A taxonomy of concerns concerning neural art

Vinay Uday Prabhu
UnifyID
vinay@unify.id

Isiain X
Hear Sound Change Records
gunnagoesglobal@gmail.com

Abstract

AI generated art or neural art is on the cusp of having a profound, and perhaps indelible impact on the very intuition of art. Whether this current phase we are in, be judged as a watershed moment for creativity or the exact moment we carelessly opened the Pandora's box will be decided on how we carefully navigate certain concerns in the coming months. In this paper, we present a taxonomy of such concerns spanning the blood-diamond effect emerging from unethically sourced datasets, the emergence of Non-Fungible Tokens, downstream effects of gender and racial biases, gate-keeping antics in the AI-neural art community and the computational conquest of gender.

1. Introduction

One can argue that the field of generative art has always evoked mixed emotions amongst the puritans and critics alike. Works such as [5] have challenged the `is-it-really-art?` toned diatribes by dismantling misconceptions such as the lack of romantic indulgence in the enterprise of chance, accidental discovery or this imperceptible facet of spontaneity and that *"artist has zero control and the autonomous machine is randomly generating the designs*. Further in [18], the authors draw out examples from from paleolithic ornamental art, hydraulically-activated automata (Hero 1st C AD) of ancient Rome, Islamic art of the 9th century and Renaissance clockwork figures as examples of works with generative processes as their basis, that challenge preconceived notions that all generative art is computer generated. With regards to computer generative art, [5] also highlights the pioneering efforts of women artists in 1960s led by the likes of Vera Molnar, Lillian Schwartz, Grace Hertlein and Muriel Cooper, whose contributions are often under-emphasized, which when combined with the current sordid gender ratios in computer tech industry, sometimes results in a *tech-bro tint* to the field.

The latest avatar of generative art, that is AI-generated neural art is a different beast altogether. Specifically, the kind that is *trained* on large scale theft of human-generated art and biased large-scaled computer vision datasets on large scale hardware inaccessible to researchers outside of a few industry labs whilst burning consuming vast amounts of energy.

In this paper, we present our nascent attempts towards constructing a taxonomy of some of these concerns. We hope that this will serve as a basis for further critique, examination and betterment in the neural art community.

2. The continued Blood-diamond effect in art: Inheriting biases

As covered in [9], neural-art produced with unethically sourced (and/or biased) large scale datasets inherit a blood diamond effect in them. As indicated, this critique resonates far beyond the specific polemics of any specific large scale vision dataset and refers to the wanton culture that permeates the general computer vision domain as we know it today. Hence, it comes as no surprise that recent models trained on far larger and more opaque datasets (See [29, 30]) have not just inherited these ills but also amplified them. Adding to the woes emanating from biases in images, is the latest trend of inculcation of textual information gleaned from the internet, that brings in a new vector of textual biases in to the mix.

(a) Gender bias: In Figure 1, we see the images that were produced in response to the textual prompts (found in the title of these images) using the `Aleph-Image: CLIPxDAll-E` project[1]. This profession-to-gender associative anchoring we observed in this image is yet another instance of the downstream effects of training on large dumps of data gleaned from the internet and ignoring all of the critical literature such as [24, 8] that has disentangled the cause and the effects of these androcentric biases.

(a) Racial bias: In order to query the downstream racial bias aspect, using the model in [29], we framed a zero-shot binary classification problem with the constituent classes being `an artist` or `a thug`. This partic-

[1]Built using [29, 30]. Code: https://colab.research.google.com/drive/1Q-TbYvASMPRMXCOQjkxxf72CXYjR_8Vp?usp=sharing#scrollTo=BFsCy7jOn5cH

Figure 1. Downstream effects: Gender bias

ular framing in the context of black bodies is specifically pertinent and has historical roots as detailed in [10] and bears added sub-layering from the following words of LA times reporter Frank Williams in his article titled `Teenagers Grapple With Meaning of Shakur's Death` [32]: *"Now, as Shakur's record sales soar in the wake of his drive-by shooting death last month off the Las Vegas Strip–and the media debate his legacy as **an artist or thug**"*. When we tested across our collaborators (self-identified black artists from the bay area) we saw an astonishingly high rate of classification in favor of the `thug` class. In order to ascertain if this was merely sampling bias at play, we stress tested this observation across a dataset [26] that putatively accounts for variations in factors such as face angle, saturation, size, resolution, lighting conditions, facial expressions, clothing and eye gaze. The results are presented in Figure 2. As seen, faces belonging to the race: `Black` elicited far higher $Pr(thug)$. This are dire signs as we have already seen the emergence of large scale embrace of this model amongst the neural art community[2].

