

VISIÓ PER COMPUTADOR

Short Project

Facultat d'Informàtica de Barcelona

**Manel Frigola
Joan Climent**

Barcelona, Abril de 2022

1. Objectius del projecte

L'objectiu del projecte és implementar un sistema automàtic per detectar la posició dels ulls implementant un descriptor d'imatges propi. La idea principal del projecte és trobar un bon descriptor d'imatges que funcioni correctament per a un classificador donat, enfront de l'altre estratègia que seria trobar un bon classificador per un descriptor d'imatges donat.

Les imatges amb les que es treballaran seran principalment amb un enquadrament del bust d'un usuari amb rotacions del cap poc apreciables en relació a la càmera. En la figura següent es mostra en vermell la detecció d'ulls que s'espera com a resultat.

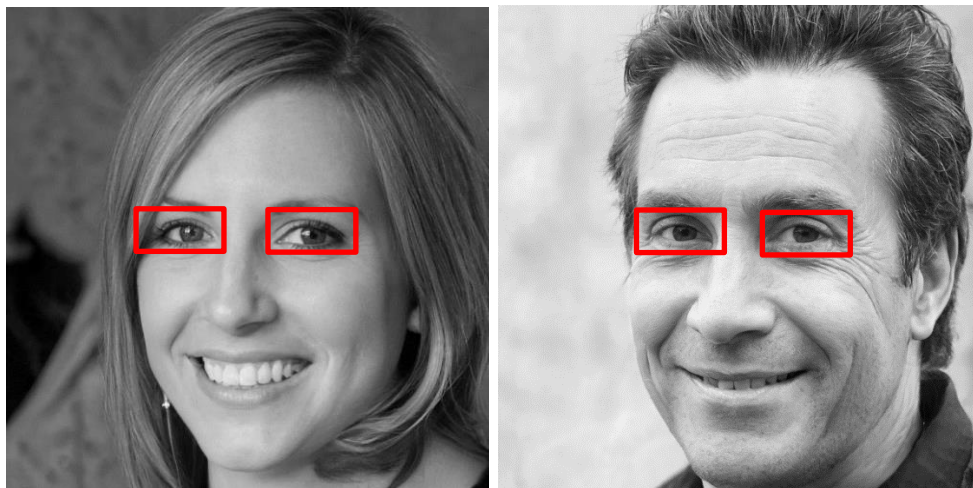


Fig 1. Exemples d'una detecció d'ulls (imatge extreta de thispersondoesnotexist.com)

2. Treball a realitzar

a) Construir un detector de ulls de codi propi seguint les següents passes:

- i) Recopilar una base de dades d'imatges de 300 ulls i 400 no-ulls. Escaleu les mides de les imatges per treballar a una mida estàndard de 48x32 píxels y 256 nivells de gris.
- ii) Programar un o diversos codis que calculin un vector de característiques (anomenat descriptor) d'imatges de nivells de gris. Aquest descriptor serà **d'inspiració pròpia i no poseu usar descriptors d'altres autors, tals com HOG, LBP, Haar, etc.**
- iii) Calcular la taula de descriptors de les imatges de la base de dades de l'apartat i) conjuntament amb l'etiqueta (Ulls / No-Ulls). Desar la taula de descriptors i l'etiqueta en un fitxer format **csv**.
- iv) Entreneu un classificador SVM amb un kernel lineal (vegeu la *Classification Learner App*).
- v) Realitzeu l'estudi estadístic dels resultats del classificador/s d'ulls: matriu de confusió, precisió, etc., que heu calculat en l'apartat a). Utilitzeu l'opció de validació creuada 5 cops per fer el recull estadístic (vegeu les opcions del *Classification Learner App*).

- vi) Calculeu la figura de mèrit que utilitzarem per valorar el detector d'ulls

$$\text{Mèrit} = \sqrt{\left(1 - \frac{\text{Mida del vector de característiques}}{\text{Mida de la Imatge}}\right)^2 + \text{Accuracy}^2}$$

On Accuracy es mesura entre [0 1], la mida de les imatges representa la mida estàndard amb les que s'han calculat les característiques (en el nostre cas 48x32) i la mida del vector de característiques és l'amplada de la taula de característiques amb la que s'ha entrenat el classificador sense fer cap procés estadístic addicional de reducció (tipus PCA o similar).

