Deep Fakes - Gefahren, Erkennung, Maßnahmen

Bei Deep Fakes handelt es sich um manipulierte Videos oder Bilder, die mithilfe von Computertechnologie so täuschend echt gemacht werden, dass sie schwer von echten Aufnahmen zu unterscheiden sind.Diese Technologie wird sowohl für humorvolle als auch für betrügerische Zwecke genutzt, z. B. um Prominente in peinliche Situationen zu bringen oder Politiker falsch darzustellen. Dadurch haben sie enormes Potential, Demokratien zu schädigen. Wie kann man sie erkennen und welche Maßnahmen sind noch notwendig?

Was ist der Unterschied zwischen Fake News und Deep Fakes?

Fake News und Deep Fakes sind zwei verschiedene Phänomene, die jedoch beide mit der Verbreitung von irreführenden oder manipulierten Informationen zu tun haben. Hier sind die Hauptunterschiede zwischen ihnen:

Fake News:
Fake News beziehen sich auf falsche Informationen, die absichtlich verbreitet werden, um Menschen zu täuschen, Meinungen zu beeinflussen oder bestimmte Ziele zu erreichen. Diese falschen Informationen können entweder komplett erfunden sein oder aus dem Kontext gerissen und verzerrt werden.

Deep Fakes:
Deep Fakes sind künstlich generierte Medien, typischerweise Videos oder Audiodateien, die mit Hilfe von fortgeschrittenen Technologien wie Deep Learning und KI erstellt werden. Diese Medieninhalte können täuschend echt aussehen oder klingen und Personen in Handlungen oder Aussagen zeigen, die sie nie getan oder gesagt haben.
Besonders bekannt sind die Face-Swap-Techniken. Face-Swap tauscht - wie der Name schon sagt - Gesichter einfach aus.

montviso
Montviso- Internetdienstleistungen
Regina Oswald
Stresemannstraße 15
DE-88400 Biberach
Mail: oswald@montviso.de
Tel.: +49 (0)7351 8295287
Fax: +49 (0)7351 3549980

Gefahren für Demokratien durch Deep Fakes

- **Manipulation von Wahlen**: Deep Fakes könnten dazu verwendet werden, politische Kandidaten in schlechtem Licht darzustellen oder falsche Aussagen zu verbreiten, um das Wahlergebnis zu beeinflussen.
- **Verbreitung von Desinformation**: Deep Fakes können dazu verwendet werden, Desinformationen zu verbreiten, indem sie falsche Inhalte erzeugen, die schwer von echten Inhalten zu unterscheiden sind.
- **Untergrabung des Vertrauens in Medien und Institutionen**: Deep Fakes können das Vertrauen der Öffentlichkeit in Medien und Institutionen untergraben, indem sie Zweifel an der Echtheit von Informationen schüren.
- **Manipulation von öffentlichen Diskussionen und Debatten**: Durch die Erstellung gefälschter Inhalte können Deep Fakes dazu verwendet werden, öffentliche Diskussionen und Debatten zu manipulieren.

Erkennung von Deep Fakes

Es gibt verschiedene Methoden, sowohl technischer Art, als auch menschliche Kontrolle, um Deep Fakes zu erkennen.

Technische Ansätze:
Es gibt Software, die Medieninhalte auf bestimmte Merkmale oder Artefakte, die auf eine Manipulation hinweisen können, analysiert. Zum Beispiel können ungewöhnliche Muster oder Anomalien im Bild oder im Audio auf ein Deep Fake hinweisen. Diese Software wird ständig weiterentwickelt, um mit den neuesten Technologien Schritt zu halten.

Möglichkeiten für Anwender:

- Überprüfe die Herkunft des Videos oder Bildes. Ist es von einer vertrauenswürdigen Quelle oder einem bekannten Medienunternehmen?
- Suche nach dem gleichen Ereignis oder der gleichen Person in anderen Medienquellen, um die Echtheit zu verifizieren.

montviso

Montviso- Internetdienstleistungen
Regina Oswald
Stresemannstraße 15
DE-88400 Biberach

