

**CEX- Detecția și recunoașterea limbajului semnelor din
imagini/secvențe video**

**Epure Carla-Maria
Velnic Vlad-Andrei**

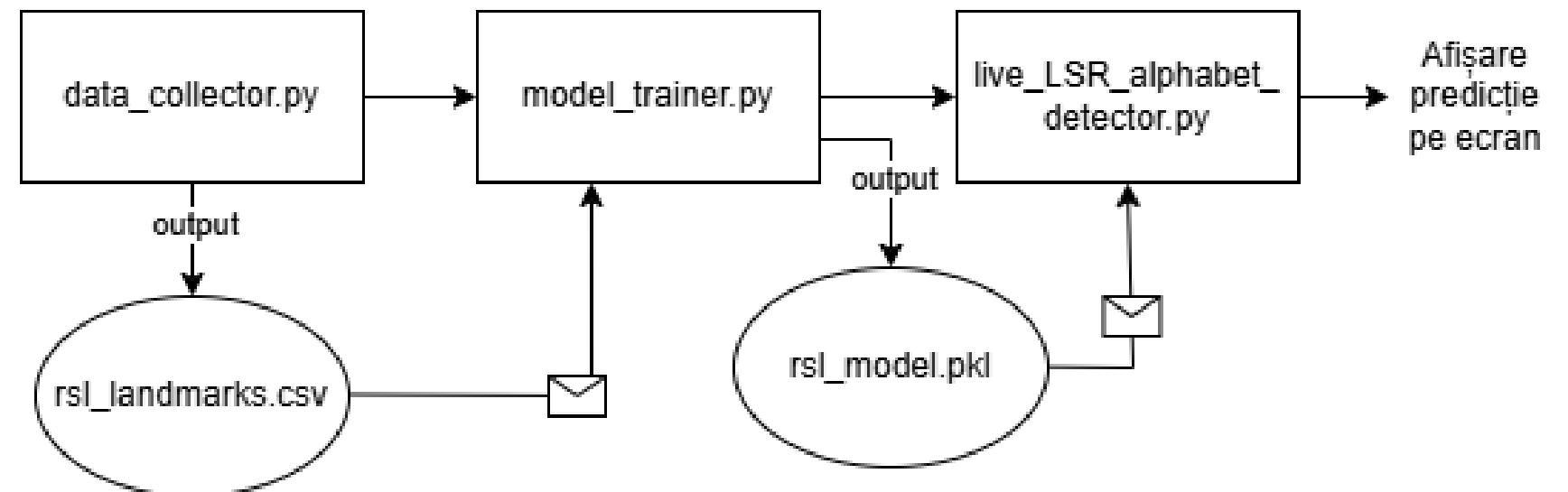
1. Context & Motivătie

- **Provocări actuale din contextul prelucrării imaginilor:**
 - Recunoașterea gesturilor în timp real
 - Detectia poziției mâinilor
 - Clasificarea semnelor în contextul conversației
- **Motivătie:**
 - Facilitează comunicarea între persoanele cu și fără deficiențe de auz
 - Sprijină incluziunea socială
- **Obiectivul proiectului:**

Dezvoltarea unui sistem inteligent capabil să recunoască limbajul ASL, urmând să fie extins pentru RSL - Romanian Sign Language.



2. Arhitectura preliminară a soluției



- **`data_collector.py` (modul colectare date):**
 - creează setul de date -> `rsl_landmarks.csv`
 - transformă tastă apăsată într-o etichetă, pe care o asociază cu cele 21 puncte de referință pentru fiecare mână
- **`model_trainer.py` (modul antrenare):**
 - împarte setul de date (80% antrenare, 20% testare)
 - antrenează un model Random Forest Clasifier -> `rsl_model.pkl`
- **`live_LSR_alphabet_detector.py` (modul detectie live):**
 - calculează punctele de referință ale mâinii și le trimite la modelul antrenat care returnează o predicție

