

Background removal in real time

**Burlacu Cristian-Ștefan
Pavel Ștefan**

1. Context & Motivație

- **Context:**

Această problemă este esențială în aplicații precum: videoconferințe, streaming sau filmare profesională , unde fundalul trebuie modificat sau eliminat fără a folosi ecran verde.

- **Motivație:**

În prezent, comunicarea video este omniprezentă în aplicații precum Zoom, Teams sau OBS folosesc tehnici similare.

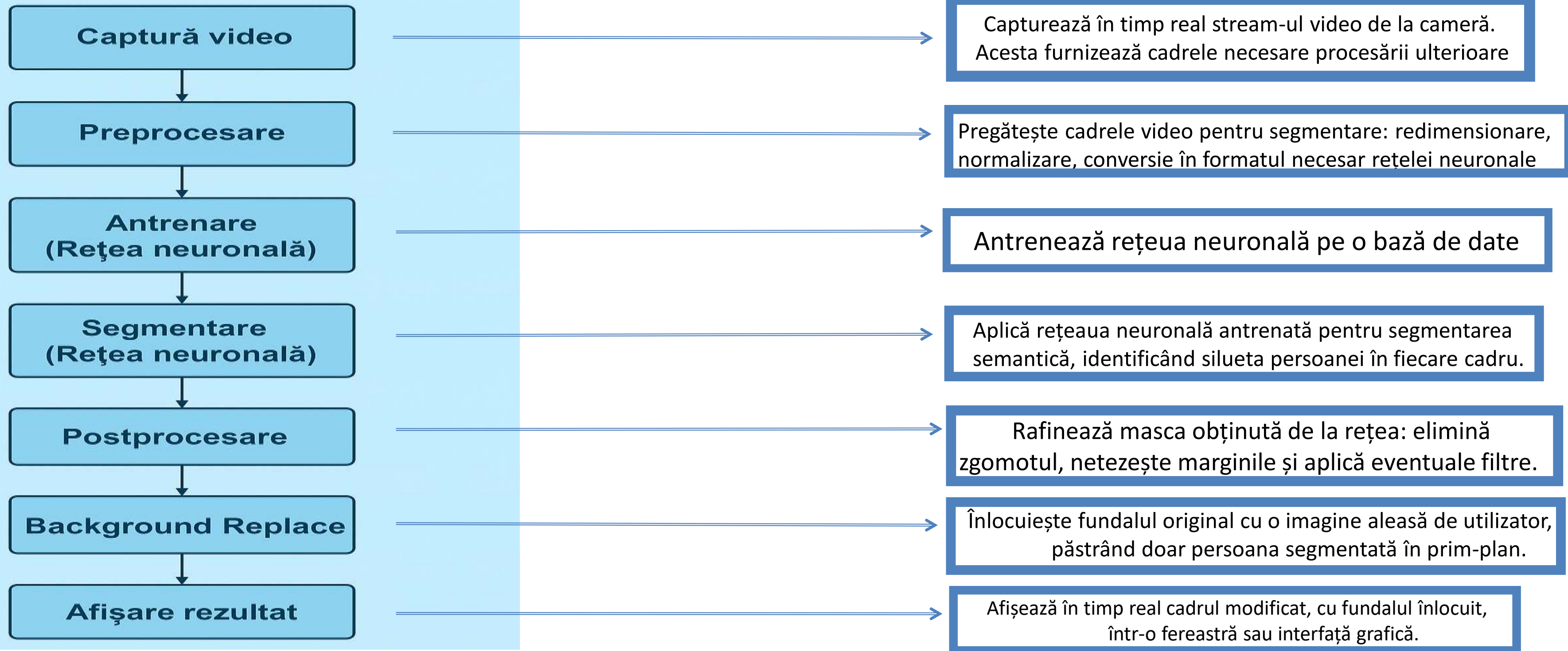
O soluție ar permite personalizarea mediului video, confidențialitate sporită și crearea de efecte vizuale profesionale fără echipament special.

- **Obiectivul proiectului:**

Dezvoltarea unui sistem automat care captează video în timp real, segmentează persoana folosind o rețea neuronală și înlocuiește fundalul conform alegerii utilizatorului, oferind un rezultat fluent și realist.

2. Arhitectura preliminară a soluției

- Schema arhitecturii:



3. Evaluarea Preliminară a Soluției

- **Metodologia de evaluare:**

Soluția a fost evaluată prin analizarea performanței modelului de segmentare pe același set de date folosit la antrenare (12.811 imagini + 12.811 măști binare).

