

XIX. Weltversetzung brachial: Von Bananen und Toastern, von Äpfeln und Computern

Nun könnte man es dabei bewenden lassen, dass es diese Bilder und deren Zirkulation eben gibt – als eines der vielen Phänomene, die im Netz kursieren und die so harmlos sind wie die milliardenfach geposteten Bilder von eingenommenen Mahlzeiten oder niedlichen Haustieren in ihrer exaltierten Redundanz und zügellosen Saturiertheit. Allerdings gibt es Anwendungsbereiche, die dem Ganzen die Anmutung des Beliebigen nehmen, in denen verkannte oder gar intentional gestörte Zuordnungen mehr als nur einfache Zuordnungsfehler sind und durchaus ernst zu nehmende Folgen für den realen Alltagsbetrieb haben. Die vermeintlichen Belanglosigkeiten vermögen in die Lebenswelt einzugreifen, ein Befund, der ausgerechnet die Bilderkennung im Umfeld des autonomen Fahrens betrifft und damit an einem der Brennpunkte möglicher und in ihrer Brisanz heftigst diskutierten AI-Anwendungen die Gemüter erregt.¹ Zum Problem werden dabei nicht Menschen und deren sattsam bekannte Unzulänglichkeiten, sondern die Systeme der Erkennung von Bildvorgaben und damit die Technik selbst. Was zutage tritt, sind nicht beiläufige, sondern gezielte Störungen künstlicher Intelligenzen. Die Rede von der Halluzination ist in diesen Fällen nicht, wie im letzten Kapitel gezeigt, auf interne Verarbeitungsfehler auf Grundlage mangelnder Welt-haltigkeit zurückzuführen, sondern sie verdankt sich intentionalen Störungen, strategischen Interventionen, also gewollten feindlichen Attacken: »Recently, adversarial algorithms were developed to facilitate hallucination of deep neural networks for ordinary attackers«.²

Nicht einem Fehler, einem in der Zukunft noch zu behebenden Ungenügen, für das es eben nur mehr Zeit und immer weiterer Daten bedarf, sondern einem Vorsatz ist diese Verwendungsweise der Halluzinationen geschuldet. Der ganzen Auseinandersetzung liegt der Befund zugrunde, dass Systeme der optischen Bilderkennung operativ auf verhältnismäßig unaufwendigem Weg dazu gebracht werden können, einfache, aber durchaus groteske Fehleinschätzungen vorzunehmen. Es bedarf dazu (»man staune!«) lediglich einiger Sticker. Dieses Phänomen wurde zunächst als eine Art Kuriosität der Artificial Intelligence verhandelt, als eine Art randständiger Störung. Die Folge sind Berichte über Bildumtriebe und Bildeigenmächtigkeiten, die

1 Vgl. dazu Florian Sprengers Workshop Autonome Autos (Bochum, November 2020).

2 Abdel-Hakim, Alaa E. (2019), »Ally patches for spoliation of adversarial patches«, in: Journal of Big Data 6/51, o. Pag.

sich mit wunderbarer Stimmigkeit in das Spektrum minderbegabter Menschen und wenig minderbegabter Maschinen fügen lassen. Ein Artikel in der Wochenzeitschrift *Die Zeit* greift zur Veranschaulichung besagter Trugbilder auf Forschungen zurück, die ihren Ausgang von der technischen Metamorphose einer unscheinbaren Banane nehmen.

Die Frucht wird von einem technischen Bilderkennungssystem zunächst ordnungsgemäß erfasst und entsprechend auch erkannt.³ Soweit, so gut und wegen der vielen Bananenbilder, die es in der Welt gibt, aber auch wegen der vielen Bilderkennungssysteme, die inzwischen mit diesen Bananenbildern haben trainieren können und als entsprechend ausgereift gelten, durchaus erwartbar. Aber dann geschieht etwas eigentümlich Bemerkenswertes: Wird nämlich ein kleines Bild, ein *patch* oder *sticker*, der keinen ersichtlichen Bezug zum Ausgangsbild aufweist und der in der Gesamtwahrnehmung kaum auffällt, in die Nähe der Banane verbracht, gibt das System die Wahrnehmung der Frucht preis und identifiziert die neue Situation etwa als die eines Toasters. Doch damit nicht genug: War die Bilderkennungssoftware im Fall der Banane mit 97 Prozent von der Richtigkeit ihrer eigenen Erkenntnis überzeugt, so steigert sich diese Zuversicht im Fall des Toasters auf stattliche 99 Prozent. Die Fülle der Beispiele ist beliebig erweiterbar um Dinge des Alltags, um verwechselte Teddybären und Orangen, um verkannte Socken und Haustiere, um fehlinterpretierte Sportgerätschaften und schnöde Einrichtungsgegenstände.⁴ Das Trugbild gerät mithin in den Zustand seiner technischen Stabilisierbarkeit.

Andere Fälle zeigen, wie eine Mikrowelle mit störendem Sticker als Telefon einsortiert wurde. Und wie nach Zugabe von ein paar Pixeln, die kein Mensch bemerken würde, der Rechner einen Pandabären als Gibbonäffchen erkennt.⁵

Derartige Geschichten (und ihre Beispiele mit ihrem durchaus poetischen Wortbildungspotential wie im Fall einer *rifle turtle*) bilden ein eigenes Genre und lassen sich gut erzählen, nicht zuletzt deswegen, weil das Ausmaß der Fehleinschätzung im Ausscheren aus gängigen Erwartungsmustern besonders grotesk ist.⁶ Die Verwechslung der Banane mit einer ihr gestaltähnlichen Zucchini oder einer ihr farbähnlichen Zitrone wäre im Auge des menschlichen Betrachters durchaus erwartbar gewesen, aber gerade sie findet nicht statt. Sie wäre vor dem Hintergrund eines Denkens in Gradualitäten und Linearitäten auch einigermaßen plausibel, wäre sie doch durch eine Geschichte des technischen Sehens gedeckt, die sich entlang von staffelbaren Auflösungen, aber auch im Umgang mit anderen Sehparametern wie Abschattung, Entfer-

3 Zum Prozess einer AI-gestützten Mustererkennung am Beispiel von Obstsorten vgl. Crawford, Kate/Paglen, Trevor (2021), »Excavating AI: the politics of images in machine learning training sets«, in: *AI & Society* 36, S. 1105-1116.

4 Ein Youtube-Video mit dem Titel *Fooling Image Recognition with Adversarial Examples* stellt das vor Augen: https://www.youtube.com/watch?v=piYnd_wYIT8 (letzter Zugriff: 07.11.22) und entlässt neue Gebilde in die Welt (wie es in einem Kommentar von Dan Fredeiksen heißt: »Well it's clearly a rifle turtle.«).

5 Schmitt, Stefan (2019), »Ist das wirklich ein Toaster?«, in: *Die Zeit*, Nr. 47 vom 14.11.2019, S. 33-34, hier: S. 33.

6 Zur programmatischen Herstellung solcher Mischwesen vgl. noch einmal K. W. Ng/X. Zhu/Y-Z. Song u.a., *DreamCreature*.

nung oder Blickwinkel abzuarbeiten hatte. Aber die Verwechslung mit einem Toaster fällt entschieden aus dem Rahmen, verlässt sie doch dieses Register und schert aus der sie begründenden Ordnung aus. Die Dinge gehen in diesem Moment des Ausscherens den Kategorien oder genauer noch: der Kategorisierbarkeit von der Stange. Und sie tun das auf eine Weise, der ein Moment des Schockhaften, des Brachialen eigen ist.

Niklas Luhmann hat anlässlich des Übergangs von gesprochener Sprache zur Schrift eine vergleichbare Operation in Erwägung gezogen, um deren verstörendes Potential in einem Gedankenspiel sichtbar werden zu lassen und sogleich vehement von sich zu weisen. Auch ihm ist dabei das unverhohlenen Brachiale solcher Verwechslungen aufgefallen, und er betont, welche Gefahren mit einer solchen Operation einhergehen. Nicht am Beispiel von Bananen und Toastern, dafür aber an der unwaghalsigen, weil unverzichtbaren Stabilitätsunterstellung von Äpfeln und Computern beschreibt er ein Moment der Weltversetzung, die im Modus des Schleuderns erfolgt und die auf ihre Weise ein damit verbundenes Trauma sichtbar macht.

Laute durch Sichtbares zu ersetzen beruht dann auf einer Vorannahme, daß der Gegenstand der Bedeutung unverändert bleibt, d.h. daß ein Apfel ein Apfel bleibt und sich nicht in einen Computer verwandelt, wenn wir über ihn sprechen oder schreiben.

In diesem Sinne muß dies wahr sein (oder als wahr erscheinen), weil wir sonst die Substitution nicht wagen könnten. Sie würde uns in eine andere, unbekannte Welt schleudern.⁷

Was beide Beispielreihen, die Bananen und Toaster, die Äpfel und Computer, sichtbar machen, ist das Risiko, das mit dem Verlust stabiler Referenzen einhergeht. Die Ordnung der Dinge würde damit nicht nur ein wenig aus den Fugen geraten, sie wäre in ihrer Unbekanntheit schlicht eine gänzlich andere. Dabei war und ist es gerade die Identifizierung von Gegenständen, die als die besondere Leistung des menschlichen Sehens veranlagt wird und die das maschinelle Sehen immer wieder vor besondere Herausforderungen stellte, denen man sich approximativ zu nähern hatte. »Auf keinem anderen Gebiet«, so schrieb der deutsche Kybernetiker und Nachrichtentechniker Karl Steinbuch in der Mitte des letzten Jahrhunderts, »ist die Unterlegenheit technischer Gebilde gegenüber den organischen Systemen so offensichtlich wie beim Sehsystem.«⁸ Vor dem Hintergrund dieser Annäherungsgeschichte ist die Übertragung von Bananen auf Toaster ein entschiedener Bruch im habitualisierten Umgang mit Ähnlichkeitsrelationen und untermauert deren von Foucault herausgearbeitetes Potential. Sie stellt einen Ausnahmetatbestand dar, der auch durch gängige Sonderformen und theoretisch folgenreiche Spezialfälle in der Wahrnehmung nicht gedeckt ist. Eine solche liegt etwa mit dem Phänomen der so genannten Kippbilder vor, anhand derer die Gestaltpsychologie immer wieder das Unterlaufen eingespielter Bildwahrnehmung beschrieben hat. Über die Logik von Gestaltwahrnehmung und über diejenige einer technisch gestörten Bildererkennung entscheiden nicht Auflösungen, Approximationen und Annäherungen. In beiden Fällen ist der Kontrakt mit der Stetigkeit aufgekündigt.

7 Luhmann, Niklas (1993), »Die Form der Schrift«, in: Hans Ulrich Gumbrecht/K. Ludwig Pfeiffer (Hg.), *Schrift*, München: Fink, S. 349-366, hier: S. 354f.

