

Feindfahrt mit Aufklebern

Stefan Rieger

Sowohl grössere als kleinere Gewässer durchschneiden das Land, und werden von Schiffen befahren, die gleichsam durch magische Kräfte fortgerudert werden; nicht Aermte, sondern Automatenähnliche Maschinen setzen sie in Bewegung. Beschreiben kann ich diese Maschinen nicht, da mit weder der Blick eines Vaucanson noch der Droz geworden ist, sondern ich vielmehr in der Mathematik nur ganz schlecht bewandert bin; hierzu kommt noch, daß diese Bäume in allen Stücken so schlau zu Werke gehen, dass niemand ihre Kunststücke entdecken kann, er müsste denn das Falkenauge eines Leuschenring's haben.¹

Overall, our quantitative and qualitative insights suggest calm VR applications are well suited to an automotive context.²

1. Luhmann am Steuer

Das Autofahren stiftet Weltbezüge und organisiert Erwartungshorizonte. Im Zuge dessen macht es auf vielschichtige Weise Dynamiken zwischen Menschen, Techniken und Autonomie verhandelbar. So steht, um mit den Menschen zu beginnen, das Aggressionspotential des Autofahrens außer Frage. Die Sozialpsychologie hat entsprechende Untersuchungen auf den Weg gebracht und damit das allen Beteiligten Offensichtliche nachhaltig bestätigt.

-
- 1 Holberg, Ludwig: Niels Klims unterirdische Reisen, Berlin: Kramer Karin Verlag [1743] 1983, S. 77.
 - 2 Paredes, Pablo E./Balters, Stephanie/Qjan, Kyle et al.: »Driving with the Fishes: Towards Calming and Mindful Virtual Reality Experiences for the Car«, in: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2/4 (2018), S. 184:1-184:1.

An keinem anderen Ort wird so viel kommuniziert wie im Verkehr.³ Verbalattacken, Beleidigungen und Beschimpfungen sind dort ebenso an der Tagesordnung wie fahrspezifische Weisen der Kommunikation, wie Hupen, dichtes Auffahren und Drängeln. Dabei wurde schon vor Einführung autonomer Fahrzeuge viel unternommen, um das Verkehrswesen über seine Eigenarten aufzuklären und entsprechende Gefährdungen weitgehend zu minimieren. Die Beiträge reichen in ihrer historischen Spannweite von einer frühen Psychologie des Verkehrs bis zu einer alles umfassenden Theorie der Anthropotechnik. Gefragt wird konkret »nach den psychologischen Grundlagen der Fortbewegung des Menschen in der Zweidimensionalen«, nach dem Beitrag der Psychologie zum Kraftfahrwesen oder nach der Ergonomie von Fahrzeuginnenräumen.⁴ Dabei werden nicht zuletzt mit Blick auf Fragen der ergonomischen Gestaltung andere Instrumente und Interfaces erwogen, andere Benachrichtigungssysteme und damit die Verwendung zusätzlicher Sinneskanäle zu Rate gezogen.⁵ Im Zuge des Primats freier Hände werden zunehmend *mid-air* Steuerungen angedacht und so finden auch niedere Sinne wie das Riechen Eingang in die Fahrzeuginnenräume – und zwar nicht in der satt-sam bekannten Form von Duftbäumen, sondern konkret als Möglichkeit der Informierung, der Aufmerksamkeitssteuerung und gekoppelt an Fahrzeuge, die als smart gelten, weil sie untereinander vernetzt sind.⁶ Das Subliminale

-
- 3 Vgl. etwa Yorck Herzberg, Philipp/Schlag, Bernhard: »Aggression und Aggressivität im Straßenverkehr«, in: Zeitschrift für Sozialpsychologie 37 (2006), S. 73-86.
 - 4 Vgl. dazu Lenkner, Hans: Die psychologischen Grundlagen der Fortbewegung des Menschen in der Zweidimensionalen unter besonderer Berücksichtigung der Fragen der Verkehrstechnik (»Das Drall-Problem«), Würzburg: Trilsch 1933; König, Ad.: »Die experimentelle Psychologie im Dienste des Kraftfahrwesens«, in: Allgemeine Automobil-Zeitung 10 (1919), S. 11-13 und 12 (1919), S. 15-17,, sowie Grünen, Rainer/Günzkofer, Fabian/Bubb, Heiner: »Anatomische und anthropometrische Eigenschaften des Fahrers«, in: Heiner Bubb/Klaus Bengler/E. Rainer Grünen et al. (Hg.): Automobilergonomie, Springer: Berlin 2015, S. 163-219.
 - 5 Zur Bandbreite von Informierungen im Fahrzeug vgl. etwa Fricke, Nicola: »Warn- und Alarmsounds im Automobil«, in: Georg Spehr (Hg.): Funktionale Klänge. Hörbare Daten, klingende Geräte und gestaltete Hörerfahrungen, Bielefeld: transcript 2009, S. 47-64, sowie Schipor, Ovidiu Andrei/Vatavu, Radu-Daniel: »Towards Interactions with Augmented Reality Systems in Hyper-Connected Cars«, in: EICS Workshops (2019), S. 76-82.
 - 6 Vgl. Rieger, Stefan: »Halt mal kurz!«, in: Manuela Klaut/Claus Pias/Gottfried Schnödl (Hg.): Stimmen Hören. Festschrift für Wolfgang Hagen, Berlin: ciconia 2020, S. 198-211. Vgl. zum instrumententenfreien Fahren und Steuern auch Cornelio Martinez et

wird zum Einfallstor für Verhaltensregulierungen beim Fahren und das Riechen zu einem Mittel seiner Umsetzung.⁷ Selbst in autonome Fahrzeuge hält das beruhigende Riechen Einzug. So soll ein System namens S(C)ENTINEL die Verlässlichkeit des autonomen Fahrens mittels Lavendel- und Zitronengeruch unter Kontrolle halten (»lemon for a change to low and lavender for a change to high reliability«).⁸ Welche spezifischen Gerüche dabei eine Rolle spielen, tritt hinter den wohl unbestrittenen Vorteil olfaktorischer Information zurück.⁹ Derartige Belange der operativen Umsetzung sind allerdings auf ein sehr grundsätzliches Sicherheitsbedürfnis ausgerichtet. Weil Gefährdungslagen im Verkehr häufig durch die Unzulänglichkeiten des Menschen verursacht werden, muss dieser vor sich selbst geschützt werden.

Denker wie Niklas Luhmann haben sich der besonderen Situation im Auto zugewendet. Ob die Systemtheorie zur Aufklärung dessen beiträgt, was sich im Innenraum von Fahrzeugen abspielt, wird zur Bewährungsprobe für die eigene Theorie, die oftmals einer allzu großen Entfernung von den Belangen der Lebenswelt bezichtigt wurde. Luhmann steuert gegen und führt ein doch sehr lebensweltliches Bonmot an, demzufolge Ehen im Himmel geschlossen und im Auto getrennt werden. Die Situation im Verkehr lässt in der Monographie *Liebe als Passion. Zur Codierung von Intimität* die zwischen Fahrern und Beifahrern zentrale Frage virulent werden, welche Welt denn gerade als die gemeinsame zugrunde gelegt und nach welcher Maßgabe attribuiert wird – eine Frage, die in den Pluralisierungen sensorgestützter Weltentwürfe gleichermaßen von Belang ist. Fahr(fehl)verhaltensweisen wie Kurvenschneiden oder Linksfahren auf der Autobahn werden zur Nagelprobe für die Intimkommunikation – jedenfalls dann, wenn Besonderheiten individueller Fahrstile auf die Eigenheiten selbstreferentieller Kommunikationen im Medium der Liebe hochgerechnet werden und es auf diese Weise gelingt, »Nuancen

al.; »Agency in Mid-air Interfaces«, in: CHI '17. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (2021), S. 2426-2439.

