

El sistema **SignGlasses** se soporta en un entorno software distribuido que contempla tres módulos principales, cada uno con una funcionalidad específica.

1- Programa de Recolección, Entrenamiento y Administración de Señas

Este módulo está diseñado para la gestión y entrenamiento del modelo de aprendizaje automático, ejecutado en una PC. Su objetivo es facilitar la administración del dataset y la actualización del modelo neuronal. Las funcionalidades principales son:

1. Recolección de Nuevas Señas

- Permite registrar nuevas clases de señas. El operador asigna un nombre identificativo (label) y el sistema inicia la captura.
- Recibe vídeos por UDP desde la Raspberry Pi, procesados en tiempo real mediante MediaPipe, extrayendo 21 landmarks por mano (42 puntos en total).
- Se generan vectores de 126 características (3 coordenadas por landmark), muestreados cada 50 ms, alcanzando un total de 5000 frames (aproximadamente 10 minutos de grabación).
- El dataset se almacena en formato “.pkl”, incluyendo los vectores y sus respectivas etiquetas, facilitando la carga posterior sin necesidad de regrabar.

2. Entrenamiento del Modelo

- Se encarga de la normalización del dataset mediante StandardScaler, codificación de etiquetas con LabelEncoder y entrenamiento del modelo.
- El modelo es un clasificador secuencial denso con capas de 64 y 32 neuronas (activación ReLU), Batch Normalization, Dropout y capa de salida Softmax.
- Optimización mediante el algoritmo Adam, guardado en “.h5” y posterior conversión a formato TensorFlow Lite (.tflite) para reducir la carga computacional en inferencia.

3. Listado de Señas

- Visualiza el inventario actual de señas registradas en el dataset junto a sus etiquetas correspondientes.

4. Eliminación de Señas

- Permite eliminar registros específicos del dataset, ya sea por errores de grabación o por bajo rendimiento durante la inferencia.

5. Generación de Audios

- Mediante pyttsx3 y pydub se generan archivos de audio .WAV asociados a cada etiqueta registrada, los cuales serán utilizados para la síntesis de voz en tiempo real.

2- Programa de Evaluación en Tiempo Real

Este programa se ejecuta en la PC y corresponde al entorno final de usuario, enfocado en la inferencia y operación en tiempo real, sin permitir la modificación del modelo ni del dataset.

1. Listado de Señas Cargadas

- Muestra la lista de señas habilitadas para inferencia con sus respectivas etiquetas.

2. Evaluación en Tiempo Real

- Recibe flujos de imágenes en fragmentos JPEG por UDP desde la Raspberry Pi, reconstruyendo los frames con OpenCV.
- Procesa los landmarks con MediaPipe, calcula la posición normalizada de la muñeca derecha y envía comandos al servo de seguimiento.
- Normaliza el vector de 126 características, ejecuta inferencia con TensorFlow Lite, aplicando StandardScaler y LabelEncoder previamente guardados.
- Si la predicción supera un umbral de confianza del 90%, muestra en pantalla la etiqueta correspondiente junto con los landmarks y barras de progreso.
- Envía la etiqueta detectada por UDP a la Raspberry Pi para síntesis de audio, además de gestionar señales de control como "HANDS_DETECTED" o "NO_HANDS" mediante heartbeat para asegurar la robustez del sistema.

3- Programa Implementado en la Raspberry

El script embebido en la Raspberry Pi cumple la función de **punto de comunicación y procesamiento ligero**, realizando tareas de captura y gestión periférica. Debe trabajar en simultaneo con uno de los dos programas de la PC para cumplir con su función:

- Captura vídeo desde la cámara y lo fragmenta en paquetes UDP de **1460 bytes** para transmisión a la PC.
- Adquiere audio desde el micrófono **INMP441** a **48 kHz**, realiza remuestreo a **16 kHz** y lo transmite cada **5 segundos**.
- Reproduce archivos de audio recibidos vía UDP por el altavoz incorporado.
- Recibe ángulos de movimiento para el **servo SG90** y realiza control en tiempo real del seguimiento.
- Muestra la transcripción de voz recibida en la **pantalla OLED** para el usuario sordo.
- Mantiene un sistema de control de estado mediante **heartbeat bidireccional**, asegurando comunicación activa y detección de caídas de conexión.