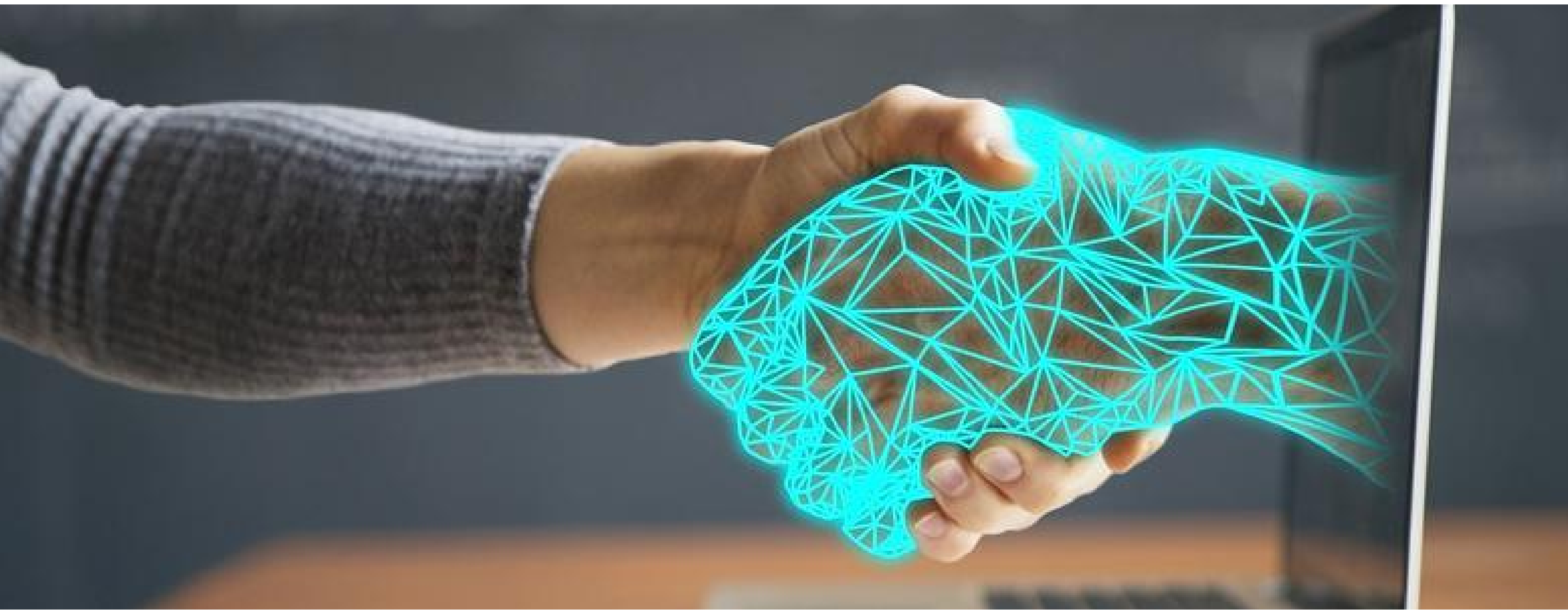
JUNE 2024

Explosion in Deepfake Videos

BROUGHT TO YOU BY SIMS IT TEAM

Shweta Mehta

WHAT ARE DEEPFAKES?



Deepfakes are synthetic media where a person in an existing image, video, or audio is replaced with someone else's likeness or voice. Created using advanced AI techniques, especially deep learning and neural networks, these forgeries can be extremely realistic and hard to distinguish, it fabricates events or statements that never occurred.

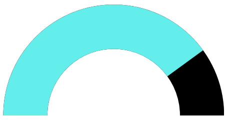
Initially, deepfake technology was developed for creative and educational purposes, but its capabilities have since expanded into various sectors, raising ethical concerns. Over the years, different types of it have emerged, such as textual deepfakes, deepfake audio, live deepfakes, etc.

THE DEEPFAKE DANGER

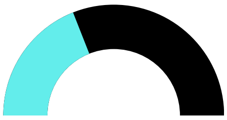
- Privacy Issues: Deepfakes can violate personal privacy, leading to harassment and reputation damage.
- Misinformation: They can spread false information, manipulating public opinion and causing societal harm.
- Legal Frameworks: Some regions have laws against malicious use of deepfakes, such as non-consensual deepfake pornography and political manipulation.
- Financial Risk: facilitating fraud, scams, and extortion.

10X
increase in the number of deepfakes detected globally from 2022 to 2024

DID YOU KNOW?

