



Descrierea mediului înconjurător cu NAO

Grosu Alexandru

Grădinaru Laurențiu-Andrei

1. Context & Motivație

- **Context:**

În domeniul prelucrării imaginilor, recunoașterea și descrierea mediului înconjurător reprezintă o provocare esențială pentru dezvoltarea roboților autonomi.

- **Motivație:**

Capacitatea unui robot de a înțelege și descrie elementele din jurul său îl face mai autonom și mai util în context reale (educaționale, casnice sau de asistență).

- **Obiectivul proiectului:**

Scopul proiectului este de a dezvolta o aplicație prin care robotul NAO să poată identifica, urmări și descrie obiectele din mediul înconjurător, utilizând informații vizuale obținute prin camera sa și exprimând rezultatele verbal.

2. Arhitectura preliminară a soluției

- **Schema arhitecturii:**

1. Nao capturează o imagine din mediul înconjurător;
2. Imaginea este codificată și transmisă prin HTTP către un server Flask;
3. Serverul rulează un model YoloV8 pentru detecția obiectelor;
4. Rezultatul (lista obiectelor) este transmisă înapoi robotului;
5. NAO anunță verbal obiectele detectate (ex: "I saw, laptop, chair, person").

- **Descrierea componentelor:**

1. Camera lui NAO – captează în timp real imaginea din mediul înconjurător;
2. Server Flask (laptopul): decodează imaginea, rulează YOLOv8, returnează rezultatul în format JSON;
3. Rețea locală (Wi-Fi): asigură comunicarea bidirecțională între robot și server;

2. Arhitectura preliminară a soluției

- **Fluxul de date:**

- **Captură imagine** – NAO obține un cadru de la camera sa RGB.
- **Pre-procesare** – imaginea este convertită și codificată în Base64.
- **Transmitere către server** – folosind o cerere HTTP POST (Flask endpoint).
- **Detecție obiecte** – serverul rulează YOLOv8 și extrage numele claselor recunoscute.
- **Răspuns JSON** – lista obiectelor este trimisă înapoi către NAO.
- **Sinteză vocală** – NAO interpretează răspunsul și spune:
„I saw: person, bottle, chair.”

3. Evaluarea Preliminară a Soluției

- **Metodologia de evaluare:**

Testul a constatat în rularea programului demo în laborator, observând cererea către server, procesarea și transmisia răspunsului.

- **Metrici:**

- Detecția corectă (rata de success ~ 90% pentru obiecte comune);
- Timpul de procesare (~2-4 secunde);
- Stabilitatea comunicării prin HTTP;

Setul de date: Datele provin din fluxul video live al camerei NAO; nu a fost folosit un set extern de imagini, ci mediul real.

4. Rezultate Preliminare

- **Rezultate obținute:**

I. Robotul Detectează corect persoana în aproximativ 80-90% din cazuri;

II. Comunicarea dintre NAO și server este stabilă;

III. Răspunsul vocal este emis clar și sincron cu rezultatul YOLO.

- **Interpretarea rezultatelor:** Integrarea dintre robot și modelul YOLOv8 oferă o detecție rapidă și fiabilă, permițând robotului să descrie mediul său în timp real.

5. Concluzii Preliminare

- **Rezumatul progresului:** Am reușit implementarea unei soluții de bază prin care robotul NAO detectează cea mai apropiată persoană și o urmărește, reacționând verbal.
- **Limitările soluției actuale:**
 - I. Necesită conexiune de rețea stabilă;
 - II. Distanțele mari afectează precizia;
 - III. Procesarea se face pe un server extern.
- **Potențiale îmbunătățiri:** Adăugarea unei componente de urmărire (Indicarea folosind mâna a obiectului detectat).

6. Direcții Viitoare

- **Pași următori:** Se va extinde funcționalitatea sistemului prin optimizarea vitezei de comunicare între robot și server, îmbunătățirea acurateței detecției YOLOv8 și adăugarea recunoașterii contextuale a obiectelor din mediul înconjurător.
- **Plan de implementare:** Se vor integra funcții de urmărire a obiectelor, se vor testa versiuni optimizate ale modelului YOLOv8 și se va dezvolta o interfață web pentru vizualizarea în timp real a imaginilor și detecțiilor.
- **Obiectivele finale:** Robotul NAO va putea detecta și descrie complet scena vizuală.