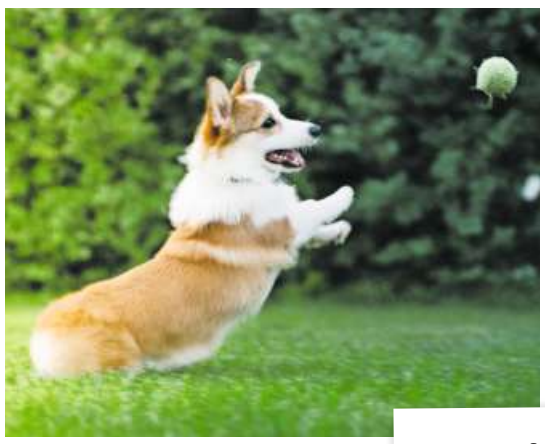
Welches ist echt?
Jedes Bildpaar besteht aus einem echten Foto und einem KI-generierten Bild, das kaum zu identifizieren ist. Erst Details verraten die Fälschungen.



Auflösung: Familie: links fake (behaartes Bein), rechts echt / Baby: oben echt, unten fake (sechs Zehen) / Hund: links echt, rechts fake (krummer Baum).
Fotos: Getty, KI-Bilder: Midjourney/Watson

Diese Tipps helfen, Fälschungen zu entlarven

Mit künstlicher Intelligenz (KI) lassen sich im Nu täuschend echt aussehende Bilder und Videos produzieren. Doch perfekt sind sie selten, weshalb sie mit ein bisschen Recherche oft enttarnt werden können.

Stephanie Schnydrig

Noch nie war es einfacher, ohne jegliche Informatik-Kenntnisse Bilder und Videos zu erstellen, die schockierend realistisch aussehen. Innert Sekunden spucken KI-Anwendungen wie «Stable Diffusion» oder «Midjourney» sogenannte Deepfakes aus, nachdem man die gewünschte Szene mit ihren Akteuren und deren Handlungen in ein paar Worten in die Befehlszeilen der Programme eingetippt hat.

Manche Werke sind lustig, wie das Bild des Papstes im Daunenmantel, das vor einiger Zeit viral ging. Andere säen Desinformation und sind nach Ansicht mancher Fachleute eine Gefahr für die Demokratie. Auch, weil Bilder und Videos viel überzeugender wirken als reiner Text.

KI-Inhalte sollten klar gekennzeichnet werden

Die Nichtregierungsorganisation Algorithm Watch CH und die Agentur Feinheit haben kürzlich das Projekt «Fake or Not» ins Leben gerufen, um zu zeigen, welche Risiken von solchen Bildern für politische Wahlkämpfe ausgehen. In einem Experiment erstellten sie mit einem KI-Generator Porträtbilder der aktuellen Ständerätinnen und Ständeräte, die sich kaum von Original-Fotos unterscheiden liessen.

Für Aufregung gesorgt hat letzthin ein Plakat der FDP, auf dem Klimaaktivistinnen und -aktivisten sitzend auf der Strasse zu sehen sind, die eine Ambulanz blockieren. Ein mit KI erstelltes Sujet, wie die Partei auf dem Plakat zwar festhält, allerdings in einem nicht auf den ersten Blick sichtbaren Hinweis.

Algorithm Watch CH beurteilt das Wahlplakat denn auch als bedenklich und schreibt auf Twitter: «Solche verfänglichen Fake-Bilder können die öffentliche Debatte und die freie Meinungsbildung auf eine Art und Weise verzerren, wie es aus demokratischer Sicht problematisch ist.» Die NGO fordert deshalb klare gesetzliche Spielregeln, wann und wofür KI-Systeme entwickelt und eingesetzt werden dürfen.

Weitherum ertönt der Ruf nach Transparenz für KI-generierte Bilder. Die EU-Kommission fordert von Unternehmen wie Google, Facebook, Youtube oder Tiktok eine Kennzeichnung für KI-Produktionen, beispielsweise durch eine Art Wasserzeichen. Die Vorgabe beruht allerdings auf Freiwilligkeit. Microsoft hat Ende Mai bereits ein Tool vorgestellt, womit die Bilder, die mit dem hauseigenen KI-Generator erstellt werden, markiert werden. Doch das allein vermag die Verbreitung von Deepfakes nicht zu stoppen: Wer Desinformation verbreiten will, tut dies ohnehin.

Aber: Noch sind die gefälschten Bilder und Videos nicht perfekt, und mit ein wenig Recherche lassen sich viele Deepfakes entlarven. Catherine Gilbert ist Faktencheckerin bei der Nachrichtenagentur Keystone-SDA und gibt Tipps für den eigenständigen Faktencheck. Sie betont: «Die Strategie muss immer sein, mehrere Dinge zu prüfen. Ein Hinweis allein ist noch kein Beweis für Fakt oder Fake.»

1 Die Nachrichtenlage checken

«Ein Alarmsignal ist, wenn ein Bild oder ein besonders schockierendes Video eine Szene darstellt, von der nirgends sonst berichtet wird», sagt Expertin Gilbert. Wenn beispielsweise ein Video eines berühmten Politikers kursiere, in dem er vor einer Menschenmenge kontroverse Statements ausspreche, müsste davon anderswo zu lesen sein. «Geschieht etwas Aussergewöhnliches in der Öffentlichkeit, erfährt dies in den meisten Fällen eine Resonanz in den klassischen und sozialen Medien sowie bei Nachrichtenagenturen.»

