AI Art and its Ethical Value

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia | Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science in Computer Science, School of Engineering

Jerry Gu

Spring 2025

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Advisor:

Sean Murray, Associate Professor of STS, Department of Engineering and Society

Introduction

Art has long been the subject of much discussion in the realm of philosophy. Questions of what art actually is in a philosophical sense has been the subject of much controversy, with some of the most famous philosophers throwing their hat in the ring such as Kant and Plato. However, with the advent of artificial intelligence models being able to reach new heights with innovations like stable diffusion, we have now come to live in the world where the thought experiment has become our new reality.

Whether machines can create art has begun to pull apart in a more salient way, our views on what constitutes art and what constitutes an artist. For instance, in this study of human reactions to art created by an AI or a human, they found that the participants were willing to see the final piece as an art piece, but were less willing to call the AI agent an artist (Coeckelbergh, 2016). The value of art in this sort of sense (especially when it comes to abstract art) comes from interpreting the artist's intentions, and when the art piece has no intentions due to being artificially generated, this seems to devalue the art. But if the value of the art is diminished somehow by being generated by AI, should we care about artists being replaced by AI? What are the engineering implications of an AI being even "good enough" to pass for not only visual arts, but the musical and performative arts as well?

Background

The primary concern that will be addressed in this paper is the implications of job displacement in the lens of an ethical framework. However, before we analyze that, it is important to gain some insight on people who actually are in the industry and the arts, and their opinions. Liz Mineo asked different instructors at Harvard University's art schools whether they

believed AI would act more as a threat, collaborator, or even a tool that can be used to fuel their imagination. The highlights of that discussion will be paraphrased below. Novelist and short-story writer Daphne Kalotay who is an instructor in the Creative Writing & Literature Program at the Harvard Extension School believes that "AI is a superb mimic and quick learner and might easily write strong works in recognizable modes…but – I think – will lack true insight and experience. Most at risk are commercial genres with easily recognizable insight and experience." (Mineo, 2023)

This seems to be in line with our understanding of AI as ultimately it all boils down to the dataset. The old adage goes for anything related to machine learning of "garbage in, garbage out." While not necessarily garbage, there are models where the users can easily tell which sources have had the most prominence in the training sets, to the point where styles can be outright replicated. This leaves commercialized and popular genres at most risk for upheaval. However, the upheaval was welcomed by Professor Matt Saunders in the Department of Art, Film and Visual Studies. He believes that "... we should be grateful to be challenged and knocked out of our habits and assumptions!" His worries mostly involve the social and ethical ramifications of AI art, rather than whether he feels truly threatened as an artist (Mineo, 2023).

This mindset is not shared as much in industry however. In the 2023 SAG-AFTRA Hollywood Strike, many artists in the industry felt that their livelihoods were being threatened if there was not any regulation. As such, there was a provision in the eventual signed agreement so that "parties acknowledge the importance of human performance in motion pictures and the potential impact on employment." (SAG-AFTRA, 2023). Overall, there are many mixed views of this even in the artist community, as artists working in Hollywood feel more threatened by AI. Note that although not every artist in the movie industry is not affiliated with SAG-AFTRA, their

organization is large enough that it certainly represents a sizable and important subset of artists in the industry. So while some artists may disagree that they may be entirely replaced by AI, there is a significant subset of artists concerned with consumer arts.

Literature Review

When referring to "art," this will refer to digital, visual forms of art unless specifically mentioned otherwise. We will also be primarily analyzing the implications of fully generated digital art and less about AI-generation tools that may or may not be used by artists. Art generated by AI versus art generated by humans has become more difficult to differentiate as time has gone on. However, some studies have shown that there is a difference, and perhaps the Turing Test has not been fully conquered, but simply made more difficult. In a study by Hong, et. al., human art was able to be rated consistently higher than AI art when asked to evaluate the specific aesthetic value. (Hong, 2019, 58:12). It is also true that artificial intelligence cannot demonstrate creativity nor can they embed meanings into paintings, whereas humans are able to do both of these things (Ch'ng, 2019, 59:15). Another study found that people tend to judge robot and human paintings similarly, but that people were less willing to consider robots as artists. (Mikalonytė & Kneer, 2022). However, a study claimed to have implemented a "Creative Adversarial Network" able to capture creativity to the point where they are indistinguishable to humans, and even consistently rated higher by humans (Elgammal et al., 2017).

