**DuocUC**

Escuela de Informática y Telecomunicaciones



**Identificación De Éxitos**

**“SignAI”**

**Equipo 1**

**Jenniffer Coñuel / Mattias Gonzalez / Ariel Silva**

Capstone 002D

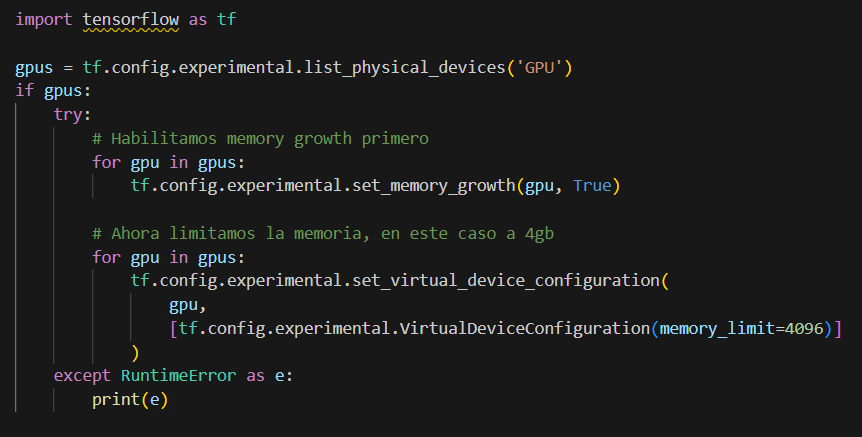
Profesora: Jazna Meza & Juan Pablo Mellado

**Fase 1: ¿Qué se define como éxito?**

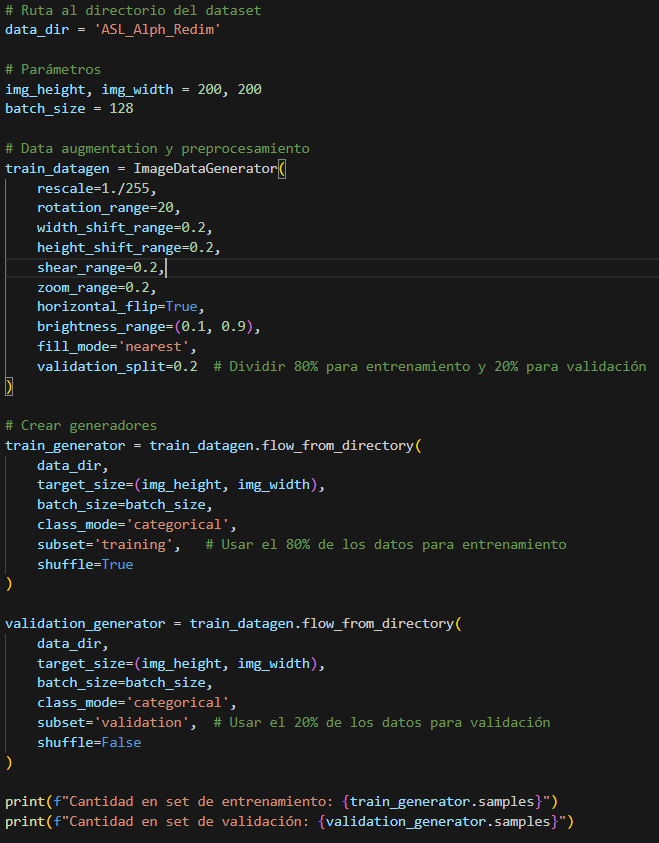
Crear un modelo de redes neuronales que cumpla un mínimo de **80% de accuracy** en set de pruebas y al menos un **50% en casos reales** al momento de identificar las señas con la cámara.

Luego de haber definido estas métricas, se logró crear un modelo que cumple con las métricas propuestas y a continuación se mostrará el paso a paso sobre cómo se llegó a este modelo exitoso.

**Fase 2: ¿Cómo se logró?**

**Limitación de recursos**

El siguiente código limita la cantidad de memoria que usará la tarjeta gráfica (RTX 2060) para reducir los errores por OOM (Out of Memory), pero no elimina completamente la posibilidad

****

Aca se declara el dataset para el entrenamiento, donde se tienen que normalizar, o sea, que todas tengan la misma cantidad de pixeles en la entrada de esta red neuronal, para eso se redimensiona todas las imágenes del dataset a 200x200 y luego se declara que el tamaño de la imagen dentro del modelo sea 200x200 para garantizar que todas las imágenes tengan el mismo tamaño y resolución.

Luego se elige el batch size de 128, este tiene que ser si o si algún número en potencias de 2, para una mejor compatibilidad con NVIDIA, se eligió 128 gracias a varios intentos de ensayo y error, donde 128 era la cantidad justa para que no lance errores de OOM.

Después de finalizar el proyecto, se importa el modelo junto con sus atributos (pesos) a la API para poder clasificar imágenes y videos con el modelo de red neuronal.

**Fase 3: ¿Qué es lo que hace?**

SignAI puede integrarse en cualquier ecosistema que permita la emisión de peticiones HTTP.

Antes de poder utilizar SignAI, cada petición debe contar con una APIKey ligada en sus cabeceras (headers).

**Para voz a señas:**

Una vez se integra SignAI y se envía la petición con un Audio o Vídeo. Se recepciona y se cambia su codificación a .wav o .mp4 dependiendo del módulo utilizado.

Se transcribe el Audio y luego se generará una serie de QUERYS en base al contenido del Audio, posteriormente se obtienen las imágenes de la Base de Datos.

**Para señas a texto:**

Se reducen los FPS del vídeo a 8 y luego se divide el vídeo por frames.

Posteriormente pasamos los frames al modelo y este clasifica las señas.

Una vez se obtienen los resultados de los procesos anteriores, se retorna hasta la aplicación en la cual se integró SignAI en formato JSON.