**UNIVERSIDAD TECNOLÓGICA BOLIVIANA**

**INGENIERÍA DE SISTEMAS**



**Propuesta de un Sistema de Reconocimiento Facial para la Validación de Usuarios de Tarjetas Preferenciales en el Transporte Público**

**Estudiantes:**

* Fabio Camacho Encinas
* Juan Carlos Marca Álvarez
* Fátima Lucia Rocha Rospigliossi

**Materia:** Inteligencia Artificial

**LA PAZ – BOLIVIA**

**2024**

**ÍNDICE GENERAL**

[1. INTRODUCCIÓN AL TEMA 1](#_Toc184625862)

[1.1. Introducción 1](#_Toc184625863)

[2. FUNDAMENTOS TEORICOS 2](#_Toc184625864)

[2.1. Desarrollo 2](#_Toc184625865)

[2.1.1. Inteligencia artificial 2](#_Toc184625866)

[2.1.1.1. ¿Cómo funciona la IA? 2](#_Toc184625867)

[2.1.2. Machine Learning 2](#_Toc184625868)

[2.1.3. AI, Machine Learning y Deep Learning 3](#_Toc184625869)

[2.1.4. Visión artificial 4](#_Toc184625870)

[2.1.4.1. Conceptos Básicos en Visión Artificial 4](#_Toc184625871)

[2.1.4.2. Componentes de un Sistema de Visión Artificia 5](#_Toc184625872)

[2.1.4.3. Aplicaciones de la Visión Artificial 6](#_Toc184625873)

[2.1.5. Librerías de Python 6](#_Toc184625874)

[2.1.5.1. Cv2 (OpenCV): 6](#_Toc184625875)

[2.1.5.2. Mediapipe: 7](#_Toc184625876)

[2.1.5.3. TKinter: 7](#_Toc184625877)

[2.1.5.4. PIL (Pillow): 7](#_Toc184625878)

[2.1.5.5. Image: 7](#_Toc184625879)

[2.1.5.6. Numpy: 7](#_Toc184625880)

[3. DESARROLLO DE LA TEMATICA 8](#_Toc184625881)

[3.1. Introducción 8](#_Toc184625882)

[3.2. Planteamiento del problema 9](#_Toc184625883)

[3.2.1. Identificación del problema 9](#_Toc184625884)

[3.2.2. Planteamiento del problema 9](#_Toc184625885)

[3.3. Objetivos 10](#_Toc184625886)

[3.3.1. Objetivo general 10](#_Toc184625887)

[3.3.2. Objetivos específicos 10](#_Toc184625888)

[3.4. Alcance 10](#_Toc184625889)

[3.5. Límites 10](#_Toc184625890)

[3.6. Desarrollo del sistema 11](#_Toc184625891)

[3.6.1. Variables 11](#_Toc184625892)

[3.6.1.1. MediaPipe Face Mesh y utilidades 11](#_Toc184625893)

[3.6.1.1.1. mp\_face\_mesh 11](#_Toc184625894)

[3.6.1.1.2. mp\_drawing 12](#_Toc184625895)

[3.6.1.2. Variables globales 12](#_Toc184625896)

[3.6.1.2.1. cap 12](#_Toc184625897)

[3.6.1.2.2. captura\_guardada 12](#_Toc184625898)

[3.6.1.2.3. reconocimiento\_activo 13](#_Toc184625899)

[3.6.1.2.4. puntos\_clave\_guardados 13](#_Toc184625900)

[3.6.1.3. Puntos clave críticos 13](#_Toc184625901)

[3.6.1.3.1. PUNTOS\_CRITICOS 13](#_Toc184625902)

[3.6.2. Interfaz gráfica de usuario 14](#_Toc184625903)

[3.6.2.1. Ventana principal 15](#_Toc184625904)

[3.6.2.2. Sección de botones 15](#_Toc184625905)

[3.6.2.3. Visualización de video 20](#_Toc184625906)

[3.6.2.4. Bucle principal 20](#_Toc184625907)

[3.6.3. Funciones 21](#_Toc184625908)

[3.6.3.1. Función encender\_webcam 21](#_Toc184625909)

[3.6.3.2. Función apagar\_webcam 22](#_Toc184625910)

[3.6.3.3. Función guardar\_captura 23](#_Toc184625911)

[3.6.3.4. Función comparar\_captura 25](#_Toc184625912)

[3.6.3.5. Función comparar\_captura 27](#_Toc184625913)

[3.6.3.6. Función comparar\_puntos\_clave 29](#_Toc184625914)

[3.6.3.7. Función iniciar\_reconocimiento 31](#_Toc184625915)

[3.6.3.8. Función detener\_reconocimiento 33](#_Toc184625916)

[3.6.3.9. Función actualizar\_frame 35](#_Toc184625917)

[3.6.3.10. Función actualizar\_frame 38](#_Toc184625918)

[4. RECOMENDACIONES Y CONCLUSIONES 40](#_Toc184625919)

[4.1. Recomendaciones 40](#_Toc184625920)

[4.2. Conclusiones 41](#_Toc184625921)

[Bibliografía 41](#_Toc184625922)

**INDICE DE ILUSTRACIONES**

[Ilustración 1 Captura de pantalla de las librerías utilizadas 6](#_Toc184367075)

[Ilustración 2 Captura de pantalla de las variables utilizadas 11](#_Toc184367076)

[Ilustración 3 Captura de pantalla del código para la interfaz grafica 14](#_Toc184367077)

[Ilustración 4 Captura de pantalla: Encender webcam 16](#_Toc184367078)

[Ilustración 5 Captura de pantalla: Guardar captura 17](#_Toc184367079)

[Ilustración 6 Captura de pantalla: Comparar captura 18](#_Toc184367080)

[Ilustración 7 Captura de pantalla: Comparación exitosa en tiempo real 19](#_Toc184367081)

[Ilustración 8 Captura de pantalla: Detener reconocimiento 20](#_Toc184367082)

[Ilustración 9 Captura de pantalla Fucion encender\_webcam 21](#_Toc184367083)

[Ilustración 10 Captura de pantalla: Función apagar\_webcam 22](#_Toc184367084)

[Ilustración 11 Captura de pantalla: Función guardar\_captura 23](#_Toc184367085)

[Ilustración 12 Captura de pantalla: Función comparar\_captura 25](#_Toc184367086)

[Ilustración 13 Captura de pantalla: Función comparar\_captura 27](#_Toc184367087)

[Ilustración 14 Captura de pantalla: Funcion comparar\_puntos\_clave 29](#_Toc184367088)

[Ilustración 15 Captura de pantalla: Función iniciar\_reconocimiento 31](#_Toc184367089)

[Ilustración 16 Captura de pantalla: Funcion detener\_reconocimiento 33](#_Toc184367090)

[Ilustración 17 Captura de pantalla: Función actualizar\_frame 35](#_Toc184367091)

[Ilustración 18 Captura de pantalla: Función actualizar\_frame 38](#_Toc184367092)

[Ilustración 19 Captura de pantalla: Funcion encender\_webcam 39](#_Toc184367093)

**INDICE DE FIGURAS**

[Figura 1 Inteligencia Artificial 3](#_Toc184367360)

[Figura 2 Utilización tarjetas de servicio publico 9](#_Toc184367361)

# INTRODUCCIÓN AL TEMA

# Introducción

La visión por computadora, también conocida como visión artificial, es una disciplina que busca dotar a los sistemas computacionales de la capacidad de interpretar y analizar imágenes de forma automática. Este campo combina conceptos avanzados de óptica, geometría, procesamiento digital de imágenes y algoritmos, permitiendo a los sistemas extraer información significativa del entorno tridimensional a través de imágenes bidimensionales. Según lo expuesto en Conceptos y Métodos en Visión por Computador, esta tecnología ha evolucionado significativamente desde su uso inicial en aplicaciones industriales hacia áreas más diversas como la seguridad, la automatización de vehículos y la interacción hombre-máquina. (Comité Español de Automática, 2016)

Una de las bases fundamentales de la visión por computadora es la representación digital de imágenes, las cuales son tratadas como matrices numéricas que contienen valores que representan niveles de intensidad de luz o color. Este proceso se inicia con la captura de imágenes utilizando sensores que transforman señales electromagnéticas en datos digitales, los cuales luego son procesados para extraer características relevantes como bordes, texturas y formas. (Comité Español de Automática, 2016).

En la actualidad, los avances en cámaras, sensores y algoritmos han ampliado las capacidades de esta disciplina. Por ejemplo, aplicaciones como el reconocimiento facial o la conducción autónoma utilizan técnicas complejas para identificar patrones, clasificar objetos y tomar decisiones en tiempo real. Estas capacidades demuestran el impacto significativo de la visión por computadora en la resolución de problemas del mundo real, mejorando procesos industriales, servicios de seguridad y experiencias cotidianas de los usuarios. (Comité Español de Automática, 2016).

Este proyecto, centrado en el reconocimiento facial como una aplicación práctica de la visión por computadora, ilustra cómo esta tecnología puede aplicarse para abordar desafíos específicos, como la autenticación biométrica en sistemas de transporte público. Al adoptar principios y métodos descritos en Conceptos y Métodos en Visión por Computador, el proyecto aprovecha herramientas modernas para demostrar el potencial de la visión por computadora en escenarios de alta relevancia social.

