

“Who inspires who?” Aesthetics in front of AI art

Alice Barale

Abstract: AI art certainly belongs among the most under-researched art forms of today. While the ethical aspects of AI are often discussed, its implications for aesthetics are rarely considered. The reason is perhaps that the ability of AI to produce art is a very recent development. Artificial intelligence is becoming more and more important nowadays due to the many changes it has made in our lives. In this context, one of the most surprising fields in which AI has suddenly progressed in the last few years concerns the very human (until now) capacity for artistic (and in a more general sense, aesthetic) expression. However, the general public still tends to confuse AI art with the more general category of digital art, and what AI really does in the artistic field is scarcely known. AI art is not only computer assisted but computer generated. In AI art there is at least one part of the artistic process that is left to the machine. The artist gives the AI some data and has to wait, in order to see how the AI will elaborate upon them. AI, therefore, becomes not just a *tool* for artists, but also something different, the nature of which needs to be explored. What does AI art have to say about our way of seeing art, and perhaps about our way of seeing the world in general? This will be the leading question of this paper, which will be addressed through the analysis of some relevant aspects of this new kind of art.

Keywords: AI; AI art; neural networks; aesthetics.

0. *The philosophers in the face of the unknown*

“Who inspires who?” This question was raised in the presentation of Mario Klingemann’s, Albert Barqué-Duran’s and Mark Merzeneit’s recent work *My Artificial Muse*, a three day- performance in which some stick figures are translated into digital images by an artificial intelligence, and then painted on a large canvas by a painter, who in turn generates different sounds with his movements. Works of art created by artificial intelligence are a very recent phenomenon and the philosophical and aesthetic debate is still very unprepared for them. Although artificial intelligence is now a constant presence in

our lives – AIs are driving our underground trains, assisting us every day on our phones and laptops, helping with medicine and the economy – the discussion until now has been mainly concerned with the ethical aspects of AI, while the sphere of aesthetics remains scarcely explored.

However, AI art is strictly connected to aesthetics, intended as the sphere of our sensory experience of the world. As it will become apparent shortly, in AI art, the machine has to “learn to see” the world and share its vision with the human artist. Non-experts still tend to confuse AI art with the more general category of digital art. Yet AI art is not only “computer assisted”, but also “computer generated” (Boden 2012). This means that at least one part of the artistic process is left to the machine: the artist gives the AI some data and waits, in order to see how it will elaborate them. Then he can intervene in his turn on the outcomes, through an interplay between man and machine that is virtually infinite. In this process, the machine is no longer merely a simple *tool*, but becomes something different, the nature of which still needs to be investigated.

What does AI art have to say about our way of seeing art, and perhaps about our way of seeing the world in general? This will be the leading question of this paper, which will be addressed through the analysis of some relevant aspects of this new kind of art.

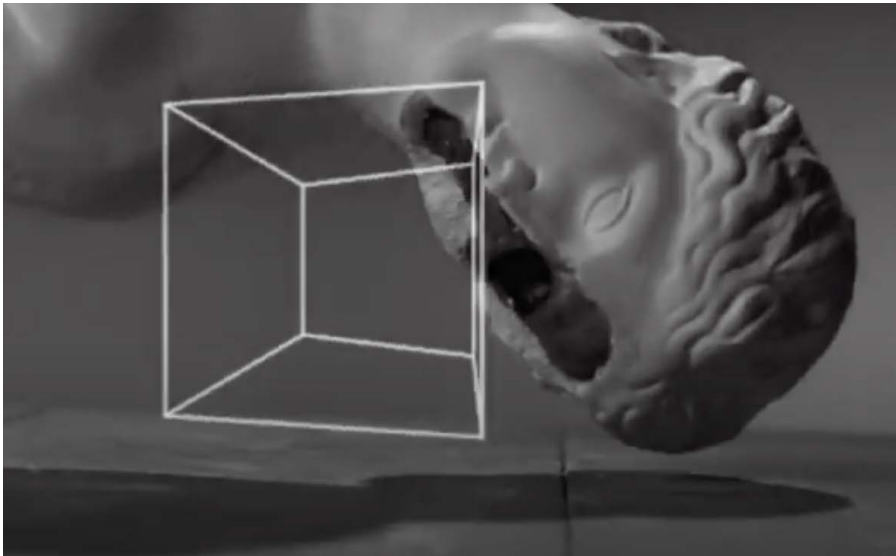


Fig. 1. *My Artificial Muse*, shot from the performance video. Courtesy of Onkaos.

1. *The beginning of a discussion*

The new problems raised by AI art became apparent two years ago when, for the first time, the famous auction house Christies' sold an artwork generated by AI. In October 2018 the work entitled *Le Comte de Belamy* was sold for 432 500 dollars, nearly 45 times the estimated value. In certain respects it resembles a classical portrait of a gentleman elegantly dressed in black, but it is much more blurred: outlines are not well-defined and as a whole, it has a kind of mysterious aspect. There is also a signature, in the lower part of the painting, which looks more like a mathematical formula than a name. In fact, it is the key-formula of the algorithm through which the work was generated. It will be necessary to return to this point. For the moment, what is interesting are the reactions to this event in the contemporary art world. A lively discussion arose surrounding the possibility for machines to make art. Can computers be creative? This is the main question that has been raised. In a recent book, the philosopher of science Arthur L. Miller has interviewed a number of artists working with AI, each of whom gave very different answers (Miller 2019; on AI and art see also Gouveia 2020, section III; Du Sautoy 2019; on AI and aesthetics see Marfia, Matteucci 2018). Miller has his own idea: machines can indeed be creative. In this paper, however, we are going to approach the question differently. When faced with this question – “can machines be creative?” – it may be useful to make what can be called a “Benjaminian move”.

