

MEMORIA DEL MINIPROYECTO

Reconocimiento de emociones

Arthur Font
María Román
Alicia Carrasco
Àngela Serrano

Índice

Objetivo	3
Contexto	3
Implementación	3
Reconocimiento basado en Haar Cascades:	4
Reconocimiento basado en Fisher, Eigen y LBPH:	4
Reconocimiento basado en gradientes orientados a histogramas (HOG):	4
Funciones personalizadas con detección de puntos de referencia:	5
Reconocedor basado en CNN:	6
Relación con el material visto en clase	6
Posibles extensiones	7
Conclusión	8
Bibliografía	8

Objetivo

Las emociones han estado siempre presentes en los humanos cumpliendo un papel muy importante: nos ayudan a dirigir nuestra conducta y a actuar rápidamente. Una emoción puede explicar a la perfección y sin necesidad de palabras nuestra posición respecto a nuestro entorno, un hecho que puede ser muy útil para múltiples aplicaciones. Fisiológicamente, las emociones organizan rápidamente las respuestas de distintos sistemas biológicos, incluidas las expresiones faciales, los músculos, la voz, la actividad del SNA y la del sistema endocrino. Nuestro principal objetivo es reconocer las emociones mediante expresiones faciales a través de las imágenes y clasificarlas en diversas categorías: felicidad, tristeza, sorpresa, miedo, ira, desprecio y neutral. Estas 7 emociones básicas aparecen durante el desarrollo natural de cualquier persona independientemente del contexto en el que se desarrolle, ya que son procesos relacionados con la evolución y la adaptación.

Contexto

El reconocimiento de las emociones faciales supone la integración de los aspectos físicos relativos a los movimientos de la musculatura facial.

El **Facial Action Coding System (FACS)** es un sistema para denominar movimientos faciales humanos por su apariencia en la cara, basado en un sistema desarrollado originalmente por un anatomista sueco llamado Carl- Herman Hjortsjö. Fue tomado después por Paul Ekman y Wallace V. Friesen, y publicado en 1978. El sistema FACS define un conjunto de **Action units (AUs)**, que son las acciones fundamentales de músculos vistas típicamente cuando se producen las expresiones faciales de una emoción particular. Para reconocer las emociones faciales, se detectan AUs individuales y el sistema clasifica la categoría facial de acuerdo con la combinación de AUs.

Para implementar la detección y clasificación de expresiones faciales hay diferentes enfoques, como veremos en el siguiente apartado.

Implementación

Para resolver este problema, podemos encontrar distintos métodos con los cuales podemos reconocer las expresiones faciales. Primero veremos algunos enfoques clásicos y luego pasaremos al aprendizaje profundo.

Reconocimiento basado en Haar Features:

Quizás el método más antiguo que podría funcionar con Haar Features. Básicamente, este método, también llamado clasificador de Viola Jones, es una técnica de detección de objetos propuesta por **Paul Viola** y **Michael Jones** en **2001**. Es un enfoque basado en el aprendizaje automático en el que se entrena una función en cascada a partir de muchas imágenes positivas y negativas. Luego se usa para detectar objetos en otras imágenes. Este algoritmo necesita muchas imágenes positivas (imágenes de caras) y negativas (imágenes de no caras) para entrenar al clasificador y extraer características de él. Luego se usa para detectar objetos en imágenes.

El uso más popular de estas cascadas es como detector facial, que todavía se utiliza hoy en día, aunque existen mejores métodos disponibles.

En lugar de utilizar la detección de rostros, se puede entrenar para detectar expresiones. Una mejor manera de hacerlo es realizar primero la detección de rostros y luego buscar diferentes características dentro del ROI de la cara, como detectar una sonrisa con esta cascada de detección de sonrisas. También puede entrenar un detector de ceño fruncido, etc.

En comparación con otros métodos, es muy débil y actualmente no se usa para este tipo de problemas, pero lo nombramos ya que se usó en el pasado.

Reconocimiento basado en Fisher, Eigen y LBPH:

El módulo de reconocimiento facial integrado de OpenCV tiene 3 algoritmos de reconocimiento facial diferentes, el reconocedor de rostros Eigenfaces, el reconocedor de rostros Fisherfaces e histogramas de patrones binarios locales (LBPH) Face Recognizer.

Estos algoritmos se basan en el reconocimiento facial, pero pueden extraer algunas características realmente interesantes como componentes principales de las caras e histogramas locales para luego, poder entrenar un clasificador de Machine Learning, por lo que en teoría, pueden ser reutilizados para el reconocimiento de emociones, sólo que, en vez de extraer identidades de las personas, pueden obtener expresiones faciales.

