

Proiect2 - CAVA - 2023

462, Ghergu Nicolae-Marius
CTI

Scopul acestui proiect a fost detectarea faciala a personajelor din desenul animat 'Familia Flintstone' (titlul original in engleza: "The Flintstones") sau aventuri in epoca de piata.

In arhiva se afla:

1. un fisier README cu instructiuni despre cum trebuie sa fie rulat script-ul si ce librarii sunt folosite in script
2. fisierele .py folosite pentru a antrena modelul pe baza datelor pozitive si negative
3. un fisier construire-date.py cu care mi-am construit datele pozitive si negative pentru model
4. fisierele .npy care constituie descriptorii pozitivi si negativi si modelul folosit pentru a evalua solutia.

Aceasta documentatie prezinta solutia propusa de mine pentru a atinge scopul acesui proiect si pasii cum am ajuns la aceasta solutie.

Pasul 1: Cum incep?

Primul pas pe care l-am inceput a fost sa ma gandesc de ce am nevoie pentru a antrena un model de machine learning. Pentru acest lucru, am inceput in a-mi construi un fisier .py prin care imi alegeam din datele pe care le aveam, date pentru modelul meu, clasificate in fisier: exemple pozitive pentru fiecare caracter din desenul animat si exemple negative pentru fiecare caracter din desenul animat (fundale ale caracterelor).

Pasul 2: Cum rezolv cerinta?

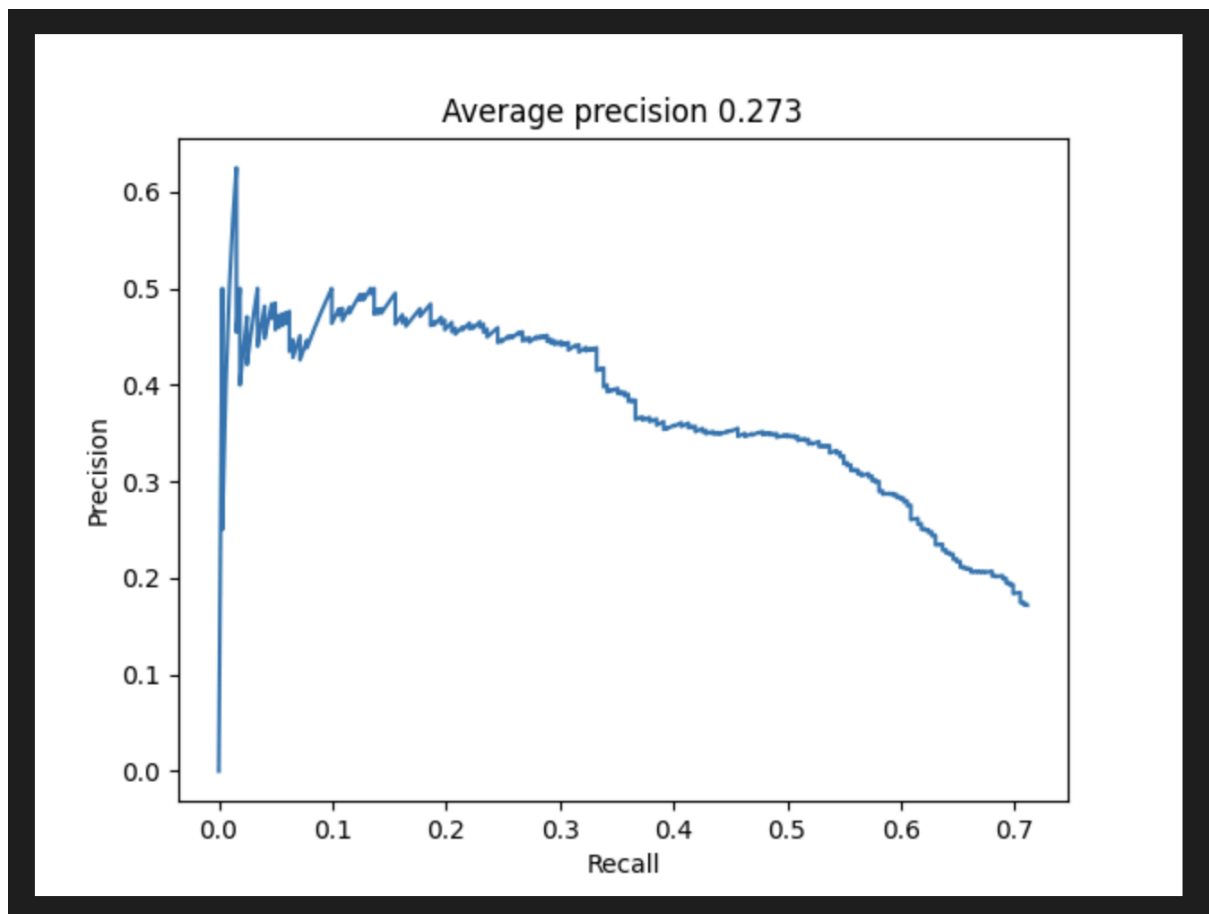
Pentru a rezolva cerinta am inceput a folosi codul folosit in Laboratorul 9+10, acesta avand la baza un model LinearSVC si care foloseste descriptori hog. Am incercat sa adaptez pe baza datelor mele pentru a

