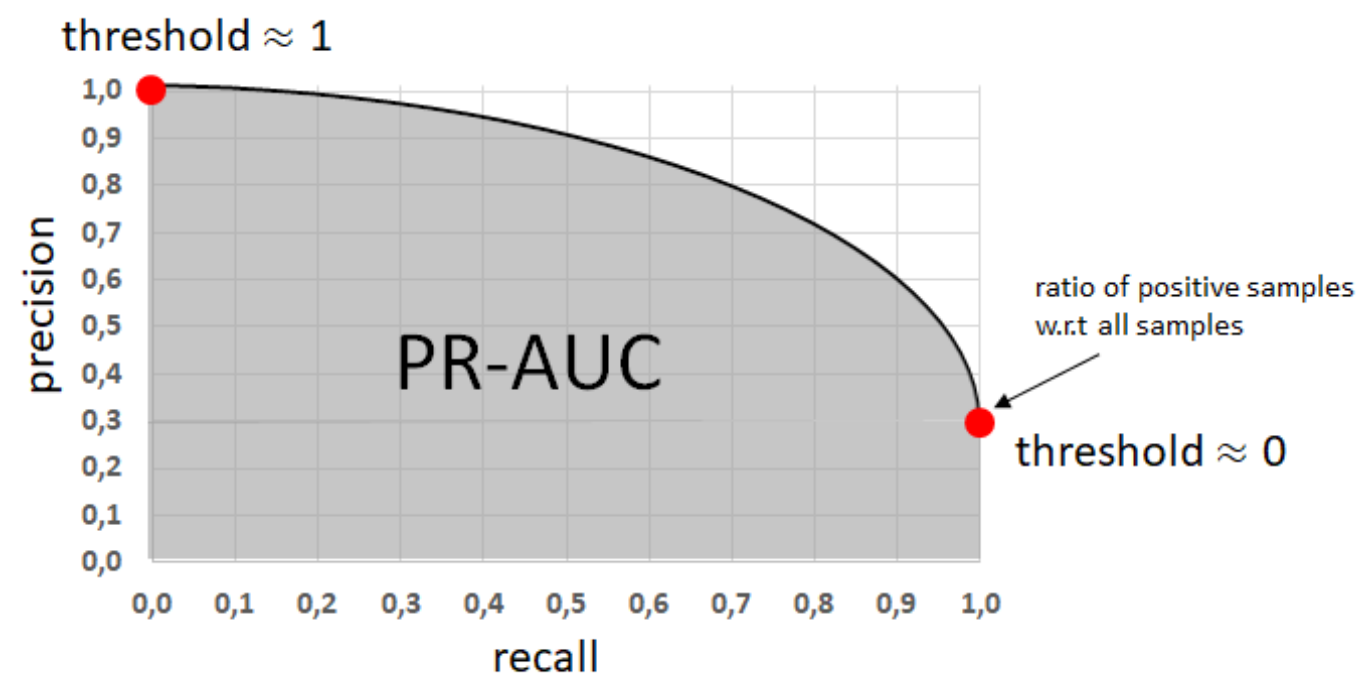
$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

## 2. Các metrics phổ biến trong bài toán object detection

---

### 2.1. Precision x Recall curve

Precision x Recall curve là một biểu đồ giúp đánh giá tương quan về đánh đổi giữa precision và recall.



## 2.2. Average Precision (AP) và Mean Average Precision (mAP)

Average Precision (gọi tắt là AP) là metrics phổ biến nhất dùng đánh giá mô hình object detection.