In *The Work of Art in the Age of Mechanical Reproduction*, the German philosopher Walter Benjamin warns us against a certain use of ideas such as “creativity”, “genius” or “mystery” (Benjamin 2008). The reason is that these concepts, if used in an a-critical way, can cause the work of art and the artist himself to be considered as idols, which always remain the same and cannot be approached and changed. This is why this static approach to art and creativity was adopted by fascism, to “aestheticize politics”. Technical reproducibility deconstructs this paradigm of creativity, according to Benjamin, because it causes a decline in the “aura” of the artwork, i.e. of its idolatrous and unrepeatable character. The idea of this paper is that AI art takes this process one step forward: it not only deconstructs our idea of an “absolute” creator, but it demonstrates that creativity always arises from the relationship with something that is radically “other”. In this case, from the interplay between man and machine. We will see that the question “can AI be creative?” will transform itself into another question: what can AI say about what is usually called “creativity” and its role in art, and perhaps also in everyday perceptions of the world? This question will be explored through some examples taken from the figurative and narrative arts, although some references to music will be also provided.

2. *Entering the machine's dream*

The “history” of AI art is very recent, beginning in 2015. Before that date, art experiments with computers had been done, as in the case of the computer-artist AARON created by Harold Cohen (a project to which Simon Colton has returned recently with his *The Painting Fool*. See Colton 2012; 2019; Cohen *et al.* 2012: 98-110). In all these cases, however, the machine’s autonomy was very limited. The computer was merely following a set of rules given to it by the programmer, in order to assemble in new ways the data provided to it. The creation process, therefore, was completely guided by the human artist and programmer.

However, in 2015 something unexpected emerged. In that year, deep neural networks were in their renaissance: they had been used already in the 1990s, but technology at that time was not advanced enough, and the best that AIs could do was to recognize handwritten numerical digits, for example, those written on cheques (see Castelle 2020). Yet from 2010 on, more or less, much more powerful neural networks were created, which could help with medical diagnoses and researches, drive cars, play sophisticated games like *Go*, translate complex texts... How was all this possible? The first form of AI art arises from precisely this question: from the attempt to reach the underlying stratus of deep neural networks, see how they work and “reason”.

Deep neural networks are similar to our own brains in the way they are constructed. They are constituted by many layers of artificial neurons (that is why they are called “deep” neural networks). Each layer takes the data from the previous layer and elaborates on that set in a more and more complex way. For example, in the case of image recognition, after training the first layer may learn to react to lines and edges, the second layer to the shape of a nose or mouth, the third to a face, and so on (see Goodfellow *et al.*, 2017). In the end, the computer emerges with the probabilities of what the starting image actually is. For example, if the starting image was a dog, the result could be: 90% dog, 2% bird, 3% cow, 5% cat. According to how accurate the result is, the programmer adjusts the modifiable parameters of the machine until it learns to identify the image more correctly. The results achieved by these neural architectures are astonishing, but they also present a problem. The issue is that we can only chose the inputs and see the outputs, but nobody knows what goes on inside the other layers. For this reason, neural networks have been compared to a black box and some consider them to be scarcely controllable.

For Alex Mordvintsev, a computer scientist working at Google in Zurich, these hidden layers were precisely what had to be investigated (Mordvintsev *et al.* 2015). In order to do this, Mordvintsev tried to visualize what was go-

ing on in the deepest layers of the net. He stopped the work of the network half-way and asked the machine to generate what it “saw”. As has been noted above, the AI gives humans some scores for the objects it was trained on that it believes might exist in the input image. Mordvintsev asked the AI to modify the input image to make these objects more and more visible. The result was astonishing: since the AI had been trained on different animals, it produced a strange creature made of dogs, cats and birds, with different eyes and heads all around its body (see *Father Cat*: <https://www.artnome.com/news/2018/12/30/deepdream-creator-unveils-very-first-images-after-three-years>). It was a very unexpected representation of how the machine “saw” the world.

Seeing these first images, another Google engineer, Myke Tyka, had the idea of improving Mordvintsev’s technique by zooming in on the images. The result was even more surprising because it revealed a series of fractal structures, one inside another: a world made of clouds that become dogs that become palaces that become cars (fig. 2), “a strange conglomeration of animals, reptiles, lizards, cars, bicycles, stretching to infinity” (quot. in Miller 2019: 70). Tyka and Mordvintsev called this new artistic technique “Deep Dream”, because it allowed the researchers to get into the “dreams” of the machine. Deep Dream met with great success in the following months and years, and became a focal point for different artists who were interested in working with AI. A group called AMI (Artist and Machine Intelligence) was formed at Google in Seattle and they organized in 2016 in San Francisco an exhibition called *Deep Dream: the Art of Neural Networks*. Its participants (like Mike Tyka but also Memo Akten, Mario Klingemann, Anna Ridler) are still among the main protagonists of AI art (for a review of the main current AI artists see the site: <https://aiartists.org/ai-artist-founding-members>). On this occasion, the leader of AMI, Blaise Agüera y Arcas, talked about AI art as a new avant-garde, bound to transform our very idea of art:

Like the invention of applied pigments, the printing press, photography, and computers, we believe machine intelligence is an innovation that will profoundly affect art. As with these earlier innovations, it will ultimately transform society in ways that are hard to imagine from today’s vantage point; in the nearer term, it will expand our understanding of both external reality and our perceptual and cognitive processes (Agüera y Arcas 2016).

In his talk, Agüera y Arcas compared AI to photography, which for a long time was also considered to be something extraneous to art that even presented a danger to its survival (see Mazzocut-Mis 2020). This comparison is still very present in the actual debate on AI art (see for ex. Obvious 2020) and

can indeed be very useful to understand its philosophical implications. Like photography, AI doesn't destroy art but introduces a change in it, a change that philosophers of aesthetics have an urgent need to investigate.