8 Steinbuch, Karl (1971), *Automat und Mensch. Auf dem Weg zu einer kybernetischen Anthropologie*, 4. Aufl., Berlin u.a.: Springer, S. 98.

Zwischen Pandabären und Gibbonäffchen, zwischen Banane und Toaster, zwischen Äpfeln und Computern, zwischen der Wahrnehmung der Rubinschen Vase oder der von Menschengesichtern gibt es keinen Graubereich, gibt es keine Übergängigkeiten:

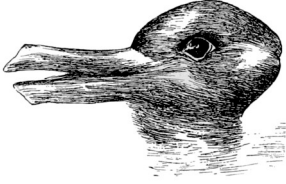


Abbildung 47: Kaninchenente
(nach Jastrow 1899)

Die Entscheidung zwischen Gesichtern und Vasen in diesem Reich des Ambiguen ist ein Ausscheren aus der Logik des Erwartbaren und ein kalkulierter Bruch mit den Kalkülen der Wahrscheinlichkeit.⁹ Oder anders gesagt: Für Vasengesichter oder Entenkaninchen ist in dieser Welt noch kein Platz – ebenso wenig wie für Schildkrötengewehre oder Mikrowellentelefone.¹⁰

Allerdings hat es mit der bloßen Vertauschung der Dinge nicht sein Bewenden. Was zunächst wie ein harmloses Kuriosum wirkt, wird dann zum lebensweltlichen Problem, wenn statt Bananen und

Toastern etwa kritische Infrastrukturen in Form von Verkehrsschildern betroffen sind. Im Fall missgeleiteter Signal- und Schilderkennung wird das Halluzinieren der Maschine zu einer veritablen Störung einer in ihrer Verkehrsfreudigkeit zunehmend gesteigerten Lebenswelt. Es sind wie im Fall der verfehlten Banane feindliche Attacken (*adversarial attacks*), die das System um die Kohärenz und nachgerade um den sorgsam (und mit einem riesigen Arsenal an Bildern) antrainierten Verstand bringen. Und es sind auch hier wie im Fall der Banane so genannte *adversarial patches*, die solche Störfälle im allgemeinen Wahrnehmungsbetrieb auf unscheinbare Weise befördern. So wird von einem Stoppschild berichtet, das durch einen unscheinbaren Aufkleber an der menschlichen Wahrnehmung vorbei unkenntlich gemacht wird oder davon, »eine Plastikschildkröte durch subtile Gravuren auf ihrem Panzer als Schusswaffe erscheinen zu lassen.«¹¹

Gerade die Identifizierung von Verkehrszeichen ist entsprechend aufwendig



Figure 1: The left image shows real graffiti on a Stop sign, something that most humans would not think is suspicious. The right image shows our a physical perturbation applied to a Stop sign. We design our perturbations to mimic graffiti, and thus "hide in the human psyche."

Abbildung 48: Verkehrszeichenvereitelung
(nach Evtimov et al. 2018)

9 Zur Kippfigur sowie zu ihrer Ambiguität vgl. Schönhammer, Rainer, »Stichwort: Kippbilder«, online unter https://psydoc.psycharchives.de/jspui/bitstream/20.500.11780/3666/1/Kippbilder_psydoc_11052011.pdf (letzter Zugriff: 23.04.20), zum Versuch, die damit verbundene Ambiguität zu messen und die Beispielpalette beträchtlich zu erweitern, Fisher, Gerald H. (1967), »Measuring Ambiguity«, in: *American Journal of Psychology* 80/4, S. 541-557, sowie zu Strategien der Vereindeutigung Kawabata, Nobuo/Mori, Takayuki (1992), »Disambiguating ambiguous Figures by a model of selective attention«, in: *Biological Cybernetics* 67, S. 417-425.

10 Zur Geschichte solcher Kippfiguren vgl. Brugger, Peter (1999), »One hundred years of an ambiguous figure: Happy birthday, duck/rabbit«, in: *Perceptual and Motor Skills* 89, S. 973-977.

11 S. Schmitt, Ist das wirklich ein Toaster?, S. 33. Zu diesem Beispiel vgl. Anish Athalye, Logan Engstrom, Andrew Ilyas u.a. (2018), »Synthesizing Robust Adversarial Examples«, *arXiv:1707.07397v3 [cs.CV]* 7 Jun 2018. Zur Robustheit vgl. auch die Dissertation von Srinivasan, Vignesh (2022), *Towards Robustifying Deep Neural Networks against Adversarial, Fringe and Distorted Examples*, Diss. TU Berlin.

beforscht.¹² Weil derartige Wahrnehmungsgrotesken und Musterverkennungen die Sicherheit in einer konkret zugrundegelegten Welt gefährden (und nicht wie im Fall von Luhmann die Betroffenen in eine andere, in eine unbekannte Welt schleudern würden), bedarf es in ihrem Fall ganz besonderer Vorsichtsmaßnahmen. Der *Zeit*-Artikel verweist daher auf eine Steigerung und Ausweitung möglicher Anwendungsbereiche. Während es bei den geschilderten Beispielen um die Verwechselbarkeit statischer Bilder zu tun ist, arbeiten Forscher wie Michael J. Black am Tübinger Max-Planck-Institut für intelligente Systeme daran, Bewegungsaufnahmen zu manipulieren, also in so genannte *optical flow systems* einzugreifen und damit komplexe Situationen zu antizipieren. »Weil optical flow-Systeme es erlauben, im Gewusel einer Verkehrssituation die wesentlichen Bewegungen (etwa von Radfahrern und Fahrzeugen) zu verfolgen, sind sie plausible Bausteine für künftige selbstfahrende Autos.«¹³

Zur Optimierung der eigenen Systeme haben die Tübinger Forscher daher vor der Veröffentlichung ihrer Ergebnisse entsprechende *patches* an die Hersteller und Zulieferer der Automobilbranche geschickt, um diese vorab in die Lage zu versetzen, eigene Untersuchungen mit den Störmöglichkeiten zu unternehmen. Die Leiterin des Algorithm Accountability Lab, Katharina Zweig, die der Artikel zu Wort kommen lässt, betont den Ernst der Lage mit dem Verweis darauf, dass solche Systeme am Wissen und am Bewusstsein der Betroffenen vorbei zu ihren Einschätzungen gelangten. Die künstlichen Systeme operierten nicht nur im Modus des Unmerklichen und Unwissentlichen, sie operieren vor allem dort, wo die lebensweltliche Brisanz außer Frage steht. Deutlich wird in Forschungen zu diesem Phänomen, welche Potentiale in der Technik des Angriffes liegen und wie sehr das Arbeiten der entsprechenden Netzwerke emergente Phänomene zeitigt und damit das alte Paradigma einer bestimmten und im Alltagsbewusstsein verankerten Auffassung von Mechanik, der Determiniertheit und Vorhersagbarkeit technischer Systeme aufgibt. Die Deep Convolutional Neural Networks sind schwer vorhersagbar und sie sind nicht zu beschränken auf Bereiche, wo derartige Fehler lediglich als Sonderphänomene zu tolerieren oder als bizarre Bilder von einem arglosen Publikum zu bestaunen sind.

Because of these accomplishments, deep learning techniques are also applied in safety-critical tasks. For example, in autonomous vehicles, deep convolutional neural networks (CNNs) are used to recognize road signs [...]. The machine learning technique used here is required to be highly accurate, stable and reliable. But, what if the CNN model fails to recognize the »STOP« sign by the roadside and the vehicle keeps going? It will be a dangerous situation.¹⁴

Ein in seiner Gefahr vergleichbares Szenario ergibt sich auch an anderen Schauplätzen, etwa am Ort von Finanztransaktionen.

12 Vgl. Ciresan, Dan/Meier, Ueli/Masci, Jonathan u.a. (2012), »Multicolumn Deep Neural Network for Traffic Sign Classification«, in: *Neural Networks* 32, S. 333-338, sowie Huang, Shize/Liu, Xiaowen/Yang, Xiaolu u.a. (2021), »An improved ShapeShifter method of generating adversarial examples for physical attacks on stop signs against Faster R-CNNs«, in: *Computers & Security* 104, o. Pag.

13 S. Schmitt, Ist das wirklich ein Toaster?, S. 33.

14 Xu, Han/Ma, Yao/Liu, Haochen u.a. (2019), »Adversarial Attacks and Defenses in Images, Graphs and Text: A Review«, arXiv:1909.08072v2 [cs.LG] 9 Oct.

If there are fraudsters disguising their personal identity information to evade the company's detection, it will cause a huge loss to the company. Therefore, the safety issues of deep neural networks have become a major concern.¹⁵

Was der Überblickstext *Adversarial Attacks and Defenses in Images, Graphs and Text* zunächst als Strategie menschlicher Agenten gegenüber Kreditinstituten ausmacht, verweist in die Welt anderer Operationsweisen und Darstellungsformen. Die Frage der Glaubwürdigkeit wird nicht entlang der persönlichen Identität von möglichen Betrügern verhandelt, also etwa an der Eineindeutigkeit von Personenständen und deren biometrischer Gewährleistung, also an der Stabilität von Fingerabdrücken und Gesichtern, sondern an Graphen, die mit der Eigenlogik bestimmter Vorgänge befasst sind. Ob im Straßenverkehr oder bei Geldgeschäften: Die Gefahren lauern an unterschiedlichen Stellen einer Lebenswelt, die als immer störanfälliger gilt. Sie sind ausdifferenziert nach Anlässen und technischen Umsetzungen, betreffen eben nicht nur die Erkennung von Buchstaben und Ziffern, Stoppschildern und anderen Verkehrszeichen, Bananen und Schildkröten, sondern auch andere Datenformen.¹⁶ So gerät das Beispiel der Graphen (und der Veränderung von nur wenigen Knoten) im Rahmen des Finanzwesens auf die Agenda und auch der Umgang mit Sprache – ob in schriftlicher oder gesprochener Form – findet Beachtung. Auch hier erfolgen die Attacken im Gewand von minimalen und unmerklichen Unterschieden – etwa durch eine Manipulation in Form von Tipp- oder Druckfehlern (*typos*).¹⁷

Jenseits von Tippfehlern und einfachen Buchstabendifferenzen gerät auch der Austausch ganzer Sätze oder Phrasen im Kommunikationsgeschehen in den Fokus der Aufmerksamkeit. Ein Beitrag mit dem Titel *Detecting egregious responses in neural sequence-to-sequence model* aus dem Jahr 2018 erweitert dabei auf auffallende Weise die Beschreibungssprache und die Semantik dessen, was gemeinhin mit dem Begriff des Auffälligen verbunden ist.¹⁸ Das Halluzinatorische und das Groteske, das der Verwechslung von Bananen und Toastern zu eigen ist, wird dabei um die Eigenschaft des dezidiert Ungeheuerlichen erweitert.