rezolva primul task, dar am obtinut rezultate foarte mici (average precision) pentru imagini de 36x36, asa ca am folosit imagini 72x72, avand un mai bun rezultat. Am avut si ceva erori de compilare care au necesitat resize-ul fiecărei imagini la 180x180 pentru a calcula descriptorii negativi.

Pasul 3: Obținerea unui scor mai bun

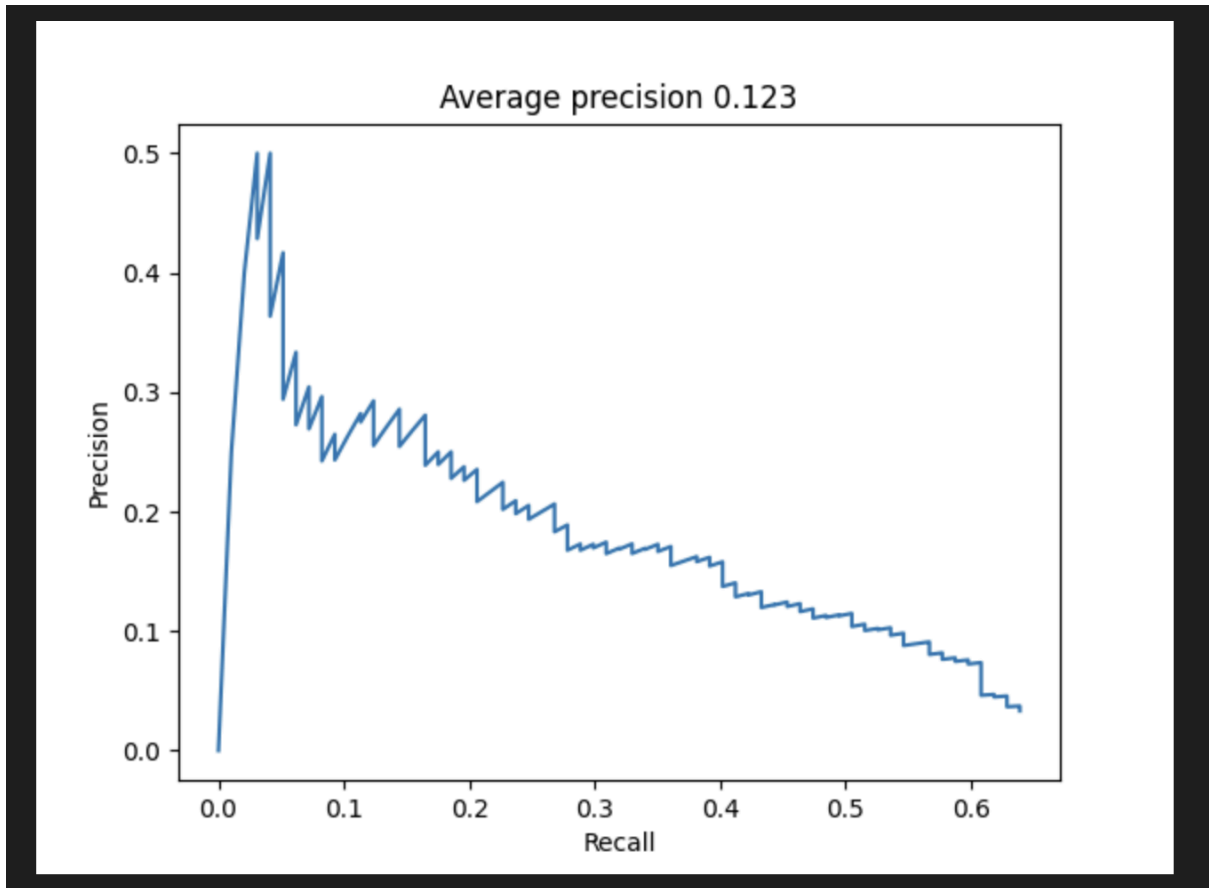
Am incercat diverse optiuni si anume redimensionarea sliding window, fara succes, caracteristici bazate pe culori, resize intre 0,5 si 2 pentru sliding window, dar fara succes, dupa multe antrenari si incercari. Solutia mea se bazeaza pe un threshold = 1 in final, folosind LinearSVC si descriptorii hog, folosind un hog_cell = 3.