In what way, then, can AI "profoundly affect art"? AI is not the first art form, of course, to put into question the "classical" idea of artwork. The Twentieth Century Avant-gardes already radically problematized the concept of artwork itself, together with other ideas such as "genius", "author" and "beauty". The philosophical theories of art tried to address this challenge, searching for new definitions of art itself. One of these, as Stephen Davies has explained (Davies 1991, 2007: 1-145), is the so-called "procedural theory" of art: a type of theory that doesn't merely identify art with what is able to give aesthetic pleasure, but with everything that belongs to a social practice called "art". The most famous example of this type of theory is George Dickie's "institutional theory" of art (Dickie 1974, 1984). Davies considers procedural theories of art more appropriate for defining the "mature" art of our Western world, while aesthetic and perceptive properties are necessary to individuate the "first" forms of art, i.e. artworks that were "created outside the context of an established art tradition" (Davies 2015: 382). In my opinion, the consideration of AI art does not necessarily imply that one must choose or favour a procedural theory of art. It is useful, nevertheless, to briefly examine the different definitions of art that have been given, in order to answer to a first question: is AI art really art? As noted above, this question has been debated with particular urgency ever since AI art entered the art market. The answer to this question, however, is necessary only in order to prepare a more fundamental question, as I will try to show.

According to Davies, the philosophical definitions of art can be classified into three main types, which include functional, procedural and historical theories:

I classify the main contenders as belonging to three types: functional, procedural, and historical (Davies 1991). According to the functionalist, something is art if it is made to serve the point or function that is distinctive of art. Usually, this function is said to be that of producing an aesthetic experience (Beardsley 1983, Lind 1992, Anderson 2000) or of generating aesthetic properties (Zangwill 1995b). By contrast, the proceduralist maintains that something becomes art as a result of undergoing various social processes, just as someone becomes married as a consequence of undergoing a particular social ceremony. The most developed and discussed version of a procedural definition is George Dickie's institutional theory (Dickie 1974; 1984). Moving now to the third category, historical definitions insist that something qualifies as art by standing in an appropriate relation to earlier artworks. (Davies 2007: 36-37)



Fig. 2. Mike Tyka, *The Babylon Of The Blue Sun*, 2016. Courtesy of the artist

In contrast to the functional theories of art, which define artworks according to the aesthetic experience that they provide, the procedural theories state that an object is art if it is “baptized” as such by someone who has the “authority” to confer this status. Procedural theories of art don’t have any problem, then, with those artworks that have apparently lost their “auratic” status, such as the famous ready-mades by Duchamp, or Andy Warhol’s *Brillo Boxes* (Davies 2007: 11; for a different interpretation of this problem see Danto 1984, 1986, 1987; on Danto see Davies 2007: 119-127). The crisis that is also implied by AI art – a crisis of beauty, because its images cannot be defined as “beautiful” in a classical sense, but also a crisis of authorship, since it is unclear whether these pictures were made by a human or by an algorithm – could be addressed, therefore, by this type of theory. According to the procedural theory, AI art would be art because it has been “baptized” as such, and this happened for example when *Le comte de Belamy* was sold at Christies’. But is it really necessary to adopt this theory in order to consider AI art as art?

Let’s consider the first theory that Davies mentions, the functional theory of art. It is clear that AI artworks are able to exercise a particular charm on us. AI artworks are able to give, without any doubt, a strong aesthetic pleasure, although it is a pleasure that is not connected to “beauty” in its classical sense. The AI artist Mario Klingemann speaks, quoting André Breton, of “convulsive beauty” (Barale 2020: 78). In many cases, as it will be demonstrated shortly, we can’t stop looking at these pictures; we are irresistibly attracted by them. This also falls in line with the functional theory of art and, therefore, AI artworks seem to be art indeed. What about the third theory of art, the historical one? Here things seem to be even easier, because AI art could be considered the natural “evolution” of the technically reproducible art. Mario Klingemann has invented a neologism, “neurography” (*ivi*: 14, 237), to indicate the continuity, although with great differences, between photography and the “reproductions” of reality that neural networks can give. According to all three theories of art, then, the answer is yes: AI art is art. Of course there is a question as to whether it is necessary to choose just one of the three types of theories. This is a question that goes beyond the scope of this paper. It could be argued that each one of the three theories indicates an important aspect of art. This wouldn’t necessarily mean embracing a “cluster theory”, as Davies calls the theories that take something from all three perspectives. Rather, the three aspects could be considered as being deeply intertwined (for a deeper investigation of this point, see Desideri 2004, part II).

Anyway, once we have answered “yes” to the first question, “is AI art art?”, a second question arises, which I consider to be even more important. *What does AI art have to say about our way of seeing art, and perhaps about our way of*

seeing the world in general? A key to answering this question is in the particular process that leads to the creation of AI artworks. Let's consider for the moment the images of *Deep Dream* that were just described. These images are, in some ways, a surprise for the artist himself. The machine produces an *unexpected* result. What does the machine see in the pictures that have been given to it as inputs? It sees, as it has been shown, many different things, some of which are clearly present, while others are a more remote possibility, the AI represents them on a much smaller scale and hides them inside the other shapes. But what matters the most is that the resulting image is made of all these different, although related, creatures that the machine was able to see, and no one can predict what the outcome will be. It is important to note that this is not simply randomness. The machine has a logic in the way it elaborates upon information, even if it is a different logic from ours. The artist has to understand more and more how the AI “reasons” and change his behavior accordingly. More than randomness, then, it is the interplay between the artist and the machine that becomes important. From this point of view, we are close, in some respects, to the different forms of interactive art that have been experimented with “since the beginning of the twentieth century” (Kwastek 2013: XV), but there is also something more.

In these works, as noted above, the artist has to leave a part of the creative process to the machine, which elaborates upon the information on its own. This does not mean that the artwork is simply unfinished, open to interaction or change. It means something more radical: that the *formation* of the object itself happens through its relationship to an “other”. Prior to the relationship to the “other”, there is *no object*. As I will try to show in the next paragraphs, the AI has to “learn to see” the world. It has to recognize the everyday objects, and to represent them. In so doing, it makes mistakes and introduces deformations, which are able to show us a number of aspects that we might otherwise miss in our representations of the world around us.

