Ethics and Information Technology (2024) 26:58
https://doi.org/10.1007/s10676-024-09796-0

ORIGINAL PAPER

A portrait of the artist as a young algorithm

Sofie Vlaad[1]

Abstract
This article explores the question as to whether images generated by Artificial Intelligence such as DALL-E 2 can be considered artworks. After providing a brief primer on how technologies such as DALL-E 2 work in principle, I give an overview of three contemporary accounts of art and then show that there is at least one case where an AI-generated image meets the criteria for art membership under all three accounts. I suggest that our collective hesitancy to call AI-generated images art stems from the lack of a clear author figure. I propose two possible complementary solutions. First, that some AI-generated images as artworks are conjunctively authored by both the developers of the AI and the prompt-giver. Second, that AI image generators can themselves be considered works of art authored by the developers. I conclude by way of suggesting that we might have separate art competitions specifically for AI-generated artworks.

Keywords Philosophy of AI · Philosophy of art · DALL-E 2 · Authorship · AI art

Can AI-generated images sometimes be considered art?

Something shocking happened at the 2023 Sony World Photography Awards. The winning submission was from a photographer named Boris Eldagsen (Parshall, 2023). Titled *The Electrician* (Fig. 1), the photograph depicts two women in black and white. However, Eldagsen did not take the photograph himself. In fact, *The Electrician* is not a photograph at all. Rather, Eldagsen used an Artificial Intelligence (AI) platform called DALL-E 2 to generate the image. Eldagsen refused the award, stating that the submission was a stunt to see if the competition could identify AI-generated images. Clearly, it did not. Eldagsen used his victory as a platform to start a discussion of the role of AI in photography—and art more broadly. The story made international headlines and public response was mixed. While some applaud and encourage the use of AI technologies in the creation of art, others question whether such creations can even be considered art at all. Eldagsen himself championed the use of AI to generate images akin to photographs as a new art form dubbed promptography.

The response to AI image generators has not been all positive however, with many taking to social media to express their disdain for the medium and to dismiss such images as art (Shaffi, 2023). As far back as 2019, Sean Dorrance Kelly, a philosophy professor at Harvard wrote an article for the *MIT Technology Review* arguing that AI cannot be considered an artist, as it lacks the creative abilities found in humans (Kelly, 2019). Contrary to Eldagsen, many people either dismiss AI-generated images as art or dismiss DALL-E 2 as an artistic collaborator.

At present, AI image generators are ubiquitous. From Midjourney to DALL-E 2, one need only log on to social media to become inundated with images created with generative AI. As the case of *The Electrician* shows, while some people believe that at least some AI-generated images ought to be considered art, others remain unconvinced. Additionally, there do not seem to be any straightforward ways of adjudicating which AI-generated images—if any—we ought to understand as art. We might expect to find some solutions by turning to philosophical definitions of art. However, in this paper I will argue that while current trends in philosophy of art have little difficulty understanding AI-generated images as art, there remains strong resistance from the general public accepting them as such. In order to explain this hesitancy, I suggest that the case of

✉ Sofie Vlaad
16sv8@queensu.ca

[1] Philosophy Department, Queen's University, Kingston, ON, Canada

Published online: 27 August 2024

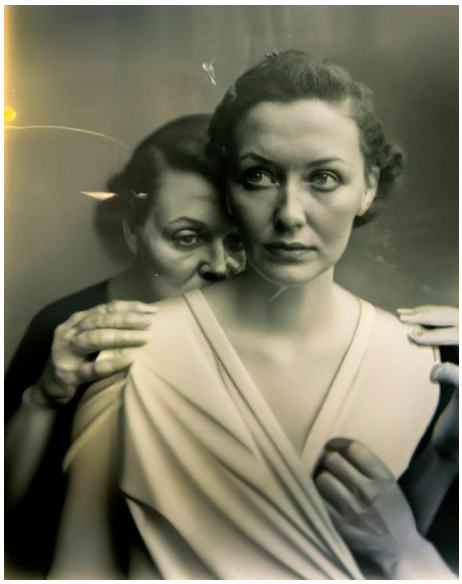

Fig. 1 *The electrician* (attributed to Boris Eldagsen and DALL-E 2) https://www.scientificamerican.com/article/how-my-ai-image-won-a-major-photography-competition/

AI-generated images—such as *The Electrician*—presents us with difficult questions about *authorship* in ways that deeply challenge our intuitions of the concept.[1]

This paper proceeds as follows. First, I give a brief explanation of how AI image generators like DALL-E 2 work. Next, following Lopes (2014), I identify three kinds of theories of art: exhibited features accounts, genetic features accounts, and Lopes' buck passing account. I then show that all of these prevailing theories can satisfactorily deal with the case of so-called AI art by applying them to *The Electrician*. I argue that despite the fact that our current accounts of art all classify *The Electrician* as a work of art, resistance to this claim speaks to important issues that arise about the locus of authorship in such cases that might make us question their status as art. I outline the issues that arise when we consider any single individual to be the author in these cases, and present two complementary solutions. First, we can conceive of AI-generated images as collaborative works of art in a conjunctive sense. Additionally, we can consider AI image generators themselves to be artworks authored by the developers. I outline both views and conclude by exploring some implications that follow.

AI image generators

Contemporary AI image generators are a subset of deep generative models called vision-language models (Yang et al., 2023). As the name suggests, a popular application of these models is Text-to-Image generation, wherein an image is generated from a corresponding descriptive text (Du et al., 2022). A user inputs text prompts that the model then translates into images. These models are trained on datasets consisting of millions or even billions of image-text pairs (Schuhmann et al., 2022). For example, one might have an image of a dog with the corresponding text "dog." These image-text pairs help train the model to associate input text with sufficiently relevant corresponding images. Typically, models use an encoder-decoder framework. First, an encoder reads the data and transforms it into a different representation. Then, a decoder transforms the new representation into output (Ridoy et al., 2024). An example of this would be an encoder transforming image-text pairs (input) into vectors (representation) which are then transformed into new images (output) by the decoder.

