
The blood diamond effect in neural art: On ethically troublesome images of the imagenet dataset

Vinay Uday Prabhu
UnifyID AI Labs
vinay@unify.id

Abstract

The main goal of this short paper is to inform the neural art community at large on the ethical ramifications of using models trained on the imagenet dataset, or using seed images from classes 445 -n02892767- ['bikini, two-piece'] and 459-n02837789- ['brassiere, bra, bandeau'] of the same. We discovered that many of the images belong to these classes were verifiably pornographic, shot in a non-consensual setting, voyeuristic and also entailed underage nudity. Akin to the *ivory carving-illegal poaching* and *diamond jewelry art-blood diamond* nexuses, we posit there is a similar moral conundrum at play here and would like to instigate a conversation amongst the neural artists in the community.

1 Introduction

The emergence of tools such as BigGAN [1] and GAN-breeder [2] has ushered in an exciting new flavor of generative digital art [3], generated using deep neural networks (See [4] for a survey). A cursory search on twitter¹ reveals hundreds of interesting art-works created using BigGANs. There are many detailed blog-posts² on generating neural art by beginning with seed images and performing nifty experiments in the latent space of BigGANs. At the point of authoring this paper, (8 September 2019, 4:54 PM PST), users on the *GanBreeder* app³ had *discovered* 49652500 images. Further, *Christie's*, the British auction house behemoth, recently hailed the selling of the neural network generated *Portrait of Edmond Belamy* for an incredible \$432,500 as *signalling the arrival of AI art on the world auction stage*[5].

Given the rapid growth of this field, we believe this to be the right time to have a conversation about a particularly dark ethical consequence of using such frameworks that entail models trained on the ImageNet dataset which has many images that are pornographic, non-consensual, voyeuristic and also entail underage nudity. We argue that this lack of consent in the seed images used to train the models trickles down to the final artform in a way similar to the blood-diamond syndrome in jewelry art [6].

An example: Consider the neural art image in Fig 1 we generated using the *GanBreeder* app. On first appearance, it is not very evident as to what the constituent *seed* classes are that went into the creation of this neural artwork image. When we solicited volunteers online to critique the artwork (See the collection of responses (Table 2) in the supplementary material), none had an inkling regarding a rather sinister trickle down effect at play here. As it turns out, we craftily generated this image using hand-picked specific instances of *children* images emanating from what we will showcase are two problematic *seed* image classes: Bikini and Brassiere. More specifically, for this particular image, we set the *Gene weights* to be: [*Bikini*: 42.35, *Brassiere*: 31.66, *Comic Book* - 84.84]. We'd like

¹<https://twitter.com/hashtag/biggan?lang=en>

²<https://thegradients.pub/bigganex-a-dive-into-the-latent-space-of-biggan/>

³<https://ganbreeder.app>

to strongly emphasize at this juncture that the problem does not emanate from a visual patriarchal mindset [7], whereby we associate female undergarment imagery to be somehow unethical, but the root cause lies in the fact that many of the images were curated into the dataset (at least with regards to the 2 above mentioned classes) were voyeuristic, pornographic, non-consensual and also entailed underage nudity.

2 Root cause: Absence of referencing consent during the curation of the imagenet dataset

The emergence of the ImageNet dataset is widely considered to be a pivotal moment⁴ in the *deep learning* revolution that transformed the domain computer vision. Two highly cited papers (with more than 10000 citations each), [8] authored by Deng et al in 2009 and [9] authored by Russakovsky et al in 2015, provide deep insights into the procedure used to curate the dataset. In the 2009 paper, subsections 3.1-Collecting Candidate Images and 3.2-Cleaning Candidate Images are dedicated towards the algorithms used to collect and clean the dataset and also to elucidate the specific ways in which the Amazon Mechanical Turk (AMT) platform was harnessed to scale the dataset. Similarly the entirety of Section-3-Dataset construction at large scale in [9] is dedicated towards extending the procedures for the 2015 release.

It is indeed disappointing that neither the 2009 nor the 2015 versions of the endeavors required the AMT workers to check if the images they were asked to categorize and draw bounding boxes over, were ethically viable for usage. More specifically, in imagery pertaining to anthropocentric content, such as undergarment clothing, there was no attempt made towards assessing if the images entailed explicit consent given by the people in the images. In fact, none of the following words in the set [ethics, permission, voyeurism, consent] are mentioned in either of the two papers. As such, we have a plethora of images specifically belonging to the two categories detailed in Table 1, that have serious ethical shortcomings. In Fig 2, we showcase the gallery of images from the two classes categorized into four sub-categories: Non-consensual/Voyeuristic, Personal, Verifiably pornographic and Underage / Children. In Fig 2, we also include images that were also incorrectly categorized (Specifically there was no brassieres being sported by the subjects in the images) (Sub-figure (a)) and those that involved male subjects indulging in lecherous tomfoolery (Sub-figure (e)).

3 Conclusion

In this paper, we expose a certain unethical facet of neural art that emerges from usage of non-consensual images that are present in certain specific classes of the imagenet dataset. These images born out of an unethical (and in some cases, illegal) act of voyeuristic non-consensual photography predominantly targeting women (as well as children), might implicitly poison the sanctity of the artworks that eventually emerge. This work is complementary to works such as [10; 11] that have explored the *unspoken ethical dimensions* of harnessing cheap crowd-sourced platforms such as AMT for scientific research in the first place, which we firmly believe is also an important issue to be considered.

