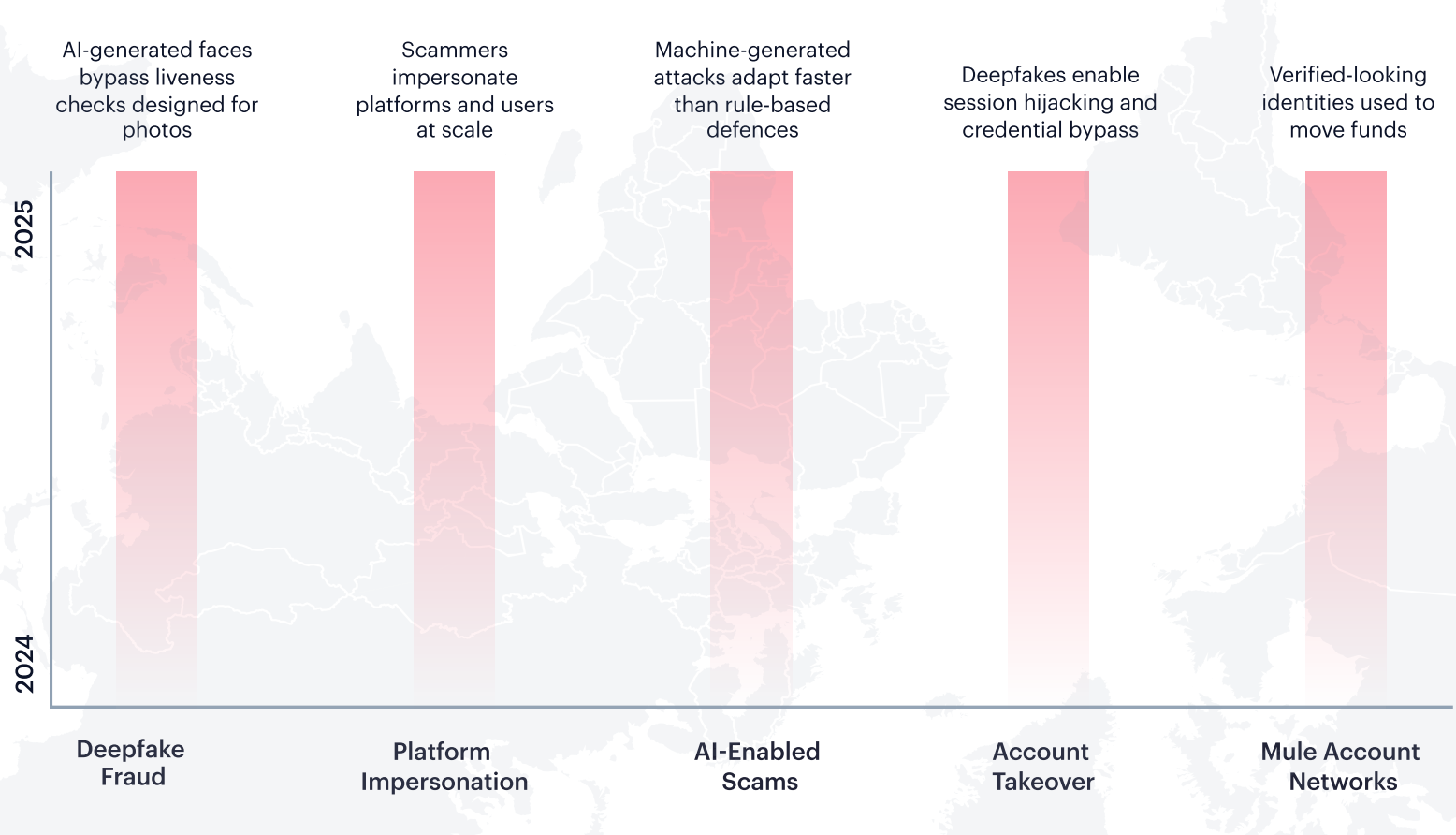
Shufti

VIDEOIDENT

Multi-Layered Defence Against Deepfakes, Synthetic Media, And Identity Fraud

The Fraud Landscape Has Changed

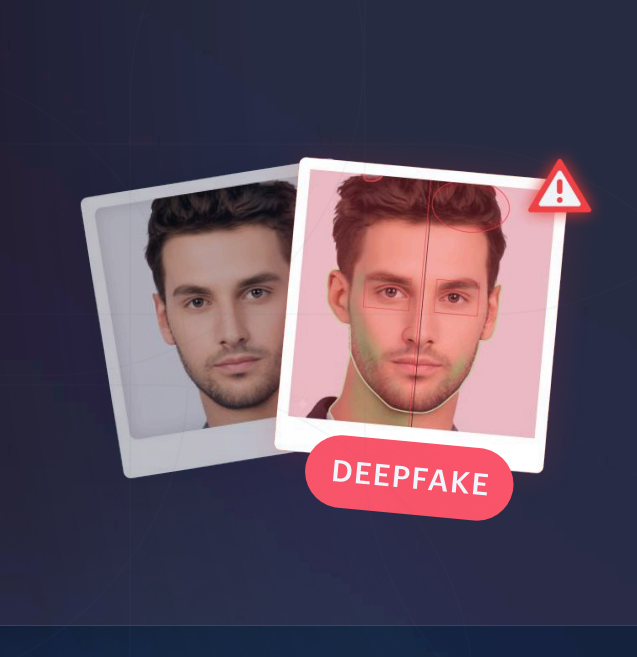


The core problem: Many verification systems, whether automated or agent-led, were designed before generative AI made synthetic media trivial to produce. They check surface signals, not forensic depth.

Attack Types Targeting VideoIdent Sessions

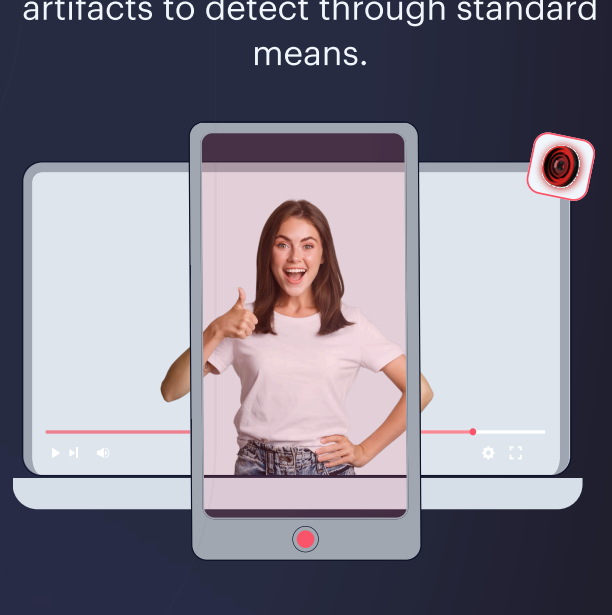
Face Swap Attacks

Attacker replaces their face with another person's face in real time. Identity is fully substituted: the