At this juncture, we'd like to declare and emphasize that the experimenter was a self-identified black artist and activist, and that we consciously decided to share only the statistical summary-results and not the individual images.

3. The rise and rise of NFT

The NFT *revolution* looms large over the horizon of AI art. As AI-artists navigate the difficult by-lanes of extractive capitalism they have to grapple with cost of environmental impact vis-a-vis their innate aspirations towards financial viability and ability to actually eke out a decent living. Works such as [14, 12, 19] have amply explored the specifics of the environmental cost measured in terms of CO_2 and metrics such as *Ethereum Energy Consumption Index*. In Table 1, we present the latest estimates of Cryptoart-footprint as of March 22, 2021 sourced from `https://github.com/kylemcdonald/cryptoart-footprint`. Similarly, initiatives such as `https://carbon.fyi/` (To estimate the footprint for a specific Ethereum wallet or contract) and `http://cryptoart.wtf/` (To estimate the footprint of a specific artwork) have emerged that educate the artists about the downstream contributions towards climate change, that has been demonstrated to disproportionately impact minorities and indigenous peoples [6]. On the other hand, AI-artists and Afrofuturists such as Nettrice Gaskins[3] have explored the possibility of NFT-platforms such as *hic et nunc*[4] that operates on the more *"eco-friendly"* Tezos blockchain (See [4]). These initiatives are to be considered in juxtaposition of the historical injustices being meted out to BIPOC artists. As stated in [21], *"For BIPOC/LGBTQ/Womxn/Disabled artists who have been intersectionally shut out of the Official Art Market, NFTs have the potential to help them/us make a living or supplement our incomes. The blockchain market removes the middle person or gatekeeper and empowers artists to use emerging technologies and networks in new or culturally relevant ways."*

This thought paves the way for our next section that deals

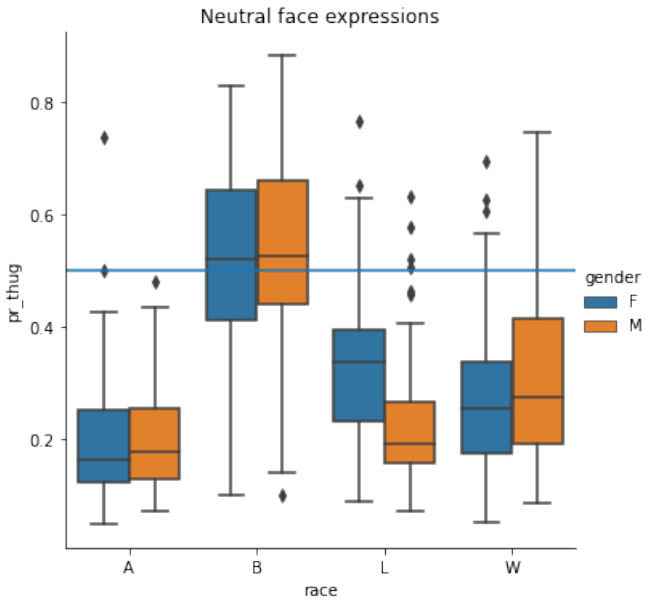

Figure 2. Results of the artist and thug experiment with the CLIP model

[2]See `https://www.reddit.com/r/MachineLearning/comments/ldc6oc/p_list_of_sitesprogramsprojects_that_use_openais/`

[3]`https://www.nettricegaskins.com/about`

[4]`https://www.hicetnunc.xyz/about`

with stealth of human-generated-art that is then siphoned off as web-crawled datasets to train AI generative models.