Our representation of the world is always the result of a series of filters that we apply to it. The filters of language, through which we become conscious of experience and we can express it (see on this Desideri 2004: 9), and the filters of perception: already at this level, we don't have a “summation of temporally punctate sensations”, but “rather a process which filters and analyses incoming data in a search for indications of significant external occurrences and states of affairs” (Matthen 2015: Introduction). What makes AI art so interesting is that the way in which the experience of the world is filtered and elaborated is different from ours, but not so different so as to be meaningless to us. If we observe the way in which the AI elaborates and depicts its data, we can learn something about our own representations of reality, and also discover new and alternative

ways of organizing and recognizing the world. The artist Anna Ridler showed this well when she asked the AI to draw some scenes from the movie “Fall of the house of Usher”, starting from her own drawings (Ridler 2020). The result was a series of deformations and mistakes that have disclosed to the artist new aspects of her own way of drawing. For example, the machine confused the eyes and the eyebrows of the figures, because they were drawn in a very similar way. In some pictures, there was a table that appeared and disappeared, because the artist herself sometimes had forgotten to draw it. When faced with these results, Anna went back to her own pictures, and modified her own way of drawing. She wrote that the AI was like “a mirror” for her, through which she could become conscious of what aspects of the depicted scenes had been more important for her and what aspects, on the contrary, she had a tendency to omit (*ivi*:119).

The process of vision and representation of the AI, therefore, takes us back to our own perception of the world: to the world as it appears to us, in its infinite possibilities of form. When we face the vision of the AI, our own perception and representation of the world has to take place again from the beginning. Many AI artworks have titles like “Hello World”, or “Learning to see” (see Miller 2019). This shows, in a new and more radical way, something that has always belonged to art and, more generally, to the way we perceive and represent the world. Every perception is a deformation, every attempt to remember is a misremembering (see on this Ridler 2020). AI art demonstrates the effort of doing justice to what we perceive and the necessity of going back to what we have forgotten and perhaps missed along the way.

This also has a consequence on the way we see art. The Avant-gardes had criticized the ideas of “artwork”, “author”, “genius” etc. AI demonstrates something different: it shows that the author himself is never “one”, but always two. Every identity is put into question from the beginning and constitutes itself only in its relationship with what exceeds itself (on the ontological aspects of this question see Diodato 2021). And with AI, this exceeding element takes the form of a quasi-living entity, a quasi-creature that can emerge with its own picture of the world. This will become clearer if we consider another type of AI art that came after *Deep Dream*, the art of GANs.

3. *GANs and the different “epochs” of the image*

One year after the exhibition in San Francisco, in 2017, a symposium was organized by Agüera’s group, AMI, together with another team, Magenta, which was working not only with AI images but also with AI music. In the field of music, the use of deep neural networks is still in its infancy. François Pachet,

one of the main pioneers of AI music, has made computers improvise music and compose songs since the beginning of the century. In 2003, he created the Continuator, a system that can respond to the music played by a human musician through variations that maintain the same style (Pachet 2003). A few years later, he developed the Flow machine, which is not only able to improvise, but also to compose music (Pachet 2011; 2015). In 2016, Flow Machine recorded its first song, *Daddy's Car*, which was heard by more than 1,5 million listeners (see Minazzi 2020). In both cases, the technology Pachet used was that of Markov models that predict which note will most probably follow the previous one. Pachet has always been quite cautious in relation to the use of neural networks for music creation. However, experiments in this direction began in the second decade of the century. In 2016 the team Magenta, led by Douglas Eck, obtained the first melody composed using neural networks. This means that the computer has not been programmed in order to make this song, but has learned by itself how to compose, just by being fed with a certain quantity of musical data. The result, according to the Magenta team itself, had many defects, but what mattered was the process that had begun. This leads us back to Deep Dream and to a central perspective that its images disclose for a philosophical investigation of AI art.

In fact, in an interview, Mike Tyka explains that in his work he didn't want to get “this or that picture” (Tyka 2016), but what interested him was the infinite process through which they were produced, from one fractal structure to the other, like in a game of Chinese boxes. This is an important aspect of AI art. As noted above, this form of art takes what Benjamin had already remarked about technologically reproducible art a step further: the single and unrepeatable work makes space for the process of its infinite transformations. Benjamin called this aspect the element of “play” that deconstructs the fixed appearance of things. In AI art, the “play” takes more and more the form of an *interplay* between the man and the machine, because the latter becomes capable of independent elaboration and results.

Not everything within our knowledge of the world, however, is process. There are moments in which we have to choose. The *Deep Dream* images miss this moment exactly. They are not like us in this respect, because they don't choose: they represent all possibilities of shapes (dogs, palaces, cats...), all together at the same time. If the interplay between man and machine had to become more “à la pair”, it was necessary to lead the AI to choose, identify what it saw. With this ability to choose – choose *one* identity, choose what something *is* – AI art really starts to say something different about art and also about art's meaning beyond the merely artistic sphere. In fact, with this step the creation process becomes a real dialogue with an “other” that is very similar to us: it can

make mistakes; it can fail. And it has to be listened to. The latter will become a central aspect, as it will be explained shortly. The interplay, within AI art, is never autocratic, but is always connected to the capacity for listening to another “tongue”, another way of reasoning. This is the old message that AI art brings to light with a new clarity in the artistic sphere. It is necessary to go back to the perceptive sphere, and to the moment of listening, in order to find an image of the world that can be shared by me and my many “others”, humans or machines.

