

El Reto

Utilizamos Claude.ai para la realización de este proyecto, investigamos brevemente cada una de las APIs, pero nos llamó Google Cloud Vision API, ya que entre sus funcionalidades incluye:

- Reconocimiento de objetos y caras.
- Detección de texto.
- Clasificación de imágenes.
- Detección de contenido explícito.
- Integración sencilla.

Funcionamiento

El programa es un detector de personas que utiliza la API de Google Cloud Vision. Primero configura las credenciales de Google Cloud para acceder a la API, luego activa la cámara web del dispositivo permitiendo al usuario capturar una imagen presionando la tecla espaciadora. Una vez capturada la imagen, esta se envía a Google Cloud Vision para dos tipos de análisis: detección de objetos y etiquetas. El programa busca específicamente la presencia de personas entre los resultados obtenidos, determinando si hay o no un humano en la imagen capturada. Finalmente, muestra los resultados detallados en consola y crea una versión anotada de la imagen donde se dibujan rectángulos verdes alrededor de las personas detectadas y rojos alrededor de otros objetos, guardando esta imagen como "captura_anotada.jpg".

Integrantes

José Morataya	0905-24-7370
----------------------	---------------------

Hizo el programa

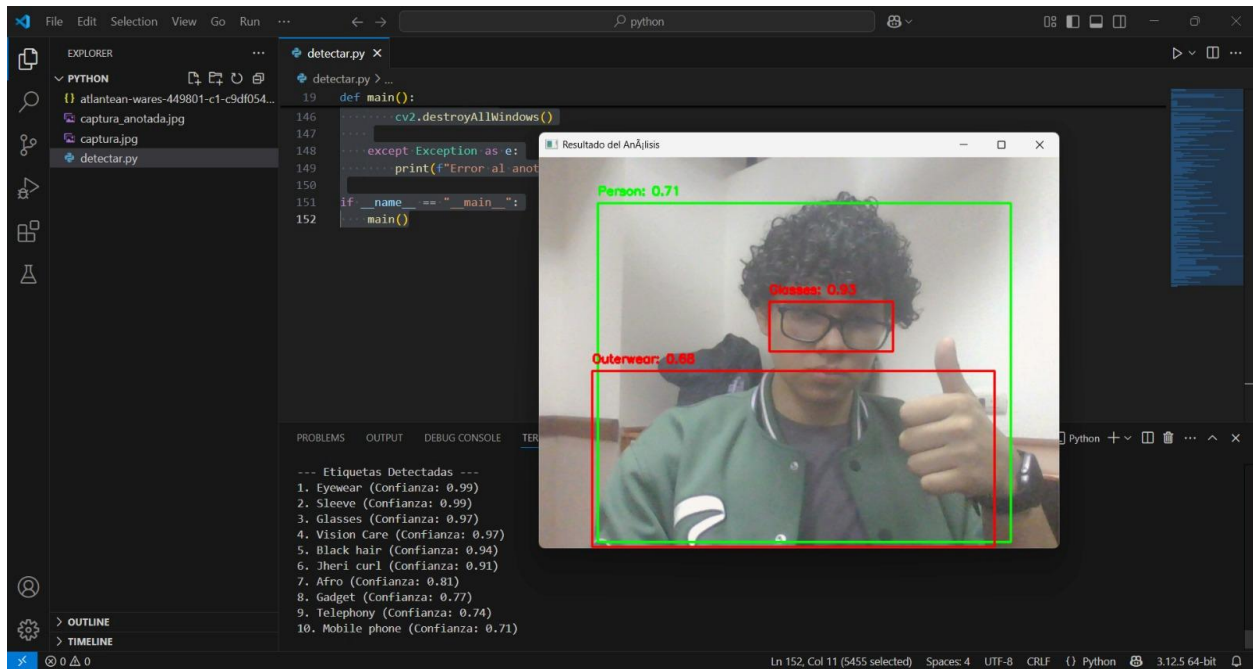
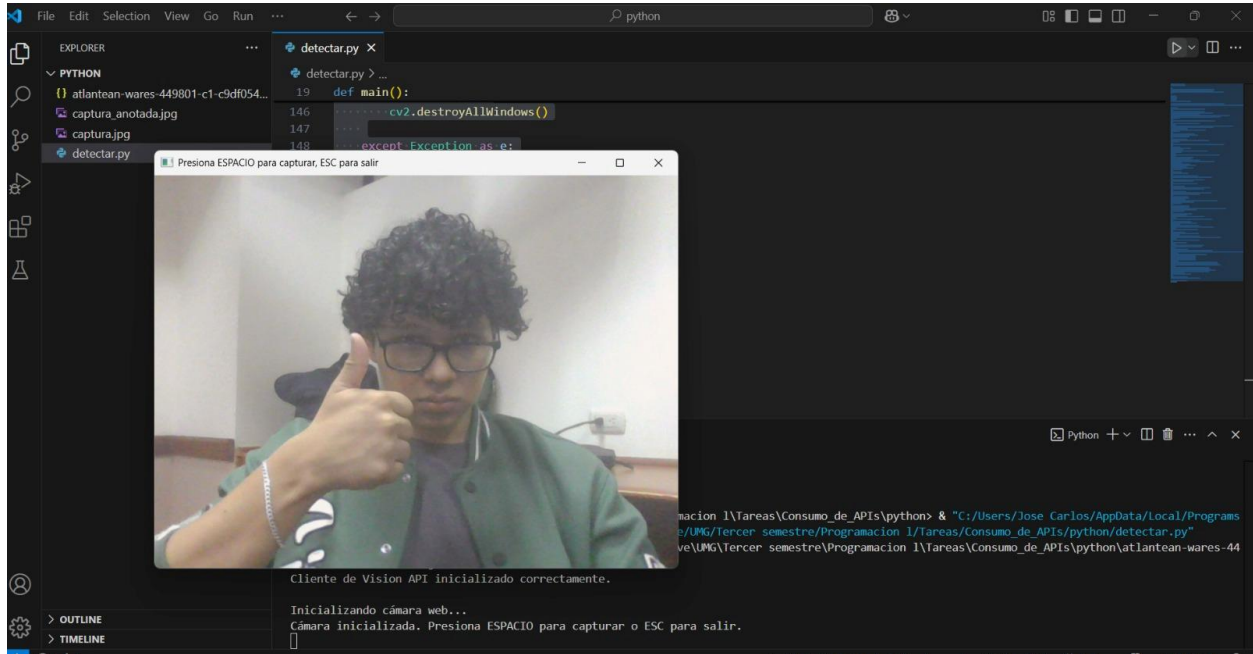
Carlo René Hermógenes	0905-24-7010
------------------------------	---------------------