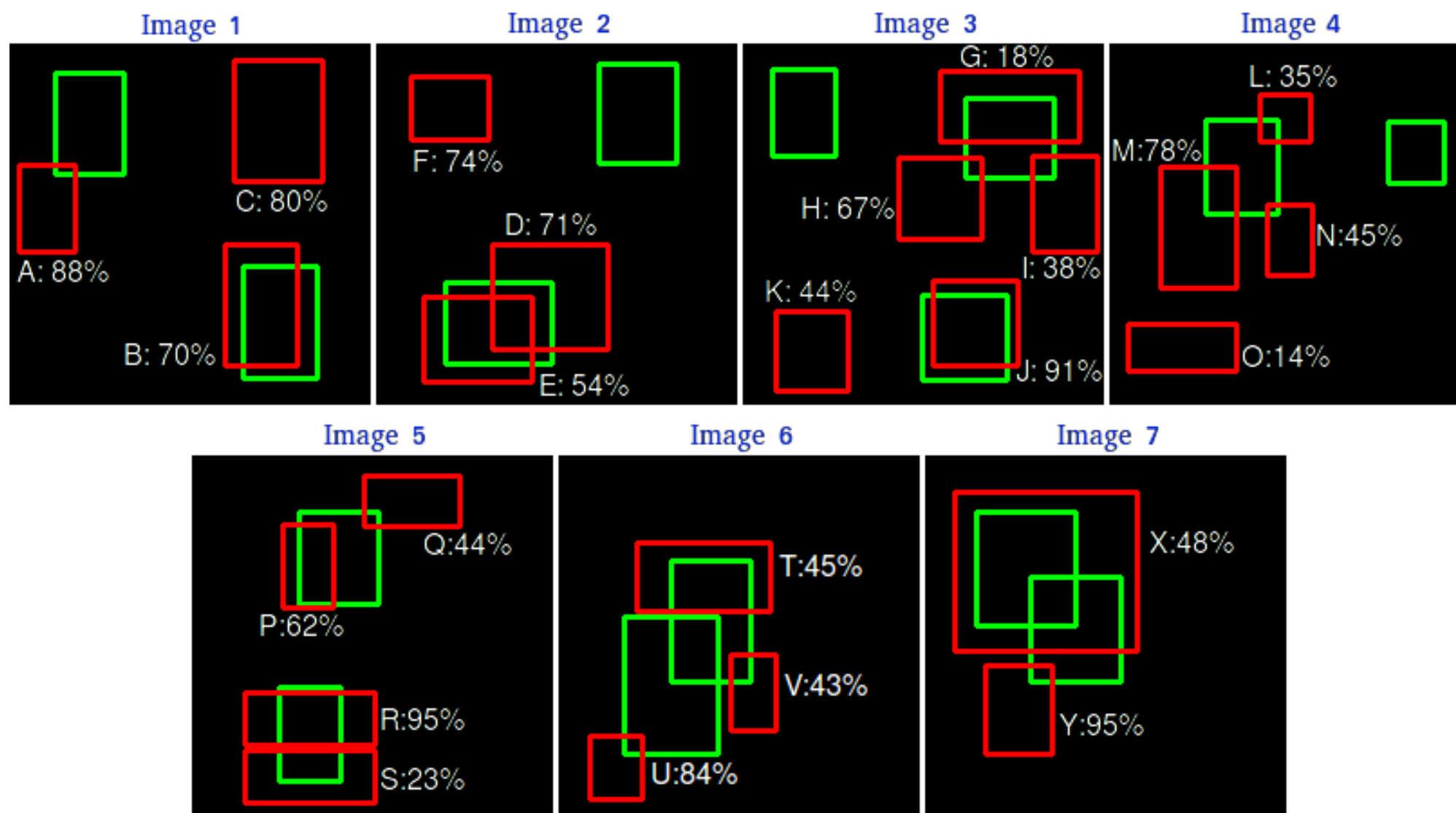
AP được tính bằng chỉ số Area under Curve (AuC) của Precision x Recall curve. Thay vì quan sát hai Precision x Recall curve xem đường nào tốt hơn thì ta tính toán ra một chỉ số để dễ dàng đánh giá được.

Trong thực tế, tên gọi Average Precision có nghĩa là Precision được tính Average trên khoảng Recall từ 0 đến 1.

Ngoài ra, đối với những mô hình object detection với nhiều class object khác nhau, ta có chỉ số Mean Average Precision (mAP) được tính bằng trung bình chỉ số AP trên tất cả các class object.

## 2.3. Ví dụ minh họa việc tính toán AP

Ta có các bbox prediction màu đỏ (được ký hiệu các ký tự A, B, C ... Y) với các chỉ số confidence bên cạnh, các bbox groundtruth màu xanh lá.



