Drag and drop: Deepfakes create a new kind of crime

The 'Nth room' case shocked the public earlier this year when details of the inhumane crime was revealed through media reports and investigation results. Although the idea of digital sex crimes wasn't new, the horrifying atrocities of what had been going on inside the anonymous chat rooms brought the issue to the surface like never before. Among the crimes are some lesser-known forms of digital sex crimes that had gone unnoticed by the public over the years but has come to light with the case. Through this two-part series, the Korea JoongAng Daily looks into the two types of crime that had not been properly addressed before: 'friend fouling' and deepfakes.

As technology advances, it's becoming much easier for people to collage faces onto different photographs with just a few clicks and drags of a mouse. You don't have to be a Photoshop expert or a tech geek anymore to throw the faces of your friends onto different photographs for a laugh.

But such technological advancement has also cultivated the growth of deepfakes, or digitally manipulated videos, that often become the source of digital sex crimes.

Digital sex crimes are a relatively new type of crime, but even newer are deepfakes, specifically those that manipulate the facial features of a video and replaces it with the face of another person. The technology has existed for years, but only since the recent revelation of digital sex crimes in the darkest corners of the "Nth room" chat rooms has the Korean public become aware of the proliferation of this particular phenomenon, especially when it comes to female celebrities.

The concept of digital sex crimes hit home for Koreans after the Nth room crimes made headlines in March, and in late April a report from the Mapo Police Precinct in western Seoul revealed that a gang of criminals had been producing deepfake pornography with celebrities' faces and distributing them through an illegal website since 2018. Over a hundred celebrities are said to be featured on the website, which had around 3,000 videos. The police are tracking down the perpetrators, but it may be hard as the website is based on a South American server.

What are deepfakes?

Contrary to "friend fouling," which usually happens between ordinary — especially younger — people, deepfakes are currently posing a major threat to female celebrities and may possibly become a threat to everyone else as the technology advances in the future due to the requirements of this particular technology.

The term deepfake is a portmanteau of "deep-learning" for artificial intelligence and "fake." Engineers and programmers birthed primitive forms of deepfake in the late 1990s in which they could change the facial features of a person on video, but the idea as we know it now was conceived in 2017 by a user nicknamed "deepfakes" who began actively uploading videos on Reddit — both pornographic and non-pornographic.

President Barack Obama's face merged onto Jordan Peele's face warned of the dangers of deepfake technology. [SCREEN CAPTURE]
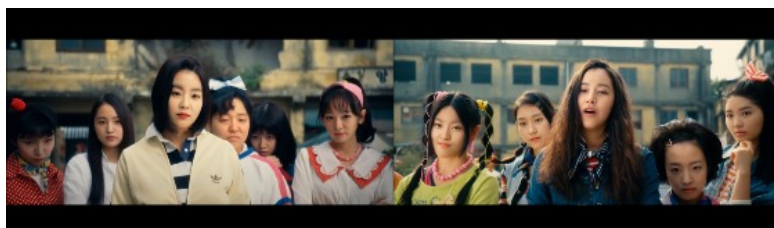
Deepfake videos take the facial features of a person and changes them to those of a different person. It takes three steps: extract, learn and merge. In the first stage of the machine learning process, the program detects the facial features and extracts them from the two faces that will be swapped. Then it learns the details of the two faces, which can take many hours depending on how much material there is. After it has learned which facial expressions correspond to each other and need to be applied, the merging takes place. Right now, deepfakes only change the eyes, nose and mouth, so using two people who have similar features brings out the most natural results.

Right now, it's actually not difficult to make deepfake videos, thanks to free apps and programs that allow even ordinary people to make them. Cost-free source codes and machine-learning algorithms are abundant online, and the only things you need to make yourself a video are time and materials. With enough photographs and videos for the program to learn from, the facial features and voice can theoretically be replaced by a different person to the point of perfection. A three-minute video containing a person's face from all angles, with words spoken throughout, can be high enough quality to fool people.

So here, the reason why celebrities' faces are targeted becomes clear: because politicians and celebrities are the only groups of people who have enough photographs and videos of them on the internet that people can access to create a high-quality deepfake video; and of those two groups of people, celebrities generally attract other people's attention.