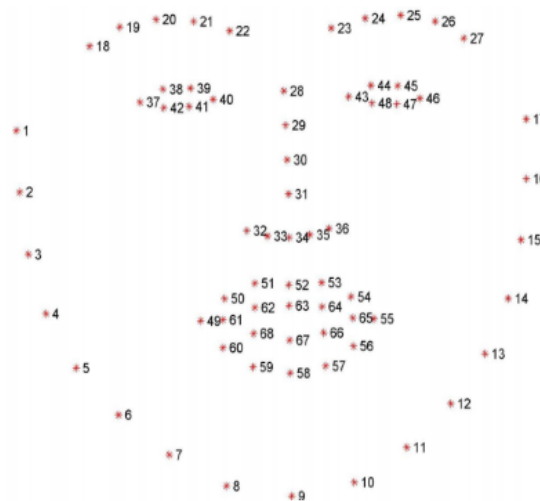
Reconocimiento basado en gradientes orientados a histogramas (HOG):

El detector de rostros HOG es bastante simple de entrenar. El detector de rostros HOG usa una ventana de detección deslizante que se mueve alrededor de la imagen. En cada posición de la ventana del detector, se calcula un descriptor HOG para la ventana de detección. Este descriptor se muestra luego al conjunto de métodos de aprendizaje supervisado entrenado, que lo clasifica como "rostro" o "no rostro". Como en el caso

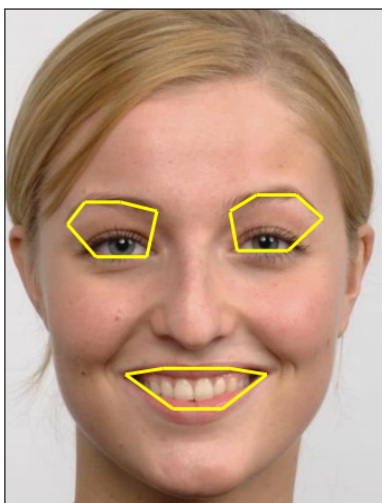
anterior, este método lo podemos reutilizar para que nos encuentre los diferentes componentes de un rostro como la boca, ojos o cejas y los pueda clasificar en distinta emociones, ya que este es muy sensible a la deformación de objetos.

Funciones personalizadas con detección de puntos de referencia:

Una de las formas más fáciles y efectivas de crear un sistema de reconocimiento de emociones es usar un detector de puntos de referencia como el de dlib que le permite detectar 68 puntos de referencia importantes en la cara.



Al usar este detector, se pueden extraer rasgos faciales como ojos, cejas, boca, etc. Puede tomar medidas personalizadas de estos rasgos, como medir la distancia entre los extremos de los labios, para detectar si la persona está sonriendo o no. Del mismo modo, puede medir si los ojos están bien abiertos o no, lo que indica sorpresa o conmoción.



Hay dos formas de hacerlo. Se puede enviar estas medidas personalizadas a un clasificador de Machine Learning y dejar que aprenda a predecir emociones en función de estas medidas o puede utilizar sus propias heurísticas para determinar cuándo llamarlo feliz, triste, etc. las medidas.

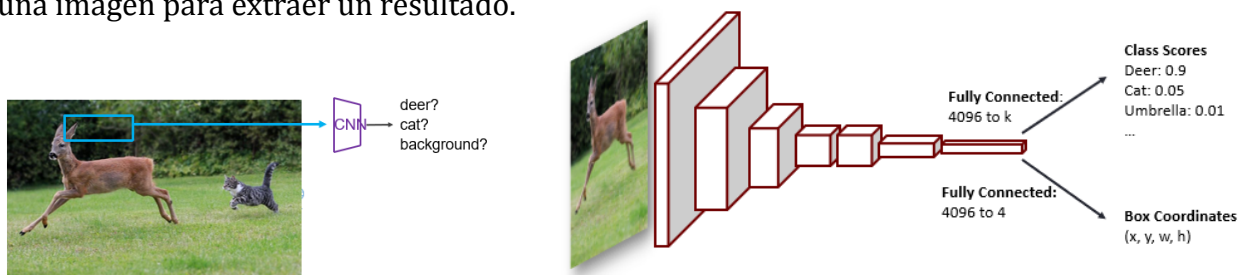
El primer enfoque es más eficaz que el segundo. Pero si solo está definida una emoción singular, como si una persona está sonriendo o no, entonces es más fácil usar la heurística.

Reconocedor basado en CNN:

El enfoque por excelencia para detectar emociones es el basado en el aprendizaje profundo. Para ello, simplemente se debe entrenar una red neuronal convolucional (CNN) con un database que contenga diferentes imágenes de expresión facial (idealmente miles de imágenes para cada clase / emoción). Dado que entrenar un modelo de cero es muy costoso, lo que se suele buscar un modelo ya entrenado con otra tarea y ajustarlo a un problema concreto. Como las características de la primera capa de la CNN, como detección de bordes o esquinas, son similares para todas las imágenes no se tienen que reemplazar por otras capas y se pueden reutilizar en este modelo. Las capas más densas de la red son las que se tienen que entrenar para poder lograr el reconocimiento de emociones.