# FUNDAMENTOS TEORICOS

# Desarrollo

# Inteligencia artificial

La inteligencia artificial es un campo de la ciencia relacionado con la creación de computadoras y máquinas que pueden razonar, aprender y actuar de una manera que normalmente requeriría inteligencia humana o que involucra datos cuya escala excede lo que los humanos pueden analizar.

La IA es un campo amplio que incluye muchas disciplinas, como la informática, el análisis y la estadística de datos, la ingeniería de hardware y software, la lingüística, la neurociencia y hasta la filosofía y la psicología.

A nivel operativo para el uso empresarial, la IA es un conjunto de tecnologías que se basan principalmente en el aprendizaje automático y el aprendizaje profundo, que se usan para el análisis de datos, la generación de predicciones y previsiones, la categorización de objetos, el procesamiento de lenguaje natural, las recomendaciones, la recuperación inteligente de datos y mucho más. (GoogleCloud, 2024)

# ¿Cómo funciona la IA?

Si bien los detalles varían según las diferentes técnicas de IA, el principio central gira en torno a los datos. Los sistemas de IA aprenden y mejoran a través de la exposición a grandes cantidades de datos, lo que permite identificar patrones y relaciones que las personas pueden pasar por alto.

Este proceso de aprendizaje suele implicar algoritmos, que son conjuntos de reglas o instrucciones que guían el análisis y la toma de decisiones de la IA. En el aprendizaje automático, un subconjunto popular de la IA, los algoritmos se entrenan con datos etiquetados o no etiquetados para hacer predicciones o categorizar información.

El aprendizaje profundo, una especialización adicional, utiliza redes neuronales artificiales con varias capas para procesar información, imitando la estructura y la función del cerebro humano. A través del aprendizaje y la adaptación continuos, los sistemas de IA se vuelven cada vez más hábiles en la realización de tareas específicas, desde el reconocimiento de imágenes hasta la traducción de idiomas y más. (GoogleCloud, 2024)

# Machine Learning

El aprendizaje máquina o Machine Learning (ML) provee a los computadores la capacidad de encontrar patrones y responder a preguntas gracias a esos patrones encontrados (o aprendidos). Idealmente, estos sistemas, además, son capaces de responder a preguntas que no se han encontrado antes, es decir, que no fueron las utilizadas durante el entrenamiento.

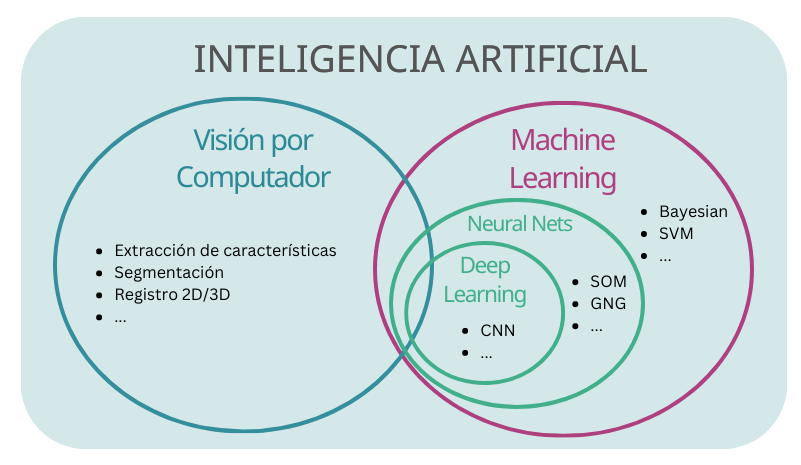
El ML puede utilizarse en numerosas situaciones como encontrar tendencias en el mercado de valores, predicciones climatológicas, o ayuda a expertos en el diagnóstico médico. Los algoritmos aprenden de forma supervisada (conociendo la respuesta esperada), no supervisada (de forma autónoma) o semi-supervisada, y pueden estar basados en diversos tipos, como las redes neuronales, arboles de decisión, algoritmos genéticos, etc. (Marcelo Saval, 2024).

# AI, Machine Learning y Deep Learning

Actualmente hay mucha confusión sobre la diferencia entre Machine Learning y Deep Learning, y donde encajan estas en el área de conocimiento de Inteligencia Artificial (IA, o AI del inglés Artificial Intelligence).

Inteligencia Artificial se entiende como el conjunto de métodos que dota a un computador de la capacidad de resolver problemas de forma inteligente o autónoma sin que sea el usuario el que tenga que darle las indicaciones precisas de como alcanzar la solución. La IA incluye multitud de sub-ámbitos, como el ML, la Visón por Computador, Procesamiento de Lenguaje Natural, etc.

**Figura 1 Inteligencia Artificial**



**Fuente:** (Marcelo Saval, 2024)

# Visión artificial

La visión artificial o visión por computadora es un campo de la inteligencia artificial que se encarga de desarrollar sistemas capaces de interpretar y comprender imágenes y videos para realizar tareas que, hasta ahora, solo los humanos podían llevar a cabo con su visión. La visión artificial permite que las computadoras "vean" y "entiendan" el mundo visual que las rodea, utilizando datos visuales provenientes de imágenes o secuencias de video. Esta disciplina abarca un conjunto de métodos y tecnologías que buscan emular o incluso superar las capacidades visuales humanas en ciertas tareas específicas, como la clasificación, detección y segmentación de objetos.

El objetivo principal de la visión artificial es permitir que las máquinas puedan interpretar y extraer información significativa de las imágenes, realizar tareas como el reconocimiento de patrones y tomar decisiones en función de esos datos visuales. En este marco, se combinan conocimientos de diferentes áreas, como la computación, el procesamiento de señales, la teoría de imágenes y el aprendizaje automático.

# Conceptos Básicos en Visión Artificial

Para comprender la visión artificial, es necesario familiarizarse con algunos conceptos clave:

* **Imagen Digital**: Una imagen digital es una representación bidimensional de información visual compuesta por píxeles, que son unidades discretas de información visual (cada píxel tiene un valor de color o intensidad).
* **Procesamiento de Imágenes**: Consiste en aplicar una serie de técnicas matemáticas y computacionales para transformar, analizar y extraer información útil de una imagen digital. Algunos procesos comunes incluyen el filtrado, la mejora de contrastes, la detección de bordes y la segmentación.
* **Segmentación**: La segmentación de una imagen es el proceso de dividir una imagen en regiones o partes de interés. En visión artificial, se busca separar los objetos dentro de una escena para facilitar su análisis.
* **Características Visuales**: Las características visuales son atributos importantes de una imagen que permiten identificar patrones o objetos. Algunas características comunes incluyen el color, la textura, la forma y el tamaño.
* **Reconocimiento de Patrones**: Es el proceso mediante el cual un sistema de visión artificial identifica patrones o estructuras en las imágenes para clasificarlas, reconocer objetos, personas o acciones específicas.
* **Modelos de Aprendizaje Automático**: En la visión artificial moderna, se utilizan modelos de machine learning (aprendizaje automático) y Deep learning (aprendizaje profundo), particularmente redes neuronales convolucionales (CNNs), para enseñar a las máquinas a reconocer patrones y realizar tareas complejas basadas en grandes volúmenes de datos visuales.

# Componentes de un Sistema de Visión Artificial

Un sistema de visión artificial consta de varios componentes fundamentales que interactúan entre sí para lograr una interpretación efectiva de las imágenes:

* **Captura de Imágenes**: El primer paso en cualquier sistema de visión artificial es la adquisición de imágenes. Esto se realiza mediante cámaras o sensores que convierten la información visual en datos digitales que pueden ser procesados por el sistema.
* **Preprocesamiento de Imágenes**: Las imágenes adquiridas suelen estar afectadas por ruido, distorsiones o malas condiciones de iluminación. El preprocesamiento mejora la calidad de las imágenes mediante técnicas como el filtrado, la corrección de color y la normalización.
* **Segmentación de Imágenes**: Una vez preprocesada la imagen, se realiza la segmentación, es decir, el proceso de dividir la imagen en regiones más pequeñas y manejables. Esto facilita la identificación de objetos o áreas de interés, como el fondo y los elementos que se desean analizar.
* **Extracción de Características**: La extracción de características es el proceso de identificar elementos clave dentro de una imagen (bordes, texturas, formas, etc.). Las características extraídas sirven como la base para la clasificación de objetos, el reconocimiento de patrones y otras tareas.
* **Reconocimiento y Clasificación**: Una vez que las características relevantes se han extraído de la imagen, los algoritmos de clasificación, como las redes neuronales profundas, se utilizan para identificar o clasificar los objetos presentes en la imagen. Esto permite asignar una etiqueta o categoría a los objetos identificados.
* **Toma de Decisiones y Acción**: En algunos casos, el sistema de visión artificial no solo tiene que identificar los objetos, sino también tomar decisiones basadas en esa información. Por ejemplo, en sistemas de conducción autónoma, el sistema debe identificar obstáculos y decidir cómo actuar (frenar, desviar, etc.).

