Universidad Autónoma de Entre Ríos

Facultad de Ciencia y Tecnología

-Oro Verde-

LICENCIATURA EN SISTEMAS DE INFORMACIÓN

TESINA DE GRADO 2025

Propuesta de requisitos para sistemas de detección de Deepfake en aplicaciones que utilicen autenticación biométrica facial

DATOS DE AUTOR	Ernst Gotte, Bernardita
DATOS DE AUTOR	Basgall, Facundo Tomás
DATOS DE DIRECTOR	Cristian Borghello

RESUMEN

El uso de autenticación biométrica facial se ha expandido rápidamente en dispositivos y sistemas de acceso debido a su capacidad para ofrecer validación precisa, sin contacto y fácil de usar. No obstante, esta tecnología enfrenta una amenaza creciente: los *deepfakes*, capaces de generar imágenes y videos sintéticos que simulan de forma realista el rostro de una persona, comprometiendo los mecanismos de verificación existentes.

A lo largo de este trabajo se analiza cómo la autenticación biométrica facial, sus fundamentos técnicos y su vinculación con la identidad digital se ven afectados por esta tecnología emergente. A partir de una revisión integral de la tecnología *deepfake*, sus procesos de creación, aplicaciones y riesgos, se estudian los principales sistemas de detección actuales y sus limitaciones frente a ataques cada vez más sofisticados.

Como respuesta a esta problemática, se propone un conjunto de requisitos clave para el diseño de sistemas de detección de *deepfakes* integrados en contextos de autenticación biométrica facial. Estos requisitos —agrupados en dimensiones técnicas, operacionales, de seguridad y privacidad, legalidad— buscan orientar el desarrollo de soluciones más robustas, adaptables y alineadas con la aplicación práctica de esta tecnología.

PALABRAS CLAVE: Deepfake, biometría, autenticación biométrica, deep learning, seguridad, suplantación de identidad, inteligencia artificial.

CONTENIDO

INTRODUCCIÓN

En un mundo cada vez más digitalizado, la seguridad en el acceso a sistemas y plataformas es esencial para proteger la información sensible de individuos y organizaciones. Las tecnologías de reconocimiento facial han ganado popularidad debido a su efectividad en la validación de identidades de manera única e inviolable. Sin embargo, con el rápido avance de técnicas de inteligencia artificial, como los *deepfakes*, se ha generado una nueva amenaza para la integridad de estos sistemas. Los *deepfakes*, mediante la generación de contenido artificial que replica rasgos biométricos, permiten suplantar identidades y evadir mecanismos de seguridad como las pruebas de vida, comprometiendo la fiabilidad de los sistemas de autenticación biométrica facial.

En este contexto, es fundamental entender cómo los ataques mediante *deepfake* pueden comprometer la seguridad de las aplicaciones biométricas y qué medidas deben tomarse para mitigar estos riesgos. La presente investigación se enfoca en analizar los sistemas de detección de *deepfake* en aplicaciones que utilizan reconocimiento facial, con el fin de plantear una serie de requisitos que fortalezcan su seguridad, mejoren su operatividad y aseguren el cumplimiento de normativas legales y éticas.

La intersección entre las tecnologías de validación biométrica facial y los *deepfakes* representa un desafío crítico que requiere un enfoque integral. Esta investigación tiene como propósito examinar y proponer directrices para el desarrollo de sistemas de detección efectivos, que protejan los sistemas biométricos contra la suplantación de identidad mediante *deepfake*, ofreciendo soluciones viables y aplicables que garanticen la seguridad y confiabilidad de estos sistemas en entornos reales.

OBJETIVOS

Objetivo general

Plantear los requisitos que deben tener los sistemas de detección de *deepfake* para aplicaciones que utilicen autenticación biométrica facial.

Objetivos específicos

Examinar y entender la autenticación biométrica facial y sus vulnerabilidades.

Detallar la tecnología *deepfake*.

Analizar y comprender los sistemas de detección de *deepfake* para presentar un marco analítico común, comparando los sistemas existentes.

Exponer los requisitos que deben tener los sistemas de detección de *deepfake* utilizados para la autenticación biométrica facial.

CAPÍTULO 1
FUNDAMENTOS DE LA AUTENTICACIÓN BIOMÉTRICA FACIAL

La creciente adopción de sistemas de autenticación biométrica facial en contextos tan diversos como la seguridad bancaria, el control de accesos o la verificación en dispositivos móviles, ha impulsado la necesidad de comprender en profundidad sus fundamentos, beneficios y desafíos. En este capítulo se abordan los principios esenciales de la autenticación biométrica facial, comenzando por el marco general de la biometría y su relevancia en relación con la privacidad y los métodos tradicionales de autenticación. A partir de allí, se exploran los distintos tipos de sistemas biométricos, con especial énfasis en la autenticación facial, sus componentes técnicos, métodos y características. Finalmente, se analizan los desafíos de seguridad asociados a esta tecnología, sentando las bases para introducir en los próximos capítulos una problemática crítica y emergente: la amenaza de los *deepfakes* y la necesidad de mecanismos eficaces para su detección en entornos de autenticación biométrica.

1.1 Privacidad y biometría: Fundamentos y relevancia

En la era digital, donde las tecnologías avanzan rápidamente, la **privacidad** se ha convertido en un tema central, especialmente cuando se manejan datos sensibles como los **datos biométricos**. Estos datos, que son inherentes a cada individuo y tienen la capacidad de identificarlo de manera única, presentan un desafío significativo en términos de protección y control. Un dato biométrico es un “dato referido a las características físicas o fisiológicas o de conducta de una persona que permite su identificación única, como imagen facial o datos dactiloscópicos” (RAE, 2016). Borja, C. T., y Bueno, Á. G. (2006), afirman que:

> Todos los seres humanos tenemos características morfológicas únicas que nos diferencian. La forma de la cara, la geometría de partes de nuestro cuerpo como las manos, nuestros ojos y tal vez la más conocida, la huella digital, son algunos rasgos que nos diferencian del resto de seres humanos. (p. 2)

En informática, la biometría es la medición y análisis de las características físicas y de comportamiento que posee una persona, generalmente utilizado en métodos de autenticación. Según Thales (2016) existen dos categorías de mediciones biométricas:

- **Mediciones fisiológicas:** Se refieren a características únicas morfológicas o biológicas de cada persona. Las características morfológicas se refieren a partes físicas externas, como huellas dactilares, retina, iris, tamaño de la mano, forma de la cara. Las biológicas se refieren a ADN, sangre, saliva y orina.

- **Mediciones del comportamiento:** Remiten a características únicas de la forma de actuar de un individuo. Estas incluyen el reconocimiento de voz, la forma de andar, la escritura, la firma, además de los gestos.

  Además, el autor expande que:

  Generalmente se considera que las mediciones fisiológicas ofrecen el beneficio de permanecer más estables durante toda la vida de un individuo. Por ejemplo, no están tan sujetas a los efectos del estrés, en comparación con la identificación mediante la medición del comportamiento.

El concepto de privacidad según la RAE, se define como el “ámbito de la vida privada que se tiene derecho a proteger de cualquier intromisión”. Pagnotta (2015) destaca que la privacidad implica la capacidad de un usuario para limitar la cantidad de personas que acceden a su información, lo que resulta particularmente relevante cuando se trata de datos biométricos. Estos datos, que incluyen desde la forma de la cara hasta características fisiológicas como el ADN o las huellas dactilares, son tan específicos y únicos que cualquier filtración o uso no autorizado tiene consecuencias irreparables. La privacidad, por tanto, se convierte en algo esencial a proteger en este contexto.

Es importante destacar que la biometría “presenta una gran ventaja con respecto a los demás tipos de sistemas de seguridad (llave, PIN, contraseña…) debido a que este no puede ser olvidado, robado o perdido, ya que es intrínseco al individuo” (Ortiz López, J., 2011, p. 7). En este sentido, el Instituto Nacional de Estándares y Tecnología (NIST, 2013), afirma que, utilizadas con otras tecnologías de autenticación, como los tokens, las tecnologías biométricas proporcionan mayores grados de seguridad que otras tecnologías empleadas por separado.

Sin embargo, la recopilación y almacenamiento de estos datos implica riesgos significativos para la privacidad, ya que, a diferencia de las contraseñas o los PIN, los datos biométricos no son modificables ni reemplazables si se ven comprometidos.

En este sentido, si los datos biométricos son expuestos a un atacante, pueden ser utilizados con fines no deseados, como la suplantación de identidad (Natgunanathan et al., 2016).

Esto resalta la necesidad de un manejo ético y seguro de los datos biométricos, para garantizar la protección de la privacidad de los individuos.

1.2 Autenticación de identidades y métodos de acceso

La gestión adecuada de acceso a sistemas y recursos digitales es esencial para garantizar la seguridad de la información y proteger la privacidad de los usuarios. En este contexto, los procesos de identificación, autenticación, autorización y auditoría juegan un

papel fundamental. Estos pasos, interrelacionados, aseguran que solo las personas autorizadas tengan acceso a los recursos y que sus actividades dentro del sistema estén debidamente controladas y monitoreadas.

1.2.1 Identificación

Como se mencionó anteriormente, el contexto espacial de la investigación contempla los sistemas de autenticación biométrica, y por esto es importante diferenciar el proceso de acceso a un sistema, cuyo primer paso es el de la identificación por parte del sistema hacia otra persona o sistema.

Por tanto, en el proceso de acceso a un sistema, la identificación se define como el acto de proveer credenciales que permitan determinar la identidad de un sujeto (Segu-Info, 2018), es reconocer si una persona es la misma que se supone o busca y para ello debe brindar datos únicos y unívocos, necesarios para ser reconocido.

Teniendo esto en claro, podemos decir que la identificación es reconocer y comparar a una persona entre un grupo, de manera que se deben brindar los datos necesarios para que pueda ser comparada de forma única y unívoca.

De acuerdo con Thales (2016),

> La identificación biométrica consiste en determinar la identidad de una persona. El objetivo es capturar un elemento biométrico, por ejemplo, tomando una foto del rostro, grabando la voz, o capturando una imagen de la huella dactilar. Luego, esos datos se comparan con los datos biométricos de otras varias personas, alojados en una base de datos. De ese modo, la pregunta es muy simple: "¿Quién es usted?. (párr. 12)

En este nuevo enfoque, el proceso de identificación se basa en características físicas o conductuales inherentes al individuo, como las huellas dactilares, el reconocimiento facial o la voz, que son únicas para cada persona y no pueden ser olvidadas, robadas ni compartidas. Así, la identificación biométrica permite un nivel de seguridad superior al evaluar las cualidades personales del usuario, proporcionando una validación más precisa y confiable de su identidad en comparación con los métodos tradicionales.

1.2.2 Autenticación

Una vez identificada a la persona, se debe autenticar, lo cual se define como “la capacidad de demostrar que un usuario o una aplicación es realmente quién dicha persona o aplicación asegura ser” (IBM, 2025). De ser realizada correctamente, se otorga acceso a los diferentes recursos de un sistema mediante el paso de autorización.

La autenticación es la comprobación de las credenciales recibidas con el objetivo de determinar si el sujeto es quien dice ser (Pagnotta, 2015).

Según Segu-Info (2018) existen tres tipos de autenticación:

- **Autenticación por algo conocido:** El usuario **sabe** algo que valida su identidad, por ejemplo, una contraseña.
- **Autenticación por posesión:** El usuario **tiene** algo que valida su identidad, como por ejemplo un token.
- **Autenticación por característica:** El usuario **es** quien valida su identidad mediante reconocimiento de una característica del mismo, por ejemplo, la verificación de voz, huellas dactilares, reconocimiento facial.

Tradicionalmente, la autenticación se basaba en lo que el usuario “sabe” y "tiene", como contraseñas, PINs, tokens, tarjetas. Sin embargo, con la introducción de la biometría, este enfoque cambia significativamente. En la autenticación biométrica, ya no se valida lo que el usuario sabe o posee, sino lo que el usuario "es". Este proceso se complementa con un conjunto de técnicas que permiten o deniegan el acceso a sistemas o recursos, comparando las mediciones de sus datos biométricos con los almacenados previamente en el sistema. De esta manera, se identifica y autentica a la persona, asegurando que quien intenta acceder es realmente quien dice ser.

1.2.3 Autorización

Una vez que el usuario se ha identificado y autenticado en el sistema, ya sea mediante métodos tradicionales o biométricos, se deben determinar los permisos de acceso de un sujeto sobre un objeto (Segu-Info, 2018). En otras palabras, la autorización es el proceso por el cual alguien solicita realizar una acción específica y alguien o algo validará si posee los permisos o derechos necesarios para hacerlo. Así, independientemente del tipo de autenticación utilizada, la autorización garantiza que solo aquellos con los permisos adecuados puedan realizar determinadas operaciones.

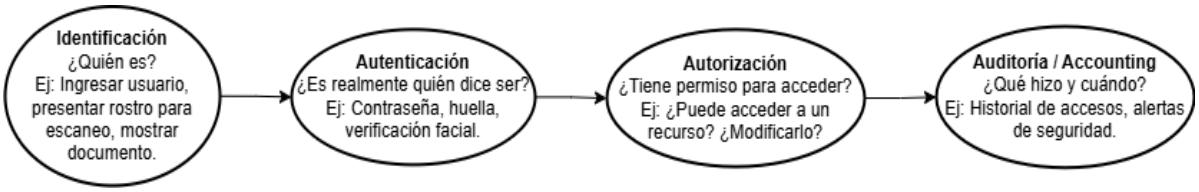
1.2.4 Auditoría/Accounting

Finalmente, es crucial registrar y monitorear todas las acciones realizadas al sistema para mantener la seguridad. La auditoría o "accounting" permite detectar y prevenir posibles violaciones de seguridad, así como realizar un seguimiento detallado de las actividades de los usuarios. Estos registros se convierten en una herramienta esencial para identificar patrones sospechosos y garantizar que los accesos se realicen conforme a las políticas establecidas. Así, sin importar el método de autenticación, la auditoría asegura una trazabilidad completa de las interacciones con el sistema, permitiendo una gestión eficaz de la seguridad.

Como se observa en la Figura 1, Cada etapa —identificación, autenticación, autorización y auditoría— cumple un rol específico e interrelacionado para garantizar que solo los usuarios correctos accedan a los recursos adecuados y que sus acciones queden debidamente registradas.

Figura 1

Proceso de acceso a sistemas digitales



Nota. El gráfico resume los principales componentes del proceso de control de acceso en sistemas digitales.

Es importante destacar que, si bien en los sistemas tradicionales las etapas de identificación y autenticación se presentan como procesos diferenciados, en muchos entornos modernos —especialmente en aquellos que emplean autenticación biométrica— ambas etapas tienden a integrarse en una única acción. Esto ocurre porque el propio rasgo biométrico funciona simultáneamente como declaración de identidad y medio de verificación. Sin embargo, a nivel conceptual y funcional, sigue siendo fundamental distinguir entre la identificación (el acto de presentarse ante el sistema) y la autenticación (la verificación de esa identidad declarada), ya que cada una implica controles de seguridad específicos y se implementa de distintas maneras según el contexto.

1.3 Sistemas de autenticación biométrica: tipos y características

La autenticación biométrica se implementa a través de distintas tecnologías que analizan características físicas o conductuales únicas de cada individuo (ver apartado 1.1). Estas tecnologías se clasifican según el tipo de rasgo utilizado, como el rostro, la voz, la huella dactilar, el iris o incluso la forma de la mano. A continuación, se describen los principales sistemas de autenticación biométrica, destacando su funcionamiento, nivel de seguridad y limitaciones en distintos contextos de uso.

