Frontiers in Computer Science and Artificial Intelligence
DOI: 10.32996/fcsai
Journal Homepage: www.al-kindipublisher.com/index.php/fcsai

FCSAI
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Machine Learning for Identifying Deepfake-Driven Identity Abuse, Authentication Evasion, and Customer Impersonation in U.S. Banking

Yusuf Oli Rahat
University of New Haven, Business Analytics
Email: yraha1@unh.newhaven.edu
ORCID: https://orcid.org/ 0009-0001-8834-7759

Md Kamrul Islam
University of New Haven, Business Analytics
Email: misla22@unh.newhaven.edu
ORCID: https://orcid.org/0009-0001-8906-630X

Shah Farhan Rabbani
University of New Haven, Business Analytics
Email: srabb2@unh.newhaven.edu
ORCID: https://orcid.org/0009-0001-8434-223X
Corresponding Author: Yusuf Oli Rahat**, E-mail:** yraha1@unh.newhaven.edu

ABSTRACT

Deepfake-enabled identity abuse has moved from a peripheral cyber-risk to an operational threat for U.S. banking, especially in remote onboarding, account recovery, contact-center authentication, and high-risk payment authorization. The problem is no longer limited to obvious synthetic media. Financial institutions increasingly face blended attacks that combine forged identity documents, face or voice cloning, injected video streams, social-engineering pressure, mule accounts, and adaptive retries across channels. This paper develops a publication-ready research framework for machine learning systems that identify deepfake-driven customer impersonation, authentication evasion, and identity abuse in U.S. banking environments. The study synthesizes regulatory guidance, public fraud data, biometric spoofing research, deepfake detection literature, and banking fraud analytics to propose a multimodal detection architecture spanning document forensics, face presentation attack detection, deepfake-video detection, speaker anti-spoofing, behavioral biometrics, device and network telemetry, graph-based entity resolution, and risk-calibrated decisioning. Real public evidence motivates the design: FTC data show reported fraud losses reached $12.5 billion in 2024, including $2.95 billion in imposter-scam losses, while FinCEN reported increasing suspicious activity narratives involving deepfake media targeting financial institutions. The proposed methodology treats identity abuse as a sequential, multimodal, and adversarial classification problem rather than a single-screen biometric check. The paper argues that the most effective defense is not one detector but an explainable ensemble with smart friction, human escalation, and governance aligned to NIST identity-proofing standards, anti-money-laundering expectations, and consumer-protection obligations. By linking technical detection with operational controls, the study provides a practical blueprint for authentication, resilient fraud prevention, and more trustworthy digital banking systems.

KEYWORDS

deepfakes; banking fraud; identity proofing; account takeover; customer impersonation; voice cloning; face anti-spoofing; graph neural networks; explainable AI

ARTICLE INFORMATION

ACCEPTED: 01 February 2026 **PUBLISHED:** 13 February 2026 **DOI:** 10.32996/fcsai.2026.4.3.7x

1. Introduction

U.S. banking is entering a phase in which identity fraud is increasingly mediated by generative artificial intelligence. The practical danger is not merely that synthetic media exist, but that they can now be inserted into everyday banking workflows that were designed around assumptions of stable visual, vocal, and documentary authenticity. Remote account opening, selfie-based know-your-customer checks, document upload, step-up verification, account recovery, call-center voice authentication, and executive approval workflows all depend on signals that can be manipulated, replayed, or fully synthesized. In November 2024, FinCEN issued a formal alert noting an increase in suspicious activity reporting describing the suspected use of deepfake media in fraud schemes targeting financial institutions and their customers, often involving altered or fabricated identity documents used to circumvent verification and authentication controls. The FBI also warned in late 2024 that criminals were using generative AI to create fraudulent identity documents, profile images, and other deceptive artifacts that support impersonation and confidence fraud. These warnings matter because banking systems increasingly operate across remote and mobile channels where the institution may have little direct contact with the genuine customer.

The economic environment reinforces the urgency. FTC data show that consumers reported more than $12.5 billion in fraud losses in 2024, up from more than $10 billion in 2023 and $8.8 billion in 2022, with imposter scams remaining one of the largest categories by reported loss. In the 2024 Consumer Sentinel Data Book, imposter scams accounted for 845,806 reports and $2.95 billion in losses, while bank transfers and payments produced the highest aggregate losses among payment methods. These numbers are broader than bank-specific deepfake cases, yet they establish the scale of impersonation-driven financial harm flowing through channels that banks must monitor. They also illustrate a strategic asymmetry: attackers need only one convincing synthetic artifact to gain entry, while banks must consistently authenticate real users without creating unacceptable friction for legitimate customers.

Traditional identity defenses remain necessary but increasingly insufficient. Rule-based fraud engines, one-time passwords, static device fingerprints, knowledge-based questions, and single-modality liveness checks can all be bypassed in specific contexts. A voice print may be vulnerable to cloned speech; facial verification may fail under replay or injection attacks; a document classifier may accept a high-quality fabricated credential if document and face are evaluated in isolation; and a contact-center agent may be influenced by urgency cues even when the underlying authentication signal is weak. The emerging threat model is therefore

multimodal and sequential. Attackers combine synthetic media with social engineering, mule infrastructure, SIM swapping, credential theft, synthetic identity building, and transaction staging. This means that machine learning must move beyond detecting "fake video" in the abstract and instead identify coordinated identity abuse across the full customer journey.

This paper addresses that need by proposing a machine-learning framework for identifying deepfake-driven identity abuse, authentication evasion, and customer impersonation in U.S. banking. The core argument is that effective defense requires signal fusion across identity proofing, authentication, and post-authentication behavior. A publishable contribution in this area must do three things. First, it must situate deepfakes inside the actual banking threat surface rather than treating them as isolated media-forensics curiosities. Second, it must translate advances in document fraud detection, face anti-spoofing, deepfake video detection, speaker anti-spoofing, behavioral analytics, and graph learning into a coherent operational pipeline. Third, it must align model design with regulatory expectations around fraud mitigation, suspicious activity monitoring, explainability, model governance, privacy, and consumer fairness.

Accordingly, the paper synthesizes approximately forty to fifty relevant peer-reviewed and institutional sources across five literatures: banking fraud analytics, biometric spoofing and deepfake detection, digital identity proofing, multimodal machine learning, and U.S. financial-sector governance. The manuscript is structured as follows. The literature review examines how prior research has addressed identity fraud, synthetic identity, face and voice spoofing, and explainable fraud analytics, while identifying the gap between generic deepfake detection benchmarks and production banking controls. The methodology then presents a multimodal ensemble architecture that combines document-level, biometric, behavioral, device, and graph features with calibrated risk scoring and human escalation. The discussion interprets the model from operational and regulatory perspectives, emphasizing that deepfake resilience depends on adaptive orchestration rather than static biometrics alone. The resulting framework is intended to help banks strengthen trust while preserving access, usability, accountability.

Table 1: Selected public indicators motivating deep-fake-focused banking controls

Indicator	2024 value	Why it matters for banking	Public source
FTC total reported fraud losses	$12.5 billion	Shows the scale of digitally enabled fraud pressure affecting customer channels	FTC Consumer Sentinel 2024 / March 2025 release
FTC imposter-scam losses	$2.95 billion	Impersonation remains a major precursor to banking abuse and customer manipulation	FTC Consumer Sentinel 2024
FTC fraud losses via bank transfers/payments	$2.09 billion	Supports payment-side smart friction after risky identity events	FTC Consumer Sentinel 2024
IC3 total losses	$16.6 billion	Cyber-enabled fraud losses continue to rise across the U.S. threat environment	FBI IC3 2024 Annual Report
IC3 identity-theft complaints	21,403	Confirms persistent identity misuse in cybercrime reporting	FBI IC3 2024 Annual Report
FinCEN deepfake alert	Increase observed in 2023-2024 SAR narratives	Directly links deepfakes to fraud against institutions and customers	FinCEN FIN-2024-Alert004

Note. Values are drawn from FTC, FBI/IC3, and FinCEN public releases and reports summarized in the reference list.