person on camera isn't who they claim to be. Modern face swaps can bypass basic liveness checks using a single stolen photo.



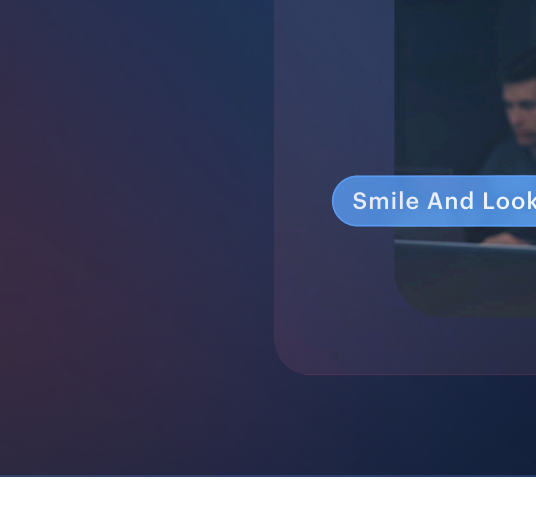
Video-Driven Attacks

Attacker animates a static face image using their own movements, creating a "puppet" effect. Turns a single stolen image into live-looking video. Highly effective against basic liveness tests.



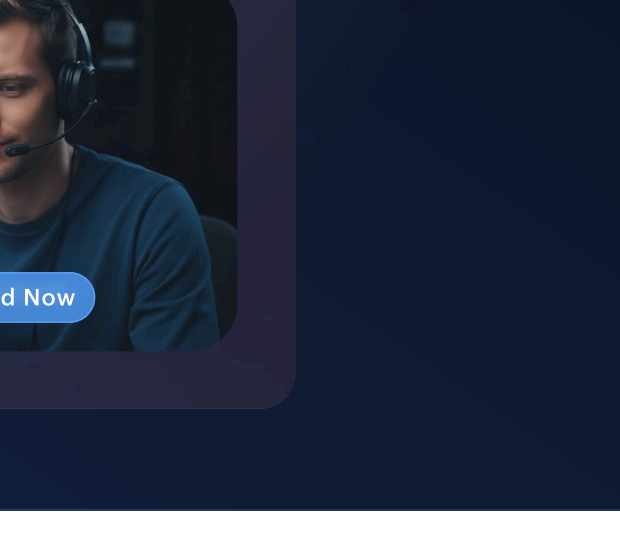
Attribute Edit Attacks

Attacker modifies specific facial attributes (age, glasses, skin tone, facial hair) to match an ID photo. Changes are subtle and help bypass appearance consistency checks.



