Artículo de Revisión

Deepfakes y autenticación biométrica: Desafíos y soluciones contra la suplantación de identidad

Deepfakes and biometric authentication: challenges and solutions against identity theft

Deysi Elvia Yuvixa Quiliche Plasencia: Universidad Nacional de Trujillo, Estudiante de Ingeniería en Sistemas,
https://orcid.org/0009-0000-7549-7048

Kelita Marilu Mauricio Saavedra: Universidad Nacional de Trujillo, Estudiante de Ingeniería en Sistemas,
https://orcid.org/0000-0003-0007-7369

Alberto Carlos Mendoza de los Santos: Universidad Nacional de Trujillo, Doctor en Ingeniería de Sistemas e Informática,
https://orcid.org/0000-0002-0469-915X
Autor de correspondencia: dquiliche@unitru.edu.pe

DOI: https://doi.org/10.64424/rcu51202698

Recibido: 19 de febrero 2026
Aprobado: 2 de junio 2026
Publicado: 19 de junio 2026

Resumen:
Los deepfakes en base a la inteligencia artificial generativa representan un desafío crítico en la autenticación biométrica, debido a que permite la creación de identidades sintéticas complicadas de detectar para los sistemas y el ojo humano. El objetivo principal es analizar los desafíos de la autenticación biométrica frente a deepfakes y las soluciones que existen para evitar la suplantación de identidad. Se realizó una revisión sistemática bajo la metodología PRISMA, empleando bases de datos como Google Académico, IEEE Xplore, ScienceDirect, Springer Link y Core, a partir de cadenas de búsqueda relacionadas con deepfake, biometric authentication y identity spoofing. De un total de 18432 resultados iniciales se seleccionaron 23 artículos publicados entre 2020 y 2025. Los hallazgos evidencian que los deepfakes afectan tanto al reconocimiento facial como de voz, generando vulnerabilidades frente a ataques sofisticados que superan técnicas tradicionales de detección. Entre los principales desafíos identificados se encuentran la dependencia de datasets reducidos y la falta de generalización de los modelos en contextos reales. Las soluciones revisadas incluyen enfoques multimodales, detección basada en biometría conductual, integración de blockchain, autenticación multifactor y modelos de aprendizaje continuo, optimizados para tiempo real. Se resalta que existen implicaciones éticas y legales que requieren marcos regulatorios internacionales y códigos de conducta para desarrolladores. Por ende, la seguridad biométrica frente a deepfakes demanda un enfoque integral que combine innovación tecnológica, regulación y educación digital, incentivando futuras investigaciones orientadas a datasets diversos y algoritmos adaptativos.

Palabras clave: Aprendizaje automático, aprendizaje multimodal, blockchain, ciberseguridad, detección de fraudes, inteligencia artificial generativa.

Abstract:
Deepfakes based on generative artificial intelligence represent a critical challenge in biometric authentication, as they allow the creation of synthetic identities that are difficult for systems and the human eye to detect. The main objective is to analyze the challenges of biometric authentication against deepfakes and the solutions that exist to prevent identity theft. A systematic review was conducted using the PRISMA methodology, using databases such as Google Scholar, IEEE Xplore, ScienceDirect, Springer Link, and Core, based on search strings related to deepfake, biometric authentication, and identity spoofing. From a total of 18,432 initial results, 23 articles published between 2020 and 2025 were selected. The findings show that deepfakes affect both facial and voice recognition, generating vulnerabilities against sophisticated attacks that surpass traditional detection techniques. Among the main challenges identified are the dependence on small datasets and the lack of generalization of the models in real-life contexts. The reviewed solutions include multimodal approaches, behavioral biometrics-based detection, blockchain integration, multifactor authentication, and continuous learning models optimized for real-time use. It is emphasized that there are ethical and legal implications that require international regulatory frameworks and codes of conduct for developers. Therefore, biometric security against deepfakes demands a comprehensive approach that combines technological innovation, regulation, and digital education, encouraging future research focused on diverse datasets and adaptive algorithms.

Keywords: Machine learning, multimodal learning, blockchain, cybersecurity, fraud detection, generative artificial intelligence

Como citar la obra: Quiliche-Plasencia, D., E., Y., Mauricio-Saavedra, K., M. y Mendoza-de los Santos, A., C. (2026). Deepfakes y autenticación biométrica: Desafíos y soluciones contra la suplantación de identidad. *Revista Científica Unanchay*, 5(1), 139-154
doi: https://doi.org/10.64424/rcu51202698

Introducción

En los últimos años, ha evolucionado exponencialmente la inteligencia artificial (IA) generativa en el ámbito de la ciberseguridad y sistemas de autenticación biométrica. Este avance ha impulsado el desarrollo de aplicaciones innovadoras, pero también ha facilitado la aparición de nuevas amenazas que comprometen la seguridad de los sistemas de verificación de identidad. En este contexto, los deepfakes emplean técnicas novedosas de la IA para crear información engañosa que representan desafíos significativos en la seguridad y entornos de autenticación y validación. Según Khan et al. (2025), esta tecnología que surgió con el propósito creativo o de entretenimiento, se ha convertido en una amenaza crítica para la integridad de los sistemas de verificación de identidad.

