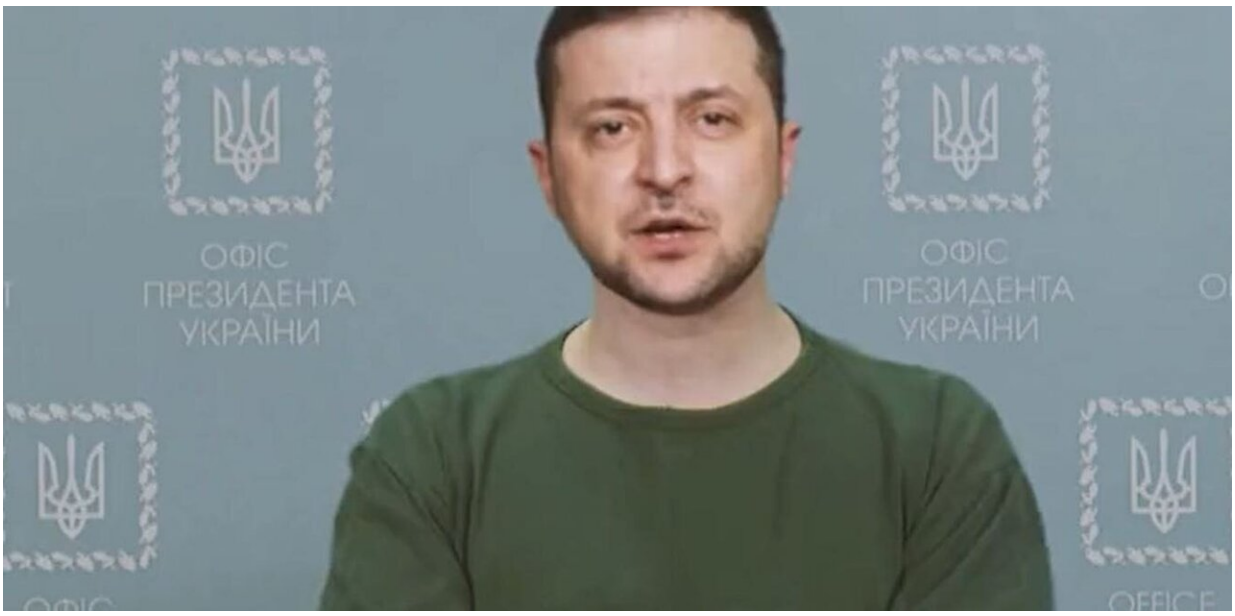
Celebrity deepfakes are all over TikTok. Here's why they're becoming common, and how you can spot them

July 19 2022, by Rob Cover



One of the world's most popular social media platforms, TikTok, is now host to a steady stream of deepfake videos.

Deepfakes are videos in which a subject's face or body has been digitally altered to make them look like someone else—usually a famous person.

One notable example is the @deeptomcriuse TikTok account, which has posted dozens of deepfake videos impersonating Tom Cruise, and attracted some 3.6 million followers.

In another example, Meta CEO Mark Zuckerberg seems to be confessing to conspiratorial data sharing. More recently there have been a number of silly videos featuring actors such as Robert Pattinson and Keanu Reeves.

Although deepfakes are often used creatively or for fun, they're increasingly being deployed in disinformation campaigns, for identity fraud and to discredit public figures and celebrities.

And while the technology needed to make them is sophisticated, it's becoming increasingly accessible, leaving detection software and regulation lagging behind.

One thing is for sure—deepfakes are here to stay. So what can we do about them?

Varying roles

The manipulation of text, images and footage has long been a bedrock of interactivity. And deepfakes are no exception; they're the outcome of a deep-seated desire to participate in culture, storytelling, art and remixing .

The technology is used extensively in the digital arts and satire. It provides more refined (and cheaper) techniques for visual insertions, compared to green screens and computer-generated imagery.

Deepfake technology can also enable authentic-looking resurrections of deceased actors and historical re-enactments. They may even play a role

in helping people grieve their deceased loved ones.

