EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

TECNOLÓGICO NACIONAL DE MÉXICO

Tecnológico Nacional de México

Centro Nacional de Investigación y Desarrollo Tecnológico

Tesis de Doctorado

Método para la detección de suplantación de identidad en imágenes digitales del rostro

presentada por

M.C. Wendy Valderrama Cardenas

Como requisito para la obtención del grado de

Doctora en Ciencias de la Computación

Directora de tesis

Dra. Andrea Magadán Salazar

Cuernavaca, Morelos, México. Octubre de 2023.

cenidet
Centro Nacional de Investigación y Desarrollo Tecnológico

ESC\FORDOC09

Cuernavaca, Morelos, 07/agosto/2023

ASUNTO: ACEPTACIÓN DEL TRABAJO DE TESIS DOCTORAL

MARÍA YASMÍN HERNÁNDEZ PÉREZ
JEFA DEL DEPARTAMENTO DE CIENCIAS COMPUTACIONALES
PRESENTE

Los abajo firmantes, miembros del Comité Tutorial de la Tesis Doctoral de la alumna **WENDY VALDERRAMA CARDENAS** manifiestan que después de haber revisado su trabajo de tesis doctoral titulado **"MÉTODO PARA LA DETECCIÓN DE SUPLANTACIÓN DE IDENTIDAD EN IMÁGENES DIGITALES DEL ROSTRO",** realizado bajo la dirección de ANDREA MAGADÁN SALAZAR**,** el trabajo se ACEPTA para proceder a su impresión.

A T E N T A M E N T E
"Excelencia en Educación Tecnológica®
"Educación Tecnológica al Servicio de México"

ANDREA MAGADÁN SALAZAR
CENIDET

RAÚL PINTO ELÍAS
CENIDET

MARÍA YASMÍN HERNÁNDEZ PÉREZ
CENIDET

DANTE MÚJICA VARGAS
CENIDET

Osslan Osiris Vergara Villegas

OSSLAN OSIRIS VERGARA VILLEGAS
UNIV. AUTÓNOMA DE CIUDAD JUÁREZ

EDUCACIÓN | TECNOLÓGICO NACIONAL DE MÉXICO
19 SEP 2023
CENTRO NACIONAL DE INVESTIGACIÓN Y DESARROLLO TECNOLÓGICO
SERVICIOS ESCOLARES
RECIBIDO
EDJ

C.c.p.: María Elena Gómez Torres / Jefa del Depto. de Servicios Escolares
Carlos Manuel Astorga Zaragoza / Subdirector Académico
Expediente

Cuernavaca, Mor., 19/septiembre/2023
No. De Oficio: SAC/156/2023
Asunto: Autorización de impresión de tesis

WENDY VALDERRAMA CARDENAS
CANDIDATA AL GRADO DE DOCTORA EN CIENCIAS DE LA COMPUTACIÓN
P R E S E N T E

Por este conducto, tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado **"MÉTODO PARA LA DETECCIÓN DE SUPLANTACIÓN DE IDENTIDAD EN IMÁGENES DIGITALES DEL ROSTRO"**, ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

Esperando que el logro del mismo sea acorde con sus aspiraciones profesionales, reciba un cordial saludo.

A T E N T A M E N T E
Excelencia en Educación Tecnológica®
"Conocimiento y tecnología al servicio de México"

EDUCACIÓN | TECNOLÓGICO NACIONAL DE MÉXICO
CENTRO NACIONAL DE INVESTIGACIÓN Y DESARROLLO TECNOLÓGICO
SUBDIRECCIÓN ACADÉMICA

CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO

EDUCACIÓN | TECNOLÓGICO NACIONAL DE MÉXICO
20 SEP 2023
CENTRO NACIONAL DE INVESTIGACIÓN Y DESARROLLO TECNOLÓGICO
SERVICIOS ESCOLARES
RECIBIDO
EBN

C. c. p. Departamento de Ciencias Computacionales
Departamento de Servicios Escolares

CMAZ/lmz

Dedicatoria

Dedico este trabajo a Dios que nunca me ha abandonado.

A mis padres Victoria y José, quienes siempre me han motivado a cumplir mis metas y me han dado las fuerzas necesarias para seguir adelante con sus palabras de aliento en los buenos y malos momentos. Este logro también es de ustedes.

A Mirsha, mi compañera de aventuras en la vida, por estar siempre a mi lado en cada paso del camino, gracias por ser mi apoyo incondicional en esta travesía.

A mis hermanos Alejandro y Elvira, por todas sus palabras de aliento y creer en mí.

A mis sobrinos Lesli, Diego, Alex, Uriel, Santi, Mirshita. Gracias los quiero mucho.

A Carlos y Nayeli por su gran amistad y apoyo. En las adversidades se fortalecen los lazos de amistad, y es allí donde se descubren los amigos leales y genuinos.

Agradecimiento

Al Consejo Nacional de Humanidades Ciencias y Tecnologías (CONAHCYT) por el apoyo económico que me brindó durante mis estudios de doctorado.

A mi asesora la Dra. Andrea Magadán Salazar por su paciencia, su tiempo y el gran apoyo brindado.

A mi comité revisor por sus comentarios y observaciones que me ayudaron a culminar este proyecto.

A mis compañeros por sus recomendaciones.

Al Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET/TecNM) por las facilidades y el apoyo proporcionado para la terminación de esta tesis.

Finalmente agradezco a todos los que de forma directa o indirecta contribuyeron a la realización de este proyecto.

"Mientras los filósofos discuten sobre la posibilidad de la inteligencia artificial, los investigadores la construyen" - Judea Pearl.

Resumen

La progresión de la tecnología ha generado en la humanidad la necesidad de contar con herramientas que brinden seguridad en los sistemas que contengan su información personal o al acceder a algún hardware como una laptop, tableta, celular, entre otros, por lo que se ha recurrido al uso de la biometría facial como una forma de solución en tiempo real sin hacer uso de hardware específico. Sin embargo, la solución no está exenta de ataques de presentación que vulneran la confiabilidad en dichos sistemas, además de los desafíos propios de los entornos no controlados en los sistemas como son: fondo dinámico, mala iluminación, resolución variable y tamaño de la imagen, entre otros aspectos.

Por lo tanto, la presente tesis tiene como objetivo el desarrollo de un método de visión artificial que detecte posibles ataques de suplantación de identidad mediante imágenes faciales en un ambiente no controlado. La presente investigación plantea una propuesta de detección de ataques de suplantación sin muestra previa de la imagen a analizar. El método propuesto se integra de seis puntos clave:

1) Adquisición de la imagen. El método de detección de suplantación facial propuesto permite la adquisición de imágenes sin imponer restricciones en cuanto a la pose específica de la persona frente a la cámara, la iluminación intensa o la resolución precisa del dispositivo de captura.
2) Respecto a la iluminación se hace uso del algoritmo *retinex multi-escala,* previo a la localización del rostro para contemplar las características generales de la imagen y lograr que la detección de la cara sea precisa.
3) Se localiza el rostro y se recorta, posteriormente, se hace uso de espacios de color para resaltar características como saturación y brillo, los cuales tienen una influencia significativa al momento de hacer la diferencia entre un rostro real y la captura del rostro replicado por algún medio sintético.
4) Las imágenes resultantes son tratadas con el algoritmo de ruido de falta de uniformidad de la foto respuesta (PRNU por sus siglas en inglés), un descriptor de ruido de la lente de la cámara que en investigaciones del 2012 a la fecha se emplea para textura.
5) Con respecto al clasificador, se realizó una comparativa de cuatro clasificadores (máquinas de vector soporte, árboles de decisión, bosque aleatorio y perceptrón multi-capa) para validar el vector de características entre ellos se utiliza el clasificador de máquinas de vector soporte que es de los más utilizados en la literatura, mientras que en la propuesta final después de obtener los resultados se optó por utilizar bosque aleatorio.
6) El sistema es evaluado mediante tres casos de experimentación en los cuales se prueba el desempeño del clasificador, la generalización de método propuesto y la robustez ante dos tipos de ataque.

El porcentaje de error de detección de ataques de suplantación fue de 8.03 ± 59.73, de acuerdo con los diferentes bancos de imágenes que se utilizaron, mostrando que es posible realizar una detección de suplantación en ambientes hostiles.

Abstract

The progression of technology has generated in humanity the need to have tools that provide security in the systems that contain their personal information or when accessing hardware such as a laptop, tablet, cell phone, among others, which is why it has been resorted to. to the use of facial biometrics as a form of real-time solution without using specific hardware. However, the solution is not exempt from presentation attacks that violate the reliability of these systems, in addition to the challenges of uncontrolled environments in systems such as: dynamic background, bad lighting, variable resolution and image size, in other aspects.

Therefore, this thesis aims to develop an artificial vision method that detects phishing attacks using facial images in an uncontrolled environment. The present investigation proposes for the detection of detecting impersonation attacks without a previous sample of the image to be analyzed. The proposed method integrates six key points:

1) Image acquisition. The proposed facial spoofing detection method allows image acquisition without imposing restrictions regarding the specific pose of the person in front of the camera, strong lighting, or the precise resolution of the capture device.
2) Regarding lighting, the multi-scale retinex algorithm is used, prior to the location of the face to contemplate the typical characteristics of the image and achieve accurate face detection.
3) The face is located and cropped, subsequently, color spaces are used to highlight characteristics such as saturation and brightness, which have a considerable influence when making the difference between a real face and the capture of the face replicated by some synthetic medium.
4) The resulting images are treated with the photo response non-uniformity noise algorithm (PRNU), a camera lens noise descriptor used in research from 2012 to date to texture.
5) According to the classifier, a comparison of four classifiers (support vector machines, decision trees, random forest and multi-layer perceptron) was made to validate the vector of characteristics between them, the support vector machine classifier is used, which It is one of the most used in the literature, while in the final proposal, after obtaining the results, it was decided to use a random forest.
6) The system is evaluated through three experimentation cases in which the performance of the classifier, the generalization of the proposed method and the robustness against two types of attack are evaluated.

The impersonation attack detection error percentage was 8.03 ± 59.73, according to the different image banks that were used, showing that it is possible to detect impersonation in hostile environments.

Contenido

Lista de Figuras

Lista de Tablas

Glosario

2D	2 dimensiones
3D	3 dimensiones
APCER	*Attack Presentation Classification Error Rate* (Tasa de error de clasificación de presentación de ataques)
BPCER	*Bona fide Presentation Classification Error Rate* (Tasa de error de clasificación de presentación de buena fe)
BSIF	*Binarized statistical image features* (Funciones de imágenes estadísticas binarizadas)
Dlib	Biblioteca de software multiplataforma de propósito general escrita en el lenguaje de programación C ++
CNN	*Convolutional neural network* (Red neuronal convolucional)
ELBP	*Extended Local Binary Pattern* (Patrón Binario Local Extendido)
FPN	*Fixed pattern noise* (Ruido de patrón fijo)
GPU	*Graphics Processing Unit (Unidad de Procesamiento Gráfico)*
HTER	*Half Total Error Rate* (Tasa de error total medio)
HSV	*Hue, Saturation, Value* (Tono, saturación, valor)
IDA	*Image distortion analysis* (Análisis de distorsión de imagen)
JCR	*Journal Citation Reports*
OpenCV	Librería de visión artificial
LBP	*Local Binary Pattern* (Patrón binario local)
MSR	*Multi-scale Retinex* (Retinex multi-escala)
MLP	*Multi-Layer Perceptron* (Perceptrón multi capa)
MVS	*Máquinas de vectores de soporte*
PA	*Presentation Attacks* (Ataques de presentación)
PAD	*Presentation attack detection* (Detección de ataques de presentación)
PAI	*Presentation attack instruments* (Instrumentos de ataques de presentación)
PNU	*Photo non-uniformity* (Falta de uniformidad de la foto)
PRNU	*Photo response non-uniformity* (Falta de uniformidad en la fotorespuesta)
RGB	*Red, Green, Blue* (Rojo, Verde, Azul)
SSR	*Single scale Retinex* (Retinex de una sola escala)
MVS	Máquinas de vectores de soporte
YCbCr	Y: luminancia; Cb, Cr: características colorimétricas del color

CAPÍTULO 1
Introducción

La autenticación biométrica ha superado sistemáticamente a los esquemas basados en contraseñas convencionales [1], debido a que se puede identificar a las personas sin necesidad de presentar sus afiliaciones, pertenencias o información confidencial. Las identificaciones impresas han sido reemplazadas por identificaciones biométricas, que permiten probar "quién eres" sin tener que llevar una tarjeta u otro documento [2]. Otra solución de autenticación convencional es la contraseña, la cual no puede distinguir entre usuarios legítimos e impostores que accedieron al sistema de manera fraudulenta [1], [3], además de que existen posibilidades de olvidarla.

La biometría se ha investigado intensamente por su automatización, accesibilidad y precisión para satisfacer las crecientes demandas de seguridad de la vida diaria, de acuerdo a lo reportado en [4], el mercado de la biometría sin contacto en 2023 alcanzaría los 37 100 millones de USD mientras que, para 2028, el reconocimiento biométrico basado en el rostro alcanzará los 12 110 millones de USD debido a la prometedora aplicación en diversas categorías.

La biometría se ha implementado de manera efectiva en diversas áreas donde la seguridad es una prioridad. Por ejemplo, tarjetas de identificación personal para el *check-in* y *check-out* de algunos aeropuertos en el mundo, datos confidenciales de personas no autorizadas y validación de tarjetas de crédito. Para el reconocimiento y la autenticación se utilizan varias funciones biométricas, como la huella dactilar, el iris, la palma de la mano y el rostro [5].

La autenticación basada en el rostro proporciona una autenticación sin contacto más segura para el usuario que las huellas dactilares y el iris [6]. A demás, no necesitan de hardware específico para su funcionamiento. Sin embargo, los sistemas de verificación de identidad no pretenden determinar si la muestra biométrica presentada al sensor es real o falsa; consecuentemente, son vulnerables a los ataques de suplantación de identidad. Por ejemplo, el rostro de una persona puede ser obtenido a través de redes sociales u otros medios de difusión y llevar a cabo la suplantación de identidad mediante la superposición de una fotografía reproducida por un objeto sintético (*Tablet*, celular, fotografía impresa, etc.) frente a la cámara de la computadora o utilizando máscaras faciales en 3D. Por ello, el rostro es vulnerable a los ataques de suplantación, también conocidos como ataques de presentación (*Presentation Attacks* PA) [7].

Por consiguiente, en la actualidad es un tema importante debido a los avances tecnológicos y su impacto en la seguridad personal, la privacidad y la confianza en los sistemas de autenticación. A continuación, se destacan algunas de las razones por las cuales el tema es relevante [8]:

- Seguridad personal: La suplantación de identidad por medio del rostro puede tener consecuencias graves para las personas afectadas. Los ataques que involucran el robo de datos biométricos faciales pueden permitir que los delincuentes accedan a información personal, realicen transacciones fraudulentas o cometan delitos en nombre de la víctima [9].

- Protección de la privacidad: Los sistemas de reconocimiento facial están cada vez más presentes en la sociedad, como en aplicaciones móviles, cámaras de seguridad y sistemas de identificación. Es importante comprender los riesgos asociados con la suplantación de identidad facial para garantizar la protección de la privacidad y evitar el uso indebido de los datos biométricos [10].

- Confianza en los sistemas de autenticación: Con el aumento de la adopción de tecnologías biométricas, como el reconocimiento facial, es fundamental asegurar la integridad y la confianza en estos sistemas. Comprender las vulnerabilidades y los métodos de suplantación de identidad puede ayudar a desarrollar mejores soluciones de autenticación y fortalecer la seguridad de los sistemas [11].

- Prevención del fraude: La suplantación de identidad por medio del rostro se utiliza en muchos casos de fraude, como el acceso no autorizado a cuentas bancarias, la obtención de documentos falsos o el engaño en procesos de verificación de identidad. Conocer las técnicas utilizadas por los delincuentes y las medidas de protección pertinentes puede contribuir a prevenir y detectar este tipo de fraudes [12].

En resumen, la importancia de la suplantación de identidad por medio del rostro radica en la protección de la seguridad personal, la preservación de la privacidad, la confianza en los sistemas de autenticación y la prevención del fraude. Aspectos fundamentales para asegurar una sociedad digital segura y confiable. Por lo que la presente tesis propone un método para la detección de suplantación por medio del rostro en condiciones adversas sin hardware especializado.

1.1 Descripción del problema

La seguridad en los sistemas de verificación de identidad por medio del rostro es un tema importante, debido a que pueden ser vulnerables en un entorno de confrontación, en el que una persona puede camuflarse como un usuario legítimo para vulnerar la seguridad del sistema utilizando diferentes instrumentos como lo son, fotografías impresas y/o digitales, así como vídeos reproducidos por medios digitales, ya que son considerados instrumentos de oportunidad y de bajo costo. Aunque a través de los años en la literatura se han propuesto métodos de detección de vida del rostro [2], [13]–[16], para distinguir entre rostros reales y falsos, son sensibles a la iluminación y a ruido diverso [17].

El problema que se aborda es reconocer ataques de suplantación de identidad mediante el análisis de imágenes sintéticas impresas y/o digitales considerando variaciones en la iluminación, resolución, pose y oclusión parcial.

1.2 Complejidad del problema

La complejidad del problema se ve reflejada en los siguientes puntos:

- Realizar la detección de suplantación de identidad en imágenes en ambientes no controlados (Iluminación, resolución, pose, oclusión parcial).

- Se considera que los fondos pueden ser estáticos y dinámicos; es decir, con imágenes (impresas o por medios electrónicos) o vídeos y además con diferente cantidad de detalles o información capturada y representada en la imagen, video o dispositivo de visualización.

- Realizar la detección de una posible suplantación de identidad sin un entrenamiento previo del rostro a analizar. En otras palabras, el sistema deberá estar entrenado para detectar que la imagen que se analiza no corresponde a una persona viva (detección de suplantación) sin importar el rostro.

- Realizar la detección de suplantación en un tiempo cercano al real, considerando como real no más de 3 segundos.

- Realizar la detección de suplantación sin utilizar hardware especializado.

1.3 Objetivos

General

Proponer un método de visión artificial para la detección de posibles ataques de suplantación de identidad mediante imágenes faciales, en un ambiente no controlado.

Específicos

- Revisar en el estado del arte los procedimientos aplicados para los distintos tipos de ataques utilizados en la detección de suplantación de identidad.

- Evaluar el nivel mínimo de iluminación y de visibilidad del rostro necesaria para realizar la detección de suplantación de identidad.

- Seleccionar las técnicas con los mejores resultados, de acuerdo con el estado del arte, para describir el rostro.

- Implementar al menos un método para abordar la suplantación de identidad de acuerdo con la literatura.

- Formular una variante de una técnica o la combinación de al menos dos, para dar solución al problema propuesto.

- Evaluar el método propuesto utilizando repositorios públicos especializados.

- Comparar con otros métodos del estado del arte usando las métricas aplicadas a la suplantación del rostro.

1.4 Alcances y limitaciones

Alcances

El sistema de visión artificial realiza la detección de suplantación de identidad ante diferentes condiciones de iluminación, incluyendo fondos complejos (entornos no controlados) e incluso dinámicos con movimiento. Las imágenes son adquiridas a través de cámaras web con diferentes resoluciones, y la respuesta del algoritmo se obtiene en un tiempo máximo de 3 segundos. Para evaluar el sistema de suplantación de identidad, se emplean las métricas APCER, BPCER, HTER y F-mesure. Además, se aborda tanto la suplantación de ataques a través de fotografías impresas como digitales, y se evalúan con al menos dos técnicas.

Limitaciones

La detección de suplantación de identidad por medio del rostro presenta limitaciones que son importantes considerar, como son:

- Variedad en condiciones de captura: La detección puede verse afectada por diferentes condiciones de captura, como iluminación nula (imagen completamente oscura o blanca), ángulos de captura inusuales o con ausencia de un rostro humano. Dichas condiciones pueden dificultar la precisión de la detección.

- Cambios en la apariencia facial: El sistema de detección puede presentar dificultades al enfrentar cambios en la apariencia facial, como el uso de maquillaje, barba, lentes, sombreros u otros elementos que modifiquen la apariencia original de la persona por lo que no se garantiza la correcta validación de suplantación si existe una oclusión mayor al 50% del rostro.

- Suplantación sofisticada: No se contemplan los ataques por medio de máscaras 3D o fotografías impresas en 3D.

- Bases de datos de referencia limitada: La precisión del sistema de detección está sujeta a la calidad de los bancos de imágenes empleados en el transcurso de la investigación.

- Eficiencia computacional: No se hace uso de hardware especializado como cámaras térmicas, infrarrojas o tarjetas gráficas que reduzcan el tiempo de procesamiento.

1.5 Propuesta de solución

Para abordar el problema planteado en el trabajo de investigación se realizaron las siguientes actividades:

- Fundamentación: Se realizó el estado del arte buscando trabajos relacionados con la suplantación de identidad.

- Delimitación: Se seleccionaron dos tipos de ataques a los que el sistema debe ser robusto como son: imágenes impresas y vídeos.

- Especificación: El desarrollo se realizó con el lenguaje de programación C++ y aplicando técnicas de visión artificial.

- Desarrollo: Se implementaron dos enfoques para evaluar la detección de suplantación de identidad: Visión artificial y *Deep learning*. Dentro del enfoque de visión artificial se manejaron técnicas como la textura y canales de color. Las cuales son técnicas utilizadas ampliamente en la literatura con buenos resultados, con ello se realizó una comparativa con el enfoque de *Deep learning* que ha tomado mayor relevancia en los últimos años con las redes neuronales convolucionales, las cuales están siendo aplicadas al problema de detección de suplantación de identidad en imágenes faciales. Además, se realizó una propuesta de un método que hace uso de visión artificial aplicado al problema.

- Pruebas: Se realizaron pruebas con bancos de imágenes públicos y un banco de imágenes propio.

- Evaluación: Se verificó que el sistema cumpla con los objetivos esperados, respecto a las métricas reportadas en el estado del arte, comparando los resultados obtenidos con el estado del arte (de manera teórica).

- Documentación: Se realizaron los reportes semestrales, se publicaron artículos en congresos internacionales, un artículo JCR, conferencias, la propuesta de tesis, el reporte de resultados obtenidos y la tesis de la investigación realizada.

1.6 Organización de la tesis

En los próximos capítulos se presenta una descripción detallada de todos los procedimientos requeridos para abordar la problemática planteada en el presente trabajo de investigación. La estructura de la tesis se compone de la siguiente manera:

Capítulo 1: Contiene el análisis del problema de este trabajo.

Capítulo 2: Contiene el estado del arte, su discusión correspondiente y el estado de la práctica.

Capítulo 3: Muestra el marco teórico.

Capítulo 4: Este capítulo comprende el diseño e implementación de la solución para el problema de tesis.

Capítulo 5: Describe el objetivo y diseño de las pruebas que se realizaron para cada módulo del desarrollo de la tesis; además del análisis de los resultados obtenidos.

Capítulo 6: Contiene las conclusiones generales, la discusión sobre el tema de tesis, trabajos futuros y aportaciones.

Anexos: Contiene la experimentación adicional realizada en el transcurso del doctorado.

CAPÍTULO 2
Estado del arte

La investigación sobre la suplantación de identidad por medio del rostro ha sido un tema de estudio en diversas disciplinas, como la informática, la seguridad, la inteligencia artificial y la biometría entre otras. Aunque no se sabe una fecha precisa, el interés en este tema se ha intensificado en las últimas décadas con el avance de las tecnologías relacionadas con el reconocimiento facial y la preocupación por la seguridad y la protección de la identidad personal.

En el presente capítulo se describen los conceptos base del tema, así como los hallazgos reportados en la literatura para dar solución a la suplantación facial. El capítulo se encuentra dividido en tres secciones que son: suplantación facial, artículos relevantes en el transcurso de la investigación y una recopilación de sistemas que están actualmente en el mercado, los cuales contemplan suplantación de identidad (estado de la práctica).

2.1 Suplantación facial

Para comprender el tema medular de la investigación que es la suplantación por medio del rostro, se describen los siguientes conceptos.

A. Instrumento de ataque de presentación (PAI por sus siglas en inglés)

De acuerdo con ISO / IEC 30107-1, la característica u objeto biométrico utilizado en un ataque de presentación se denomina Instrumento de Ataque de Presentación (PAI, por sus siglas en inglés) [18]. El PAI puede clasificarse en dos tipos:

1) Artificial: Se refiere a un medio artificial para generar el PAI. Lo cual, a su vez, se puede clasificar como:
   a) Completo, en referencia a la generación de un PAI artificial completo; por ejemplo, un vídeo de una cara, una máscara facial 3D, una impresión facial 2D, etc.
   b) Parcial, implica un PAI artificial que puede mostrar características biométricas parciales; por ejemplo, un vídeo facial con gafas de sol o una cara parcialmente visible.
2) Características humanas: Implica el uso de humanos como un PAI y puede ser:
   a) Sin vida: por ejemplo, una parte de la cara de un cadáver.
   b) Alterado: incluyendo la mutación de caras y cirugía estética.
   c) No conforme: esto incluye el uso de expresiones faciales.
   d) Forzado: esto incluye el uso de la cara de un humano inconsciente.
   e) Conforme: esto incluye intentos de impostor de esfuerzo cero.

De estos diferentes tipos de PAI, el artificial es el más utilizado en la literatura para estudiar las vulnerabilidades de los sistemas de reconocimiento facial.

B. Ataque de presentación (PA por sus siglas en inglés) o *spoofing facial*

Comúnmente conocido como suplantación de identidad, se produce cuando una persona intenta hacerse pasar por otra persona, falsificando datos e intentando obtener acceso y ventajas ilegítimos [7].

C. Detección de ataques de presentación (PAD por sus siglas en inglés) o *anti-spoofing*

También conocidos como contramedidas o métodos *anti-spoofing*, pueden detectar y mitigar ataques dirigidos de suplantación de identidad [18].

D. Tipos de ataques de suplantación de identidad

Enfocándose en los ataques al sensor existen diferentes tipos de suplantación de identidad, los primeros de los cuales se tiene reporte en la literatura son de imágenes impresas; sin embargo, con el paso del tiempo y el avance de la tecnología se han implementado otros instrumentos de suplantación. En la Tabla 2.1 se muestra una recopilación de los diferentes tipos de ataques de suplantación de identidad encontrados en la literatura, la tabla contiene el nombre y una breve descripción de los ataques.

Tabla 2.1 Recopilación de tipos de ataques de suplantación de identidad reportados en la literatura.

Tipo de ataque	Descripción
Fotografía impresa plana	El uso de una fotografía impresa plana es la más común, con un gran potencial, ya que la mayoría de las personas tienen imágenes faciales disponibles en Internet (por ejemplo, redes sociales) o podrían ser fotografiadas sin colaboración o permiso.
Fotografía plana por medios digitales	Al igual que la fotografía impresa plana, las imágenes faciales, pueden ser adquiridas por internet. A diferencia de las fotografías impresas, éstas tienen un menor costo y la resolución que proporcionan los medios digitales al momento de mostrar la imagen puede ser mayor.
Fotografía impresa plana con corte ocular	En el ataque fotográfico de corte ocular. Las regiones oculares de una fotografía impresa se cortan para exhibir un comportamiento de parpadeo.
Foto deformada	Los ataques de fotografías deformadas consisten en doblar una fotografía impresa en cualquier dirección para simular el movimiento facial.
Vídeo por medios digitales	Un ataque a través de la reproducción de video muestra casi todos los comportamientos similares a las caras reales, con muchas de las características intrínsecas de los movimientos válidos del usuario. El ataque por video tiene signos fisiológicos de vida que no se presentan en fotos, como parpadeo, expresiones faciales y movimientos en la cabeza y la boca; y se puede realizar con tabletas o teléfonos inteligentes grandes.
Máscara	Los ataques de máscara son de dos tipos: con una máscara portátil de tamaño natural y con una máscara cortada en papel. Estos ataques están dirigidos a sistemas *anti-spoofing* que analizan estructuras faciales en 3D, siendo uno de los ataques más complejos que se detectan. La fabricación de máscaras es mucho más difícil y costosa que los otros tipos de ataques, ya que requiere escaneo 3D e impresión de dispositivos especiales.

En la Figura 2.1 se muestran ejemplos de los diferentes tipos de ataques, cabe resaltar que, en el caso de las fotografías impresas, existen variaciones en la calidad dependiendo de la calidad de impresión y el tipo de papel en que se imprime. De igual manera las fotografías y vídeo por medios digitales resultan un desafío debido a la calidad de las imágenes del medio por el que se visualizan, así como, si la pantalla tiene reflejos o no.

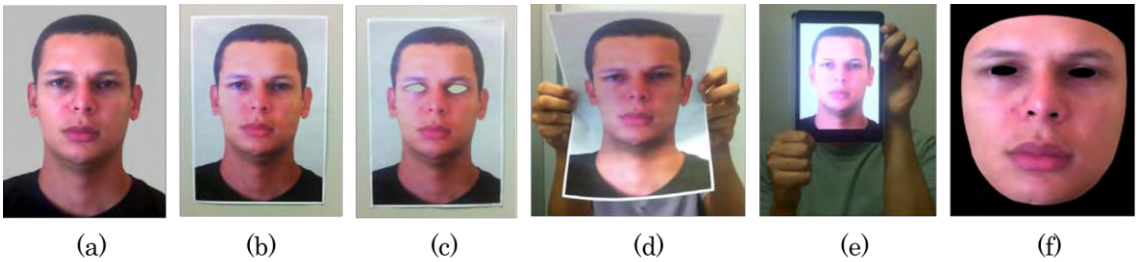

Figura 2.1 Tipos de instrumentos de suplantación de identidad: (a) Imagen real, (b) fotografía impresa plana, (c) fotografía impresa plana con recorte ocular, (d) fotografía deformada, (e) fotografía por medios digitales, (f) máscara [19].

E. Tipos de detección de ataques de suplantación

Los sistemas de reconocimiento facial son vulnerables a varios tipos de instrumentos (o PAI) que pueden generarse de manera rentable. Lo que exige la necesidad de detectar y mitigar los ataques para mejorar tanto la seguridad como la confiabilidad de los sistemas de reconocimiento facial por medio de un sistema PAD o detección de vida; sin embargo, estrictamente hablando, la detección de vida se define como la medición y el análisis de características anatómicas o reacciones involuntarias o voluntarias, para determinar si se está capturando una muestra biométrica de un sujeto vivo presente en el punto de captura [18]. Por consiguiente con base en la definición estandarizada del término, la detección de vida se puede considerar como un subconjunto de PAD y no como un sinónimo de PAD en sí [20].

Los tipos de detección de suplantación se pueden clasificar en dos tipos, los basados en hardware [7], [21]–[23] y los basados en software [21], [22], en la Figura 2.2 se muestra un desglose de ellos. En los basados en hardware se contempla que los sensores ocupados para la adquisición de la imagen tengan alguna característica que pueda ayudar a la detección de suplantación, mientras que los basados en software consideran que las imágenes sean adquiridas por cualquier medio digital. En el presente proyecto solo se usan los basados en software, por su bajo costo.

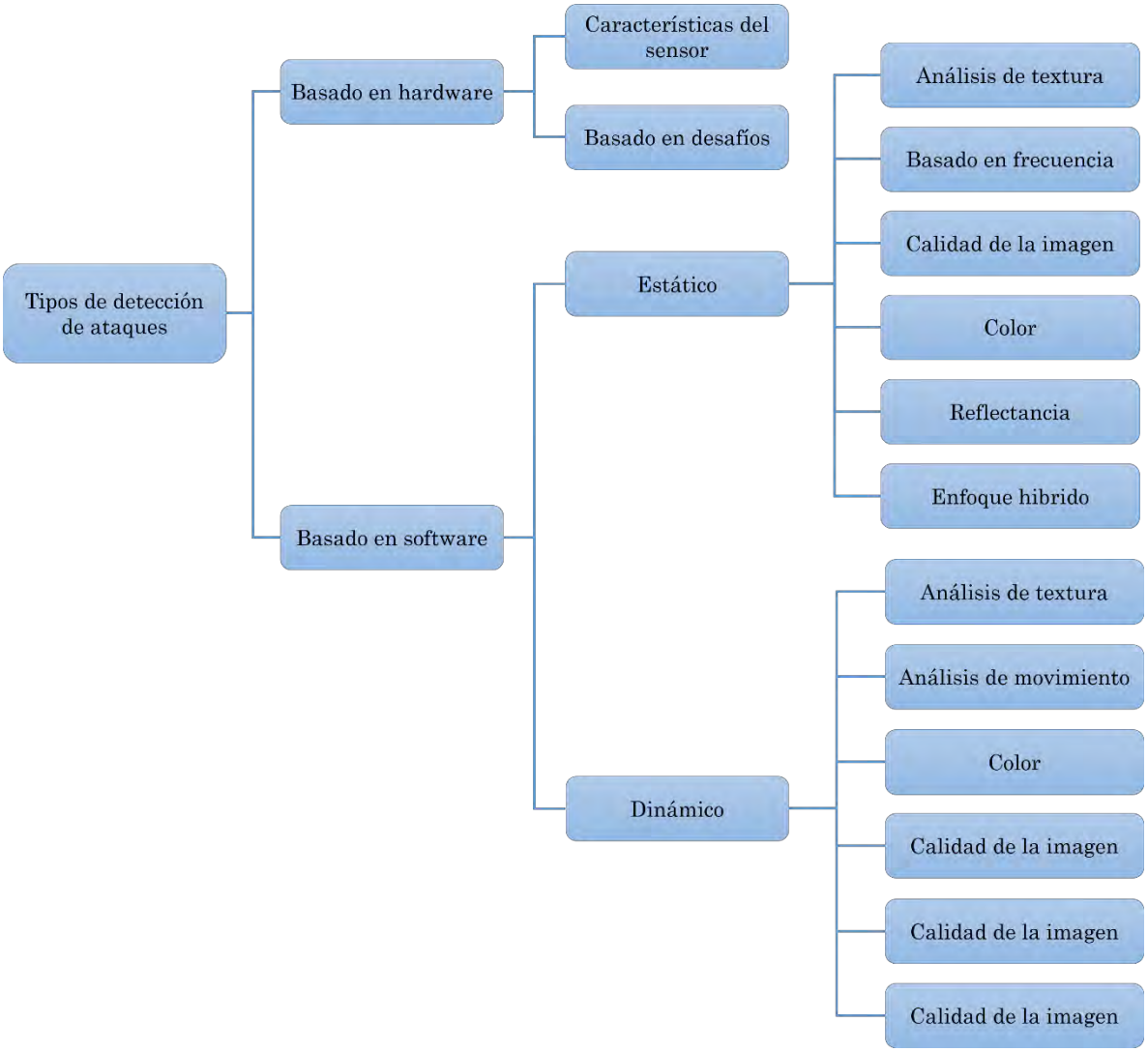

Figura 2.2 Tipos de detección de ataques de suplantación (Imagen basada en [24] y complementada con artículos de la literatura).

