

# **CUIA 22/23**

## **PROYECTO FINAL:**

### **APP CON HCI & CC**

#### **(ver.1.2)**



Marina Jun Carranza Sánchez  
Computación Ubicua e Inteligencia Ambiental  
Prácticas - Grupo A1  
Grado en Ingeniería Informática

## ÍNDICE

1/ Objetivos.....	3
2/ Propuesta.....	3
3/ Utilidad.....	3
4/ En qué consiste.....	3
5/ Funcionalidad.....	3
6/ Estructura del proyecto.....	4
7/ Librerías usadas.....	4
8/ Instrucciones de uso.....	4

### **PARTE IMPORTANTE: SECCIÓN 8 (INSTRUCCIONES DE USO)**



## 1/ Objetivos

La realización de una o varias aplicaciones en las que se incorpore Interacción Hombre-Máquina implícita (Procesado de lenguaje natural y Realidad Aumentada) y Consciencia de Contexto (Reconocimiento e identificación de imágenes).

## 2/ Propuesta

Aplicación de testeo de dibujos/cuadros/pósters (en cualquier formato de imagen) en Realidad Aumentada, con reconocimiento de voz y facial incluidos.

## 3/ Utilidad

En una escala mayor, podría usarse una versión muy mejorada de esta app para que las instituciones que deben otorgar permiso para hacer graffitis o murales en zonas públicas, y los propios artistas, puedan visualizar con antelación una aproximación del resultado final e impacto de la obra sobre el entorno.

En este caso, debido a las restricciones de software y hardware (idóneamente, sería una app para dispositivos Android, pero se va a realizar para ordenadores, con menor resolución), puede servir para hacerse una idea de cómo quedarían cuadros o pósters en una sala de museo o habitación.

## 4/ En qué consiste

El usuario dispondrá de una galería donde podrá subir imágenes, y a través del uso de la cámara, podrá enfocar y seleccionar una pared o cualquier superficie lisa (preferiblemente de un color homogéneo y definido), que la app reconocerá para poder “pintar” la imagen en cuestión, mostrando cómo quedaría en el entorno real mediante Realidad Aumentada.

## 5/ Funcionalidad

- Interacción Hombre-Máquina implícita
  - Procesado de lenguaje natural → procesamiento de texto que ha sido obtenido mediante el reconocimiento de voz.
  - Realidad Aumentada → el “pintar” las imágenes en superficies de colores lisos mediante la webcam, actualizándose en tiempo real.
- Consciencia de Contexto
  - Reconocimiento e identificación de imágenes → reconocimiento facial para identificar a los usuarios y mostrar sus galerías personales.

## 6/ Estructura del proyecto

- `appInterfaz.py` → todos los aspectos de la interfaz de la app (disposición de las ventanas y sus elementos, funciones relacionadas con la galería, etc). Fichero de ejecución de la app.
- `appCamara.py` → funciones relacionadas con la webcam (superposición del logo, código de detección de superficies).
- `appImages.py` → funciones para cargar imágenes en la galería, reconocer/identificarlas, hacer búsquedas de imagen en internet, etc.
- `appSpeech.py` → reconocimiento de voz para la obtención de texto que será procesado con el fin de llevar a cabo ciertas tareas.
- `appRecFacial.py` → funciones relativas al reconocimiento facial (guardar imágenes de usuarios, detección de caras en la webcam...)
- `/galerias` → donde se almacenan las galerías personales de cada usuario.
- `/faces` → donde se guardan las imágenes de los usuarios registrados.
- `/icons` → carpeta con los iconos de los elementos de la interfaz.

## 7/ Librerías usadas

- `speech_recognition` → Reconocimiento de voz
- `face_recognition` → Reconocimiento facial
- `tkinter (ttk, filedialog, simpledialog...)` → Aspectos de la interfaz
- `PIL (Image, ImageTk)` → Tratamiento de imágenes en tkinter
- `os` → Gestión de diálogos y archivos (lectura, escritura, abrir, guardar, etc)
- `cv2, numpy` → Para la implementación de Realidad Aumentada
- `tensorflow.keras (ResNet50, preprocess_input, decode_predictions...)`  
→ Modelos de reconocimiento e identificación de imágenes
- `requests` → Gestión de peticiones HTTP/HTTPS, consultas del navegador
- `urllib.parse` → Para el formato de los elementos de las urls

## 8/ Instrucciones de uso

***Para iniciar la app, se ejecuta en la terminal: "python .\appInterfaz"***

Primero se lanzará una ventana con dos botones: el de "Eliminar usuarios" abrirá la lista de usuarios registrados, pudiendo eliminar cualquier perfil registrado y el de "Adelante :D", procederá a abrir la webcam para el reconocimiento facial del usuario.