But they're also available for misuse

At the same time, deepfake technology is thought to present several social problems such as:

- deepfakes being used as "proof" for other fake news and disinformation
- deepfakes being used to discredit celebrities and others whose livelihood depends on sharing content while maintaining a reputation
- difficulties providing verifiable footage for political communication, health messaging and electoral campaigns
- people's faces being used in deepfake pornography.

The last point is of particular concern. In 2019, deepfake detection software firm Deeptrace found 96% of 14,000 deepfakes were pornographic in nature. Free apps such as the now-defunct DeepNude 2.0 have been used to make clothed women appear nude in footage, often for revenge porn and blackmail.

In Australia, deepfake apps have even allowed perpetrators to circumvent "revenge porn" laws—an issue expected to soon become more severe.

Beyond this, deepfakes are also used in identity fraud and scams, particularly in the form of video messages from a trusted "colleague" or "relative" requesting a money transfer. One study found identity fraud using digital manipulation cost US financial institutions US$20 billion in 2020].

A growing concern

The creators of deepfakes stress the amount of time and effort it takes to make these video look realistic. Take Chris Ume, the visual effects and AI artist behind the @deeptomcruise TikTok account. When this account made headlines last year, Ume told The Verge "you can't do it by just pressing a button."

But there's good evidence deepfakes are becoming easier to make. Researchers at the United Nation Global Pulse initiative have demonstrated how speeches can be realistically faked in just 13 minutes.

As more deepfake apps are developed, we can expect lesser-skilled people to increasingly produce authentic-looking deepfakes. Just think about how much photo editing has boomed in the past decade.

Legislation, regulation and detection software are struggling to keep up with advances in deepfake technology.

In 2019, Facebook came in for criticism for failing to remove a doctored video of American politician Nancy Pelosi, after it fell short of its definition of a deepfake.

In 2020, Twitter banned the sharing of synthetic media that may deceive, confuse or harm people (except where a label is applied). TikTok did the same. And YouTube banned deepfakes related to the 2020 U.S. federal election.

But even if these are well-meaning policies, it's unlikely platform moderators will be able to react to reports and remove deepfakes fast enough.

In Australia, lawyers at the NSW firm Ashurst have said existing copyright and defamation laws could fall short of protecting Australians against deepfakes.

And while attempts to develop laws have begun overseas, these are focused on political communication. For example, California has made it illegal to post or distribute digitally manipulated content of a candidate during an election—but has no protections for non-politicians or celebrities.

How to detect a deepfake

One of the best remedies against harmful deepfakes is for users to equip themselves with as many detection skills as they can.

Usually, the first sign of a deepfake is that something will feel "off." If so, look more closely at the subject's face and ask yourself:

- is the face too smooth, or are there unusual cheekbone shadows?
- do the eyelid and mouth movements seem disjointed, forced or otherwise unnatural?
- does the hair look fake? Current deepfake technology struggles to maintain the original look of hair (especially facial hair).

Context is also important:

- ask yourself what the figure is saying or doing. Are they disavowing vaccines, or performing in a porn clip? Anything that seems out of character or contrary to public knowledge will be relevant here
- search online for keywords about the video, or the person in it, as many suspicious deepfakes will have already been debunked
- try to judge the reliability of the source—does it seem genuine? If you're on a social media platform, is the poster's account verified?

A lot of the above is basic digital literacy and requires exercising good

judgment. Where common sense fails, there are some more in-depth ways to try to spot deepfakes. You can:

- search for keywords used in the video to see if there's a public transcript of what's being said—outlets often cover quotes by high-profile politicians and celebrities within 72 hours
- take a screenshot of the video playing and do a Google reverse image search. This can reveal whether an original version of the video exists, which you may then compare to the dubious one
- run any suspicious videos featuring a "colleague" or "relative" by that individual directly.

Finally, if you do manage to spot a deepfake, don't keep it to yourself. Always hit the report button.

Provided by The Conversation

Citation: Celebrity deepfakes are all over TikTok. Here's why they're becoming common, and how you can spot them (2022, July 19) retrieved 23 April 2024 from https://techxplore.com/news/2022-07-celebrity-deepfakes-tiktok-theyre-common.html