



**UNIVERSIDAD NACIONAL DE CHIMBORAZO**  
**FACULTAD DE INGENIERÍA**  
**CARRERA DE TELECOMUNICACIONES**

Creación de un sistema de reconocimiento facial basado en inteligencia artificial para la verificación de identidad de estudiantes en entornos de aprendizaje, mediante la implementación de técnicas de aprendizaje profundo.

**Trabajo de Titulación para optar al título de:**  
Ingeniero en Telecomunicaciones

**Autor:**

Robalino Silva Danilo Joel

**Tutor:**

Phd. Juan Carlos Cepeda Pacheco

**Riobamba, Ecuador. 2026**

## DECLARATORIA DE AUTORÍA

Yo, Danilo Joel Robalino Silva, con cédula de ciudadanía 0605955558, autor del trabajo de investigación titulado: Creación de un sistema de reconocimiento facial basado en inteligencia artificial para la verificación de identidad de estudiantes en entornos de aprendizaje, mediante la implementación de técnicas de aprendizaje profundo, certifico que la producción, ideas, opiniones, criterios, contenidos y conclusiones expuestas son de mi exclusiva responsabilidad.

Asimismo, cedo a la Universidad Nacional de Chimborazo, en forma no exclusiva, los derechos para su uso, comunicación pública, distribución, divulgación y/o reproducción total o parcial, por medio físico o digital; en esta cesión se entiende que el cesionario no podrá obtener beneficios económicos. La posible reclamación de terceros respecto de los derechos de autor (a) de la obra referida, será de mi entera responsabilidad; librando a la Universidad Nacional de Chimborazo de posibles obligaciones.

En Riobamba, (05 de Junio de 2026).



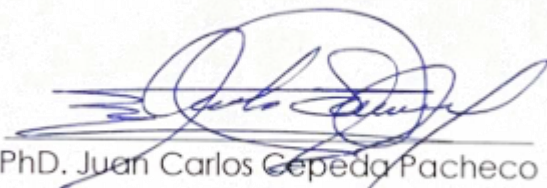
---

Danilo Joel Robalino Silva  
C.I: 0605955558

## **DICTAMEN FAVORABLE DEL PROFESOR TUTOR**

Quien suscribe, Juan Carlos Cepeda Pacheco catedrático adscrito a la Facultad de Ingeniería, por medio del presente documento certifico haber asesorado y revisado el desarrollo del trabajo de investigación titulado: Creación de un sistema de reconocimiento facial basado en inteligencia artificial para la verificación de identidad de estudiantes en entornos de aprendizaje, mediante la implementación de técnicas de aprendizaje profundo, bajo la autoría de Danilo Joel Robalino Silva; por lo que se autoriza ejecutar los trámites legales para su sustentación.

Es todo cuanto informar en honor a la verdad; en Riobamba, a los 18 días del mes de febrero de 2026.



PhD. Juan Carlos Cepeda Pacheco  
**TUTOR(A)**

# CERTIFICADO DE LOS MIEMBROS DEL TRIBUNAL

Quienes suscribimos, catedráticos designados Miembros del Tribunal de Grado para la evaluación del trabajo de investigación “**Creación de un sistema de reconocimiento facial basado en inteligencia artificial para la verificación de identidad de estudiantes en entornos de aprendizaje, mediante la implementación de técnicas de aprendizaje profundo**”, presentado por **Danilo Joel Robalino Silva** con cédula de identidad número **0605955558**, bajo la tutoría de Phd. Juan Carlos Cepeda Pacheco; certificamos que recomendamos la APROBACIÓN de este con fines de titulación. Previamente se ha evaluado el trabajo de investigación y escuchada la sustentación por parte de su autor; no teniendo más nada que observar.

De conformidad a la normativa aplicable firmamos, en Riobamba 05 de junio de 2026

Dr. Marlon Basantes Valverde, Ph.D  
**PRESIDENTE DEL TRIBUNAL DE  
GRADO**



Ing. Leonardo Rentería  
**MIEMBRO DEL TRIBUNAL DE  
GRADO**



Ing. Marco Nolivos  
**MIEMBRO DEL TRIBUNAL DE  
GRADO**





Dirección  
Académica  
VICERRECTORADO ACADÉMICO



UNACH-RGF-01-04-08.17  
VERSIÓN 01: 06-09-2021

## CERTIFICACION

Que, **Danilo Joel Robalino Silva** con CC: **0605955558**, estudiante de la Carrera **Telecomunicaciones**, Facultad de **Ingeniería**; ha trabajado bajo mi tutoría el trabajo de investigación titulado ” “Creación de un sistema de reconocimiento facial basado en inteligencia artificial para la verificación de identidad de estudiantes en entornos de aprendizaje, mediante la implementación de técnicas de aprendizaje profundo.”, cumple con el **2% de similitud y 6% de Inteligencia Artificial**, de acuerdo con el reporte del sistema Anti plagio **Compilation Magister**, porcentaje aceptado de acuerdo a la reglamentación institucional, por consiguiente autorizo continuar con el proceso.

Riobamba, 29 de mayo de 2026



Phd. Juan Carlos Cepeda Pacheco  
**TUTOR**

## **DEDICATORIA**

Dedicado a Heine

## **AGRADECIMIENTO**

Agradecimientos primero quiero expresar mi gratitud a Dios por brindarme salud inteligencia y fortaleza necesaria para culminar esta etapa de mi vida profesional.

A mis padres Irene y Byron quienes fueron el pilar fundamental para continuar con su apoyo incondicional esfuerzo constante y los valores que me han enseñado siendo el principal motor que me impulso alcanzar esta meta. A mis hermanos Rony, Pamela, mis primos Kevin y Enya todos han formado parte importante en mi crecimiento personal y académico quienes brindándome ánimo y confianza me han acompañado en las diferentes etapas de este camino les manifiesto mi más profundo agradecimiento este logro también les pertenece.

A mi novia Lisbeth B. por estar a mi lado por su paciencia, por ayudarme y motivarme, con su amor, compañía y apoyo incondicional formo parte importante para alcanzar este objetivo.

A mis amigos, JLuis, Kevin, Teo, Anthony, Angie, Olmedito mi reconocimiento por su valiosa colaboración participando como parte de la base de datos empleada en este estudio recordándome que este camino no lo recorrí solo demostrando el verdadero significado de la amistad.

Finalmente, a todas esas personas que de diferentes formas me brindaron su apoyo para la culminación de este objetivo académico.

## ÍNDICE GENERAL

**DECLARATORIA DE AUTORÍA**

**DICTAMEN FAVORABLE DEL PROFESOR TUTOR**

**CERTIFICADO DE LOS MIEMBROS DEL TRIBUNAL**

**CERTIFICADO ANTIPLAGIO**

**DEDICATORIA**

**AGRADECIMIENTO**

**ÍNDICE DE TABLAS**

**ÍNDICE DE FIGURAS**

**RESUMEN**

**ABSTRACT**

**CAPÍTULO I..... 14**

1. INTRODUCCIÓN..... 14

1.1. Antecedentes..... 14

1.2. Planteamiento del problema ..... 15

1.3. Justificación ..... 16

1.4. Objetivos..... 17

1.4.1. Objetivo general ..... 17

1.4.2. Objetivo específico ..... 17

**CAPÍTULO II..... 18**

2. MARCO TEÓRICO ..... 18

2.1. Estado del arte ..... 18

2.2. Términos y definiciones ..... 19

2.2.1. Fundamentos de la inteligencia artificial..... 19

2.2.2. Aprendizaje automático ..... 19

2.2.3. Programación tradicional..... 20

2.2.4. Tipos de aprendizaje..... 20

2.2.4.1. Supervised Learning (Aprendizaje supervisado)..... 20

2.2.5. Ramas y subcampos de la IA..... 20

2.2.6. Aprendizaje y razonamiento ..... 21

2.2.6.1. Simulación de procesos cognitivos humanos ..... 21

2.2.7. Cognición artificial en la implementación de sistemas de reconocimiento facial. ....	22
2.2.7.1. Simulación de procesos cognitivos.....	22
2.2.7.2. Interdisciplinariedad y optimización Algorítmica.....	23
2.2.7.3. Modelos Cognitivos en reconocimiento facial .....	23
2.2.8. Principios del Reconocimiento Facial .....	23
2.2.8.1. Espacio de características multidimensional .....	23
2.2.8.2. Vectores de características .....	24
2.2.8.3. Principio de disparidad .....	24
2.2.8.4.Reducción de dimensiones .....	24
2.2.9. Modelos de embeddings faciales .....	24
2.2.9.1. Modelo FaceNet .....	25
2.2.9.2. Principios de triplet loss .....	25
2.2.10. Técnicas de Fine-Tuning en modelos preentrenados .....	26
2.2.11. Parámetros de ajustes en el Fine-Tuning.....	26
<b>CAPÍTULO III.....</b>	<b>28</b>
3. METODOLOGÍA.....	28
3.1. Tipo de Investigación .....	28
3.2. Diseño de Investigación.....	28
3.3. Técnicas de recolección de Datos.....	28
3.4. Población de estudio y tamaño de muestra.....	32
3.5. Hipótesis (de ser el caso).....	32
3.6. Métodos de análisis, y procesamiento de datos .....	33
<b>CAPÍTULO IV.....</b>	<b>36</b>
4. RESULTADOS Y DISCUSIÓN .....	36
<b>CAPÍTULO V .....</b>	<b>42</b>
5. CONCLUSIONES Y RECOMENDACIONES .....	42
5.1. Conclusiones.....	42
5.2. Recomendaciones .....	42
<b>BIBLIOGRAFÍA .....</b>	<b>44</b>
<b>ANEXOS.....</b>	<b>48</b>

## ÍNDICE DE TABLAS

Tabla 1. Tiempo estimado para redimensionar imágenes de rostros.....	24
Tabla 2. Resultados ANOVA .....	36
Tabla 3. Decisión de hipótesis para el factor brillo .....	37
Tabla 4. Correlaciones de Spearman brillo.....	38
Tabla 5. Decisión de hipótesis para el factor brillo .....	38
Tabla 6. Correlaciones de Spearman contraste.....	39
Tabla 7. Decisión de hipótesis para el factor brillo .....	39
Tabla 8. Correlación de Spearman para el factor oclusión.....	40

## ÍNDICE DE FIGURAS

Fig.1 Enfoque del machine learning.....	19
Fig.2 Tipos de aprendizaje en machine learning .....	20
Fig.3 Técnicas de inteligencia artificial para Simulación de procesos cognitivos humanos .....	22
Fig.4 Rostros de estudiantes utilizados para entrenamiento de modelo de reconocimiento facial.....	29
Fig.5 Conjunto de imágenes de estudiantes para entrenamiento de modelo de reconocimiento facial .....	29
Fig.6 Fragmento de código en Jupyter Notebook para la generación de embeddings faciales con FaceNet.....	31
Fig.7 Significancia estadística ANOVA.....	36
Fig.8 FAR promedio por brillo .....	37
Fig.9 FAR promedio por contraste. ....	39
Fig.10 FAR promedio por oclusión. ....	40
Fig.11 Fuerza de correlación con FAR.....	40

## RESUMEN

El sistema de reconocimiento facial que se desarrolló fue examinado considerando como medida principal la tasa de falsos positivos (false acceptanse rate FAR) con el fin de analizar cómo se comporta el modelo en ambientes reales.

Se desarrolló diferentes pruebas en el que se utilizó el modelo base pre entrenado y el modelo ajustado mediante fine-tuning usando el mismo conjunto de datos que se conforma por 11 personas y 400 imágenes. También fueron tomadas como variables independientes las condiciones de oclusión brillo e iluminación debido a la repercusión directa en la calidad de captura del rostro

Los resultados que se obtuvieron permitieron observar la manera en la que se comporta el sistema bajo diferentes ambientes, demostrando que el modelo ajustado mediante fine-tuning registra una variación en la tasa de falsos positivos en contraste con el modelo base.

En un entorno controlado (rostro frontal, buena iluminación y sin oclusión), el sistema se mantiene estable alcanzando una óptima identificación de las personas que se encontraban registradas, pero al introducir cambios en la iluminación o la parencia de oclusiones parciales (por ejemplo, manos que cubren parte del rostro) se pueden visualizar cambios en la tasa de falsos positivos.

El modelo con fine-tuning a través de la implementación de restricciones de calidad (como filtros de desenfoque y estabilidad temporal hizo posible controlar el comportamiento del sistema en estos escenarios logrando así eludir en varios casos que se asigne una incorrecta identidad cuando las condiciones en las que se realizaba la captura no eran las adecuadas.

En general los resultados que se obtuvieron hicieron viable analizar como las variables de oclusiones, brillo y contraste influyen en el desempeño del sistema, además del efecto del proceso de fine-tuning sobre la representación de los embeddings faciales y la adopción de decisiones del modelo.

**Palabras claves:** reconocimiento facial; FaceNet; fine-tuning; oclusión; iluminación; tasa de falsos positivos (FAR)

## ABSTRACT

This study presents an AI-based facial recognition system for student identity verification in educational environments. The system is based on the FaceNet model and is implemented using the Inception-ResNet V1 architecture, pre-trained on the VGGFace2 dataset.

A custom dataset was created, comprising a total of 11 subjects, with 40 images per subject. Image preprocessing was performed by face detection and alignment using the MTCNN algorithm, after which each face was represented as a 128-dimensional embedding vector.

The pre-trained model was fine-tuned using a triplet loss and an reduced learning rate, which allowed it to adapt to the specific characteristics of the dataset. Additionally, to enhance recognition stability, real-time constraints, including blur, illumination, and motion filters, were introduced.

The False Acceptance Rate (FAR) was used as the primary performance metric for the system. The effects of independent variables such as occlusion, illumination, and contrast on the system's performance were determined.