Why are they dangerous?

There's not much harm to deepfake videos, as long as the technology is used only for fun.

Deepfake videos by Hallyubyte feature girl groups Red Velvet, left, and Twice, right. [SCREEN CAPTURE]

For instance, the YouTube channel HallyuBytes puts the faces of K-pop stars onto the faces of actors in popular movies and dramas, which many fans get a laugh from. In one clip, the members of girl groups Red Velvet and Twice look as though they are in the famed gang-fight scene from the movie "Sunny" (2011).

When BuzzFeed uploaded a clip of Barack Obama's face being acted out by comedian and director Jordan Peele, the comments on the video were about how funny it was and how impressive the technology seemed to be. The clip was actually made to warn people about the dangers of deepfakes a
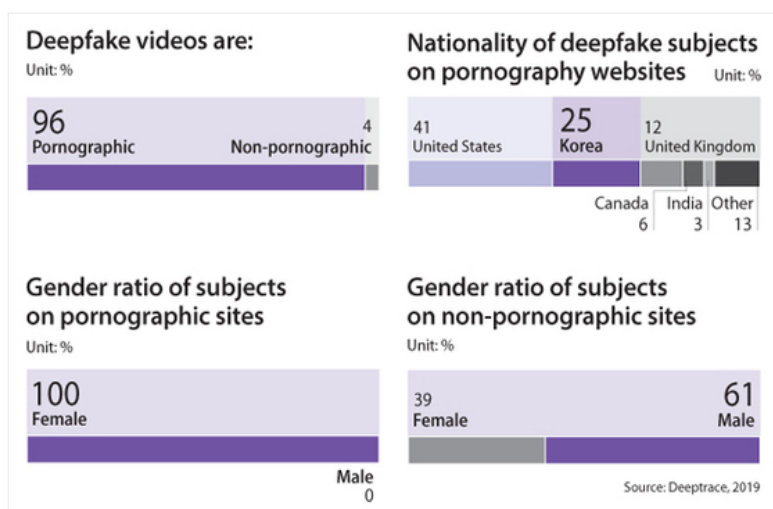
nd fake news.

The problem is, however, that most deepfakes are far from harmless.

According to a 2019 report by Deeptrace, an Amsterdam-based company dedicated to detecting online synthetic media, 96 percent of online deepfake videos were pornographic in nature, while only 4 percent were content depicting anything other than porn, including politics and humor — in September of last year, when the report was published. Deeptrace counted 14,678 videos on the internet, "an almost 100 percent increase based on our previous measurement (7,964) taken in December 2018" according to the report. But since a vast majority of illegal contents are shared through the dark web, it's hard to know at what pace it is spreading.

The most shocking finding from the report was that among all the deepfake pornography websites' posts, 25 percent featured Korean K-pop singers, while 41 percent of videos featured Americans celebrities.

In fact, of the 10 most frequently targeted individuals on deepfake pornography, Korean musicians were listed at No. 2, 3 and 10, though their names were not mentioned. Most deepfake pornography clips were uploaded on websites solely dedicated to making deepfake porn, as opposed to "mainstream" porn sites that offer pornography in general.
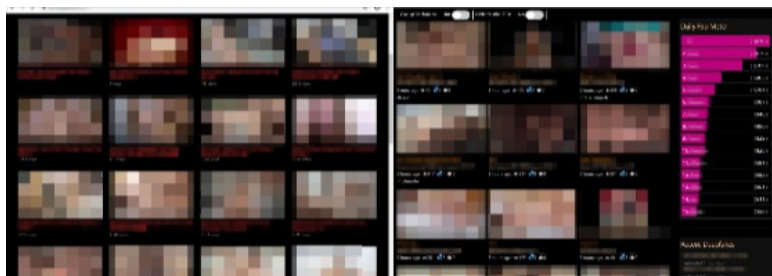

Deepfake graphic

"Deepfake pornography is by far the most prevalent kind of deepfake currently being created and circulated. The damaging effects of deepfake pornography have already affected a significant number of women, including celebrities and private individuals," the report said.