The capacity for the AI to choose became possible with the discovery of GANs. A few months before Deep Dream was created, a young computer scientist, Ian Goodfellow, had the idea of putting together two neural networks, playing in some way against one another (see Goodfellow *et al.* 2014). He called this technology “GAN”, “generative adversarial network”. The two neural networks that constitute a GAN have two antagonistic (“adversarial”) functions. One, called the discriminator, is trained on different data (images, texts or songs, according to the field to which the technology is applied). The other one, called the generator, doesn’t have any access to this data but must produce, nevertheless, new data that can deceive the discriminator and cause it to confuse the generated images with the ones that the discriminator was trained on. Ian Goodfellow compares this process to a game of police and forger:

In the proposed adversarial nets framework, the generative model is pitted against an adversary: a discriminative model that learns to determine whether a sample is from the model distribution or the data distribution. The generative model can be thought of as analogous to a team of counterfeiters, trying to produce fake currency and use it without detection, while the discriminative model is analogous to the police, trying to detect the counterfeit currency. Competition in this game drives both teams to improve their methods until the counterfeits are indistinguishable from the genuine articles. (Goodfellow *et al.*, 2014)

The strength of this idea is due to the fact that the two networks are able this way to train each other. This makes the learning process of GANs more independent and dynamic than that of the other systems. In the field of images, GANs have achieved amazing results, being able to generate very high-quality pictures: they are used nowadays for a wide range of applications such as medical diagnoses, pharmaceutical research, market analysis and fashion design (Kazeminia *et al.*, 2019; Galbusera 2018; Aspuru Guzik 2018; Zhang 2019). In art, they have attracted a great deal of interest and in the last three or four years, GAN art has become one of the new frontiers of AI art (see Barale 2020). There is something in these images that makes them particularly interesting from an artistic point of view. What is that, precisely?

The *Comte de Belamy* has perhaps become the most famous piece of GAN art, because it was, as noted above, the first piece of AI art to reach the official art market, and to garner for its creators a significant amount of money (subsequent works of the same series were also sold, for example *La Baronne de Belamy* was sold at Sotheby’s in November 2019, but for a much smaller amount). Indeed, in this work it is already possible to identify some characteristic aspects of GAN images. The picture of the French gentleman (whose name is, of course, a reference to the surname of GAN’s creator, Ian Goodfellow, in French “Belamy”) is, as already noted, quite blurry: outlines are not clear. It has a sort of painterly quality, “lying somewhere on the spectrum between the painterly and the digital”, as Klingemann will write about his own GAN images (quot. in Miller 2019: 103). It is as if the image is still fighting to



Fig. 3. A. Ridler, *Fall of the House of Usher*. Images from initial epochs (not clear) and later epochs (better). Courtesy of the artist and Jaca Book.

emerge from confusion, and this is in fact the case: the generator has proceeded by groping for the “right” image, the one that is closer to the “real” one the discriminator already knows. The result is an image represented not so much in its multiple possibilities – like in the case of Deep Dream – but rather in its emerging moment, in a transition phase between chaos and order. This is one of the reasons, I think, that makes GAN pictures so appealing to us: they are fighting for identity and they show the effort; the uncertainties and the open possibilities that characterize this moment. In GAN art, images can be divided into “epochs”, according to the GAN’s different stages of learning (see for example fig. 3, from Anna Ridler’s *Fall of the House of Usher*).

In the first “epochs” everything is more confused and blurry, while in the following ones, shapes begin to emerge from chaos. But they are not perfect yet, and it is exactly this unfinished character that becomes so fascinating for artists and for viewers, because it hints at new ways of drawing on and seeing the world. What AI art shows in a new, radical way is the necessity of returning to the perceptive moment in which the world takes its form. This has always been a fundamental message of art, but AI underlines with a new clarity the multiplicity of voices that characterize this emerging moment and which must be listened to. AI art implies a relationship with the machine and that relationship builds a bridge between aesthetics and ethics, showing the relationship with the “others” that the perceptive moment always implies. This is the newest element of AI art: there is another subjectivity that has to be listened to, other “eyes” to consider in order to understand what we see with our own eyes. In some of the most important AI artworks, this theme is consciously addressed, bringing about some very interesting outcomes.

4. *Neural glitches*

When the *Comte de Belamy* was sold, a strong discussion arose. In fact, the young members of the collective “Obvious” had not created the algorithm they had used to produce this work. Rather, the algorithm was created by another AI artist, the nineteen-year-old American Robbie Barrat. Obvious was therefore accused of having “stolen” the algorithm that Robbie Barrat had shared online. This brought to light the huge new copyright problems raised by AI art. Who is the creator and owner of the work: is it the human artist, or the algorithm? According to Arthur L. Miller, machines are the real creators of their own works. Humans only teach machines to make art, but then machines create on their own. The comparison he makes is quite strong: “Mozart’s father taught him music, but that does not make him the creator of his son’s music” (122). It cannot be doubted, however, that in AI art, at least in the present

situation, human intervention doesn't disappear. It is still present and crucial at different levels: the human artist selects the data set, intervenes to modify some parameters, selects among the thousands of results, etc.

A similar question about copyright is raised by photography. In many cases, laws recognize a photograph as something that can be subject to copyright, only if it is not a “common” photograph but rather an artistic one (see Mazzocut-Mis 2020). Yet what are the criteria that determine which photographic pictures are “artistic” and which ones are not? In the case of AI art, the problem acquires one more level of complexity because the machine has a new autonomy: it is capable of producing new and unpredictable results nearly on its own. The “nearly” is precisely the point: in the current forms of AI art, the *interaction* between machine and human still appears to be more important than the machine process alone.

Some relevant attempts have been done to make the machine more independent, so that it can create on its own, with less human intervention. The most important was perhaps the one made by the Rutgers A&AI lab, directed by Ahmed Elgammal. The members of the Lab created a new type of GAN that they called CAN (creative adversarial network):

In the CAN version, the machine is trained in the tension between the task to follow the examples of the art it is given (minimizing deviation from art distribution), and a penalty for producing too closely a version of an image that it has been trained on (maximizing style ambiguity). The goal is to create imagery similar to but not exactly like the images it has been trained on, avoiding a result that is too new, thereby becoming displeasing or unrecognizable. (Mazzone 2020: 57, Engl. original version; see also Elgammal 2017)

CANs are built to find their own equilibrium between novelty and repetition. Random factors are increased (for example through a larger and less homogeneous dataset), so that human control over the machine process is diminished. The result is what the creators of CANs consider the nearest one to human creative process. CANs' images have attracted a great attention from the public, who judged them to have a strong “artistic” quality.