While some innovations in the realm of computer science have claimed that computers are capable of creative reasoning as mentioned in Elgammel et al., the jury is still out in the realm of philosophy. We will assume that AI artists are not agential when they create art.

However, because this move may be controversial I will give a brief summary here of one argument to illustrate the contention between whether AIs themselves can be agents.

The Strong AI Hypothesis posits that a computer with a sophisticated enough set of inputs and outputs would be the same as humans, in that they actually understand the inputs and outputs in the same way as a human would (Dennett, 1991). Pushback against this is Searle's Chinese Room thought experiment which is summarized as follows: Imagine there is a room in which a person sits who does not know Chinese, but they have a book that perfectly guides them into taking Chinese inputs, and outputting intelligible Chinese characters. When the person is given a set of Chinese characters, they simply use the book to look up the syntactic rules and appropriate responses, and to the outside world, it is as if the room can have a perfect and fluent comprehension of the Chinese language. However, it would be incorrect to say that the man or the room has any comprehension or understanding of Chinese at all (Searle, 1986). Searle believes that the artificial intelligence is acting in this way, that they are merely executing a perfect set of instructions, and have no understanding of the semantic meanings of both inputs and outputs. Just as the person inside the room has no understanding of the Chinese character inputs, nor the outputs that the instructions tell them to give (Searle, 1986).

One reply is The Other Mind's reply. If an alien species comes to earth and observes Chinese speakers, how can they know that humans can speak Chinese and understand the semantic meanings of their own words? It is only by their behavior and the evidence that they take in Chinese inputs and return appropriate Chinese outputs. The only evidence we have of other humans having cognition is by their behavior. Because computers can have this same behavior, we must attribute cognition to them as well (Searle, 1986).

The question of whether artificial intelligence can have minds in the same sense that human beings have minds is out of the scope of this paper, but still a controversial topic. Moving forward, we will assume against the Strong AI Hypothesis in that AI has no understanding nor creative process embedded in the works they generate.

Conceptual Framework

We will analyze the impacts of artificial intelligence through the lens of utilitarianism to determine the ethical merits and limits of artificial intelligence. Utilitarianism is the idea that there exists some measure of utility like "happiness," and we ultimately ought to maximize that utility. Imagine each person has an allocation of happiness points that they may accumulate across their entire life. A utilitarian would say whatever decision we make, we ought to aggregate everyone's happiness points and make the decision that maximizes happiness the most for everyone, *in total* (Maxwell & Driver, 2009). This means that it does not matter who stands to lose utility so long as the overall population utility is a net positive. We will analyze through the lens of utilitarianism whether we can say with confidence that AI and namely large tech companies are giving a net benefit to society as a whole, and explore whether artists may or may not just be getting the short end of the stick when it comes to societal gains or losses.

We will juxtapose utilitarianism with Rawls's idea of Social Justice theory and whether artists' complaints about job displacement in the system of society are justified or not. This question about justice is mostly asking "what is the fairest way to distribute the benefits and burdens of society." Rawls believes in two main principles:

> 1. Each person participating in a practice or affected by it has an equal right to the extensive liberty leading to life and liberty for all." (Rawls, 1999)

2. Inequalities are indefensible unless they are reasonably expected to provide benefits for everyone, and have opportunities for all. (Rawls, 1999)

The difference between these two lies in the "baseline" that Rawls provides in these main principles. He believes that it would be wrong to anything that would infringe upon these two principles, no matter how much utility it may provide. He uses the illustration of the "Veil of Ignorance" for this point. In this illustration you are put behind a veil of ignorance where you are ignorant of your current social status, religion, etc. If you were to then construct the society's rules from there, you would be more cognizant of the society as a whole. Imagine you are deciding on whether or not your society should have laws protecting religious freedom. If you go behind the veil of ignorance, where you do not know what religion you hold and it is effectively randomized, you would want religious freedom (Rawls, 1999).