Actualmente, los desarrollos basados en aprendizaje automático provocan que los deepfakes sean altamente creíbles y difíciles de diferenciar entre lo que es real y no, aquello afecta a los sistemas automatizados de reconocimiento biométrico. Como señalan Firc et al. (2023), los deepfakes representan una amenaza constante y emergente, si bien es cierto esta tecnología ha remodelado aplicaciones de manera positiva con asistentes virtuales personalizados y herramientas de clonación de voz para contenido creativo, también se les ha dado un mal uso planteando severas amenazas a la seguridad y confianza, en especial en sistemas financieros y de ciberseguridad.

Los avances y complejidad en la tecnología de los deepfakes aborda nuevos desafíos en el ámbito biométrico. Los sistemas de autenticación biométrica emplean reconocimiento facial o análisis de voz donde se ha evidenciado vulnerabilidades significativas frente ataques engañosos. En el ámbito de la detección forense según Qureshi et al. (2024), los deepfakes representan un conjunto de técnicas que emplean redes neuronales profundas (CNN) que generan contenido multimedia que es indistinguible en comparación con el contenido auténtico. Esta característica ha llevado a un incremento en el fraude de deepfakes a nivel mundial entre el 2022 y 2023.

Las implicaciones son diversas y complejas en los sistemas biométricos. Los nuevos desafíos presentes en la seguridad biométrica se evidencian en especial en las Redes Adversariales Generativas (GANs) y los Autocodificadores Variacionales (VAEs). Así establece Khan et al. (2025), mientras la investigación en IA ha mejorado los métodos para la autenticación biométrica, también estas herramientas producen identidades sintéticas avanzadas y capaces de derrotar las medidas de seguridad. Los sistemas biométricos, aunque altamente efectivos bajo muchas condiciones, enfrentan riesgos significativos debido a las capacidades evolutivas de los deepfakes que pueden explotar vulnerabilidades inherentes en sistemas diseñados para reconocer y autenticar identidades basadas en características físicas o comportamentales de individuos.

Por otro lado, Alazani et al. (2024) explica que el contenido multimedia sintético creado mediante técnicas avanzadas de IA representa una amenaza significativa en varios ámbitos de la seguridad. De igual forma Srivastava et al. (2025), detalla que los enfoques de detección han evolucionado desde métodos simples hacia técnicas sofisticadas que emplean redes neuronales profundas, con transformadores de visión (ViT) logrando precisiones del 94.3% al 95.1% en conjuntos de datos estándar.

Asimismo, Xing et al. (2025), sostiene que los sistemas de anti-suplantación facial enfrentan desafíos significativos ante ataques realistas como máscaras 3D y contenido generado por IA. A su vez Jbara et al. (2024), comenta que las contramedidas actuales incluyen detección y análisis de artefactos neurales, mientras que los deepfakes presentan aplicaciones tanto maliciosas como fraude y robo de identidad, son también benéficas como entretenimiento y educación, creando desafíos únicos para los desarrolladores de seguridad.

El objetivo principal es analizar los desafíos de la autenticación biométrica frente a deepfakes y las soluciones que existen para evitar la suplantación de identidad. La investigación aborda tanto los aspectos técnicos de los ataques como las soluciones tecnológicas emergentes.

Metodología

El presente artículo se llevó a cabo mediante una revisión sistemática de la literatura siguiendo la metodología PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), con el objetivo de responder a la pregunta de investigación: ¿Cuáles son los desafíos que enfrentan los sistemas de autenticación biométrica frente a los deepfakes y qué soluciones existen para evitar la suplantación de identidad?

Según Page et al. (2021), PRISMA brinda un marco metodológico estandarizado, garantizando la transparencia, reproducibilidad y rigor científico en la síntesis de evidencia. Esta metodología ha sido validada y aplicada en múltiples disciplinas, entre ellas ingeniería e informática, donde la revisión sistemática resulta ser fundamental para analizar literatura homogénea y en constante evolución (Moher et al., 2009; Liberati et al., 2009).

Estudios recientes en ingeniería de software y seguridad informática confirman que las revisiones sistemáticas son un enfoque confiable para identificar tendencias, desafíos y vacíos de investigaciones, así como sintetizar evidencia de forma reproducible (Wohlin et al., 2020; Garousi et al., 2020). En ese contexto, la adopción de PRISMA es pertinente debido al carácter multidisciplinario del problema abordado, ya que integra IA, biometría y ciberseguridad y a la necesidad de garantizar un seguimiento metodológico frente a la rápida evolución de los deepfakes.

Para la recolección de la información se emplearon términos como: deepfake, biometric authentication, attacks, etc, estos términos fueron incluidos en diversas cadenas de búsqueda y ejecutados en las siguientes bases de datos: Google Académico, Springer Link, IEEE Xplore, Sciencedirect, Core. La cantidad de resultados obtenidos, número de estudios seleccionados, así como las cadenas de búsqueda empleados se presenta en la Tabla 1.

