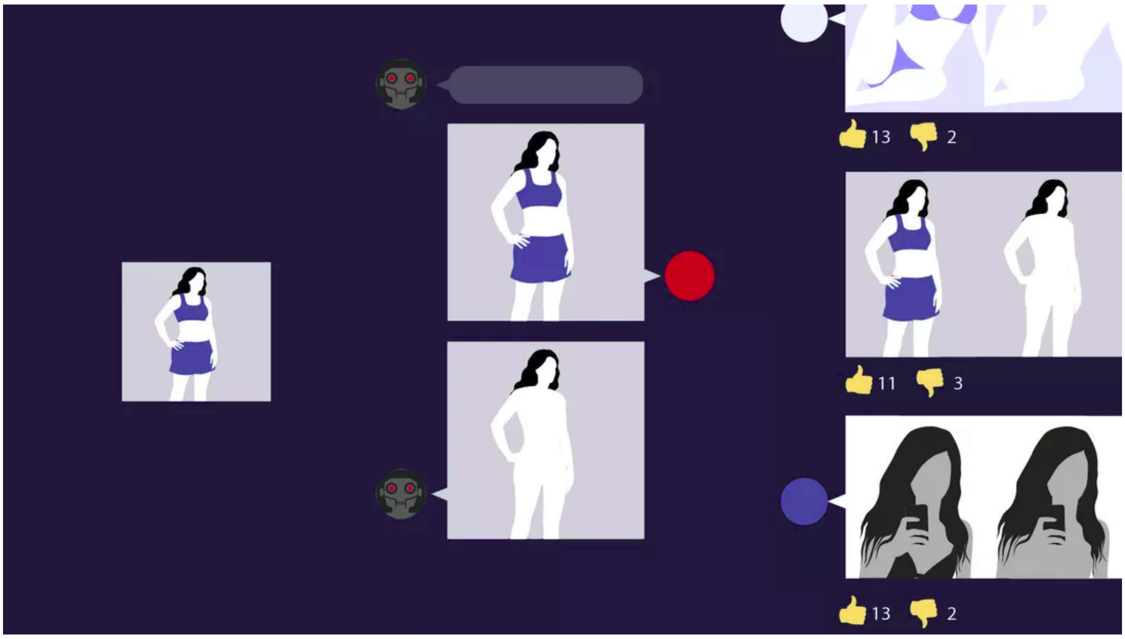
CDC touts masks indoors again Simone Biles pulls out of team finals Activision Blizzard employees walkout Disney's cryptic The S

cnet | Tech

FEATURED MOBILE COMPUTING HOME ENTERTAINMENT SERVICES



Deepfake bot on Telegram is viola women by forging nudes from reg

Free, easy and requiring just a single still photo, the deepfake produced more than 100,000 fake pornographic images publi for anyone to see.

Joan E. Solsman Oct. 22, 2020 8:06 a.m. PT

LISTEN - 09:19

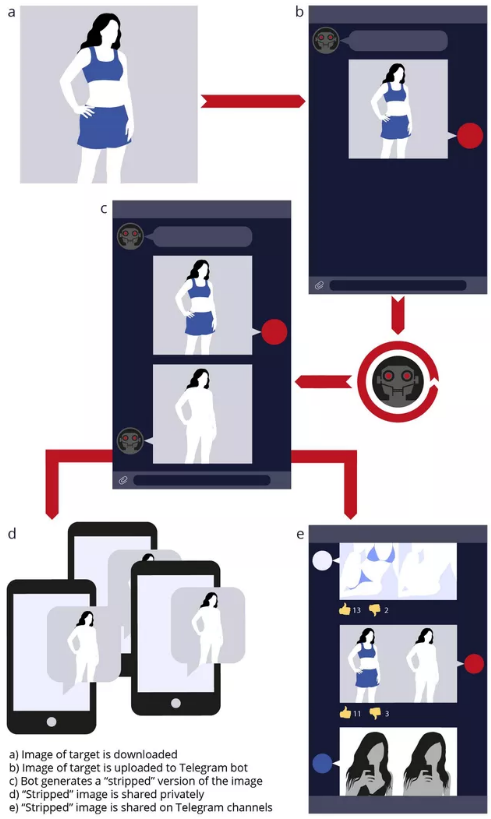
A free, easy-to-use deepfake bot found on the Telegram messenger app wearing simple T-shirts and shorts. Some were visibly underage. All are women.

Deepfake porn isn't new. Deepfake technology -- artificial intelligence that makes sophisticated media forgeries -- has been used early and often to fabricate pornography. But this Telegram bot takes the ease and access of this technology to a new level.

"The innovation here is not necessarily the AI in any form," said Giorgio Patrini, CEO of deepfake-research company Sensity and coauthor of the report. "It's just the fact that it can reach a lot of people, and very easily."

Computer manipulation of media has existed for decades, and sexual imagery has been weaponized online for as long as the internet could host photos. Whether it's nude photos posted without consent or crudely doctored forgeries, sexual images have been weaponized to extort, threaten, humiliate and harass victims.

But only in the last few years has deepfake tech intensified the threat of manipulated sexual media, posing frightening implications for what may come.



a) Image of target is downloaded
b) Image of target is uploaded to Telegram bot
c) Bot generates a "stripped" version of the image
d) "Stripped" image is shared privately
e) "Stripped" image is shared on Telegram channels

"The deepfake phenomenon is even more alarming abuse much more difficult."