# Aplicaciones de la Visión Artificial

* La visión artificial tiene un amplio espectro de aplicaciones en diversos campos. Algunas de las más destacadas son:
* **Reconocimiento Facial**: Se utiliza en sistemas de seguridad y verificación de identidad, como en el desbloqueo de dispositivos móviles y la autenticación en entornos bancarios.
* **Vehículos Autónomos**: Los coches autónomos utilizan visión artificial para reconocer señales de tráfico, detectar obstáculos y entender el entorno de conducción.
* **Medicina**: En la medicina, la visión artificial se emplea para analizar imágenes médicas, como radiografías y resonancias magnéticas, para ayudar en el diagnóstico de enfermedades y la planificación de tratamientos.
* **Robótica**: Los robots utilizan visión artificial para percibir su entorno y realizar tareas como la manipulación de objetos, la navegación y la interacción con seres humanos.
* **Inspección Industrial**: En el ámbito industrial, los sistemas de visión artificial se emplean para detectar defectos en productos fabricados, asegurando la calidad en la producción.
* **Reconocimiento de Texto**: La visión artificial también se usa en el reconocimiento óptico de sus principales caracteres para convertir texto impreso en texto digital, lo que facilita la digitalización de documentos.

# Librerías de Python

**Ilustración 1 Captura de pantalla de las librerías utilizadas**



**Fuente: Elaboración propia**

# Cv2 (OpenCV):

Cv2 es el módulo de OpenCV para Python. OpenCV es una librería de computación visual para el procesamiento de imágenes en Python. Esta biblioteca proporciona herramientas para realizar operaciones de procesamiento de imágenes, como el filtrado, la detección de bordes, el reconocimiento de características, el seguimiento de objetos, etc. Estas herramientas nos permiten desarrollar aplicaciones de visión artificial, como el reconocimiento facial, el seguimiento de objetos, etc. (Imagina, 2024)

# Mediapipe:

Es un entorno de código abierto para construir aplicaciones que realicen inferencia de visión por computadora sobre datos sensoriales como video, imágenes o audio.

Fue creado para equipos de inteligencia artificial, aprendizaje automático y desarrolladores de software con aplicaciones listas para producción o prototipos como parte de trabajos de investigación.

El entorno Mediapipe permite implementar soluciones para detección facial, manos, gestos, iris, etc. (viso.ai, 2024)

# TKinter:

Es una biblioteca de lenguaje de programación Python que se utiliza para la creación y desarrollo de aplicaciones de escritorio. Esta librería proporciona soluciones para el posicionamiento de interfaces graficas de usuario en aplicaciones de escritorio.

# PIL (Pillow):

Es una biblioteca de procesamiento de imágenes en Python que proporciona funcionalidades para abrir, manipular y guardar imágenes en varios formatos. Permite realizar tareas como: (pillow, 2024)

* Cargar imágenes desde archivos.
* Modificar el tamaño y la resolución de una imagen.
* Aplicar filtros y efectos a las imágenes.
* Realizar operaciones de recorte y rotación.
* Guardar imágenes en diferentes formatos, como JPEG, PNG, BMP, etc.

# Image:

El módulo de Image proporciona una clase con el mismo nombre que se utiliza para representar una imagen PIL. El módulo también proporciona una serie de funciones de fábrica, incluidas funciones para cargar imágenes desde archivos y crear nuevas imágenes.

# Numpy:

Es una librería de Python especializada en el cálculo numérico y el análisis de datos, especialmente para un gran volumen de datos.

Incorpora una nueva clase de objetos llamados arrays que permite representar colecciones de datos de un mismo tipo en varias dimensiones, y funciones muy eficientes para su manipulación.

La ventaja de Numpy frente a las listas predefinidas en Python es que el procesamiento de los arrays se realiza mucho más rápido (hasta 50 veces más) que las listas, lo cual la hace ideal para el procesamiento de vectores y matrices de grandes dimensiones. (ax, 2024)

# DESARROLLO DE LA TEMATICA

# Introducción

En un mundo donde la movilidad urbana se ha convertido en un factor esencial para la calidad de vida, el transporte público debe ser accesible y equitativo para todos. Las tarifas preferenciales son un claro ejemplo de políticas inclusivas que buscan beneficiar a sectores vulnerables como las personas mayores, estudiantes y personas con discapacidad. Sin embargo, estas iniciativas, diseñadas con un propósito noble, se enfrentan a desafíos significativos derivados del abuso y la suplantación de identidad. Este problema no solo impacta las finanzas de los sistemas de transporte, sino que también erosiona la confianza de los usuarios legítimos en el sistema.

El avance de la tecnología ha abierto puertas a soluciones innovadoras que podrían transformar la gestión de estos beneficios. Entre ellas, el reconocimiento facial destaca como una herramienta poderosa y confiable. Esta tecnología biométrica, que permite identificar de manera única a las personas mediante el análisis de sus rasgos faciales, representa un cambio paradigmático en la forma en que se autentica la identidad en contextos cotidianos. Su aplicación en el transporte público no solo garantiza la seguridad y la eficiencia, sino que también redefine la experiencia del usuario, eliminando la necesidad de controles manuales engorrosos.

El presente proyecto propone un sistema de reconocimiento facial como solución a la problemática de la suplantación de identidad en el uso de tarjetas preferenciales. Mediante el desarrollo de un software accesible y funcional, basado en tecnologías de código abierto como Python, OpenCV y Mediapipe, se busca implementar un mecanismo que valide automáticamente la identidad del usuario en tiempo real. Este sistema no solo refuerza la integridad del programa de tarifas preferenciales, sino que también contribuye a la optimización operativa de los servicios de transporte.

En esencia, este proyecto se erige como un puente entre la innovación tecnológica y la responsabilidad social, demostrando cómo el reconocimiento facial puede ser una herramienta clave para garantizar la equidad en el acceso a recursos públicos. Más allá de un simple sistema, esta iniciativa plantea un modelo replicable para la modernización de políticas inclusivas, posicionándose como un ejemplo de cómo la tecnología puede ser un motor para el cambio positivo.

# Planteamiento del problema

# Identificación del problema

Los sistemas de transporte público con tarifas preferenciales enfrentan problemas de suplantación de identidad y uso indebido de las tarjetas destinadas a grupos específicos de usuarios. Esto resulta en pérdidas económicas y una desventaja para los usuarios que realmente necesitan el servicio. Los métodos tradicionales de control, como revisiones manuales, son ineficaces y poco prácticos en escenarios de alta demanda. Por lo tanto, es necesario un sistema automatizado que autentifique la identidad de los usuarios preferenciales de manera rápida y precisa para evitar abusos.

**Figura 2 Utilización tarjetas de servicio publico**



**Fuente:** (Pariste, 2024)

# Planteamiento del problema

¿Qué solución se propone para abordar el problema de la suplantación de identidad en el uso de tarjetas preferenciales en el transporte público?

# Objetivos

# Objetivo general

Desarrollar un sistema de reconocimiento facial para validar la identidad de los usuarios preferenciales en los servicios de transporte público para garantizar el buen uso del medio de transporte utilizado.

# Objetivos específicos

* Analizar las necesidades y características del sistema de reconocimiento facial para identificar los requerimientos funcionales que aseguren su efectividad en la validación de identidad
* Diseñar un módulo de reconocimiento facial en tiempo real que integre técnicas de procesamiento de imágenes para autenticar usuarios de manera eficiente
* Desarrollar una interfaz gráfica intuitiva que permita a los operadores y administradores del transporte interactuar con el sistema de manera sencilla y práctica
* Evaluar la precisión y funcionalidad del sistema mediante pruebas en tiempo real, verificando su capacidad para prevenir suplantaciones y garantizar un desempeño confiable.

# Alcance

* Implementará un reconocimiento facial en tiempo real utilizando técnicas accesibles con Python y bibliotecas como OpenCV y Mediapipe.
* Funcionará como una herramienta demostrativa sin depender de bases de datos para almacenamiento o recuperación de registros.
* Proveerá alertas en caso de detección de discrepancias, permitiendo a los operadores tomar medidas inmediatas.

# Límites

* El sistema no integrará un manejo de bases de datos para almacenamiento masivo o gestión de múltiples usuarios.
* La precisión del reconocimiento puede verse afectada en condiciones adversas, como iluminación insuficiente o ángulos de captura desfavorables.
* No se considerarán métodos avanzados de aprendizaje profundo para el reconocimiento facial; se utilizarán técnicas básicas accesibles con herramientas de Python.
* Se presenta un prototipo funcional y no es apto para su implementación directa en redes de transporte a gran escala.

# Desarrollo del sistema

# Variables

**Ilustración 2 Captura de pantalla de las variables utilizadas**



**Fuente: Elaboración propia**

# MediaPipe Face Mesh y utilidades

# mp\_face\_mesh

* + **Tipo:** Objeto de *MediaPipe* *Face* *Mesh*.
  + **Descripción:**  
    Inicializa el modelo de *MediaPipe* para la detección y seguimiento de puntos faciales.  
    Configuración de parámetros:
    - *min\_detection\_confidence*: Nivel mínimo de confianza para detectar un rostro (0.2 significa que al menos el 20% de probabilidad es necesario para considerarlo válido).
    - *min\_tracking\_confidence*: Nivel mínimo de confianza para rastrear puntos faciales tras la detección inicial.
    - El objeto permite procesar imágenes y extraer coordenadas faciales en tiempo real.

# mp\_drawing

* + **Tipo:** Utilidad de *MediaPipe* *Drawing*.
  + **Descripción:**  
    Herramienta que facilita la visualización de puntos clave y conexiones faciales en imágenes o videos procesados.
    - Proporciona funciones predefinidas para dibujar puntos y mallas en imágenes, lo que es útil para depuración o demostración.

