THE GESTALT OF AI: BEYOND THE HOLISM-ATOMISM DIVIDE

By Hannes Bajohr

"One could therefore argue that neural networks do not only produce outputs that humans perceive as Gestalten, but that, as statistical models, they internally already operate according to a Gestalt logic."

Suggested citation:
Hannes Bajohr, The Gestalt of AI: Beyond the Holism-Atomism Divide. *Interface Critique* 3 (2021), pp. XX–XX.
DOI: XX.

Fig. 1, left to right: Output and input of Mario Klingemann's application of the Pix2Pix deep learning model (2017, https://twitter.com/quasimondo/status/934709314375372801) compared to Leonardo da Vinci's *La Gioconda* (*Mona Lisa*, ca. 1502/03).

Let me start with a face: In the fall of 2017, media artist Mario Klingemann posted a picture on Twitter that, at first glance, looked quite familiar, indeed clichéd (fig. 1, left). The smile was there, as was the slight turn of the head to the right and the familiar look directed at or slightly behind the viewer. And yet, a closer look raised doubts whether this was indeed the often- and over-reproduced likeness of the Gioconda. In fact, it was Klingemann's own creation, the result of applying the recently published deep-learning model Pix2Pix to a blurry image file of Leonardo's painting.

Klingemann's version of Pix2Pix achieves what until then seemed possible only in science fiction movies like *Bladerunner*, in which a character points to a grainy CCTV image and orders the computer to "enhance" it.[1] What was previously fictitious is now reality: image enhancement makes it possible to compensate for the loss of data that occurs when an image is down-sampled to a lower resolution, and to highlight details that were not visible in the original. Indeed, Klingemann's input did not consist of the real *Gioconda*, but a blurred black-and-white version of the painting (fig. 1, center). From *this* image, in which the features of the face are all but invisible, Pix2Pix produced the output. The direct comparison with the original clearly shows the differences between Klingemann's and Leonardo's *Mona Lisa* (fig. 1, right) – notice the over-exact details, the glitchy eyelids, the flowing hair more reminiscent of a shampoo commercial than the painted likeness of a sixteenth-century Florentine woman. The output is not truly an *enhancement* of

1 This text first appeared in German as "Die 'Gestalt' der KI: Jenseits von Atomismus und Holismus," *Zeitschrift für Medienwissenschaft* 23, No. 2 (2020), pp. 168–181. It is based on a lecture I gave at the conference *Things Beside Themselves: Mimetic Existences* at the IKKM Weimar in March 2020. I would like to thank the participants for their comments, Mario Klingemann for the kind permission to reproduce his images, and Julia Pelta Feldman, Florian Sprenger, and Jana Mangold for helpful suggestions.

The architectures of neural networks are usually published as PDFs on the open access repository arXiv.org and are not peer reviewed. This allows for the fastest possible publication and makes the provided exemplary results and the code available via GitHub an additional basis for evaluation. For Pix2Pix, see: Phillip Isola et al., Image-to-Image Translation with Conditional Adversarial Networks. *ArXiv*, November 21, 2016; arxiv.org/abs/1611.07004, access: June 6, 2020.

Fig. 2: Generated portraits, *thispersondoesnotexist.com*, 2019.

the original, but a *new creation* based on a few features of its overall appearance. Pix2Pix thus does not restore details that were blurred out – by the principle of entropy, lost information remains lost – but rather, it plausibly interpolates a face from the input image by drawing on its knowledge of what faces usually look like.[2] It does this not by being fed explicit rules about which elements constitute a face – where the eyes go, what an eyebrow looks like and so on – but by learning, without any guidance, what likely constitutes "face-ness." Thus, Klingemann's painting is not a composite picture. It is not a collage, made of isolated elements, nor is it simply a mean of other faces, in the style of Francis Galton, that converges towards the features that the majority of objects of a class share by way of linear regression.[3] This is made obvious by another face-generating artificial neural network (ANN) called StyleGAN2, developed by the graphics card manufacturer NVIDIA, that is on display on the website thispersondoesnotexist.com (fig. 2). As the URL indicates, these images do not show real people. Rather, the "photos" are generated anew each time I refresh my browser window. These faces have enough individual features to suggest that they are neither a collage nor a mere collection of the most common features in a series. Whatever generates the faces in this process prioritizes the *whole* over its constituent *parts*. It seems as if Pix2Pix and StyleGAN2 have learned and then reproduced the *Gestalt* of a face.

In what follows, I would like to take up the concept of Gestalt, but neither as a technical description nor as a phenomenon of human perception, which is usually the focus of Gestalt psychology. Instead, I will talk about the conceptual preconditions that play a role in the representation and emergence of non-derivable entities in discrete systems. The term Gestalt is intended to help to

2 In this case, however, the training data set only included female faces according to Klingemann; https://twitter.com/quasimondo/status/934546438507376640, access: June 6, 2020.

3 See Suzanne Bailey, Francis Galton's Face Project: Morphing the Victorian Human. *Photography and Culture* 5, No. 2 (2012), pp. 189–214.

discuss some of the assumptions that underlie the theorization of a particular type of artificial intelligence, which is summarized under the term "deep learning" and implemented by means of multi-layer perceptrons.[4] I will argue that we should place the conceptualization of current state-of-the-art machine learning technology *beyond* or maybe *beside* the two philosophical lineages that usually are mobilized to explain the possibility or impossibility of artificial intelligence: What one can very broadly call atomistic theories on the one side and holistic theories on the other are both unfit to describe the particular type of 'Gestalt' effects Klingemann's Gioconda and the technology of its production display. Instead, it is more illuminating to speak either of a "mixed type" or of something different altogether. To make this argument, I will briefly outline the division between atomism and holism, show its appearance in the competing approaches to AI, and delve into the mechanics of deep neural nets themselves. In a final step, I will suggest that one may use the mixed type as a conceptual tool for nontechnical domains – as an intuition pump, as Daniel Dennett calls it.[5]

My use of the face as an example of Gestalt effects is no coincidence. "[T]he human face with its unequalled situative meaning"[6] is – from Georg Simmel's aesthetic unifying function, to Emmanuel Lévinas' constitutive connection of "face and ethics," to Hans Belting's image anthropology – an object of investigation with its own philosophical, art historical, and cultural genealogy.[7] As a prime example of maximally irreducible meaning, and even as an "anthropogenetic primal type" of significance in general,[8] it is particularly well suited for investigating the possible correspondences and incongruences of technical structures and life-worldly expectations of meaning.

4 Helpful for the recent discussion: Christoph Engemann and Andreas Sudmann (eds.), *Machine Learning: Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz* (Bielefeld 2018), as well as the special issues of *Zeitschrift für Medienwissenschaft* 11, No. 2 (2019) and *spheres* 5, No. 5 (2019).

5 Daniel C. Dennett, *Consciousness Explained* (New York 1991), p. 440.

6 Hans Blumenberg, Prospect for a Theory of Nonconceptuality, in: *History, Metaphors, Fables: A Hans Blumenberg Reader*, ed. Hannes Bajohr, Florian Fuchs and Joe Paul Kroll (Ithaca, NY 2020), p. 242.

7 Georg Simmel, *The Aesthetic Significance of the Face, in: Essays on Art and Aesthetics*, ed. Austin Harrington (Chicago 2020), pp. 231–235; Emmanuel Lévinas, *Totality and Infinity: An Essay on Exteriority*, trans. Alphonso Lingis (Dordrecht 1991), pp. 194–219; Hans Belting, *Face and Mask: A Double History*, trans. Thomas S. Hansen and Abby J. Hansen (Princeton 2017). For a cultural studies approach, see also Thomas Macho, *Vorbilder* (Munich 2011), and Sigrid Weigel (ed.), *Gesichter: Kulturgeschichtliche Szenen aus der Arbeit am Bildnis des Menschen* (Paderborn 2013) – by the same author on the digital aspect of faciality, see Der konventionelle Code als buckliger Zwerg im Dienste der Emotion Recognition. Überlegungen zu einer Urgeschichte der digitalen Kultur, in: *Internationales Jahrbuch für Medienphilosophie* 6, no. 1 (2020), pp. 47–79; for an evolutionary approach see Terry Landau, *About Faces* (New York 1989).