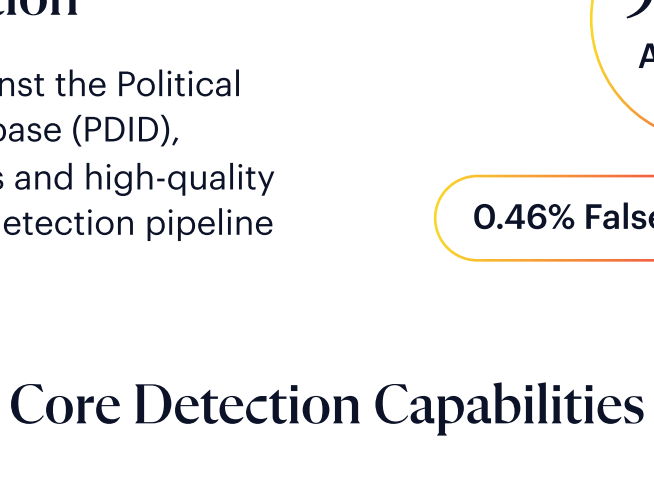
Injection Attacks

Pre-recorded or manipulated video injected directly into the session, bypassing the device camera entirely. Attacker has unlimited time to perfect the deepfake. No real-time generation artifacts to detect through standard means.



Coached Impersonation

A real person completes verification while being directed by a fraudster, through visible prompts, audio coaching, or coercion. A genuine person passes all technical checks; the fraud is social engineering.



How Shufti Detects And Prevents These Attacks

Layer 1

Forensic AI Detection

In benchmark testing against the Political Deepfakes Incidents Database (PDID), including both cheapfakes and high-quality synthetic media, Shufti's detection pipeline achieved:

99.54% Accuracy

0.46% False Acceptance Rate

Core Detection Capabilities

Multi-Stream Architecture
Processes images through parallel RGB and high-frequency DCT (Discrete Cosine Transform) streams. Captures manipulation artifacts invisible to standard analysis. Detects compression-resistant "scars" in the frequency domain that persist through screenshots and re-encoding.

Metadata-Independent Detection
Does not rely on EXIF data, device signatures, or file provenance. Analyses intrinsic content, effective even when metadata is stripped, altered, or never existed.

Active and Passive Liveness
Active liveness prompts users to position their face and respond to challenges. Passive liveness runs silently, assessing light reflection patterns, skin texture, and depth to confirm genuine presence.

Device Forensics
Tracks micro-movements: blink frequency, pupil dilation, lip sync timing. Flags unnatural stillness, erratic eye movements, or mechanical motion patterns that may indicate synthetic generation.

Multi-Model Fusion
Combines outputs from multiple specialised detection models, each analysing different facial regions and artifact types. Weighted scoring produces unified risk assessment.

Behavioural Biometrics (AI Analysis)
Processes images through parallel RGB and high-frequency DCT (Discrete Cosine Transform) streams. Captures manipulation artifacts invisible to standard analysis. Detects compression-resistant "scars" in the frequency domain that persist through screenshots and re-encoding.

Continuous Adaptation
Automated pipeline monitors new deepfake generators. Retraining triggered when new methods exceed detection thresholds. Models updated without service disruption.

Noise-Signature Detection
Identifies generator-specific noise patterns left by GANs and diffusion models. Detects structural inconsistencies and frequency-domain artifacts that persist even when watermarks and compression traces are removed.

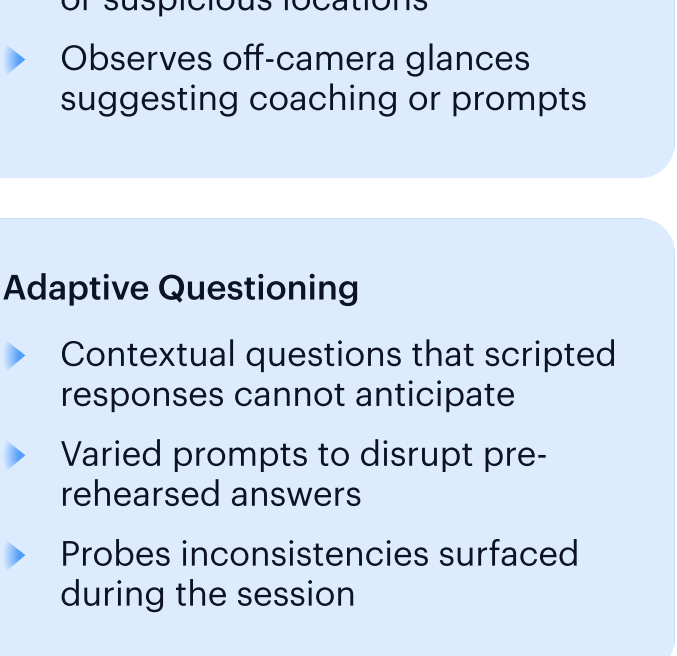
Detection Coverage

Attack Type	Primary Detection Method
Face Swap	Frequency analysis, blending boundary detection, texture inconsistency
Facial Reenactment	Motion analysis, lip sync verification, expression transfer artifacts
Attribute Edit	Facial region comparison, localised manipulation detection
Injection Attack	Device forensics, sensor signature verification, virtual camera detection
Cheapfakes	Multi-resolution analysis, editing tool signature detection