# Variables globales

# cap

* + **Tipo:** Objeto de *OpenCV* *VideoCapture*.
  + **Descripción:**  
    Representa la conexión con la cámara web o dispositivo de captura de video.
    - Inicia con un valor *None* hasta que se llame a la función *encender*\_*webcam* para inicializarlo.
    - Se usa para capturar cuadros (fotogramas) del flujo de video en tiempo real.

# captura\_guardada

* + **Tipo:** Variable de imagen (*ndarray* de *NumPy* o similar).
  + **Descripción:**  
    Almacena una captura de pantalla o un cuadro procesado del flujo de video para referencia posterior.
    - Se utiliza como punto de comparación en el proceso de reconocimiento facial.

# reconocimiento\_activo

* + **Tipo:** Booleano (bool).
  + **Descripción:**  
    Indica si el reconocimiento facial en tiempo real está activado o no.
    - Valor inicial: False.
    - Cambia a True cuando se llama a la función *iniciar*\_*reconocimiento* y a False al detener el proceso mediante la función *detener*\_*reconocimiento*.

# puntos\_clave\_guardados

* + **Tipo:** Lista de tuplas (*list* *de* *tuple*).
  + **Descripción:**  
    Almacena las coordenadas tridimensionales (x, y, z) de los puntos clave faciales extraídos de la captura guardada.
    - Estos puntos clave se usan para comparar similitudes con otros cuadros en el flujo de video.
    - Se obtiene mediante la función *obtener*\_*puntos*\_*clave*.

# Puntos clave críticos

# PUNTOS\_CRITICOS

* + **Tipo:** Lista de enteros (*list* de *int*).
  + **Descripción:**  
    Conjunto de índices que representan puntos clave específicos de la malla facial generada por *MediaPipe*.
    - Incluye puntos importantes de las siguientes áreas:
      * **Ojo izquierdo:** Índices del 33 al 42.
      * **Ojo derecho:** Índices del 43 al 52.
      * **Nariz:** Índices del 1 al 5.
      * **Boca:** Índices del 61 al 72.
      * **Contorno del rostro:** Índices del 0, 17 al 23.
    - Estos puntos son cruciales para reducir la complejidad del procesamiento al enfocar solo en áreas relevantes del rostro.

# Interfaz gráfica de usuario

**Ilustración 3 Captura de pantalla del código para la interfaz grafica**

 **Fuente: Elaboración propia**

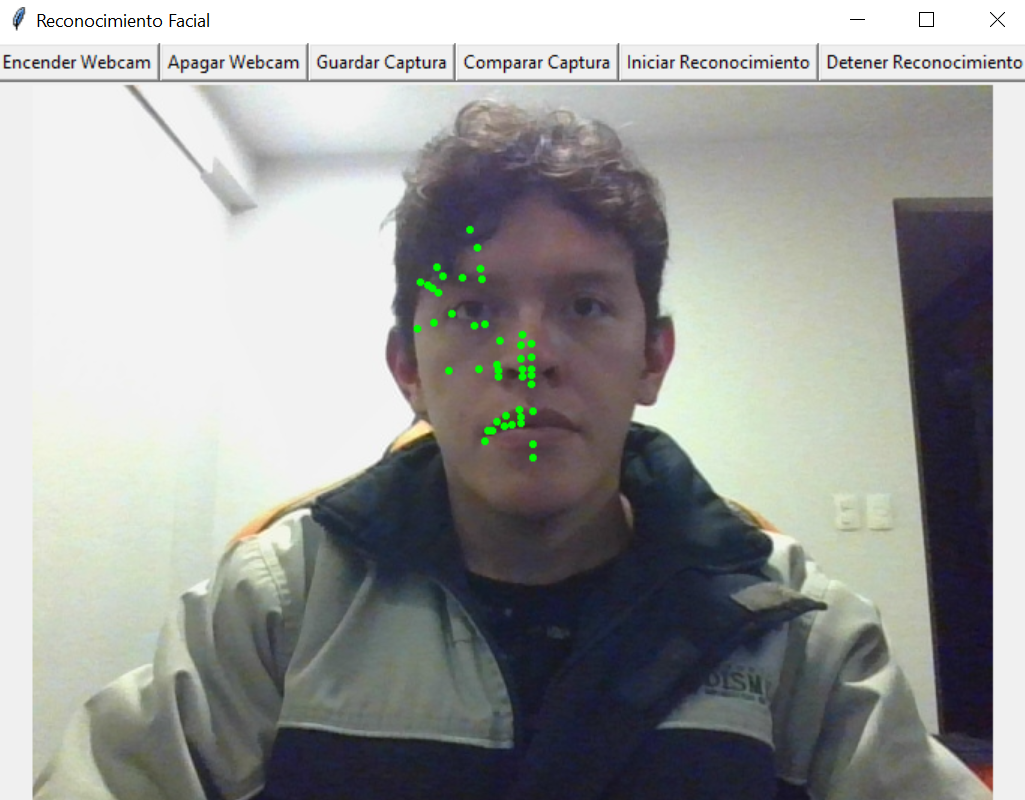
# Ventana principal

* **Título:**
  + La ventana principal lleva el título "Reconocimiento Facial" por medio al método *root.title().*
* **Tamaño:**
  + No se especifica un tamaño fijo, pero los elementos de la ventana (botones y video) se ajustan automáticamente al tamaño del contenido.

# Sección de botones

* **Frame de botones (*frame*\_*botones*):**
  + Ubicado en la parte superior de la ventana mediante *pack(side=tk.TOP, fill=tk.X).*
  + Contiene una fila horizontal de botones, cada uno con una funcionalidad específica.
* **Botones dentro del frame:**
  + Los botones están dispuestos horizontalmente con *pack(side=tk.LEFT, fill=tk.X, expand=True),* lo que permite que cada botón se ajuste proporcionalmente al ancho del frame.
  + **Lista de botones y sus funciones:**
    - **Encender Webcam:**
      * Texto: "Encender Webcam".
      * Acción: Llama a la función *encender*\_*webcam* para activar la cámara (**Ilustración *4***).

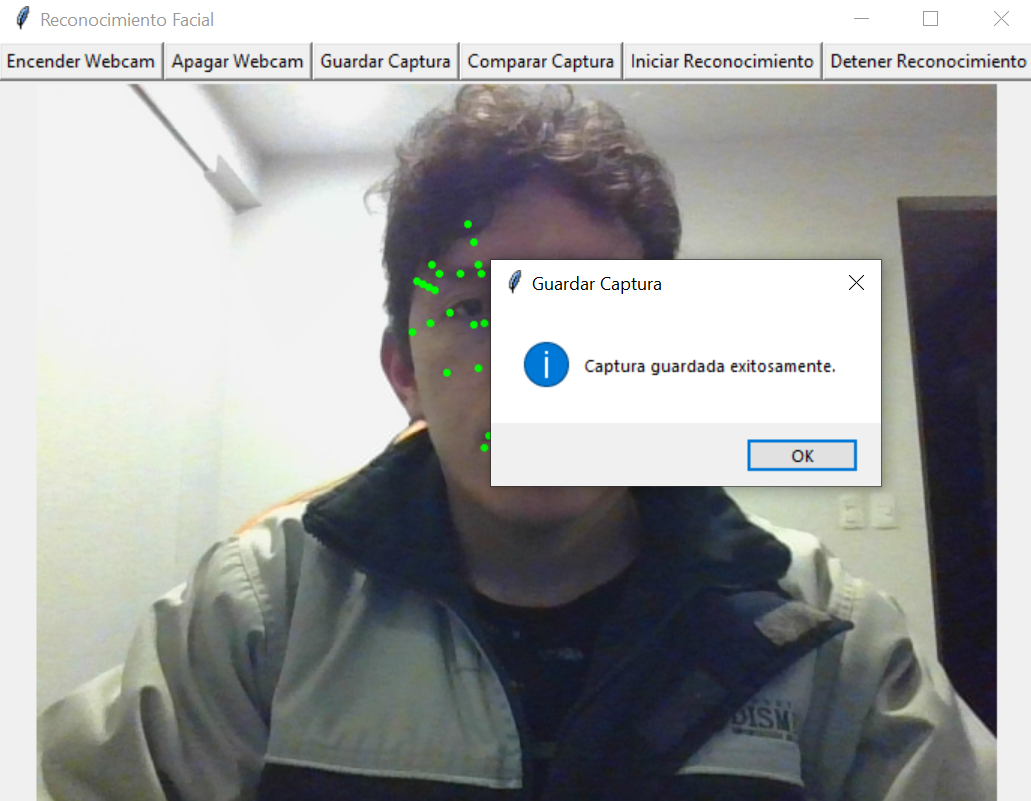
**Ilustración 4 Captura de pantalla: Encender webcam**



**Fuente: Elaboración propia**

* + - **Apagar Webcam:**
      * Texto: "Apagar Webcam".
      * Acción: Llama a la función *apagar*\_*webcam* para detener la cámara.
    - **Guardar Captura:**
      * Texto: "Guardar Captura".
      * Acción: Llama a la función *guardar*\_*captura* para capturar y guardar la imagen actual (**Ilustración *5***).