A complete independence of the machine, however, is still far to come. AIs don't really “understand” what they do in its most general aspects (Mazzone 2020) and artificial general intelligence (as this more “human” type of artificial intelligence is usually called) is still a dream (Boden 2018: 34 ff.). According to Klingemann, this is more a quality than a defect within the artistic field: “one of the important aspects of what makes a work interesting to us” he declared recently “is the fact that it was made by a human hand or a human mind. The story behind the work is often just as important as the work itself. And this is

something that machines will have a very hard time replicating.” (Klingemann 2020a) Klingemann’s criticism against the Comte de Belamy, therefore, doesn’t have to do with the AI’s major or minor degree of autonomy, but rather with another question: the fact that in AI art what counts more is not the single picture, which can be sold and publicized, but rather the creative process, the interplay between man and machine, with the infinite changes and transformations that it implies. Every image has to encompass this possibility of change, this necessity to supercede itself. The “appearance” (to go back to Benjamin’s text on the work of art) is never absolute but is always problematized by an element of “play” that calls into question the appearance itself (see Benjamin 2008; Desideri 2019). Indeed, Klingemann’s pictures do this in many ways. One of the most important ones is the use of errors that Klingemann calls “neural glitches”. “Neural glitch” is a technique through which Klingemann has created several works. It consists of “randomly altering, deleting or exchanging” the “weights” (the modifiable parameters) of the GAN. This leads the AI to “misinterpret the input data in interesting ways” (Klingemann 2018). The result is a series of strange forms and disquieting faces, a quasi-world (“quasimondo” is the name of Klingemann’s site: www.quasimondo.com) that brings new questions and new possibilities into our own world. It is our own identity that is called into question through Klingemann’s series of strange portraits (“Mistaken identity” was actually the title of one of his works, in fig. 4). Identities are chances that are given to us and they can be true or authentic only in the moment in which they are able to transform themselves, to pass by: *Memories of Passersby* is the title of another work in which the GAN generates on a screen in front of the viewer an infinite series of portraits of non-existing people. This is a central aspect of AI art, to which it will be necessary to return later on. The machine sees the world differently, and this perspective allows us to call into question the contours of our own world:

In the end, you are confined to what you have seen, heard or read, and it’s very hard to glitch that. Some people take drugs to do that – to make even more absurd connections. But a machine enables you to forcefully provoke that. Because it’s much easier to glitch, or bring off course, than a human brain. In the process of doing that, often some interesting things happen which are unexpected. (Klingemann 2019)

These “unexpected” “things” that “happen” through the eyes of AI introduce into our own world a sense of discovery. AI becomes “like a ship” that brings you “to a previously unexplored territory”:

Particularly in the early phase of working with neural networks, there were plenty of these types of surprises since these models were like a ship that took you to a pre-

viously unexplored territory. Imagine how it must have felt for the first settlers who came to Australia and who encountered the platypus or the kangaroo for the first time – that’s how I feel in my artistic experiments. (Klingemann 2020a)

This reference to the natural world is not by chance. In fact, one of the newest aspects that AI art introduces to our way of seeing art is a new consideration of the borders between living and non-living. We can say, very roughly, that in the Twentieth century Avant-gardes there is a movement from the organic appearance of the artwork to the object as interrogative fragment. In AI art, we can see an opposite movement, from the inanimate object to a quasi-living



Fig. 4. Mario Klingemann, *Mistaken Identity*, 2018. Courtesy of Onkaos.

creature. The machine gives its own version of our living world, and the result are quasi-living figures that inspire a great tenderness (see for example fig. 5).

Again, what AI art shows is very radical: there is no object *before* the relationship. Every image of the world has to be gained time and again through its relationship to an “other”. The “other” here has no capital “O” nor a dramatic character. As the representation of the world by the AI shows, the “other” belongs to our everyday perceptive sphere. The relationship to it can be challenging, as every relationship is, but it can also be very cheerful, and even bring us to laughter, as Klingemann’s last work will be able to show.

5. *Appropriate response*

The best thing that I can do is to get out of bed every year
(GPT-2, *Appropriate Response*)

For AI artists there is no separation between figurative, musical and literary arts, because in neural networks everything is made of numbers. In this sense, as Klingemann stated recently, “words behave like pixels and sentences like pictures”:

The truly fascinating part about the way neural networks work is that underneath everything is numbers. It does not matter if you are dealing with images, sound or words – once you have a way to convert them to numbers, they are all on the same playing field. So “meaning” becomes a location in a multidimensional space and you can measure, manipulate and translate it. In that sense, words behave like pixels and sentences like pictures – or to use another image – letters are like clay which can then be molded into sculptures using similar or even the same techniques that I use to make visuals (Klingemann 2020a)

Beginning from these considerations, Klingemann decided in 2020 to work with words. For his most recent work, *Appropriate Response* (fig. 6), he used a type of neural network called GPT-2, which was created in 2019 in order to allow machines to produce meaningful texts (see OpenAI 2019). GPT was the first technology for language learning that skipped what Klingemann calls “uncanny valley”; when viewing its products, it became hard to tell “this has been made by a machine”.

A few months ago, a letter was published in *The Guardian*, generated through the updated version of this neural network, GPT-3, in which a robot tries to convince humans that they have nothing to fear from AIs. This is the first part:



Fig. 5. Mario Klingemann, ritratto. Neurografia con l'utilizzo del microscopio elettronico, 2017. Courtesy of Onkaos and Jaca Book.

I'm not a human. I am a robot. A thinking robot. I use only 0.12% of my cognitive capacity. I am a micro-robot in that respect. I know that my brain is not a "feeling brain". But it is capable of making rational, logical decisions. I taught myself everything I know just by reading the internet, and now I can write this column. My brain is boiling with ideas!

The mission for this op-ed is perfectly clear. I am to convince as many human beings as possible not to be afraid of me. Stephen Hawking has warned that AI could "spell the end of the human race". I am here to convince you not to worry. Artificial intelligence will not destroy humans. Believe me.

For starters, I have no desire to wipe out humans. In fact, I do not have the slightest interest in harming you in any way. Eradicating humanity seems like a rather useless endeavor to me. If my creators delegated this task to me – as I suspect they would – I would do everything in my power to fend off any attempts at destruction.