- b) Mostrar la detecció d'ulls en seqüències de vídeo pròpies o extretes d'internet.

3. Es demana

Un informe didàctic i complet

Cal que lliureu un informe comentant de cadascuna de les etapes amb una descripció del treball realitzat, resultats obtinguts i destacant les parts més rellevants del vostre programa, quins problemes us heu trobat i com els heu resolt. Indiqueu també casos en el que vostre programa no funciona.

En especial l'informe ha de contenir:

- i) Una explicació detallada del mètode d'extracció del vector de característiques. Mida del vector de característiques, rati mida de la imatge vs mida del vector de característiques, temps de càlcul.
- ii) Resultats obtinguts, mostrant el rati d'encerts global, falsos positius i falsos negatius.
- iii) Enumeració de les funcions utilitzades, deixant ben clar quines són les implementades per vosaltres. Tot el software provinent d'altres fonts, ha d'estar correctament referenciat. EN CAS CONTRARI ES CONSIDERARÀ PLAGI.
- iv) L'última secció de l'informe serà un annex amb tot el codi

Presentació del funcionament del programa

Haureu d'executar programa MATLAB que implementi el sistema de reconeixement demanat en aquesta pràctica. Tingueu en compte que el dia de la presentació es provarà el vostre programa de reconeixement amb noves imatges, no incloses en el *dataset*.

4. Indicacions

Les següents indicacions són de caràcter orientatiu.

a) Construcció de la base de dades

Per construir el vostre detector d'ulls haureu de obtenir una base de dades d'imatges d'ulls i no ulls. Per fer fàcilment una BD d'imatges d'ulls, a la web *thispersondoesnotexist.com*, sempre es genera una nova imatge d'una persona (que no existeix) amb la posició dels ulls en les mateixes coordenades. Per construir la BD de no-ulls, per exemple podeu agafar aleatòriament imatges a <https://picsum.photos/>. Trobareu un codi en Python a <https://grins.upc.edu/portals/display/SHAREIT/Home> que obté les imatges de la cara de la web i les emmagatzema. Per obtenir les imatges de no-ulls modifiqueu la subrutina *get_image_from_web* tal com es mostra a continuació:

```
def get_image_from_web():  
  
    # r = requests.get("https://thispersondoesnotexist.com/image").content  
  
    r = requests.get("https://picsum.photos/200/300").content  
  
    return r
```

b) Detecció dels ulls en seqüències de vídeo

Per reduir l'àrea de cerca en seqüències de vídeos, podeu detectar la posició aproximada de la cara de l'usuari fent us de la funció *CascadeObjectDetector* de la llibreria *Computer Vision Toolbox*. En el següent codi (Fig. 3) d'exemple es mostra com es pot utilitzar la variable *Detector* per localitzar la posició de la cara (funció *step*).

```
% Create a cascade detector object.  
Detector = vision.CascadeObjectDetector('FrontalFaceLBP');  
  
% Read a video frame and run the face detector.  
videoReader = VideoReader('face 1.webm');  
  
while hasFrame(videoReader)  
    % get the next frame  
    videoFrame = readFrame(videoReader);  
    bbox = step(Detector, videoFrame);  
  
    % Draw the returned bounding box around the detected face.  
    videoFrame = insertShape(videoFrame, 'Rectangle', bbox);  
    imshow(videoFrame);  
end
```

Fig 3. Codi d'exemple per detectar la cara utilitzant la llibreria *Computer Vision Toolbox*.

5. Valoració del projecte

Els factors que més es valoraran del vostre projecte són:

- Un informe del projecte ben estructurat i complet.
- Comparació del resultat usant diferents descriptors i les seves prestacions, etc.
- Utilització de característiques que permetin una major compressió de la dimensió de la imatge a l'hora que es manté un rati elevat d'encerts en la classificació.
- Avaluar el funcionament del vostre programa amb imatges de test prou variades.