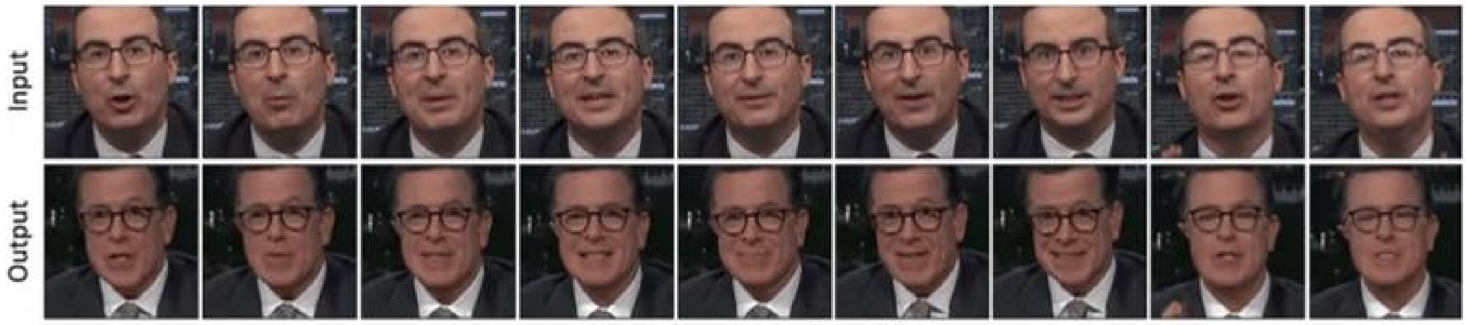


75%
of Indians present online have seen some form of deep fake content over the last 12 months.



38%
of the respondents have encountered a deep fake scam.

https://m.economictimes.com/tech/technology/75-indians-have-viewed-some-deepfake-content-in-last-12-months-says-mcafee-survey/articleshow/109599811.cms



John Oliver to Stephen Colbert

In 2019, DeepTrace found that 96% of deepfake videos online consisted of pornographic content. The main victims of deepfake pornography are predominantly females in which celebrities from the entertainment industry make up 99% of these videos, while individuals from news and media make up 1%.

EXAMPLE EVENTS

> A deepfake video featured President Nixon's resignation speech and his draft speech for a failed moon landing. Created for artistic purposes, it convinced many people it was real.

> On April 17, 2018, a deepfake was posted on YouTube, depicting Barack Obama cursing Donald Trump . The intent was to portray the dangerous consequences & power of deepfakes.

> Prior to the 2024 United States presidential election, phone calls imitating the voice of the incumbent Joe Biden were made to dissuade people from voting for him.

> In May 2019, two artists created a deepfake video of Facebook founder Mark Zuckerberg talking about harvesting and controlling data from billions of people.

IT CLUB

TECHNOLOGY BEHIND DEEPFAKES

STEP 01 DATA COLLECTION

Large datasets of videos and audio recordings of the target person are collected. These datasets are crucial for training the AI model to understand the person's facial expressions, movements, and voice patterns.

STEP 02 DEEP NEURAL NETWORKS (DNNs)

Deepfake creation starts with (DNNs), a type of artificial intelligence. Here's how it works:

Learning Facial Features: The DNN studies many pictures and videos of a person from different angles. It learns how their face looks and moves.

Applying to Another Person: Once it understands the person's facial features, it replaces the person's face onto someone else in a different video, making it look like the original person is doing or saying something they didn't actually do.

STEP 03 GENERATIVE ADVERSARIAL NETWORKS (GANs)

GANs improve the deepfake to make it look even more realistic:

GANs are two part system:
Generator: This creates fake images
Discriminator: Checks if the images are real or fake.

This back-and-forth process continues until the fake images are so good that it's hard to tell they are fake.

STEP 04 SYNTHESIS & REFINEMENT

The AI synthesizes the learned features to produce the final deepfake. Manual adjustments may be made to ensure the deepfake looks as realistic as possible, fixing any imperfections or anomalies.

The goal is to create a deepfake that is difficult to distinguish from genuine content, often requiring careful refinement and attention to detail.