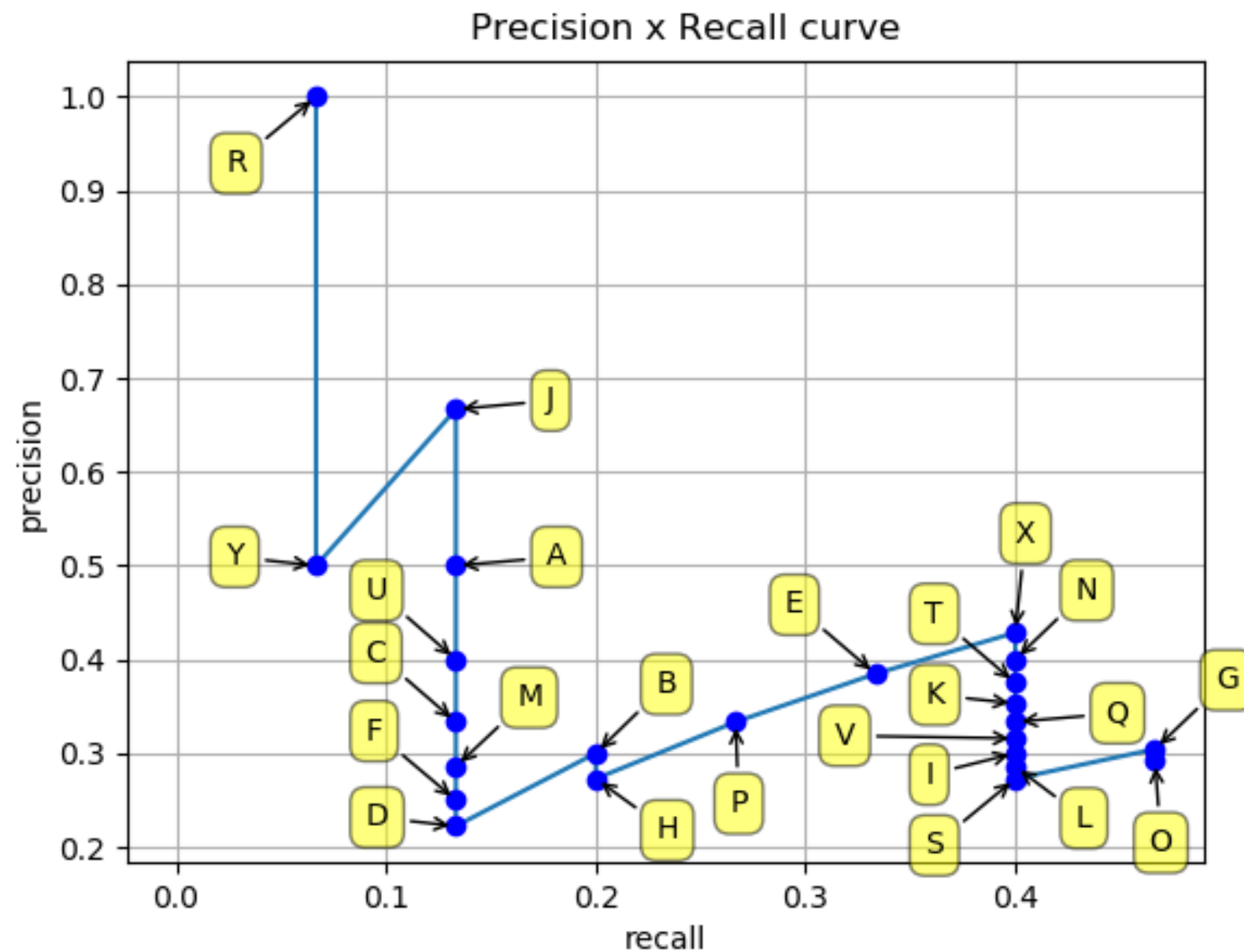
Với việc lấy  $\text{IoU} \geq 0.3$ , ta có từng bbox prediction là các bbox TP hay FN như bảng dưới

| Images  | Detections | Confidences | TP or FP |
|---------|------------|-------------|----------|
| Image 1 | A          | 88%         | FP       |
| Image 1 | B          | 70%         | TP       |
| Image 1 | C          | 80%         | FP       |
| Image 2 | D          | 71%         | FP       |
| Image 2 | E          | 54%         | TP       |
| Image 2 | F          | 74%         | FP       |
| Image 3 | G          | 18%         | TP       |
| Image 3 | H          | 67%         | FP       |
| Image 3 | I          | 38%         | FP       |
| Image 3 | J          | 91%         | TP       |
| Image 3 | K          | 44%         | FP       |
| Image 4 | L          | 35%         | FP       |
| Image 4 | M          | 78%         | FP       |
| Image 4 | N          | 45%         | FP       |
| Image 4 | O          | 14%         | FP       |
| Image 5 | P          | 62%         | TP       |
| Image 5 | Q          | 44%         | FP       |
| Image 5 | R          | 95%         | TP       |
| Image 5 | S          | 23%         | FP       |
| Image 6 | T          | 45%         | FP       |
| Image 6 | U          | 84%         | FP       |
| Image 6 | V          | 43%         | FP       |
| Image 7 | X          | 48%         | TP       |
| Image 7 | Y          | 95%         | FP       |

