# DEEP/FAKES

## Warum man keinem Video trauen kann – und was das Ganze mit Pornos zu tun hat

**Anders als Fotos ließen sich Videos bislang nur sehr aufwendig manipulieren. Spätestens seit dem Siegeszug des Machine Learnings ist das vorbei: Per „Deepfake“-Skript lässt sich ein Gesicht in einem Video einfach durch ein anderes austauschen – zumindest, wenn man genug Fotos hat. Populär wurde die fragwürdige Technik ausgerechnet durch Pornos.**

**Von Jan-Keno Janssen**

Mit „Komprimierte Charakterköpfe“ war im April 2002 ein Artikel in der c’t überschrieben: Ein neuer Videokompressions-Codec mache es möglich, Schauspielergesichter in Filmen einfach auszutauschen. Man müsse lediglich „die Facelet-Datei von Julia Roberts in arnold_schwarzenegger.m4f“ umbenennen, schon kann man sich „Notting Hill“ mit dem österreichischen Haudegen in der Hauptrolle anschauen. Das Ganze war so absurd, dass die meisten c’t-Leser den Artikel korrekterweise als Aprilscherz erkannten. Heute, 16 Jahre später, wird der harmlose Technikspaß Wirklichkeit – und zwar unter deutlich ernsteren Vorzeichen.

Es begann im Dezember letzten Jahres: Ein Reddit-Nutzer mit dem Pseudonym „deepfakes“ (offenbar ein Kofferwort aus „Deep Learning“ und „Fake“) veröffentlichte einen Porno-Clip, in dem „Wonder Woman“-Hauptdarstellerin Gal Gadot zu sehen war. Er hatte das Gesicht der Schauspielerin mithilfe der Machine-Learning-Bibliothek Keras (plus Googles TensorFlow als Backend) auf den Körper der Pornodarstellerin montiert. Den Algorithmus trainierte er zuvor mit Pornovideos und tausenden Fotos von Gal Gadots Gesicht, die er automatisiert aus Google Photos und YouTube-Videos gezogen hatte.

## Wie bei Frankenstein

Kurze Zeit später tauchte mit „FakeApp“ ein einigermaßen leicht zu bedienendes Frontend für den Machine-Learning-Gesichtertausch auf. Zeitgleich fluteten weitere Fake-Pornos mit bekannten Schauspielerinnen das Netz. Dass die Gesichter falsch waren, konnte man bei genauem Hinschauen zwar erkennen; dennoch war die Realitätsnähe in vielen Videos beeindruckend. Solche Tricks beherrschten bislang nur Hollywood-Effektschmieden mit viel Know-how und Handarbeit.

Das blieb auch den Medien nicht verborgen. Mit einer Mischung aus Faszination und Abscheu berichteten etliche Zeitungen und Websites über das Thema – unter anderem auch heise online. Hin und wieder schwang ein bisschen Technikbegeisterung mit („Was heutzutage alles geht!“); generell standen aber wie bei der alten Diskussion über Bildmontagen und Photoshop ethische Fragen im Mittelpunkt: Darf ein Algorithmus frei verfügbare Video- und Fotodaten so zusammensetzen, dass hinterher ein Gesicht auf einem fremden Kopf sitzt? Ist das überhaupt legal? Und was kann man heutzutage überhaupt noch glauben?

## Mobbing per AI

Die Leidtragenden der Technik sind nicht nur die Medienkonsumenten, die künftig bei „Beweisvideos“ noch genauer hinschauen müssen, sondern vor allem die unfreiwilligen (Porno-)Darsteller. „Es zeigt, dass einige Männer Frauen ausschließlich als Objekte sehen, die sie manipulieren und zu allem zwingen können“, zitiert das Vice-Magazin die langjährige Pornodarstellerin Alia Janine: „Diesen Männern fehlt der Respekt vor Schauspielerinnen und Pornodarstellerinnen.“

Weil ja theoretisch nicht nur Prominente in Pornovideos montiert werden können, sondern auch Ex-Freundinnen oder Kolleginnen, sieht „Die Welt“ noch ganz andere ethische Probleme. Die Technik würde „perfidem Cybermobbing“ Tür und Tor öffnen. So etwas wäre allerdings definitiv illegal: Nicht nur wegen des Rechts am eigenen Bild, sondern auch wegen der klaren Verletzung des Persönlichkeitsrechts. Wer unfreiwillig in peinlichen Videos landet, kann die Produzenten obendrein wegen Beleidigung, Verleumdung und übler Nachrede anzeigen – Deepfakes können also sogar strafrechtlich relevant sein.

Das Online-Magazin „The Outline“ macht ein noch größeres Fass auf: „Experten befürchten, Gesichtstausch-Technik könnte einen internationalen Konflikt auslösen“ lautet die Überschrift eines Artikels zum Thema. Noch sei es einfach, verändertes Videomaterial zu erkennen. Aber das könne sich schnell ändern, hieß es bei einer Medien-Forensik-Konferenz der US-amerikanischen DARPA (Defence Advanced Research Agency). Der Informatik-Professor Hany Farid, einer der Konferenzteilnehmer, entwarf gegenüber Outline ein regelrechtes Horrorszenario: Nämlich dass ein Fake-Video, in dem Donald Trump den Abschuss von Nuklearwaffen gegen Nordkorea ankündigt, einen Atomkrieg auslöst. Schließlich lassen sich nicht nur Videos manipulieren, sondern auch Sprache (siehe S. 110). „Ich neige nicht zu Hysterie oder Übertreibung, aber wir können uns vermutlich darauf einigen, dass so etwas nicht komplett ausgeschlossen ist“, so Hany Farid.