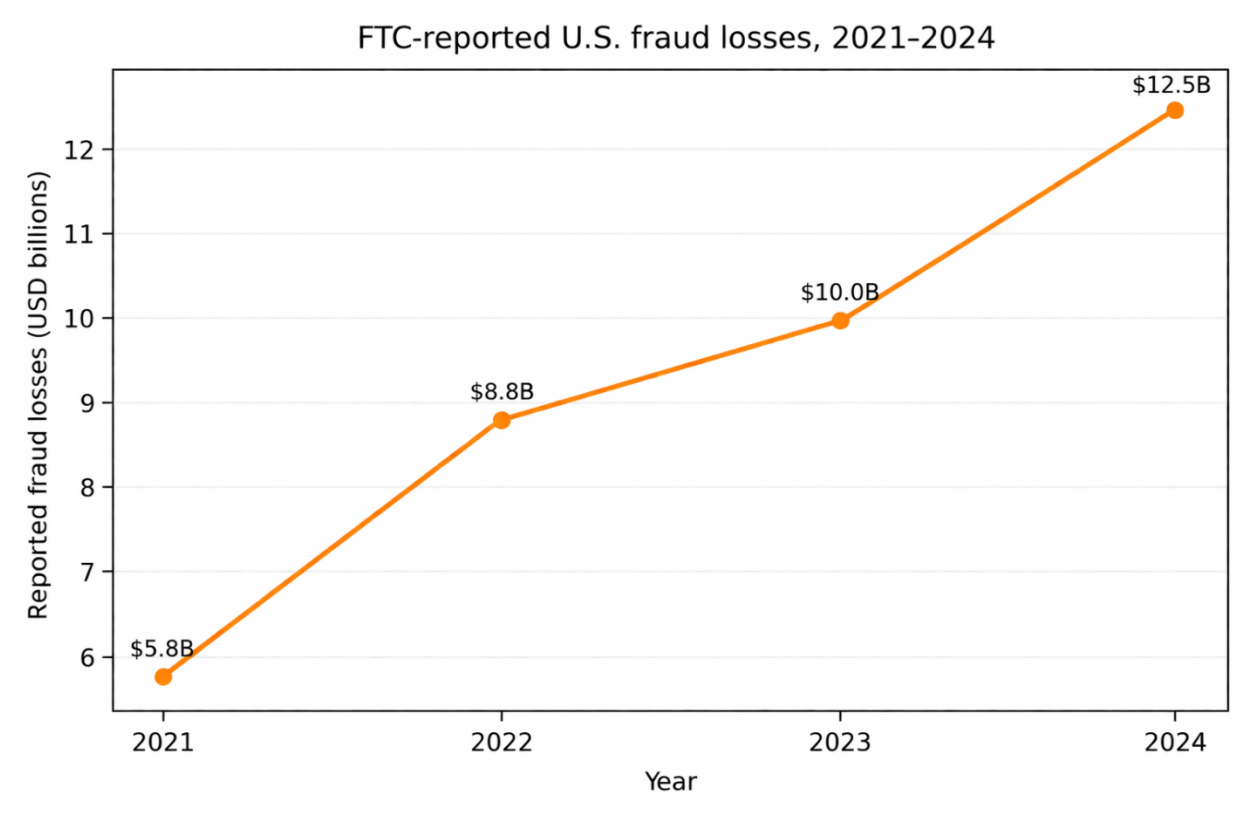

Figure 1. FTC-reported U.S. fraud losses, 2021-2024.

Source. FTC public releases for 2021 through 2025 and the 2024 Consumer Sentinel Data Book.

2. Literature Review

Research on deepfake-driven identity abuse in banking sits at the intersection of several mature but only partially integrated literatures. One stream examines financial fraud analytics, including anti-money-laundering monitoring, anomalous transaction detection, account takeover, and early warning systems for irregular activity. Another examines digital identity proofing and authentication, with particular attention to remote onboarding, synthetic identity fraud, and device-based risk controls. A third investigates biometric spoofing, including face presentation attacks, morphing, replay attacks, injected media, and voice attacks against speaker verification. A fourth focuses on deepfake detection in image, video, and audio domains. A fifth addresses governance, explainability, and operational resilience in regulated environments. The main gap is the lack of integrated frameworks tailored to banking workflows in which adversaries blend multiple attack modes over time.

Within banking and payments, fraud research has long emphasized that identity abuse is both an onboarding problem and a lifecycle problem. Federal Reserve materials on synthetic identity fraud argue that early detection at account opening is critical because once a synthetic or compromised identity is booked as a customer, subsequent activity may mimic normal behavior and become harder to distinguish from legitimate use. That insight is highly relevant to deepfakes. A falsified selfie, forged document, or cloned voice is rarely the final objective; it is a gateway to account opening, account recovery, credential reset, social engineering, payment release, or mule account use. Studies on predictive analytics in financial infrastructure and fraud governance similarly show that effective controls depend on combining front-end screening with downstream behavioral monitoring and explainable escalation logic (Fahim et al., 2023; Ibrahim et al., 2024; Mahmud et al., 2025). Cross-domain work supplied by the client, although not specific to deepfakes, is useful because it frames predictive analytics as a governance problem rather than a technical optimization problem.

The identity-proofing literature provides the operational backbone for this topic. NIST SP 800-63A conceptualizes identity proofing as resolution, validation, verification, and enrollment, and explicitly treats fraud mitigation as part of the process. The 800-63 revision also raises the bar for remote biometric assurance by requiring presentation attack detection conformant to ISO/IEC 30107-3 and a low impostor attack presentation accept rate threshold for remote biometric collection. This shift matters

because banking institutions increasingly rely on remote proofing vendors, mobile capture, and self-service channels. ENISA's 2024 remote ID proofing guidance, while European, usefully documents replay, mask, morphing, injection, and remote-session fraud patterns that mirror risks present in U.S. financial onboarding. The implication is that a banking-grade deepfake defense must be embedded into a broader proofing workflow and audit trail, not bolted onto a single selfie step.

Face anti-spoofing research is highly developed and offers several lessons. Early systems focused on hand-crafted texture cues for print and replay attacks; recent work employs convolutional and transformer-based models, depth supervision, rPPG signals, reflection cues, frequency-domain artifacts, and domain-generalization methods. Yu et al. (2023) note that the field has progressed from binary liveness toward richer supervision, including depth, physiological, and temporal constraints. Yet the literature also documents a recurrent weakness: methods trained on a narrow attack distribution often fail on unseen sensors, lighting conditions, and spoof types. For banking, this means that excellent benchmark performance on public datasets does not guarantee robustness against low-cost mobile-device injection, composited selfie videos, or adversaries who iteratively test vendor thresholds. Studies of deepfake video detection make a similar point. Detection based on compression artifacts or generator-specific fingerprints may degrade rapidly as generation quality improves. Reviews by Verdoliva (2020), Mirsky and Lee (2021), and later surveys emphasize the challenge of generalization, especially in real-world, cross-platform settings. Banking applications therefore require continuous evaluation under distribution shift.