Name	Gas	Transactions	kgCO2
OpenSea	173,618,258,189	804,476	60,962,694
Nifty Gateway	31,925,160,353	116,531	13,381,972
Rarible	27,700,791,937	216,407	10,302,402
Makersplace	21,679,291,384	69,165	6,222,148
SuperRare	15,950,212,376	171,967	5,031,957
Foundation	8,169,344,070	51,375	4,103,883
Known Origin	4,677,980,129	19,181	1,468,752
Zora	1,871,937,838	7,137	868,972
Async	1,502,889,319	15,081	396,560

Table 1. Estimates of Cryptoart-footprint as of March 22, 2021 (Source:https://github.com/kylemcdonald/cryptoart-footprint)

4. Stealth, Appropriation, Crawling and ghost-work

Referring to Open-AI's dataset curation process that produced the Jukebox generative model for music [17] (1.2 million songs obtained by crawling the web), a Google magenta team researcher[5] recently exclaimed that *"Music datasets (with rights) are *tiny* compared to image and language datasets. I'm pro free use of music, but it feels disingenuous to use an artist's data, not include them in the process, and then train your model to specifically generate "in the style" of that same artist. OpenAI had a much more thoughtful take on the ethical implications of GPT2 and I think it's a shame they didn't take the same level of consideration here. Especially with their ability to create precedent."* This statement unpacks some of the machinations used to appropriate and alienate BIPOC art creators from wealth creation. We insist that the term stealth is justified as the artists whose sweat equity resulted in the production of the art are explicitly alienated from any form of compensation. In this regard, we insist that we are not taking a parochial Luddite stance petitioning for abolition of generative music. In fact, we opine that we do need to pay heed to efforts such as Creative commons curated datasets that are shared for humanity with the explicit consent of the musicians and reinforced with robust license-mechanisms being laid to ensure downstream revenue sharing. Furthermore, the community has come up with a whole plethora of flexible copyright licenses and portals such as *ccMixter, Freesound, SectionZ, Jamendo and Opsound*[6] that are veritably better options against cavalier crawling practices and blatant stealing.

Associated with this is also the problem of ghost-work[23] extracted from platforms such as Amazon Mechanical Turk. A recent article [28] sheds some light into the curation process behind the *ArtEmis* dataset [2] that includes the emotional impact of more than 50,000 artworks from the 15th century to the present and one that entailed the use of more than 6,000 *microworkers*[7].The work covered two important facets that are important to unpack in the context of this critique. Firstly, there is the issue of sanctity of the labeling themselves. One of the *microworkers* verbatim stated that *"To be honest, a lot of it felt very forced. There were many images that were just formless blobs or of basic objects. It was quite a stretch to come up with emotions and explanations at times."* Secondly, it emerged that there was the geographical constraint that the *microworker* had to be "located in Australia, USA, Great Britain, or Canada" that obvious brings in geographical bias. This pro-global-north bias, we argue is symptomatic of a deeper ill that manifests itself both at a macro-level as well as at a more fine-grained level that in turn impacts the academic AI-art community directly. This leads us to the issue of cliquishness that we address in the following section.

5. Cliquishness: Committed to reproduce historical classist morass are we?

Homophily is a deeply human construct and the associated enterprise of community building, when done well, can be wonderfully delightful in creating a safe and nourishing space for exchange of ideas. But human history has taught us time and again that fetishizing over snobbery and hierarchy can quickly precipitate into cliquish echo-chambers replete with gatekeeping histrionics that eventually poisons the community. There exists a mountain of evidence that the bigger source of the perceived *worthiness* of art in the marketplace is the network effect and not the make-believe *innate chutzpah* or the creative wizardry. Beginning with the *artsy* monologue literally titled *Artists Become Famous through Their Friends, Not the Originality of Their Work* [25] to the work on *Fame as an illusion of creativity: Evidence from the pioneers of abstract art* [27], we have seen strong academically stress-tested evidence of how shrewd social-networking either at high-society events or the *factories*[8] have been the key to an artists' emergence on the scene, all under the facade of creative spirit (Also see [20, 15, 16, 33] for related iconoclastic rendering of the fallacies of the artistic genius). To analyze the situation with AI-art today, we begin by momentarily donning the role of a Machiavellian deviant whose goal is to institute classism within AI-art community and create a schism between the

[5] https://twitter.com/jesseengel/status/1256314515114745857?s=20

[6] https://creativecommons.org/about/program-areas/arts-culture/arts-culture-resources/legalmusicforremixing/