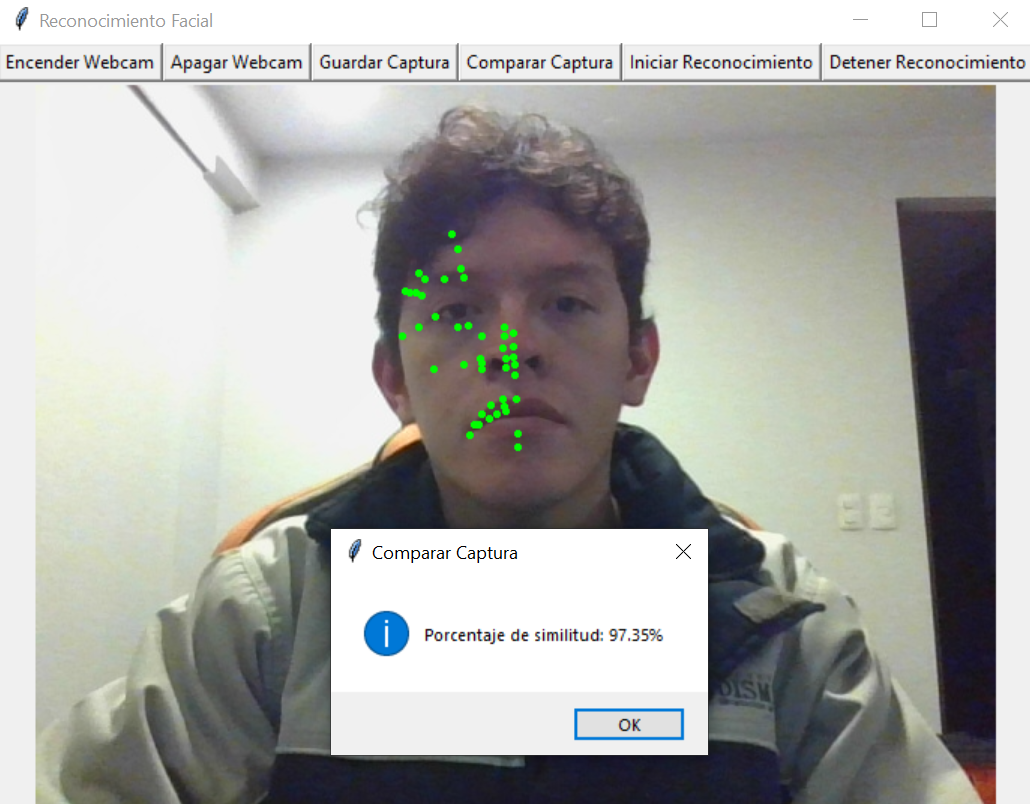
**Ilustración 5 Captura de pantalla: Guardar captura**



**Fuente: Elaboración propia**

* + - **Comparar Captura:**
      * Texto: "Comparar Captura".
      * Acción: Llama a la función *comparar\_captura* para comparar la imagen guardada con el cuadro actual (**Ilustración 6**).

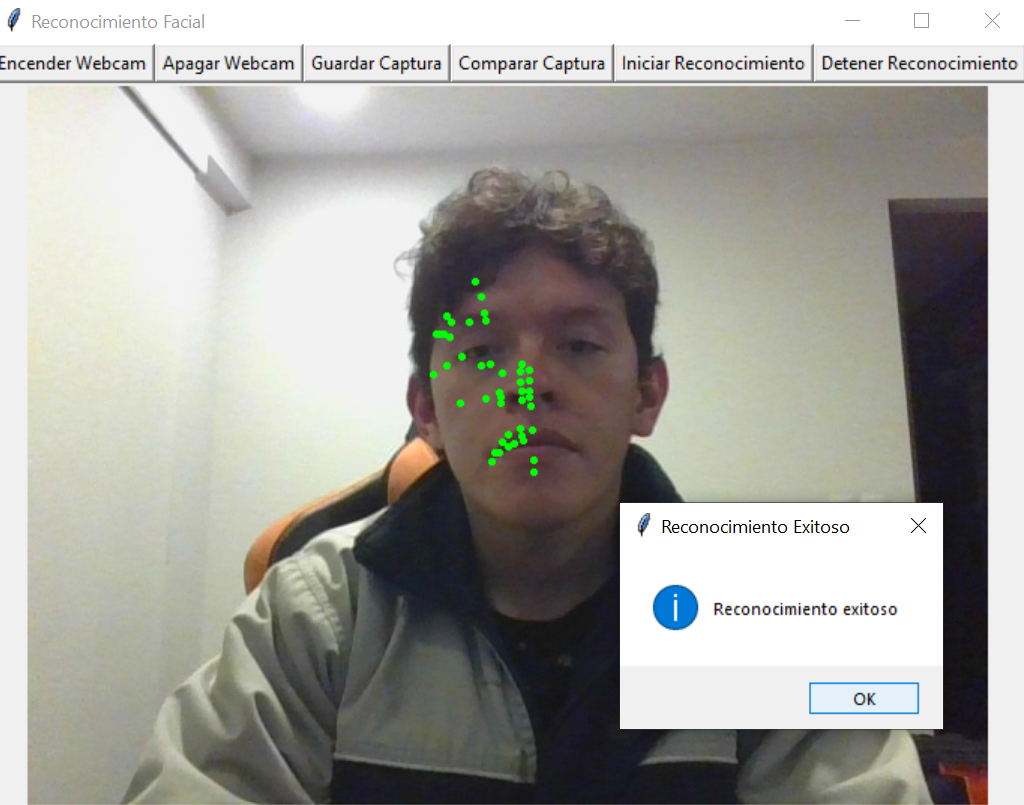
**Ilustración 6 Captura de pantalla: Comparar captura**



**Fuente: Elaboración propia**

* + - **Iniciar Reconocimiento:**
      * Texto: "Iniciar Reconocimiento".
      * Acción: Llama a la función *iniciar*\_*reconocimiento* para comenzar a analizar el video en tiempo real.
      * Si la identidad de la persona transmitida por la cámara corresponde a la captura guardada mientras el reconocimiento esta activo, entonces se mostrara una ventana de dialogo con el texto “Reconocimiento exitoso” (**Ilustración *7***).

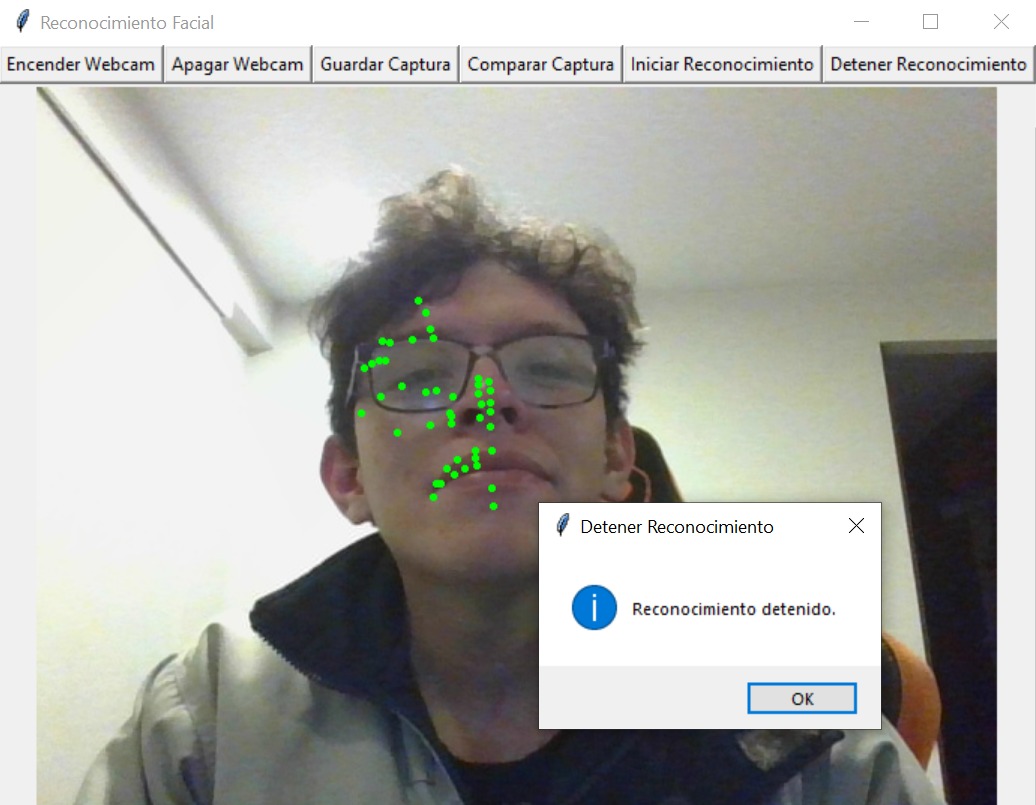
**Ilustración 7 Captura de pantalla: Comparación exitosa en tiempo real**



**Fuente: Elaboración propia**

* + - **Detener Reconocimiento:**
      * Texto: "Detener Reconocimiento".
      * Acción: Llama a la función *detener*\_*reconocimiento* para finalizar el análisis (**Ilustración *8***).

**Ilustración 8 Captura de pantalla: Detener reconocimiento**



**Fuente: Elaboración propia**

# Visualización de video

* **Etiqueta para video (*lbl*\_*video*):**
  + Ubicada debajo del *frame* de botones mediante *lbl\_video.pack().*
  + Sirve como contenedor para mostrar los cuadros de video procesados en tiempo real.
  + Se actualizará continuamente con la función *actualizar*\_*frame* durante el reconocimiento facial o al encender la cámara.

# Bucle principal

* ***root*.*mainloop*():**
  + Mantiene la interfaz gráfica activa y en constante escucha de eventos.
  + Gestiona las interacciones del usuario, como clics en los botones.

