

Facebook trained AI to fool facial recognition systems, and it works on live video

Facebook remains embroiled in a multibillion-dollar judgement lawsuit over its facial recognition practices, but that hasn't stopped its artificial intelligence research division from developing technology to combat the very misdeeds of which the company is accused. According to *VentureBeat*, Facebook AI Research (FAIR) has developed a state-of-the-art "de-identification" system that works on video, including even live video. It works by altering key facial features of a video subject in real time using machine learning, to trick a facial recognition system into improperly identifying the subject.

This de-identification technology has existed in the past and there are entire companies, like Israeli AI and privacy firm D-ID, dedicated to providing it for still images. There's also a whole category of facial recognition fooling imagery you can wear yourself, called adversarial examples, that work by exploiting weaknesses in how computer vision software has been trained to identify certain characteristics. Take for instance this pair of sunglasses with an adversarial pattern printed onto it that can make a facial recognition system think you're actress Milla Jovovich.

AI software can trick facial recognition systems by altering key facial features

But that type of thwarting of facial recognition usually means altering a photograph or a still image captured from a security camera or some other source after the fact. Or in the case of adversarial examples, preemptively setting out to fool the system. Facebook's research supposedly does similar work in real time and on video footage, both pre-captured and live.

That's a first for the industry, FAIR claims, and good enough to combat sophisticated facial recognition systems. You can see an example of it in action in this YouTube video, which, because it's de-listed, cannot be embedded elsewhere.



Image: Facebook AI Research

"Face recognition can lead to loss of privacy and face replacement technology may be misused to create misleading videos," reads the paper explaining the company's approach, as cited by *VentureBeat*. "Recent world events concerning the advances in, and abuse of face recognition technology invoke the need to understand methods that successfully deal with de-identification. Our contribution is the only one suitable for video, including live video, and presents quality that far surpasses the literature methods."

Facebook apparently does not intend to make use of this technology in any of its commercial products, *VentureBeat* reports. But the research may influence future tools developed to protect individuals' privacy and, as the research highlights with "misleading videos," prevent someone's likeness from being used in video deepfakes.



Image: Facebook AI Research

The AI industry is currently working on ways to combat the spread of deepfakes and the increasingly sophisticated tools used to create them. This is one method, and both lawmakers and tech companies are trying to come up with other tools, like deepfake detection software, and regulatory frameworks for how to control the spread of fake videos, images, and audio.

The other concern FAIR's research addresses is facial recognition, which is also unregulated and causing concern among lawmakers, academics, and activists who fear it may violate human rights if it continues to be deployed without oversight by law enforcement, governments, and corporations.

<https://www.theverge.com/2019/10/25/20932879/facebook-ai-facial-recognition-live-video-de-identification-deepfakes>