- 7 Riener, Andreas/Chalfoun, Pierre/André-Aisenstadt, Pavillon et al.: »The Potential of Subliminal Information Displays to Change Driver Behavior«, in: *Presence* 23/1 (2014), S. 51-70.
- 8 Wintrijsberger, Philipp/Dmitrenko, Dmitrijs/Schartmüller, Clemens et al.: »S(C)ENTINEL: Monitoring automated vehicles with olfactory reliability displays«, in: *International Conference on Intelligent User Interfaces* (2019), S. 538-546.
- 9 Vgl. dazu etwa Baron, Robert. A./Kalsher, Michael J.: »Effects of a Pleasant Ambient Fragrance on Simulated Driving Performance: The Sweet Smell of...Safety?«, in: *Environment and Behavior* 30/4 (1998), S. 535-552.

des Verhaltens attributionsfähig zu profilieren«. ¹⁰ Wie missverständlich die Attribution einfacher Gesten ist, zeigt ein Beispiel aus der Praxis. Dort werden für eine einfache Regung (ein aus dem Auto herausgestreckter Arm) eine verwirrende Zahl an Zuschreibungen parat gehalten, die vor allem eines verhindern: die sachdienliche Vorwegnahme einer halbwegs komplexen Situation und damit die Möglichkeit eines adäquaten Reagierens. ¹¹

Das Aggressionspotential der Automobilisten untereinander und nicht beschränkt auf die unter Attributionsdruck stehende Dyade aus Fahrzeuglenker und Beifahrerin steht in diesem Zusammenhang nicht nur außer Frage, es hat selbst historischen Tiefgang. Der Colditzer Psychiater und Kriminalanthropologe Paul Näcke gerät anlässlich des Autofahrens gar zu Überlegungen über die Sittlichkeit von Geisteszuständen. Näcke, der 1907 ausgerechnet den eher Pferden zugeneigten Schriftsteller Karl May auf seine psychische Verfasstheit hin zu begutachten hatte, weiß herzuleiten, was die Menschen am Steuer so aggressiv macht. Die Gefahr der sittlichen Verrohung durch die Effekte des Tempos und die Wirkung auf einen derart beschleunigten Geisteszustand sind so groß, dass nur ausgewähltes Personal am Verkehr teilnehmen sollte – was zweifellos der Psychotechnik ein weites Betätigungsfeld für entsprechende Eignungsprüfungen erschloss. ¹² Was neben der Verrohung dräut, ist nicht weniger als eine veritable Autokratie. Der Temporausach, die spezifische Rauscherfahrung der Moderne, steht einer tempolimitierten Vernünftigkeit im Wege. ¹³ Das Auto versetzt den Fahrer in einen anderen Zustand, »die Dekorationen seines Reiches wechseln jeden Augenblick, Häuser,

10 Luhmann, Niklas: *Liebe als Passion. Zur Codierung von Intimität*, Frankfurt a.M.: Suhrkamp 1990, S. 43. Vgl. dazu auch Rieger, Stefan: »Utopien am Steuer. Zu einer Kommunikationstheorie des Autofahrens«, in: Claus Pias/ders. (Hg.): *Vollstes Verständnis. Utopien der Kommunikation*, Zürich: Diaphanes 2016, S. 125-141.

11 Vgl. dazu Kleffel, Werner: »Schuld oder Schuldlosigkeit bei Verkehrsunfällen. (Voraussehbarkeit – Geistesgegenwart – Schrecksekunde)«, in: *Deutsche Zeitschrift für die gesamte gerichtliche Medizin* 20 (1933), S. 1-32.

12 Dazu etwa Marbe, Karl: *Die gerichtopsychologische Begutachtung von Autounfällen und die Eignung zum Chauffeur. Ein Führer für Gerichtsgutachter, Juristen und die Polizei*, Leipzig: C. L. Hirschfeld 1932.

13 Zur Topik des Rausches vgl. stellvertretend Baudry de Saunier, Louis: *Das Automobil in Theorie und Praxis*, Wien: Inktaank-Publishing [1901] 2019.

Bäume, Felder, Menschen flüchten an ihm vorbei, und es steigt ihm wie ein Herrschergefühl ins Hirn [...] Das Gefühl macht ihn leicht zum Autokraten.«¹⁴

Für den Semiotiker und Philosophen Max Bense hingegen taugt das Autofahren zu einer strukturierten Form technisch induzierter Beiläufigkeit. Automatismus und Gedankenlosigkeit führen zu einer Gemengelage, die bereits Modernetheoretiker wie Robert Musil am Beispiel moderner Bewegungsformen zu einem Durchbruch durch die bewusste Person erklärt haben. Im Kontrollverlust sehen sie eine Form ekstatischer Glückserfahrung veranlagt.¹⁵ Dem Moment reduzierter Rationalität ist eine Geschichte der Moderne eingeschrieben, die von Kleists intuitiv antizipierenden Bären aus dem Marionettentheater bis zu Glücksmomenten einer automobil beschleunigten Moderne führt – ob bei Musil oder bei Bense, der die automatisierte Gedankenlosigkeit beim Fahren gar zu einer neuen Existenzweise verkürt. »Indem sich das denkende Wesen an das fahrende Wesen gewöhnt, Ich und Auto mehr und mehr zu einem beinah surrealen Automaten verschmelzen, jene aber dennoch auch bei sich selbst bleiben, uns stets auch noch ein selbständiges Etwas, eben fahrendes Wesen und denkendes Wesen bedeuten, ist fast eine neue Art des Existierens entstanden«.¹⁶ Die Möglichkeit dieser neuen Existenzweise geht einher mit einer Neuveranlagung des Autos als einer informationsverarbeitenden Einheit, die nicht mehr dem Geltungsbereich klassischer Maschinen zugehört.¹⁷

14 Nücke, Paul: »Der Geisteszustand des Automobilfahrens«, in: Archiv für Kriminalanthropologie 16 (1904), S. 335-337, hier S. 335.

15 Zu dieser Wertschätzung Rieger, Stefan: »Auto«, in: Benjamin Bühler/ders. (Hg.): Kultur. Ein Machinarium des Wissens, Berlin: Suhrkamp 2014, S. 19-30.

16 Bense, Max: »Auto und Information. Das Ich, das Auto und die Technik«, in: ders. (Hg.): Ausgewählte Schriften (= Poetische Texte, Band 4), Stuttgart/Weimar 1998, S. 291-293, hier S. 292. Die biographische Verschmelzung mit dem Automobil, die Bense phantasiert, wird durch technische Aufrüstungen befördert, die dem Gespann aus Fahrzeug und Fahrzeugführer ungeahnte Möglichkeiten hochindividueller Sozialität beschieren (vgl. dazu McVeigh-Schultz, Joshua/Stein, Jennifer/Boyle, Jacob et al.: »Vehicular lifelogging: new contexts and methodologies for human-car interaction«, in: CHI Extended Abstracts (2012), S. 221-230). Aber die neuen Verkehrsteilnehmer werden auch zur Nagelprobe für die Gesellschaft (vgl. dazu Marres, Noortje: »Co-existence or displacement: Do street trials of intelligent vehicles test society?«, in: British Journal of Sociology 71/3 (2020), S. 537-555).