The study's results demonstrate the performance characteristics of the modified model and the impact of constraints under various conditions, including changes in illumination, partial occlusion, and facial motion.

**Keywords:** facial recognition; FaceNet; fine-tuning; occlusion; lighting; false acceptance rate (FAR)



Reviewed by:

Ms.C. Ana Maldonado León

ENGLISH PROFESSOR

C.I.0601975980

# CAPÍTULO I

## 1. INTRODUCCIÓN

### 1.1. Antecedentes

El reconocimiento facial como tecnología biométrica ha evolucionado significativamente durante los últimos años. Los métodos tradicionales, basados en la geometría facial y en la extracción manual de características, han sido progresivamente reemplazados por sistemas avanzados fundamentados en técnicas de aprendizaje profundo (deep learning) y redes neuronales. Modelos como FaceNet y conjuntos de datos como VGGFace2 han demostrado que el uso de embeddings faciales permite representar cada rostro mediante un vector dentro de un espacio métrico. En este espacio, la similitud entre rostros puede calcularse mediante métricas como la distancia euclidiana o la similitud coseno. De esta manera, se han alcanzado altos niveles de precisión tanto en escenarios controlados como en condiciones variables de captura [1] [2].

En el contexto universitario, la necesidad de optimizar procesos administrativos y académicos ha impulsado la implementación de sistemas biométricos. Entre sus principales aplicaciones se encuentran el control de asistencia, la verificación de identidad durante evaluaciones y la autenticación de usuarios en plataformas virtuales. Diversos estudios han evidenciado que los sistemas de reconocimiento facial permiten reducir los tiempos operativos y disminuir los errores asociados con los procedimientos manuales. Además, contribuyen a mejorar el registro de los procesos y la seguridad de la información académica [3][4].

En el año 2020, la pandemia de COVID-19 impulsó la virtualización de las actividades académicas y de las evaluaciones remotas. Esta situación generó una mayor demanda de mecanismos confiables para la verificación de identidad. En respuesta a esta necesidad, instituciones de educación superior de Asia, Europa y América Latina desarrollaron soluciones basadas en reconocimiento facial para el control de exámenes en línea y la autenticación continua de los usuarios. Estas herramientas permitieron reducir los casos de suplantación de identidad y fraude académico. Sin embargo, las investigaciones también demostraron que el desempeño de estos sistemas puede verse afectado por factores como las variaciones de iluminación, las oclusiones parciales, el desenfoque y los cambios en la posición del rostro [5][6].

En América Latina, diferentes estudios desarrollados en instituciones educativas han evidenciado la transición desde métodos manuales hacia plataformas automatizadas de control biométrico. Estas investigaciones destacan mejoras en la eficiencia operativa y una reducción en los tiempos de registro. No obstante, también señalan la necesidad de adaptar los modelos preentrenados a las características específicas de cada entorno mediante procesos de *fine-tuning*, especialmente cuando se dispone de una cantidad limitada de imágenes por usuario [7][8].

En este contexto, investigaciones recientes han demostrado que el ajuste de modelos preentrenados puede mejorar considerablemente la capacidad del sistema para rechazar intentos de suplantación de identidad en escenarios abiertos (*open set*). Como resultado, es posible reducir la tasa de falsa aceptación (*False Acceptance Rate*, FAR), considerada una métrica crítica en aplicaciones relacionadas con la seguridad. Asimismo, se ha determinado que la calibración del umbral de decisión y la adaptación del modelo a condiciones reales de captura pueden optimizar el rendimiento del sistema sin necesidad de entrenarlo desde cero [9].

Los antecedentes científicos permiten establecer que, aunque los modelos de reconocimiento facial alcanzan altos niveles de precisión en bases de datos estandarizadas, su desempeño en entornos reales depende de varios factores. Entre ellos se encuentran la calibración del umbral de similitud, la calidad de las imágenes y la adaptación del modelo a las condiciones específicas de funcionamiento. Por esta razón, resulta necesario analizar experimentalmente la influencia de variables como el brillo, el contraste y la oclusión, las cuales pueden afectar indicadores de seguridad como la FAR dentro de un contexto universitario.

## **1.2. Planteamiento del problema**

Durante la última década, las instituciones universitarias y politécnicas han incorporado diversas tecnologías con el propósito de optimizar sus procesos administrativos y académicos. Entre estas tecnologías se encuentran los sistemas biométricos, empleados en actividades como el control de asistencia y la verificación de identidad de los estudiantes. Dentro de este campo, el reconocimiento facial constituye una alternativa relevante debido a que permite automatizar los procesos de identificación mediante un mecanismo de verificación no invasivo. Además, su desarrollo se ha visto favorecido por los avances alcanzados en los modelos de aprendizaje profundo [10].

A pesar de los avances registrados, estudios recientes han demostrado que el rendimiento de los sistemas de reconocimiento facial puede disminuir cuando se implementan en escenarios reales. Entre los factores que pueden afectar su desempeño se encuentran las variaciones de iluminación, los cambios en la calidad de la imagen, los movimientos involuntarios de los usuarios y las oclusiones parciales provocadas por gafas, manos u otros objetos. Estas condiciones inciden directamente en métricas como la tasa de falsa aceptación (FAR) y la tasa de falso rechazo (FRR), las cuales son fundamentales para evaluar la seguridad y la confiabilidad de los sistemas biométricos [11][12][13].

La tasa de falsa aceptación (FAR) representa la probabilidad de que una persona no autorizada sea reconocida incorrectamente como un usuario legítimo. En procesos educativos que involucran a un número considerable de estudiantes, un valor elevado de FAR puede comprometer la confiabilidad del sistema y facilitar casos de suplantación de identidad. Por este motivo, la evaluación de un sistema de reconocimiento facial no debe limitarse únicamente a la precisión general del modelo, sino que también debe considerar su capacidad para minimizar los errores asociados con la seguridad [14][15].

Los modelos de reconocimiento facial preentrenados, como FaceNet y otras arquitecturas basadas en redes neuronales profundas, suelen ser entrenados con conjuntos de datos extensos y diversos. Sin embargo, estos modelos no necesariamente se encuentran adaptados a las características de una población específica ni a las condiciones particulares del entorno en el que serán implementados. Como consecuencia, su desempeño puede disminuir cuando se aplican directamente en escenarios reales sin un proceso adicional de ajuste [16][17].

Diversas fuentes bibliográficas señalan que la aplicación de técnicas de *fine-tuning*, es decir, la adaptación de modelos preentrenados mediante datos propios del entorno de aplicación, puede mejorar el rendimiento de los sistemas de reconocimiento facial. Esta estrategia permite ajustar el modelo a las características particulares de los usuarios registrados y reducir tanto la tasa de falsas aceptaciones como los errores de clasificación en condiciones reales de operación [18].

En las instituciones de educación superior, la eficiencia operativa constituye un aspecto fundamental debido al número de estudiantes y a la necesidad de agilizar los procesos académicos. En este contexto, la evaluación, optimización y adaptación de los sistemas de reconocimiento facial representan un desafío relevante. Su implementación debe garantizar tiempos de respuesta adecuados, seguridad frente a posibles intentos de suplantación de identidad y estabilidad ante diferentes condiciones de captura.

En consecuencia, se identifica una brecha entre el rendimiento reportado por los sistemas de reconocimiento facial en condiciones controladas y su funcionamiento en un entorno universitario real. Esta problemática plantea la necesidad de analizar si la adaptación de un modelo preentrenado mediante técnicas de *fine-tuning*, junto con la incorporación de controles de calidad de imagen, permite mejorar la confiabilidad del proceso de verificación de identidad de los estudiantes.

### **1.3. Justificación**

En la presente investigación se seleccionó la tasa de falsa aceptación (FAR) como variable dependiente para evaluar el rendimiento del sistema de reconocimiento facial. Esta métrica constituye un indicador relevante para analizar la seguridad de los sistemas biométricos implementados en entornos educativos, debido a que mide la proporción de casos en los que una identidad no autorizada es aceptada incorrectamente como legítima. En el ámbito académico, este tipo de error puede ocasionar accesos indebidos, suplantación de identidad o registros incorrectos de asistencia.

Aunque existen otras métricas de evaluación, como el área bajo la curva (AUC) o la puntuación F1 (*F1-score*), estas permiten analizar principalmente el rendimiento general del clasificador. En cambio, la FAR se enfoca directamente en el riesgo de seguridad asociado con la aceptación errónea de usuarios no autorizados. Por esta razón, su utilización se relaciona de manera directa con los objetivos planteados en la investigación.

El estudio busca mejorar la robustez del modelo mediante la aplicación de *fine-tuning* y la incorporación de controles asociados con las condiciones de captura. Para ello, se consideran

factores como el brillo, el contraste y la oclusión parcial del rostro. El análisis de estas variables permite determinar de qué manera las perturbaciones presentes en escenarios reales pueden afectar la tasa de falsas aceptaciones del sistema.

La FAR se encuentra directamente relacionada con el umbral de decisión ( $\tau$ ). Por consiguiente, su análisis permite estudiar la sensibilidad del sistema frente a diferentes configuraciones. Estas características convierten a la FAR en una métrica adecuada para la aplicación de métodos estadísticos, como el análisis de varianza (ANOVA), mediante los cuales se puede determinar el efecto de las variables independientes sobre el comportamiento del sistema.

En esta investigación se consideraron como variables dependientes cuantitativas el nivel de brillo, el nivel de contraste y el grado de oclusión. Estas variables fueron seleccionadas debido a que representan condiciones que pueden presentarse durante la captura de imágenes en entornos académicos. Por tanto, su evaluación permite analizar la estabilidad del sistema de reconocimiento facial ante situaciones cercanas a un escenario real de implementación.

En conclusión, la selección de la FAR como variable dependiente no solo mantiene coherencia con la literatura relacionada con los sistemas biométricos, sino que también responde directamente al problema de investigación. Su análisis permite evaluar la confiabilidad de un sistema de reconocimiento facial orientado a la verificación de identidad de estudiantes en un entorno universitario.

## **1.4. Objetivos**

### **1.4.1. Objetivo general**

- Crear un sistema de reconocimiento facial basado en inteligencia artificial para la verificación de identidad de estudiantes en entornos de aprendizaje, mediante la implementación de técnicas de aprendizaje profundo.

### **1.4.2. Objetivo específico**

- Seleccionar la base de datos adecuada a través del análisis de datasets para el sistema de reconocimiento facial.
- Desarrollar un modelo de reconocimiento facial mediante la adaptación de técnicas de aprendizaje profundo, configurando los parámetros necesarios para el entrenamiento del sistema.
- Evaluar el rendimiento del sistema de reconocimiento facial mediante métricas de evaluación estándar.

## CAPÍTULO II

### 2. MARCO TEÓRICO

#### 2.1. Estado del arte

El reconocimiento facial basado en aprendizaje profundo constituye una de las principales aplicaciones de la visión por computadora. Los sistemas modernos suelen estar conformados por tres etapas principales: la detección y alineación facial, la extracción de características mediante redes neuronales profundas y la comparación de representaciones vectoriales o embeddings mediante métricas de similitud. Las investigaciones recientes sobre reconocimiento facial profundo destacan la relevancia de las arquitecturas preentrenadas y de las funciones de pérdida discriminativas para mejorar la separación entre clases. Este enfoque permite generar representaciones faciales robustas incluso cuando se dispone de un número limitado de muestras [19].

Diversos estudios han demostrado que el rendimiento de estos sistemas no depende únicamente de la arquitectura del modelo, sino también de la calidad de las imágenes de entrada. Factores como la iluminación, el desenfoque, la resolución, la orientación del rostro y la oclusión pueden incidir directamente en las tasas de error biométrico, incluida la tasa de falsa aceptación (FAR). Por esta razón, el análisis de la calidad facial previo al reconocimiento constituye una estrategia relevante para regular las decisiones del sistema y reducir los errores críticos de seguridad [20].

En el ámbito educativo y en otros escenarios de aplicación real, las investigaciones recientes señalan que la evaluación de los sistemas de reconocimiento facial no debe limitarse a la precisión general. También se deben considerar métricas vinculadas con la seguridad y con el rechazo de usuarios impostores. Estudios desarrollados desde el año 2020 destacan la necesidad de minimizar la FAR en entornos administrativos y académicos, debido a que la aceptación incorrecta de un usuario no autorizado puede representar un riesgo significativo para la confiabilidad del sistema [21].

En América Latina y Ecuador, el reconocimiento facial se ha implementado principalmente en sistemas de control de asistencia y verificación de identidad de estudiantes. En la Universidad Técnica de Machala, por ejemplo, se desarrolló un sistema de control de asistencia basado en visión artificial. En este estudio se identificó que las condiciones ambientales y la calidad de las capturas pueden repercutir significativamente en el desempeño del modelo. De manera similar, trabajos académicos realizados en universidades ecuatorianas han incorporado modelos preentrenados adaptados a entornos locales, lo que ha evidenciado la necesidad de realizar ajustes específicos para mejorar la estabilidad y reducir los errores durante el funcionamiento del sistema [22][23].

En la Universidad Nacional de Chimborazo (UNACH) existen antecedentes de trabajos relacionados con el desarrollo de modelos de reconocimiento facial para el control de acceso y la verificación de identidad de los estudiantes. Estas investigaciones evidencian la viabilidad de implementar soluciones biométricas dentro de la institución. Sin embargo, la

mayor parte de las propuestas se centra en la precisión general del sistema y no profundiza en métricas de seguridad como la reducción de la FAR [24].

