Deepfakes in warfare: New concerns emerge from their use around the Russian invasion of Ukraine

October 29 2023, by John Joseph Twomey, Conor Linehan and Gillian Murphy

Credit: Markus Winkler from Pexels

Visitors to a Ukrainian news website on the evening of February 18

2022, were greeted by a familiar sight, a video of their president giving a speech. While the resemblance was there, the face seemed slightly out of sync with the head of the Ukrainian president.

In the video, Volodymyr Zelensky announced that the war was over, a fact that the majority of Ukrainian people knew was false. It was a deepfake video. While this was happening online, the ticker at the bottom of the screen on the channel's live television feed read the same message. It claimed—again, falsely—that Ukraine was surrendering.

Our team at the Lero research center in University College Cork has just published a first-of-its-kind study examining the ways in which deepfake videos were presented and discussed on Twitter during the early months of the Russian invasion of Ukraine.

Deepfake technology is a recent technological development that essentially allows people to create videos of events that never happened. It seems particularly well suited to the spreading of disinformation, misinformation and "fake news" on social media platforms and elsewhere online. Deepfakes are also very suited to being used in cyberwarfare.

The videos are manipulated using AI technology and usually involve mixing real and fake content. This makes them appear more realistic and convincing than videos that have been generated entirely using AI. For example, deepfake technology can take a real video and swap the faces of two people in that video or change the lip movements so that a person appears to say something different to what they originally said.

Commentators and academics have pointed out that fake videos can be made much more easily and quickly using deepfake technology than with previous methods.

Loss of trust

Our study found many instances where the presence of deepfakes caused doubt or confusion. Strikingly, our data demonstrated instances where people accused real videos of being fake. We also found evidence of people losing faith in all videos from the conflict, with some people endorsing theories that world leaders were dead and had been replaced by deepfakes.

Of the many examples of deepfakes used online during the Russo-Ukrainian war, the example of Zelensky claiming that the war was over was perhaps the most frightening. This is because it highlighted how deepfakes could be used, along with hacked media services, to spread messages that were counterfactual. The net result of this incident was that false information was distributed from a trustworthy source. Similar deepfake videos of Russian president Vladimir Putin surrendering also emerged during the war.

Our research is the first academic study into the effect of deepfakes in the war. In order to uncover how deepfakes were used in the early days of the Russian invasion of Ukraine, we first created a timeline of the various deepfakes that were disseminated early on in the invasion. While we could not capture every deepfake that emerged, we tried to find the most notable examples and those with most impact. We then analyzed how those videos were being discussed on Twitter (now called X).

Humor, confusion and skepticism

Many of the deepfakes made during the conflict were humorous in nature, some involved inserting Putin into films such as Downfall (Der Untergang), or Charlie Chaplin's 1940 film The Great Dictator. There were also deepfake (and CGI) videos made by the Ukrainian government

to educate people on the conflict.

Interestingly, demonstrating the ability of Ukraine to create false videos, even though they were educational, might have been counterproductive. They created distrust and suspicion among viewers towards real media.

Much of the discussion of deepfakes online involved healthy skepticism, such as advice for fact checking and detecting deepfakes. However, we also found numerous examples of videos that people falsely claimed were deepfakes. There were two categories of such videos. First, many videos turned out to be low-tech fakes, such as those with false subtitles or videos of events from other wars presented as evidence of events in Ukraine.

Second, we found many instances of real videos of events in Ukraine, which commentators falsely accused of having been deepfaked. Losing trust in real media is a serious consequence of deepfakes, and only fuels the creation of deepfake-centered conspiracy theories.

Learning the lessons

What lessons can the average social media user take away from this research? The prevalence of deepfake videos online has increased over the last five years and the technological detection of deepfake videos is not currently accurate enough to be a solution in itself. The important thing is to encourage good media literacy in individuals in balancing healthy and unhealthy skepticism.

There is the need to view highly inflammatory media with skepticism and to wait for such news stories to be verified by multiple trustworthy sources. On the other side of the coin, people should be careful not to falsely accuse videos of being a deepfake.

It is important not to lose trust in every single piece of media we encounter, especially while deepfakes are not particularly prevalent online. One thing is for certain, the question of deepfake misinformation will be on everyone's mind as global conflicts develop throughout this decade and beyond.

Provided by The Conversation