State of the art image generation often uses diffusion models (Ho et al., 2020; Yang et al., 2023). Introduced by Sohl-Dickstein et al. (2015) and improved upon by others (Ho et al., 2020), Denoising Diffusion Probabilistic Models (DDPMs) work by introducing signal noise into the data and slowly reversing the process to generate new data samples. In the case of Text-to-Image generation we might imagine the text prompt "dog" gets matched with a corresponding image of a dog from the dataset. The image-text pair is then transformed into a vector by an encoder. The diffusion model then introduces noise interference. The interference is slowly reversed by a decoder that then transforms the vector into a new image reconstructed based on the likelihood of certain elements being closer in proximity to others (see Fig. 2).

For example, the brown pixel of a dog's ear has a low probability of being located next to a white pixel of fur from the dog's tail, so the model is unlikely to generate an image where this is the case. However, there is a high probability that two brown ear pixels will be next to each other and thus it is highly likely that the model will generate an image concordant with this high probability. Diffusion models tend to achieve better results over time and as such are used in many vision-language models—including DALL-E 2.

Elsewhere, Radford et al. (2021) have proposed a vision-language model called CLIP (Contrastive Language-Image Pre-training) designed to classify images by giving corresponding text. Originally conceived as a tool for automated image captioning, CLIP uses both an image encoder and a text encoder. During pre-training, image-text pairs are fed to CLIP. The image encoder translates the images into vector representations which are then arranged in a 1xN matrix. The text encoder does the same for the text, arranging the vectors in an Nx1 matrix. The two matrices are multiplied to generate an NxN vector space (see Fig. 3). The elements that comprise the space are image-text pairs. Image-text

[1] Professor Kelly's (2019) arguments against AI as artists serves as a saliant example of this phenomenon.

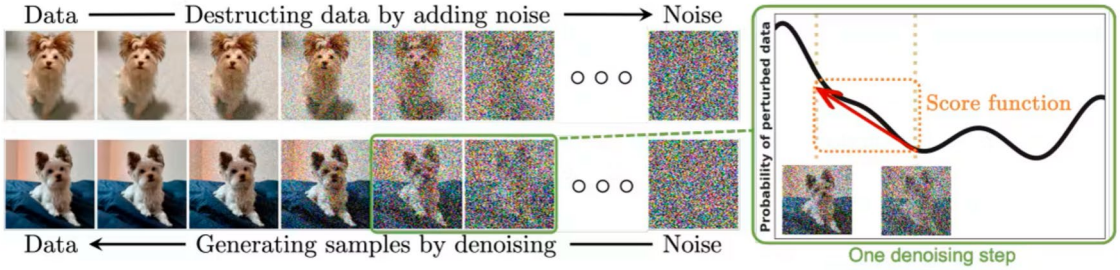

Fig. 2 The diffusion process (taken from Yang et al. (2023) Fig. 2)

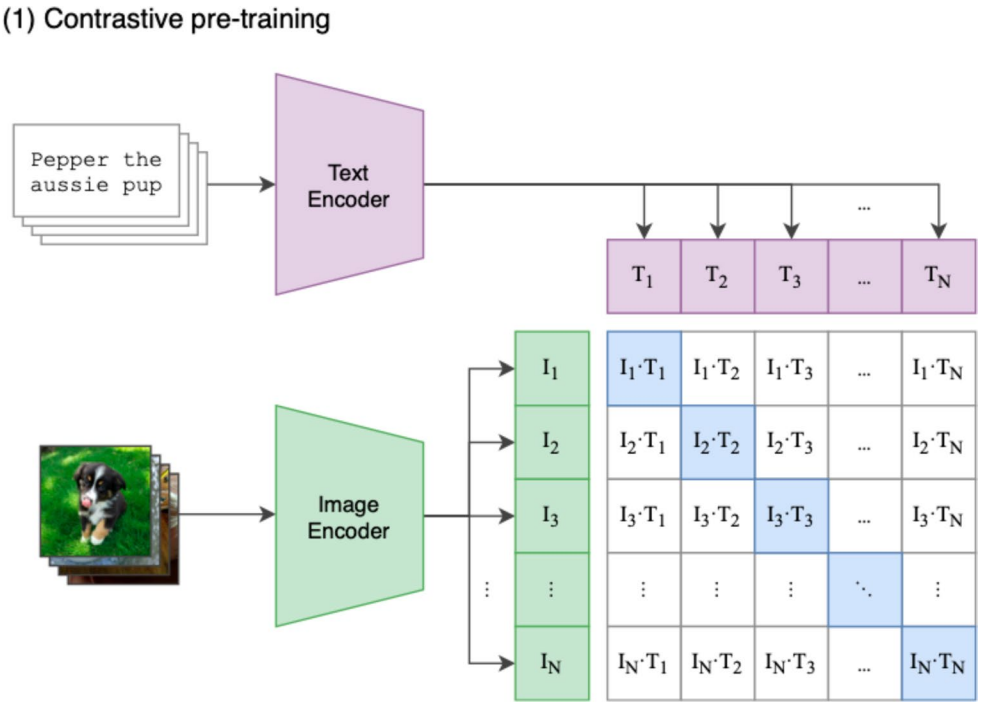

Fig. 3 CLIP pre-training (abbreviated Fig. 1 from Radford et al. (2021)

pairs that are similar to one another are closer to each other in the vector space. When fed an image I_n, CLIP matches it with the closest image-text pair I_nT_n and produces the corresponding text T_n as output.

For example, we might have many image-text pairs of dogs, but only some of those dogs are running. An image of a running dog is closer to the image-text pair of a running dog with the caption "running dog" than it would be to an image-text pair of a sitting dog with the caption "sitting dog." Thus, when fed an image of a running dog, CLIP generates a caption along the lines of "running dog."

Ramesh et al. (2022) have suggested a two-step image generation process involving a prior model that produces CLIP-based image embeddings and a diffusion decoder that transforms the embeddings into images. This technique is called unCLIP—as it inverts the CLIP process—and is used by DALL-E 2. In step one, a text prompt is transformed into an associated CLIP-based image embedding by the prior model. For example, the prompt "running dog" will generate embeddings based on training set images for which "running dog" is an appropriate caption. In step two, the diffusion decoder transforms the image embeddings into new images. The user can then choose from a set of images produced which one they like best, or they can refine their prompt to produce a different set of images. Whereas CLIP generates text from image inputs, unCLIP generates images from text inputs. Figure 4 illustrates the training process above the dotted line and the generation process below the dotted line.