A nivel internacional, diversas investigaciones han analizado la aplicación de técnicas de fine-tuning con cantidades limitadas de datos (few-shot fine-tuning) en modelos preentrenados. Los resultados han demostrado que es posible adaptar redes neuronales profundas a contextos específicos, mejorar la separación entre identidades conocidas y reducir el riesgo de aceptación de usuarios impostores. Este enfoque resulta relevante en escenarios educativos, donde se dispone de un número limitado de imágenes por estudiante y las condiciones de captura pueden variar [25].

En síntesis, el estado del arte permite identificar tres aspectos principales: la utilidad de los modelos preentrenados adaptados mediante *fine-tuning*, la importancia de validar la calidad de las imágenes antes de realizar el reconocimiento y la necesidad de evaluar métricas de seguridad como la FAR en entornos educativos reales. En este contexto, la presente investigación busca adaptar un sistema de reconocimiento facial a las condiciones de la UNACH, con énfasis en la reducción de la tasa de falsas aceptaciones y en el fortalecimiento de la confiabilidad del proceso de verificación de identidad.

## 2.2. Términos y definiciones

### 2.2.1. Fundamentos de la inteligencia artificial

La inteligencia artificial (IA) es una rama de la informática orientada al desarrollo de sistemas capaces de ejecutar tareas que tradicionalmente han requerido capacidades humanas. Como se muestra en la fig.1, estos sistemas buscan reproducir funciones como el razonamiento, el aprendizaje, la interpretación de información y la toma de decisiones a partir de datos o experiencias previas. Entre sus aplicaciones se encuentran el reconocimiento del habla, el procesamiento del lenguaje natural y la automatización de decisiones.

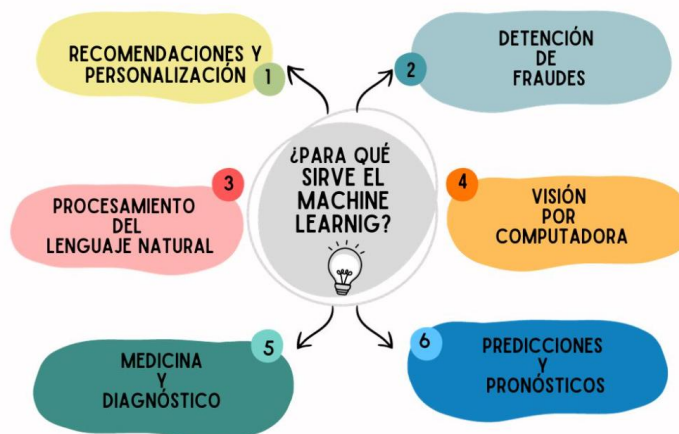


Fig.1 Enfoque del machine learning

### 2.2.2. Aprendizaje automático

El aprendizaje automático (*machine learning*) es una rama de la inteligencia artificial que permite a los sistemas identificar patrones y relaciones a partir de datos, sin que cada

procedimiento deba ser programado de manera explícita. A partir de los datos de entrada, los algoritmos generan modelos capaces de realizar predicciones o tomar decisiones sobre información nueva [26][27].

Los algoritmos de aprendizaje automático pueden mejorar su precisión y eficiencia mediante el procesamiento de ejemplos. Para ello, se proporcionan datos de entrenamiento y, en determinados casos, los resultados esperados. Posteriormente, el sistema identifica patrones que le permiten responder ante nuevas entradas [27][28].

A diferencia de la programación tradicional, el comportamiento del modelo depende principalmente de los datos utilizados durante el entrenamiento y no únicamente de reglas explícitas definidas por el desarrollador [29].

### 2.2.3. Programación tradicional

En la programación tradicional, el desarrollador establece reglas y directrices específicas para que el sistema procese los datos y genere un resultado. Por tanto, el comportamiento de la aplicación depende de la lógica implementada mediante instrucciones previamente definidas [27][29].

### 2.2.4. Tipos de aprendizaje

Existen diferentes tipos de aprendizaje automático, como se muestra en la fig.2. Entre ellos, el aprendizaje supervisado resulta especialmente relevante para el reconocimiento facial.

TIPOS DE MACHINE LEARNING			
Aprendizaje supervisado	Aprendizaje no supervisado	Aprendizaje semi-supervisado	Aprendizaje por refuerzo
Entrenamiento con datos etiquetados (entrada + salida)	Entrenamiento con datos no etiquetados	Combinación de datos etiquetados y datos no etiquetados	Aprendizaje a través de la interacción con el entorno y el feedback (recompensa/castigo)

Fig.2 Tipos de aprendizaje en machine learning

#### 2.2.4.1. Supervised Learning (Aprendizaje supervisado)

El aprendizaje supervisado emplea datos etiquetados para entrenar modelos capaces de clasificar o reconocer patrones. En el reconocimiento facial, las imágenes de entrada se encuentran asociadas con identidades específicas. De esta manera, el sistema aprende a relacionar los rasgos presentes en una imagen con una persona determinada.

Este tipo de aprendizaje resulta adecuado para tareas de identificación y verificación facial, debido a que permite distinguir entre usuarios registrados y analizar si una imagen corresponde a una identidad previamente almacenada [30].

### 2.2.5. Ramas y subcampos de la IA

La inteligencia artificial busca desarrollar sistemas capaces de ejecutar tareas asociadas con el razonamiento, el aprendizaje y la toma de decisiones. Entre sus principales campos se encuentran los sistemas expertos, las redes neuronales artificiales, los algoritmos genéticos,

la lógica difusa, los sistemas híbridos, el aprendizaje automático, el procesamiento del lenguaje natural, la robótica, la visión por computadora y la representación del conocimiento. Para el desarrollo de sistemas de reconocimiento facial, resultan especialmente relevantes los siguientes campos:

- **Sistemas expertos:** son programas diseñados para reproducir procesos de decisión similares a los realizados por una persona especializada en un área determinada.
- **Redes neuronales artificiales:** son modelos computacionales inspirados en el funcionamiento de las redes biológicas. Permiten identificar y aprender patrones complejos a partir de los datos.
- **Aprendizaje automático:** comprende métodos que permiten a los sistemas mejorar su desempeño a partir de la información disponible, sin necesidad de definir manualmente todas las reglas de funcionamiento. Incluye el aprendizaje supervisado, el aprendizaje no supervisado, el aprendizaje profundo y el aprendizaje por refuerzo.
- **Representación del conocimiento y razonamiento automático:** reúne métodos destinados a organizar, almacenar y procesar información para resolver problemas y apoyar la toma de decisiones.

Existen dos enfoques principales para representar el conocimiento mediante IA: el simbólico y el conexionista. La IA simbólica utiliza reglas explícitas y símbolos comprensibles para representar conceptos y establecer relaciones entre ellos. Este enfoque emplea bases de conocimiento desarrolladas por especialistas [31][32].

Por otro lado, la IA conexionista se fundamenta en representaciones distribuidas, como las utilizadas por las redes neuronales artificiales. En este caso, el conocimiento se almacena mediante conexiones entre nodos. A diferencia de la IA simbólica, su interpretación puede resultar más compleja debido a que las relaciones aprendidas no se expresan mediante reglas fácilmente comprensibles.

### 2.2.6. Aprendizaje y razonamiento

En la IA simbólica, el razonamiento se basa en reglas lógicas y conocimientos previamente definidos. Este enfoque permite formular hipótesis y aplicar procedimientos deductivos. En cambio, la IA conexionista emplea un proceso inductivo y asociativo fundamentado en la experiencia. Mediante algoritmos de optimización, el sistema ajusta las conexiones entre nodos para identificar patrones y establecer relaciones a partir de los datos [31][33].

Ambos enfoques presentan ventajas y limitaciones. La IA simbólica ofrece un mayor nivel de interpretabilidad y resulta adecuada para tareas que requieren razonamiento lógico o causal. Por su parte, la IA conexionista presenta una mayor flexibilidad y capacidad para aprender a partir de grandes volúmenes de datos. Por esta razón, se utiliza ampliamente en tareas como el reconocimiento de patrones, la clasificación de imágenes y el procesamiento del lenguaje natural [33].

Los dos enfoques pueden complementarse mediante arquitecturas híbridas, en las que se integran mecanismos de razonamiento simbólico y métodos de aprendizaje conexionista. De esta manera, se busca reducir las limitaciones que presenta cada alternativa cuando se utiliza de forma independiente [33].

#### 2.2.6.1. Simulación de procesos cognitivos humanos

La inteligencia artificial busca reproducir funciones asociadas con la percepción, la identificación de patrones, el aprendizaje, el razonamiento y la toma de decisiones. Para ello,

se emplean algoritmos capaces de analizar grandes cantidades de datos, reconocer regularidades y generar respuestas de manera autónoma. Estos procesos permiten simular, en determinados contextos, la forma en que los seres humanos identifican estímulos y toman decisiones [34][35].



Fig.3 Técnicas de inteligencia artificial para Simulación de procesos cognitivos humanos

Como se muestra en la fig.3, existen diferentes técnicas y enfoques aplicados a la simulación de procesos cognitivos. Entre los principales se encuentran los siguientes:

- **Aprendizaje automático y aprendizaje profundo:** permiten que los sistemas aprendan a partir de los datos y mejoren su desempeño mediante la experiencia [36].
- **Redes neuronales:** son modelos inspirados en el funcionamiento del cerebro humano y se utilizan en tareas como el reconocimiento de imágenes, sonidos y otros estímulos sensoriales [35].
- **Procesamiento sensorial:** utiliza sensores y algoritmos para analizar información visual, auditiva o táctil, con el propósito de reproducir determinados procesos de percepción [37].

Estas técnicas permiten desarrollar aplicaciones orientadas al reconocimiento visual y auditivo, la detección de objetos camuflados y el aprendizaje autónomo. En el campo de la robótica, por ejemplo, se integran sensores y algoritmos de aprendizaje para mejorar la percepción del entorno, aunque todavía existen limitaciones frente a las capacidades humanas [30][38][39].

## 2.2.7. Cognición artificial en la implementación de sistemas de reconocimiento facial.

La cognición artificial lleva a cabo un pilar fundamental en el desarrollo de sistemas de reconocimiento facial, destacándose en aspectos tales como:

### 2.2.7.1. Simulación de procesos cognitivos

Los sistemas de inteligencia artificial buscan integrar determinadas capacidades asociadas con la cognición humana. En el reconocimiento facial, estas capacidades incluyen la percepción de rostros, la identificación de expresiones y la toma de decisiones en contextos variables. Su implementación permite desarrollar sistemas capaces de procesar información visual y responder ante escenarios complejos [40].

### **2.2.7.2. Interdisciplinariedad y optimización Algorítmica**

El desarrollo de sistemas de reconocimiento facial requiere la integración de disciplinas como la inteligencia artificial, la ciencia cognitiva y la psicología. Estas áreas aportan fundamentos para mejorar los algoritmos utilizados en la identificación de rostros y en el análisis de expresiones faciales [30].

### **2.2.7.3. Modelos Cognitivos en reconocimiento facial**

Existen modelos que representan el reconocimiento de identidades y el análisis de expresiones como procesos diferenciados dentro de un mismo sistema cognitivo. Estos enfoques permiten codificar las características del rostro e identificar a una persona sin confundir dicho proceso con el reconocimiento de sus expresiones faciales [41].

Estos modelos pueden aplicarse en áreas como la interacción humano-computadora, la educación, la seguridad y la medicina. Mediante el desarrollo de mecanismos cognitivos, los sistemas pueden analizar con mayor precisión la información visual y responder de forma adecuada dentro de contextos específicos [27].

### **2.2.8. Principios del Reconocimiento Facial**

La manera para el reconocimiento facial se encuentra formada por etapas, las cuales son repetidas en varios estudios y son:

- **Detección de rostro:** Aquí se valida el rostro humano en una imagen o en un video, este paso toma relevancia al momento de determinar la región facial en la cual se trabajará después ya que sin un reconocimiento preciso las siguientes etapas no se pueden desarrollar de forma correcta [42].
- **Alineación de rostro:** Aquí se trata la imagen facial obtenida anteriormente para normalizarla, lo cual puede incluir un proceso de cambiar la imagen a una escala de grises, corrección de tamaño y manejar la orientación del rostro, buscando así que los rasgos faciales estén en posiciones coherentes para facilitar la extracción de sus particularidades faciales [42].
- **Extracción de características:** Se adquiere las características faciales como distancia entre ojos, la extensión de la nariz, tamaño de la mandíbula incluso la profundidad de las cuencas oculares, entre otros; dando como respuesta un vector de  $n$  características que refleja la imagen facial, siendo esta una etapa determinante del proceso, ya que dependiendo de la calidad de estas características obtenidas se va a determinar el éxito del reconocimiento o no [42].
- **Comparación o reconocimiento:** Se compara los vectores obtenidos en la extracción de características con los datos almacenados en una base de datos, con el objetivo de identificar y así validar la identidad de la persona [42].

Para representar el rostro de manera matemática se tienen unos enfoques principales como son:

#### **2.2.8.1. Espacio de características multidimensional**

Un rostro puede representarse como un punto dentro de un espacio multidimensional. En este espacio, cada dimensión corresponde a una característica extraída por el modelo. La

distancia entre dos puntos permite estimar el grado de similitud entre dos rostros: cuanto menor sea la distancia entre sus representaciones, mayor será su semejanza [43].

### 2.2.8.2. Vectores de características

Los vectores de características permiten representar numéricamente la información relevante de un rostro y facilitan su comparación mediante procedimientos matemáticos. Existen diferentes métodos para generar estas representaciones. Entre ellos se encuentran *eigenfaces* y *fisherfaces*, que proyectan las imágenes en espacios de menor dimensión; *local binary patterns*, que codifica texturas locales; y FaceNet, que genera *embeddings* faciales mediante redes neuronales profundas [40].

### 2.2.8.3. Principio de disparidad

Este principio busca representar un rostro mediante un conjunto reducido de características relevantes. El objetivo es conservar la información necesaria para diferenciar una identidad de otra, evitando el uso de datos redundantes o poco significativos [40].