## Was stimmt?

Aber ist die Situation wirklich so dramatisch? Ist es wirklich so einfach, solche Videos selbst zu basteln? Und vor allem: Klappt das nur mit Gesichtern, die man im Netz in allen erdenklichen Ausdrücken,

**Die Dame sieht aus wie „Wonder Woman“-Darstellerin Gal Gadot – in Wirklichkeit war ein Algorithmus am Werk, der das Gesicht von Gadot auf den Körper einer Pornodarstellerin montiert hat.**

Positionen und Beleuchtungssituationen findet? Oder geht das auch einem Otto-Normalverbraucher-Gesicht?

In der c't-Redaktion haben wir lange darüber diskutiert, ob wir uns mit „Deepfakes" beschäftigen wollen. Am Ende war klar: Ja, wollen wir. Nicht nur, weil die Technik für sich genommen völlig unschuldig ist. Sie kann nichts dafür, dass man sie für unethische Zwecke verwendet. Sondern auch, weil man die technischen Hintergründe kennen muss, um sich über das Thema eine differenzierte Meinung bilden zu können – und um gefälschte Videos zu erkennen. Wir sind uns auf alle Fälle sicher: Das Thema wird die Welt auch in Zukunft beschäftigen. Ob wir wollen oder nicht. *(jkj@ct.de)* c't

# Meinungen zu Deepfake-Videos

**Jörg Wirtgen**

Für die Betroffenen sind Porno-Fälschungen mit ihren hereingerechneten Köpfen natürlich eine Katastrophe, ein Trauma, ein persönlicher, intimer Angriff. Das gilt auch für Promis, die woanders ihren Körper gerne zeigen – dort allerdings selbstbestimmt.

Aber wenn man gesamtgesellschaftlich auf alle Anwendungsgebiete der Deepfake-Videos schaut? Das sendet sich doch alles weg, wer hat denn überhaupt die Zeit, sich diesen ganzen Blödsinn anzusehen? Die sowieso meist völlig unrealistischen Hollywood-Produktionen werden durch die Fake-Köpfe auch nicht glaubwürdiger.

Gefälschte Videos erfüllen vielleicht sogar einen bitter nötigen Schritt, nämlich eine wachsende Achtsamkeit gegenüber der digitalen Un-Authentizität. Es muss halt nicht stimmen, nur weil es im Internet steht – jetzt gilt auch für Videos, was für Fotos und Texte schon lange gilt. Je mehr offensichtlich gefälschte Videos auftauchen, desto stärker könnte die Skepsis vor vermeintlichen Wahrheiten werden, desto mehr könnten Leser auf vertrauenswürdige Quellen zurückkommen oder selbst recherchieren. Es bleibt nur zu hoffen, dass es nicht gerade massenweise gefälschte Pornos zu dieser Einsicht braucht.

**Helga Hansen**

Neue Technik – und die erste Anwendung ist nicht nur Porno, sondern sie verletzt auch noch Persönlichkeitsrechte. Deepfakes ist ein weiteres Symptom für Frauenverachtung im Internet.

Einst war das Netz ein Sammelplatz der Utopisten und ihrer Hoffnung, dass die Welt endlich gerechter würde: Gleichberechtigung für alle! Dabei warnen insbesondere Forscherinnen schon lange vor Diskussionsdynamiken im Netz, die unter dem Deckmantel der Freiheit auch Missbrauch und Gewalt erlauben. Eine der ältesten Untersuchungen verweist auf das Usenet der 80er Jahre: Neben ermüdenden Debatten mit Männern plagten gefälschte Nachrichten eine öffentliche Frauengruppe so, dass schließlich eine strikt moderierte Mailingliste aufgesetzt wurde. Seither wiederholt sich das Problem in den Online-Kommentarspalten. Moderation und Konsequenzen kosten Zeit, Geld und Nerven – deshalb gibt es sie meist nicht, was zu einer Normalisierung der sprachlichen Gewalt führt.

Inzwischen sind 40 Jahre mit immer neuen Dingen im Internet vergangen – und jedes Mal wird der Rahmen des Sag- und Machbaren nach unten ausgetestet. Warum also sollte es bei Video-Fakes anders sein?

**Jan-Keno Janssen**

Deepfakes finde ich aus technischer Sicht faszinierend. Was man damit alles machen kann! Zum Beispiel Freunden zum Geburtstag den Trailer ihres Lieblingsfilms schenken, in dem sie alle Rollen übernehmen. Und was macht die Internetgemeinde? Die tolle Technik für Pornofilmchen verwenden.

Was mich noch mehr erschüttert: Dass es offenbar bei vielen Deepfake-„Produzenten" keinerlei Unrechtsbewusstsein gibt. Es scheint für sie nicht nur das Selbstverständlichste der Welt zu sein, das Gesicht einer Pornodarstellerin ungefragt mit dem einer Schauspielerin auszutauschen. Es erscheint ihnen auch als vollkommen legitim, ihre Arbeitskollegin in solche Filmchen einbauen zu wollen – zumindest gab es entsprechende Anfragen im Deepfakes-Subreddit, als dieser noch nicht gebannt war. Die anderen Nutzer antworten darauf nicht etwa mit „Bist du noch zu retten?" – sondern erklärten geduldig, dass es schwierig sei, genug Fotos von der Zielperson zu sammeln, damit der Machine-Learning-Algorithmus ordentlich funktioniert.

Das Problem ist nicht die Technik: Die ist nach wie vor faszinierend. Das Problem sind ungesunde männliche Anspruchshaltung und mangelndes Einfühlungsvermögen.