Wichtig sei auch, sich zu fragen, wann und wo das Bild oder Video aufgenommen wurde – und ob dies überhaupt möglich gewesen wäre. Also, ob es am besagten Tag beispielsweise überhaupt sonnig war oder ob ein anderes Ereignis stattfand zum selben Zeitpunkt, an dem der Politiker gleichzeitig gewesen sein soll.

2 Umgebung des Bildmittelpunkts studieren

Der Hintergrund verrät oft, ob es manipuliert wurde. Denn, wie Gilbert festhält, konzentriere sich der KI-Generator oftmals auf den Bildmittelpunkt. «Rundherum sind Objekte dann aber häufig deformiert, beispielsweise Bäume oder Strassenlaternen.» Oder aber die gezeigten Objekte existieren gar nicht, zum Beispiel Bücher mit frei erfundenem Titel oder Fantasie-Logos. «Häufig verraten auch Beleuchtung, Spiegelungen oder Schattenwurf, dass etwas nicht stimmen kann», sagt die Faktencheckerin.

3 Typische KI-Fehler kennen

Zu schaffen machen KI-Generatoren unter anderem Hände. Es ist keine Seltenheit, dass das Programm eine Person mit zu vielen oder zu wenigen Fingern ausstattet. Dies war auch der Fall im gefälschten Papst-Bild: Sieht man sich seine rechte Hand genau an, scheint es, als hielte er den Kaffeebecher nur mit vier Fingern. Auch sehen die Finger der linken Hand ungewöhnlich lang aus.

Andere häufige Fehler in KI-generierten Bildern sind Menschen mit viel zu vielen Zähnen, Brillengestelle, die seltsam deformiert sind, oder Ohren, die unrealistische Formen haben.

4 Auf Details im Gesicht achten

Auch im Gesicht können Schattenwürfe stutzig machen, etwa rund um Augen und Augenbrauen. Häufig machen KI-Generatoren auch eine zu geglättete oder zu gerunzelte Haut, nichts dazwischen. Die Dreidimensionalität wird nicht korrekt wiedergegeben. Auch einzelne, abstehende Haarsträhnen gibt es bei KI-Werken manchmal.

Die Faktencheck-Profis des Massachusetts Institute of Technology (MIT) raten zudem, bei Videos das Zwinkern zu beobachten, denn häufig zwinkerten Personen bei Deepfakes zu viel oder zu wenig. Auch Lippenbewegungen passen mit dem Gesagten häufig nicht überein.

5 Körperproportionen studieren

Oftmals treten in KI-generierten Werken Unstimmigkeiten in Bezug auf Körperproportionen auf: etwa zu kleine Hände, zu lange Beine für einen kurzen Oberkörper oder ein zu grosser Kopf. Manchmal passen auch die verschiedenen Merkmale nicht zusammen. Beispielsweise wenn eine Person viele Runzeln, aber kein einziges graues Haar am Kopf hat.

6 Bild bis ins kleinste Detail zoomen

Schaut man sich ein Bild nur flüchtig und im Kleinformat an, bleiben die Ungereimtheiten und Fehler häufig unentdeckt. «Mögliche Deepfakes sollten immer auf einem grossen Desktop-Bildschirm angeschaut werden, nicht bloss auf dem Smartphone», sagt Gilbert. Sie empfiehlt, die höchstmögliche Auflösung des Bildes im Internet zu suchen und anschliessend so nah wie möglich hineinzuzoomen.

7 Bild-Rückwärtssuche durchführen

Es gibt Anwendungen, die eine sogenannte Bild-Rückwärtssuche erlauben. Damit kann man prüfen, was die ursprüngliche Quelle des Bildes ist und ob das Bild so oder in ähnlicher Fassung überhaupt schon einmal im Netz verbreitet wurde. «Wenn das Bild nirgends auffindbar ist, kann es natürlich sein, dass es sich um eine brandneue Aufnahme handelt», sagt Faktencheckerin Gilbert. Aber es könne auch ein wichtiger Hinweis sein, dass das vermeintliche Foto so gar nicht existiere und demzufolge auch auf keinem sozialen Netzwerk und keinem Medienbericht auffindbar ist.

Gute Bilder-Rückwärtssuchprogramme sind die auf Bilder spezialisierte Suchmaschine Tineye, die Browsererweiterung RevEye oder die Bildsuche von Google, der Google Lens.

8 Vorsicht mit Entlarvungsprogrammen

Es existieren auch spezifische Programme, die KI-Inhalte erkennen. «Sie sind allerdings nicht perfekt», warnt Gilbert. Die Tools sind nicht ausgereift, sodass selbst echte Fotos als Fälschungen deklariert werden und andersherum. Als Baustein der Faktenüberprüfung können sie dennoch hilfreich sein. ZeroGPT verspricht etwa, mit KI generierte Texte zu entlarven. Andere Programme wie etwa ein Tool von Hive Moderation wollen vorhersagen, ob ein KI-Bild vorliegt.

Fazit: Es gibt keine einfache Lösung für die Entlarvung von Deepfakes. Es ist daher wichtig, dass wir als Medienkonsumentinnen und -konsumenten sensibilisiert werden für das Problem von täuschend echt aussehenden Fotos und Videos. Doch die Technologie dahinter wird zweifelsohne immer besser und Fehler werden in Zukunft seltener. Die Pflicht zu einer klar sichtbaren Kennzeichnung würde immerhin für Transparenz sorgen.