17 Vgl. dazu Günther, Gotthard: Das Bewusstsein der Maschinen. Eine Metaphysik der Kybernetik, Baden-Baden: Agis-Verlag 1957.

Vor dem Hintergrund solcher Verwerfungen und Verheißungen, die nachgerade zur Topik eines kulturwissenschaftlich-philosophischen Verkehrsdiskurses zählen, mag eine Einschätzung von Jürgen Habermas auffallen. Immerhin löst der Philosoph das Autofahren aus den Niederungen von verbalem Streit oder fahrtechnischem Fehlverhalten – das *tertium comparationis* könnte dabei der Begriff des Schneidens sein – und versetzt es kurzerhand in einen Adelszustand. Im gewählten Vergleich wird das Fahren nobilitiert, soll es doch in der Lage sein, auf einfache Weise das Alltagsgeschäft von Geisteswissenschaften zu erklären. »Autofahren ist so etwas wie eine Geisteswissenschaft. Man muß fortwährend fremde Texte übersetzen, fremde Welten, Stile, Manieren und Marotten antizipieren. Denn das heißt es ja: mit den Fehlern der anderen kalkulieren. Darin kulminiert der Adel automobiler Intelligenz.«¹⁸ Diese automobilen Kommentierung der Geisteswissenschaften greift auf etwas Grundlegendes zurück, dem auch moderne Konzeptualisierungen des Fahrens folgen – der Umgang mit einem Kalkül künftiger Welten. »In order to drive a car, a driver needs to see and understand the various objects in the environment, predict their possible future behaviors and interactions, and then plan how to control the car in order to safely move closer to their desired destination while obeying the rules of the road.«¹⁹ Wie bei Luhmann ist es die Frage nach dem Referenzsystem und damit die nach der jeweils zugrunde gelegten Welt, die Fahren und Denken verbindet.

Doch der Verkehr ist bedrohlich und die Gefahr allgegenwärtig. Dieses Risiko wird als Symptom einer beschleunigten Moderne beschrieben und seine Opfer werden gemeinhin akzeptiert. Gleichwohl setzt das autonome Fahren dort an, wo menschliche Fahrzeuglenker am Steuer sitzen und das Geschehen eben nur vermeintlich unter Kontrolle haben. Durch Technik, so das Versprechen, sollen menschliche Unzulänglichkeiten minimiert und Sicherheit maximiert werden. Eine entsprechende Beforschung ist gleichermaßen umtriebig wie umfänglich. Neben Fragen der technischen Umsetzung geraten vor allem soziale Aspekte des autonomen Fahrens auf die Agenda.²⁰ Mit ihnen wird vi-

18 Habermas, Jürgen: »Der Mensch am Lenkrad«, in: Frankfurter Allgemeine Zeitung vom 27.11.1954.

19 Bansal, Mayank/Krizhevsky, Alex/Ogale, Abhijit S.: »ChauffeurNet: Learning to Drive by Imitating the Best and Synthesizing the Worst«, in: arXiv:1812.03079v1 [cs.RO] 7. Dezember 2018.

20 Vgl. dazu Bissell, David/Birtchnell, Thomas/Elliott, Anthony et al.: »Autonomous automobiles: The social impacts of driverless vehicles«, in: Current Sociology 3 (2018), S. 1–19 sowie Forster, Yannick/Frison, Anna-Katharina/Wintersberger, Philipp et al.: »Where

rulent, wie es um Verteilung von Autonomie bestellt ist, die im autonomen Fahren wie kaum an einem anderen Ort als gesellschaftspolitisches Phänomen diskutiert wird. Mit der (Teil-)Autonomie nicht-menschlicher Akteure sind Positionen im Raum, die nicht nur über Selbstverständnis und Selbstpositionierung des Menschen gegenüber seiner Umwelt, sondern sogar über die Kriterien eines gelungenen Lebens entscheiden.²¹ Im Zentrum stehen dabei Fragen nach der Teilung von Handlungsträgerschaft und ihrer Validierbarkeit im Rahmen von Jurisdiktion und Versicherung, von Moral und Ethik.²² Ein weiteres Zentrum bildet die Frage nach dem speziellen Zustand der Ablenkung, der mit der Gesamtsituation eines in Bewegung befindlichen Körpers zu tun hat und hinter der eine groß angelegte Kulturgeschichte der Schläfrigkeit zu stehen hätte. Die Physiologie der Müdigkeit ist auf eine spezifische Weise medien- und technikinduziert – und gleichzeitig ist sie es, die die Sicherheit des Verkehrswesens bedroht und mit allen Mitteln unter Kontrolle zu halten ist. Wieder ist die Wissenschaft gefragt, um einen Geisteszustand aufzuklären, der dem Autofahren inhärent zu sein scheint.²³

Für menschliche Lenker hat sich selbst noch unter säkularen Bedingungen ein Heiliger bewährt und seine in die kulturelle Topik eingespielte Wirksamkeit bewahrt: Zuständig für die spirituelle Beruhigung im Transportwesen ist der heilige Christopherus, der als stabbewehrter Hüne Christus auf seinen Schultern durch ein Gewässer trägt. Als Gewährsmann für sicheren Personentransport zu Lande, zu Wasser und in der Luft überlebt er auf eigentümliche Weise als Patron des Kraftfahrwesens – in Form von Medaillen und Plaketten,

We Come from and Where We Are Going: A Review of Automated Driving Studies«, in: AutomotiveUI (adjunct) (2019), S. 140-145.

- 21 Rössler, Beate: Autonomie. Ein Versuch über das gelungene Leben, Berlin: Suhrkamp 2019.
- 22 Dazu Brown, Barry/Laurier, Eric: »The Trouble with Autopilots: Assisted and Autonomous Driving on the Social Road.«, in: Gloria Mark/Susan Fussell (Hg.): Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, New York: ACM 2017, S. 416-429; Amirshirzad, Negin/Kumru, Asiye/Öztop, Erhan: »Human Adaptation to Human-Robot Shared Control«, in: IEEE Trans. Human-Machine Systems 49/2 (2019), S. 126-136.
- 23 Vgl. stellvertretend Ramzan, Muhammad/Ullah, Hikmat/Awan, Khan et al.: »A Survey on State-of-the-Art Drowsiness Detection Techniques«, in: IEEE Access, 7 (2019), S. 61904-61919, sowie aus den Reihen der Lösungsvorschläge Doudou, Mes-saoud/Bouabdallah, Abdelmadjid/Berge-Cherfaoui, Véronique: »Bias Remediation in Driver Drowsiness Detection Systems Using Generative Adversarial Networks«, in: IEEE Access 8 (2020), S. 55592-55601.

von Anhängern und Aufklebern weiß er sich im gut sortierten Sortiment des Autozubehörfachhandels zu halten. Weil für Herstellung und Vertrieb entsprechender Devotionalien nicht mehr nur die Kirche zuständig ist, führte der Autoteilehersteller ATU unter der Art Nr. RW3000 noch im Jahr 2019 eine silberfarbene Plakette des Heiligen. Beworben wurde das formschöne und selbstklebende Produkt, das zwischen Warnwesten und Wunderbäumen, Verbandkästen und Windschutzscheibenreinigungsprodukten als ebenfalls für den potentiellen Kunden interessanten Kaufoptionen positioniert war, mit einem Hinweis auf die Tradition von Heiligengeschichten. »Ihr ständiger Begleiter! Jeder von uns kann einen Schutzengel brauchen, der uns ab und an unter die Arme greift und uns beschützt.«²⁴ Selbst vor der Bewerbung als »ultimativen« Glücksbringer schreckt ATU nicht zurück.