Tabla 1

Criterios de búsqueda

Base de datos	Cadena de búsqueda	Resultados obtenidos	Estudios seleccionados
Google académico	(biometrics OR biometric authentication) AND (deepfake OR deep fake OR synthetic media) AND (identity theft OR impersonation OR spoofing OR attacks)	17600	13
Springer Link	("deepfake" OR "deepfakes" OR "synthetic media") AND ("biometric authentication" OR "facial recognition" OR "voice authentication" OR "fingerprint recognition") AND ("identity spoofing" OR "impersonation" OR "presentation attack") AND ("security" OR "threats" OR "vulnerabilities") AND ("countermeasures" OR "solutions" OR "detection methods")	86	4
IEEE Xplore	deepfake biometric authentication	52	2
Science Direct	deepfake biometric authentication identity spoofing facial recognition security detection	61	3
Core	("biometric authentication" OR biometrics) AND ("deepfake" OR "deep fake" OR "synthetic media") AND ("identity theft" OR spoofing OR impersonation OR attacks)	637	1

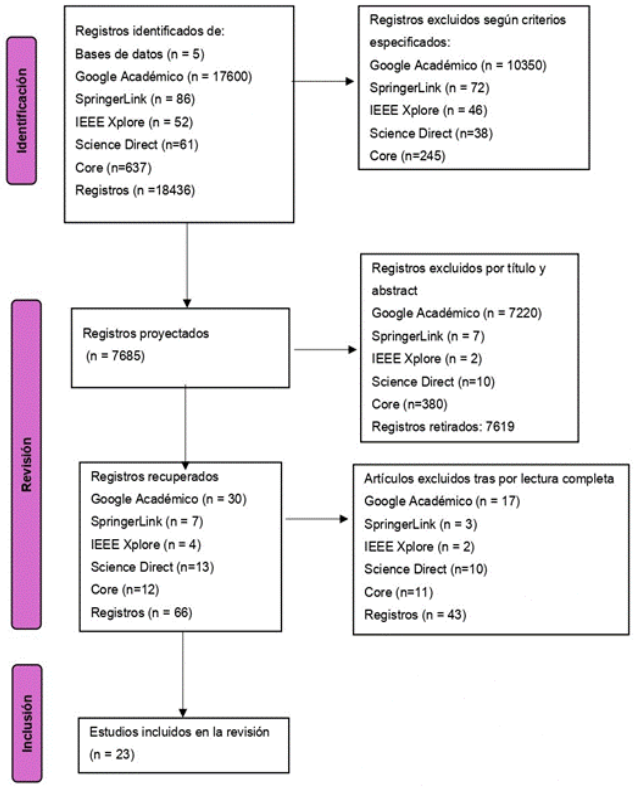
Nota. Autores, (2026)

Con el objetivo de obtener estudios relevantes que permitan realizar un análisis completo y actualizado se aplicaron filtros de inclusión y exclusión. Se incluyeron publicaciones entre los años 2020 y 2025, de libre acceso y en los idiomas de inglés o español. Por otro lado, se excluyeron aquellos artículos incompletos o que no sean de libre acceso, publicaciones fuera del rango temporal y redactados en otro idioma

El siguiente gráfico permite visualizar y comprender el proceso de recolección de los estudios a analizar a través de un flujo PRISMA en el cual se describen los filtros aplicados y la cantidad final de estudios incluidos en la investigación.

Figura 1

Flujo del proceso de extracción de datos



Nota. Autores, (2026)

ISSN: 2953-6707

Resultados

En base a los artículos seleccionados, se muestran los resultados detallados de cada uno de ellos

Tabla 2

Síntesis de los resultados obtenidos

Autor	Desafíos	Soluciones
Felouat, Nguyen, Yamagishi & Echizen, 2025	Los atacantes pueden crear rostros falsos que sean hiperrealistas tanto en 2D como en 3D para burlar a los controles de reconocimiento facial, además de una falta de datasets 3D amplios y diversos.	Se propone usar mallas faciales 3D incorporadas con rasgos geométricos difíciles de falsificar, esto en conjunto del modelo tipo mesh-MLP con self-attention y TabTransformer/ensemble entrenados con dataset 3DDGD para clasificar mallas real vs. fake.
Duhan & Kajal (2025)	Muchos detectores son entrenados con un conjunto limitado y específico de datos y fallan cuando se enfrentan a técnicas no vistas. Los deepfakes suelen difundirse rápidamente siendo complicada una respuesta temprana.	Se emplea modelos basados en CNN y GAN implementando datasets como FaceForensics++, Celeb-DF y DFDC, con métricas de precisión y AUC para identificar deepfakes como *face swap*, manipulación de expresiones, *lip-sync*, clonación de voz y alteraciones hiperrealistas.
Khan, Laghari, Inam, Ullah, Shahzad & Syed (2025)	Ataques imperceptibles en la información que engañan a los sistemas de detección con falsos positivos. Los modelos funcionan bien en los datasets entrenados, pero en otros ambientes disminuye su eficiencia.	Se busca integrar blockchain (BDLT) para un registro de huellas criptográficas y asegura la detección efectiva de manipulaciones, con aprendizaje autosupervisado y un algoritmo eficiente para operar en tiempo real con recursos limitados.
Singh, Panem, Charanarur & Chaudhary (2025)	Se cuenta con dependencia a datos de entrenamiento que pueden ser costosos y difíciles de obtener. Falta de mecanismos de autoaprendizaje para que mantengan al modelo actualizado ante los deepfakes.	Se desarrolla un modelo que combina visión por computador, análisis de voz y metadatos que detectan datos imperceptibles por humanos, utilizando redes neuronales comprimidas para que pueda operar en dispositivos móviles y sistemas más amplios.
Pakina, Kejriwal & Goel (2023)	Uso de deepfakes para falsificar rostros y voces ficticias en sistemas biométricos. Dependencia de controles estáticos que tienen mayor probabilidad de ser vulnerables ante ataques.	Implementación de modelos de biometría facial, de voz y análisis de comportamiento, con autenticación multimodal y dinámica, en base a algoritmos de detección de deepfakes para identificar patrones en los ataques.
Alazwari, Alsamri, Alamgeer, Alabdan, Alzahrani, Rizwanullah & Osman (2024)	Deepfakes emplean técnicas innovadoras de GANs y autoencoders mostrando menos artefactos visibles y complica su detección. Los modelos convencionales emplean gran cantidad de volúmenes de datos en su entrenamiento y no generalizan bien las nuevas técnicas de manipulación	Se plantea un modelo optimizado con Artificial Rabbits Optimization (ARO) que permite ajustar hiper parámetros de transfer learning incrementando la precisión sin aumentar el coste computacional. Además, implementarlo a la par de modelos como ResNet, VGG y EfficientNet para reutilizar características significativas.