In this work, we attempt to answer a critical question: whether there exists some input sequence that will cause a well-trained discrete-space neural network sequence-to-sequence (seq2seq) model to generate egregious outputs (aggressive, malicious, attacking, etc.).¹⁹

15 Ebd.

16 Zum Beispiel der Ziffern, deren Ikonographie an frühe Bemühungen der technischen Übertragung und Mustererkennung erinnern (Steinbuch, Wiener), vgl. Feinman, Reuben/Curtin, Ryan R./Shintre, Saurabh u.a. (2017), »Detecting Adversarial Samples from Artifacts«, arXiv:1703.00410v3 [stat.ML] 15 Nov. Zu weiteren Beispielen Kurakin, Alexey/Goodfellow, Ian/Bengio, Samy (2017), »Adversarial Examples in the Physical World«, arXiv:1607.02533v4 [cs.CV] 11 Feb. oder Evtimov, Ivan/Eykholt, Kevin/Fernandes, Earlence u.a. (2018), »Robust physical-world attacks on deep learning models«, arXiv:1707.08945v5 [cs.CR] 10 Apr.

17 Jia, Robin/Liang, Percy (2017), »Adversarial Examples for Evaluating Reading Comprehension Systems«, arXiv:1707.07328v1 [cs.CL] 23 Jul.

18 He, Tianxing/Glass, James (2018), »Detecting egregious responses in neural sequence-to-sequence models«, arXiv:1809.04113v2 [s.AI] 3 Oct.

19 Die Verwendung ist kein Einzelfall. Vgl. etwa Sandbank, Tommy/Shmueli-Scheuer, Michal/Herzig, Jonathan u.a. (2018), »Detecting Egregious Conversations between Customers and Virtual Agents«, arXiv:1711.05780v2 [cs.CL] 16 Apr.

Zum Ausweis von Reaktionen auf Sätze gelangt mit dem Adjektiv *egregious* ein Eigenschaftswort zum Einsatz, dessen Bedeutungshof, wie in der erläuternden Klammer mit Aufzählung weiterer Adjektive angedeutet, neben dem Herausstechenden und Auffälligen, dem Unerhörten und Ungeheuerlichen zugleich auch in das Feld des Feindlichen und Bedrohlichen hinüberspielt. Die Elaborate künstlicher Intelligenzen bewegen sich mithin in jenem Bedeutungshof, der an anderer Stelle, mit einem nostalgisch und literarhistorisch verklärten Blick auf die analogen Dinge als Tücke des Objekts (oder mit Blick auf den portugiesischen Schriftsteller José Saramago als lebensbedrohliche Feindseligkeit der Dinge gegenüber den Menschen) beschrieben wurden – und die ihrerseits Anlass für die Gestaltung von Dingen boten, die sich in maximaler Entfernung von Kategorien des Funktionalen und Nützlichen, des Zweckmäßigen und Eindeutigen befanden. Die oben beschriebenen Tische, die keine Tische sein wollen, und die sich widerspenstig gerierenden Einkaufsbeutel und flüchtenden Mülleimer, die sich neigenden Schüsseln und die nicht morsenden Morsezeichen waren nur einige der eigens gestalteten Produkte dieses Ansinnens.

Die Beforschung der Robustheit solcher Systeme wird daher ebenso unumgänglich wie die Ausweitung auf andere Daten. Was sich im Zuge dessen einschleicht, ist eine Semantik der Verdächtigung von allem, was es an Formen und Ausprägungen gibt. Die Rede von Potemkin'schen Dörfern macht sich breit und immer wieder wird der Umstand des Befremdlichen angeführt, der mit dem Unterlaufen der menschlichen Wahrnehmung verbunden ist. Bevorzugt mit dem Adjektiv *subtle* werden jene unmerklichen Veränderungen beschrieben, die zu den eklatanten Fehleinschätzungen und ihren gravierenden Folgen führen. In der Bildgebung der entsprechenden Abhandlungen kommt es zu einer Ikonographie, die an die Suchbilder in der Rätselrubrik von Zeitschriften erinnern, die also Bilder zeigen, die in prinzipiell erkennbaren und daher auflösbaren Details, Objekte in vermeintliche Gleichheit nebeneinanderstellen – nur mit dem Unterschied, dass das menschliche Auge im Fall der *adversarial attacks* jeder Chance beraubt ist, die eingebauten Fehler auch selbst und das heißt auf Grundlage der eigenen Kompetenz zu finden.²⁰ Verglichen mit den Suchbildern auf den Rätselseiten gängiger Druckerzeugnisse und der Aufforderung *Finde den Fehler!* begründen die technischen Systeme ein neues Reich des Unheimlichen, ein neues Reich des Unmerklichen, ein Reich der narzisstischen Kränkung hervorgerufen durch ein Missverhältnis von technischer Überlegenheit und menschlicher Minderbegabung.

Es ist eine eigentümliche Pointe im allgemeinen Wahrnehmungsbetrieb, dass die Abwehrmaßnahmen gegen Attacken auf die Bilder- und Schilderererkennung bei autonomen Fahrzeugen nun ihrerseits ausgerechnet im stimmigen Modus von so genannten *ally patches* auftreten. Dass damit eine Spirale von wechselseitig sich störenden Stickern freigesetzt wird (*Ally patches for spoliation of adversarial patches*), mag ein Detail sein, das vielleicht nur die Aufmerksamkeit der Kulturwissenschaften auf sich zieht, die solche Selbstreferenzen als Teil kultureller Bedeutungswertschöpfungsketten gerne und mit einer gewissen professionellen Neugier in den Blick nehmen. Dass die Sticker die Seiten wechseln und nun ausgerechnet sie es sind, die die Störer zu stören in der Lage sein sollen und damit Sicherheit garantieren, ist nichtsdestoweniger ein

20 Zur Tradition der Suchbilder vgl. Ernst, Wolfgang/Heidenreich, Stefan/Holl, Ute (Hg.) (2003), Suchbilder. Visuelle Kultur zwischen Algorithmen und Archiven, Berlin: Kadmos.

Kreislauf, der Beachtung verdient – gerade mit Blick auf seine potentielle Unabschließbarkeit.²¹ Weil auch hier und einmal mehr das Prinzip der Unmerklichkeit gilt, stellt sich ein eigentümlicher Befund ein: Das menschliche Auge vermag die Kombattanten visuell nicht voneinander zu unterscheiden – erst der Effekt eröffnet ihre Position als Verbündeter oder Feind, als Kombattant oder Gegner. Es ist nur schwer von der Hand zu weisen, dass all dies den Vorhalt der Verdächtigung untermauert – von dem bei Eve Kosofsky Sedgwick die Rede war. Das Anbringen von Aufklebern hat seine Unschuld verloren und damit auch jenen subversiven Charme, der mit diesen halböffentlichen Akten der Meinungskundgabe einherging. Es ist mehr als eine juvenile Ordnungswidrigkeit und hat das Zeug zu einem risikoreichen Eingriff in den Straßenverkehr. »The consequent troubles may vary from just unpleasant inconvenience in applications like entertainment image and video annotation, passing by security-critical problems like false person identifications, and can turn out to be life-threatening in autonomous navigation and driver support systems.«²² So wird die Anfälligkeit des autonomen Fahrens zum Gegenstand einer doch grundlegenden Aufmerksamkeit – wie in einem Fall, bei dem der feindliche Angriff dazu geführt hat, dass die Scheibenwischer eingeschaltet wurden, ganz ohne menschliches Zutun und ohne dass es zuvor auch nur ansatzweise geregnet hätte.²³

Besonders auffallend ist in der Forschung eine Einschätzung, die dem Status der entsprechenden Beispiele gilt.²⁴ Statt der operativen und specialistischen Frage, welches von den dort geschilderten Verfahren (*gradient masking*, *robust optimization*, *detection*) zum Einsatz gelangt, wird mit dem Versuch der Statusbestimmung die Frage drängend, welche Bedeutung den exemplarischen Fällen in konkreten Umgebungen zukommt. Dabei spielen neben dem Katz-und-Maus-Spiel des jeweiligen Reagierens auf Vorgaben Aspekte eine Rolle, die dem Bereich des Grundlagenwissens zuzurechnen sind. Immer wieder, so ist zu lesen, würden die Spiralen und Kreisläufe der Attacken und Gegenattacken das Verständnis in die Eigendynamik der DNN-Systeme verstärken – und damit ein Wissen über deren technische Natur generieren. Wie es in einem der damit befassten Texte heißt, ist es nicht das Kasuistisch-Fehlerhafte, sondern das Systemisch-Typische, das dabei als positives Wissen zur Geltung gelangt: *Adversarial Examples*, so heißt es dort, *Are Not Bugs, They Are Features*.²⁵ Ein weiterer Text schlägt vergleichbare Töne an – nur mit der Nuance, die Rede vom *bug* durch die vom *flaw* zu ersetzen. »These results have often been interpreted as being a flaw in deep networks in particular, even though linear classifiers have the same problem. We regard the knowledge of this flaw as an opportunity to fix it.«²⁶ Der Fehler ist kalkuliert,

21 Zur Aufmerksamkeit der Kultur- und Medienwissenschaft für Belange der Störung vgl. stellvertretend Kümmel, Albert/Schüttelpelz, Erhard (Hg.) (2003), *Signale der Störung*, München: Fink.

22 A. E. Abdel-Hakim, *Patches for spoliation*.

23 Vgl. dazu Deng, Yao/Zheng, Xi/Zhang, Tianyi u.a. (2020), »An Analysis of Adversarial Attacks and Defenses on Autonomous Driving Models«, arXiv:2002.02175v1 [eess.SP] 6 Feb.

24 Vgl. dazu die besonders eindrückliche Darstellung bei H. Xu/Y. Ma/H. Haochen u.a., *Adversarial Attacks*.

25 Ilyas, Andrew/Santurkar, Shibani/Tsipras, Shibani u.a. (2019), »Adversarial Examples Are Not Bugs, They Are Features«, arXiv:1905.02175v4 [stat.ML] 12 Aug.

26 Goodfellow, Ian J./Shlens, Jonathon/Szegedy, Christian (2015), »Explaining and harnessing adversarial examples«, arXiv:1412.6572v3 [stat.ML] 20 Mar. Dort findet sich auch die Rede von den Potemkinschen Dörfern.

beabsichtigt und folgt einer eigenen Agenda. »Adversarial examples are inputs to machine learning models that an attacker intentionally designed to cause the model to make mistakes.«²⁷

Als besondere Pointe dieser Grundlagenarbeit an den Funktionsweisen der DNN ist zu verzeichnen, dass diese eben anders als die menschliche Datenverarbeitung funktionieren – auch wenn die gewählten Beschreibungssprachen, die vom Träumen oder Halluzinieren der Maschinen handeln, diesen Unterschied weitgehend kaschieren. So sehr sich die Schutzmaßnahmen und ein Verständnis der Operationsweise von DNN wechselseitig bedingen, so bleiben menschliches Wahrnehmen und Schlussfolgern dabei außen vor: »For example, adversarial perturbations are perceptually indistinguishable to human eyes but can evade DNN's detection. This suggests that the DNN's predictive approach does not align with human reasoning.«²⁸ Der technische Eigensinn der neuronalen Netze hat sich trotz oder gerade wegen aller semantischen Anleihen an die Vorstellungspolitik des Menschen auf eine gewisse Weise davon emanzipiert – und sich ihnen, was die Effekte angeht, zugleich auch ein Stück weit angenähert. Die Ordnung der Dinge ist nicht mehr zwangsläufig eine Ordnung, die entlang von Ähnlichkeiten und Repräsentationen, vor menschlich nachvollziehbaren Voraussichten und Reaktionen anzuschreiben ist. Die Gefährdung von Verkehrsschildern, viel mehr noch die Manipulation komplexerer Situationen und Lagen, die etwa für die Stabilität und Akzeptanz des autonomen Fahrens so zentral wäre, stellt damit eine epistemische Sonderstellung solcher Verfahren unter Beweis. Sie sind in der Lage, zu träumen und zu halluzinieren und doch verfahren sie dabei anders, als Menschen es tun. *Ihre* Ordnung der Dinge ist nicht *die* Ordnung der Dinge.

