New technique to spot deepfakes

Prelims: *Current events of national and international importance | Science & Technology*

Why in news?

Researchers at the University of California, Los Angeles, have found a way to tell if a video is a deepfake based on the light in it.

- **New Technique** – A ***hybrid optical-electronic computer*** that uses the physics of light to detect deepfake videos with remarkable speed and accuracy.

  *A **hybrid optical-electronic computer** combines light-based (photonic) processing with traditional electricity-based (electronic) processing to achieve ultra-fast speeds, high bandwidth, and low energy consumption.*

- **Need for a detection system** – It was designed to address the massive energy, speed, and security limitations of traditional digital deepfake detectors.
- Because humans only catch deepfakes about 50% of the time, this light-powered technology provides a critical first line of defence for social media platforms struggling against a flood of AI-generated misinformation.

  ***Deepfakes** – It is synthetic media (images, video, or audio) created or altered using AI and deep learning techniques to make it look or sound like a real person is saying or doing something they never did.*

- **Working –** Unlike traditional digital processors that evaluate data frame-by-frame using silicon chips, this hybrid system offloads the heaviest computational work onto photons (light).
  - **Frame Selection** – A lightweight digital encoder selects 12 specific frames from a video and maps unique facial features.
  - **Light Modulation** – These features are converted into a phase pattern and projected onto a programmable spatial light modulator.
  - **Diffraction Analysis** – The light travels through a passive optical decoder. The machine splits the light wavefront toward sensor spots and analyses the differences in brightness caused by light diffraction.
  - Based on the difference in brightness between these spots, the system could say

whether a video was fake.
    - **Instant Classification** – If the brightness pattern matches known anomalies typical of generative AI, it instantly flags the video as synthetic.
- **Performance & Capabilities** –
    - 97.79% average accuracy while processing 15 videos simultaneously.
    - 99.86% sensitivity, meaning it correctly identifies almost all fake videos.
    - Only about 0.14% false negatives.
- **Effective against advanced AI videos** – When tested on videos generated using Google's VEO-3 model, which produces highly realistic synthetic footage, the system achieved 94.8% accuracy with minimal additional training.

Reference

[The Hindu | New technique uses light to spot deepfakes in seconds]

SHANKAR IAS PARLIAMENT
Information is Empowering