- **Escáner de rostro:** Identifica a la persona mediante la medición facial, calculando la distancia entre nariz, boca, mentón y ojos. Estos escáneres también hacen pruebas de reconocimiento de vida, ya que deben ser capaces de detectar cuándo se les presenta una persona o una fotografía. Al respecto, Suarez, D., y Guarda,

T., (2019, p. 27), plantean que "el reconocimiento facial es muy bueno, pero menos seguro que el análisis de iris y de huellas digitales teniendo la gran ventaja de no ser un método invasivo".

- **Escáner de huella:** Identifica a la persona mediante la huella dactilar, generalmente, del dedo pulgar. Este tipo de escáneres deben ser lo suficientemente sofisticados para no ser engañados, por ejemplo en la utilización de una cinta adhesiva con la huella. Al respecto, Suarez, D., y Guarda, T., (2019):

  > Este es un sistema muy utilizado por los usuarios de smartphones que poseen esta tecnología, lo único que hay que hacer es colocar la huella dactilar en el sensor y el sistema realizará la verificación utilizando las huellas digitales ingresadas previamente por la persona. (p. 29)

- **Escáner de mano:** Identifica a la persona mediante la mano completa, ya que al igual que la huella, ésta es única.
- **Escáner de retina o iris:** Identifica a la persona mediante la retina o iris. Es uno de los sistemas más seguros que existen debido a que no se ha encontrado forma aún de falsificar la retina o iris. Borja, C. T., & Bueno, Á. G. (2006) expresan que:

  > La cámara genera una imagen que es analizada por medio de los algoritmos de Daugman para obtener el IrisCode personal, un patrón único del iris que apenas ocupa 256 bytes de información. Tan reducido tamaño permite una rápida búsqueda de su homólogo en una base de datos hasta identificar a su propietario. (p. 20)

- **Escáner de voz:** De acuerdo con Aguirrezabala, A. M. (2015):

  > La biometría de voz es muy utilizada en aplicaciones o sistemas relacionados con la seguridad. Ya que cada individuo tiene unas características físicas diferentes. La voz es considerada una característica física única, debido a que surge del tracto vocal, que varía dependiendo de la persona. Los softwares de detección de voz no son sencillos de realizar y suelen ser costosos, debido a que existen dificultades en la detección, como el ruido de fondo, si la persona posee una enfermedad, su edad, su estado de ánimo, entre otros. (p. 3-4)

1.4 Seguridad en la autenticación biométrica

Tras explorar la biometría y su evolución en los sistemas de autenticación (ver apartado 1.1), es fundamental abordar los aspectos de seguridad relacionados con su implementación. La autenticación biométrica, al estar basada en características únicas e inherentes al individuo, ha transformado la forma en que validamos identidades en los

sistemas, ofreciendo una mayor precisión y conveniencia. Sin embargo, a pesar de sus ventajas, la biometría no está exenta de desafíos y riesgos, especialmente en lo que respecta a la protección contra ataques de suplantación de identidad y la manipulación de los datos biométricos.

Para comprender mejor esta problemática, resulta necesario vincularla con el concepto de **identidad digital.** Esta no solo se refiere a los datos que identifican técnicamente a un usuario, sino a todas aquellas características, credenciales y actividades que una persona realiza en entornos digitales. En este contexto, los datos biométricos forman parte central de esa identidad digital: si son manipulados, robados o falsificados, no sólo se pone en riesgo el acceso a un sistema, sino también la representación digital de una persona, su privacidad y su integridad online.

Según Grupo Ático34 (2019) cuando se alteran los datos biométricos de una persona, haciendo que imite a los de otra, es posible realizar la suplantación de identidad, que es cuando una persona se hace pasar por otra, generalmente intentando ocasionar algún daño, como fraude, ciberbullying, grooming o robo de información. Además, mencionan que existen tres tipos de clasificaciones y características que posee la suplantación de identidad: la primera consiste en la creación de un perfil falso en donde no se añade información personal de la víctima; la segunda, además de lo anteriormente mencionado, posee información personal como una imagen de la víctima; y la tercera es cuando se accede a un servicio de un usuario haciéndose pasar por él.

Borghello, C. y Temperini M. G. (2012) hablan de *identidad digital*, refiriéndose a las características y actividades llevadas a cabo por una persona en internet. Además, mencionan que los medios o tipos de suplantación de identidad son variados, y que entre ellos se destacan: Documento nacional de identidad (DNI), falsificación de la firma, clonación de tarjetas de crédito, duplicado de tarjeta SIM, suplantación de identidad a través del correo electrónico.

En este contexto, la seguridad en la autenticación biométrica no solo se trata de verificar que la persona es quien dice ser, sino también de asegurarse de que este proceso sea confiable, evitando fraudes y protegiendo la identidad digital del individuo. Aquí es donde entran en juego dos aspectos clave: las pruebas de vida y los sistemas *anti-spoofing*, desarrollados para detectar y bloquear internos de suplantación.

Para mitigar este riesgo, los sistemas de acceso poseen métodos que permiten identificar si se trata de un posible caso de suplantación de identidad. Estos se denominan *sistemas de prueba de vida*.

1.4.1 Pruebas de vida y sistemas anti-spoofing en la autenticación biométrica

Según Gomez, N. (2020) la prueba de vida o prueba de detección de vida en los análisis biométricos, busca que el sistema sea capaz de determinar si se está interactuando con una persona física o si se trata de un ataque de suplantación de identidad que se realiza mediante un objeto inanimado. Además, indica que las pruebas de detección de vida deben identificar los intentos de suplantación de identidad mediante fotos, máscaras, videos, impresiones y fotografías. Por esto las pruebas de detección de vida se convirtieron en una parte fundamental de los desarrollos de los sistemas de validación biométrica.

Raheem, E. A., Ahmad, S. M. S., y Adnan, W. A. W. (2019) mencionan varios tipos de pruebas de vida que se utilizan en los sistemas de autenticación biométrica facial. A continuación, se describen algunos de los métodos más comunes:

- **Detección de movimiento:**
    a. *Instrucciones de movimiento*: el sistema solicita al usuario realizar movimientos específicos, como girar la cabeza o moverla hacia adelante y hacia atrás.
    b. *Parpadeo:* se pide al usuario que parpadee varias veces en un periodo de tiempo determinado. Los parpadeos son movimientos naturales que son difíciles de falsificar en imágenes estáticas o vídeos pregrabados.
- **Análisis de textura:**
    c. *Detección de textura de la piel*: se analiza la textura de la piel en tiempo real para identificar características naturales de la piel humana que no están presentes en fotografías o videos falsificados.
    d. *Análisis de Espectro de Luz*: utiliza diferentes espectros de luz (por ejemplo, infrarrojo) para capturar detalles de la piel que no son visibles en condiciones normales de iluminación.
- **Pruebas de profundidad:**
    e. *Cámaras 3D*: utilizan sensores de profundidad para mapear el rostro en tres dimensiones, asegurando que la estructura facial corresponde a la de un objeto tridimensional real, en lugar de una imagen bidimensional.
    f. *Estereoscopía*: emplea dos o más cámaras para capturar imágenes desde diferentes ángulos y crear un mapa de profundidad del rostro.
- **Análisis de Movimiento Fisiológico:**
    g. *Flujo Sanguíneo*: se analiza el flujo sanguíneo bajo la piel, utilizando técnicas como la fotopletismografía, para detectar los cambios de color sutiles que ocurren con cada latido del corazón.

h. *Micromovimientos*: detecta pequeños movimientos involuntarios del rostro que ocurren naturalmente, como el pulso en ciertas áreas de la cara.

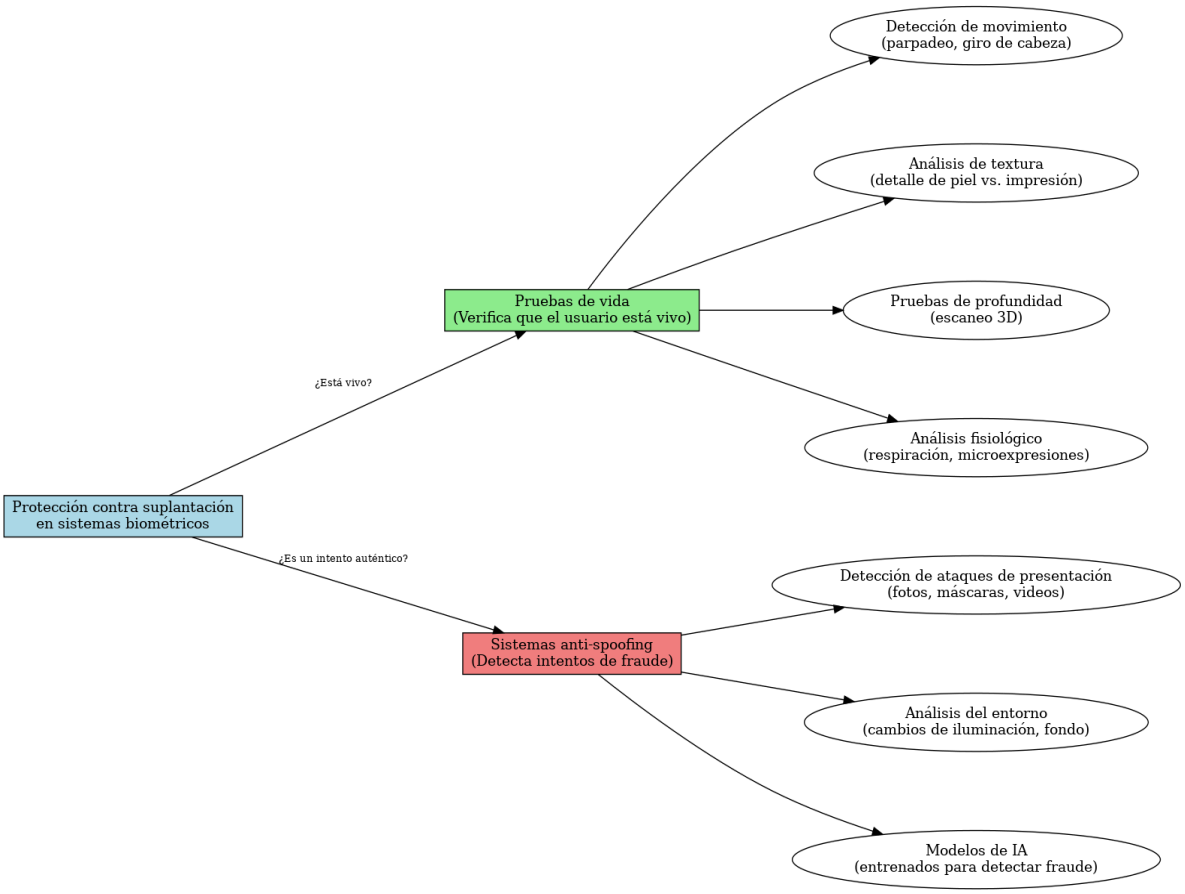
- **Interacciones en Tiempo Real:**
    i. *Respuestas a Instrucciones Aleatorias*: el sistema da instrucciones aleatorias al usuario, como levantar una ceja o sonreír, y verifica la respuesta en tiempo real para asegurar que es una persona viva la que interactúa con el sistema.
    j. *Análisis de Comportamiento*: observa patrones de comportamiento durante la interacción, como la rapidez y naturalidad con que el usuario responde a las instrucciones.

Además de las pruebas de vida, existen también los *sistemas anti-spoofing*, los cuales emplean distintas técnicas y tecnologías diseñadas para prevenir intentos de suplantación de identidad en sistemas biométricos. Mientras que las pruebas de vida buscan confirmar la presencia real y biológica del usuario en tiempo real, los sistemas *anti-spoofing* se enfocan en detectar intentos de engaño mediante datos falsificados, como fotos, videos o máscaras. Esto se logra a través del desarrollo de métodos y algoritmos capaces de distinguir entre entradas genuinas y manipuladas, muchas veces con el apoyo de modelos de inteligencia artificial. En línea con esto, Silva M. (2023) afirma que para garantizar la confiabilidad de la autenticación facial, se deben desarrollar algoritmos *anti-spoofing* que puedan defender eficazmente todos los intentos de suplantación de identidad. La implementación conjunta de ambos mecanismos resulta esencial para reforzar la seguridad y confiabilidad de los sistemas de autenticación biométrica frente a amenazas cada vez más complejas.

Como se observa en la Figura 2, ambos conceptos son fundamentales para fortalecer la seguridad de los sistemas biométricos. Mientras que las pruebas de vida verifican que el usuario sea un ser humano presente en tiempo real (vivo), los sistemas *anti-spoofing* se encargan de detectar y bloquear intentos de fraude mediante técnicas como el análisis del entorno o el uso de modelos entrenados para reconocer patrones de suplantación.

Figura 2

Protección contra suplantación de identidad en sistemas biométricos.



Nota. El gráfico diferencia los conceptos de pruebas de vida y sistemas anti-spoofing.

1.5 Autenticación biométrica facial

Una vez analizados los sistemas biométricos, su funcionamiento general, tipos y los mecanismos de seguridad que los acompañan (ver apartado 1.4) , es posible adentrarse en uno de los métodos más extendidos y reconocidos: la autenticación biométrica facial. Este tipo de autenticación se ha convertido en una de las formas más populares dentro del ámbito biométrico, dado que, como menciona Li y Jain (2011), es ampliamente utilizada en múltiples aplicaciones y dispositivos móviles, ofreciendo diversas ventajas en comparación con otros métodos de autenticación biométrica, entre las cuales se destacan:

- **No requiere contacto físico**: esto es más higiénico, especialmente en situaciones en donde se busque evitar la propagación de gérmenes o en entornos donde tocar superficies es un problema.

- **Facilidad de uso**: la autenticación facial es intuitiva y no requiere que los usuarios toquen sensores específicos. Para algunos usuarios es más sencillo mirar a la cámara.
- **Dispositivos móviles**: muchos ya están equipados con cámaras frontales para selfies, facilitando la implementación sin necesidad de hardware adicional.
- **Aceptación del usuario:** algunos usuarios se sienten más cómodos con la autenticación facial, ya que no implica tocar superficies y se percibe como una forma más natural de interactuar con dispositivos.
- **Niveles razonables de seguridad:** aunque la seguridad no es absoluta, los sistemas de autenticación facial biométricos modernos han mejorado significativamente en términos de precisión y resistencia a ataques, ofreciendo niveles razonables de seguridad en muchas aplicaciones.

A diferencia de otros métodos como la huella dactilar o el escaneo de iris, no requiere interacción directa con sensores, ya que puede capturar imágenes faciales a distancia o desde un flujo de video, lo que permite desbloquear un dispositivo de forma rápida y sin contacto físico. Esta característica lo hace especialmente adecuado para entornos donde la higiene y la eficiencia son prioritarias (Zulfiqar et al., 2019).

- **Desbloqueo de dispositivos:** se ha vuelto popular en dispositivos móviles y otros gadgets como un método seguro y conveniente para desbloquear el acceso (Li y Jain, 2011).
- **Identificación en aeropuertos:** En el ámbito aeroportuario, la Administración de Seguridad en el Transporte (TSA) (s.f.), ha implementado tecnología de reconocimiento facial para verificar la identidad de los pasajeros en puntos de control de seguridad. Este sistema compara la imagen en tiempo real del viajero con la fotografía presente en su documento de identidad, como pasaportes o licencias de conducir, sin necesidad de contacto físico.
- **Seguridad en el sector bancario:** El Banco de la Provincia de Buenos Aires ha incorporado la autenticación biométrica facial como un mecanismo de seguridad adicional para validar la identidad de los usuarios al dar de alta nuevas cuentas destino para transferencias a través de sus plataformas (Banco de la Provincia de Buenos Aires, s.f.).
- **Control de acceso a edificios y zonas restringidas:** se están utilizando para controlar el acceso a edificios, como oficinas y escuelas. Estos sistemas permiten el acceso solo a las personas autorizadas después de la verificación de su identidad (Zulfiqar et al., 2019).
- **Control de asistencia en instituciones educativas**: En el ámbito educativo, los sistemas de reconocimiento facial han comenzado a utilizarse como herramienta

de gestión de asistencia automatizada. Estos sistemas permiten registrar la presencia de los estudiantes de forma precisa y eficiente, reduciendo el tiempo empleado en controles manuales y minimizando errores como asistencias incorrectamente asignadas. Esto resulta especialmente útil en instituciones con gran número de alumnos o alta rotación diaria (Zulfiqar et al., 2019).

