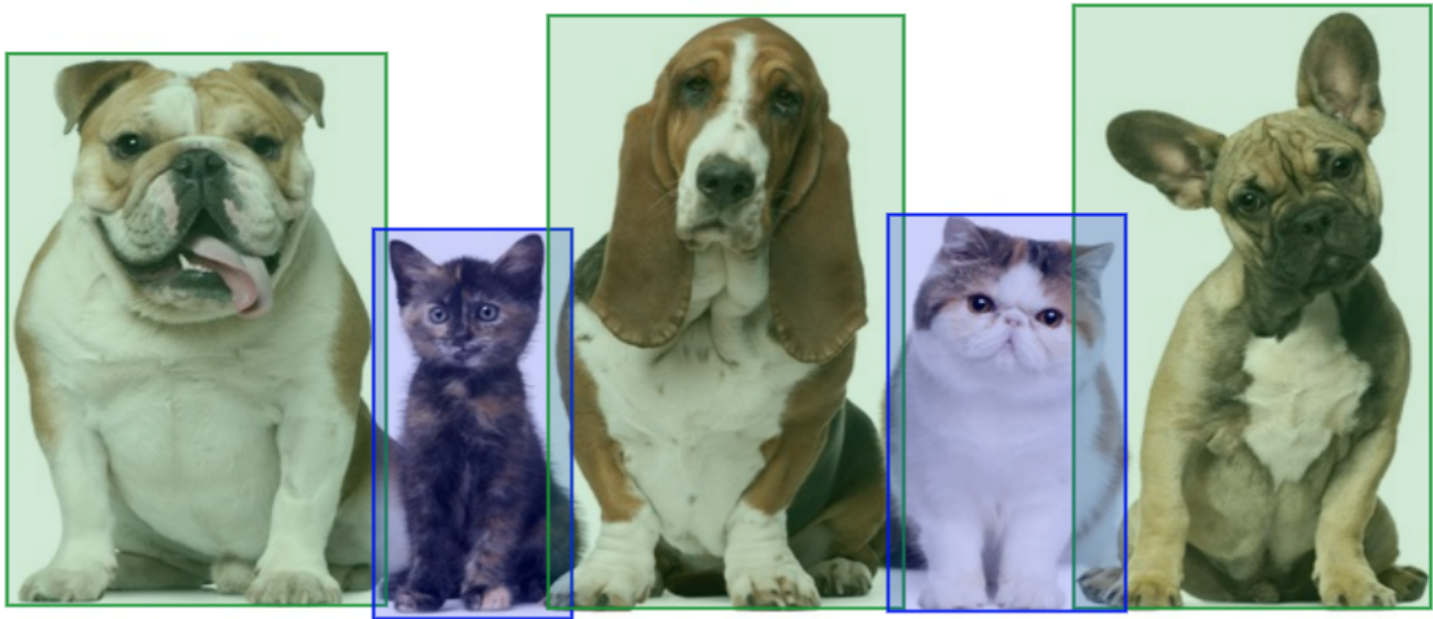
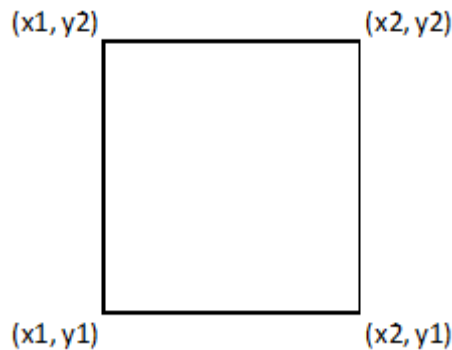


