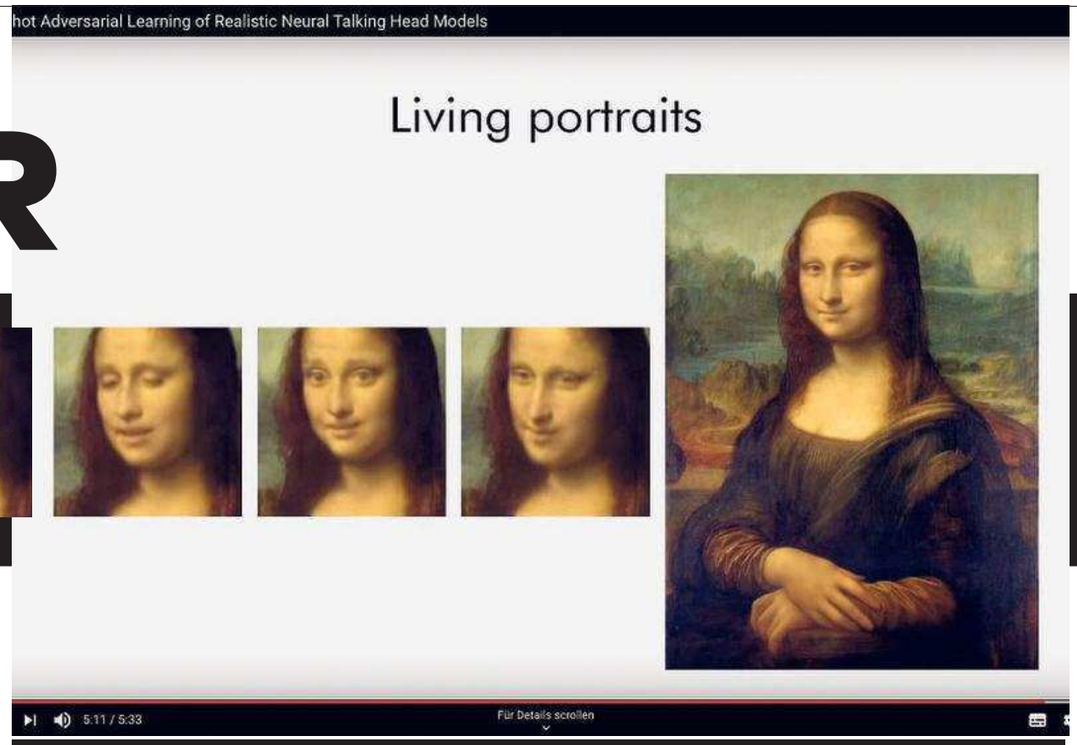
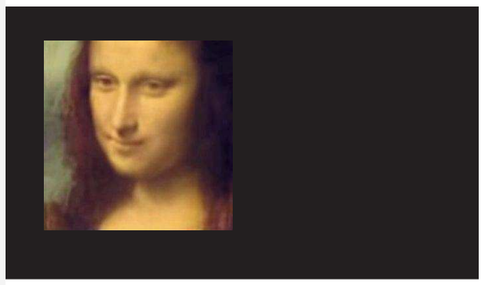
axel springer
syndication

Welt am Sonntag Nr. 49 vom 08.12.2019 – Seite 67

FÄLSCHER

Auf Basis von Leonardo da Vincis Gemälde erstellten Forscher ein Video, in dem die berühmte Frau spricht





auf Speed

Mit künstlicher Intelligenz erstellen Fälscher mittlerweile täuschend echte Videos von Prominenten. Die Folgen sind verheerend

VON JENS LUBBADEH

Ein Freitagnachmittag im März 2019. In einem britischen Konzern klingelt das Telefon eines Managers. Am Apparat: Johannes X., der CEO des Mutterkonzerns aus Deutschland. Der Chef macht Druck, eine dringende Zahlung über 220.000 Euro müsse umgehend an einen Lieferanten in Ungarn rausgehen. Er selbst könne es nicht veranlassen, die Banken in Deutschland hätten bereits zu. Großbritannien liege eine Stunde zurück, der britische Manager soll das Geld umgehend überweisen.

Der wundert sich, aber es besteht kein Zweifel: Es ist die Stimme des deutschen Chefs Johannes. Also überweist er 220.000 Euro nach Ungarn. Das Geld landet auf dem Konto von Betrügern. Sie haben ihn mit der Stimmen-Imitationssoftware „Lyrebird“ hereingelegt. Lyrebird nutzt künstliche Intelligenz (KI) und benötigt zum Trainieren zwar authentisches Audiomaterial, doch das muss nur wenige Minuten lang sein. Die Stimme wird geklont und kann dann jeden beliebigen Text sprechen.