Mail: oswald@montviso.de
Tel.: +49 (0)7351 8295287
Fax: +49 (0)7351 3549980

- Verwende Google Lens, um nach weiteren Fundstellen des Bildes und dessen Urheber zu fahnden.
- Untersuche den Kontext, in dem das Video auftaucht. Passt es zu anderen Informationen oder Ereignissen, die zu diesem Zeitpunkt bekannt sind?
- Achte auf unnatürliche Bewegungen der Personen im Video
- Achte auf Unregelmäßigkeiten oder Inkonsistenzen in den Gesichtszügen der Personen im Video. Sind die Augen, die Lippen oder andere Merkmale unnatürlich? Fehlendes Blinzeln oder ein leerer Blick sind verdächtig.
- Schau dir den Hintergrund des Videos an. Gibt es Anzeichen dafür, dass er nachträglich bearbeitet wurde, z.B. unnatürliche oder verschwommene Elemente?
- Achte darauf, ob die Lippenbewegungen mit den gesprochenen Worten übereinstimmen.
- Achte auf verpixelte Stellen, Unschärfe oder Artefakte im Gesicht, besonders an den Rändern von Haaren oder um die Augen herum. .
- Achte auf unnatürlich wirkende Haare oder Haut, z.B. fehlende Textur oder unrealistische Haarsträhnen.
- Überprüfe, ob die Lichtquellen im Video oder Bild mit der dargestellten Umgebung übereinstimmen. Deep Fakes können Beleuchtungsmängel aufweisen.
- Höre dir die Audioqualität des Videos an. Gibt es Anzeichen von Audiomanipulationen oder ungewöhnliche Hintergrundgeräusche?
- Prüfe die einzelnen Bilder (Frames) des Videos auf Unterschiede, Anomalien oder Inkonsistenzen, die auf eine Manipulation hinweisen könnten.
- Suche nach Analysen oder Einschätzungen von Experten, um eine unabhängige Bestätigung der Authentizität des Videos zu erhalten.
- Überprüfe, ob die im Video gezeigten Personen, Orte und Ereignisse zeitlich und räumlich konsistent sind.

montviso

Montviso- Internetdienstleistungen
Regina Oswald
Stresemannstraße 15
DE-88400 Biberach

Mail: oswald@montviso.de
Tel.: +49 (0)7351 8295287
Fax: +49 (0)7351 3549980

- Prüfe die Metadaten des Videos, um festzustellen, ob sie Anzeichen von Bearbeitung oder Manipulation enthalten.
- Es gibt Online-Tools und Apps, die bei der Erkennung von Deep Fakes helfen können.

Die Technologie entwickelt sich ständig weiter, und manche Deep Fakes sind so gut gemacht und Manipulationen so gut verborgen, dass sie selbst für Experten schwer zu erkennen sind.

Deepfake-Erkennungstools

- https://github.com/dfaker/dfaker: Dieses Tool nutzt maschinelles Lernen, um Deep Fakes anhand von Merkmalen wie Hauttextur, Haar und Augen zu erkennen.

- Sensity: https://www.sensity.ai/: Sensity bietet verschiedene Tools zur Erkennung von Deep Fakes und anderen Formen von synthetischen Medien.
- Truepic: https://www.truepic.com/: Truepic ist eine App, die Fotos und Videos mit einem kryptografischen Siegel versieht, um Manipulationen zu verhindern.
- https://www.spektrum.de/news/wie-man-sich-gegen-das-klonen-der-eigenen-stimme-durch-ki-wehrt/2199860

Links zum Thema:

- https://www.computerbild.de/artikel/cb-News-Internet-Intel-FakeCatcher-Tool-Erkennung-Deepfakes-Analyse-Blutfluss-34821201.html
- https://www.unite.ai/de/Die-besten-Deepfake-Detektor-Tools-und--Techniken/
- https://www.spektrum.de/news/deepfake-wie-lassen-sich-ki-generierte-bilder-enttarnen/2127222

Maßnahmen zur Bekämpfung von Deep Fakes

- **Gesetzliche Maßnahmen:**
  Wegen der Auswirkungen bis hin zu Gefahr für Globale Sicherheit, denken

montviso
Montviso- Internetdienstleistungen
Regina Oswald
Stresemannstraße 15
DE-88400 Biberach
Mail: oswald@montviso.de
Tel.: +49 (0)7351 8295287
Fax: +49 (0)7351 3549980
Impressum | Datenschutz

Regierungen auf der ganzen Welt darüber nach, wie sie Deep Fakes regulieren können. Das kann Verbote für die Verbreitung von Deep Fakes oder rechtliche Haftung für deren Erstellung umfassen. Durch klare Gesetze sollen diejenigen zur Verantwortung gezogen werden, die Deep Fakes erstellen oder verbreiten.

