WIE KI-SYSTEME MANIPULIEREN KÖNNEN

Schon lustig, diese gefälschten Fotos. Aber auch gruselig, wenn man seinen eigenen Augen nicht mehr trauen kann, oder?

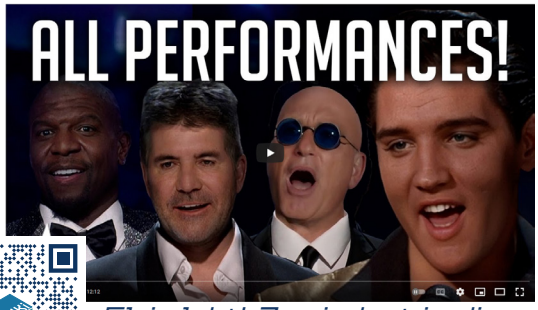
Technisch möglich ist das mit künstlichen neuronalen Netzen. Zuerst werden viele Bilder, Ton- oder Videoaufnahmen von Dir und der Person, die ersetzt werden soll, untersucht: Welche Gesichts- oder Stimmmerkmale habt ihr? Dabei ist es wichtig, viele unterschiedliche Gesichtsausdrücke, Lichtverhältnisse und Aufnahmebedingungen beziehungsweise Tonsequenzen mit unterschiedlicher Aufnahmequalität zu untersuchen, damit das KI-System möglichst genau weiß, wie das Gesicht in verschiedenen Situationen aussieht.

Anschließend vergleicht das KI-System das Original und Deine Aufnahmen, um die Aufnahmen von Dir so anzupassen, dass es diese auf die Originale legen kann. Wie das funktioniert, wird in Station 25 (Fotobox) genauer erklärt. Je öfter das KI-System das macht und je mehr Daten ihm zur Verfügung stehen, desto realistischer werden die Ergebnisse. Am besten funktioniert der Lernprozess mit einem GAN (engl. Generative adversarial network), bei dem ein KI-System die Fakes erstellt und ein zweites KI-System diese entlarven muss. Die KI-Systeme lernen also voneinander, da sie jeweils besser als das andere sein wollen.

Deepfakes lassen sich sogar mit unscharfen Fotos erstellen.

Die Erzeugung von Deepfakes hat **Vorteile**:

Personen, die nie lesen gelernt haben oder blind sind, können sich mithilfe dieser KI-Systeme Inhalte von ihrer eigenen oder einer vertrauten Stimme vorlesen lassen. Wer durch einen Unfall seine Stimme verliert, kann ein Hilfssystem mit der eigenen Stimme nutzen. Und bekannte Künstler:innen können Besucher:innen selbst durch ihre Ausstellungen führen und von ihren Kunstwerken berichten. Stell dir einmal vor, Mozart erzählt Dir selbst von der Zauberflöte oder Goethe liest aus Faust vor. Wäre das nicht cool?



Elvis lebt! Zumindest in dieser Deepfake-Performance bei America's Got Talent.

Doch es lauern auch viele **Gefahren**:

Wie würdest Du Dich fühlen, wenn im Klassenchat ein Bild auftaucht, dass Dich auf einer Veranstaltung zeigt, auf der Du nie warst? Wie würde Deine Großmutter reagieren, wenn sie ein Video per Messenger von Dir erhält, in welchem Du sie um Geld bittest, obwohl Du dieses Video nicht aufgenommen und verschickt hast? Was passiert, wenn Diktatoren Videoaufnahmen so manipulieren, dass sie Wahlen gewinnen oder gar Kriege entfachen? Ist dann die Demokratie in Gefahr? Und was kann man dagegen tun?

Im Video ist eindeutig Barak Obama zu sehen. Doch es spricht nicht der ehemalige US-Präsident, der seinen Nachfolger Donald Trump als „vollkommenen Dummkopf" bezeichnet, sondern ein Comedian.

In einigen Ländern ist das Verbreiten von Deepfakes bereits strafbar und auch in Europa wird darüber diskutiert. Verletzungen der Persönlichkeitsrechte oder von Urheberrechten sind schon verboten.

Doch was kannst Du konkret tun?

Sei aufmerksam und versuche, Deepfakes zu entlarven!

Häufig sind die Übergänge zwischen dem eingefügten Teil und dem realen Bild verschwommen. Auch können Gesichter und deren Konturen fehlerhaft sein. Stimmen Licht und Farben der Vorlagen nicht überein, muss das KI-System improvisieren. Da es aber nur darauf trainiert wurde, Gesichter zu manipulieren, hat es häufig Probleme mit den Hintergründen, wodurch es dort zu Fehlern kommt (falsche Schatten, Gegenstände an unnatürlichen Stellen usw.). Bei Videos kannst Du auf die Augen achten. Wirken diese, als ob jemand in die Ferne schaut oder blinzelt die Person häufig oder sehr selten, ist dies ein Zeichen, dass es sich um einen Fake handeln könnte. Audiofakes sind hingegen meistens fast perfekt. Während Menschen Füllwörter nutzen, das Tempo beim Sprechen verändern oder eben nicht immer flüssig sprechen, muss ein KI-System diese Unregelmäßigkeiten erst lernen.

Solange ein KI-System noch lernt, haben wir Menschen die Möglichkeit, es aufgrund seiner Fehler zu enttarnen. Wenn Du Dir nicht sicher bist, ob Du es mit realer Information oder einem Fake zu tun hast, versuche immer, andere Quellen zu finden, welche die Information auch enthalten.

QUELLEN

Graphik „Deepfakes lassen sich sogar mit unscharfen Fotos erstellen.“
https://the-decoder.de/geschichte-der-deepfakes-so-rasant-geht-es-mit-ki-fakes-voran

Deepfake-Video Barak Obama
https://youtu.be/cQ54GDm1eL0

Deepfake-Performance America‘s Got Talent
https://youtu.be/_pwpdc6oqaE

QUELLEN ALLES FAKE - APP

Alle Deepfake-Videos und Bilder wurden erstellt auf: https://facehub.live