### 2.2.8.4. Reducción de dimensiones

La reducción de dimensiones consiste en proyectar una señal visual de alta dimensionalidad en un espacio de menor tamaño, conservando las características esenciales para realizar tareas de reconocimiento o identificación. El tiempo de procesamiento puede variar según el método empleado, el nivel de optimización y las características del equipo utilizado, como se muestra en la tabla 1 [28].

Tabla 1. Tiempo estimado para redimensionar imágenes de rostros.

Sistema / Método	Tiempo promedio por imagen
MTCNN	1.17 s
Detección de ojos/boca	0.26 s
Detección facial (sin optimizar)	0.23 s
Procesamiento por cuadro	0.047 s
Detector de puntos faciales	0.08 s
Detección facial (región de interés)	0.035 s
Detección facial (región de interés)	0.5 s (0.2-0.8 s)
Procesamiento total (landmarks, parpadeo)	<0.02 s

Los tiempos de detección facial varían considerablemente entre los diferentes métodos como muestra en la Tabla 1. Algunos sistemas requieren menos de 20 ms por imagen, mientras que otros pueden superar un segundo. Estas diferencias dependen del método utilizado, del nivel de optimización y del hardware disponible. Los sistemas que procesan únicamente regiones de interés pueden alcanzar tiempos inferiores a 50 ms por imagen, lo que favorece su aplicación en tiempo real. En cambio, los métodos menos optimizados pueden requerir entre 0,2 y 1,17 segundos por imagen [44].

### 2.2.9. Modelos de embeddings faciales

Los *embeddings* faciales son representaciones numéricas de alta dimensión que codifican las características distintivas de un rostro. Estos vectores permiten comparar diferentes

identidades dentro de un espacio métrico. Su dimensión puede variar, generalmente, entre 128 y 512 componentes [45].

Para generar los *embeddings*, se emplean redes neuronales convolucionales profundas (CNN o DCNN) previamente entrenadas para extraer características faciales. La imagen del rostro se introduce en la red y atraviesa varias capas de procesamiento hasta obtener un vector que representa sus rasgos distintivos. El entrenamiento puede realizarse mediante técnicas de aprendizaje métrico, como la función de pérdida *triplet loss*. Esta función busca aproximar los vectores correspondientes a una misma persona y separar aquellos pertenecientes a identidades diferentes [45].

El incremento excesivo del número de dimensiones puede elevar el costo computacional sin aportar una mejora proporcional en el rendimiento. Por esta razón, la cantidad de componentes debe seleccionarse de acuerdo con el propósito del sistema, el conjunto de datos disponible y los requerimientos computacionales [46].

### 2.2.9.1. Modelo FaceNet

Es un modelo ampliamente usado para la generación de los embeddings, este modelo crea vectores con dimensiones de 128 o de 512 dimensiones y a su vez usa la distancia euclidiana o la similitud coseno para realizar la comparación de las caras siguiendo los siguientes pasos:

- Detección facial: Aquí localiza y recorta la cara de la imagen
- Preprocesamiento: Este proceso normaliza las características de la imagen y el tamaño
- Extracción de embeddings: Este paso pasa a la imagen por una CNN para obtener su vector de características
- Comparación: Este proceso final el cual usan las distancias cosenoidal o euclidiana para comparar embeddings entre rostros

Facenet cuenta con ventajas frente a otros modelos ya que cuenta con una precisión superior a otros algoritmos, de manera importante al contar con más imágenes por persona; se tiene que con 40 imágenes por persona se alcanza un 97% de precisión, mientras que otros algoritmos tienen una precisión por debajo del 50% con esta misma cantidad de fotos [47].

Facenet logra ser versátil y de fácil uso ya que los embeddings generados por este se pueden usar directamente en tareas de verificación, identificar e incluso en agrupamiento facial empleando otros métodos adicionales como el K-NN o algoritmos de clustering y estas no tienen la necesidad de implementar capas intermedias específicas a cada tarea logrando así simplificar su integración en diferentes aplicaciones [48].

### 2.2.9.2. Principios de triplet-loss

La función de pérdida triplet loss se fundamenta en las distancias entre muestras dentro de un espacio de embeddings. Su objetivo es mejorar la separación entre identidades, de manera que las representaciones correspondientes a una misma persona permanezcan próximas entre sí y aquellas pertenecientes a personas diferentes se mantengan alejadas [37].

Se encuentra formada por tres elementos: anchor, positivo y negativo; las cuales representan la imagen de referencia, la imagen del mismo individuo en un ancho referencial y la imagen de una persona diferente respectivamente; con lo cual se obtiene una función que obliga al

modelo que la distancia entre el anchor y el positivo sea mucho menor que entre el anchor y el positivo, por un margen que defina el usuario [37].

Para optimizar este aprendizaje de distancia facial se tiene una discriminación de identidades en la que el triplet loss mantiene separadas las identidades de unas personas con otras logrando así optimizar la discriminación entre identidades faciales, logrando generalizar de mejor manera a nuevos datos que este no depende de clases fijas sino depende de las relaciones de distancia y es utilizado por el modelo FaceNet [32].

#### **2.2.10. Técnicas de Fine-Tuning en modelos preentrenados**

El *fine-tuning* es una técnica de aprendizaje por transferencia que consiste en adaptar un modelo previamente entrenado mediante un proceso adicional de entrenamiento con datos específicos. Este procedimiento permite aprovechar el conocimiento adquirido por el modelo durante su entrenamiento inicial y ajustarlo a las características de una nueva tarea [49].

Esta técnica resulta especialmente útil cuando se dispone de un conjunto de datos reducido. En lugar de entrenar un modelo desde cero, se modifican determinados parámetros del modelo preentrenado. De esta manera, se reduce el tiempo de entrenamiento y se optimiza el uso de los recursos computacionales. Además, el modelo puede aprender características más específicas relacionadas con el problema analizado [48][51].

En el ámbito del reconocimiento facial, el *fine-tuning* permite adaptar modelos generales a tareas concretas. Algunos ejemplos incluyen el reconocimiento de identidades, el análisis de expresiones y la identificación de determinados patrones faciales. Modelos como MobileFaceNet, FaceNet e IRSE-50 pueden beneficiarse de este proceso cuando se aplican en escenarios específicos [52].

#### **2.2.11. Parámetros de ajustes en el Fine-Tuning**

Con el fine-tuning se va a ajustar algunos parámetros de modelos ya entrenados anteriormente incluye tanto hiperparámetros como también los parámetros que el modelo tiene internamente; los más comunes suelen ser: pesos y biases, algunas capas determinadas, learning rate, Batch size, número de epochs, algoritmos de optimización y los function [37].

- Pesos y bases de las capas: Mientras se ejecuta esta técnica se puede modificar los pesos y bases de cualquiera de las capas que el modelo tenga, pero es recomendado modificar solo las capas superiores o solo la capa de salida dependiendo de que estrategia se emplee y para ello lo más usual es congelar las capas de extracción de características y ajustar las capas finales [37].
- Capas específicas: Al modificar capas superiores se le conoce como Top\_K, a las capas finales se le conoce como Task heads y al cambiar todo un conjunto de parámetros se le conoce como Full fine-tune [37].

En el área de los hiperparámetros de entrenamiento se tiene:

Learning rate: Esta es la tasa de aprendizaje y determina el tamaño de paso en la actualización de pesos, cuando se hace fine-tune con una base de datos menor se recomienda disminuir este parámetro.

- Batch size: Es el tamaño de lotes y nos indica la cantidad de muestras empleadas al momento de actualizar los parámetros.

- Número de epochs: Nos indica cuantas veces el modelo recorrerá un conjunto de datos completo mientras se realiza el entrenamiento.
- Algoritmo de optimización: Aquí se elige que optimizadores se van a emplear, pueden ser el SGD o el Adam que sirven para actualizar los parámetros.
- Loss function: Es la función de pérdidas que es ajustada para adaptarse a la tarea que se va a realizar, y nos indica que tan mal el modelo se encuentra prediciendo [45][53][54].

Aplicar esta técnica en una base de datos personalizada mejora incluso el accuracy promedio con mejoras que van del 0.5% al 2.4% y dependiendo de los parámetros ajustados se ha logrado mejorar hasta un 18% [45].

## CAPÍTULO III

### 3. METODOLOGÍA

#### 3.1. Tipo de Investigación

El presente estudio adoptó un enfoque cuantitativo y experimental. Se evaluó el rendimiento del sistema de reconocimiento facial mediante la tasa de falsa aceptación (FAR). Para ello, se modificó el modelo preentrenado FaceNet mediante la aplicación de la técnica de *fine-tuning*. A partir de este proceso, se obtuvo un modelo ajustado que permitió analizar el efecto de las variaciones controladas en las condiciones de captura sobre el rendimiento del sistema.

#### 3.2. Diseño de Investigación

Para llevar a cabo la experimentación se estableció un modelo experimental factorial de tres factores. Evaluando la incidencia de las variables independientes sobre el rendimiento del sistema de reconocimiento facial (modelo ajustado). Específicamente se evaluó el efecto del brillo, contraste, y oclusión sobre la métrica FAR. Para esto se utilizaron distintos niveles de cada variable, simulando las condiciones de operación reales.

El proceso se llevó a cabo en tres fases principales:

- **Línea base:** en esta etapa se utilizó el modelo de reconocimiento facial (basado en redes neuronales convolucionales) de Google llamado FaceNet. A partir de este se generaron los embeddings de las imágenes del conjunto de prueba y se obtuvo una métrica FAR, permitiendo establecer una referencia inicial del desempeño del sistema sin fine tuning.
- **Intervención experimental:** el modelo base fue reentrenado con el fin de lograr que identificara de manera específica los rostros de los estudiantes. Para esto se modificaron los pesos del modelo, parámetros de alineamiento facial, umbrales, filtros de calidad, detección y estabilidad temporal. Una vez obtenido el modelo ajustado se procedió a aplicar las variaciones controladas en las condiciones de captura de las imágenes: brillo, contraste, y oclusión.
- **Evaluación comparativa:** finalmente se generaron nuevamente los embeddings faciales a partir del modelo ajustado y de las variaciones controlados de captura. Con estos datos se analizó la variación que producía cada factor experimental en la métrica de rendimiento FAR. Los resultados obtenidos fueron analizados mediante técnicas estadísticas, incluyendo análisis de varianza (ANOVA) y correlación de Spearman.

#### 3.3. Variables de estudio

- **Variable Dependiente:** Tasa de Falsa Aceptación (FAR)
- **Variables Independientes:** Nivel de brillo, Nivel de contraste, Nivel de oclusión.

#### 3.4. Técnicas de recolección de Datos

La recolección de datos durante la investigación se estableció con el fin de generar un data set robusto con fotografías de calidad, que permitieran reentrenar de manera adecuada el modelo base de reconocimiento facial.

##### 3.4.1. Protocolo de captura (Data set fotográfico)

Con el fin de garantizar la calidad de las imágenes y controlar el sesgo en el proceso de recolección se estableció un protocolo para la toma de las fotografías. En primera instancia

el proceso se llevó a cabo en un ambiente con condiciones de iluminación controladas como: luz blanca natural, uniforme, y frontal. Evitando las sombras y reflejos, lo que permitió una captura óptima de los rasgos faciales como en la Fig.4.



*Fig.4* Rostros de estudiantes utilizados para entrenamiento de modelo de reconocimiento facial

De igual manera, se utilizó un fondo blanco y se aleatorizó la distancia entre la cámara y los estudiantes con el objetivo de simular condiciones reales de uso. Esta implementación de múltiples distancias permite que el modelo aprenda representaciones faciales más robustas ante cambios de escala, mejorando su desempeño en escenarios prácticos. Finalmente, se conservó el data set en una definición de 1080 p.

### 3.4.2. Variabilidad controlada de las capturas

Por otro lado, para la obtención de un data set representativo se realizaron varias fotografías a los estudiantes en distintas posturas Fig.5. Esto con el fin de poder simular condiciones de operación reales. Las fotografías se generaron con variaciones controladas en las posiciones (rango de 15 grados de rotación en cada eje) y distintas expresiones faciales (neutra, sonriendo, gesticulando). Además, a los estudiantes que generalmente usan gafas se les realizó la captura de imagen con y sin el uso de estas.



*Fig.5* Conjunto de imágenes de estudiantes para entrenamiento de modelo de reconocimiento facial

### 3.4.3. Almacenamiento y organización

Las imágenes fueron almacenadas en formato JPEG sin compresión adicional para preservar la calidad original. Cada archivo fue etiquetado siguiendo una nomenclatura alfanumérica sistemática: [ID\_estudiante]\_[número\_secuencial]\_[variación].jpg, facilitando así la trazabilidad y organización del dataset.

### **3.5. Fine-tuning del modelo de FaceNet**

El modelo utilizado para el proceso fue el Inception-ResNet V1, en su configuración preentrenada en VGGFace2. Debido a la finalidad del estudio y al tamaño de data set establecido se recurrió a la técnica de aprendizaje por transferencia. A través de esta técnica durante el proceso de fine tuning se modificaron las capas finales (encargadas de producir el embedding facial) de la arquitectura del modelo, mientras que las capas iniciales se mantuvieron sin cambios.

Para el ajuste fino de estas capas finales se utilizó el optimizador Adam (tasa de aprendizaje de 0.0001) y un weight decay (0.00001). De esta manera se garantizó que el modelo no sufra una destrucción del conocimiento previo durante el entrenamiento. Además, se utilizó la función Triplet Loss ( $\alpha = 0.2$ ) forzando al modelo a organizar el espacio de embeddings, de manera que las representaciones de una misma identidad converjan entre sí, mientras que las representaciones de identidades distintas se mantengan separadas por al menos el margen establecido.