Voice authentication presents a related but distinct challenge. Automatic speaker verification can improve user convenience, but it is vulnerable to replay, voice conversion, and synthetic speech. The ASVspoof initiatives have become a central benchmark for this domain, showing both major progress and persistent weaknesses under channel variation, unseen attacks, and real acoustic conditions. The 2021 challenge in particular incorporated logical-access, physical-access, and deepfake scenarios, illustrating that spoofing resilience requires anti-spoofing countermeasures rather than speaker verification alone. Recent surveys of threats against voice authentication show that the risk is not limited to laboratory deepfakes; modern attacks include interactive voice cloning, adversarial perturbations, and telephony-transmitted synthetic speech. For banks still using voice biometrics in contact centers or high-value account servicing, the literature implies that voice should function as one feature within a multimodal decision framework, not as a stand-alone identity proof.

Deepfake detection research in image, video, and audio has expanded rapidly. Methods include spatial artifact analysis, temporal inconsistency modeling, physiological-signal extraction, frequency analysis, self-supervised representation learning, multimodal fusion, and large-model approaches. Altuncu et al. (2024) argue that definitions, performance metrics, and standards remain inconsistent across studies, complicating operational comparison. This observation is crucial for banking because institutions do not simply need high area-under-the-curve values on public benchmarks; they need thresholding strategies, explainable outputs, false-positive cost control, adversarial testing, and evidentiary logging suitable for customer disputes and regulatory review. Academic work increasingly recognizes the importance of robustness, provenance, and calibration, but most studies still optimize for media-level classification rather than journey-level fraud prevention.

Graph-based and entity-resolution methods add a powerful dimension that pure media forensics often miss. Fraud rings reuse devices, phone numbers, mailing addresses, IP blocks, mule beneficiaries, employer fields, and behavioral patterns across applications. Graph neural networks and heterogeneous graph methods have shown promise in financial fraud, anti-money-laundering detection, and credit-risk modeling because they capture relational structure that tabular models discard. The paper Beyond FICO demonstrates how graph architectures can fuse behavioral and contextual signals beyond traditional scalar scores (Rasel et al., 2023). In the present domain, graph learning is valuable because a deepfake event may appear plausible in isolation yet become suspicious when linked to prior failed attempts, shared infrastructure, or anomalous relationship topology. That relational perspective converts media detection from an isolated filter into part of an identity-abuse network.

Explainability and governance form the final critical literature. Financial institutions must justify adverse actions, manage model risk, document controls, and avoid discriminatory or inaccessible outcomes. Work on algorithmic accountability in U.S. consumer FinTech, explainable AI for medical debt forecasting, and fraud-focused predictive analytics stresses the need for traceable features, human review, and governance controls around automated decisioning (Fahim et al., 2023; Fahim et al., 2025; Rasel et al., 2026). In practice, this means that a bank should be able to explain whether a session was escalated because of document tampering signals, face-injection anomalies, voice anti-spoofing risk, graph proximity to known mule patterns, or impossible customer-journey behavior. The literature further shows that "explainable" should not mean simplistic; rather, it should mean that complex multimodal systems are translated into intelligible operational reasons and validated fairness metrics.

Taken together, the literature supports five conclusions. First, deepfake-enabled fraud should be modeled as identity abuse across a customer journey, not as isolated media manipulation. Second, single-modality biometric defenses are brittle against adaptive attackers. Third, multimodal fusion improves resilience, but only if models are calibrated under real operational constraints. Fourth, graph analytics and temporal sequencing are essential because fraudsters reuse infrastructure and stage

attacks. Fifth, governance, auditability, and consumer-protection considerations are central to deployment quality. What remains underdeveloped is a banking-specific framework that unifies these insights into a rigorous, explainable, and operationally feasible detection architecture. The present study aims to fill that gap.

An additional literature strand concerns synthetic identity, identity theft, and account-takeover detection in the payment system. Federal Reserve and industry working groups repeatedly stress that synthetic identities are difficult to measure because they often perform as "good" customers before bust-out or mule activation. This matters because many banking attack chains are hybrid: fraudsters may use stolen identity elements, fabricated documents, and synthetic media to cross the final threshold of proofing or recovery. Thus, deepfakes do not replace legacy identity fraud; they amplify it by making compromised or synthetic identities more operationally credible.

Another relevant body of work examines human factors. Studies in social engineering and media trust suggest that people often overweight confidence, fluency, and contextual plausibility when judging authenticity. This matters in branches and contact centers, where a human reviewer may interpret a smooth video call or emotionally persuasive voice as authenticity even when machine indicators are weak. FTC warnings about voice cloning reinforce that harmful synthetic media work partly because they exploit human urgency and familiarity, not only technical vulnerabilities.

The literature also highlights tension between privacy and detection depth. Behavioral biometrics, device telemetry, and graph linkage can materially improve fraud detection, yet they raise governance questions about data minimization, retention, and permissible use. Financial institutions therefore need architectures that extract security value while maintaining role-based access, purpose limitation, and retention schedules aligned with legal and regulatory obligations. This helps explain why some academic models remain difficult to deploy. In short, the literature does not support a one-size-fits-all detector. It supports a layered, governable system built for remote identity proofing, adaptive fraud operations, and accountable decisioning. These findings collectively motivate multimodal, governed, and operationally aware banking fraud research agendas. future.

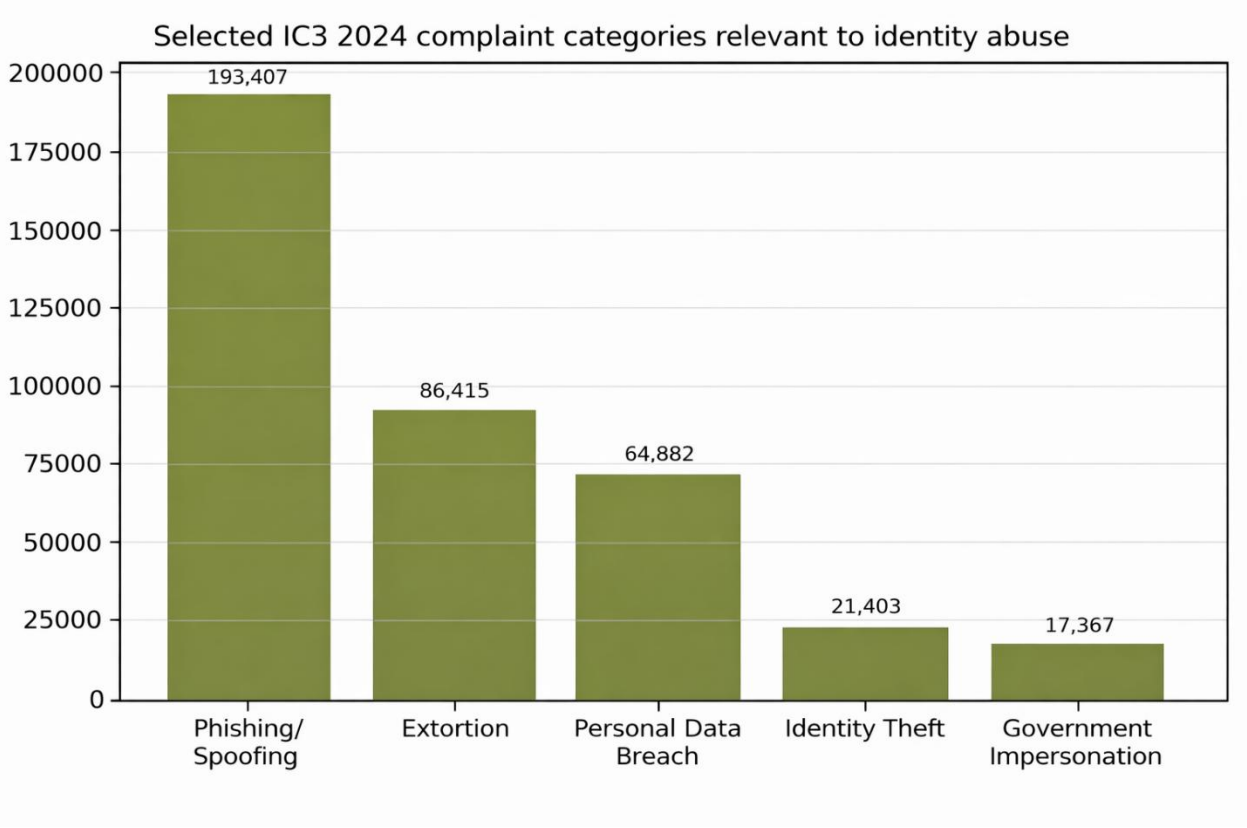

Figure 2. Selected IC3 2024 complaint categories relevant to identity abuse.