While DALL-E 2 has the capacity to generate fascinating images such as *The Electrician*, it remains an open question as to whether some of these images ought to be considered art. In the next section I give an overview of three leading contemporary accounts of art to guide this discussion.

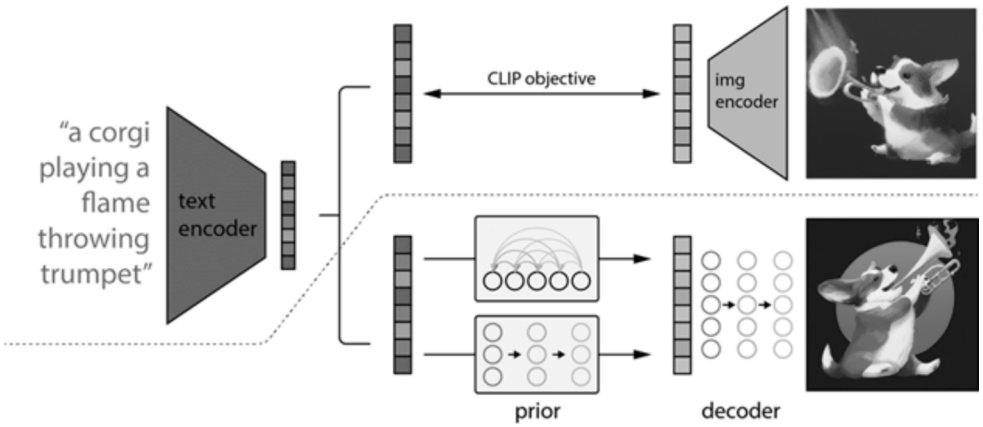

Fig. 4 unCLIP (taken from (Ramesh et al., 2022) Fig. 2)

Three accounts of art

Philosophers and art critics have long debated what constitutes art. For Kant (1790/2009), fine art is the product of genius. It is both novel and exemplary of the form. The Beauty of art thus differs from the Beauty of nature as it has no final end but itself. Some thinkers (Meynell, 1986) claim that aesthetic value in the form of Beauty is a necessary property of good art. This view has been contested (Danto, 2003; Steiner, 2001), with scholars arguing both that we ought to consider antiaesthetics (Ziff, 1984) and negative aesthetic value (Brady, 2011). The former holds that there is art without any aesthetic value present at all, whereas the latter holds that some aesthetic value is negative in quality. Rather than judge a work of art as exemplary in virtue of its Beauty, we might judge it as exemplary in virtue of its Ugliness. Others still reject the concept of art entirely (Richter, 1965), a move common among practitioners of Dada. Recently some scholars have attempted to reclaim Beauty (Lopes, 2018; Nehamas, 2007), but the conversation has largely shifted to defining art in ways other than the presence or absence of aesthetic value.

In *Beyond Art* (2014), Dominic Lopes addresses a problem that arises in contemporary trends in analytic philosophy of art.[2] Lopes takes there to be two main kinds of theory of art. On the one hand, there are exhibited features accounts, that hold that we can identify art by its properties. On the other hand, there are genetic features accounts, that hold that we can identify art by the context of its genesis. As there are paradigm examples of art that either exhibited features or genetic features accounts fail to classify as art, neither account can provide a comprehensive understanding of art as such. Lopes proposes the buck passing account in order to solve this problem. While the buck passing account is able to solve the problem of classifying art, it does not solve the demarcation problem of art—nor does Lopes claim it can do so.

I explore each theory in depth, beginning with exhibited features accounts, followed by genetic features accounts, and concluding with the buck passing account.

Exhibited features accounts

Exhibited features accounts define art as possessing a set of properties. While these properties are often aesthetic in nature, that need not necessarily be the case (Lopes, 2014). For example, Ziff (1953), proposes seven sufficient properties of a work of art:

1. It is recognizable as a work of art.
2. It was made by an artist deliberately with skill and care.
3. It was intended by the artist to be treated like a work of art.
4. It is presented to others by whom it is discussed, studied, contemplated, admired, and criticized.
5. It contains subject matter.
6. It has a complex formal structure.
7. It is good of its kind. (Ziff, 1953, pp. 60–61)

It is important to note that in Ziff's account a work of art need not possess all seven properties. If a work were to possess all seven, then we know for certain that it is a work of art. Accordingly, if a work does not exhibit any of the features listed, then it is not a work of art. However, if a work exhibits some of the features in question, then we might make a case for the work as art.

Consider Van Gogh's *The Starry Night* as an example of a work that meets all seven criteria:

[2] There are other ways of carving out current trends in analytic philosophy of art, such as the inclusion of hybrid exhibited/genetic features accounts as distinct kinds of accounts. See Longworth and Scarantino (2010) for an example of such an account. However, given that these accounts are ultimately comprised of exhibited and genetic features accounts as Lopes describes them, and given the uniqueness of Lopes' own account of art, it is reasonable to follow Lopes as other views are both reducible to Lopes' taxonomy, and would have to discuss the buck passing account regardless.

1. *The Starry Night* is recognizable as a painting.
2. It was made by Van Gogh deliberately with skill and care.
3. It was intended by Van Gogh to be treated like a work of art.
4. *The Starry Night* has been presented to others by whom it has been discussed, studied, contemplated, admired, and criticized for over a century.
5. *The Starry Night* contains subject matter—it depicts a night sky in Paris.
6. *The Starry Night* has a complex formal structure—Van Gogh used a myriad of impressionist techniques in its creation.
7. *The Starry Night* is a good painting.

As *The Starry Night* is a paradigm case of art, it is unsurprising that it meets all seven of Ziff's criteria. While exhibited features accounts such as Ziff's provide useful heuristics for identifying many—if not most—works of art, they struggle to deal with certain difficult cases such as Marcel Duchamp's *Fountain*.

