Synthetic media and deepfakes are here–but our economy isn't ready

BY MOUNIR IBRAHIM

July 27, 2022 at 1:01 PM EDT

Until now, generating a synthetic piece of media or even editing an existing image required knowledge, expertise, and access to tools. New A.I.-powered tools are removing these barriers.

FABRICE COFFRINI—AFP/GETTY IMAGES

In our digitized world, users can order, buy, claim, or rent anything in a matter of seconds. This is the future of the global economy.

Visuals, images, and videos are being increasingly used to inform decision-making. Everything from a food delivery app, to an online dating platform, or an insurance company paying out a virtual claim, usually begins with a picture or video.

However, the great shift toward digital systems coincides with the proliferation of the synthetic generation of images, videos, and audio. Though experts have been warning about synthetic media and deepfakes for several years now, it appears the accessibility of these tools are reaching an inflection point.

Recently, notable A.I. and synthesis achievements propelled the conversation forward and highlighted the alarming power already available to the layperson. Google's Pixel phone (6) delivers its "magic eraser" as part of the native camera. Though not wholly a synthetic image, this impressive feature leverages A.I. on devices to erase any aspect of an image and recreate pixels based on machine learning. While great for countering "photo bombs," it also gives bad actors the ability to create and disseminate certain types of sophisticated cheapfakes instantly.

Open A.I.'s DALL-E-2 and Google Brain's Imagen are remarkable steps toward the proliferation of synthetic media with text-to-image synthesis tools. These platforms can turn any descriptive sentence, such as "surveillance footage of Homer Simpson running in a mall, low quality, black and white," into a hyper-realistic photo in less than 20 seconds.

The most astounding aspect of these tools? They require absolutely no skill or knowledge to generate the images. Until now, generating a synthetic piece of media or even editing an existing image required knowledge, expertise, and access to tools. These barriers are being lifted. It's the start of a new paradigm shift in the accessibility of image alteration and synthesis to the masses.

It would be foolish to assume that as these technologies become accessible, they will not be used for fraud and deception. There have already been glimpses of image synthesis used to defraud and deceive the general public, most recently in Ukraine, where a deepfake of President Zelensky proliferated

online. Last year, warnings by the FBI and other authorities about the use of synthetic media in corporate espionage and sophisticated crimes helped raise awareness in private industry. Illicit networks, state sponsors, and sophisticated actors are worrisome enough, but this problem becomes even more concerning when access to this technology is available to practically anyone.

Two years ago, a bizarre story emerged about a Pennsylvania mother who was arrested for allegedly cyberbullying a group of teenage girls. Initial reports suggested the mother used synthetic videos depicting the young girls smoking to discredit them. The allegation was shocking. Those following closely found it hard to conceive that a non-expert could successfully access and deploy deepfake technology for such use. Most experts ruled out that Generative Adversarial Networks (GAN) (neural networks that make deepfakes) were used. However, this scenario could be a harbinger of things to come as synthetic media becomes widely accessible. That day is almost here.

DALL-E 2 has important guardrails such as a limit on the number of users (in beta and up to 1 million users now) and strict content synthesis policies to limit misuse. Google also publicly states that this technology can be seriously misused and "decided not to release code or a public demo."

However, text-to-image synthesis is a popular topic with many research papers and GitHub repositories available for less ethical actors to proceed without precaution. Even less sophisticated technologies pose a threat. The lesser DALL-E mini is publicly available and can still produce impressive synthetics in less than one minute on most textual prompts.

These images of a damaged Jeep Cherokee were created in roughly 70 seconds by simply entering "Damaged Bumper Jeep Grand Cherokee," on the DALL-E mini platform. Today, consumers can synthetically create images of defective, broken, or sub-optimal items and take advantage of consumer-friendly "no-questions-asked" corporate strategies or government policies that favor digital submissions.

It is vitally important to consider how courts and legal proceedings will address the use of visual media as proof in civil or even criminal claims. Without established and authenticated visual evidence, courts too will find it hard to differentiate fact from fiction. Deceptive imagery could be admitted as evidence, or all media could stop being used out of the inability to determine what is authentic.

Businesses, society, and our legal and governmental institutions will need to start learning, engaging with, and adopting digital content provenance open standards to help distinguish reality from fabrication.

Open standards are led by the Coalition for Content Provenance and Authenticity (C2PA) but content provenance is also supported and tested by other consortiums such as the Content Authenticity Initiative (CAI), Project Origin, and projects like Starling Labs at Stanford University.

Without a way to separate fact from fiction, the future of our digitized economy will be in serious jeopardy.

Mounir Ibrahim is the vice president of public affairs and impact at Truepic. Truepic is a steering committee member of the C2PA and a member of the CAI.

The opinions expressed in Fortune.com commentary pieces are solely the views of their authors and do not reflect the opinions and beliefs of Fortune.