- **Metrici folosiți pentru evaluare:**

IoU (Intersection over Union) - măsoară suprapunerea dintre masca prezisă de model și masca reală, valorile mai mari indică performanță mai bună

Loss - “cât de greșit a prezis” modelul față de masca reală

Dacă pixelul real = 1 și modelul a prezis 0.9 → greșeală mică

Dacă pixelul real = 1 și modelul a prezis 0.1 → greșeală mare

Exemple de date folosite pentru antrenare:



«..... Imagine originală»

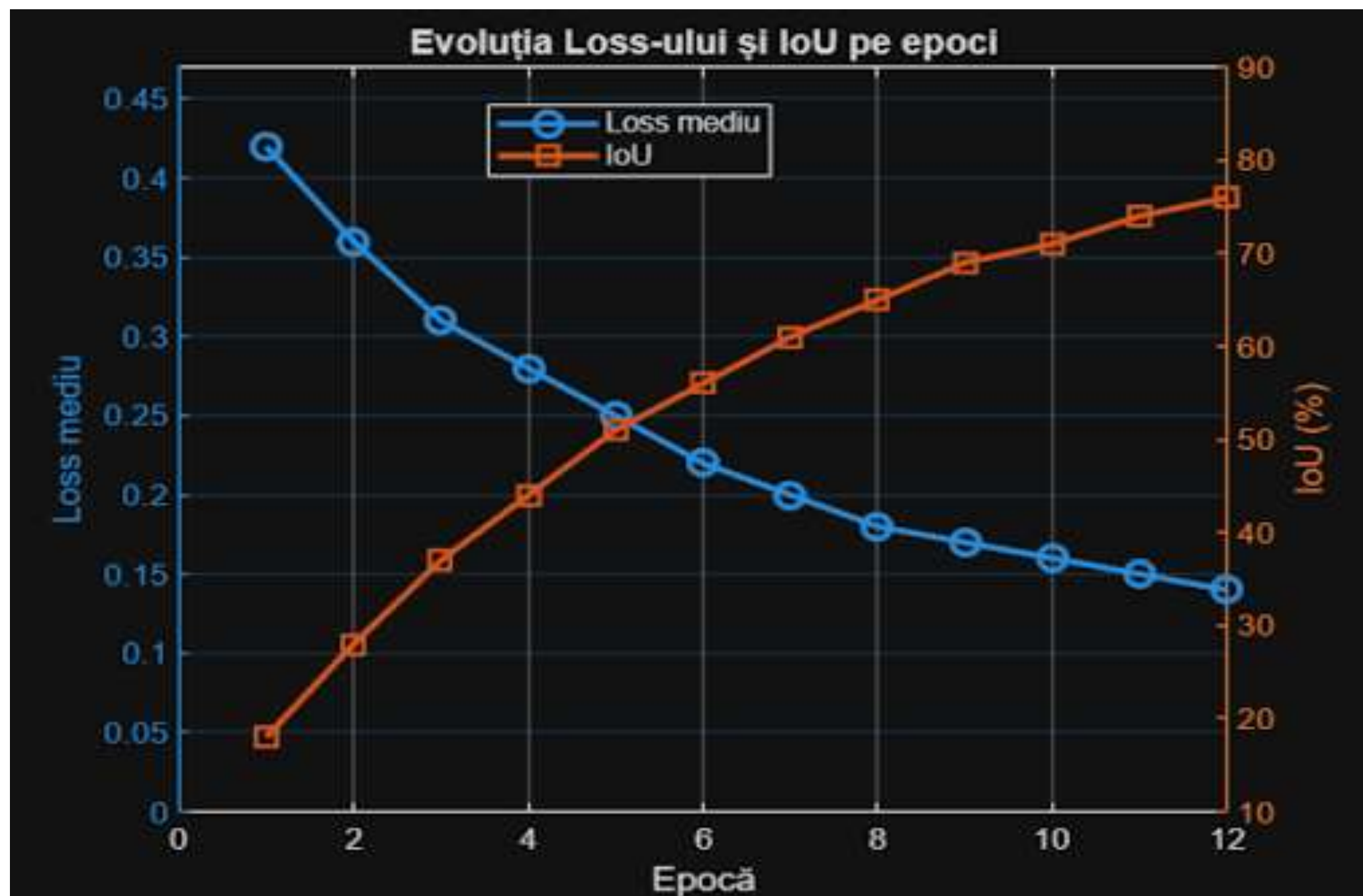


«..... Mască binară»



4. Rezultate Preliminare

- Rezultate obținute după antrenarea pe 12.811 imagini:



