artnet®

Has Artificial Intelligence Given Us the Next Great Art Movement? Experts Say Slow Down, the 'Field Is in Its Infancy'

Christie's decision to sell a work of AI art this fall has sparked a debate about the status of the movement.

Tim Schneider & **Naomi Rea**, September 25, 2018

Theo Triantafyllidis, *How to Everything* (2016). Copyright Theo Triantafyllidis, courtesy of the artist.

The news that Christie's would sell an artwork made by artificial intelligence this October captured worldwide headlines and imaginations alike. *Portrait of Edmond de Belamy* (2018), an uncanny, algorithm-created rendering of an aristocratic gentleman, will hit the auction block in New York with an estimate of $7,000 to $10,000. But the piece's inclusion in such a high-profile sale is creating controversy far ahead of the auction itself.

While images generated using AI technology have been circulating relatively widely since Google's pattern-finding software DeepDream roared onto the scene in 2015, the field was still young, and the artworks produced via AI were neither aesthetically nor conceptually rich enough to hold the attention of the art world for long. But after the heavyweight auction house announced it was ready to sell this latest work, the mysterious portrait—and the even more mysterious algorithm behind it—were cast by many in the media as the new standard-bearers for the genre.

There is, however, a man behind the curtain—or rather, three men. The algorithm responsible for the portrait was developed by a French art collective called Obvious, which created its AI artist using a model called a GAN, short for "generative adversarial network."

GANs entered the AI conversation in 2014 when researcher Ian Goodfellow published an article theorizing that they represented the next step in the evolution of neural networks: the interconnected layers of processing nodes, modeled loosely on the human brain, that have driven many recent developments in artificial intelligence. The concept of GANs captured people's attention partly because, unlike the repetitious results produced by Google DeepDream—all of which were simply pre-existing images that appeared to be run through what our critic Ben Davis described as the same "psychedelic Instagram filter"—GANs could be trained to produce completely new, and dramatically different, images.

In order to get their algorithm to produce a legitimately classical-looking portrait, Obvious's members say they fed it a training data set of more than 15,000 portraits created between the 14th and 20th centuries. Using these images, the algorithm was able to "generate" new images similar to the ones it had been fed. These new portraits were then presented to another algorithm (the "adversarial" part of the GAN acronym) that was trained to distinguish between images produced by humans versus those by machines—a Turing-like test for artworks—until the generated portraits could fool this discriminator into thinking they were "real," too.

Since the announcement, many in the traditional art world have been losing their minds over this new movement, which Obvious has christened "GAN-ism." But other artists making work via AI think the hype about what the technology can do on its own is premature. Mario Klingemann, a pioneer of AI art who has been working with GANs since their inception, says that while this model can create images that look new and seem to define the genre of AI-made art, they are actually just one tool in the neural art toolbox.

"Because they create instant gratification even if you have no deeper knowledge of how they work and how to control them, they currently attract charlatans and attention seekers who ride on that novelty wave," Klingemann says.

Specifically, Klingemann doesn't think that Obvious deserves its newfound legitimacy. "Pretty much everyone who is working seriously in this field is shaking their head in disbelief about the lack of judgment when it comes to featuring Obvious, and of course Christie's decision to auction them out of all artists who work with neural networks," he says.

"The work isn't interesting, or original," agrees Robbie Barrat, a young artist who works with AI. "They try to make it sound like they 'invented' or 'wrote' the algorithm that produced the works," but says the collective actually uses a pre-existing implementation of the model to generate low-resolution outputs, which they then run through an online enhancer to get the final image. "People have been working with low resolution GANs like this since 2015," says Barrat, adding that last year, when he was 17 years old, he did a project using the exact same type of neural network and an identical data set. "No one in the AI and art sphere really considers them to be artists—they're more like marketers."