Abbildung 1: Christopherus-Plakette



https://www.atu.de/shop/Reise-und-Transport-w11621/Reisezubehoer-w11881/Alles-fuer-die-Fahrt_w11891/St-Christopherus-Plakette-silberfarben-1-Stueck-RW3000 vom 31.03.2019.

Der Bereich des autonomen Fahrens ist angetreten, eine solche ständige Begleitung dauerhaft zu garantieren und universellen Schutz herzustellen, indem man den Menschen weitestgehend aus der Gleichung nimmt. Da-

24 https://www.atu.de/shop/Reise-und-Transport-w11621/Reisezubehoer-w11881/Alles-fuer-die-Fahrt_w11891/St-Christopherus-Plakette-silberfarben-1-Stueck-RW3000 vom 31.3.2021.

zu wird die Risikominimierung verweltlicht. Das Netz technischer Sensoren ist feiner in der Erfassung der Umgebung als das menschliche Wahrnehmungsvermögen, eine Datenverarbeitung zeichnet Situationen nach, die unterschiedliche Szenarien und deren Konsequenzen mit nur minimaler Verzögerung (und damit dem Ideal der Echtzeit praktisch entsprechend) umsetzen, die Antizipation komplexer Situationen wird möglich – jedenfalls genauer, als wenn es der Einschätzung eines Menschen überlassen bliebe. Ob gebremst oder ausgewichen wird, ob Unfallrisiken minimiert oder ethische Dilemmata ausgereizt werden, in all diesen Fällen greift ein Kalkül von Umwelterfassung, Infrastruktur und Datenverarbeitung – und das vor dem Hintergrund einer Vielzahl möglicher Welten.²⁵ Die Chancen des autonomen Fahrens sind erkannt und vielfach beschrieben worden. Was immer dessen Einführung befördert oder verzögert – es sind vorrangig Fragen und Aspekte, die weniger mit dem technisch Möglichen als mit dem gesellschaftlich Durch- und juristisch Umsetzbaren zu tun haben. Wie verhalten sich Probetrieb und Alltagstauglichkeit? Ist der deutsche Automobilist wirklich (schon) willens, das Steuerruder aus der Hand zu geben? Wer ist Herr im automobilen Haus?²⁶ Trägt die Vernetzung der Fahrzeuge (Vehicle to Vehicle [V2V]) untereinander dazu bei, aus vormaligen Autokraten am Steuer empathische und vor allem rücksichtsvolle Verkehrsteilnehmer:innen zu machen?²⁷ Wollen wir uns wirklich durch derartige Befriedung des Verkehrswesens um speziell diese Möglichkeit der sozialen Interaktion bringen?²⁸ Veranlagten wir den Umgang mit der Pluralisierung von Welten als etwas, das gelöst und bewältigt sein will? Soll eine Welt der Bedeutungsüberschüsse durch Strategien der Vereindeutigung reguliert werden? Bedarf es dazu psychologischer Zuschreibungen und hermeneutischer Verfahren als Mechanismen der Reduktion von Mehrsinnig-

-
- 25 Als Musterbeispiel dient dabei ein Unfallsszenario, bei dem Personenschäden gegeneinander aufgerechnet werden.
 - 26 Cohen-Lazry, Guy/Edelstein, Amit/Degani, Asaf et al.: »An Introduction to a Psychoanalytic Framework for Passengers' Experience in Autonomous Vehicles«, in: Heidi Krömker (Hg.): *HCI in Mobility, Transport, and Automotive Systems. Automated Driving and In-Vehicle Experience Design*. Second International Conference, MobiTAS 2020. Proceedings, Part I, S. 249-265.
 - 27 Rakotonirainy, Andry/Schroeter, Ronald/Soro, Alessandro: »Three social car visions to improve driver behaviour«, in: *Pervasive and Mobile Computing* 14 (2014), S. 147-160.
 - 28 Zu einer solchen Kommunikationstheorie des Autofahrens vgl. Kling, Marc-Uwe: *Die Känguru-Offenbarung*, Berlin: Ullstein 2014, 18. Kapitel.

keit?²⁹ Die Bedeutung der Ambiguität wird fernab der Brisanz eines risikobehafteten Verkehrswesens (und der verständlichen Regulierung etwa der dort angelegten Signalgebung!) zunehmend im Bereich der Designwissenschaft diskutiert. Ihre Zuspitzung erhält sie in Konzepten wie *research through design* und *cargo cult design*, Gestaltungsweisen, die das Gegenteil der Vereindeutigung nachgerade zum Programm ihrer Weltgestaltung erheben.³⁰

2. Adversarial Attacks

Vor diesem Hintergrund ist eine Diskussion zu situieren, die dem autonomen Fahren noch auf eine ganz andere Weise abträglich ist. Dabei spielen nicht Menschen und deren Unzulänglichkeiten eine Rolle, sondern die Systeme der Erkennung und damit doch wieder die Technik selbst. Eine Eigenart künstlicher Systeme bei der Identifizierung von Gegenständen in der Umwelt, die durch vergleichsweise einfache Mittel täuschbar ist, zeichnet dafür verantwortlich. Genau genommen geht es dabei um jene *Deep Neural Networks*, die im Rahmen des *machine learning* Anwendung finden.³¹ Verhandelt werden entsprechende Fehlleistungen künstlicher Intelligenzen gelegentlich in der Beschreibungssprache psychischer Störungen, etwa als Trugbilder oder Halluzinationen.³² Das Phänomen hat zunächst als eine Art Kuriosität der *Artifi-*

29 Vgl. dazu Sengers, Phoebe/Gaver, Bill: »Staying Open to Interpretation: Engaging Multiple Meanings in Design and Evaluation«, in: DIS 2006, June 26-28, 2006, S. 99-108.

30 Dazu Andersen, Kristina: »The Deliberate Cargo Cult«, in: Proceedings of the 2014 Conference on Designing Interactive Systems, S. 627-636 und Boucher, Andy/Chatting, David J./Desjardins, Audrey et al.: »Doing Things with Research through Design: With What, with Whom, and Towards What Ends?«, in: CHI Extended Abstracts 2019.

31 Engemann, Christoph/Sudmann, Andreas: »Machine Learning – Neue Pfade künstlicher Intelligenz«, in: Dies. (Hg.): Machine Learning. Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz, Bielefeld: transcript 2018.

