

Norbert Elias und Deepfakes

Vom Sehen in der künstlichen Realität

»Erst als die Natur wie ein dem Menschen gegenüberliegendes Bild angesehen wurde, konnte man auf den Gedanken kommen, sie im Bilde wiederzugeben, den dreidimensionalen Naturraum auf die Bildfläche zu übertragen, den Bildraum perspektivisch zu erweitern und seit dieser Zeit sah man nun, umgekehrt den Naturraum als Bildraum deutend, die Natur als Landschaft.«

Norbert Elias,
»Vom Sehen in der Natur« (2002: 11)

Norbert Elias hat sich bereits in einer seiner ersten Veröffentlichungen mit der menschlichen Wahrnehmung und ihrer Einbettung in soziale Strukturen beschäftigt. Die dort gewonnenen Erkenntnisse und seine späteren Analysen eines Gemäldes dienen als Basis der nachfolgenden Betrachtung einer neuen Art des digitalen Bildes – nämlich jenes, das durch Künstliche Intelligenz (KI) verändert oder vollständig hergestellt wird. Sogenannte *Deepfakes* bezeichnen dabei eine Form, bei der die Darstellung von Menschen und ihren Körpern im Fokus stehen. Deepfakes lassen sich in einen Entwicklungsprozess einordnen, der die generelle Veränderung der Wahrnehmung von Bildern im Alltag beschreibt.

Der Anfang dieses Prozesses kann im Sehen in der Natur gefunden werden, das von Elias als eine von der Erziehung abhängige, den kulturellen Gegebenheiten ihrer Zeit folgende Fähigkeit charakterisiert wird. Es muss erst ein (Selbst-)Bewusstsein über die Individualität des Menschen entstehen, damit letzterer die Natur als etwas von ihm Abgegrenztes wahrnehmen kann. Allmählich erfasst er dann die Landschaft als etwas, in dem sich »Baum und Fluß, Wiese und Berg und Straße und Dorf [...] zu einer bildhaften Einheit ergänzen«. Mit der Entwicklung dieser Betrachtungsweise beginnt der Mensch, die Landschaft in Kunstwerken festzuhalten – schafft es aber vielfach nicht, ihre Einheit darzustellen, sodass »wir doch unzählige Bilder [...] als wertlos verdammen müssen« (Elias 2002: 10). Der Wert der Bilder ergibt sich durch die Prüfung ihrer einzelnen Darstellungselemente im Zusammenhang mit der gesamten Landschaft; wie man umgekehrt die Einzelheiten in der Gesamtheit dieses überwältigenden Kunstwerks der Natur *sehen* kann, muss der Mensch zunächst jedoch lernen.

Die Vermittlung des entsprechenden richtigen Sehens ist Elias zufolge Teil der Erziehung. Diese ist als ein Kreislauf zu verstehen, der aus dem Stellen von Fragen durch die zu erziehenden »Ungelernte[n]«, einer wissenschaftlich begründeten Beantwortung durch den Erzieher¹ sowie daraus erneut entstehenden Fraglichkeiten besteht (vgl. ebd.: 20). Verschiedene wissenschaftliche Disziplinen haben dabei ihre je eigenen Antworten, bleiben aber trotz »aller reinlichen Scheidung dennoch enger oder weiter aufeinander bezogen« (ebd.: 23). Charles Goodwin beschreibt beispielsweise, was das Sehen in der Archäologie bedeutet: Der Boden wird auf besondere Merkmale untersucht, die etwa auf Spuren von menschlicher Architektur hindeuten. Damit Archäologen diese erkennen, nutzen sie spezifische Methoden, mit denen das Gesehene erfasst und kategorisiert werden kann. Für Laien sind diese Kennzeichen solange unsichtbar, bis sie markiert werden. Mit einer Linie im Boden wird das erkannte Objekt hervorgehoben: »this line in the sand has very powerful persuasive consequences« (Goodwin 1994: 610). Mit diesem Vorgehen wird die Wahrnehmung außenstehender Beobachter geleitet und gleichzeitig das Objekt verdinglicht. Durch den Einsatz spezieller Berufspraktiken werden die Entdeckungen auf Karten festgehalten. Diese grafischen Repräsentationen erweitern sprachliche Beschreibungen, indem sie »records of a range of disparate events onto a single visible surface« (ebd.: 611) binden. Innerhalb der Archäologie werden diese Denkmuster und Methoden weitergegeben, sodass erwartet werden kann, dass die meisten Archäologen die Kategorisierungen nachvollziehen können – dass sie also dasselbe sehen.

Dass das Sehen aber nicht immer eindeutig ist, zeigt Goodwin anhand des Gerichtsprozesses im Fall Rodney King. King wurde 1992 in Los Angeles wegen eines Verkehrsverstoßes von mehreren Polizisten gestoppt und in diesem Zusammenhang zum Opfer schwerer Polizeigewalt. Ein Beobachter hat die Szene gefilmt; die Aufnahme wurde im Prozess zum zentralen Beweisstück. »The violence it showed was so graphic that many people assumed that a conviction was almost automatic« (ebd.: 615) – tatsächlich aber wurden die Angeklagten im ersten Prozess freigesprochen, was in einigen Stadtteilen zu aufstandsartigen öffentlichen Unruhen führte. Sowohl die Anklagevertretung als auch die Verteidigung nutzten das Video als Beweisstück. Was sie darin gesehen bzw. wie sie das Material jeweils interpretiert und in der Verhandlung präsentiert haben, widerspricht sich jedoch. Die Verteidigung ließ das Bildmaterial nicht für sich selbst stehen, sondern band die gezeigten Ereignisse im Rahmen ihrer Argumentationsrhetorik in die Berufspraktiken der Polizei ein. Die

1 Aufgrund der besseren Lesbarkeit wird im Text das generische Maskulinum verwendet. Sofern nicht anders gekennzeichnet, sind jedoch alle Geschlechter gemeint.

Ereignisse wurden umgedeutet und bis ins kleinste Detail zerlegt, in etwa so, wie Elias es beim Sehen in der Natur für erforderlich hält. Durch die Hervorhebung der einzelnen Aspekte und deren Einbettung in den Rahmen professioneller Polizeiarbeit verloren die Laienbeobachter der Jury sukzessive das Gesamtbild aus den Augen. Sie ließen sich durch die ihnen vorgetragene Expertise leiten und sahen das Video anders als die Öffentlichkeit. Dasselbe Bildmaterial offenbarte sich ihnen also auf eine andere Weise. Mehr noch: In Verbindung mit der sprachlichen Darstellung wurde die Aufzeichnung sogar zum sichtbaren Beweis für die Unschuld der Polizisten (vgl. ebd.: 620). Erst als die Vertreter der Anklage in einem zweiten, zivilrechtlichen Verfahren eine eigene Perspektive zur Interpretation des Bildmaterials vorstellten, wurden zwei der Beschuldigten schließlich doch verurteilt. Wie in der Archäologie wurden im Rahmen des Gerichtsprozesses sinnbildliche Linien im Video gezogen, die Einzelheiten aus dessen Gesamtheit hervorheben und gezielt ein Objekt, eine bestimmte Vorstellung in der Wahrnehmung der Betrachter erzeugen. Sicherlich hat nicht jeder Betrachter dieselben Eindrücke gewonnen. Auch ließen sich nicht alle der Geschworenen im ersten Prozess von der Darstellung der Verteidigung überzeugen, es müssten also weitere Faktoren zur Beurteilung der individuellen Positionen betrachtet werden. Hier genügt jedoch die Feststellung, dass Uneinigkeit über den Inhalt des sichtbaren Materials erzeugt werden kann: »Reasonable doubt can be created by muddying the water with a plausible alternative.« (Ebd.: 616)

Mit Deepfakes ist heutzutage möglicherweise eine Entwicklungsstufe des Bildes erreicht, die über die schwierige Frage nach der Lesart von Aufzeichnungen hinaus ebenfalls ›berechtigte Zweifel‹ an der generellen Verbindung von Bildern zur Realität nahelegt. In den letzten Jahren wird der Begriff immer häufiger (insbesondere medial) aufgegriffen, dabei wird aber selten genauer bestimmt, was Deepfakes überhaupt sind. Man kann sie etwa als »die Manipulation oder auch synthetische Erzeugung von Abbildungen, Videos und/oder Audiospuren menschlicher Gesichter, Körper, oder Stimmen, zumeist mithilfe von KI« (Pawelec/Bieß 2021: 23) definieren.

