# Tackling Deepfakes - A novel approach

In this digital era, it has increasingly become difficult to judge whether a piece of multimedia content is authentic or not. **A special case and recent example is deepfake videos. With recent advancements in artificial intelligence, we are nearing a future where humans would be unable to differentiate between a computer generated video and an authentic camera captured one.**

Deepfake is an artificial intelligence technique that can realistically generate fake videos. Prominent techniques work by replacing the audio clip of a person with a **forged one and then change the facial expression to match that audio, or by swapping the face altogether.** Earlier it would be considered that videos were untamperable and considered as hard evidence. Now that the video can be tampered with, it is hard to trust what we see.

**Deepfake videos have already been used in politics to spread misinformation and hate:**

One such case was of Belgian Prime Minister Sophie Wilmès. Her fake video was circulated by Belgian branch of Extinction Rebellion, which linked COVID-19 and deforestation together. The video got 100,00 views in 24 hours. Many who watched the video considered it to be genuine.

**In another case, deepfake videos were used in election campaigning:**

In case of the 2020 Delhi elections, candidate Manoj Tiwari's deepfake videos with messages translated from English to Haryanvi were used to target Haryanvi voters. The voiceover was provided by an actor and the lip-sync was done by the AI model trained on Manoj Tiwari's speeches.

The above examples show the potential of the use/misuse of deep fake videos. Since the software for making such videos is freely available, it can be easily exploited for various purposes.

**Some benign applications could be in cinema:**

Faceswap can be used by producers to add the face of a certain actor on a stunt double. One could potentially add own face in the movie too.

Deepfake can also be used in **therapeutic treatment of patients,** who have lost a loved one

This type of therapy can make use of deep fake videos created of the lost relative in order to treat the patient.

**Deepfakes used in robots and digital assistants** that can give a better experience when we are interacting with them as they can mimic facial expression to the most delicate detail.

**Interaction Request:**

**IIIT-Delhi (Indraprastha Institute of Information Technology Delhi), Delhi's leading technology research institute established in 2008 as a State University by the Delhi Government,** has introduced an algorithm which follows principles of anomaly detection **to help identify fake videos.** Since the amount of data that we have on fake videos is very less today, along with the challenge of deepfake algorithms always evolving to create better fake videos, **there is a need to learn facial expression patterns directly from the vast collection of original videos that is already available.**

Incase of fake videos, there are always some subtle differences between the input and its reconstruction. Their anomaly detection algorithm picks up this anomaly and determines whether the video is authentic or fake.

**However, its detection algorithm has provided promising results on fake and original video datasets that are available.** It performs particularly well on Faceswap GAN synthesised deep fake videos and gives satisfactory results on lip sync fake videos. It will be interesting to see how well the algorithms perform as deepfake techniques keep on evolving.

**Please do let us know if you would like to speak with the concerned person at IIIT-Delhi to know more details on how this algorithm works to detect fake deepfake videos.**