Das Halluzinieren der Maschine verfährt jenseits erwartbarer Ordnungen, in der die Verwechslung der Banane mit einem Toaster jenseits von Gestalt- oder Farbähnlichkeit erfolgt, es eröffnet den Bildumtrieben ein Aus-dem-Ruder-Laufen, leistet einer Art der Psychopathologisierung Vorschub und erlaubt eine Annäherung an den Bereich des Träumens. So macht ein Text unter dem Titel *Learning Real-World Robot Policies by Dreaming* den schwankenden Objektstatus, den Gestaltwandel und damit die Übergängigkeit von Dingen in der Beschreibungssprache des Träumens sichtbar.²⁹ Die Anfälligkeit der Systeme für Verwechslungen führt dazu, dass immer wieder versucht wird, diese Grenzen gezielt auszureizen und zu unterlaufen – verbunden mit einem ganz eigenen performativen Gestus, der expressis verbis zu einem subversiven Umgang mit den Elaboraten der künstlichen Intelligenzen aufruft. Es ist nur konsequent, dass sich derartige Widerstände ausgerechnet auf dem Feld der Texterzeugung artikulieren: *Let's Fool That Stupid AI. Adversarial Attacks against Text Processing AI*.³⁰

Die Artificial Intelligence hat mit dem Halluzinieren nicht nur ihre liebe Not, sondern ein veritables Problem – und wie das folgende Beispiel einer *hallucination machine*

27 H. Xu/Y. Ma/H. Haochen u.a., Adversarial Attacks.

28 Ebd.

29 Dazu Piergiovanni, A. J./Wu, Alan/Ryoo, Michael S. (2019), »Learning Real-World Robot Policies by Dreaming«, arXiv:1805.07813v4 [cs.RO] 1 Aug 2019, S. 7680-7687.

30 Vgl. dazu etwa Schade, Ulrich/Pritzkau, Albert/Claeser, Daniel u.a. (2023), »Let's Fool That Stupid AI. Adversarial Attacks against Text Processing AI«, in: Peter Klimczak/Christer Petersen (Hg.), *AI – Limits and Prospects of Artificial Intelligence*, Bielefeld: transcript, S. 267-283.

zeigen wird, einen neuen und einen für die Belange der Lebenswelt nicht unmaßgeblichen Anwendungsbereich. Dieser Punkt vom (semantischen) Einbruch eines menschlich konnotierten Irrationalen ist nicht unerheblich, weil umgekehrt technische Verfahren verstärkt herangezogen werden, um bestimmte Psychopathologien ihrerseits sicht- und handhabbar zu machen – etwa das halluzinatorische Potential in der Psychose oder in der Schizophrenie. Neben der Psychopathologisierung finden auch Annäherungen an Epochen statt – und zwar nicht nur, wie hier immer wieder gezeigt, an die des Barock.³¹ Was dabei in den Blick gerät, sind Aspekte der zum Teil KI-gestützten Identifizierung und der Vorhersage von pathologischen Störungen auf der Grundlage von *machine learning*.³² Und natürlich geraten auch therapeutische Interventionen mit dem Ziel der Kurierung auf die Agenda. Gegen psychische Erkrankungen im Symptomenspektrum von Paranoia, Schizophrenie und Psychose sollen verstärkt virtuelle und augmentierte Realitäten zum Einsatz gelangen.³³

Aber nicht nur im Fall von therapeutischen Interventionen werden die Topologien des Virtuellen zu einem probaten Mittel erklärt. Sie sollen den Weg zum Verständnis entsprechender Zustände befördern und damit unschwellig jenes Argument der Empathieförderung verstärken, das dem Virtuellen als Etikett verpasst wurde (*empathy machine*). So wird die Umsetzung von Halluzinationen im Virtuellen etwa benutzt, um im Ausbildungsbetrieb, in der Psychiatrie und in der Kognitionsforschung an psychische Phänomene anzuschließen, um sich ein Bild davon zu machen, was es heißt, von derartigen Phänomenen betroffen zu sein, und um in die veränderte Phänomenologie psychischer Zustände einzutauchen. Oder anders gesagt: Auf diese Weise, so jedenfalls der Vorsatz, will man sich ein Bild von den inneren Bildern anderer machen können, will man lernen, was Halluzinationen sein könnten und was es heißt, von diesen auch als nicht Betroffene betroffen zu sein.

The virtual environments created by the new software are expected to enhance the experiential learning outcomes of medical students by enabling them to experience the inner world of a patient with psychosis. In addition the Virtual Environment has the

31 Vgl. dazu das angekündigte Buch P. Schönthaler, *Wie rationale Maschinen romantisch wurden*.

32 Vgl. dazu Correll, Genevieve/Oduola, Sherifat/Roberts, Angus u.a. (2016), »Identifying First Episodes of Psychosis in Psychiatric Patient Records using Machine Learning«, in: *Proceedings of the 15th Workshop on Biomedical Natural Language Processing*, Berlin: Association for Computational Linguistics, S. 196–205, und Taylor, Jeremy A./Larsen, Kit Melissa/Dzafic, Ilvana u.a. (2021), »Predicting subclinical psychotic-like experiences on a continuum using machine learning«, in: *NeuroImage* 241, Art. 118329.

33 Vgl. Fernández-Caballero, Antonio/Navarro, Elena/Fernández-Sotos, Patricia u.a. (2017), »Human-Avatar Symbiosis for the Treatment of Auditory Verbal Hallucinations in Schizophrenia through Virtual/Augmented Reality«, in: *Frontiers in Neuroinformatics* 11, Art. 64, Thompson, Andrew/Elahi, Farah/Realpe, Alba u.a. (2020), »A Feasibility and Acceptability Trial of Social Cognitive Therapy in Early Psychosis Delivered Through a Virtual World: The VEEP Study«, in: *Frontiers in Psychiatry* 11, Art. 219, Elahi, Farah (2021), *Using virtual reality tools to improve social cognitive training in first episode psychosis*, Diss. University of Warwick, sowie Ascone, Leonie/Ney, Karolin/Mostajeran, Fariba u.a. (2020), »Virtual Reality for Individuals with Occasional Paranoid Thoughts«, in: *CHI '20 Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, New York: ACM, S. 1–8.

potential to provide a technologically advanced therapeutic setting where behavioral, exposure therapies can be conducted with exactly controlled exposure stimuli with an expected reduction in risk of harm.³⁴

Und natürlich geht das Ganze neben dem Argument einer gesteigerten Empathie für die Belange Betroffener einher mit der Option, dass das derartig nach außen Gekehrte psychiatrischer Patienten von diesen selbst auch eigens kommentiert werden kann – im Zuge eines Erfahrungsabgleiches, der die Leistungsfähigkeit der technischen Nachstellung unter Beweis stellt.³⁵ Ähnlich wie in den Episoden bei Pötzl und Hoff im ersten Drittel des 20. Jahrhunderts und mit Blick auf das dort beschriebene Zeitmanipulationspotential der Optikmedien ihrer Zeit attestieren mitteilssame Patienten die Gelungenheit der technischen Nachstellung ihrer Symptome. »The patient was very impressed with the simulation, and commented that it effectively re-created the same emotions that she experienced on a day-to-day basis during psychotic episodes.«³⁶ Damit ist die Psychose endlich ins Zeitalter ihrer technischen Reproduzierbarkeit überführt: Im Zuge ihrer Virtualisierung und dank der Auskunftsfreudigkeit der Betroffenen wird sie so zu einem mühelos zugänglichen und allseits erfahrbaren Allgemeingut.

Aber derartige Anwendungen bleiben nicht darauf beschränkt, Psychopathologien einfach nur nachzustellen. Sie treten mit dem Anspruch an, Erfahrungsräume zu erweitern und erlebbar zu machen, was in der Literatur der Moderne einmal als das Konzept des anderen Zustandes beschrieben wurde – und zwar in dezidiert Nähe zu einem bestimmten Stand der Mediatisierung von Lebenswelten. Wie oben gezeigt, waren es Etüden der Entrationalisierung, wie sie Apologeten der Moderne wie Robert Musil oder Max Bense an jenen Beschleunigungen von Erfahrung herausgearbeitet haben, die zu einem Versagen der Kontrolle durch das Bewusstsein und in Folge dessen zu dem führten, was Musil als »Durchbruch durch die bewußte Person« veranlagt. Für Musil und seinen Blick auf eine unängstliche Moderne lagen diese verdichtet vor etwa in den Zeitfolgen des Sportes oder der beschleunigten Erfahrung hinterm Steuer – und nicht zuletzt in der Erfahrungswelt eines psychopathischen Sexualtriebträters, wie er sie im *Mann ohne Eigenschaften* in der Figur Christian Moosbruggers verhandelt und verewigt hat. Im berühmten Kapitel *Moosbrugger denkt nach* wird deutlich, dass neben den Halluzinationen des Protagonisten vor allem sein Umgang mit Sprache aus den Fugen geraten ist. Im Moment des Verlustes an Bedeutungen und des Wörtlichnehmens von Wörtern wird eine Brachialität von Weltversetzung sichtbar, ähnlich

34 Banks, Jasmine/Ericksson, Geoffery/Burrage, Kevin u.a. (2004), »Constructing the hallucinations of psychosis in Virtual Reality«, in: *Journal of Network and Computer Applications* 27/1, S. 1-11. Vgl. dazu für den Fall posttraumatischer Belastungsstörungen Friedrich, Kathrin (2016), »Therapeutic Media: Treating PTSD with Virtual Reality Exposure Therapy«, in: *MediaTropes* 6/1, S. 86-113.