Eine verbreitete Variante eines solchen synthetischen Mediums ist ein *Faceswap*, ein Gesichtertausch unter Zuhilfenahme einer KI. Hierbei extrahiert das System² die Eigenschaften eines Gesichts aus einem oder mehreren Bildern und überträgt sie auf das Zielbild einer weiteren Person.³ Dafür stehen spezifische Webtools und Apps zur freien Verfügung,

- 2 Als ›System‹ ist hier die KI selbst als trainiertes Modell zu verstehen, das man über ein Benutzerinterface auf einem Computersystem für die vorgesehenen Aufgaben einsetzen kann.
- 3 Ein solcher Tausch ist auch beispielsweise mit Tieren, Gegenständen oder Zeichentrickfiguren möglich, was in diesem Rahmen aber nicht näher betrachtet werden soll.

mit denen man relativ einfache Deepfakes erstellen kann; sie sind häufig noch ohne tiefere Analyse anhand von Bildfehlern (wie etwa unpassender Farbgebung, falschen Schatten oder Unstimmigkeiten in der Perspektive) zu erkennen.



Abb. 1 und 2: Faceswaps der Portraitfotos von Norbert Elias und Thorsten Benkel (erstellt mit Face Swapper von Icons8)

Abb. 1 und 2 zeigen solche Bilder: In Abb. 1 wird das Gesicht von Thorsten Benkel in ein Portrait von Norbert Elias übertragen, in Abb. 2 sind Ursprungs- und Zielbild vertauscht.⁴ Die KI fügt dabei nicht nur die Gesichter den Körper(andeutunge)n hinzu, sondern gleicht auch die Farbgebung an und verändert bzw. entfernt sogar die Brille aus Elias' ursprünglichem Bild. Erst bei genauerer Betrachtung erkennen Uneingeweihte einige Unstimmigkeiten, die die Fälschung vermuten lassen (in Abb. 1 etwa die Einfärbung einzelner Pixel bei einem sonst in Graustufen erscheinenden Foto und in Abb. 2 die Proportionen von Gesicht und Kopf). Diese Form der Deepfakes kann sogar am Smartphone erstellt werden, nicht zuletzt deshalb, weil es bereits für diese Aufgaben trainierte KI-Modelle gibt, die frei zugänglich sind. Die Ergebnisse variieren qualitativ recht stark und sind, wenn man es darauf anlegt, meistens schnell zu enttarnen. Aufgrund der Art, wie ein Netzwerk zur Erstellung von Deepfakes aufgebaut wird, findet man vor allem manipulierte Bilder von Personen des öffentlichen Lebens. Für komplexere Deepfakes mit weniger leicht ersichtlichen Bildfehlern, gibt es Open Source-Quellcodes

- 4 Die Bilder wurden vom Autor mit Hilfe des Face Swapper von Icons8 (2023) erstellt. Dieses System funktioniert direkt im Browserfenster und gibt innerhalb von Sekunden Ergebnisse aus. Das Unternehmen bietet zudem KI-unterstützte Programme an, mit denen Bilder anonymisiert, optimiert, Bildhintergründe verändert sowie Fotos oder Gesichter vollständig künstlich generiert werden können.

sowie vortrainierte Modelle, deren Einsatz aber Programmierkenntnisse und Einarbeitungszeit erfordern.

Um insbesondere bei bewegten Bildern glaubhafte Ergebnisse zu erreichen, die einer detaillierten Prüfung standhalten, ist zudem deutlich mehr Ausgangsmaterial als in den gezeigten Beispielen erforderlich. Während die einfacheren Faceswaps üblicherweise der Unterhaltung dienen, ist den aufwändigeren Synthetika, die sich nicht mehr als solche enttarnen lassen, ein großes Risiko inhärent. Das fingierte Bild kann als ein sichtbarer Beweis für etwas gesehen werden, das in der abgebildeten Form nie existiert hat. Es kann potenziell missbräuchlich genutzt werden, etwa durch die Generierung pornografischen Bildmaterials einer Person, die tatsächlich nie in einer solchen Situation aufgenommen wurde. Sicherlich kann auch dieses Material der einvernehmlichen Unterhaltung dienen, es kann aber ebenso zur gezielten Diskreditierung einer Person eingesetzt werden. Die Wahrnehmung des Betrachters wird jedenfalls gelenkt; er sieht eine bildhafte ›Realität‹, die durch eine KI erzeugt wurde. Es wird somit ein Objekt geschaffen, dessen Existenz nicht nur wie in Goodwins Beschreibung archäologischer Objektkonstruktionen über Linien im Sand suggeriert wird, sondern auf einer »single visible surface« (Goodwin 1994: 611) deutlich zu sehen ist.

Stellen synthetische Medien wie Deepfakes eine Evolution des Bildes dar, die das Potenzial hat, zu verändern, wie der Mensch die Realität ausdrücklich *sieht* (nicht: wahrnimmt)? Zur Lebzeit von Elias gab es eine solche Technologie noch nicht, aber dennoch lassen sich manche seiner Überlegungen für die soziologische Einordnung der Deepfake-Thematik fruchtbar machen. In dem schon zitierten Erstlingsaufsatz schreibt er beispielsweise: »Beinahe jeder Künstler [...], zum mindesten jede Kulturepoche der neueren Zeit hob eine besondere Landschaft aus der Fülle des Natürlichen heraus und pflegte ihre Vorstellung mit besonderer Liebe, weil sie der besonderen kulturellen Haltung am meisten entsprach.« (Elias 2002: 25) Wie die Wirklichkeit gesehen wird, ist abhängig von den kulturellen Gegebenheiten der Zeit, und diese stehen permanent im Wandel. Ein neutrales Sehen ist unmöglich, denn der Einsatz der Sinne ist stets in soziale Strukturen eingebunden (Loenhoff 2001). Mit den Deepfakes existiert vielmehr eine neue Art des Bildes: Das durch KIs veränderte oder vollständig erstellte Bild – bzw. hier im Speziellen: die glaubhafte Darstellung von Menschen in verschiedenen Szenen, die so nie stattgefunden haben – bereichert die Menge der vorhandenen Abbildungen um nur mehr lose an Wirklichkeitsaspekte gekoppelte Artefaktbilder. Welche Rückschlüsse lässt dies nun auf die kulturelle Haltung der Gegenwartsgesellschaft zu, wenn nicht nur Künstler den menschlichen Körper aus der Fülle des Natürlichen herausnehmen und nach ihren Vorstellungen gestalten können? Und was bedeutet dies für die traditionelle Kunst der Wirklichkeitsimitation, wie sie in der Ästhetik seit Jahrtausenden verfolgt wird (dazu: Benkel 2008)?



Abb. 3: Antoine Watteau: *Pèlerinage à l'Île de Cythère*. 1719. Schloss Charlottenburg, Berlin.

Elias selbst hat ein spezifisches Bild, in diesem Fall als ein klassisches Gemälde verstanden, genau untersucht und dessen Bedeutung in den Kosmos seines theoretischen Denkens gerückt. Unter dem Titel *Watteaus Pilgerfahrt zur Insel der Liebe* (Elias 2000) betrachtet er das gleichnamige Werk von Antoine Watteau, das heute als *Pèlerinage à l'île de Cythère* (*Die Einschiffung nach Kythera*) bekannt ist und in insgesamt drei Versionen existiert, die im frühen 18. Jahrhundert entstanden sind.⁵ Elias untersucht die Bedeutung dieses Gemäldes im Zeitverlauf und erkennt dabei schon in der mehrmaligen Veränderung des Titels eine Wandlungsfacette. Das Bild steht für etwas, es verweist auf reale Vorkommnisse und Ereignisse, die mithin verfremdet werden; und unterschwellig zeigt es soziale Strukturen an. Wie die Benimmbücher, die Elias im *Prozess der Zivilisation* so ausführlich diskutiert hat (Elias 1976), ist auch Watteaus Gemälde ein Dokument, das Informationen über das Alltagsleben

- 5 Als Abb. 3 ist hier das letzte der drei Bilder zu sehen; eine frühere Fassung mit dem Titel *L'Embarquement pour Cythère* gehört zum Bestand des Louvre in Paris. Elias' Rede vor dieser Version des Werks im Schloss Charlottenburg gab Anlass zum Verfassen seines Texts. Hermann Korte (2013: 69) berichtet, dass der zu dieser Zeit bereits fast erblindete Elias das Gemälde zunächst für eine Fälschung hielt, da bestimmte Details nicht passten; er erinnerte dabei jedoch das Gemälde im Louvre, das den beiden anderen sehr ähnlich ist. Der Irrtum konnte aufgeklärt werden, und Elias bezieht bei seiner Untersuchung folglich alle drei Werke mit ein.

der Menschen beinhaltet. Zugleich bildet es die vorherrschenden ästhetischen Strategien der Zeit ab: »[I]n der Entwicklung des Kunstgeschmacks [findet der] Wandel der Machtstrukturen seinen Ausdruck« (Elias 2000: 27).