Der Fall des „falschen Johannes“ ist ein sogenannter „Deep Fake“, eine digitale Fälschung mithilfe von KI. Was mit menschlicher Manipulation von Text und Bild begann und als Fake News derzeit die sozialen Medien durchseucht, wird durch die maschinelle Perfektion auf ein neues Level gehoben. Falsche Videos von Barack Obama, Mark Zuckerberg oder Kim Kardashian kursieren im Netz, in denen die Prominenten Sachen sagen, die sie in Wirklichkeit nie von sich gegeben haben. Das macht Deep Fakes zu einer gefährlichen Waffe. Gerade jetzt, in Zeiten, in denen ein einziges Video eine Regierung kollabieren lassen kann wie im Fall Österreichs.

Die künstlichen Intelligenzen sind mittlerweile so gut, dass man sie nur noch mit Audio- und Videomaterial füttern muss, um die Bewegungen und die Stimme eines Menschen zu synthetisieren. Das Deep-Fake-Video von Barack Obama hatten Computerwissenschaftler der University of Washington 2017 erstellt. Das neuronale Netz hatte mit Stunden von Audio- und Videomaterial des Ex-Präsidenten trainiert, seine Mimik und Sprache zu kopieren gelernt. Danach war es in der Lage, zu beliebigen Sprachaufzeichnungen Obamas täuschend echt aussehende Videos zu generieren. Mittlerweile reichen zur Erstellung von synthetischen Videos sogar noch Fotovorlagen, wie der russische Wissenschaftler Egor Zakharov vom Samsung AI Center in Moskau zeigte. Er verwandelte Leonardo da Vincis stumm dasitzende Mona Lisa in eine sprechende Frau vor der Kamera. Die Videoqualität nahm zu, je mehr Fotos der Forscher einsetzte. Beeindruckend ist auch die Face2Face-Technologie, die Matthias Nießner von der TU München bereits im Jahr 2016 demonstrierte: Er übertrug mithilfe einer KI die Mimik eines lebenden Schauspielers auf das Gesicht eines Prominenten. George W. Bush oder Wladimir Putin wurden so zu digitalen Handpuppen (siehe Fotos).

Deep-Fake-Videos von Barack Obama (l.) und Facebook-Chef Mark Zuckerberg (r.)

YOUTUBE (6)

Die bislang weitreichendste Möglichkeit der Deep-Fake-Manipulation präsentierten im Sommer 2019 Forscher des Max-Planck-Instituts für Informatik in Saarbrücken, Stanford und Princeton zusammen mit der Photoshop-Hersteller-Firma Adobe: textbasiertes Editieren von Videos mit dazugehöriger Mimik und Stimme. Damit lassen sich im Transkript des Gesagten Wörter einfügen oder löschen, und die KI erstellt fehlende Videoabschnitte. So kann man einer Person im Video beliebig Sachen unterjubeln. Noch geht das nicht in Echtzeit, „mit weiterem Fortschritt bei Deep-Fake-Videos ist diese Technik als nächste Evolutionsstufe denkbar“, warnt jedoch der Versicherungskonzern Euler Hermes, der den Fall des „falschen Johannes“ bekannt gemacht hatte. Mit dieser Technik hätten die Betrüger den falschen CEO sogar per Skype in England anrufen lassen können.

Fast sehnt man sich angesichts solcher Entwicklungen zurück in Zeiten, wo „nur“ Bilder gefälscht wurden. Wie das von Wladimir Putin, lässig im Sessel sitzend, zu seiner Rechten Donald Trump, zur Linken Recep Tayyip Erdogan. Dieser Bild-Fake vom G20-Gipfel in Hamburg machte 2017 auf Twitter die Runde. Schien er doch zu bestätigen, was viele vermuteten: Der US-Präsident als Marionette Putins. In Wirklichkeit saßen Trump und Erdogan vor einem leeren Stuhl, Wladimir Putin wurde in das Bild hineinkopiert. „Die Menge an gefälschten Bildern hat dramatisch zugenommen“, sagt Mauro Barni, Multimedia-Forensiker an der Universität von Siena, „vor allem aufgrund simpel anzuwendender Bildbearbeitungs-Apps auf Smartphones.“ Seit vielen Jahren beschäftigt er sich mit Bildfälschungen, genauso wie sein Kollege Jakob Hasse vom Dresdner Unternehmen Dence, das im Auftrag von Versicherungen, Gerichten und Medien Bilder und Videos auf Manipulation hin untersucht. Vor allem Medien seien anfällig für Täuschungen, so Hasse, „weil sie meist kaum Zeit haben, um ein Bild auf Echtheit zu prüfen“. Doch in Zeiten von Deep Fakes ist man vorsichtig geworden, vor der Veröffentlichung des brisanten Enthüllungsvideos über Österreichs Ex-Vizekanzler Strache zogen „Spiegel“ und „Süddeutsche Zeitung“ Multimedia-Forensiker zurate.

