

# Aplicarea tehnicilor de detecție și recunoaștere a obiectelor în separarea stilurilor arhitecturale

## I. INTRODUCTION

În acest document sunt prezentate etapele parcurse și rezultatele obținute în cadrul unui proiect de detectare și recunoaștere a obiectelor de mobilier, având ca scop demonstrarea aplicabilității tehnicilor de procesare a imaginilor și a inteligenței artificiale în domeniul imobiliar. Pentru a demonstra utilitatea proiectului, au fost abordate două direcții de experimentare: una dedicată detectării obiectelor de mobilier și cealaltă axată pe identificarea stilului de mobilier al acestor obiecte.

Abordările de învățare a rețelelor neuronale profunde pentru detectarea obiectelor au atras recent o atenție semnificativă, demonstrând predicții promițătoare pentru aplicațiile în timp real. Metodele de îmbunătățire pe bază de CNN în două etape, cum ar fi R-CNN, Faster R-CNN și Cascade R-CNN, au fost evaluate pentru viteza lor limitată de procesare, ceea ce restricționează utilizarea acestora în scenarii în timp real. În schimb, metodele DNN într-o singură etapă, cum ar fi YOLO și RetinaNet, și-au demonstrat eficiența în detectarea obiectelor. YOLO, în special, excelează simultan în clasificare și regresia casetelor de delimitare [1], făcându-l extrem de potrivit pentru aplicațiile în timp real. În consecință, extensiile și rafinările YOLO, și anume Yolo-v3, Yolo-v4 și Yolo-v5, au demonstrat o performanță îmbunătățită în ceea ce privește acuratețea, viteza și rata scăzută de cadre [2].

## II. UTILIZARE

În acest punct proiectul este dezvoltat doar sub forma unei aplicații demonstrative cu scopul de a pune accentul pe principiul de functionare și pe rezultatele obținute pentru a lua în considerare dacă ar fi sau nu utilă o astfel de aplicație. Cu toate acestea, se poate preciza că utilizarea unei astfel de aplicații nu va necesita cunoștințe avansate. În principiu, utilizatorul va introduce o imagine a unei camere, iar în urma procesării imaginii, aplicația va oferi o listă de piese de mobilier sau va determina stilul în care se încadrează încăperea, pe baza obiectelor aflate în ea. Stilurile arhitecturale cunoscute de sistem sunt medieval și modern.

## III. CONSIDERENTE GENERALE

În cadrul acestui proiect, imaginile procesate au dimensiunea de 640 x 640 pixeli, iar pentru detectarea obiectelor a fost utilizat modelul YOLOv5. Anotarea datelor a fost realizată prin platforma Roboflow, iar Google Colab a fost folosit pentru antrenarea și utilizarea modelului. De asemenea, au fost utilizate două tipuri de seturi de date. Primul set de date se referă la diverse piese de mobilier, incluzând: scaun de

calculator, masă de calculator, canapea, scaun de sufragerie, masă de călcat, scaun de metal, masă de birou, scaun de plastic, scaun de restaurant, masă de studiu, scaun de lemn, canapea de lemn și masă de lemn. Al doilea set de date se concentrează pe stiluri ale unor piese de mobilier, cum ar fi: scaun modern, pat modern, masă modernă, scaun medieval, masă medievală și pat medieval.

## IV. SETURI DE DATE

### A. Datele utilizate

În proiectul în discuție, s-au folosit două seturi de date distincte: unul dedicat identificării pieselor de mobilier și celălalt concentrat pe stiluri de mobilier. Primul set de date, predefinit și disponibil pe platforma Roboflow, poate fi accesat la această adresă. Cel de-al doilea set de date a fost creat manual, folosind funcționalitățile oferite de Roboflow. Imaginile utilizate în acest set au fost selectate de pe internet în scopuri exclusiv academice, fără a exista drepturi de autor asociate lor. Obiectele prezente în aceste imagini au fost conturate folosind trei funcții principale. Prima funcție permite delimitarea obiectelor în cadrul unor cadre prestabilite, cea de-a doua oferă posibilitatea creării de forme neregulate utilizând poligoane, iar ultima metodă se bazează pe inteligența artificială pentru delimitarea obiectelor. În cadrul proiectului, s-au utilizat aceste două ultime metode pentru a diversifica modalitatea de identificare a obiectelor din setul de date. De precizat este faptul că am lucrat cu dataseturi de dimensiuni reduse de 600 - 700 imagini.