Das Gemälde stellt konkret eine Reisegruppe dar, die sich im Aufbruch zu einer Schiffsreise in Richtung der griechischen Insel Kythera befindet. Dort werden den Göttern Opfer entrichtet, man geht aber auch dem nach, was heute als Tempelprostitution bezeichnet wird (vgl. ebd.: 12). Es handelt sich gemäß der Bilddarstellung um eine sagenhafte Insel der Freuden, an die utopische Vorstellungen geknüpft werden, die nach der beschwerlichen Fahrt auf die Reisenden zu warten versprechen. In späteren Beschreibungen ändert sich diese positive Interpretation jedoch zeitweise ins Negative; Elias spricht von einem Pendel, das mit der Zeit immer wieder in die entgegengesetzte Richtung ausschlägt (vgl. ebd.: 30). Er erkennt, dass sich seit der Erschaffung des Werks das Verhältnis von Künstlern und Publikum gewandelt hat. Insbesondere hat sich verändert, wer Künstler werden kann und wie Kunst interpretiert wird.⁶ Das Gemälde von Watteau hat sich in seiner Form über Jahrhunderte hinweg nicht verändert, aber wofür es stehen soll, wie man es *sieht*, und was man folglich darin erkennt, unterliegt, wie die Gesellschaft, in der es besteht, einem permanenten Wandel. Das Bild zeigt zunächst eine Liebesinsel, wie sie vermutlich auch in anderer zeitgenössischer Kunst verarbeitet wird. Schnell ändert sich die Deutung aber hin zu einem Fest des Aufbruchs, das die Freude über eine neue Ära nach dem Tod des alten Königs in Frankreich zum Ausdruck bringt. Bis zu diesem Zeitpunkt wird das Gemälde positiv aufgefasst. Man sieht darin einen Rückzug aus der politisch angespannten Lage, bis es mit der Revolution zu einem Politikum und nun negativ gelesen wird. Aufgrund seiner Bedeutungsverankerung im Ancien Regime wird es aus der Öffentlichkeit verbannt. Nach der revolutionären Zeit schlägt das Pendel abermals um: Das Gemälde geht verloren und wird wiederentdeckt; nun gibt es eine Stimmung der Sehnsucht nach der vergangenen Zeit, die man etwas später wieder vergessen möchte. Mit der Transformation der Machtverhältnisse wechselt, wie bereits angedeutet, auch der Kunstgeschmack. Der Umschwung »von der Vorherrschaft von Wunschtraumutopien zu der von Furchttraumutopien und [sein Zusammenhang] mit Struktureigentümlichkeiten der Gesellschaftsentwicklung« (ebd.: 45) wird deutlich. Die Kunst des Rokoko erfreut sich im Laufe des 19. Jahrhunderts zwar neuer Begeisterung, die Idee der Utopie wird aber schnell wieder entzaubert. Für Elias ist hierfür ein Erlebnis des Schriftstellers Gérard de Nerval bezeichnend: Er träumt

6 Pierre Bourdieu, der sich am Beispiel des Malers Édouard Manet eingehend mit der entsprechenden Thematik befasst, versteht einen solchen Wandlungsprozess als eine »symbolische Revolution« (Bourdieu 2015).

von den galanten Festen, der Liebeslust und den Freuden auf der Insel, die im Bild dargestellt sind, und reist schließlich dorthin. »Was er fand, war eine kahle, hässliche Insel, die nun unter britischer Oberherrschaft Cérigo hieß. Was ihm entgegenstarrte, waren nackte Felsen und als Zeichen der menschlichen Grausamkeit ein dreiarmiger Galgen. An einem der Arme hing eine Leiche.« (Ebd.: 46)

Diese Anekdote zeigt zum einen, wie die vorherrschenden Verhältnisse das beeinflussen, was Menschen in dem Gemälde sehen. Sie verdeutlicht zum anderen aber auch die Diskrepanz zwischen der vorgeprägten Vorstellung und der Wirklichkeit. Über das Medium des Bildes erschafft der Betrachter eine eigene Realität, die auf kognitiver Ebene die eigene Prägung und Sozialisation widerspiegelt. Dennoch werden Bilder – zumindest in Form von Foto- und Videografien – gemeinhin als präzise Abbilder der Wirklichkeit wahrgenommen. Damit besteht zugleich das Risiko der Bildmanipulation – ein historisch mehrfach verwendetes Manöver. Fingierte Bilder haben Macht, indem sie die subjektive Realität der Betrachter in einer Weise prägen, die nicht der objektiven Wirklichkeit entspricht, obwohl genau dies – die Beziehung zwischen einem Akteur und einem von ihm selbst unabhängigen Ausschnitt der Realität – impliziert wird. Das sogenannte »Thomas-Theorem« besagt, dass die subjektive Definition einer Situation als Realität ausreicht, um tatsächlich reale Handlungen auszulösen (vgl. Thomas/Thomas 1928: 572). Demzufolge können gefälschte Bilder das Handeln leiten und die Fälscher Einfluss auf die Wirklichkeit menschlicher Handlungen nehmen.

Heute denkt man im Alltag beim *Bild* nicht mehr zuerst an ein Gemälde, sondern vielmehr an eine Fotografie.⁷ Solchen Bildern unterstellt man stärker, dass sie eine objektive Wahrheit festhalten, und man nutzt sie aufgrund der ihnen zugeschriebenen Aussagekraft in vielerlei Kontexten als einen für sich selbst stehenden Beweis. Dass es sich dabei indes nicht um eine exakte Repräsentation der Wirklichkeit handeln kann, wird bereits aufgrund ihrer Zweidimensionalität und der notwendigen Begrenzung auf ganz bestimmte Ausschnitte einer Szene evident. Bei der Erfassung des »dreidimensionalen Naturraum[s] auf [der] Bildfläche« (Elias 2002: 11) gehen Informationen aus der Gesamtheit verloren. Bilder umgeben Menschen im Alltag permanent und sind prägend für die Konstruktion von Wirklichkeit.⁸ Dabei sind die Möglichkeiten ihrer Manipulation vielfältig. Sie existieren, seit Künstler damit begonnen haben,

- 7 Der Stellenwert der Fotografie in der Soziologie Bourdieus wird in Bourdieu 2023 näher beleuchtet.
- 8 Zur Bedeutung des Bildes als Abbildung einer vergangenen »objektiven« Realität, die scheinbar unveränderlich festgehalten wird, siehe Benkel 2022: 35ff. Foto- und Videografien dienen beispielsweise bei Hinterbliebenen als »materielle Erinnerungsanker, unbeweglicher, dafür auch unbestechlicher als persönliche Erinnerungsinhalte« (ebd.: 37).

ihre eigenen Werke zu signieren sowie ihr Kennzeichen unter Werke ihrer Schüler zu setzen, um sie besser verkaufen zu können (vgl. Deussen 2007: 30). Elias betrachtet die Veränderung der Wahrnehmung von Bildern aufgrund von strukturellen Wandlungsprozessen, nicht aber die Veränderung von Bildern selbst. Dabei ist es bereits im 20. Jahrhundert verbreitet gewesen, durch die Manipulation von Bildern eine bestimmte Lesart induzieren zu wollen. Elias (2000: 46f.) stellt fest, dass in bestimmten Staatsgesellschaften die Machtstrukturen den Kulturproduzenten – Künstlern, Schriftstellern aber auch Wissenschaftlern – eine bestimmte Vorstellung bzw. Darstellungsweise ihrer Gegenstände vorgeben. Eine Abweichung wird hierbei »als Gefahr für die bestehende Ordnung empfunden« (ebd.: 47). Kulturgüter sind somit stets im Kontext bestimmter Machtstrukturen zu sehen, ihnen wird aber auch das Potenzial zugeschrieben, diese verändern zu können. Dabei sind sie möglicherweise nicht die Ursache eines Wandels, sondern zeigen einen bereits stattfindenden Wandel an. Elias erinnert in diesem Kontext an die fehlende Neutralität bei der Betrachtung von Bildern: »[J]edes Kunstwerk, also auch jede gemalte oder literarische Utopie, mit künstlerischen Funktionen [kann] zugleich auch, aktuell oder potentiell, ideologische Funktionen haben« (ebd.: 33).

