



Künstliche Intelligenz: Wenn Deepfakes zu "News" werden

Es wird immer einfacher, Fake-Videos und Fake-Bildmaterial zu erstellen

Deepfakes sind mit Hilfe künstlicher Intelligenz (KI) erstellte, realistisch wirkende Medieninhalte - das können Bilder, Videos oder Tonaufnahmen sein.

Während eine KI vor wenigen Jahren noch viele Bilddateien einer Person benötigte, um gute Deepfakes herzustellen, reicht heute ein einzelnes Foto aus einem öffentlich zugänglichen Social-Media-Account aus, um **täuschend echt aussehende Fälschungen** zu erstellen. Durch kostenlose Face-Apps und frei verfügbare Browser-Anwendungen wird es immer leichter, solche Fakes herzustellen und diese anschließend zu verbreiten.

Face-Apps nutzen eine Technik namens **Face-Swapping**, bei der das eigene Gesicht oder das Gesicht anderer Personen in Videoclips und Filmszenen eingesetzt wird. In Sekundenschnelle können solche Clips in überraschend guter Qualität erstellt und verbreitet werden.

Browserbasierte KI-Anwendungen, sogenannte **Bildsynthesizer**, erstellen (**foto**realistische) **Bilder anhand von Textbeschreibungen**, sogenannten **Prompts**, und bieten ein breites Spektrum an Bildstilen. Dabei kann z. B. auf Bilder von prominenten Personen, bekannten Bauwerken und Landschaften zurückgegriffen werden. Diese technischen Möglichkeiten heben die Bildmanipulation auf ein neues Level.

Jüngst sorgten mehrere Journalisten in sozialen Medien für einige Verwirrung. So machten beispielsweise **KI-generierte Bilder des Papstes als Influencer** oder von der **vermeintlichen Verhaftung des ehemaligen US-Präsidenten Donald Trump** die Runde, die viele Nutzer offenbar nicht als Fakes erkannten - und so gingen diese "viral", obwohl sie als solche gekennzeichnet waren.

Wofür werden Deepfakes genutzt?

Deepfakes können zu verschiedenen Zwecken eingesetzt werden. Die **Anwendungsmöglichkeiten reichen von Satire und Unterhaltung über bewusste Desinformation und Propaganda** bis zur gezielten Diskreditierung einzelner Personen. Wie die obigen Beispiele zeigen, sind können Deepfakes besonders gefährlich im Zusammenhang mit Fake News werden, da sie diese noch realistischer gestalten können.

Auch für Cybermobbing werden Deepfakes benützt und beispielsweise die Köpfe von Betroffenen in Porno-Videos montiert und per Messenger verbreitet.

Betrug mit Deepfakes ist ebenfalls möglich

Es ist auch **denkbar, dass KI mit betrügerischer Absicht genutzt wird**. **Love-Scammer** könnten beispielsweise diese Technik in Dating-Apps nutzen oder in Online-Partnerbörsen versuchen, andere Personen dazu zu bringen, erotisches Material mit ihnen zu teilen, um sie damit zu erpressen. Dabei geben sich die Täter als attraktive Männer oder Frauen aus und nutzen **Deepfake-Videochats**, um das Vertrauen von möglichen Opfern zu erlangen. Dieses Vorgehen ist als Sextortion bekannt.

Für die Gesellschaft bedeutet das, dass Betrüger immer einfacher Bilder, Stimmen oder auch Texte verfremden und zu ihrem Nutzen ändern können. **Es wird somit schwieriger, echtes von gefaktem zu unterscheiden**. In der Prävention muss dies in Zukunft stärker bedacht und vermehrt darauf hingewiesen werden. Doch KI wird nicht nur von Tätern und Täterinnen genutzt, auch die Polizei und Strafbehörden setzen bei Ermittlungen vermehrt auf KI-basierte Analyse-Tools.

Kann man Deepfakes erkennen?

Mittlerweile gibt es Programme, die ebenfalls **auf Basis Künstlicher Intelligenz** versuchen, **die Echtheit von Videos, Bildern und Audioinhalten zu prüfen und Fälschungen zu identifizieren**.

Inzwischen ist es jedoch teils sehr schwierig, Fakes von echten Inhalten zu unterscheiden. Es gibt aktuell immer noch **Bereiche, bei denen sich die KI-Bildgeneratoren schwertun**, z. B. bei der Darstellung von Ohren und Händen oder von Gegenständen, die in Händen gehalten werden. Auch Lichtfall und Schatten sowie Perspektiven oder inkonsistente Bewegungen können **Anhaltspunkte für Fakes** sein.

- Achten Sie auf **Artefakte**: Hände haben oft zu wenige oder zu viele Finger, Arme fehlen oder sind abgeschnitten, Ohren haben eine schlecht ausdifferenzierte Form, ein "leerer" Blick.
- Bei Videos können inkonsistente oder allzu statische Bewegungsabläufe ein Hinweis sein (z. B. verzerrte Gesichter/Mimik oder ein zu rhythmisches Blinzeln).
- Prüfen Sie die Perspektive: Passen die **Lichteinstrahlung und der Schattenwurf** zusammen? Gibt es Verzerrungen?
- Prüfen Sie die **Auflösung** bzw. die Qualität des Bildes: Hat das Gesicht beispielsweise dieselbe Auflösung wie der Körper oder der Hintergrund des Bildes?
Ist der **Inhalt des Bildes oder Videos plausibel**? Prüfen Sie andere Quellen. Lässt sich der dargestellte Inhalt unabhängig auf den Wahrheitsgehalt hin prüfen?

Hilfreich können bei der Prüfung auch seriöse News-Seiten, Faktencheck-Portale und die Bilder-Rückwärtssuche von Suchmaschinen sein.

Wenn Sie Ihr Konterfei in einem Fake entdecken

Deepfakes verletzen, juristisch gesehen, ohne entsprechende Berechtigung durch die betroffenen Personen, die **Persönlichkeitsrechte** oder können Beleidigungen darstellen.

Wenn Sie Ihr Abbild in einem künstlich erzeugten Bild oder Video entdecken, das ohne Ihre Zustimmung erstellt und verbreitet wurde, können Sie den Inhalt der Plattform, auf der dieser veröffentlicht wurde, melden (in der Regel über einen "Meldebutton" oder eine entsprechende Funktion).

Zusätzlich können Sie **Anzeige bei der Polizei erstatten oder sich juristische Hilfe holen**. Speichern Sie dazu die URL des Fundortes ab sowie das Veröffentlichungsdatum und die Uhrzeit des Abrufs. Sie können auch Screenshots des Posts als Beweis erstellen. Bitte beachten Sie jedoch, dass es je nach Inhalt des Fakes strafbar sein kann, selbst in Besitz dieses Inhalts zu gelangen.

Erstellen Sie deshalb keine Screenshots, wenn es sich beispielsweise um Darstellungen sexualisierter Gewalt oder Kinderpornografie handelt.

Wenn Sie sich unsicher sind, wie Sie die Beweise am besten sichern, können Sie sich bei Ihrer Polizeidienststelle vor Ort erkundigen.