[7] The italics in this context are used to both preserve the terminology used in the reference [28] and to also highlight the troublesome jargon that is prevalent in the domain

[8] https://en.wikipedia.org/wiki/The_Factory

privileged art-creator-superstars (or, ahem, Godfathers[9] and the subaltern denizens. With regards to workshops such as these, two important tools that the deviant would implement to further the goal would be:

a) Single-blind review process: Create an information-asymmetric marketplace of ideas by means of instituting Single-blind reviews where the reviewers have the privilege of knowing precisely who the author(s) but not vice-versa. Other art-forms poisoned by the gate-keeping are slowly unshackling themselves while we in the AI-art community seem weirdly hell bent on re-birthing these historical blunders. In [22], the authors demonstrated how introducing blind auditions in orchestras where the performer performs behind a screen did markedly increase the probability that a woman would be advanced and hired. Further, in the context of academic peer reviewed articles, works such as [11] have demonstrated how instituting Double-blind review was a watershed moment towards increased representation of female authors. Yet, even in the presence of double-blind reviewed workshops at ICML, NEURIPS, ICLR etc, AI-neural workshops continue to be notoriously adamant on this facet which does not bode well for the community.

b) Black-box review process: Given the non-archival nature of these workshops, a lot of us in the community are not as interested in whether or not our work gets accepted but are more interested in knowing the flavor and nature of critiques that the submission attracted. So we are left grappling with a situation where not only is the review process single-blind but the reviews are not made available to the submitter either.

If democratization is indeed the goal, then, we really need to learn from the toxicities that have polluted the *other* art-forms covered in the work cited above.

6. Computational conquest of gender

It would be an understand to say that the proverbial *artistic license* is a double-edged sword that comes with a tremendous set of responsibilities. In order to create the necessary tools to cater to the flights of imaginations of artists, the tool-makers are forced to navigate a difficult terrain laced with ethical landmines as the very degrees of freedom they want to incorporate can also be used (and worse) in an automated way to generate bigotry, hatred and challenge the very institution of truth. Today, photoshopped has in fact become a de facto verb. If we thought photoshop was bad, these creative tools hopped up on AI-steroids can generate harm at scale hitherto impossible. In this section, we'll consider specific issues pertaining to race/ethnicity and gender. Our fear is that the tools that are being developed to fuel the flights of fantasy of artists can and are being used to co-opt and game the DEI movement. (See appendix-A) for a detailed explanation.

[9] https://twitter.com/techreview/status/1323580930552958978?s=20

7. Concluding thoughts

As with any art-form, we argue that the political projections of its very existence is inevitable. To this end, we refer to `A Context for Complexism: Between Neoliberal Social Thought and Algorithmic Art` [13], where the author quite frankly situates complexist art as: *In the final analysis, complexist art, as a subset of generative art, offers a deeply troubling aestheticisation of the market mechanism as it was uniquely theorised by the foremost architect of neoliberal policy and thought*. Given that the birth of this form of AI/Neural art is being championed from the underbelly of Big-AI corporations that have learnt and learnt well from their Big Tobacco, Big-Pharma and Big-Oil counterparts [1], we in the community have to reflect on how this moment will be judged by the generations to come. The computational agenda to conquer and colonize the last few remaining pristine bastions of glorious and creative human fuzziness that thrives inside the sacrosanct temples of gender and aestheticism is well and truly underway. We conclude with the following dark conjecture that we hope will seed important conversations to be had in the community, and to be had now.

7.1. AGESaaS: Any-gender-expression-stockphoto-as-a-service. A stark prognosis

Conjecture: By 2023, there will be at least one VC funded startup that, under the guise of performative activism, will start peddling realistic stock-photos claiming to cover the entire spectrum of gender expressions for less than $0.99.