Sau khi xác định các bbox prediction là TP hay FN, ta sắp xếp chúng theo thứ tự về confidence để tính toán Precision và Recall.

| Images  | Detections | Confidences | TP | FP | Acc TP | Acc FP | Precision | Recall |
|---------|------------|-------------|----|----|--------|--------|-----------|--------|
| Image 5 | R          | 95%         | 1  | 0  | 1      | 0      | 1         | 0.0666 |
| Image 7 | Y          | 95%         | 0  | 1  | 1      | 1      | 0.5       | 0.0666 |
| Image 3 | J          | 91%         | 1  | 0  | 2      | 1      | 0.6666    | 0.1333 |
| Image 1 | A          | 88%         | 0  | 1  | 2      | 2      | 0.5       | 0.1333 |
| Image 6 | U          | 84%         | 0  | 1  | 2      | 3      | 0.4       | 0.1333 |
| Image 1 | C          | 80%         | 0  | 1  | 2      | 4      | 0.3333    | 0.1333 |
| Image 4 | M          | 78%         | 0  | 1  | 2      | 5      | 0.2857    | 0.1333 |
| Image 2 | F          | 74%         | 0  | 1  | 2      | 6      | 0.25      | 0.1333 |
| Image 2 | D          | 71%         | 0  | 1  | 2      | 7      | 0.2222    | 0.1333 |
| Image 1 | B          | 70%         | 1  | 0  | 3      | 7      | 0.3       | 0.2    |
| Image 3 | H          | 67%         | 0  | 1  | 3      | 8      | 0.2727    | 0.2    |
| Image 5 | P          | 62%         | 1  | 0  | 4      | 8      | 0.3333    | 0.2666 |
| Image 2 | E          | 54%         | 1  | 0  | 5      | 8      | 0.3846    | 0.3333 |
| Image 7 | X          | 48%         | 1  | 0  | 6      | 8      | 0.4285    | 0.4    |
| Image 4 | N          | 45%         | 0  | 1  | 6      | 9      | 0.4       | 0.4    |
| Image 6 | T          | 45%         | 0  | 1  | 6      | 10     | 0.375     | 0.4    |
| Image 3 | K          | 44%         | 0  | 1  | 6      | 11     | 0.3529    | 0.4    |
| Image 5 | Q          | 44%         | 0  | 1  | 6      | 12     | 0.3333    | 0.4    |
| Image 6 | V          | 43%         | 0  | 1  | 6      | 13     | 0.3157    | 0.4    |
| Image 3 | I          | 38%         | 0  | 1  | 6      | 14     | 0.3       | 0.4    |
| Image 4 | L          | 35%         | 0  | 1  | 6      | 15     | 0.2857    | 0.4    |
| Image 5 | S          | 23%         | 0  | 1  | 6      | 16     | 0.2727    | 0.4    |
| Image 3 | G          | 18%         | 1  | 0  | 7      | 16     | 0.3043    | 0.4666 |
| Image 4 | O          | 14%         | 0  | 1  | 7      | 17     | 0.2857    | 0.4666 |

Với mỗi cặp giá trị Precision và Recall vừa tính toán được, ta thu được Precision x Recall curve