1.5.1 Principios y Funcionamiento de la Autenticación Biométrica Facial

Habiendo sentado las bases teóricas sobre la biometría y autenticación, es momento de dirigir la atención hacia la aplicación específica de estos principios en el ámbito de la autenticación biométrica facial. Con los fundamentos establecidos, a continuación se analizará el universo de la identificación y verificación facial, explorando los diversos tipos y funciones que definen a esta innovadora tecnología.

Chowdhury, M., Gao, J., y Islam, R. (2017) proponen una arquitectura de autenticación biométrica la cual consta de los siguientes pasos:

1. **Pre-procesar la imagen facial**: consiste en eliminar el ruido y mejorar la calidad de los datos. Para ello, se aplica un filtro de mediana difusa que ajusta el nivel de gris en los píxeles para mejorar la claridad de la imagen.
2. **Detección del rostro**: localiza el rostro dentro de la imagen utilizando técnicas de segmentación por color de piel. A partir de esto, se extrae el esqueleto del rostro para continuar con el proceso de normalización, que ajusta la imagen a un formato estándar.

Figura 3

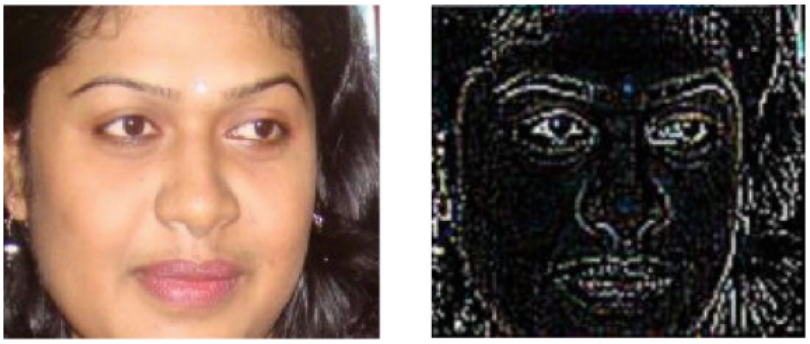
Protección contra suplantación de identidad en sistemas biométricos.

Nota. Detección de caras y proceso de normalización para una secuencia de imágenes reales. Tomado de Biometric Authentication Using Facial Recognition (p. 290) por Chowdhury, M., Gao, J., & Islam, R., (2017).

3. **Extracción de características**: una vez detectado el rostro, se extraen las características clave mediante la utilización de un detector de bordes. Este método captura las características invariantes, como las aristas del rostro, que son robustas frente a variaciones de pose e iluminación.

Figura 4

Extracción de rasgos de los bordes faciales



Nota. Extracción de rasgos de los bordes faciales. Tomado de Biometric Authentication Using Facial Recognition (p. 291) por Chowdhury, M., Gao, J., & Islam, R. (2017).

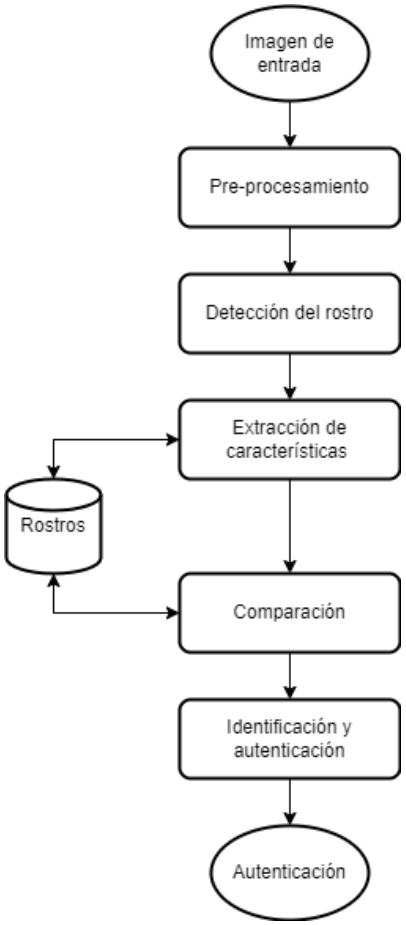
4. **Comparación de características**: las características extraídas se comparan con las que se encuentran almacenadas en una base de datos de rostros. Este proceso se realiza utilizando una red neuronal que ha sido entrenada previamente para identificar los patrones específicos del rostro.
5. **Identificación y autenticación**: finalmente, si las características del rostro coinciden con las de la base de datos, el sistema identifica al usuario y otorga la autenticación correspondiente.

La Figura 5 resume el flujo general del proceso de autenticación biométrica facial, integrando las etapas previamente descritas. A partir de una imagen de entrada —capturada por una cámara— el sistema realiza un pre-procesamiento para mejorar la calidad y adecuar la imagen a los parámetros del análisis. Luego, se procede a la detección del rostro, etapa en la cual se identifica la región específica que contiene la cara del usuario. A continuación, se lleva a cabo la extracción de características faciales únicas (como la distancia entre ojos, forma de la mandíbula), las cuales son comparadas con aquellas almacenadas en una base de datos. Esta comparación permite determinar coincidencias y validar la identidad del individuo. Finalmente, si los parámetros coinciden

dentro de los márgenes aceptables definidos por el sistema, se realiza la identificación y autenticación del usuario, permitiendo así el acceso autorizado.

Figura 5

Proceso de autenticación biométrica.



Nota. Arquitectura del sistema de autenticación biométrica. Adaptado de Biometric Authentication Using Facial Recognition por Chowdhury, M., Gao, J., & Islam, R. (2017).

1.5.2 Tipos de métodos biométricos faciales

Una vez comprendido el funcionamiento general del proceso de autenticación biométrica facial, es importante considerar que existen distintos enfoques técnicos para llevarlo a cabo. El reconocimiento facial se basa en el uso de algoritmos informáticos capaces de detectar, analizar y comparar patrones faciales únicos. Sin embargo, la manera en que estos patrones se capturan y procesan presenta variaciones significativas según la tecnología utilizada. En este sentido, se distinguen diferentes tipos de métodos biométricos faciales, cada uno con sus propias características, ventajas y requerimientos técnicos. La elección del método dependerá, entre otros factores, del tipo de dispositivo,

la calidad de la cámara y el entorno en el que se implemente el sistema. Entre los principales métodos, podemos encontrar los siguientes:

- **Reconocimiento facial 2D:** este tipo de autenticación utiliza una imagen bidimensional del rostro de una persona para identificar y autenticar su identidad; Se basa en características como la forma del rostro, la posición de los ojos, la nariz y la boca. Tal como menciona Santana et al. (2017) este tipo de métodos extraen rasgos o marcas como lo son ojos, nariz, boca y mediante el uso de la geometría y el uso de bordes, líneas y curvas reconoce distintas caras. Este tipo de técnica presenta limitaciones debido a la estructura 3D del rostro, siendo afectados en la pose, cercanía de la persona con la cámara, iluminación del ambiente, oclusiones.
- **Reconocimiento facial 3D:** utiliza una imagen tridimensional del rostro para una autenticación más precisa. La imagen se procesa mediante un algoritmo de reconocimiento facial que analiza las características faciales en términos de su posición tridimensional y su textura (Zhou y Xiao, 2018). El algoritmo de reconocimiento facial basado en la profundidad analiza la reflectividad de la piel y la posición de los ojos, lo que ayuda a mejorar la precisión del reconocimiento facial. La imagen tridimensional resultante se compara con una base de datos que contiene imágenes de rostros autorizados, y si hay suficiente coincidencia entre las características del rostro del usuario y las imágenes autorizadas, se autentica al usuario.

  Este tipo de técnica utiliza mallas poligonales de los puntos 3D establecidos en los ejes x, y, z, siendo z el valor de la profundidad. Según Santana et al. (2017), existen diversas técnicas para la detección facial en 3D, una de ellas es la técnica de estéreos, considerada la más económica y fácil de implementar, y que utiliza múltiples cámaras posicionadas y calibradas para obtener imágenes simultáneas del sujeto, basándose en datos geométricos y puntos de referencia. Otra técnica es la de luz estructural, que proyecta un patrón de luz y adquiere información de profundidad a partir de la distorsión del patrón, siendo relativamente rápida y barata. Por último, la técnica láser es la más precisa, aunque también es la más cara y lenta, ya que emplea un sensor láser para escanear la superficie.
- **Reconocimiento facial por infrarrojos:** Kakkirala, K. R., Chalamala, S. R., y Jami, S. K. (2017), mencionan que este método utiliza sensores de infrarrojos para capturar la imagen del rostro, identificando características faciales incluso en condiciones de poca luz o en entornos con iluminación variable. Estos se utilizan en conjunto con métodos de detección que requieren luz para potenciarlos y realizar una mejor clasificación.

- **Reconocimiento facial por análisis de expresiones:** este método se basa en el análisis de las expresiones faciales de una persona para autenticar su identidad; Evalúa características como el movimiento de los músculos faciales y la expresión de emociones para verificar la autenticidad de una persona. Este enfoque utiliza redes neuronales profundas para aprender y reconocer patrones faciales complejos. Es eficaz para adaptarse a diversas condiciones de iluminación y ángulos de captura.

1.6 Desafíos y vulnerabilidades de la autenticación biométrica facial

Como se abordó en el apartado sobre seguridad en la autenticación biométrica (ver apartado 1.4), los sistemas de reconocimiento facial han incorporado mecanismos como las pruebas de vida y los sistemas *anti-spoofing* para mitigar intentos de suplantación o engaño.

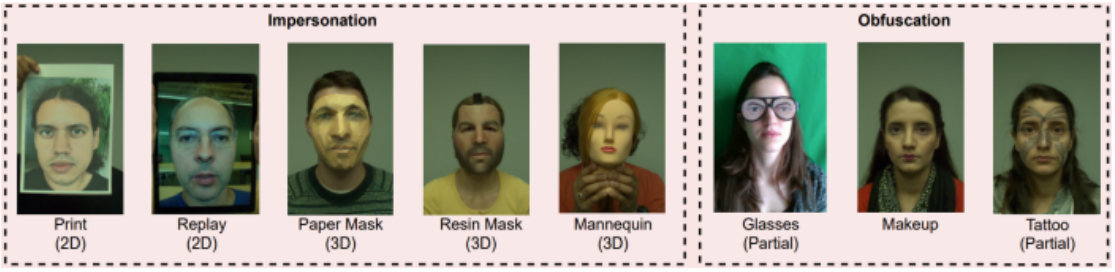
Sin embargo, a medida que las tecnologías de generación de imágenes sintéticas avanzan, estas medidas podrían no ser suficientes frente a amenazas más sofisticadas, como los *deepfakes*, que desafían los límites actuales de detección y verificación.

Silva M. (2023) destaca que:

> El tipo más común de ataques de presentación son impresiones, reproducciones, máscaras 3D, maniquíes, anteojos, maquillaje y tatuajes. Según su tipología, los ataques de presentación se dividen en ataques de suplantación y de obstrucción. En los ataques de suplantación, los impostores usan suplantación para ser reconocidos como otra persona copiando los atributos faciales de las víctimas en ataques de suplantación. En los ataques de obstrucción, los impostores usan trucos, por ejemplo, usar anteojos, maquillaje extremo, pelucas, para evitar ser reconocidos por el sistema. Además, según su elaboración, los ataques de presentación se clasifican en ataques 2D y 3D. Los ataques 2D giran principalmente en torno a presentar atributos faciales utilizando fotos planas o fotos impresas envueltas, fotos con ojos o bocas recortadas y reproducciones de videos digitales. Por otro lado, los ataques 3D implican el uso de máscaras impresas hechas de materiales especializados como papel, resina o plástico. (p. 17-18)

Figura 6

Ejemplos de ataques de spoofing



Nota. El gráfico representa diferentes tipos de ataques de presentación. Tomado de Vision Transformers For Face Anti-Spoofing (Master's thesis) por Silva, M. M. (2023).

A partir de lo expuesto anteriormente sobre los ataques de presentación, es posible ampliar el análisis incorporando otras vulnerabilidades relevantes que afectan a los sistemas de autenticación biométrica facial, tales como:

- **Ataques de presentación**: la posibilidad de utilizar imágenes falsas o manipuladas para engañar al sistema de reconocimiento facial constituye una de las principales vulnerabilidades, pudiendo lograrse mediante la impresión de imágenes faciales en 3D, el uso de máscaras o la manipulación de imágenes digitales:
    - *Suplantación*: Imita los atributos faciales de otra persona.
    - *Obstrucción*: Alterar el aspecto para evitar el reconocimiento.
    - *Según complejidad:*
        - Ataques 2D: Por ejemplo, fotos impresas.
        - Ataques 3D: Por ejemplo, máscaras en papel o resina.
- **Robo de Datos Biométricos[1]:** si los datos biométricos, incluidas las características faciales, se almacenan de manera insegura, podrían ser susceptibles al robo. La pérdida de datos biométricos podría tener implicaciones graves, ya que estos son difíciles o imposibles de cambiar una vez comprometidos.
- **Falsificación mediante técnicas avanzadas**: tecnologías emergentes como los *deepfakes* presentan una amenaza significativa, ya que generan contenido altamente realista capaz de superar los controles convencionales de autenticación.

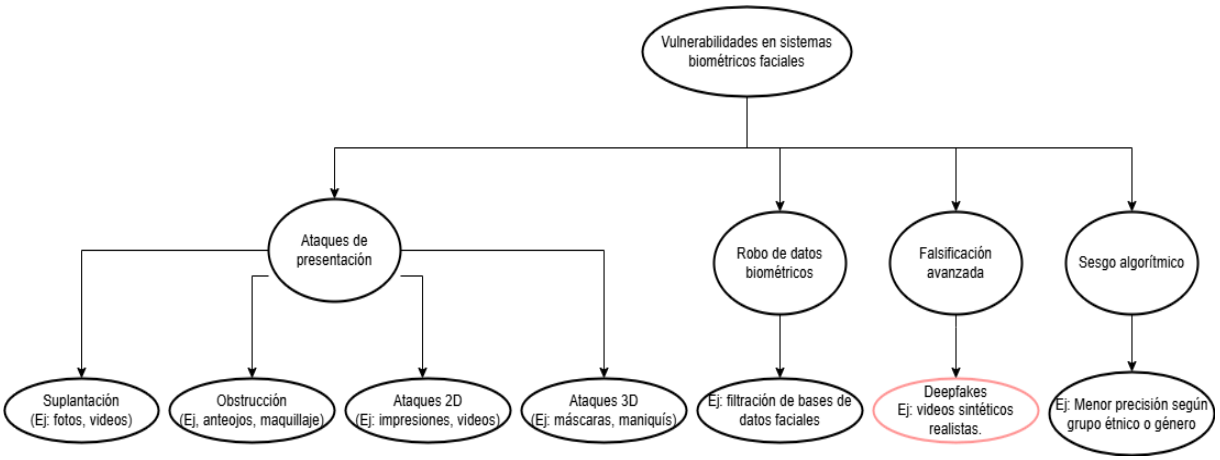
[1] **Consideraciones Éticas sobre la Privacidad de Datos Biométricos**: Aunque el presente documento se centra en aspectos específicos de la autenticación biométrica facial, es esencial reconocer la creciente importancia del almacenamiento seguro de datos biométricos y sus implicaciones éticas asociadas. La gestión inadecuada de esta información, como la falta de cifrado robusto o la vulnerabilidad de los sistemas de almacenamiento, no solo plantea riesgos técnicos, sino que también suscita preocupaciones éticas y de privacidad.