Detecție automată câini și pisici

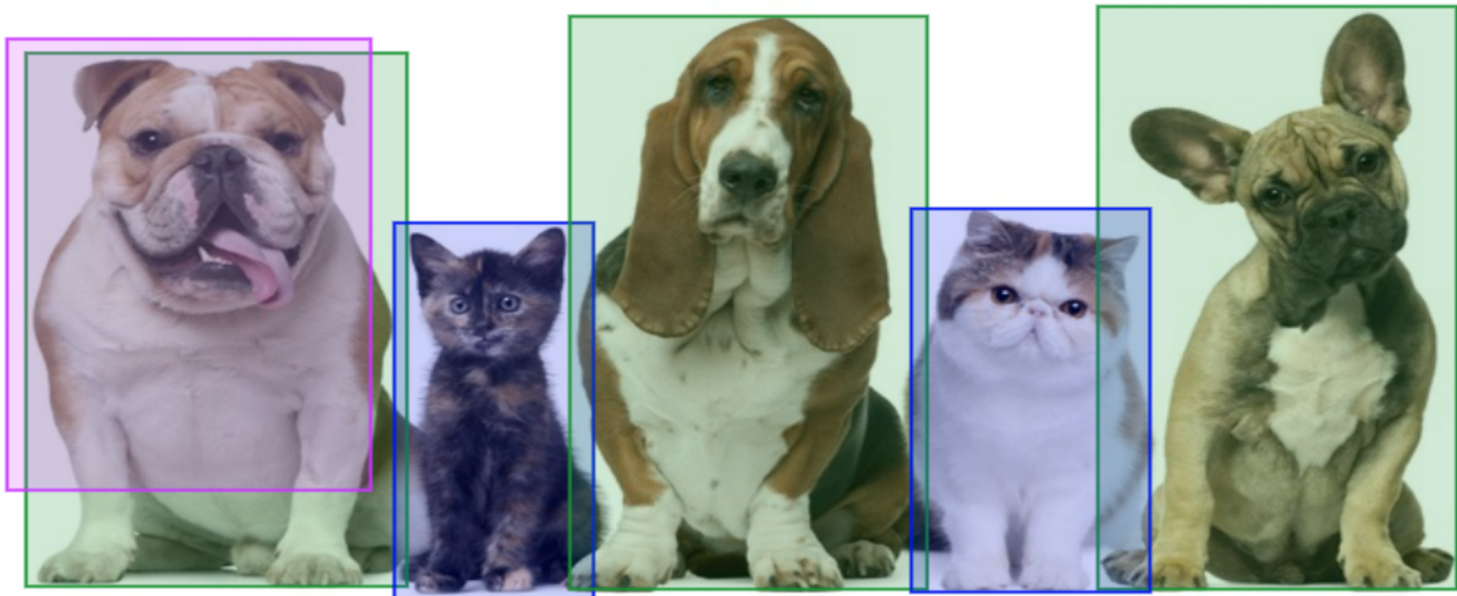
Detecția automată a obiectelor într-o imagine presupune încadrarea tuturor obiectelor cunoscute într-o casetă de încadrare (bounding box) și asocierea lor cu etichete care să clasifice obiectele. Un astfel de exemplu se găsește în imaginea de mai jos.



Casetele verzi sunt folosite pentru eticheta “câine”, iar cele albastre pentru eticheta “pisică”. O casetă de încadrare este caracterizată complet de o etichetă și 4 coordonate întregi (x_1, y_1, x_2, y_2) astfel:

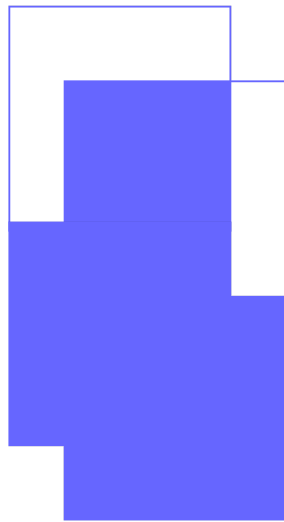


Acest exemplu este adnotat de un utilizator uman și reprezintă referința sistemului (groundtruth), adică idealul pe care poate să îl atingă un sistem automat de detecție a obiectelor. De cele mai multe ori, însă, detecția sistemului automat nu este perfectă, erorile apărând în poziționarea casetei de încadrare sau în clasificarea obiectului (sau în ambele cazuri, concomitent). În imaginea de mai jos, caseta roz reprezintă predicția oferită de un sistem de detecție automată a obiectelor care a clasificat imaginea încadrată în acest chenar drept “câine”. Eroarea vine din plasarea casetei de încadrare, care nu a reușit să încadreze complet câinele din imagine.