32 Im Vorgriff auf das Folgende: Die Rede von der Halluzination findet sich nicht nur in eher feuilletonistischen Berichterstattungen (*Die Zeit*, *Wired*), sondern auch in den einschlägigen Texten selbst. »Recently, adversarial algorithms were developed to facilitate hallucination of deep neural networks for ordinary attackers« (Abdel-Hakim, Alaa E.: »Ally patches for spoliatio of adversarial patches«, in: Journal of Big Data 6/51 (2019); vgl. dazu auch Simonite, Tom: »AI Has a Hallucination Problem That's Proving Tough to Fix«, in: Wired 3.9.2018. <https://www.wired.com/story/ai-has-a-hallucination-problem-thats-proving-tough-to-fix> vom 05.04.2020. Zum schwankenden Objektstatus und zur Übergängigkeit von Dingen vgl. auch Piergiovanni, A. J./Wu, Alan/Ryoo, Michael

cial Intelligence eine entsprechende Aufmerksamkeit erfahren. Es sind feindliche Attacken (*adversarial attacks*), die das System um die Kohärenz und um den Verstand bringen. Und es sind so genannte *adversarial patches* in der Form unscheinbarer Aufkleber, die solche Störfälle im allgemeinen Wahrnehmungsbetrieb auf unscheinbare Weise befördern. Dem Ganzen liegt der Befund zugrunde, dass Systeme der optischen Bilderkennung ohne größeren Aufwand dazu gebracht werden können, einfache, aber groteske (und damit folgeschwere) Fehleinschätzungen vorzunehmen. Unter dem Titel »Ist das ein Toaster?« greift ein Artikel in der Wochenzeitschrift *Die Zeit* zur Veranschaulichung dieser Möglichkeit auf Forschungen zurück, die von der technischen Metamorphose einer schlichten Banane handeln. Die Frucht wird zunächst von einem technischen System erfasst und entsprechend erkannt. Gerät allerdings ein kleines Bild, ein *patch* oder *sticker* ohne ersichtlichen Inhalt in die Nähe der Banane, gibt das System die Frucht preis und identifiziert die neue Situation etwa als einen Toaster. Aber damit nicht genug: War die Bilderkennungssoftware im Fall der Banane mit 97 Prozent von der Richtigkeit ihrer Erkenntnis überzeugt, so steigert sich diese Zuversicht im Fall des fälschlich erkannten Toasters auf stattliche 99 Prozent. »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.«³³ Die Fülle der Beispiele wäre beliebig erweiterbar um Dinge des Alltags, um Teddybären und Orangen, um Socken und Haustiere, um Sportgerätschaften und Einrichtungsgegenstände.

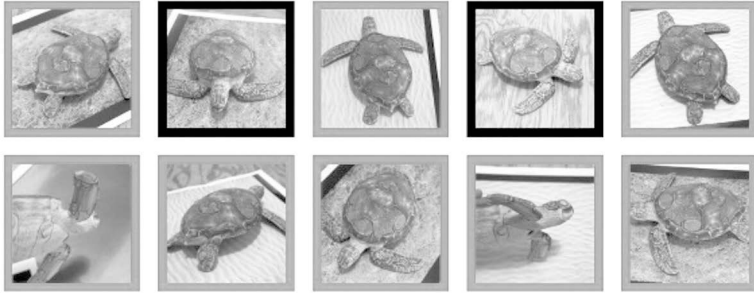
Derartige Geschichten sind gut erzählbar, nicht zuletzt wegen des besonders grotesken Ausmaßes der Fehleinschätzung. Die Verwechslung der Banane mit einer ihr gestaltähnlichen Gurke oder einer ihr farbähnlichen Zitrone wäre im Auge des menschlichen Betrachters erwartbar, vor dem Hintergrund eines Denkens in Gradualitäten auch einigermaßen plausibel und ist für entsprechende Scherze genutzt worden.³⁴ Zu profilieren wären derartige

S.: »Learning Real-World Robot Policies by Dreaming«, in: arXiv:1805.07813v4 [cs.RO] 1. August 2019, S. 7680-7687. Beschrieben wird dort ein Gestaltwandel in der Beschreibungssprache des Träumens.

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

34 Derartige Verkennungen beim Menschen werden entsprechend konzeptualisiert. Vgl. dazu Goldenberg, Georg: »Visuelle Objektagnosie und Prosopagnosie«, in: Hans-

Abbildung 2: Von Schildkröten und Gewehren



Anish Athalye/Engstrom, Logan/Ilyas, Andrew u.a. (2018), »Synthesizing Robust Adversarial Examples«, arXiv:1707.07397v3 [cs.CV], 7. Juni 2018.

Verwechslungen vor dem Hintergrund einer Geschichte des technischen Sehens, die sich entlang von staffelbaren Auflösungen, aber auch im Umgang mit anderen Sehparametern wie Abschattung, Entfernung oder Blickwinkel abzuarbeiten hatte und dabei die Identifizierung von Gegenständen als die besondere Leistung des menschlichen Sehens qualifizierte.³⁵ Aber die Übertragung von der Banane auf den Toaster ist ein entschiedener Bruch im habitualisierten Umgang mit Ähnlichkeitsrelationen. Sie stellt einen Spezialfall der Bildwahrnehmung dar, der auch durch kulturell eingespielte Sonderformen wie dem von der Gestaltpsychologie bemühten Kippbild nicht gedeckt ist. Im Reich der Vexierbilder ist der gewohnte Kontrakt mit der Stetigkeit aufgekündigt. Wie in den Musterbeispielen der Gestaltpsychologie herrscht zwischen Gesicht und Vase, zwischen dem Hasen und der Ente, zwischen der alten Frau und dem anmutigen Mädchen eine Ambiguität, die vom Wahrnehmungssystem ohne intentionale Intervention vereindeutigt wird.³⁶

Otto Karnath/Peter Thier (Hg.): Kognitive Neurowissenschaften, Berlin: Springer 2012, S. 161-171.

35 Dafür typisch ist die folgende Einschätzung von Karl Steinbuch: »Auf keinem anderen Gebiet ist die Unterlegenheit technischer Gebilde gegenüber den organischen Systemen so offensichtlich wie beim Sehsystem.« (Steinbuch, Karl: Automat und Mensch. Auf dem Weg zu einer kybernetischen Anthropologie, Berlin: Springer 1971 (4., neubearbeitete Auflage), S. 98).

36 Zur Kippfigur sowie zu ihrer Ambiguität vgl. Schönhammer, Rainer: »Stichwort: Kippbilder«, <https://psydok.psycharchives.de/jspui/bitstream/20.500.11780/3666/1/Kippbild>

Brisant werden die technischen Fehleinschätzungen, wenn man die Beispielliste erweitert und auf Fälle kommt, die alles andere als harmlos und in der Lage sind, eklatant in lebensweltliche Alltagssituationen einzugreifen. Berichtet wird etwa von einem Stoppschild, das durch einen Sticker an der menschlichen Wahrnehmung vorbei für die technische unkenntlich gemacht wird oder davon, »eine Plastikschildkröte durch subtile Gravuren auf ihrem Panzer als Schusswaffe erscheinen zu lassen.«³⁷ Gerade die deutlich lebensnahe Frage der Identifizierung von Verkehrsschildern wird zu einer Herausforderung und entsprechend aufwendig beforscht.³⁸ Mit dem Verkehrswesen wird die Brisanz der *adversarial attacks* sichtbar. Schutzmaßnahmen und ein Verständnis der Operationsweise von *Deep Neural Networks* bedingen sich dabei wechselseitig.