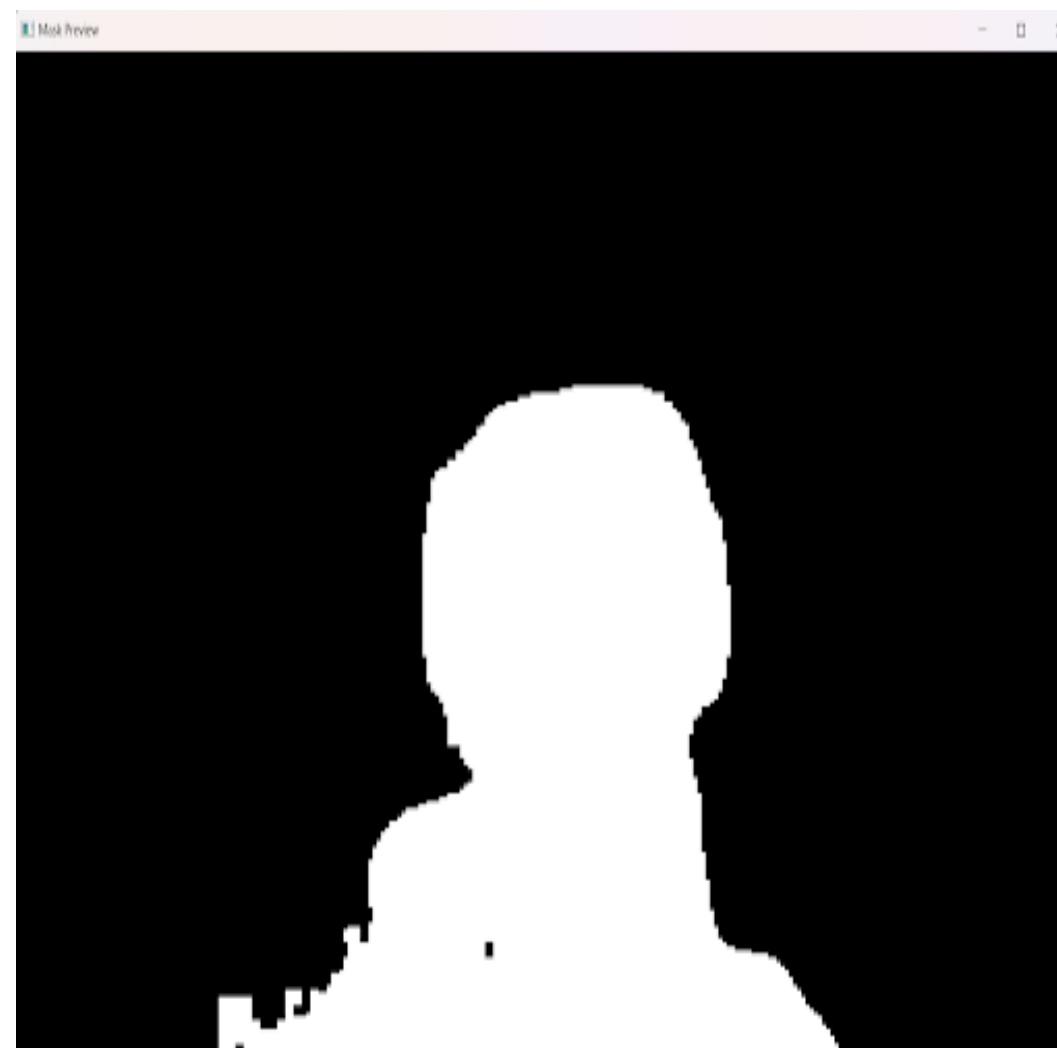
Pe măsură ce numărul epocilor crește, loss-ul scade constant, semn că modelul învață să facă predicții tot mai precise. În același timp, IoU crește progresiv, indicând o îmbunătățire clară a calității segmentării realizate de model.

4. Rezultate Preliminare

Vizualizare soluție actuală:



Input



Mască binară



Output

5. Concluzii Preliminare

- **Limitările soluției actuale:**

Modelul este încă simplu și poate greși în zone complexe (lumini, obiecte apropiate), iar performanța lui depinde puternic de calitatea setului de date.

Nu există încă un sistem real de evaluare, doar metrici în timpul antrenării.

- **Potențiale îmbunătățiri:**

Trecerea la un model mai avansat.

Utilizarea unui set de date mai mare și mai divers plus antrenare completă pe GPU pentru execuție mai rapidă.

6. Direcții Viitoare

- **Pași următori:**

Antrenarea modelului pe un set de date mai solid

Implementarea unui sistem de evaluare și verificare

Implementarea unui UI

Adaugarea funcționalității de background replace

- **Plan de implementare:**

Căutare strașnică a unor seturi de date mai mari, complexe și eficiente pentru o soluție mai bună.

Împărțirea setului de date pentru implemetarea unui sistem de validare după fiecare epocă și o testare finală pe un set de date separat.

Construirea unui UI folosind Tkinter ce afișează în timp real rezultatul final cu opțiunea de a schimba la alegere fundalul fie cu o imagine presetată fie prin încărcarea uneia noi de către utilizatorului.