Tools have made deepfake creation accessible to a wider audience, including those with limited technical skills.
Eg: DeepFace Lab, Faceswap, Reface, Zoa

IT CLUB

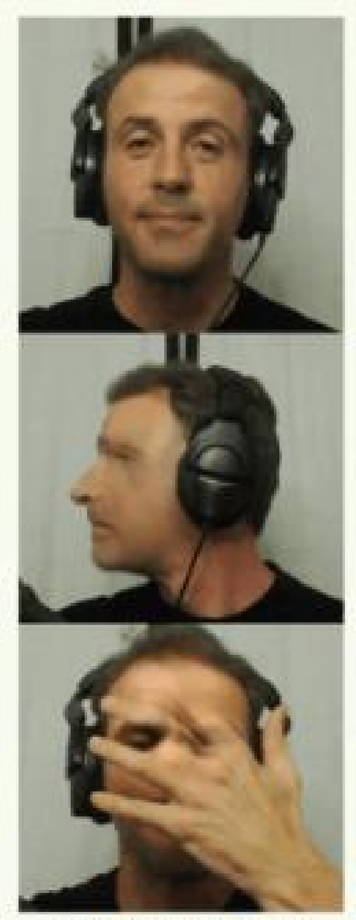


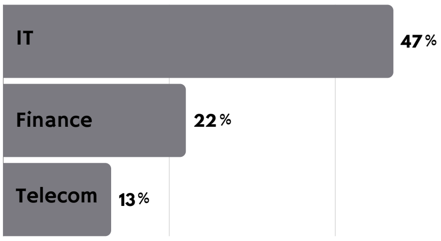
Image source: Metaphysics

How to spot Deepfakes?

- Bad lip-syncing
- Blur around the face
- Uneven facial skin tone
- Unnatural facial expression
- Odd positioning of facial features
- Unusual blinking or motion
- Inconsistent lighting & audio quality
- Misaligned or blurry face when obstructed

43% of people can't tell the difference between a real video and a deepfake

Combating Deepfakes



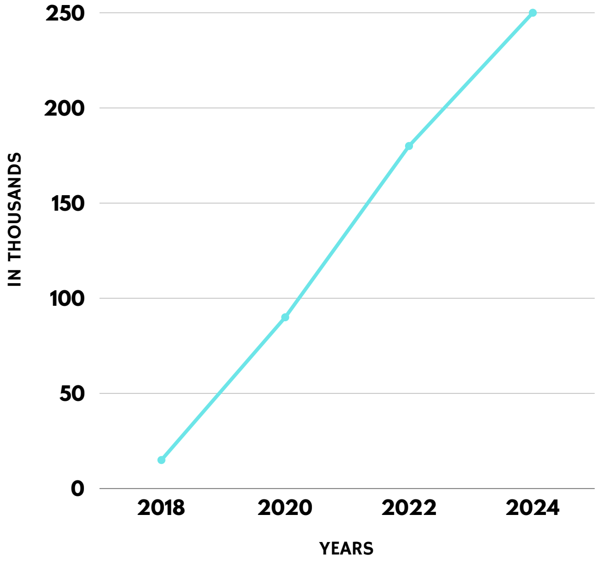
TARGETED INDUSTRIES

(https://images.app.goo.gl/cXhRPcdJr5CEcVKZ9)

- Facebook and Microsoft have launched initiatives to detect and remove deepfake content.
- Even LinkedIn is using AI-powered “deep-learning-based model” to determine whether LinkedIn profile photos are AI-generated deepfake facial images.
- Using a layered approach when planning your internet security strategy.
- Deploying basic and robust security protocols to stop fraudsters from using deepfakes to harm the business.
- Governments are creating laws and policies to address the misuse of deepfakes, such as banning non-consensual deepfake pornography and political deepfakes.
- Using blockchain, you can create digital fingerprints for your videos to establish your video authenticity, ensuring content can be traced back to its original source.

What is next for Deepfakes?

By 2025, 8 out of 10 people will likely encounter a deepfake.



SURGE IN DEEPFAKE CONTENT

https://images.app.goo.gl/Ld47eBWsfJki5IwG8

In the coming years, deepfakes will become more common, penetrating areas that we can't even guess now. As per a report of recent deepfake attacks, there have been instances where fraudsters have started using deepfake technology to conduct interviews for remote jobs. Even an increase in the use of synthetic videos and images of other people has been observed to dupe authentication systems.

Continued improvements in AI and machine learning will make deepfakes even more realistic and harder to detect. Innovations may also lead to new applications in fields like virtual reality and telepresence. According to the VPNranks deepfake report, the prevalence of deepfakes, is expected to rise substantially by the end of 2024.

Since deepfake technology makes it extremely difficult to differentiate real pictures from fake ones, enterprises, organizations, and users must stay updated about the changing technologies like AI, Blockchain, etc., to protect themselves from deepfake attacks.

2X As per Sensity, a cybersecurity firm, deepfakes are doubling every 6 months.