 ISSN: 2953-6707

Min-Jen & Cheng-Tao (2024)	Incremento en la vulnerabilidad de autenticación biométrica basada en huella palmar debido a imágenes sintéticas generadas por deepfakes. Alta tasa de falsos positivos y negativos cuando se emplean datasets reducidos.	Se busca implementar un modelo en base a redes neuronales profundas CNNs enfocadas en aprender características distintivas entre palmas reales y deepfakes, empleando a su vez un diseño de pipeline de verificación automática para reducir los fraudes biométricos.
Maria Pawelec (2025)	El desarrollo de deepfakes involucra temas éticos y responsabilidad profesional, donde el uso indebido deriva en fraudes biométricos sin marcos regulatorios claros. Carencia de leyes uniformes que aborden la creación y uso de deepfakes en contextos de autenticación y privacidad.	Se propone establecer códigos de conducta para desarrolladores que utilizan los deepfakes, con marcos legales, normativas y regulación internacional. Además, de la implementación de tecnologías como blockchain para el rastreo de datos biométricos auténticos.
Sourav Mandal (2025)	Los deepfakes se involucran en la creación de identidades falsas empleadas en delitos financieros y vulneran los sistemas KYC (Know Your Customer).	Se plantea desarrollar marcos normativos que sancionen el uso de deepfakes en fraude de identidad y delitos cibernéticos, incorporando la IA y la integración biométrica.
Felipe Romero-Moreno (2025)	La velocidad de avance de los generadores de deepfakes supera a la de las herramientas de detección. Vacíos legales en el marco jurídico para regular el uso de deepfakes en la identidad digital y biométrica.	Sistemas de autenticación que integren algoritmos de IA generativa para la detección de deepfakes para mitigar los riesgos, además de una inclusión de un marco legal integral para sancionar los deepfakes en la suplantación de identidad y fraude biométrico.
Firc, Malinka & Hanáček (2023)	Los deepfakes pueden imitar expresiones y rasgos vulnerando los sistemas biométricos y evadiendo controles KYC. Diversidad de vectores de ataque dificultando el diseño en las defensas universales.	Se propone un modelo con herramientas híbridas que combinen el análisis visual, auditivo y forense digital con una verificación multimodal para la detección de deepfakes en sistemas biométricos.
Dudykevych, Yevseiev, Mykytyn, Ruda & Hulak (2024)	Los deepfakes pueden alterar imágenes faciales afectando la confidencialidad, integridad y autenticidad de la identidad. La detección y clasificación requieren un procesamiento complejo de datos biométricos con alta carga computacional.	Se plantea un sistema basado en redes neuronales con preprocesamiento de segmentación, detección y normalización de imágenes, además de extracción de características y entrenamiento de clasificación binaria.
Norman & Farid (2025)	Técnicas actuales como FaceFusion generan sustituciones faciales casi perfectas. Detectores existentes basados en datos requieren volúmenes masivos de entrenamiento.	La implementación de un modelo basado en anomalías biométricas para detectar las variaciones anómalas propias de deepfakes.
Rabbi, Haque, Huri & Tabassum (2024)	Los sistemas de autenticación por voz no están preparados y son muy vulnerables a voces sintéticas realistas. Los métodos de detección actuales fallan en escenarios reales al no ser muy robustos.	Se plantea un modelo de aprendizaje automático y aprendizaje profundo, extrayendo características de la señal de audio para identificar las sutilezas de la voz humana que los deepfakes no replican por completo.
He, Sun, Lei, Li, Ye, Zhang & Zhang (2025)	Deepfakes avanzados pueden imitar señales sutiles como parpadeo de ojos y microexpresiones. Los métodos de detección actuales carecen de robustez. La creación de deepfakes se ha vuelto muy accesible lo que aumenta el riesgo de ataques.	Uso de señales biométricas dinámicas como movimientos oculares o secuencias de microexpresiones faciales. Un enfoque de múltiples capas, entrenamiento de modelos de detección, tecnologías que preserven la privacidad y mecanismos de autenticación continua.