- Si no está registrado, lanzará un diálogo para introducir el nombre del usuario. Automáticamente, se guardará la imagen y se creará su galería.
- Si ya se ha identificado anteriormente, reconocerá al usuario, "recordando" su nombre de usuario y los datos de su galería personal.

A continuación, se abrirá el menú principal, con los dos botones de modo de uso:

- Sin audio: la interacción será a través de botones o eventos del ratón y teclado.
- Con audio (micrófono): se utilizarán comandos de voz para interactuar con la app.

## Modo sin audio

Los botones con los comandos disponibles son:

- 1) Cargar imágenes nuevas en la galería de la app desde el dispositivo, mostrando luego la galería actualizada → Cargar imagen
- 2) Se abre la galería para seleccionar una imagen, de la que se mostrarán predicciones para “identificarla” → Identificar imagen
- 3) Se pide escribir el texto con el que se buscará la imagen en el navegador, descargándola y mostrando luego la galería → Buscar imagen
- 4) Se abre la galería para elegir la imagen, luego se abre la webcam. Para seleccionar dónde mostrarla en realidad aumentada, se hace doble click sobre la superficie deseada (**se recomienda usar una cartulina de color llamativo como azul, rojo, verde, naranja...**), y para cerrar la webcam, se presiona la barra espaciadora → Abrir cámara
- 5) Abrir la galería de la app, con los nombres de los archivos y la posibilidad de eliminar cualquier elemento → Abrir galería
- 6) Volver al menú para poder identificarse de nuevo → Cerrar sesión
- 7) Eliminar los datos del usuario “logueado” → Eliminar usuario

Las ventanas deben cerrarse manualmente en este modo.

## Modo con audio

Se activará tras pulsar el botón del micrófono tachado. Después, se le pueden dar los siguientes comandos u órdenes por voz:

- 1) “Cargar imagen”: se abrirá un diálogo para seleccionar una imagen del dispositivo que se guardará en la galería del usuario (este proceso de selección debe hacerse manualmente). Después, se abrirá la **galería en modo visualización** para poder eliminar las imágenes que se quiera.
- 2) “Identificar imagen”: se abre la **galería en modo selección** para elegir la imagen sobre la que realizar las predicciones.
- 3) “Buscar imagen”: la app vuelve a escuchar al usuario, quien deberá decir el texto que se usará para la búsqueda de la imagen en el navegador, que procederá a descargarse. Luego, se abrirá la **galería en modo visualización**.
- 4) “Abrir cámara”: se abre la **galería en modo selección** para elegir la imagen que se mostrará en realidad aumentada. Una vez abierta la webcam, para seleccionar dónde imprimirla, se debe hacer doble click sobre la superficie homogénea deseada. Para cerrarla, se pulsa la barra espaciadora.
- 5) “Abrir galería”: se abre la **galería en modo visualización**.
- 6) “Cerrar sesión”: volver al menú para poder identificarse de nuevo
- 7) “Eliminar usuario”: eliminar los datos del usuario “logueado”.

Para más detalles sobre el modo de visualización y selección de la galería (usando comandos de voz), se recomienda continuar leyendo la siguiente página.

***Galería en modo visualización:*** se muestra la galería durante 7 segundos, informando al usuario que, una vez pasado ese tiempo, la galería se cerrará y la app escuchará para que el usuario diga en voz alta el nombre de la imagen (sin la extensión) que desea borrar. Si se elimina un fichero, se volverá a abrir la galería y se repetirá el proceso hasta que el usuario diga “Nueva orden”, en cuyo caso podrá indicarle cualquiera de los 7 comandos en la lista anterior para realizar una nueva acción.

***Galería en modo selección:*** se sigue el mismo proceso que el del modo de visualización, solo que en este caso, una vez se haya seleccionado una imagen, llevará a cabo una de las dos acciones posibles: mostrar las predicciones de la imagen durante 7 segundos (“Identificar imagen”) o abrir la webcam para mostrar la imagen en realidad aumentada (“Abrir cámara”). Cuando se cierren las ventanas secundarias, se escuchará la “Nueva orden” del usuario.

Si se tienen dudas de cuándo está escuchando la app, qué texto espera (“¿Nombre del archivo?”, “¿Nueva orden?”) o simplemente se quiere visualizar el texto que reconoce la app, se pueden mirar los mensajes de la terminal.

En este modo, es posible que la app se sobrecargue si se proporcionan comandos o nombres de ficheros no válidos varias veces seguidas, o si se queda escuchando sin recibir texto por mucho rato. En caso de que se quede “Escuchando...” pero no reconozca ningún texto, habrá que reiniciar la app.