# Funciones

# Función encender\_webcam

**Ilustración 9 Captura de pantalla Fucion encender\_webcam**

 **Fuente: Elaboración propia**

**Propósito general:** La función encender\_webcam tiene como objetivo inicializar la captura de video desde la cámara web del sistema y comenzar el flujo de procesamiento de los fotogramas en tiempo real.

**Declaración de variable global:**

La línea global cap indica que la variable cap, utilizada para manejar la captura de video, será accedida y modificada a nivel global. Esto permite que otras funciones del programa interactúen con la misma instancia de la cámara web.

**Inicialización de la captura de video:**

La línea cap = cv2.VideoCapture(0) crea un objeto de captura de video usando OpenCV (cv2), asignado a la variable global cap.

El parámetro 0 indica que se utiliza la cámara web predeterminada del sistema. Si existieran múltiples cámaras conectadas, podrían utilizarse otros índices como 1, 2, etc., para seleccionar una cámara específica.

**Inicio del procesamiento de fotogramas:**

La línea actualizar\_frame() llama a la función actualizar\_frame, lo que inicia el procesamiento continuo de los fotogramas capturados por la cámara. Esta función se encargará de realizar tareas como la detección de rostros y la actualización de la interfaz gráfica con los datos de video.

**Comportamiento esperado:**

Una vez ejecutada, la función activa la cámara web y pone en marcha el flujo de procesamiento en tiempo real.

Si la cámara ya estaba activa o en uso, el código no incluye una verificación explícita para manejar esa condición, por lo que podría ser necesario agregar validaciones adicionales para evitar conflictos.

# Función apagar\_webcam

**Ilustración 10 Captura de pantalla: Función apagar\_webcam**



**Fuente: Elaboración propia**

**Propósito general:** La función apagar\_webcam tiene como objetivo detener la captura de video desde la cámara web y liberar los recursos asociados, además de limpiar el contenido visual del widget que muestra el video en la interfaz gráfica.

**Declaración de variable global:**

La línea global cap declara que la variable cap es global, permitiendo que la función acceda y modifique la instancia de captura de video inicializada previamente en encender\_webcam.

**Liberación de recursos de la cámara:**

La condición if cap: verifica si cap ha sido inicializada, es decir, si la cámara está activa. Si esta condición es verdadera:

La línea cap.release() libera los recursos asociados a la cámara web, deteniendo efectivamente la captura de video.

La línea cap = None asigna None a la variable global cap, indicando que la cámara ya no está en uso y evitando futuras operaciones no deseadas con esta variable.

**Limpieza del widget de video:**

* La línea lbl\_video.imgtk = None elimina la referencia a cualquier imagen previa asignada al widget lbl\_video, asegurándose de que no se muestre un fotograma anterior o residual.
* La línea lbl\_video.configure(image='') establece la propiedad de imagen del widget a un valor vacío, limpiando visualmente el área donde se mostraba el video en la interfaz gráfica.

**Comportamiento esperado:**

Cuando se ejecuta esta función, la cámara web se detiene y su recurso es liberado para que pueda ser utilizado por otros programas. Además, la interfaz gráfica deja de mostrar el contenido capturado por la cámara, proporcionando un estado limpio para el widget de video.

**Relación con otras funciones:**

Esta función complementa a encender\_webcam, ya que asegura que los recursos de la cámara se manejen correctamente cuando la cámara ya no es necesaria. Es especialmente útil para evitar problemas como conflictos en el acceso a la cámara por parte de otros programas o errores al intentar reiniciar la captura de video.

# Función guardar\_captura

**Ilustración 11 Captura de pantalla: Función guardar\_captura**

 **Fuente: Elaboración propia**

**Propósito general:**

La función *guardar\_captura* permite capturar un fotograma de video desde la cámara activa, guardarlo como imagen en el disco y extraer los puntos clave del rostro para su posterior comparación.

**Revisión del estado de la cámara:**

La función verifica si la cámara está activa utilizando la variable global cap. Si *cap* es válida, significa que la cámara está lista para capturar un fotograma.

**Captura del fotograma:**

Se utiliza el método *read()* para obtener el fotograma actual de la cámara.

* Si la captura es exitosa, el fotograma se almacena en la variable *captura\_guardada*.
* La imagen capturada se guarda en el archivo captura.png en el directorio de trabajo utilizando *OpenCV*.

**Extracción de puntos clave:**

Se llama a la función *obtener\_puntos\_clave*, que procesa el fotograma capturado y extrae los puntos clave específicos del rostro definidos previamente en PUNTOS\_CRITICOS. Estos puntos clave se guardan en la variable global *puntos\_clave\_guardados* para su uso en funciones de comparación.

**Notificación al usuario:**

Tras guardar la captura y los puntos clave, se muestra un cuadro de diálogo informativo para notificar al usuario que la captura se ha realizado correctamente. Esto se realiza mediante la función *messagebox.showinfo* de *Tkinter*.

**Relación con otras funciones:**

* Los datos generados por esta función son fundamentales para las operaciones de reconocimiento facial, ya que proporcionan los puntos clave almacenados en *puntos\_clave\_guardados*, que se comparan con los puntos de fotogramas en tiempo real.

**Comportamiento esperado:**

Si la cámara está activa y se captura correctamente un fotograma, se guarda la imagen y se registran los puntos clave del rostro. Si la cámara no está activa, la función no realiza ninguna acciónreferencia para realizar comparaciones o reconocimiento facial.

# Función comparar\_captura

**Ilustración 12 Captura de pantalla: Función comparar\_captura**

 **Fuente: Elaboración propia**

**Propósito general:**

La función *obtener\_puntos\_clave* procesa una imagen para detectar un rostro y extraer puntos clave específicos, utilizando la tecnología *FaceMesh* de *MediaPipe*. Estos puntos clave son necesarios para el reconocimiento facial y su comparación.

**Entrada:**

La función recibe un único parámetro llamado *frame*, que representa el fotograma capturado por la cámara. Este fotograma es un arreglo de tipo *NumPy* que contiene los datos de imagen en formato BGR (azul, verde, rojo).

**Conversión de formato de color:**

Antes de procesar la imagen, esta se convierte del formato BGR (usado por *OpenCV*) al formato RGB (requerido por *MediaPipe*). Esto se logra mediante la función *cv2.cvtColor*.

**Procesamiento con MediaPipe FaceMesh:**

El fotograma convertido se pasa al modelo *mp\_face\_mesh* para detectar rostros y obtener sus características faciales.

El método *process* de *FaceMesh* devuelve un objeto *results* que contiene los puntos de referencia (*landmarks*) de los rostros detectados, si los hay.

**Extracción de puntos clave:**

La función verifica si se detectaron puntos de referencia en el rostro mediante *results.multi\_face\_landmarks.*  
 **Si se detectan rostros:**

* Se inicializa una lista vacía llamada *puntos\_clave.*
* Para cada conjunto de puntos de referencia detectados, se seleccionan los índices especificados en *PUNTOS\_CRITICOS*.
* Los puntos seleccionados se agregan a la lista *puntos\_clave* en formato tridimensional (coordenadas X, Y, Z), donde las coordenadas están normalizadas entre 0 y 1.

**Retorno de datos:**

* Si se detectan y extraen puntos clave, la función devuelve una lista de tuplas con las coordenadas tridimensionales de los puntos clave seleccionados.
* Si no se detectan rostros, la función devuelve *None.*

**Relación con otras funciones:**

Esta función es esencial para *guardar\_captura*, *comparar\_captura*, y *comparar\_en\_tiempo\_real*, ya que proporciona los puntos clave del rostro necesarios para el reconocimiento facial y la comparación.

**Comportamiento esperado:**

* Si el fotograma contiene un rostro visible y detectable, los puntos clave del rostro se extraen correctamente.
* Si no se detecta ningún rostro o el fotograma está vacío, la función no genera puntos clave, retornando *None*.

# Función comparar\_captura

**Ilustración 13 Captura de pantalla: Función comparar\_captura**

 **Fuente: Elaboración propia**

**Propósito general:**

La función *comparar\_captura* permite comparar el rostro capturado previamente (y guardado junto con sus puntos clave) con el rostro detectado en el fotograma actual. Su objetivo es calcular un porcentaje de similitud entre ambos conjuntos de puntos clave.

**Condiciones previas:**

* **Requisitos:** Para que esta función se ejecute correctamente, deben existir una captura previa y puntos clave asociados, almacenados en las variables globales *captura\_guardada* y *puntos\_clave\_guardados*.
* **Manejo de errores:** Si no hay una captura previa o puntos clave disponibles, muestra un mensaje de advertencia indicando que no hay datos con los cuales comparar.

**Flujo de ejecución:**

1. **Validación inicial:**

Verifica si las variables *captura\_guardada* y *puntos\_clave\_guardados* contienen datos válidos.

* + Si no están inicializadas, se muestra un cuadro de advertencia con el mensaje:  
    *"No hay captura guardada para comparar."*
  + En este caso, la función finaliza sin realizar ninguna acción adicional.