 ISSN: 2953-6707

Salawudeen Olalekan Usman (2025)	El audio deepfake puede replicar de forma convincente la voz, permitiendo a los atacantes evadir sistemas de autenticación por voz. Estas voces sintéticas pueden permitir acceder a cuentas, datos sensibles o iniciar transacciones fraudulentas.	Explorar técnicas como la incrustación de firmas digitales invisibles en los archivos multimedia originales para verificar su autenticidad. Se recomienda el uso de Autenticación multifactor (MFA).
Sheikh Inam Ul Mansoor. (2024)	Los deepfakes amenazan la privacidad al explotar los sistemas de identidad digital, generando preocupaciones en la seguridad de los datos. La tecnología sofisticada dificulta demostrar la creación de contenido deepfake.	Se plantean reformas regulatorias y la promulgación de nuevas regulaciones adaptadas para combatir delitos relacionados con los deepfakes. Las estrategias deben centrarse en innovaciones tecnológicas para detectar deepfakes.
Nicole Murillo Fonseca (2025)	Los deepfakes pueden explotar las debilidades de los sistemas de reconocimiento facial y de voz. La tecnología puede imitar comportamientos humanos como parpadeo y movimientos de cabeza para engañar a los sistemas biométricos.	Métodos basados en análisis de movimientos faciales sutiles como la detección de parpadeo. Métodos de análisis de texturas ya que los medios falsificadores carecen de detalles de alta frecuencia. Uso de aprendizaje profundo para identificar patrones anómalos.
Khan & Mahmood Malik (2023)	Algunos modelos de aprendizaje profundo fallan al detectar ataques de deepfake de voz que tienen características diferentes a los datos con los que fueron entrenados. Los avances en IA generativa representan una amenaza para los sistemas de verificación automática de voz.	Técnicas One-Shot Learning y Metric Learning para comparar muestras de voz mediante una matriz de similitud. El uso de una arquitectura de red siamesa con capas LSTM para capturar dependencias temporales y patrones a largo plazo, permitiendo detectar rasgos que indican una suplantación de voz.
Salko, Firc & Malinka (2024)	El uso malintencionado de deepfakes para suplantar la identidad en sistemas de reconocimiento facial automatizados amenaza aspectos como la seguridad financiera y el acceso a lugares seguros. Se ha prestado más atención a los sistemas de reconocimiento de voz que a los sistemas de reconocimiento facial.	Se propone el enfoque de aprendizaje profundo para determinar si una imagen o video es deepfake mediante el aprendizaje de características distintivas. Otro enfoque es el seguimiento de muchas señales fisiológicas como el parpadeo o un pequeño cambio en el color de piel.
Bodepudi & Reddy (2020)	Los atacantes pueden fabricar muestras artificiales de huellas dactilares, patrones de iris y rasgos faciales para engañar al sistema. Se utilizan imitaciones o réplicas de los rasgos biométricos de un usuario como una fotografía o máscara 3D realista.	Implementa mecanismos para diferenciar muestras biométricas en vivo y falsificadas mediante la recopilación de un conjunto de datos diversos reales y falsificados entrenadas con un modelo de aprendizaje automático. Utilizar algoritmos de aprendizaje automático para analizar patrones continuamente y detectar cualquier actividad sospechosa.

 ISSN: 2953-6707

Abbas & Taeihagh (2024)	Es muy difícil distinguir entre imágenes o videos reales y falsos porque todo el mundo tiene acceso a tecnología avanzada. La calidad de la creación de deepfakes está mejorando. Casi todo el mundo puede crear deepfakes utilizando herramientas predominantes, lo que hace que sea más difícil contrarrestarlos.	Se plantea utilizar métodos de detección enfocados en aprendizaje profundo y aprendizaje automático con modelos de detección como faceswap y face reenactment. Se mencionan varias técnicas como el uso de una red 3D-Xception con la transformada de Fourier para detectar videos falsos.
Gupta, Patil & Guido (2024)	Los ataques de suplantación utilizan técnicas de conversión de voz para generar voz falsa y eludir los sistemas ASV. Se pueden montar ataques aprovechando fallas en las implementaciones de hardware de los algoritmos de seguridad.	Se plantea el uso de sistemas de detección de voz falsa en sistemas ASV para identificar un ataque y denegar la entrada de la voz falsa al sistema ASV. Se propone utilizar diversas técnicas de detección de extracción de características y estrategias basadas en el procesamiento de señales. Métodos basados en aprendizaje profundo.

Nota. Autores, (2026)

Discusión

Los avances en la generación de deepfakes están alcanzando un nivel de realismo desafiando seriamente la seguridad de los sistemas de autenticación biométrica (reconocimiento facial, verificación de voz, huellas palmares y otras modalidades), y las respuestas tanto técnicas como operativas aún están en desarrollo.

Síntesis de los desafíos principales

La literatura analizada muestra patrones recurrentes de vulnerabilidad, varios trabajos destacan que los detectores entrenados con datasets limitados tienen incapacidad para generalizar frente a técnicas deepfake nuevas o no consideradas (Dubana & Kajal, 2025; Norman & Farid, 2025; Abbas & Taeihagh, 2024). Esta limitación se intensifica cuando los deepfakes multimodales y 3D elevan la complejidad, ya que generan artefactos difíciles de detectar con los enfoques unimodales tradicionales, exigiendo datasets más robustos y diversos que capturen múltiples modalidades biométricas simultáneamente (Felouat et al., 2025; Šalko, Firc & Malinka, 2024). Alazwari et al. (2024) evidencian que los modelos entrenados con datasets unimodales presentan tasas de error significativamente más altas al enfrentar ataques multimodales, mientras que Felouat et al. (2025) demuestran que la generación de deepfakes 3D basados en mallas faciales requiere conjuntos de datos tridimensionales amplios para su detección efectiva, lo que representa un desafío adicional para los sistemas de autenticación que operan con datos limitados.

