

***Autores:*** Carla Navarro Ramirez i Berta Expósito Mateu

***Assignatura:*** Processament de Senyal, Imatge i Vídeo

***Professor:*** Felipe Lumbreras

**Títol:** Detecció d'emocions per a persones cegues

### **Esquema General i objectius:**

Una de les limitacions que més afecta a la comunitat de persones amb ceguera és la incapacitat de llenguatge no verbal per comunicar-se. L'objectiu principal d'aquest projecte és crea un detector de cares que rebí d'entrada una imatge facial, la processi i analitzi i retorni una emoció com a resultat.

Per la realització d'aquest projecte, enfocarem la solució al problema des de tres punts de vista diferents i compararem els seus resultats a l'entrada d'imatges i la sortida d'emocions corresponent per analitzar quines limitacions incorporen cada model. Aquests tres models de detecció d'imatges són:

- **Bàsic:** model aplicat amb filtrat i descriptors.
- **Intermedi:** permet una etapa intermèdia de detecció de punts interessants de la cara (llavis, nas, ulls, boca) i en base aquesta informació vectorial (coordenades x-y en una malla d'arestes que imposen una veïnat) passar les dades per un classificador que endevini l'emoció.
- **Directe:** xarxa neuronal que de la imatge d'entrada retorni directament l'emoció. Per aquesta xarxa ens caldrà un conjunt d'imatges d'entrenament per aproximar al màxim la resposta correcta i un conjunt de imatges test per comprovar que la xarxa creada funciona correctament (perceptró multicapa).

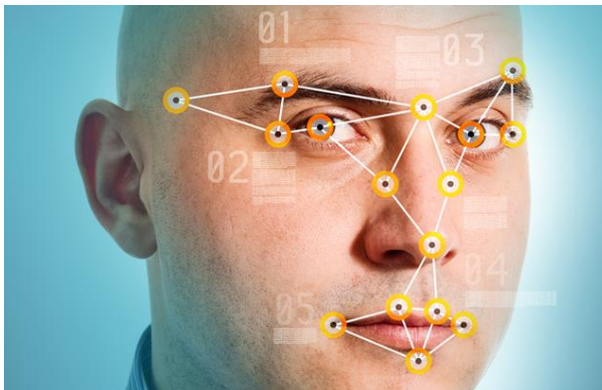
### **Idees pels diferents models**

#### **Model bàsic**

Localització de característiques locals (punts, contorns, regions) que permetin detectar i extraure característiques (algorismes concrets). Aquestes característiques es localitzen en punts interessants de la imatge prenent mesures de magnituds com la intensitat, color i textura al voltant d'aquests punts. Aquestes característiques han de ser invariants a transformacions (rotació, escala i afinitat) per poder detectar emocions en imatges capturades en

qualsevol situació, i robustes (capaces de gestionar imatges amb soroll, compressió, desenfocament,...). També haurem de tenir un gran nombre de característiques per poder detectar amb més precisió una emoció (a més punts interessants i per tant més característiques a comparar, millor serà la cerca de l'emoció correcta).

Els punts que més ens interessin per comparar amb les emocions prefixades són les cantonades. Per detectar aquestes cantonades utilitzarem el model de Harris-Affine, que és invariant a rotació, intensitat, escala i afinitat.