Durante cada época de entrenamiento se generaron dinámicamente combinaciones de triplets a partir del dataset disponible, y cada imagen fue previamente procesada mediante MTCNN para detección, alineación y normalización facial, seguida de un redimensionamiento a  $160 \times 160$  píxeles acorde con los requisitos de entrada de la arquitectura FaceNet.

### **3.6. Implementación del sistema de reconocimiento facial**

Una vez obtenido el modelo ajustado se procede a establecer un sistema de reconocimiento facial funcional, el cual, permitirá el reconocimiento de los rostros de los estudiantes. De igual manera, esta implementación se realizó en conjunto con un mecanismo de actualización automático con el fin de mantenerlo en el tiempo.

#### **3.6.1. Diseño del sistema de reconocimiento facial**

Se estableció un pipe line (utilizando Python 3.8) para el procesamiento de las imágenes con el fin de estandarizarlas. Cada fotografía pasó a través de una red neuronal convolucional multitarea (MTCNN) que detectaba los puntos clave del rostro. Posteriormente se implementó un proceso de filtrado en base a criterios de calidad. Estos fueron: puntaje de confianza mínimo de 0.95 en la detección MTCNN, nivel de nitidez adecuado (varianza Laplaciana para medir el desenfoque), y el índice de iluminación para eliminar imágenes subexpuestas y sobreexpuestas. Se establecieron criterios numéricos claros: un nivel mínimo de nitidez mayor o igual a 120, un índice mínimo de iluminación de 35 y un tamaño mínimo del recorte facial de  $150 \times 150$  píxeles, con el fin de asegurar suficiente información antes de redimensionar la imagen.

Además, se aplicaron criterios cualitativos como seleccionar el rostro de mayor tamaño en la imagen, descartar detecciones parciales o cercanas al borde del encuadre y eliminar imágenes con movimiento excesivo. Finalmente, las fotografías que pasaron los filtros de calidad fueron alineadas geométricamente utilizando las coordenadas de los ojos como punto de anclaje, esto con el fin de normalizar la posición de los rostros. Además, estas fueron recortadas a una resolución de  $160 \times 160$  píxeles, ya que, este tamaño es el requerido para el modelo de FaceNet.

### 3.6.2. Generación de los vectores (embeddings)

Una vez obtenido el data set estandarizado se procesaron todas fotografías a través de la red neuronal. Para esto se utilizó la arquitectura de FaceNet implementada en TensorFlow/Keras con el fin de vectorizar las imágenes. Es así como, las características faciales de alto nivel fueron obtenidas por el modelo y sus vectores con características de 128 dimensiones.

Posteriormente estos vectores fueron normalizados mediante L2 para asegurar comparaciones consistentes. Para esto se aplicó mediante (1).

$$\hat{x} = \frac{x}{\|x\|_2} \quad (1)$$

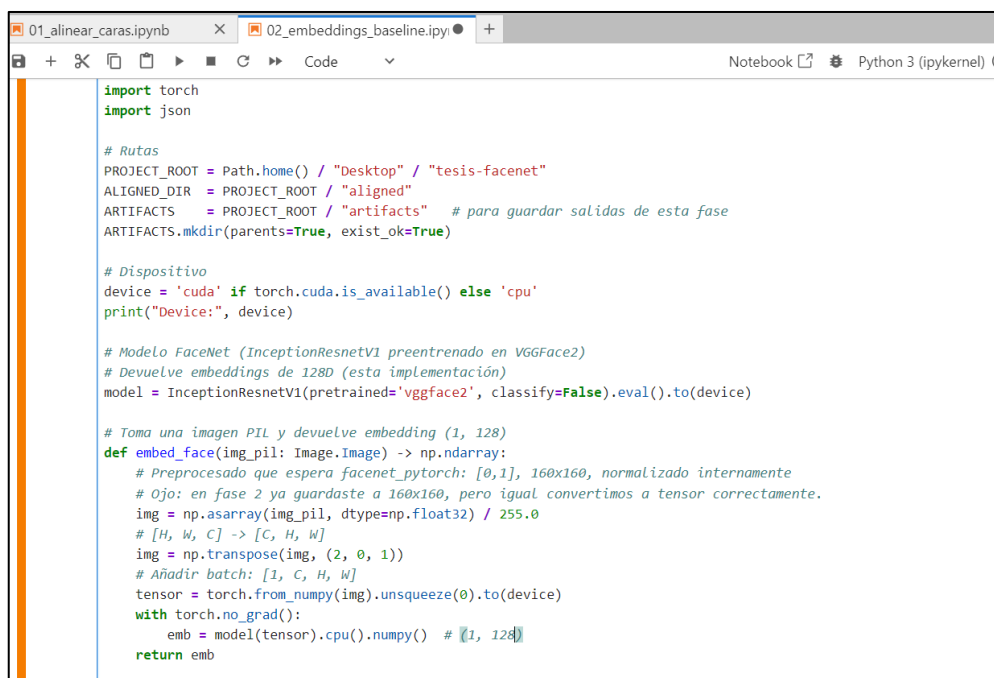
Donde:

$\hat{x}$  = *embedding normalizado*

$x$  = *embedding original de 128 dimensiones*

$\|x\|_2$  = *norma euclidiana (L2) del vector \*\*x\*\*, calculada como  $\sqrt{\sum_{i=1}^{128} x_i^2}$*

Esta normalización asegura que todos los vectores tengan magnitud unitaria, permitiendo que las comparaciones subsecuentes se basen exclusivamente en el ángulo entre vectores (similitud coseno) en lugar de su magnitud, lo cual mejora la estabilidad y consistencia de las mediciones de similitud. En total se generaron 210 embeddings para el modelo ajustado y el fragmento del código se muestra en la Fig.6.



```
import torch
import json

# Rutas
PROJECT_ROOT = Path.home() / "Desktop" / "tesis-facenet"
ALIGNED_DIR = PROJECT_ROOT / "aligned"
ARTIFACTS = PROJECT_ROOT / "artifacts" # para guardar salidas de esta fase
ARTIFACTS.mkdir(parents=True, exist_ok=True)

# Dispositivo
device = 'cuda' if torch.cuda.is_available() else 'cpu'
print("Device:", device)

# Modelo FaceNet (InceptionResnetV1 preentrenado en VGGFace2)
# Devuelve embeddings de 128D (esta implementación)
model = InceptionResnetV1(pretrained='vggface2', classify=False).eval().to(device)

# Toma una imagen PIL y devuelve embedding (1, 128)
def embed_face(img_pil: Image.Image) -> np.ndarray:
    # Preprocesado que espera facenet_pytorch: [0,1], 160x160, normalizado internamente
    # Ojo: en fase 2 ya guardaste a 160x160, pero igual convertimos a tensor correctamente.
    img = np.asarray(img_pil, dtype=np.float32) / 255.0
    # [H, W, C] -> [C, H, W]
    img = np.transpose(img, (2, 0, 1))
    # Añadir batch: [1, C, H, W]
    tensor = torch.from_numpy(img).unsqueeze(0).to(device)
    with torch.no_grad():
        emb = model(tensor).cpu().numpy() # [1, 128]
    return emb
```

Fig.6 Fragmento de código en Jupyter Notebook para la generación de embeddings faciales con FaceNet

### **3.6.3. Automatización de actualización del sistema**

Una vez que el sistema de reconocimiento se encuentra en funcionamiento se vuelve necesario la implementación de un sistema de actualización automático. Para esto se implementó un mecanismo de automatización basado en monitoreo de carpetas. De esta manera al detectar un cambio o adición el sistema ejecuta de manera autónoma la reprocesamiento de imágenes, la regeneración de embeddings y el recálculo de las representaciones de referencia asociadas a cada identidad registrada.

### **3.6.4. Estrategia de verificación y control en tiempo real**

Durante la fase de verificación en tiempo real, el sistema aplica una serie de restricciones antes de aceptar una identidad como válida. Estas condiciones fueron implementadas con el objetivo de mejorar la confiabilidad del proceso de reconocimiento facial y reducir la probabilidad de falsas aceptaciones.

En primer lugar, se valida la calidad del rostro detectado evaluando parámetros como nitidez, iluminación, tamaño del rostro y posición dentro del encuadre. Posteriormente se genera el embedding correspondiente y se calcula la similitud entre el vector obtenido y las representaciones almacenadas en la base de datos.

El proceso de verificación se basa en la comparación de la similitud obtenida con un umbral de decisión previamente definido. Adicionalmente, se evalúa la diferencia entre la mejor coincidencia y la segunda mejor coincidencia para evitar ambigüedades en la identificación.

### **3.6.5. Integración del sistema**

Finalmente, todos los módulos desarrollados fueron integrados en un sistema funcional capaz de operar en tiempo real. La arquitectura final del sistema incorpora los siguientes componentes principales:

- Captura de video en tiempo real.
- Detección y alineación facial.
- Filtrado de calidad de imágenes.
- Generación de embeddings faciales mediante FaceNet (modelo ajustado).
- Verificación de identidad mediante comparación de embeddings.
- Actualización automática del registro facial.

## **3.7. Población de estudio y tamaño de muestra**

Para armar el data set se utilizaron 440 fotografías de estudiantes de octavo semestre de la carrera de Ingeniería en Telecomunicaciones de la Universidad Nacional de Chimborazo. Por cada uno de ellos se obtuvo aproximadamente 40 fotos.

## **3.8. Hipótesis**

- Hipótesis Nula ( $H_0$ ): Las variaciones en las condiciones de captura (brillo, contraste y oclusión) no producen diferencias estadísticamente significativas en la Tasa de Falsa Aceptación (FAR) del sistema de reconocimiento facial.

- Hipótesis Alternativa ( $H_1$ ): Al menos una de las condiciones de captura (brillo, contraste u oclusión) produce diferencias estadísticamente significativas en la Tasa de Falsa Aceptación (FAR) del sistema.

### 3.9. Métodos de análisis, y procesamiento de datos.

#### 3.9.1. Cálculo de distancia entre embeddings

Para determinar el grado de similitud entre rostros se recurrió a la medición de la similitud coseno entre embeddings. Específicamente, para el estudio se utilizó la similitud coseno, la cual, permite establecer el grado de orientación compartida entre dos vectores dentro de un espacio multidimensional, independiente de su magnitud. El protocolo utilizado fue la validación cruzada k-fold con  $k = 5$ , con pares de anclaje positivo y negativo de las fotografías. Esto permitió la medición de la similitud coseno entre vectores, y el barrido de todos los umbrales posibles entre los vectores desde el cero hasta el uno con incrementos de 0.001. La primera se ejecutó a través de (2).

$$s(e_1, e_2) = \frac{e_1 \cdot e_2}{\|e_1\| \cdot \|e_2\|} = \frac{\sum_{i=1}^{128} e_{1i} \cdot e_{2i}}{\sqrt{\sum_{i=1}^{128} e_{1i}^2} \cdot \sqrt{\sum_{i=1}^{128} e_{2i}^2}} \quad (2)$$

Donde:

$s$  = similitud coseno entre dos embeddings

$e_1, e_2$  = embeddings normalizados de 128 dimensiones

$e_{1i}, e_{2i}$  = componente  $i$ -ésima de cada embedding

$i$  = índice dimensional (1 a 128)

A través de la medición de esta distancia se logró establecer si un embedding pertenecía o no a una persona específica. Para esto el resultado fue comparando con un umbral de decisión definido. En general, este procedimiento permitió obtener la distribución completa de distancias entre pares genuinos (misma identidad) y pares impostores (identidades diferentes), información que posteriormente fue utilizada para calcular la tasa de falsa aceptación (FAR).

#### 3.9.2. Análisis de métrica de rendimiento (FAR)

Una vez obtenido la distribución de los todos los umbrales se procedió a analizar la tasa de falsa aceptación (FAR). La cual permite cuantificar la proporción de pares impostores que fueron incorrectamente aceptados como genuinos para cada umbral evaluado. Esta se llevó a cabo mediante la aplicación de (3).

$$FAR = \frac{FP}{FP + VP} \quad (3)$$

Donde:

FAR = tasa de falsa aceptación

FP = número de falsos positivos (pares impostores incorrectamente aceptados)

VN = número de verdaderos negativos (pares impostores correctamente rechazados)

El cálculo de la FAR se realizó para cada uno de los umbrales evaluados dentro del proceso de validación cruzada. Los resultados obtenidos para cada pliegue fueron almacenados en un archivo en formato CSV (Comma-Separated Values), constituyendo el conjunto de datos final sobre el cual se realizó el análisis estadístico para la prueba de hipótesis.

### 3.9.3. Análisis estadístico ANOVA

Se aplicó el estadístico ANOVA con el fin de determinar si las condiciones de captura generaron un impacto significativo sobre la tasa de falsa aceptación (FAR) del modelo. A través de esta prueba se analizaron las diferencias significativas en las medias producidas de la variable dependiente (para el proceso se utilizó un nivel de significancia  $\alpha = 0.05$ ).

El modelo estadístico del ANOVA puede expresarse según (4):

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad (4)$$

Donde:

$Y_{ij}$  = representa el valor observado de la variable dependiente (FAR)

$\mu$  = corresponde a la media global del conjunto de datos.

$\alpha_i$  = representa el efecto del  $i$  –ésimo nivel del factor analizado.

$\varepsilon_{ij}$  = corresponde al término de error aleatorio asociado a cada observación.

El análisis se realizó bajo un diseño experimental factorial en el que cada factor fue evaluado en múltiples niveles previamente definidos. De esta manera se determinó la variabilidad total, entendiéndola como variabilidad atribuible a los factores experimentales y la variabilidad residual o error. A partir de esta descomposición se calculó el estadístico F, comparando la variabilidad entre grupos.