Pentru a determina **magnitudinea erorii de încadrare** se calculează o măsură denumită “Intersecție supra Reuniune” (Intersection over Union - IoU), astfel:

$$IoU = \frac{\text{Suprafața Intersecției}}{\text{Suprafața Reuniunii}}$$



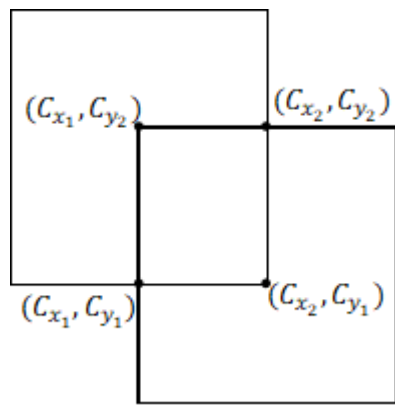
Dacă cele 2 casete de încadrare sunt caracterizate de coordonatele $(A_{x_1}, A_{y_1}, A_{x_2}, A_{y_2})$, respectiv, $(B_{x_1}, B_{y_1}, B_{x_2}, B_{y_2})$, atunci intersecția lor va avea coordonatele:

$$C_{x_1} = \max(A_{x_1}, B_{x_1})$$

$$C_{x_2} = \min(A_{x_2}, B_{x_2})$$

$$C_{y_1} = \max(A_{y_1}, B_{y_1})$$

$$C_{y_2} = \min(A_{y_2}, B_{y_2})$$



Notă: suprafețele se calculează în pixeli. Prin urmare, un segment orizontal între $x_1 = 4$ și $x_2 = 7$, va avea lungimea $7 - 4 + 1 = 4$ (pixeli).

Să se scrie o aplicație care să evalueze performanța unui sistem de detecție automată a câinilor și pisicilor în imagini. Evaluarea se va face comparând rezultatele prezise de un astfel de sistem, denumite și **ipoteză**, cu rezultatele reale, denumite și **referință**. Atât ipoteza, cât și referința **conțin cel mult 10 casete de încadrare**, iar o casetă de încadrare poate să conțină **doar una dintre cele 2 etichete**: “caine” sau “pisica”. Se citesc de la tastatură următoarele date, de pe rânduri diferite:

- un întreg n , reprezentând numărul de casete de încadrare ale referinței;
- datele pentru n casete de încadrare, în următoarea ordine:
 - 4 valori întregi, reprezentând coordonatele casetei în ordinea: x_1, y_1, x_2, y_2
 - un (singur) cuvânt, reprezentând eticheta obiectului
- o valoare întreagă în intervalul $[1, 4]$, în funcție de care se vor realiza diferite operații:
 1. afișare referință - se vor afișa pe ecran datele casetelor de încadrare ale referinței, în același format și în aceeași ordine în care au fost citite;
 2. calcul IoU - se citesc de la tastatură suplimentar datele unei casete de încadrare:
 - a. 4 valori întregi, reprezentând coordonatele.
 - b. un (singur) cuvânt, reprezentând eticheta obiectului.Se va afișa pe ecran cu 4 zecimale valoarea IoU calculată între prima casetă din referință și caseta ale cărei coordonate au fost introduse de la tastatură.
 3. verificare predicție casetă - se citesc de la tastatură suplimentar următoarele date:
 - a. 4 valori întregi, reprezentând coordonatele unei casete de încadrare prezisă de sistemul automat (ipoteză):.
 - b. un (singur) cuvânt, reprezentând eticheta obiectului prezis.
 - c. o valoare reală, reprezentând pragul IoU.

Să se verifice dacă predicția acestei casete este corectă. Pentru acest lucru se va calcula IoU între **caseta prezisă și caseta de pe aceeași poziție din referință**. Pentru perechea de casete ipoteză - referință cu valoarea IoU obținută se verifică dacă sunt îndeplinite 2 condiții simultan: etichetele coincid și valoarea IoU depășește pragul citit de la tastatură.

- În caz afirmativ (True Positive - TP), se afișează valoarea 1 pe ecran.
- În caz negativ (False Positive - FP), se afișează valoarea 0 pe ecran.