8 Blumenberg, Prospect, 242.

1. Holism, Atomism, Gestalt

To apply the term "Gestalt" to Klingemanns Gioconda means, first, to call up the definition of Gestalt psychology, and I will talk about it in a moment. However, the relationship between parts and whole is also used metonymically beyond Gestalt theory to describe a bifurcation in the history of philosophy between two traditions or schools of thought that are usually called holism and atomism.[9]

Atomism is the belief that every object and its specific properties can be explained by breaking that object down to its constitutive elements, and that such an explanation is exhaustive. Modern adherents of atomism stand in the tradition of Gottlob Frege; Bertrand Russell's logical atomism, the picture theory of the early Wittgenstein, the logical positivism of the Vienna Circle, and the sense data theory of G. E. Moore and A. J. Ayers are the most important positions of this tradition. Atomism is, on the whole, objectivist, reductionist and empiricist. It tends to look at the semantic rather than the pragmatic dimension of knowledge, at "knowing-that" rather than "knowing-how," as Gilbert Ryle put it.[10]

Holism is the reverse belief that the properties of a thing cannot exhaustively be explained by the properties of its constitutive elements. In this line of tradition, the whole is conceptually or causally prior to its parts. Related are terms like "structure" or, in Kant's case, "system" as opposed to the atomistic "aggregate."[11] Holism in the 20th century is represented above all by the hermeneutic phenomenology of Heidegger and Merleau-Ponty as well as by the late Wittgenstein and his followers.[12] Such a holism engages less with explicit propositional than with implicit pragmatic and world-constitutive knowledge.[13]

The notion of Gestalt is holistic in

9 In the following I will limit myself to a (general) discussion of atomism and holism in philosophy. However, these terms have also been applied to many other fields (physics, biology), see Michael Esfeld, Holismus und Atomismus in den Geistes- und Naturwissenschaften: Eine Skizze, in: *Holismus und Individualismus in den Wissenschaften*, ed. Alexander Bergs and Soelve I. Curdts (Frankfurt/Main 2003), pp. 7–21; Georg Toepfer, Ganzheit, in: *Historisches Wörterbuch der Biologie* (Stuttgart 2011), pp. 693–728.

10 Gilbert Ryle, Knowing How and Knowing That. *Proceedings of the Aristotelian Society* 46, No. 1 (1946), pp. 1–16.

11 Immanuel Kant, *Prolegomena to Any Future Metaphysics*, trans. Gary Hatfield (Cambridge 2004), p. 74 (4: 322).

12 The distinction atomistic/holistic is not to be equated with the questionable distinction analytical/continental. The analytic tradition, too, has a strong anti-atomistic current, most importantly in the criticism of the sense data theory as a "Myth of the Given" in Wilfrid Sellars, *Empiricism and the Philosophy of Mind* (Cambridge, MA 1997), pp. 68–79.

13 Charles Taylor formulated an influential application of the atomism/holism separation for the theory of meaning. He contrasts the theory of meaning of the atomistic tradition, which he calls the *enframing* theory, with the theory of the holistic lineage, which he calls the *expressive-constitutive* theory. For the latter, to articulate something means to make possible the perception of this feature in the first place. One is constitutive for the other, however only because the context of this operation is not neutral, but rather already shaped by a complex background knowledge that is not propositional in nature, but is rooted in an expressive practice, in forms of life, Charles Taylor, Theories of Meaning, in: *Human Agency and Language: Philosophical Papers I* (Cambridge 1985), pp. 247–292; see also Charles Taylor, *The Language Animal: The Full Shape of the Human Linguistic Capacity* (Cambridge, MA 2016).

this sense.[14] Christian von Ehrenfels, who coined the term "Gestalt qualities" in 1890, famously noted that the perception of what constitutes a melody is a unit that cannot be reduced to the sequence of individual notes. He turned away from a mere psychology of association, which argued in a purely atomistic and causal manner.[15] The sense of sight quickly advanced to become the central field of investigation of 20th century Gestalt psychology, as it was repeatedly explained, above all, in the Berlin School around Wolfgang Köhler, Kurt Koffka, and Max Wertheimer, as well as by their second generation students, e.g., Wolfgang Metzger.[16] Köhler remarked, similar to Heidegger at the same time, that we "do not perceive an undifferentiated mosaic; rather, it is characteristic of our seeing, hearing, etc., that it constantly shows units and groups that, being in themselves solid, appear relatively isolated from their surroundings."[17] Such units, the "Gestalten," exhibit an inner coherence and stability that Wertheimer called "concisiveness" (*Prägnanz*).[18] These are not based on "independent elementary sensations" – an aggregate of atomistic sensory data – but are made up of "local conditions," which are "dependent on their affiliation, position and role in the Gestalten."[19] Further, "insight" (*Einsicht*), the sudden perception of Gestalt configurations, became a measure of intelligence in Köhler's ape experiments.[20] Such Gestalten, both as visual figures and as constellations of insight, are non-derivable units of significance that must be understood *holistically*.

One of these non-derivable units is the face. Metzger thus states that in order to perceive a face in its expressive significance, one must look at it as a whole. And while it may help to attend to the movement of the brows or the mouth, to "zoom in" any further and isolate parts is detrimental to the perception of this whole:

> *The individual pores, hairs, wrinkles of the lips, freckles, etc., which the further focusing of attention brings to light, contribute nothing more to the understanding of the face. Each of these details could also be different without changing the face. And none of them says anything about what a face is really about in life; whether it is, for example, an arrogant, domineering, hard, closed, hostile or a soft, warm, open-minded and compassionate face. These decisive features become most clearly visible, or only visible*

14 Gestalt psychology is part of the holistic line, but not all holists are followers of Gestalt theory, see for example, despite some undeniable influences, Merleau-Ponty's critique of Husserl's reception of Gestalt psychology and of Gestalt psychology itself, Maurice Merleau-Ponty, *Phenomenology of Perception* (London 2005), pp. 58–9.

15 Christian von Ehrenfels, On "Gestalt Qualities" [1890], in: *Foundations of Gestalt Theory*, ed. Barry Smith (Munich 1988), pp. 82–116.

16 On the history of Gestalt psychology (including its history of emigration and collaboration under National Socialism), see Mitchell G. Ash, *Gestalt Psychology in German Culture, 1890–1967: Holism and the Quest for Objectivity* (Cambridge 1995). Exemplary for the focus on the sense of sight is Wolfgang Metzger, *Laws of Seeing* [1936], trans. Lothar Spillmann (Cambridge, MA 2006).

17 Wolfgang Köhler: Bemerkungen zur Gestalttheorie. *Psychologische Forschung* 11, no. 1 (1928), pp. 188–189.

18 Max Wertheimer, Untersuchungen zur Lehre von der Gestalt II. *Psychologische Forschung* 4, no. 1 (1923), pp. 301–350.

19 Köhler, Bemerkungen, p. 189.

20 Wolfgang Köhler, *Mentality of Apes* (New York 1927).

> *at all, when viewed as a whole [als ganzes] from a sufficient distance.*[21]

If Pix2Pix and StyleGAN2 are able to derive faces as wholes, as Gestalten seen from a distance, might it make sense to assume that they are based on a holistic logic? But how should a digital system that is based on the symbolic operation of discrete signs and, not least of all, makes use of a discrete pixel matrix (which is nothing other than Köhler's "mosaic") create non-derivable units? To answer this question, one has to take a look at the history of AI systems, of which ANNs are just one paradigm, and at the concepts underlying them.

2. Gestalt vs. AI

The classic account of the history of AI highlights its emergence as a research field in the 1940s and 1950s in the United States. It appeared in a climate of neobehaviorist, reductionist, and empiricist psychology, which entered into a productive confluence with the methods and concerns of cybernetics. The Macy Conferences between 1946 and 1952, the Hixon Symposium in 1948, and particularly the Dartmouth Workshop on Artificial Intelligence in 1956 are important milestones in this history. Especially the last, organized by Marvin Minsky and John McCarthy, established AI research as an independent field and determined the parameters under which it would be pursued in the years to follow.[22]

The distinction between two types of AI, symbolic and subsymbolic – which is still in use today – has its origin in this workshop.[23] The symbolic approach, favored heavily at Dartmouth,[24] is the most classically atomist attempt at creating AI. It conceives reasoning as the manipulation of symbols representing atomic facts. The symbolic approach was implemented in so-called expert systems, which combine a knowledge base of such facts with an inference engine containing rules that allow it to draw conclusions from the combination of these facts.[25] Expert systems initially showed great promise, but their development came to a standstill in the 1970s during the first "AI winter," in which AI research

21 Wolfgang Metzger, Was ist Gestalttheorie?, in: *Gestalttheorie und Erziehung*, ed. Kurt Guss (Darmstadt 1975), pp. 2–3.