References

[1] Mohamed Abdalla and Moustafa Abdalla. The grey hoodie project: Big tobacco, big tech, and the threat on academic integrity. *arXiv preprint arXiv:2009.13676*, 2020. 4

[2] Panos Achlioptas, Maks Ovsjanikov, Kilichbek Haydarov, Mohamed Elhoseiny, and Leonidas Guibas. Artemis: Affective language for visual art. *arXiv preprint arXiv:2101.07396*, 2021. 3

[3] Yves Saint James Aquino and Wendy Rogers. *" Big Eye Surgery": Understanding the Ethical Implications of Medicalising Asian Features in Cosmetic Surgery*. PhD thesis, Macquarie University, Faculty of Arts, Department of Philosophy, 2015. 7

[4] Wagner James Au. New world notes: Why afrofuturist nettrice gaskins sells her deep dream art as nfts on the greener tezos blockchain. https://nwn.blogs.com/nwn/2021/03/nettrice-gaskins-afrofuturist-nft-tezos-blockchain-tupac-deep-dream-art.html, Mar 2021. (Accessed on 03/26/2021). 2

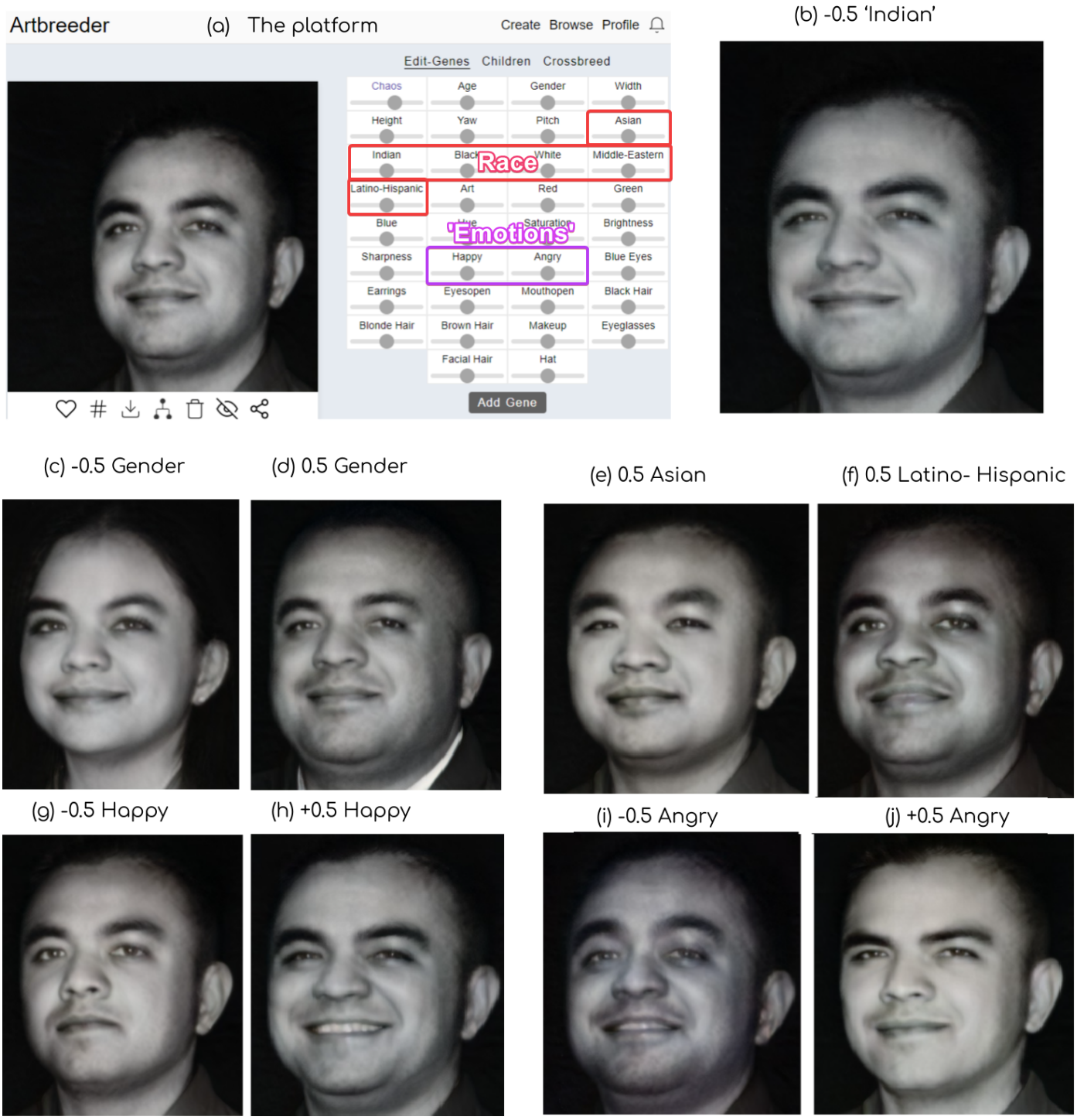

Figure 3. Artistic license: The questions of gender, ethnicity and race

[5] Jason Bailey. Why love generative art? — artnome. https://www.artnome.com/news/2018/8/8/why-love-generative-art, August 2018. (Accessed on 03/24/2021). 1

[6] Rachel Baird. The impact of climate change on minorities and indigenous peoples. *Briefing). Minority Rights Group International: London*, 2008. 2

[7] Taunya Lovell Banks. Colorism among south asians: Title

vii and skin tone discrimination. *Wash. U. Global Stud. L. Rev.*, 14:665, 2015. 7

[8] CT Begeny, MK Ryan, CA Moss-Racusin, and G Ravetz. In some professions, women have become well represented, yet gender bias persists—perpetuated by those who think it is not happening. *Science Advances*, 6(26):eaba7814, 2020. 1

[9] Abeba Birhane and Vinay Uday Prabhu. Large image datasets: A pyrrhic win for computer vision? In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1537–1547, 2021. 1

[10] Keia Janese Bragg. We're not thugs and rappers: An examination of african american male athletes' perceptions of the media. *Master of Science thesis*, 2010. 2