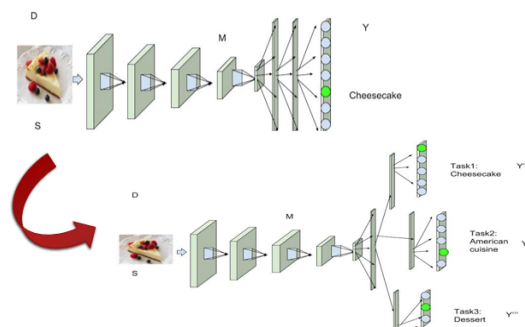
Relación con el material visto en clase

Este uso de la visión artificial lo podemos relacionar con el tema 11, visto nuevamente en clase, que trata sobre la detección de objetos y sus clasificaciones. En este tema vemos cómo a través de redes convolucionales podemos extraer las características de una imagen para extraer un resultado.



Inspirándose también en HOG podemos definir un patrón en la imagen para una mejor comparación con el objetivo, tal como se utiliza en el proceso de reconocimiento de expresiones faciales.

Por ejemplo, en las diapositivas de teoría podemos ver como aplica esta función al reconocimiento de alimentos en las comidas:



También podemos tomar como referencia el tema 8, en el cual se explica cómo podemos implementar la detección facial. Para ello usa el algoritmo Viola & Jones, un algoritmo rápido, de detección en tiempo real y preciso que divide la imagen en un conjunto de ventanas y las clasifica en “cara” o “no cara” basándose en las características locales.

Posibles extensiones

El principal objetivo del reconocimiento de expresiones faciales se centra en la **medicina**, concretamente en la realización de estudios de comportamiento o la detección de enfermedades para su posterior rehabilitación. Algunos ejemplos serían la detección de fase en el alzheimer o medir el grado de dolor en las contracciones de los músculos faciales en pacientes pediátricos operados de apendicitis.

También podemos distinguir otras aplicaciones de esta inteligencia artificial, como serían:

- **Docencia:** reconocimiento expresiones alumnos clase virtual, para así suponer, por ejemplo, el porcentaje de alumnos que está entendiendo la explicación.
- **Marketing y ventas:** analizar por webcam qué emociones transmiten los productos en los potenciales consumidores (sabor, envoltorio, comparación de productos).
- **Política:** sondeos electorales y encuestas de consumo. Según las neurociencias falla la metodología tradicional, porque en el proceso de toma de decisiones la emoción se sitúa al inicio y es lo que impulsa la acción. Las encuestas tradicionales entienden el proceso como frío y racional.
- **Videojuegos:** reconocimiento de emociones dentro de los videojuegos.

Conclusión

Vistas las diferentes formas de implementar el reconocimiento de emociones faciales, creemos que la más eficaz, dadas las diferentes fuentes estudiadas, es la que utiliza redes neuronales convolucionales.

De todas maneras no descartamos que más adelante surjan nuevas formas de implementar esta aplicación y que por tanto sean mejores, ya que como hemos ido viendo a lo largo de esta asignatura, muchos de los algoritmos todavía no funcionan al 100% y tienen potencial para mejorar.

Por lo que a las aplicaciones respecta, creemos que en un futuro, a medida que dicha tecnología avance más, estas aumentarán exponencialmente.

Bibliografía

132088914.pdf. (s. f.). Recuperado 8 de diciembre de 2021, de

<https://core.ac.uk/download/pdf/132088914.pdf>

AI Emotion and Sentiment Analysis With Computer Vision in 2022. (2021, septiembre 26).

Viso.Ai. <https://staging2.viso.ai/deep-learning/visual-emotion-ai-recognition/>

Anwar, T. (s. f.). *Emotion / Facial Expression Recognition with OpenCV. / Bleed AI*. Recuperado

8 de diciembre de 2021, de

<https://bleedai.com/facial-expression-recognition-emotion-recognition-with-opencv/>

Este software puede detectar las emociones de los estudiantes de manera virtual | Video |

CNN. (s. f.). Recuperado 8 de diciembre de 2021, de

<https://cnnespanol.cnn.com/video/software-inteligencia-artificial-guillermo-arduino-clix/>

Las aplicaciones detectando emociones a través de reconocimiento facial. (2016, enero 28).

<https://www.whatsnew.com/2016/01/28/las-aplicaciones-reconociendo-emociones-a-traves-de-reconocimiento-facial/>

Use la red neuronal convolucional (CNN) para el reconocimiento facial—Programador clic. (s.

f.). Recuperado 8 de diciembre de 2021, de

<https://programmerclick.com/article/5903876978/>