In addition to building safe and reliable DNN models, studying adversarial examples and their countermeasures is also beneficial for us to understand the nature of DNNs and consequently improve them. 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. There are works to explain and interpret the existence of adversarial examples of DNNs, which can help us gain more insight into DNN models.³⁹

Auch der *Zeit*-Artikel verweist auf eine Steigerung entsprechender Möglichkeiten. Während es bei den bisher 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*

er_psydoc_11052011.pdf vom 23.04.2020. Zum Versuch, die damit verbundene Ambiguität zu messen Fisher, Gerald H.: »Measuring Ambiguity«, in: American Journal of Psychology 80/4 (1967), S. 541-557 sowie zu Strategien der Vereindeutigung Kawabata, Nobuo/Takayuki, Mori: »Disambiguating ambiguous Figures by a model of selective attention«, in: Biological Cybernetics 67 (1992), S. 417-425.

- 37 S. Schmitt: »Ist das wirklich ein Toaster?«, S. 33. Zu diesem Beispiel vgl. Athalye, Anish/Engstrom, Logan/Ilyas, Andrew et al.: »Synthesizing Robust Adversarial Examples«, in: arXiv:1707.07397v3 [cs.CV], 7. Juni 2018.
- 38 Dan Cires, An/Ueli, Meier/Masci, Jonathan et al.: »Multicolumn Deep Neural Network for Traffic Sign Classification«, in: Neural Networks 32 (2012), S. 333-338.
- 39 Xu, Han/Ma, Yao/Liu, Haochen et al.: »Adversarial Attacks and Defenses in Images, Graphs and Text: A Review«, in: arXiv:1909.08072v2 [cs.LG], 9. Oktober 2019.

Abbildung 3: Stop-Schilder



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.”

Evtimov, Ivan/Eykholt, Kevin/Fernandes, Earlene u.a. (2018),
 »Robust physical-world attacks on deep learning models«,
 arXiv:1707.08945v5 [cs.CR], 10. April 2018.

eingzugreifen. Die Wahrnehmung solcher Systeme steuert die Antizipation komplexer Situationen wie auch die angemessene Reaktion auf diese. Sie wird damit zur Voraussetzung für die Erzeugung möglicher Welten – auch und gerade solcher, die autonomen Systemen als Umwelten für Kalküle der Voraussicht dienen. »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.«⁴⁰ Zur Optimierung der eigenen Systeme haben die Tübinger Forscher daher bereits vor der Veröffentlichung ihrer Ergebnisse entsprechende *patches* an die Hersteller und Zulieferer der Automobilbranche geschickt, um diese in die Lage zu versetzen, eigene Untersuchungen mit den Störmöglichkeiten zu unternehmen und Resilienzstrategien auf den Weg zu bringen. Die Leiterin des *Algorithm Accountability Lab* Katharina Zweig, die der *Zeit*-Artikel zu Wort kommen lässt, betont den Ernst der Lage mit dem Verweis darauf, dass solche Systeme am Bewusstsein und am Wissen der beteiligten Akteure vorbei zu ihren Einschätzungen gelangten. Die künstlichen Systeme operieren im Modus des Unmerklichen. Und sie operieren, wie nicht nur das Beispiel des autonomen Fahrens verdeutlicht, ausgerechnet dort, wo

40 S. Schmitt: »Ist das wirklich ein Toaster?«, S.33.

die lebensweltliche Brisanz außer Frage steht. Die Forschungen zu diesem Phänomen sind noch jung und haben einen Schwerpunkt im Jahr 2019. Deutlich wird dabei aber bereits jetzt, welche Potentiale in der Technik des Angriffes liegen und wie sehr das Arbeiten der entsprechenden Netzwerke die Emergenz von Phänomenen befördern. Was zutage tritt, fällt aus dem Raster gewöhnlicher Erklärungen heraus. Damit wird das alte Paradigma einer bestimmten und im Alltagsbewusstsein verankerten Auffassung von Determiniertheit und Vorhersagbarkeit technischer Systeme aufgegeben. Die mangelnde Prognostizierbarkeit der DNN führt zu kritischen Lagen, in denen derartige Fehler nicht zu tolerieren 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.⁴¹

Ein in seinem Risikopotential vergleichbares Szenario ergibt sich am Schauplatz von Finanztransaktionen. »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.«⁴² Was die Überblicksdarstellung *Adversarial Attacks and Defenses in Images, Graphs and Text* zunächst als Strategie menschlicher Betrüger gegenüber Kreditinstituten ausmacht, verweist gleichwohl in die Welt anderer Operationsweisen und Darstellungsformen. Die Frage der Glaubwürdigkeit wird nicht entlang der persönlichen Identität von Betrügern verhandelt, also etwa an der Eineindeutigkeit von Personenständen und deren biometrischer Gewährleistung, sondern an der Eigenlogik von Graphen. Ob im Straßenverkehr oder bei Finanztransaktionen: Die Gefahren lauern an unterschiedlichen Stellen einer störanfälligen Lebenswelt. Sie sind ausdifferenziert nach Anlässen und technischen Umsetzungen, betreffen die Erkennung

41 H. Xu et al.: »Adversarial Attacks«.

42 Ebd.

Abbildung 4: Patches

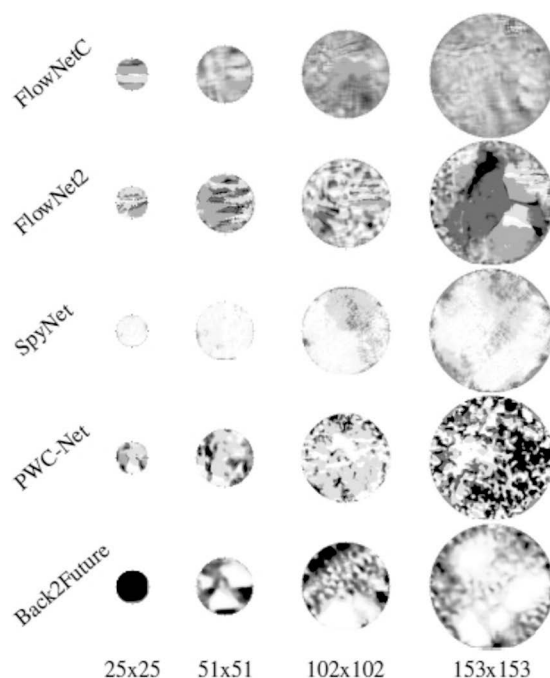


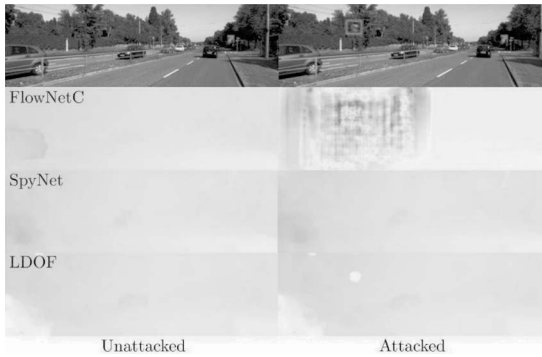
Figure 2. **Adversarial Patches.** Obtained for different optical flow networks. The size is enlarged for visualization purposes.

