Shedding light on deepfake detection

July 15 2024, by David Bradley

Credit: AI-generated image

The emergence of so-called deepfake technology, which commonly involves the generation of fake images, video, and sound that seem so authentic as to confuse even expert viewers and listeners, is at the point where it can influence important aspects of our lives, such as politics, finance, and beyond.

This new era of deception sees sophisticated image and video forgeries

making the headlines. Often these deepfakes are identified quickly, but sometimes the damage may well already be done once the deepfakes are called out, especially given the rate at which videos and other digital media can go viral on social media.

Digital manipulations that alter or completely synthesize faces, have become alarmingly convincing, contributing to fake news and eroding public trust in digital media. Research published in the *International Journal of Autonomous and Adaptive Communications Systems* describes a new approach that spots illumination inconsistencies within images and so can identify whether a video or photograph is a deepfake.

According to Fei Gu, Yunshu Dai, Jianwei Fei, and Xianyi Chen of Nanjing University of Information Science and Technology in Nanjing, China, deepfakes can be classified into four main types: identity swaps, expression swaps, attribute manipulations, and entire face synthesis. Each poses different threats and risks.

Identity swaps, where one person's face is replaced with another's, and expression swaps, which transfer facial expressions from one individual to another, are particularly worrying. Deepfakes can cause serious harm to the reputation and perception of the individuals or groups that are being deepfaked.

The usual approach to deepfake detection is to take a binary classification approach. However, the conventional approach can fail if the video or images are highly compressed or of poor quality. Compression and quality can obfuscate facial features and reduce the trust in deepfake detection.

Even expert deepfakers can falter when it comes to getting the lighting matched perfectly between altered and unaltered regions in an image or video. It is this issue that Gu and colleagues have focused on in their

detection method, which uses a [neural network](#) to spot illumination discrepancies.

Citation: Shedding light on deepfake detection (2024, July 15) retrieved 16 July 2024 from https://techxplore.com/news/2024-07-deepfake.html