22 Proceedings are available for the former: Claus Pias (ed.), *Cybernetics: The Macy-Conferences 1946–1953* (Zurich 2003); Lloyd A. Jeffress (ed.), *Cerebral Mechanisms in Behavior: The Hixon Symposium* (New York 1951). The Dartmouth Workshop did not see a publication of its own, however see Ronald R. Kline, Cybernetics, Automata Studies, and the Dartmouth Conference on Artificial Intelligence. *IEEE Annals of the History of Computing* 33, no. 4 (2011), pp. 5–16. For a, somewhat limited, historical overview, see Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge 2010); Steve Joshua Heims, *The Cybernetics Group 1946–1953: Constructing a Social Science for Postwar America* (Cambridge, MA 1991).

23 For a very useful overview, see Melanie Mitchell, *Artificial Intelligence: A Guide for Thinking Humans* (New York 2019), pp. 17–34.

24 However, the conference proposal already envisaged research on "neuron nets," see John McCarthy et al., A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. *AI Magazine* 27, no. 4, 2006, pp. 12–14.

25 The most influential system of this kind was the General Problem Solver (GPS), Allen Newell, J. C. Shaw and H. A. Simon, Report on a General Problem-Solving Program, in: *Proceedings of the International Conference on Information Processing* (Paris1959), pp. 256–264.

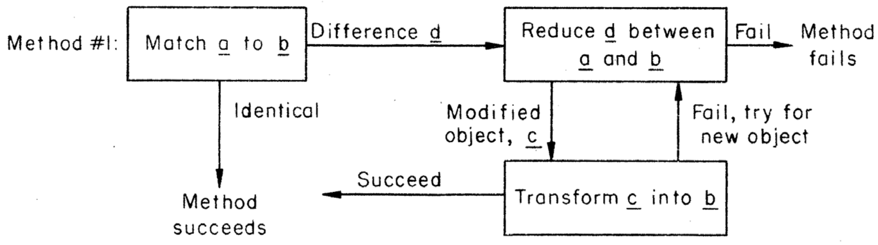

Fig. 3: Flow chart of a method for "means-ends analysis" in Allen Newell, John C. Shaw and Herbert A. Simon, Report on a General Problem-Solving Program, in: *Proceedings of the International Conference on Information Processing* (Paris: UNESCO, 1959), pp. 256–264.

virtually ground to a halt.[26]

Today's ANNs like Pix2Pix and StyleGAN2 do not belong to the symbolic but to the subsymbolic family of AI, which is, abstractly, based on the model of the brain as a network of neurons and synapses. Building on the preliminary work of Warren McCulloch and Walter Pitts, Frank Rosenblatt designed the perceptron in 1958, the first artificial neural network capable of recognizing simple visual patterns.[27] While "learning" in expert systems means the expansion of the knowledge base, perceptrons are dependent on repetitions within the domain to be learned; whereas the expert system follows linear if-then structures (fig. 3), the architecture of the perceptron has a parallel structure and does not require the separation of facts and rules (fig. 4). Already in its structure, the paradigm of the symbolic AI follows an atomistic logic, while the paradigm of the subsymbolic AI approaches a holistic or Gestalt logic.[28]

This difference was obvious from the start. As David Bates and Steve Joshua Heims have shown, American AI research, after a brief initial interest, soon became hostile to the Gestaltist ideas.[29] In a 1951 review of Norbert Wiener's book *Cybernetics*, Wolfgang Köhler – who had participated in the 1948 Hixon Symposium alongside AI pioneer Warren McCulloch – spoke out against the idea of the computer as a useful analogy for human intelligence because the former, as a discretely operating system, lacked the creative "insight" of the latter.[30] He

26 See Pamela McCorduck, *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence* (Natick 2004), pp. 417–521.

27 Frank Rosenblatt, The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review* 65, no. 6 (1958), pp. 386–408; Nilsson, T*he Quest for Artificial Intelligence*, 64–74; see also Matteo Pasquinelli, Machines that Morph Logic: Neural Networks and the Distorted Automation of Intelligence as Statistical Inference. *Glass Bead* 1, no. 1 (2017); www.glass-bead.org/article/machines-that-morph-logic, access: June 6, 2020.

28 Taking the technical difference between the two approaches as a starting point, I also develop aesthetic criteria for comparing the artworks they produce, see Hannes Bajohr, Algorithmic Empathy: On Two Paradigms of Digital Generative Literature and the Need for a Critique of AI Works. *BMCCT working papers* 1, no. 4 (2020); https://doi.org/10.5451/unibas-ep79106.

29 See in particular the excellent study by David Bates: Creating Insight: Gestalt Theory and the Early Computer, in: *Genesis Redux. Essays in the History and Philosophy of Artificial Life*, ed. Jessica Riskin (Chicago 2007), pp. 237–260; Heims, *The Cybernetics Group 1946–1953*, 201–247; Wolfgang Köhler, Relational Determination in Perception, in: Jeffress (ed.): *Cerebral Mechanisms in Behavior*, pp. 200–243.

30 Wolfgang Köhler, review of *Cybernetics, or Control and Communication in the Animal and the Machine*, by Norbert Wiener. *Social*

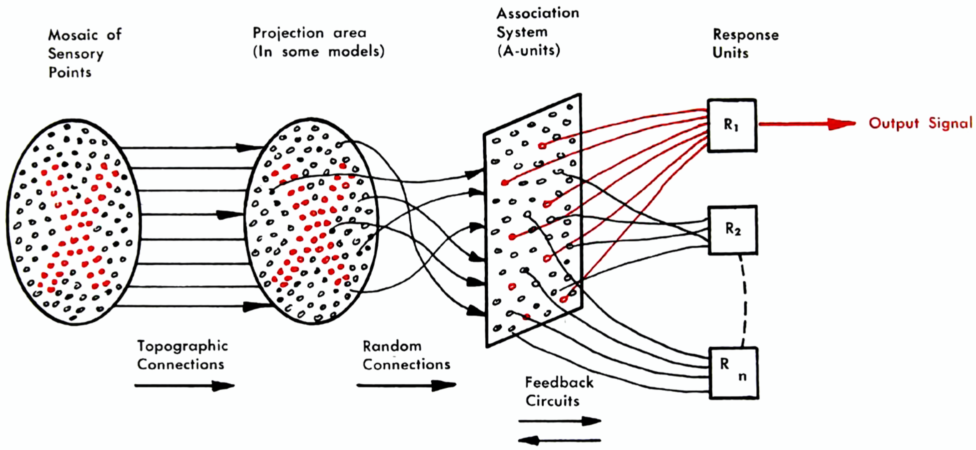

Fig. 4: Frank Rosenblatt, The Design of an Intelligent Automaton. *Research Trends* 6, no. 2 (1958), pp. 1–7.

saw the atomistic psychology of the AI researchers as a return to the empiricist psychology of the 19th century, which Gestalt psychology had initially sought to replace.[31] Nevertheless, the subsymbolic models, such as Rosenblatt's perceptron, soon lost out to the symbolic approach. Marvin L. Minsky and Seymour Papert, two of the most important symbolists, published a (factually incorrect) critique of the perceptron in 1969 that cast it aside until the 1980s.[32] They formulated their objections as a defense of an atomistic against an "unscientific" holistic theory of knowledge – a reversal, as it were, of Köhler's reservations – intending "to dispel what we feared to be the first shadows of a 'holistic' or 'Gestalt' misconception that would threaten to haunt the fields of engineering and artificial intelligence as it had earlier haunted biology and psychology."[33]

One of the first *philosophers* to formulate a holistically informed critique of the symbolists was the late Hubert L. Dreyfus. In a series of essays and in his book *What Computers Can't Do* (1972), he argued that the symbolic approach is fundamentally incapable of producing human-like intelligence.[34] Dreyfus mobilized a number of holistic arguments. His central point was that humans not only possess an embodied intelligence, but

Research 18, no. 1 (1951), pp. 125–130.