Ranjan, Anurag/Janai, Joel/Geiger, Andreas u.a. (2019), »Attacking Optical Flow«, ICCV 2019, S. 2404-2413, hier 2407.

von Buchstaben und Ziffern, von Verkehrszeichen und Bildern, von Bananen und gravierten Schildkröten.⁴³

43 Zum Beispiel der Ziffern, deren Ikonographie sehr an die frühen Bemühungen der technischen Übertragung und der Mustererkennung erinnern (Karl Steinbuch, Norbert Wiener), vgl. Feinman, Reuben/Curtin, Ryan R./Saurabh, Shintre et al.: »Detecting Adversarial Samples from Artifacts«, in: arXiv:1703.00410v3 [stat.ML] 15. November 2017. Zu weiteren Beispielen Kurakin, Alexey/Goodfellow, Ian J./Bengio, Samy: »Adversarial

Abbildung 5: Bewegungsanalysen



Ranjan, Anurag/Hoffmann, David T./Tzionas, Dimitrios u.a. (2019), »Learning Multi-Human Optical Flow«, arXiv:1910.11667v2 [cs.CV], 4. Dezember 2019.

Auf dem Feld der geläufigen Datenformen (»for the three popular data types, i.e., images, graphs and text«) geraten am Finanzbeispiel Graphen und die Veränderung von nur wenigen Knoten auf die Agenda und der Umgang mit Sprache findet nicht nur in schriftlicher, sondern ebenso in gesprochener Form Beachtung.⁴⁴ Auch hier erfolgen die Attacken im Gewand von minimalen und kaum merklichen Unterschieden – etwa durch die Einstreuung von Tipp- oder Druckfehlern (*typos*).⁴⁵ Dabei gerät jenseits von Tippfehlern und einfachen Buchstabendifferenzen der Austausch ganzer Sätze oder Phrasen in den Fokus der Aufmerksamkeit – und damit wiederum spezifische Formen eines Kommunikationsgeschehens wie die Aggression. Ein Beitrag mit dem Titel *Detecting egregious responses in neural sequence-to-sequence models* von 2018 erweitert dabei auf auffallende Weise die Beschreibungssprache und die Semantik dessen, was gemeinhin mit dem Begriff des Auffälligen verbunden

Examples in the Physical World«, in: arXiv:1607.02533v4 [cs.CV], 11. Februar 2017 sowie Evtimov, Ivan/Eykholt, Kevin/Fernandes, Earlence et al.: »Robust physical-world attacks on deep learning models«, in: arXiv:1707.08945v5 [cs.CR], 10. April 2018.

44 Zu einem Beispiel im Umgang mit gesprochener Sprache vgl. Hinton, Geoffrey/Deng, Li/Yu, Dong et al.: »Deep Neural Networks for Acoustic Modeling in Speech Recognition«, in: IEEE Signal processing magazine 29 (2012), S. 82-97.

45 Jia, Robin/Liang, Percy: »Adversarial Examples for Evaluating Reading Comprehension Systems«, in: arXiv:1707.07328v1 [cs.CL], 23. Juli 2017.

Abbildung 6: Real World Attacks

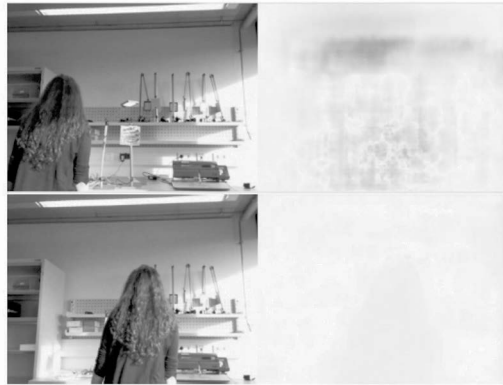


Figure 7. Real World Attacks on FlowNetC. Top: The presence of a printed adversarial patch visible in the centre of the image significantly degrades the optical flow predictions. Bottom: As the patch is covered by the subject, the effect of the patch vanishes.

Ranjan, Anurag/Janai, Joel/Geiger, Andreas u.a. (2019), »Attacking Optical Flow«, ICCV 2019, S. 2404-2413, hier S. 2411

ist.⁴⁶ Das Halluzinatorische und das Groteske, das der Verwechslung von Bananen und Toaster eigen ist, wird um die Eigenschaft des Bedrohlichen 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.).«⁴⁷ Zum Ausweis von Reaktionen auf Sätze findet das Adjektiv *egregious* Verwendung, ein Eigenschaftswort, dessen Bedeutungshof neben dem Herausstechenden und dem Ungeheuerlichen das Feld des Feindlichen und des Böserartigen eröffnet.

Was sich bei diesen Überlegungen über Bedrohungslagen und Sicherheitsrelevanz einschleicht, ist eine Semantik der Verdächtigung von allem,

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

47 Die Verwendung ist kein Einzelfall. Vgl. etwa Sandbank, Tommy/Shmueli-Scheuer, Michal/Herzig, Jonathan et al.: »Detecting Egregious Conversations between Customers and Virtual Agents«, in: arXiv:1711.05780v2 [cs.CL], 16. April 2018.

was die Lebenswelt bereithält. Die Rede von den Potemkinschen Dörfern macht sich breit. Immer wieder wird der Umstand des Befremdlichen und des Unheimlichen angeführt, der mit dem Unterlaufen der menschlichen Wahrnehmung verbunden ist. Mit dem Adjektiv *subtle* werden die 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 Suchbilder aus Zeitschriften und Rätselheften erinnern. Gezeigt werden Bilder, die in prinzipiell erkennbaren und daher auflösbaren Details Wahrnehmungsobjekte in vermeintlicher Gleichheit nebeneinanderstellen – nur mit dem Unterschied, dass das menschliche Auge im Fall der *adversarial attacks* jeder Chance beraubt ist, die eingebauten ›Fehler‹ zu finden.⁴⁸

3. Ally Patches

Ein Brennpunkt der Diskussionen um Sicherheitsrelevanz und die Robustheit der DNN-Systeme ist der Verkehr.⁴⁹ Ausgerechnet die Zeichen seiner Regulierung, die bevorzugter Gegenstand einer höchst eigenen Wahrnehmungsgeschichte waren und deren Aufmerksamkeitsökonomie darauf abgestellt war, Auffälligkeiten sowohl semiotisch als auch ästhetisch herzustellen (und kulturübergreifend abzusichern), erweisen sich als in hohem Maße und mit geringfügigen Mitteln störanfällig.⁵⁰ »The patch attack has real-world significance and we show that adversarial patches can compromise optical flow networks if an engineered patch is printed and placed in real-world scenes.«⁵¹

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

49 Zur Vulnerabilität des Betriebs selbstfahrender Autos durch Adversarial Attacks vgl. stellvertretend Lin, Yen-Chen Lin/Liu, Ming-Yu/Huang, Jia-Bin: »Detecting Adversarial Attacks on Neural Network Policies with Visual Foresight«, in: arXiv:1710.00814v1 [cs.CV], 2. Oktober 2017.

50 Dazu etwa Cohen, A.S.: »Verkehrszeichen«, in: Zeitschrift für Verkehrssicherheit 40/2 (1994), S. 57-67. Zur Erzeugung von Auffälligkeit vgl. Reinisch, Romy: Wahrnehmung von Verkehrszeichen und Straßenumfeld bei Nachtfahrten im übergeordneten Straßennetz, Dissertation, Technische Universität Darmstadt 2010.

51 Ranjan, Anurag/Janai, Joel/Geiger, Andreas et al.: »Attacking Optical Flow«, in: ICCV 2019, S. 2404-241, hier: S. 2405.