Christie's international head of prints and multiples, Richard Lloyd, justified the choice to showcase Obvious by pointing to the limited human intervention in their creative process, evinced by the fact that the collective chose to credit the artwork to the algorithm rather than themselves. "This particular work was chosen in part because of the process," Lloyd says. "Obvious tried to limit the human intervention as much as possible, so that the resulting work reflects the 'purist' form of creativity expressed by the machine."
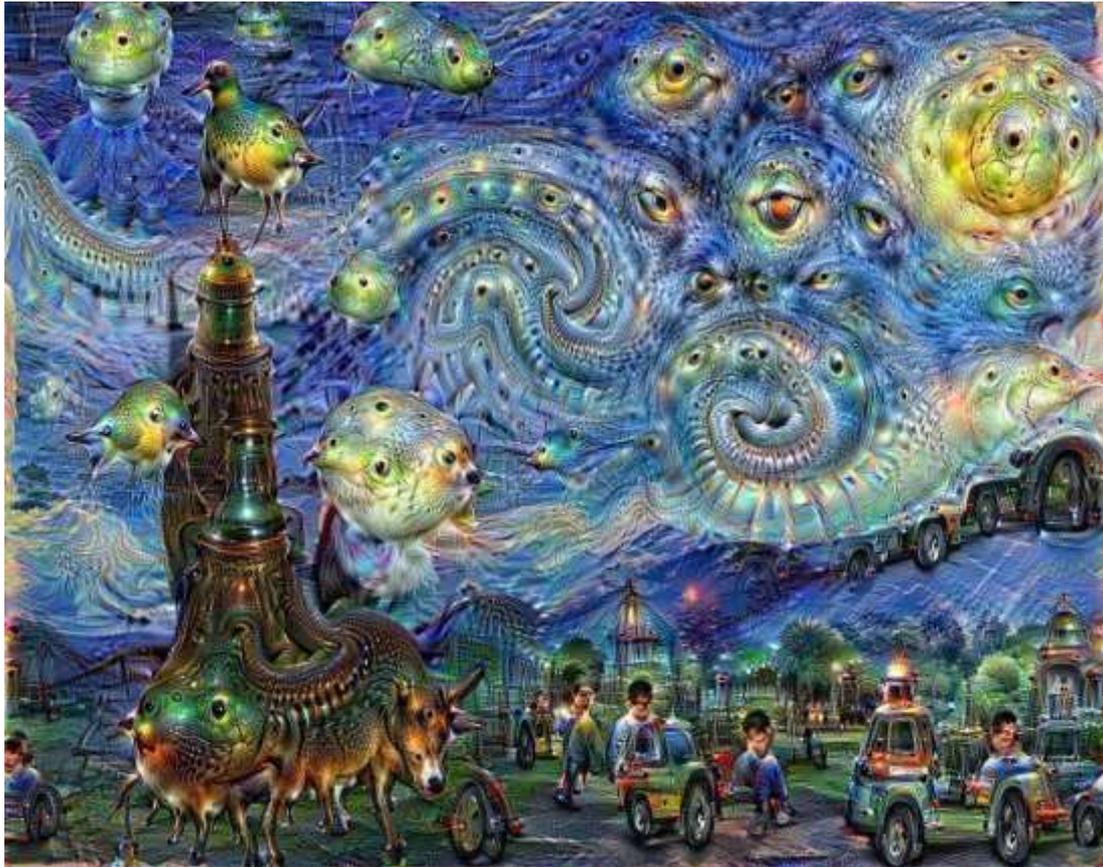
A Vincent van Gogh-inspired Google DeepDream painting. Photo courtesy of Google.

History Lessons

To get a better understanding of how excited (or alarmed) we should be about AI-influenced art, we have to look beyond what artists have been doing with GANs and to the history of artificial intelligence itself.

At present, much of the debate about any use of AI, including art, is colored by a grand misconception of the technology. Decades of science fiction have conditioned the general public to associate any invocation of AI with what technologists and researchers would call "artificial general intelligence," or a machine with the abilities to think freely, learn without specific training, and perhaps even experience emotions. It's the AI of Hollywood classics like C-3PO in *Star Wars* (the benevolent version) and HAL-9000 in *2001: A Space Odyssey* (the malevolent version).

Researchers are actively working toward artificial general intelligence, and recent breakthroughs at behemoth tech companies like Google make its eventual achievement seem plausible, though probably not imminent. But these are not the types of AI being used by Obvious, Klingemann, or any other artist making use of a neural network, for the simple reason that they do not exist yet.

A closer look at GANs reveals the long, winding road between their current capabilities and the sentient robots that the phrase "artificial intelligence" conjures for the uninitiated. Karthik Kalyanaraman, an economist and the curator of the first major gallery show devoted to art made via AI, explains that after Goodfellow's essay about GANs was published in 2014, raw and untrained GANs were made open source by different companies, including Google (TensorFlow), Facebook (Torch), and the Dutch NPO radio broadcaster (pix2pix).

The artistic part actually comes in when these algorithms are trained to produce specific images, which is a process executed by human beings. "Artists come at these algorithms that are raw, like blank slates that are like children ready to be trained, and then train them in a particular visual style," Kalyanaraman says. It can take up to two months to train GANs to produce images in line with the desired concept.

A lot separates GAN-ism from world-changing AI artists unassisted by humans. The issue isn't just that GANs can only learn about the specific labeled data they're fed ("These images are all portraits"), or that the boundaries of the tasks they can perform are very narrow ("Now that you've seen all those portraits, make me a portrait"). It's also that only one of the most advanced neural networks ever developed by a tech colossus has ever learned to perform *more than one* cognitive task. Outside of this techno-financial edge case, educating a GAN to do something new requires a complete erasure of its progress on the original task—a process known in the AI field by the vivid phrase "catastrophic forgetting."