Given these two philosophical frameworks, we will analyze the value of AI art within these two frameworks to determine whether the utilitarian would believe that AI art produces more value and whether AI art aligns with Rawls' theory of justice.

Analysis – Utilitarianism

We will start first with utilitarianism. Given the literature, it seems that there is still a gap between artists and AI, though that gap is closing. For the utilitarian, AI seems to be a benefit for society as a whole (Soroushian, 2024), as it seems to allow us to be more productive when it comes to many different tasks, such as software development (Peng et al., 2023), writing tasks, (Noy & Zhang, 2023) and even taxi driving (Kanazawa et al., 2022). For these people, it seems as if AI can be a huge boon to their daily lives, helping them be a more productive member of

society. However, when the AI becomes so good at its task that it can completely replace the person involved, it seems to suddenly have an extremely negative impact on that person's life who has been displaced. But what about the case of AI art? Perhaps if we were evaluating the comparisons between autonomous vehicles, we could look at the hard numbers and see whether the autonomous vehicle would cause more crashes than humans. For instance, if we fully replaced all vehicles on the road with autonomous vehicles, we could measure the number of crashes compared to when there were no autonomous vehicles on the road. The value of autonomous vehicles can be measured in some meaningful way, but art seems to be more difficult to measure. As mentioned in the study before, we can measure the value of art as perceived by other people, but there seems to be a different quality to it.

Without semantic understanding and agency, AI simply cannot use art as a method of communication. Take for instance, the man in the Chinese room once more. Imagine we asked the man not only to give an appropriate response in Chinese, but we asked the man to give us a poem. While an appropriate poem may appear, the truly deep meaning behind the poem may be completely lost. I believe that this poem would be worth less than a poem that is taken from the depths of a poet's creativity and emotions. Both poems are created by people, but in the former case, there is an important aspect that is missing. In fact, I argue that the former provides less utility than the latter. The knowledge itself of the origins of the poetry and the circumstances behind it add some amount of utility to its creation, even if the words themselves may be indistinguishable on the surface. Take for instance, the famous poem *O Captain! My Captain!* by Walt Whitman. While a brilliant poem in itself, the genre of the poem in its form is an extended metaphor about the death of Abraham Lincoln. The added context of when it was written, and the emotions of the poet upon receiving the news of Lincoln's assassination add value to the

meaning behind the words. This is something that simply cannot be replicated by the man in the Chinese room.

Rather than written arts like poetry, visual arts are the same. AI art does not serve as a form of communication under our assumptions that AI is not a being that does not have any understanding of its creation.

A way to advocate for the overall net positive utility of AI art is if AI art has at least a modicum of utility, then the rapid ability to generate a large quantity of art quickly would generate more overall utility. If we could generate a seemingly infinite amount of smaller utilities, would we not achieve orders of magnitude more overall utility than the limited number of artists? If this was the case, the overall utility for art is still limited by those who can look at it. In other words, even if AI art can be generated at a rapid pace, there simply aren't enough people to appreciate it and give it the modicum of utility. The demand for art that we have currently can and has been satisfied by humans already, and to replace that demand with a lower quality art would reduce the overall utility in how art is appreciated.

A pushback I am more sympathetic to is that AI art can be used as a means to achieve something I will call "higher order art." A prominent current example of this would be art assets used in video game production, as the game distributor Steam already has required publishers to mark when AI art assets are used in development (Steamworks Development, 2024). While the AI art assets themselves may not be of the highest quality, they can contribute to create a game with a strong artistic value. This democratizes the ability to create higher order art, as no longer will realistic art assets be gated behind hiring artists. Now it is inexpensive to generate art assets through AI and let people create their own documentaries, movies, and videogames using those assets. So while the works of art themselves may not communicate as much utility, the use of

them as a means to communicate something on a higher order may have creative value that might not have been able to have been realized without AI art.