31 Bates, *Creating Insight*, pp. 239–249.

32 Marvin L. Minsky and Seymour Papert, *Perceptrons: An Introduction to Computational Geometry* (Cambridge, MA 1969).

33 Minsky and Papert, *Perceptrons*, pp. 19–20.

34 Hubert L. Dreyfus: *What Computers Can't Do: A Critique of Artificial Reason* (New York 1972). Symbolism, according to Dreyfus, is based on a number of atomistic assumptions: the *biological* assumption that the brain can be identified with a digital computer, the *ontological assumption* that the world consists of isolatable facts, and the *epistemological assumption* that the mind processes such facts. For such an atomism, thinking can be formalized by explicit rules – knowing-how can be expressed as knowing-that, Dreyfus, *What Computers Can't Do*, 67–142. See in general on Dreyfus's approach Setargew Kenaw, Hubert L. Dreyfus's Critique of Classical AI and its Rationalist Assumptions. *Minds and Machines* 18, no. 2 (2008), pp. 227–238.

also draw on a tacit, implicit background knowledge that is constitutive for action – they perceive and cognize from their embeddedness within a given situation. Here, Dreyfus (not by chance one of America's most important interpreters of Heidegger) appropriated Being-in-the-world as "being-in-a-situation."[35] The way humans intelligently interact with the world is more often, as Charles Taylor (an important ally of Dreyfus) interprets Heidegger, based on "task-rightness," situational appropriateness, than on semantic rightness, the logically formalizable congruence of data and abstract world model.[36] One could only attribute intelligence to a computer possessing this implicit background knowledge, which can only be gained through actually encountering the world via "being-in-a-situation." For Dreyfus the conclusion is that "being-in-a-situation turns out to be unprogrammable in principle using presently conceivable techniques."[37]

This objection was convincing as long as atomistic assumptions formed the basis of the "presently conceivable techniques" of AI research. However, Dreyfus was less certain in his criticism when it came to the architecture of the perceptron. When ANNs, which are in essence multi-layered perceptrons, regained popularity in the 1980s,[38] Dreyfus also admitted that they came closer to a holistic notion of perception. But he remained skeptical as to whether they were really suitable as a building block of artificial machine intelligence.[39]

However, if the capabilities of today's ANNs are taken into account, even Dreyfus would have had to admit that progress has been made in principle, not just in degree. In *What Computers Can't Do* he had listed a series of tasks that a system would have to master in order to be considered intelligent. One of the most important of these was a version of the Gestalt problem that he called "perspicuous grouping," by which he meant the ability to form series of objects on the basis of similarities between them – to grasp their collective Gestalt, so to speak. In addition to Wittgenstein's concept of family resemblance, he used Köhler's notion of insight to describe such group schemes. Family resemblances as well

35 Dreyfus, *What Computers Can't Do*, p. 200.

36 Charles Taylor, Heidegger on Language, in: *A Companion to Heidegger*, ed. Hubert L. Dreyfus and Mark A. Wrathall (London 2007), pp. 435–436.

37 Dreyfus, *What Computers Can't Do*, p. 215.

38 They did so after the publication of David E. Rumelhart, James L. McClelland, PDP Research Group: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Cambridge, MA 1986), which corrected Minsky's and Papert's misrepresentations – especially the alleged inability of perceptrons to model the exclusive disjunction (XOR). Dreyfus' brother Stuart was partly responsible for allowing this new generation of neural nets to "learn" by co-developing the backpropagation algorithm that efficiently calculates the gradients of the loss function through which the weights of the network can be adjusted iteratively.

39 Hubert L. Dreyfus and Stuart E. Dreyfus, Making a Mind versus Modeling the Brain: Artificial Intelligence Back at a Branchpoint. *Daedalus* 117, no. 1 (1988), pp. 15–43: "Neural network modeling may simply be getting a deserved chance to fail, as did the symbolic approach" (37). One contemporary position stating exactly can be found in Brian Cantwell Smith, *The Promise of Artificial Intelligence: Reckoning and Judgment* (Cambridge, MA 2019). Most recently, Dreyfus gets an update in Ragnar Fjelland, Why General Artificial Intelligence Will Not Be Realized. *Humanities and Social Sciences Communications* 7, no. 10 (2020); https://doi.org/10.1057/s41599-020-0494-4, access: October 10, 2020.

as collective Gestalten cannot be grasped by counting up atomistic properties and comparing lists of characteristics:

> *Patterns as complex as artistic styles and the human face reveal a loose sort of resemblance which seems to require a special combination of insight, fringe consciousness, and ambiguity tolerance beyond the reach of digital machines.*[40]

This brings us back to the face and again to the question: If the Gestalt of a face, its family resemblance with other faces, cannot be conceptualized by a digital machine or summarized as a list of features, how is an ANN, executed on a digital machine, capable of doing this?

3. Gestalt as latent space

At this point it is necessary to take a step back again. For Dreyfus's overriding focus – human-like artificial intelligence – is irrelevant to answering the question posed. As Matteo Pasquinelli points out, ANNs are aimed not so much at simulating cognition but perception.[41] Today, it is mainly "deep" ANNs that show their capabilities in pattern recognition tasks. Indeed, Dreyfus's "perspicuous grouping" is only a particularly demanding type of such pattern recognition able to detect family resemblance without having explicit knowledge about it. Rosenblatt's perceptron was modeled after the optic nerve of the eye, not the cerebrum, and was composed of three main elements: the input layer, a hidden layer and the output layer. Modern ANNs, deep neural networks, still follow this structure, but possess a multitude of hidden layers, which consist of artificial "neurons" that act as logical gates and are connected by "synapses." These synapses in turn have an influence on the activation strength of the next neuron by being weighted in the training process.[42] The goal of a neural net is to create a function that fits the input data onto a desired output, and apply this function to future inputs to predict their outputs. As far as StyleGAN2 is concerned, the ANN is here trained to output variations of its input: its input being a large set of faces, while its output consists of new faces. In fact, StyleGAN2 and, to a certain extent, Pix2Pix, use a special architecture of ANNs, a generative adversarial network (GAN), in which these processes are separated but the basic structure is the same.[43]

40 Dreyfus, *What Computers Can't Do*, 32. It is astonishing how rarely the proximity of Gestalt psychology to neural networks in particular is investigated. An exception is Uwe Seifert, Randolph Eichert, Lüder Schmidt, Logic, Gestalt Theory, and Neural Computation in Research on Auditory Perceptual Organization, in: Marc Leman (ed.), *Music, Gestalt, and Computing: Studies in Cognitive and Systematic Musicology* (Berlin 1997), pp. 70–88.

41 Pasquinelli, Machines that Morph Logic. ANNs are, using John Searle's distinction, still examples of a weak, not a strong AI, John R. Searle, Minds, Brains, and Programs. *Behavioral and Brain Sciences* 3, no. 3 (1980), pp. 417–457.

42 See for this and in the following: Pedro Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World* (New York 2015), pp. 93–120; Ethem Alpaydin, *Machine Learning: The New AI* (Cambridge, MA 2016), pp. 85–110.

43 See for the original formulation of the GAN architecture Ian Goodfellow et al., Generative Adversarial Networks. *ArXiv*, June 6,

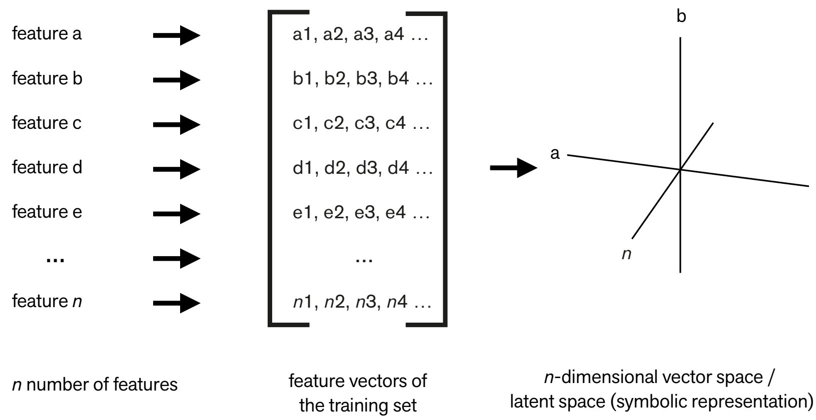

Fig. 5: Feature extraction and latent space mapping.