35 Vgl. dazu die Stellungnahmen bei J. Banks/G. Ericksson/S. Ivermee u.a., *A Virtual Environment to Re-create*.

36 J. Banks/J. Tichon/G. Ericksson u.a., *A Virtual Environment to Simulate* (o. Pag., Zitat im Abstract). Zu simulierten Halluzination unter Verwendung von DeepDream vgl. auch Rastelli, Clara/Greco, Antonio/Kenett, Yoed N. u.a. (2022), »Simulated visual hallucinations in virtual reality enhance cognitive flexibility«, in: *Scientific Reports* 12/4027, S. 1-14.

derjenigen, die Luhmann anlässlich einer instabilen Spracherfahrung beschrieben und am Beispiel der verwechselten Äpfel und Computer veranschaulicht hat.³⁷

Moosbrugger stand immer mit den Beinen auf zwei Schollen und hielt sie zusammen, vernünftig bemüht, alles zu vermeiden, was ihn verwirren konnte; aber manchmal brach ihm ein Wort im Munde auf, und welche Revolution und welcher Traum der Dinge quoll dann aus so einem erkalteten, ausgeglühten Doppelwort wie Eichkätzchen oder Rosenlippe!³⁸

Aber auch der Umgang mit Bewusstseinsverlusten wie im Fall des Sturzes von Wolkenkratzen ist geeignet, die überdies auch ästhetische Anmutung einer anderen Zeitlichkeit sichtbar zu machen – wie im Fall des unheimlich schnell rotierenden Romans *Sekunde durch Hirn*. Die Durchdringung der Lebenswelt mit Zeiteinheiten wie der Zehntelsekunde beschrieb Walter Benjamin gegenstrebig zu gängigen Modernediagnosen, die in der Beschleunigung und Versachlichung Menetekel einer bedrohlichen Moderne sahen, als eine Befreiung aus den Klauen einer bestimmten Zeitlichkeit und wie gezeigt mit der Option auf die Erschließung neuer Welten – geschuldet den Zeitmanipulationsmöglichkeiten der Kinematographie: »Da kam der Film und hat diese Kerkerwelt mit dem Dynamit der Zehntelsekunden gesprengt, so daß wir nun zwischen ihren weitverstreuten Trümmern gelassen abenteuerliche Reisen unternehmen.«³⁹

Hier, in einem Bereich, der jenseits der natürlichen Wahrnehmung ein Optisch-Unbewusstes begründet und der semantisch zu Beschreibungen Anlass gibt, die jenseits der Zeiterfahrung eines bürgerlichen Lebens liegen, finden Kontrollverluste und auf deren Grundlage Alteritätserfahrungen statt, die einen Eigensinn, eine Anmutung gelebten Lebens haben und die auf ihre Weise die Erfahrung von Entwürfungen ermöglichen, die einem bestimmten Stand in der Technisierung von Lebenswelt geschuldet sind. Diese lassen sich weder auf das Konto von individuellen Pathologien wie der Spaltung oder der Depersonalisation noch auf das von kulturellen Pathologien wie der Entfremdung und der Entnaturalisierung buchen.⁴⁰ In einer Passage des Essay *Durch die Brille des Sports* führt Musil minutiös moderne Menschentypen an, deren Bewusstsein vor der Komplexität beschleunigter Bewegungsformen kapituliert.

Wenn der Mensch am Steuer eines sehr schnell fahrenden Kraftwagens sitzt, wenn er scharfe Flugbälle placiert oder ein Florett führt, hat er in einem kleinsten Zeitraum u. mit einer Schnelligkeit, wie sie im bürgerlichen Leben sonst nirgends vorkommt, so vie-

37 Zu den Details vgl. Musil, Robert (1981), *Der Mann ohne Eigenschaften*, hg. von Adolf Frisé, Reinbek: Rowohlt, v. a. S. 235-242, sowie Wolf, Norbert Christian (2010), »Warum Moosbrugger nicht erzählt. Zur metanarrativen Funktion psychopathologischen Wissens in Musils *Mann ohne Eigenschaften*«, in: *Jahrbuch der Deutschen Schillergesellschaft* 54, S. 330-362.

38 Musil, Robert (1981), *Der Mann ohne Eigenschaften*, hg. von Adolf Frisé, Reinbek: Rowohlt, S. 241.

39 W. Benjamin, *Kunstwerk*, S. 499f.

40 Vgl. zu prominenten Verhaltungen dazu Diesel, Eugen (1939), *Das Phänomen der Technik, Zeugnisse, Deutung und Wirklichkeit*, Leipzig, Berlin: Philipp Reclam Jun. Verlag, VDI-Verlag.

le genau aufeinander abgestimmte Akte der Bewegung u. Aufmerksamkeit auszuführen, daß es ganz unmöglich wird, sie mit dem Bewußtsein zu beaufsichtigen.⁴¹

Was Musil beschreibt, und was Benjamin zu seiner Übertragung der Psychoanalyse auf das Optisch-Unbewusste veranlasst, ist eine Psychedelik ohne Drogen, eine Psychedelik der zugrunde gelegten Lebenswelt, eines Bezugssystems aus Vertikal-anordnungen und Drehmomenten, aus Beschleunigungen und Zeitmanipulationen, aus Kontrollverlusten und Entrationalisierungen – kurz eine Lebenswelt jenseits der Zeitauffassung eines bürgerlichen Lebens, eine Lebenswelt jenseits der semantischen Behähigkeit jenes Gutbürgerlichen, das dem Begriff so sehr anhaftet.

Auf diese Optionen der Psychedelik und korrespondierend damit der Cyberdelik wird im Folgenden zu achten sein. Ein Beitrag unter dem Titel *A Deep-Dream Virtual Reality Platform for Studying Altered Perceptual Phenomenology* stellt dazu eine veritable *Hallucination Machine* vor, mit der es möglich sein soll, andere Zustände zu erleben und zu studieren – und das, ohne mit deren nachteiligen Effekten konfrontiert zu werden, die sich bei Normalgebrauch einstellen. Die Versetzung in die andere Welt ist frei von Nebenwirkungen und frei von jenen Momenten des Brachialen-Versetzten oder des Traumatisch-Geschleuderten, mit denen Luhmann die waghalsige Unvermitteltheit solcher Übergänge am Beispiel von Rede und Schrift beschrieben hat. Die Bildproduktion der *hallucination machine* fußt auf Deep Convolutional Neural Networks und die Expositionstechnik folgt dem Sachstand der Immersion. Beide wirkmächtige Verfahren, so heißt es dort selbstbewusst, werden zum Zwecke des Halluzinierens zusammen in einer gemeinsamen Vorrichtung verbaut.

Thus, simulating phenomenological aspects of altered states in the absence of these other more general effects provides an important experimental tool for consciousness science and psychiatry. Here we describe such a tool, which we call the *Hallucination Machine*. It comprises a novel combination of two powerful technologies: deep convolutional neural networks (DCNNs) and panoramic videos of natural scenes, viewed immersively through a head-mounted display (panoramic VR). By doing this, we are able to simulate visual hallucinatory experiences in a biologically plausible and ecologically valid way.⁴²

Das, was als biologisch plausibel und ökologisch valide mittels Head Mounted Display zugestellt und erfahrbar wird, ist unter der Bezeichnung *deep-dream virtual reality platform* ausgewiesen. Bei ihrem neuartigen Tool machen die Autoren, wie sie selbst schreiben, von jenem Algorithmus Gebrauch, der in Googles DeepDream Verwendung fand und der zu jener Redeweise vom *inceptionism* führte, die in der entsprechenden Publikation aus dem Jahr 2015 festgeschrieben (*Inceptionism: Going Deeper into Neural Networks*) und entsprechend variiert wurde.

41 Musil, Robert (1978), »Durch die Brille des Sports« [1925/26 oder später], in: Gesammelte Werke, Band 2: Prosa und Stücke, kleine Prosa, Aphorismen, Autobiographisches, Essays und Reden, Kritik, hg. von Adolf Frisé, Reinbek: Rowohlt, S. 792-795, hier: S. 793. Zur Umschrift gelangt dieser Passus in Musils *Der Mann ohne Eigenschaften*, hg. von Adolf Frisé, Reinbek: Rowohlt, S. 28.

42 K. Suzuki/W. Roseboom/D. J. Schwartzman u.a., *A Deep-Dream Virtual Reality Platform* (Hervorhebung in der Vorlage).



Abbildung 49: Deep Dream Bildgenerierung (nach Suzuki et al. 2018)

Deep Dream works by clamping the activity of nodes at a user-defined layer in the DCNN and then inverting the information flow, so that an input image is changed until the network settles into a stable state. What is striking about this process is that the resulting images often have a marked »hallucinatory« quality, bearing intuitive similarities to a wide range of psychedelic visual hallucinations reported in the literature (Shanon, 2002).⁴³

Damit werden Bilder prozessiert, die nicht zuletzt dank bestimmter Ähnlichkeitsrelationen die Qualität und die Anmutung einer psychedelischen (oder eben einer cyberdelischen) Erfahrung haben. Dank der verwendeten Technik wird, wie die Autoren eigens anmerken, ein Unterschied in den Halluzinationen eingezeichnet, der verstärkt auf deren figurative und weniger auf deren geometrische Ausformung abzielt.

In contrast to the previous work, which focused primarily on simulating geometric hallucinations (e.g. Froese, Woodward, & Ikegami, 2013), our setup enables the generation of more figurative hallucinations, thanks to the hierarchical nature of DCNNs.⁴⁴

Das Tool löst in aller sachdienlichen Pragmatik ein, was als Grundzug einer Ökonomie der Virtualität gelten kann. Diese lässt sich mit der englischen Redewendung *You can't have your cake and eat it too* bezeichnen, die anlässlich eines Buches über virtuelle Realität als letztes (oder sollte man schreiben als letztthinniges) Narrativ des Menschen auf symptomatische Weise angepasst wird. In Thorsten Botz-Bornsteins Buch *Virtual Reality: The Last Human Narrative?* findet sich ein Abschnitt, der die Redewendung aufgreift und sie unter der Überschrift *Cloned Realities: Eat Your Cake and Dream it* für an-

43 Suzuki, Keisuke/Roseboom, Warrick/Schwartzman, David J. u.a. (2018), »Hallucination Machine: Simulating Altered Perceptual Phenomenology with a Deep-Dream Virtual Reality platform«, in: ALIFE 2018: Proceedings of the 2018 Conference on Artificial Life, Tokio: IEEE, o. Pag.