Analysis – Rawls

First let us address what I believe to be the most interesting question to ask, whether AI art falls under the "inequality [that] provides benefit for all." While many people complain about wealth, if that person has provided an extreme good to society, perhaps it is okay if those people became wealthy. So if somebody in the future creates a cure for cancer, I would be perfectly okay if that person became the richest person in the world, because they have ultimately provided an extraordinary benefit to society. Take Bill Gates, whose work in the computing space has made him extremely wealthy, and contributed positively to my life as a PC user. I have no problem with the astronomical inequality, because he has made an unfathomably large contribution to society that was ultimately rewarded by that wealth. His inventions in the computing space has allowed people to get PCs into their homes and we can be more productive as a society as a result. I argue that AI art does not fall into this category. Due to the differences in human art and AI art, I believe that there is little benefit to everyone that people can get.

The benefit to generating content that is lower in quality to save money exacerbates wealth inequality between artists and tech companies. Tech companies similar to Open AI monetize the generation of art (often by training their models on artists' work without permissions), which gives them an inequality in wealth. This inequality in wealth seems to give little opportunities to those who use that technology because the technology itself seems to give a lower quality than real artistry. Rawls emphasized that these developments ought to provide

benefits to everybody and opportunities for all. We are clearly eliminating opportunities for the artists themselves in the event of significant job displacement, but I will argue that we are providing limited benefits not only to artists, but to the population as a whole.

However, what if we assume that it is possible for further research of AI art to eventually reach a level at which it may provide equal utility to humans? Let us take the veil of ignorance, and we are deciding whether to make a decision on AI art research. The veil of ignorance strips away our current religious beliefs, societal status, etc. such that we are unaware of where we sit once the veil is taken away. So once we step behind the veil, we would not have any knowledge of whether we are the artist whose job security may be threatened or the CEO of OpenAI. We ought to choose the most equitable legislation, as the resulting inequality would not result in benefits for everyone, and opportunities for all according to Rawls's theory of justice.

I argue that further innovations in the AI art sphere do not align with Rawls's Theory of Justice because technological advances would lead to an inequality in society that does not provide benefits to all. Not only does it limit the benefits for a group of people, but it even affects their equal rights to life and liberty. Thus, the benefit that people may receive from this would be indefensible due to the infringement upon the artists' lives. Innovations in AI art research is not analogous to Bill Gates's innovations in personal computing because of this.

A pushback may be that innovations in AI art is analogous to the industrial revolution, in which artisans were replaced by machines. The artisans are artists in this case, and the industry is now just the tech corporations. While I agree that the industrial revolution did raise standards of living and had societal benefits, I believe the same thing cannot be said of AI art. While the industrial revolution lifted many out of poverty and raised their standards of living, I remain unconvinced that the same benefits could be provided by AI companies.

Another possible pushback may go as follows: Under my arguments of the difference in utility between AI generated artworks versus human created artworks, I should not need to worry about artist displacement as the market forces of capitalism would give rise to the works with the greatest utility. I also would not need to worry about there being a supply or demand issue, as I have also argued that the current supply of artists is enough to fit the demand of those who need art. So this cannot be the case of a high quality artisan chairmaker being put out of business by a lower quality, mass produced chair. I believe Rawls's theory of justice may then be applied to higher order works of art. While there may not be as much of a threat to artists creating art for pure aesthetic value, the same may not be able to be said for commercial art, or art being used to create a higher order piece. Video games or movies in the SAG-AFTRA case have real worries about worker replacement, for example (SAG-AFTRA, 2023). In these cases, if I concede that the AI art may ultimately provide more utility for a higher order work of art, it is unjust due to the innovation putting artists out of work.

