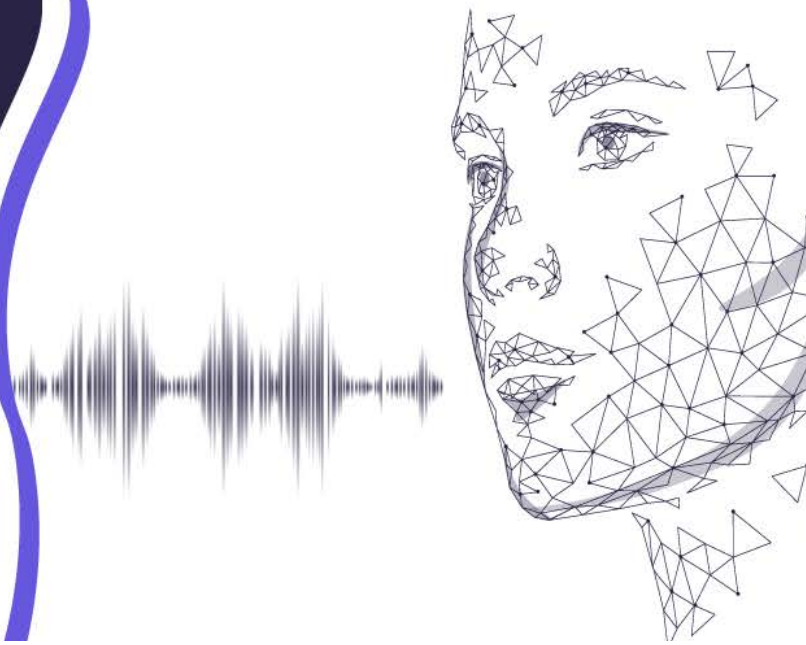
validsoft

Deepfake Audio Detection Solution:

ValidSoft *Voice Verity*™



Deepfake Audio – A New Threat

Deepfake audio, also known as synthetic audio, is a type of artificial intelligence-generated audio that can mimic the voice of a real person to a high degree of accuracy.

Deepfake audio is created by training a neural network on datasets of audio recordings of the target speaker. This technology has a wide range of potential applications, but it also has the potential to be used for malicious purposes, such as committing financial fraud. In this white paper, we will explore the concept of synthetic audio and the potential risks associated with it in the context of financial fraud. We will also discuss how ValidSoft's AI-based Deepfake Detection product – Voice Verity™ – can be used by any enterprise to detect synthetic audio and prevent deception, identity theft, financial fraud and reputational damage.

Deepfake audio, is an emerging method for perpetrating fraud and identity theft. Recently publicized examples include using deepfake technology from startup company ElevenLabs; security professionals fooling their own banks' voice biometric solution (with a major UK and USA bank), or agent handling the call, and another example whereby the actress Emma Watson appearing to be reading *Mein Kampf*. The first published incident of significant fraudulent use of deepfake audio occurred in 2019[1]

According to the Wall Street Journal, the CEO of a UK energy firm received a phone call from (what sounded like) the CEO of the company's German parent company, requesting an urgent transfer of about €220,000 ($243,000) to a Hungarian supplier. As the UK CEO recognized the German CEO's voice, including "the hint of a German accent and the same melody", he complied with the request. The money was subsequently moved from Hungary to a Mexican account and further disbursed. The details of the attack, but not the company, were shared with the WSJ by the company's insurer.