In 1917 Duchamp took a readymade sculpture of a urinal and entered it in an art gallery under a pseudonym. This caused a great deal of controversy and sparked debates about whether *Fountain* ought to be considered art at all, as it was merely an average urinal presented as a provocation. According to exhibited features accounts such as Ziff's, *Fountain* does not qualify for art status, as it meets none of the requisite criteria. However, contemporary scholars *do* consider *Fountain* to be a work of art, perhaps being an early example of what would later come to be known as conceptual art. That exhibited features accounts fail to classify avant-garde works such as *Fountain* as art seems to be a shortcoming of such accounts.

Genetic features accounts

Genetic features accounts, however, do not share this shortcoming. While exhibited features accounts focus on properties expressed by a work, genetic features accounts define art in terms of the context of its creation. It is to these accounts that I now turn.

Genetic features accounts come in two variations. The first are institutional accounts. These hold that for a work of art to be considered as such, it must be presented to an artworld (Danto, 1964; Dickie, 1974). This means that the work in question must be presented to an audience of artists, critics, and aficionados. For a work to be a work of art, it must have an intentional connection to art as an institution. Note that this does not mean that the artist or creator of the work must intentionally understand it as art. Rather, anyone can take the work and present it to an artworld.

Dickie (1969) himself gives an example in the form of works made by chimpanzees. According to Dickie, if chimpanzees were to make paintings—an event which has been known to occur on occasion—then what matters is what is done with the paintings. If the paintings were to be displayed in the study of a primatologist, then we ought not consider them art. Rather we might conceive of the paintings as artifacts providing a glimpse into the cognitive and motor capacities of chimpanzees. However, if the paintings were instead displayed in an art gallery, and thus presented to an artworld, then they would have a claim to art status and likely be understood as works of art. We might similarly think of a state of affairs wherein *The Starry Night* remained in Van Gogh's attic and was never presented to an artworld. If that were the case, then *The Starry Night* would not count as art according to this account. Levinson (1979) presents this as a reductio of Dickie's account, and develops a historical account of art in response.

The second variation of genetic features accounts are historical accounts. These hold that for a work of art to be considered as such it must stand in relation to a history of art (Levinson, 1990). For example, a Modernist poem might be responding to a tradition of Romantic poetry, thus standing in historical relation to earlier works of art. Tom Stoppard's play *Rosencrantz and Guildenstern are Dead* is an example of a more explicit response to histories of art, as it is a retelling of Shakespeare's *Hamlet* from the perspective of the minor characters Rosencrantz and Guildenstern.

An upshot of both variations of genetic features accounts is that they can classify cases of avant-garde works as art. Whereas exhibited features accounts fail to understand *Fountain* as a work of art, for example, genetic features have no such issue. *Fountain* was both presented to an artworld when it was displayed in a gallery, and it exists in relation to a history of avant-garde and experimental art. However, exhibited features accounts would not struggle with the case of *The Starry Night* in the attic, as it would meet most of the sufficient criteria. It is especially important that it meets the first of Ziff's criteria in that it is recognizable as a painting. If the inability of exhibited features accounts to classify works such as *Fountain* as art counts as a weakness of these accounts, then the inability of genetic features accounts to classify *The Starry Night* in the attic as art would seem to be a weakness of those accounts as well.

While neither exhibited features nor genetic features accounts can easily deal with challenging cases, Lopes provides a solution in the form of the buck passing account, to which I now turn.

The buck passing account

In *Beyond Art* (2014) Lopes provides an alternative to both exhibited features and genetic features accounts of art. According to Lopes, the buck passing theory of art "passes the buck" from a general theory of art to specific theories of the arts. In this account, "x is a work of art = x is a work of K, where K is an art" (Lopes, 2014, p. 49). "The Wasteland" by T. S. Eliot is a work of poetry, where poetry is one of the arts. Therefore, "The Wasteland" is a work of art. The buck passing account of art displaces the burden of explanation from a general theory of art to specific theories of the arts.

Lopes defends the buck passing account of art against both exhibited features and genetic features accounts. According to Lopes, the buck passing account of art "… deals more effectively with the hard cases than its competition" (Lopes, 2014, p. 46). By "hard cases" Lopes has in mind works such as *Fountain*. Artists and critics have since developed vocabulary to deal with the case of *Fountain*, with many—including Lopes—considering it to be a work of conceptual art.

Lopes argues that exhibited features accounts are forced to deny hard cases like *Fountain* as art, on the grounds that there will exist non-art counterparts to the hard cases that exhibit the same features. Put simply: there is no discernable difference in features between *Fountain* and any given urinal one might find in a public restroom. Genetic features accounts take the opposite position and accept *Fountain* as art. However, for Lopes, both genetic features accounts and exhibited features accounts—what he terms buck stopping accounts—fail to take the hard cases seriously. For Lopes,

> buck stopping theories of art cannot cope effectively with the hard cases because the theories themselves and the criteria for choosing among the theories echo—and so cannot adjudicate—intuitions about the hard cases. (Lopes, 2014, p. 63)

According to Lopes, exhibited features accounts do not take the hard cases seriously because they unilaterally dismiss them as art. This dismissal runs contrary to how works such as *Fountain* are understood by artists and critics themselves. On the other hand, genetic features accounts accept the hard cases as art a priori, leaving no room for discussion.[3]

The foreclosure of contesting the status of the hard cases also runs contrary to how these works are understood. While *Fountain* is considered art by most professional artists today, this was not the case upon its unveiling in 1917. The buck passing account is appealing because it allows room for discussion and contestation, whereas the other types of accounts do not.

Additionally, the buck passing account can also deal with the case of the chimpanzee paintings stored in the primatologist's study. Applied to the buck passing account, the chimpanzee paintings are works of painting, where painting is one of the arts. Therefore, the chimpanzee paintings are works of art. Likewise for *The Starry Night* that has been left in the attic. The fact that the buck passing account can classify challenging cases that either exhibited features or genetic features accounts struggle with is a definite upshot of the account.

However, there is a downside to the buck passing account. Namely, in passing the buck to theories of the arts, it does nothing to solve the demarcation problem of art. Furthermore, if the case in question is not currently legible as one of the arts, then the buck passing account has nothing to say about it. It is easy to classify works like *Fountain* as conceptual art when one has the vocabulary and framework available to do so. It is another situation entirely where no such vocabulary or frameworks presently exist.