However, the data on the basis of which this function is generated must first be prepared for processing and a limited number of features must be selected from all possible features present. In supervised learning techniques, this is done manually by the programmer, for example when images in a data set of portraits are marked depending on whether a person is smiling or not. With unsupervised learning, on the other hand, such features are found automatically by the ANN. Each of the layers of the ANN is tasked with extracting salient patterns – Prägnanz, one could say – from the input of the previous layer. Since this happens progressively between layers, a process of abstraction is at work: The first layer may look at a combination of a few pixels, and then pass the result on to the next layer, which now looks at a combination of a combination of pixels, and so on. And while the first layer may only detect edges (high contrast between pixels), the second layer already groups edges into simple shapes (straight lines or curves), the third into parts of objects, the fourth into objects, and so on.[44] The n features derived from the input are mapped onto an n-dimensional vector space (fig. 5). In this vector space, it is possible to calculate the interdependence of all features and to reduce the number of features to a lower-dimensional space or "latent space." This process of abstraction is called dimensionality reduction and "it reduces a large number of visible [or explicit] dimensions (the pixels) to a few implicit ones (expression, facial features)."[45]

In the case of faces, mapped onto two dimensions, one could imagine the latent space as shown in fig. 6. The model can now recreate the explicit dimensions of the input from the implicit ones available in the model. This is what Pix2Pix does when it interpolates the facial features of

2014; arxiv.org/abs/1406.2661, access: June 6, 2020.

44 See Yann LeCun, Yoshua Bengio and Geoffrey Hinton, Deep Learning. *Nature* 521 (May 2015), pp. 436–444; Alpaydin, *Machine Learning*, pp. 75 and 99–100.

45 Domingos, *The Master Algorithm*, p. 211.

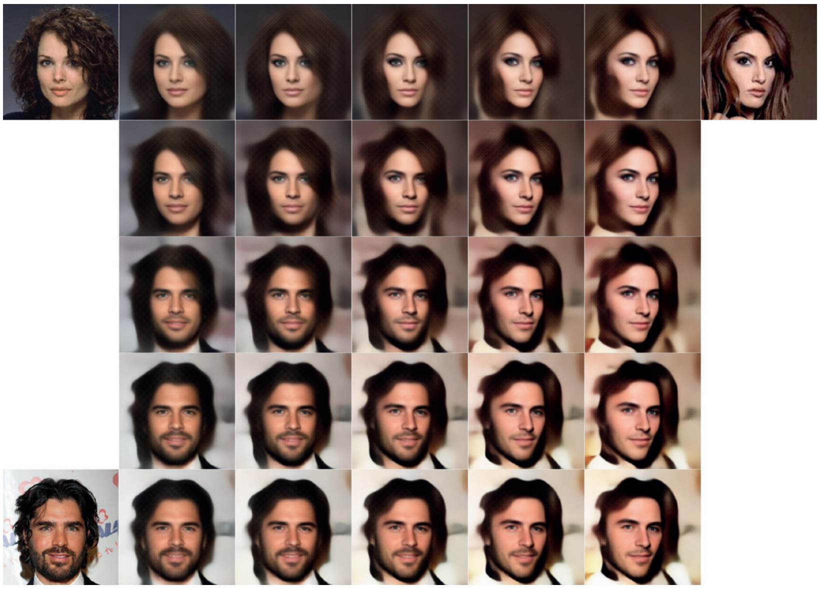

Fig. 6: Example for latent space interpolation. Top left: the input image; top right: the first target image; bottom left: the second target image; the remaining images are the interpolations of the model. Visualization as J-Diagram in a paper by Tom White, 2016, https://arxiv.org/pdf/1609.04468.pdf

the blurred input image and fills its gaps with the probable states from its model. It can also, simply by randomly selecting feature values, produce new outputs from the dimensions of the latent space. In this way, StyleGAN2 generates a new face every time I refresh thispersondoesnotexist.com. What both operations have in common is that their basis is the *overall structure* of what is modeled – in this case, a face.[46]

46 Another way to imagine the latent space is as a series of slide controls for changing any feature of the face – from the overall color tone to the direction of the light source to the facial expression – containing all degrees of abstraction that have taken place during feature extraction. It is also possible to "travel" through this high-dimensional latent space ("manifold traversal") and to interpolate any possible configuration, Robert Luxemburg, StyleGAN2 Interpolation Loop, December 12, 2019; youtube.com/watch?v=6E1_dgYlifc, access: June 5, 2020.

4. "Quasi-analog" and "quasi-holistic"

With regard to Dreyfus's criterion of "perspicuous grouping," these capabilities of deep neural nets seem impressive indeed. While no single one of the extracted features here represents a face as such, one may argue that "face-ness" – the abstraction that is the overall Gestalt of a face – is located in the totality of the latent space itself.

There are a few points that support this observation: First, as indicated, an ANN does not contain any explicit knowledge. Unlike in the case of knowledge base and inference engine, a neural network's "knowledge" is not localized in some particular place, but is distributed throughout the whole system as a statistical dependency.[47] ANNs thus generalize without forming concepts.[48] Secondly, a neural network does not follow the paradigm of logical deduction or explicitly stated rules that are executed sequentially; rather, it operates by statistical induction, and it is the system as a whole that does the computing.[49] Third, one could therefore argue that neural networks do not only produce outputs that humans perceive as Gestalten, but that, as statistical models, they internally already operate according to a Gestalt logic – producing neither a mere collage nor just a mean of existing images.

From all this, it appears that ANNs cannot plausibly be described as atomistic. However, they possess a number of characteristics that disqualify them from being considered truly holistic. First, the network, at least on the operational level, is still hierarchically structured: Even if the resulting latent space contains more or less abstract features, in the process of abstraction it nevertheless proceeds from parts to wholes, and not the other way around. Second, the fact that neural networks do not use symbolic representation does not mean that they do not use any representation at all; statistical models, too, represent. Third, the latent space can show any possible interpolation of a face. However, what is possible is dependent on the breadth of features present in the training set as well as on the selection of features, which may exclude some that could be relevant; dimensionality reduction is, after all, reductive. Thus, beyond the "face-ness" stored in the latent space, there would still be faces that humans would recognize as such which the model could not interpolate. And finally, one may argue that a central feature of the Gestalt concept is that it implies an understanding rather than just a recognition of the thing. Metzger held that a Gestalt says something "about what a face is really about in life," whether it is arrogant or friendly and so on; of course, this knowledge is not represented in the latent space of the model. At this point the distinction between perception and cognition becomes blurry again.[50] This either brings us back to Dreyfus's doubt about whether a strong AI is possible at all, or draws our attention to the fact that even with a weak AI, and even if it learns unsupervised, there is cognition at play that is encoded in the parameters set by humans – that no training is truly unsupervised.[51]

ANNs can therefore neither be conceptualized as completely atomistic nor

47 Pasquinelli, Machines that Morph Logic; Andreas Sudmann, Szenarien des Postdigitalen: Deep Learning als MedienRevolution, in: Engemann, Sudmann (ed.), *Machine Learning*, pp. 66–68.

48 If each layer provides an abstraction of the features of the previous layer, this is not yet conceptual work and does not result in a theory, Dreyfus and Dreyfus, Making a Mind versus Modeling the Brain, p. 36.

49 Pasquinelli points out that Rosenblatt himself in *Principles of Neurodynamics* already considered this whole as an emergent quality in the sense of Gestalt.

50 This is also pointed out in Smith, *Promise*, pp. 7, 24–27 and 56–7.

51 See Matteo Pasquinelli: How a Machine Learns and Fails – a Grammar of Error for Artificial Intelligence. *spheres* 5, no. 5 (2019), spheres-journal.org/how-a-machine-learns-and-fails-a-grammar-of-error-for-artificial-intelligence; access: June 5, 2020.

as completely holistic. They seem to be something in between. John von Neumann called this in-between a "mixed system." Von Neumann coined the term in 1958 when, in his posthumous book *The Computer and the Brain*, he discussed the differences and similarities between the titular nervous system and the digital automaton. Von Neumann believed that the brain transmits information digitally between synapses, and that its make-up was fundamentally one of discrete states. But while the computer, as an instance of Turing's Universal Machine, is serial and deterministic, the brain has a parallel structure and its operations are based on statistical states. The brain is therefore a "mixed system."[52] While neuroscientists today are cautious about calling the operations of the brain digital, von Neumann's "mixed system" describes ANNs rather well.