What Deepfake Exploitation Looks Like

As well as demonstrating that this form of deepfake attack is now in the wild and no longer just theoretical, it also confirms that fraudsters are capable of using synthetic "voice skin" technology, as distinct from creating static deepfake recordings. A synthetic voice skin allows a fraudster to speak in a conversational manner with the target, with the fraudster's voice being converted by the "skin" to sound like the impersonated voice. Whereas previously, voice skins have been of lower quality than static deepfake recordings, the technology used in this incident was clearly good enough to fool a CEO who would easily recognize his colleague's voice. A "voice skin" is one type of deepfake attack, and as with any fraud vector that is shown to work, it will only become more common and be used in ever more original ways. In the case of deepfakes, the technology will evolve, allowing them to become ever more realistic. The rapid advances in usage of Large Language Models (LLMs) in AI applications such as ChatGPT have created opportunities for fraudsters/hackers, and we are already seeing the impact of the convergence of such technologies with Deepfake Audio to create very sophisticated virtual bad actor impersonations/attacks.

Unified Communications and Omni-Channel strategies mean organizations, including banks, will increasingly communicate with their customers using browser-based video/audio, for instance. This could be with a human agent, but in the future, also Artificial Intelligence (AI) based agents.

Imagine, therefore, an audio conversation between a high net-worth client and their private banker. If the client sounds authentic and, of course, can provide the answers to some security questions (as they invariably would), why would the banker not acquiesce to any instructions the "client" gives?

1. see this link for the Wall Street Journal article

How ValidSoft Stops Deepfakes

ValidSoft has been at the forefront of synthetic speech detection for years, and developed its first such applications to strengthen its voice biometric systems in around 2012. In addition, ValidSoft participated in the multi-Billion Euro, EU-funded, Horizon 2020 project since 2015, focusing on detecting synthetically generated speech. ValidSoft collaborated with some of the world's leading academic institutions on this cutting R&D work, including working directly with the institutions behind the most famous series of academic challenges for fake audio detection, ASVspoof.

Our anti-spoof detection capabilities, particularly our Deepfake detection solutions, are based on years of experience and research in speech science, signal processing and voice biometrics and are based on complex machine learning, AI, and large-scale Deep Neural Network (DNN) techniques.

Users of ValidSoft's latest-generation of voice biometric technologies are already protected against deepfake audio, as Deepfake detection is integral to those solutions. However, for the first time, enterprises can now deploy Deepfake detection to organizations that don't use voice biometrics at all, as a fully standalone solution.

The Solution: Voice Verity™

ValidSoft's patent protected Voice Verity™ Deepfake detection solution has a compelling value proposition, and is immediately available in multiple deployment configurations, including: Cloud, Private Cloud, On-premise, Hosted and SaaS. The technology can be integrated with any customer engagement channel that supports audio, either real-time streamed audio or "at rest". Moreover, because Voice Verity™ is not a biometric solution, it requires no user enrollments, no consent workflows, and does not hold or store any Personally Identifiable Information or "personal data" under GDPR or similar privacy frameworks, so can be deployed and operational immediately (of course, enterprises can and typically do store audio recordings in their normal course of business, and for future evidence and auditing purposes, but this is not a requirement of Voice Verity™).

For users of legacy biometric solutions with no/inadequate deepfake detection capability, ValidSoft's Deepfake detection solution can run as a standalone service in parallel, requiring just an audio feed, providing a strong and additional layer of protection for any enterprise.

Key features of Voice Verity™:

- Immediately available to all enterprises, and the technology is already deployed in real world live deployments with some of the world's largest enterprises and financial services institutions.
- Totally standalone product, decoupled from any existing voice biometric deployments.
- Highly accurate, and not limited to (or dependent on) specific watermarking or specific synthetic audio tools.
- Can process and search for Deepfake audio in real time, or as a retrospective batch process.
- 100% privacy compliant, including with GDPR, CCPA, BIPA, and all major privacy law frameworks.
- Can be deployed to provide visual real-time Alerts to (contact center) agents or users in a fraud control panel for escalation handling.

Conclusion

Deepfake audio is here to stay, and exploitation, both good and bad, will increase. Staying ahead of the threat is key, and ValidSoft can help any organization gain a competitive advantage and protect its reputation and its customers by leading the field in the detection and prevention of such attacks.

For more information, please visit www.validsoft.com or contact ValidSoft at info@validsoft.com.