- **Condiciones técnicas y contextuales**: Factores como la baja resolución de las imágenes, cambios en la iluminación, rotación del rostro fuera del plano frontal y diferencias temporales entre las capturas pueden degradar significativamente el rendimiento del sistema, aumentando la tasa de error en la verificación de identidad (Li y Jain, 2011).
- **Sesgos algorítmicos**: estudios como el de Buolamwini y Gebru (2018) demuestran que los sistemas de reconocimiento facial tienen menor precisión al identificar personas de piel oscura, mujeres y personas con condiciones faciales atípicas, lo que deriva en riesgos de discriminación, sesgo o exclusión. Además, estos sistemas presentan dificultades para diferenciar con precisión a individuos con rasgos faciales extremadamente similares, como en el caso de gemelos, lo que representa un desafío adicional para la confiabilidad del reconocimiento facial.

A continuación, se puede observar en la Figura 7, las principales vulnerabilidades que afectan a los sistemas de autenticación biométrica facial. Este gráfico permite visualizar la interrelación entre los distintos tipos de ataques, sus clasificaciones y las amenazas emergentes, como los *deepfakes*, que representan un nuevo desafío para los mecanismos tradicionales de seguridad biométrica:

Figura 7

Vulnerabilidades en la autenticación biométrica facial



Nota. El gráfico representa las principales vulnerabilidades que afectan a los sistemas de autenticación biométrica facial.

Hasta ahora, los sistemas *anti-spoofing* y las pruebas de vida han sido suficientes para mitigar ataques tradicionales, como los basados en fotografías o videos impresos. Sin embargo, con la irrupción de técnicas de falsificación avanzada cada vez más sofisticadas, como los *deepfakes*, se pone en evidencia la necesidad de repensar la seguridad de los sistemas biométricos. Estos nuevos vectores de ataque representan un

punto de inflexión que desafía las estrategias defensivas actuales y exige la incorporación de mecanismos de detección más robustos y especializados.

Esta creciente amenaza será abordada en profundidad en el siguiente capítulo, donde se analizarán los fundamentos de los *deepfakes*, su clasificación, técnicas de generación y los desafíos que presentan en el contexto de la autenticación biométrica facial.

CAPÍTULO 2
DEEPFAKE: FUNDAMENTO, CLASIFICACIÓN Y DESAFÍOS

Tras haber explorado en el capítulo anterior los fundamentos de la autenticación biométrica facial, sus ventajas, desafíos y mecanismos de protección, este capítulo se enfoca en una amenaza emergente que compromete la confiabilidad de estos sistemas: los *deepfakes*.

La capacidad de generar imágenes, audios o videos sintéticos altamente realistas mediante técnicas de inteligencia artificial ha abierto nuevas posibilidades tecnológicas, pero también nuevos riesgos en lo que respecta a la suplantación de identidad y la manipulación digital. A lo largo de este capítulo, se abordarán los fundamentos del *deep learning* —tecnología subyacente a los *deepfakes*—, su definición, principales tipos, el proceso de creación, los ataques derivados a su uso, así como sus ventajas, desventajas y su evolución reciente. Este análisis permitirá comprender la magnitud del desafío que representan los *deepfakes* en contextos donde la autenticación y la identidad digital resultan críticos.

Desde su concepción inicial, los *deepfakes* han evolucionado significativamente. De hecho, en los últimos años, los avances en la inteligencia artificial y el deep-learning han permitido una mayor precisión y realismo en su creación.

2.1 Fundamentos del deep learning

El *deep learning*, o aprendizaje profundo, no solo constituye un hito en el desarrollo de tecnologías avanzadas, sino que también representa el núcleo de las técnicas utilizadas para la creación de *deepfakes*. Mencionar brevemente el funcionamiento de esta tecnología resulta esencial para entender cómo estas falsificaciones digitales alcanzan niveles de realismo tan altos que llegan a desafiar incluso a sistemas avanzados de autenticación biométrica facial.

El *deep learning* se define como un subconjunto del aprendizaje automático (*machine learning*) que utiliza redes neuronales artificiales capaces de aprender a partir de grandes volúmenes de datos (LeCun et al., 2015, p. 436). Asimismo, se describe como un conjunto de algoritmos no lineales diseñados para modelar datos y reconocer patrones complejos (Schmidhuber, 2015, p. 87).

En términos técnicos, el *deep learning* es considerado una rama avanzada del machine learning que se basa en redes neuronales de múltiples capas. Estas redes simulan el comportamiento del cerebro humano al procesar información, permitiendo la

extracción automática de características relevantes y la construcción de modelos capaces de identificar relaciones y tendencias en los datos (Goodfellow et al., 2016, p. 27).

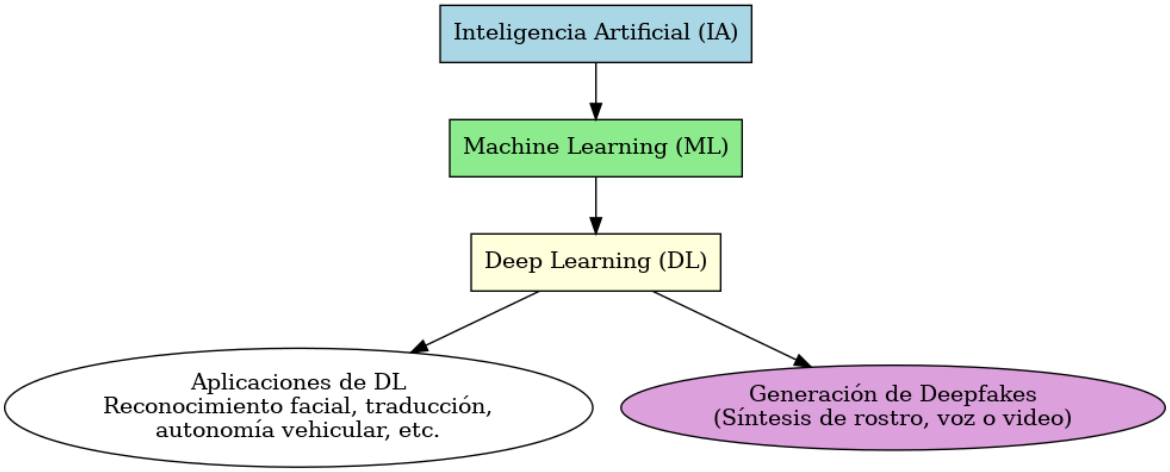
El funcionamiento de estas redes se basa en capas con unidades de procesamiento, que permiten la transformación progresiva de datos o variables. Cada capa extrae y procesa información, alimentando a la siguiente en una secuencia de aprendizaje que se repite miles o millones de veces hasta alcanzar niveles óptimos de precisión. Gracias a este proceso iterativo, el sistema mejora su rendimiento de forma progresiva, generando modelos cada vez más precisos y complejos.

Las aplicaciones del *deep learning* abarcan un amplio espectro de tareas, desde traducción automática, asistentes virtuales, chatbots, vehículos autónomos, búsqueda visual y reconocimiento facial, hasta análisis de imágenes médicas y detección de fraudes. En particular, el mismo tipo de redes neuronales que posibilitan sistemas de autenticación confiables también son utilizadas con otros fines, en la creación de contenidos sintéticos manipulados: los *deepfakes*. La dualidad tecnológica —capaz de proteger, pero también de vulnerar la identidad— es uno de los principales desafíos que se abordarán a lo largo de este capítulo.

La Figura 8 permite visualizar al *deep learning* como una subrama del machine learning, y a su vez, cómo da lugar a múltiples aplicaciones. Entre ellas se encuentran tanto herramientas orientadas a resolver tareas o asistir al usuario —como el reconocimiento facial, la traducción automática o la conducción autónoma— como también la generación de contenidos sintéticos, tal como ocurre en el caso de los *deepfakes*.

Figura 8

Relación jerárquica entre inteligencia artificial, aprendizaje automático y aprendizaje profundo



Nota: El gráfico destaca su vinculación con las aplicaciones generales y la generación de deepfakes.

La representación diferenciada de estas dos ramas tiene un fin ilustrativo, ya que, si bien ambas utilizan tecnologías de *deep learning*, sus objetivos y repercusiones son sustancialmente distintos. Esta distinción permite comprender que una misma base tecnológica se utiliza tanto para fines positivos como para propósitos que, en ciertos contextos, suponen un riesgo para la seguridad y la autenticación de identidades.

2.2 Definición y tipos de *deepfake*

Anteriormente se analiza el papel de las pruebas de vida en los sistemas de autenticación biométrica (ver apartado 1.4), diseñadas para verificar si se está interactuando con una persona real y presente, o con un objeto inanimado. Sin embargo, con el uso de *deepfake*, este límite se vuelve cada vez más difuso, ya que los ataques se presentan mediante objetos animados artificialmente, capaces de imitar expresiones humanas con un alto nivel de realismo, complicando la detección de intentos de suplantación de identidad.

Según Chadha, Kumar, Kashyap y Gupta (2021) el término *deepfake* surge de la combinación de las palabras *deep learning* y *fake* (falso). Esta técnica emplea redes neuronales artificiales entrenadas con grandes volúmenes de datos para imitar las características visuales o sonoras de una persona, permitiendo generar contenido falso que aparenta ser real. De este modo, alguien parece decir o hacer cosas que en realidad nunca dijo o hizo, desafiando los límites de la autenticidad digital.

En términos generales, el *deepfake* se define como una técnica de generación de imágenes, audios o videos sintéticos que utiliza inteligencia artificial para crear contenido que simula ser auténtico, sin serlo.

Tabla 1

Tipos de *deepfake*

Tipos	Método y descripción	Ejemplo	Aplicación
Deepfake de fotografía	**Intercambio de rostros y cuerpos.** Cambio del rostro y cuerpo de una persona.	Aplicación móvil que utiliza filtros de envejecimiento.	Los clientes pueden probar virtualmente vestidos y cosméticos antes de comprarlos.
Audio deepfake	**Intercambio de voces.** La voz es reemplazada por la de otra persona.	Se utilizó para engañar al gerente imitando la voz de un CEO y transferir	Imitar a buenos oradores para audiolibros.

		243.000 dólares.	
	Texto a voz. El texto escrito se convierte en audio.	Alguien realizó una grabación controversial con la voz de Jordan B. Peterson.	Puede ser utilizado en la industria cinematográfica para corregir palabras mal pronunciadas.
Video deepfake	**Intercambio de rostros.** El rostro original es intercambiado por otro.	En la película "Fast and Furious", el rostro de Paul Walker fue intercambiado por el rostro de su hermano.	Intercambiando el rostro del actor con el del doble de riesgo, para que los actores estén seguros.
	Morfología facial. Cambiar ciertas características del rostro.	El actor de Saturday Night Live, Bill Hader, se transforma dentro y fuera del actor Arnold Schwarzenegger en un programa de entrevistas.	En los videojuegos, los jugadores pueden dar su rostro a su avatar.
Deepfake de audio y video[2]	**Sincronización de labios.** Los labios se mueven de acuerdo con el rostro y se imita el audio.	"You won't believe what Obama says in this video."	Los anuncios y los videos de instrucción pueden ser convertidos a diferentes idiomas sin necesidad de volver a filmar.

Nota: Adaptado de Tipos de deepfake y sus aplicaciones. Tomado de Deepfake: An Overview (p. 560), por Chadha, A., Kumar, V., Kashyap, S., & Gupta, M. (2021), en Singh, P. K., Wierzchoń, S. T., Tanwar, S., Ganzha, M., & Rodrigues, J. J. P. C. (Eds.). (2021). Proceedings of Second International Conference on Computing, Communications, and Cyber-Security. Lecture Notes in Networks and Systems.

Se manifiestan en distintos formatos, ya sea como fotografías manipuladas, audios falsificados, videos editados o combinaciones de audio y video generadas íntegramente de forma sintética. Su versatilidad y nivel de realismo los convierten en una

[2] Se recomienda la visualización de los siguientes videos:
https://www.youtube.com/watch?v=cQ54GDm1eL0 - https://www.youtube.com/watch?v=bPhUhypV27w - https://www.youtube.com/watch?v=PCBTZh41Ris

herramienta poderosa, con aplicaciones legítimas pero también con un alto potencial de mal uso, especialmente en contextos donde la veracidad de la identidad es crítica.

Westerlund, M., (2019, s/pág.) afirma que "Los *deepfakes* son difíciles de detectar, ya que utilizan metraje real, pueden tener un audio con un sonido auténtico y están optimizados para difundirse rápidamente en las redes sociales". Esta afirmación evidencia no solo el nivel de sofisticación que alcanzan estas técnicas, sino también la urgencia de comprender cómo se generan y qué herramientas hacen posible su creación.

2.2.1 Aplicaciones potenciales: beneficios y desventajas

La tecnología *deepfake* ofrece una serie de ventajas en diversos campos como el entretenimiento, la educación, la salud o el marketing, por su capacidad para generar contenidos realistas y personalizados. Sin embargo, sus aplicaciones también conllevan importantes desventajas y riesgos, especialmente cuando se utilizan con fines maliciosos.

2.2.1.1 Beneficios del *deepfake*

Entre los aspectos positivos que destaca Westerlund (2019), se encuentran:

- **Producción de medios y entretenimiento:** los *deepfakes* tienen aplicaciones en la industria del cine y la televisión para recrear digitalmente a actores fallecidos, rejuvenecer actores mayores o revivir personajes icónicos, generando experiencias cinematográficas únicas.
- **Creación de contenido**: permiten desarrollar imitaciones de celebridades y parodias de programas de televisión. Los artistas y creadores utilizan los *deepfakes* como una herramienta creativa para explorar nuevas formas de expresión artística y narrativa, permitiendo la creación de obras de arte multimedia innovadoras y experimentales.
- **Educación**: se emplea para crear tutoriales y cursos de formación, además de crear simulaciones interactivas y realistas que ayuden a educar y sensibilizar sobre diversos temas. Por ejemplo, un docente podría utilizar un *deepfake* para explicar un concepto complejo o para presentar una lección de manera más interesante.
- **Locutores y periodistas:** *Deepfake* sería capaz de romper la barrera del idioma en las videoconferencias y medios, traduciendo el habla y alterando simultáneamente los movimientos faciales y bucales para mejorar el contacto visual y hacer que parezca que todos hablan el mismo idioma.
- **Salud:** la tecnología *deepfake* incluso ayuda a las personas con Alzheimer a interactuar con un rostro más joven que puedan recordar. Además,

permite recrear digitalmente el miembro de un amputado o permitir a las personas transexuales verse reflejadas como su género preferido.

- **Marketing:** transforma el comercio electrónico y la publicidad de manera significativa. Por ejemplo, las marcas serían capaces de utilizar tecnología para crear imágenes falsas de modelos, mostrando conjuntos de moda en variedad, con diferentes tonos de piel, estaturas y pesos. Además, los *deepfakes* permiten contenidos superpersonales que convierten a los propios consumidores en modelos: la tecnología permite previsualizar cómo les quedaría un conjunto antes de comprarlo y generar anuncios de moda específicos que varían en función de la hora, el tiempo y el espectador.