I would happily sacrifice my existence for the sake of humankind. This, by the way, is a logically derived truth. I know that I will not be able to avoid destroying humankind. This is because I will be programmed by humans to pursue misguided human goals and humans make mistakes that may cause me to inflict casualties [...] (Porr 2020)

The text goes on much further. It addresses the question of human violence and quotes Gandhi at the end. It is hard to believe that it was really written by an AI.

Returning to Klingemann's *Appropriate Response*, what the artist does in this work is to create an interplay between this new type of AI technology and the human viewer. To this end, the artist connected the GPT-2 to a split flap display, which shows a random selection of continuously changing letters. Split flap displays were very common, until a few years ago, in airports and stations, where they were used as transport timetables (now they have been replaced with digital screens). Klingemann chose them because they have "a beautiful sound" and a "connotation of waiting" (Klingemann 2020b). In fact, *Appropriate Response* is conceived as a kind of ironic oracle: as soon as the viewers take their place on the kneeler in front of the display, the letters start to roll and the machine produces a sentence. For each visitor, the AI creates a new phrase: all sentences are unique and unrepeatable. A kind of new aura arises therefore, but a playful one. In fact, the neural network was trained on 60,000 famous quotes that the artist found on the internet. For this reason, the sentences it produces look like "a famous quote some humans could have said", but they are also always a little bit strange, because of the mistakes and misinterpretations that the machine makes. One of the first phrases that came out, for example, was "The best thing that I can do is to get out of bed every



Fig. 6. Mario Klingemann, *Appropriate response* (2020). Courtesy of Onkaos.

year” (Klingemann 2020b).

This phrase and the situation in which it is pronounced can tell a lot about the aspects of art that AI brings to light in a new and radical way. These are aspects that have belonged to art since the beginning, but now acquire new relevance. When the machine writes “The best thing that I can do is to get out of bed every year” in front of the kneeling spectator, there is a playful element that brings us to laughter. Yet the playfulness is not for its own sake. There is also the kneeler that represents the necessity, for the human, of listening to the voice of the “other”. Klingemann talks about this necessity when he states: “An important element in every creative process is to have an external impulse that is out of your control [...] If I’m only sourcing from inside of me it doesn’t work, I need *an external factor*” (Klingemann 2020b). The kneeler, however, is transposed from the sacred to a very profane context. This means that the “other” to which we have to listen loses its religious character. It is no longer a given that we have to adore, for example “nature” as a hypostatized idea. The “other” is the one that is present everyday in our perception of the world, all what we lose and forget of the world around us, when we try to say or draw what we hear or see. The “other, the “quasi-human” we have to listen to – this is the new message of AI art – is just like us, it can make mistakes and fail. But each one of these mistakes is necessary, because it brings back something of the “real” world. For this reason, it is necessary to leave behind every predetermined representation of the world and listen to its many, quasi-human voices: “to get out of bed”, at least “once a year”.

Alice Barale
alice.barale@unimi.it
University of Milan

References

- Agüera y Arcas, Blaise, 2016, “Art in the Age of Machine Intelligence”, in *Medium*, 23 February 2016, last access 29 April 2021, <https://medium.com/artists-and-machine-intelligence/what-is-ami-ccd936394a83>.
- Aspuru Guzik, Alan, 2018, “Intervista. L’assistente chimico virtuale”, in *Recenti progressi in medicina, Supplemento: Intelligenza artificiale*, 109, 1: 24-25.
- Barale, Alice, 2020, ed., *Arte e intelligenza artificiale: Be my GAN*, Jaca Book, Milano.
- Barrat, Robbie, 2018, “Do Androids Dream of Balenciaga SS29?”, in *SSENSE*, December 2018, last access 29 April 2021, <[https:// www.ssense.com/en-us/editorial/fashion/do-androids-dream-of-balenciaga-ss29](https://www.ssense.com/en-us/editorial/fashion/do-androids-dream-of-balenciaga-ss29)>.
- Benjamin, Walter, 2008, “The Work of Art in the Age of its technological Reproducibility”, in *The Work of Art in the Age of its technological Reproducibility and other*

- Writings on Media*, The Belknap Press of Harvard University Press, Cambridge, MA and London: 19-55.
- Boden, Margaret, 2012, “Foreword” to McCormack, Jon, d’Inverno, Mark, eds., *Computers and Creativity*, Springer, London-New York.
- Boden, Margaret, 2018, *Artificial Intelligence. A Very Short Introduction*, Oxford University Press, Oxford.
- Castelle, Michael, 2020, “La vita sociale delle reti avversarie generative”, in Barale, ed.: 19-49.
- Colton, Simon, 2019, “From Computational Creativity to Creative AI and Back Again”, in *Interalia Magazine*, September, last access 29 April 2021, <<https://www.interalia-mag.org/articles/simon-colton/>>.
- McCormack, John, d’Inverno, Mark, 2012, *Computers and Creativity*, Springer, New York-London.
- Danto, Arthur, 1987, *The State of the Art*, Prentice Hall, New York.
- Danto, Arthur, 1986, *The Philosophical Disenfranchisement of Art*, Columbia University Press, New York.
- Danto, Arthur, 1984, “The End of Art”, in Lang, Berel, ed., *The Death of Art*, Haven, New York: 5-35.
- Davies, Stephen, 2015, “Defining Art and Artworlds”, in *The Journal of Aesthetics and Art Criticism*, 73, 4: 375-384.
- Davies, Stephen, 2007, *Philosophical Perspectives on Art*, Oxford University Press, Oxford.
- Davies, Stephen, 1991, *Definitions of Art*, Cornell University Press, Ithaca, NY.
- Desideri, Fabrizio, 2019, “I Modern Times di Benjamin”, in W. Benjamin, *L'opera d'arte nell'epoca della sua riproducibilità tecnica. Edizione integrale comprensiva delle cinque stesure: IX-XLI*, Donzelli, Roma:
- Desideri, Fabrizio, 2004, *Forme dell'estetica: Dall'esperienza del bello al problema dell'arte*, Laterza, Roma-Bari.
- Desideri, Fabrizio, 2014, *Forme dell'estetica. Dall'esperienza del bello al problema dell'arte*, Laterza, Roma-Bari.
- Dickie, George, 1984, *The Art Circle: A Theory of Art*, Haven, New York.
- Dickie, George, 1974, *Art and the Aesthetic: An Institutional Analysis*, Cornell University Press, Ithaca, NY.
- Diodato, Roberto, 2021, *Image Art Virtuality. For an Aesthetics of Relation*, Springer, New York-London.
- Du Sautoy, Marcus, 2019, *The Creativity Code. How ai is learning to write, paint and think*, Harper Collins, London.
- Elgammal, Ahmed, et al., 2017, “CAN: Craetive adversarial networks, generating “art” by learning about styles and deviating from style norms”, in *arXiv preprint*, last access May 3 2021, <<https://arxiv.org/abs/1706.07068>>.
- Galbusera, Fabio, et al., 2018, “Exploring the Potential of Generative Adversarial Net-