Source. FBI Internet Crime Complaint Center, 2024 Annual Report.

Table 2: Threat taxonomy for deep fake-driven identity abuse in banking

Attack class	Primary target	Typical attack assets	Core defensive signals
Document forgery + face mismatch	Remote onboarding	Edited IDs, portrait substitution, synthetic selfies	MRZ/OCR consistency, template conformity, portrait-live similarity
Replay or injection attack	Selfie/liveness check	Screen replay, virtual camera, stream injection	Sensor noise mismatch, timing artifacts, challenge-response failure
Voice cloning / synthesized speech	Contact center or phone verification	TTS, voice conversion, replayed prompts	Anti-spoofing score, channel mismatch, semantic risk, caller continuity
Synthetic identity with deepfake support	Account opening and bust-out fraud	Stolen SSN elements, fabricated attributes, polished media artifacts	Cross-source identity checks, graph linkage, delayed payment controls
Customer impersonation for profile changes	Account recovery / step-up flows	Stolen PII, cloned voice, remote social engineering	Out-of-band confirmation, behavioral drift, device continuity

3. Methodology

This study proposes a multimodal machine-learning framework for detecting deepfake-driven identity abuse across four banking stages: digital onboarding, login and step-up authentication, account recovery, and assisted-service impersonation in contact-center or branch-supported workflows. Rather than presenting a narrow benchmark experiment, the methodology is designed as a realistic, production-oriented research blueprint grounded in public data, published technical literature, and current U.S. identity and fraud-control expectations. The objective is to classify sessions and entities into ordered risk states—allow, allow with monitoring, step up, manual review, or block—while generating interpretable evidence for fraud operations, model risk management, and suspicious activity investigations.

The analytical unit is a "customer-journey event bundle." Each bundle contains all machine-observable data generated during a discrete identity-sensitive interaction: submitted document images, selfie or video frames, audio samples when voice is involved, device metadata, network telemetry, session logs, keystroke or touch dynamics where lawfully collected, prior account and application history, and entity-linkage information from connected applications or beneficiaries. A bundle also stores outcome labels when available, such as successful onboarding, confirmed fraud, customer dispute, manual-review disposition, or SAR escalation. This journey-level design reflects the reality that a single frame or utterance rarely determines the final fraud outcome; fraud emerges from combinations of signals over time.

Feature engineering proceeds across six layers. The first layer is document and image forensics. Candidate features include OCR consistency, field-layout conformity, font irregularity, hologram or texture anomalies when visible, EXIF inconsistencies, compression signatures, boundary artifacts, MRZ validation, portrait-to-document-face similarity, and document-to-applicant attribute coherence. Because FinCEN specifically noted the use of altered or fabricated identity documents in deepfake-related schemes, document integrity features are treated as first-order signals rather than peripheral metadata. The second layer is face verification and presentation attack detection. This includes face-embedding similarity between live capture and document portrait; temporal liveness cues; eye, lip, and head-motion coherence; depth or parallax estimators where available; screen-reflection and replay signatures; and injection-specific inconsistencies such as implausible sensor noise or pipeline timing. The third layer is audio and voice anti-spoofing. Features include CQCC or LFCC spectral representations, phase information, prosodic stability, vocoder artifacts, speaker-embedding mismatch, challenge-response latency, and channel-informed spoof scores adapted from ASVspoof-style pipelines.

The fourth layer captures behavioral and environmental context. Examples such as include device reputation, app attestation state, OS and browser consistency, IP geolocation distance, VPN or emulator indicators, time-of-day anomalies, session velocity, retry counts, typing cadence, pointer movement, touchscreen dynamics, and cross-channel inconsistency between customer history and current session behavior. The fifth layer is relational or graph structure. A heterogeneous graph is constructed with nodes representing customers, applications, devices, phone numbers, emails, addresses, employers, bank accounts, beneficiaries, IP clusters, and review agents. Edges represent shared attributes, transaction flows, communication events, and temporal co-occurrence. Graph features include degree, community membership, suspicious-neighbor counts, temporal motif frequencies,

and embeddings generated by graph neural networks. The sixth layer contains supervised meta-features, including previous model scores, review outcomes, chargeback or return history, and prior fraud rule triggers.

The proposed architecture is an ensemble with both modality-specific encoders and a fusion stage. For documents and images, a vision backbone extracts forgery and consistency embeddings. For face-video streams, a temporal encoder models liveness and manipulation signatures. For audio, a spoof-detection model produces both frame-level and utterance-level risk scores. Tabular environmental data are modeled with gradient-boosted trees because they handle mixed data types, missingness, and nonlinear threshold effects well. Graph relationships are modeled with a heterogeneous graph neural network that propagates risk across shared infrastructure while preserving edge types. Outputs from these subnetworks are concatenated into a fusion layer that can be implemented as a calibrated neural meta-classifier or regularized stacking model. The final output is not simply a binary fraud label; it is a calibrated risk distribution over decision states, accompanied by reason codes and uncertainty estimates.

Model training requires hierarchical labels because banking fraud outcomes are noisy and delayed. Positive labels may include confirmed document fraud, confirmed spoof attempt, mule-linked onboarding, first-party fraud reclassified after investigation, account takeover, customer-verified impersonation, and deepfake-related SAR narrative tags. Negative labels include verified legitimate sessions, but with safeguards against contamination from undetected fraud. Ambiguous cases remain in a semi-supervised pool for contrastive or self-training methods rather than being forced into binary ground truth. Class imbalance is addressed with focal loss, cost-sensitive weighting, and event-based sampling. Temporal splits are mandatory so that training predates validation and test sets, reducing leakage from repeated identities or fraud rings. Because adaptive attacks evolve quickly, rolling-window backtesting is included to evaluate drift.

Evaluation is organized around operational rather than purely academic metrics. For each stage, the model is assessed using AUROC, precision-recall AUC, equal-error rate for biometric spoofing components, false reject rate for legitimate customers, false accept rate for attacks, and expected cost under bank-specific action policies. Additional metrics include manual-review rate, customer-friction rate, escalation yield, time-to-decision, and marginal fraud-loss reduction. Calibration is assessed through reliability curves and Brier score because a well-ranked but poorly calibrated model can produce costly operational thresholds. Cross-domain robustness tests evaluate performance under changes in camera quality, compression, telephony channels, document templates, demographic distribution, and attack novelty. Adversarial tests simulate replay attacks, voice cloning, injected video streams, and stitched synthetic identity applications using red-team protocols consistent with NIST and emerging sector guidance.

Explainability is designed at both local and global levels. Global explainability includes feature-group importance, cohort performance, stability over time, and graph-community diagnostics. Local explainability includes session-level reason codes such as “document portrait-live face mismatch,” “video-injection signature,” “voice anti-spoofing alert,” “device previously linked to denied applications,” or “identity cluster linked to mule beneficiaries.” These explanations are not treated as cosmetic outputs; they determine which cases route to specialized queues, such as document-forensics review, biometric challenge, contact-center script escalation, or AML investigation. Governance controls include challenger models, threshold committees, bias testing, periodic relabeling audits, immutable event logging, and restrictions on using protected or proxy variables in ways that could create unjustified disparate impact.