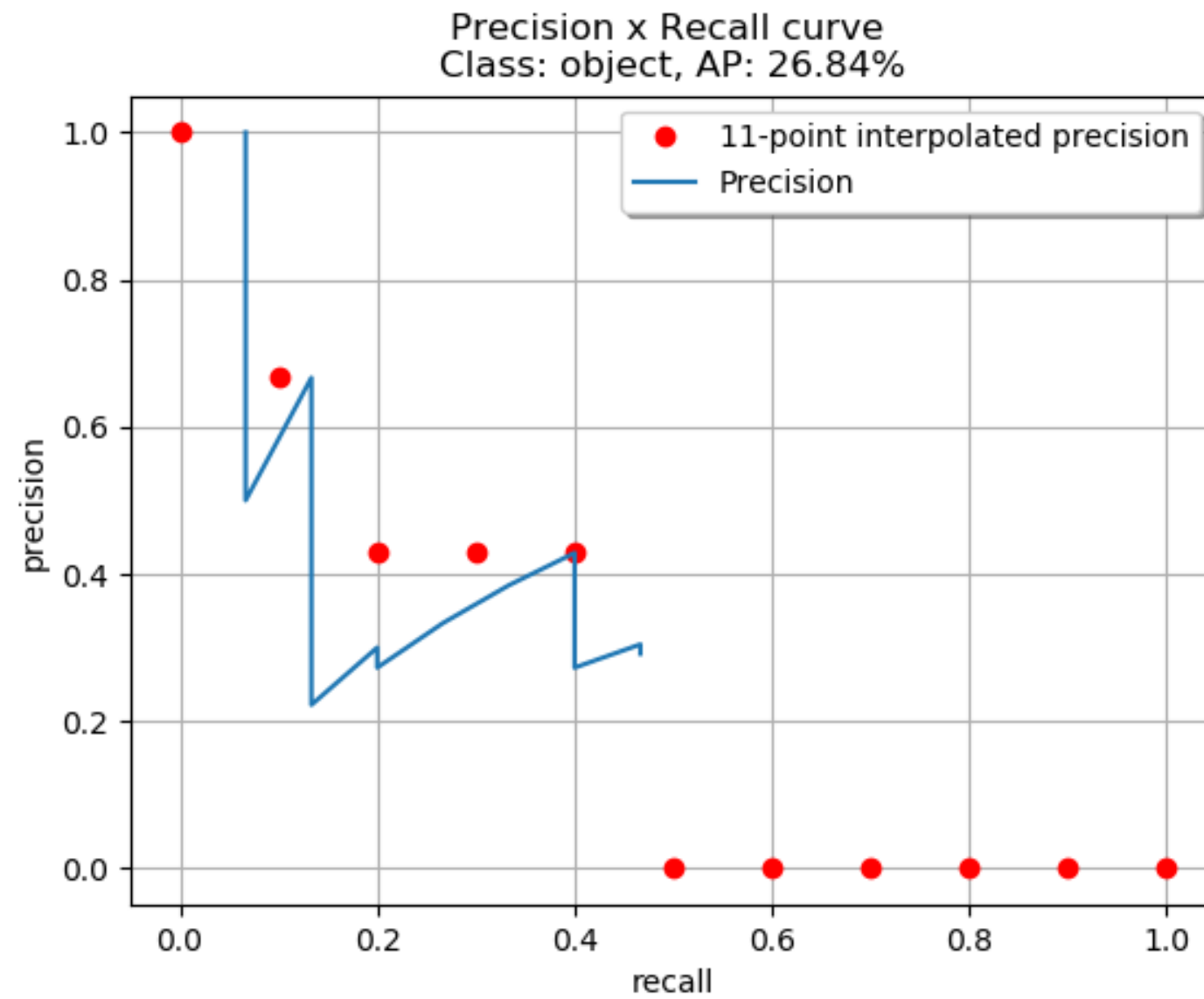


Tiếp theo, ta xấp xỉ đường Precision trên bằng một đường gọi là Interpolated Precision. Đến đây, ta có hai cách để tính đường interpolated precision.

#### 11-point interpolation

Cách đầu tiên được gọi là 11-point interpolation, ta chia trục recall của Precision x Recall curve thành 10 phần với 11 mốc recall (0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1). Với mỗi mốc recall, ta lấy chỉ số precision cao nhất mà nhận chỉ số recall lớn hơn hoặc bằng mốc recall đang xét.