 ISSN: 2953-6707

En el ámbito de autenticación de voz, Gupta, Patil y Capobianco Guido (2024) resaltan que los sistemas ASV y KYC son vulnerables ante voces sintéticas convincentes, y que los corpus públicos facilitan la obtención y utilización de voz para ataques (Khan & Malik, 2023; Saqlain Rabbi et al., 2024). Además, existe una brecha importante entre el rendimiento de los detectores reportados en condiciones controladas y la degradación observada en escenarios reales (Laishram et al., 2025; Fakhar Abbas & Taeihagh, 2024; He et al., 2025).

Existen problemas adicionales que surgen fuera del rendimiento técnico como aspectos legales, éticos y de confianza social que afectan la adopción y respuestas contra deepfakes (Pawelec, 2025; Mandal, 2025; Romero-Moreno, 2025; Mansoor, 2024). La insuficiente regulación y accesibilidad de herramientas de generación de deepfakes reducen las barreras para actuar contra estas herramientas maliciosas (Murillo Fonseca, 2025; Pawelec, 2025).

Evaluación comparativa de las soluciones propuestas

Las soluciones planteadas en los hallazgos pueden agruparse de la siguiente manera:

Detección biométrica y análisis de anomalías: Métodos que exploran vectores faciales, texturas de alta frecuencia o anomalías biométricas muestran una alta precisión en condiciones controladas (Norman & Farid, 2025; Dudykevych et al., 2024; Salko et al., 2024). Sin embargo, He et al. (2025) y Abbas & Taeihagh (2024) sostienen que el rendimiento cae frente a técnicas de falsificación no vistas y cuando los deepfakes reproducen microexpresiones y patrones dinámicos.

Autenticación multimodal y dinámica: La combinación de señales visuales, auditivas, así como la incorporación de de liveness y secuencias dinámicas parece ser la estrategia más robusta frente a suplantaciones sofisticadas (Firc, Malinka y Hanáček, 2023; Pakina, Kejriwal y Goel, 2023; Laishram et al., 2025; Usman, 2025). Estudios en fintech y KYC muestran que cuando se integran modalidades y análisis de comportamiento se reducen los falsos negativos en detección de deepfakes (Pakina et al., 2023).

Técnicas de entrenamiento y generalización: Transfer learning combinado con optimizadores y arquitecturas diseñadas para mejorar la generalización frente a ataques no vistos son prometedores (Alazwari et al., 2024; Khan y Malik, 2023; Duhana y Kajal, 2025). Sin embargo, varios estudios coinciden en que el principal factor que limita su efectividad son datasets reducidos, poco diversos o desactualizados y no representan adecuadamente la diversidad de deepfakes generados por técnicas emergentes (Felouat et al., 2025; Duhana y Kajal, 2025; Abbas y Taeihagh, 2024).

 ISSN: 2953-6707

Esta limitación se da porque la mayoría de los modelos son entrenados con datos cerrados y controlados, provocando un sobreajuste a patrones específicos y una degradación significativa del rendimiento al enfrentarse a ataques no vistos o escenarios reales (Duhan & Kajal, 2025). Para mitigar este problema, la literatura propone incorporar procesos de actualización continua del modelo a traves de learning-based pipelines, usar datasets multimodales más amplios, asi como estrategias de aprendizaje incremental y autosupervisado que permitan adaptar los detectores a nuevas tecnicas de manipulacion sin requerir entrenamiento completo (Felouat et al., 2025; Alazwari et al., 2024).

Integridad, trazabilidad y marcos legales: Propuestas que incorporan firmas digitales, watermarking y blockchain facilitan verificar el origen y trazabilidad de biometría multimodal, contribuyendo a prevenir el uso de contenido manipulado en procesos de autenticación (Khan et al., 2025; Laishram et al., 2025; Pawelec, 2025). Además, desde el ámbito jurídico, Mandal (2025), Romero-Moreno (2025) y Mansoor (2024) insisten en la formulación de marcos específicos para deepfakes.

La revisión resalta ideas urgentes de investigación: mejorar la generalización frente a ataques no vistos, desarrollar benchmarks multimodales y adversariales, avanzar en estándares legales y técnicos internacionales (Mandal, 2025; Romero-Moreno, 2025; Felouat et al., 2025). Como autores creemos que enfoques interdisciplinarios como técnico, legal y social serán los más efectivos para diseñar medidas sostenibles y aceptables.

Los trabajos revisados revelan que hay herramientas prometedoras para mitigar la suplantación biométrica basada en deepfakes, sin embargo, la amenaza evoluciona con rapidez, exigiendo respuestas integrales y adaptativas (Pawelec, 2025; Abbas y Taeihagh, 2024). Por eso, no basta con detectar los deepfakes, es necesario diseñar sistemas biométricos que integren multimodalidad, actualización continua, protección de la privacidad y gobernanza regulatoria. Caso contrario, el riesgo además de permanecer, podría escalar hasta minar la confianza pública y operativa en la autenticación biométrica (Mandal, 2025; Romero-Moreno, 2025; Bodepudi y Reddy, 2020).