Hoy en día, existen diversas aplicaciones tanto en computadora como en móviles que hacen uso de la biometría por medio del rostro para dar acceso a programas o al dispositivo. Sin embargo, todavía falta seguridad en la implementación, debido a que no todos son robustos a los ataques de suplantación y los que cuentan con la funcionalidad anti-spoofing requieren la captura de diversas imágenes del rostro en un entorno iluminado, sin oclusión y en algunos casos sin lentes.

Dentro de los sistemas de reconocimiento facial se pueden encontrar diferentes vulnerabilidades [20]. La Figura 2.3 muestra un diagrama de bloques de un sistema de reconocimiento facial genérico con nueve vulnerabilidades, como se indica en ISO / IEC 30107-1: 2016 [25].

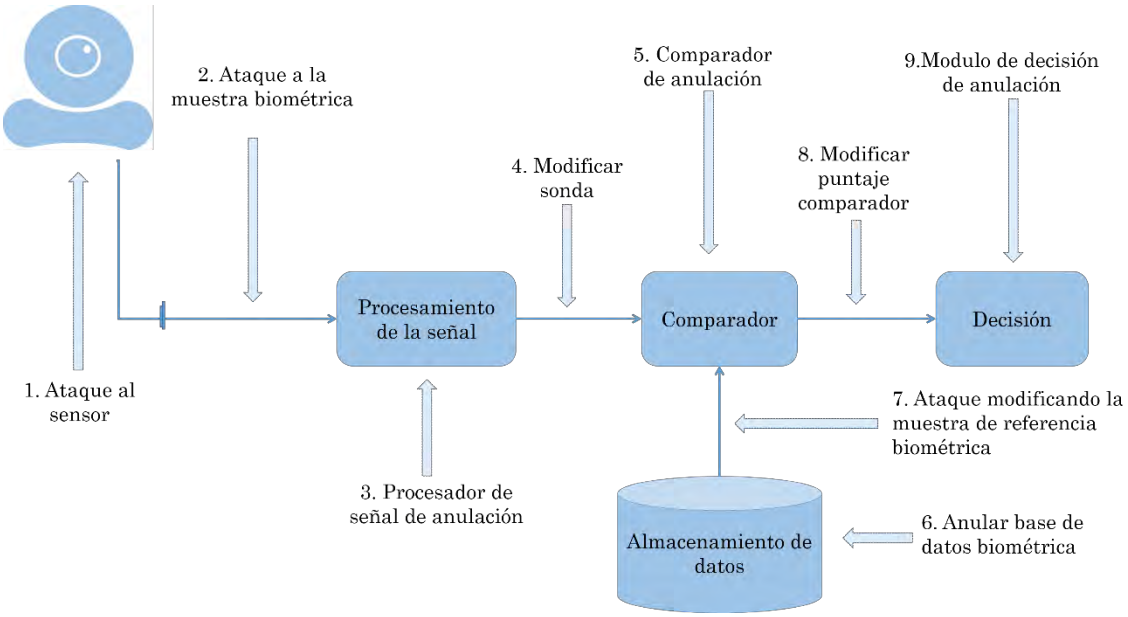

Figura 2.3 Vulnerabilidad de un sistema de reconocimiento facial (Figura inspirada en [18]).

Las vulnerabilidades observadas son:

1. En el sensor (es decir, el subsistema de captura de datos) implica presentar un instrumento biométrico facial del usuario legítimo como una entrada al sensor.
2. Interceptación de la muestra biométrica que fue capturada por el sensor. Éste ataque básicamente implica reemplazar la muestra biométrica de la cara capturada con una muestra falsa.
3. Anular el módulo de procesamiento de señal. Esto podría implicar la modificación de la funcionalidad del extractor de características, por ejemplo, utilizando un caballo de Troya.
4. Permite al atacante reemplazar las características extraídas de la muestra de la sonda con características de destino.
5. Anular el proceso de validación para que produzca una puntuación de comparación requerida por el atacante.
6. Reemplazar la plantilla de referencia de modo que la ID autorizada esté asociada con la plantilla del atacante.
7. Modificación de la plantilla de referencia en el canal de comunicación.
8. Intercepción y corrupción de la salida del clasificador.
9. Anular el módulo de decisión para generar la respuesta deseada.

De las nueve vulnerabilidades, sólo la primera implica un ataque al sensor mismo. Todas las demás están relacionadas con la integridad del sistema en general, en la presente investigación solo se contempló el ataque al sensor.

2.2 Trabajos relacionados

Existen diversos usos y aplicaciones en las que, la detección de suplantación facial es importante, por ejemplo, en aplicaciones forenses y de seguridad pública, autenticación para acceso a sistemas bancarios hasta sistemas de educación a distancia como exámenes en línea. En los escenarios mencionados la seguridad es una parte esencial para los usuarios, dado que la educación en línea permite a los estudiantes realizar actividades de aprendizaje desde cualquier ubicación, existe el riesgo de que un estudiante pueda hacer trampa o hacer que otra persona tome el examen en su lugar.

Por lo que se han desarrollado métodos de análisis de comportamiento para detectar si el estudiante está presente y es quien dice ser durante el examen en línea como son la verificación del usuario, detección de texto, detección de voz, detección de ventana activa, estimación de la mirada y detección del teléfono para identificar si hay una suplantación o actividad sospechosa [26].

Otras medidas involucran hardware especializado, como la autentificación y monitoreo por medio de huella digital [27][28], si bien es un método muy seguro para la verificación del usuario, tiene la limitante de necesitar hardware especializado, en donde aplicarlo a esquemas de evaluaciones en línea a distancia involucra un costo de hardware específico que no todos los estudiantes o instituciones podrían costear, además, se debe considerar dos escenarios

a) El proceso de captura y verificación de las huellas puede ser complejo y requiere capacitación adecuada para su correcta utilización.

b) Aunque en general las huellas dactilares son estables, hay situaciones en las que pueden dañarse o ser difíciles de leer. Por ejemplo, ciertos trabajos manuales o exposición prolongada al agua pueden afectar la calidad de las huellas dactilares, lo que dificultaría su detección precisa.

La detección de suplantación facial se puede utilizar en la educación en línea como una medida de seguridad para verificar la identidad de los estudiantes durante exámenes o evaluaciones en línea en combinación con analizar la atención del usuario durante el examen [29] o tecnologías de análisis de comportamiento para detectar si el estudiante está presente y es quien dice ser durante el examen en línea. Las cuales pueden monitorear los movimientos faciales, los patrones de parpadeo[30], los cambios en la expresión facial y otros indicadores para identificar si hay una suplantación o actividad sospechosa [31].

Al utilizar la detección de suplantación facial, las instituciones educativas en línea pueden aumentar la integridad y la confiabilidad de los procesos de evaluación y garantizar que los estudiantes sean evaluados de manera justa y equitativa [32]. Sin embargo, es importante tener en cuenta las implicaciones de privacidad y asegurarse de obtener el consentimiento adecuado de los estudiantes antes de implementar estas medidas de monitoreo [33], para evitar posibles robos de identidad [34] [35].

Por lo tanto, la implementación de contramedidas que involucren datos biométricos puede ser una estrategia eficaz para mejorar la confianza y la seguridad de los sistemas en línea [36]. Los datos biométricos, como las huellas dactilares, el reconocimiento facial o el escaneo de iris, pueden proporcionar una capa adicional de autenticación y verificación de identidad, ya que son características únicas y difíciles de replicar [37].

Algunas razones por las cuales las contramedidas biométricas pueden ser beneficiosas son:

- Mayor seguridad: Los datos biométricos están asociados a rasgos físicos únicos de cada individuo, lo que hace que sea difícil de falsificar o suplantar. Lo que proporciona un nivel más alto de seguridad en comparación con otros métodos de autenticación, como contraseñas o tarjetas de acceso, que pueden ser robados o compartidos.

- Mejor experiencia del usuario: Las contramedidas biométricas pueden proporcionar una experiencia de usuario más conveniente y fluida. En lugar de recordar contraseñas complicadas o llevar consigo tarjetas de acceso, los usuarios pueden utilizar sus características biométricas inherentes para acceder a los sistemas en línea de manera rápida y sencilla.

- Reducción del riesgo de robo de identidad: Los datos biométricos no se pueden cambiar o alterar fácilmente, lo que ayuda a reducir el riesgo de robo de identidad. En comparación con información personal como nombres, direcciones o números de seguro social, los datos biométricos son menos susceptibles de ser comprometidos y utilizados de manera fraudulenta.

Sin embargo, también es importante considerar las siguientes consideraciones:

- Privacidad y seguridad de los datos: Al utilizar datos biométricos, es fundamental garantizar la protección de la privacidad y la seguridad de dichos datos. Las organizaciones deben implementar medidas sólidas para almacenar, transmitir y proteger los datos biométricos, así como obtener el consentimiento informado de los usuarios para su uso.

- Posibles limitaciones técnicas: La implementación de contramedidas biométricas puede requerir la adquisición de hardware y software especializado, así como la integración con sistemas existentes. Además, algunos métodos biométricos pueden tener limitaciones técnicas en términos de precisión, tiempo de procesamiento y capacidad de adaptación a diferentes condiciones.

Por lo tanto, las contramedidas biométricas pueden mejorar la confianza y la seguridad de los sistemas en línea al proporcionar una autenticación más sólida y difícil de falsificar. Sin embargo, es importante equilibrar los beneficios con las preocupaciones de privacidad y seguridad, y garantizar que se implementen medidas adecuadas para proteger los datos biométricos de los usuarios. Por lo que, la inteligencia artificial es útil en la detección de suplantación de identidad por medio del rostro debido a su capacidad para analizar y reconocer patrones complejos en imágenes y videos. Al aplicar técnicas de aprendizaje profundo y visión por computadora, los algoritmos de inteligencia artificial pueden extraer características distintivas de los rostros y detectar señales de manipulación o generación falsa.

2.2.1. Visión artificial

La visión artificial se utiliza en la detección de suplantación de identidad por medio del rostro debido a su capacidad para analizar y comprender las imágenes y videos de manera similar a como lo hacen los seres humanos. A lo largo del tiempo, los investigadores han explorado y desarrollado diferentes técnicas y enfoques basados en visión artificial para mejorar la detección de suplantación facial.

Inicialmente, se utilizaron enfoques basados en características faciales lo que se remonta a los primeros trabajos en reconocimiento facial y biometría por la década de 1960. En aquel entonces, los métodos se centraban principalmente en el análisis de características geométricas de la cara, como la posición de los ojos, la nariz y la boca, para identificar individuos [38]. Al emplear estos enfoques en la detección de suplantación los métodos eran limitados y susceptibles a ataques más sofisticados como máscaras y videos.

Luego, se introdujeron métodos basados en el análisis de patrones de textura facial, la cual ha sido objeto de investigación durante varios años, ya que proporciona información detallada sobre la piel y las características faciales de una persona al extraer características y encontrar anomalías, lo que puede contribuir a identificar señales de manipulación o generación artificial [39].

LBP (*Local Binary Patterns* o Patrones Binarios Locales) [40] fue uno de los primeros enfoques utilizados para extraer características de textura y detectar suplantaciones faciales desde principios de los años 2000, debido a que se basa en las variaciones locales de la intensidad de los píxeles en una imagen [41]. Por lo que es menos sensible a las variaciones globales en la iluminación y el contraste, además, es invariante ante las transformaciones espaciales, lo que significa que la detección de suplantación facial puede funcionar bien incluso cuando la posición o el tamaño de la cara varían en la imagen, por lo tanto lo hace más robusto frente a ataques de suplantación facial como el uso de máscaras, maquillaje o cambios en la iluminación [42], aunque puede presentar dificultades para capturar detalles finos de la imagen del rostro.

Más tarde, surgieron variantes como BSIF (*Binarized statistical image features o* Funciones de imágenes estadísticas binarizadas) la cual es relativamente nueva en comparación con LBP ya que fue propuesta por primera vez en el 2012 por Kannala y Rahtu como una forma de extraer características de textura binarias y patrones sutiles en imágenes [43], su principal ventaja es la captura de detalles de textura más finos en comparación con LBP [44]. Lo que significa que puede identificar patrones sutiles en una imagen facial, además es robusto ante cambios locales de textura, permite utilizar diferentes filtros en la etapa de extracción de características lo que brinda flexibilidad para adaptar el enfoque a diferentes tipos de suplantación o condiciones de captura de la imagen [45].

Sin embargo, con el avance tecnológico la resolución de las cámaras, la calidad de las imágenes y sofisticación de los ataques, dio paso a emplear técnicas como PRNU (*Photo response non-uniformity o* Falta de uniformidad en la fotorespuesta) que en sus inicios (2006), se utilizaba principalmente en el campo de la autenticación de imágenes forenses y la identificación de cámaras individuales [46]. El PRNU se refiere a las imperfecciones inherentes en los sensores de las cámaras digitales, que crean patrones de ruido únicos en las imágenes capturadas. Estos patrones se deben a pequeñas variaciones en la respuesta de cada píxel del sensor y son prácticamente imposibles de eliminar o alterar. Por lo tanto, el PRNU se considera una "huella digital" única de la cámara y puede utilizarse para verificar la autenticidad de las imágenes capturadas por esa cámara en particular. Su naturaleza única y su capacidad para abordar los desafíos específicos asociados con la manipulación de imágenes faciales es lo que ha hecho que los investigadores la utilicen para aprovechar las características intrínsecas de las cámaras y su robustez ante manipulaciones. Y de esta manera poder ayudar a verificar la autenticidad de las imágenes faciales [47].

Por lo tanto, cada uno de los enfoques de textura tiene sus fortalezas y limitaciones en la detección de suplantación facial. LBP es rápido y robusto, pero puede tener dificultades con suplantaciones sofisticadas.

BSIF captura características de textura más finas, pero puede ser más costoso computacionalmente. PRNU se basa en características únicas de la cámara, pero puede requerir una etapa de calibración y ser menos efectivo en ciertos escenarios. Lo que ha dado pie a investigar otros enfoques o técnicas que contribuyan a una detección más precisa.

En años recientes se han propuesto métodos que exploran las características de frecuencia de la imagen facial para la detección de suplantación. El uso de características de frecuencia se basa en la premisa de que las imágenes faciales reales y las imágenes manipuladas tienen diferencias estadísticas en el dominio de frecuencia. Estas diferencias se pueden explorar mediante técnicas de transformación de Fourier u otras técnicas de análisis espectral [48].

Es importante mencionar que, en el transcurso de la evolución de las técnicas del análisis de patrones, textura y la frecuencia. Fueron surgiendo características complementarias a las técnicas como el uso del color [49]. Las cuales son usadas para aprovechas la información visual proporcionada por las características cromáticas de la piel y otros elementos faciales [50]. Al analizar el color y buscar discrepancias o alteraciones inusuales, se puede mejorar la capacidad de detectar suplantaciones y manipulaciones faciales en imágenes.

En resumen, la evolución de las técnicas de visión artificial en la detección de suplantación facial ha pasado por diferentes enfoques. Esta evolución continúa en curso con el objetivo de mejorar la precisión y la capacidad de detección en escenarios de suplantación facial cada vez más desafiantes. En la Tabla 2.2 se muestra un concentrado de artículos relevantes para la investigación con enfoque de visión artificial que se recabaron en el transcurso de la investigación.

Tabla 2.2 Resumen de artículos del estado del arte con visión artificial.

Año y referencia	Ataque	Método de detección	Banco de imágenes	Métrica	Resultado (%)
2019 [49]	Imágenes	Bioinspirada en el funcionamiento del ojo, por medio de la fusión de los canales RGB y textura	CASIA, MFSD Y PROPIA(FRAV)	FAR	CASIA=2.857; FRAV=2.98; MFSD=9.64
				FRR	CASIA=13.9; FRAV=17.34; MFSD=39.07
				ACER	CASIA=8.37; FRAV=10.16; MFSD=24.34
2019 [48]	Video	Análisis de las diferencias de frecuencia espacial (Fourier) entre los videos reales y los falsos	*Replay Attack Database, CASIA-Face, 3D Mask Attack Dataset y Unicamp Video Attack DB*	ERR	*Replay-attack*=0; CASIA=0; UVAD=26
				HTER	*Replay-Attack*=0; 3DMAD=2.44
2019 [42]	Imágenes planas y capturas de videos	Patrón binario local completo uniforme (UCLBP) con ocho vecindades	NUAA	ACC	98.89
2019 [41]	Imágenes	Variación del patrón binario local (LBPV)	NUAA	ACC	87.22
2019 [50]	Imágenes extraídas de video (*frames*)	*Local ternary pattern* (LTP)	NUAA, CASIA FASD y *Replay-attack*	HTER	NUAA=0.0216; CASIA-FASD= .0722;
				AUC	NUAA=0.9849; *Replay- attack*=0.9952
2020 [51]	Videos	Calidad de la imagen	*Replay attack* →CASIA	HTER	30.2
2020 [52]	Video	Representación de relatividad en la variedad de Riemann con características de *Haralick*	*Replay attack* → CASIA	HTER	27.59
2021 [53]	Videos	Tres filtros de paso alto y tres filtros de paso bajo para crear los mapas de HF y los mapas de LF. Posteriormente, se implementa una red pseudo-siamesa para extraer las características de suplantación de HF y LF.	MSU → CASIA	HTER	35.7
			CASIA → MSU		18.8
			Replay attack → CASIA		27.2
			Replay attack → MSU		24.3

Tabla 2.3 Continuación resumen de artículos del estado del arte con visión artificial.

Año y referencia	Ataque	Método de detección	Banco de imágenes	Métrica	Resultado (%)
2022 [54]	Imágenes	Características de evaluación de calidad de imagen perceptiva multi-escala	*Replay attack* → CASIA	EER	12.7
				APCER	75.83
				BPCER	2.22
				HTER	39.03
			UVAD → CASIA	APCER	97.50
				BPCER	8.89
				HTER	53.19
2022 [55]	Imágenes	Parches faciales y progresión lineal	OULU-NPU & MSU & Replay attack → CASIA	HTER	9.81
2023 [56]	Imágenes	Operador de gradiente aprendible (*learnable gradient operator* LGO) basado en sobel	Replay&SiW&OULU→ CASIA	HTER	22.78
			Replay attack → CASIA	HTER	31.9

2.2.2. Aprendizaje profundo

Las redes neuronales convolucionales (CNN) comenzaron a utilizarse en la detección de suplantación facial en torno al año 2015 [57]. En ese momento, el aprendizaje profundo y las CNN ya habían demostrado su eficacia en diversas tareas de visión por computadora, como el reconocimiento de objetos y el reconocimiento facial [58].

Desde entonces, se han realizado numerosas investigaciones y desarrollos en el campo de la detección de suplantación facial utilizando CNN y otras arquitecturas de aprendizaje profundo [59] [60] [61] [62] [63]. El objetivo de las investigaciones es mejorar la precisión y robustez de los sistemas de detección al extraer características relevantes de las imágenes faciales y han explorado nuevas arquitecturas y técnicas para abordar desafíos específicos en la detección de manipulaciones faciales [64].

Es importante tener en cuenta que la evolución del aprendizaje profundo en la detección de suplantación facial es un campo de investigación activo y en constante desarrollo. Los investigadores continúan explorando nuevas arquitecturas, técnicas y enfoques para mejorar la precisión y la resistencia de los sistemas de detección frente a manipulaciones cada vez más sofisticadas [65].

Sin embargo, si bien el aprendizaje profundo ha demostrado ser eficaz en muchos problemas en los cuales se aplicaban técnicas de visión por computadora [66], también tiene algunas desventajas en comparación con los métodos tradicionales en la detección de suplantación facial. Algunas de estas desventajas incluyen:

- Requiere grandes cantidades de datos de entrenamiento: Los modelos de aprendizaje profundo requieren grandes conjuntos de datos de entrenamiento para aprender patrones y características discriminativas [67]. Obtener y etiquetar un gran conjunto de datos de imágenes faciales auténticas y manipuladas puede ser costoso y requerir mucho tiempo. Esto puede ser una desventaja en escenarios donde la disponibilidad de datos es limitada [68].

- Necesita recursos computacionales potentes: El entrenamiento y la ejecución de modelos de aprendizaje profundo son computacionalmente intensivos y requieren recursos significativos, como unidades de procesamiento gráfico (GPU) y memoria. Esto puede limitar la aplicabilidad del aprendizaje profundo en entornos con recursos limitados o en dispositivos con capacidades computacionales más modestas [69], [70].

- Requiere tiempo de entrenamiento prolongado: El entrenamiento de modelos de aprendizaje profundo puede llevar mucho tiempo, especialmente cuando se utilizan conjuntos de datos grandes y complejos. Esto puede ser una limitación en aplicaciones donde se requiere una respuesta en tiempo real o una rápida adaptación a cambios en los datos de entrada [71].

- Necesita optimizar una gran cantidad de parámetros: Los modelos de aprendizaje profundo suelen tener una gran cantidad de parámetros que deben ser ajustados durante el proceso de entrenamiento. Esto puede aumentar el riesgo de sobreajuste (*overfitting*) a los datos de entrenamiento, lo que significa que el modelo puede tener dificultades para generalizar a nuevos datos y puede mostrar un rendimiento deficiente en escenarios del mundo real [72].

- Dificultad para interpretar los resultados: Los modelos de aprendizaje profundo, especialmente los modelos más complejos como las redes neuronales profundas, pueden ser difíciles de interpretar. La naturaleza "caja negra" de estos modelos puede dificultar la comprensión de las razones detrás de las decisiones del modelo, lo que puede ser problemático en aplicaciones donde se requiere una explicación clara y transparente de los resultados [73].

Es importante tener en cuenta que las desventajas encontradas no invalidan el uso del aprendizaje profundo en la detección de suplantación facial, pero destacan algunos de los desafíos y consideraciones que deben tenerse en cuenta al utilizar esta técnica en comparación con los métodos tradicionales. En la Tabla 2.4 se muestra un concentrado de los artículos con enfoque de CNN que son relevantes en la investigación.

Tabla 2.4 Resumen de artículos del estado del arte con aprendizaje profundo.

Año y referencia	Ataque	Método de detección	Banco de imágenes	Métrica	Resultado (%)
2019 [64]	Imágenes	Histogramas aplicados a imágenes RGB, NIR y regresión logística (LR)	*Wide Multi-Channel Presentation Attack (WMCA)*	APCER	NIR + LBP+LR = 7.1
2019 [59]	2D	Red neuronal convolucional de dos flujos (TSCNN) que funciona en dos espacios complementarios: espacio RGB	CASIA-FASD, *REPLAY-ATTACK* y OULU	HTER	*Train* CASIA y *test replay attack*=30; *train replay attack y test* CASIA=33.4
2019 [61]	2D y Video	*MMD to the multilayer full connected layers (ML-MMD)*	CA- SIA *Face Anti-Spoofing Database* (CBSR), *Replay- Attack Database*	HTER	Replay=0.6
				ERR	Replay=0.3; CBSR=3.7
2019 [62]	2D	Fotopletismografía remota (rPPG) y una red neuronal convolucional basada en parches contextuales (CP-CNN)	3DMAD, HKBU-Mars V1, MSU-MFSD, CASIA-FASD y OULU-NPU	ERR	3DMAD=0; HKBU-Mars=0; MSU-MFSD=0; CASIA-FASD=1.8;
				HTER	3DMAD=0; HKBU-Mars=0;
				TPR	MSU-MFSD=100; CASIA-FASD=98.7;
				APCER	OULU-NPU=4.7
2019 [63]	2D	Combinación de análisis de textura y una red neuronal convolucional (CNN)	NUAA	ACC	100
2019 [60]	2D	Red neuronal convolucional poco profunda con incrustación laplaciana (shallowCNN-LE)	CASIA FASD, *Replay attack* y MSU USSA	HTER	CASIA=9.6; *Replay Attack*=3.7
				TPR	CASIA=90.5; MSU=92.16
				EER	CASIA=4.0; MSU=8.41
2021 [74]	Imágenes y videos	Extracción de características discriminatorias basadas en una recomposición de los componentes aprendidos de baja / alta frecuencia por medio de CNN	MSU → CASIA	HTER	27.3
2021 [75]	Imágenes y videos	Red de destrucción y combinación (DCN)	CASIA → MSU	HTER	17.5
			CASIA & *Replay attack* → MSU		14.8
			MSU & *Replay attack* → CASIA		32.3
			Replay attack → CASIA		29.4

Tabla 2.5 Continuación resumen de artículos del estado del arte con aprendizaje profundo.

Año y referencia	**Ataque**	**Método de detección**	**Banco de imágenes**	**Métrica**	**Resultado**
2021 [76]	Imágenes y videos	ResNet50	*Replay attack* → CASIA	HTER	31.3
2021 [77]	Imágenes y videos	Aprendizaje adaptativo de representación normalizada (ANRL)	MSU & *Replay attack* → CASIA	HTER	31.06
2021 [78]	Imágenes y videos	Red de Transferencia de Dominio (DTN)	OULU-NPU & CASIA & *Replay attack* → MSU	HTER	19.4
			OULU-NPU & MSU & *Replay attack* → CASIA		22.03
2022 [79]	Imágenes y videos	Red de destrucción y combinación (*Destruction and Combination Network* DCN)	*Replay attack* → CASIA	HTER	29.4

2.2.3. Comentarios

Después de revisar cómo han evolucionado los enfoques para dar solución a la detección de suplantación facial, se puede determinar que los métodos tradicionales pueden ser preferibles en algunos casos en comparación con el aprendizaje profundo basado en los siguientes puntos:

- Disponibilidad limitada de datos de entrenamiento: El aprendizaje profundo requiere grandes cantidades de datos de entrenamiento para lograr un rendimiento óptimo. En entornos no controlados, puede ser difícil obtener conjuntos de datos suficientemente grandes y variados para entrenar modelos de aprendizaje profundo de manera efectiva. En cambio, los métodos tradicionales pueden aprovechar características y algoritmos diseñados específicamente para abordar el problema de la suplantación facial con conjuntos de datos más limitados.

- Eficiencia computacional: El aprendizaje profundo es computacionalmente intensivo y requiere recursos computacionales significativos, como GPUs y grandes cantidades de memoria. En entornos no controlados, donde la detección de suplantación facial puede ser necesaria en dispositivos con recursos limitados o en tiempo real, los métodos tradicionales pueden ser más adecuados debido a su menor demanda computacional.

- Interpretabilidad y transparencia: Los modelos de aprendizaje profundo suelen ser "cajas negras" en el sentido de que pueden ser difíciles de interpretar y comprender cómo toman sus decisiones. Esto puede ser problemático en aplicaciones donde se requiere una explicación clara y transparente de los resultados, especialmente en entornos no controlados donde la transparencia y confiabilidad y la son importantes. Los métodos tradicionales, en cambio, a menudo se basan en reglas y características más interpretables, lo que facilita la comprensión de los resultados y la interpretación de los casos difíciles.

- Robustez frente a la variabilidad de los entornos no controlados: Los entornos no controlados pueden presentar desafíos adicionales como iluminación variable, cambios en la pose facial, oclusiones parciales y otros factores que pueden dificultar la detección precisa de suplantación facial. En algunos casos, los métodos tradicionales pueden ofrecer una mejor capacidad de adaptación y robustez ante estas variabilidades, ya que pueden ser diseñados para abordar específicamente estos desafíos sin depender exclusivamente de grandes conjuntos de datos de entrenamiento.

Por lo tanto, los métodos tradicionales pueden ser preferibles en la detección de suplantación facial en entornos no controlados debido a la disponibilidad limitada de datos, la eficiencia computacional, la interpretabilidad y la robustez requerida en dichos entornos.

2.3 Estado de la práctica

Existen varios sistemas que utilizan la detección de suplantación facial como parte de sus soluciones, muchos teléfonos móviles y tabletas en la actualidad incorporan sistemas de reconocimiento facial para desbloquear el dispositivo o acceder a ciertas aplicaciones. Los cuales utilizan algoritmos de detección de suplantación facial para garantizar que el rostro detectado sea auténtico y no una imagen o máscara. Algunos ejemplos de teléfonos móviles que han incorporado funcionalidades anti-spoofing son:

- iPhone X y modelos posteriores: Los modelos de iPhone X y posteriores de Apple han implementado una función llamada "*Face ID*" que utiliza sensores infrarrojos y una cámara *TrueDepth* para crear un mapa tridimensional del rostro del usuario, para prevenir suplantaciones utilizando imágenes estáticas o máscaras [80].

- Samsung Galaxy S10 y modelos posteriores: La serie Galaxy S10 y los modelos posteriores de Samsung han incorporado un sensor de huella dactilar ultrasónico en la pantalla, junto con su sistema de reconocimiento facial pretenden aumenta la seguridad y dificulta la suplantación facial[81].

- Google Pixel 4 y modelos posteriores: Los modelos de Google Pixel 4 y posteriores utilizan un sistema de reconocimiento facial llamado "*Face Unlock*" [82]. Utiliza una combinación de hardware y software para analizar y autenticar el rostro del usuario, lo que ayuda a prevenir suplantaciones.

- Huawei Mate 20 Pro y modelos posteriores: La serie Mate 20 Pro y los modelos más recientes de Huawei han implementado un sistema de reconocimiento facial en 3D que utiliza un sensor infrarrojo y una cámara frontal para autenticar el rostro del usuario. Lo que aumenta la seguridad y reduce la posibilidad de suplantación facial [83].

Es importante destacar que son solo algunos ejemplos de teléfonos móviles que han incorporado funcionalidades *anti-spoofing*. La lista de dispositivos con características de detección de suplantación facial está en constante evolución, y muchas otras marcas y modelos también pueden ofrecer medidas de seguridad adicionales para evitar suplantaciones faciales.

Además de los teléfonos móviles la detección de suplantación facial se ha investigado para agencias como aeropuertos y fronteras para mejorar la seguridad y la identificación precisa de los viajeros. Estos sistemas utilizan cámaras de alta resolución y algoritmos avanzados para detectar posibles suplantaciones y comparar las características faciales de los individuos con bases de datos de identidad. Muestra de ellos son los aeropuertos de *Schiphol* (Ámsterdam, Países Bajos) y Barajas (Madrid)[84], los cuales ha estado probando y utilizando tecnologías de reconocimiento facial para mejorar los controles de seguridad y la experiencia de los pasajeros.

Otros sistemas que han incursionado en soluciones *anti-spoofing* son los de video vigilancia y seguridad en entornos públicos como estaciones de tren, centros comerciales y edificios gubernamentales, las soluciones utilizan la detección de suplantación facial como una capa adicional de seguridad, alertando a los operadores de cuando se detecta una suplantación facial o un comportamiento sospechoso. A continuación, se mencionan algunos sistemas de video vigilancia que utilizan algoritmos anti-spoofing para mejorar la detección de suplantaciones faciales:

- *Dahua Technology* ofrece sistemas de video vigilancia avanzados que utilizan algoritmos *anti-spoofing* para detectar intentos de suplantación facial. Sus sistemas incorporan tecnología de inteligencia artificial y aprendizaje profundo para identificar características faciales genuinas y distinguirlas de imágenes falsas [85].

- *Hikvision* es otro fabricante líder de sistemas de video vigilancia que ha implementado algoritmos anti-spoofing en sus soluciones. Sus sistemas utilizan tecnología de reconocimiento facial y algoritmos avanzados para detectar intentos de suplantación mediante el análisis de características faciales dinámicas [86].

- *Axis Communications* ofrece cámaras de video vigilancia de alta calidad que incluyen funciones de detección de suplantación facial. Sus cámaras utilizan algoritmos de aprendizaje automático para analizar las características faciales y detectar intentos de suplantación mediante el uso de imágenes o máscaras [87].

- *NEC Corporation* ha desarrollado sistemas de video vigilancia que incorporan algoritmos anti-spoofing. Sus soluciones utilizan tecnología de reconocimiento facial avanzada, incluido el análisis de textura y la detección de movimientos, para identificar intentos de suplantación y mejorar la precisión en la autenticación facial [88].

Los sistemas de control de acceso en entornos corporativos y residenciales también han hecho uso de algoritmos y técnicas de detección de suplantación facial para garantizar que solo las personas autorizadas puedan ingresar a determinadas áreas o edificios. Los sistemas pueden incluir el uso de cámaras y algoritmos de detección de suplantación facial para verificar la identidad de los individuos y evitar la suplantación de rostros. Algunos ejemplos de sistemas de control de acceso que cuentan con tecnología *anti-spoofing* son:

- IDEMIA es una empresa líder en tecnología de identificación y seguridad. Ofrece soluciones de control de acceso que utilizan algoritmos de detección de suplantación facial para garantizar una mayor seguridad y prevenir ataques de spoofing [89].

- Suprema es un proveedor de soluciones de control de acceso biométrico. Sus sistemas incorporan algoritmos anti-spoofing para detectar intentos de suplantación facial mediante el análisis de características faciales dinámicas, como movimientos y patrones de parpadeo [90].

Por último, pero no menos importante están las soluciones de prevención de fraudes financieros, algunas instituciones utilizan sistemas comerciales de detección de suplantación facial para prevenir el fraude y garantizar la autenticidad de las transacciones. Estos sistemas pueden requerir la autenticación facial de los usuarios durante las transacciones en línea, utilizando algoritmos de detección de suplantación para evitar el uso de imágenes o videos falsos, algunos ejemplos son:

- *BioCatch* es una empresa que utiliza el comportamiento biométrico y algoritmos *anti-spoofing* para detectar y prevenir el fraude financiero. Su tecnología analiza el comportamiento del usuario durante las transacciones para identificar patrones sospechosos y detectar intentos de suplantación [91].

- *Featurespace* es una empresa que utiliza el aprendizaje automático y algoritmos *anti-spoofing* para prevenir el fraude financiero en tiempo real. Sus soluciones analizan el comportamiento del usuario, así como los patrones de actividad, para identificar anomalías y detectar actividades fraudulentas [92].

Los sistemas mencionados son solo algunos ejemplos de sistemas comerciales que utilizan la detección de suplantación facial. La tecnología avanza constantemente y se espera que aparezcan nuevos sistemas y soluciones en el futuro, impulsados por avances en el aprendizaje automático, la inteligencia y la visión artificiales.

2.3.1 Comentarios

Después de revisar algunos sistemas comerciales que ofrecen soluciones a la suplantación facial, se destacan los siguientes puntos:

- Los sistemas comerciales pueden ser sensibles a factores ambientales como la iluminación, la calidad de la imagen y los ángulos de captura. Dichas condiciones pueden afectar la precisión de la detección y aumentar la tasa de falsos positivos o falsos negativos.

- Limitaciones en la base de datos de referencia, es decir dependen de varias capturas previas de la imagen que se pretende analizar para comparar las características faciales y detectar la suplantación. Sin embargo, si la base de datos no es lo suficientemente amplia o diversa, el sistema puede tener dificultades para identificar suplantaciones que no estén presentes en la base de datos.

- Costo y accesibilidad, algunos sistemas comerciales pueden ser costosos y no estar ampliamente disponibles para todos los usuarios o empresas. Lo que limita su implementación y uso en entornos donde los recursos financieros son limitados.

- Actualizaciones y adaptabilidad, pueden requerir actualizaciones periódicas para mantenerse al día con las nuevas técnicas de suplantación facial. Si los proveedores no proporcionan actualizaciones regulares, los sistemas pueden volverse obsoletos y menos efectivos con el tiempo.