### B. Augmentarea seturilor de date

Augmentarea seturilor de date poate crește semnificativ performanțele sistemului, în principiu folosind metode artificiale sunt create imagini noi, derivate din cele existente, având trasaturi modificate, printre aceste trasaturi se afla: rotirea, luminozitate, contrast, translație, inversare etc. În cadrul proiectului am utilizat în prima etapă *shear*<sup>1</sup> pentru augmentarea setului de date, ce a triplat dimensiunea setului de date. Într-o etapă de optimizare am adăugat și augmentarea prin luminozitate  $\pm 20\%$  *zgomot*  $\pm 2\%$ .

<sup>1</sup>În contextul augmentării seturilor de date pentru prelucrarea imaginilor, termenul "shear" se referă la o transformare geometrică aplicată unei imagini. Shear implică deplasarea sau înclinarea pixelilor imaginii într-o direcție specifică, păstrând în același timp proporționalitatea și dimensiunile relative ale obiectelor din imagine.

## V. ETAPA DE IMPLEMENTARE

Se clonează repository-ul YOLOv5 din GitHub și se instalează toate dependențele necesare pentru a rula proiectul. Acest lucru asigură că toate pachetele și bibliotecile necesare sunt disponibile în mediu.

Se utilizează platforma Roboflow pentru a descărca datasetul de antrenament necesar pentru antrenarea modelului. Acest dataset poate fi accesat printr-un API și este găzduit pe platforma Roboflow. Datasetul este descărcat și pregătit pentru antrenare.

Modelul YOLOv5 este antrenat folosind datasetul descărcat. Acest proces de antrenare implică ajustarea parametrilor modelului pentru a învăța să detecteze obiecte în imagini. Antrenarea se efectuează pe multiple epoci pentru a îmbunătăți performanța modelului.

După antrenarea modelului, se evaluează performanța acestuia folosind metrici precum pierderea și acuratețea. Aceste metrici oferă o înțelegere a modului în care modelul se comportă pe setul de date de test.

Utilizând modelul antrenat, se efectuează detecția obiectelor pe un set de imagini de test. Acest lucru implică identificarea și localizarea obiectelor în imagini, precum și clasificarea lor în clase specifice.

Rezultatele detecției sunt vizualizate prin afișarea imaginilor de test cu obiectele detectate evidențiate. Acest lucru oferă o perspectivă vizuală asupra performanței modelului și a calității detecțiilor.

## VI. REZULTATE

Am început antrenarea sistemului utilizând 100 de epoci de antrenare, apoi am crescut treptat la 150 respective 300 pentru a identifica numărul de epoci aproximativ ce determină cea mai bună precizie. Acest lucru ar fi putut fi realizat și automat însă am dorit să obținem și o vizualizare grafică a diferențelor obținute prin modificarea numărului de epoci.

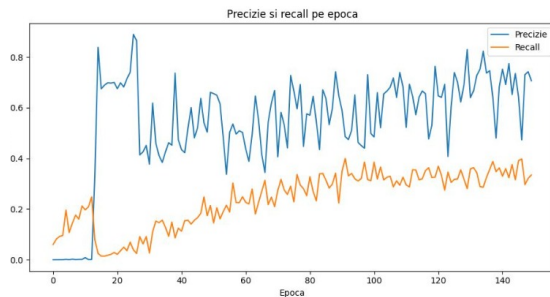


Fig. 1.

În urma comparării Figurilor 1,2 cu 3,4 am ajuns la concluzia că 150 de epoci nu sunt destule pentru a atinge un punct de saturatie al precizie drept pentru care am optat pentru o antrenare bazată pe 300 de epoci.

Din tabelul I se poate observa evoluția optimizării sistemului prin adăugarea de augmentari setului de date.