### 3.9.4. Análisis de correlación de Spearman

Para el análisis de la incidencia en cuanto a fuerza y dirección de las variables independientes (brillo, oclusión, contraste) en la métrica FAR se utilizó el método no paramétrico de correlación de Spearman. Este se calcula mediante (5).

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (5)$$

Donde:

$\rho$  = representa el coeficiente de correlación de Spearman.

$d_i$  = corresponde a la diferencia entre los rangos de las observaciones

$n =$  *representa el número total de observaciones analizadas.*

La significancia estadística de cada coeficiente se evaluó con un nivel de significancia  $\alpha = 0.05$ ; valores  $p < \alpha$  indican una relación estadísticamente significativa. Este análisis complementó los resultados del ANOVA al proporcionar información sobre la dirección y magnitud del efecto de cada factor sobre el rendimiento del sistema biométrico.

## CAPÍTULO IV

### 4. RESULTADOS Y DISCUSIÓN

En este capítulo se presentan los resultados obtenidos del análisis estadístico aplicado que se realizó a nuestra Tasa de Falsa Aceptación (FAR) del sistema de reconocimiento facial basado en el modelo FaceNet (VGG2), tomando como factores experimentales el brillo, el contraste y la oclusión. Para esta evaluación se utilizó un conjunto de 625 comparaciones faciales con un umbral de similitud de 0.25, dando como resultado un FAR global de 16.16%, con el cual se pudo observar el impacto directo de las condiciones analizadas.

#### 4.1. Resultados generales del análisis ANOVA

Se utilizó un análisis de varianza ANOVA de tres factores para evaluar el efecto del brillo, contraste y oclusión sobre la Tasa de Falsa Aceptación (FAR) del sistema de reconocimiento facial basado en FaceNet (VGG2).

Tabla 2. Resultados ANOVA

Fuente	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
Brillo	18.6464	4	4.6616	60.9552167	2.41E-43
Contraste	12.6304	4	3.1576	41.2888691	1.18E-30
Oclusion	6.5984	4	1.6496	21.5702174	1.15E-16
Error	46.8032	612	0.07647582		
Total	84.6784	624			

Los resultados de la Tabla 2 nos muestran que, los tres factores afectan de manera estadísticamente significativa el FAR ( $p < 0.001$ ). El brillo presentó el mayor valor  $F = 60.96$ , siendo el factor con mayor impacto en la variabilidad del FAR, seguido del contraste con  $F = 41.29$  y la oclusión con  $F = 21.57$ .

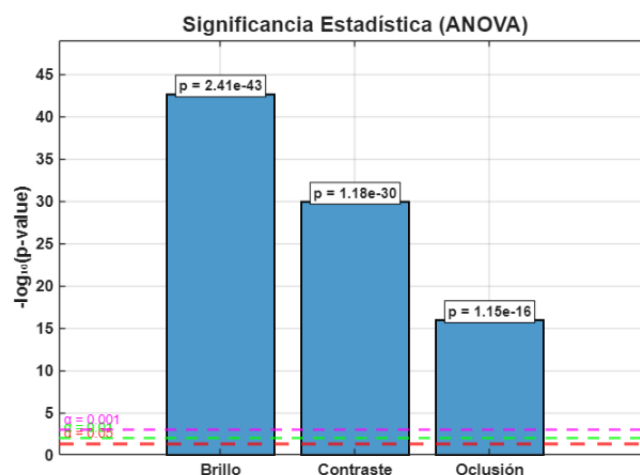


Fig.7 Significancia estadística ANOVA.

La Fig.7 muestra la significancia estadística de cada factor en escala logarítmica, dado que los valores  $p$  son extremadamente pequeños y no se podían observar de manera adecuada en escala lineal. Los tres factores superan ampliamente el nivel de significancia representado

por la línea entrecortada roja ( $\alpha = 0.05$ ), confirmando su importancia en el rendimiento del sistema.

## 4.2. Factor Brillo

Para el factor brillo se planteó la siguiente hipótesis:

$H_0$ : El nivel de brillo no afecta la Tasa de Falsa Aceptación (FAR). Las medias de FAR son iguales en todos los niveles de brillo.

$H_1$ : El nivel de brillo sí afecta el FAR. Al menos una media de FAR no es igual entre los niveles de brillo.

Con base en los resultados del ANOVA y el criterio de decisión establecido  $\alpha = 0.05$ , para el factor brillo se obtuvo lo siguiente:

*Tabla 3. Decisión de hipótesis para el factor brillo*

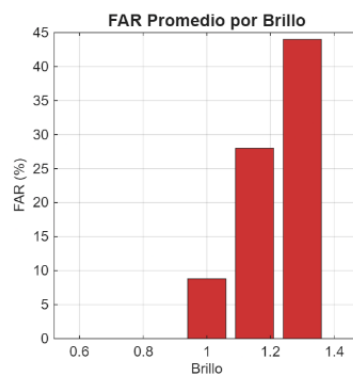
Factor: Brillo
Valor p: $2.41 \times 10^{-43}$
Decisión: Se rechaza la $H_0$
Conclusión: Se rechaza la hipótesis nula $H_0$ . El brillo sí afecta significativamente el FAR.

El valor p es menor que 0.05 ( $p < 0.001$ ), por lo tanto se rechaza la hipótesis nula y se acepta la hipótesis alternativa, confirmando que el brillo tiene un efecto significativo sobre la Tasa de Falsa Aceptación del sistema de reconocimiento facial basado en FaceNet.

El brillo evidenció ser el factor más dominante sobre el FAR. El FAR aumenta progresivamente con niveles más altos de brillo:

- Brillo bajo (0.70 – 0.85): FAR = 0%
- Brillo medio (1.00): FAR = 8.8%
- Brillo alto (1.15): FAR = 28%
- Brillo muy alto (1.30): FAR = 44%

Niveles elevados de brillo vistos en la Fig.8 producen sobreexposición en las imágenes faciales, reduciendo así la apreciación de las características distintivas y provocando que rostros diferentes presenten embeddings más similares, incrementando con esto la probabilidad de falsa aceptación.



*Fig.8 FAR promedio por brillo*

Además, se calculó la correlación de Spearman para cuantificar la fuerza y dirección de la relación entre el brillo y el FAR dando como resultado

*Tabla 4. Correlaciones de Spearman brillo*

Factor	$\rho$ (Rho Spearman)	Valor p
Brillo	0.445683621	$7.87 \times 10^{-32}$

La Tabla 4 confirma que el brillo es el factor más influyente sobre el FAR con  $\rho = 0.446$ , presentando una correlación positiva que indica que condiciones de mayor brillo incrementan la probabilidad de falsa aceptación en el sistema.

### 4.3. Factor Contraste

Para el factor contraste se planteó la siguiente hipótesis:

$H_0$ : El nivel de contraste no afecta la Tasa de Falsa Aceptación (FAR). Las medias de FAR son iguales en todos los niveles de contraste.

$H_1$ : El nivel de contraste sí afecta el FAR. Al menos una media de FAR no es igual entre los niveles de contraste.

Con base en los resultados del ANOVA y el criterio de decisión establecido  $\alpha = 0.05$ , para el factor contraste se obtuvo lo siguiente:

*Tabla 5. Decisión de hipótesis para el factor brillo*

Factor: Contraste
Valor p: $1.18 \times 10^{-30}$
Decisión: Se rechaza la $H_0$
Conclusión: Se rechaza la hipótesis nula $H_0$ . El contraste sí afecta significativamente el FAR.

El valor p es menor que 0.05 ( $p < 0.001$ ), por lo tanto se rechaza la hipótesis nula y se acepta la hipótesis alternativa, confirmando que el contraste tiene un efecto significativo sobre la Tasa de Falsa Aceptación del sistema de reconocimiento facial basado en FaceNet.

El contraste mostró un efecto significativo sobre el FAR, con un patrón similar al del brillo:

- Contraste bajo (0.800 – 0.975): FAR = 0 – 6.4%
- Contraste medio (1.150): FAR = 8.8%
- Contraste alto (1.325): FAR = 28%
- Contraste muy alto (1.500): FAR = 37.6%

Valores extremos de contraste mostrados en la Fig.9 exageran que las sombras y luces en el rostro, distorsionando las características faciales y dificultando la correcta discriminación entre individuos diferentes.

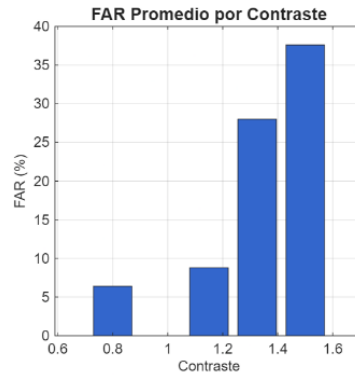


Fig.9 FAR promedio por contraste.

Además, se calculó la correlación de Spearman para cuantificar la fuerza y dirección de la relación entre el contraste y el FAR.

Tabla 6. Correlaciones de Spearman contraste

Factor	$\rho$ (Rho Spearman)	Valor p
Contraste	0.3473	$3.69 \times 10^{-19}$

La Tabla 6 confirma que el contraste es un factor influyente sobre el FAR con  $\rho = 0.347$ , presentando una correlación positiva que indica que condiciones de mayor contraste incrementan la probabilidad de falsa aceptación en el sistema.

#### 4.4. Factor Oclusión

Para el factor oclusión se planteó la siguiente hipótesis:

$H_0$ : El nivel de oclusión no afecta la Tasa de Falsa Aceptación (FAR). Las medias de FAR son iguales en todos los niveles de oclusión.

$H_1$ : El nivel de oclusión sí afecta el FAR. Al menos una media de FAR no es igual entre los niveles de oclusión.

Con base en los resultados del ANOVA y el criterio de decisión establecido  $\alpha = 0.05$ , para el factor oclusión se obtuvo lo siguiente:

Tabla 7. Decisión de hipótesis para el factor brillo

Factor: Oclusión
Valor p: $1.15 \times 10^{-16}$
Decisión: Se rechaza la $H_0$
Conclusión: Se rechaza la hipótesis nula $H_0$ . La oclusión sí afecta significativamente el FAR.

El valor p es menor que 0.05 ( $p < 0.001$ ), por lo tanto se rechaza la hipótesis nula y se acepta la hipótesis alternativa, confirmando que la oclusión tiene un efecto significativo sobre la Tasa de Falsa Aceptación del sistema de reconocimiento facial basado en FaceNet.

La oclusión presentó un comportamiento inverso al de los factores anteriores:

- Sin oclusión (0.0 – 0.1): FAR = 25.6 – 28.8%
- Oclusión baja (0.2): FAR = 13.6%
- Oclusión media (0.3): FAR = 12.8%

- Oclusión alta (0.4): FAR = 0%

De manera contraintuitiva, se observa en la Fig.10 que a mayores niveles de oclusión se asociaron con menores tasas de FAR. Esto se debe a que las oclusiones severas (más del 30% del rostro cubierto) impiden la extracción de suficientes características faciales discriminativas, resultando en valores de similitud más bajos que facilitan el rechazo correcto de impostores.

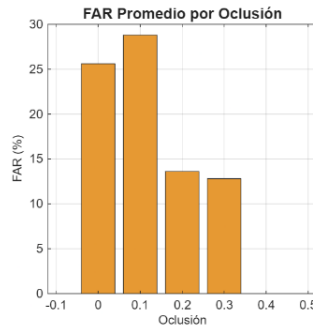


Fig.10 FAR promedio por oclusión.

Además, se calculó la correlación de Spearman para cuantificar la fuerza y dirección de la relación entre la oclusión y el FAR.

Tabla 8. Correlación de Spearman para el factor oclusión

Factor	$\rho$ (Rho Spearman)	Valor p
Contraste	0.3473	$3.69 \times 10^{-19}$

La correlación negativa con  $\rho = -0.258$  confirma el comportamiento observado en el ANOVA: a mayor oclusión, menor FAR. Este efecto se explica porque las oclusiones extremas dificultan la extracción de características faciales, resultando en menores valores de similitud que facilitan el rechazo de impostores.

#### 4.5. Comparación general de correlación

Para observar de manera general la fuerza de relación entre cada factor y el FAR, se presenta la siguiente figura comparativa.

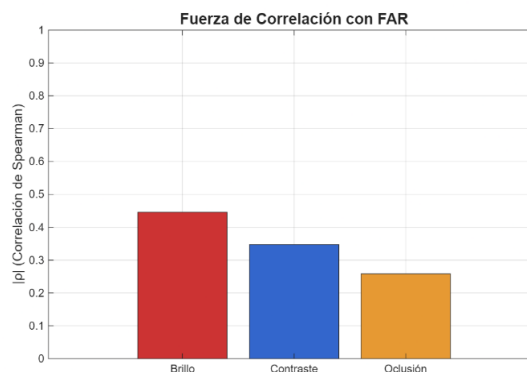


Fig.11 Fuerza de correlación con FAR.

La Fig.11 muestra que nuestro factor brillo es el factor con mayor correlación con el FAR, luego le sigue el contraste y finalmente la oclusión. Esto confirma que las condiciones de iluminación tienen un impacto más fuerte sobre el sistema que la cobertura parcial del rostro.