In der Geschichte der Fotografie gibt es zahlreiche Fälle, in denen Bilder manipuliert werden, um eine bestimmte Botschaft zu vermitteln. Insbesondere in der Sowjetunion könnte man die Bearbeitung von Fotografien als eine eigene Kunstform bezeichnen, die »der systematischen Verfälschung der Geschichte selbst« (King 1997: 7) dienen sollte und einige berühmte Beispiele hervorgebracht hat. So wurden etwa die später als politische Gegner wahrgenommenen Leo Kamenew und Leo Trotzki aus einem Gemälde von ihrem Platz neben Lenin auf dem Rednerpult entfernt (mehr zu diesem und anderen Beispielen ebd.: 67ff.). Dies geschah noch in vielen weiteren Fällen und diente der *damnatio memoriae* – der »Verurteilung des Andenkens [...] als Sanktion für ein besonders verabscheuungswürdiges Verhalten« (Ries 2001: 238), also dem Versuch, die Erinnerung an unliebsame Personen aus der Geschichte zu tilgen. Die Fälscher arbeiteten zu dieser Zeit noch mit Tusche und Pinsel am Original, und man benötigte Experten, die die Bearbeitung vornahmen (vgl. Jaubert 1989: 10ff.). Fälschungen dieser Art sind aufwändig, ihre Anzahl daher begrenzt. Die zugrundeliegenden Methoden werden aber bis heute angewandt und entwickeln sich gewissermaßen parallel zur Fotografie. Nach dem Massaker von Luxor 1997 hat etwa die Schweizer Boulevard-Zeitung *Blick* eine Wasserpflütze in der Aufnahme des Tempels rot eingefärbt, sodass der Anschein einer Blutlache entstand. Mit der noch einfacheren Methode des Ausschneidens verschiedener Bildteile verändert die Künstlerin Ursula Dahmen (siehe *Abb. 4*) die Grundaussage des Bilds zweier US-amerikanischer und eines irakischen

Soldaten völlig: Das eine Bild legt die Interpretation von Hilfeleistung nahe, während das andere das Drohen mit einer Waffe impliziert. Durch den Einsatz neuer digitaler Technologien erfordern solche alten Methoden der Retusche, Abhebung, des Ausschneidens oder der Vergrößerung von Ausschnitten deutlich weniger Expertise. Die Veränderungen am Bild können so sehr schnell vorgenommen werden und bei ungewolltem Anblick ebenso rasch wieder rückgängig gemacht werden, bis das gewünschte Resultat vorliegt.



Abb. 4: Fotomontage eines irakischen Soldaten, umgeben von US-Militärs (Ursula Dahmen/AP, nach: Der Spiegel 2008)

Glaubhafte Ergebnisse zu erzeugen, bedingt dennoch Übung im Umgang mit den verfügbaren Werkzeugen. Die Unterstützung durch KIs bei der Bearbeitung von Bildern kann die Anforderungen weiter reduzieren, indem dieser Prozess der inkrementellen Veränderung dem Nutzer abgenommen und automatisiert wird. Fälschungen dieser Art benötigen kein Expertenwissen, sodass sie nicht mehr, wie in der Sowjetunion, auf ideologisch ›bedeutungsvolle‹ Bilder begrenzt bleiben, sondern vielmehr bei jeder alltäglichen Aufnahme vorgenommen werden können. Deep-fakes sind dabei ein spezieller und relativ junger Bereich von KI-veränderten Bildern. Der Begriff setzt sich schließlich aus *Deep Learning* und *Fake* zusammen; Deep Learning bezeichnet dabei die Art der KI, die auf einem künstlichen Neuronalen Netzwerk basiert. Ein solches Netzwerk umfasst eine hohe Anzahl künstlicher Neuronen, die in mehreren Schichten angeordnet sind. Zwischen den einzelnen Schichten nimmt das Netzwerk verschiedene Operationen vor, um die ihm gegebenen Daten zu

verarbeiten. Eine KI dieser Art ist nicht klassisch programmiert, sondern trainiert (vgl. Chollet 2018: 23): Auf der Grundlage eines dem System übergebenen Datensatzes und fixierter Zielsetzungen lernt die KI über viele Trainingsiterationen selbstständig, welche Verarbeitungsschritte sie benötigt, damit die gestellte Aufgabe bestmöglich ausgeführt werden kann. Für den Nutzer sind lediglich die Ein- und Ausgabeschichten ersichtlich, das eigentliche Netzwerk bleibt eine Black-Box (dazu näher Spiegel 2022: 33ff.).⁹

Mit dieser Technologie können Bilder vollständig synthetisch erzeugt werden. Abhängig vom Datenmaterial kann etwa ein täuschend echtes Gemälde im Stil eines Künstlers produziert werden, das dieser selbst nie gesehen, geschweige denn geschaffen hat (vgl. Volland 2018: 12). Der US-amerikanische Technologiekonzern NVIDIA zeigt, dass es ebenso möglich ist, vollständige Portraits von nicht existierenden Menschen zu erzeugen (vgl. Spiegel 2022: 109). In einer Datenbank sind mehrere Millionen solcher Bilder sowie der zugrundeliegende Code für jeden zur Nutzung bereitgestellt (*Abb. 5* zeigt einen Ausschnitt) und in einem Generator können neue Bilder nach gewünschten Parametern erzeugt werden (Generated Media 2019). Die Personen existieren nicht, können aber dennoch beispielsweise in der Werbung auftreten, ohne dass der Betrachter ihre ›künstliche Natur‹ erkennen kann. Auf Basis des gleichen Modells demonstrierte die University of Washington, dass es bei einer Gegenüberstellung oft schwer ist, die Fälschungen von einem Original zu unterscheiden (West/Bergstrom 2019). Mittlerweile werden solche Netzwerke auch im Bereich der Pornografie eingesetzt, um beliebige Personen oder Fantasiewesen in ebenso beliebigen Situationen zu generieren.¹⁰ Der

9 Beim Programmieren werden alle Verarbeitungsschritte bereits im Voraus festgelegt; der Programmcode wird vollständig durch den Programmierer geschrieben. Beim Deep Learning ist lediglich der ›Rahmen‹ des Systems auf diese Weise erzeugt; welche statistischen Berechnungen zum Füllen dieses Rahmens nötig sind, ›entscheidet‹ bzw. lernt die KI über viele Versuche selbst. Der Blick in diese Art ›Black Box‹ ist zwar möglich, aber die Bedeutungen der einzelnen Werte erschließen sich dem Menschen (noch) nicht. So erklärt sich auch, warum, wie bereits angeführt, aufwändigere Deepfakes Programmierkenntnisse erfordern, obwohl trainierte Netzwerke frei verfügbar sind. Sie müssen in eine programmierte Umgebung implementiert werden, die dem Nutzer beispielsweise den Zugriff über ein Interface ermöglicht.

10 So nutzt etwa die Website Pornpen.ai (Pornpen 2023) nach eigenen Angaben speziell angepasste Algorithmen, mit denen jederzeit Bilder erzeugt werden können. Der Generator operiert mit erwachsenen Personen jeden Geschlechts – der permanente Live Feed der Website zeigt aber überwiegend Bilder von Frauen, eine Tendenz, die auch bei den Faceswaps sehr deutlich zu erkennen ist.

hier gesetzte Fokus liegt jedoch auf der Manipulation bereits bestehenden Bildmaterials; die vorgestellten Werkzeuge können dabei nicht nur zur Produktion synthetischer Bilder genutzt werden, sondern auch der Veränderung von Videomaterial dienen.

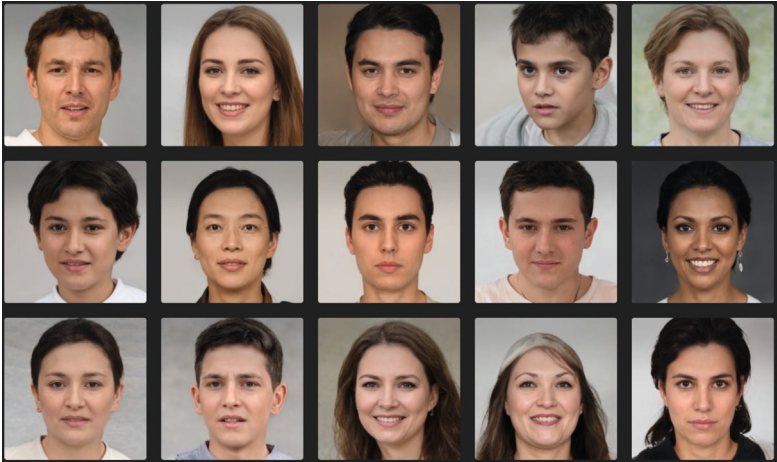


Abb. 5: Generated Media: 2.675.749 von einer künstlichen Intelligenz generierte Bilder (Generated Media 2023)