Understanding *the electrician* as art

I now show how the contemporary accounts of art discussed in this article come to understand *The Electrician* as a work of art. I begin with exhibited features accounts (exemplified by Ziff), then turn to genetic features accounts (both institutional and historical) and conclude with Lopes' buck passing account. I argue that each kind of account classifies *The Electrician* as art—albeit for different reasons. This section should be understood as a small case study rather than a comprehensive survey on how all accounts of art would understand *The Electrician*.

The electrician and exhibited features accounts

I begin with exhibited features accounts. Recall Ziff's (1953) seven sufficient properties for whether a work qualifies for art status:

1. It is recognizable as a work of art.
2. It was made by an artist deliberately with skill and care.
3. It was intended by the artist to be treated like a work of art.
4. It is presented to others by whom it is discussed, studied, contemplated, admired, and criticized.
5. It contains subject matter.
6. It has a complex formal structure.

[3] There are many examples of contested cases of demarcation problems in the arts, such as photography (Lopes, 2016; Scruton, 1981) and videogames (Ensslin, 2012; Smuts, 2005). After developing the buck passing account in *Beyond Art* (2014), Lopes gives an analysis of videogames using his account.

7. It is good of its kind. (Ziff, 1953, pp. 60–61)

It seems to be the case that *The Electrician* meets most of these conditions. Taken one at a time, we see that:

1. *The Electrician* is clearly recognizable as a work of art, insofar as it is recognizable as a photograph. It need not actually *be* a photograph for this criteria to obtain. Note that it need only be *recognizable* as such.
2. It is unclear whether we can say that *The Electrician* was made by an artist with skill and care, given that we would first have to determine to whom we ought to attribute the piece. The work is credited to both Eldagsen and DALL-E 2. It seems strange to attribute the qualities of skill and care to AI, and it is also unclear whether the prompt curation done by Eldagsen counts as such.
3. Eldagsen intended for *The Electrician* to start a conversation about art, serving a role similar to *Fountain*.
4. *The Electrician* was submitted to the Sony World Photography Awards, where it was discussed, studied, contemplated, admired, and criticized by an audience of judges to whom it was presented.
5. Whether *The Electrician* contains subject matter depends on which account of depiction one endorses. Resemblance accounts such as those endorsed by Abell (2009) and Lopes (2004) would hold that *The Electrician* contains subject matter insofar as it depicts two women posing. Those who reject resemblance accounts of depiction, such as Goodman (1976), would deny that *The Electrician* contains subject matter.
6. Given the realistic quality and high resolution of *The Electrician*, it seems fair to say that it has a complex formal structure.
7. *The Electrician* was well received by the judges, winning the competition. This seems to be evidence that it is good of its kind (it is a good work of art).

Given that *The Electrician* clearly meets four out of the seven sufficient criteria provided by Ziff—with criteria 2, 3, and 5 being ambiguous or contentious—it seems safe to say that it has a reasonable claim to art status.

The electrician and genetic features accounts

Recall that genetic features accounts come in two varieties. First, there are institutional accounts that hold that a given work must be presented to an artworld for it to count as art. Second, there are historical accounts that hold that a given work must stand in a historical relation to art for it to count as such. I will argue that *The Electrician* counts as art according to both forms of genetic features accounts.

As *The Electrician* was submitted to the Sony World Photography Awards as a work of art, it was indeed presented to an artworld. In this case the artworld is comprised of judges, although the international news coverage that followed its victory arguably exposed it to an even larger artworld in the form of the general public.[4]

The Electrician also has ties to a history of art, as it seem to be an example of generative art (G-art) (Boden & Edmonds, 2009). According to Boden and Edmonds, "G-art works are generated, at least in part, by some process that is not under the artist's direct control" (Boden & Edmonds, 2009, p. 37). This would include any kind of art generated through the use of generative systems. Aleatoric music—sometimes called chance music—comes to mind as a salient example. Composers such as John Cage would sometimes craft pieces of music through the use of chance operations. The paradigm example here is using dice as a tool in composition. Suppose you have a twelve-sided die (D12). Each number on the die corresponds to a note of the chromatic scale (1 = C, 2 = $C^{\#}$, 3 = D, etc.), and rolling the die determines the next note in the sequence. This approach to art-making is also seen in the generation of *The Electrician*. While DALL-E 2 is far more complex than a D12, both are forms of generative systems.

The electrician and the buck passing account

In order for the buck passing account to classify *The Electrician* as a work of art, it must belong to one of the arts. In this section I argue that *The Electrician* is a work of computer generative art (CG-art), a subset of G-art. As such, *The Electrician* is a work of art.

In "What is Generative Art?" Margaret Boden and Ernest Edmonds (2009) identify a taxonomy of what they call generative art. While Boden and Edmonds identify eleven distinct but related types of generative art—sometimes called computer art—the relevant definition here is CG-art. According to Boden and Edmonds, "CG-art is produced by leaving a computer program to run by itself, with minimal or zero interference from a human being" (Boden & Edmonds, 2009, p. 37). This definition certainly maps on to Eldagsen's role in creating *The Electrician*. While Eldagsen provided the prompt that ultimately generated *The Electrician*, refining it along the way, it was DALL-E 2 that actually created the image.

One could argue that prompt curation is not constitutive of minimal interference. Rather, the very existence of *The Electrician* is dependent on Eldagsen's prompts. Here we might turn to the analogy of building a house, where Eldagsen takes on a role equivalent to the homeowner whose

[4] For an alternative perspective on how Danto would interpret AI art, see Cascales (2023).

house is being built. While both Eldagsen and the homeowner might guide the contractors and the AI by giving them suggestions—such as what color to paint the walls, or how many women to include in the image—it is ultimately the contractors and the AI that do the heavy lifting. If we would say that the homeowner in this case is minimally involved in building the house, then the same should hold of Eldagsen's relationship with the creation of *The Electrician*.