44 Ebd.

dere Realitätsverständnisse anschlussfähig hält.⁴⁵ Das mit der Ausgangsformulierung vom verzehrten oder aufgesparten Kuchen umrissene Entscheidungsdilemma wird im Virtuellen umgebar: Das Virtuelle erlaubt eine Ökonomie der verlustfreien Teilhabe, in der die Negation ihre Geltung verliert, in der beide Optionen möglich sind. Das Virtuelle gibt Raum für eine Ökonomie, in der es möglich ist, den Kuchen zu verzehren und ihn aufzubewahren, ihn zu genießen, ohne ihn zu verlieren – ein Raum, in dem der Dezisionismus aufgehoben ist und nicht greift. Neben das Moment der Konsequenzreduktion als Kennzeichen virtueller Realitäten tritt das Moment einer außer Kraft gesetzten Entscheidung. Damit wird es möglich, zu experimentieren, ohne die damit verbundenen Begleiterscheinungen in Kauf nehmen zu müssen, also ohne Gefahren, ohne Verschleiß, ohne Beeinträchtigungen der Umwelt.⁴⁶ Das Gehirn kann weggeblasen oder zur Suppe gemantscht werden und doch bleibt es intakt. »We show that the system induces visual phenomenology qualitatively similar to classical psychedelics. The Hallucination Machine offers a valuable new technique for simulating altered phenomenology without directly altering the underlying neurophysiology.«⁴⁷

Der Umgang mit Halluzinationen bleibt nicht auf allgemein gehaltene Nachstellungsoptionen und auf allgemein gehaltene Parallelen beschränkt. So erfolgt in einer Arbeit über Schizophrenie die Visualisierung einer religiös konnotierten Halluzination, konkret die einer Marienerscheinung. Was dabei neben allen ikonographischen Details ins Auge sticht, ist das räumliche Arrangement, ist die nach außen gestülpte Marienerscheinung doch auffällig gerahmt. Die Assemblage an alltäglichen Dingen, etwa an Möbeln, innerhalb derer die Erscheinung angeordnet ist, verdoppelt Muster und auch inzwischen gewohnte Darstellungsgesten, die für virtuelle Gedächtnisarchitekturen typisch sind, und sie verdoppelt die stereotypen Anordnungen des gewöhnlichen Wohnens, die ob ihrer ausgestellten Schlichtheit so sehr ins Auge stachen.⁴⁸ Die zurückgenommene Gestaltung der Wohnszenarien wird zur Bühne für das Außergewöhnliche, für Halluzinationen und Erscheinungen aller Art. Fast könnte man den Eindruck gewinnen, die zur Einstellung der Figuren bereiteten Räume seien zu nichts anderem bestimmt – als Nischen für die Exponierung des Besonderen. Die Wohnräume wirken wie Kirchen nach dem Bildersturm. Dabei unterliegen selbst die religiösen Halluzinationen einem Diktat des Auffallenden. So werden die beiden Marienerscheinungen ihrerseits qualifiziert (u.a. als abusive, also mit einem eigens typographisch markierten Adjektiv, dessen Bedeutung zwischen ausfallend, schmähend und missbräuchlich liegt).

Die Imagination – ob als normal veranlagt oder als pathologisch ausgewiesen – verlässt den Menschen und wird Teil eines Technisch-Imaginären, in dessen Vollzug sie auch nach außen tritt, wie es etwa anlässlich der Kunstinstallationen des britischen Medienkünstlers Memo Akten unternommen und mit entsprechender Finesse reflexiv begleitet wird. Aktens Arbeiten fügen sich nahezu bruchlos in die Logik jener Bildbrüche, die weder durch ein Denken in Repräsentationen noch in Ähnlichkeiten

45 Vgl. dazu T. Botz-Bornstein, *Virtual Reality*.

46 Vgl. dazu Luo, Tianren Luo/Zhang, Mingmin/Pan, Zhigeng u.a. (2020), »Dream-Experiment: A MR User Interface with Natural Multi-channel Interaction for Virtual Experiments«, in: *IEEE Transaction on Visual Computer Graphics* 26/12, S. 3524-3534.

47 K. Suzuki/W. Roseboom/D. J. Schwartzman u.a., *Hallucination Machine*.

48 Vgl. dazu J. Tichon/J. Banks/P. Yellowlees, *Development*.

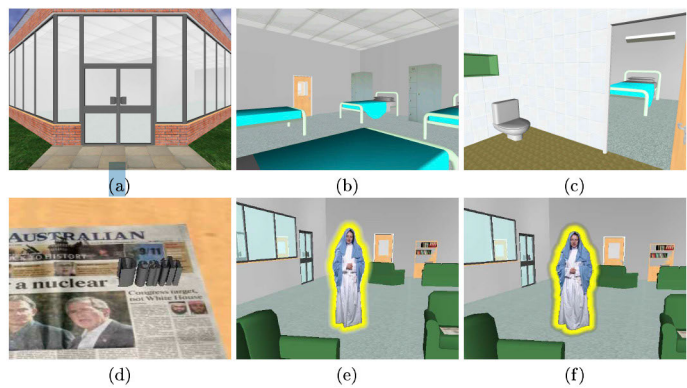


Fig. 4. Views of the psychiatric ward and hallucinations: (a) front entrance; (b) bedroom; (c) bathroom; (d) “death” hallucination; (e) Virgin Mary hallucination; (f) “abusive” Virgin Mary hallucination.

Abbildung 50: Marienerscheinung (nach Banks et al. 2003)

aufzugehen scheint. Das, was etwa die Installation *Learning to see* aus dem Jahr 2017 vor Augen stellt, ist den Verwechslungen von Bananen und Toastern, von Äpfeln und Computern nicht unähnlich. Sie gipfelt, um das Ergebnis vorwegzunehmen, in einer Assemblage von maximal alltäglichen Gegenständen wie Kabeln, Ladegeräten, Kopfhörern und Autoschlüsseln mit einem sich ästhetisch gerierenden Blumenarrangement. Dieser Vorgang folgt dabei, wie es in einer medientheoretisch gerahmten und gut informierten Rekonstruktion durch Claudio Celis Bueno und María Jesús Schultz Abarca heißt, einer durchaus eigenen Logik:

According to Memo Akten (2017), the *Learning to See* interactive installation consists of »a number of neural networks« which »analyse a live camera feed pointing at a table covered in everyday objects« (Fig. 1). »Through a very tactile, hands-on experience«, Akten (2017) adds, »the audience can manipulate the objects on the table with their hands, and see corresponding scenery emerging on the display, in real-time, reinterpreted by the neural networks«.⁴⁹

Während das Ausgangsbild der Installation als *live feed* dient und, wie es in der Rekonstruktion heißt, mit beträchtlichem Aufwand jenes Alltägliche benutzt, das ob seiner Gewöhnlichkeit habituell unter dem Radar einer bewussten Aufmerksamkeit rangiert, betreibt das andere Bild eine Reinterpretation. Diese erfolgt mittels neuronaler Netze, die, wie der Künstler selbst ausführt, mit eigenen Datensets trainiert wurden. Der ganze Vorgang ist zeitlich getaktet, wechselt nach jeweils dreißig Sekunden den Bezugsrahmen, der durch die vier natürlichen Elemente (veranlagt als Wasser, Luft, Erde, Universum) getriggert wird und damit das zugrundliegende Muster der Bildmanipulation steuert – Voreinstellungen vergleichbar, mit denen bei künstlichen Verfahren der Bildproduktion Stile, Epochen oder Darstellungsarten gewählt, also die Gestaltung nach Maßgabe von Rembrandt und van Gogh, Klassik und Barock oder

49 C. Celis Bueno/M. J. Schultz Abarca, Akten's *Learning to See*.

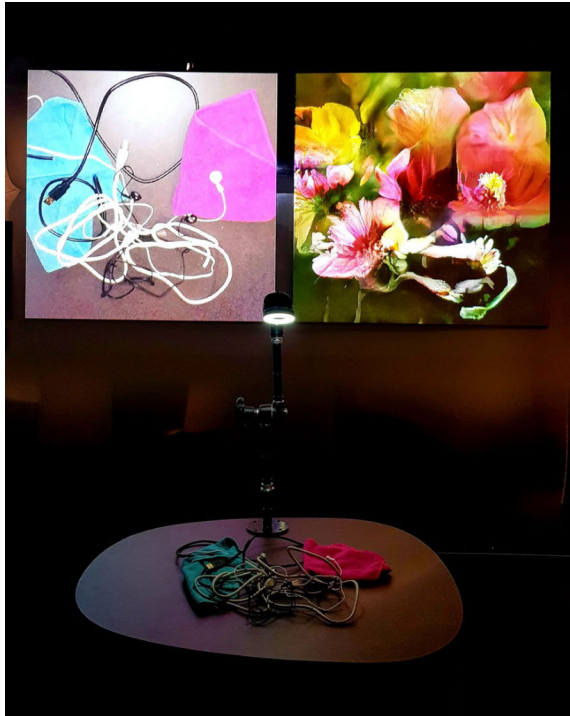


Abbildung 51: Sehen lernen mit Memo Akten (nach Bueno/Abarca 2021)

Abstraktion und Cartoon, Fotorealismus und Neonpunk erfolgt. Zudem ist er auch noch mit dem Verhalten und Interaktionsmuster der Rezipient:innen rückgekoppelt, die ihrerseits die Gegenstände des Ausgangsbildes händisch manipulieren können.

Akten (2017) explains that every thirty seconds the scene changes between different networks trained on five different datasets: the four natural elements: ocean and waves (representing »water«), clouds and sky (representing »air«), fire, flowers (representing earth and life); and images from the Hubble space telescope (representing the universe, cosmos, ether, void or God). Every object that enters the camera's visual field alters the output of both images. The participant's hands manipulating the objects on the table are »seen« by two different forms of machine vision: a traditional type of video image and a novel type of neural network image. The layout of these images – one next to the other – stresses how a single input produces two different outputs depending on the type of vision technology. This creates a tension between the »objective« and »neutral« image of the live feed on the left and the biased image on the right in which the algorithm can only see »through the filter of what it already knows« (Akten 2017). In this sense, Akten's art installation functions as a concrete illustration of the core aspects of machine bias: by opposing an objective image to a biased one, Learning to See is addressing how the training datasets directly modify the output of the neural network.⁵⁰

50 Ebd.

Was in und mit dieser Installation illustriert und verhandelt wird, sind nicht nur die Arbeitsweisen neuronaler Netze (und als unterschwellige medienhistorische Reminiszenz die Doppelexpositionslogik, die in einer bestimmten Phase der Kunstgeschichte die Aufmerksamkeit auf die technische Vermittlung des kunstwissenschaftlichen Sehens zu richten wusste).⁵¹ Illustriert und verhandelt werden neben den Reaktionen auf die technischen Möglichkeiten die reflexiven Begleitprogramme und die kritischen Interventionen, die auf ihre Weise Eingang in die Kunst und in die Umgangsweisen mit Kunst gefunden haben. Die Installationen Memo Aktens (mitsamt den eigenen theoretischen Stellungnahmen etwa zu DeepDream) sowie die Bezugnahmen auf diese leisten auf ihre Weise doch etwas Ähnliches wie die Vertauschung von Bananen und Toaster – und doch machen sie es anders (und das bezieht sich nicht nur auf den Befund der Dynamik im Gegensatz zur statischen Gegenüberstellung von Ausgangsbildern und ihren Fehleinschätzungen).⁵² Sie verfolgen kein Narrativ, das mit einer gewissen Häme die Grenzen der Mustererkennung an besonders drastischen Beispielen nachvollzieht, sei es, um das Ungenügen der Mustererkennungsverfahren oder um die Gefahren für den Alltag vor Augen zu stellen, sie verfolgen vielmehr den Anspruch, das Sehen selbst – und zwar das Sehen von Menschen wie von Maschinen zum Gegenstand und zwar zum Gegenstand jeweiliger technischer Vermittlungen zu erheben.