1. **Captura del fotograma actual:**
   * Verifica si la cámara está activa (*cap* no es *None*).
   * Si la cámara está activa, lee un nuevo fotograma de video con el método *cap.read().*
   * Si la lectura del fotograma es exitosa, este se pasa a la función *obtener\_puntos\_clave* para extraer los puntos clave del rostro en el fotograma actual.
2. **Extracción de puntos clave:**
   * Si *obtener\_puntos\_clave* devuelve un conjunto de puntos clave válidos (*puntos\_clave\_frame*), se procede a comparar estos puntos con los almacenados en *puntos\_clave\_guardados*.
3. **Cálculo de similitud:**
   * Llama a la función *comparar\_puntos\_clave*, pasando como argumentos los puntos clave guardados y los del fotograma actual.
   * La función *comparar\_puntos\_clave* calcula el porcentaje de similitud entre los dos conjuntos de puntos clave.
4. **Resultado:**
   * Muestra un cuadro de diálogo con el porcentaje de similitud calculado, utilizando el mensaje:  
     *"Porcentaje de similitud: XX.XX%"*, donde XX.XX representa el valor numérico obtenido.

**Relación con otras funciones:**

* Utiliza *obtener\_puntos\_clave* para extraer puntos clave del rostro en el fotograma actual.
* Se apoya en *comparar\_puntos\_clave* para calcular la similitud entre los puntos clave guardados y los detectados.

**Comportamiento esperado:**

* Si ambos conjuntos de puntos clave (captura guardada y fotograma actual) están disponibles, la función calcula y muestra un porcentaje de similitud.
* Si no hay captura guardada, muestra un mensaje de advertencia y termina sin errores.
* Si no se detecta un rostro en el fotograma actual, no realiza ninguna comparación ni muestra resultados.

# Función comparar\_puntos\_clave

**Ilustración 14 Captura de pantalla: Funcion comparar\_puntos\_clave**



**Fuente: Elaboración propia**

**Propósito general:**

La función *comparar\_puntos\_clave* calcula la similitud entre dos conjuntos de puntos clave faciales. Esta comparación se basa en la distancia euclidiana promedio entre los puntos correspondientes en los dos conjuntos, y la similitud resultante se expresa como un porcentaje.

**Parámetros de entrada:**

1. **puntos\_guardados** - Lista de coordenadas tridimensionales (x, y, z) correspondientes a los puntos clave del rostro guardados previamente.
2. **puntos\_frame** - Lista de coordenadas tridimensionales (x, y, z) de los puntos clave extraídos del fotograma actual.

**Flujo de ejecución:**

1. **Validación inicial:**
   * Comprueba si el número de puntos clave en *puntos\_guardados* es diferente del número de puntos clave en *puntos\_frame*.
   * Si las listas no tienen la misma longitud, devuelve inmediatamente un valor de similitud del 0.0%, ya que no es posible realizar una comparación válida.
2. **Cálculo de la distancia promedio:**
   * Itera sobre ambos conjuntos de puntos clave en paralelo utilizando la función zip.
   * Para cada par de puntos clave (p1 y p2), calcula la distancia euclidiana entre ellos utilizando la fórmula:
   * Suma todas las distancias obtenidas y las almacena en una variable llamada *distancia\_total*.
3. **Normalización de la similitud:**
   * Divide la *distancia\_total* entre el número total de puntos clave para obtener la distancia promedio.
   * Calcula la similitud como un porcentaje inversamente proporcional a la distancia promedio, utilizando la fórmula:
   * Asegura que el valor de similitud nunca sea negativo utilizando *max(0, similitud)*.
4. **Retorno del resultado:**
   * Devuelve el porcentaje de similitud calculado.

**Relación con otras funciones:**

* Es llamada por la función *comparar\_captura* para comparar los puntos clave de una captura guardada con los del fotograma actual.
* También es utilizada en *comparar\_en\_tiempo\_real* para evaluar similitudes continuamente mientras el reconocimiento facial está activo.

**Comportamiento esperado:**

* Si ambos conjuntos de puntos clave tienen el mismo tamaño, devuelve un porcentaje de similitud basado en la distancia promedio entre los puntos.
* Si los conjuntos tienen tamaños diferentes, devuelve una similitud del 0.0%.
* Es robusta para manejar casos en los que la similitud pueda ser mínima (distancias grandes) o máxima (distancias pequeñas).

# Función iniciar\_reconocimiento

**Ilustración 15 Captura de pantalla: Función iniciar\_reconocimiento**



**Fuente: Elaboración propia**

**Propósito general:**

La función *iniciar\_reconocimiento* activa un proceso de comparación continua entre los puntos clave de una captura facial previamente guardada y los puntos clave extraídos en tiempo real de la transmisión de la cámara. Es una preparación para realizar el reconocimiento facial continuo.

**Flujo de ejecución:**

1. **Validación inicial:**
   * Comprueba si la captura guardada (*captura\_guardada*) y sus puntos clave (*puntos\_clave\_guardados*) están disponibles.
   * Si cualquiera de estos elementos está ausente, muestra un mensaje de advertencia al usuario indicando que no hay datos disponibles para realizar la comparación, utilizando *messagebox.showwarning*. La ejecución se detiene en este punto.
2. **Activación del reconocimiento:**
   * Si la validación inicial es exitosa, establece la variable global *reconocimiento\_activo* como True. Esto indica que el sistema debe comenzar a comparar puntos clave en tiempo real.
3. **Llamada a la comparación en tiempo real:**
   * Llama a la función *comparar\_en\_tiempo\_real* para iniciar el proceso de comparación continua mientras *reconocimiento\_activo* esté habilitado.

**Parámetros de entrada:**

La función no recibe parámetros externos, pero utiliza las siguientes variables globales:

* *captura\_guardada*: Imagen guardada previamente desde la webcam.
* *puntos\_clave\_guardados*: Lista de puntos clave faciales extraídos de la captura guardada.
* *reconocimiento\_activo*: Indicador booleano que controla si el reconocimiento en tiempo real está habilitado.

**Relación con otras funciones:**

* Se complementa con *comparar\_en\_tiempo\_real*, que ejecuta el proceso de comparación continua en fotogramas sucesivos de la cámara.
* Requiere que las funciones *guardar\_captura* y *obtener\_puntos\_clave* hayan guardado correctamente los datos de referencia.

**Comportamiento esperado:**

* Si los datos necesarios (captura guardada y puntos clave) están disponibles, el sistema activa el reconocimiento facial en tiempo real.
* Si faltan datos, el usuario es informado mediante una advertencia y no se activa el reconocimiento.

**Errores manejados:**

* Muestra un mensaje claro al usuario si no se ha guardado previamente una captura o los puntos clave asociados.
* Evita iniciar procesos innecesarios cuando no hay datos válidos para comparar.

Esta función es el punto de entrada para iniciar el reconocimiento facial continuo del sistema.

# Función detener\_reconocimiento

**Ilustración 16 Captura de pantalla: Funcion detener\_reconocimiento**

 **Fuente: Elaboración propia**

**Propósito general:**

La función *detener\_reconocimiento* permite al usuario detener el proceso de reconocimiento facial en tiempo real, interrumpiendo la ejecución continua de la función *comparar\_en\_tiempo\_real*.

**Desactivación del reconocimiento:**

La función modifica la variable global *reconocimiento\_activo* para detener el proceso de comparación en tiempo real:

* *global reconocimiento\_activo*
* *reconocimiento\_activo = False*
* *reconocimiento\_activo*: Se asigna el valor False, indicando que el reconocimiento facial ya no está activo.

**Notificación al usuario:**

Una vez que se ha desactivado el reconocimiento, se muestra un mensaje al usuario para confirmar la acción:

* *messagebox.showinfo("Detener Reconocimiento", "Reconocimiento detenido.")*
* *messagebox.showinfo*: Utiliza un cuadro de diálogo informativo de Tkinter para comunicar que el reconocimiento ha sido detenido correctamente.

**Relación con otras funciones:**

* Se utiliza para finalizar el proceso iniciado por *iniciar\_reconocimiento*.
* Detiene la ejecución cíclica de *comparar\_en\_tiempo\_real* al cambiar el estado de *reconocimiento\_activo*.

**Comportamiento esperado:**

* Interrumpe inmediatamente cualquier comparación en curso.
* Garantiza que no se realicen nuevas evaluaciones de similitud hasta que el reconocimiento sea activado nuevamente.

# Función actualizar\_frame

**Ilustración 17 Captura de pantalla: Función actualizar\_frame**



**Fuente: Elaboración propia**

**Propósito general:**

La función *actualizar\_frame* gestiona la visualización en tiempo real del flujo de video capturado por la cámara web en la interfaz gráfica. Además, aplica un procesamiento facial mediante *MediaPipe* para detectar puntos clave en el rostro y dibujarlos directamente sobre el fotograma.