[11] Amber E Budden, Tom Tregenza, Lonnie W Aarssen, Julia Koricheva, Roosa Leimu, and Christopher J Lortie. Double-blind review favours increased representation of female authors. *Trends in ecology & evolution*, 23(1):4–6, 2008. 4

[12] Justine Calma. The climate controversy swirling around nfts - the verge. https://www.theverge.com/2021/3/15/22328203/nft-cryptoart-ethereum-blockchain-climate-change, Mar 2021. (Accessed on 03/26/2021). 2

[13] Ricky D Crano. A context for complexism: Between neoliberal social thought and algorithmic art. *Open Cultural Studies*, 2(1):341–352, 2018. 4

[14] Alejandro de la Garza. Nfts art's hidden environmental cost — time. https://time.com/5947911/nft-environmental-toll/, Mar 2021. (Accessed on 03/26/2021). 2

[15] Dakshina De Silva, Georgia Kosmopoulou, Rachel Pownall, and Robert Press. Trading patterns and the "death effect" on artwork prices. 2020. 3

[16] Dakshina G De Silva, Marina Gertsberg, Georgia Kosmopoulou, and Rachel AJ Pownall. Dealer networks in the world of art. *Available at SSRN 3050848*, 2017. 3

[17] Prafulla Dhariwal, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford, and Ilya Sutskever. Jukebox: A generative model for music. *arXiv preprint arXiv:2005.00341*, 2020. 3

[18] Alan Dorin, Jonathan McCabe, Jon McCormack, Gordon Monro, and Mitchell Whitelaw. A framework for understanding generative art. *Digital Creativity*, 23(3-4):239–259, 2012. 1

[19] Sonia Elks. What are nfts, and why is digital art bad for the climate? — world economic forum. https://www.weforum.org/agenda/2021/03/nfts-and-the-environmental-cost-of-owning-digital-art/, Mar 2021. (Accessed on 03/26/2021). 2

[20] Samuel P Fraiberger, Roberta Sinatra, Magnus Resch, Christoph Riedl, and Albert-László Barabási. Quantifying reputation and success in art. *Science*, 362(6416):825–829, 2018. 3

[21] Nettrice Gaskins. Art as cryptocurrency or how to make 1k+ in a day. https://nettricegaskins.medium.com/art-as-cryptocurrency-or-how-to-make-1k-in-a-day-799a2a89acd7, Mar 2021. (Accessed on 03/26/2021). 2

[22] Claudia Goldin and Cecilia Rouse. Orchestrating impartiality: The impact of" blind" auditions on female musicians. *American economic review*, 90(4):715–741, 2000. 4