- Hardware especializado, en la mayoría de los casos los sistemas comerciales hacen uso de cámaras de alta resolución, cámaras térmicas, infrarrojo o lectores de huella digital, para garantizar la eficacia de los sistemas.

CAPÍTULO 3
Marco teórico

En el presente capítulo se describen algoritmos y técnicas para la comprensión de la presente investigación, así como las métricas para evaluar los resultados de los algoritmos utilizados para el desarrollo de este trabajo.

3.1 Iluminación

La iluminación de las imágenes es importante en la detección de suplantación facial porque permiten resaltar características faciales, revelar detalles sutiles en una cara, como las arrugas, los poros y los cambios en el color de la piel; distinguir materiales y texturas, por ejemplo, una cara real y una máscara hecha de silicona tendrán propiedades de reflexión y absorción de luz distintas; detectar anomalías en las sombras y proporcionar información contextual relevante sobre el entorno y las condiciones en las que se capturó una imagen, lo que puede incluir la dirección de la luz, la intensidad, la temperatura de color y los reflejos. Al analizar estos aspectos, los expertos en seguridad y los sistemas de detección pueden identificar posibles casos de suplantación facial y proteger la integridad de los sistemas biométricos [93], [94].

Existen diferentes algoritmos y técnicas relacionadas con la iluminación en el procesamiento de imágenes. Los cuales se utilizan para abordar diferentes aspectos de la iluminación, como la normalización, la corrección, la manipulación y la estimación de la iluminación en una imagen, algunos ejemplos son:

- Ecualización de Histograma [95]: Se utiliza para redistribuir los niveles de intensidad en una imagen con el objetivo de mejorar el contraste y la visibilidad de los detalles.

- Ecualización de Histograma Adaptativa [39]: Similar a la ecualización de histograma, pero se adapta localmente a diferentes regiones de la imagen para evitar la amplificación del ruido.

- Retinex multi-escala [96]: Se basan en el modelo Retinex, que busca separar la reflectancia y la iluminación en una imagen. El objetivo es mejorar la apariencia de la imagen y reducir los efectos de variaciones en la iluminación.

La elección del algoritmo a utilizar depende de varios factores, como la naturaleza de los ataques de suplantación que se desean detectar, las características de las imágenes o videos disponibles y los recursos computacionales disponibles. En la presente investigación se determinó el uso de retinex multi-escala debido a tres características que son:

- Sensibilidad a las variaciones de iluminación, el algoritmo puede ayudar a mitigar las diferencias en la iluminación y resaltar anomalías. Lo que puede ser especialmente relevante en la detección de suplantación facial, donde los atacantes pueden manipular las condiciones de iluminación para ocultar sus falsificaciones.

- Conservación de detalles, el objetivo del algoritmo es preservar los detalles y las características en una imagen, lo cual es importante ya que las manipulaciones pueden afectar la calidad y la consistencia de los detalles faciales. Al preservar los detalles, el algoritmo puede ayudar a identificar cambios o manipulaciones sospechosas.

- Robustez frente a ciertas transformaciones, el algoritmo ha demostrado ser relativamente robusto frente a ciertas transformaciones de imagen, como los cambios de iluminación global, lo cual puede ser beneficioso en la detección de suplantación facial.

Retinex multi-escala

El algoritmo utiliza una técnica de descomposición de imagen en múltiples escalas espaciales para separar la reflectancia del objeto y la iluminación. Al considerar múltiples escalas, el algoritmo es capaz de capturar detalles de diferentes tamaños y texturas en la imagen facial, lo que contribuye a mitigar los efectos de variaciones en la iluminación y detectar diferencias sutiles entre una cara genuina y una suplantación [97] [98] [99] [100].

Como primer paso se aplica un filtro gaussiano de tamaño variable a la imagen original para obtener versiones suavizadas en diferentes escalas. Los resultados en la literatura han demostrado que es suficiente con usar tres escalas, una con valor gaussiano bajo (<20), otra con valor alto (>200) y una tercera como valor intermedio [101], [102]. Para el caso del presente trabajo de investigación, después de realizar pruebas con combinaciones de diferentes tamaños se determinó hacer uso de tres escalas con valores de 20, 160 y 300.

Posterior a la definición de la escala con los filtros gaussianos se calcula la reflectancia $R(x,y)$ para cada escala y se realiza una suma ponderada, la cual se define por:

$$R(x,y) = \sum_{n=1}^{N} \omega_n \{log[I(x,y)] - log[F(x,y) * I(x,y)]\} \qquad 1$$

Donde $\sum_{n=1}^{N} \omega_n$ representa la suma ponderada de los filtros gaussianos con diferentes escalas N y un factor de ponderación ω_n. Donde I es la imagen de entrada, $F(x, y)$ es un filtro lineal con núcleo gaussiano y * indica convolución [103].

Para comprender visualmente lo que se busca con el algoritmo de retinex multi-escala, en la Figura 3.1 se muestra un mapa visual con los colores resultantes reflejados en la imagen, después de ser procesada con el algoritmo de acuerdo con el contraste e iluminación que presente.

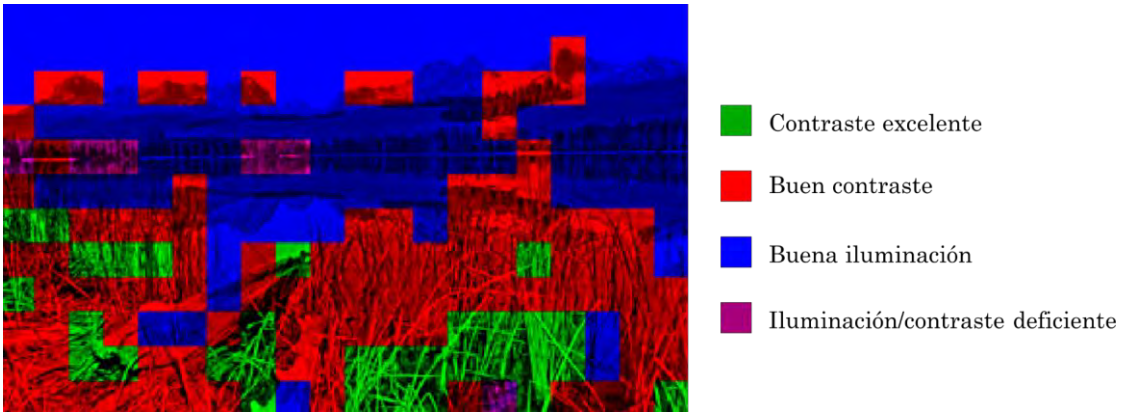

Figura 3.1 Mapa visual que muestra las clases regionales con retinex multi-escala [104].

3.2 Detección del rostro

La localización del rostro es un área específica de la visión por computadora y el procesamiento de imágenes que es de vital importancia en el reconocimiento facial y recientemente en la detección de ataques de suplantación por medio del rostro. Pero la detección facial no es 100% certera, hay muchas diferencias presentes en los rostros humanos, como pose, expresión, posición, orientación y en lo que respecta a la imagen hay factores que le afectan como la iluminación y la resolución [105].

Los algoritmos de detección del rostro utilizan diferentes técnicas y enfoques, algunos de los métodos más comunes son:

- Viola-Jones en 2001 [106]: Segmenta la foto en múltiples subsecciones e intenta encontrar características similares a Haar dentro de cada sub área.

- Características basadas en Haar-like [107]: Se utilizan características simples y rectangulares llamadas Haar-like para detectar regiones faciales en base a la diferencia de intensidad en diferentes partes de la cara.

- Redes Neuronales Convolucionales (CNN) [108]: Estas redes aprenden automáticamente características relevantes para la detección del rostro a través del entrenamiento con grandes conjuntos de datos etiquetados. Las CNN se han vuelto muy populares y efectivas en la detección precisa de rostros.

- *Histogram of Oriented Gradients* (HOG) [109]: Se extraen características locales de las imágenes y se buscan patrones que correspondan a la presencia de rostros, el algoritmo se puede aplicar mediante la biblioteca Dlib [110], la cual es de código abierto.

En la Tabla 3.1 se muestra una comparativa de las características importantes para el presente trabajo de investigación con referencia a los algoritmos de detección del rostro. En la Tabla 3.1 se observa que la diferencia que puede determinar el uso de los métodos convencionales o las redes neuronales radica esencialmente en la cantidad de datos necesarios para el entrenamiento, debido a que si no se cuenta con grandes cantidades de imágenes para el entrenamiento las redes neuronales carecerían de precisión y robustez antes entornos no controlados. Por lo tanto, tomando en cuenta que el presente trabajo cuenta con una cantidad limitada de datos, se considera no hacer uso de hardware especializado y se busca tener una eficiencia computacional cercana al tiempo real (igual o menor a 3 segundos), se determinó utilizar el algoritmo de Haar-like.

Tabla 3.1 Características de localizadores del rostro.

Características	Viola-Jones	Haar-like	CNN	HOG
Robusto a variaciones de iluminación	✓	✓	✓	✓
Robusto a oclusión	✘	✘	✓	✘
Robusto a cambios de escala	✘	✘	✓	✘
Robusto a la orientación	✘	✘	✓	✘
Baja complejidad computacional	✓	✓	✘	✓
Alta eficiencia computacional	✓	✓	✘	✘
Tiempos de entrenamiento cortos	✓	✓	✘	✓
Entrenamiento con datos limitados	✓	✓	✘	✓
Bajo riesgo de sobreajuste	✓	✓	✘	✓
Precisión en condiciones adversas	✘	✘	✓	✘
Disponibilidad de modelos pre-entrenados	✘	✓	✘	✓

Haar-like [107]

La detección del rostro utilizando Haar-like es un método basado en características visuales que busca patrones específicos en una imagen para identificar la presencia de un rostro. Hacen uso del cambio en los valores de contraste entre grupos de píxeles rectangulares adyacentes en lugar de usar los valores de intensidad de un píxel. Las variaciones de contraste entre los grupos de píxeles se utilizan para determinar las áreas claras y oscuras relativas.

Dos o tres grupos adyacentes con una variación relativa de contraste forman una característica Haar. Las características de tipo Haar, como se muestra en la Figura 3.2, se utilizan para detectar una imagen; además, se pueden escalar fácilmente aumentando o disminuyendo el tamaño del grupo de píxeles que se está examinando, lo que permite tener una gran flexibilidad en su aplicación [111].

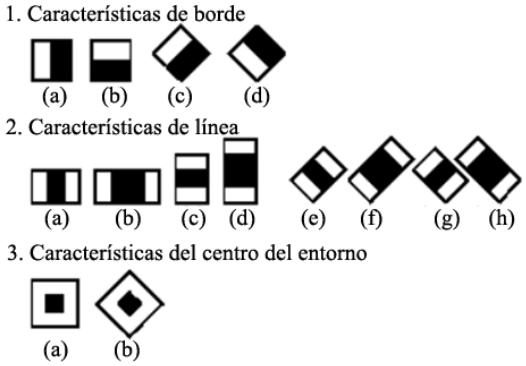

Figura 3.2 Características comunes de Haar [107].

Las características rectangulares simples de una imagen se calculan utilizando una representación intermedia de una imagen, llamada imagen integral [106]. La imagen integral es una matriz que contiene las sumas de los valores de intensidad de los píxeles ubicados directamente a la izquierda de un píxel y directamente encima del píxel en la ubicación *(x, y)*. Entonces, si *A [x, y]* es la imagen original y *AI [x, y]* es la imagen integral, se tiene que la imagen integral se calcula como se muestra en la ecuación 2.

$$AI[x,y] = \sum_{x' \le x, y' \le y} A(x', y') \qquad 2$$

3.3 Modelos de color

Los modelos de color son de gran ayuda en la detección de suplantación facial debido a que contribuyen a identificar anomalías cromáticas, extraer características faciales relevantes, discriminar texturas, detalles, y segmentar las regiones de interés. Los modelos contribuyen a analizar los componentes cromáticos de una imagen facial y detectar discrepancias o alteraciones inusuales que podrían indicar una suplantación. Además, facilitan la extracción de características distintivas del rostro, ayudando a distinguir entre un rostro genuino y uno falsificado. Al combinar los modelos de color con otros enfoques de detección, se puede mejorar la precisión y efectividad de los sistemas de detección de suplantación fácil [45].

En el transcurso de la investigación se encontró que se utilizan principalmente tres modelos de color en lo que respecta a la detección de suplantación facial: RGB (*Red, Green, Blue*), YCbCr (Luminancia, Crominancia Azul y Roja) y HSV (*Hue, Saturation, Value*). El modelo RGB es el más común dentro de los modelos y se basa en la combinación de tres componentes primarios de color: rojo, verde y azul. Es utilizado en imágenes digitales y captura información detallada de color y textura en los rostros [112]. Por otro lado, el modelo YCbCr es especialmente útil para detectar cambios en la crominancia, ya que separa la información de luminancia (brillo) de los componentes de color. Lo que permite identificar alteraciones en los colores de la piel y detectar posibles manipulaciones o suplantaciones en las imágenes faciales [113]. El modelo HSV se basa en la percepción humana del color y divide la imagen en tres componentes: matiz, saturación y valor. El matiz representa el tipo de color, la saturación indica la intensidad o pureza del color, y el valor se refiere al brillo de la imagen. Por lo que al detectar cambios en el matiz y la saturación, se puede indicar manipulaciones en la apariencia del rostro [112]. Por lo tanto, se determinó el uso de los canales de color YCbCr y HSV en el presente trabajo debido a sus características.

YCbCr [114], [115]

El uso del modelo de color YCbCr en la detección de suplantación facial es ampliamente investigado debido a que contribuye en la separación de la información de luminancia y crominancia. Al separarlos, se pueden detectar cambios sutiles en los componentes de color que pueden indicar manipulación o alteración en la imagen facial. Además, el modelo YCbCr es resistente a los cambios en la luminancia, lo que ayuda a minimizar los efectos de variaciones en la iluminación en la detección de suplantación facial. Además, el modelo de color YCbCr es utilizado en la industria de la imagen, lo que facilita la comparación y el intercambio de datos con otros sistemas. En conjunto, estas características hacen que el modelo YCbCr sea una opción eficaz y relevante para la detección de suplantación facial.

YCbCr se compone de Y que es el brillo, Cb y Cr son la cromaticidad, Cb es el componente azul y el componente Cr es el componente rojo. El brillo y la cromaticidad se pueden separar, de modo que el color de la piel se junte en un rango pequeño en el espacio de color. En gran medida, la influencia del brillo se elimina y la eficiencia de cálculo es relativamente alta [114]. Las ecuaciones del canal de color YCbCr se presentan a continuación:

$$Y = \left(\frac{1}{256}\right) * [(16 * 256 + 129 * G) + (66 * R + 25 * B)] \quad 3$$

$$Cb = \left(\frac{1}{256}\right) * [(128 * 256 + 112 * B) - (38 * R + 74 * G)] \quad 4$$

$$Cr = \left(\frac{1}{256}\right) * [(128 * 256 + 112 * R) - (94 * G + 18 * B)] \quad 5$$

Donde R'G'B' son valores RGB con corrección de gamma y las señales de entrada y salida son de valores de 8 bits.

HSV [115]

El espacio de color HSV (tono, saturación, valor) pertenece a el sistema de coordenadas de color orientado al tono que se corresponden más estrechamente con la percepción humana del color. Este espacio de color orientado al usuario se basa en el atractivo intuitivo del tinte, la sombra y el tono. El espacio de color HSV, propuesto originalmente en Smith [116], es cilíndrico y está convenientemente representado por el modelo de cono hexagonal que se muestra en la Figura 3.3 [117].

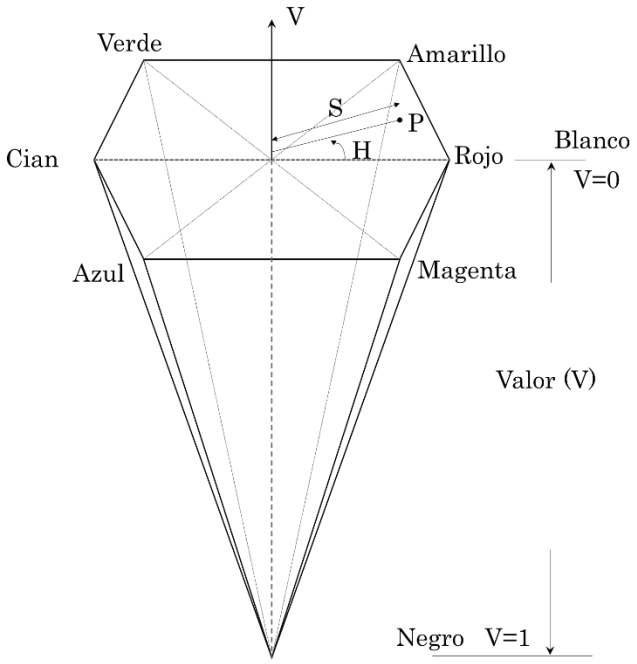

Figura 3.3 Espacio de color HSV [117]

La transformación del color se realiza con las siguientes ecuaciones:

$$R' = \frac{R}{255}; G' = \frac{G}{255}; B' = B/255 \quad (6)$$

$$C_{max} = MAX(R', G', B') \quad (7)$$

$$C_{min} = MIN(R', G', B') \quad (8)$$

$$\Delta = C_{max} - C_{min} \quad (9)$$

$$H = \begin{cases} 60° * \left(\frac{G' - B'}{\Delta} mod 6\right), C_{max} = R' \\ 60° * \left(\frac{B' - R'}{\Delta} + 2\right), C_{max} = G' \\ 60° * \left(\frac{R' - G'}{\Delta} + 4\right), C_{max} = B' \end{cases} \quad (10)$$

$$S = \begin{cases} 0, \Delta = 0 \\ \frac{\Delta}{C_{max}}, \Delta <> 0 \end{cases} \quad (11)$$

$$V = C_{max} \quad (12)$$

Donde H es un tipo de color y el rango es de 0 a 360 grados. S es la intensidad del color y el rango es 0-100%. V es el brillo de un color y el rango es 0-100%.

3.4 Algoritmos de textura

La textura se refiere a los patrones de distribución de píxeles en una imagen, como arrugas, poros, patrones de piel y características distintivas del rostro. Dichos detalles son únicos y difíciles de falsificar o replicar de manera precisa. Los algoritmos de textura son útiles para abordar las limitaciones de otros métodos de detección de suplantación facial que se centran en características globales, como la forma o la apariencia general del rostro. La textura proporciona información adicional y complementaria que puede ayudar a mejorar la precisión y robustez de la detección de suplantación facial. A continuación de describen tres algoritmos de especial relevancia para la investigación.

LBP

El operador LBP (*Local Binary Patterns*) fue propuesto por Ojala et al. [40], como una técnica de descripción de texturas locales en imágenes. El objetivo principal de LBP es capturar patrones locales de píxeles en una imagen y representarlos mediante un código binario.

LBP caracteriza la estructura espacial de un parche de imagen local codificando las diferencias entre el valor de píxel del punto central y los de sus vecinos, considerando solo los signos para formar un patrón binario. El valor decimal resultante del patrón binario generado se usa para etiquetar el píxel seleccionado [118].

Por lo tanto, dado un píxel x_c en la imagen, la respuesta LBP se calcula comparando su valor con los de sus p píxeles vecinos $\{x_{r,p,n}\}_{n=0}^{p-1}$, distribuidos uniformemente en ángulo en un círculo de radio r con centro en x_c como:

$$LBP_{r,p}(x_c) = \sum_{n=0}^{p-1} s(x_{r,p,n} - x_c)2^n, s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad 13$$

Donde $s()$ es la función de signo. Si las coordenadas de x_c son (0, 0), entonces las coordenadas de $x_{r,p,n}$ vienen dadas por $(-r\ \sin(2\pi\ n/p), r\ \cos(2\pi\ n/p))$. En la Figura 3.4 se ilustran conjuntos de vecinos circularmente simétricos para varios (p, r), siendo (8,1) los valores por defecto [40]. Los valores de gris $x_{r,p,n}$ de los vecinos que no caen exactamente en el centro de los píxeles se estiman por interpolación.

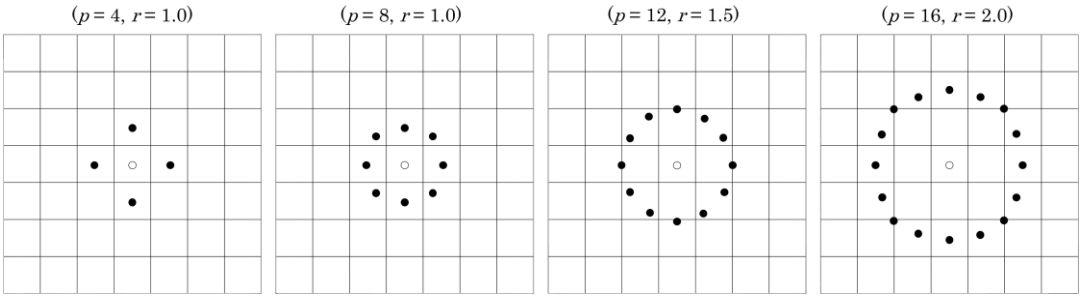

Figura 3.4 Conjuntos de vecinos circularmente simétricos para diferentes (p, r) (basada en [40]).

En una imagen real LBP examina las regiones de interés, como la textura de la piel, los ojos o los labios, y calcula los patrones locales binarios que representan las variaciones de intensidad de los píxeles en esas regiones. Mediante la comparación de los patrones de una imagen sospechosa con los de imágenes reales de referencia, se pueden identificar diferencias significativas que indiquen la presencia de suplantación facial. Las diferencias pueden manifestarse como patrones atípicos, ausencia de patrones naturales o presencia de patrones incoherentes en la imagen falsa, como es el caso en la Figura 3.5, en donde se puede observar las diferencias en los patrones visuales entre una imagen real y una réplica de la imagen por medio de un celular, donde las áreas negras en la réplica se pueden interpretar como ausencia de textura, en tanto el rostro si bien contiene textura de manera visible, se puede notar que existe diferencia con la imagen real.

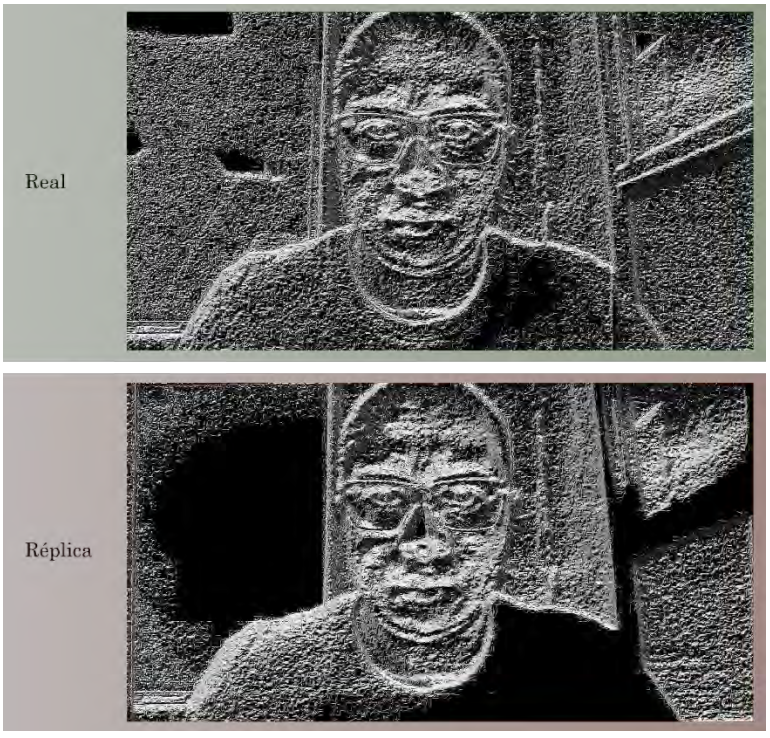

Figura 3.5 Comparativa visual de $LBP_{8,1}$ entre una imagen real y una réplica.

ELBP [118]

El uso de ELBP (*Extended Local Binary Patterns*) en lugar de LBP (*Local Binary Patterns*) en la detección de suplantación facial presenta varias ventajas. ELBP amplía la capacidad de capturar información espacial al considerar las relaciones entre los puntos vecinos en diferentes direcciones y distancias, permitiendo una representación más completa y detallada de la textura facial, lo que puede ser especialmente útil para detectar manipulaciones sutiles o complejas en la imagen. Al incorporar las relaciones espaciales extendidas, ELBP proporciona una mayor discriminación entre características auténticas y falsificadas, mejorando así la precisión y confiabilidad de la detección de suplantación facial. Además, ELBP es más robusto frente a variaciones en la iluminación, rotación y escala, lo que lo convierte en una opción más adecuada para lidiar con condiciones adversas [118].

El funcionamiento de ELBP es similar al LBP, mientras que LBP codifica sólo la relación entre un punto central y sus vecinos, ELBP está diseñado para codificar relaciones espaciales distintivas en una región local y, por lo tanto, contiene más información espacial. ELBP consta de tres descriptores similares a LBP ELBP_CI, ELBP_NI y ELBP_RD que exploran información de la intensidad del píxel central de sus píxeles vecinos, diferencias radiales y diferencias angulares, respectivamente.

La intensidad del píxel central se establece como umbral frente a β que es la media de toda la imagen.

$$ELBP\_CI(x_c) = s(x_c - \beta) \qquad 14$$

donde x_c son las coordenadas del píxel central, de tal caso que si las coordenadas de x_c son (0,0), entonces las coordenadas de $x_{r,p,n}$ en la 15, vienen dadas por $(-r \sin(2\pi n/p), r \cos(2\pi\, n/p))$. Mientras que los valores de gris $x_{r,p,n}$ de los vecinos que no caen exactamente en el centro de los píxeles se estiman por interpolación.

En lugar de utilizar el valor de gris del píxel central como el valor de umbral, como se usa en LBP, ELBP_NI utiliza el promedio de las intensidades de los píxeles vecinos para generar el patrón binario. ELBP_NI se define como el umbral frente a la media local $\beta_{r,p} = \frac{1}{p}\sum_{n=0}^{p-1} x_{r,p,n}$.

$$ELBP\_NI_{r,p}(x_c) = \sum_{n=0}^{p-1} s\left(x_{r,p,n} - \beta_{r,p}\right)2^n \qquad 15$$

En paralelo a los descriptores basados en intensidad ELBP_NI y ELBP_CI, ELBP_RD se deriva de las diferencias de píxeles en direcciones radiales:

$$ELBP\_RD_{r,r-1,p}(x_c) = \sum_{n=0}^{p-1} s(x_{r,p,n} - x_{r-1.,p.n})2^n \qquad 16$$

El operador ELBP permite un análisis multi-escala mediante la variación de los parámetros (r, p); es decir, cualquier radio y número de píxeles en la vecindad. Para el presente trabajo se modificaron los valores por default (radio =1, vecinos=8) de ELBP y se optó por utilizar 16 vecinos con un radio de tamaño 2. Al emplear un tamaño mayor de 2 y 16 en lugar de 1 y 8, respectivamente, se logra una representación más robusta y detallada de las texturas faciales. Lo que es crucial en la detección de suplantación facial, ya que los patrones de textura pueden variar significativamente entre imágenes genuinas y falsificadas. Al considerar áreas amplias y un mayor número de puntos vecinos, se pueden capturar características complejas y sutiles, permitiendo una mejor discriminación entre características auténticas y manipuladas en el rostro. Además, el uso de valores de 2,16 ofrece una mayor resistencia a las variaciones en la iluminación, rotación y escala, lo que mejora la capacidad de detección en diferentes condiciones [119].

La Figura 3.6 muestra un ejemplo del resultado de aplicar a una imagen $\mathrm{ELBP}_{8,1}$ y $\mathrm{ELBP}_{16,2}$. En donde se observa que, al comparar visualmente una imagen real con una réplica de la misma con valores de 8,1, existe una diferencia de textura, debido a que en apariencia la réplica tiene zonas (círculos rojos), que presentan una mayor pérdida de textura en comparación con la real. En comparación cuando se aplican valores de 16,2 la diferencia en textura se hace mayor, como se observa en la Figura 3.6, el rostro de la réplica presenta zonas (círculos rojos) que de forma visual contienen menos información de textura que la imagen real. Por lo que utilizar valores de 16, 2 puede favorecer la detección de ataques de suplantación facial.

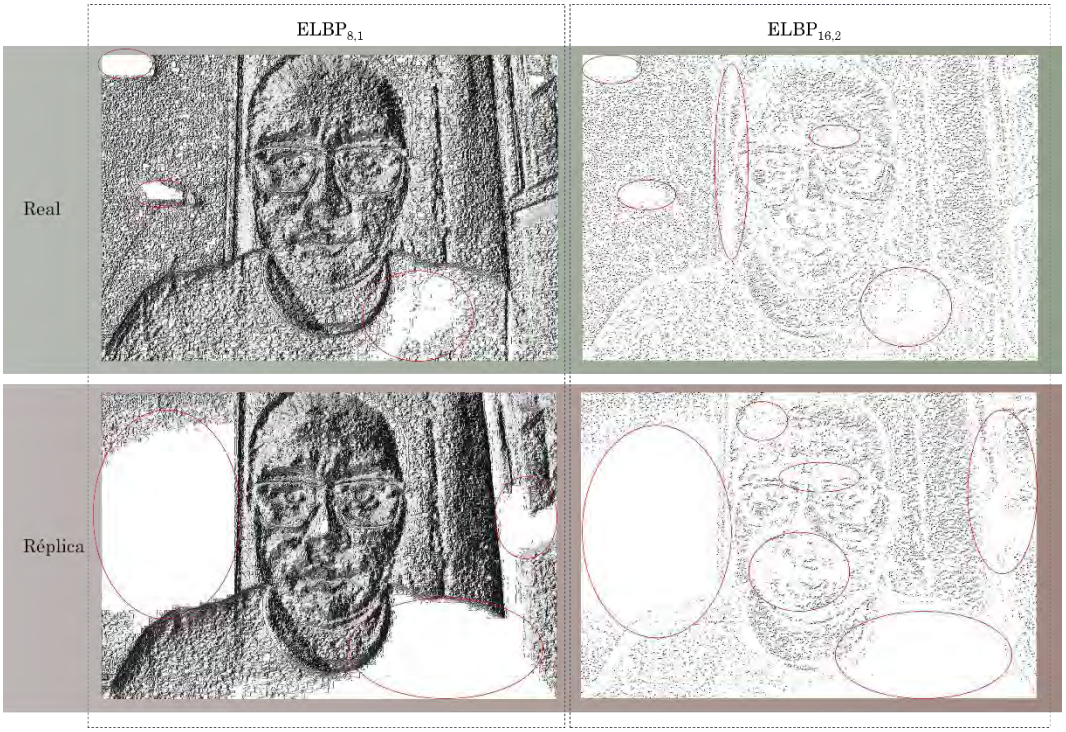

Figura 3.6 Comparativa de $ELBP_{8,1}$ y $ELBP_{16,2}$

BSIF [43]

Es un descriptor similar a LBP, la diferencia radica en la forma en que se aprenden los filtros. En el caso de BSIF, los filtros se aprenden de las imágenes naturales mientras que los filtros LBP se predefinen manualmente [120]. Por lo tanto, dada la imagen del rostro $I_f(m,n)$ y un filtro $BSIF_i^{k\times k}$, la respuesta del filtro se obtiene de la siguiente manera:

$$r_i = \sum_{m,n} I_f(m,n) * W_i^{k\times k}(m,n) \qquad 17$$

Donde * denota la operación de convolución, m y n denotan el tamaño del parche de la imagen en el rostro y $W_i^{k\times k}, \forall i = \{1,2,\dots,L\}$ denota la longitud del filtro BSIF e $k \times k$ indica el tamaño del filtro BSIF, cuya respuesta se puede binarizar para obtener la cadena de la siguiente manera:

$$b_i = \begin{cases} 1, & \text{Si } r_i > 0 \\ 0, & \text{Si no} \end{cases} \qquad 18$$

Finalmente, las características de BSIF se extraen considerando cada píxel (m, n) como un conjunto de valores binarios obtenidos del número L de filtros lineales. Matemáticamente, para un píxel dado (m, n) y su correspondiente representación binaria $b_i(m, n)$, las características codificadas por BSIF se obtienen de la siguiente manera:

$$BSIF_i^{k\times k}(m,n) = \sum_{i=1}^{L} (b_i(m,n) \times (2^{i-1})) \qquad 19$$

En una imagen real, las características de textura capturadas por los filtros serán consistentes y seguirán patrones naturales presentes en rostros reales, loa cuales incluyen detalles finos, texturas suaves y estructuras coherentes. Por otro lado, en una imagen falsa o réplica, es probable que haya discrepancias en las características de textura capturadas por los filtros de BSIF. Lo que se puede deber a los objetos sintéticos con los cuales se realiza la suplantación, lo que afecta la respuesta de los filtros de BSIF.

Al comparar las características de textura extraídas por el algoritmo BSIF de una réplica con las de una imagen real de referencia, es posible identificar diferencias significativas como la falta de detalles finos o texturas inconsistentes en la imagen falsa. En la Figura 3.7 se puede observar como la imagen real presenta texturas más suaves a diferencia de la réplica, lo que se puede ver con mayor claridad al realizar un acercamiento al ojo en donde las texturas en la imagen real a manera visual son sutiles mientras que en la réplica los bordes y la textura son particularmente pronunciados.

Figura 3.7 Comparativa visual de BSIF entre una imagen real y una réplica.

PRNU [46]

Durante el proceso de adquisición de las imágenes existen diversos factores que influyen en la generación de imperfecciones y ruido. Incluso si el sensor de imágenes toma una imagen de una escena iluminada de manera absolutamente uniforme, la imagen digital resultante aún exhibirá pequeños cambios en la intensidad entre los píxeles individuales, esto se debe en parte al ruido de disparo (también conocido como ruido fotónico [121], [122]). Debido a esta propiedad, el patrón de ruido está presente en cada imagen que toma el sensor por lo que, al ser una distorsión sistemática, puede parecer que no es correcto llamarlo ruido. No obstante, el patrón de ruido es un término establecido en la literatura sobre sensores de imagen [121], [122]. Al utilizar PRNU en la detección de suplantación facial, se busca aprovechar esta huella única dejada por la cámara para distinguir entre imágenes reales y falsas.

Los dos componentes principales del patrón de ruido son: el ruido de patrón fijo (FPN por sus siglas en inglés) y el ruido de falta de uniformidad de la foto respuesta (PRNU por sus siglas en inglés).

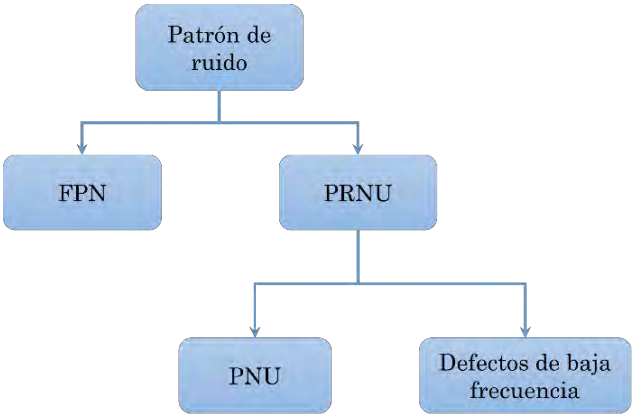

Figura 3.8 Patrón de ruido de sensores de imágenes (basada en [46]).