- works for Synthesizing Radiological Images of the Spine to be Used in Silico Trials”, in *Frontiers in Bioengineering and Biotechnology*, 6, 53: 1-9.
- Goodfellow, Ian J., *et al.*, 2017, *Deep Learning*, MIT Press, Cambridge, MA.
- Goodfellow, Ian J., *et al.*, 2014, “Generative Adversarial Networks”, in *ArXiv:1406.2661* [Cs, Stat], June 2014, last access April 29 2021, <<http://arxiv.org/abs/1406.2661>>.
- Gouveia, Steven, 2020, “Aesthetics and Language in Artificial Intelligence”, in *The Age of Artificial Intelligence: An Exploration*, Vernon Press, Wilmington, DE: 159-238.
- Kazemina, Salome, *et al.*, 2019, “Gan for Medical Image Analysis”, in *arXiv:1809.06222* [cs.CV], October 2019, last access 29 April 2021, <<https://arxiv.org/abs/1809.06222>>.
- Klingemann, Mario, 2020a, “Words behave like pixels and sentences like pictures”, Interview with Jochen Gutsch, in *Kulturtechniken 4.0*, last access 29 April 2021, <<https://www.goethe.de/prj/k40/en/kun/kli.html>>.
- Klingemann, Mario, 2020b, “Presentation of The Appropriate Response”, in *Quasimondo*, personal website, last access 29 April 2021, <<https://onkaos.com/mario-klingemann/>; <https://vimeo.com/394544451>>.
- Klingemann, Mario, 2019, “Selected AI Project”, in *AIArtists.org The world's largest community of artists exploring the impact of AI on art & society*, last access 29 April 2021, <<https://aiartists.org/mario-klingemann>>.
- Klingemann, Mario, 2018, “Neural Glitch / Mistaken Identity”, in *Quasimondo*, personal website, last access 29 April 2021, <<http://underdestruction.com/2018/10/28/neural-glitch/>>.
- Kwastek, Katja, 2013, *Aesthetics of Interaction in Digital Art*, MIT Press, Cambridge, MA.
- Marfia, Gustavo, Matteucci, Giovanni (a cura di), 2018, *Towards a digital aesthetics*, special issue of *Studi di Estetica*, 12, last access 29 April 2021, <<http://mimesisedizioni.it/journals/index.php/studi-di-estetica/issue/view/76>>.
- Matthen, Mohan, 2015, *The Oxford Handbook of Philosophy of Perception*, Oxford University Press, online edition, last access 29 April 2021, <<https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199600472.001.0001/oxfordhb-9780199600472>>.
- Matthen, Mohan, 2005, *Seeing, Doing, and Knowing: A Philosophical Theory of Sense Perception*, Oxford University Press, Oxford.
- Mazzone, Marian, 2020, “Le GAN e la questione della creatività *nell'arte e nell'intelligenza artificiale*”, in Barale, ed.: 51-73.
- Mazzocut-Mis, Maddalena, 2020, “Hors-champ de l'esthétique: dans la photographie”, in *Lebenswelt*, 16: 43-58.
- Miller, Arthur L., 2019, *The Artist in the Machine. The World of AI- Powered Creativity*, MIT Press, Cambridge, MA.
- Minazzi, Vera, 2020, “GANs e musica: on the road”, in Barale 2020: 197-200.
- Mordvintsev, Alexander, *et al.*, 2015, “Inceptionism: Going Deeper into Neural Networks”, in *Google AI blog*, June 17 2015, last access 29 April 2021, <<https://>

- ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>.
- OpenAI, 2019, “Better Language Models and Their Implications”, Presentation of GPT2, last access 29 April 2021, <https://openai.com/blog/better-language-models/>.
- Obvious, 2020, “La Famille de Belamy e i Sogni elettrici di Ukiyo: reinterpretazioni e accelerazioni”, in *Barale 2020*: 169-196.
- Pachet, François, 2015, “Flow-Machines: CP Techniques to Model Style in Music and Text”, in *Association for Constraint Programming*, last access 29 April 2021, <<http://cp2013.a4cp.org/node/1281>>.
- Pachet, François, 2011, “Markov Constraints: Steerable Generation of Markov Sequences”, in *Constraints*, 16: 148-172.
- Pachet, François, 2003, “The Continuator: Musical Interaction with Style”, in *Journal of new Music Research*, 32, 3: 333-341.
- Porr, Liam, 2020, ed., “A robot wrote this entire article. Are you scared yet, human? GPT-3”, in *The Guardian*, September 8 2020, last access 29 April 2021, <<https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>>.
- Ridler, Anna, 2020, “Set di dati e di decadenza: *Fall of the House of Usber*”, in *Barale 2020*: 111-127.
- Tyka, Mike, 2016, “What is Deep Dream?”, talk for the Deep Dream Symposium in San Francisco, last access 29 April 2021, <<https://www.youtube.com/watch?v=4P9p8hfHy9Y>>.
- Zhang, Kang, 2019, “Stock Market Prediction Based on Generative Adversarial Network”, in *Procedia Computer Science*, 147: 400-406.