Conclusiones

El análisis realizado evidencia que los deepfakes representan desafíos para los sistemas de autenticación biométrica debido al potencial que cuentan al replicar rasgos físicos y conductuales acercándose cada vez a datos más realistas. Si bien es cierto se han intensificado los avances en inteligencia artificial, lo cual ha permitido un avance en el desarrollo de nuevas técnicas de detección como el aprendizaje automático, multimodal, el uso de biometría conductual y optimización de modelos por el aprendizaje transferido, también persiste la necesidad constante de actualizar datasets y que se generalice los datos en diferentes ambientes para así lograr cubrir la evolución de los generadores de contenido sintético.

Es por ello, que resulta claro que las soluciones no solo deben limitarse a lo técnico, sino que también enfocarse a fortalecer marcos normativos, incorporar principios éticos y considerar una educación digital hacia los usuarios que son los pilares fundamentales para proteger la confianza de los sistemas biométricos. Asimismo, la integración de estrategias como la autenticación dinámica y multifactorial logra mitigar en gran medida los riesgos que presenta la suplantación de identidad.

No obstante, se identifican ciertas limitaciones en esta investigación debido a la disponibilidad escasa en el acceso a base de datos diversos, así como la creciente obsolescencia que presentan las herramientas de detección en sistemas biométricos en contextos reales y actuales. En este sentido, se recomienda para futuras investigaciones explicar nuevos enfoques que se centren en la creación de datasets más robustos y generalizados, la evaluación de técnicas de aprendizaje continuo aplicadas a escenarios de ataque no vistos, y el diseño de modelos híbridos que integren seguridad técnica con regulación y gobernanza internacional.

De este modo, de cara al futuro en la autenticación biométrica frente a deepfakes requiere innovación tecnológica constante para hacerle frente a los ataques novedosos de esta herramienta, además de políticas regulatorias y conciencia social para proteger la confiabilidad de los sistemas y garantizar su seguridad y resiliencia ante la creciente sofisticación de las amenazas digitales.

 ISSN: 2953-6707

Referencias

Abbas, F., & Taeihagh, A. (2024). Unmasking deepfakes: A systematic review of deepfake detection and generation techniques using artificial intelligence. *Expert Systems with Applications*, *252*(124260), 124260. https://doi.org/10.1016/j.eswa.2024.124260

Alanazi, F., Ushaw, G., & Morgan, G. (2024). Improving Detection of DeepFakes through Facial Region Analysis in Images. *Electronics*, *13*(1), 126. https://doi.org/10.3390/electronics13010126

Alazwari, S., Alsamri, M. O. J., Alamgeer, M., Alabdan, R., Alzahrani, I., Rizwanullah, M., & Osman, A. E. (2024). Artificial rabbits optimization with transfer learning based deepfake detection model for biometric applications. *Ain Shams Engineering Journal*, *15*(12), 103057. https://doi.org/10.1016/j.asej.2024.103057.

Bodepudi, A., & Reddy, M. (2020). Spoofing attacks and mitigation strategies in biometrics-as-a-service systems. *Eigenpub Review of Science and Technology*, *4*(1), 1-14. https://www.researchgate.net/publication/372134385_Spoofing_Attacks_and_Mitigation_Strategies_in_Biometrics-as-a-Service_Systems

Dudykevych, V., Yevseiev, S., Mykytyn, H., Ruda, K., & Hulak, H. (2024). Detecting Deepfake modifications of biometric images using neural networks. *Cybersecurity Providing in Information and Telecommunication Systems 2024*, *3654*, 391-397. https://elibrary.kubg.edu.ua/id/eprint/48581

Felouat, H., Nguyen, H. H., Yamagishi, J., & Echizen, I. (2025). 3DDGD: 3D Deepfake Generation and Detection Using 3D Face Meshes. *IEEE access: practical innovations, open solutions*, *13*, 107429–107441. https://doi.org/10.1109/access.2025.3580950

Firc, A., Malinka, K., & Hanáček, P. (2023). Deepfakes as a threat to a speaker and facial recognition: An overview of tools and attack vectors. *Heliyon*, *9*(4), e15090. https://doi.org/10.1016/j.heliyon.2023.e15090

Garousi, V., Felderer, M., & Mäntylä, M. V. (2020). Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. Information and Software Technology, 106–121. https://doi.org/10.1016/j.infsof.2018.09.006

Gupta, P., Patil, H.A. & Guido, R.C. Vulnerability issues in Automatic Speaker Verification (ASV) systems. *J AUDIO SPEECH MUSIC PROC.* 2024, 10 (2024). https://doi.org/10.1186/s13636-024-00328-8

 ISSN: 2953-6707

He, S., Lei, Y., Zhang, Z., Sun, Y., Li, S., Zhang, C., & Ye, J. (2025). Identity deepfake threats to biometric authentication systems: Public and expert perspectives. En *arXiv [cs.HC]*. http://arxiv.org/abs/2506.06825

Jbara, W. A., Hussein, N. A.-H. K., & Soud, J. H. (2024). Deepfake detection in video and audio clips: A comprehensive survey and analysis. *Mesopotamian Journal of CyberSecurity*, *4*(3), 233–250. https://doi.org/10.58496/mjcs/2024/025

Khan, A., & Malik, K. M. (2023, December). Securing voice biometrics: One-shot learning approach for audio deepfake detection. In *2023 IEEE International Workshop on Information Forensics and Security (WIFS)* (pp. 1-6). IEEE. https://doi.org/10.48550/arXiv.2310.03856

Khan, A.A., Laghari, A.A., Inam, S.A. *et al.* A survey on multimedia-enabled deepfake detection: state-of-the-art tools and techniques, emerging trends, current challenges & limitations, and future directions. *Discov Computing* 28, 48 (2025). https://doi.org/10.1007/s10791-025-09550-0

Krity Duhan, Abhishek Kajal, A Comparative Analysis of Deep Learning Based Approaches for DeepFake Identification, Procedia Computer Science, Volume 259, 2025, Pages 482-493, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2025.03.350.