En imágenes naturales, la parte dominante del patrón de ruido es el PRNU, el cual es causado principalmente por la falta de uniformidad de píxeles (PNU), que se define como una sensibilidad diferente de los píxeles a la luz causada por la falta de homogeneidad de las ondas de silicio y las imperfecciones durante el proceso de fabricación del sensor.

El origen del ruido PNU hace que sea poco probable que incluso los sensores que vienen de la misma fabrica contengan patrones PNU correlacionados. Como tal, el ruido de la PNU no se ve afectado por la temperatura o la humedad ambiental; sin embargo, la refracción de la luz en partículas de polvo y superficies ópticas y la configuración del *zoom* si contribuyen al ruido PRNU, lo cual es una ventana de oportunidad al utilizarlo en suplantación de identidad por medio de objetos sintéticos.

Para el cálculo de PRNU se denotan los conteos de fotones que idealmente serían registrados por el sensor debido a la luz entrante como $x = (x_{ij}), i = 1, \dots, m, j = 1, \dots, n$, donde $m \times n$ es la resolución del sensor que denota el ruido de disparo como $\eta = (\eta_{ij})$, el ruido aditivo aleatorio como $\varepsilon = \varepsilon_{ij}$, y la corriente oscura como $c = c_{ij}$, la salida del sensor $y = (y_{ij})$ se puede expresar de la siguiente forma:

$$y_{ij} = f_{ij}(x_{ij} + \eta_{ij}) + c_{ij} + \varepsilon_{ij} \qquad 20$$

Donde los factores f_{ij} suelen estar cercanos al 1 y capturan el ruido PRNU, que es un ruido multiplicativo.

Cuando se realiza una suplantación facial y se captura una imagen falsa con la misma cámara, la huella PRNU de la cámara seguirá presente, pero la imagen manipulada puede tener otras características distintivas debido a la falsificación. En la Figura 3.9 se muestra cómo se visualiza el ruido PRNU en una imagen real y su réplica capturadas con la misma cámara (Toshiba *web cam* HD 0.92 MP). En donde se observa que existen diferencias a nivel de ruido y ausencia del mismo a pesar de ser usar la misma cámara.

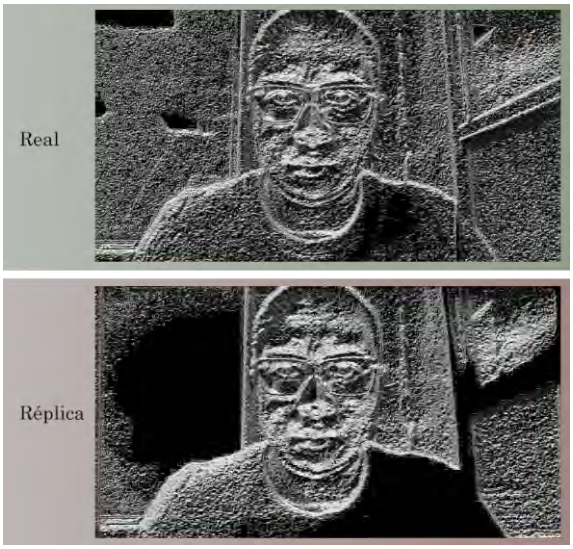

Figura 3.9 Comparativa visual del resultado de aplicar PRNU entre una imagen real y una réplica con la misma cámara.

FeatherNetB [123]

Las redes neuronales convolucionales son un tipo de redes neuronales multicapa de retroalimentación que, consisten en una serie de capas convolucionales y agrupadas seguidas de una o más capas completamente conectadas. En una capa convolucional, las neuronas se organizan en varios parches rectangulares separados, comúnmente conocidos como mapas de características [124]. Esta configuración los hace aptos para procesar datos multidimensionales, como imágenes en color 2D y señales de video 3D. Las capas convolucionales están diseñadas para explotar la correlación espacial en una capa neuronal de entrada, empleando un patrón de conectividad local entre neuronas en capas consecutivas; en otras palabras, en una capa convolucional cada neurona está conectada sólo a pequeños parches de neuronas, comúnmente denominados campos receptivos locales, que residen en diferentes mapas de características de la capa convolucional anterior. En la Figura 3.10 se observa la imagen de la estructura de una red convolucional típica.

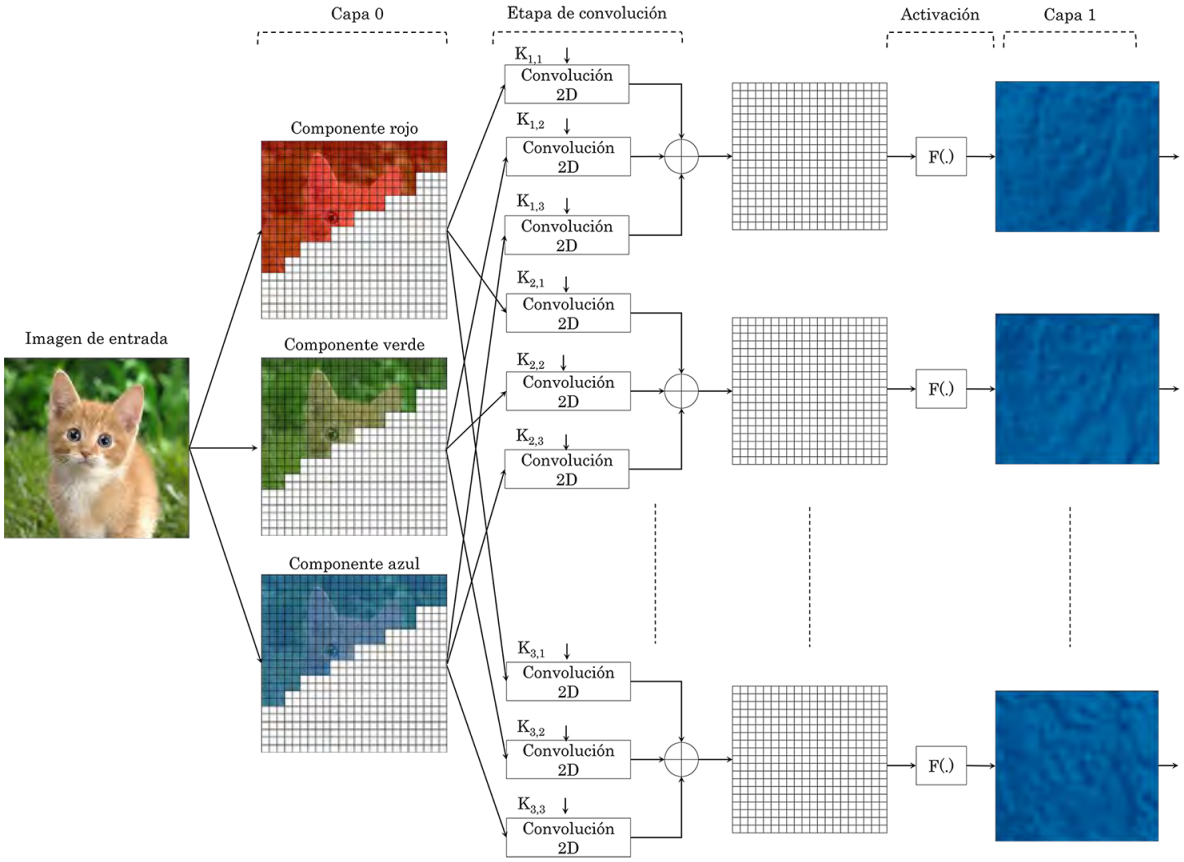

Figura 3.10 Esquema de una red convolucional típica (basada en [125]).

En donde, a partir de una imagen de entrada en formato RGB, la red separa los colores en capas convolucionales. En primer lugar, se realiza esta separación para luego proceder con el intercambio de parámetros, donde todas las neuronas que pertenecen a la misma capa convolucional comparten el mismo conjunto de pesos.

FeatherNet-B al igual que otras CNN, funciona extrayendo características a través de capas de convolución, se basa en la arquitectura de MobileNetV2 [126], su principal características es la optimización para aplicaciones con restricciones de recursos computacionales, como dispositivos móviles o sistemas embebidos. La idea detrás de FeatherNet-B es lograr un equilibrio entre la precisión del modelo y la eficiencia computacional, reduciendo al mismo tiempo la cantidad de parámetros y la carga computacional requerida para ejecutar la red. El objetivo principal es proporcionar una opción liviana pero efectiva para tareas de clasificación de imágenes en dispositivos con recursos computacionales limitados sin comprometer demasiado la precisión. Por ello es una opción viable al utilizarla en sistemas de detección facial con limitaciones en el hardware.

En la Figura 3.11 se muestra el esquema de funcionamiento de la red FeatherNet-B donde c es el número de canales de entrada, el bloque A contiene los bloques residuales invertidos propuestos en MobilenetV2 [126] y se utiliza como bloque de construcción principal. Los factores de expansión son los mismos que en MobilenetV2 [126]. El bloque B es el módulo de muestreo descendente de FeatherNetB. En Zhang et. al [127] se demostró que la agrupación promedio (AP) beneficia el rendimiento, debido a su capacidad para incorporar información de múltiples escalas y agregar características en diferentes campos receptivos. Por lo tanto, la agrupación promedio (núcleo 2 × 2 con salto = 2) se introduce en el Bloque B. Además, en la red ShuffleNet [128], el módulo de muestreo descendente se une a la capa de agrupamiento promedio 2 × 2 con un salto = 2 para obtener un excelente rendimiento. Li y col. [129] sugirieron que aumentar la capa de agrupación promedio funciona bien e impacta poco en el costo computacional.

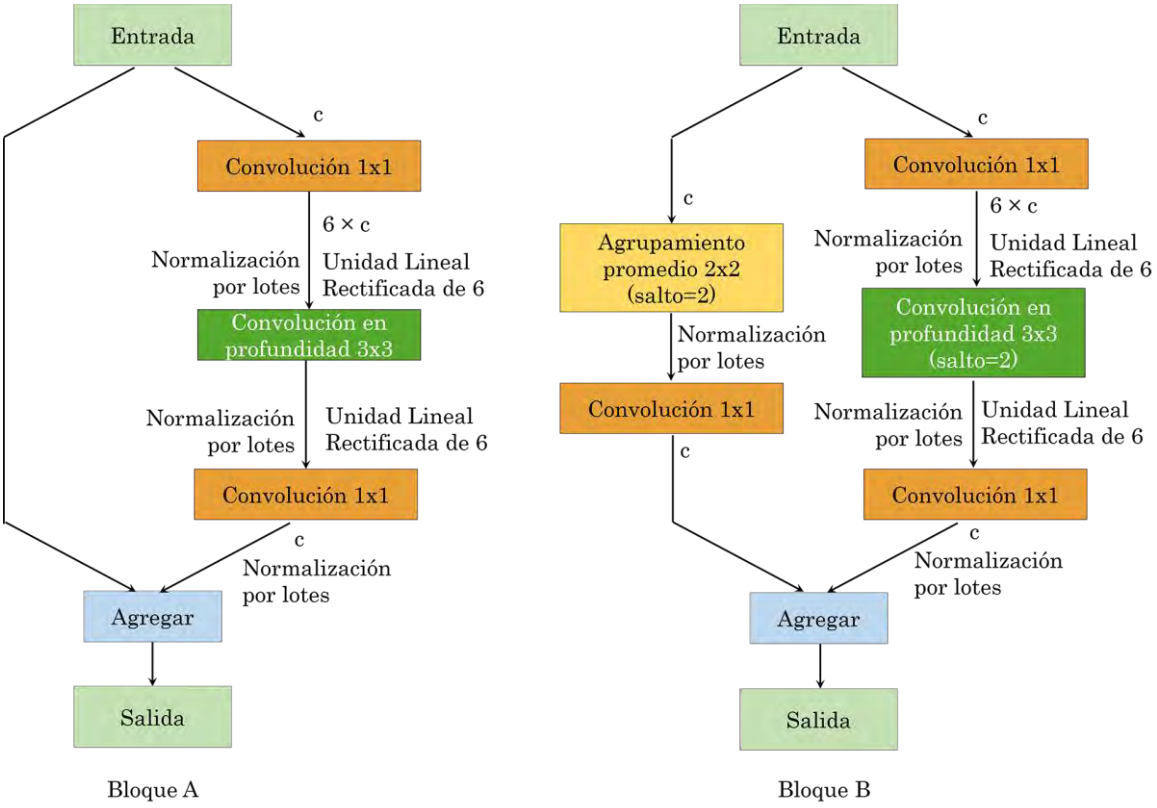

Figura 3.11 Bloques A y B de FeatherNet-B (basada en [123]).

La principal ventaja de FeatherNet-B en comparación con las redes neuronales clásicas es que ofrece una mayor eficiencia computacional al ser ligera y requerir menos recursos, lo que la hace adecuada para implementaciones en tiempo real y en dispositivos con recursos limitados.

FeatherNet-B también cuenta con un modelo entrenado específicamente para detección de suplantación facial con un conjunto de datos faciales multimodales (*Multi-Modal Face Dataset* MMFD) [130] el cual está conformado por 15415 imágenes reales y 28438 imágenes falsas de 15 sujetos bajo diversas condiciones de captura. Además, la red ha demostrado ser resistente a ataques de suplantación avanzados, como el uso de máscaras, fotografías impresas o videos manipulados [123], lo que la convierte en una opción adecuada para comparar con enfoques de visión clásicos.

3.5 Algoritmos de clasificación

Los algoritmos de clasificación son importantes en los sistemas de visión por computadora, ya que permiten identificar, categorizar y comprender la información visual en imágenes y videos. Desempeñan un papel importante en el reconocimiento, detección y segmentación de objetos, así como en la clasificación de patrones y características visuales. Además, son clave en la toma de decisiones y acciones basadas en la interpretación de la información visual. Por lo que en los sistemas de detección de suplantación facial son de suma importancia para lograr distinguir entre rostros genuinos y falsificados. En la Figura 3.12 se ilustra el objetivo de la etapa de clasificación dentro del método de detección de suplantación facial. En la etapa de entrenamiento se utilizan imágenes del banco de creación propia para generar los modelos en archivos .xml, que serán utilizados para la verificación de las imágenes a analizar.

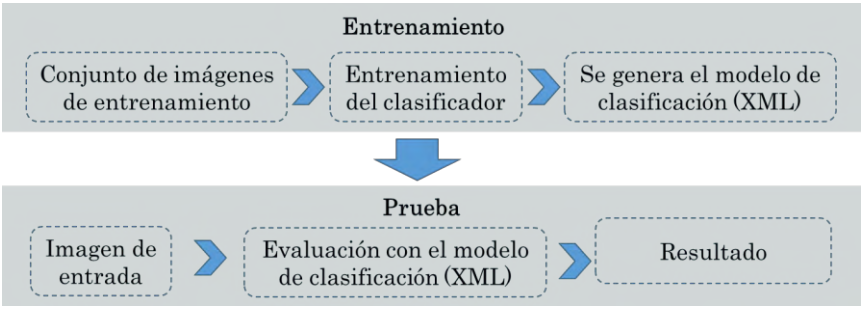

Figura 3.12 Objetivo del algoritmo de clasificación en el sistema.

En la literatura se encontró que se emplean diferentes clasificadores para abordar el problema de la verificación de identidad, algunos de los comúnmente empleados son:

- Máquinas de vectores de soporte (MVS) [131]: Es un algoritmo de aprendizaje supervisado utilizado para clasificar datos en dos categorías. En la detección de suplantación facial, es usado debido a su capacidad para realizar clasificaciones precisas y manejar eficientemente problemas con un gran número de características o variables [132], [133].

- Bosque aleatorio (*random forest*) [134]: Los bosques aleatorios son un conjunto de árboles de decisión que trabajan en paralelo para realizar la clasificación. El clasificador es utilizado en la detección de suplantación facial debido a su capacidad para manejar características complejas, su resistencia al ruido y las variaciones en los datos [51], [135].

- Perceptrón Multicapa (Multilayer Perceptron, MLP) [136]: Es una arquitectura de red neuronal artificial compuesta por múltiples capas de neuronas interconectadas. Cada neurona en una capa está conectada a las neuronas de la capa siguiente, y utiliza una función de activación no lineal para realizar la clasificación. En la detección de suplantación facial, el perceptrón multicapa puede modelar relaciones no lineales entre las características de las imágenes faciales lo que facilita su capacidad de generalizar y clasificar correctamente nuevas imágenes [1], [137].

- Árboles de Decisión [138]: Son estructuras de clasificación que dividen el espacio de características en ramas y nodos, basándose en reglas de decisión. En la detección de suplantación facial, se pueden construir árboles de decisión para clasificar los rostros en categorías de genuinos o falsificados, basándose en características como texturas, bordes y detalles faciales. La diferencia respecto a los bosques aleatorios es que los árboles de decisión son clasificadores individuales basados en reglas de decisión mientras que los bosques aleatorios son un conjunto8ensamble) de árboles que trabajan juntos para realizar la clasificación final [4].

En la presente investigación se utilizan los cuatro clasificadores (MVS, bosque aleatorio, MLP y árboles de decisión) para determinar qué algoritmo es más efectivo en diferentes escenarios de suplantación facial. Además, brindar una comprensión profunda de los factores que influyen en la clasificación, como las características utilizadas y la capacidad de generalización. De esta manera poder seleccionar el clasificador adecuado para el problema abordado en el trabajo.

Un aspecto importante en los algoritmos de clasificación es el cálculo de los parámetros, ya que permite mejorar el rendimiento del modelo, adaptarlo al problema específico, controlar el sobreajuste, subajuste, y comprender mejor su comportamiento. Ajustar los parámetros puede conducir a una clasificación precisa y generalizar de manera efectiva. A continuación, se explican los parámetros utilizados en la experimentación en los cuatro clasificadores.

MVS [131]

Las máquinas de vectores de soporte utilizan una función de kernel para mapear los datos de entrada a un espacio de características de mayor dimensión, donde se busca encontrar el hiperplano de separación óptimo. Los tipos comunes de kernels son el lineal, polinomial, radial (RBF) y sigmoidal. En la presente investigación se utiliza un kernel polinomial debido a que se destaca por su capacidad para capturar relaciones no lineales entre las características, lo cual es fundamental en la detección de suplantación facial, donde las transformaciones pueden generar relaciones complejas. Además, su flexibilidad en la representación de características y su resistencia al ruido, lo hacen una opción viable. En comparación, los kernels lineal, radial y sigmoidal pueden ser menos efectivos para capturar estas relaciones no lineales y ser más sensibles al ruido [139].

Además del kernel, es necesario encontrar un equilibrio adecuado de los parámetros de las MVS mediante experimentación para evitar el sobreajuste o el subajuste. En la presente investigación se realizó una búsqueda exhaustiva al realizar combinaciones mediante experimentación de los parámetros considerando lo siguiente:

- Factor de penalización (C), un valor más alto de C penalizará más los errores de clasificación, lo que puede llevar a un modelo más complejo y ajustado a los datos de entrenamiento.

- Grado polinomial (D), determina la complejidad y la flexibilidad del modelo. Un grado bajo permitirá un modelo más complejo capaz de capturar relaciones no lineales más complicadas en los datos. Sin embargo, un grado muy alto puede llevar al sobreajuste.

- Coeficiente de sesgo (COEF), un valor alto de COEF puede permitir una flexibilidad adicional al modelo.

- Parámetro de suavizado (GAMMA), Un valor bajo de gamma hará que la influencia sea amplia, mientras que un valor alto hará que la influencia sea más localizada.

Por lo tanto, para las máquinas de vectores de soporte se utilizan los siguientes parámetros:

- Kernel: Polinomial
- D: 1
- GAMMA: 0.006
- COEF: 0.1
- C: 3000

Bosque aleatorio o *random forest* [134]

Es un algoritmo de aprendizaje automático que utiliza múltiples árboles de decisión para realizar la clasificación o regresión. Los parámetros que se deben ajustar son:

- Árboles en el bosque, un mayor número de árboles puede mejorar la precisión del modelo, pero también puede aumentar el tiempo de entrenamiento y la complejidad del modelo.

- Profundidad máxima, controla la complejidad de los árboles y evita el sobreajuste. Si se establece en *None*, los árboles se expandirán hasta que todas las hojas sean puras o hasta que las muestras se dividan completamente.

- Mínimo de muestras requeridas para formar una hoja, un valor alto evita la creación de hojas.

- Máximo de características, es el número de características que se consideran al buscar la mejor división en cada nodo. Puede ser un número entero, un valor flotante o una cadena.

En la presente investigación se realizó una búsqueda exhaustiva de los valores para los parámetros al realizar combinaciones mediante experimentación, obteniendo los siguientes valores:

- Número de árboles en el bosque: 100
- Profundidad máxima: INT_MAX
- Mínimo de muestras requeridas: 10
- Máximo de características: 50

Perceptrón multicapa [136]

Es una arquitectura de red neuronal artificial compuesta por múltiples capas de neuronas, incluyendo una capa de entrada, una o varias capas ocultas y una capa de salida. Los parámetros que se ajustaron son:

- Número de capas ocultas, una mayor cantidad de capas ocultas puede permitir al modelo capturar representaciones complejas de los datos, pero también puede aumentar la complejidad y el tiempo de entrenamiento.

- Número de neuronas por capa, el número de neuronas en las capas ocultas puede influir en la capacidad del modelo para aprender representaciones más detalladas y complejas de los datos.

- Función de activación, especifica la función que se utiliza para activar las neuronas en cada capa, incluyendo las capas ocultas y la capa de salida. Algunas funciones de activación comunes son la función sigmoide, la función ReLU (*Rectified Linear Unit*) y la función tangente hiperbólica.

- Tasa de aprendizaje, controla la velocidad a la que se actualizan los pesos de las conexiones durante el entrenamiento del modelo. Una tasa de aprendizaje demasiado alta puede hacer que el modelo se ajuste demasiado rápido y no converja, mientras que un valor de tasa de aprendizaje demasiado baja puede hacer que el modelo tarde mucho en converger.

Por lo tanto, después de realizar una búsqueda exhaustiva mediante experimentación, se utilizan los siguientes parámetros:

- Capas ocultas: 3
- Neuronas por capa: 100
- Función de activación: Sigmoide
- Tasa de aprendizaje: 0.1

Árboles de decisión [138]

Es un algoritmo de clasificación y regresión que se basa en la creación de una estructura de árbol donde cada nodo representa una característica o atributo y cada rama representa una regla de decisión. Los parámetros que se ajustaron son:

- Criterio de división, determina cómo se evalúa la calidad de una división en el árbol. Algunos criterios comunes incluyen la ganancia de información (*entropy*) y la impureza de Gini.

- Profundidad máxima, especifica la profundidad máxima del árbol, lo que limita la cantidad de divisiones que se pueden hacer.

- Número mínimo de muestras en una hoja, establece el número mínimo de muestras requeridas en una hoja del árbol. Si una división resulta en un número menor de muestras en una hoja, la división no se realizará.

- Máximo número de características consideradas especifica el número máximo de características que se considerarán al buscar la mejor división en cada nodo. Esto ayuda a controlar la dimensionalidad y la complejidad del árbol.

Por lo tanto, después de realizar una búsqueda exhaustiva mediante experimentación, se utilizan los siguientes parámetros:

- Criterio de división: Debido a que se utiliza el módulo de árboles de decisión de la librería OpenCV no es necesario establecer un parámetro para especificar el criterio de división, ya que la librería utiliza por defecto el índice Gini como criterio de división.

- Profundidad máxima: INT_MAX (No existe límite máximo de profundidad establecido), lo que significa que el árbol puede crecer hasta su máxima capacidad sin restricciones en términos de profundidad.

- Mínimo de muestras en una hoja: 5

- Máximo de características consideradas:10

3.6 Métricas

Las métricas desempeñan un papel fundamental en los sistemas de detección de suplantación, ya que proporcionan una evaluación cuantitativa del rendimiento del sistema. Las métricas permiten medir la efectividad de los algoritmos de detección y comparar los resultados con diferentes enfoques. Algunas de las métricas más comunes son:

Precisión [140]
Mide la proporción de casos correctamente identificados como genuinos o fraudulentos. Un porcentaje alto indica que hay una baja tasa de falsos positivos y falsos negativos.

$$Precisión = \frac{VP}{VP + FP} \quad 21$$

***Recall* (Recuperación)** [140]
También conocido como sensibilidad o tasa de verdaderos positivos, mide la proporción de casos positivos que son correctamente identificados. Un valor alto indica que se detectan la mayoría de los casos positivos.

$$Recall = \frac{VP}{VP + FN} \quad 22$$

Exactitud [140]
Es la proporción de casos correctamente clasificados en relación con el total de casos. Es una medida global de la precisión del sistema.

$$Exactitud = \frac{VP + VN}{VP + VN + FN + FP} \quad 23$$

F1-Score [140]
Es una medida que combina la precisión y la recuperación en un solo valor. Es útil cuando hay un desequilibrio entre las clases de imágenes reales y falsificaciones.

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad 24$$

Tasa de error [140]
Mide la proporción de casos clasificados incorrectamente en relación con el total de casos

$$Tasa\ de\ error = \frac{FP + FN}{VP + VN + FN + FP} \quad 25$$

Donde:

- **VP** es la cantidad de positivos que fueron clasificados correctamente como positivos por el modelo.
- **VN** es la cantidad de negativos que fueron clasificados correctamente como negativos por el modelo.
- **FN** es la cantidad de positivos que fueron clasificados incorrectamente como negativos.

- **FP** es la cantidad de negativos que fueron clasificados incorrectamente como positivos.

Las métricas presentadas permiten evaluar el rendimiento del sistema; sin embargo, un sistema de detección de suplantación de identidad tiene que lidiar con dos tipos de eventos:

- La persona que reclama una identidad dada es la que dice ser (en cuyo caso, la imagen corresponde a una persona real).

- No lo es (en cuyo caso, es un impostor).

Por lo tanto, una parte importante dentro de los sistemas PAD es verificar que tan eficientes son los sistemas de detección de suplantación facial, para ello se utilizan tres métricas propuestas en el estándar ISO / IEC 30107-3 [25], que evalúan el rendimiento en términos de detección de ataques de presentación, las cuales se describen brevemente a continuación.

APCER [141]

La tasa de error de clasificación de presentación de ataques (*Attack Presentation Classification Error Rate*), mide la tasa de error de clasificación para los ataques de presentación (ataques de suplantación facial). Es la proporción de ataques de presentación clasificados incorrectamente como casos genuinos. Un porcentaje bajo de APCER indica una buena capacidad del sistema para detectar y clasificar correctamente los ataques de presentación. Se calcula de la siguiente manera:

$$APCER = \frac{1}{N_{PAIS}} \sum_{i=1}^{N_{PAIS}} (1 - RES_i) \qquad 26$$

Donde N_{PAIS} es el número de imágenes detectadas como ataques, RES_i toma el valor de 1 si el valor de clasificación marca la imagen como falsa y valor de 0 si la clasifica como real.

BPCER [141]

La tasa de error de clasificación de presentación de buena fe (*Bona fide Presentation Classification Error Rate*), mide la tasa de error de clasificación para las presentaciones genuinas. Es la proporción de presentaciones genuinas clasificadas incorrectamente como ataques de presentación. Un porcentaje bajo de BPCER indica una buena capacidad del sistema para distinguir entre presentaciones genuinas y ataques de suplantación. Se calcula de la siguiente manera:

$$BPCER = \frac{\sum_{i=1}^{N_{BF}} RES_i}{N_{BF}} \qquad 27$$

Donde N_{BF} es el número de imágenes detectadas como reales mientras RES_i toma el valor de 1 si el valor de clasificación marca la imagen como falsa y valor de 0 si la clasifica como real.

HTER [142]

La tasa de error total medio (*Half Total Error Rate*), es el promedio de APCER y BPCER. Representa la tasa de error total del sistema, considerando tanto las presentaciones genuinas como los ataques de presentación. Un porcentaje bajo de HTER indica un buen equilibrio entre la detección precisa de ataques y la aceptación de imágenes reales. Se calcula de la siguiente manera:

$$HTER = \frac{APCER + BPCER}{2} \qquad 28$$

De acuerdo con las características presentadas en cada métrica se determinó el uso de cuatro que son, F1-score, APCER, BPCER y HTER debido a su capacidad para evaluar de manera completa y precisa el rendimiento de los sistemas. El F1-score permite considerar tanto la precisión como la recuperación, lo que es especialmente relevante en escenarios de desequilibrio de clases. Por otro lado, el APCER y el BPCER se centran específicamente en la detección de ataques y la clasificación de presentaciones genuinas, respectivamente, lo cual es crucial en el contexto de la suplantación facial. Además, el HTER ofrece una visión general del rendimiento del sistema, considerando tanto los errores de clasificación de ataques como los de presentaciones genuinas. Al utilizar las cuatro métricas se proporciona una evaluación integral del rendimiento de los sistemas de detección de suplantación facial y permiten comparar diferentes enfoques y algoritmos, facilitando así la toma de decisiones.

CAPÍTULO 4
Método de solución

En el presente capítulo se describe la evolución del método planteado para el reconocimiento de una posible suplantación de identidad por medio del rostro sin información previa del rostro a analizar.

En el transcurso de la investigación, se ha llevado a cabo una amplia experimentación con diversos métodos de solución basados en la evolución actual de los algoritmos de suplantación facial. En este contexto, se han explorado y aplicado diferentes enfoques y técnicas para aprovechar los avances más recientes en el campo de la suplantación facial y aplicarlos en imágenes de entornos no controlados. Los algoritmos de solución utilizados se basan en la evolución y adaptación de técnicas existentes, con el propósito de superar o detectar limitaciones en la detección de ataques de suplantación facial.

Los resultados obtenidos a través de los diferentes métodos de solución evaluados durante la investigación han proporcionado información valiosa sobre las fortalezas y limitaciones de las técnicas actuales, permitiendo identificar áreas de mejora y posibles direcciones futuras de investigación. A continuación, se presenta una descripción de la evolución del método para la detección de suplantación facial propuesto en el presente trabajo.

4.1 Evolución del método propuesto

A lo largo del tiempo, se han llevado a cabo investigaciones exhaustivas con el objetivo de detectar de manera efectiva la manipulación de rostros en imágenes y videos. En este contexto, el uso del LBP (*Local Binary Patterns*) con radio 1 y 8 vecinos ha surgido como una opción prometedora que proporciona una base sólida para explorar y comprender la detección de suplantación facial. Este enfoque se eligió como punto de partida en la investigación por varias razones.

En primer lugar, $LBP_{8,1}$ se destaca por su simplicidad y eficiencia computacional. Su implementación y ejecución no requieren recursos computacionales significativos, lo que lo hace adecuado para su aplicación en sistemas en tiempo real o con limitaciones de recursos. Esta característica es especialmente relevante en el campo de la detección de suplantación facial, donde la eficiencia y velocidad son cruciales para procesar grandes volúmenes de datos de manera oportuna.

Asimismo, $LBP_{8,1}$ funciona como una línea de base sólida en comparación con técnicas avanzadas y sofisticadas. Al establecer este punto de referencia inicial, es posible evaluar la efectividad y el rendimiento de enfoques más complejos en relación con LBP.

En donde a la imagen completa de entrada se le aplica el algoritmo $LBP_{8,1}$. Lo que se busca es considerar el contexto visual global en el que aparece el rostro. Además, con ello, se tienen en cuenta detalles adicionales relevantes, como el cabello, el fondo y la iluminación. Para el vector de características se calcula el histograma de la imagen, el cual es utilizado en la clasificación con las maquinas se soporte de vectores (la especificación de los parámetros esta descrita en el capítulo 3).

En la Figura 4.1 se muestra el diagrama del método inicial de experimentación (versión 1), en donde, se emplea el algoritmo $LBP_{8,1}$ (es decir P=8 y R=1, en donde P es el número de vecinos y R el radio) para analizar la imagen completa de entrada. El objetivo principal es considerar el contexto visual global en el que el rostro se encuentra, tomando en cuenta no solo las características del rostro en sí, sino también detalles adicionales como el cabello, el fondo y la iluminación. Una vez obtenido el vector de características mediante el cálculo del histograma de la imagen, se pasa a la etapa de clasificación, en donde se emplean máquinas de soporte de vectores (los parámetros utilizados están descritos en el capítulo 3 de la investigación) para determinar si una imagen es real o no.

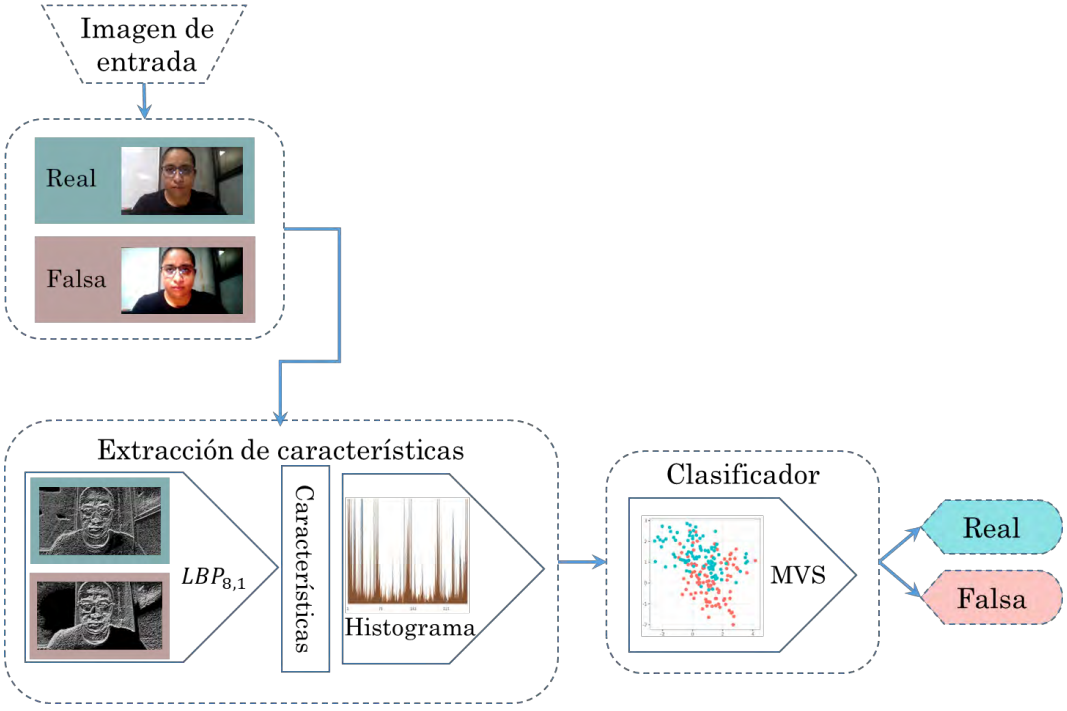

Figura 4.1 Versión 1 del método experimental.

