Ursprünglich erschienen am 13.12.2022 unter
https://www.wzb.eu/de/artikel/der-wahrheit-ins-gesicht-sehen

Der Wahrheit ins Gesicht sehen

Neue Dummheiten in der Künstlichen Intelligenz

A.K.M. Skarpelis

Jeder Mensch sieht anders und unverwechselbar aus – und lässt sich anhand seines Bildes oder einer Aufnahme identifizieren. Auch die automatische Gesichtserkennung nutzt diesen Umstand. Politisch fragwürdig ist, dass die Verfahren automatischer Gesichtserkennung zunehmend von Regierungen und Unternehmen genutzt werden. Diese Entwicklung wird vielfach kritisiert. Anders sieht es bei der technischen Seite der Gesichtserkennung aus, die von Kritikern weniger in Frage gestellt wird. Die Fotografie als Wiedergabe der Wirklichkeit genießt nach wie vor einen Vertrauensvorschuss. Doch nicht nur ist dieses Vertrauen im Zeitalter der massenhaften Retusche nicht mehr gerechtfertigt. Uns entgeht darüber hinaus, dass in der Künstlichen Intelligenz Techniken zur "Rassenerkennung" entwickelt worden sind, die neue ethnische Kategorien bilden. Sie sind im Vergleich zu denen des 19. und 20. Jahrhunderts nicht weniger fragwürdig.

Wenn wir über automatisierte Gesichtserkennung sprechen, gehen wir, ohne darüber nachzudenken, vom Verweis-Charakter der Fotografie aus: dass das, was man sieht, auch das ist, was ist. Die Glaubwürdigkeit von Fotografien beruht auf ihrer Objektivität, ihrer Nähe zu dem Ereignis, das sie abbilden. Diese Repräsentation kann auch um die Ecke funktionieren: Roland Barthes beginnt Camera Lucida mit dem Ausruf angesichts eines Bildes von Napoleons jüngstem Bruder Jérôme: "Ich sehe die Augen, die den Kaiser ansahen!". Entscheidend ist hier die Authentifizierungsleistung der Fotografie: Man geht

davon aus, dass die Fotografie, auch die der Gesichtserkennung, sich mit ihrem Referenten deckt, mehr als es Gemälde oder Zeichnungen tun.

Aber zum einen sind die Dinge wie immer nicht so einfach; und zum anderen können wir die gegenwärtigen Entwicklungen in der künstlichen Intelligenz nicht verstehen, ohne auf ältere Erkenntnisse zu Fragen der Repräsentation zurückzugreifen. Daher nähert sich dieser Beitrag dem Thema Wahrheit in der künstlichen Intelligenz aus der Perspektive, was mit den Menschen passiert, die in die Maschinerie algorithmischer Logiken geraten. Denn die Binärform „wahr/falsch" eignet sich nicht dazu, uns bei der Klärung ethischer Probleme in der künstlichen Intelligenz, insbesondere bei digitalen Bildern, zu helfen. Vielmehr haben wir es mit einer komplexeren Reihe von Fragen zu tun, die die grundlegende Einheit der sozialwissenschaftlichen Analyse zu sprengen drohen: Was ist eine Person, und wie erkennen wir sie?

Aufsätze über die Wahrheit in Bildern beginnen häufig mit Verweis auf eines der ältesten Schlachtfeldfotos der Welt: Roger Fentons Valley of the Shadow of Death. Das Foto aus dem Jahr 1855 zeigt eine trostlose Landschaft und eine sich über einen Hügel windende Straße, die mit Kanonenkugeln übersät ist. Wir sind aufgefordert, das trostlose Nachspiel des Krieges zu betrachten: Zurück bleiben Trümmer. Doch wie mehrere Kritiker aufgezeigt haben, handelt sich bei dem Bild wahrscheinlich um eine Fälschung. Es wird vermutet, dass Fenton die Kanonenkugeln aus dem Graben auf die Straße verlegt hat, um einen dramatischeren Effekt zu erzielen.

Fentons Kritikern geht es darum, den Aussagewert von Fotografien zu hinterfragen, da der Fotograf die soziale Welt durch Hinzufügen oder Streichung in verschiedenen Phasen des Produktionsprozesses verändern kann. Diese Kritik am Zeugnis- und Repräsentationswert von Fotografien ist später von verschiedenen Seiten aufgegriffen worden. In seinem Video Facial Weaponization Suite baut der Künstler Zach Blas amorphe Masken aus gesammelten Gesichtsdaten, die die Identität des Trägers verschleiern. Think Tanks warnen vor dem erheblichen Risiko von Deepfakes - digital veränderte Videos, die mit Hilfe von Deep Learning erstellt werden und realistisch erscheinen, aber Dinge darstellen, die nicht stattgefunden haben. Videos galten im Vergleich zu Fotografien lange als fälschungssicher, aber nun ist auch ihre Zeugnisfunktion in Frage gestellt. Dadurch können sie von politischen Akteuren strategisch zur Desinformation genutzt werden. Fentons Fotografien, Blas' Masken und die inzwischen allgegenwärtigen Deepfakes zeigen, wie die kausale Verbindung von Referent und Abbild aufgebrochen werden kann.