Is it worse in Korea?

While the damage caused by deepfakes are many — spanning from identity theft to the overall decrease in the level of people's trust in the media — Lee So-eun, a senior researcher at the Korea Press Foundation's Media Research Center, pointed out that defamation and insult are the two most damaging aspects of deepfake videos, especially for the pornographic deepfakes made using female celebrities' faces, regardless of whether the viewer believes the situation to be actually true or not.

"Deepfakes use an existing person and then synthesizes them into a situation or body for the purpose of 'leisure,'" said Lee. "Deepfakes are problemati
c not only because they twist the truth and steal other people's identity, but because they amplify a certain emotional effect. With 'friend fouling' or de
epfakes, people objectify a person that they have admired and put them in situations that would otherwise be impossible — against their will — and u
se those videos for their pleasure, which can be quite problematic morally and ethically."

What aggravates the situation is that since this is a relatively new type of crime and Korean society already has a low level of awareness of digital sex cr
imes, people aren't properly aware of the devastating effects that this content can cause.

Rep. Jeong Jeom-sig of the opposition Unified Future Party said that "these videos that were only made for self-satisfaction really shouldn't go as far as
being punished," and National Court Administration's Vice Minister Kim In-kyeom said that "some people could make them thinking it's art."

In fact, some online have likened deepfake porn to "fan fic," or "fan fiction" as it is commonly called outside of Korea, based on the fact that the subjec
ts in both are celebrities. Fan fic refers to the fictional stories fans write using the members of their favorite K-pop groups and some have claimed that
deepfake pornography isn't much different. But Lee Taek-gwang, culture critic and professor of global communication at Kyung Hee University, says th
at the two are nowhere close and comparing the two is ignoring the problems the deepfakes present.

"Fan fic is a genre of popular culture, created by fans," said Lee. "But deepfakes are created with a clear goal to target a certain person and sell them. It
goes beyond the level of private individuals. And whether that goal be sexual or commercial or both, those who make such videos form an online sex i
ndustry and make profit by running their 'business.'"

What has changed?

A revision of the Act on Special Cases Concerning the Punishment, Etc. of Sex Crimes passed the National Assembly and goes into effect on June 25. A
ccording to the new law, those who produce or distribute videos that fabricate, manipulate, edit or copy a person's face, body or voice against their wil
l can be punished with up to five years in prison or charged a fine of up to 50 million won ($40,500). If the act was done for the purposes of earning
money, then the prison sentence can be extended up to seven years.

Until the amendment, deepfake pornography was not dealt with as a sex crime in the eyes of the law and those behind the producing of the content
were accused of either breaching the Act on Promotion of Information and Communications Network Utilization and Information Protection, Etc. for di
splaying or distributing "obscene content" or of libel. The amendment does not, however, punish anyone for owning or watching deepfake pornograp
hy. And if the video was produced overseas, law enforcement officers have little they can do.

"The amendment only punishes the act of producing and spreading deepfakes," the Korean Bar Association said in a statement.

**"But there are limits to punishing crimes after they occur to eradicate content that sexually exploits victims through online platfor
ms such as Telegram. There needs to be governmental restrictions on internet services both in and outside of Korea and new reg
ulations must be put in place to include the diverse and complicated ways through which digital sex crimes take place."**

According to an entertainment agency home to some celebrities frequently targeted for deepfake pornography, one of the biggest difficulties is that th
e videos are shared in the deepest corners of the internet and cannot be tracked down easily.

Lee Soo-jung, a forensic psychologist and a professor at Kyonggi University, says that the amendment is just the beginning.

"The amendment is our society's first official agreement that digital sex crimes need to be punished," said Lee. "But right now, there aren't guidelines o
n how to actually find the criminals. How are the investigations going to catch people if [the content is] based on foreign servers? How are the police g
oing to gather evidence in the first place? They need to go undercover, but not all undercover cases are legally protected in Korea. We need to set up
proper guidelines on how to investigate the cases, before talking about what punishments to give them."

BY YOON SO-YEON   [yoon.soyeon@joongang.co.kr]