The methodology also embeds smart friction. Instead of treating every elevated score as grounds for denial, the system can invoke adaptive controls such as higher-assurance challenge-response prompts, out-of-band confirmation to a validated address, forced camera recapture, delay-and-review for high-value disbursements, or branch-assisted verification. This matters because a core banking objective is to detect abuse without excluding legitimate customers, especially those with poor camera quality, accents, speech disabilities, or atypical documents. The proposed research design therefore treats deepfake defense as a human-machine orchestration problem. Machine learning identifies risk patterns at scale, but final control quality depends on escalation design, staff training, and the integrity of audit evidence. That orientation makes the framework suitable not only for academic evaluation but also for implementation, validation, and supervisory discussion in U.S. banking environments.

Data assembly for empirical implementation should follow a staged labeling protocol. Historical fraud cases, manual-review outcomes, chargebacks, account closures for fraud, customer attestations, and SAR-related internal investigations should be joined at the event-bundle level using privacy-preserving identifiers. To reduce hindsight bias, labels should be time-stamped according to when the institution first obtained reasonably reliable evidence of fraud. Negative samples should include legitimate events and “resolved legitimate” cases that passed friction or review. A quarantine set of unresolved cases should be maintained for sensitivity analysis because forcing uncertain outcomes into clean classes can overstate model performance.

A rigorous experimental design would estimate performance at three layers: component, fusion, and intervention. Component evaluation tests each detector family separately, such as document fraud models, face PAD models, voice anti-spoofing models,

or graph models. Fusion evaluation measures whether combined signals improve discrimination and calibration beyond the best single modality. Intervention evaluation estimates business impact by simulating decision policies. For example, researchers can test whether routing the top 0.5 percent of risk scores to manual review or applying step-up verification to the top 2 percent reduces fraud losses enough to justify friction and staffing costs. This policy-centric design helps convert machine-learning metrics into operational recommendations.

Model maintenance is also part of the methodology. The framework assumes scheduled retraining, champion-challenger comparison, drift monitoring, and feature-store governance. Drift indicators include shifts in camera metadata, document types, language mix, attack retry patterns, graph-community emergence, and divergence between predicted and observed fraud rates. A deepfake-resilient system should trigger investigation when these indicators move materially, even before classic accuracy metrics collapse. The methodology further recommends adversarial data augmentation, including replayed screen recordings, recompressed voice samples, partial face occlusions, and scripted social-engineering prompts. Such augmentation does not perfectly replicate attacker behavior, but it can reduce overfitting to clean lab conditions.

Finally, the methodology treats the human reviewer as an explicit subsystem. Manual-review interfaces should show the minimum necessary evidence for a decision: document anomalies, liveness confidence, graph links, prior attempts, and recommended next actions. Reviewer feedback should be captured in structured form to improve future models and identify policy inconsistencies. Inter-reviewer disagreement is itself a valuable label because it can reveal ambiguous edge cases, training gaps, or model explanations that are not sufficiently actionable. Incorporating reviewer reliability into the research design makes the framework more realistic and more aligned with how fraud operations actually function. Operationally.

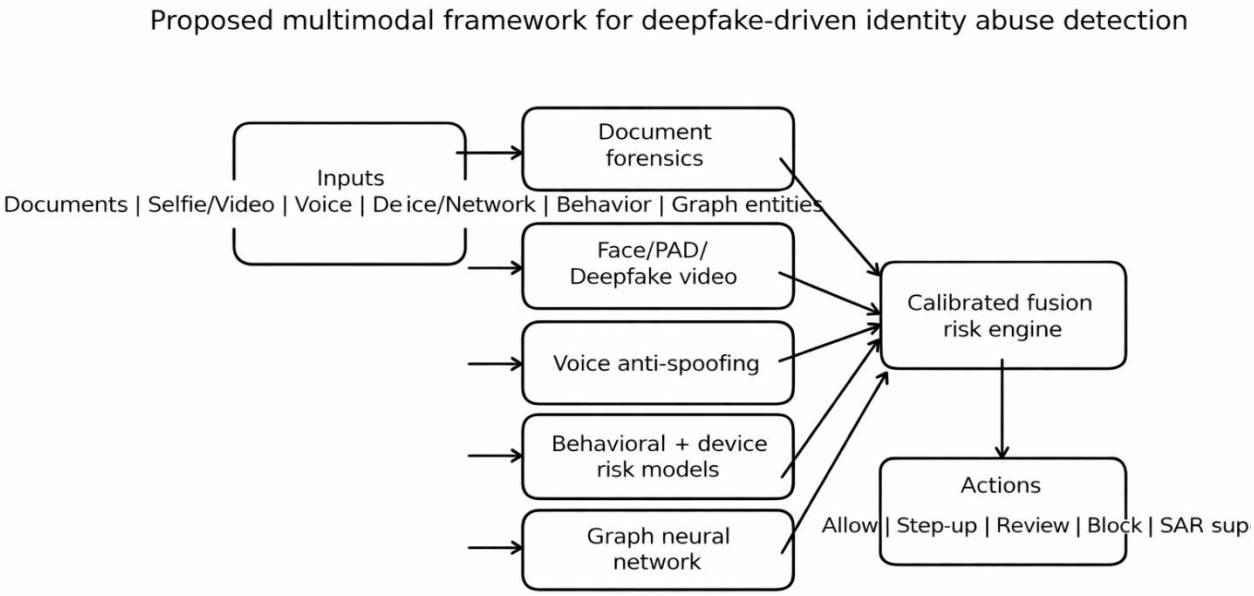

Figure 3. Proposed multimodal framework for deep fake-driven identity abuse detection.

Table 3: Evaluation design for deployment-oriented banking models

Layer	Representative metrics	Decision question answered	Operational owner
Component models	AUROC, EER, FRR/FAR, precision-recall AUC	Does each modality detect the attack class it was built for?	Fraud data science / identity engineering
Fusion model	Calibration, expected cost, top-k precision	Does signal fusion improve action quality beyond single detectors?	Enterprise fraud analytics
Policy simulation	Manual-review rate, friction rate, fraud-loss reduction	Which thresholding and smart-friction policy creates best business value?	Fraud operations / product
Governance	Stability, drift, cohort fairness, explanation quality	Can the model be defended, monitored, and remediated in production?	Model risk / compliance

4. Discussion

The proposed framework yields several implications for how banks should conceptualize and operationalize deepfake risk. The first is strategic: deepfakes should not be treated as a media-integrity niche owned solely by an identity vendor or cybersecurity team. In banking, synthetic voice, altered identity documents, replayed selfies, and injected video feeds are useful because they help attackers move money, open accounts, reset credentials, or socially manipulate frontline staff. The deepfake itself is therefore a means, not an end. This is why a journey-level architecture is superior to a stand-alone detector. A face model may flag an anomalous liveness session, but if the same device, IP range, mailing address, and beneficiary graph also connect to prior denials and mule accounts, the institution can act with far greater confidence. Conversely, a slightly suspicious media artifact may not justify friction when graph, behavior, and customer-history signals are otherwise clean.

A second implication concerns remote identity proofing. NIST's identity-proofing framework and PAD requirements make clear that verification is not only about matching a face to a document; it is about confirming that the applicant is the genuine owner of valid evidence. Deepfakes challenge exactly this linkage. An institution that outsources onboarding to a vendor but fails to collect session telemetry, event-level confidence scores, and provenance data will struggle to reconstruct why a fraudulent account passed. The proposed system addresses this by preserving intermediate evidence across modalities. In a post-incident investigation, the bank can review whether the failure originated in document validation, biometric capture quality, liveness classification, graph screening, or escalation logic. This is valuable for control remediation and for defensible conversations with auditors and supervisors.