Manipulationen mit einfacheren Bearbeitungswerkzeugen können bereits trotz ihrer teils niedrigen Qualität einen großen Einfluss auf die Wahrnehmung des Betrachters haben (vgl. Ajder et al. 2019: 11). Die oben beschriebenen Faceswaps auf Basis von Deep Learning-Algorithmen sind eine Form von komplexeren KI-Fälschungen, deren Anteil unter den Deepfakes aber am größten ist (vgl. Pawelec/Bieß 2021: 38f.). Andere Formen sind *Lipsyncs*, bei denen die Gesichtsbewegungen einer Person an eine andere Audiospur angepasst werden. Diese kann eine existierende Tonaufnahme sein oder vollständig künstlich hergestellt werden und die Stimme einer gewünschten Person imitieren. Beim *Facial Reenactment* werden die Muskelbewegungen eines Schauspielers aufgezeichnet und mit einer anderen Aufnahme verbunden. Beim *Puppeteering* wird dasselbe mit den Körperbewegungen eines Menschen gemacht; das Dalí Museum in Florida hat dies gezeigt, indem Salvador Dalí im Bildschirm »wiedererweckt« wurde und eine Führung durch das ihm gewidmete Museum gab (Dalí Museum 2019). Dies ist jedoch lediglich eine Demonstration der technischen Möglichkeiten in einem Kunstprojekt. Der Filmregisseur Jordan Peele zeigt im Video darüber hinaus eine machtvollere Anwendungsmöglichkeit: Mit seinem Deepfake von Barack Obama macht er darauf aufmerksam, dass man Bildern aus dem

Internet nicht mehr trauen dürfe – denn jeder könne im Video dazu gebracht werden, alles zu sagen. Man müsse lernen, mit Quellen umzugehen, um nicht in einer »fucked up dystopia« (Peele/BuzzFeedVideo 2018) zu enden. Deepfakes können ferner genutzt werden, um demokratische Prozesse zu beeinflussen, weswegen ihnen ein großes Bedrohungspotenzial innewohnt (etwa durch verfälschte Auftritte politischer Kandidaten, um Wahlen zu beeinflussen oder um das Vertrauen in politische Institutionen zu untergraben; vgl. Farid/Schindler 2020: 24ff.). Manipulationen müssen dabei nicht immer durch den Einsatz von KIs erfolgen, aber glaubhafte KI-Veränderung von Bildern bis hin zur Echtzeit-Videoübertragung sind bereits möglich. Dabei wird deutlich, worauf Peele hinweist: Durch die bloße Existenz der Technologie kann eine generelle Unsicherheit dem Bild gegenüber entstehen; niemand kann sich sicher sein, ob es sich um eine echte oder eine manipulierte Aufnahme handelt.

Experten der Digitalen Forensik nutzen verschiedene Methoden, um Deepfakes zu enttarnen. Sie reichen von der manuellen Überprüfung auf Bearbeitungsspuren über verschiedene Low-Level-Ansätze, bei denen vor allem Bildartefakte erkannt werden sollen, die bei der Synthese einer Fälschung entstehen, bis zu High-Level-Ansätzen, die semantische Merkmale in den Fokus nehmen (vgl. ebd.: 18ff.). Diese Prüfungen können ebenfalls durch KIs unterstützt werden. Für den Laien ist es im Alltag heutzutage kaum noch möglich, hochwertige Manipulationen zu erkennen. Zudem befassen sich mehr Forscher mit der Erstellung von Deepfakes und der Weiterentwicklung der Technologie als mit dem Erkennen der Manipulationen (vgl. ebd.: 20). Die wenigen entwickelten Techniken zur Erkennung dienen den Fälschern letztlich immer wieder bei der Verbesserung der eigenen Systeme. Bei sogenannten Generative Adversarial Networks wird dieses Prinzip sogar noch erweitert, indem sich zwei KIs gegenseitig trainieren: Die erste KI erzeugt ein Bild, das die zweite auf seine Echtheit untersucht. Die Ergebnisse dieser KI werden wiederum genutzt, um die Synthetika des Ersten zu verbessern (vgl. Pawelec/Bieß 2021: 37). Durch eine Vielzahl an Iterationen optimiert sich das Netzwerk somit selbst immer weiter. Die Fälschungen werden immer besser, die Systeme sind immer einfacher einzusetzen und die benötigte Rechenleistung dafür nimmt ab, sodass theoretisch jeder Deepfake erstellen kann – sie aber kaum noch jemand sicher erkennt. Hinzu kommt, dass Bilder und Videos selbst bei der Feststellung einer Manipulation weiter im Umlauf sind und die Wahrnehmung bzw. Beurteilung der Betrachter beeinflussen.

Elias bezieht sich auf das Verhalten von Menschen, auf Abhängigkeitsgeflechte und das Ausformen eines Über-Ichs entgegen den Trieben sowie zum Teil auch auf die Veränderung der Physis im Zusammenhang mit der Psyche, wenn er schreibt, dass »der Mensch [...] ein außerordentlich modellierbares und variables Wesen [ist]« (Elias 1982: 377). Im Kontext

der Deepfakes erhält diese Äußerung noch eine weitere Bedeutung: Der Mensch wird in der Realität der Bilder zu einem frei modellierbaren Wesen. Der Ursprung des Begriffs Deepfake weist auf den größten Anwendungsbereich dieser Technologie hin (vgl. Ajder et al. 2019: 1): 2017 veröffentlicht ein Nutzer auf der Plattform *Reddit* manipulierte Videos mit pornografischem Inhalt von Schauspielerinnen und zugleich den Code, mit dem er diese erstellt hat. (Sein Nutzernamen *deepfakes* wurde im anschließenden Diskurs für das Phänomen übernommen; vgl. Pawelec/Bieß 2021: 15f.). Wie Elias konstatiert, trägt die Zivilisationsentwicklung dazu bei, dass der nackte Körper im Alltag an Selbstverständlichkeit verliert. Dafür aber »gewinnt die Darstellung des nackten Körpers in der Kunst eine neue Bedeutung: Sie wird in stärkerem Maße als bisher Traumbild und Wunscherfüllung« (Elias 1981: 225). Traumbild und Wunscherfüllung wiederum können mit Deepfakes grafisch umgesetzt werden, indem vor allem durch Faceswaps realistisch wirkende pornografische Bilder und Videos erzeugt werden, ohne dass die Betroffenen davon wissen, damit einverstanden sind – oder jemals nackt aufgenommen wurden. In den Foren können überdies Wünsche zu Zielpersonen und Zielvideos geäußert werden.¹¹ Die Schöpfer der Deepfakes entscheiden im Anschluss, ob sie die Nutzerwünsche verwirklichen wollen; Regeln geben sie sich dabei nur selbst.¹² Abb. 6 zeigt einen (bereits mehrere Jahre alten) Ausschnitt eines Videos, bei dem das Gesicht der Schauspielerin Natalie Portman auf den Körper der Pornodarstellerin Tori Black gesetzt wurde – in den Kommentaren wird die Qualität der Fälschung beispielsweise folgendermaßen gelobt: »Jesus, this one is scary good.«

Theoretisch ist alles Abbildbare möglich, de facto sind jedoch Grenzen durch das verfügbare Datenmaterial gesetzt. Die KI muss auch hier zunächst trainiert werden. Dafür benötigt das Netzwerk möglichst viele Aufnahmen der Zielperson aus verschiedenen Perspektiven und bei verschiedenen Lichtverhältnissen; je mehr Ausgangsmaterial zur Verfügung

11 So finden sich beispielsweise unter einer einschlägigen Deep-Fake-Webseite Anfragen zu verschiedenen Personen des öffentlichen Lebens (MrDeepFakes 2023a). Das Forum bezeichnet sich auf der Website selbst als die größte aktive Deepfake Community und verspricht, dass sich dort alles in Bezug auf Deepfakes finden lässt, inklusive Videos sowie Anleitungen zur Nutzung und Entwicklung der KIs.

12 Im angeführten Fall der MrDeepFakes-Foren sind keine Anfragen zu minderjährigen Prominenten möglich; über eine Blacklist wird der Umgang mit gerade erst volljährig gewordenen Personen geregelt. Ferner ist auch der Umgang mit nicht-prominenten Personen nicht zugelassen (der Status wird Persönlichkeiten des öffentlichen Lebens mit einer bestimmten, in »followern« quantifizierten Reichweite auf sozialen Medien ab- bzw. zuerkannt). Außerdem sind die Bereiche pornografischer und nicht-pornografischer sowie bezahlter und kostenloser Anfragen voneinander getrennt.

steht, desto besser lernt die KI die Kennzeichen der Zielperson und kann sie genauer modellieren. Die auszutauschende Person im Zielvideo muss dabei möglichst ähnliche Körper- und Gesichtsproportionen haben, damit das Ergebnis glaubhaft erscheint. So erklärt sich auch, warum Deepfakes vor allem von Personen des öffentlichen Lebens, insbesondere von Schauspielerinnen, Politikerinnen oder Musikerinnen, erstellt werden: Es gibt schlicht mehr Datenmaterial von ihnen. Die weibliche Sprachform wurde gezielt verwendet, da insbesondere im Kontext der Pornografie überwiegend Frauen betroffen sind (vgl. Ajder et al. 2019: 2; Pawelec/Bieß 2021: 53f.).



