

AI can now create fake porn, making revenge porn even more complicated

March 1 2018, by Nicola Henry, Anastasia Powell And Asher Flynn



Credit: CC0 Public Domain

In January this year, a new app was released that gives users the ability to swap out faces in a video with a different face obtained from another photo or video – similar to Snapchat's "face swap" feature. It's an

everyday version of the kind of high-tech computer-generated imagery (CGI) we see in the movies.

You might recognise it from the [cameo of a young Princess Leia](#) in the 2016 Star Wars film *Rogue One*, which used the body of another actor and footage from the first Star Wars film created 39 years earlier.

A Reddit user created another version of young [#PrincessLeia](#) from [#RogueOne](#) in [#FakeApp](#), which looks more realistic. This tech is crazy! <https://t.co/FftqcepVYp>
[pic.twitter.com/fUOad5k9ft](https://t.co/FftqcepVYp)

— 80Level (@EightyLevel) [January 30, 2018](#)

Now, anyone with a high-powered computer, a graphics processing unit (GPU) and time on their hands can create realistic fake videos – known as "deepfakes" – using artificial intelligence (AI).

Sounds fun, right?

The problem is that these same tools are accessible to those who seek to create non-consensual pornography of friends, work colleagues, classmates, ex-partners and complete strangers – and post it online.

The evolution of deepfakes

In December 2017, Motherboard broke [the story](#) of a Reddit user known as "deep fakes", who used AI to swap the faces of actors in pornographic videos with the faces of well-known celebrities. Another Reddit user then created the desktop application called FakeApp.

It allows anyone – even those without technical skills – to create their own fake videos using [Google's TensorFlow](#) open source machine learning framework.

The technology uses an AI method known as "[deep learning](#)", which involves feeding a computer data that the computer then uses to make decisions. In the case of fake porn, the computer will assess which facial images of a person will be most convincing as a face swap in a pornographic video.

Known as "morph" porn, or "parasite porn", fake sex videos or photographs are not a new phenomenon. But what makes deepfakes a new and concerning problem is that AI-generated pornography looks significantly more convincing and real.

Another form of image-based sexual abuse

Creating, distributing or threatening to distribute fake pornography without the consent of the person whose face appears in the video is a form of "image-based sexual abuse" ([IBSA](#)). Also known as "non-consensual pornography" or "revenge porn", it is an invasion of privacy and a violation of the right to dignity, sexual autonomy and freedom of expression.

[In one case](#) of morph porn, an Australian woman's photos were stolen from her social media accounts, superimposed onto pornographic images and then posted on multiple websites. She described the experience as causing her [to feel](#) "physically sick, disgusted, angry, degraded, dehumanised."

Yet responses to this kind of sexual abuse remain inconsistent. Regulation is lacking in Australia, and elsewhere.

Recourse under Australian criminal law

South Australia, NSW, Victoria and the ACT have specific criminal offences for image-based sexual abuse with penalties of up to four years imprisonment. South Australia, NSW and the ACT explicitly define an "intimate" or "invasive" image as including images that have been altered or manipulated.

Jurisdictions without specific criminal offences could rely on more general criminal laws. For example, the federal telecommunications offence of "using a carriage service to menace, harass or cause offence", or state and territory offences such as unlawful filming, indecency, stalking, voyeurism or blackmail.

But it is unclear whether such laws would apply to instances of "fake porn", meaning that currently, the criminal law provides inconsistent protection for image-based [sexual abuse victims](#) across Australia.

Recourse under Australian civil law

Victims have little recourse under copyright law unless they can prove they are the owner of the image. It is unclear whether that means the owner of the face image or the owner of the original video. They may have better luck under defamation law. Here the plaintiff must prove that the defendant published false and disparaging material that identifies them.

Pursuing civil litigation, however, is time-consuming and costly. It will do little to stop the spread of non-consensual nude or sexual images on the internet. Also, Australian civil and criminal laws will be ineffective if the perpetrator is located overseas, or if the perpetrator is an anonymous content publisher.

Addressing the gap in legislation

The Australian Parliament is currently debating the [Enhancing Online Safety \(Non-Consensual Sharing of Intimate Images\) Bill 2017](#). This bill, which is yet to become law, seeks to give the Office of the eSafety Commissioner the power to administer a complaints system and impose formal warnings, removal notices or civil penalties on those posting or hosting non-consensual intimate images.

Civil penalties are up to A\$105,000 for "end-users" (the individuals posting the images) or A\$525,000 for a social media, internet service or hosting service provider.

Importantly, the proposed legislation covers images which have been altered, and so could apply to instances of deepfakes or other kinds of fake porn.

Prevention and response beyond the law

While clear and consistent laws are crucial, online platforms also play an important role in preventing and responding to fake porn. Platforms such as Reddit, Twitter and PornHub have already [banned deepfakes](#).

However, at the time of writing, the clips [continue to be available](#) on some of these sites, as well as being posted and hosted on other websites.

Read more: Facebook wants your nude photos to prevent 'revenge porn' – here's why you should be sceptical

A key challenge is that it is difficult for online platforms to distinguish between what is fake and what is real, unless victims themselves discover their images are online and contact the site to request those images be removed.

Yet victims may only become aware of the fake porn when they start receiving harassing communications, sexual requests, or are [otherwise alerted to the images](#). By then, the harm is often already done. Technical solutions, such as better [automated detection of altered imagery](#), may offer a way forward.

To adequately address the issue of fake porn, it is going to take a combination of better laws, cooperation from online platforms, as well as technical solutions. Like other forms of image-based sexual abuse, support services as well as prevention education are also important.

This article was originally published on [The Conversation](#). Read the [original article](#).

Provided by The Conversation

Citation: AI can now create fake porn, making revenge porn even more complicated (2018, March 1) retrieved 23 April 2024 from <https://phys.org/news/2018-03-ai-fake-porn-revenge-complicated.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.