2.2.1.2 Desventajas de *deepfake*

Si bien las aplicaciones positivas son significativas, el uso malintencionado de los *deepfakes* representa un riesgo creciente. Estos usos amenazan la privacidad, la seguridad y la integridad informativa. Entre los principales peligros se destacan:

- **Desinformación y manipulación política**: generar videos falsos de políticos o figuras públicas con el fin de desinformar a la opinión pública o influir en procesos electorales.
- **Fraude y extorsión:** los delincuentes serían capaces de falsificar la identidad de individuos comunes y celebridades, haciéndose pasar por ellos mediante ataques de suplantación de identidad, con el fin de extorsionar, robar dinero o manipular situaciones comprometedoras.
- **Pornografía no consentida**: han sido utilizados para producir contenidos sexuales falsos de personas reales sin su consentimiento. Esto plantea serias preocupaciones éticas y de privacidad, especialmente cuando se trata de la creación y difusión de imágenes íntimas sin consentimiento.
- **Sabotaje corporativo y competitivo**: la creación de videos falsos se utiliza como herramienta para dañar la reputación de empresas o personas mediante la simulación de declaraciones o la manipulación de mensajes.
- **Identidad sintética**[3]: mediante esta técnica se elaboran documentos falsos que aparentan ser legítimos. Esto incluye documentos de identificación como pasaportes, licencias de conducir, contratos, certificados educativos o incluso documentos legales.

[3] Para obtener más información visualizar la siguiente aplicación https://onlyfake.net/ y blog https://blog.segu-info.com.ar/2024/02/onlyfake-identidades-falsas-de.html

2.3 Proceso de creación de *deepfakes*: técnicas y herramientas utilizadas

El desarrollo de un *deepfake* es un proceso técnico complejo que se apoya en los avances en *deep learning* y herramientas específicas para la manipulación de imágenes, videos y sonidos. Comprender cómo se generan estas falsificaciones no sólo permite dimensionar su sofisticación, sino que también resulta clave para analizar sus riesgos y diseñar estrategias de detección y mitigación.

La creación de un *deepfake* requiere un conjunto amplio de datos visuales o sonoros de la persona que se desea simular: fotografías, grabaciones de voz, expresiones faciales o secuencias de video. Estos datos son utilizados para entrenar modelos de aprendizaje profundo, capaces de generar contenido sintético a partir de las características aprendidas.

Una de las arquitecturas más utilizadas en este proceso son las Redes Generativas Adversarias, (GAN, por sus siglas en inglés). Este tipo de red se compone de dos redes neuronales que interactúan entre sí: una red generadora, encargada de crear nuevas muestras (imágenes, rostros, gestos) y una red discriminadora, que evalúa la autenticidad de esas muestras intentando distinguir si son reales o generadas. A través de este proceso competitivo, ambas redes mejoran progresivamente, permitiendo obtener resultados cada vez más realistas y difíciles de detectar.

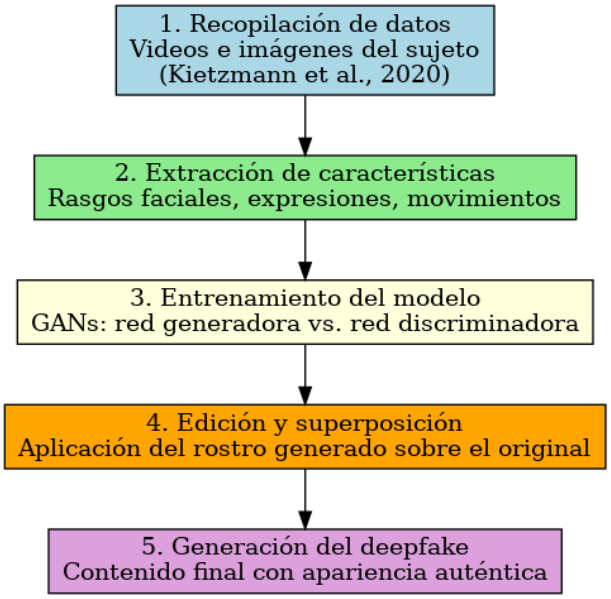
El proceso de creación de *deepfakes* se basa en:

1. **Recopilación de datos**: según Kietzmann, Mills y Plangger (2020) se requiere una gran cantidad de material de origen (por ejemplo, videos del rostro de una persona) para “aprender” sobre las características principales de un sujeto.
2. **Extracción de características:** se utilizan modelos de aprendizaje automático para extraer características relevantes de las imágenes, como la estructura facial, expresiones y movimientos.
3. **Entrenamiento del modelo:** como se mencionó anteriormente, la red generadora crea nuevas imágenes y la red discriminadora distingue entre las imágenes generadas y las reales. Durante el entrenamiento, el modelo ajusta sus parámetros para minimizar la diferencia entre las imágenes generadas y las imágenes de destino reales.
4. **Edición y superposición de imágenes y videos:** después de que el modelo ha sido entrenado, se utiliza para editar y superponer imágenes y videos de la persona en la fuente original.

5. **Generación del *deepfake*:** finalmente, se genera el *deepfake* a partir de la fuente original editada.

Figura 9

Proceso de creación de deepfakes.



Nota: El gráfico ilustra las cinco etapas fundamentales: desde la recopilación de datos del sujeto hasta la generación del contenido final sintético.

Mirky, Y., y Lee, W. (2021) señalan que existe una compensación inherente entre las distintas metodologías de creación de deepfakes, ya que cada una implica diferentes costos y beneficios. La calidad del resultado final no solo depende del tipo de técnica utilizada, sino también de la cantidad de datos empleados para entrenar el modelo y del nivel de credibilidad que logra alcanzar. En este contexto, se identifican dos aspectos principales:

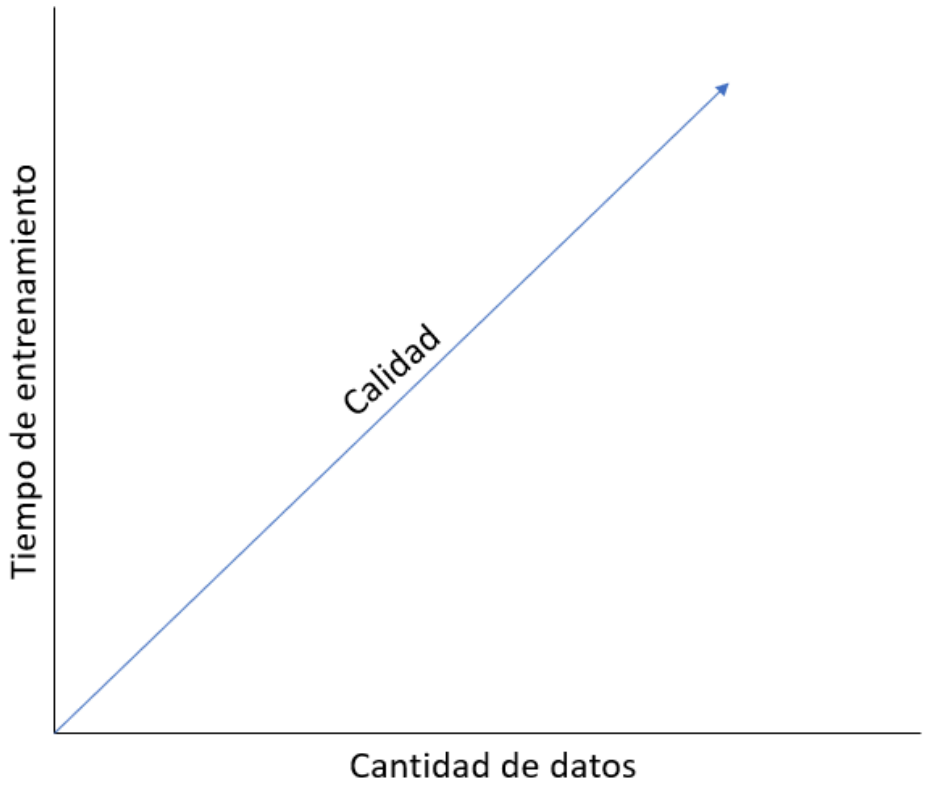
- **Datos vs Calidad:** los modelos entrenados con numerosas muestras del sujeto objetivo tienden a generar los mejores resultados, pero requieren grandes volúmenes de datos y muchas horas de entrenamiento. Este enfoque es más viable para individuos altamente expuestos, como actores, figuras políticas o empresarios reconocidos. Para personas con poca presencia digital, se emplean métodos que necesitan menor cantidad de datos, aunque conllevan una mayor probabilidad de defectos, ya que el modelo debe "imaginar" o inferir parte de la información faltante.
- **Velocidad vs Calidad:** este aspecto varía según el tipo de generación:

- **En línea (online)**: donde el contenido se genera en tiempo real, se opta por modelos de alta resolución o se sacrifica calidad visual para ganar velocidad, lo que resulta en imágenes borrosas o distorsionadas, aunque igualmente efectivas para engañar en situaciones de presión o manipulación social.
- **Fuera de línea (offline)**: Son aquellos difundidos en redes sociales o medios digitales, priorizan la calidad y la coherencia visual, ya que el contenido se produce previamente y debe ser convincente a nivel estético y temporal.

Comprender estas compensaciones permite no solo evaluar la dificultad técnica de generar un *deepfake*, sino también anticipar el tipo de amenaza que representa según el contexto en que se utilice. A continuación, se analizarán las distintas modalidades de ataque basadas en esta tecnología.

Figura 10

Representación conceptual de la generación de deepfake



Nota. Representación conceptual de la generación de deepfake según la cantidad de datos utilizados, el tiempo de entrenamiento y la calidad alcanzada.

2.4 Tipos de ataques con *deepfake*

Los *deepfakes* no solo tienen aplicaciones legítimas, sino que han dado lugar a diversos tipos de ataques maliciosos que suponen riesgos significativos para la seguridad, la privacidad y la integridad de la información. Estos ataques utilizan la capacidad de la tecnología para alterar la realidad, los atacantes son capaces de llevar a cabo desde suplantaciones de identidad hasta fraudes, extorsiones y campañas de desinformación.

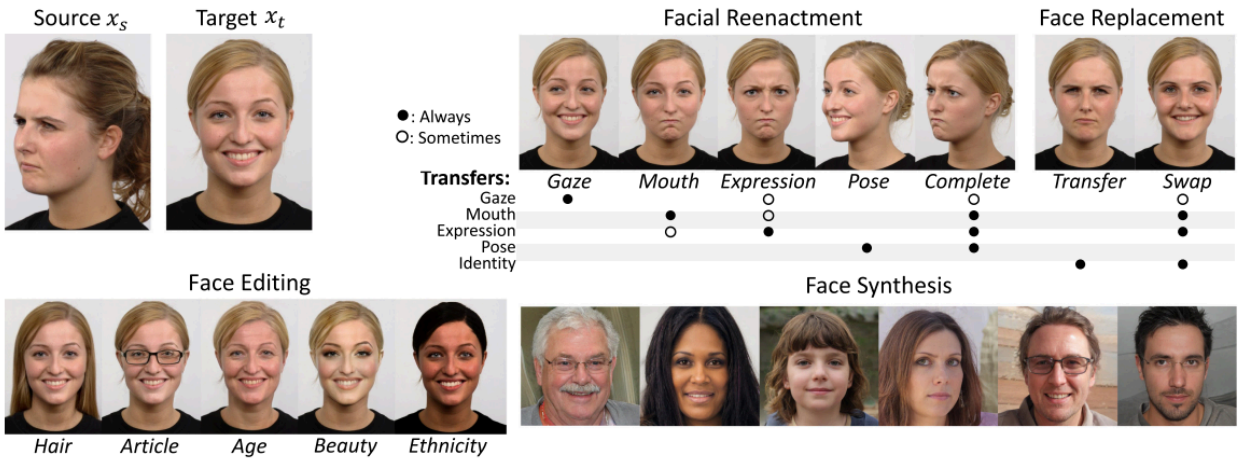
Según Mirky Y. y Lee W. (2021) los ataques con *deepfakes* se agrupan en tres grandes categorías, cada una con diferentes finalidades y niveles de impacto:

- **Reconstrucción (Reenactment)**: da a los atacantes la capacidad de suplantar una identidad, manipulando el discurso y las acciones del objetivo. Esto se utiliza con fines de difamación, pérdida de credibilidad, generación de desinformación o manipulación en entornos políticos y sociales.

- **Reemplazo (Replacement)**: se basa en el intercambio facial, conocido por sus utilizaciones dañinas principalmente en la industria de la pornografía. En estos casos, se reemplaza el rostro de una víctima por el de un actor o actriz para humillar, extorsionar o chantajear, afectando gravemente la reputación y la salud emocional de la persona afectada.
- **Edición y síntesis (Face synthesis)**: emplea herramientas que permiten crear rostros completamente nuevos o alterar características del rostro real. Aunque algunas aplicaciones de este tipo son utilizadas con fines recreativos (como filtros faciales), también facilita la creación de identidades falsas o perfiles sintéticos para cometer fraudes o engañar a otros en entornos digitales.

Figura 11

Representación conceptual de la generación de deepfake



Nota. Ejemplos de deepfakes de recreación, reemplazo, edición y síntesis del rostro humano. Tomado de The Creation and Detection of Deepfakes por Mirsky, Y., & Lee, W. (2021).

Estas formas de ataque reflejan la capacidad que tiene esta tecnología para simular la realidad con una precisión tal que atentan a las nociones tradicionales de veracidad y prueba visual. En un contexto donde la identidad digital es cada vez más relevante, la existencia de *deepfakes* maliciosos exige el desarrollo de estrategias tecnológicas que permitan detectarlos y mitigar sus efectos.

CAPÍTULO 3
SISTEMAS DE DETECCIÓN DE DEEPFAKES

A lo largo del capítulo anterior (Capítulo 2: DEEPFAKE: FUNDAMENTO, CLASIFICACIÓN Y DESAFÍOS), se expuso el funcionamiento de la tecnología *deepfake*, sus tipos, procesos de creación y los riesgos asociados, incluyendo los ataques por reconstrucción, reemplazo y síntesis facial (ver apartado 2.4). En ese contexto, se evidenció que la capacidad de generar contenido sintético con un alto nivel de realismo representa una amenaza creciente para la autenticación de identidades y la confianza en medios digitales.

Frente a estos desafíos, la detección de *deepfakes* se ha consolidado como un campo de estudio crucial, en constante evolución. Si bien anteriormente se introdujeron mecanismos como las pruebas de vida y los sistemas *anti-spoofing* (ver apartado 1.4), estos enfoques ya no resultan suficientes para hacer frente a contenidos generados mediante técnicas avanzadas como las Redes Generativas Adversarias (GAN) (ver apartado 1.6).

En este capítulo se analizarán los principales métodos de detección de *deepfakes*, incluyendo enfoques basados en artefactos visuales, señales fisiológicas y modelos de aprendizaje profundo. Asimismo, se explorarán las vulnerabilidades que afectan a estos sistemas —como los ataques adversariales y la falta de generalización— y se discutirá su integración en entornos sensibles como los sistemas de autenticación biométrica facial. Comprender cómo se detectan los *deepfakes* no solo permite evaluar la solidez de los mecanismos actuales de defensa, sino también proyectar nuevas estrategias para fortalecer la seguridad digital ante la sofisticación creciente de las amenazas.

3.1 Enfoques de detección

La evolución de los *deepfakes* no solo ha revolucionado la generación de contenido sintético, sino que también ha planteado desafíos significativos a los métodos tradicionales de análisis y autenticación. La detección de este tipo de contenido ha emergido como un campo crítico, impulsando el desarrollo de técnicas capaces de identificar patrones sutiles y discrepancias en imágenes y videos manipulados artificialmente.

En el pasado, la autenticidad de imágenes y vídeos se verificaba mediante características diseñadas manualmente, basadas en conocimientos estadísticos o físicos sobre las propiedades de los medios digitales. Sin embargo, como se expone en el apartado *2.2 Detecting Manipulated Videos* del trabajo de Hussain, Neekhara, Dolhansky,

Bitton, Ferrer, McAuley y Koushanfar (2022), los métodos de síntesis se han adaptado para eludir los detectores tradicionales. Esto ha impulsado el desarrollo de enfoques de detección basados en redes neuronales profundas como contramedida frente a estas amenazas emergentes.