Abb. 6: Deepfake mit dem Gesicht von Natalie Portman (MrDeepFakes 2023b)

Neben Faceswaps ist das Entfernen von Kleidung aus Bildern und die Generierung des nackten Körpers der vormals bedeckten Stellen (vgl. Ajder et al. 2019: 8) eine weitere Möglichkeit, nicht einvernehmliches (pornografisches) Material herzustellen. Solche Bilder und Videos können sehr realistisch erscheinen; kennt man die Zielperson nicht und vermutet man keinen KI-Zusammenhang, so entstehen bei oberflächlicher Betrachtung keine Zweifel an der Echtheit einer Szene. Um eine solche Qualität zu erreichen, müssen bislang noch einige Voraussetzungen erfüllt werden. So benötigt das Training der KI große Mengen an Daten, der Nutzer muss über grundlegende Kenntnisse im Programmieren verfügen und leistungsfähige Hardware besitzen. Die Fortschritte in der recht kurzen Zeit seit dem Aufkommen der Deepfakes sind jedoch beeindruckend, und es kommt hinzu, dass die Anforderungen an Technik und Nutzer immer weiter sinken.

Aus Elias' Perspektive lassen sich einige Schlussfolgerungen zu dieser Entwicklung ziehen: »So wird es also auch notwendig sein, daß die

Erfassung der Natur, gleich der jeden anderen Kulturgutes, in den Bereich der Erziehung tritt.« (Elias 2002: 15) Elias betrachtet nicht nur Kunstwerke, sondern bezieht sich auch auf andere Kulturgüter, die sich in die sozialen Strukturen ihrer (Entstehungs-)Zeit einordnen lassen (vgl. Elias 2000: 35). Deepfakes können als solche bezeichnet werden, selbst wenn man sie nicht im Bereich der Kunst verorten möchte; die Erkenntnisse aus Elias' Analyse von Watteaus Gemälde können diesbezüglich genutzt werden. Im weitesten Sinne befasst sich das Gemälde inhaltlich sogar mit dem gleichen Thema: Die fiktionale Vorstellung von Sexualität auf der Liebesinsel in Form von Tempelprostitution im Gemälde steht der fiktionalen Sexualität in der Pornografie bzw. in Deepfakes gegenüber. Es sind Wunschbilder und Traumvorstellungen derer, die sie erstellen (hierbei sind über nicht-einvernehmliche Manipulationen hinaus auch bewusste Veränderungen des eigenen Körpers im Bild einzubeziehen) oder, wie beispielhaft anhand der Foren gezeigt, für spezifische Personen nachfragen. Sie haben wenig mit der Wirklichkeit gemein, obwohl sie selbst – als Bilder – real sind.

Zudem untersucht Elias primär die Oberschicht und den Adel. Auch bei Deepfakes ist das heutige ›Adelsäquivalent‹ in Form von Politikern und Personen aus der Populärkultur betroffen.¹³ Es ist vorstellbar, dass sich dies von oben nach unten durchsetzt, wenn die Systeme weiterhin dem Trend folgen, dass die Anforderungen an Datenmaterial, Hardware und Fähigkeiten abnehmen. Eine breite Anwendungsmöglichkeit kann man in Form von »Deepfake-Rachepornografie« (Pawelec/Bieß 2021: 55) sehen, die nach gescheiterten Beziehungen oder generell mit dem Ziel angefertigt werden, die Reputation einer Person zu beschädigen. Über das Bild lassen sich Machtstrukturen erkennen, die einzelne Personen betreffen: Die Zielperson eines Deepfakes wird zu einer Reaktion genötigt. Sie ist dazu gezwungen, die Öffentlichkeit darüber aufzuklären, dass ein Video nicht echt ist, Schaden verursachen kann und gelöscht werden muss (Goldberg 2019). Schauspielerinnen wie Scarlett Johansson, Kristen Bell oder Bella Thorne haben dies versucht, erreichen nach dem *Streisand-Effekt*¹⁴ aber womöglich das Gegenteil. Ursprünglich wurden

- 13 Inwiefern hierbei von einer Äquivalenz gesprochen werden kann, bedarf einer genaueren Prüfung, die hiernicht vorgenommen werden kann. Legt man die oben beispielhaft angeführten Kriterien für ›Prominenz‹ der MrDeepFakes-Foren zugrunde, ist die Zahl der hiermit bezeichneten Personen stark begrenzt und schließt viele Personen aus, die sich der breiten Öffentlichkeit nicht zeigen.
- 14 Das hiermit bezeichnete Phänomen geht auf eine Klage der Schauspielerin Barbra Streisand zurück, die von dem Betreiber einer Website forderte, illegal veröffentlichte Luftaufnahmen ihres Hauses zu entfernen. Das Foto wäre zwischen tausenden anderer Aufnahmen von Häusern an der kalifornischen Küste untergegangen, durch den Prozess hat das Bild jedoch eine

die Deepfakes nur in einzelnen Foren und bestimmten Gruppen geteilt, sobald jedoch Stellung dazu bezogen wird, verbreiten sich die Videos schnell weiter und immer mehr Menschen werden darauf aufmerksam. Auch wenn es gelingen sollte, dass das ›Original‹ – das künstlich erstellte Bild, verstanden als die originäre Fälschung – von einer Plattform gelöscht wird, verschwindet es nicht endgültig aus dem Internet. Zudem können neue Videos jederzeit hergestellt werden, sodass es bislang unmöglich ist, sich effektiv dagegen zu wehren. Auch die Bundesregierung hat die Risiken durch Deepfakes in diesem Bereich anerkannt und sieht fragwürdiges Potenzial darin, dass sie »nicht nur massiv die Persönlichkeitsrechte Betroffener [...] verletzen, sondern auch tiefgreifende persönliche Schäden [...] verursachen« (Bundesregierung 2019: 2). Wie man in Zukunft mit diesem Problem umgehen soll, wird jedoch nur vage beantwortet: Man zählt auf »die Stärkung der Medienkompetenz« (ebd.: 5) und hofft auf weitere Forschungserkenntnisse der digitalen Forensik. Auch der Verfassungsschutz verweist aktuell nur auf einige Low-Level-Ansätze zur Erkennung von Deepfakes im Alltag – im Fall von Echtzeit-Übertragungen wirkt der Hinweis, man solle im Zweifel das Gegenüber darum bitten, »sich kurz an Nase oder Wange« zu tippen (Bundesamt für Verfassungsschutz 2021: 37), geradezu hilflos.

Auf dieser Basis lässt sich eine mögliche Verschiebung von Scham- und Peinlichkeitsschwellen erwarten: Jeder kann nackt im Bild zu sehen sein und niemand kann letztgültig sagen, ob es manipuliert ist. Denkt man dies konsequent weiter, dann könnte man sich jederzeit darauf berufen, dass es sich hierbei jeweils um ein Deepfake handle, sodass man sich, strenggenommen, nicht mehr für die nackte oder gar pornografische Darstellung der eigenen Person zu schämen bräuchte. Die Folge daraus wäre eine gewisse ›Normalisierung‹ des eigenen nackten (bzw. sexuellen) Körpers im Bild. Dies entspräche durchaus der Dynamik, von der Elias spricht: »Es ändert sich die Art, in der die Menschen miteinander zu leben gehalten sind; deshalb ändert sich ihr Verhalten; deshalb ändert sich ihr Bewußtsein und ihr Triebhaushalt als Ganzes.« (Elias 1982: 377)

Eine andere mögliche Entwicklung ist die Veränderung des Verhältnisses zum Bild. »Graphic representations constitute a prototypical example of how human beings build external cognitive artifacts for the organization and persuasive display of relevant knowledge.« (Goodwin 1994: 628) Für Goodwin ist die Schöpfung von Bildern eine zentrale Eigenschaft des Menschen. In einen sprachlichen Rahmen eingefügt, haben sie machtvollen Einfluss auf die Wahrnehmung des Betrachters.

stark erhöhte Aufmerksamkeit erhalten, wurde vielfach gezielt gesucht und angeschaut und ist auch heute noch ohne größeren Aufwand online zu finden (vgl. Kleineberg 2014: 67).