Third, the results implied by public fraud data support the use of "smart friction" rather than blanket hardening. FTC data show that phone calls, texts, social media, websites, and email all remain significant fraud contact channels, while bank transfers and payments account for the largest reported losses by payment method. The bank's objective is to interrupt the conversion of impersonation into funded fraud. Smart friction can include delaying first outbound transfers after remote onboarding, imposing stricter thresholds for credential resets followed by beneficiary changes, or requiring dual-channel confirmation when voice-authenticated requests seek profile changes or urgent wires. In this framing, deepfake detection tunes downstream controls. The approach is more practical than a binary "deepfake yes/no" gate because many fraud losses occur only after a sequence of events that institutions can still interrupt.

Fourth, contact-center authentication deserves far more attention in the literature than it currently receives. Academic deepfake work often focuses on media benchmarks, but banks face a distinct operational challenge: agents may trust a caller not only because of biometric signals but also because of persuasive storytelling, urgency, and partial account knowledge. A cloned or converted voice can be especially dangerous when combined with stolen personal data and emotional pressure. The methodology therefore treats voice as one component in a layered system including device continuity, prior call patterns, transcript semantics, challenge-response difficulty, and post-call transaction monitoring. No realistic institution should rely on passive voiceprint matching as a sole gate for high-risk actions.

Fifth, graph analytics are likely to determine whether banks can move from isolated detection to ring disruption. Fraud operations reuse infrastructure because scale requires reuse. Devices are recycled across account applications; addresses and phone numbers appear in clusters; beneficiaries receive funds from apparently unrelated customers; and synthetic identities are often maintained in cohorts. Deepfake evidence strengthens, rather than replaces, this graph perspective. A convincing forged video may pass once. It is much less likely to pass repeatedly when each attempt enriches a graph of suspicious reuse. The

paper's emphasis on heterogeneous graphs is not ornamental. It reflects the operational reality that identity abuse is relational. Future empirical studies should compare multimodal-only models against multimodal-plus-graph models under rolling fraud-ring scenarios, because that comparison is likely to matter more to banks than incremental improvements on a single public media dataset.

Sixth, explainability is a practical necessity, not a regulatory afterthought. In a banking context, deepfake detection can trigger account denial, delayed access to funds, rejection of a dispute, or SAR filing. Such outcomes are too consequential to be justified by opaque model scores without supporting evidence. Yet explainability must be handled carefully. A bank should not disclose thresholds or attack signatures that enable attacker adaptation. The right balance is operational explainability: enough transparency for reviewers, model validators, compliance personnel, and affected customers, without publishing a playbook for evasion. Reason-code hierarchies, confidence bands, and queue-specific evidence views can achieve this balance. The user-supplied literature on algorithmic accountability and explainable AI, while not centered on deepfakes, helps illuminate why controlled interpretability is essential in regulated machine learning.

Seventh, fairness and accessibility concerns are central. Deepfake countermeasures can burden legitimate users with poor lighting, low-end devices, unstable network conditions, accents, speech impairments, facial differences, age-related appearance changes, or limited familiarity with digital capture instructions. Overly aggressive anti-spoofing may therefore reproduce a familiar financial-services problem: stronger controls that disproportionately inconvenience already vulnerable customers. The solution is not to relax fraud defenses indiscriminately. It is to implement multi-path verification and measure friction outcomes by cohort. For example, a customer who fails an automated selfie check should have a documented alternate path, such as branch-assisted verification, trusted-referee review, or asynchronous document re-capture. Bias testing should examine error rates not only by demographic proxies where legally permissible, but also by device class, channel, language, and disability-related accessibility scenarios. A publishable banking study should include such operational fairness analysis because exclusion risk is part of system quality.

Eighth, the governance burden of deepfake detection will likely increase as institutions incorporate generative tools of their own. Many banks are experimenting with AI assistants, automated customer communication, and synthetic content for internal productivity. This creates a subtle governance challenge: institutions defending against malicious synthetic media may simultaneously generate benign synthetic or AI-enhanced artifacts inside their own environments. Provenance, watermarking, signing, content-origin metadata, and internal model inventories therefore become part of fraud defense. If customers cannot distinguish official communication from scammer communication, the attack surface expands. Banking strategy should thus connect deepfake detection with broader trust architecture, including outbound communication controls, call-back policies, signed messaging, and customer education.

Ninth, the proposed framework has implications for SAR quality and AML integration. FinCEN's deepfake alert specifically requests the key term FIN-2024-DEEPFAKEFRAUD in relevant SAR narratives. A system that logs reasoned evidence across document, biometric, and transaction layers can improve the specificity of such reporting. Instead of vague narratives about "suspected identity fraud," banks can describe the sequence: altered credential presented during onboarding, liveness anomalies suggestive of replay or injection, device linked to prior denied applications, and rapid funds movement to a shared beneficiary network. In other words, machine learning contributes not only to prevention but also to institutional memory and financial-intelligence quality.

Tenth, the main technical challenge is adversarial adaptation. Deepfake detectors often degrade not because they were badly trained, but because the attack surface changes faster than the training distribution. Better generative models reduce artifact-based cues. Attackers can proxy human behavior through real-time operator steering. Injection tools can bypass camera-capture assumptions. Fraud rings can stage "warm-up" activity to look legitimate. For this reason, the proposed architecture relies on diversity: modality-specific detectors, graph linkage, temporal monitoring, challenge-response, and human review. Diversity raises attacker cost because defeating one layer no longer guarantees success. This is a classic resilience principle, but it is especially important in banking, where the institution must continue to serve millions of legitimate sessions while only a small fraction are adversarial. The system must therefore identify rare, evolving threats without destabilizing routine customer experience.

Finally, the broader significance of this research is that it reframes digital trust in banking. Historically, institutions could assume that seeing a face, hearing a voice, or reviewing a familiar document provided meaningful evidence of authenticity. Deepfake-capable adversaries erode those assumptions. The way forward is not to abandon remote banking but to redesign it around corroboration. Authenticity should emerge from the agreement of multiple, independently difficult-to-forge signals: evidence validity, biometric liveness, behavioral continuity, relational context, and downstream conduct. When those signals disagree, the institution should slow down, gather more evidence, and document its reasoning. That principle is both technologically realistic

and institutionally defensible. It acknowledges that no model will permanently "solve" deepfakes, but well-governed multimodal machine learning can materially reduce the probability that synthetic identity abuse becomes funded fraud in the U.S. banking system.

An eleventh implication involves third-party dependency risk. Many institutions use external vendors for document capture, facial matching, liveness checks, sanctions screening, or call analytics. A bank may therefore operate a fragmented control stack in which no single party has full visibility into the customer journey. Deepfake defense under these conditions requires careful orchestration and contracting. Banks need event-level data rights, service-level expectations for adversarial updates, model documentation, drift alerts, and evidence retention. Otherwise, they may inherit a black-box defense that performs well in vendor demos yet fails under institution-specific attack patterns. The issue is especially acute when multiple vendors each assume another layer is compensating for risk.

Twelfth implication concerns incident response and customer trust after failure. Even a strong multimodal framework will not eliminate all successful impersonation events. Institutions therefore need response protocols tailored to synthetic-media incidents. These include rapid containment of affected accounts, provenance review of onboarding or recovery sessions, reversal or recall attempts for outbound payments, targeted outreach to potentially affected customers, and structured learning loops into model governance. A bank that cannot explain how a deepfake-enabled account was opened or how a cloned-voice request bypassed controls will struggle to restore trust.