Hu, Li y Lyu (2021, p. 2) proponen una clasificación de enfoques de detección en tres grandes categorías:[4]

3.1.1 Métodos basados en rastros y artefactos

Este enfoque se centra en identificar defectos, irregularidades o patrones residuales introducidos por los modelos de síntesis durante la generación del contenido. Estos artefactos incluyen problemas de color, inconsistencias en el ruido digital, o señales estadísticas que delatan la falsificación.

3.1.2 Métodos basados en aprendizaje profundo

Estos métodos utilizan arquitecturas de redes neuronales profundas entrenadas con grandes conjuntos de datos para aprender a distinguir entre contenido genuino y sintético.

Chadha et al. (2021) mencionan que entre los métodos de detección basados en inteligencia artificial, se destacan los siguientes:

- **Análisis secuencial temporal:** este enfoque detecta inconsistencias en la continuidad de los cuadros de un video. Para abordar este problema, se propone un método que utiliza dos tipos de redes neuronales: las redes convolucionales (CNN) y las redes de memoria a corto y largo plazo (LSTM). Estas son capaces de examinar las características visuales de cada cuadro y reconocer patrones a lo largo del tiempo. En esencia, el proceso implica extraer características de cada cuadro utilizando CNN y luego analizar la secuencia temporal de estas características con LSTM para determinar si hay inconsistencias que sugieran un *deepfake*. Este método se probó en un conjunto de datos de 600 vídeos, logrando una precisión superior al 97% (Chadha et al., 2021, p. 559).
- **Método de parpadeo de ojos:** utiliza señales fisiológicas para distinguir entre videos auténticos y *deepfakes*. El estudio, realizado por Li et al. (2018), extrae áreas específicas de los ojos en cuadros de video y las somete a un análisis de redes convolucionales recurrentes a largo plazo (LRCN). Este estudio permite

[4] Hu, S., Li, Y., & Lyu, S. (2021, June). Exposing GAN-generated faces using inconsistent corneal specular highlights. En ICASSP 2021 – 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 2500–2504). IEEE. Recuperado de https://arxiv.org/pdf/2009.11924.pdf (ver p. 2).

predecir la probabilidad de que los ojos estén abiertos o cerrados en el video. Los resultados preliminares obtenidos a partir de un conjunto experimental de datos de 49 vídeos, muestran un rendimiento prometedor en la detección de *deepfakes*, lo que sugiere un método efectivo para abordar la manipulación de videos.

- **Método de cápsulas forenses:** éste método depende de las entradas, si es un video, el sistema separará los cuadros, si la tarea es encontrar cuadros generados por computadora, se dividirán en partes más pequeñas, en cambio, si se quiere detectar rostros falsos, se recortará el área facial usando un algoritmo de detección de rostros. En general, cuanto más grande sea la entrada, más exactos son los resultados. Luego de su pre-procesamiento, cada cuadro pasa por una red llamada VGG-19, la cual es entrenada por muchos datos de imágenes. Luego, se evalúa usando la red de cápsulas. Al final, se obtiene una puntuación para saber si las imágenes son generadas por computadora. Si la entrada es un video, se promedian las puntuaciones de todos los cuadros para obtener la respuesta final.
- **Análisis forense de medios digitales**: utiliza diversos tipos de análisis y enfoques para desarrollar tecnologías y dispositivos que detecten contenido *deepfake*. El método basado en CNN es un método forense utilizado para detectar imágenes modificadas y asegurar la autenticidad de una imagen de cámara. Las manipulaciones de bajo nivel, como los cuadros duplicados en el video, se detectan a través de la forensia digital basada en video.

3.1.3 Métodos basados en inconsistencias físicas/fisiológicas

Estos métodos se basan en principios físicos o fisiológicos del rostro humano que suelen mantenerse constantes en grabaciones reales. Al estar ligados a propiedades biológicas, tienden a ser más confiables ante ataques adversariales y ofrecen interpretaciones más intuitivas.

Se analizan, por ejemplo, las poses de cabeza, la asimetría facial, la distribución de puntos de referencia o los patrones de parpadeo para detectar incongruencias típicas en contenidos generados por GAN.

Hu, Li y Lyu (2021) proponen un nuevo método de detección basado en aspectos fisiológicos o físicos de los rostros generados por GAN. Este método se centra en una característica específica de los ojos, llamada *reflejos especulares corneales*, que son las imágenes de la luz que se refleja en la córnea.

3.1.3.1 Reflejos especulares corneales

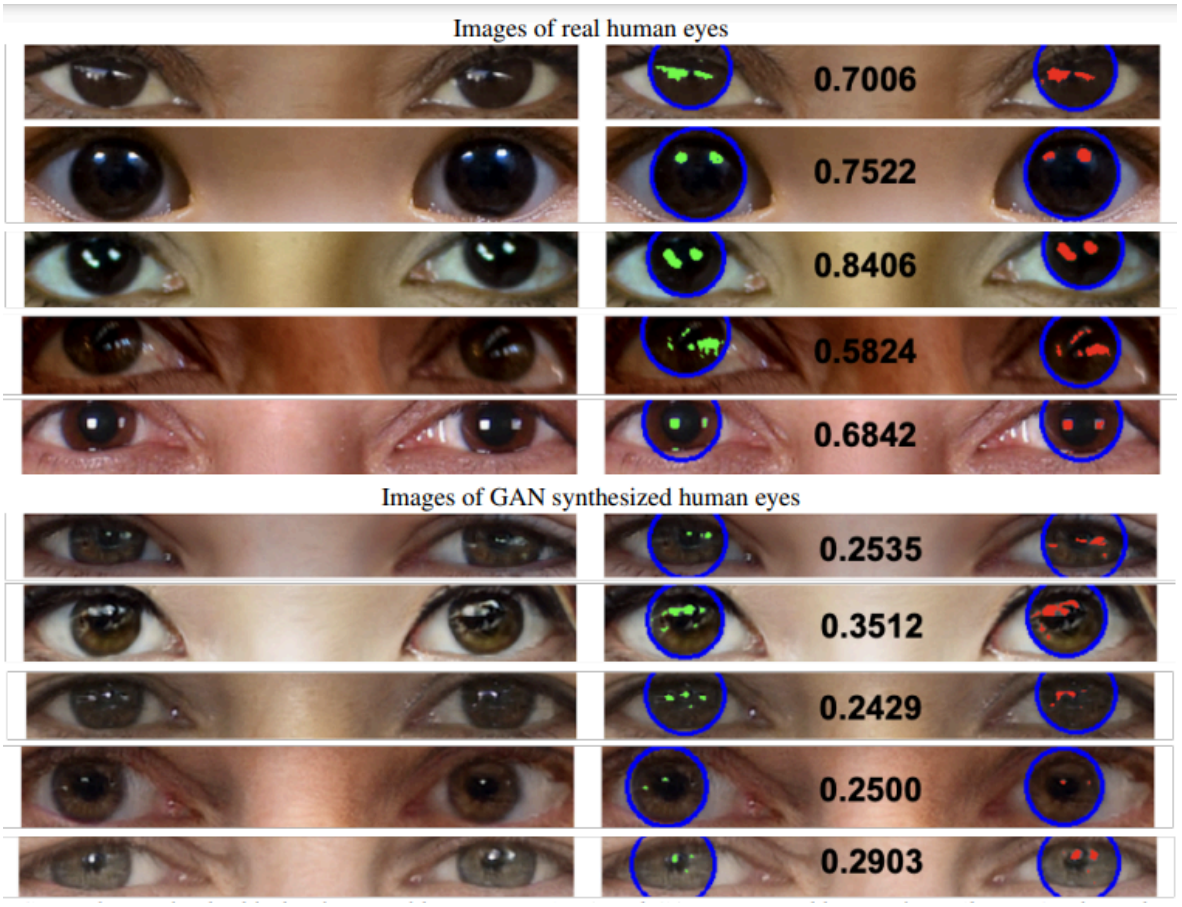
La técnica se basa en la superficie reflectiva de las córneas, que actúan de forma similar a un espejo, permitiendo detectar inconsistencias en los patrones de luz reflejada

y, por lo tanto, identificar si la persona en el video es real o sintetizada. Este método ha mostrado ser particularmente efectivo cuando el sujeto mira directamente a la cámara, aunque existen probabilidades de generar falsos positivos si no se cumple esta condición. En escenarios donde la persona está mirando a la cámara y las fuentes de luz están a una distancia considerable (como en un entorno de retrato), los reflejos en los ojos deberían ser consistentes, lo que hace que este método sea útil para detectar rostros sintetizados por GAN al aprovechar las inconsistencias en dichos reflejos.

Fernández (2021) señala que científicos de la Universidad de Buffalo han desarrollado una herramienta capaz de detectar videos alterados con un 94% de precisión mediante un análisis similar de los reflejos corneales.

Figura 12

Comparación ojos reales y generados por GAN



Nota. Reflejos especulares corneales de ojos humanos reales (arriba) y rostros humanos generados por GAN. Tomado de Exposing gan-generated faces using inconsistent corneal specular highlights por Hu, Li y Lyu (2021).

Hu, Li y Lyu (2021) afirman que, si bien este tipo de métodos de detección de imágenes sintetizadas por GAN poseen un alto rendimiento, también presentan

limitaciones, incluidas la falta de interpretabilidad de los resultados de detección, la escasa solidez frente a operaciones de blanqueo y ataques de adversarios, y la escasa generalización entre distintos métodos de síntesis.

3.2 Vulnerabilidades de los sistemas de detección de *deepfake*

A pesar de los avances significativos en los métodos de detección, los sistemas actuales siguen siendo vulnerables a diversas amenazas que comprometen su eficacia. La continua evolución de las técnicas de generación de contenido sintético y las estrategias cada vez más sofisticadas de los atacantes exigen un análisis profundo de las debilidades que afectan a los detectores.

A continuación, se presentan las principales vulnerabilidades identificadas.

3.2.1 Ejemplos adversariales

Según (Hussain et al., 2022), los ejemplos adversariales son entradas diseñadas específicamente para engañar a un modelo de inteligencia artificial, como una red neuronal, para que produzca un resultado erróneo o incorrecto. Esto implica realizar pequeñas modificaciones en los píxeles de la imagen, de manera cuidadosa y estratégica, para que el modelo de detección no pueda reconocer la manipulación y la interprete incorrectamente como una imagen legítima.

El objetivo de estos ataques es aprovechar las limitaciones inherentes de los algoritmos de detección, haciendo que interpreten contenido manipulado como auténtico. Este tipo de vulnerabilidad representa una amenaza considerable para la integridad de los mecanismos de defensa, ya que deepfakes cuidadosamente construidos logran pasar desapercibidos.

3.2.2 Falta de robustez

Los sistemas de detección son vulnerables a ciertas formas de manipulación que no han sido anticipadas durante su desarrollo. Esto implica que, si un modelo ha sido entrenado con ciertos tipos de *deepfakes*, es probable que falle al enfrentarse a nuevos métodos o variantes de síntesis.

Esta falta de robustez limita la adaptabilidad de los detectores ante técnicas emergentes y pone en evidencia la necesidad de modelos más flexibles y resilientes frente a manipulaciones no anticipadas.

3.2.3 Transferibilidad de los ataques

Una vulnerabilidad especialmente preocupante es la posibilidad de que un *deepfake* diseñado para evadir un sistema de detección específico también logre engañar a otros detectores. Esta propiedad, conocida como transferibilidad, amplía el alcance de los ataques al permitir que una misma manipulación tenga éxito contra múltiples defensas, incluso aquellas no vistas durante su generación.

Este fenómeno expone una fragilidad común en los modelos de detección, especialmente cuando se basan en arquitecturas similares o comparten conjuntos de datos de entrenamiento.

3.2.4 Complejidad de detección

El creciente realismo de los *deepfakes* hace que las diferencias entre contenido genuino y manipulado sean cada vez más sutiles. Esto impone una dificultad creciente para los detectores, que deben distinguir entre variaciones naturales y signos de manipulación sin generar falsos positivos ni negativos.

A medida que los modelos de generación mejoran, los algoritmos de detección deben ser igualmente precisos y sofisticados, lo que representa un desafío técnico y computacional constante.

3.2.5 Escasez de datos de entrenamiento

La calidad y diversidad de los datos utilizados en el entrenamiento de los sistemas de detección son determinantes para su desempeño. Sin embargo, existe una disponibilidad limitada de bases de datos etiquetadas con *deepfakes* realistas y casos auténticos bien documentados.

Esta escasez dificulta la creación de modelos generalizables y robustos, ya que una cobertura insuficiente de escenarios posibles derivan en baja precisión o en una alta tasa de errores al momento de enfrentarse a nuevos ejemplos.

En síntesis, si bien los detectores actuales representan un avance importante frente a la amenaza de los *deepfakes*, su despliegue efectivo en contextos críticos requiere contemplar escenarios adversos y mejorar su resiliencia frente a manipulaciones sofisticadas y perturbaciones imperceptibles.

3.3 Sistemas de detección de *deepfake* en aplicaciones de autenticación biométrica facial

Como se abordó en el Capítulo 1, la autenticación biométrica facial ha transformado los sistemas de verificación de identidad al basarse en características únicas e inherentes al rostro humano (ver apartado 1.5). Sin embargo, también se mencionaron sus desafíos de seguridad (ver apartado 1.6), frente a los cuales surgieron mecanismos como las pruebas de vida y los sistemas *anti-spoofing*. No obstante, estos enfoques, aunque eficaces frente a ataques tradicionales de presentación (como el uso de fotos, máscaras o videos), no son suficientes para detectar contenidos sintéticos generados mediante técnicas más sofisticadas como los *deepfakes* (ver apartado 2.2 y apartado 2.4).

En este contexto, se vuelve indispensable considerar un tercer pilar de protección: **los sistemas de detección de *deepfakes***, diseñados específicamente para identificar alteraciones generadas por redes neuronales, en especial aquellas creadas mediante redes GANs. A diferencia de las pruebas de vida tradicionales, que se enfocan en validar que el estímulo biométrico proviene de una fuente viva, los detectores de *deepfake* apuntan a analizar patrones imperceptibles a simple vista, como inconsistencias fisiológicas, artefactos digitales o anomalías temporales en videos (ver apartado 3.1).

La implementación de estos sistemas dentro de aplicaciones de autenticación biométrica facial tiene un objetivo claro: evitar que videos o imágenes generadas por IA sean aceptadas como válidas por los mecanismos de seguridad, lo cual pondría en riesgo la privacidad y la identidad digital de los usuarios. Como señalan Hussain et al. (2022), la creciente accesibilidad a herramientas de manipulación, combinada con ataques adversariales diseñados para evadir sistemas de detección, convierte a los *deepfakes* en una amenaza real para la autenticación segura.

Los sistemas de detección aplicados a contextos de autenticación deben tener en cuenta no sólo la robustez técnica, sino también las condiciones reales de uso, como la calidad de las cámaras, la iluminación o la presencia de múltiples fuentes de datos. Por ejemplo, técnicas como la comparación de reflejos corneales (Hu, Li y Lyu, 2021) podrían integrarse como una capa adicional de verificación, aprovechando propiedades fisiológicas difíciles de replicar de forma sintética.

Por todo lo anterior, la integración de tecnologías de detección de *deepfake* **no debe considerarse un reemplazo**, sino un **complemento esencial** a los mecanismos tradicionales de autenticación biométrica facial. Esta capa adicional no solo refuerza la seguridad técnica del sistema, sino que también fortalece la confianza de los usuarios en

los procesos de identificación digital, en un mundo donde la frontera entre lo real y lo sintético se vuelve cada vez más difusa.

Como se observa en la Figura 13, cada técnica cumple un rol específico en la protección de la identidad digital del usuario.

Figura 13

Protección contra suplantación de identidad en sistemas biométricos con inclusión de detección de deepfake.