Die visuelle Darstellung gilt auf dieser Ebene, wie schon erwähnt, als ein Beweis für etwas, sie ist eine sichtbare Repräsentation. Für Archäologen ist die Karte ein Beweis für vergangene Bauwerke, vor Gericht ist die Aufnahme, wie anhand des Falls Rodney King gezeigt, im bewegten Bild ein Beweis für gute Ermittlungsarbeit und steht dabei gleichzeitig für erlebte Polizeigewalt. Sogar das Gemälde ist für de Nerval, wie oben beschrieben, die wahrheitsgetreue Darstellung der Liebesinsel, bis er es selbst überprüft hat. Prüfungen des Gegebenen erfolgen auf verschiedenen Ebenen (vgl. ebd.: 622), sind aber mit den heutigen Bildern nicht so einfach durchzuführen, wie es noch für de Nerval möglich gewesen ist – und auch zu dieser Zeit dürfte es nicht jedem möglich gewesen sein, die Reise zur Insel anzutreten, um die durch das Bild erzeugte Vorstellung zu verifizieren. Mit den Deepfakes ist es hingegen jedem möglich, ein Bild zu erzeugen, das kaum noch einen Bezug zur Wirklichkeit besitzt, ohne dass der Betrachter dies auch nur vermuten würde (sofern Deepfakes keine generellen Zweifel an der Repräsentationskraft von Bildern auslösen).¹⁵ Durch das Wissen über die bloße Existenz von Deepfakes kann das Bild nun aber seine Bedeutung als Evidenzgenerator verlieren und als reines Kunstobjekt gesehen werden. Es besteht immerzu die Möglichkeit, dass es manipuliert ist; Zweifel können jederzeit aufgeworfen werden, die stets ›berechtigt‹ sind, da es immer eine alternative Sichtweise und immerzu die Gefahr der fingierten Bildebene gibt. Für Elias muss die Natur erst als etwas vom Menschen Abgegrenztes wahrgenommen werden, damit man auf den Gedanken kommt, sie bildlich darzustellen. Mit dem heutigen Wandel des Bildes kann nicht mehr angenommen werden, dass stets eine Verbindung zwischen Abbildung und dargestelltem Objekt vorliegt, sodass erneut eine Abgrenzung in der Vorstellung des Menschen nötig ist.

Sicherlich ist in Deepfakes nicht nur die Furchttraumutopie zu sehen, die diese Schlüsse nahelegen. Die Technologie kann für positiv konnotierte Zwecke, etwa in der Kunst, zur Unterhaltung oder eingängigeren Vermittlung von öffentlichen Botschaften genutzt werden (vgl. Farid/Schindler 2020: 29). Auch in den Bereichen der Medizin (vgl. Bundesregierung 2019: 2), der Bildung (Silbey/Woodrow 2019) oder der Trauerkultur in Form der visuellen Repräsentation Verstorbener in Verbindung mit einem »Thanabot«¹⁶ (Henrickson 2023) könnten sie einen Nutzen finden. Die Vermengung von Traumbild und Wunscherfüllung unter

15 Jean Baudrillard bezeichnet eine solche Welt als Hyperrealität (etwa in: Baudrillard 1982: 113 f.): Eine simulierte Realität, in der alle Dinge ihren Bezug zu einer zugrundeliegenden Wirklichkeit verloren haben und sich nur noch auf ihr eigenes Modell beziehen (dazu ausführlich Spiegel 2022).

16 Dieser Anwendungsbereich ist noch kaum verbreitet, und abgesehen von ethischen Abwägungen darüber, inwieweit solche Technologien generell eingesetzt werden sollten (Heesen 2022), ist bisher ungeklärt, ob und in

Zuhilfenahme von KIs ist nicht per se negativ. Aufgrund der Möglichkeit, dass jeder zu einem Nutzer dieser Technologie werden kann, liegt es jedoch nahe, dass sie zunehmend missbräuchlich und zum Schaden anderer eingesetzt wird. In Verbindung mit der zunehmenden Bekanntheit der Deepfakes und der Unmöglichkeit der Prüfung auf Echtheit erklärt sich, weshalb sich die Betrachtungsweise von Bildern verändern kann. Eine wirksame Kontrolle kann jedoch nicht durch Einzelmaßnahmen erreicht werden (vgl. ebd.: 29f.). Die Systeme haben keine bzw. nicht ausschließlich spezifische Hersteller, die adressiert werden können, sondern werden auch und vor allem von Privatpersonen auf der ganzen Welt erstellt, geteilt und weiterentwickelt.

Der Fall von *DeepNude*, einem Netzwerk, mit dem – wie angedeutet – die Kleidung von Personen aus Bildern entfernt werden kann, verdeutlicht die Problematik. Über die Website der Anbieter konnten kostenfreie und bezahlte Deepfakes erstellt werden; nach der Veröffentlichung des kritischen Berichts einer Journalistin über diese Praxis stiegen die Zugriffszahlen stark an, was zur Überlastung und zum Absturz der Server der Anbieter führte, sodass die Website offline ging und ihr Angebot infolgedessen nicht weiter existierte (vgl. Ajder et al. 2019: 8). Obwohl die ursprünglichen Entwickler das System nicht wieder bereitstellen wollen, ist es über andere Anbieter weiterhin verfügbar, ohne dass sie noch darauf Einfluss nehmen können. Es »will likely continue to spread and mutate like a virus« (ebd.). Die populäre Plattform *Github*, auf der viele Programmcodes, Anleitungen und Netzwerke dieser Art geteilt werden, hat *DeepNude* mittlerweile verbannt (Robertson 2019), trotz dieser Anstrengungen lässt sich dieses und andere Netzwerke seiner Art weiterhin leicht finden. Die Verbreitung der Systeme und der von ihnen erstellten Synthetika lassen sich, wie dieses Beispiel demonstriert, nur schwer kontrollieren und ein generelles Verbot des Programmierens selbst lässt sich nicht umsetzen. Die Rückverfolgung von Manipulationen wird zudem deutlich erschwert, wenn man nicht mehr erkennen kann, was ein Original und was eine Fälschung ist. Überdies kommen womöglich gar keine Zweifel an einem Bild auf, wenn eine betroffene Person nicht weiß, dass sie Zielperson eines Deepfakes ist und somit auch nicht darauf hinweisen kann. Deepfakes sind zumindest auf einer bildrechtlichen und moralischen Ebene insofern eine Bedrohung für die bestehende Ordnung, als sie das Verhältnis zum Bild oder zur Sexualität sukzessive verändern.

Womöglich ist die gesamte Deepfake-Thematik aber auch nur ein singuläres Ereignis, ein kurzfristiger Ausschlag, der langfristig keine signifikante Veränderung hervorruft. Die KI-Forschung ist aktuell auf einem (vorläufigen) Höhepunkt und macht große Fortschritte; in ihrer

welchem Maß sie Menschen bei der Bewältigung eines Verlustes helfen können oder genau dies nicht vielmehr behindern (Lindemann 2023).

Geschichte ist dies jedoch nicht das erste Mal. Seit ihrer Gründung auf der Dartmouth-Konferenz 1956 wurde ihr mehrfach großes Potenzial vorausgesagt und mehrfach wurde dieser Anspruch nicht eingehalten, sodass die Investitionsbereitschaft abnahm und die Forschung im Anschluss in sogenannte *KI-Winter* fiel. Aktuell verbreiten sich Deepfakes schnell; es ist aber nicht auszuschließen, dass sie nur ein kurzer Trend sind, der bald wieder vergessen sein wird, da man sich womöglich erneut zu viel von der Technologie verspricht und ihre noch bestehenden Grenzen nicht wahrhaben will. Es können schließlich nicht *jegliche* Wünsche durch Deepfakes (oder KIs im Allgemeinen) erfüllt werden. Auch daran erinnert, mit etwas anderer Blickrichtung, Elias: »Der Prozeß der Zivilisation vollzieht sich keineswegs geradlinig. Man kann [...] den allgemeinen Trend der Veränderung herausarbeiten. Im Einzelnen gibt es auf dem Wege der Zivilisation die mannigfachsten Kreuz- und Querbewegungen, Schübe in dieser und jener Richtung.« (Elias 1981: 256) Im Prozess der Manipulation von (bewegten) Bildern lässt sich – wie in vielen anderen Bereichen – ein Trend zum Einsatz von KIs erkennen. Sie ermöglichen einer breiten Nutzergruppe die Erfüllung bestimmter Wünsche in einer bildhaften ›Realität‹; in welche Richtung der Prozess aber führt, ist nicht absehbar. Deepfakes haben keine lange Geschichte, die man zurückverfolgen kann, entscheidend sind aber die korrespondierenden sozialen Phänomene, die womöglich einen strukturellen Wandel in der Wahrnehmung von Bildern kennzeichnen. Die genannte Forderung von Peele ist dabei bedeutsam: Auch der Laie muss wissen, was technisch möglich ist, um das Gesehene hinterfragen zu können. Das Sehen ist Teil der Erziehung, mit der Fragen evoziert werden sollen. Elias stellt diese Aufgabe der Wissenschaft, denn es muss Experten geben, die das Phänomen untersuchen, um die Fragen beantworten können, die sich beim bloßen Anblick des Phänomens eventuell bereits aufdrängen.