Die Daten sind der Tatort, an dem digitale Forensiker auf Spurensuche gehen. Es beginnt mit der Dateistruktur: „Eine Digitalkamera speichert eine Bilddatei anders ab als eine Bildbearbeitungssoftware“, sagt Jakob Hasse. „Wir können sehen, ob ein Bild nach der Aufnahme mit einer Software bearbeitet wurde.“ In den Pixeln stecken weitere Hinweise: „Der Sensorchip der Kamera hinterlässt einen charakteristischen Fingerabdruck im Pixelmuster“, sagt Mauro Barni. Anhand dessen lässt sich zum einen sagen, welche Kamera für die Aufnahme verwendet wurde. Und wenn ein Fälscher in eine Aufnahme oder in einzelne Video-Frames etwas einfügt, verändert er dieses Muster. „Der fremde Fingerabdruck des aus einem anderen Bild hineinkopierten Materials verrät die Manipulation“, sagt Mauro Barni. Weitere Anhaltspunkte sind die Schattenwürfe, die Perspektiven und das Licht in einem Bild generell. „Natürliches Licht erzeugt immer zufällige Muster“, sagt Jakob Hasse. „Die Muster einer Software haben Eigenschaften, die es in der Natur so nicht gibt.“

Was aber ist mit Bildern und Videos, die von einer KI verändert wurden? Können Forensiker dies auch entlarven? Martin Steinebach, Forensiker vom Fraunhofer SIT, unter anderem beteiligt an der Prüfung des Strache-Videos, hat keine Angst vor KI: „Sie ist letztlich auch nur ein Synthesizer, der Daten erzeugt.“ Auch Jakob Hasse sagt: „Bei KI-Material greifen die gleichen forensischen Methoden wie bisher.“ Eine sehr grundlegende Spur für die Forensiker ist beispielsweise das 50-Hertz-Signal des Stromnetzes. Es findet sich unbeachtet in sehr vielen Audiodaten, da fast immer irgendwo ein elektrisches Gerät läuft oder ein Strommast steht. Forensiker spüren dieses Signal in den zu prüfenden Daten auf und vergleichen es mit dem aufgezeichneten Referenzmuster des 50-Hertz-Signals des Stromnetzes in Deutschland. „Durch den Abgleich beider Muster können wir einerseits erkennen, wann die Datei erzeugt wurde“, sagt Martin Steinebach. „Und wenn das Signal in den Daten von der Referenz abweicht, wissen wir, dass an den Daten herumgedoktert wurde.“

Doch wäre es nicht denkbar, dass KI eines Tages auch noch lernt, dieses Signal perfekt zu fälschen? „Prinzipiell ist das möglich“, sagt Hasse. „Aber letztlich ist es ein Duell zwischen Fälscher und Forensiker.“ Eine KI ist immer nur so gut wie der Mensch, der sie programmiert und trainiert hat. „Um den Forensiker auszutricksen, muss der Fälscher alle seine Methoden kennen. Das erfordert eine Menge Spezialwissen.“ Die Datenanalysten brauchen Zeit, wie aber schützt man Leute vor Deep Fakes, die sich schnell auf Facebook und Twitter verbreiten? Die Organisation AI Foundation arbeitet an einem Browser-Plug-in namens „Reality Defender“, das diese automatisch erkennen soll. Auch Matthias Nießner, Erfinder der Face2Face-Technik, trainiert eine Forensik-KI, die Fälschungen entlarven soll. Aber am Ende stehen hinter der KI immer noch Menschen. Und die machen Fehler. So auch im Fall des „falschen Johannes“. Nach seinem ersten Erfolg wurde der Täter gierig. Doch bei einem zweiten Anruf nutzte er eine österreichische statt einer deutschen Nummer. Dadurch flog er auf.