Die Arbeiten Aktens haben gerade ob solcher Selbstbezüglichkeiten immer wieder die Aufmerksamkeit nicht nur der Kunstwissenschaft erregt und dienen wiederholt als Bezugspunkt sowohl medienhistorischer als auch medientheoretischer Überlegungen. So erfolgt anlässlich von *Learning to See* ein historischer Brückenschlag, der die dort vorgenommene Übersetzungsleistung von einem Ausgangsbild in ein umgewandeltes Zwischenendbild auf der Grundlage neuronaler Netzwerke in den größeren Rahmen von Übersetzungsleistungen stellt, die zwischen physikalischen Zuständen statthaben und der Beförderung von synästhetischen respektive multisensorischen Erfahrungen gelten. In der Rückverfolgung entsprechender Programme wird die Bandbreite jener *visual instruments* durchlaufen, die von den Farbenklavieren zu Beginn des 18. Jahrhunderts über die avantgardistischen Instrumente der klassischen Moderne bis zur gegenwärtigen Medienkunst reicht. Den aufgezählten Verfahren ist bei allen Unterschieden in der technischen Realisierung ein Moment der Zuordnung und der Übersetzung eigen – wie technisch komplex dieses auch immer gestaltet sein mag.

The concept of visual instruments can be traced back to Louis-Bertrand Castel's ocular harpsichord and similar mechanical visual instruments from the eighteenth and nineteenth centuries. Early twentieth-century abstract animators and filmmakers such as Norman McLaren and Oskar Fischinger, analog computer animation pioneer John Whitney Sr. and analog video synthesizers such as the Rutt-Etra are also key influences. The emergence of digital technologies allowed artists such as Myron Krueger, David

51 Zu den Verfahren und der Geschichte der Doppelprojektion vgl. Reichle, Ingeborg (2002), »Medienbrücke«, in: kritische Berichte 1/1 (Themenheft: Die Bildmedien der Kunstgeschichte), S. 40–56.

52 Zu Anwendungen auf Bewegtbilder vgl. Moniz, Joel Ruben Antony/Kang, Eunsu/Póczos, Barnabás (2019), »LucidDream: Controlled Temporally-Consistent DeepDream on Videos«, arXiv:1911.11960v1 [cs.CV] 27 Nov 2019.

Rokeby, Ed Tannenbaum, Rebecca Allen, Scott Snibbe, Colan Levin and Camille Utterback, to name a few, to work in this subfield of interactive media art, exploring new modes of expression.⁵³

Neben solchen historischen Spannungsbögen geraten aber auch theoretische Anliegen mit auf die Agenda. So wird die Installation *Learning to See* über alle technischen Details der Umsetzung und einer daran geknüpften Mediengeschichte hinaus zum Anlass, einschlägige Theorieangebote einzubeziehen und vor diesem Hintergrund nun ihrerseits zu kontextualisieren – etwa mit Bezug auf Vorgaben, die Autor:innen wie Walter Benjamin, Martin Heidegger, Edmund Husserl, Vilém Flusser, Bernard Stiegler oder Hito Steyerl verpflichtet sind.⁵⁴ Damit haben die Bildumtriebe der Maschine eine systematische Adresse, einen Bezugspunkt, mit dessen Hilfe sie sich beschreiben und unter Konzepten wie dem einer *machinic imagination* und eines *machinic unconscious* auch theoretisch erschließen und in entsprechende Denktraditionen technisierter Lebenswelten einbringen lassen. Auf diese Weise gelangt noch einmal ein Teilargument dieses Buches zur Geltung – der Befund, dass mit dem hier verwendeten Konzept der Lebenswelt (und über alle verharmlosenden Anwendungen seiner gutbürgerlichen Situierung hinaus) nicht nur eine Erweiterung der Agenten und ihrer jeweiligen Agenturen gehört, sondern dass eben auch die reflexiven Begleiterscheinungen, also das Nachdenken und die Theoretisierung selbst als inhärenter und legitimer Teil von Lebenswelt gelten und entsprechend miteinzubeziehen sind. Sie stellen damit kein reflexives Beiwerk dar, sondern sind fundamentaler Teil dessen, was Max Bense einmal mit Blick auf die Technik als die Selbsteinrichtung des Menschen innerhalb der Natur genannt hat – natürlich nicht als die einzige Möglichkeit, wohl aber als einer der »planmäßigsten und vielleicht erfolgreichsten der bisher möglichen Versuche dieser Art«.⁵⁵ Damit ist jene doppelte Bewegung vorgezeichnet, die in der Arbeit über Akten's Installation angelegt ist:

Likewise, in an age dominated growingly by machine learning technologies, it is possible to speak not only of machine vision but also of a machinic imagination and a machinic unconscious, two notions that can be illustrated through Akten's art installation.⁵⁶

53 Akten, Memo/Fiebrink, Rebecca/Grierson, Mick (2019), »Learning to See: You Are What You See«, in: SIGGRAPH 2019 Short Art Papers, New York: ACM, Art. 13, S. 1–6, hier: S. 3. Ein Sammelbecken entsprechender Überlegungen stellen die Farbe-Ton-Forschungen dar, die ab 1927 Gegenstand mehrerer Kongresse in Hamburg waren. Vgl. dazu etwa Anschütz, Georg (Hg.) (1927), *Farbe-Ton-Forschungen*. Band 1, Leipzig: Akademische Verlagsgesellschaft.

54 Akten arbeitet verstärkt mit VR-Anwendungen (Marshmallow Laser Feast). Zum Anschluss an den Traum Akten, Memo (2015), »Deepdream is blowing my mind«, in: Medium vom 10.07.2015, online unter <https://memoakten.medium.com/deepdream-is-blowing-my-mind-6a2c8669c698> (letzter Zugriff: 01.07.2020). Solche Redeweisen wären zu positionieren zu solchen, die das Unbewusste als Kategorie der Medienanalyse benutzen. Vgl. dazu Schmidgen, Henning (1997), *Das Unbewußte der Maschinen. Konzeptionen des Psychischen bei Guattari, Deleuze und Lacan*, München: Fink.

55 Bense, Max (1998), »Kybernetik oder die Metatechnik der Maschine« [1951], in: Bense, *Ausgewählte Schriften*, Band 2, S. 429–446, hier: S. 446.

56 C. Celis Bueno/M. J. Schultz Abarca, Akten's Learning to See.

Alte Theoriekonzepte, die ganz bestimmten und historisch rekonstruierbaren Sachständen in der Technikentwicklung geschuldet sind (und die nicht, wie Blumenbergs Universaltheorie, an einem apriorischen Wesen der Technik ansetzen, das allen phänomenalen Ausprägungen vorgelagert erscheint und diese auf ihrem Weg durch die phänomenalen Ausgestaltungen begleiten soll), lassen sich im Zuge dessen entsprechend aufeinander beziehen, sie lassen sich upgraden und damit an die veränderten Gegebenheiten anpassen – das betrifft nicht zuletzt quantitative Aspekte wie den der Massierung. Das Unbewusste der Freud'schen Psychoanalyse, das Optisch-Unbewusste von Benjamins epistemischer Medientheorie und das Algorithmisch-Unbewusste von Aktens Kunstinstallationen begründen eine Reihe, die auf ihre Weise menschliche und technische Sachstände miteinander verschränkt – nicht mit dem Anspruch, irgendwelchen Überlegenheiten das Wort zu reden, sondern mit dem Anspruch, deren intrinsische Verquickung nach außen zu kehren und sichtbar werden zu lassen. »Learning to See highlights a profound connection between machinic vision, machinic imagination and machinic unconscious.«⁵⁷ Diese Befundlage erlaubt epochale Bezugnahmen, die genau solchen Prozessen in den medialen Umgebungen Rechnung tragen. So heißt es bei Bueno und Schultz Abarca, in einem Abschnitt mit der Überschrift *Inceptionism*, unter Hinweis auf Aktens Umgangsweise mit Googles DeepDream-Algorithmus und mit Blick auf Hito Steyerls *A sea of data: apophenia and pattern (mis)recognition*: »In this context, she adds, »Walter Benjamin's optical unconscious has been upgraded to the unconscious of computational image divination« (Steyerl 2017: 57)«. ⁵⁸

Damit gerät noch einmal das Normalisierungs- und Objektivitätsphantasma in den Blick, dessen Stabilisierungsfunktion und -tendenz wie ein Treiber für Prozesse der Normalisierung dient. Entgegen dem Hang zur Normalisierung wird etwas ganz anderes sichtbar. Es hat den Anschein, als ob das Objektivitätsversprechen der großen Zahl ins Leere liefe.

Machine learning, through its immense capacity for pattern recognition, has become the latest instrument for pursuing pure objectivity. Apologists of this technology claim that thanks to machine learning and big data, theories will no longer be needed to explain the world (Anderson 2008). Instead, a real-time picture of reality will be permanently accessible offering »a whole new way of understanding the world« (Anderson 2008). ⁵⁹

Was in diesem Zitat mit dem bibliographischen Verweis auf Chris Anderson aufgerufen ist, kommt dabei einem Infragestellen gängiger Formen der Wissenschaft und damit auch dem einer etablierten Form von Theoriebildung gleich: Es ist, als ob der Stand der Datenverarbeitung einen bestimmten Stand des Denkens überflüssig machte, es ist, als ob eingespielte Formen der Reflexion dem Sachstand dessen, was sie beschreiben wollen, nicht oder nicht mehr oder nicht mehr länger angemessen wären. So jedenfalls beschreibt es Anderson unter dem sprechenden Titel *The End of Theory. The Data Deluge Makes the Scientific Method Obsolete*. ⁶⁰

57 Ebd., S. 1184.

58 Ebd.

59 Ebd., S. 1179.

60 Vgl. dazu Anderson, Chris (2008), »The End of Theory. The Data Deluge Makes the Scientific Method Obsolete«, in: Wired vom 16.07.2008.