With this Telegam bot, any woman who's ever posted a selfie of herself from the waist up could be a potential victim. Even women out walking could be victimized if surreptitiously snapped by the wrong stranger.

And in one of the most disturbing forms of abuse with this bot, photographs of children have been uploaded to the bot's AI, automatically manipulated to sexualize the child and then shared publicly.

Neither Sensity's report nor this article are disclosing the name of the bot, to avoid amplifying it. CNET viewed galleries of images with the bot's watermark posted online and interacted with the bot itself, stopping short of uploading any photos for it to manipulate.

Telegram's tenacious commitment to free speech and privacy may make bots like this challenging to stamp out. Telegram has been criticized for hosting terrorist propaganda and coordination, facilitating piracy and copyright infringement, and harboring varieties of predatory pornography. But the service has also taken actions to remove abuse, such as kicking off groups for violent extremists like neo-Nazis and ISIS.

"There's clearly value in encrypted platforms" like Telegram, said Sam Gregory, a program director with human-rights video organization Witness, who also advised Sensity on its report. "That doesn't mean they shouldn't be thinking about the use of their platform for things that have nothing to do with free expression."

Sensity reached out to Telegram multiple times over the last six months about its findings. Telegram didn't respond to Sensity's outreach, nor did Telegram respond to CNET's messages seeking comment.

Deepfake nudes

Deepfake technology is like a high-speed Photoshop conveyor belt on steroids. Using a kind of artificial intelligence known as neural networks, deepfake tech can generate media forgeries that make people appear to be doing or saying things they never did. The term deepfake is used most often with videos, but deepfakes can refer to any so-called "synthetic" media produced by deep machine learning, including pornographic still photos.

If this bot on Telegram sounds disturbingly familiar, a similar technology

> "The deepfake phenomenon is even more alarming because it doesn't look Photoshopped. It's much more easy for somebody without the technical knowledge to make one."
>
> *Mary Anne Franks, president, Cyber Civil Rights Initiative*

The bot is also designed to make it easy for abusers to share the manipulated images by posting them in chats and other online forms.

These nonconsensual sexual images have "been put out there to be found," Patrini said. "They're completely open, without any login, without any passwords, on the internet. Those are actually exposed completely."

Sensity found 104,852 images of women that were victimized by the bot and then shared publicly, as of the end of July. While each image may not be of a unique individual, Patrini said instances of the same woman being victimized, or the same photo being manipulated repeatedly, were rare.

The 100,000-plus total number of images is limited to manipulated photos that were publicly posted and that Sensity was able to track down. Sensity doesn't know the scope of material that is not shared, Patrini added. "But definitely we are talking about some multiplier of that 100,000."

The bot's promotional website suggests that as many as 700,000 images have been manipulated by the bot.

And the bot is growing in popularity. A year ago, about 1,000 images manipulated by the bot were posted in channels in a month. In July, that number had swelled to at least 24,168 images, according to Sensity.

And while deepfake pornography has long fixated on victimizing actresses, models and other celebrity women, 70% of this bot's targets were private individuals, according to a self-reported survey of the bot's users in Sensity's report.

About 100,000 people are members of channels linked to the bot, Sensity found. These members overwhelmingly come from Russia and former-USSR countries, about 70% of those surveyed. Telegram is used seemingly up to five a day. But "paid premium" features include sending multiple pics, skipping the line of free users and removing watermarks from the pornographic images they get in return.

Watch this: We're not ready for the deepfake revolution

▶ 7:07

But the bot's business strategy is also ambitious, inspired by strategies from gaming and classic promotional tropes.

In a gamefied turn, the premium features are paid for with virtual "coins." These coins can be purchased cheap with real currencies, and coin lotteries appear to distribute some for free. They can also be earned, as rewards.

One of the rewarded behaviors is recruiting new users. And because the app says that its virtual coins can be paid back in rubles, it effectively creates a system that pays abusers money in a government-issued currency for bringing in new abusers.

Fortunately, the payouts are presumably meager: The value of the bot's coins are cheap, roughly five cents each.

The bot's designer has also adopted classic promotional tactics. You can get a deeper discount on coins with the more of them you buy. The bot pitches new users with a one-time "beginner rate" special on coins.

'This terrible technology'

The bot also underscores how a fixation on electoral deepfakes misses wider damage caused by pornographic ones, which are much more common and already devastating victims.

"So much of the focus on deepfakes is in an electoral context," Gregory said. A preoccupation with "the perfect deepfake" of a political candidate or world leader is the kind of disinformation that tends to stoke

Gregory added. "That doesn't mean we shouldn't be extremely vigilant."

But even vigilance is unlikely to result in justice for victims, Franks said, who pointed to a historical failure of our legal systems to address weaponized sexual imagery years ago.

"We wouldn't be in this position, where we have technology capable of releasing this kind of malicious content at such a scale ... if we paid attention before. We need to do better now," she said. "If there's any good to come of this terrible technology, it's that people may take this more seriously."

First published on Oct. 20, 2020 at 7:00 a.m. PT.

Digital Media | Facial Recognition | Privacy | Cybersecurity