4. calcul precizie clasă - se citesc de la tastatură suplimentar următoarele date:

- un întreg m , reprezentând numărul de casete de încadrare prezise de sistem (ipoteza)
- datele pentru m casete de încadrare:
 - 4 valori întregi, reprezentând coordonatele.
 - un (singur) cuvânt, reprezentând eticheta obiectului.
- o valoare reală, reprezentând pragul IoU.

Să se calculeze precizia sistemului pentru perechile detectate, citite de la tastatură. Fiecare casetă de încadrare din ipoteză este comparată (IoU) cu caseta de pe aceeași poziție din referință. Pentru fiecare pereche se decide dacă predicția reprezintă TP sau FP, conform punctului 3. Precizia se calculează ca $\frac{TP}{TP+FP}$, și se calculează cumulat pentru toate casetele din ipoteză. Se va afișa pe ecran precizia obținută, cu 4 zecimale exacte.

Exemplu:

1. Afișare casete referință

Input	Output
5 26 199 12 282 caine 193 283 8 197 pisica 272 433 12 301 caine 510 672 13 306 caine 426 534 11 204 pisica 1	26 199 12 282 caine 193 283 8 197 pisica 272 433 12 301 caine 510 672 13 306 caine 426 534 11 204 pisica

2. Calcul IoU

Input	Output
5 26 199 12 282 caine 193 283 8 197 pisica 272 433 12 301 caine 510 672 13 306 caine 426 534 11 204 pisica 2 16 181 88 292 caine	0.5992

Explicație:

$\text{IoU}(\text{caseta}(26\ 199\ 12\ 282), \text{caseta}(16\ 181\ 88\ 292)) = 0.599243558$, conform formulelor prezentate

3. Verificare predicție v1

Input	Output
-------	--------

5 26 199 12 282 caine 193 283 8 197 pisica 272 433 12 301 caine 510 672 13 306 caine 426 534 11 204 pisica 3 16 181 88 292 caine 0.5	1
Explicație: 1. $IoU = 0.5992 > 0.5$ (prag) 2. Etichetele coincid	

4. Verificare predictie v2

Input	Output
5 26 199 12 282 caine 193 283 8 197 pisica 272 433 12 301 caine 510 672 13 306 caine 426 534 11 204 pisica 3 16 181 88 292 caine 0.6	0
Explicație: 1. $IoU = 0.5992 < 0.6$ (prag) 2. Etichetele coincid	

5. Verificare predictie v3

Input	Output
5 26 199 12 282 caine 193 283 8 197 pisica 272 433 12 301 caine 510 672 13 306 caine 426 534 11 204 pisica 3 16 181 88 292 pisica	0

0.5	
Explicație: 1. $IoU = 0.5992 > 0.5$ (prag) 2. Etichetele nu coincid	

6. Calcul precizie clasă

Input	Output
5 26 199 12 282 caine 193 283 8 197 pisica 272 433 12 301 caine 510 672 13 306 caine 426 534 11 204 pisica 6 5 16 181 88 292 caine 204 290 30 180 caine 272 413 11 300 caine 600 672 13 306 caine 426 534 11 204 pisica 0.5	0.6
Explicație: 1. $IoU(caseta(26\ 199\ 12\ 282), caseta(16\ 181\ 88\ 292)) = 0.599243558 > 0.5$ si <code>caine == caine</code> => TP 2. $IoU(caseta(193\ 283\ 8\ 197), caseta(204\ 290\ 30\ 180)) = 0.658418 > 0.5$ si <code>pisica != caine</code> => FP 3. $IoU(caseta(272\ 433\ 12\ 301), caseta(272\ 413\ 11\ 300)) = 0.870888 > 0.5$ si <code>caine == caine</code> => TP 4. $IoU(caseta(510\ 672\ 13\ 306), caseta(600\ 672\ 13\ 306)) = \mathbf{0.447853} < \mathbf{0.5}$ si <code>caine == caine</code> => FP 5. $IoU(caseta(426\ 534\ 11\ 204), caseta(426\ 534\ 11\ 204)) = 1 > 0.5$ si <code>pisica == pisica</code> => TP Rezultat final: $3/5 = 0.6$	