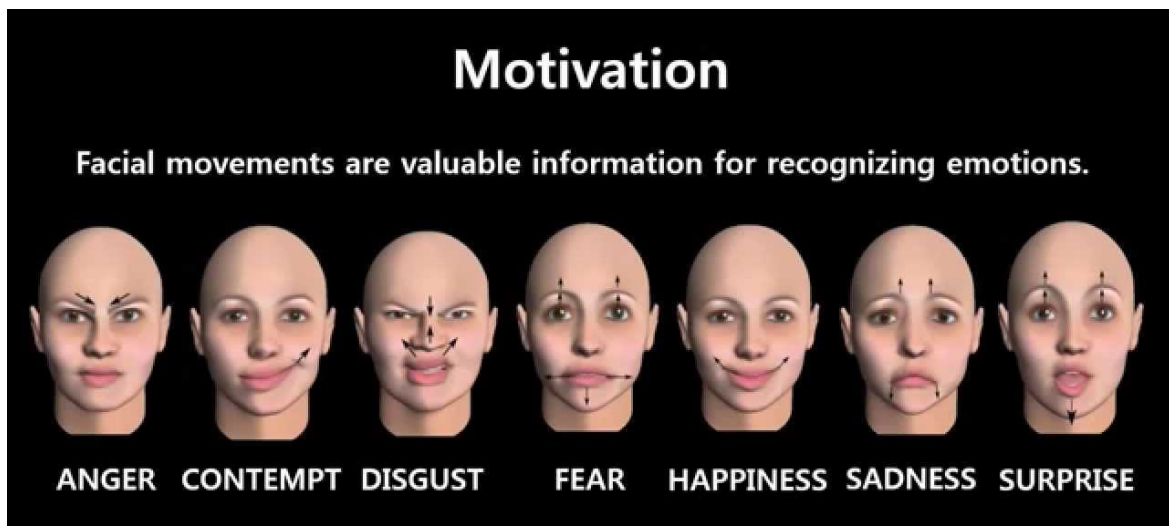
#### Model intermedi

Triarem uns punts d'interès que formin estructures geomètriques i les compararem en el classificador perquè retornin una emoció.

p. e. un diamant en la comissura dels llavis. En funció de la posició dels extrems podrem saber si està content o trist, o si està sorprès o enfadat.

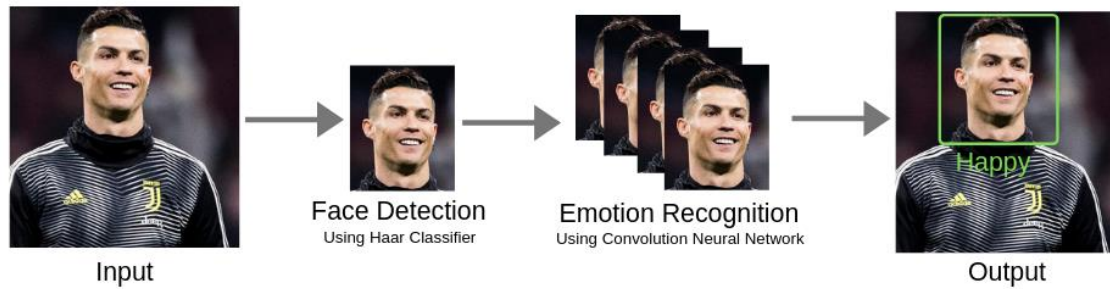
Amb els punts de les celles i l'extrem de l'ull podem saber quina expressió fa amb la part superior de la cara i suposar que està sorprès si les aixeca molt o enfadat i les arronsa molt. També tindrem un estat de neutralitat dins el ventall d'emocions.

#### Ventall d'emocions dels nostres models de reconeixement



#### Model directe

Crearem una Convolutional Neural Network, que depenent de la mida de les imatges d'entrada (número de píxels) realitzarà un seguit de convolucions de les imatges de forma iterativa, on cada convolució aplica un conjunt de kernels diferents i d'aquestes imatges convolucionades farem una reducció de píxels amb (Max-Pooling) agafant el valor més gran de cada sub-matriu de píxels d'aquestes imatges. A aquest conjunt d'imatges de sortida li aplicarem de nou una convolució fins que la mida de píxels ja no ho permeti.



### **Webgrafia**

[https://www.youtube.com/watch?v=rmz1pzfDZUo&list=PLAnA8FVrBI8AWkZmb\\_swwWiF8a\\_52dQ3JQ&index=2](https://www.youtube.com/watch?v=rmz1pzfDZUo&list=PLAnA8FVrBI8AWkZmb_swwWiF8a_52dQ3JQ&index=2) (canal de youtube sobre xarxes neuronals)

<https://www.aprendemachinelearning.com/como-funcionan-las-convolutional-neural-networks-vision-por-ordenador/> (convolution neural network)