**Flujo de ejecución:**

1. **Captura de un nuevo fotograma:**
   * Comprueba si la cámara está encendida (*cap* no es *None*).
   * Usa *cap.read()* para capturar un fotograma de la transmisión en vivo.
   * Si la captura no es válida (ret es False), la función termina su ejecución para ese ciclo.
2. **Procesamiento facial:**
   * Convierte el fotograma capturado de BGR (formato predeterminado de *OpenCV*) a RGB, ya que *MediaPipe* requiere imágenes en formato RGB.
   * Procesa el fotograma con *MediaPipe Face Mesh* (*mp\_face\_mesh.process*), que detecta puntos clave en los rostros presentes en la imagen.
3. **Dibujar puntos clave:**
   * Si *MediaPipe* detecta rostros en el fotograma, itera a través de los puntos clave definidos en *PUNTOS\_CRITICOS*.
   * Para cada punto, calcula las coordenadas en píxeles utilizando el ancho y alto del fotograma y las proporciones proporcionadas por *MediaPipe*.
   * Dibuja un pequeño círculo en la posición de cada punto clave usando cv2.circle. Los círculos son de color verde (RGB: (0, 255, 0)) y tienen un tamaño de 2 píxeles.
4. **Conversión para la interfaz gráfica:**
   * Convierte el fotograma procesado de nuevo a formato RGB usando *PIL* (*Python Imaging Library*), para que sea compatible con la visualización en *Tkinter*.
   * Convierte la imagen a un objeto *ImageTk.PhotoImage*, que puede usarse como imagen en una ventana de *Tkinter*.
5. **Actualización de la interfaz:**
   * Asigna la imagen procesada a la ventana *lbl\_video* que se encarga de mostrar el video en la ventana principal de la aplicación.
   * Programa la ejecución de la función *actualizar\_frame* nuevamente después de un breve intervalo de 10 milisegundos, utilizando *lbl\_video.after(10, actualizar\_frame)*. Esto crea un bucle de actualización constante para la transmisión de video.

**Parámetros de entrada:**

La función no recibe parámetros, pero utiliza las siguientes variables globales:

* **cap:** Objeto que representa la conexión con la cámara web.
* **lbl\_video:** Ventana de etiqueta (*label*) en el cual se muestra el flujo de video procesado.
* **PUNTOS\_CRITICOS:** Lista de índices que representan los puntos clave específicos que deben destacarse en el rostro.

**Relación con otras funciones:**

* Complementa a *encender\_webcam*, que inicia la captura de video al encender la cámara.
* Utiliza la funcionalidad de *MediaPipe* (instanciada como *mp\_face\_mesh*) para detectar puntos faciales en tiempo real.

**Comportamiento esperado:**

* Muestra el flujo de video en tiempo real dentro de la ventana principal de la aplicación.
* Dibuja puntos clave críticos en los rostros detectados, proporcionando una visualización clara de las áreas de interés procesadas por el sistema.
* Continúa actualizándose constantemente mientras la cámara esté encendida.

**Errores manejados:**

* Si no se puede capturar un fotograma (por ejemplo, si la cámara está desconectada o apagada), no se genera ningún error visible y la función simplemente no actualiza la ventana.
* Si no se detectan rostros o puntos clave en el fotograma, no intenta dibujar los puntos y sigue mostrando el video original.

Esta función es esencial para proporcionar retroalimentación visual al usuario durante el uso de la aplicación, mostrando tanto el flujo de video en vivo como los resultados de las detecciones faciales en tiempo real.

# Función actualizar\_frame

**Propósito general:** La función actualizar\_frame tiene como objetivo principal capturar y procesar los fotogramas de video desde la cámara web en tiempo real, utilizando Mediapipe para la detección de rostros, y actualizar el contenido visual en una ventana de la interfaz gráfica.

**Ilustración 18 Captura de pantalla: Función actualizar\_frame**

 **Fuente: Elaboración propia**

**Referencia a la variable global cap:**

La función comienza declarando que la variable cap es global, lo que permite acceder a la captura de video inicializada previamente por la función encender\_webcam.

**Ilustración 19 Captura de pantalla: Funcion encender\_webcam**

 **Fuente: Elaboración propia**

**Verificación de captura activa:**

La línea if cap: asegura que cap ha sido correctamente inicializada, es decir, que la cámara está activa y lista para capturar fotogramas. Si esto es cierto, se ejecutará el código contenido dentro de esta condición.

**Captura del fotograma:**

La línea ret, frame = cap.read() realiza una operación de asignación doble. La función cap.read() devuelve dos valores:

* ret: un valor booleano que indica si la captura del fotograma fue exitosa (True) o si falló (False).
* frame: un arreglo de tipo numpy que contiene los datos en formato de imagen del fotograma capturado.

**Conversión de color del fotograma:**

La línea img\_rgb = cv2.cvtColor(frame, cv2.COLOR\_BGR2RGB) convierte el fotograma capturado (almacenado en frame) del formato de color BGR (Blue-Green-Red) al formato RGB (Red-Green-Blue). Esta conversión es necesaria porque Mediapipe y la biblioteca PIL esperan las imágenes en formato RGB para procesarlas correctamente.

**Procesamiento de detección de rostros con Mediapipe:**

La línea results = mp\_face\_detection.process(img\_rgb) utiliza la instancia de Mediapipe Face Detection (mp\_face\_detection) para procesar el fotograma convertido a RGB. Este procesamiento devuelve un objeto results que contiene la información de detección, incluyendo las coordenadas de los rostros encontrados, si es que existen.

**Dibujado de las detecciones en el fotograma original:**

El bloque if results.detections: verifica si se detectaron rostros en el fotograma. Si el atributo detections del objeto results contiene datos, se ejecutará el bloque de código interno.

Dentro de este bloque, for detection in results.detections: recorre cada detección individual en el fotograma.

La línea mp\_drawing.draw\_detection(frame, detection) utiliza Mediapipe Drawing Utilities (mp\_drawing) para dibujar un marco y otros elementos visuales directamente en el fotograma original (frame) sobre los rostros detectados.

**Conversión del fotograma para compatibilidad con Tkinter:**

La línea img = Image.fromarray(cv2.cvtColor(frame, cv2.COLOR\_BGR2RGB)) convierte nuevamente el fotograma procesado a formato RGB, utilizando OpenCV (cv2) para la conversión, y luego lo convierte a un objeto Image de la biblioteca PIL. Esto es necesario para que el fotograma pueda ser utilizado en Tkinter.

**Creación de imagen para Tkinter:**

La línea imgtk = ImageTk.PhotoImage(image=img) crea un objeto PhotoImage, que es el formato que Tkinter requiere para mostrar imágenes en su interfaz.

**Actualización del widget de video:**

La línea lbl\_video.imgtk = imgtk asigna la imagen procesada a la propiedad imgtk del widget lbl\_video, asegurándose de que la referencia a la imagen no se pierda.

La línea lbl\_video.configure(image=imgtk) establece la imagen en el widget lbl\_video, actualizando así el contenido visual mostrado en la interfaz.

**Programación de la siguiente actualización:**

La línea lbl\_video.after(10, actualizar\_frame) programa la ejecución de la función actualizar\_frame nuevamente después de un intervalo de 10 milisegundos. Esto crea un ciclo continuo que permite la actualización en tiempo real de los fotogramas de video en la interfaz gráfica.

# RECOMENDACIONES Y CONCLUSIONES

# Recomendaciones

* Se recomienda explorar métodos más avanzados, como el uso de modelos de aprendizaje profundo y redes neuronales, para incrementar la precisión y robustez del sistema en condiciones adversas.
* Para ampliar la funcionalidad y la capacidad del sistema, sería beneficioso incorporar un manejo de bases de datos que permita almacenar, consultar y gestionar múltiples registros faciales. Esto facilitaría su aplicación a gran escala.
* Considerar el uso de cámaras con mejores especificaciones técnicas.
* Evaluar cómo el sistema podría integrarse en redes de transporte más amplias, adaptándose a diferentes contextos y escalas, para convertirse en una solución integral a nivel regional o nacional.

# Conclusiones

* Se logró identificar los requerimientos funcionales esenciales para la implementación del sistema de reconocimiento facial, como la precisión en la detección de rasgos faciales, la capacidad de procesar datos en tiempo real y la compatibilidad con el entorno de transporte público. Este análisis permitió establecer una base sólida para las siguientes fases del desarrollo del proyecto.
* El diseño del módulo cumplió con los estándares establecidos durante la fase de análisis, implementando técnicas como la detección de puntos clave faciales para garantizar autenticación precisa. Esto permitió optimizar la detección y validación de usuarios, integrando herramientas accesibles y robustas para el procesamiento de imágenes.
* La interfaz gráfica creada resultó funcional y amigable, simplificando la interacción con el sistema. Las pruebas realizadas confirmaron que los operadores pueden manejar el sistema sin necesidad de capacitación técnica avanzada, cumpliendo con el objetivo de facilitar su uso y garantizar su integración en operaciones diarias.
* Las pruebas realizadas en condiciones simuladas y reales demostraron que el sistema es eficaz en la detección de suplantaciones y la autenticación de usuarios. Aunque se identificaron limitaciones bajo condiciones adversas de iluminación, el desempeño general validó la confiabilidad y precisión del sistema en su entorno de aplicación.

# 

# Bibliografía

(dicienbre de 2024). Obtenido de Marcelo Saval: https://marcelosaval.com/what-is-machine-learning/

(diciembre de 2024). Obtenido de viso.ai: https://viso.ai/computer-vision/mediapipe/

(2024). Obtenido de ax: https://aprendeconalf.es/docencia/python/manual/numpy/

Comité Español de Automática. (2016). *Conceptos y métodos en Visión por Computador.* España.

*GoogleCloud*. (diciembre de 2024). Obtenido de https://cloud.google.com/learn/what-is-artificial-intelligence?hl=es-419

*Imagina*. (diciembre de 2024). Obtenido de https://imaginaformacion.com/tutoriales/opencv-en-python

*Pariste*. (diciembre de 2024). Obtenido de https://www.pariste.net/es/transporte-en-paris-8-billetes-y-tarjetas-de-transporte/

*pillow*. (2024). Obtenido de https://pillow.readthedocs.io/en/stable/reference/Image.html