# Deepfake technology poses rising threat as criminals exploit AI



Generative artificial intelligence tools, known for their rapid improvement, are increasingly being exploited by criminals for nefarious purposes, notably in the realm of deepfake technology. These tools produce deceptively realistic audio and video content, an alarming development that has led to a rise in scams targeting vulnerable individuals and businesses alike.

Debby Bodkin shared a distressing incident involving her 93-year-old mother, who received a phone call featuring a cloned voice claiming, "It's me, mom... I've had an accident." The impersonator provided details that matched those of a hospital, attempting to extract money under the guise of an emergency. Fortunately, the call was intercepted by Bodkin's granddaughter, who contacted her to confirm her safety. "It's not the first time scammers have called grandma," Bodkin remarked in an interview with AFP. "It's daily."

Such deepfake scams often manipulate victims into funding fictitious medical emergencies. This tactic extends beyond individual cases; criminal organisations also leverage deepfakes to infiltrate companies. In a notable instance, Hong Kong police earlier disclosed that an employee of a multinational firm was deceived into transferring HK$200 million (approximately US$26 million) to fraudsters, who utilised AI-generated avatars of his colleagues during a fake videoconference.

Research conducted by identification start-up iBoom revealed a startling statistic: only a tenth of one percent of Americans and Britons could accurately identify a deepfake image or video. Vijay Balasubramaniyan, CEO of voice authentication company Pindrop Security, noted that advancements in generative AI have significantly reduced the time required to replicate a voice. "Before, it took 20 hours (of voice recording) to recreate your voice," he explained, adding, "Now, it's five seconds."

In response to the growing prevalence of deepfakes, technology firms are investing in detection solutions. Intel has developed "FakeCatcher," a tool that discerns legitimate from altered images by analysing changes in facial blood vessel colour. Additionally, Pindrop's technology examines audio characteristics meticulously to identify discrepancies indicative of artificial creation. Nicos Vekiarides, chief of the Attestiv platform, emphasised the need for ongoing innovation in this area: "You have to keep up with the times," he stated. He remarked on the initial clear flaws in deepfakes, such as the inclusion of impossible features, but noted the technology's progression has made these forgeries increasingly convincing.

The escalation of deepfake technology has led experts to label it a "global cybersecurity threat." Vekiarides cautioned, "Any company can have its reputation tarnished by a deepfake or be targeted by these sophisticated attacks." With a shift toward telecommuting, there are increased opportunities for scammers to masquerade as legitimate personnel and manipulate access to sensitive information.

Consumers are also turning to innovative solutions to guard against deepfake fraud. In January, the Chinese company Honor introduced its Magic7 smartphone, which features a built-in deepfake detector powered by AI. Meanwhile, British start-up Surf Security released a web browser aimed at helping businesses identify synthetic audio and video content.

Looking ahead, Siwei Lyu, a professor of computer science at the State University of New York at Buffalo, anticipates that deepfakes will eventually become as ubiquitous as spam on the internet. He projected that detection algorithms would evolve to function similarly to spam filters currently used in email systems, although he acknowledged that the technology has yet to reach this level of maturity.

As the landscape of generative AI continues to evolve, both individuals and corporations are poised to navigate the challenges presented by deepfake technology, contemplating measures to safeguard against its misuse.

Source: [Noah Wire Services](https://www.noahwire.com)

## References

* <https://www.ic3.gov/PSA/2024/PSA241203> - This FBI warning highlights how criminals exploit generative AI for fraud, including the use of deepfakes and AI-generated content to deceive victims. It underscores the increasing believability and scale of these scams.
* <https://www.trmlabs.com/post/the-rise-of-ai-enabled-crime-exploring-the-evolution-risks-and-responses-to-ai-powered-criminal-enterprises> - This report from TRM Labs discusses the evolution of AI-enabled crimes, including deepfakes used for extortion and social manipulation, emphasizing the challenges in detecting and countering these sophisticated attacks.
* <https://unicri.org/generative-ai-new-threat-online-child-sexual-exploitation-and-abuse-september-2024> - While not directly related to deepfakes, this study highlights the broader risks associated with generative AI, including its misuse in various criminal activities, which parallels the concerns around deepfake technology.
* <https://pmc.ncbi.nlm.nih.gov/articles/PMC10311201/> - This article discusses the role of digital evidence in criminal investigations, which is relevant to the context of deepfakes and AI-generated content being used in crimes, as these technologies increasingly involve digital forensics.
* <https://www.intel.com/content/www/us/en/architecture-and-technology/fakecatcher.html> - Intel's FakeCatcher technology is mentioned as a solution to detect deepfakes by analyzing facial blood vessel color changes, highlighting efforts by tech companies to combat AI-generated fraud.
* <https://www.pindrop.com/press-release/pindrop-security-announces-new-voice-technology-to-detect-deepfakes/> - Pindrop Security's technology is designed to identify deepfakes by examining audio characteristics, reflecting the industry's response to the growing threat of AI-generated audio scams.