Manipulation ist leicht möglich

Das Zeitalter von Big Data hat uns zu Recht zu denken gegeben, wie leicht diese Manipulation möglich ist und wie wahrscheinlich es ist, dass sie bereits weithin ausgeübt wird. Wissenschaftler:innen, die sich mit Fehlinformationen befassen, leisten bewundernswerte Arbeit, wenn sie klare Grenzen zwischen wahr und falsch, authentisch und gefälscht ziehen. Es gibt jedoch Fälle von Manipulation, die schwieriger zu beurteilen sind. Nehmen wir als Beispiel die persönliche Identifizierung für zugeschriebene Identitätskategorien wie Geschlecht oder Rasse. Hier hat die Forschung gezeigt, dass die

Künstliche Intelligenz Probleme damit hat, Grenzfälle oder Mehrdeutigkeiten zu erkennen. In dem Moment in dem eine Maschine ein Foto „liest“ und eine ethnische Klassifizierung als Output liefert, stellt sich die Frage: Was produzieren Bilder? Empirisch gefragt: was passiert, wenn die künstliche Intelligenz auf fotografische Bilder von menschlichen Gesichtern trifft?

Antworten, die der „Wahr-falsch“-Logik von Fentons Kanonenkugeldiskussion folgen, führen uns in eine Sackgasse. Weil eine Kategorie wie „Rasse“ sozial konstruiert ist, kann es keine „wahrheitsgetreue“ Abbildung von ihr geben. Probieren wir es also anders und machen eine Zeitreise ins Jahr 1842, als der berühmte Fotograf Nadar den sich sträubenden französischen Romancier Honoré de Balzac zu einer Daguerreotypie überreden musste; Balzac war bis zum Schluss überzeugt davon, dass Fotografien seinen Körper Schicht um Schicht abtragen würden. Diese Angst vor dem Entzug von etwas Wesentlichem zieht sich durch die Geschichte der Fotografie. Mehr noch, die Geschichte der Fotografie ist mit einer Geschichte der Gewalt verwoben, und die Entwicklung des Mediums Fotografie ist verflochten mit der des Kapitalismus und des Kolonialismus.

Balzacs Metapher des Körpers, der Schicht für Schicht entkleidet wird, mag uns heute hoffnungslos abergläubisch erscheinen. Aber zieht man den Aspekt des Aberglaubens ab, so behält sie in doppelter Hinsicht Gültigkeit. Die fehlende Einwilligung ist bis heute ein Problem, und wie Balzac dürfen wir uns sorgen, was eigentlich von der Person übrigbleibt, sobald sich unsere Ebenbilder in Vektor und Matrix umwandeln. Dies führt uns weg von der Frage, ob Bilder inhärent akkurate Repräsentationen der Welt sind, und hin zu einer Auseinandersetzung damit, wie die Statistik Gesichtern Bedeutungsebenen zuordnet. Und somit auch zu dem bereits vorgestellten sozialwissenschaftlichen Erkenntnisinteresse: Was bestimmt eine Person, und wie können wir ihre Charakteristika durch wissenschaftliche Methoden erfassen?

Wie Algorithmen bei der Gesichtserkennung wirken

Die automatisierte Gesichtserkennung wird von Unternehmen und staatlichen Einrichtungen zunehmend eingesetzt, beispielsweise bei verdachtsunabhängigen Polizeikontrollen, aber auch an Grenzen zum Aussieben von Geflüchteten, denen man unterstellt, über ihre Herkunft gelogen zu haben. Aber nicht immer finden sich in den Anwendungen ausdrücklich rassistische Motivationen: Die Gesichtserkennung kann auch ganz profan Marktforschungszwecken zur Umsatzsteigerung dienen. Ihr Reiz besteht aus Sicht seiner Hersteller nicht zuletzt daraus, dass keine Einwilligung eingeholt werden muss: Informatiker:innen loben sie als Identifizierungsmethode gerade deshalb, weil sie aus der Ferne, ohne Zustimmung und sogar ohne Wissen der fotografierten Person durchgeführt werden kann. Sie ist zwar bei der Eins-zu-eins-Identitätsüberprüfung weniger effizient als ein Iris-Scan, doch wird dies durch die einfache Anwendung wettgemacht. Die fehlende Einwilligung der Beteiligten bei der Erhebung von Bilderdatensätzen gilt im Übrigen auch in anderen Bereichen, wenn beispielsweise polizeiliche Verbrecherfotos zum Zweck des Trainings neuer Algorithmen zweitverwertet werden und wiederum bei der Gesichtserkennung Anwendung finden.