Layer 2

Trained Expert Verification

Technology catches forensic anomalies. Experts catch human anomalies.
Even with high detection accuracy, sophisticated fraud may present borderline signals, or attack vectors that are primarily social.



The Trained Expert Layer Adds

Real-Time Behavioural Assessment
- Observes hesitation, nervousness, or scripted responses
- Notes inconsistences between claimed identity and verbal cues
- Assesses document handling, familiarity with their own ID

Environment and Context Evaluation
- Identifies unusual lighting or studio-like setups
- Notes background inconsistencies or suspicious locations
- Observes off-camera glances suggesting coaching or prompts

Coercion Signal Recognition
- Trained to recognise signs of distress or duress
- Identifies coached behaviour patterns
- Flags situations where the person may not be acting freely

Adaptive Questioning
- Contextual questions that scripted responses cannot anticipate
- Varied prompts to disrupt pre-rehearsed answers
- Probes inconsistences surfaced during the session

Challenge-Response Verification
- Random prompts designed to disrupt pre-rendered video
- Requests requiring genuine real-time response

What Trained Experts Add to the Detection Layer

Signal Type	Primary Detection Method
Off-camera glances or audio cues	May indicate coaching or prompting
Hesitation before "personal" answers	May indicate scripted responses
Unusual document handling	Person unfamiliar with their own ID
Environmental anomalies	Studio setup, unusual lighting, suspicious background
Verbal inconsistencies	Answers don't match claimed identity or context
Signs of distress	Potential coercion, duress, or fraud victimisation

Evidence That Supports Investigation

Full Session Recording
What It Contains: Video of entire verification interaction
Purpose
Records behaviour for review

Forensic Analysis Summary
What It Contains: Detection outputs, confidence indicators
Purpose
Documents what AI analysis flagged

Captured ID Images
What It Contains: Front, back, security features, MRZ
Purpose
Documents what was inspected

Face Comparison
What It Contains: Side-by-side user face vs. document
Purpose
Confirms matching was performed

Consent Capture
What It Contains: Timestamped explicit consent
Purpose
GDPR compliance; dispute defence

TAN/OTP Confirmation
What It Contains: Session binding verification
Purpose
Documents authenticated user completed session

Expert Assessment
What It Contains: Decision, risk flags, notes
Purpose
Audit trail; investigation starting point

Timestamps and Metadata
What It Contains: Duration, location signals, device info
Purpose
Forensic context for review

Retention: Configurable per your policy. Supports 5+ year retention for GwG/BaFin compliance requirements.

The Cost Of Getting It Wrong

Direct Loss
€1,000 - €100,000+ per incident
Funds stolen, chargebacks, remediation

Regulatory Fine
Reported penalties: €5M - €500M
Delayed SARs, inadequate controls, AML failures

Operational Burden
Staff time, consultants
Investigation, reporting, remediation

Customer Friction
Lost revenue, support costs
False positives block legitimate users

Reputation Damage
Hard to quantify; long recovery
Media coverage, customer churn

The math: At **€5 per session**, helping block one **€10,000** fraud event pays for **2,000** verifications.

Why Shufti for Fraud Prevention

Capability	What You Get
99.54% Benchmark Accuracy	Forensic AI validated against real-world deepfake datasets
Multi-Stream Analysis	RGB + frequency domain + device forensics + behavioural biometrics
Injection Detection	Camera signature database designed to detect virtual camera
iBeta Level 2 PAD Certified	Independent presentation attack detection validation
Trained Expert Network	Specialists trained on fraud indicators, coercion recognition
24/7 Availability	Matches fraud activity patterns across time zones
Continuous Adaptation	Automated retraining as new generators emerge
Flexible Deployment	Shufti agents, your agents, or hybrid
BaFin/FMA Aligned	Evidence supports DACH regulatory audit requirements
As low as €5 per verification	Apply fraud defence broadly without budget constraints

Ready To See It In Action?

Forensic AI detection. Trained expert verification. Evidence designed for audit and investigation.

www.shufti.com
sales@Shufti.com

Book a Demo