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. (2009) The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. PLoS Med 6(7): e1000100. https://doi.org/10.1371/journal.pmed.1000100

Mandal, Sourav, Deep Fake Technology and Identity Theft: An Emerging Challenge for Cyber Laws in India (January 16, 2025). Available at SSRN: https://ssrn.com/abstract=5161545 or http://dx.doi.org/10.2139/ssrn.5161545

Mansoor, S. I. U. (2024). Legal implications of deepfake technology: In the context of manipulation, privacy, and identity theft. *Central University of Kashmir Law Review*, *4*, 65-92. https://www.researchgate.net/publication/387499036_Legal_Implications_of_Deepfake_Technology_In_the_Context_of_Manipulation_Privacy_and_Identity_Theft

Min-Jen, T., & Cheng-Tao, C. (2024). Convolutional Neural Network for Detecting Deepfake Palmprint Images. *IEEE access: practical innovations, open solutions*, *12*, 103405–103418. https://doi.org/10.1109/access.2024.3433497

 ISSN: 2953-6707

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. https://doi.org/10.1371/journal.pmed.1000097

Murillo Fonseca, N. (2025). Redefining identity and security: Spoofing in the age of artificial intelligence. https://repositorio.uniandes.edu.co/server/api/core/bitstreams/408c83e6-a08a-47b7-b32a-0ff6cc79db5e/content

Norman, J. D., & Farid, H. (2025). Detecting Deepfake Talking Heads from Facial Biometric Anomalies. *arXiv preprint arXiv:2507.08917*. https://doi.org/10.48550/arXiv.2507.08917

Page M J, McKenzie J E, Bossuyt P M, Boutron I, Hoffmann T C, Mulrow C D et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews BMJ 2021; 372 :n71 doi: https://doi.org/10.1136/bmj.n71

Pakina, Anil Kumar & Kumar, Dk & Goel, Anshul & Pujari, Tejaskumar. (2023). AI-Generated Synthetic Identities in Fin Tech: Detecting Deep fakes KYC Fraud Using Behavioral- https://www.researchgate.net/publication/390947564_AI-Generated_Synthetic_Identities_in_Fin_Tech_Detecting_Deep_fakes_KYC_Fraud_Using_Behavioral_Biometrics

Pawelec, M. Decent deepfakes? Professional deepfake developers' ethical considerations and their governance potential. *AI Ethics* 5, 2641–2666 (2025). https://doi.org/10.1007/s43681-024-00542-2

Qureshi, S. M., Saeed, A., Almotiri, S. H., Ahmad, F., & Al Ghamdi, M. A. (2024). Deepfake forensics: a survey of digital forensic methods for multimodal deepfake identification on social media. *PeerJ. Computer Science*, *10*(e2037), e2037. https://doi.org/10.7717/peerj-cs.2037

Rabbi, B. M., Haque, E., Huri, H. S., & Tabassum, N. (2024). *From impersonation to authentication: techniques for identifying deep fake voices* (Doctoral dissertation, Brac University). http://hdl.handle.net/10361/25249

Romero-Moreno, F. (2025). Deepfake detection in generative AI: A legal framework proposal to protect human rights. Computer Law & Security Review, 58(106162), 106162. https://doi.org/10.1016/j.clsr.2025.106162

Salko, M., Firc, A., & Malinka, K. (2024, April). Security implications of deepfakes in face authentication. In *Proceedings of the 39th ACM/SIGAPP symposium on applied computing* (pp. 1376-1384). https://doi.org/10.1145/3605098.3635953

 ISSN: 2953-6707

Singh, L.H., Charanarur, P. & Chaudhary, N.K. Advancements in detecting Deepfakes: AI algorithms and future prospects − a review. *Discov Internet Things* 5, 53 (2025). https://doi.org/10.1007/s43926-025-00154-0

Srivastava, A., Badal, T., Agarwal, M., Dwivedi, K., & Naik, A. H. (2025). Enhancing deepfake detection accuracy with vision transformers: A generative AI approach. En *Communications in Computer and Information Science* (pp. 68–78). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-84602-1_5

Usman, S. O. (2025). IMPACT OF DEEPFAKE TECHNOLOGY ON TELECOM SECURITY AND USER TRUST. https://www.researchgate.net/publication/395043194_IMPACT_OF_DEEPFAKE_TECHNOLOGY_ON_TELECOM_SECURITY_AND_USER_TRUST

Wohlin, C., Mendes, E., Felizardo, K. R., & Kalinowski, M. (2020). Guidelines for the search strategy to update systematic literature reviews in software engineering. Information and Software Technology, 127, 106366. https://doi.org/10.1016/j.infsof.2020.106366

Xing, H., Tan, S. Y., Qamar, F., & Jiao, Y. (2025). Face Anti-Spoofing Based on Deep Learning: A Comprehensive Survey. *Applied Sciences*, *15*(12), 6891. https://doi.org/10.3390/app15126891

 ISSN: 2953-6707