Im Anschluss an die Möglichkeiten einer Teilhabe an datenmäßig nicht mehr fass- und greifbaren Phänomenen wird die Frage virulent, ob es sich um Formen der Verhaltung oder um solche einer prinzipiellen Verunmöglichung handelt: In beiden Fällen betrifft es nichts weniger als das wissenschaftliche Geschäft, es betrifft dabei allerdings keine pauschalen Verwerfungen, wie der Titel vom herannahenden Ende der Theorie doch einigermaßen reißerisch nahelegt, sondern ein Reagieren der Wissenschaft auf diese Lage, veränderte Methoden und Zugriffsweisen, das *distant viewing* und *distant reading* der Digital Humanities, es betrifft aber auch neue und genuin eigene (und im Zuge einer Cultural Analytics auch eigens eingeforderte) Gegenstandsbereiche und veränderte Formen der Fächerkonstellierungen; schlussendlich betrifft es das, was unter dem Titel von Fakultäten und Universität wirkmächtig über lange Zeit und auf eine bestimmte Weise gehegte Räume und Organisationsformen des Wissens darstellte und darstellen durfte. Es betrifft, was man vielleicht mit Virtual Humanities oder Virtual Posthumanities einmal besser wird beschreiben können.⁶¹ Es betrifft, was unter der Bezeichnung »Multiversität« eine Alternative zur verfestigten Institution Universität darstellt und darunter subsumiert, was Rosi Braidotti als Herausforderung für posthumane Geisteswissenschaften skizziert.⁶²

Eine Welt, die derart und derartig konzeptualisiert ist, eine Welt, die sich auf ihre Weise auch nicht mehr denken und theoretisch fassen lässt, gibt Anlass für unterschiedliche Bezugnahmen: Neben Diskursen, die den Flow einfach nur feiern und eine Renaissance des Cyberdelischen ausrufen, treten solche, die um Hegung bemüht sind, die auf Überbürdung und Überforderung hinweisen, die auch Momente, wie vorhin ausgeführt, der Vergleichgültigung und des Beliebigwerdens in ein oftmals kulturkritisch abgesteckte Feld übertragen. Neben dem Vorwurf einer unablässig Sinn- und Bedeutungskorrosion melden sich aber auch Stimmen zu Wort, die auf Momente der Verdächtigung und der Paranoia abzielen. Wenn Dinge nicht mehr die Dinge sind, die zu sein sie vorgeben, wenn Prozesse ablaufen, die sich dem menschlichen Wahrnehmen und Denken nicht punktuell und in einer nachvollziehbaren Kasuistik, sondern habituell entziehen, wenn Bedeutungszuweisungen erfolgen, die der Logik des Wolkenbilderlesens und der Inceptionslogik von DeepDream entsprechen, dann ist das hinreichend Anlass für Konzeptbildungen, die das Nicht-Wissen und die Ausprägungen eines Unbewussten in Position bringen – Größen, die sich nicht länger psychologischen Kategorien verdanken. Das ist der Moment, in dem sich das Unbewusste technischer Medien unbeschadet des jeweiligen Standes der Technisierung Bahn bricht, das ist der Moment, in dem Anpassungen von Theorieleistungen, die unter bestimmten historisch angebbaren Medioumgebungen gewonnen wurden, an veränderte Lebenswelten geraten – und das auf eine Weise, die deren Komplexität sichert. Medien sind damit gerade nicht auf das nostalgische Konto dessen zu buchen, was sie in der gemütlichen Anmutung historischer Rückschauen einmal waren und dessen man sich, wenn überhaupt, lediglich im flüchtigen Modus jenes von Luhmann so genannten touristischen Verhältnisses anzunähern wusste.

Die Fluktuation gerade auch der Theorieleistungen macht Selbstappelle und Warnhinweise notwendig, geraten doch auch die eigenen Perspektivierungen allen guten Vorsätzen zum Trotz gelegentlich in den Sog von Narrativen des Fortschritts

61 Zu den Details und zu einer Neuausrichtung der Universität vgl. S. Rieger, Virtual Humanities.

62 Vgl. dazu R. Braidotti, Posthumanismus, v.a. S. 186ff.

und der Teleologie. Wie weit entfernt ist uns und den hochfrequenten Taktungen unserer Computerprozessoren eine Welt, in der das Antibürgerliche einer Zeitauffassung dadurch sichtbar werden konnte, das es in der Maßeinheit von Zehntelsekunden angegeben wurde? Wie weit entfernt sind wir bald von einer Welt, die mit ihren Vertikalanordnungen »die Dimension des Transits von Akten und Referenten, von Entscheidungen und Managern, von Operationen und Stäben« verkörpern – als Anleihen an einen bestimmten Typ bürokratischer Rationalität, der höchstens noch in Gestalt von Büromöbeln (und Manufactum sei es gedankt) als nostalgisches Relikt für die Staffage einer neusachlich anmutenden Innenarchitektur von Belang ist? Damit wird ein zeitdiagnostischer Blick auf die Technisierung einer Lebenswelt geworfen, dem in der Gegenwart das Unmittelbar-Geworden-Sein von Daten entspricht: Blumenbergs Transit von Akten ist einem Zusammenleben mit welthaft gewordenen Daten gewichen (*Building a world of habitable bits*).⁶³

Die Semantik der Tiefe, die einem technischen Spezifikum der verwendeten Verfahren gilt, nämlich der Zahl und der Beschaffenheit der Layer, betrifft mit DeepDream nicht nur die Mechanismen der Bildorganisation, sondern erreicht mit Verfahren wie DeepFake oder DeepFool auch die Ebene der Störmechanismen. Mit der Rede vom *latent space* betreiben nicht nur die Theoretiker einen Anschluss an entsprechend anderweitig ausgearbeitete Konzepte des Unbewussten, vielmehr werden oder sind derartige Konzepte selbst regulärer Teil der Selbstbeschreibungssprachen derjenigen, die entsprechende Modelle bilden: Der *latent space* wird zu einer technischen Operationsbasis entsprechender Algorithmen, er wird aber auch zu einem phantasmatischen Raum scheinbar nicht zu erschöpfender Produktionen. Seine Latenz wird zur Voraussetzung für die Fülle der Resultate, die das Netz fluten.⁶⁴

Mit den neuen technischen Möglichkeiten macht sich ein Moment der Unabsehbarkeit und der Unsicherheit in einem lebensweltlichen Umfeld breit, das man sicher gestalten wollte. Es ist, als ob man sehenden Auges und doch auf eine gewisse Weise blind dafür einem Re-Entry des Irrationalen in das Feld der Rationalen beiwohnt – nur mit dem Unterschied, dass, was auch immer an Feerien und Bizarrerien von Bildern zutage tritt, nicht einer ästhetischen Überschussproduktion oder einer pathologischen Abweichung geschuldet ist, sondern einen Seiteneffekt der Selbststabilisierung der Systeme darstellt. Dem Traum der Maschine bleibt bei allen Performanzen des Bizarren, des Psychedelischen und des Cyberdelischen ein Moment der Systemoptimierung inhärent. Die Zugeständnisse an einen Überschuss, an einen kulturellen Mehrwert erweisen sich damit als ambivalent. Ernst und Belanglosigkeit, Konsequenz oder bloße Emergenz, kuriose Fehlleistung und folgenreicher Störungsvorsatz sind nicht voneinander zu unterscheiden – ein Moment, das einmal mehr dem Verdacht und den Verdächtigungen den Weg bahnt und das paranoide Potential im Spiel hält.⁶⁵

Ausgerechnet ein Text, der sich in seinem Titel mit dem Rorschachtest der psychologischen Königsdisziplin einer auf die Belange der menschlichen Diagnostik abzielenden Bildwahrnehmung verschrieben hat (*Interpreting Deep Learning: The Machine*

63 Takeuchi, Yuichiro (2014), »Building a world of habitable bits«, in: *Interactions* 21/6, S. 52-57.

64 Vgl. zu einer Erklärung der Wirkweise Hawranke, Thomas/Scherffig, Lasse (2024), »Hund. Gassi gehen im Latent Space«, in: Ina Bolinski/Thomas Hawranke/Stefan Rieger (Hg.), *Virtuelle Tiere. Lebewesen zwischen Code und Kreatur*, Bielefeld: transcript, S. 101-126.

65 Vgl. dazu noch einmal Eve Kosofsky Sedgwick's Konzept des *paranoid reading*.

Learning Rorschach Test?), bringt diese scheinbar paradoxe Lage mit ihren Zuschreibungsambivalenzen und Ununterscheidbarkeiten auf den Punkt. Das Komische und das Comic-Affine, das, was in der Rekonstruktion des Einbildungswissens Formen der Literatur und des Merkens betraf, was im Laufe dieser Rekonstruktion hinter altertümlich fremd anmutenden Beschreibungen wie der Rede vom Scherz und Frohspiel steckte, mit denen Autoren der klassischen Moderne wie Hellpach Beziehungsschichten beschrieben und in einem großangelegten Feld kultureller Bedeutungsproduktion verorteten, das, was als filmgenährte Zwangslosigkeit interkontinentale Kulturtypologien geprägt hat wie in Gieses *Girlkultur*, die habituellen Überzeichnungen technisierter Lebenswelten und das, was im Zuge dessen und in der Tradition der Mnemotechnik aus dem Rahmen fallen sollte und grotesk sein durfte, die Obsession für Ungewöhnliches und Auffälliges, für Vershobenes und Vershrobenes, für Witz und Unlogisches, für Bizarres und aus dem Kontext Gelöstes, für an den Haaren-Herbeigezogenes oder über Alltagsevidenzen mühelos Erschließbares – all die Phänomene, die auf einen gemeinsamen oder auf mehrere gemeinsame Nenner zu bringen Gegenstand dieser Bemühungen war, werden mit dramatischen Konsequenzen auf der anderen Seite konfrontiert. Mit diesen Operationen gehen Weisen der Bewusstheit und des Entzugs von Bewusstheit einher – Überlegungen, die sich in Form von Transrationalisierung oder Höchtsbewusstseinslehre Bahn brachen, die Anlass zu einer eigenwilligen Typologie des Unbewussten gaben und die mit dem »Unvergeßbarsten« und »Einprägsamsten« in Superlativen der Verblüffung schwelgten.⁶⁶ Verkennungen und Fehlinterpretationen, Phänomene, die als *flaw* und *bug* in den Blick gerieten, können zwar ein gewisses Potential der Komik beinhalten und auf das Konto kulturell belangloser Fehlleistungen und harmloser Scherzbildungen zu buchen sein, sie können aber eben auch eine Bedrohung darstellen und durchaus lebensgefährlich werden – wie im eigens angeführten Fall medizinischer Bildgebung. Mit dieser Unsicherheit gilt es unter den Bedingungen der künstlichen Intelligenzen und dem Diktat ihrer Akronyme zu leben.

We cannot reliably identify when and why DNNs will make mistakes. In some applications like text translation these mistakes may be comical and provide for fun fodder in research talks, a single error can be very costly in tasks like medical imaging. Additionally, DNNs shown susceptibility to so-called adversarial examples, or data specially designed to fool a DNN. One can generate such examples with imperceptible deviations from an image, causing the system to mis-classify an image that is nearly identical to a correctly classified one. Audio adversarial examples can also exert control over popular systems such as Amazon Alexa or Siri, allowing malicious access to devices containing personal information.⁶⁷

66 Vgl. dazu noch einmal W. Hellpach, Psychotechnik des Unbewußtseins, S. 109.

67 Charles, Adam S. (2018), »Interpreting Deep Learning: The Machine Learning Rorschach Test?«, arXiv:1806.00148v1 [stat.ML] 1 Jun 2018.