In summary, *The Electrician* is a work of CG-art, where CG-art is one of the arts. Therefore, *The Electrician* is a work of art. Note that we could also say that *The Electrician* is a work of promptography if we so desire—bearing in mind that the term is new and contested. According to all three kinds of accounts of art—exhibited features accounts, genetic features accounts, and the buck passing account—*The Electrician* ought to be considered a work of art. Note that it does not follow that *The Electrician* is a work of photography even if it is a work of art. However, there remains popular resistance to accepting works such as *The Electrician* and other AI-generated images as art. In the next section I suggest that we have unique reasons for being skeptical of AI-generated content as art given questions that arise about authorship in these cases.

Who is the artist?

While the term author has diverse connotations within and outside the field of literature, in this context I am using the term to refer to creator of a given work. For this reason, I use the terms author and artist interchangeably. It is not readily apparent to whom we ought to attribute authorship in cases of AI-generated images. Recall that for Kant, the difference between the Beauty of fine art and the Beauty of nature is that the former is the product of genius. Regardless of how we interpret the term genius, it seems clear that Kant intends for it to refer to a person. If we cannot identify an artist this seems to be grounds for remaining skeptical of AI-generated images possessing art status, as it seems conceptually incoherent to have art in the absence of a corresponding artist. It is also relevant that one of the seven of the criteria proposed by Ziff that does not clearly obtain in the case of *The Electrician* concerns whether it was made by an artist with skill and care.

In this section I provide an overview of some accounts of authorship and then identify two possible solutions to the case of *The Electrician*. First, I consider that authorship is conjunctive in these cases. As such, both the developers and the prompt-giver are responsible for the art generated by DALL-E 2. Second, I consider that the images generated by DALL-E 2 are not always themselves art. Rather, we can understand DALL-E 2 itself as a work of computer art (Lopes, 2010). In this case, the images generated are the results of the audience—the prompt-givers—interacting with the artwork. Taken together, we can understand how some AI-generated images achieve art status while still understanding the entire system as a work of art.

Theories of authorship

The philosophy of authorship has largely been concerned with ascribing author status in the case of literary texts, rather than visual artworks. In this section I give a brief overview of some of these accounts and how they might apply to the case of *The Electrician*.

Roland Barthes (1967/1977) famously argues that the author as such is dead. Literary works do not acquire their meanings from an all-knowing author-figure, rather, it is up to the audience or the readers to imbue a work with meaning. If we take the "author" to be the one who imbues a work with meaning, then perhaps the author is dead in the case of AI. However, this is not how we have been using the term. For our purposes, the "author" or "artist" is synonymous with the "creator" of a given work. Determining the locus of meaning making is somewhat beside the point in this case. For example, it is logically consistent that we might identify a "creator" (or "creators") in the case of *The Electrician*, and this individual or group is not responsible for imbuing the work with meaning.

Elsewhere, for Foucault (1969/1998), all texts have writers, but not all writers are authors. Private correspondence is an example of texts with writers but not authors. Foucault conceives of an author-function that sets the terms of the discourse surrounding a given text. According to Foucault, this author-function must be abolished in order to transform the kinds of conversations we can have and questions we can ask of a text. On the other hand, Nehamas (1986) argues against Foucault and holds that the author is co-constructed with a text and is themselves subject to interpretation. A consequence of this view is that the author cannot be abolished as they always emerge through the text.

Abolishing the notion of authorship might seem appealing in the case of *The Electrician*, as it would mitigate some of the hesitancy we have with attributing art status to the work. However, this once again seems to sidestep the issue of ascribing creator status. If we could abolish the author (or the author-function) we would still be left with a "writer" or "creator." With this in mind, we can clearly see that the question we should be asking in the case of AI-generated art is not "who is the author" or "who is the artist" but "to whom do we attribute creator status?"

To this end I turn to work on collective authorship that takes the terms "author" and "creator" to be largely synonymous. We can add "artist" to that list as well.

Conjunctive authorship

If we consider AI-generated images such as *The Electrician* to be works of art in their own right, attributing authorship to a single individual or class of individuals seems to be an untenable position. We cannot consider the prompt-giver to be the sole artist, as they neither generate the images nor design the system that does. In other cases of generative art, such as aleatoric music, the composer both designs and enacts the systems that generates the art. In the case of AI-generated images, the prompt-giver only enacts the system.

Likewise, we cannot consider the developers to be the sole artists, as the input of the prompt-giver is *necessary* to generate the artwork. For similar reasons, the model cannot be the sole author as it can only perform its functions thanks to human involvement—both from the developers and the prompt-giver. As such, if we are to consider AI-generated images to be art, then a conjunctive view of authorship seems to be required.

The notion that authorship can be distributed amongst a group is not new. Marcel Duchamp (1957/1975) argues that all art is multi-authored, given that it is the audience that deciphers and interprets the art and thus contributing to its meaning. Barthes (1967/1977) gives a similar argument when he declares that the author is not the sole locus of meaning-making. Rather, the audience are active participants in the creation of a work: "the birth of the reader must be at the cost of the death of the author" (Barthes, 1967/1977, p. 148). This is not to say that the writer of a text plays no role, rather the act of authorship is a group endeavor.

We need not go as far as to say the audience co-authors every work of art. However, in the case of AI-generated images there are clear candidates for authorship such that we can identify the individual contributing artists. We might say that the AI-generated images are the work of at least two artists. These are the developers and the prompt-giver. We could possibly claim that the model also counts as a contributing artist. However, this would commit us to also assigning the role of contributing artist to the D12 in the case of aleatoric music. I remain agnostic with respect to the status of the model as artist.

Building on work on collective authorship in cases such as films, wherein works are co-created by multiple individuals (Sellors, 2007; Stillinger, 1991), Mag Uidhir, (2011) gives an account of authorship capable of accounting for attributing authorship in collaborations. What makes this account unique is that instead of considering authorship a two-place relation—an agent **a** is an author of a work **w**—Mag Uidhir takes authorship to be a three-place relation such that **a** is an author of **w** as a work-description **F**. For example, Charlie Kaufman is the author of *Being John Malkovich* as a script, and Spike Jonze is the author of *Being John Malkovich* as a film. Mag Uidhir goes one step further and posits conjunctive authorship such that "Only [**A** and **B**…] is an author of **w** as an **F**, but neither **A** nor **B**… taken alone is an author of **w** as an **F**." (Mag Uidhir, 2011, p. 377). This account of conjunctive authorship maps neatly onto the case of *The Electrician*.