## 4.6. Discusión de Resultados

El análisis estadístico demuestra que las condiciones de control las que más se destacan son el brillo y el contraste ya que estos factores son los más importantes al momento que afectan significativamente el rendimiento del sistema FaceNet en términos de seguridad. Los hallazgos principales son:

1. **El brillo es el factor más crítico:** Con un valor F de 60.96 y una correlación de  $\rho = 0.446$ , el brillo tiene el mayor impacto en el FAR. Niveles altos de brillo ( $>1.15$ ) incrementan el FAR hasta un 44%, comprometiendo seriamente la seguridad del sistema.
2. **El contraste también tiene un impacto significativo:** Aunque menor que el brillo, el contraste extremo ( $>1.325$ ) puede elevar el FAR hasta un 37.6%, indicando que la sobreexposición afecta negativamente el rendimiento del sistema.
3. **La oclusión presenta un efecto protector:** Contrario a lo esperado, mayores niveles de oclusión reducen el FAR. Sin embargo, esto no implica una mejora en el rendimiento global del sistema, ya que las oclusiones también incrementarían la Tasa de Falso Rechazo (FRR), afectando la usabilidad.
4. **Todos los factores son estadísticamente significativos:** Con valores  $p < 0.001$ , se confirma que los tres factores deben ser considerados en el diseño e implementación de sistemas de reconocimiento facial en entornos reales.

Al comparar el modelo base con el modelo fine-tuneado, utilizando un umbral de similitud de 0.25, se observó una reducción en la Tasa de Falsa Aceptación, pasando de 17.28% a 16.16%. Esto indica que el proceso de fine-tuning permitió una mejor adaptación del modelo a las condiciones del conjunto de datos utilizado, logrando embeddings faciales más discriminativos y reduciendo la probabilidad de aceptar erróneamente a un impostor.

Estos resultados sugieren que, para aplicaciones prácticas de reconocimiento facial, es necesario implementar mecanismos de control de calidad de imagen que detecten y rechacen capturas con condiciones de iluminación inadecuadas antes de realizar la comparación biométrica.

# CAPÍTULO V

## 5. CONCLUSIONES Y RECOMENDACIONES

### 5.1. Conclusiones

Desarrollar, implementar y examinar un sistema de reconocimiento facial que se basó en un modelo pre entrenado configurado mediante técnicas de fine-tuning que se aplicó al entorno universitario para verificar la identidad de los usuarios, en base a este análisis experimental realizado se pudo comprobar que la calibración del modelo la selección óptima del umbral de decisión tiene influencia directa en cómo se comporta la tasa de falsa aceptación (False Acceptance Rate, FAR), indicador principal utilizada en este estudio.

Obtener resultados que demostraron que la optimización del modelo en conjunto con la implementación de restricciones relacionadas con la calidad de imagen (brillo, contraste, desenfoque y oclusión), permitió mejorar el control sobre la aceptación de impostores. El análisis estadístico ANOVA demostró que algunas variables independientes influyen significativamente el FAR, confirmando que la condición de captura no puede ignorarse en situaciones reales.

Concluir que el FAR es una métrica adecuada para poder evaluar la seguridad del sistema en entornos académicos reales ya que mide de forma directa el riesgo de suplantación de identidad. En contraste de métricas generales como AUC o F1 score el FAR permitió analizar de forma específica el impacto de como las condiciones del entorno afectan el rendimiento del modelo.

Determinar que la optimización del umbral de semejanza constituye un factor esencial para ajustar la seguridad y la funcionalidad. Umbrales demasiado bajos aumentan el riesgo de aceptación falsa, mientras que umbrales demasiados altos pueden cambiar la experiencia del usuario.

Confirmar que los modelos pre entrenados pueden ajustarse de forma efectiva a entornos locales al realizar el fine-tuning y el control de variables sin que sea necesario entrenar redes neuronales desde cero lo que evidencia una alternativa eficiente y variable para instituciones de educación superior.

### 5.2. Recomendaciones

Se recomienda que en futuras implementaciones del sistema se incorpore procesos sistemáticos de calibración del umbral de decisión antes de la implementación final, examinando e comportamiento del FAR bajo diferentes entornos reales de operación.

Se sugiere integrar métricas adicionales como false rejection rate (FRR) y Equal Error Rate (ERR) en futuras investigaciones, con la finalidad de analizar el balance entre seguridad y usabilidad del sistema.

También explorar técnicas innovadoras de mejorar de calidad de imagen como el ajuste automático de iluminación o detección de oclusiones mediante modelos de apoyo, para así de esta manera se reduce los cambios provocados por el entorno.

Finalmente se recomienda que la institución establezca políticas claras sobre la protección de datos biométricos que garanticen el cumplimiento de normativas de privacidad y así asegurar que la información facial que se almacena sea tratada bajo estándares de seguridad óptimos.

## BIBLIOGRAFÍA

- [1] R. X. Gao, J. Krüger, M. Merklein, H.-C. Möhring, y J. Váncza, «Artificial Intelligence in manufacturing: State of the art, perspectives, and future directions», *CIRP Ann.*, vol. 73, n.o 2, pp. 723-749, ene. 2024, doi: 10.1016/j.cirp.2024.04.101.
- [2] R. Gonzales Martinez y D.-M. van Dongen, «Deep learning algorithms for the early detection of breast cancer: A comparative study with traditional machine learning», *Inform. Med. Unlocked*, vol. 41, p. 101317, ene. 2023, doi: 10.1016/j.imu.2023.101317.
- [3] S. Li et al., Discrimination of Pb-Zn deposit types using the trace element data of galena based on deep learning, *Ore Geol. Rev.*, vol. 170, p. 106133, jul. 2024, doi: 10.1016/j.oregeorev.2024.106133.
- [4] J. F. AlSamhori et al., «Artificial intelligence for breast cancer: Implications for diagnosis and management», *J. Med. Surg. Public Health*, vol. 3, p. 100120, ago. 2024, doi: 10.1016/j.glmedi.2024.100120.
- [5] E. Halford y A. Webster, «Using chat GPT to evaluate police threats, risk and harm», *Int. J. Law Crime Justice*, vol. 78, p. 100686, sep. 2024, doi: 10.1016/j.ijlcj.2024.100686.
- [6] Q. Xu, Y. Zhou, y L. Wu, «Advancing tea detection with artificial intelligence: Strategies, progress, and future prospects», *Trends Food Sci. Technol.*, vol. 153, p. 104731, nov. 2024, doi: 10.1016/j.tifs.2024.104731.
- [7] Soto et al., “Implementación de reconocimiento facial en entornos académicos latinoamericanos,” *Revista Iberoamericana de Tecnología Educativa*, 2021.
- [8] Vargas et al., “Sistema biométrico para control de asistencia universitaria basado en redes neuronales,” *IEEE Latin America Transactions*, 2022.
- [9] Ranjan et al., “Fast and Accurate Deep Face Recognition Using Fine-Tuned Models,” *Pattern Recognition Letters*, 2020.
- [10] Liu, X., Wang, Y., & Huang, D. (2020). A Survey of Human Face Recognition Based on Deep Learning. *IEEE Access*.
- [11] de Abreu, R. A., Oliveira, L. S., & de Souza, A. O. (2021). Face Recognition Under Unconstrained Conditions: Advances and Perspectives. *Electronics*.
- [12] Akinlar, C., & Topal, C. (2022). Challenges in real-world face recognition: Illumination, pose and occlusion effects. *Signal Processing*.
- [13] Wang, M., & Yin, R. (2021). Biometric Performance Evaluation: A Review of Metrics and Measures. *IEEE Transactions on Biometrics, Behavior, and Identity Science*.
- [14] Castillo, C. D., & Sapiro, G. (2021). Reducing False Accept Rates in Face Recognition via Adaptive Thresholds. *Pattern Recognition*.
- [15] Ying, X., & Dai, F. (2022). Performance Optimization in Face Authentication Systems. *ACM Computing Surveys*.
- [16] Guo, Y., Zhang, L., & Lei, Z. (2021). A Review of Deep Learning in Face Recognition: Datasets, Models and Algorithmic Challenges. *IEEE Transactions on Neural Networks and Learning Systems*.
- [17] Smith, J. D., & Patel, K. (2022). Domain Adaptation for Deep Face Recognition in Real-World Applications. *International Journal of Computer Vision*.

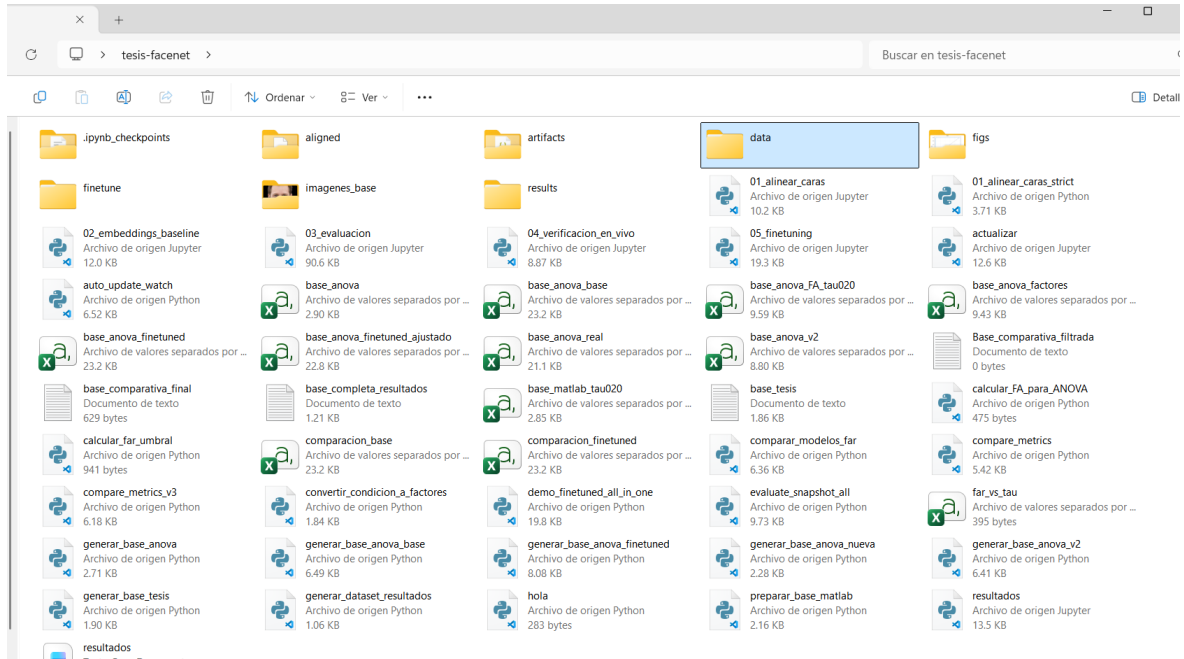
- [18] Singh, A., & Apende, A. (2020). Fine-tuning Deep Neural Networks for Face Recognition with Limited Training Data. *Journal of Visual Communication and Image Representation*.
- [19] M. Wang y W. Deng, “Deep Face Recognition: A Survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp. 9126–9157, 2022.
- [20] T. Schlett et al., “Face Image Quality Assessment: A Literature Survey,” *ACM Computing Surveys*, vol. 54, no. 8, 2022.
- [21] A. Drozdowski et al., “Demographic Bias in Biometrics: A Survey on an Emerging Challenge,” *IEEE Transactions on Technology and Society*, vol. 1, no. 2, pp. 89–103, 2020.
- [22] Universidad Técnica de Machala, “Aplicación de reconocimiento facial para control de asistencia,” *Repositorio Institucional UTMachala*, 2023.
- [23] Universidad Técnica del Norte, “Sistema de control de asistencia mediante reconocimiento facial,” *Repositorio UTN*, 2024.
- [24] Universidad Nacional de Chimborazo, “Diseño de un prototipo de reconocimiento facial para control de acceso,” *Repositorio UNACH*, 2023.
- [25] Y. Deng et al., “ArcFace: Additive Angular Margin Loss for Deep Face Recognition,” *IEEE TPAMI*, vol. 44, no. 10, pp. 5962–5979, 2022.
- [26] W. Duan, L. Zhang, D. Cao, X. Sun, X. Zhang, y L. Huang, «Reconstruction of significant wave height distribution from sparse buoy data by using deep learning», *Coast. Eng.*, vol. 194, p. 104616, dic. 2024, doi: 10.1016/j.coastaleng.2024.104616.
- [27] S. K. Baduge et al., «Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications», *Autom. Constr.*, vol. 141, p. 104440, sep. 2022, doi: 10.1016/j.autcon.2022.104440.
- [28] O. I. Odufisan, O. V. Abhulimen, y E. O. Ogunti, «Harnessing artificial intelligence and machine learning for fraud detection and prevention in Nigeria», *J. Econ. Criminol.*, vol. 7, p. 100127, mar. 2025, doi: 10.1016/j.jeconc.2025.100127.
- [29] A. Das y A. Bhattacharya, «Design and validation of a DC–DC converter-based inductive power transfer system with increased efficiency and reduced voltage stress across switches», *Measurement*, vol. 236, p. 115030, ago. 2024, doi: 10.1016/j.measurement.2024.115030.
- [30] F. Zhao, J. Li, L. Zhang, Z. Li, y S.-G. Na, «Multi-view face recognition using deep neural networks», *Future Gener. Comput. Syst.*, vol. 111, pp. 375-380, oct. 2020, doi: 10.1016/j.future.2020.05.002.
- [31] Deng et al., “ArcFace: Additive Angular Margin Loss for Deep Face Recognition,” *IEEE TPAMI*, 2021.
- [32] Wang & Deng, “Deep Face Recognition: A Survey,” *Neurocomputing*, 2021.
- [33] Nguyen et al., “Face Recognition-Based Attendance System,” *IEEE Access*, 2020.
- [34] Kaur et al., “Biometric-Based Student Attendance System Using Deep Learning,” *Sensors*, 2022.
- [35] Ullah et al., “Deep Learning-Based Identity Verification for Online Examinations,” *IEEE Access*, 2021.
- [36] Patel & Yagnik, “A Review on Face Recognition for Online Proctoring Systems,” *Multimedia Tools and Applications*, 2022.