Rezultate Task 1

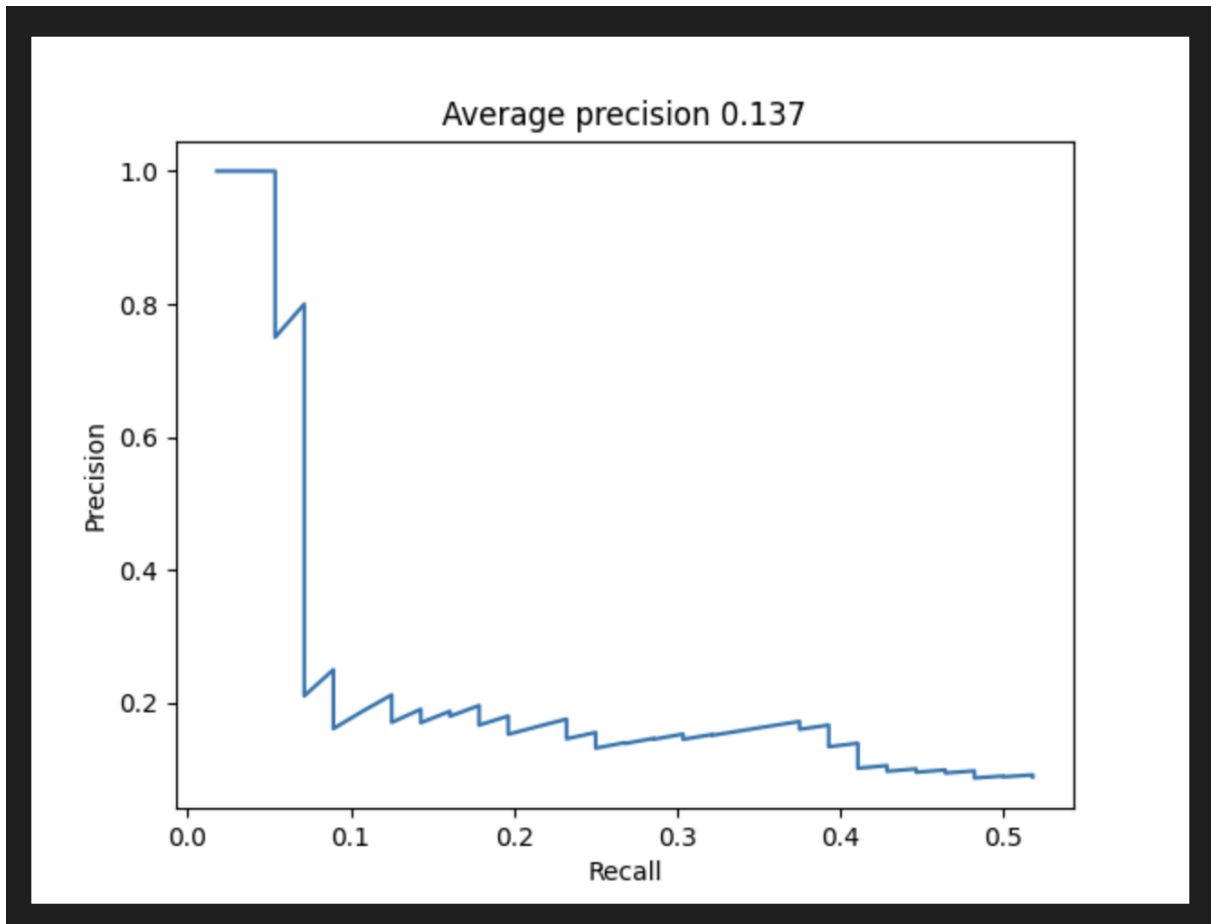


Rezultate Task 2

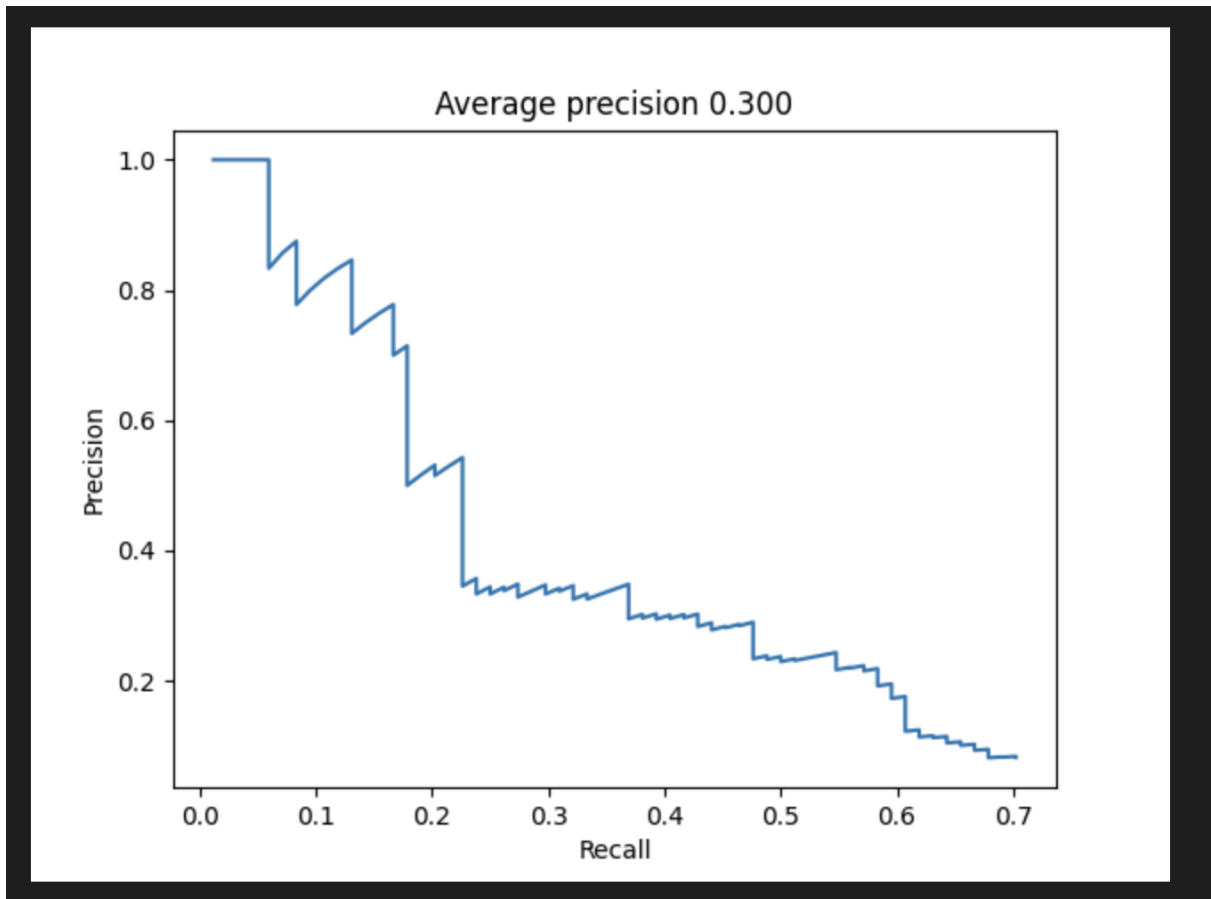
Barney



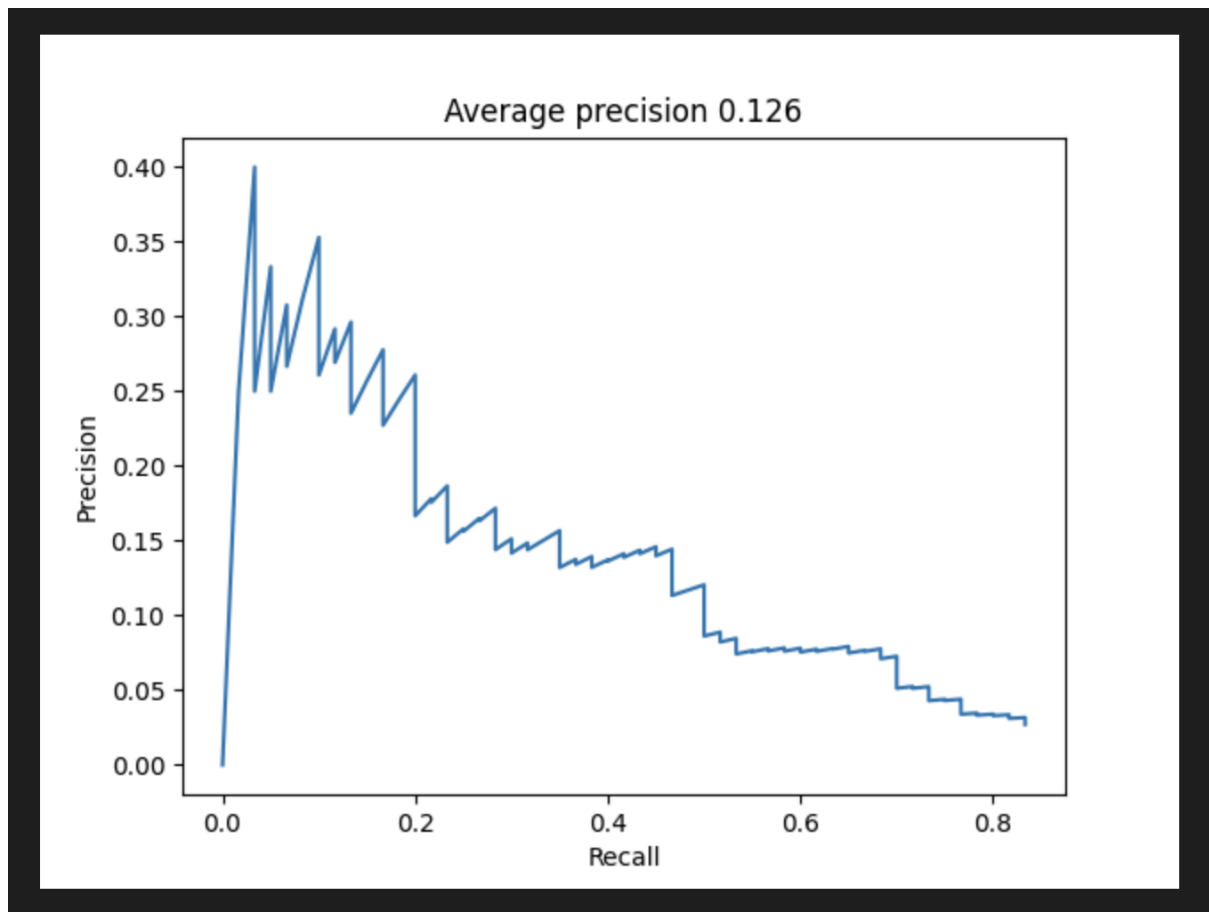
Betty



Fred



Wilma



Aceste rezultate le-am obtinut mutand corespunzator in fiecare fisier de descriptori pozitiv: descriptori pozitiv caracter , iar la descriptorii negativi: descriptori negativi fiecare caracter + descriptori pozitiv celelalte caractere.

Ex barney: Descriptori totali pozitivi: descriptori pozitivi barney ,
descriptori totali negativi: descriptori pozitiv betty, fred, wilma si toti
descriptorii negativi de la fiecare caracter.