La inclusión de técnicas de color también permite adaptarse a diversos escenarios y condiciones de iluminación, lo que resulta especialmente útil en situaciones de entornos complicados. Por lo que la incorporación de técnicas de color, específicamente utilizando espacios como YCbCr y HSV, en el método basado en LBP proporcionan información adicional sobre la distribución de colores y tonalidades en el rostro, lo que puede ser importante para distinguir entre rostros auténticos y falsos. En la Figura 4.2 se muestra con verde la incorporación de los espacios de color previo a la extracción de características (versión 2).

Con la incorporación de los canales de color se busca aprovechar las características que aportan los canales de color YCbCr y el canal HSV, los cuales se utilizan de manera secuencial para capturar la información de color en la imagen. Primero, se convierte la imagen de entrada al espacio de color YCbCr para resaltar detalles finos, a continuación, se realiza una conversión adicional al espacio de color HSV para resaltar características como variaciones en el tono de piel o regiones con colores anómalos que puedan indicar manipulación.

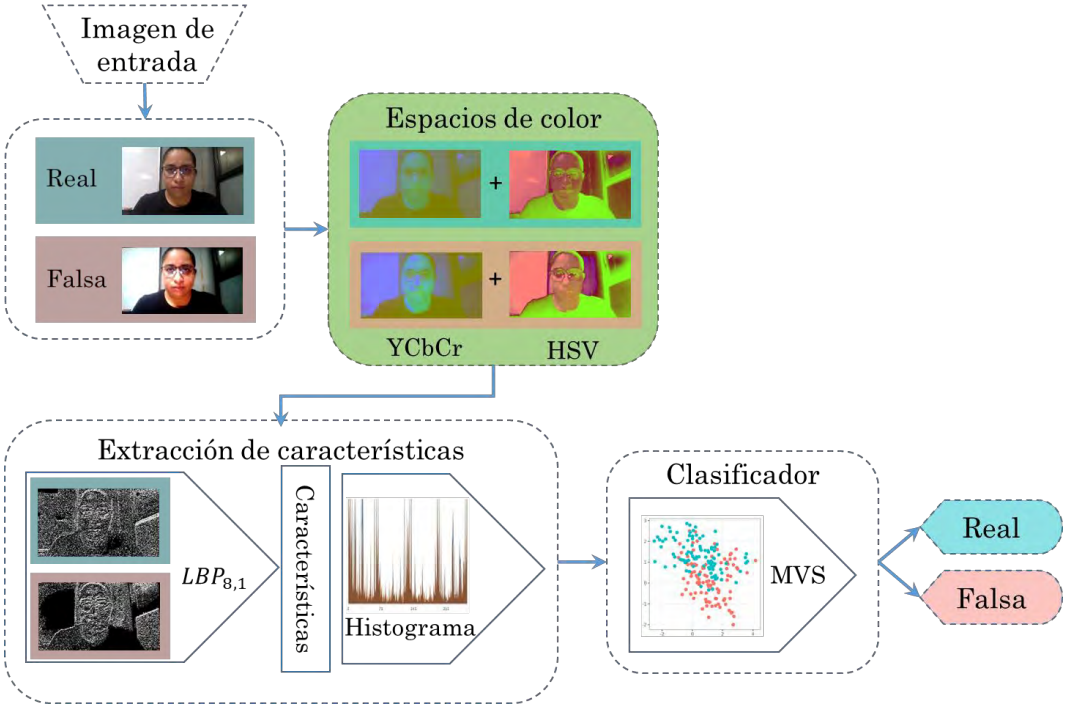

Figura 4.2 Versión 2 del método experimental.

Al continuar con la experimentación de las técnicas utilizadas para la detección de suplantación facial, se decidió ampliar el enfoque cambiando de la técnica LBP a su variante denominada ELBP (*Extended Local Binary Patterns*). Con $ELBP_{16,2}$ se busca mejorar la capacidad de discriminación y robustez al considerar regiones más amplias alrededor de cada punto de interés facial.

La variante permite capturar información contextual adicional y proporcionar una descripción más completa de las texturas presentes en la imagen, lo que contribuye a mejorar la precisión en la detección de rostros genuinos y falsificados. En la Figura 4.3 se observa que el proceso del método sigue siendo igual con la diferencia del cambio de LBP por ELBP el cual se encuentra resaltado en verde.

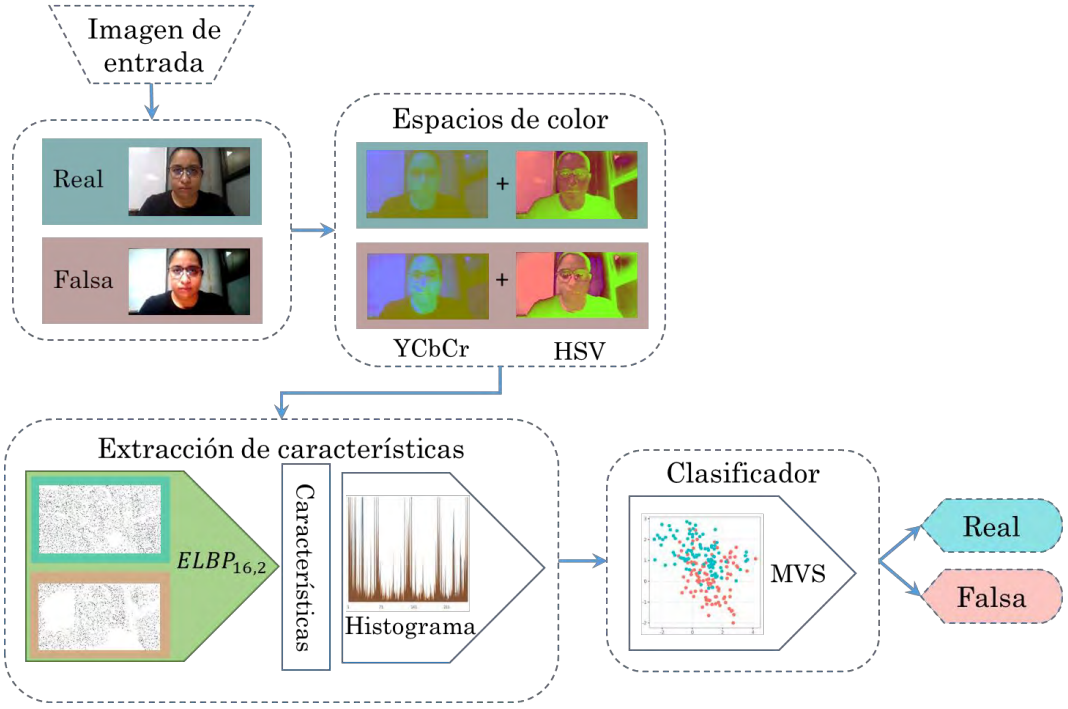

Figura 4.3 Versión 3 del método experimental.

En la evolución de la experimentación, se introdujo el uso del algoritmo de Retinex multi-escala para mejorar la iluminación de la imagen de entrada y optimizar la localización del rostro. Una vez seleccionado y recortado el rostro, se procedió a redimensionarlo a un tamaño de 200 x 200 píxeles para estandarizar la escala de las imágenes. El algoritmo de Retinex permite realizar ajustes adaptativos en la iluminación, realzando los detalles y mejorando la visibilidad de la zona facial. La mejora en la iluminación del rostro se puede apreciar visualmente a través de un aumento en la claridad y el contraste de la imagen.

Además, se realizó un cambio en el descriptor de textura utilizado, reemplazando ELBP por BSIF (*Binarized statistical image features*). BSIF es un enfoque que se basa en el cálculo de características de textura utilizando filtros binarios. Con su uso se buscó proporcionar una descripción más eficiente y robusta de las texturas presentes en la imagen del rostro.

Asimismo, se realizó una selección específica del área de la frente como zona a analizar. La elección de esta región se basa en la premisa de que la frente es una parte del rostro que tiende a ser menos modificada o afectada en casos de suplantación facial. Al enfocarse en esa área, se buscó obtener una mejor representación de los rasgos faciales genuinos y detectar de manera más precisa cualquier intento de suplantación. En la Figura 4.4 se muestra con verde los pasos agregados al método de experimentación, así como el cambio del algoritmo ELBP a BSIF.

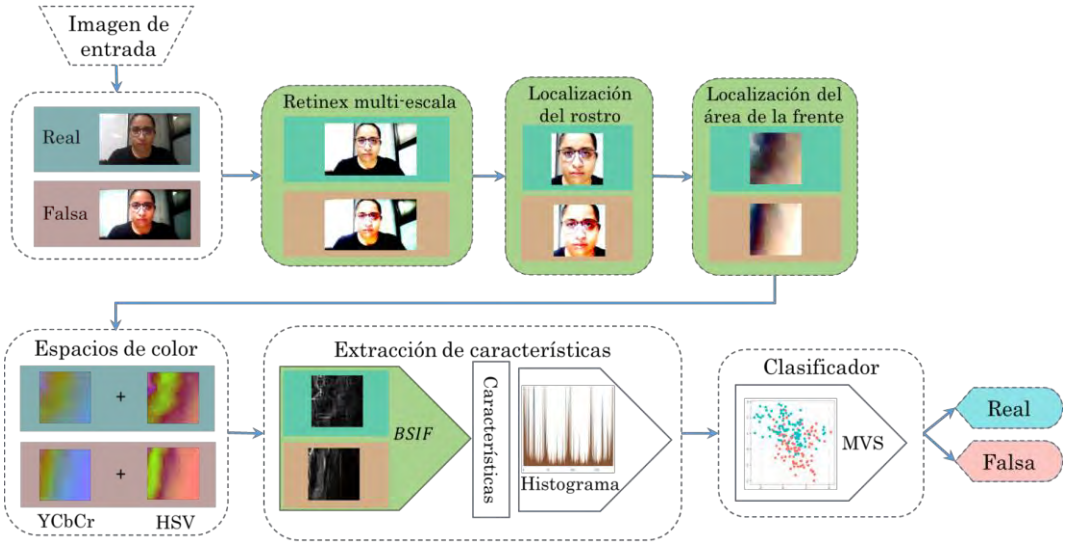

Figura 4.4 Versión 4 del método experimental.

Además de experimentar con técnicas de visión tradicionales, se realizó una exploración en el campo del aprendizaje profundo para abordar el problema de la detección de suplantación facial. La red utilizada en este contexto fue Feathernet-B [123]. La cual es una red neuronal profunda diseñada específicamente para la detección de suplantación facial. Fue desarrollada con el objetivo de evitar la dependencia de hardware especializado, lo cual es una ventaja significativa en términos de accesibilidad y costos.

La elección de utilizar aprendizaje profundo, y en particular la red Feathernet-B, se basa en la capacidad de las redes para aprender y representar características complejas de las imágenes de manera automática. En la Figura 4.5 se observa que, para la experimentación con aprendizaje profundo, se eliminó el uso del algoritmo Retinex multi-escala, la localización de la zona de la frente y los algoritmos de textura para la caracterización. En cambio, se destaca la adición de la red FeatherNet-B.

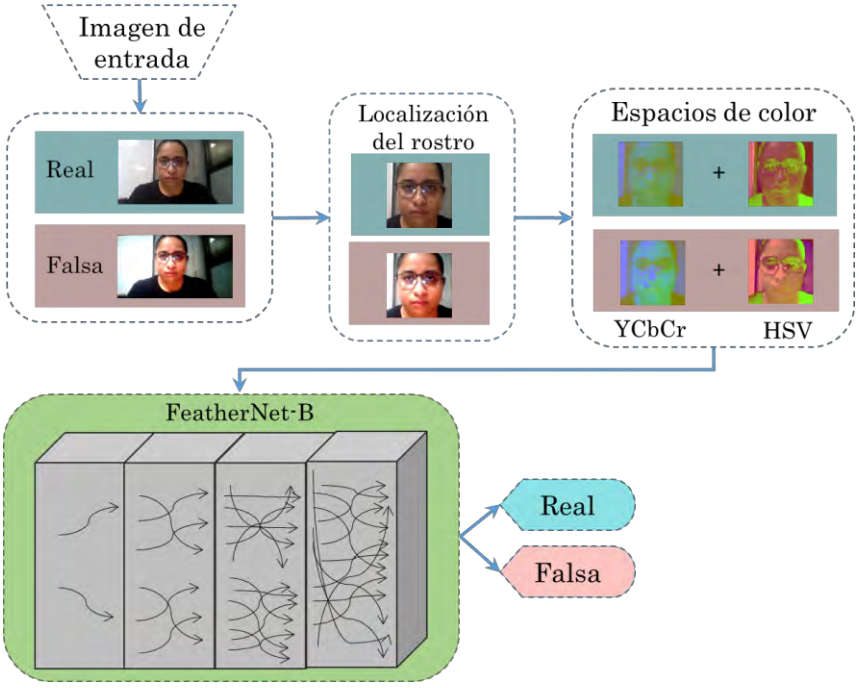

Figura 4.5 Versión 5 del método experimental.

Para explorar los posibles beneficios de utilizar modelos de color antes de introducir una imagen en la red neuronal, se tomó la decisión de eliminarlos en esta etapa y mantener únicamente la localización del rostro como información de entrada. La eliminación de los modelos de color implica trabajar directamente con la imagen en escala de grises o en un formato de color estándar, en lugar de utilizar representaciones de color, con el propósito de simplificar el proceso y evaluar si la información de localización del rostro por si sola es suficiente para lograr resultados satisfactorios. En la Figura 4.6 se muestra el proceso del método con el rostro como entrada de la red neuronal FeatherNet-B.

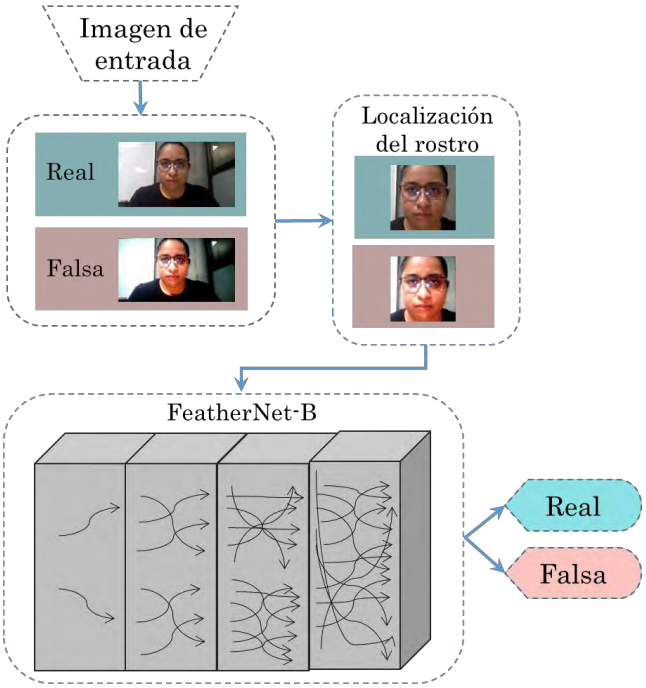

Figura 4.6 Versión 6 del método experimental.

Con el objetivo de comparar y evaluar si la combinación de enfoques de visión tradicional y aprendizaje profundo puede mejorar la detección de suplantación facial, se implementó un método de votación. En el cual se utilizaron tres enfoques para obtener una decisión conjunta.

El primer enfoque utilizado es un método de visión tradicional que emplea el rostro, los canales de color y $ELBP_{16,2}$. Los componentes se utilizan en conjunto para extraer características relevantes de la imagen y detectar posibles suplantaciones faciales. Además, se incorporó la red neuronal Feathernet-B, que es un modelo de aprendizaje profundo específicamente diseñado para la detección de suplantación facial. Adicionalmente, se introdujo la detección de movimiento como parte del sistema de votación. Utilizando la lectura de un video, se realiza la localización del rostro mediante el análisis de las distancias faciales. El enfoque permite detectar cambios o movimientos inusuales en la posición o estructura del rostro, lo cual puede indicar la presencia de una suplantación.

El sistema de votación mostrado en la Figura 4.7, combina las salidas del enfoque de visión tradicional, aprendizaje profundo y detección de movimiento para llegar a una decisión conjunta sobre la autenticidad de un rostro. Al utilizar diferentes enfoques y combinar sus resultados, se busca mejorar la precisión y robustez de la detección de suplantación facial.

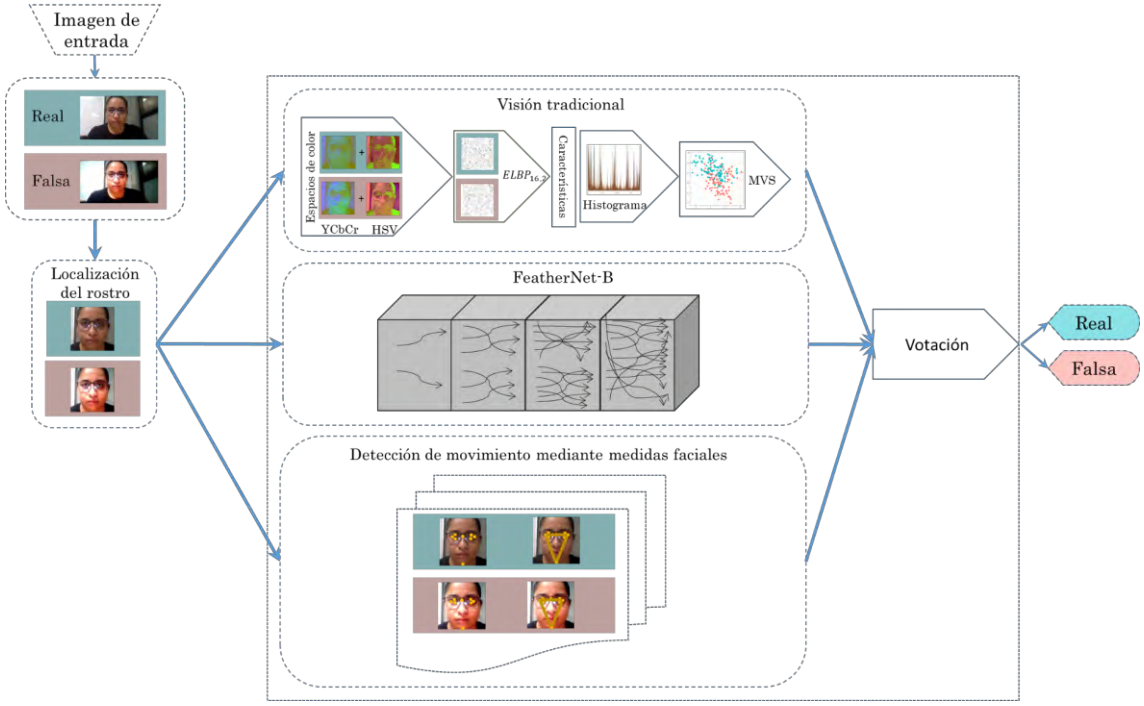

Figura 4.7 Versión 7 del método experimental.

Con base en las experimentaciones llevadas a cabo durante la evolución del método, se optó por utilizar algoritmos de visión tradicional. En consecuencia, se propone un método final que combina diferentes componentes para lograr una detección robusta de suplantación facial, el diagrama se muestra en la Figura 4.8.

En primer lugar, se utiliza el algoritmo de Retinex multi-escala para mejorar la iluminación de la imagen del rostro. El algoritmo permite realizar ajustes adaptativos en la iluminación, lo que resulta en una imagen con mayor claridad y contraste, de forma visual. Una vez realizada la mejora de la iluminación, se procede a la selección del rostro. Posteriormente, se emplean los modelos de color YCbCr y HSV, la imagen resultante se procesa con el algoritmo PRNU para la extracción de características. Los modelos de color permiten capturar información valiosa sobre la distribución de los componentes de color en la imagen, mientras que el algoritmo PRNU se utiliza para identificar patrones únicos e intransferibles presentes en los sensores de la cámara.

Por último, los resultados obtenidos se evalúan mediante la utilización de cuatro algoritmos de clasificación. Los cuales se aplican al vector de características extraído anteriormente y se comparan para determinar cuál de ellos proporciona una mayor robustez en la detección de suplantación facial.

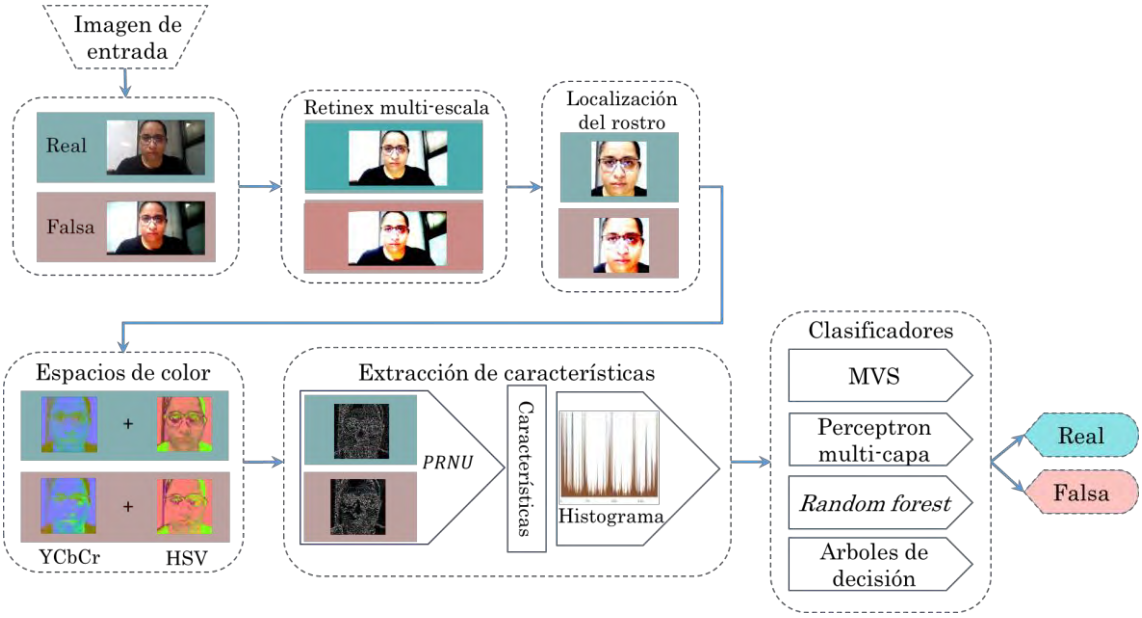

Figura 4.8 Método de detección de suplantación facial propuesto.

4.2 Método de detección de suplantación facial

La finalidad del método propuesto es abordar la detección de posibles ataques de suplantación de identidad en imágenes faciales, independientemente de factores como la calidad de iluminación, resolución o tamaño de la imagen. El enfoque se diseñó con el objetivo de ser robusto y capaz de identificar suplantaciones incluso en condiciones desafiantes. En la Figura 4.9 se muestra el diagrama del proceso del método considerando el entrenamiento.

Una característica destacada del método es que se busca lograr un entrenamiento completamente independiente respecto a la identidad o características específicas del rostro que se analiza. Lo que significa que el modelo se entrena de manera generalizada, sin estar específicamente adaptado a un individuo o conjunto de individuos en particular. Lo que proporciona una mayor flexibilidad y capacidad para detectar suplantaciones en rostros desconocidos.

El método se ha desarrollado teniendo en cuenta la necesidad de abordar escenarios de suplantación de identidad de manera amplia y sin restricciones. Se busca superar las limitaciones asociadas con la iluminación, la resolución de la imagen y el tamaño de la misma, con el fin de brindar una solución robusta y eficaz.

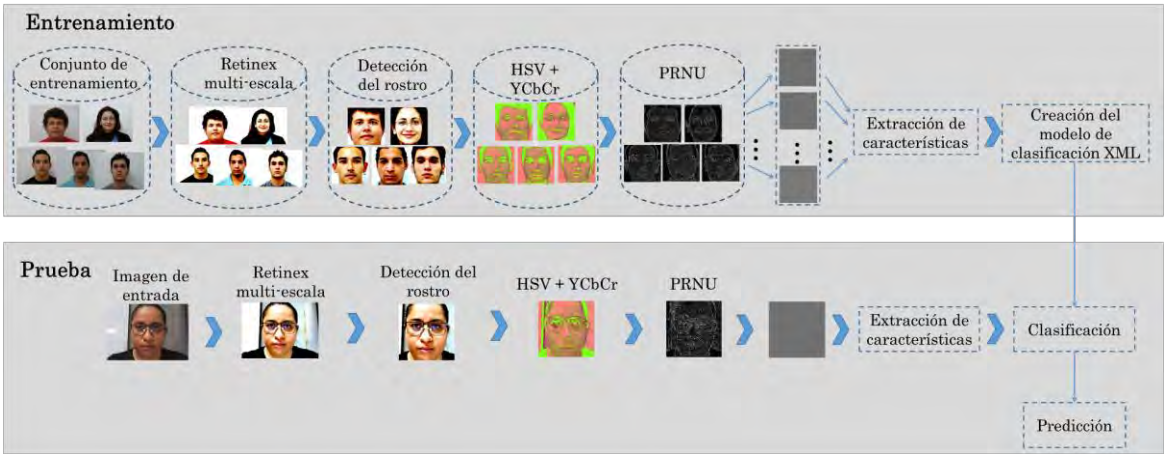

Figura 4.9 Diagrama del método de detección de suplantación.

El proceso de detección de suplantación facial se divide en dos partes: el entrenamiento y la etapa de pruebas. Ambas partes desempeñan un papel importante en el funcionamiento y la efectividad del método.

Por un lado, el entrenamiento es una fase de vital importancia, ya que en esta etapa el modelo de detección de suplantación facial se construye y se ajusta a partir de un conjunto de datos de entrenamiento. Durante el entrenamiento, el modelo aprende a reconocer patrones distintivos entre imágenes faciales genuinas y falsificadas. Por otro lado, la etapa de pruebas se refiere al momento en el que una imagen se introduce en el proceso para ser validada como real o falsa. Durante la fase de prueba, el modelo entrenado se utiliza para analizar y evaluar la autenticidad de la imagen facial. El modelo aplica las características y conocimientos aprendidos durante el entrenamiento para realizar una predicción y determinar si la imagen es genuina o si muestra signos de suplantación facial.

El proceso de entrenamiento utilizado en el método de detección de suplantación facial se realiza utilizando un banco de imágenes creado específicamente para este propósito. Los detalles y características del banco de imágenes se describen en el capítulo 5 del documento. Una vez que se disponen de las imágenes en el banco de entrenamiento, se inicia el proceso de entrenamiento del modelo. Asimismo, la etapa de prueba se inicia una vez que el modelo ha sido entrenado y se introduce la imagen que se desea analizar.

Es importante destacar que la imagen de entrada no está condicionada a tener una iluminación óptima, una alta resolución o que la persona se encuentre perfectamente orientada hacia la cámara. Sin embargo, se asume que la imagen contendrá al menos un rostro.

Es importante destacar que la clasificación de una imagen en cuanto a su iluminación se realiza mediante un procedimiento de análisis visual. Se evalúa la calidad de iluminación en base a la apreciación subjetiva de la imagen. No se realiza una medición objetiva específica ni se establecen rangos predefinidos para clasificar la iluminación como buena, media o mala. Se realiza una observación visual de las características de iluminación presentes en la imagen y se realiza una clasificación subjetiva en base a la percepción del evaluador. Lo que permite identificar distintos niveles de iluminación. La Figura 4.10 ilustra un ejemplo que muestra tres posibles situaciones de iluminación al ingresar al sistema, pero no establece una métrica o rango específico para la clasificación de la iluminación.

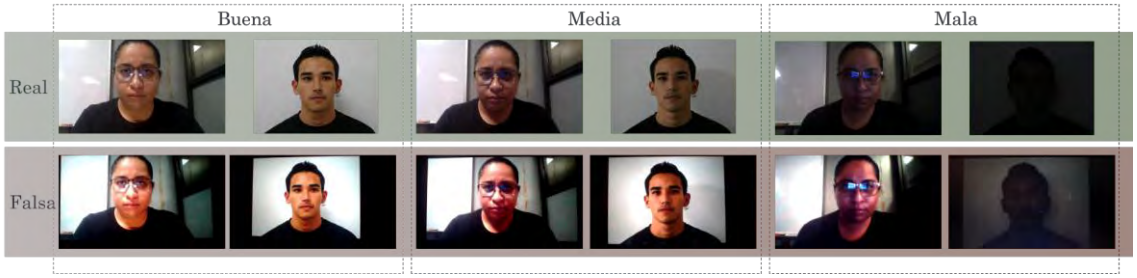

Figura 4.10 Muestra de diferentes situaciones de iluminación.

Una vez iniciadas cada una de las etapas, tanto el entrenamiento como las pruebas por separado, se aplica el algoritmo de retinex multi-escala a cada imagen con el objetivo de mejorar la iluminación y resaltar las características de color presentes en ellas. La mejora en la iluminación se determina visualmente al observar los cambios en la claridad y el contraste de la imagen. El algoritmo de retinex multi-escala ajusta adaptativamente la iluminación de la imagen, lo que resulta en una imagen con una iluminación más equilibrada y realce de detalles. Al mejorar la iluminación, se pueden apreciar mejor los rasgos faciales y las características relevantes en la imagen. En cuanto al resaltado de las características de color, el algoritmo de retinex multi-escala también contribuye a este aspecto. Al ajustar la iluminación de manera adaptativa, se resaltan los detalles y las variaciones de color presentes en la imagen facial. Lo que permite una mejor visualización y análisis de las características de color, como tonalidades, saturación y texturas presentes en el rostro.

En la Figura 4.11 se presentan las imágenes que se muestran en la Figura 4.10 después de aplicar el procesamiento de iluminación utilizando el algoritmo de retinex multi-escala. Para determinar si la iluminación ha sido visualmente corregida, se realiza una comparación visual entre las imágenes originales y las imágenes procesadas. Se busca observar cambios en la claridad, el contraste y la visibilidad de los detalles faciales.

Figura 4.11 Resultado de aplicar retinex multi-escala a las imágenes de entrada.

A continuación, se localiza el rostro en cada imagen mediante el algoritmo Haar descrito en el capítulo 3 y se utiliza la técnica de OpenCV *resize()*, de redimensionamiento para ajustar el área del rostro a un tamaño estándar de 500 x 500 píxeles (ver Figura 4.12). Lo que asegura que todas las imágenes utilizadas en el entrenamiento y pruebas tengan la misma escala y tamaño, lo que facilita el proceso de extracción y comparación de características. Antes del redimensionamiento, el tamaño del área del rostro puede variar en función de la imagen original y las características específicas del rostro detectado.

Figura 4.12 Muestra de imágenes con selección del rostro.

Cuando las condiciones de iluminación extremas y el rostro no se detecta, se considera que la imagen es falsa al no contener un rostro. En la Figura 4.13 se muestran las imágenes originales con condiciones de iluminación extremas tanto en la imagen real como en la falsa. En este contexto, se considera iluminación oscura como aquella que impide ver correctamente a la persona en la imagen. Como resultado, la imagen resultante después de aplicar el algoritmo de retinex multi-escala no contiene la información suficiente que el módulo de detección del rostro necesita para determinar la presencia de un rostro en la imagen.

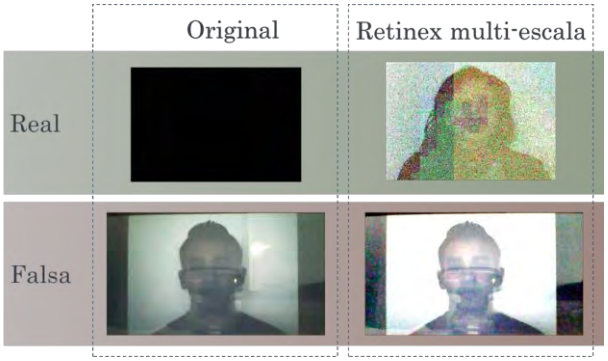

Figura 4.13 Imágenes con iluminación oscura extrema.

Después de recortar el rostro, se realiza una conversión a dos canales de color, el espacio de color YCbCr y el espacio de color HSV, con el objetivo de resaltar las características de brillo y saturación presentes en la imagen. La saturación es una propiedad del color que se refiere a la intensidad o pureza de un color específico en una imagen. Indica cuánto está presente el color puro en relación con el blanco o el gris neutro.

Para resaltar las características de saturación, se utiliza el canal de crominancia (CbCr en YCbCr) del espacio de color YCbCr. Al utilizarlo, se pueden identificar variaciones en la intensidad y la pureza del color, lo que ayuda a resaltar las características de color presentes en el rostro. Al realizar la conversión al canal de color YCbCr (ver Figura 4.14), las tonalidades en las imágenes del rostro pueden proporcionar información importante, el componente Y representa la luminancia (brillo) de la imagen, mientras que los componentes Cb y Cr representan las diferencias de color en las direcciones azul-amarillo y rojo-verde, respectivamente.

En el caso del canal Cb, las tonalidades azules y amarillas pueden destacar ciertos detalles y características de la piel. Por ejemplo, áreas más oscuras o con tonalidades azuladas pueden indicar sombras o imperfecciones en la piel. Por otro lado, áreas más claras o con tonalidades amarillentas pueden resaltar zonas con mayor brillo o saturación. En cuanto al canal Cr, las tonalidades rojas y verdes pueden revelar detalles relacionados con los tonos de piel y la presencia de enrojecimiento o alteraciones de color. Por ejemplo, tonos rojizos pueden indicar áreas con mayor enrojecimiento o enfoque en los labios, mientras que tonos verdes pueden resaltar ciertas características de la piel.

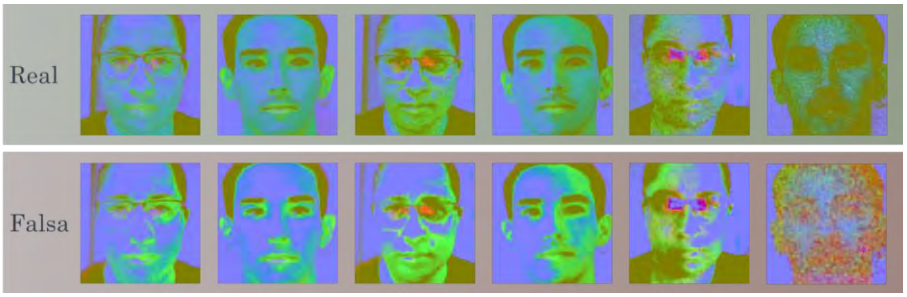

Figura 4.14 Muestra de imágenes después de la conversión al canal de color YCbCr.

Por otro lado, para resaltar las características de brillo, se utiliza el canal de luminosidad (V en HSV) del espacio de color HSV. El canal aísla la información de brillo en la imagen y permite resaltar las variaciones de intensidad presentes en el rostro. Al realizar la conversión al canal, se pueden identificar diferencias significativas en la iluminación del rostro (ver Figura 4.15).

El componente de tonalidad (H) indica el tipo de color, como rojo, verde, azul, etc. Se representa en forma de un ángulo circular, donde diferentes valores de tonalidad corresponden a diferentes colores. El componente de saturación (S) indica la pureza o vivacidad del color. Valores más altos de saturación representan colores más intensos y vibrantes, mientras que valores más bajos de saturación resultan en colores más desaturados o menos intensos.