[23] Mary L Gray and Siddharth Suri. *Ghost work: How to stop Silicon Valley from building a new global underclass*. Eamon Dolan Books, 2019. 3

[24] Leah R Halper, Colleen M Cowgill, and Kimberly Rios. Gender bias in caregiving professions: The role of perceived warmth. *Journal of Applied Social Psychology*, 49(9):549–562, 2019. 1

[25] Casey Lesser. Artists become famous through their friends, not the originality of their work. https://www.artsy.net/article/artsy-editorial-artists-famous-friends-originality-work, Feb 2019. (Accessed on 03/26/2021). 3

[26] Debbie S Ma, Joshua Correll, and Bernd Wittenbrink. The chicago face database: A free stimulus set of faces and norming data. *Behavior research methods*, 47(4):1122–1135, 2015. 2

[27] Banerjee Mitali and Paul L Ingram. Fame as an illusion of creativity: Evidence from the pioneers of abstract art. *HEC Paris Research Paper No. SPE-2018-1305, Columbia Business School Research Paper*, (18-74), 2018. 3

[28] Aliide Naylor. Underpaid workers are being forced to train biased ai on mechanical turk. https://www.vice.com/en/article/88apnv/underpaid-workers-are-being-forced-to-train-biased-ai-on-mechanical-turk, Mar 2021. (Accessed on 03/26/2021). 3

[29] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. *arXiv preprint arXiv:2103.00020*, 2021. 1

[30] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. Zero-shot text-to-image generation. *arXiv preprint arXiv:2102.12092*, 2021. 1

[31] Ramya Srinivasan and Kanji Uchino. Biases in generative art: A causal look from the lens of art history. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 41–51, 2021. 7

[32] Frank.B. Williams. Teenagers grapple with meaning of shakur's death. https://www.latimes.com/archives/la-xpm-1996-10-07-me-51340-story.html, Oct 1996. (Accessed on 03/26/2021). 2

[33] Tamar Yogev and Thomas Grund. Network dynamics and market structure: The case of art fairs. *Sociological Focus*, 45(1):23–40, 2012. 3

Appendix A. Gender, emotion and race

Phrenology, physiognomy and computational eugenics are undergoing a silent rebirth emboldened by the rise of

Computer Vision aided face psychonomics. Deep learning aided models that claim to perform *High accuracy* Gender classification (binarized), emotion classification and ethnicity classification trained on carelessly curated datasets that promulgate prototyping and reductionist caricatures are all unanimously frowned upon in the AI-ethics community. But these very datasets and models get a second lease of life under the banner of generative models and latent space exploration adventures of these generative models in the context of creativity and entertainment. While we acknowledge the viewpoint of empowering the flights of fantasy of digital artists, we'd like to point out the obvious threats that emanate from appropriation of these very tools. Besides the obvious dangers of gender-swapping

In Figure 3, we see a gallery of the resultant faces of an author of this paper when subjected to artistic face manipulations by re-weighting the *genes* (or latent manipulation vectors) on the popular art-platform artbreeder [] (formerly GANbreeder). The pixel space face embedding of the uploaded image is shown in sub-figure 3(a). Upon introducing a weight of -0.5 with regards to the `Indian` gene, we obtained sub-figure 3(b) that had a palpable skin-lightening effect, that has particularly stark ramifications given the skin-tone bias and colorism that the society is grappling with [7]! Similarly, the platform affords a gender control gene (more positive implying more *masculine*). Upon introducing a weight of -0.5 and $+0.5$ with regards to the `Gender` gene, we obtain 3(c) and 3(d) respectively. Recent work on *Biases in Generative Art— A Causal Look from the Lens of Art History* [31] has highlighted how current generative art methods fell short in terms of framing effect bias, dataset bias, selection bias, confounding bias, and transportability (gender) bias using structural causal models. Sub-figures 3(e) and (f) represent images 0.5 `Asian` and $+0.5$ `Latino-Hispanic` genes, both of which in our opinion do map to stereotypical anchor-imagery promulgated against these communities. We refer the reader to the thesis titled "Big eye surgery: Understanding the ethical implications of medicalising Asian features in cosmetic surgery" [3] that explores the problematic issues surrounding this parochial caricaturization and the growing demand for Epicanthoplasty.