German media theorist Andreas Sudmann recently elaborated this conclusion in more detail. He emphasizes that neural networks are still based on the digital architecture named after von Neumann, which is based on analog structures. They also function in parallel rather than serially. While their "neurons," as logical gates, indeed operate discretely, the weights distributed in their "synapses" are not binary states but are rather stored as floating-point numbers. Because of this, Sudmann proposes to call ANNs neither completely digital nor completely analog, but "postdigital" or – more clearly, as I think – "quasi-analog."[53]

If "quasi-analog" denotes the technical structure of modern artificial neural networks, "quasi-holistic" would be the term to describe their conceptualization. Neural networks neither completely follow the atomistic paradigm, nor are they really holistic. They are, conceptually speaking, a quasi-holistic mixed system and combine properties of both paradigms. ANNs clearly show that the distinction between atomism and holism is too rigid to really capture this phenomenon adequately. This would require a third class.

One concept that comes into question for such a third option is that of "assemblage." Following Gilles Deleuze and Félix Guattari's notion of "agencement" – which has been translated into English as "assemblage" – Manuel DeLanda proposed this term to conceptualize something between the atomistic and the holistic approach. If "atomism" describes wholes as mechanical aggregates of isolated elements, "holism" denotes "relations of interiority" in which each element is in an organic constitutive relationship to the whole. In contrast to both, DeLanda interprets assemblages as characterized by "relations of exteriority."[54] An assemblage still forms a whole with properties that are not necessarily present in its el-

52 John von Neumann, *The Computer and the Brain* (New Haven 1958), pp. 58–60.

53 Sudmann, Szenarien des Postdigitalen, p. 66. The term "post-digital" for quasi-analog structures is an unhappy choice, I believe, since it already means so many other things that a further extension is not desirable, see Hannes Bajohr, Experimental Writing in its Moment of Digital Technization: Post-Digital Literature and Print-on-Demand Publishing, in: *Publishing as Artistic Practice*, ed. Annette Gilbert (Berlin 2016), pp. 100–115.

54 Manuel DeLanda, *A New Philosophy of Society: Assemblage Theory and Social Complexity* (London 2006), pp. 9–10.

ements – but the elements also retain a degree of autonomy for which the holistic view does not allow. For this reason, according to DeLanda, assemblages exhibit nonlinear causalities that are rather statistical than deterministic in nature. While his example is a chemical process like catalysis, the weight model of a neural network would be another case of nonlinear and statistical behavior.[55] Although one may not agree with DeLanda in everything, one must concede that the term assemblage is at least *one* candidate for describing a third option between holism and atomism that is already available. Thus understood, an artificial neural network is, as Deleuze and Guattari put it in *A Thousand Plateaus*, quite literally a "machine … to produce faces."[56]

5. Addendum: Reversing the Perspective – ANNs as "intuition pumps"

In my discussion of the Gestalt properties of artificial neural networks, I have tried to extract conceptual assumptions from a technical system, which I called quasi-holistic assemblages. However, I find it likewise possible and possibly productive to take the reverse route: The assemblage-like, quasi-analog, quasi-holistic view of ANNs can serve as an "intuition pump,"[57] as Daniel Dennett called it, to rethink some more traditional problems of non-atomistic concepts.

The face serves as the leitmotif of this essay, but in the citation quoted above, Hubert L. Dreyfus also assigned *style* to the domain of Gestalt recognition.[58] In fact, style – where it is not understood in a purely formalistic way – is often thought of as an irreducibly holistic phenomenon, which may be hermeneutically accessible but is opposed to the atomistic listing of features.[59] However, StyleGAN2 already brings ANNs and at least visual style together. It not only encodes a quasi-holistic statistical model of faces, but also makes it possible to

55 Ibid., 13–15.

56 Gilles Deleuze and Félix Guattari, *A Thousand Plateaus: Capitalism and Schizophrenia* (Minneapolis 1987), pp. 173.

57 Dennett, *Consciousness Explained*, p. 440.

58 Dreyfus, *What Computers Can't Do*, p. 32.

59 Style as a perceptual phenomenon with an irreducible subjective quality that resists quantification is described ironically but concisely by George Kubler: Style, he writes, is "like a rainbow. It is a phenomenon of perception governed by the coincidence of certain physical conditions. We can see it only briefly while we pause between the sun and the rain, and it vanishes when we go to the place where we thought we saw it." George Kubler, *The Shape of Time: Remarks on the History of Things* (New Haven 1962), p. 129. For a historical overview see Hans Ulrich Gumbrecht, Schwindende Stabilität der Wirklichkeit: Eine Geschichte des Stilbegriffs, in: *Stil: Geschichten und Funktionen eines kulturwissenschaftlichen Diskurselements*, ed. Hans Ulrich Gumbrecht and K. Ludwig Pfeiffer (Frankfurt/Main 1986), pp. 726–788. The semiotic theory of style emphasizes its constitutive expressiveness, which is rooted in life forms, see for example Dick Hebdige, *Subculture: The Meaning of Style* (London 1979). The difficulty of defining and cataloguing style is best illustrated in Susan Sontag, Notes on Camp, in: *Against Interpretation and Other Essays* (New York 1978), pp. 275–292.

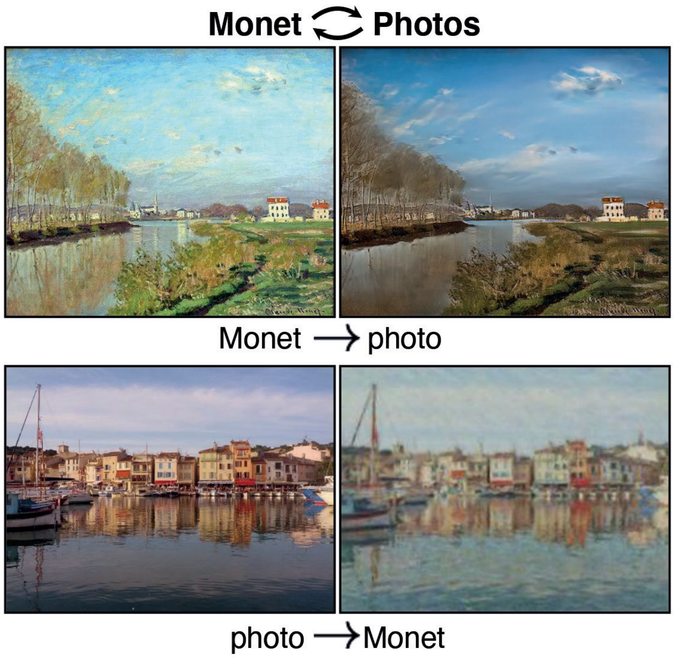

Fig. 7: Style transfer with CycleGAN; from a paper by Jun-Yan Zhu et al., 2018. https://arxiv.org/abs/1703.1059.

transfer a certain configuration of a latent space – a "style" – to another image by first extracting the specific feature distribution of the input and then mapping it to the feature vectors of the reference image. This method is also used by another GAN implementation called CycleGAN, which transposes styles of certain painters onto photos and vice versa (fig. 7).[60]

This process suggests that "style," understood as a Gestalt-like family resemblance, need not necessarily be a purely irreducible, holistic affair. This is not to say that style can be quantified entirely in an atomistic, empirical manner, as stylometry attempted to do in the 1970s and as the digital humanities do today, nor does it mean an easy separation of form and content in the sense of a mere *ornatus*. But to rethink style not as a truly holistic, but only as a quasi-holistic concept – as an assemblage with *some* moving parts that to *some extent* can be pinned down – could make it possible to reassess this concept so often treated as suspect.[61]

60 Jun-Yan Zhu et al., Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. *ArXiv*, March 30, 2017, arxiv.org/abs/1703.10593; access: June 6, 2020. Corresponding attempts for textual style transfer are less powerful, see Xiangyang Li et al., Review of Text Style Transfer Based on Deep Learning. *ArXiv*, May 6, 2020, arxiv.org/abs/2005.02914; access: June 6, 2020, although the more advanced approaches in GPT-2 and, although not open-source, GPT-3 suggest that great strides are possible here.