Thirteenth implication is that public-private information sharing should become more operationally granular. Today, many fraud typologies are shared as broad advisories. Deepfake defense would benefit from more structured exchange of attack indicators, such as replay-tool fingerprints, document-template abuse patterns, challenge-response failure modes, and graph motifs associated with coordinated onboarding attacks. Sector organizations and regulators can help by promoting standardized taxonomies that separate document forgery, biometric replay, biometric injection, voice synthesis, and composite attacks. Without such standardization, institutions will continue to mix heterogeneous incidents under broad labels like "identity theft" or "account takeover," making learning slower than attacker innovation.

Fourteenth implication is that customer education needs to evolve. Fraud prevention messages have traditionally focused on phishing, password hygiene, and suspicious links. Deepfake-era banking education should also normalize call-back procedures, distrust of urgency, caution around remote screen sharing, and awareness that a familiar voice or realistic document image may still be fraudulent. This layer is not peripheral. If customers are trained to expect out-of-band confirmation and temporary holds after risky identity events, smart friction becomes easier to deploy and less likely to be interpreted as institutional incompetence. Behavioral legitimacy can therefore reinforce technical controls.

A fifteenth implication is methodological humility. There is a temptation to present deepfake detection as a race toward ever-higher classifier accuracy. But in banking, the decisive question is often whether a control stack meaningfully changes attacker economics. A model that reduces successful impersonation by forcing attackers into slower, more expensive, and more human-intensive pathways may create substantial security value even if it never reaches laboratory-style perfection. This perspective aligns the research agenda with resilience rather than benchmark vanity. It also explains why publication in this space should report operational lift, review efficiency, and fraud-loss mitigation, not only media-classification scores.

From a scholarly standpoint, the proposed framework contributes by integrating research areas that are usually studied separately. Media-forensics papers often abstract away institutional workflow. Banking fraud papers often treat identity signals as static attributes rather than contested evidence. Governance papers often discuss explainability without specifying how multimodal detectors should route cases through operations. By combining these strands, the paper suggests a more realistic direction for future empirical work: evaluate fraud prevention as an end-to-end socio-technical system in which signals, people, vendors, policies, and attackers co-evolve. That integrated view is important in U.S. banking, where harms arise not merely from false media but from false media translated into account access and financial loss. Strategically.

5. Conclusion

Deepfake-driven identity abuse has become a material banking risk because synthetic media now interact directly with remote onboarding, account recovery, call-center servicing, and payment authorization. This paper argues that effective defense requires a shift from single-modality authentication toward multimodal, journey-level machine learning. Drawing on public U.S. fraud evidence, regulatory guidance, biometric spoofing research, and fraud-analytics literature, the study proposes an ensemble framework that fuses document forensics, face and voice anti-spoofing, behavioral telemetry, and graph-based entity resolution. The central conclusion is that banks will identify deepfake-enabled impersonation more reliably when they evaluate whether identity signals corroborate one another across time, channel, and relational context. The paper also shows that technical performance alone is insufficient: smart friction, human escalation, explainability, fairness safeguards, and SAR-ready audit trails

are integral to deployment quality. In practice, the goal is not to classify media perfectly, but to prevent synthetic identity abuse from becoming account access, profile change, payment release, or mule-account proliferation. A resilient U.S. banking response will therefore depend on continuous adversarial testing, strong governance, and layered control design rather than trust in any single biometric or vendor score. The future of secure digital banking will depend on corroboration, not appearance alone. Overall resilience matters.

6. Limitations and Future Directions

This paper has several limitations. First, it is an applied research framework rather than a proprietary bank data study, so model architecture and evaluation design are grounded in public evidence, benchmark literature, and institutional guidance rather than a labeled production dataset. Second, public fraud statistics from the FTC, FBI, and related sources do not isolate deepfake incidents with precision, meaning that the empirical motivation is necessarily broader than "confirmed deepfake banking losses." Third, biometric-spoofing benchmarks and public deepfake datasets may not reflect the exact sensor, lighting, compression, and customer-behavior conditions present in real U.S. banking channels. Fourth, fraud labels in financial services are often delayed, disputed, or partially observed, which complicates supervised learning and fairness evaluation. Fifth, some user-supplied references are adjacent-domain studies rather than direct evidence on deepfake abuse in banking, so they are used selectively for predictive-analytics, explainability, and governance framing.

Future research should move in four directions. One is consortium-based empirical testing using privacy-preserving or federated methods across institutions, vendors, and channels. A second is red-team evaluation against injected video, interactive voice cloning, and cross-channel impersonation playbooks rather than static benchmark attacks. A third is better measurement of customer-friction and accessibility outcomes under smart-friction policies. A fourth is the development of standardized banking taxonomies, labels, and reporting protocols for deepfake-enabled fraud so that institutions, regulators, and researchers can compare results more consistently. These steps would make the field more cumulative, more operationally relevant, and more useful for protecting customers and the U.S. banking system. An additional limitation is that public regulatory and complaint data are reported under evolving category schemes, so longitudinal comparisons can carry definitional noise. Future studies should also investigate standardized outcome labeling for deepfake-related fraud and compare detection architectures across onboarding, recovery, and contact-center contexts rather than assuming one thresholding strategy fits all channels.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.

References

[1]. Altuncu, E., Güran, M., & Alptekin, G. İ. (2024). Deepfake: Definitions, performance metrics and standards. Frontiers in Big Data, 7, Article 1400024. https://doi.org/10.3389/fdata.2024.1400024
[2]. Arman, M., & Fahim, A. S. M. (2023). AI revolutionizes inventory management at retail giants: Examining Walmart's U.S. operations. Journal of Business and Management Studies, 5(6), 145-148. https://doi.org/10.32996/jbms.2023.5.6.15
[3]. Arman, M., Hasan, M. N., & Rasel, I. H. (2024). Clean energy transition in USA: Big data analytics for renewable energy forecasting and carbon reduction. Journal of Management World, 2024(3), 192-206. https://doi.org/10.53935/jomw.v2024i4.1196
[4]. Arman, M., Rasel, I. H., Razib, M. N. H., & Fahim, A. S. M. (2024). Big data and machine learning for sustainable waste reduction. Journal of Posthumanism, 4(2), 448-467. https://doi.org/10.63332/joph.v4i2.3361
[5]. Arman, M., Fahim, A. S. M., Razib, M. N. H., & Rasel, I. H. (2025). Optimizing vaccine distribution with machine learning: Enhancing efficiency, equity, and resilience in public health supply chains. International Journal of Innovative Research and Scientific Studies, 8(6), 2944-2953. https://doi.org/10.53894/ijirss.v8i6.10230
[6]. ENISA. (2024). Remote ID proofing good practices. European Union Agency for Cybersecurity.
[7]. Fahim, A. S. M., Ibrahim, M., Pritty, A. A., & Tania, T. A. (2023). Algorithmic accountability in U.S. consumer FinTech: Governance mechanisms for credit risk, fair lending, and financial stability. Journal of Economics, Finance and Accounting Studies, 5(4), 80-93. https://doi.org/10.32996/jefas.2023.5.4.8
[8]. Fahim, A. S. M., Pritty, A. A., Ibrahim, M., & Tania, T. A. (2024). Real-time payments and real-time fraud: A U.S. FinTech risk framework for RTP rails and consumer protection. Journal of Economics, Finance and Accounting Studies, 6(6), 134-149. https://doi.org/10.32996/jefas.2024.6.6.11
[9]. Fahim, A. S. M., Pritty, A. A., Ibrahim, M., & Tania, T. A. (2025). Explainable AI for medical debt forecasting: Integrating healthcare and FinTech data for risk prediction. Journal of Management World, 2025(6), 92-103. https://doi.org/10.53935/jomw.v2024i4.1253
[10]. Federal Bureau of Investigation, Internet Crime Complaint Center. (2024). Criminals use generative artificial intelligence to facilitate financial fraud.
[11]. Federal Bureau of Investigation, Internet Crime Complaint Center. (2024). 2024 IC3 annual report.
[12]. Federal Trade Commission. (2024). Fighting back against harmful voice cloning.
[13]. Federal Trade Commission. (2025a). Consumer Sentinel Network data book 2024.
[14]. Federal Trade Commission. (2025b, March 10). New FTC data show a big jump in reported losses to fraud to $12.5 billion in 2024.