- **Technologische Lösungen:**
  Es gibt eine ganze Reihe von Technologien, die helfen können, Deep Fakes zu erkennen und zu verhindern. Dazu gehören zum Beispiel Wasserzeichen, die in Bilder oder Videos eingebettet werden können, um ihre Echtheit zu bestätigen. Auch die Blockchain-Technologie wird erforscht, um die Integrität von Medieninhalten zu schützen und deren Manipulation zu verhindern.

- **Ethische Überlegungen:**
  Zu den genannten Auswirkungen auf die Demokratie, können Deep Fakes massiv das Persönlichkeitsrecht von Personen verletzen und Personen oder Institutionen schädigen, indem sie sie in falsche oder kompromittierende Situationen bringen.
  Deep Fakes könnten dazu verwendet werden, Ungerechtigkeit und Ungleichheit zu verstärken, indem sie bestimmte Gruppen oder Personen gezielt diskreditieren oder diffamieren.
  Der Grad zwischen berechtigtem Eingreifen und Zensur ist dabei schmal, nicht alle gefälschten "Werke" sind ethisch fragwürdig oder gesetzwidrig.
  Aber wie meistens bei KI steht auch das Thema Urheberrecht im Raum.
  Dabei muss auch diskutiert werden, wo Verantwortlichkeiten liegen, z.B. wenn Deep Fake in sozialen Netzwerken geteilt wird.

- **Medienkompetenz:**
  Viele Aspekte zur Medienkompetenz sind nicht neu. Allgemeine Fähigkeit zu kritischem Denken und ein ethisches Bewusstsein sind Grundvoraussetzung. Dazu kommen viele technische Fähigkeiten, die notwendig sind, um Deep Fakes zu erkennen. Auch wer grundsätzlich KI ablehnt, sollte sich mit den ständig wachsenden Möglichkeiten beschäftigen und bei den Rechtsgrundlagen auf dem Laufenden halten.
  Damit kann in den Schulen nicht früh genug begonnen werden. Da die Methoden der Deep Fake Generierung immer besser werden und gleichzeitig die Anwendung und Erzeugung kinderleicht ist, lässt es sich kaum vermeiden,

montviso

Montviso- Internetdienstleistungen
Regina Oswald
Stresemannstraße 15
DE-88400 Biberach

Mail: oswald@montviso.de
Tel.: +49 (0)7351 8295287
Fax: +49 (0)7351 3549980

dass auch Kinder Opfer von Deep Fakes und gefälschten Bildern werden. Ein gesundes Selbstbewusstsein und Resilienz muss trainiert werden, um psychische Schäden gering zu halten, bei gleichzeitiger Sensibilisierung dafür, dass es sich um kein Kavaliersdelikt handelt, wenn Mitschüler*innen mit solchen Bildern gemobbt werden.
Auch Erwachsene tun gut daran, sich zu vernetzen und gemeinsam zu lernen und sich vor neuen Möglichkeiten zu warnen bzw. gemeinsam abzuwägen, ob ein Werk echt ist, oder Deep Fake. Dies bedarf ständigen Trainings.

Links zum Thema:

- https://t3n.de/news/fake-ki-erkennen-blockchain-near-1593849/
- https://www.spektrum.de/news/wie-nationen-vorhaben-kuenstliche-intellig…
- https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Inform…
- https://datascientest.com/de/deep-fake-gefahren-massnahmen-und-rechtsla…
- https://www.spektrum.de/news/ai-safety-summit-ki-braucht-regeln-aber-we…
- https://taz.de/Forscherin-ueber-Fake-News/!5996038/
- https://deutsches-schulportal.de/schule-im-umfeld/wie-man-schuelerinnen…

montviso
Montviso- Internetdienstleistungen
Regina Oswald
Stresemannstraße 15
DE-88400 Biberach
Mail: oswald@montviso.de
Tel.: +49 (0)7351 8295287
Fax: +49 (0)7351 3549980
Impressum | Datenschutz