61 It would be possible, for example, to understand Ernst Gombrich's answer to "the riddle of style," the "schema," as a quasi-holistic assemblage. Gombrich, too, argues nonatomistically, since the knowledge of schemata is still restricted to humans, i.e. to systems that already process Gestalten, see Ernst H. Gombrich: *Art and Illusion* (London 1961), pp. 3, 60. I thank Jana Mangold

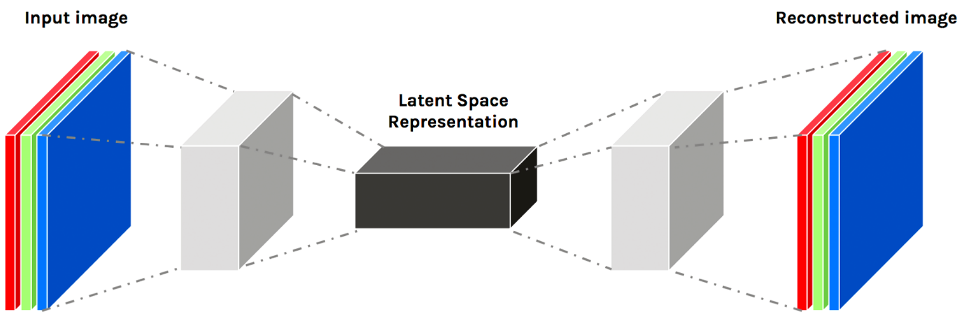

Fig. 8: Symbolic depiction of an autoencoder; from Grietzer, A Theory of Vibe, https://www.glass-bead.org/article/a-theory-of-vibe

Another of those seemingly irreducible holistic terms is "mood."[62] The fact that it also plays a role in Heidegger's existential analysis and is thus firmly anchored in the holistic tradition makes this case particularly interesting. In his essay "A Theory of Vibe," the literary scholar and mathematician Peli Grietzer has approached ANNs as conceptualization models. Grietzer uses a specific neural network, the autoencoder (fig. 8), as an intuition pump for the holistic concept of mood – or, as he prefers to call it, "vibe."[63] An autoencoder is a very simple type of neural net that is trained to exactly output its inputs. However, since its hidden layers are "smaller" than the outer layers, and the data passes through a bottleneck of maximum dimensional reduction, the auto-encoder compresses the characteristics of the input to an extreme degree before restoring them.

If one now imagines, according to Grietzer's thought experiment, that a set of aesthetic objects is compressed in this way so that it can be reproduced without error, the compression model would simply consist of a list of possible variations of its general features. This, as a collective, Grietzer describes as "vibe." The vibe is present in all aesthetic objects of this set, but never in its pure form – while they are complex on their own, they are collectively simple. Although we can imagine the vibe in this way through the metaphor of the autoencoder, it is never encountered on its own: "A vibe is … an abstractum that cannot be separated from its concreta." This is, he argues, a reversal of Goethe's symbol understood as concretum that expresses an abstraction. For Grietzer, vibe describes a particularly modernist, materialist quality of aesthetic works: instead of representing something abstract concretely (symbolism), the "canon" of a modernist work discloses a vibe and represents abstrac-

for the reference. For a practical application of a "neural reading" of poetry and its style, see Boris Orekhov and Frank Fischer, Neural Reading: Insights from the Analysis of Poetry Generated by Artificial Neural Networks. *Orbis Litterarum* 75, no. 5 (2020), pp. 230–46.

62 See Hans Ulrich Gumbrecht, *Atmosphere, Mood, Stimmung: On a Hidden Potential of Literature* (Stanford 2012).

63 Peli Grietzer, A Theory of Vibe. *Glass Bead* 1, no. 1 (2017), www.glass-bead.org/article/a-theory-of-vibe; access: June 6, 2020.

tion through the repetition of structural similarities present in all of the canon's works. Conversely, grasping a real-world vibe through its idealization as the vibe of a literary work – an obvious example is the Kafkaesque – is itself a type of mapping that takes place in an autoencoder; the reader, too, encodes.[64]

Like the double face of the Gioconda with which this text began, style and mood are examples of how phenomena usually understood in a holistic way can also be thought of as quasi-holistic assemblages. Even though I could only give a sketch of this idea here, these forms point to a whole host of areas in which the strange third position of ANNs – beyond the dichotomy of holism and atomism – not only denotes a technology, but can be a conceptual tool in its own right.

64 Grietzer, "A Theory of Vibe." See in more detail: Peli Grietzer, *Ambient Meaning: Mood, Vibe, System*, PhD Diss., Harvard University, 2017.

References

Alpaydin, Ethem, *Machine Learning: The New AI* (Cambridge, MA: MIT Press, 2016), pp. 85–110.

Ash, Mitchell G., *Gestalt Psychology in German Culture, 1890–1967: Holism and the Quest for Objectivity* (Cambridge: Cambridge University Press, 1995).

Bailey, Suzanne, Francis Galton's Face Project: Morphing the Victorian Human. *Photography and Culture* 5, no. 2 (2012), pp. 189–214.

Bajohr, Hannes, Algorithmic Empathy: On Two Paradigms of Digital Generative Literature and the Need for a Critique of AI Works. *BMCCT working papers* 1, no. 4 (2020); https://doi.org/10.5451/unibas-ep79106.

Bajohr, H., Die "Gestalt" der KI: Jenseits von Atomismus und Holismus. *Zeitschrift für Medienwissenschaft* 23, no. 2 (2020), pp. 168–181.

Bajohr, H., Experimental Writing in its Moment of Digital Technization: Post-Digital Literature and Print-on-Demand Publishing, in: *Publishing as Artistic Practice*, ed. Annette Gilbert (Berlin: Sternberg, 2016), pp. 100–115.

Bates, David, Creating Insight: Gestalt Theory and the Early Computer, in: *Genesis Redux. Essays in the History and Philosophy of Artificial Life*, ed. Jessica Riskin (Chicago: University of Chicago Press, 2007), pp. 237–260.

Belting, Hans, *Face and Mask: A Double History*, trans. Thomas S. Hansen and Abby J. Hansen (Princeton: Princeton University Press, 2017).

Blumenberg, Hans, Prospect for a Theory of Nonconceptuality, in: *History, Metaphors, Fables: A Hans Blumenberg Reader*, ed. Hannes Bajohr, Florian Fuchs and Joe Paul Kroll (Ithaca, NY: Cornell University Press, 2020).

DeLanda, Manuel, *A New Philosophy of Society: Assemblage Theory and Social Complexity* (London: Continuum, 2006).

Deleuze, Gilles, and Félix Guattari, *A Thousand Plateaus: Capitalism and Schizophrenia* (Minneapolis: University of Minnesota Press, 1987).

Dennett, Daniel C., *Consciousness Explained* (New York: Little, Brown, and Co., 1991).

Domingos, Pedro, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World* (New York: Penguin, 2015).

Dreyfus, Hubert L., and Stuart E. Dreyfus, Making a Mind versus Modeling the Brain: Artificial Intelligence Back at a Branchpoint. *Daedalus* 117, no. 1 (1988), pp. 15–43.

Dreyfus, H. L., *What Computers Can't Do: A Critique of Artificial Reaso*n (New York: Harper & Row, 1972).

von Ehrenfels, Christian, On "Gestalt Qualities" [1890], in: *Foundations of Gestalt Theory*, ed. Barry Smith (Munich: Philosophia, 1988), pp. 82–116.

Engemann, Christoph, and Andreas Sudmann (eds.), *Machine Learning: Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz* (Bielefeld: Transcript, 2018)

Esfeld, Michael, Holismus und Atomismus in den Geistes- und Naturwissenschaften: Eine Skizze, in: *Holismus und Individualismus in den Wissenschaften*, ed. Alexander Bergs and Soelve I. Curdts (Frankfurt/Main: Lang, 2003), pp. 7–21

Fjelland, Ragnar, Why General Artificial Intelligence Will Not Be Realized. *Humanities and Social Sciences Communications* 7, no. 10 (2020); https://doi.org/10.1057/s41599-020-0494-4.

Gombrich, Ernst H., *Art and Illusion* (London: Phaidon, 1961).

Goodfellow, Ian et al., Generative Adversarial Networks. *ArXiv*, June 6, 2014; arxiv.org/abs/1406.2661, access: June 6, 2020.

Grietzer, Peli, A Theory of Vibe. *Glass Bead* 1, no. 1 (2017), www.glass-bead.org/article/a-theory-of-vibe; access: June 6, 2020.

Grietzer, P., *Ambient Meaning: Mood, Vibe, System*, PhD Diss., Harvard University, 2017.

Gumbrecht, Hans Ulrich, *Atmosphere, Mood, Stimmung: On a Hidden Potential of Literature* (Stanford: Stanford University Press, 2012).