[15]. Federal Reserve Banks. (2019). Synthetic identity fraud in the U.S. payment system.
[16]. Federal Reserve Banks. (2021). Identifying a synthetic at account opening.
[17]. Hasan, M. N., Papel, M. S. I., Rasel, I. H., Akter, S., Aktar, M. K., Abedin, M. Z., & Mani, L. (2025). Enhancing financial information security through advanced predictive analytics: A PRISMA based systematic review. Edelweiss Applied Science and Technology, 9(7), 2222-2245. https://doi.org/10.55214/2576-8484.v9i7.9142
[18]. Hasan, N., Rasel, I. H., Rahman, M., Islam, K., Arman, M., & Jahan, N. (2022). Securing U.S. healthcare infrastructure with machine learning: Protecting patient data as a national security priority. International Journal of Computational and Experimental Science and Engineering, 8(3). https://doi.org/10.22399/ijcesen.3987
[19]. Ibrahim, M., Rahman, M. M., Razib, M. N. H., & Jahan, N. (2022). Climate risk, financial stability, and global capital allocation: A predictive analytics approach to assessing climate-related financial risk in international investment markets. Journal of Business and Management Studies, 4(4), 264-276. https://doi.org/10.32996/jbms.2022.4.4.34
[20]. Ibrahim, M., Mahmud, S., Zadid, M. U., Jahan, N., Rahman, M. M., & Fahim, A. S. M. (2024). AI-driven predictive analytics framework for anti-money laundering risk management and financial infrastructure protection in U.S. banking systems. Journal of Economics, Finance and Accounting Studies, 6(1), 155-166. https://doi.org/10.32996/jefas.2024.6.6.12
[21]. Ibrahim, M., Fahim, A. S. M., Zadid, M. U., & Pritty, A. A. (2025). FinTech for climate resilience: Measuring insurance gaps, mortgage stress, and household credit risk in the United States. Journal of Economics, Finance and Accounting Studies, 7(4), 190-205. https://doi.org/10.32996/jefas.2025.7.4.15
[22]. Ibrahim, M. N. H. R., & Rasel, I. H. (2025). The role of data analytics in enhancing ESG transparency in the corporate sector of Bangladesh. Global Journal of Engineering and Technology Advances, 22(1), 81-93. https://doi.org/10.30574/gjeta.2025.22.1.0245
[23]. Jahan, N., Pritty, A. A., Ibrahim, M., Zadid, M. U., Fahim, A. S. M., & Mahmud, S. (2024). Machine learning-driven early warning analytics for identifying market manipulation, irregular trading activity, and suspicious market signals in U.S. stock markets. Journal of Computer Science and Technology Studies, 6(2), 257-283. https://doi.org/10.32996/jcsts.2024.6.2.26
[24]. Khan, F. A., Ali, R., & coauthors. (2025). Generative AI and deepfake detection in biometric systems. Cognitive Computation. https://doi.org/10.1007/s12559-025-10469-3
[25]. Liu, X., Sahidullah, M., Kinnunen, T., Yamagishi, J., Todisco, M., Delgado, H., & Evans, N. (2022). ASVspoof 2021: Towards spoofed and deepfake speech detection in the wild. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 31, 2507-2522.
[26]. Mahmud, S., Fahim, A. S. M., Rahman, M. M., Jahan, N., & Ibrahim, M. (2025). Artificial intelligence and predictive machine learning for financial fraud detection, cyber risk management, and infrastructure resilience in the U.S. banking industry. British Journal of Multidisciplinary Studies, 3(1), 58-77. https://doi.org/10.32996/bjmss.2025.4.1.6
[27]. Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. ACM Computing Surveys, 54(1), Article 1.
[28]. National Institute of Standards and Technology. (2024). Digital identity guidelines: Identity proofing and verification (SP 800-63A, revision 4 web edition).
[29]. National Institute of Standards and Technology. (2025). Adversarial machine learning: A taxonomy and terminology of attacks and mitigations (AI 100-2e2025).
[30]. Pritty, A. A., Ibrahim, M., Fahim, A. S. M., & Zadid, M. U. (2024). Generative AI and U.S. financial reporting integrity: Detecting narrative manipulation, risk disclosure gaming, and fraud signals in 10-K filings. Journal of Economics, Finance and Accounting Studies, 6(4), 113-129. https://doi.org/10.32996/jefas.2024.6.4.11
[31]. Rasel, I. H., Arman, M., Hasan, M. N., & Bhuyain, M. M. H. (2022). Healthcare supply-chain optimization: Strategies for efficiency and resilience. Journal of Medical and Health Studies, 3(4), 171-182. https://doi.org/10.32996/jmhs.2022.3.4.26
[32]. Rasel, I. H., Ibrahim, M., Pritty, A. A., Fahim, A. S. M., & Jahan, N. (2023). Beyond FICO: Enhancing mortgage default forecasting and inclusive lending via multimodal graph neural networks and urban mobility analytics. Frontiers in Computer Science and Artificial Intelligence, 2(2), 62-81. https://doi.org/10.32996/fcsai.2023.2.2.5
[33]. Rasel, I. H., Razib, M. N. H., & Zadid, M. U. (2026). Explainable AI for institutional fraud decisions: A cross-sector empirical study using public healthcare and financial transaction data. Journal of Computer Science and Technology Studies, 8(1), 97-106. https://doi.org/10.32996/jcsts.2025.8.1.7
[34]. Razib, M. N. H., Ibrahim, M., & Rasel, I. H. (2025). Predictive analytics and its role in optimizing sustainable supply chain performance. International Journal of Business Management, 8(1), 12-27. https://doi.org/10.35409/IJBMER.2025.3640
[35]. Sunil, R., Mer, P., Diwan, A., Mahadeva, R., & Sharma, A. (2025). Exploring autonomous methods for deepfake detection: A detailed survey on techniques and evaluation. Heliyon, 11(8), e42273. https://doi.org/10.1016/j.heliyon.2025.e42273
[36]. U.S. Department of the Treasury, Financial Crimes Enforcement Network. (2024a). FinCEN alert on fraud schemes involving deepfake media targeting financial institutions (FIN-2024-Alert004).
[37]. U.S. Department of the Treasury, Financial Crimes Enforcement Network. (2024b). Notice on the use of counterfeit U.S. passport cards to perpetrate identity theft and fraud schemes at financial institutions.
[38]. Verdoliva, L. (2020). Media forensics and deepfakes: An overview. IEEE Journal of Selected Topics in Signal Processing, 14(5), 910-932.
[39]. Yamagishi, J., Kinnunen, T., Todisco, M., Delgado, H., Sahidullah, M., Wang, W., Evans, N., Singh, J., & Lee, K. A. (2021). ASVspoof 2021: Accelerating progress in spoofed and deepfake speech detection. arXiv. https://arxiv.org/abs/2109.00537
[40]. Yu, Z., Qin, Y., Zhao, X., Li, C., Gedeon, T., & Harandi, M. (2023). Deep learning for face anti-spoofing: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(5), 5609-5631.