

# Concepte și aplicații în Vederea Artificială - Tema 2

## Detectarea și recunoașterea facială a personajelor din serialul de desene animate Familia Flintstone

Cioclov Maria-Simona

21 ianuarie 2024

### Task 1 - Detectare facială

- **Etapa de antrenare:**

1. **Obținerea exemplelor pozitive:** Am creat un script cu ajutorul căruia să extrag, pe baza fișierului cu adnotări, patchurile cu fețele personajelor, pe care apoi le-am convertit la grayscale și le-am adus la diverse dimensiuni: 36x36, 42x42, 58x58, 64x64.
2. **Obținerea exemplelor negative:** Am modificat imaginile de antrenare convertite la grayscale, tot cu ajutorul fișierelor cu adnotări, astfel încât să elimin fețele de personaje din poze, punând în locul lor un patch alb. Astfel, voi putea extrage exemple negative din aceste imagini, luând din fiecare mai multe patchuri de dimensiunile exemplelor pozitive, cu siguranța că nu va exista, printre aceste exemple, niciunul care să conțină fața unui personaj.
3. **Obținerea descriptorilor pentru exemple pozitive și negative:** Descriu conținutul vizual al unei imagini folosind histograme de gradienti orientați, sub forma funcției *hog* din biblioteca *Sklearn*. Va urma să experimentez cu diverse valori pentru dimensiunea celulei *hog* și a blocului de celule.
4. **Antrenarea unui clasificator liniar:** Am folosit un model SVM pentru clasificarea imaginilor ca fiind exemple pozitive sau negative. Acesta primește ca input vectori reprezentând descriptorii *hog*.

- **Etapa de testare:**

1. **Metoda ferestrei glisante:** Pentru o imagine de test, glisez o fereastră de dimensiunea exemplelor pozitive și negative, de sus în jos și de la stânga la dreapta, cu pași de dimensiunea unei celule *hog*. Pentru fiecare poziție a ferestrei extrag patchul corespunzător, obțin descriptorii *hog* și aflui, cu ajutorul modelului liniar antrenat, dacă aceștia reprezintă un exemplu pozitiv sau negativ.

După ce am terminat de parcurs întreaga imagine, aceasta este redimensionată cu un anumit factor, iar procesul este reluat. Reiau acești pași pentru mai mulți factori de

redimensionare (1.2, 1, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1), cu scopul de a detecta în poză fete de diferite marimi.

2. **Localizarea corectă a unei fețe:** O localizare este considerată corectă atunci când zona identificată de către detector se intersectează cu cel puțin 30% din suprafața ferestrei adnotate.
3. **Eliminarea non-maximelor:** Deoarece doar detecția ce acoperă cel mai mult exemplul adnotat va fi considerată corectă, iar celelalte vor fi considerate fals pozitive, pentru un scor mai bun al performanței, trebuie eliminate detecțiile care se suprapun cu o altă detecție cu scor mai mare.
4. **Experimente:** Am testat cu ferestre de diverse dimensiuni: 36x36, 42x42, 58x58, 64x64. Pentru fiecare dintre acestea, am încercat să obțin descriptorii hog pentru diferite valori ale dimensiunii celulei hog (5, 6, 7, 8) și a blocului de celule (2, 3, 4, 5). Am variat, apoi, numărul de exemple negative folosite la antrenarea clasificatorului (10000, 50000, 70000, 90000). Observații:
  - cele mai bune rezultate au fost obținute cu ferestre de dimensiune 64x64
  - cele mai bune rezultate au fost obținute cu dimensiunea unei celule hog de 6 pixeli, iar dimensiunea unui bloc de celule de dimensiune 3x3
  - cu cât crește numărul de exemple negative, crește și acuratețea detectorului: 10000 - 0.505; 50000 - 0.578; 70000 - 0.599; 90000 - 0.617
  - cele mai bune rezultate au fost obținute cu thresholdul = 0

## Task 2 - Recunoaștere facială

**Dimensiunile ferestrelor glisate:** Fețele personajelor diferă ca formă, astfel că dreptunghiurile ce încadrează chipurile acestora în adnotări pot fi ori mai alungite (în cazul lui Fred și al Wilmei), ori mai apropiate de forma unui pătrat (ca în cazul lui Barney și al lui Betty). O precizie mai mare ar fi oferită dacă, pentru fiecare personaj, fereastra glisantă ar avea o formă potrivită feței respectivului personaj.

Pentru a afla care este cel mai potrivit raport dintre lățime și lungime pentru dreptunghiul corespunzător fiecărui personaj, am creat un script care calculează media rapoartelor tuturor chenarelor adnotate pentru personajul respectiv. Am obținut următoarele rezultate: Fred - 1.05, Wilma - 1.13, Barney - 0.91, Betty - 1.0469.

Rezolvarea taskului 2 constă în rularea codului de la taskul 1 pentru fiecare personaj în parte, cu parametrii coprespunzători, aleși în urma mai multor experimente și observații:

- **Fred:**
  - numărul de exemple negative: 20000
  - dimensiunea blocului hog: 3
  - raportul laturilor ferestrei glisante: 1.05
- **Wilma:**
  - numărul de exemple negative: 50000
  - dimensiunea blocului hog: 4

- raportul laturilor ferestrei glisante: 1.13

- **Barney:**

- numărul de exemple negative: 50000
- dimensiunea blocului hog: 3
- raportul laturilor ferestrei glisante: 0.91

- **Betty:**

- numărul de exemple negative: 50000
- dimensiunea blocului hog: 4
- raportul laturilor ferestrei glisante: 1.0469

Pentru toate personajele s-au folosit exemple de dimensiune 64x64, celule hog de dimensiune 6 și thresholdul 0.