Al aplicar el canal de color HSV al rostro en una imagen, las tonalidades presentes en diferentes áreas pueden proporcionar información sobre los colores predominantes en esa región. Por ejemplo, tonalidades más rojizas pueden indicar la presencia de piel, tonalidades más verdes pueden indicar la presencia de sombras.

Figura 4.15 Modelo de color HSV aplicado a las imágenes del rostro.

Como resultado la conversión del canal de color YCbCr seguida de la conversión al canal de color HSV en el rostro puede ayudar a resaltar características relevantes, detectar manipulaciones en el color y garantizar una mayor robustez ante variaciones de iluminación, lo cual contribuye a mejorar la detección de suplantación facial (ver Figura 4.16).

Figura 4.16 Muestra de imágenes resultantes de la conversión del canal de color YCbCr seguido de la conversión al canal de color HSV en un rostro.

Para abordar las características mediante textura, se aplica el algoritmo de PRNU a la imagen, como se muestra en la Figura 4.17. El algoritmo se utiliza para resaltar el ruido generado por la cámara en la imagen facial. PRNU aprovecha las propiedades únicas del ruido inherente a cada cámara, el cual es producido debido a las variaciones en la sensibilidad de los píxeles y otros factores relacionados con el sensor de imagen. Al aplicar el algoritmo de PRNU, se extrae y resalta el ruido generado por la cámara, se obtiene una representación única e intransferible de la imagen facial.

Figura 4.17 Muestra de imágenes con PRNU.

Cuando se obtiene la imagen de PRNU, que se encuentra en formato de datos de punto flotante, se realiza la conversión a un formato de datos de 8 bits (uint8) para obtener el histograma de 256 valores, el cual conforma el vector de características. El formato original de la imagen de PRNU suele ser de punto flotante, lo que permite valores de píxeles con decimales y una mayor precisión. No obstante, con el propósito de generar el histograma de 256 valores, se requiere convertir la imagen al formato de datos de 8 bits (uint8). La conversión implica redondear y escalar los valores de los píxeles, asegurándose de que se ajusten al rango permitido en el formato uint8 (ver Figura 4.18).

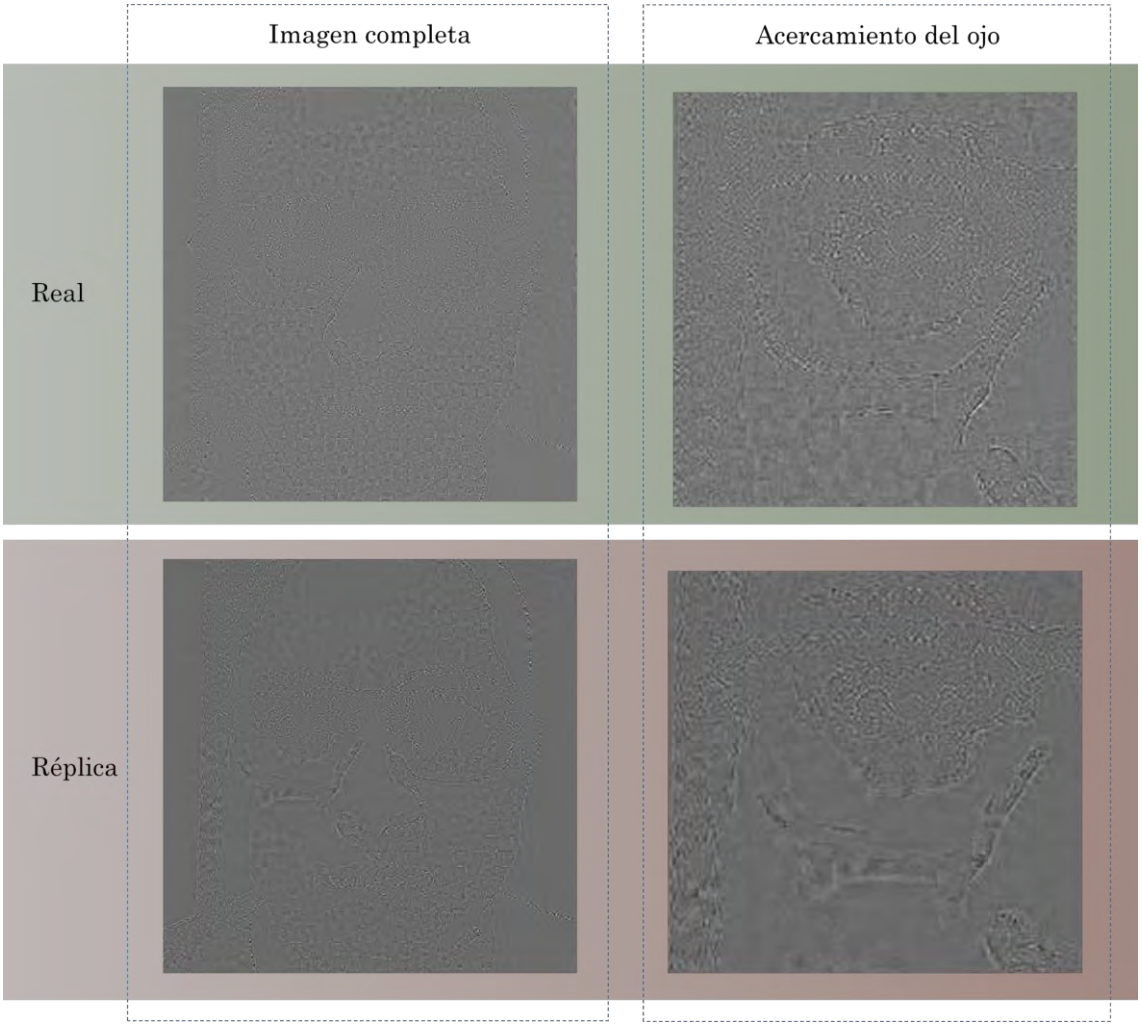

Figura 4.18 Muestra de imágenes con textura resultado de aplicar PRNU.

Una vez que se ha realizado la conversión, se puede generar el histograma de la imagen, que cuenta la frecuencia de ocurrencia de cada valor de intensidad en el rango de 0 a 255. El histograma se utiliza como el vector de características que captura información relevante de la imagen de PRNU y se utiliza en etapas posteriores del proceso de detección de suplantación facial.

Finalmente, se lleva a cabo el entrenamiento del clasificador utilizando el vector de características y las etiquetas correspondientes que indican si la imagen es genuina o si es un posible ataque de suplantación. El resultado de este entrenamiento es un modelo clasificador en formato XML, que se utilizará para la clasificación de imágenes desconocidas y la detección de posibles ataques de suplantación durante la etapa de pruebas.

CAPÍTULO 5
Pruebas y Resultados

El presente capítulo describe las pruebas realizadas para evaluar el método descrito en el capítulo 4 sección 2, con cuatro clasificadores que son: máquinas de vector soporte (MVS), perceptrón multicapa, bosque aleatorio y árboles de decisión. Los resultados son evaluados mediante las métricas de: APCER, BPCER, HTER y F1-score.

El objetivo de realizar las pruebas es evaluar el rendimiento y la efectividad del método propuesto. Lo que permite analizar el comportamiento del sistema en diferentes situaciones, determinar su precisión y capacidad para distinguir entre imágenes auténticas y falsificadas. Además, las pruebas ayudan a identificar posibles limitaciones y mejorar la calidad del método, asegurando que sea confiable y eficaz en la detección de suplantaciones faciales en escenarios del mundo real.

5.1 Herramientas utilizadas (software y equipo)

El objetivo de las herramientas seleccionadas para la investigación es maximizar el rendimiento y la eficiencia del método a pesar de las limitaciones de recursos.

En cuanto al software, se utilizan algoritmos y técnicas de procesamiento de imágenes para ejecutarse en entornos con recursos limitados, sin comprometer la precisión y la confiabilidad de la detección. Las herramientas de software utilizadas son:

- C++ en Visual Studio 2015 [143].
- OpenCV 3.4.0 [144].
- Librería Dlib C++ [145].

En cuanto al hardware, aunque puede haber limitaciones en términos de capacidad de procesamiento y memoria, Se seleccionaron componentes que se ajustan a los requisitos mínimos del sistema y que son capaces de ejecutar las tareas de detección en tiempo real. El hardware empleado para el desarrollo y pruebas del método cuenta con las siguientes características:

- Laptop *Toshiba Satellite*.
    - Procesador Intel® Core™ i5-3230M CPU 2.60GHz
    - 4 Gb de memoria RAM

La importancia de las herramientas utilizadas en el método de detección de suplantación facial propuesto en la presente investigación radica en lograr un equilibrio entre el rendimiento, la eficiencia y la precisión. Al utilizar herramientas y técnicas adaptadas a las limitaciones de recursos, se puede lograr una detección efectiva de suplantaciones faciales incluso en entornos con hardware limitado, lo que resulta clave para aplicaciones en dispositivos móviles, sistemas embebidos o escenarios donde se disponga de recursos limitados.

5.2 Bancos de imágenes

El proceso para seleccionar los bancos de imágenes utilizados en el proyecto se basó en varios criterios. En primer lugar, se consideró la disponibilidad y el acceso a los bancos de imágenes, priorizando aquellos a los que se podía acceder de manera gratuita. Además, se consideró la diversidad y representatividad de los bancos de imágenes en términos de escenarios de captura, variaciones de iluminación, expresiones faciales, poses y condiciones ambientales. Se buscó incluir bancos de imágenes que abarcaran una amplia gama de situaciones para entrenar y evaluar el método de detección de suplantación facial en distintos contextos.

En este sentido, se consideraron varios bancos de imágenes ampliamente utilizados en la comunidad de investigación, como NUAA [146], *Print-Attack Database* [147], CASIA [148], *Replay-Attack Database* [149], MSU [150], OULU-NPU[151], *REPLAY-MOBILE* [152] y LCC_FASD [153]. Estos bancos de imágenes son reconocidos por su variedad y cantidad de muestras, así como por sus diferentes condiciones de captura y escenarios controlados. Sin embargo, después revisar la disponibilidad de los bancos de imágenes, se decidió descartar algunos.

En última instancia, se seleccionaron los bancos de imágenes disponibles que mejor se ajustaban a los criterios de diversidad y representatividad, y se incluyó un banco de imágenes propio para entrenamiento (BIPE) y otro en entornos no controlados (ENC) con el fin de agregar mayor variedad y desafío al método de detección de suplantación facial. A continuación, se describen los diferentes bancos de imágenes.

1) Entrenamiento

El banco de entrenamiento BIPE se compone de tres bancos, dos de los cuales se utilizan con el objetivo de proporcionar una amplia variedad de imágenes de rostros genuinos para entrenar el método de detección de suplantación facial, en la Figura 5.1 se muestra un ejemplo de las imágenes. Los bancos son una valiosa fuente de imágenes auténticas que permiten al sistema familiarizarse con características faciales y variaciones. Al utilizarlos, se busca mejorar la capacidad del método para distinguir entre rostros reales y suplantados, proporcionando una base sólida para su entrenamiento y desarrollo con 9977 imágenes totales, los bancos de imágenes son:

- *FEI Face Database* [154]: Es una colección de imágenes faciales utilizada en investigaciones y desarrollos relacionados con el reconocimiento facial y la visión por computadora. Fue creada por el Laboratorio de Visión por Computadora y Robótica de la Universidad Estatal de Campinas (FEI) en Brasil.

El banco contiene imágenes de 200 individuos (100 hombres y 100 mujeres) de diferentes edades y orígenes étnicos, capturadas en condiciones controladas de iluminación y pose. FEI es utilizada en la comunidad académica para el desarrollo y evaluación de algoritmos de reconocimiento facial, detección de emociones y otras aplicaciones relacionadas con el análisis de rostros, debido a su disponibilidad y características.

- *Human faces* [155]: Es una colección de imágenes faciales utilizada en investigaciones y aplicaciones relacionadas con el reconocimiento facial. El banco de imágenes contiene más de 7200 imágenes para múltiples casos de uso, como reconocimiento facial.

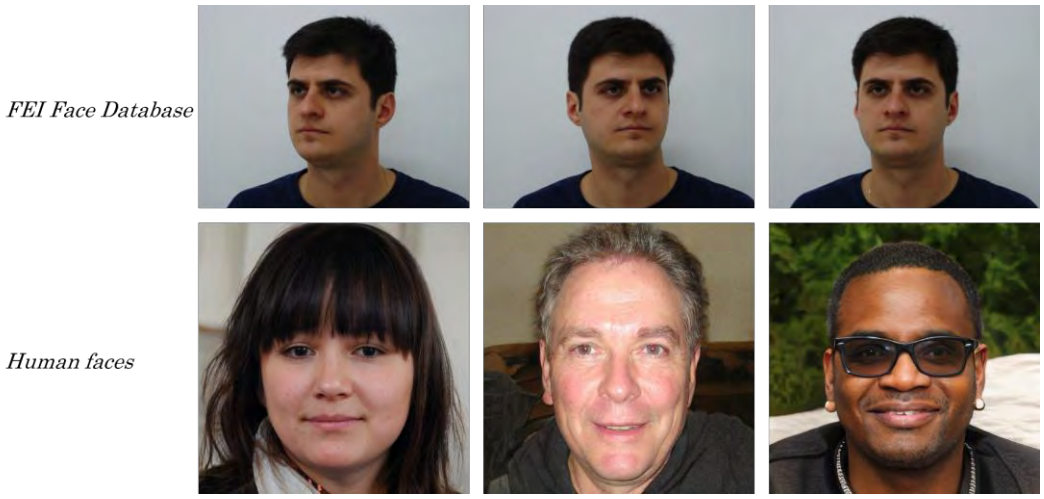

Figura 5.1 Muestra de imágenes reales utilizadas para entrenamiento.

Por otro lado, para las imágenes falsas se utiliza el banco GI4E, el cual es usado para simular ataques de suplantación. Se emplean tres dispositivos de captura de imágenes y cuatro objetos de suplantación para generar una diversidad de escenarios y condiciones de suplantación. Lo que permite recrear diferentes tipos de ataques y ayudan a mejorar la capacidad del sistema para detectar suplantaciones.

- *Gaze Interaction for Everybody* (GI4E) [156]: Es utilizado en investigaciones relacionadas con la interacción visual y el seguimiento de la mirada. El banco de imágenes contiene una amplia variedad de fotografías que representan diferentes direcciones y movimientos oculares, lo que permite estudiar el comportamiento de la mirada en diversas situaciones, así como la pose del rostro frente a la cámara. GI4E se utiliza como recurso para desarrollar y evaluar algoritmos y sistemas de seguimiento ocular, así como para investigar aplicaciones en áreas como la interacción hombre-máquina, la detección de atención y la realidad virtual.

El proceso de selección de los dispositivos de captura para ataques de suplantación se basó en varios criterios. En primer lugar, se tuvo en cuenta la calidad de imagen y la resolución de las cámaras para garantizar una captura detallada y nítida de los rostros. Se buscó utilizar cámaras con resoluciones variadas para evaluar el rendimiento del método en diferentes condiciones de captura. Además, se consideraron factores como la disponibilidad y accesibilidad de los dispositivos. Se seleccionaron cámaras comúnmente utilizadas y disponibles en el mercado, lo que facilita la replicación de los experimentos y la obtención de resultados comparables. También se valoró la diversidad en términos de las características técnicas de las cámaras, como la calidad del sensor y las opciones de configuración disponibles. Lo que permite evaluar el desempeño del método en diferentes entornos y situaciones de captura. En base a estos criterios, se eligieron los siguientes dispositivos:

- Cámara web AONI C98 4K HD con resolución de 3840 x 2160 píxeles.
- Cámara Logitech B525 con resolución de 1920 x 1080 píxeles.
- Cámara web de una laptop TOSHIBA *satellite* con resolución de 1280 x 720 píxeles.

El proceso de selección de los objetos utilizados para generar ataques de suplantación se basó en diferentes criterios. En primer lugar, se tuvo en cuenta la diversidad en los tipos de objetos utilizados en ataques de suplantación, que son diferentes dispositivos o medios utilizados comúnmente en la vida cotidiana, como tabletas, teléfonos celulares y fotografías impresas. Además, se consideraron aspectos técnicos de los objetos, como la resolución de pantalla en el caso de las tabletas y los teléfonos celulares. Se buscó utilizar dispositivos con resoluciones variadas para evaluar el impacto de la calidad de imagen en la detección de suplantación. Asimismo, se valoró la disponibilidad de los objetos en el mercado. En base a estos criterios, se eligieron los siguientes objetos:

- Tablet HUAWEI MediaPad M5 lite 10 con resolución FullHD de 1920 x 1200 píxeles.
- Teléfono celular Motorola G4 play, con una resolución de 1280 x 720 píxeles.
- Fotografía impresa por impresora BROTHER Business Smart pro de inyección de tinta en papel fotográfico inkjet matte de 220 gr.
- Fotografía impresa por impresora BROTHER Business Smart pro de inyección de tinta en hojas blancas de 75 gr.

Otros objetos que podrían haberse considerado incluyen diferentes modelos de tabletas, teléfonos celulares u otros medios impresos. Sin embargo, se descartaron debido a limitaciones de disponibilidad o accesibilidad. La selección final se basó en encontrar un equilibrio entre diversidad, representatividad y factibilidad para realizar los ataques de suplantación de manera efectiva. En la Figura 5.2 se muestra el ejemplo de las imágenes por ataque obtenidas. El total de imágenes de ataques es 7974.

Figura 5.2 Muestra de imágenes de ataques de suplantación para el banco de entrenamiento.

2) *NUAA Photograph Imposter Database* [146]

Es un banco de imágenes utilizado en investigaciones sobre detección de suplantación facial. Desarrollada por la *Northwestern Polytechnical University* (NUAA), contiene una amplia colección de imágenes de rostros genuinos y de impostores, es decir, personas que intentan suplantar la identidad de alguien más. Las imágenes genuinas son capturadas como fotografías normales, mientras que las imágenes de impostores simulan intentos de suplantación facial utilizando impresiones en papel fotográfico. En la Figura 5.3 se muestra un ejemplo del banco de imágenes:

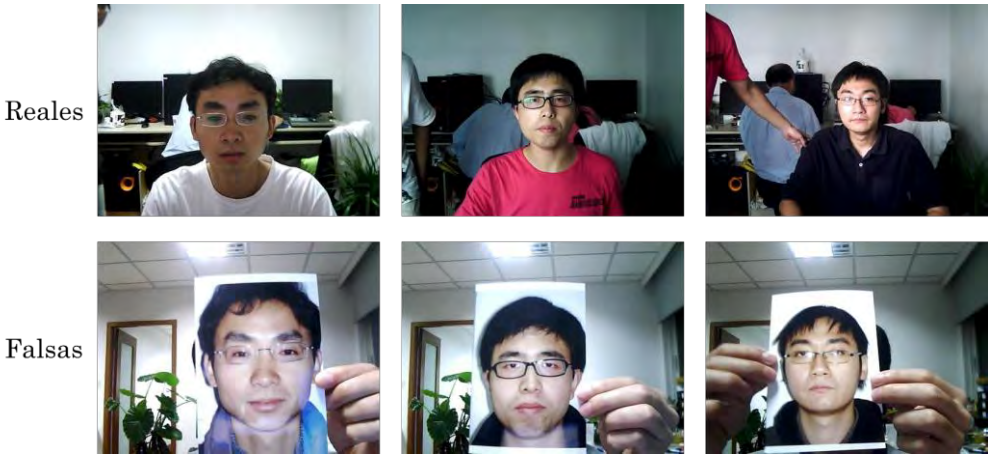

Figura 5.3 Muestra de imágenes contenidas en el banco de imágenes NUAA.

3) CASIA-FASD [148]

El banco de imágenes CASIA (*Chinese Academy of Sciences Institute of Automation*) es popularmente utilizado en investigaciones sobre reconocimiento facial y detección de suplantación facial. Fue desarrollado por el Instituto de Automatización de la Academia China de Ciencias y contiene una amplia colección de imágenes faciales de alta calidad. El banco CASIA proporciona imágenes de rostros genuinos, así como imágenes que simulan ataques de suplantación facial, (a) impresas en papel fotográfico, (b) impresiones con recorte de ojos y (c) reproducción por medio de tableta. En la Figura 5.4 se puede observar un ejemplo de los diferentes ataques.

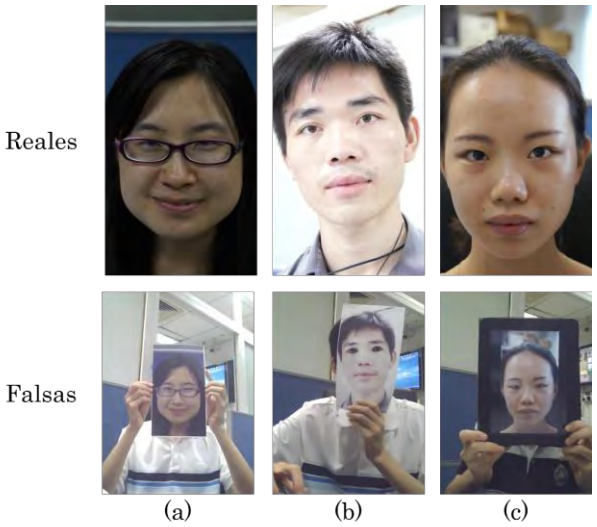

Figura 5.4 Imágenes de ejemplo del banco CASIA-FASD.

4) MSU [150]

El banco de imágenes MSU (Multimedia Security Lab) es un conjunto utilizado en investigaciones relacionadas con la seguridad multimedia, específicamente en el campo de la detección de manipulación y suplantación de imágenes. El banco de imágenes MSU-MFSD, ver Figura 5.5, consta de 440 videoclips de intentos de ataque con fotografías y video de 55 sujetos. Se utilizaron dos tipos de cámaras: i) cámara incorporada en *MacBook Air* 13, denominada cámara de portátil; ii) cámara frontal en el teléfono *Android Google Nexus* 5, denominada cámara Android. Para la cámara de la computadora portátil, los videos se capturan usando el marco QuickTime en la plataforma *Mac OS X Mavericks*, con una resolución de 640 × 480. Para la cámara de Android, los videos se capturaron usando el software de cámara integrado de Google en Android 4.4.2, con una resolución de 720 × 480.

Figura 5.5 Imágenes de ejemplo del banco de imágenes MSU MFSD capturadas con la cámara del teléfono inteligente Google Nexus 5 (fila superior) y MacBook Air 13" cámara portátil (fila inferior). (a) Caras reales; (b) Caras falsas generadas por iPad para ataques de reproducción de video; (c) caras falsas generadas por el iPhone para el ataque de reproducción de video; (d) Caras falsas generadas para el ataque fotográfico impreso. [150]

5) LCC_FASD

LCC_FASD (*Large-Scale CelebFaces Attributes for Forgery Detection*) es un banco de imágenes utilizado en el campo de la detección de suplantación facial. Se compone de una colección de imágenes faciales de celebridades recopiladas de *Youtube*, *Amazon Mechanical Turk* (https://www.mturk.com/) y *Yandex Toloka* (https://toloka.yandex.com/) y abarca diferentes expresiones faciales. Además de las imágenes de rostros genuinos, el banco también incluye imágenes reproducidas desde las pantallas de diferentes calidades. Ver Figura 5.6.

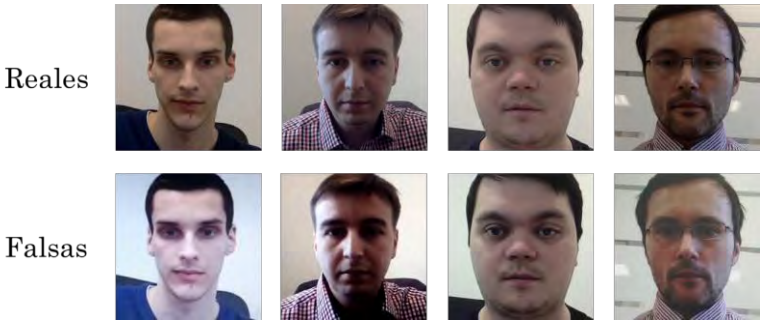

Figura 5.6 Muestra del banco de imágenes LCC FASD.

6) Entornos no controlados (ENC)

Consta de imágenes de 136 personas en entornos no controlados (variación de iluminación) con resoluciones que van desde los 320 x 240 hasta los 4032 x 3024, las cuales fueron adquiridas durante diversos exámenes de admisión realizados en el CENIDET, ver Figura 5.7.

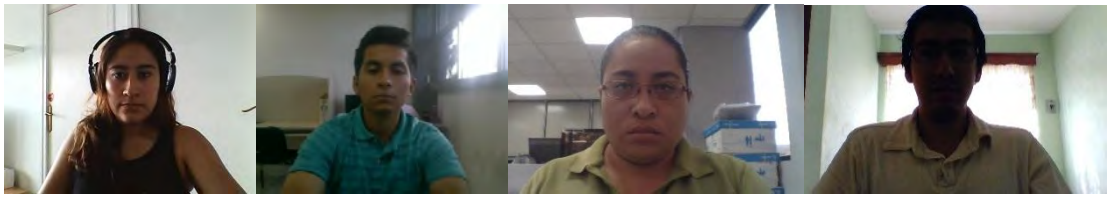

Figura 5.7 Muestra de imágenes reales de ENC.

Para llevar a cabo los ataques de suplantación, se emplearon tres dispositivos de captura de imágenes y cuatro objetos de suplantación. Los dispositivos de captura son: una cámara web AONI C98 4K HD con una resolución de 3840 x 2160 píxeles, una cámara Logitech B525 con una resolución de 1920 x 1080 píxeles, y una cámara web de una laptop TOSHIBA Satellite con una resolución de 1280 x 720 píxeles. Estos dispositivos de captura permitieron obtener imágenes de alta calidad y variedad de características.

Por otro lado, los objetos utilizados para la suplantación consistieron en una tablet HUAWEI MediaPad M5 Lite 10 con una resolución FullHD de 1920 x 1200 píxeles, un teléfono celular Motorola G4 Play con una resolución de 1280 x 720 píxeles, una fotografía impresa en papel fotográfico inkjet matte de 220 gr utilizando una impresora BROTHER Business Smart Pro de inyección de tinta, y una fotografía impresa en hojas blancas de 75 gr también utilizando la misma impresora. Los objetos proporcionaron diferentes tipos de ataques de suplantación para evaluar la robustez del sistema de detección.

La utilización de los dispositivos de captura y objetos de suplantación permitió simular una amplia gama de escenarios y condiciones para evaluar la efectividad del método de detección de suplantación facial. Con lo que se buscó una mayor representatividad y robustez en el proceso de evaluación. En la Figura 5.8 se muestra un ejemplo de los ataques de suplantación con los diferentes objetos y sensores.

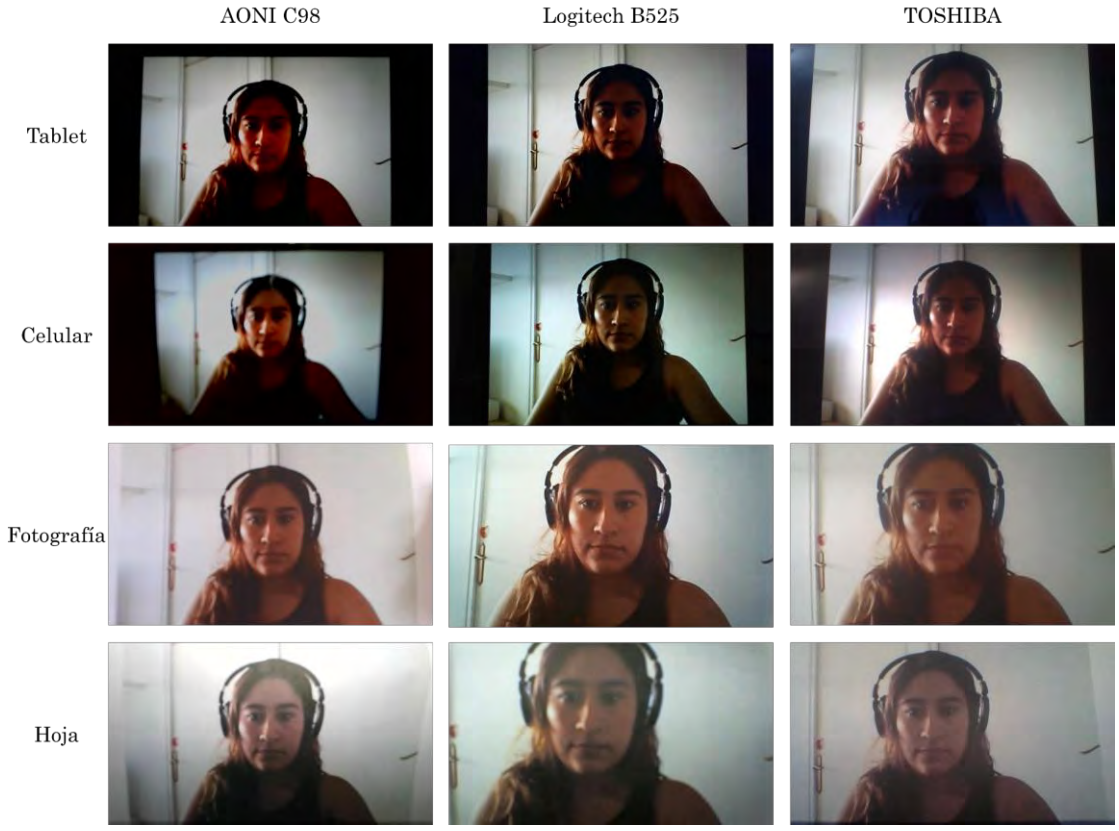

Figura 5.8 Muestra de imágenes de ataques en el banco ENC.

En la Tabla 5.1 se muestra el número de imágenes utilizadas por banco para entrenamiento y predicción. Para las pruebas de validación cruzada se utiliza el banco BIPE para entrenamiento y la columna de predicción de cada banco de imágenes para pruebas.

En el contexto del problema de desbalance de clases, la selección de imágenes se llevó a cabo mediante el uso de muestreo estratificado. La técnica se basa en una selección equilibrada de muestras de cada clase, con el objetivo de asegurar que todas las clases tengan una representación similar en el conjunto de datos de entrenamiento. Para lograr esto, se tomaron muestras en proporciones iguales de cada clase.

Tabla 5.1 Imágenes utilizadas por banco.

Nombre	Entrenamiento			Predicción		
	Falsas	Reales	Total	Falsas	Reales	Total
NUAA	5293	2166	7459	2216	1987	4203
CASIA	135	135	270	135	135	270
MSU	180	180	360	240	240	480
LCC_FASD	1363	1223	2586	719	719	1438
ENC	300	300	600	566	1791	2357

5.3 Experimentación

Se realizaron cinco diferentes pruebas, las cuales tienen por objetivo evaluar y validar el desempeño del método de detección de suplantación facial. Cada prueba se diseñó con propósitos específicos, que incluyen:

A) **Entrenamiento y pruebas con el mismo banco de imágenes**
La finalidad es examinar que tan discriminantes son las características obtenidas del procesamiento con el método propuesto, contemplando que las características del banco de imágenes de entrada son similares al banco de imágenes de prueba.

B) **Entrenamiento con el banco de imágenes propio**
Se tiene como objetivo evaluar los resultados con el banco de imágenes propuesto al presentar condiciones hostiles.

C) **Pruebas por tipo de ataque**
El objetivo es analizar qué tipo de ataque presenta un mayor desafío ante el entrenamiento propuesto.

D) **Comparativa con la literatura**
El objetivo de la prueba es comparar, de manera teórica, los resultados obtenidos respecto a los reportados en la literatura.

E) **Tiempo de procesamiento**
El objetivo es revisar el tiempo estimado que tarda el algoritmo partiendo desde el momento que se ingresa la imagen hasta la salida con el resultado.

5.4 Análisis de resultados

5.4.1 Entrenamiento y pruebas con el mismo banco de imágenes

Se realizaron pruebas con cada banco de imágenes los cuales contienes una parte de entrenamiento y otra para las pruebas, el número de imágenes se muestra en la Tabla 5.1. Las pruebas tienen como finalidad revisar los resultados del método propuesto en bancos de imágenes con características similares en entrenamiento y validación en la Tabla 5.2, se observan los resultados por banco de imágenes con los cuatro clasificadores y cuatro métricas.

Tabla 5.2 Resultados de la experimentación con cada banco de imágenes expresado en porcentajes.

	Clasificador	APCER	BPCER	HTER	F-mesure
CASIA	MVS	60	**38.52**	49.26	55.52
	MLP	45.19	49.63	**47.41**	51.52
	Bosque aleatorio	**38.52**	57.78	48.15	46.72
	Árboles de decisión	52.59	46.67	49.63	**51.8**
LCCC_FASD	MVS	42.42	65.79	54.1	38.74
	MLP	43.25	**47.29**	45.27	**53.8**
	Bosque aleatorio	**21.56**	60.78	**41.17**	48.79
	Árboles de decisión	45.76	48.54	47.15	52.19
NUAA	MVS	55.64	**42.17**	48.91	52.6
	MLP	23.38	44.74	**34.06**	**60.95**
	Bosque aleatorio	**9.97**	84.05	47.01	25.11
	Árboles de decisión	33.21	45.9	39.56	56.61
MSU	MVS	**26.25**	73.75	50	34.43
	MLP	43.75	**42.08**	42.92	57.44
	Bosque aleatorio	35	47.08	41.04	56.32
	Árboles de decisión	35.42	43.33	**39.38**	**59**
ENC	MVS	41.6	63.95	52.77	24.37
	MLP	25.01	40.56	32.79	46.52
	Bosque aleatorio	**22.84**	**33.91**	**28.37**	**52.07**
	Árboles de decisión	24.46	46.78	35.62	43.06

Comentarios

Con la evaluación realizada, se puede concluir, que el clasificador bosque aleatorio muestra un mejor desempeño con las características de entrada proporcionadas por el método propuesto. Como se observa en el penúltimo renglón, con en el banco de imágenes ENC (entornos no controlados) en las 4 métricas el bosque aleatorio logra superar a los otros 3 clasificadores, lo que podría indicar que para el problema planteado tiene un mejor funcionamiento. Sin embargo, para asegurar que el bosque aleatorio es la mejor opción, es necesario realizar más análisis y pruebas adicionales en otros bancos de imágenes.

Por otro lado, el clasificador con el peor desempeño es la máquina de vectores de soporte, lo que es contrario a lo reportado en la literatura debido a que es el clasificador más utilizado en las investigaciones. Con el banco de imágenes NUAA la diferencia entre los resultados con el bosque aleatorio y máquinas de vector soporte es de 45.67% respecto a la métrica APCER, lo que indica que en la detección de imágenes sintéticas con el bosque aleatorio tiene un mejor desempeño.

5.4.2 Entrenamiento con el banco de imágenes propio

Se llevó a cabo el proceso de entrenamiento utilizando el banco de imágenes propio denominado BIPE. Para evaluar la tasa de aprendizaje y clasificación, se realizó el entrenamiento y las pruebas utilizando el mismo banco de imágenes designado para el entrenamiento. Lo que permitió validar el rendimiento del sistema al evaluar su capacidad para reconocer y clasificar correctamente las imágenes del banco. La validación interna brinda una medida confiable del desempeño del método en relación con las imágenes utilizadas durante el proceso de entrenamiento. Los resultados se muestran en la Tabla 5.3.