- [37] Y. Qian, H. Cheng, J. Wang, J. Liang, W. Pedrycz, y C. Dang, «Grouping granular structures in human granulation intelligence», *Inf. Sci.*, vol. 382-383, pp. 150-169, mar. 2017, doi: 10.1016/j.ins.2016.11.024.
- [38] Q. Zou, Y. Wu, Z. Liu, W. Xu, y S. Gao, «Intelligent CAD 2.0», *Vis. Inform.*, vol. 8, n.o 4, pp. 1-12, dic. 2024, doi: 10.1016/j.visinf.2024.10.001.
- [39] G. Mai et al., «Towards the next generation of Geospatial Artificial Intelligence», *Int. J. Appl. Earth Obs. Geoinformation*, vol. 136, p. 104368, feb. 2025, doi: 10.1016/j.jag.2025.104368.
- [40] W. T. Lima Junior, R. A. Welter, W. P. Ferreira, R. F. Souza, y T. Eduardo, «Human robot interaction (HRI): An artificial cognitive autonomy approach to enhance Decision-Making», *Cogn. Syst. Res.*, vol. 91, p. 101336, jun. 2025, doi: 10.1016/j.cogsys.2025.101336.
- [41] N. N. Singh, D. P. Oswald, G. E. Lancioni, C. R. Ellis, M. Sage, y J. R. Ferris, «The neuropsychology of facial identity and facial expression in children with mental retardation», *Res. Dev. Disabil.*, vol. 26, n.o 1, pp. 33-40, ene. 2005, doi: 10.1016/j.ridd.2004.02.003.
- [42] J. Yang, R. Wang, X. Guan, M. M. Hassan, A. Almogren, y A. Alsanad, «AI-enabled emotion-aware robot: The fusion of smart clothing, edge clouds and robotics», *Future Gener. Comput. Syst.*, vol. 102, pp. 701-709, ene. 2020, doi: 10.1016/j.future.2019.09.029.
- [43] A. Phippen, «Artificial Intelligence», en *Encyclopedia of Libraries, Librarianship, and Information Science (First Edition)*, D. Baker y L. Ellis, Eds., Oxford: Academic Press, 2025, pp. 3-11. doi: 10.1016/B978-0-323-95689-5.00098-5.
- [44] J. Inbanathan, R. Anitha, P. G. Biswas, M. R. Sanjana, K. Ponnazhagan, y D. Thirumal Kumar, «Chapter 5 - Exploring imaging technologies and computational resources in stem cell research for regenerative medicine: A comprehensive review», en *Computational Biology for Stem Cell Research*, P. K. Raghav, R. Kumar, A. Lathwal, y N. Sharma, Eds., Academic Press, 2024, pp. 51-64. doi: 10.1016/B978-0-443-13222-3.00022-8.
- [45] J. C. Hung y J.-W. Chang, «Multi-level transfer learning for improving the performance of deep neural networks: Theory and practice from the tasks of facial emotion recognition and named entity recognition», *Appl. Soft Comput.*, vol. 109, p. 107491, sep. 2021, doi: 10.1016/j.asoc.2021.107491.
- [46] M. E. Sughrue, J. T. Yeung, y N. B. Dadario, «Chapter 8 - Machine learning and its utility in connectomic medicine», en *Connectomic Medicine*, M. E. Sughrue, J. T. Yeung, y N. B. Dadario, Eds., Academic Press, 2024, pp. 167-175. doi: 10.1016/B978-0-443-19089-6.00004-5.
- [47] L. Chen, Z. Tang, y X. Fan, «Enhancing video segmentation with contrastive self-supervised learning of distinctive class features for visually homogeneous frames», *Expert Syst. Appl.*, vol. 271, p. 126594, may 2025, doi: 10.1016/j.eswa.2025.126594.
- [48] A. Phippen, «Artificial Intelligence», en *Encyclopedia of Libraries, Librarianship, and Information Science (First Edition)*, D. Baker y L. Ellis, Eds., Oxford: Academic Press, 2025, pp. 3-11. doi: 10.1016/B978-0-323-95689-5.00098-5.

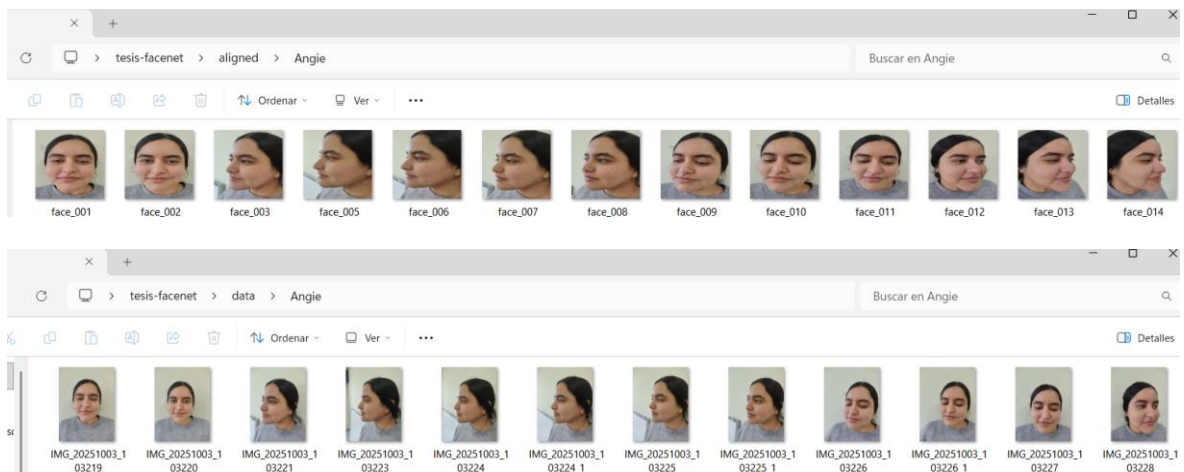
- [49] H. J. Park, K. W. Kim, y S. S. Lee, «Chapter 3 - Artificial intelligence in radiology and its application in liver disease», en *Artificial Intelligence, Machine Learning, and Deep Learning in Precision Medicine in Liver Diseases*, T.-H. Su y J.-H. Kao, Eds., Academic Press, 2023, pp. 53-79. doi: 10.1016/B978-0-323-99136-0.00002-7.
- [50] B. A. Kumar et al., «Hybrid CMNV2: DeepFake faces classification and recognition using deep learning methods», *Results Eng.*, vol. 28, p. 107513, dic. 2025, doi: 10.1016/j.rineng.2025.107513.
- [51] A. Ahmad, S. Jabbar, S. Akram, A. Paul, U. Raza, y N. M. Alshuqayran, «Enhancing ChatGPT's Querying Capability with Voice-Based Interaction and CNN-Based Impair Vision Detection Model», *Comput. Mater. Contin.*, vol. 78, n.o 3, pp. 3129-3150, mar. 2024, doi: 10.32604/cmc.2024.045385.
- [52] J. Yoon, A. Panizo-LLedot, D. Camacho, y C. Choi, «Triple-modality interaction for deepfake detection on zero-shot identity», *Inf. Fusion*, vol. 109, p. 102424, sep. 2024, doi: 10.1016/j.inffus.2024.102424.
- [53] M. Li, Z. Wang, R. Xu, X. Zhang, Z. Chen, y Q. Wang, «Advances in plasma-assisted ignition and combustion for combustors of aerospace engines», *Aerosp. Sci. Technol.*, vol. 117, p. 106952, oct. 2021, doi: 10.1016/j.ast.2021.106952.
- [54] H. Zhang, Q. Peng, J. Zhang, y P. Gu, «Planning for automatic product assembly using reinforcement learning», *Comput. Ind.*, vol. 130, p. 103471, sep. 2021, doi: 10.1016/j.compind.2021.103471.

# ANEXOS

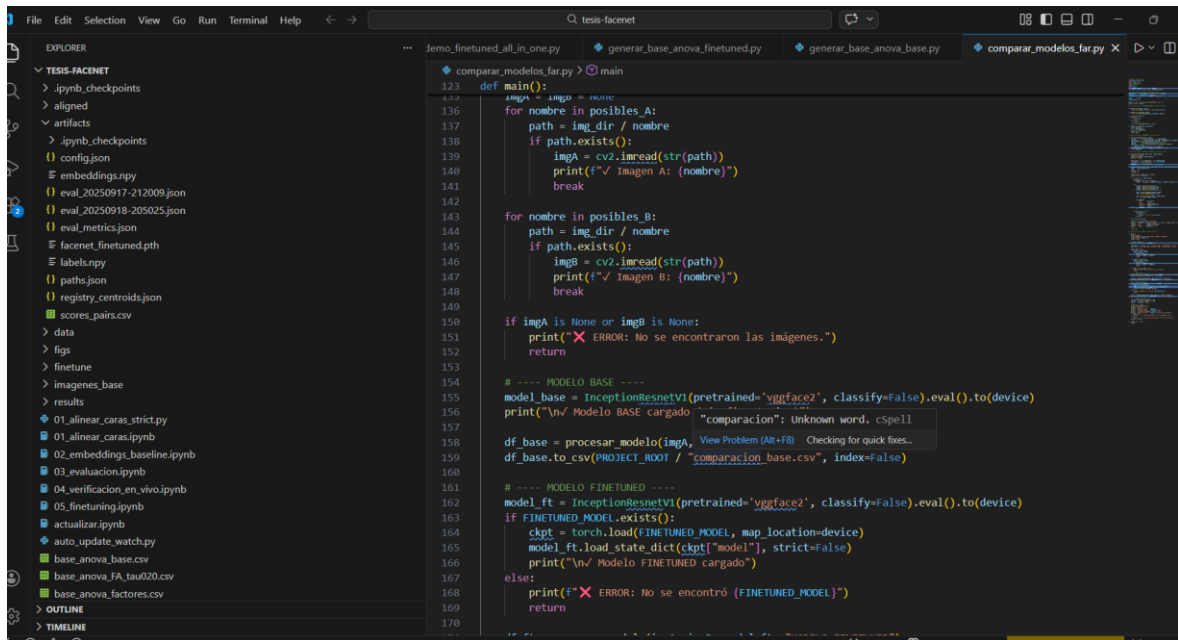
## Anexo 1. Carpeta con todas las dependencias y archivos necesarios para correr el programa.



## Anexo 2. Base de datos de los rostros humanos normalizados por MCNN.

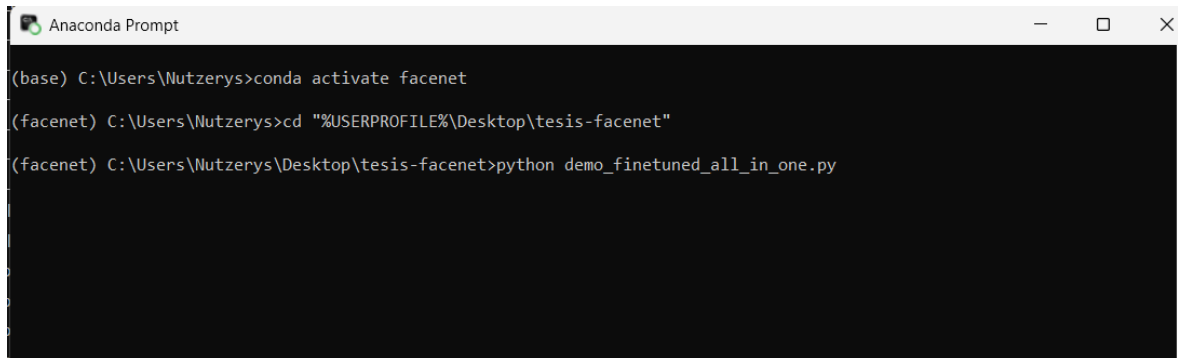


### Anexo 3. Fragmento del código integrando todas las tres fases del código en un solo archivo py.



```
def main():
    123
    124 imgA = None
    125 for nombre in posibles_A:
    126     path = img_dir / nombre
    127     if path.exists():
    128         imgA = cv2.imread(str(path))
    129         print(f"✓ Imagen A: {nombre}")
    130         break
    131
    132 for nombre in posibles_B:
    133     path = img_dir / nombre
    134     if path.exists():
    135         imgB = cv2.imread(str(path))
    136         print(f"✓ Imagen B: {nombre}")
    137         break
    138
    139 if imgA is None or imgB is None:
    140     print("✗ ERROR: No se encontraron las imágenes.")
    141     return
    142
    143 # ---- MODELO BASE ----
    144 model_base = InceptionResnetV1(pretrained='vggface2', classify=False).eval().to(device)
    145 print("\n✓ Modelo BASE cargado")
    146
    147 df_base = procesar_modelo(imgA, imgB, model_base)
    148 df_base.to_csv(PROJECT_ROOT / "comparacion_base.csv", index=False)
    149
    150 # ---- MODELO FINETUNED ----
    151 model_ft = InceptionResnetV1(pretrained='vggface2', classify=False).eval().to(device)
    152 if FINETUNED_MODEL.exists():
    153     ckpt = torch.load(FINETUNED_MODEL, map_location=device)
    154     model_ft.load_state_dict(ckpt["model"], strict=False)
    155     print("\n✓ Modelo FINETUNED cargado")
    156 else:
    157     print(f"✗ ERROR: No se encontró {FINETUNED_MODEL}")
    158     return
    159
    160
    161
    162
    163
    164
    165
    166
    167
    168
    169
    170
```

### Anexo 4. Comandos necesarios para utilizar el programa y funcionar la cámara en vivo.



```
(base) C:\Users\Nutzerys>conda activate facenet
(facenet) C:\Users\Nutzerys>cd "%USERPROFILE%\Desktop\tesis-facenet"
(facenet) C:\Users\Nutzerys\Desktop\tesis-facenet>python demo_finetuned_all_in_one.py
```