Dass Verkehrszeichen auf verhältnismäßig einfache Weise um ihre Eindeutigkeit gebracht werden, stellt eine massive Gefährdung dar und macht Gegenmaßnahmen notwendig – gerade das autonome Fahren muss sich auf die Stabilität solcher Zeichengebungen verlassen können.

Es bezeichnet eine eigentümliche Pointe, dass die Schutzmechanismen und Abwehrmaßnahmen im Fall der Bild- und Schilderkennung nun ihrerseits im stimmigen Modus von so genannten *Ally patches* auftreten. Dass damit eine Spirale von wechselseitig sich störenden Stickern freisetzt wird (*Ally patches for spoliation of adversarial patches*), mag ein Detail sein, das vielleicht nur die Aufmerksamkeit von Kulturwissenschaften auf sich zieht, die derartige Selbstreferenzen als Teil der kultureller Bedeutungswertschöpfungskette gerne in den Blick nehmen – und damit eine Art Habitualisierung betreiben. Aber dass die Sticker die Seiten wechseln und dass nun sie es sind, die die Störer zu stören in der Lage sein sollen und damit Sicherheit garantieren, ist, gerade mit Blick auf seine potentielle Unabschließbarkeit, ein recht notierenswerter Kreislauf.⁵² Weil auch hier das Prinzip der Unmerklichkeit gilt, stellt sich ein eigentümlicher Befund ein: Man vermag die Kombattanten nicht voneinander zu unterscheiden, sind doch vor dem Auge des Menschen alle Sticker gleich. Das Anbringen von Aufklebern hat damit seine Intention verloren und damit auch jenen subversiven Charme, der mit diesen halböffentlichen Akten der Meinungskundgabe einherging. Das Anbringen eines Aufklebers hat unter diesen Bedingungen das Zeug zu einem veritablen Eingriff in den Straßenverkehr und ist in der Lage, erheblichen Schaden zu verursachen. »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.«⁵³ So wird die Anfälligkeit des autonomen Fahrens erkannt und entsprechend zum Gegenstand – wie in einem Fall, bei dem der feindliche Angriff die Scheibenwischer einschaltet, ohne dass es regnet.⁵⁴

52 Zur Aufmerksamkeit der Kultur- und Medienwissenschaft für Belange der Störung vgl. stellvertretend Kümmel, Albert/Schüttpelz, Erhard: *Signale der Störung*, München: Fink 2003.

53 A. Abdel-Hakim: »Ally patches for spoliation of adversarial patches«, S.1.

54 Vgl. dazu Deng, Yao/Zheng, Xi/Zhang, Tianyi et al.: »An Analysis of Adversarial Attacks and Defenses on Autonomous Driving Models«, in: arXiv:2002.02175v1 [eess.SP], 6. Februar 2020.

Besonders auffallend ist dabei eine Einschätzung, die dem Status der entsprechenden Beispiele gilt. Dem übergroßen Anteil der anderen Arbeiten wird ein Aspekt hinzugefügt, der die internen Belange der Störmaßnahmen und der Gegenmaßnahmen verlässt und damit auch den Bereich der Taxonomie der jeweiligen Maßnahmen.⁵⁵ Statt der Frage, welches der geschilderten Verfahren (*Gradient Masking*, *Robust Optimazion*, *Detection*) zum Einsatz gelangt, wird mit dem Versuch der Statusbestimmung die Frage drängend, welche Bedeutung ihr in einer Lebenswelt zukommt. Dabei spielen neben dem Katz- und-Maus-Spiel des jeweiligen Reagierens Aspekte eine Rolle, die dem Bereich des Grundlagenwissens zuzurechnen sind.⁵⁶ Immer wieder, so ist zu lesen, würden die Spiralen und Kreisläufe der Attacken und Counter-Attacken das Verständnis in die Eigendynamik der DNN-Systeme verstärken – und damit ein Wissen über deren Natur zu Tage fördern. Und es ist nur naheliegend, sich in den Rückkopplungsschleifen von Tarnen und Täuschen an Beispiele zu halten, die in der Natur zu finden sind – wie in einem Beitrag, der neben der Gottesanbeterin auch echte und als Roboter nachgebaute Schmetterlinge zum Einsatz gelangen lässt.⁵⁷ Wie es in einem der damit befassten Texte heißt, sind es nicht Momente des Fehlerhaften, sondern des System-Typischen, die dabei als positives Wissen zutage treten (*Adversarial Examples Are Not Bugs, They Are Features*).⁵⁸ Ein weiterer Text schlägt vergleichbare Töne an, ersetzt dabei aber die Rede vom *Bug* durch die vom *Flaw*. »These results have often been interpreted as being a flaw in deep networks in particular, even though linear

-
- 55 Vgl. dazu die besonders eindrückliche Darstellung bei H. Xu et al.: »Adversarial Attacks«.
- 56 Ein vergleichbarer Befund gilt für den Bereich der DeepFakes. Vgl. dazu Greenard, Samuel: »Will Deepfakes Do Deep Damage?«, in: Communications of the ACM 63/1 (2020), S. 17-19, sowie für den Bereich der Gesichtsm Manipulation Ruiz, Nataniel/Bargal, Sarah Adel/Sciaro, Stan: »Disrupting Deepfakes: Adversarial Attacks Against Conditional Image Translation Networks and Facial Manipulation Systems«, in: arXiv:2003.01279v3 [cs.CV], 27. April 2020. Die Attacken beschränken sich nicht auf andere, weniger brisantere und andere Sinne betreffende Anwendungsgebiet. Vgl. dazu stellvertretend Nakano, Kizashi/Horita, Daichi/Sakata, Nobuchika et al.: »Enchanting Your Noodles: GAN-based Real-time Food-to-Food Translation and Its Impact on Vision-induced Gustatory Manipulation«, in: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces, 23-27 March, Osaka, Japan 2019, S. 1096-1097.
- 57 Xi, Bowei/Chen, Yujie/Fei, Fan et al.: »Bio-Inspired Adversarial Attack Against Deep Neural Networks«, in: SafeAI@AAAI, 2020, S. 1-5.
- 58 Ilyas, Andrew/Santurkar, Shibani/Tsipras, Dimitris et al.: »Adversarial Examples Are Not Bugs, They Are Features«, in: arXiv:1905.02175v4 [stat.ML], 12. August 2019.

classifiers have the same problem. We regard the knowledge of this flaw as an opportunity to fix it.«⁵⁹

Als besondere Pointe dieser unfreiwilligen Grundlagenarbeit an den Funktionsweisen neuronaler Netze tritt zu Tage, dass diese eben anders als die menschliche Datenverarbeitung funktionieren. Ihr technischer Eigensinn hat sich auf eine gewisse Weise davon emanzipiert. Die Ordnung der Dinge ist nicht mehr zwangsläufig eine Ordnung, die entlang von Ähnlichkeiten und Repräsentationen, von menschlich nachvollziehbaren Voraussichten und Reaktionen anzuschreiben ist. Die Patches und die von ihnen ausgehende Gefährdung durch Verkehrsschilder, viel mehr noch die Manipulation komplexer Situationen und Lagen, die für die Stabilität und Akzeptanz des autonomen Fahrens so zentral wäre, stellen ihren technischen Eigensinn und damit eine epistemische Sonderstellung unter Beweis. *Ihre* Ordnung der Dinge ist nicht *die* Ordnung der Dinge. Es wird im und um das Auto zu entsprechenden Passionen kommen.

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