Tabla 5.3 Resultados de validación del entrenamiento expresado en porcentajes.

Clasificador	APCER	BPCER	HTER	F-mesure
MVS	57.51	20.51	39.01	69.71
MLP	**0.45**	1.44	0.94	99.09
Bosque aleatorio	2.68	10.95	6.82	93.09
Árboles de decisión	0.64	**0.89**	**0.77**	**99.28**

Como se observa en la Tabla 5.3 el mejor aprendizaje se alcanza con el algoritmo de árboles de decisión seguido por perceptrón multicapa. Cabe resaltar que en las métricas de APCER, BPCER y HTER cuando el resultado es cercano al 0 es mejor debido a que el porcentaje expresa el porcentaje de error al momento de clasificar. En el caso contrario, la máquina de vector soporte no logra tener resultados cercanos al 0, lo que indica que con dicho clasificador no se logra un correcto aprendizaje.

Una vez que se tienen los modelos de clasificación, se realizaron las pruebas con los bancos de imágenes NUAA, CASIA, MSU, LCC_FASD y ENC. A modo de comparación en la Tabla 5.4 se muestran los resultados obtenidos con los cuatro clasificadores y las cuatro métricas.

Tabla 5.4 Resultados del entrenamiento con el banco de imágenes propio expresado en porcentajes.

	Clasificador	APCER	BPCER	HTER	F-mesure
NUAA	MVS	64.44	**38.95**	51.7	**52.42**
	MLP	17.51	79.77	48.64	28.95
	Bosque aleatorio	**8.03**	91.85	49.94	13.92
	Árboles de decisión	17.33	74.84	**46.08**	34.83
CASIA	MVS	79.26	**20.74**	50	**61.32**
	MLP	18.52	80.74	**49.63**	27.96
	Bosque aleatorio	**12.59**	91.11	51.85	14.63
	Árboles de decisión	26.67	72.59	**49.63**	35.58
MSU	MVS	60.83	**23.33**	42.08	**64.56**
	MLP	25.83	58.75	42.29	49.38
	Bosque aleatorio	**12.92**	70.42	41.67	41.52
	Árboles de decisión	21.25	60.42	**40.83**	49.22
LCC_FASD	MVS	63.42	**28.65**	46.04	**60.78**
	MLP	33.8	58.55	46.18	47.3
	Bosque aleatorio	**26.15**	57.86	**42**	50.08
	Árboles de decisión	37.27	49.93	43.6	53.45
ENC	MVS	77.04	**28.07**	**52.56**	**37.64**
	MLP	**59.73**	82.51	71.12	12.47
	Bosque aleatorio	60.59	88.51	74.55	8.3
	Árboles de decisión	64.08	78.85	71.47	14.29

Como se observa en la Tabla 5.4 para la métrica APCER los mejores resultados se obtienen con el clasificador de bosque aleatorio, mientras que para la métrica BPCER se obtienen con las máquinas de vector soporte sin que ninguno de los clasificadores logre sobresalir en las cuatro métricas.

Comentarios

Con base en la evaluación realizada, se puede concluir que los clasificadores como árboles de decisión, bosque aleatorio y perceptrón multi-capa pueden ser alternativas viables al uso de máquinas de vector soporte, que son ampliamente reportadas en la mayoría de las investigaciones en detección de suplantación facial. Debido a que se observa una diferencia significativa en los resultados, como en el caso del banco de imágenes CASIA, donde se alcanza un error del 12.59% con el clasificador de bosque aleatorio, mientras que con la máquina de vector soporte se obtiene un error del 79.26%, lo que representa una diferencia del 66.67%.

Sin embargo, para poder ofrecer una solución contundente, es necesario realizar más análisis y evaluaciones exhaustivas. Se deben considerar otros factores, como el tamaño y la diversidad del conjunto de datos. Además, es fundamental realizar pruebas en diferentes escenarios y con conjuntos de datos más amplios y variados para validar la eficacia y la generalización del enfoque propuesto. Lo que permitirá obtener resultados más sólidos y confiables, respaldando la solución propuesta.

5.4.3 Pruebas por tipo de ataque

Para las pruebas por tipo de ataque se utilizaron los bancos de imágenes propios BIPE para el entrenamiento y ENC para validación, para lo cual se dividen las pruebas en dos partes que son: por tipo de ataque (digital y papel) y por calidad de la imagen (buena y mala).

El objetivo de realizar la prueba por tipo de ataque es analizar cuál de los dos tiene mayor complejidad al momento de realizar una detección de suplantación en la Tabla 5.5 se muestran los resultados.

Tabla 5.5 Resultados de pruebas por tipo de ataque expresado en porcentajes.

	Clasificador	APCER	BPCER	HTER	F-mesure
Digitales	MVS	65.44	**27.94**	**46.69**	**59.61**
	MLP	**40.32**	82.51	61.41	21.81
	Bosque aleatorio	41.18	88.51	64.84	14.79
	Árboles de decisión	50.61	78.72	64.67	24.29
Papel	MVS	88.84	**28.07**	**58.46**	**56.4**
	MLP	78.24	82.51	80.37	18.34
	Bosque aleatorio	77.68	88.51	83.1	12.47
	Árboles de decisión	**76.43**	78.85	77.64	21.95
Totales	MVS	77.04	**28.07**	**52.56**	**37.64**
	MLP	**59.73**	82.51	71.12	12.47
	Bosque aleatorio	60.59	88.51	74.55	8.3
	Árboles de decisión	64.08	78.85	71.47	14.29

Como se observa en la Tabla 5.5 se tiene un mejor resultado con las imágenes digitales ya que los resultados de los 4 clasificadores respecto a la métrica APCER oscila entre 40.32% y 65.44% mientras que para los ataques por medio de papel impreso oscila entre 76.43% y 88.84%. La diferencia en el resultado se debe a que el método propuesto en la presente investigación trabaja con brillo y saturación para lograr hacer una diferencia entre textura lo cual beneficia a las imágenes digitales ya que el reflejo de las pantallas incrementa el brillo que disminuye a la textura. Las imágenes en papel no muestran reflejo extra que proporcione un brillo y con ello se favorezca a la detección, por lo cual la textura del papel provoca que el método no logre detectar la diferencia entre una imagen real y una falsa.

Otra prueba de interés es evaluar cómo el método y los clasificadores se comportan frente a imágenes que tienen una buena calidad visualmente, como se puede apreciar en la Figura 5.9. Además, también es importante analizar cómo se desempeñan ante imágenes que presentan problemas, como mala iluminación y baja resolución en entornos reales. Cuando se habla de problemas de resolución en entornos reales, se hace referencia a situaciones en las que las imágenes presentan una baja resolución debido a factores como la baja calidad de la cámara, la distancia entre la cámara y el sujeto, el movimiento del sujeto durante la captura, entre otros. Lo que puede resultar en imágenes borrosas, pixeladas o con poca definición, por lo tanto, se dificulta la extracción de características precisas y la detección confiable de suplantaciones faciales. Un ejemplo de este tipo de imágenes se muestra en la Figura 5.10. Las pruebas permiten examinar la capacidad del sistema de detección para manejar situaciones desafiantes y determinar si es capaz de detectar suplantaciones faciales de manera precisa y confiable incluso en condiciones adversas.

Figura 5.9 Imágenes con buena iluminación y resolución.

Figura 5.10 Imágenes con problemas en iluminación y resolución.

Los resultados de la experimentación se muestran en la Tabla 5.6, en donde se observa que no existe una diferencia significativa cuando se trata de problemas de iluminación o calidad de la imagen. Se puede concluir que el método está siendo robusto ante este problema, aun así, las métricas reportan valores elevados lo que podría deberse a lo mencionado en la experimentación por tipo de ataque, que indica que el tipo de ataque sí influye en el resultado.

Tabla 5.6 Resultados por calidad de imagen expresado en porcentajes.

	Clasificador	APCER	BPCER	HTER	F-mesure
Buena	MVS	79.9	**29.61**	**54.75**	**68.01**
	MLP	**56.46**	84.87	70.66	21.46
	Bosque aleatorio	60.77	87.72	74.24	17.53
	Árboles de decisión	64.59	78.07	71.33	28.94
Mala	MVS	76.13	**30.1**	**53.11**	**28.13**
	MLP	59.04	80.58	69.81	10.41
	Bosque aleatorio	**58.81**	87.38	73.09	6.91
	Árboles de decisión	64.08	74.76	69.42	12.6
Totales	MVS	77.04	**28.07**	**52.56**	**37.64**
	MLP	**59.73**	82.51	71.12	12.47
	Bosque aleatorio	60.59	88.51	74.55	8.3
	Árboles de decisión	64.08	78.85	71.47	14.29

Comentarios

Con la evaluación realizada, se puede concluir, que el tipo de ataque si es relevante al momento de realizar la detección de suplantación, debido a las características de cada uno de ellos, sin embargo, la iluminación y resolución si bien presentan retos no provocan una gran diferencia al momento de introducirlos al método propuesto.

5.4.4 Comparativa con la literatura

El objetivo de esta prueba es realizar una comparativa entre los resultados reportados en esta investigación y los que se reportan en la literatura que son lo más apegados al tema planteado. Las tablas se dividen en dos: en la Tabla 5.7 se muestra la comparación con literatura que utiliza validación cruzada con el banco de imágenes CASIA y, en la Tabla 5.8 se muestran los resultados con el banco de imágenes MSU.

Tabla 5.7 Comparación de la literatura en el banco de imágenes CASIA expresado en porcentajes.

Referencia	Método	Banco de imágenes Entrenamiento → Pruebas	APCER	BPCER	HTER
2020 [51]	Calidad de la imagen	*Replay attack* →CASIA			30.2
2020 [52]	Representación de relatividad en la variedad de Riemann con características de *Haralick*	*Replay attack* → CASIA			27.59
2021 [53]	Uso de filtros paso alto y paso bajo con una red pseudo-siamesa para extraer las características de suplantación de HF y LF.	MSU → CASIA			35.7
		Replay attack → CASIA			27.2
2022 [54]	Características de evaluación de calidad de imagen perceptiva multi-escala	*Replay attack* → CASIA	75.83	**2.22**	39.03
		UVAD → CASIA	97.50	8.89	53.19
2022 [55]	Parches faciales y progresión lineal	OULU-NPU & MSU & Replay attack → CASIA			**9.81**
2023 [56]	Operador de gradiente aprendible (*learnable gradient operator* LGO) basado en sobel	Replay&SiW&OULU→ CASIA			22.78
		Replay attack → CASIA			31.9
2021 [74]	Extracción de características discriminatorias basadas en una recomposición de los componentes aprendidos de baja / alta frecuencia por medio de CNN	MSU → CASIA			27.3
2021 [75]	Red de destrucción y combinación (DCN)	MSU & Replay attack → CASIA			32.3
		Replay attack → CASIA			29.4
Propuesta	Método propuesto	MSU→ CASIA	70.73	40.74	55.56
		BIPE→ CASIA	**12.59**	91.11	51.85

Como se observa en la Tabla 5.7 en la mayoría de las investigaciones solo se reportan los resultados de la métrica HTER, lo que deja en duda el comportamiento de los métodos que reportan con las métricas APCER y BPCE; sin embargo, que lo que sí se puede observar es que cuando se trata de entrenar con un banco de imágenes diferente al de pruebas los resultados que reportan no son tan cercanos al 0 que sería lo ideal.

A diferencia con el estado del arte cuando se entrena con un banco de imágenes diferente a CASIA, el método propuesto mejora significativamente con un 12.59% en APCER en comparación con otras investigaciones que superan el 70% de error. Con respecto a la métrica BPCER se reportan resultados menores al 9% en comparación del método propuesto que supera el 91% de error, lo que indica que la propuesta funciona mejor para la detección de imágenes falsas que para las reales lo contrario a las investigaciones que tienen un mejor funcionamiento para las imágenes reales que las falsas.

Por otro lado, en la Tabla 5.8 se muestra la comparativa con el banco de imágenes MSU y de la misma manera los resultados que se reportan en su mayoría son con la métrica HTER logrando resultados por debajo del 20%. Sin embargo, en esta comparativa las investigaciones carecen de resultados con las métricas APCER y BPCER lo que impide una comparación con las dos, aun así, se puede percibir que el porcentaje de error con el método propuesto para la métrica APCER es bueno con un 12.92% de error y de BPCER con un 70.42%, característica que prevalece en el banco de imágenes CASIA y MSU; es decir, tener un mejor desempeño con las imágenes falsas sobre las reales.

Tabla 5.8 Comparación de la literatura en el banco de imágenes MSU expresado en porcentajes.

Referencia	**Metodología**	**Banco de imágenes Entrenamiento → Pruebas**	**APCER**	**BPCER**	**HTER**
2021 [53]	Tres filtros de paso alto y tres filtros de paso bajo para crear los mapas de HF y los mapas de LF. Posteriormente, se implementa una red pseudo-siamesa para extraer las características de suplantación de HF y LF.	CASIA → MSU			18.8
		Replay attack → MSU			24.3
2021 [75]	Red de destrucción y combinación (DCN)	CASIA → MSU			17.5
		CASIA & *Replay attack* → MSU			**14.8**
Propuesta	Método propuesto	CASIA→ MSU	34.17	**67.92**	51.04
		BIPE→ MSU	**12.92**	70.42	41.67

Comentarios

Con base en la evaluación realizada, se puede concluir que, a pesar de los avances relevantes y las numerosas investigaciones sobre el tema, los métodos existentes presentan dificultades en la detección precisa de posibles ataques de suplantación cuando se realizan pruebas de validación cruzada. Lo que indica que aún queda un camino considerable por recorrer para comprender plenamente los factores que influyen en lograr una detección asertiva tanto en imágenes falsas como reales. El recorrido de este camino requiere la colaboración y el esfuerzo continuo de investigadores, académicos, profesionales y la comunidad científica en general para desarrollar y mejorar las técnicas de detección de suplantación facial y abordar los desafíos asociados con este campo en constante evolución.

5.4.5 Tiempo de procesamiento

Una parte importante de los sistemas es el tiempo que se requiere para proporcionar una respuesta, especialmente en aquellos que operan en tiempo real. Por lo tanto, es fundamental revisar el tiempo que el método propuesto tarda en procesar la imagen, desde su entrada hasta la salida, determinando si la imagen analizada es o no un posible ataque de suplantación. Un sistema que trabaja en tiempo real es aquel que procesa y responde a los eventos o datos de entrada de manera inmediata, dentro de un límite de tiempo definido, lo que significa que deben cumplir con ciertas características clave:

- Tiempo de respuesta determinista: Los sistemas en tiempo real deben garantizar un tiempo de respuesta predecible y acotado. Lo que implica que se conoce y cumple un límite máximo de tiempo para procesar las solicitudes y generar una respuesta.

- Condiciones de carga variable: Los sistemas deben ser capaces de manejar fluctuaciones en la carga de trabajo sin comprometer su tiempo de respuesta.

En la Figura 5.11 se muestra una gráfica con el comportamiento de los tiempos promedios que se tarda el sistema en dar respuesta por banco de imágenes siendo un promedio de 0.56 segundos en total, si se considera como tiempo real de 0 a 3 segundos, el método propuesto podría ser implementado en sistemas de verificación de identidad sin tener repercusiones significativas en los tiempos de respuesta.

Es importante mencionar que la variación del tiempo depende de la calidad de la imagen, así como de la resolución, ya que a mayor calidad más tarda el sistema en dar respuesta. El número de imágenes totales por banco de imágenes se puede consultar en la Tabla 5.1, sección 2.

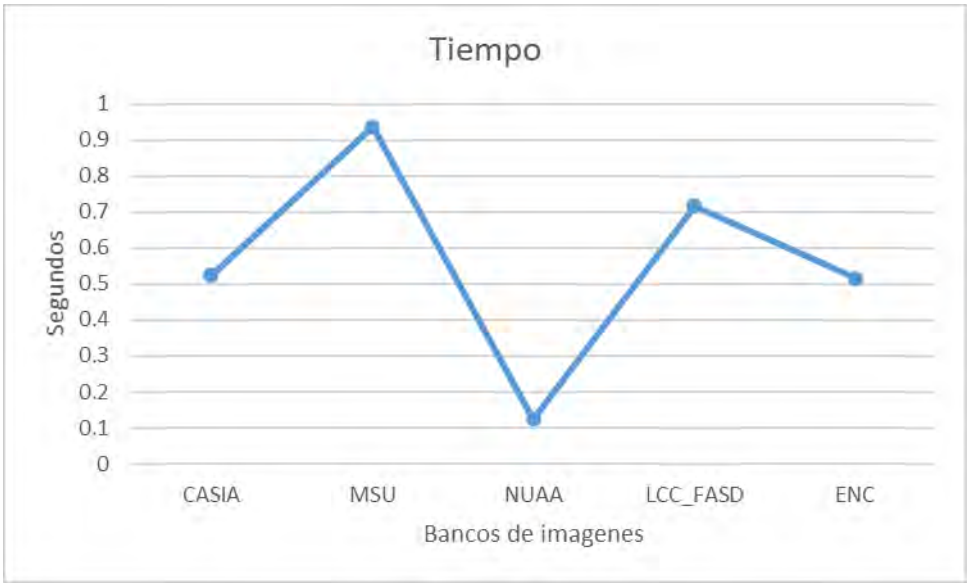

Figura 5.11 Gráfica de comportamiento de tiempo.

5.5 Comentarios finales

Después de realizar las diferentes pruebas se concluye lo siguiente: tomando como referencia la métrica APCER, el método propuesto alcanza buenos resultados con la combinación de algoritmos propuesta considerando un escenario hostil, con porcentajes por debajo del 50%; sin embargo, en la métrica BPCER los porcentajes de error se elevan dejando una ventana de posibilidad para investigar cuáles son los factores que repercuten al momento de hacer una detección de imágenes reales en ambientes no controlados.

Los resultados de la experimentación confirman que el clasificador de bosque aleatorio es el que muestra mejores resultados en la etapa de clasificación. Es interesante destacar que el clasificador no es ampliamente mencionado en la literatura como una opción común para abordar el problema de suplantación. Sin embargo, es importante tener en cuenta que la literatura científica puede tener un enfoque más establecido en ciertos algoritmos o clasificadores ampliamente conocidos y utilizados en las investigaciones de suplantación. Lo que no significa necesariamente que otros clasificadores, como el bosque aleatorio, sean menos efectivos, sino que pueden haber recibido menos atención o estudio en el contexto de la detección de suplantación facial. Motivo por el cual se decidió incluirlo en la experimentación del método propuesto. En cuanto a tiempos, el algoritmo propuesto reporta tiempo aceptables al no superar los 0.59 segundos en promedio en la detección de suplantación siendo una opción viable para implementarse en sistemas de tiempo real.

CAPÍTULO 6
Conclusiones

La investigación en detección de suplantación facial ha logrado avances significativos, pero aún existen áreas de mejora y desafíos por enfrentar. La selección adecuada de bancos de imágenes, la elección de características y clasificadores efectivos, son aspectos fundamentales para garantizar una detección de suplantación facial precisa y confiable.

6.1 Conclusiones

En este trabajo se presentó un método para llevar a cabo las tareas de detección de posibles ataques de suplantación, así como la experimentación y los resultados.

Se presenta un método para abordar el problema planteado en el tema de la tesis, al cual se llegó mediante una evolución con diferentes técnicas. El objetivo principal fue desarrollar un enfoque basado en visión artificial tradicional que pudiera reducir el uso de recursos al prescindir de hardware específico.

En términos de recursos, se logró una reducción significativa en varios aspectos clave. En primer lugar, en cuanto a los recursos computacionales, el método final requiere un menor poder de procesamiento en comparación con enfoques más complejos basados en técnicas avanzadas de aprendizaje automático. Lo que permite ejecutarlo en equipos de cómputo convencionales sin la necesidad de utilizar *clústers* de alta capacidad o aceleradores de hardware especializados.

En cuanto a los recursos financieros, el método basado en visión artificial tradicional también presenta una ventaja significativa al no requerir la adquisición de costosos dispositivos de hardware específicos. En lugar de depender de tecnologías sofisticadas y costosas, se aprovecharon algoritmos y técnicas tradicionales disponibles en bibliotecas y herramientas ampliamente utilizadas, lo que resulta en un enfoque más accesible desde el punto de vista económico.

Se presenta un método de visión artificial final que utiliza técnicas de iluminación, como el *retinex multi-escala* y los canales de color (HSV, YCbCr), con el fin de destacar las diferencias en términos de textura entre una imagen real y una falsa en entornos no controlados. El enfoque propuesto se caracteriza por su capacidad para trabajar con imágenes de entrada que no presentan limitaciones en cuanto a resolución, calidad de la imagen, iluminación y pose. Sin embargo, es importante tener en cuenta que, aunque el método ha mostrado resultados prometedores, no se puede garantizar que funcione correctamente con todas las imágenes. La detección de suplantación facial es un desafío complejo debido a la diversidad de factores que pueden afectar la apariencia de las imágenes, como cambios en la iluminación, la calidad de la imagen y las poses faciales.

Se propone el uso de PRNU, algoritmo que calcula la huella digital de las cámaras en términos de ruido, para hacer uso del mismo en términos de textura, debido a que el ruido que se genera tiene variaciones con base en el brillo del objeto que se captura, característica que se resalta con los canales de color y así logra una verificación lo más asertiva posible.

Para comprobar la robustez del método haciendo uso del vector de características se compararon cuatro clasificadores que son: máquinas de vector soporte (MVS), perceptrón multicapa, bosque aleatorio y árboles de decisión; los cuales son evaluados considerando: APCER, BPCER, HTER y F-mesure. El objetivo de buscar comprobar la robustez de las máquinas de vector soporte como parte de la evaluación de los clasificadores se debe a su amplio uso y reconocimiento en la literatura científica como una opción común para abordar problemas de clasificación, incluyendo la detección de suplantación facial. Sin embargo, a pesar de su popularidad, es fundamental analizar su desempeño y compararlo con otros clasificadores como se realizó en la presente investigación. Después de los experimentos se concluye que el método propuesto en combinación con el clasificador bosque aleatorio alcanza el mejor rendimiento en comparación con los otros clasificadores; siendo las máquinas de vector soporte uno de los algoritmos más utilizados dentro de la literatura.

En conclusión, el trabajo presenta un método de detección de suplantación facial basado en visión artificial tradicional, con reducción de recursos en términos computacionales y financieros. El método mostró resultados prometedores en la detección de posibles ataques de suplantación, aunque se reconoce la necesidad de continuar mejorándolo y considerar las limitaciones inherentes a la diversidad de factores que pueden afectar la apariencia de las imágenes.

6.2 Objetivos logrados

En la Tabla 6.1 se muestra una breve descripción de cómo se cumplió con el objetivo general y cada objetivo específico.

Tabla 6.1 Logros por objetivo.

Objetivo general	**Solución del objetivo**
Proponer el desarrollo de un sistema de visión artificial que detecte posibles ataques de suplantación de identidad mediante imágenes faciales, en un ambiente no controlado.	Se implementaron diferentes técnicas para la investigación de la detección de suplantación facial en entornos no controlados, basado en la experimentación se propone un método final, el cual aborda los entornos no controlados y sus desafíos.
Objetivos específicos	**Solución del objetivo**
Revisar en el estado del arte las técnicas utilizadas para los distintos tipos de ataques utilizados en la detección de suplantación de identidad	Se revisaron 150 artículos de relevancia para la presente investigación, los cuales abarcan desde posibles aplicaciones, enfoque de visión artificial tradicional y aprendizaje profundo.

Tabla 6.2 Continuación de logros por objetivo.

Objetivos específicos	Solución del objetivo
Evaluar el nivel mínimo de iluminación y de visibilidad del rostro necesaria para realizar la detección de suplantación de identidad.	El método propuesto no impone restricciones en cuanto a la iluminación o visibilidad del rostro en las imágenes de entrada. Debido a que el método considera características como la saturación, el brillo y el ruido de la cámara para analizar la imagen, sin depender de rasgos faciales específicos. Por lo tanto, el método es capaz de funcionar incluso en condiciones de baja iluminación o visibilidad, lo que lo hace más robusto y adecuado para entornos no controlados.
Seleccionar e implementar las técnicas con los mejores resultados, de acuerdo con el estado del arte, para describir el rostro.	Se experimentó con técnicas de visión tradicional y aprendizaje profundo reportadas en la literatura de lo cual se seleccionaron algoritmos de iluminación, color, textura y clasificadores.
Formular una variante de una técnica o la combinación de varias, para dar solución al problema propuesto.	Se propone un método enfocado en visión artificial tradicional que contempla técnicas de selección del rostro, canales de color y PRNU para la extracción de características.
Evaluar el método propuesto utilizando repositorios públicos especializados.	En el transcurso de la investigación se recopilaron cuatro bancos de imágenes de dominio público y se generaron dos bancos, uno para entrenamiento y otro para pruebas. Los resultados fueron evaluados con las métricas APCER, BPCER, HTER y F1-Score
Comparar con otros métodos del estado del arte usando las métricas aplicadas a la suplantación del rostro	En el transcurso del doctorado se han reportado las comparativas de los resultados parciales obtenidos con el estado del arte.

6.3 Aportaciones

Las aportaciones que se obtuvieron en la realización del proyecto de investigación son:

1. Algoritmo de detección de suplantación de identidad basado en el rostro.
   Se propone un algoritmo de detección de posibles ataques de suplantación de identidad utilizando técnicas de visión artificial. El algoritmo presenta varias ventajas, ya que no requiere que la imagen de entrada cumpla con requisitos estrictos de iluminación, pose específica del rostro frente a la cámara o una resolución determinada. Además, tiene la capacidad de analizar tanto imágenes individuales como secuencias de video sin necesidad de realizar un entrenamiento previo del rostro que se desea analizar. Una de las principales fortalezas del algoritmo es su flexibilidad para trabajar con diferentes condiciones de captura de imágenes, lo que lo hace adecuado para entornos no controlados. No se limita a imágenes con buena iluminación o poses específicas, lo que amplía su aplicabilidad en situaciones del mundo real donde las condiciones de captura pueden variar ampliamente.

2. Evaluación de clasificadores
   Se llevó a cabo una comparativa de diferentes clasificadores utilizados en la literatura científica para abordar el problema de la detección de suplantación de identidad. Además, se exploraron clasificadores que no son ampliamente utilizados en este contexto. El objetivo fue determinar qué clasificador ofrece mejores resultados en términos de las métricas APCER, BPCER, HTER y F1-Score. Los clasificadores evaluados incluyeron máquinas de vector soporte, árboles de decisión, perceptron multi-capa y bosque aleatorio. Los clasificadores fueron seleccionados debido a su relevancia en la literatura y su potencial para abordar problemas de clasificación. Los resultados obtenidos en la experimentación revelaron que el uso exclusivo de máquinas de vector soporte no es la única opción para lograr una detección efectiva de suplantación de identidad. Los clasificadores de árboles de decisión y bosque aleatorio demostraron un mejor desempeño en términos de precisión y capacidad de detección.

3. Respuesta en tiempo real
   Después de llevar a cabo las pruebas de detección de suplantación facial utilizando diferentes bancos de imágenes, se ha comprobado que el método propuesto presenta un tiempo de ejecución promedio por imagen de 0.59 segundos, lo cual es considerado aceptable para su implementación en sistemas de tiempo real. Es importante destacar que los resultados se lograron sin la necesidad de utilizar hardware específico, lo que demuestra la viabilidad y eficacia del enfoque basado en visión artificial tradicional. Los hallazgos abren nuevas oportunidades de investigación en las cuales se podría suponer que, al agregar hardware especializado, los resultados podrían mejorar significativamente.

6.4 Trabajo futuro

Como trabajos futuros se tienen los siguientes puntos:

- Creación de un banco de imágenes con variación en la iluminación y resolución de una manera controlada para descubrir en que punto de iluminación la detección de suplantación es viable.
- Investigar cuales son las mejores características para el banco de entrenamiento el cual logre mejores resultados tanto en imágenes falsas como reales.
- Integración del método de detección de suplantación facial y el de verificación de identidad desarrollado en maestría. Actividad no contemplada en los alcances de este proyecto.

6.5 Productos académicos adicionales

- Presentación del artículo titulado "Estudio comparativo de la descripción con puntos faciales para reconocimiento del rostro" en la escuela de Inteligencia Artificial y Robótica 2019, realizada en la Universidad Tecnológica Emiliano Zapata (UTEZ), los días 22 y 25 de octubre.

- Presentación del artículo titulado "Suplantación de identidad en imágenes faciales" en el XII Seminario Internacional de Ciencias de la Computación, SICC 2019, realizado en la Universidad de Medellín Colombia, los días 23 y 25 de octubre.

- Presentación del artículo titulado "Comparación de OpenCV y DLIB para detección del rostro en ambientes no controlados" en la 3A jornada de Ciencia y Tecnología aplicada, realizada en el Centro Nacional de Investigación y Desarrollo Tecnológico, los días 14 y 15 de noviembre de 2019 [105].

- Aceptación del proyecto TecNM titulado "Detección de suplantación de identidad en imágenes faciales, para evaluaciones en línea" en la convocatoria 2020: proyectos de desarrollo tecnológico e innovación para estudiantes con una vigencia del 01 de enero de 2020 al 31 de diciembre de 2020.

- Publicación y presentación del artículo titulado "*Comparison of Gabor Filters and LBP descriptors applied to spoofing attack detection in facial images*" en el tercer Congreso Internacional de Informática aplicada ICAI 2020 que se llevó a cabo del 29 al 31 de octubre en Ota, Nigeria [157] y el cual se presentó de manera virtual. Además de ser publicado en el libro "*Applied Informatics*" de la editorial *Springer* con índice SCOPUS.

- Aceptación del proyecto TecNM titulado "Detección de suplantación de identidad en imágenes faciales, para evaluaciones en línea" en la convocatoria 2021: proyectos de desarrollo tecnológico e innovación para estudiantes con una vigencia del 01 de enero de 2021 al 31 de diciembre de 2021.

- Ponencia del tema "Suplantación de identidad en la educación a distancia" en las jornadas académicas de innovación, tecnología, liderazgo y sostenibilidad 2021 que se llevó a cabo el 28 y 29 de octubre en el Instituto Tecnológico de Zacatepec de forma virtual.

- Se realizó una ponencia con el tema “La importancia de la inteligencia artificial en la seguridad por medio del rostro” en el congreso nacional de innovación, tecnología, liderazgo y sostenibilidad 2022 que se llevó a cabo el 6 de abril en el Instituto Tecnológico de Zacatepec de forma virtual.

- Se publicó el artículo titulado "*Detection of facial spoofing attacks in uncontrolled environments using ELBP and color models*" en la revista *IEEE Latin America Transactions* indizada en el *Journal Citation Report* (JCR) [137].

Referencias

[1] S. Marcel, M. S. Nixon, J. Fierrez, y N. Evans, *Handbook of Biometric Anti-Spoofing: Presentation Attack Detection*. 2012. [En línea]. Disponible en: http://www.springerlink.com/index/10.1007/978-1-4471-2458-0

[2] S. Minaee, A. Abdolrashidi, H. Su, M. Bennamoun, y D. Zhang, "Biometrics recognition using deep learning: a survey", *Artif. Intell. Rev.*, pp. 1–49, 2023.

[3] E. Verbitskiy, P. Tuyls, D. Dentenner, y J. P. Linnartz, "Reliable Biometric Authentication with Privacy Protection", *Proc. 24th Benelux Symp. Inf. theory*, vol. 902, p. 19, 2023, doi: 10.1007/978-981-19-2004-2_21.

[4] S. Khairnar, S. Gite, K. Kotecha, y S. D. Thepade, "Face Liveness Detection Using Artificial Intelligence Techniques: A Systematic Literature Review and Future Directions", *Big Data Cogn. Comput.*, vol. 7, núm. 1, p. 37, 2023, doi: 10.3390/bdcc7010037.

[5] D. Vázquez, "Empleo de sistemas biométricos faciales aplicados al reconocimiento de personas en aeropuertos", Universidad rey Juan Carlos de España, 2006.

[6] J. Galbally, S. Marcel, y J. Fierrez, "Biometric antispoofing methods: A survey in face recognition", *IEEE Access*, vol. 2, pp. 1530–1552, 2014.

[7] L. Li, P. Lobato, y A. Hadid, "Face recognition under spoofing attacks: countermeasures and research directions", *IET Biometrics*, vol. 7, núm. 1, pp. 3–14, 2017, doi: 10.1049/iet-bmt.2017.0089.

[8] M. Romero, G. Figueroa, D. Vera, C. Alava, A. Murillo, y M. Castillo, *Introducción a la seguridad*, Primera ed. Alicante, 2018. doi: http://dx.doi.org/10.17993/IngyTec.2018.46.

[9] E. Aguilar, "Suplantación de la identidad digital con fines de trata de personas en facebook", INFOTEC Centro de investigación e innovación en tecnologías de la información y comunicación, 2019.

[10] A. García, I. García, y F. Guillen, "Retos y posibilidades del software de reconocimiento facial como herramienta de autenticación en los entornos virtuales de aprendizaje", *EDUTEC. Rev. Electrónica Tecnol. Educ.*, pp. 1–15, 2015.

[11] G. Quintanilla, "Legislation, risks and challenges of biometric systems", *Rev. Chil. Derecho y Tecnol.*, vol. 9, núm. 1, pp. 63–91, 2020, doi: 10.5354/0719-2584.2020.53965.

[12] F. Villavicencio, "Delitos Informáticos Cybercrimes", *IUS Verit.*, vol. 49, pp. 284–304, 2014.

[13] E. Raheem, S. Mumtazah, y W. Azizun, "Insight on face liveness detection: A systematic literature review", *Int. J. Electr. Comput. Eng.*, vol. 9, núm. 6, pp. 5165–5175, 2019, doi: 10.11591/ijece.v9i6.pp5165-5175.

[14] Kavita, G. Singh, y R. Rohilla, "A contemporary survey of unimodal liveness detection techniques: Challenges opportunities", *Third Int. Conf. Intell. Sustain. Syst.*, pp. 848–855, 2020, doi: 10.1109/ICISS49785.2020.9316059.

[15] S. Purnapatra *et al.*, "Face Liveness Detection Competition (LivDet-Face) - 2021", en *IEEE International conference on biometrics*, 2021, p. 10.

[16] Y. Xin *et al.*, "A survey of liveness detection methods for face biometric systems", *Sens. Rev.*, vol. 37, núm. 3, pp. 346–356, 2017, doi: 10.1108/SR-08-2015-0136.

[17] Z. Yu, J. Komulainen, X. Li, y G. Zhao, "Review of Face Presentation Attack Detection Competitions", *Adv. Comput. Vis. Pattern Recognit.*, pp. 287–336, 2023, doi: 10.1007/978-981-19-5288-3_12.

[18] *ISO/IEC 30107-1 and Biometrics (2016) Information Technology Biometric Presentation Attack Detection*. 2016.