Given that neither Eldagsen nor the developers of DALL-E 2 are solely responsible for generating *The Electrician*, and yet both are necessary for its genesis, we can say the following. Only [Boris Eldagsen and the developers of DALL-E 2] is an author of *The Electrician* as a work of art, but neither Boris Eldagsen nor the developers of DALL-E 2 taken alone is an author of *The Electrician* as a work of art. While this view is compatible with the suggestion that not all images generated by DALL-E 2 constitute art, I go one step further. In the next section I argue that we can conceive of DALL-E 2 as a work of computer art authored by the developers, whose output via audience interaction sometimes becomes art in its own right.

DALL-E 2 as computer art

According to Lopes (2010), "an item is a computer art work just in case (1) it's art, (2) it's run on a computer, (3) it's interactive, and (4) it's interactive because it's run on a computer" (Lopes, 2010, p. 27). Let's unpack this in relation to DALL-E 2. Condition (1) will be the most difficult to satisfy, so I will leave it for last. Condition (2) easily obtains—DALL-E 2 is run on a computer. Condition (3) also easily obtains—DALL-E 2 is interactive insofar as it requires text-prompts from its users. Condition (4) requires that DALL-E 2's interactivity be necessarily linked to its being run on a computer. Given that DALL-E 2 is AI, there is no way it could be interactive unless it were run on some kind of hardware. Ergo, Condition (4) obtains. With respect to Condition (1), we have pragmatic reasons for understanding DALL-E 2 as a work of (computer) art, as it changes how we understand both the images it generates and how we are to attribute authorship.

On the surface however, DALL-E 2 does not appear to be a work of art itself. We might appeal to the three accounts of art provided earlier to see how it fares under each. Exhibited features accounts present the most resistance to understanding DALL-E 2 as art. Some conditions clearly obtain. DALL-E 2 was made with skill and care, it has a complex formal structure, and it was presented to others by whom it was discussed, studied, contemplated, admired, and criticized. For some conditions it is more ambiguous if they obtain. It is unclear if DALL-E 2 is good of its kind, or what it would mean for it to have content. Furthermore, there is at least one condition that does not obtain at all—DALL-E

2 is not immediately recognizable as art. Even still, we can certainly say that DALL-E 2 has been embraced as a source of entertainment by many.

Genetic features accounts fare better—DALL-E 2 was presented to the public with an invitation to engage as artists. We might be able to interpret this as DALL-E 2 being presented to an artworld. Additionally, DALL-E 2 sits in a historical relationship to other kinds of computer art insofar as interacting with it as a prompt-giver yields outcomes similar to other forms of computer art. We might think of DALL-E 2 as being similar to some video games where the agency of the player determines—to an extent—how the world of the game develops. Lopes himself uses the example of video games in his discussion of computer art.

Finally, the buck passing account seems to have little difficulty understanding DALL-E 2 as art, as it can be classified under several different arts. We might say that DALL-E 2 is a work of computer art, given that its interactivity is due to its being on a computer. We might say that it is a work of conceptual art, with the concept being that anyone can use it to become an artist. Finally, we might say that it is interactive art—similar to performance art—given that it is an object whose form and content is affected by the behavior of the audience. All these ways of classifying DALL-E 2 render it a work of art according to the buck-passing account.

Conclusion

These results coupled with our pragmatic reasons seem to be at least promising. If DALL-E 2 is an artwork, then it follows that the developers are the artists. Not all interactions with the audience—the users—will generate images that are artworks in their own right. Indeed, many images generated will be discarded in the refining process. However, some images are able to achieve art status. Such cases can be explained by appealing to a conjunctive account of authorship.

Given that AI-generated images such as *The Electrician* can be considered works of art, there is a strong case to be made that at least in principle, they ought to be allowed in art competitions. However, the framework of conjunctive authorship holds that such images ought to be attributed to both the prompt-giver and the developers. Furthermore, identifying what kind of art AI-generated images count as will determine what kinds of art competitions they ought to be allowed into. For example, since *The Electrician* is not a photograph, the Sony World Photography Awards is an inappropriate venue to which it could be submitted.

One solution would be to have separate competitions for AI-generated images as art. This would both provide a venue for sharing and appreciating such works as art without taking away from other forms of art such as photography. This would also discourage practitioners of AI art to disguise their works as paintings or photographs and might even encourage them to push the boundaries of AI-generated art into strange new directions unique to the medium.

This might even encourage developers to rethink how they design models in the first place. If AI-generated images constitute a unique form of art, what are the aesthetics and compositional strategies unique to this new artform? Such questions will surely be answered as these technologies develop and become even more widespread.

Funding This research was supported by SSHRC funding.

Data availability Data sharing is not applicable, as the manuscript proceeds from philosophical foundations and does not analyse or generate datasets.

References

Abell, C. (2009). Canny resemblance. *The Philosophical Review*, *118*(2), 183–223. https://doi.org/10.1215/00318108-2008-041

Barthes, R. (1977). The Death of the Author. In S. Heath (Trans.), *Image, Music, Text* (pp. 142–148). Fontana. (Original work published 1967).

Boden, M. A., & Edmonds, E. A. (2009). What is Generative Art? *Digital Creativity*, *20*(1–2), 21–46.

Brady, E. (2011). The Ugly Truth: Negative aesthetics and Environment. *Royal Institute of Philosophy Supplement*, *69*, 83–99.

Cascales, R. (2023). Interpreting AI-Generated art: Arthur Danto's perspective on intention, authorship, and creative traditions in the age of Artificial Intelligence. *The Polish Journal of Aesthetics*, *71*, 17–29.

Danto, A. (1964). The Artworld. *The Journal of Philosophy*, *61*(19), 571–584.

Danto, A. (2003). *The abuse of Beauty: Aesthetics and the Concept of Art*. Open Court.

Dickie, G. (1969). Defining art. *American Philosophical Quarterly*, *6*(3), 253–256.