Dass diese Verfahren überaus problematisch sind und das, was wir für „wahr“ halten, untergraben, liegt auf der Hand. Zum einen führt die fehlende Einvernehmlichkeit zu einer ethischen und moralischen „Distanz“ zwischen Kamera und Person; zum anderen zeigt sich besonders am Beispiel der automatisierten „rassischen“ Gesichtserkennung die wachsende Distanz zur Person als Person: Im Vergleich zu anderen Bereichen der künstlichen Intelligenz bleibt die Detailschärfe unterentwickelt. Obwohl Algorithmen Massen von Bildern verwerten, tun sie dies auf eine Art und Weise, die oberflächlicher ist als vergangene Methoden der Klassifizierung. Wenn diese Behauptung widersprüchlich erscheint, dann deshalb, weil wir bei den großen Datenmengen von Input in künstlicher Intelligenz, maschinellem Lernen und Trainingsdatensätzen als Begleiterscheinung eine gleichzeitige Komplexität der Analyse voraussetzen. Doch wie ich an anderer Stelle zeigen werde, sind die heute genutzten Computer-Vision-Technologien deutlich detailärmer als frühere Technologien der Rassenverfolgung, wie zum Beispiel „rassische Eignungsprüfungen“ im Nationalsozialismus, da sie allein auf der Basis von 2D-Repräsentationen des Gesichtes funktionieren und alle möglichen weiteren Daten, die der Klassifizierung dienen könnten, außen vor lassen. Bevor man nun auf die Idee kommt, dass dies eine gute Nachricht sein könnte: Die Tatsache, dass die Techniken flacher und weniger offensichtlich ideologisch bestimmt sind heißt nicht, dass rassistische Auswahlprozesse dadurch weniger wirksam wären.

Die zweite Balzac’sche Sorge hängt mit den Bedenken zusammen, die in der Wirtschaftssoziologie hinsichtlich der Auswirkungen von Algorithmen auf Kategorien der Person geäußert werden. Hier wird argumentiert, dass ein Individuum seinen Status auf Grundlage seiner Gruppenzugehörigkeit erlangt; die algorithmische Logik unterbricht jedoch diese Verbindung von gemeinsamen Klassenmerkmalen. Es verwässert sich der Blick auf Bevölkerungsgruppen, die vorher gesellschaftlich erkennbar waren. Dadurch geht nicht nur das Mobilisierungspotential von Seiten der Betroffenen verloren, die mangelnde „Lesbarkeit“ von Bevölkerungsgruppen kann sich auch konkret auf Umverteilungsprozesse auswirken. Die Soziologen Greta Krippner und Dan Hirschman haben dies mit Hinweis auf Kreditmärkte und Gender gezeigt: Wohingegen frühere „versicherungsmathematische“ Logiken Risiken nach bestimmten Gruppenkriterien zusammenfassten, so führen algorithmische Logiken dazu, dass jede Person eine einzigartige Kennung bekommt, auf deren Basis die Kreditwürdigkeit aber im wahren Leben keinerlei soziologisch sinngebende Charakteristiken teilen. Diese abstrakte Form der Beurteilung bildet mathematische und nicht soziale Logiken ab. Dadurch entsteht eine Form der Intransparenz von Kreditentscheidungen, die es für benachteiligte Gruppen noch schwieriger macht als bisher, ihre Diskriminierung nachzuweisen. In der Computer-Vision wird der Mensch sowohl statistisch als auch im übertragenen Sinne Schicht für Schicht auseinandergenommen, allerdings mit anderen Konsequenzen, als Balzac befürchtet hatte. Das Resultat: eine perfide Neuerfindung von visuellen Stereotypen im digitalen Raum, die erstaunlicherweise weniger komplex ist – und damit: noch dümmer – als ihr Pendant aus dem vorherigen Jahrhundert.

Literatur

Barthes, Roland. 1981. Camera Lucida: Reflections on Photography. London: Hill and Wang.

Coleman, Kevin, and Daniel James. 2021. Capitalism and the Camera. New York: Verso.

Lee-Morrison, Lila. 2019. Portraits of Automated Facial Recognition: On Machinic Ways of Seeing the Face. Bielefeld: transcript-Verlag.

Rini, Regina. 2020. "Deepfakes and the Epistemic Backstop." Philosophers' Imprint 20 (24).

Skarpelis, A.K.M. 2022. "What do Computer Vision Engineers Do All Day? On the Making of Ethnoracial Categorization in Computer Vision Practice."

—. forthcoming. "Horror Vacui: Racial Misalignment, Symbolic Repair and Imperial Legitimation in German National Socialist Portrait Photography." American Journal of Sociology.

Sontag, Susan. 2008. On Photography. London: Penguin Books.

--

13.12.2022

Zur Bildbeschreibung: Zur Illustration des Beitrags über automatisierte Gesichtserkennung sind wir mithilfe der künstlichen Intelligenz noch einen Schritt weiter gegangen, denn diese Personen gibt es nicht. Ihre Gesichter sind geschaffen von der Fotografin und Bildredakteurin Gesine Born mit Hilfe der Software DALL-E 2. Ihre Eingabe: “portrait photograph of {... description; z.B. woman with a baby}, looking worried, street photography, Leica style, 35 mm, warm colors”. Mehr zu den Fotos finden Sie hier (PDF).

Der Text steht unter der Creative-Commons-Lizenz Namensnennung 4.0 International (CC BY 4.0: https://creativecommons.org/licenses/by/4.0/).