[19] L. Souza, L. Oliveira, M. Pamplona, y J. Papa, "How far did we get in face spoofing detection?", *Engineering Applications of Artificial Intelligence*, vol. 72, pp. 368–381, 2018. doi: 10.1016/j.engappai.2018.04.013.

[20] R. Ramachandra y C. Busch, "Presentation Attack Detection Methods for Face Recognition Systems : A Comprehensive Survey", *ACM Comput. Surv.*, vol. 50, núm. 1, p. 8, 2017.

[21] Z. Boulkenafet, J. Komulainen, y A. Hadid, "Face Spoofing Detection Using Colour Texture Analysis", *IEEE Trans. Inf. Forensics Secur.*, vol. 11, núm. 8, pp. 1818–1830, 2016, doi: 10.1109/TIFS.2016.2555286.

[22] T. Edmunds y A. Caplier, "Motion-based countermeasure against photo and video spoofing attacks in face recognition", *J. Vis. Commun. Image Represent.*, vol. 50, pp. 314–332, 2018, doi: 10.1016/j.jvcir.2017.12.004.

[23] L. Li, X. Feng, Z. Xia, X. Jiang, y A. Hadid, "Face spoofing detection with local binary pattern network", *J. Vis. Commun. Image Represent.*, vol. 54, pp. 182–192, 2018, doi: 10.1016/j.jvcir.2018.05.009.

[24] V. Patel, N. Ratha, y R. Chellappa, "Cancelable Biometrics: A review", *IEEE Signal Process. Mag.*, vol. 32, pp. 54–65, 2015, doi: 10.1109/MSP.2015.2434151.

[25] *ISO/IEC DIS 30107-3 and biometrics (2017) Part 3- Information Technology Biometric Presentation Attack Detection*. 2017.

[26] Y. Atoum, L. Chen, A. Liu, S. Hsu, y X. Liu, "Automated Online Exam Proctoring", *IEEE Trans. Multimed.*, vol. 19, núm. 7, pp. 1609–1624, 2017.

[27] R. Bawarith, A. Basuhail, A. Fattouh, y S. Gamalel-din, "E-exam Cheating Detection System", *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, núm. 4, pp. 176–181, 2017.

[28] S. Ketab, N. Clarke, y P. Dowland, "The Value of the Biometrics in Invigilated E-Assessments", *EDULEARN16 Proc.*, vol. 1, núm. July, pp. 7648–7658, 2016, doi: 10.21125/edulearn.2016.0687.

[29] A. Dewan, M. Murshed, y F. Lin, "Engagement detection in online learning : a review", *Smart Learn. Environ.*, pp. 1–20, 2019.

[30] C. Aravena, D. Pasmino, J. Tapia, y C. Busch, "Impact of Face Image Quality Estimation on Presentation Attack Detection", *arXiv2209.15489, 2022*, pp. 1–9, 2022, [En línea]. Disponible en: http://arxiv.org/abs/2209.15489

[31] Z. Ming, M. Visani, M. Muzzamil, y J.-C. Burie, "A survey on anti-spoofing methods for facial recognition with rgb cameras of generic consumer devices", *J. Imaging*, vol. 6, núm. 12, 2020, doi: 10.3390/jimaging6120139.

[32] J. Valera, J. Valera, y Y. Gelogo, "A Review on Facial Recognition for Online Learning Authentication", *2015 8th Int. Conf. Bio-Science Bio-Technology*, pp. 16–19, 2015, doi: 10.1109/BSBT.2015.15.

[33] S. Rodchua, G. Yiadom-Boakye, y D. R. Woolsey, "Student Verification System for Online Assessments: Bolstering Quality and Integrity of Distance Learning", *J. Ind. Technol.*, vol. 27, núm. 3, 2011.

[34] L. Yu y K. Li, "Application of Face Recognition Technology in the Exam Identity Authentication System", *DEStech Trans. Soc. Sci. Educ. Hum. Sci.*, pp. 684–688, 2017.

[35] Z. Khanam y M. N. Ahsan, "Implementation of the pHash algorithm for face recognition in a secured remote online examination system", *Comput. Sci.*, vol. 4, núm. 11, pp. 1–5, 2018, doi: 10.31695/IJASRE.2018.32917.

[36] A. K. Singh y A. Mohan, *Handbook of multimedia information security: Techniques and applications*. 2019. doi: 10.1007/978-3-030-15887-3.

[37] J. Galbally, S. Marcel, y J. Fierrez, "Image quality assessment for fake biometric detection: Application to Iris, fingerprint, and face recognition", *IEEE Trans. Image Process.*, vol. 23, núm. 2, pp. 710–724, 2014, doi: 10.1109/TIP.2013.2292332.

[38] K. Gates, "The Past Perfect Promise of Facial Recognition Technology", *ACDIS Occas. Pap.*, vol. 21, núm. 2, pp. 125–125, 2004, doi: 10.1192/pb.21.2.125.

[39] R. Gonzalez y R. Woods, *Digital Image Processing*, Second edi., vol. 44, núm. 8. 1987. doi: 10.1088/1751-8113/44/8/085201.

[40] T. Ojala, M. Pietikäinen, y T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, núm. 7, pp. 971–987, 2002, doi: 10.1109/TPAMI.2002.1017623.

[41] A. Kartika, I. B. Kusuma, T. Agung, B. Wirayuda, y K. Nur, "Image Spoofing Detection Using Local Binary Pattern and Local Binary Pattern Variance", *Int. J. Inf. Commun. Technol.*, vol. 4, núm. December, pp. 11–18, 2019, doi: 10.21108/indojc.2018.42.134.

[42] S. A. Angadi y V. C. Kagawade, "Detection of Face Spoofing using Multiple Texture Descriptors", en *International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, 2019, pp. 151–156. doi: 10.1109/ctems.2018.8769129.

[43] J. Kannala y E. Rahtu, "BSIF: Binarized statistical image features", *Proc. - Int. Conf. Pattern Recognit.*, pp. 1363–1366, 2012.

[44] S. Rahimzadeh y J. Kittler, "Face Spoofing Detection Based on Multiple Descriptor Fusion Using Multiscale Dynamic Binarized Statistical Image Features", *IEEE Trans. Inf. Forensics Secur.*, vol. 10, núm. 11, pp. 2396–240, 2015.

[45] Z. Boulkenafet, J. Komulainen, y A. Hadid, "On the generalization of color texture-based face anti-spoofing", *Image Vis. Comput.*, vol. 77, pp. 1–9, 2018, doi: 10.1016/j.imavis.2018.04.007.

[46] J. Luká, J. Fridrich, y M. Goljan, "Digital Camera Identification From Sensor Pattern Noise", *IEEE Trans. Inf. Forensics Secur. Forensics Secur.*, vol. 1, núm. 2, pp. 205–214, 2006.

[47] A.-T. Phan-Ho y F. Retraint, "A Comparative Study of Bayesian and Dempster-Shafer Fusion on Image Forgery Detection", *IEEE Access*, vol. 10, pp. 99268–99281, 2022, doi: 10.1109/ACCESS.2022.3206543.

[48] Z. Ali y U. Park, "Face Spoofing Attack Detection Using Spatial Frequency and Gradient-Based Descriptor", *KSII Transactions on Internet and Information Systems*, vol. 13, núm. 2, pp. 892–911, 2019.

[49] A. Tsitiridis, C. Conde, B. Gomez, y E. Cabello, "Bio-Inspired Presentation Attack Detection for Face Biometrics", *Front. Comput. Neurosci.*, vol. 13, núm. May, pp. 1–17, 2019, doi: 10.3389/fncom.2019.00034.

[50] L. Song y H. Ma, "Face Liveliness Detection Based on Texture and Color Features", *2019 IEEE 4th Int. Conf. Cloud Comput. Big Data Anal.*, pp. 418–422, 2019.

[51] A. Bakshi, S. Gupta, A. Gupta, S. Tanwar, y K. F. Hsiao, "3T-FASDM: Linear discriminant analysis-based three-tier face anti-spoofing detection model using support vector machine", *Int. J. Commun. Syst.*, vol. 33, núm. 12, pp. 1–22, 2020, doi: 10.1002/dac.4441.

[52] C. Yao, Y. Jia, H. Di, y Y. Wu, "Face Spoofing Detection Using Relativity Representation on Riemannian Manifold", *IEEE Trans. Inf. Forensics Secur.*, vol. 15, núm. c, pp. 3683–3693, 2020, doi: 10.1109/TIFS.2020.2998956.

[53] B. Chen, W. Yang, y S. Wang, "Generalized Face Anti-spoofing by Learning to Fuse Features from High and Low Frequency Domains", *IEEE Multimed.*, núm. c, pp. 1–1, 2021, doi: 10.1109/mmul.2021.3053698.

[54] H.-H. Chang y C.-H. Yeh, "Face anti-spoofing detection based on multi-scale image quality assessment", *Image Vis. Comput.*, vol. 121, p. 104428, 2022, doi: 10.1016/j.imavis.2022.104428.

[55] Z. Wang, Q. Wang, W. Deng, y G. Guo, "Face Anti-Spoofing Using Transformers with Relation-Aware Mechanism", *IEEE Trans. Biometrics, Behav. Identity Sci.*, vol. 4, núm. 3, pp. 439–450, 2022, doi: 10.1109/TBIOM.2022.3184500.

[56] C. Wang, B. Yu, y J. Zhou, "A Learnable Gradient operator for face presentation attack detection", *Pattern Recognit.*, vol. 135, p. 109146, 2023, doi: 10.1016/j.patcog.2022.109146.

[57] D. Gragnaniello, G. Poggi, C. Sansone, y L. Verdoliva, "An Investigation of Local Descriptors for Biometric Spoofing Detection", *IEEE Trans. Inf. Forensics Secur.*, vol. 10, núm. 4, pp. 849–863, 2015, doi: 10.1109/TIFS.2015.2404294.

[58] S. Zafeiriou, C. Zhang, y Z. Zhang, "A survey on face detection in the wild: Past, present and future", *Comput. Vis. Image Underst.*, vol. 138, núm. March, pp. 1–24, 2015, doi: 10.1016/j.cviu.2015.03.015.

[59] H. Chen, G. Hu, Z. Lei, Y. Chen, N. M. Robertson, y S. Z. Li, "Attention-Based Two-Stream Convolutional Networks for Face Spoofing Detection", *IEEE Trans. Inf. Forensics Secur.*, vol. PP, núm. 4, p. 1, 2019, doi: 10.1109/TIFS.2019.2922241.

[60] X. Qu, J. Dong, y S. Niu, "shallowCNN-LE : A shallow CNN with Laplacian Embedding for face", en *14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, 2019, pp. 1–8.

[61] F. Zhou, C. Gao, F. Chen, C. Li, y X. Li, "Face anti-spoofing based on multi-layer domain adaptation", en *IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, 2019, pp. 192–197. doi: 10.1109/ICMEW.2019.00-88.

[62] B. Lin, Z. Yu, X. Li, y G. Zhao, "Face Liveness Detection by rPPG Features and Contextual Patch-Based CNN", en *3rd International Conference on Biometric Engineering and Applications*, 2019, pp. 61–68.

[63] R. Koshy y A. Mahmood, "Optimizing Deep CNN Architectures for Face Liveness Detection", *Entropy*, vol. 21, p. 423, 2019, doi: https://doi.org/10.3390/e21040423.

[64] O. Nikisins, A. George, y S. Marcel, "Domain Adaptation in Multi-Channel

Autoencoder based Features for Robust Face Anti-Spoofing", en *International Conference on Biometrics 2019, IEEE*, 2019, pp. 1–8.

[65] M. Wang y W. Deng, "Deep face recognition: A survey", *Neurocomputing*, vol. 429, pp. 215–244, 2021, doi: 10.1016/j.neucom.2020.10.081.

[66] D. Bhatt *et al.*, "Cnn variants for computer vision: History, architecture, application, challenges and future scope", *Electron.*, vol. 10, núm. 20, pp. 1–28, 2021, doi: 10.3390/electronics10202470.

[67] W. Jingying, "A Survey on Crowd Counting Methods and Datasets", *Adv. Intell. Syst. Comput.*, vol. 1158, pp. 851–863, 2021, doi: 10.1007/978-981-15-4409-5_76.

[68] J. Sirignano, "Deep learning for limit order books", *Quant. Financ.*, vol. 19, núm. 4, pp. 549–570, 2019, doi: 10.1080/14697688.2018.1546053.

[69] K. Asanovic *et al.*, "Training Deep Neural Networks on GPUs", *Commun. ACM*, vol. 49, núm. 4, pp. 109–119, 2006.

[70] L. Wang *et al.*, "SuperNeurons: Dynamic GPU Memory Management for Training Deep Neural Networks", *ACM SIGPLAN Not.*, vol. 53, núm. 1, pp. 41–53, 2018, doi: 10.1145/3178487.3178491.

[71] K. Simonyan y A. Zisserman, "Very deep convolutional networks for large-scale image recognition", *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–14, 2015.

[72] C. Zhang, S. Bengio, M. Hardt, B. Recht, y O. Vinyals, "Understanding deep learning (still) requires rethinking generalization", *Commun. ACM*, vol. 64, núm. 3, pp. 107–115, 2021, doi: 10.1145/3446776.

[73] C. Molnar, *Interpretable Machine Learning. A Guide for Making Black Box Models Explainable.* 2020.

[74] B. Chen, W. Yang, H. Li, S. Wang, y S. Kwong, "Camera Invariant Feature Learning for Generalized Face Anti-spoofing", *IEEE Trans. Inf. Forensics Secur.*, vol. 6013, núm. c, pp. 1–1, 2021, doi: 10.1109/tifs.2021.3055018.

[75] K. Y. Zhang *et al.*, "Structure destruction and content combination for face anti-spoofing", *IEEE Int. Jt. Conf. Biometrics, IJCB*, 2021, doi: 10.1109/IJCB52358.2021.9484395.

[76] Z. Yu, X. Li, J. Shi, Z. Xia, y G. Zhao, "Revisiting Pixel-Wise Supervision for Face Anti-Spoofing", *IEEE Trans. Biometrics, Behav. Identity Sci.*, vol. 3, núm. 3, pp. 285–295, 2021, doi: 10.1109/tbiom.2021.3065526.

[77] S. Liu *et al.*, "Adaptive Normalized Representation Learning for Generalizable Face Anti-Spoofing", *MM 2021 - Proc. 29th ACM Int. Conf. Multimed.*, pp. 1469–1477, 2021, doi: 10.1145/3474085.3475279.

[78] Y. Wang, X. Song, T. Xu, Z. Feng, y X. J. Wu, "From RGB to Depth: Domain Transfer Network for Face Anti-Spoofing", *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 4280–4290, 2021, doi: 10.1109/TIFS.2021.3102448.

[79] C. Hu, J. Cao, K. Y. Zhang, T. Yao, S. Ding, y L. Ma, "Structure Destruction and Content Combination for Generalizable Anti-Spoofing", *IEEE Trans. Biometrics, Behav. Identity Sci.*, vol. 4, núm. 4, pp. 508–521, 2022, doi: 10.1109/TBIOM.2022.3220406.

[80] Apple, "iPhone X", 2022. https://support.apple.com/es-lamr/HT208108

[81] Samsung, "Samsung Galaxy S10", 2019. https://news.samsung.com/cl/10-caracteristicas-que-mejoraran-tu-experiencia-con-el-galaxy-s10

[82] Google, "Face Unlock", 2012. http://www.android.com/ about/ice-cream-sandwich/

[83] Huawei, "Huawei Mate 20 Pro", 2018. https://consumer.huawei.com/latin/phones/mate20-pro/specs/

[84] J. Sanchez del Rio, D. Moctezuma, C. Conde, I. Martin de Diego, y E. Cabello, "Automated border control e-gates and facial recognition systems", *Comput. Secur.*, vol. 62, pp. 49–72, 2016, doi: 10.1016/j.cose.2016.07.001.

[85] "FACT'S 4", *Dahua Technology*. https://www.dahuasecurity.com/sa/products/All-Products/Access-Control--Time-Attendance

[86] "Hikvision", *Hangzhou Hikvision Digital Technology*. https://www.hikvision.com/en/products/Access-Control-Products/Face-Recognition-Terminals/

[87] "Axis Communications". https://help.axis.com/en-us/axis-camera-station-system-hardening-guide

[88] "NEC". https://www.nec.com/en/press/201902/global_20190225_01.html

[89] B. Pad, P. A. Detection, P. A. Detection, y N. Nvlap, "IDEMIA's facial recognition device VisionPass gets best results from iBeta antispoofing evaluation", *IDEMIA*, pp. 2–3, 2021.

[90] "Suprema". https://www.supremainc.com/en/solutions/facial-recognition-system.asp

[91] "BioCatch". https://www.biocatch.com/behavioral-biometrics-solution?utm_term=biocatch&utm_source=adwords&utm_medium=ppc&utm_campaign=Biocatch-Brand-Search-LATAM-ENG&utm_content=g&hsa_acc=8099910299&hsa_net=adwords&hsa_cam=19266156640&hsa_tgt=kwd-384804394534&hsa_kw=b

[92] "Featurespace". https://www.featurespace.com/solutions/

[93] S. Shan, W. Gao, B. Cao, y D. Zhao, "Illumination normalization for robust face recognition against varying lighting conditions", *IEEE Int. Work. Anal. Model. Faces Gestures, AMFG 2003*, núm. January 2015, pp. 157–164, 2003, doi: 10.1109/amfg.2003.1240838.

[94] B. Peixoto, C. Michelassi, y A. Rocha, "Face liveness detection under bad illumination conditions", *Proc. - Int. Conf. Image Process. ICIP*, núm. September 2011, pp. 3557–3560, 2011, doi: 10.1109/ICIP.2011.6116484.

[95] A. Weeks, "Histogram modification for contrast enhancement", *Comput. Graph. Image Process.*, vol. 9, núm. 4, pp. 366–374, 1979.

[96] E. H. Land y J. J. McCann, "Lightness and retinex theory.", *J. Opt. Soc. Am.*, vol. 61, núm. 1, pp. 1–11, 1971, doi: 10.1364/JOSA.61.000001.

[97] M. A. Ochoa-villegas, J. A. Nolazco-flores, O. Barron-cano, y I. A. Kakadiaris, "Addressing the illumination challenge in two- dimensional face recognition : a survey", *IET Computer Vision*, vol. 9, pp. 978–992, 2015. doi: 10.1049/iet-cvi.2014.0086.

[98] J. H. Shah, M. Sharif, M. Raza, y M. Murtaza, "Robust Face Recognition Technique under Varying Illumination", *Journal of Applied Research and Technology*, vol. 13, núm. 1, pp. 97–105, 2015. doi: 10.1016/S1665-6423(15)30008-0.

[99] A. Lumini, L. Nanni, y S. Brahnam, "Ensemble of texture descriptors and classifiers for face recognition", *Comput. Informatics*, vol. 13, núm. 1, pp. 79–91,

2017, doi: 10.1016/j.aci.2016.04.001.

[100] F. Juefei-xu y M. Savvides, "Encoding and Decoding Local Binary Patterns for Harsh Face Illumination Normalization", *IEEE Int. Conf. Image Process.*, pp. 3220–3224, 2015, doi: 10.1109/ICIP.2015.7351398.

[101] S. Parthasarathy y P. Sankaran, "An automated multi scale Retinex with color restoration for image enhancement", *2012 Natl. Conf. Commun. NCC 2012*, núm. February, 2012, doi: 10.1109/NCC.2012.6176791.

[102] D. J. Jobson, Z. U. Rahman, y G. A. Woodell, "Properties and performance of a center/surround retinex", *IEEE Trans. Image Process.*, vol. 6, núm. 3, pp. 451–462, 1997, doi: 10.1109/83.557356.

[103] A. B. Petro, C. Sbert, y J.-M. Morel, "Multiscale Retinex", *Image Process. Line*, vol. 4, pp. 71–88, 2014, doi: 10.5201/ipol.2014.107.

[104] D. J. Jobson, "Retinex processing for automatic image enhancement", *J. Electron. Imaging*, vol. 13, núm. 1, p. 100, 2004, doi: 10.1117/1.1636183.

[105] W. Valderrama y A. Magadán, "Comparación de OpenCV y DLIB para detección del rostro en ambientes no controlados", *Jorn. Cienc. y Tecnol. Apl.*, vol. 2, pp. 22–27, 2019.

[106] P. Viola y M. Jones, "Rapid object detection using a boosted cascade of simple features", *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, núm. July 2014, pp. I-511-I–518, 2005, doi: 10.1109/cvpr.2001.990517.

[107] P. Wilson y J. Fernandez, "Facial feature detection using Haar classifiers", *J. Comput. Sci. Coll.*, vol. 21, núm. 4, pp. 127–133, 2006.

[108] C. Intel, "The OpenCV Tutorials 2.3", 2011.

[109] N. Dalal y B. Triggs, "Histograms of Oriented Gradients for Human Detection", *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 886–893, 2010.

[110] D. E. King, "Dlibml: A Machine Learning Toolkit", *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009. doi: 10.1145/1577069.1755843.

[111] R. Lienhart y J. Maydt, "An Extended Set of Haar-like Features for Rapid Object Detection", *Proceedings. Int. Conf. image Process.*, vol. 1, pp. 0–3, 2002.

[112] Z. Boulkenafet, J. Komulainen, y A. Hadid, "Face anti-spoofing based on color texture analysis", *Proc. - Int. Conf. Image Process. ICIP*, vol. 2015-Decem, pp. 2636–2640, 2015, doi: 10.1109/ICIP.2015.7351280.

[113] Z. Boulkenafet, J. Komulainen, y A. Hadid, "Face Spoofing Detection Using Colour Texture Analysis", *IEEE Trans. Inf. Forensics Secur.*, vol. 11, núm. 8, pp. 1818–1830, 2016, doi: 10.1109/TIFS.2016.2555286.

[114] J. He y J. Luo, "Face Spoofing Detection Based on Combining Different Color Space Models", *2019 IEEE 4th Int. Conf. Image, Vis. Comput. ICIVC 2019*, pp. 523–528, 2019, doi: 10.1109/ICIVC47709.2019.8981232.

[115] Z. Boulkenafet, J. Komulainen, y A. Hadid, "Face anti-spoofing based on color texture analysis", *2015 IEEE Int. Conf. image Process.*, pp. 2636–2640, 2015.

[116] J. R. Smith y S. F. Chang, "Single color extraction and image query", *IEEE Int. Conf. Image Process.*, vol. 3, pp. 528–531, 1995, doi: 10.1109/icip.1995.537688.

[117] K. Plataniotis, "Color image processing and applications", *Springer Sci. Bus. Media*, 2000.

[118] L. Liu, S. Lao, P. W. Fieguth, Y. Guo, X. Wang, y M. Pietikäinen, "Median Robust Extended Local Binary Pattern for Texture Classification", *IEEE Trans. Image*

Process., vol. 25, núm. 3, pp. 1368–1381, 2016, doi: 10.1109/TIP.2016.2522378.

[119] D. Huang, C. Shan, M. Ardabilian, Y. Wang, y L. Chen, "Local binary patterns and its application to facial image analysis: A survey", *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 41, núm. 6, pp. 765–781, 2011, doi: 10.1109/TSMCC.2011.2118750.

[120] R. Raghavendra y C. Busch, "Texture based features for robust palmprint recognition: a comparative study", *Eurasip J. Inf. Secur.*, vol. 2015, núm. 1, 2015, doi: 10.1186/s13635-015-0022-z.

[121] G. C. Holst, *CCD arrays, cameras, and displays, 2nd ed.*, 2nd ed. WA:JCD & SPIE, 1998.

[122] J. Janesick, *Scientific Charge-Coupled devices*. 2001.

[123] P. Zhang *et al.*, "FeatherNets : Convolutional Neural Networks as Light as Feather for Face Anti-spoofing", *IEEE Conf. Comput. Vis. Pattern Recognit. Work.*, 2019.

[124] Y. Lecun, Y. Bengio, y G. Hinton, "Deep learning", *Nature*, vol. 521, núm. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[125] R. Morcel *et al.*, "Feathernet: An accelerated convolutional neural network design for resource-constrained FPGAs", *ACM Trans. Reconfigurable Technol. Syst.*, vol. 12, núm. 2, 2019, doi: 10.1145/3306202.

[126] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, y L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks", *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.

[127] G. Zeng, Y. He, Z. Yu, X. Yang, R. Yang, y L. Zhang, "Preparation of novel high copper ions removal membranes by embedding organosilane-functionalized multi-walled carbon nanotube", *J. Chem. Technol. Biotechnol.*, vol. 91, núm. 8, pp. 2322–2330, 2016, doi: 10.1002/jctb.4820.

[128] N. Ma, X. Zhang, H. T. Zheng, y J. Sun, "Shufflenet V2: Practical guidelines for efficient cnn architecture design", *Comput. Sci.*, vol. 11218 LNCS, pp. 122–138, 2018, doi: 10.1007/978-3-030-01264-9_8.

[129] T. He *et al.*, "Bag of Tricks for Image Classification with Convolutional Neural Networks", *Comput. Vis.*, pp. 558–567, 2021.

[130] J. Wan, G. Guo, S. Escalera, H. J. Escalante, y S. Z. Li, *Multi-Modal Face Presentation Attack Detection*, vol. 9, núm. 1. 2020. doi: 10.2200/s01032ed1v01y202007cov017.

[131] V. Vapnik, S. E. Golowich, y A. Smola, "Support vector method for function approximation, regression estimation, and signal processing", *Adv. Neural Inf. Process. Syst.*, pp. 281–287, 1997.

[132] P. Van Ngoan y L. H. Trang, *Deep and Wide Features for Face Anti-Spoofing*, vol. 1, núm. 1. Association for Computing Machinery, 2022. doi: 10.1145/3561801.3561808.

[133] Y. C. Wang, C. Y. Wang, y S. H. Lai, "Disentangled Representation with Dual-stage Feature Learning for Face Anti-spoofing", *Proc. - 2022 IEEE/CVF Winter Conf. Appl. Comput. Vision, WACV 2022*, pp. 1234–1243, 2022, doi: 10.1109/WACV51458.2022.00130.

[134] L. Breiman, "Random forests", *Mach. Learn.*, pp. 5–32, 2001, doi:

10.1023/A:1010933404324.

[135] Y. Zhang, R. K. Dubey, G. Hua, y V. L. L. Thing, "Face Spoofing Video Detection Using Spatio-Temporal Statistical Binary Pattern", *TENCON 2018 - 2018 IEEE Reg. 10 Conf.*, núm. October, pp. 309–314, 2018.

[136] D. E. Rumelhart, G. E. Hinton, y R. J. Williams, "Learning representations by back-propagating errors", *Nature*, vol. 323, núm. 6088, pp. 533–536, 1986, doi: 10.1038/323533a0.

[137] W. Valderrama, A. Magadan, O. Vergara, J. Ruiz, R. Pinto, y G. Reyes, "Detection of Facial Spoofing Attacks in Uncontrolled Environments Using ELBP and Color Models", *IEEE Lat. Am. Trans.*, vol. 20, núm. 6, pp. 875–883, 2022.

[138] L. Breiman, *Classification and Regression Trees*, 1st Editio. 1984. doi: https://doi.org/10.1201/9781315139470.

[139] X. Zhu, T. Hua, F. Yang, G. Tu, y X. Chen, "Global positioning system spoofing detection based on Support Vector Machines", *IET Radar, Sonar Navig.*, vol. 16, núm. 2, pp. 224–237, 2022, doi: 10.1049/rsn2.12178.

[140] D. Basso, "Propuesta de Métricas para Proyectos de Explotación de Información", *Rev. Latinoam. Ing. Softw.*, vol. 2, núm. 4, p. 157, 2015, doi: 10.18294/relais.2014.157-218.

[141] C. Busch, *Standards for biometric presentation attack detection*. Springer International Publishing, 2019. doi: 10.1007/978-3-319-92627-8_22.

[142] I. Chingovska, A. R. Dos Anjos, y S. Marcel, "Biometrics evaluation under spoofing attacks", *IEEE Trans. Inf. Forensics Secur.*, vol. 9, núm. 12, pp. 2264–2276, 2014, doi: 10.1109/TIFS.2014.2349158.

[143] Microsoft Corp., "Visual Studio", *Visualstudio.Com*, 2015. https://www.visualstudio.com/

[144] Intel, "OpenCV", 2017. https://www.opencv.org/

[145] D. King, "dlib C++ Library", 2015. Www.Dlib.Net

[146] J. L. and L. J. X.Tan, Y.Li, "Face Liveness Detection from A Single Image with Sparse Low Rank Bilinear Discriminative Model", *Proc. 11th Eur. Conf. Comput. Vis.*, 2010.

[147] A. Anjos y S. Marcel, "Counter-Measures to Photo Attacks in Face Recognition: a public database and a baseline", *Int. Jt. Conf. Biometrics*, pp. 1–7, 2011.

[148] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, y S. Z. Li, "A Face Antispoofing Database with Diverse Attacks", *2012 5th IAPR Int. Conf. Biometrics*, pp. 26–31, 2012.

[149] I. Chingovska, A. Anjos, y S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing", *Proc. Int. Conf. Biometrics Spec. Interes. Group, BIOSIG 2012*, pp. 183–194, 2012.

[150] D. Wen, H. Han, y A. K. Jain, "Face spoof detection with image distortion analysis", *IEEE Trans. Inf. Forensics Secur.*, vol. 10, núm. 4, pp. 746–761, 2015, doi: 10.1109/TIFS.2015.2400395.

[151] Z. Boulkenafet, J. Komulainen, X. Feng, y A. Hadid, "OULU-NPU : A mobile face presentation attack database with real-world variations", *2017 12th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG 2017)*, pp. 612–618, 2017, doi: 10.1109/FG.2017.77.

[152] A. Costa-pazo, S. Bhattacharjee, E. Vazquez-fernandez, y S. Marcel, "The REPLAY-MOBILE Face Presentation-Attack Database", 1920.

[153] D. Timoshenko, K. Simonchik, V. Shutov, P. Zhelezneva, y V. Grishkin, "Large Crowdcollected Facial Anti-Spoofing Dataset", *12th Int. Conf. Comput. Sci. Inf. Technol. CSIT 2019*, pp. 123–126, 2019, doi: 10.1109/CSITechnol.2019.8895208.

[154] D. C. E. Thomaz, "FEI Face Database", *Image Processing Laboratory Department of Electrical Engineering Centro Universitario da FEI, São Bernardo do Campo, São Paulo, Brazil*, 2012. https://fei.edu.br/~cet/facedatabase.html

[155] A. Gupta, "Human Face", *kaggle*, 2020. https://www.kaggle.com/datasets/ashwingupta3012/human-faces

[156] A. Villanueva, V. Ponz, L. Sesma-Sanchez, M. Ariz, S. Porta, y R. Cabeza, "Hybrid method based on topography for robust detection of iris center and eye corners", *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 9, núm. 4, pp. 1–20, 2013, doi: 10.1145/2501643.2501647.

[157] W. Valderrama, A. Magadán, R. Pinto, y J. Ruiz, "Comparison of Gabor Filters and LBP Descriptors Applied to Spoofing Attack Detection in Facial Images", *Int. Conf. Appl. Informatics*, pp. 395–408, 2020, doi: 10.1007/978-3-030-61702-8_27.

Anexos

A. Resultados de los métodos desarrollados en el transcurso del doctorado

En la Tabla A.1, Tabla A.2 y Tabla A.3 se observan los resultados obtenidos con los diferentes métodos desarrollados en el transcurso del doctorado, con rojo se marcan los resultados con porcentajes superiores el 50%, con anaranjado los resultados entre el menor y mayor porcentaje y con color verde los porcentajes más bajos, visualmente se pude observar la evolución de los resultados en donde se muestra que con el método final(Multiescale retinex+Rostro+HSV+YCbCr+PRNU) se obtienes mejores resultados en comparación con los otros.

Tabla A.1 Resultados con la métrica HTER de los métodos desarrollados en el transcurso del doctorado expresado en porcentajes.

Banco de imágenes	NUAA		CASIA		LCC FASD	
Método	SVC	CVC	SVC	CVC	SVC	CVC
Imagen completa+YCbCr+HSV+LBP	37.13	34.07	44.81	49.26	49.79	51.67
Imagen completa+YCbCr+HSV+ELBP$_{16,2}$	50	42.24	45.19	55.93	46.59	59.39
Rostro+FeatherNetB	48.32	48.32	50.95	50.54	44.46	44.46
YCbCr+HSV+Rostro+ FeatherNetB	55.82	56.33	39.98	39.98	58.9	58.9
VDLM (Visión, Deep Learning y detección de Movimiento)	54.57	56.33	38.1	39.98	59.58	58.9
MRYHB (*Multiscale Retinex*, Frente+YCbCr, HSV y BSIF)	50	29.69	40.7	54.42	49.67	50.09
Multiescale retinex+Rostro+HSV+YCbCr+PRNU	47.01	49.94	48.15	51.85	41.17	42

Tabla A.2 Resultados con la métrica APCER de los métodos desarrollados en el transcurso del doctorado expresado en porcentajes.

Banco de imágenes	NUAA		CASIA		LCC FASD	
Método	SVC	CVC	SVC	CVC	SVC	CVC
Imagen completa+YCbCr+HSV+LBP	24.7	15.56	51.11	94.81	37.97	53.82
Imagen completa+YCbCr+HSV+ELBP$_{16,2}$	0	15.93	41.48	48.89	12.66	52.71
Rostro+FeatherNetB	65.77	61.13	95.56	87.8	75.94	75.58
YCbCr+HSV+Rostro+ FeatherNetB	82.55	82.15	47.15	47.1	78.61	78.6
VDLM (Visión, Deep Learning y detección de Movimiento)	82.05	82.55	47.08	47.15	78.35	78.61
MRYHB (*Multiscale Retinex*, Frente+YCbCr, HSV y BSIF)	100	13.4	57.26	73.5	0.59	21.94
Multiescale retinex+Rostro+HSV+YCbCr+PRNU	9.97	8.03	38.52	12.59	21.56	26.15

Tabla A.3 Resultados con la métrica BPCER de los métodos desarrollados en el transcurso del doctorado expresado en porcentajes.

Banco de imágenes	**NUAA**		**CASIA**		**LCC FASD**	
Método	**SVC**	**CVC**	**SVC**	**CVC**	**SVC**	**CVC**
Imagen completa+YCbCr+HSV+LBP	49.57	52.57	38.52	3.7	61.61	49.51
Imagen completa+YCbCr+HSV+ELBP$_{16,2}$	100	68.54	48.89	62.96	80.53	66.06
Rostro+FeatherNetB	35.51	31.5	13.28	18.3	13.34	14.3
YCbCr+HSV+Rostro+ FeatherNetB	30.1	30.5	32.81	31.8	39.19	39
VDLM (Visión, Deep Learning y detección de Movimiento)	31	32.1	38.1	38.2	31.9	32.8
MRYHB (*Multiscale Retinex*, Frente+YCbCr, HSV y BSIF)	0	45.92	24.14	35.34	98.74	78.23
Multiescale retinex+Rostro+HSV+YCbCr+PRNU	84.05	91.85	57.78	91.11	60.78	57.86