Dickie, G. (1974). *Art and the aesthetic: An institutional analysis*. Cornell University Press.

Du, Y., Liu, Z., Li, J., & Zhao, W. X. (2022). *A Survey of Vision-Language Pre-Trained Models* (arXiv:2202.10936). arXiv. https://doi.org/10.48550/arXiv.2202.10936

Duchamp, M. (1975). The Creative Act. In M. Sanouillet, & E. Peterson (Eds.), *The essential writings of Marcel Duchamp* (pp. 138–140). Thames and Hudson. (Original work published 1957).

Ensslin, A. (2012). Computer gaming. In J. Bray, A. Gibbons, & B. McHale (Eds.), *The Routeledge Companion to Experimental Literature* (pp. 497–511). Routledge.

Foucault, M. (1998). What is an Author? In J. D. Faubion (Ed.)R. Hurley (Trans.), *Aesthetics, Method, and Epistemology* (pp. 205–222). The New Press. (Original work published 1969).

Goodman, N. (1976). *Languages of Art: An Approach to a Theory of Symbols* (2d ed.). Hackett.

Ho, J., Jain, A., & Abbeel, P. (2020). *Denoising Diffusion Probabilistic Models* (arXiv:2006.11239). arXiv. https://doi.org/10.48550/arXiv.2006.11239

Kant, I. (2009). In N. Walker, & J. C. Meredith (Eds.), *The critique of judgement*. Trans.). Oxford University Press. (Original work published 1790).

Kelly, S. (2019). *A philosopher argues that an AI can't be an artist*. *MIT Technology Review*. https://www.technologyreview.com/2019/02/21/239489/a-philosopher-argues-that-an-ai-can-never-be-an-artist/

Levinson, J. (1979). Defining art historically. *The British Journal of Aesthetics*, *19*(3), 232–250. https://doi.org/10.1093/bjaesthetics/19.3.232

Levinson, J. (1990). *Music, art, and Metaphysics: Essays in Philosophical aesthetics*. Cornell University Press.

Longworth, F., & Scarantino, A. (2010). The disjunctive theory of art: The Cluster Account Reformulated. *The British Journal of Aesthetics*, *50*(2), 151–167. https://doi.org/10.1093/aesthj/ayq001

Lopes, D. (2004). *Understanding pictures*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199272037.001.0001

Lopes, D. (2010). *A philosophy of Computer Art*. Routledge.

Lopes, D. (2014). *Beyond art*. Oxford University Press.

Lopes, D. (2016). *Four arts of photography: An essay in Philosophy*. Wiley.

Lopes, D. (2018). *Being for Beauty: Aesthetic Agency and Value*. Oxford University Press.

Mag Uidhir, C. (2011). Minimal authorship (of sorts). *Philosophical Studies*, *154*(3), 373–387. https://doi.org/10.1007/s11098-010-9525-0

Meynell, H. (1986). *The nature of aesthetic value*. SUNY.

Nehamas, A. (1986). What an author is. *The Journal of Philosophy*, *83*(11), 685–691.

Nehamas, A. (2007). *Only a Promise of Happiness: The place of Beauty in a world of art*. Princeton University Press.

Parshall, A. (2023, April 21). How This AI Image Won a Major Photography Competition. *Scientific American*. https://www.scientificamerican.com/article/how-my-ai-image-won-a-major-photography-competition/

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). Learning transferable visual models from Natural Language Supervision. *arXiv*. arXiv:2103.00020. http://arxiv.org/abs/2103.00020

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). *Hierarchical text-conditional image generation with clip latents* (arXiv:2204.06125v1). arXiv.

Richter, H. (1965). *Dada: Art and Anti-art*. McGraw-Hill.

Ridoy, M. A. R., Hasan, M. M., & Bhowmick, S. (2024). *Compressed Image Captioning using CNN-based Encoder-Decoder Framework* (arXiv:2404.18062). arXiv. https://doi.org/10.48550/arXiv.2404.18062

Schuhmann, C., Beaumont, R., Vencu, R., Gordon, C., Wightman, R., Cherti, M., Coombes, T., Katta, A., Mullis, C., Wortsman, M., Schramowski, P., Kundurthy, S., Crowson, K., Schmidt, L., Kaczmarczyk, R., & Jitsev, J. (2022). *LAION-5B: An open large-scale dataset for training next generation image-text models* (arXiv:2210.08402). arXiv. https://doi.org/10.48550/arXiv.2210.08402

Scruton, R. (1981). Photography and representation. *Critical Inquiry*, *7*(3), 577–603.

Sellors, C. P. (2007). Collective authorship in Film. *Journal of Aesthetics and Art Criticism*, *65*(3), 263–271. https://doi.org/10.1111/j.1540-594X.2007.00257.x

Shaffi, S. (2023, January 23). 'It's the opposite of art': Why illustrators are furious about AI. *The Guardian*. https://www.theguardian.com/artanddesign/2023/jan/23/its-the-opposite-of-art-why-illustrators-are-furious-about-ai

Smuts, A. (2005). Are Video games Art? *Contemporary Aesthetics*, *3*, 1–11.

Sohl-Dickstein, J., Weiss, E. A., Maheswaranathan, N., & Ganguli, S. (2015). *Deep Unsupervised Learning using Nonequilibrium Thermodynamics* (arXiv:1503.03585). arXiv. https://doi.org/10.48550/arXiv.1503.03585

Steiner, W. (2001). *Venus in Exile: The rejection of Beauty in Twentieth-Century Art*. University of Chicago Press.

Stillinger, J. (1991). *Multiple authorship and the myth of solitary genius*. Oxford University Press.

Yang, L., Zhang, Z., Song, Y., Hong, S., Xu, R., Zhao, Y., Zhang, W., Cui, B., & Yang, M-H. (2023). Diffusion models: A comprehensive survey of methods and applications. *arXiv*. arXiv:2209.00796v10

Ziff, P. (1953). The Task of defining a work of art. *The Philosophical Review*, *62*(1), 58–78.

Ziff, P. (1984). *Antiaesthetics: An appreciation of the cow with the Subtle nose*. D. Reidel Publishing Company.