Making Deepfakes Gets Cheaper and Easier Thanks to A.I.

Meme-makers and misinformation peddlers are embracing artificial intelligence tools to create convincing fake videos on the cheap.

By Stuart A. Thompson
Stuart Thompson writes about online information flows.

March 12, 2023

It wouldn't be completely out of character for Joe Rogan, the comedian turned podcaster, to endorse a "libido-boosting" coffee brand for men.

But when a video circulating on TikTok recently showed Mr. Rogan and his guest, Andrew Huberman, hawking the coffee, some eagle-eyed viewers were shocked — including Dr. Huberman.

"Yep that's fake," Dr. Huberman wrote on Twitter after seeing the ad, in which he appears to praise the coffee's testosterone-boosting potential, even though he never did.

A.I. Fakes

An advertisement on TikTok combined a real video of Andrew Huberman with an altered version of Joe Rogan's voice that experts say was likely generated with artificial intelligence.

Note: Advertiser has been obscured.

The ad was one of a growing number of fake videos on social media made with technology powered by artificial intelligence. Experts said Mr. Rogan's voice appeared to have been synthesized using A.I. tools that mimic celebrity voices. Dr. Huberman's comments were ripped from an unrelated interview.

Making realistic fake videos, often called deepfakes, once required elaborate software to put one person's face onto another's. But now, many of the tools to create them are available to everyday consumers — even on smartphone apps, and often for little to no money.

The new altered videos — mostly, so far, the work of meme-makers and marketers — have gone viral on social media sites like TikTok and Twitter. The content they produce, sometimes called cheapfakes by researchers, work by cloning celebrity voices, altering mouth movements to match alternative audio and writing persuasive dialogue.

The videos, and the accessible technology behind them, have some A.I. researchers fretting about their dangers, and have raised fresh concerns over whether social media companies are prepared to moderate the growing digital fakery.

Disinformation watchdogs are also steeling themselves for a wave of digital fakes that could deceive viewers or make it harder to know what is true or false online.

“What’s different is that everybody can do it now,” said Britt Paris, an assistant professor of library and information science at Rutgers University who helped coin the term “cheapfakes.” “It’s not just people with sophisticated computational technology and fairly sophisticated computational know-how. Instead, it’s a free app.”

Reams of manipulated content have circulated on TikTok and elsewhere for years, typically using more homespun tricks like careful editing or the swapping of one audio clip for another. In one video on TikTok, Vice President Kamala Harris appeared to say everyone hospitalized for Covid-19 was vaccinated. In fact, she said the patients were unvaccinated.

Graphika, a research firm that studies disinformation, spotted deepfakes of fictional news anchors that pro-China bot accounts distributed late last year, in the first known example of the technology’s being used for state-aligned influence campaigns.

But several new tools offer similar technology to everyday internet users, giving comedians and partisans the chance to make their own convincing spoofs.

Last month, a fake video circulated showing President Biden declaring a national draft for the war between Russia and Ukraine. The video was produced by the team behind “Human Events Daily,” a podcast and livestream run by Jack Posobiec, a right-wing influencer known for spreading conspiracy theories.

In a segment explaining the video, Mr. Posobiec said his team had created it using A.I. technology. A tweet about the video from The Patriot Oasis, a conservative account, used a breaking news label without indicating the video was fake. The tweet was viewed more than eight million times.

Many of the video clips featuring synthesized voices appeared to use technology from ElevenLabs, an American start-up co-founded by a former Google engineer. In November, the company debuted a speech-cloning tool that can be trained to replicate voices in seconds.

ElevenLabs attracted attention last month after 4chan, a message board known for racist and conspiratorial content, used the tool to share hateful messages. In one example, 4chan users created an audio recording of an anti-Semitic text using a computer-generated voice that mimicked the actor Emma Watson. Motherboard reported earlier on 4chan’s use of the audio technology.

ElevenLabs said on Twitter that it would introduce new safeguards, like limiting voice cloning to paid accounts and providing a new A.I. detecting tool. But 4chan users said they would create their own version of the voice-cloning technology using open source code, posting demos that sound similar to audio produced by ElevenLabs.

“We want to have our own custom AI with the power to create,” an anonymous 4chan user wrote in a post about the project.

In an email, a spokeswoman for ElevenLabs said the company was looking to collaborate with other A.I. developers to create a universal detection system that could be adopted across the industry.

Videos using cloned voices, created with ElevenLabs' tool or similar technology, have gone viral in recent weeks. One, posted on Twitter by Elon Musk, the site's owner, showed a profanity-laced fake conversation among Mr. Rogan, Mr. Musk and Jordan Peterson, a Canadian men's rights activist. In another, posted on YouTube, Mr. Rogan appeared to interview a fake version of the Canadian prime minister, Justin Trudeau, about his political scandals.

"The production of such fakes should be a crime with a mandatory ten-year sentence," Mr. Peterson said in a tweet about fake videos featuring his voice. "This tech is dangerous beyond belief."

In a statement, a spokeswoman for YouTube said the video of Mr. Rogan and Mr. Trudeau did not violate the platform's policies because it "provides sufficient context." (The creator had described it as a "fake video.") The company said its misinformation policies banned content that was doctored in a misleading way.

Experts who study deepfake technology suggested that the fake ad featuring Mr. Rogan and Dr. Huberman had most likely been created with a voice-cloning program, though the exact tool used was not clear. The audio of Mr. Rogan was spliced into a real interview with Dr. Huberman discussing testosterone.

The results are not perfect. Mr. Rogan's clip was taken from an unrelated interview posted in December with Fedor Gorst, a professional pool player. Mr. Rogan's mouth movements are mismatched to the audio, and his voice sounds unnatural at times. If the video convinced TikTok users, it was hard to tell: It attracted far more attention after it was flagged for its impressive fakery.

TikTok's policies prohibit digital forgeries "that mislead users by distorting the truth of events and cause significant harm to the subject of the video, other persons or society." Several of the videos were removed after The New York Times flagged them to the company. Twitter also removed some of the videos.

A TikTok spokesman said the company used "a combination of technology and human moderation to detect and remove" manipulated videos, but declined to elaborate on its methods.

Mr. Rogan and the company featured in the fake ad did not respond to requests for comment.

Many social media companies, including Meta and Twitch, have banned deepfakes and manipulated videos that deceive users. Meta, which owns Facebook and Instagram, ran a competition in 2021 to develop programs capable of identifying deepfakes, resulting in one tool that could spot them 83 percent of the time.

Federal regulators have been slow to respond. One federal law from 2019 requested a report on the weaponization of deepfakes by foreigners, required government agencies to notify Congress if deepfakes targeted elections in the United States and created a prize to encourage the research on tools that could detect deepfakes.

“We cannot wait for two years until laws are passed,” said Ravit Dotan, a postdoctoral researcher who runs the Collaborative A.I. Responsibility Lab at the University of Pittsburgh. “By then, the damage could be too much. We have an election coming up here in the U.S. It’s going to be an issue.”