Literatur

- Ajder, Henry/Patrini, Giorgio/Cavalli, Francesco/Cullen, Laurence (2019): *The State of Deepfakes. Landscape, Threats, and Impact*, Amsterdam.
- Baudrillard, Jean (1982): *Der symbolische Tausch und der Tod*, München.
- Benkel, Thorsten (2008): *Soziale Welt und Fiktionalität. Chiffren eines Spannungsverhältnisses*, Hamburg.
- Benkel, Thorsten (2022): »Erinnern an das Vergessen. Über Gedächtnis und Gedenken«, in: ders./Oliver Dimbath/Matthias Meitzler (Hg.): *Sterblichkeit und Erinnerung*, Baden-Baden, S. 27–51.
- Bourdieu, Pierre (2015): *Manet. Eine symbolische Revolution*, Berlin.
- Bourdieu, Pierre (2023): *Konversion des Blicks. Die Kamera im Dienst soziologischer Erkenntnis*, Wien/Berlin.

- Bundesamt für Verfassungsschutz (2021): »Deepfake. Auf dem Weg zu Social Engineering 2.0?«, in: *Single Point of Contact* 1, S. 35–37.
- Bundesregierung (2019): »Beschäftigung der Bundesregierung mit Deepfakes. Drucksache 19/15210, Drucksache 19/15657«, <https://dserver.bundestag.de/btd/19/156/1915657.pdf> (2. August 2023).
- Chollet, François (2018): *Deep Learning mit Python und Keras. Das Praxis-Handbuch*, Frechen.
- Dalí Museum (2019): »Dalí lives. Museum Brings Artist Back to Life with AI«, <https://thedali.org/press-room/dali-lives-museum-brings-artists-back-to-life-with-ai/> (29. August 2023).
- Deussen, Oliver (2007): *Bildmanipulation. Wie Computer unsere Wirklichkeit verzerren*, Heidelberg.
- Elias, Norbert (1976): *Über den Prozeß der Zivilisation. Soziogenetische und psychogenetische Untersuchungen*, 2 Bde., Frankfurt am Main.
- Elias, Norbert (2000): *Watteaus Pilgerfahrt zur Insel der Liebe*, Frankfurt am Main/Leipzig.
- Elias, Norbert (2002): »Vom Sehen in der Natur«, in: ders., *Gesammelte Schriften*, Bd. 1: *Frühschriften*, Frankfurt am Main, S. 9–28.
- Farid, Hany/Schindler, Hans-Jakob (2020): *Deepfakes. Eine Bedrohung für Demokratie und Gesellschaft*, Berlin.
- Generated Media (2019): »Generated Photos«, <https://generated.photos/> (29. August 2023).
- Generated Media (2023): »2,675,674 AI Generated Photos«, <https://generated.photos/faces> (29. August 2023).
- Goldberg, Carrie (2019): *Nobody's Victim. Fighting Psychos, Stalkers, Pervs, and Trolls*, New York.
- Goodwin, Charles (1994): »Professional Vision«, in: *American Anthropologist* 96 (3), S. 606–633.
- Heesen, Jessica (2022): »Verstorbene als Medienprodukt. Die Programmierung von Unendlichkeit als ethische Herausforderung«, in: Wolfgang George/Karsten Weber (Hg.): *Fehlendes Endlichkeitsbewusstsein und die Krisen im Anthropozän*, Gießen, S. 161–172.
- Henrickson, Leah (2023): »Chatting with the Dead. The Hermeneutics of Thanabots«, in: *Media, Culture and Society* 45 (5), S. 949–966.
- Icons8 (2023): »Tausche Gesichter in höchster Qualität«, <https://icons8.de/swapper> (29. August 2023).
- Jaubert, Alain (1989): *Fotos, die lügen. Politik mit gefälschten Bildern*, Frankfurt am Main.
- King, David (1997): *Stalins Retuschen. Foto- und Kunstmanipulationen in der Sowjetunion*, Hamburg.
- Kleineberg, Christoph (2014): »Shitstorm-Attacken. Digitaler Orkan oder Sturm im Wasserglas?«, in: Ulrich Dittler/Michael Hoyer (Hg.): *Social Network. Die Revolution der Kommunikation. Kundenkommunikation, Facebook-Freundschaften, digitale Demokratie und virtuelle Shitstorms unter medienpsychologischer und mediensoziologischer Perspektive*, München, S. 61–70.

- Korte, Hermann (2013): »Watteaus Pilgerfahrt zur Insel der Liebe«, in: ders., *Biographische Skizzen zu Norbert Elias*, Wiesbaden, S. 67–73.
- Lindemann, Nora F. (2022): »The Ethics of ›Deathbots‹«, in: *Science and Engineering Ethics* 28 (6), S. 1–15.
- Loenhoff, Jens (2001): *Die kommunikative Funktion der Sinne. Theoretische Studien zum Verhältnis von Kommunikation, Wahrnehmung und Bewegung*, Konstanz.
- MrDeepFakes (2023a): »Paid Requests«, <https://mrdeepfakes.com/forums/forums/video-requests.10/> (29. August 2023).
- MrDeepFakes (2023b): »Natalie Portman Fingering«, <https://mrdeepfakes.com/video/4234/natalie-portman-fingering> (29. August 2023).
- Pawelec, Maria/Bieß, Cora (2021): *Deepfakes. Technikfolgen und Regulierungsfragen aus ethischer und sozialwissenschaftlicher Perspektive*, Baden-Baden.
- Peele, Jordan/BuzzFeedVideo (2018): »You won't Believe what Obama says in this Video!«, in: *youtube.com*, 17. April 2018, <https://youtu.be/cQ54GDm1eLo> (29. August 2023).
- Pornpen (2023): <https://pornpen.ai/> (29. August 2023).
- Ries, Gerhard (2001): »Damnatio memoriae. Die Vernichtung des Andenkens an Verstorbene in Politik und Strafrecht«, in: Markwart Herzog (Hg.): *Totengedenken und Trauerkultur. Geschichte und Zukunft des Umgangs mit Verstorbenen*, Stuttgart, S. 237–248.
- Robertson, Adi (2019): »Github is Banning Copies of ›Deepfakes‹ Porn App DeepNude«, in: *theverge.com*, 9. Juli 2019, <https://www.theverge.com/2019/7/9/20687902/github-bans-deepnude-deepfakes-ai-nudity-app-copies> (29. August 2023).
- Silbey, Jessica/Hartzog, Woodrow (2019): »The Upside of Deep Fakes«, in: *Maryland Law Review* 78 (4), S. 960–966.
- Der Spiegel (2008): »Bildmanipulationen. Mehr Blut, mehr Rauch, weniger Speck«, in: *Spiegel online*, 11. Juli 2008, <https://www.spiegel.de/fotostrecke/manipulierte-bilder-fotostrecke-107186.html> (29. August 2023).
- Spiegel, Mirco (2022): *Hyperrealität und Transhumanismus. Der Mensch in der simulierten Gesellschaft*, Wiesbaden.
- Thomas, William I./Thomas, Dorothy S. (1928): *The Child in America. Behavior Problems and Programs*, New York.
- Volland, Holger (2018): *Die kreative Macht der Maschinen. Warum Künstliche Intelligenzen bestimmen, was wir morgen fühlen und denken*, Weinheim/Basel.
- West, Jevin/Bergstrom, Carl (2019): »Which Face Is Real?«, <https://www.whichfaceisreal.com/index.php> (29. August 2023).

Abbildungen

Abb. 1 und 2: Faceswaps der Portraitfotos von Norbert Elias und Thorsten Benkel (erstellt mit Face Swapper von Icons8)

Abb. 3: Antoine Watteau: Pèlerinage à l'Île de Cythère. 1719. Schloss Charlottenburg, Berlin.

Abb. 4: Fotomontage eines irakischen Soldaten, umgeben von US-Militärs (Ursula Dahmen/AP, nach: Der Spiegel 2008)

Abb. 5: Generated Media: 2.675.749 von einer künstlichen Intelligenz generierte Bilder (Generated Media 2023)

Abb. 6: Deepfake mit dem Gesicht von Natalie Portman (MrDeepFakes 2023)