

Detección de expresiones faciales

Hoy en día, el aprendizaje profundo es una técnica que se lleva a cabo en muchas aplicaciones y en estudios relacionados con la visión por computadora. En este proyecto se tratan ambos campos para solventar el problema de la detección de emociones en seres humanos.



Resumen del proyecto

Este estudio tiene como objetivo construir un sistema de reconocimiento de expresión facial basado en un modelo de red neural convolucional (CNN), para clasificar automáticamente las expresiones presentadas en la base de datos de FER2013 y AffectNet.

Tecnologías a usar

Para llevar a cabo este proyecto se utilizará el entorno de desarrollo Jupyter Notebook en lenguaje Python. Se necesitarán bibliotecas específicas de Deep Learning como Tensorflow y Keras. Además de librerías para procesamiento de datos como Numpy, Scipy, Matplotlib y Pandas. Se tiene pensado usar si fuera necesario también la biblioteca para visión artificial OpenCV.



** Es posible tener que contratar algún servicio para aumentar el procesamiento*

<https://www.ovhcloud.com/es-es/public-cloud/gpu/>

Datasets

Para entrenar los modelos se podrán utilizar estos dos conjuntos de datos:

AffectNet es un gran conjunto de datos de expresiones faciales con alrededor de 0.4 millones de imágenes etiquetadas manualmente con la presencia de ocho expresiones faciales (neutral, feliz, enfado, tristeza, miedo, sorpresa, disgusto y desprecio)

Fer2013 que contiene aproximadamente 30.000 imágenes de diferentes expresiones y sus etiquetas principales se pueden dividir en (neutral, feliz, enfado, tristeza, miedo, sorpresa, y asco)

Objetivo Principal

El principal objetivo es entrenar un modelo que permita clasificar las 7 emociones que están etiquetadas. Existe mucho margen para mejorar con los estudios que se han realizado en el pasado, ya que los mejores resultados son de un 62% aproximadamente en el dataset de AffectNet.

Posibles problemas y soluciones

El dataset puede que no contenga suficientes datos, pero se utilizarían técnicas de aumentado de datos (Data Augmentation). Es necesario también utilizar técnicas de limpieza y procesamiento de datos con imágenes.

Se va a intentar utilizar ambos datasets en el problema, pero si no funcionase el modelo correctamente existen muchos datasets relacionados con lo cual sería posible modificarlo. Es posible que esta temática no consiga mejorar los resultados obtenidos, si ocurre este caso se intentaría buscar algo que permita diferenciar el proyecto de los que ya estén publicados.

Posibles mejoras

Existen numerosas variaciones que se pueden realizar si una vez entrenado no se consiguiera mejorar los resultados actuales:

- Crear una aplicación que a través de la composición de imágenes analizara un vídeo y realizase gráficas de la emoción de la persona que está siendo grabada.
- Podríamos tapar la zona de los ojos y tratar de analizar el efecto que tendría en nuestro modelo.
- El impacto del uso de mascarillas durante el periodo de pandemia en el reconocimiento de expresiones faciales.
- También se podría comprobar el funcionamiento del modelo añadiendo ruido y oclusiones.
- Otra mejora podría consistir en añadir audio. De esta manera además de tratar con técnicas de procesamiento de imágenes, se aprendería a tratar con procesamiento de audio.