In practice, catastrophic forgetting means the same GAN that Obvious used to create *Portrait of Edmond de Belamy* cannot also be used to create a landscape

painting, let alone go rogue and liquefy life on Earth. (Oxford AI theorist <u>Nick Bostrum</u> famously proposed this possibility as a <u>legitimate danger of a super-AI programmed only to manufacture paper clips</u>.) For now, the potential of the technology vastly outpaces the reality. Artist Theo Triantafyllidis, whose work will appear in the Athens Biennial this October, grounds the discussion in relatable terms: "Even state-of-the-art AI today is somewhere between a really clumsy child and a really smart pet."

Theo Triantafyllidis, *Seamless* (2017). Copyright Theo Triantafyllidis, courtesy of the artist.

Adversarial Approaches to GAN-ism

In that light, what Obvious has done in training the raw algorithm on a data set of existing portraits might seem less original than was previously understood, as others could replicate the training process and achieve similar results.

Unsurprisingly, then, the subjects in Obvious's work share many similar, surreal distortions with the subjects of early work by Barrat, Klingemann, and other artists using GANs. Klingemann even came up with a name for the phenomenon: the "<u>Francis Bacon</u> effect."

This repetitiveness shows that humans still have an important role to play in shaping how an AI develops creatively. Although the tech was made publicly available shortly after the Goodfellow paper was published, Kalyanaraman says it took until early 2017 for artists to really begin to recognize the potential for aesthetic diversity and conceptual richness that AI art can provide. "I think it's just a sign of the fact that this field is in its infancy," Kalyanaraman says, explaining that in the past year and a half there have been new developments on the technical side of things, as well as more traditional artists getting involved in the field.

"They've started to think about this stuff beyond the 'selling point' of it being art produced by AI, but have also started to think about how conceptually rich can they

make it," Kalyanaraman says. He adds that the more artists start working with AI, the more obvious it will be when the work being produced is unoriginal.

Indeed, AI has captured the attention of artists both more engaged in the conceptual depths of the technology and more deeply embedded in the structures of the traditional art market. Take Trevor Paglen and his recent series, "Adversarially Evolved Hallucinations." Rather than train a GAN on portraits to produce an approximation of a human-rendered portrait, he trained GANs on images labeled only with symbolic or metaphorical terms, then directed them to produce images of these subjects that evade the (seemingly) simple literalism of training images labeled "portrait" or "cat."

For instance, one work in the series arose from a GAN fed a training set (or "corpus") labeled only with the open-ended classification "interpretations of dreams," with specific images including Freudian classics like hypodermic needles, windows, and puffy faces. When asked to produce its own contribution from what it had reviewed, the GAN created a work that Paglen titled *False Teeth (Corpus: Interpretations of Dreams)* (2018), a creepy, pseudo-figurative image dominated by bloody reds, gummy pinks, and a fragmentary central form evoking dental implants.

Trevor Paglen's *False Teeth (Corpus: Interpretations of Dreams)* "Adversarially Evolved Hallucination" (2017). ©Trevor Paglen, courtesy of Metro Pictures.

In an interview this summer, Paglen compared working with AI to the way Sol LeWitt worked with installers for his *Wall Drawings*. In both cases, the job is to construct a set of rules descriptive enough to guide assistants to a general formal output that also reveals the open spaces and breakdowns within those rules.

The difference is that LeWitt's assistants were other humans with interpretive capabilities, making judgment calls when necessary. Paglen's are neural networks

with deliberately complicated data sets, visualizing the human opinions embedded in their training. Instead of showing how independent and advanced GANs are, as Obvious's "purist" work purports to do, Paglen's work shows that GANs are dependent, limited, and hugely vulnerable to perpetuating the hidden biases held by the developers and researchers guiding them.

Artists' horizons expand further in light of the fact that GANs are not even the only type of AI infrastructure available. Theo Triantafyllidis's practice, for example, avoids neural networks in favor of the types of decision-tree-based AI that powers non-player characters in video games—technology that he has put to use in fixed-screen simulations and even virtual reality works. Other artists, such as Lynn Hershman Leeson, Martine Syms, and Martine Rothblatt, have created pieces involving chatbots, the communicative manifestations of natural language processing—the branch of AI that powers familiar digital "assistants" like Siri and Alexa.

Ultimately, the mini-controversy over Christie's and Obvious offers a window onto an important lesson here in the early life of AI art: To identify truly path-breaking work, we would do better to stop asking where the boundary line lies between human artists' agency and that of AI toolsets, and instead start asking whether human artists are using AI to plumb greater conceptual and aesthetic depths than researchers or coders. Anything less and we're simply elevating a technology demo to the status of enlightened studio practice.