Gumbrecht, H. U., Schwindende Stabilität der Wirklichkeit: Eine Geschichte des Stilbegriffs, in: *Stil: Geschichten und Funktionen eines kulturwissenschaftlichen Diskurselements*, ed. Hans Ulrich Gumbrecht and K. Ludwig Pfeiffer (Frankfurt/Main: Suhrkamp, 1986), pp. 726–788.

Hebdige, Dick, *Subculture: The Meaning of Style* (London: Routledge, 1979).

Heims, Steve Joshua, *The Cybernetics Group 1946–1953: Constructing a Social Science for Postwar America* (Cambridge, MA: MIT Press, 1991).

Isola, Phillip et al., Image-to-Image Translation with Conditional Adversarial Networks. *ArXiv*, November 21, 2016; arxiv.org/abs/1611.07004, access: June 6, 2020.

Jeffress, Lloyd A. (ed.), *Cerebral Mechanisms in Behavior: The Hixon Symposium* (New York: Hafner, 1951).

Kant, Immanuel, *Prolegomena to Any Future Metaphysics*, trans. Gary Hatfield (Cambridge: Cambridge University Press, 2004).

Kenaw, Setargew, Hubert L. Dreyfus's Critique of Classical AI and its Rationalist Assumptions. *Minds and Machines* 18, no. 2 (2008), pp. 227–238.

Kline, Ronald R., Cybernetics, Automata Studies, and the Dartmouth Conference on Artificial Intelligence. *IEEE Annals of the History of Computing* 33, no. 4 (2011), pp. 5–16.

Klingemann, Mario, https://twitter.com/quasimondo/status/934546438507376640, access: June 6, 2020.

Köhler, Wolfgang, *Mentality of Apes* (New York: Harcourt, Brace, 1927).

Köhler, W., Relational Determination in Perception, in: *Cerebral Mechanisms in Behavior: The Hixon Symposium*, ed. Lloyd A. Jeffress (New York: Hafner, 1951). pp. 200–243.

Köhler, W., review of *Cybernetics, or Control and Communication in the Animal and the Machine*, by Norbert Wiener. *Social Research* 18, no. 1 (1951), pp. 125–130.

Köhler, W., Bemerkungen zur Gestalttheorie. *Psychologische Forschung* 11, no. 1 (1928), pp. 188–189.

Kubler, George, *The Shape of Time: Remarks on the History of Things* (New Haven: Yale University Press, 1962).

Landau, Terry, *About Faces* (New York: Doubleday, 1989).

LeCun, Yann, Yoshua Bengio and Geoffrey Hinton, Deep Learning. *Nature* 521 (May 2015), pp. 436–444.

Lévinas, Emmanuel, *Totality and Infinity: An Essay on Exteriority*, trans. Alphonso Lingis (Dordrecht: Kluwer, 1991).

Li, Xiangyang et al., Review of Text Style Transfer Based on Deep Learning. *ArXiv*, May 6, 2020, arxiv.org/abs/2005.02914; access: June 6, 2020.

Luxemburg, Robert, StyleGAN2 Interpolation Loop, December 12, 2019; youtube.com/watch?v=6E1_dgYlifc, access: June 5, 2020.

Macho, Thomas, *Vorbilder* (Munich: Fink, 2011).

McCarthy et al., John, A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. *AI Magazine* 27, no. 4, 2006, pp. 12–14.

McCorduck, Pamela, *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence* (Natick: A K Peters/CRC Press, 2004).

Merleau-Ponty, Maurice, *Phenomenology of Perception* (London: Routledge, 2005).

Metzger, Wolfgang, *Laws of Seeing* [1936], trans. Lothar Spillmann (Cambridge, MA: MIT Press, 2006).

Metzger, W., Was ist Gestalttheorie? in: *Gestalttheorie und Erziehung*, ed. Kurt Guss (Darmstadt: Steinkopff, 1975), pp. 1–17.

Minsky, Marvin L., and Seymour Papert, *Perceptrons: An Introduction to Computational Geometry* (Cambridge, MA: MIT Press, 1969).

Mitchell, Melanie, *Artificial Intelligence: A Guide for Thinking Humans* (New York:

Farrar, Straus and Giroux, 2019).

von Neumann, John, *The Computer and the Brain* (New Haven: Yale University Press, 1958).

Newell, Allen, J. C. Shaw and H. A. Simon, Report on a General Problem-Solving Progra, in: *Proceedings of the International Conference on Information Processing* (Paris: UNESCO, 1959), pp. 256–264.

Nilsson, Nils J., *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge: Cambridge University Press, 2010).

Orekhov, Boris and Frank Fischer, Neural Reading: Insights from the Analysis of Poetry Generated by Artificial Neural Networks. *Orbis Litterarum* 75, no. 5 (2020), pp. 230–46.

Pasquinelli, Matteo, Machines that Morph Logic: Neural Networks and the Distorted Automation of Intelligence as Statistical Inference. *Glass Bead* 1, no. 1 (2017); www.glass-bead.org/article/machines-that-morph-logic, access: June 6, 2020.

Pasquinelli, M., How a Machine Learns and Fails – a Grammar of Error for Artificial Intelligence. *spheres* 5, no. 5 (2019), spheres-journal.org/how-a-machine-learns-and-fails-a-grammar-of-error-for-artificial-intelligence; access: June 5, 2020.

Pias, Claus (ed.), *Cybernetics: The Macy-Conferences 1946–1953* (Zurich: Diaphanes, 2003).

Rosenblatt, Frank, The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review* 65, no. 6 (1958), pp. 386–408.

Rumelhart, David E. , and James L. McClelland, PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Cambridge, MA: MIT Press, 1986).

Ryle, Gilbert, Knowing How and Knowing That. *Proceedings of the Aristotelian Society* 46, No. 1 (1946), pp. 1–16.

Searle, John R., Minds, Brains, and Programs. *Behavioral and Brain Sciences* 3, no. 3 (1980), pp. 417–457.

Seifert, Uwe, Randolph Eichert and Lüder Schmidt, Logic, Gestalt Theory, and Neural Computation in Research on Auditory Perceptual Organization, in: *Music, Gestalt, and Computing: Studies in Cognitive and Systematic Musicology*, ed. Marc Leman (Berlin: Springer, 1997), pp. 70–88.

Sellars, Wilfrid, *Empiricism and the Philosophy of Mind* (Cambridge, MA: Harvard University Press, 1997).

Simmel, Georg, The Aesthetic Significance of the Face, in: *Essays on Art and Aesthetics*, ed. Austin Harrington (Chicago: The University of Chicago Press 2020), pp. 231–235.

Smith, Brian Cantwell, *The Promise of Artificial Intelligence: Reckoning and Judgment* (Cambridge, MA: MIT Press, 2019).

Sontag, Susan, Notes on Camp, in: *Against Interpretation and Other Essays* (New York: Octagon, 1978), pp. 275–292.

Sudmann, Andreas, Szenarien des Postdigitalen: Deep Learning als Medien-Revolution, in: *Machine Learning: Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz*, ed. Christoph Engemann and Andreas Sudmann (Bielefeld: Transcript, 2018), pp. 66–68.

Taylor, Charles, Heidegger on Language, in: *A Companion to Heidegger*, ed. Hubert L. Dreyfus and Mark A. Wrathall (London: Blackwell, 2007), pp. 435–436.

Taylor, C., *The Language Animal: The Full Shape of the Human Linguistic Capacity* (Cambridge, MA: Cambridge University Press, 2016).

Taylor, C., Theories of Meaning, in: *Human Agency and Language: Philosophical Papers I* (Cambridge: Cambridge University Press, 1985), pp. 247–292.

Toepfer, Georg, Ganzheit, in: *Historisches Wörterbuch der Biologie* (Stuttgart: Metzler, 2011), pp. 693–728.

Weigel, Sigrid (ed.), *Gesichter: Kulturgeschichtliche Szenen aus der Arbeit am Bildnis des Menschen* (Paderborn: Fink, 2013)

Weigel, S., Der konventionelle Code als buckliger Zwerg im Dienste der Emotion Recognition. Überlegungen zu einer Urgeschichte der digitalen Kultur, in: *Internationales Jahrbuch für Medienphilosophie* 6, no. 1 (2020), pp. 47–79.

Wertheimer, Max, Untersuchungen zur Lehre von der Gestalt II. *Psychologische Forschung* 4, no. 1 (1923), pp. 301–350.

Zhu, Jun-Yan et al., Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. *ArXiv*, March 30, 2017, arxiv.org/abs/1703.10593; access: June 6, 2020.