3. Evaluarea Preliminară a Soluției

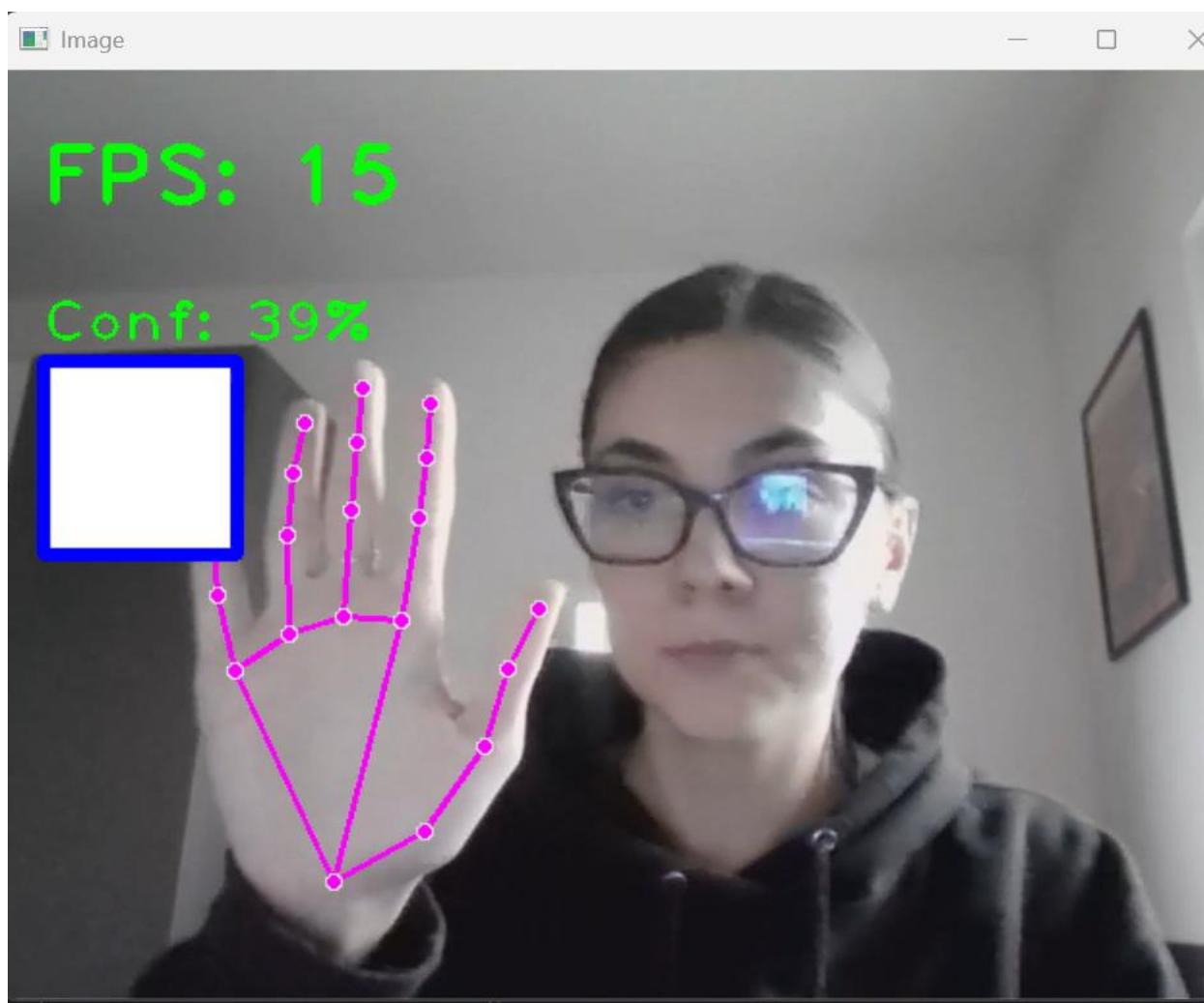
Evaluare cantitativă	Evaluare calitativă
<ul style="list-style-type: none">- Realizată automat de scriptul <code>model_trainer.py</code>- Modelul este antrenat pe 80% din date și testat pe 20% din date- Metrici folosite:<ul style="list-style-type: none">• precizia• sensibilitatea• scorul f1	<ul style="list-style-type: none">- Realizată practic, rulând scriptul <code>live_LSR_alphabet_detector.py</code>- Metrici folosite:<ul style="list-style-type: none">• stabilitatea predicției• viteza de răspuns• pragul de încredere (0.65)

Classification Report:			
	precision	recall	f1-score
a	1.00	0.99	0.99
b	0.97	1.00	0.98
c	0.99	0.97	0.98
d	0.98	1.00	0.99
e	0.96	0.98	0.97
f	1.00	0.98	0.99
g	1.00	1.00	1.00
h	1.00	1.00	1.00

- **Cazuri de test**
 - litere asemănătoare vizual (U vs V, M vs N vs T, A vs S)
 - variații de unghiuri și distanțe
 - lipsa unui gest cunoscut (nicio mâncă în cadru, mâncă în repaus)

4. Rezultate Preliminare

- Modelul are o acuratețe rezonabilă pentru fluxul video de date.
- Scopul inițial a fost să creăm un prototip funcțional care să detecteze real-time literele alfabetului.
- Necesitatea antrenării modelului pentru eliminarea confuziei între literele similare: U, V, M, N, T.
- Pragul relativ scăzut de încredere setat (0.65) oferă cîteodată rezultate incerte.



5. Concluzii Preliminare

- **Rezumatul progresului:**

- Implementarea a 3 module funcționale pentru recunoașterea alfabetului în limbajul semnelor

- **Limitările soluției actuale:**

- Modelul este antrenat doar pe datele noastre- modul propriu de gesticulare, condiții de iluminare
- Conține doar semne statice- nu recunoaște litere care implică mișcare (J, Z)
- Limitat doar la litere, nu înțelege cuvinte sau concepte

- **Potențiale îmbunătățiri:**

- Diversificarea setului de date
- Extinderea vocabularului și a semnelor dinamice

6. Directii Viitoare

- **Pași următori:**

- Diversificarea setului de date și reantrenarea modelului
- Cercetarea metodelor pentru integrarea semnelor dinamice

- **Plan de implementare:**

- Înregistrarea datelor de alte persoane
- Trecerea de la cadre individuale la secvențe de cadre pentru învățarea tiparelor temporale ale literelor
- Extinderea vocabularului cu fraze/ cuvinte de bază

- **Obiectivele finale:**

- Realizarea unui model antrenat pe date diversificate, cu utilizatori mulți și o acuratețe de peste 90%
- Recunoașterea alfabetului complet
- Stabilitate crescută