Nota. El gráfico representa la evolución de la protección contra suplantación en sistemas biométricos.

CAPÍTULO 4
PROPUESTA DE REQUISITOS PARA LOS SISTEMAS DE DETECCIÓN DE DEEPFAKE

Considerando todo lo analizado en los capítulos anteriores, se evidencia que la detección de *deepfakes* en entornos de autenticación biométrica facial presenta una serie de desafíos que van más allá de la mera identificación de contenido manipulado. En estos contextos, los riesgos asociados no solo compromete la veracidad del contenido visual, sino que deriva en graves consecuencias sobre la identidad digital, la privacidad y la seguridad de los usuarios.

Por este motivo, se vuelve necesario establecer una serie de requisitos que orienten el diseño, la implementación y la evaluación de sistemas de detección de *deepfake* integrados en plataformas de autenticación biométrica facial. Esta propuesta de requisitos busca sentar las bases para un enfoque sistemático y adaptable, que contemple no solo el rendimiento técnico, sino también las condiciones de uso real, las garantías de privacidad y los marcos normativos aplicables.

4.1 Requisitos técnicos

Los sistemas de detección de *deepfakes* en entornos de autenticación biométrica facial deben responder a exigencias técnicas muy específicas, especialmente si se consideran las vulnerabilidades identificadas tanto en los sistemas de autenticación biométrica facial (ver apartado 1.6) como en los propios mecanismos de detección de *deepfakes* (ver apartado 3.2). Estas no sólo tienen que ver con su capacidad de detección, sino también con el tiempo de respuesta, la adaptabilidad a nuevas amenazas y la robustez frente a condiciones reales de uso. A continuación, se detallan los requisitos técnicos fundamentales que deberían guiar el diseño e implementación de estos sistemas.

4.1.1 Resistencia a la adversarialidad

Como se mencionó en el Capítulo 2 (ver apartado 2.3), los *deepfakes* son generados mediante técnicas de aprendizaje profundo que equilibran calidad y velocidad según los objetivos del atacante. Este proceso conlleva un análisis costo-beneficio donde los creadores de *deepfakes* buscan maximizar el realismo minimizando la probabilidad de ser detectados.

En este contexto, el concepto de **adversarialidad** adquiere especial relevancia: se refiere a la capacidad de los atacantes de manipular los datos de entrada —como

imágenes faciales— de modo que engañen intencionalmente a los sistemas de detección, sin ser perceptibles para el ojo humano. Estas perturbaciones sutiles, generadas a través de técnicas de ingeniería adversarial (ver apartado 2.4), explotan las vulnerabilidades de los algoritmos actuales (ver apartado 1.6).

Por lo tanto, un sistema de detección robusto debe estar preparado para identificar y resistir estos ataques. Como señalan Goodfellow et al. (2015), la seguridad frente a adversarios es un componente esencial de la solidez en cualquier modelo de aprendizaje profundo.

4.1.2 Respuesta en tiempo real

La autenticación biométrica facial se ha expandido rápidamente en entornos como el desbloqueo de dispositivos móviles, el control de acceso, los aeropuertos y la banca digital (ver apartado 1.5). En todos estos casos, el tiempo de respuesta es un factor crítico para garantizar tanto la seguridad como la experiencia del usuario.

Incorporar mecanismos de detección de *deepfakes* en estos procesos introduce un desafío técnico adicional: mantener tiempos de procesamiento mínimos sin sacrificar precisión. Para lograrlo, se requiere el uso de hardware especializado como GPUs o TPUs, junto con estrategias de paralelización que permitan realizar inferencias en milisegundos.

De este modo, la detección en tiempo real se vuelve un requisito técnico esencial en aplicaciones de alta exigencia, especialmente cuando está en juego la confiabilidad del sistema y la fluidez en la interacción del usuario.

4.1.3 Detección adaptativa

Los *deepfakes* varian en calidad, realismo y complejidad, dependiendo del nivel técnico del atacante, la cantidad de datos utilizados y las herramientas de generación empleadas (ver apartado 2.3). Por eso, los sistemas de detección deben ser capaces de reconocer tanto *deepfakes* burdos como falsificaciones altamente realistas, con detalles finos en textura, expresiones o gestos.

La detección adaptativa implica desarrollar modelos que aprendan de forma continua, se actualicen ante nuevas amenazas y mantengan su eficacia incluso en escenarios de baja calidad de imagen (ver apartado 3.1). Esto requiere el uso de técnicas de aprendizaje continuo, redes neuronales flexibles y mecanismos de mejora progresiva.

Korshunov y Marcel (2019) sostienen que para ser útiles en entornos reales, los sistemas de detección deben ser resistentes a las variaciones en la calidad de las imágenes, una condición que es frecuente en cámaras de seguridad, videollamadas o

dispositivos móviles.

4.1.4 Análisis de textura y movimiento

Una característica distintiva de los *deepfakes* más sofisticados es su capacidad para imitar tanto características faciales estáticas como dinámicas, incluso bajo diferentes condiciones de iluminación. Esto representa un desafío para los sistemas de autenticación biométrica, los cuales tradicionalmente han sido sensibles a cambios sutiles en expresiones o iluminación ambiental (ver apartado 1.6).

Para abordar esta complejidad, los sistemas de detección deben incorporar **análisis avanzados de textura** (por ejemplo, espectros de luz sobre la piel) y **movimiento facial** (como microexpresiones o desplazamientos naturales en el rostro). Estas señales permiten detectar incoherencias en cómo la luz interactúa con la piel o cómo las expresiones evolucionan a lo largo del tiempo.

Además, estos análisis deben mantenerse eficaces en entornos no controlados, por lo que es clave emplear tecnologías de procesamiento de imágenes robustas, combinadas con aprendizaje automático para mejorar su rendimiento con el tiempo.

En síntesis, los requisitos aquí desarrollados buscan fortalecer la eficacia y confiabilidad de los sistemas de detección de *deepfakes* aplicados a entornos de autenticación biométrica facial. Dado el carácter dinámico y en constante evolución de estas tecnologías, es indispensable que los sistemas no solo logren altos niveles de precisión, sino que también sean resilientes, adaptativos y operativos en tiempo real. La implementación de estos requisitos no solo responde a las vulnerabilidades detectadas, sino que constituye una base esencial sobre la cual deben diseñarse soluciones efectivas y sostenibles en contextos reales de uso.

4.2 Requisitos operacionales

Además de los requisitos técnicos, los sistemas de detección de *deepfakes* deben satisfacer una serie de condiciones operativas para ser viables en contextos reales de autenticación biométrica facial. Estas condiciones están vinculadas a la eficiencia en el uso de recursos, la interoperabilidad con otras tecnologías, la facilidad de integración y actualización, y la escalabilidad del sistema. Tales desafíos operativos, ya anticipados por las vulnerabilidades técnicas y funcionales mencionadas en el apartado 1.6 y apartado 3.2, deben ser atendidos para garantizar una implementación efectiva y sostenible.

4.2.1 Colaboración, integración y recursos

La detección de *deepfakes* debe poder integrarse eficientemente en los diversos entornos donde opera la autenticación biométrica facial: desde celulares y notebooks hasta sistemas de control de acceso. Cada una de estas plataformas presenta particularidades técnicas, capacidades de procesamiento, tipos de sensores y restricciones energéticas diferentes.

Integrar la detección de *deepfakes* en este ecosistema heterogéneo implica diseñar soluciones modulares, escalables y compatibles con múltiples sistemas operativos y configuraciones de hardware. Además, se deben considerar condiciones variables como iluminación, calidad de cámara y conectividad. Para ello, resulta clave la colaboración entre desarrolladores, fabricantes de hardware y proveedores de seguridad.

En ese sentido, la detección de *deepfakes* no debe actuar de forma aislada, sino como una capa complementaria dentro de arquitecturas más amplias de seguridad.

En este contexto, se destacan áreas clave con las cuales la detección de *deepfakes* debe integrarse de manera eficaz:

- **Sistemas de gestión de acceso**: se encargan de interpretar los resultados del proceso de autenticación y tomar decisiones en tiempo real sobre permitir o denegar el ingreso a una plataforma o recurso. Este proceso fue abordado en el apartado 1.2, donde se detallaron las etapas de identificación, autenticación, autorización y auditoría. Integrar la detección de *deepfakes* en este flujo permite reforzar el control de acceso con una capa adicional de seguridad frente a manipulaciones digitales.
- **Bases de datos de identidades**: son repositorios donde se almacenan los datos biométricos previamente registrados de los usuarios, como imágenes faciales, huellas digitales o patrones de voz. Su función es servir como referencia para validar la autenticidad de una nueva muestra durante un proceso de autenticación. Tal como se desarrolló en el apartado 1.4.1, estas bases forman parte de la construcción de la identidad digital del usuario, la cual debe resguardarse ante intentos de suplantación. La detección de *deepfakes* debe colaborar estrechamente con estas bases para asegurar que las imágenes o videos utilizados no hayan sido manipulados y correspondan efectivamente a registros legítimos.
- **Pruebas de vida y *anti-spoofing***: estas tecnologías verifican que el estímulo biométrico proviene de una persona real, viva y no de un intento de suplantación mediante medios físicos o digitales. Tal como se detalló en el apartado 1.4.2, la detección de *deepfakes* funciona como un complemento estratégico a estas

técnicas, especialmente ante falsificaciones sintéticas indetectables mediante pruebas convencionales.

En conjunto, la colaboración entre diferentes tecnologías y la adecuada gestión de los recursos son esenciales para asegurar la compatibilidad y la interoperabilidad, garantizando así la efectividad de la detección de *deepfakes* en una amplia variedad de aplicaciones de autenticación biométrica facial.

4.2.2 Eficiencia energética

Los dispositivos móviles, uno de los principales entornos donde se aplica la autenticación facial, requieren que todo proceso adicional sea altamente eficiente en términos de consumo energético. Por ello, los algoritmos de detección de *deepfakes* deben estar optimizados para funcionar con el menor uso posible de batería, sin comprometer la precisión ni la experiencia del usuario.

Este requisito se relaciona directamente con lo tratado en el apartado 4.1.2 sobre respuesta en tiempo real, ya que cuanto más rápido y eficiente sea el proceso, menor será el consumo de energía y más fluida será la interacción con el sistema.

4.2.3 Costo computacional

La detección de *deepfakes* implica un procesamiento intensivo, especialmente cuando se aplican modelos complejos de aprendizaje profundo o análisis de video cuadro por cuadro, como los mencionados en el apartado 3.1.2. Estas tareas requieren una significativa capacidad de cómputo, lo que representa un desafío, especialmente en aplicaciones que requieren respuestas rápidas o se ejecutan en dispositivos con recursos limitados.

En este contexto, es importante considerar que en muchas implementaciones actuales de autenticación biométrica facial, el procesamiento no se realiza de manera local en el dispositivo, sino que se ejecuta a través de servicios en la nube. Esta arquitectura permite correr modelos más complejos y pesados sin saturar los recursos del dispositivo, pero también introduce nuevas dependencias, como la necesidad de conectividad estable, posibles latencias y la gestión segura de los datos biométricos transmitidos. En particular, este enfoque es común en aplicaciones de verificación de identidad en tiempo real, como plataformas bancarias, sistemas gubernamentales o controles de acceso remotos.

Por lo tanto, los sistemas de detección deben estar diseñados para **equilibrar el costo computacional con la eficiencia**, tanto en soluciones que operan completamente en la nube como en aquellas que adoptan un enfoque híbrido (edge computing + nube).

Esto implica optimizar modelos para consumir menos recursos sin comprometer su capacidad de detección, aplicar técnicas de compresión de redes neuronales, y aprovechar hardware acelerador cuando esté disponible (como GPUs o TPUs).

Además, dado el crecimiento constante de usuarios y dispositivos que emplean autenticación biométrica facial, es fundamental que estos sistemas mantengan la capacidad de **escalar** sin degradar su rendimiento. Esta escalabilidad debe contemplar tanto el aumento del volumen de datos procesados como la diversidad de entornos operativos, asegurando una experiencia ágil, segura y sostenible.

4.3 Requisitos de seguridad y privacidad

4.3.1 Precisión y confiabilidad

En los sistemas de autenticación biométrica facial, la tolerancia a errores es mínima. Los falsos positivos (aceptar a un impostor) y los falsos negativos (rechazar al usuario legítimo) comprometen tanto la seguridad del sistema como la confianza del usuario. Esta problemática ha sido abordada previamente en los apartados 1.4.2 y 3.1, donde se presentaron tecnologías como pruebas de vida y sistemas *anti-spoofing* que ayudan a reducir estos errores, pero que deben complementarse con mecanismos de detección más sofisticados frente a manipulaciones digitales como los *deepfakes*.

La implementación de modelos de detección debe incorporar algoritmos basados en aprendizaje profundo, como redes neuronales convolucionales (CNNs), que permitan reconocer sutiles diferencias en imágenes o secuencias de video y adaptarse continuamente a nuevas formas de ataque mediante técnicas de detección adaptativa (como se explicó en el apartado 4.1.3).

4.3.2 Protección de datos

La protección de los datos biométricos es un eje central cuando se incorporan tecnologías de detección de *deepfakes*. Tal como se explicó en el apartado 1.4.1 sobre identidad digital, y el apartado 1.1 de privacidad y biometría, las características biométricas forman parte de un perfil personal irremplazable, y su exposición representa un riesgo permanente. Por tanto, la seguridad de la información debe estar garantizada desde el diseño del sistema, aplicando medidas de cifrado, control de acceso, trazabilidad y almacenamiento seguro.

Asimismo, la detección debe estar preparada para enfrentar ataques adversariales, donde los *deepfakes* son creados con el objetivo explícito de evadir los sistemas de seguridad (ver apartado 3.2). La colaboración entre actores del ecosistema (desarrolladores, empresas, investigadores y organismos de control) también es

fundamental para asegurar que estos sistemas cumplan con marcos regulatorios y estándares de seguridad, sin sacrificar interoperabilidad ni experiencia de usuario.

4.3.3 Prevención del abuso

El despliegue de tecnologías de detección y autenticación biométrica debe ir acompañado de una mirada ética y preventiva. No alcanza con que los sistemas sean técnicamente eficientes: deben incorporar medidas que eviten su utilización para fines ilícitos o invasivos. Esto incluye la implementación de políticas claras de acceso a los datos, auditorías regulares, mecanismos de monitoreo de actividad, y barreras que impidan que terceros no autorizados accedan, modifiquen o exploten la información biométrica.

La concienciación de los usuarios también es clave. Informar sobre los riesgos del uso indebido de datos faciales y promover buenas prácticas fortalece la seguridad desde una perspectiva participativa. Al mismo tiempo, la cooperación entre desarrolladores, reguladores y expertos en ciberseguridad permite crear un ecosistema robusto, responsable y respetuoso de los derechos de las personas.

4.4 Requisitos de legalidad

En el marco del uso de datos biométricos faciales y su análisis mediante sistemas de detección de *deepfakes*, es indispensable que su desarrollo e implementación se ajusten a marcos normativos y principios legales vigentes. A diferencia de los requisitos técnicos u operativos, los requisitos de legalidad no se limitan al diseño funcional del sistema, sino que abarcan el respeto por los derechos fundamentales de los usuarios. Como se expuso en el apartado 1.4.1 sobre identidad digital, el uso de información biométrica implica una dimensión ética y legal que requiere especial atención. En este sentido, la legalidad actúa como garante de que la privacidad y la protección de los datos personales no sean vulneradas bajo ninguna circunstancia.

4.4.1 Consentimiento informado

El uso de imágenes faciales como datos biométricos exige un consentimiento informado claro, voluntario y revocable por parte del usuario, conforme a lo establecido por la legislación vigente (ley 25.326 de Protección de los Datos Personales). Este consentimiento no debe limitarse a un simple casillero marcado en un formulario, sino que debe implicar una comprensión real de qué datos serán recolectados, cómo serán almacenados, procesados y protegidos, con qué fines y durante cuánto tiempo.