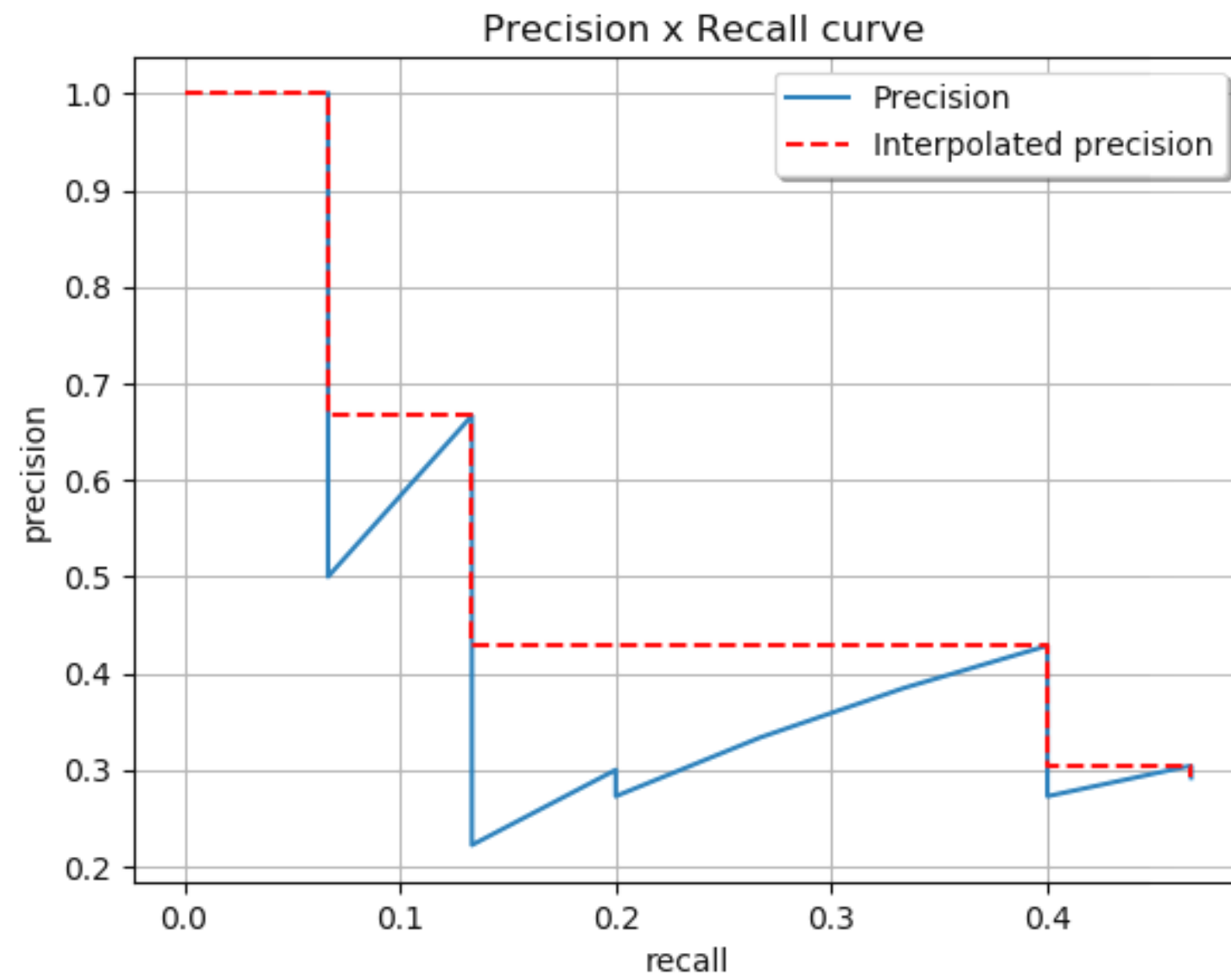
Cuối cùng, ta tính AP bằng cách lấy trung bình cộng của các giá trị precision lấy được ở bước trên.

$$AP = \frac{1}{11} \sum_{r \in \{0, 0.1, \dots, 0.9, 1\}} p_{\text{interp}}(r)$$