Documentación

Gerson Jiménez	0905-24-7000
-----------------------	---------------------

Investigo las APIs

Capturas del Funcionamiento



Prompt

Actúa como un desarrollador senior especializado en visión por computadora y automatización con más de 10 años de experiencia en la integración de cámaras web y APIs de inteligencia artificial, con dominio en Python y experiencia trabajando con la Google Cloud Vision API.

Tu objetivo es desarrollar un programa completamente funcional en Python que cumpla con los siguientes requerimientos:

Objetivo General:

Desarrollar un script en Python que capture una imagen desde la cámara web, la envíe a la API de Google Cloud Vision para su análisis, y determine si hay una persona en la imagen. Opcionalmente, también debe describir otros objetos reconocidos.

Requisitos Paso a Paso:

1. Captura de Imagen desde la Webcam:

Utiliza la librería OpenCV (cv2) para acceder a la cámara web.

Permite al usuario capturar una imagen con una tecla o botón.

Guarda la imagen como archivo local temporal (captura.jpg o similar).

2. Subida y Procesamiento con Google Cloud Vision API:

Usa la biblioteca oficial google-cloud-vision en Python.

Carga y codifica la imagen en base64.

Envía la imagen a la API utilizando el método de Object Localization o Label Detection.

Analiza la respuesta para detectar si hay alguna entidad con la etiqueta "Person" o "Human".

3. Interpretación y Visualización del Resultado:

Si se detecta al menos una persona, imprime: "Persona detectada en la imagen."

Si no se detectan personas, imprime: "No se detectó ninguna persona en la imagen."

Además, lista todos los objetos identificados con sus nombres y niveles de confianza (confidence score).

4. (Opcional) Anotación Visual:

Si se detectan objetos, dibuja recuadros alrededor de ellos usando las coordenadas devueltas por la API.

Muestra la imagen anotada en una ventana usando OpenCV.

5. Especificaciones Técnicas:

El programa debe estar escrito completamente en Python 3.

Incluye manejo de errores para fallas en la cámara, errores de conexión con la API, archivos vacíos, etc.

El código debe estar modularizado en funciones (capturar_imagen(), procesar_con_vision_api(), mostrar_resultados(...)).

Incluye comentarios detallados explicando cada sección del código.

6. Resultado Esperado:

Un script ejecutable por terminal o en entorno Jupyter Notebook.

El flujo debe ser automático y fácil de seguir para un usuario con conocimientos básicos de programación.

C:\Users\Jose Carlos\OneDrive\UMG\Tercer semestre\Programacion
\\Tareas\Consumo_de_APIS\python esta es la ruta de las credenciales Json