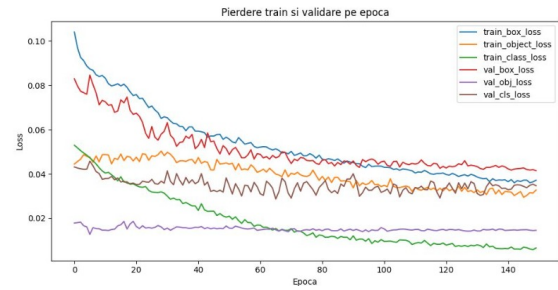


Fig. 2.

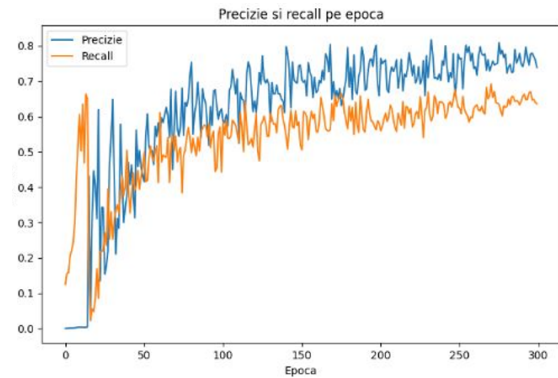


Fig. 3.

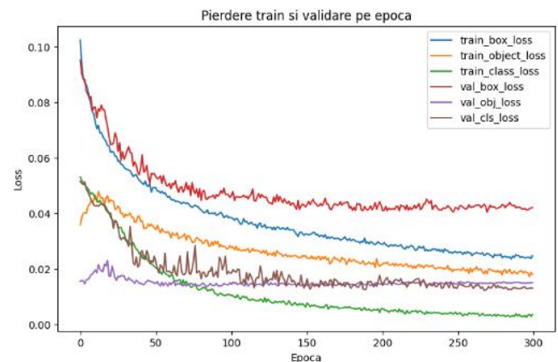


Fig. 4.

TABLE I  
METRICI PERFORMANȚĂ

Augmentare	Precizie	Recall	MAP @0.5
sher	0.62	0.57	0.58
luminozitate și zgomet	0.64	0.53	0.57
sher,luminozitate și zgomet	0.77	0.67	0.68

## VII. CONCLUZII

În cadrul proiectului am încercat să obținem rezultate ce pot confirma faptul că tehnicile de detectare și identificare a obiectelor pot fi aplicate pentru a diferenția piesele de mobilier, ba mai mult, stilurile de care acestea aparțin.

Am lucrat cu seturi de date de dimensiuni reduse peste care am aplicat diverse augmentări pentru a îmbunătăți precizia.

Astfel, după ajustarea parametrilor rețelei, am ajuns la o precizie de 77%. Această precizie este suficientă pentru a confirma ipoteza că tehnicile de detectare și identificare a obiectelor pot fi folosite atât în separarea stilurilor arhitecturale, cât și în diferențierea pieselor de mobilier. În cadrul studiului am considerat doar două stiluri arhitecturale: modern și medieval. Acest studiu poate motiva cercetări mai ample ce iau în considerare stiluri arhitecturale ce au trăsături în comun.

#### REFERENCES

- [1] T. Dean, M. A. Ruzon, M. Segal, J. Shlens, S. Vijayanarasimhan, and J. Yagnik, "Fast, accurate detection of 100,000 object classes on a single machine," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1814–1821, New York, NY, USA, 2013.
- [2] X. Yu, T. W. Kuan, Y. Zhang and T. Yan, "YOLO v5 for SDSB Distant Tiny Object Detection," 2022 10th International Conference on Orange Technology (ICOT), Shanghai, China, 2022, pp. 1-4, doi: 10.1109/ICOT56925.2022.10008164.
- [3] Jacob Murel, Eda Kavlakoglu, *Object Detection*, Published: 3 January 2024, <https://www.ibm.com/topics/object-detection>.
- [4] Ultralytics, *Train Custom Data*, <https://github.com/ultralytics/yolov5/wiki/Train-Custom-Data>.