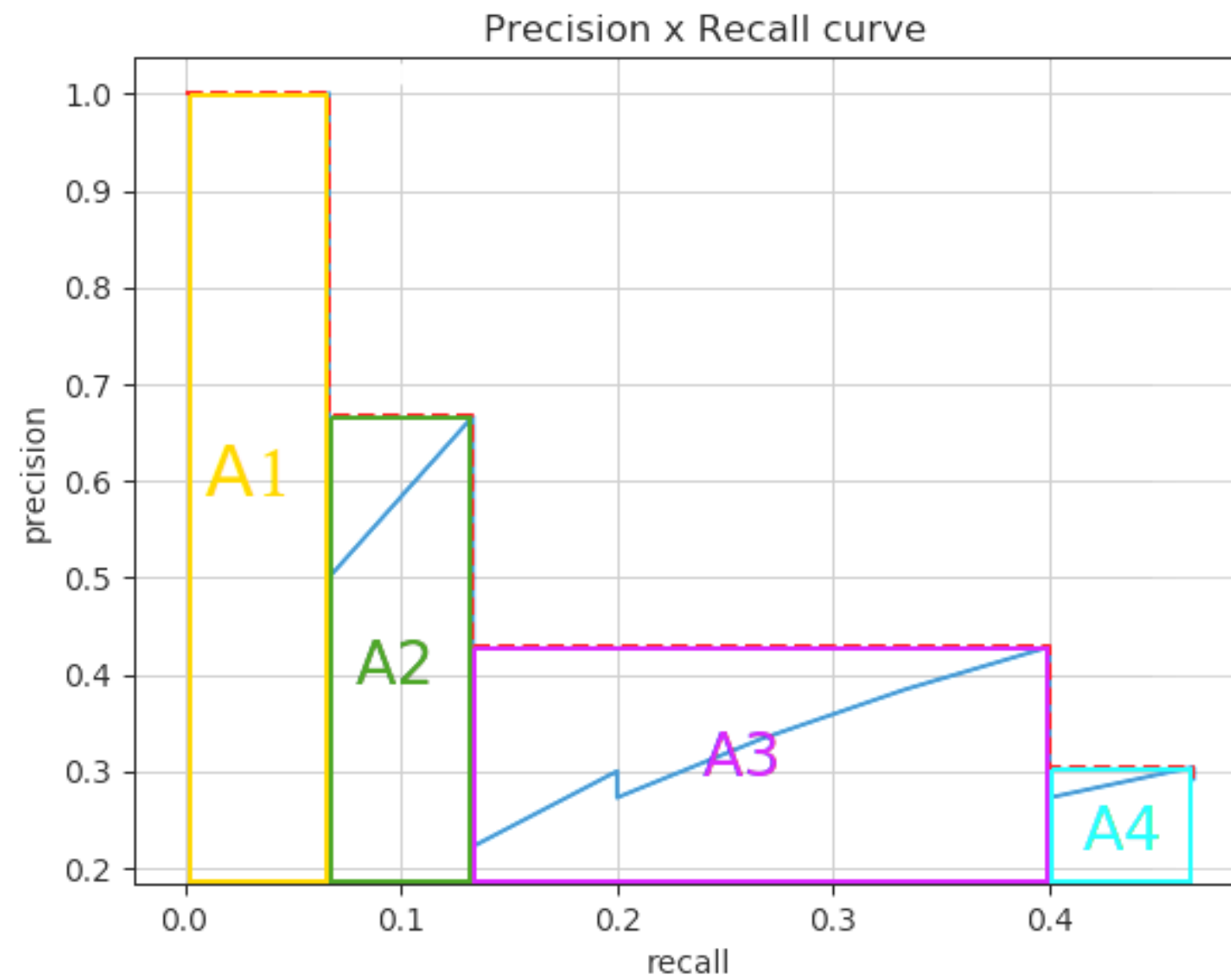
$$AP = \frac{1}{11} (1 + 0.6666 + 0.4285 + 0.4285 + 0.4285 + 0 + 0 + 0 + 0 + 0 + 0) = 0.268372 \approx 26.84\%$$

#### All-point interpolation

Cách thứ hai được gọi là all-point interpolation, ta tính interpolated precision với chính xác các mốc recall của Precision x Recall curve.



Cuối cùng, ta tính toán AuC của đường interpolated precision để thu được chỉ số AP cần tính



$$AP = A1 + A2 + A3 + A4$$

$$A1 = (0.0666 - 0) \times 1 = 0.0666$$

$$A1 = (0.1333 - 0.0666) \times 0.0666 = 0.04446222$$

$$A1 = (0.4 - 0.1333) \times 0.4285 = 0.11428095$$

$$A1 = (0.4666 - 0.4) \times 0.3043 = 0.02026638$$

$$AP = 0.0666 + 0.04446222 + 0.11428095 + 0.02026638 = 0.24560955 \approx 24.56\%$$

Hai cách tính chỉ số AP khác nhau là 11-point interpolation và All-point interpolation cho ta hai giá trị AP khác nhau một chút là 26.84% và 24.56%.