GetReal

Deepfakes:

You’re Detecting Fraud After It’s Already Authorized

“Deepfake-enabled fraud occurs outside what today’s fraud controls are trained to identify.”

What Fraud Leaders Need to Know Now

The controls you rely on to prevent fraud are being undermined by AI-generated voice and video that socially engineer legitimate customers, employees, and executives during live interactions. These attacks happen before transactions occur, before fraud detection models see the data, and before rules engines can intervene. They exploit trust in conversations rather than weaknesses in payment systems or transaction monitoring.

When identity is assumed during a phone call, video conference, or customer service interaction, fraud controls activate too late — after authorization has already been granted and money has already moved. Losses occur even when detection systems, authentication protocols, and monitoring models function exactly as designed.

This is not a technology failure. It is a control design failure. Fraud programs were built on the assumption that voice and video provide reliable identity verification. AI has invalidated that assumption.

Why This Is a Fraud Leader-Level Issue

Your fraud prevention program is designed to stop losses before they occur. Deepfakes introduce a timing problem: authorization happens during live interactions, in many cases before fraud detection systems have visibility or proper telemetry on deepfake attacks.

- **Loss attribution and performance metrics:** Fraud incidents involving deepfakes may be classified as preventable losses under your program, regardless of how sophisticated the attack
- **Executive and board scrutiny:** CFOs and Chief Risk Officers will question why known risks were not addressed and why existing controls failed to prevent material losses
- **Regulatory and audit exposure:** Regulators and auditors may challenge whether controls remained "reasonable and appropriate" given widely understood deepfake capabilities
- **Customer trust and reimbursement pressure:** Authorized Push Payment (APP) fraud involving deepfakes drives complaint volume, reimbursement demands, and erosion of customer trust in your ability to protect them
- **Program credibility:** Repeated deepfake-enabled fraud incidents lead to scrutiny of the fraud prevention program and raise questions about the function's ability to adapt to emerging threats

A successful deepfake attack doesn't just create a loss. It creates accountability for failing to anticipate a known and growing threat vector.

What "Reasonable Preparedness" Looks Like

Leading fraud prevention programs are implementing controls designed to verify identity and intent beyond what customers and employees see and hear:

- **Enhanced authentication for high-risk transactions:** Implement step-up verification that uses channels separate from the initial contact (e.g., callback to verified number, step up to video, in-app confirmation, physical token)
- **Behavioral and contextual analysis:** Deploy models that flag anomalies in transaction patterns, device fingerprinting, location, timing, and interaction behavior — not just the content of the request
- **Call center and customer support hardening:** Train agents to use verification protocols that go beyond voice recognition and knowledge-based authentication (KBA), particularly for account changes, MFA resets, and fund movements
- **Real-time deepfake detection:** Invest in voice and video analysis tools that identify synthetic media indicators during live interactions, providing agents with risk signals
- **Multi-party authorization for high-value actions:** Require dual approval or out-of-band confirmation for large transfers, account ownership changes, or security setting modifications
- **Customer education and awareness:** Proactively inform customers about deepfake risks and establish verification expectations for unusual requests

Real Consequences

Fraud losses from deepfake-enabled attacks are already occurring across industries, with documented incidents showing that traditional fraud controls cannot prevent these schemes:

- In 2024, a multinational firm's finance director was tricked into transferring $25 million after joining a video conference call with deepfaked senior executives. The attack bypassed all authorization controls by exploiting visual and audio trust. Fraud detection systems had no opportunity to intervene before the funds were transferred. (Source)
- Financial institutions throughout 2023-2025 reported significant losses from AI voice-based vishing attacks, where fraudsters passed call-center authentication checks using cloned voices to reset credentials, change contact details, or unlock accounts.
- Research from Gartner in 2025 revealed that 30% of surveyed enterprises experienced deepfake video used against automated face biometrics or identity verification in the prior 12 months, while 32% experienced deepfake audio used against automated voice biometrics. (Source)
- Fraud rings are using synthetic identities, full identity packs (ID images, matching selfies, personal data), and real-time impersonation to defeat customer verification during account opening, live support interactions, and authorization workflows.
- Industry reports claim that 10% of banks have experienced deepfake vishing losses of more than $1 million, with an average loss of $600,000 per incident. (Source)

Bottom Line

Deepfakes exploit the gap between when authorization is granted (during live interaction) and when fraud controls activate (during transaction processing). Fraud prevention must move upstream into identity verification, not just transaction monitoring.

Recommendation

Treat deepfake-enabled fraud as a systematic control gap requiring investment in multichannel deepfake detection and real-time identity verification. Fraud programs that rely solely on post-authorization detection will continue to experience preventable losses.

About GetReal Security

GetReal Security is the cybersecurity leader specializing in the detection and mitigation of threats posed by malicious generative AI content including deepfakes and impersonation attacks. Its technology serves multinational corporations, financial institutions, media organizations, government agencies, and social media companies.

The company was incubated by Ballistic Ventures, the venture capital firm dedicated exclusively to funding and incubating entrepreneurs and innovations in cybersecurity, and Dr. Hany Farid, the preeminent expert in media forensics.

getrealsecurity.com