Conclusion

AI art is a concern in today's world as it threatens to displace artists and replace them with computer generated content (SAG-AFTRA, 2023). Assuming that AI Agents have no semantic knowledge of the works that they create, I have examined the value of AI art through two different ethical lenses, utilitarianism and Rawls's Social Contract Theory. In utilitarianism, the main weakness of the utility of AI art is its inability to be more inherently more aesthetically pleasing to humans as an art form. Because of this disparity, AI art ultimately has less utility, and its use to replace artists would result in a net negative utility due to the loss in aesthetic value. The weakness of AI art innovation is that its continuation threatens to be inequitable to artists,

thus making the inequality of benefit untenable according to Rawls's theory of justice. Any further innovation, should it close the gap, would be unjust. In these two frameworks, I have argued the progression of the visual arts would not fit into their frameworks of utility and justice.

References

Ch'ng, E. (2019, July 3). Art by Computing Machinery: Is Machine Art Acceptable in the Artworld? *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, *15*(2s), 1-17. https://doi.org/10.1145/3326338

Coeckelbergh, M. (2016, September 24). Can Machines Create Art? *Philosophy & Technology*, *30*, 285-303. https://doi.org/10.1007/s13347-016-0231-5

Dennett, D. C. (1991). *Consciousness Explained*. Little, Brown and Company.

Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017, June 21). *[1706.07068] CAN: Creative Adversarial Networks, Generating "Art" by Learning About Styles and Deviating from Style Norms*. arXiv. Retrieved March 4, 2025, from https://arxiv.org/abs/1706.07068

Hong, J.-W. (2019, July 19). Artificial Intelligence, Artists, and Art: Attitudes Toward Artwork Produced by Humans vs. Artificial Intelligence. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, *15*(2s), 1-16. https://doi.org/10.1145/3326337

Kanazawa, K., Kawaguchi, D., Shigeoka, H., & Watanabe, Y. (2022, October). *NBER WORKING PAPER SERIES AI, SKILL, AND PRODUCTIVITY: THE CASE OF TAXI DRIVERS Kyogo Kanazawa Daiji Kawaguchi Hitoshi Shigeoka*. National Bureau of

Economic Research. Retrieved March 4, 2025, from https://www.nber.org/system/files/working_papers/w30612/w30612.pdf

Maxwell, J., & Driver, J. (2009, March 27). *The History of Utilitarianism (Stanford Encyclopedia of Philosophy)*. Stanford Encyclopedia of Philosophy. Retrieved April 18, 2025, from https://plato.stanford.edu/entries/utilitarianism-history/

Mikalonytė, E. S. M., & Kneer, M. (2022, September 08). Can Artificial Intelligence Make Art?: Folk Intuitions as to whether AI-driven Robots Can Be Viewed as Artists and Produce Art. *ACM Transactions on Human-Robot Interaction (THRI)*, *11*(4), 1-19. https://doi.org/10.1145/353087

Mineo, L. (2023, August 15). *Is art generated by artificial intelligence real art?* Harvard Gazette. Retrieved October 26, 2024, from https://news.harvard.edu/gazette/story/2023/08/is-art-generated-by-artificial-intelligence-real-art/

Noy, S., & Zhang, W. (2023, July 13). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, *381*, 187-192. 10.1126/science.adh2586

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023, February 13). *The Impact of AI on Developer Productivity: Evidence from GitHub Copilot arXiv:2302.06590v1 [cs.SE] 13 Feb 2023*. arXiv. Retrieved March 4, 2025, from https://arxiv.org/pdf/2302.06590

Rawls, J. (1999). *A Theory of Justice*. Harvard University Press.

SAG-AFTRA. (2023, November 16). *TV/ THEATRICAL CONTRACTS 2023*. SAG-AFTRA. Retrieved November 9, 2024, from https://www.sagaftra.org/sites/default/files/sa_documents/TV-Theatrical_23_Summary_Agreement_Final.pdf

Searle, J. R. (1986). *Minds, Brains and Science*. Harvard University Press.

Soroushian, J. (2024, July 11). *Is AI Making the Workforce More Productive?* Bipartisan Policy Center. Retrieved March 4, 2025, from https://bipartisanpolicy.org/blog/is-ai-making-the-workforce-more-productive/

Steamworks Development. (2024, January 9). *AI Content on Steam*. Steam Community. Retrieved March 30, 2025, from https://steamcommunity.com/groups/steamworks/announcements/detail/3862463747997849619