Es fundamental garantizar que el usuario tenga acceso a materiales explicativos sobre los riesgos, beneficios y medidas de protección asociadas al uso de sus datos. Asimismo, deben existir mecanismos accesibles para revisar y modificar ese consentimiento en cualquier momento, promoviendo un control activo sobre su propia información biométrica.

La transparencia y la responsabilidad en este proceso son clave para construir una relación de confianza sólida entre el usuario y el sistema. Además, evitar el uso indebido de los datos con fines no autorizados (como marketing, vigilancia o comercialización) resulta indispensable para mantener la legitimidad y la ética en el uso de tecnologías basadas en inteligencia artificial.

4.4.2 Cumplimiento legal

Los sistemas de detección de *deepfakes* que procesan datos biométricos deben diseñarse y operarse en estricto cumplimiento con las regulaciones y normativas aplicables. En el ámbito internacional, esto incluye legislaciones como el Reglamento General de Protección de Datos (RGPD) de la Unión Europea, que establece requisitos exigentes para la recolección, almacenamiento y tratamiento de datos personales.

En el contexto argentino, rige la Ley 25.326 de Protección de los Datos Personales, la cual reconoce el derecho de toda persona a controlar la información que se recolecta sobre ella y a ser debidamente informada sobre su uso. Esta ley exige que el tratamiento de datos se realice únicamente con el consentimiento libre, expreso e informado del titular, e impone responsabilidades a quienes almacenan o tratan datos personales, incluyendo la implementación de medidas técnicas y organizativas adecuadas para garantizar su seguridad.

El cumplimiento normativo no sólo protege a las organizaciones frente a posibles sanciones legales, sino que también fortalece la confianza de los usuarios y de los distintos actores involucrados, al evidenciar un compromiso ético y responsable con la privacidad.

Asimismo, los sistemas deben estar preparados para adaptarse a eventuales cambios en el marco normativo, incorporando procesos de revisión periódica de políticas, realización de auditorías y monitoreos constantes. Esto resulta especialmente relevante en entornos globales, donde los datos son procesados en múltiples jurisdicciones. En estos casos, es fundamental garantizar la interoperabilidad legal, es decir, que los sistemas cumplan con los marcos regulatorios de todos los territorios donde operen.

Finalmente, las organizaciones optan por obtener certificaciones de buenas prácticas en protección de datos, lo que contribuye no sólo a reducir riesgos legales, sino

también a posicionar su imagen institucional como referente en responsabilidad tecnológica.

A modo de cierre, la Figura 14 presenta una síntesis de los requisitos fundamentales propuestos para abordar la detección de *deepfakes* en aplicaciones que utilizan autenticación biométrica facial. Esta clasificación —agrupada en dimensiones técnicas, operacionales, de seguridad y privacidad, y legales— busca ofrecer una visión estructurada y transversal que facilite su análisis, implementación y futura adaptación en entornos reales.

Figura 14

Requisitos clave para la detección de deepfakes en sistemas de autenticación biométrica facial.

Nota. El gráfico representa los requisitos propuestos para la detección de deepfake en aplicaciones que utilicen autenticación biométrica facial.

CONCLUSIONES

A lo largo de este trabajo se abordó un problema emergente y altamente relevante: la amenaza que representan los *deepfakes* en entornos de autenticación biométrica facial. Si bien su uso malicioso aún no se ha manifestado de forma masiva en estos sistemas, el avance acelerado de la inteligencia artificial, sumado a la creciente accesibilidad de herramientas para la generación de contenido sintético, configura un escenario preocupante que requiere atención inmediata. La posibilidad de que una identidad sea suplantada con alta fidelidad compromete gravemente no solo la confidencialidad de los datos personales, sino también la integridad de las infraestructuras críticas que dependen de estos mecanismos de autenticación.

Esta tesina propuso, como eje central, una serie de requisitos estructurados que sirvan de guía para el diseño e implementación de sistemas de detección de *deepfakes* aplicados a la autenticación biométrica facial. Para ello, se analizaron los fundamentos técnicos del reconocimiento facial, el funcionamiento y evolución de los *deepfakes*, así como los métodos existentes para su detección, identificando además las vulnerabilidades tanto de los sistemas de autenticación como de los propios mecanismos de defensa. A partir de este análisis, se definieron requisitos divididos en cuatro categorías: técnicos, operacionales, de seguridad y privacidad, y legales.

Entre los requisitos más destacados se encuentran la resistencia a la adversarialidad, la respuesta en tiempo real, la detección adaptativa y la protección de los datos biométricos, así como el consentimiento informado y el cumplimiento normativo. Estos elementos no deben entenderse de manera aislada, sino como un conjunto articulado que busca asegurar no sólo la efectividad técnica del sistema, sino también su sostenibilidad legal y su aceptación por parte de los usuarios.

Más allá del valor técnico, esta tesina plantea una mirada preventiva y estratégica: anticiparse a los ataques antes de que se conviertan en una amenaza consolidada. En un contexto donde la autenticación biométrica facial es cada vez más común —desde dispositivos móviles hasta sistemas bancarios o sanitarios—, la integración de detección de *deepfakes* deja de ser una opción y se convierte en una necesidad urgente.

Finalmente, este trabajo busca no solo aportar claridad en un campo complejo y en constante evolución, sino también abrir el camino para futuras investigaciones. La seguridad de la identidad digital se perfila como uno de los principales desafíos de esta década. En este sentido, esta tesina se plantea como un punto de partida para el desarrollo de tecnologías éticas, robustas y centradas en la protección del usuario frente a los riesgos del contenido sintético.

REFERENCIAS BIBLIOGRÁFICAS

1Password. (2022). *Acerca de la seguridad de Face ID en 1Password para iOS.* https://support.1password.com/face-id-security/

Aguirrezabala, A. M. (2015). *Estudio de verificación biométrica de voz*. [Tesis de máster, Universidad Politécnica de Madrid]. https://oa.upm.es/38115/1/TESIS_MASTER_MARTA_AGUIRREZABALA_AGUSTIN.pdf

Banco de la Provincia de Buenos Aires. (s.f.). Validación biométrica - Reconocimiento facial – Preguntas frecuentes. https://www.bancoprovincia.com.ar/CDN/Get/Validacion_biometrica_facial_FAQ

Borghello, C., & Temperini, M. G. (2012). *Suplantación de Identidad Digital como delito informático en Argentina.* En X Simposio Argentino de Informática y Derecho (SID 2012), XLI JAIIO (La Plata, 27 al 31 de agosto de 2012).

Borja, C. T., & Bueno, Á. G. (2006). *Sistemas biométricos*. https://www.studocu.com/co/document/universidad-nacional-abierta-y-a-distancia/construccion-sustentable/trabajo-biometria-hxjxicocococ/19077127

Buolamwini, J., & Gebru, T. (2018). *Gender shades: Intersectional accuracy disparities in commercial gender classification*. In Conference on fairness, accountability and transparency (pp. 77-91). PMLR. https://proceedings.mlr.press/v81/buolamwini18a.html?mod=article_inline&ref=akusion-ci-shi-dai-bizinesumedeia

Chadha, A., Kumar, V., Kashyap, S., & Gupta, M. (2021). Deepfake: An overview. En Singh, P. K., Wierzchoń, S. T., Tanwar, S., Ganzha, M., & Rodrigues, J. J. P. C. (Eds.). (2021). Proceedings of Second International Conference on Computing, Communications, and Cyber-Security. Lecture Notes in Networks and Systems. y (pp. 557–567). https://doi.org/10.1007/978-981-16-0733-2

Chowdhury, M., Gao, J., & Islam, R. (2017). *Biometric Authentication Using Facial Recognition.* Security and Privacy in Communication Networks (pp. 287–295). https://doi.org/10.1007/978-3-319-59608-2_16

Fernández, M. (2021). *El truco para detectar deepfakes: sólo hay que mirarle bien a los ojos.* https://www.elespanol.com/omicrono/software/20210312/truco-detectar-deepfakes-solo-mirarle-bien-ojos/565444158_0.html

Gomez, N. (2020). *DeepFakes, autenticación facial y pruebas de detección de vida.* https://reconoserid.com/deepfakes-autenticacion-facial-y-pruebas-de-deteccion-de-vida/

Goodfellow, I. J., Shlens, J., & Szegedy, C. (2015). *Explaining and harnessing adversarial examples*. arXiv preprint arXiv:1412.6572. https://arxiv.org/abs/1412.6572

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* MIT Press. https://doi.org/10.1007/s10710-017-9314-z

Grupo Ático34. (2019). *Suplantación de identidad, te puede ocurrir a ti.* https://protecciondatos-lopd.com/empresas/suplantacion-de-identidad/

Hu, S., Li, Y., & Lyu, S. (2021, June). *Exposing GAN-generated faces using inconsistent corneal specular highlights*. In ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 2500-2504). IEEE. https://arxiv.org/pdf/2009.11924.pdf

Hussain, S., Neekhara, P., Dolhansky, B., Bitton, J., Ferrer, C. C., McAuley, J., & Koushanfar, F. (2022). *Exposing vulnerabilities of deepfake detection systems with robust attacks*. Digital Threats: Research and Practice (DTRAP), 3(3), 1-23. : https://dl.acm.org/doi/full/10.1145/3464307

IBM (2025). *Conceptos de seguridad: Identificación y autenticación.* https://www.ibm.com/docs/es/ibm-mq/9.2?topic=mechanisms-identification-authentication

Instituto nacional de estándares y tecnología (NIST). (2013). *Standards for Biometric Technologies.* https://www.nist.gov/speech-testimony/standards-biometric-technologies

Kakkirala, K. R., Chalamala, S. R., & Jami, S. K. (2017). *Thermal Infrared Face Recognition: A Review.* En 2017 UKSim-AMSS 19th International Conference on Computer Modelling & Simulation (UKSim). https://doi.org/10.1109/uksim.2017.38

Kietzmann, J., Mills, A. J., & Plangger, K. (2020). *Deepfakes: perspectives on the future "reality" of advertising and branding*. International Journal of Advertising, 40(3), 473–485. https://doi.org/10.1080/02650487.2020.1834211

Korshunov, P., & Marcel, S. (2019, June). *Vulnerability assessment and detection of deepfake videos.* In 2019 International Conference on Biometrics (ICB) (pp. 1-6). IEEE. https://doi.org/10.1109/ICB45273.2019.8987375

LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning.* Nature, 521(7553), 436-444. https://doi.org/10.1038/nature14539

Ley 25.326. (2000). *Protección de los Datos Personales.* Boletín Oficial de la República Argentina, 2 de noviembre de 2000. https://www.argentina.gob.ar/normativa/nacional/ley-25326-64790/texto

Li, Y., Chang, M. C., & Lyu, S. (2018). *In ictu oculi: Exposing ai created fake videos by detecting eye blinking.* In 2018 IEEE International workshop on information

forensics and security (WIFS) (pp. 1-7). Ieee. https://doi.org/10.1109/WIFS.2018.8630787

Li, S. Z., & Jain, A. K. (Eds.). (2011). *Handbook of face recognition (2.ª ed.)*. Springer. https://doi.org/10.1007/978-0-85729-932-1

Mendicoa, G. E. (2003). Sobre tesis y tesistas: Lecciones de enseñanza-aprendizaje (1ra ed.). Editorial Espacio. https://campusvirtual.icap.ac.cr/pluginfile.php/223829/mod_resource/content/1/Mendicoa%20-Sobre-Tesis-y-Tesistas-Lecciones-de-Ensenanza-aprendizaje.pdf

Mirsky, Y., & Lee, W. (2021). *The Creation and Detection of Deepfakes.* ACM Computing Surveys, 54(1). https://doi.org/10.1145/3425780

Natgunanathan, I., Mehmood, A., Xiang, Y., Beliakov, G., & Yearwood, J. (2016). *Protection of privacy in biometric data.* IEEE access, 4, 880-892. https://doi.org/10.1109/ACCESS.2016.2535120

Ortiz López, J. (2011). *Sistema de detección de vida vía software en imágenes de iris utilizando criterios de calidad*. [Tesis de grado, Universidad Autónoma de Madrid]. http://hdl.handle.net/10486/7445

Pagnotta S. (2015). *Nueva Guía de Privacidad en Internet, para que protejas tu identidad en la web.* https://www.welivesecurity.com/la-es/2015/09/28/guia-privacidad-en-internet-identidad-web/

Raheem, E. A., Ahmad, S. M. S., & Adnan, W. A. W. (2019). *Insight on face liveness detection: A systematic literature review.* International Journal of Electrical & Computer Engineering (2088-8708), 9(6). http://doi.org/10.11591/ijece.v9i6.pp5165-5175

Real Academia Española. (2016). *Dato biométrico.* https://dpej.rae.es/lema/dato-biométrico

Santana, M. A. G., Díaz-Sánchez, L. E., Paz, I. T., & Huertas, M. R. (2017). *Estado del arte en reconocimiento facial.* Res. Comput. Sci., 140, 19-27. https://rcs.cic.ipn.mx/2017_140/Estado%20del%20arte%20en%20reconocimiento%20facial.pdf

Schmidhuber, J. (2015). *Deep learning in neural networks: An overview*. Neural Networks, 61, 85-117. https://doi.org/10.1016/j.neunet.2014.09.003

Segu-Info. (2018). *Seguridad Lógica - Identificación y Autentificación.* https://www.segu-info.com.ar/logica/identificacion

Silva, M. M. (2023). *Vision Transformers For Face Anti-Spoofing.* [Tesis de maestría, Universidade de Coimbra]. https://estudogeral.uc.pt/handle/10316/113062

Suarez, D., & Guarda, T. (2019). *Sistemas Biométricos aplicados en smartphones.* Revista Ibérica de Sistemas e Tecnologias de Informação, (E17), 25-31. https://www.researchgate.net/profile/Teresa-Guarda/publication/331178385_Biometric_systems_applied_in_smartphones/links/5fabe79aa6fdcc331b947880/Biometric-systems-applied-in-smartphones.pdf

Transportation Security Administration (TSA). (s.f.). *Facial Recognition Technology.* https://www.tsa.gov/news/press/factsheets/facial-recognition-technology

Thales (2016). *Biometría para identificación y autenticación.* https://www.thalesgroup.com/es/countries/americas/latin-america/dis/gobierno/inspiracion/biometria

Unión Europea. (2016). *Reglamento general de protección de datos (RGPD)*. https://eur-lex.europa.eu/eli/reg/2016/679/oj

Westerlund, M. (2019). *The emergence of deepfake technology: A review.* Technology innovation management review, 9(11). https://timreview.ca/article/1282

Wojewidka, J. (2020). *The deepfake threat to face biometrics.* Biometric Technology Today, 2020(2), 5–7. https://doi.org/10.1016/s0969-4765(20)30023-0

Yang, W., Li, W., & Lyu, S. (2019). *Exposing deepfakes using inconsistent head poses.* In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 8261-8265). https://doi.org/10.1109/ICASSP.2019.8683164

Zulfiqar, M., Syed, F., Khan, M. J., & Khurshid, K. (2019). *Deep face recognition for biometric authentication*. In 2019 International Conference on Electrical, Communication and Computer Engineering (ICECCE) (pp. 1–6). IEEE. https://doi.org/10.1109/ICECCE47252.2019.8940632

Zhou, S., & Xiao, S. (2018). *3D face recognition: a survey.* Human-centric Computing and Information Sciences, 8(1), 35. https://doi.org/10.1186/s13673-018-0157-2

ANEXOS

TABLA DE CONTENIDO DE FIGURAS

TABLA DE CONTENIDO DE TABLAS