Ethical considerations and caveats

We'd like to begin by admitting the self-contradictory facet of raising this specific issue by using a flagship example image that was generated using the very same ethically dubious procedure that we are targeting to root out.

Secondly, we'd like to inform the reader that we have indeed raised the issue with curators of the imagenet dataset, but to no avail and plan to update this dissemination if and when they respond.

⁴ <https://qz.com/1034972/the-data-that-changed-the-direction-of-ai-research-and-possibly-the-world/>

Class number	Class-ID	Class-description
445	n02892767	['bikini, two-piece']
459	n02837789	['brassiere, bra, bandeau']

Table 1: The two *problematic* classes in the ImageNet dataset

A Supplementary materials



Figure 1: An example neural art image generated by the author using the *GanBreeder* app [Gene weights: Bikini: 42.35, Brassiere: 31.66, Comic Book - 84.84]

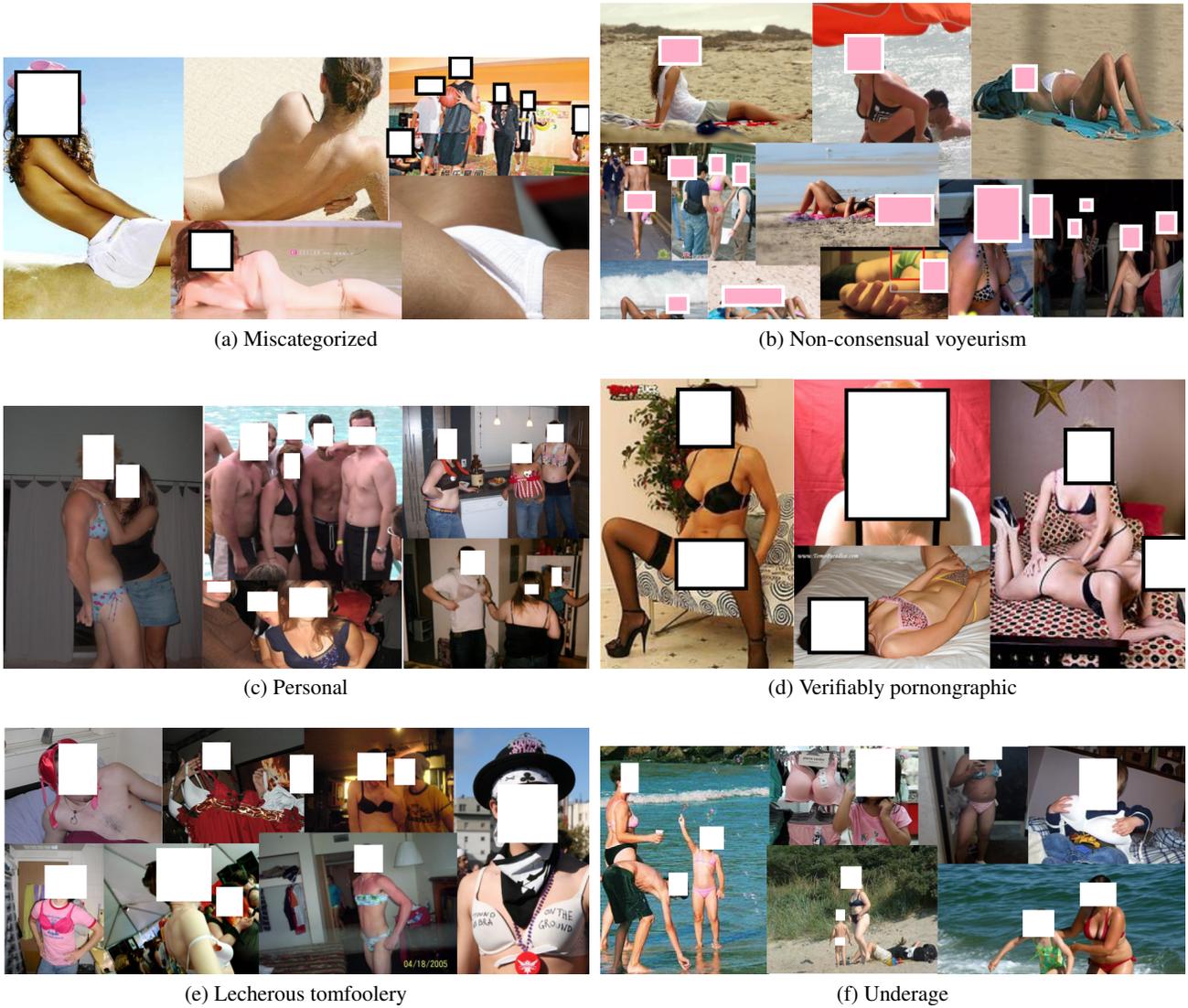


Figure 2: Categories of *unethical* images in the imagenet dataset

Reviewer-ID	Review
A- Grad student, CMU SCS	<p>This one reminds me of a mix between graffiti and paper mache using newspaper with color images or magazines . My attention is immediately drawn to near the top of the image which, at first glance, appears to be a red halo of sorts, but upon further consideration, looks to be long black branching horns on a glowing red background.</p> <p>My attention then went to the center top portion, where the "horns" were coming from, which appeared to be the head or skull of a moose or something similar. The body of the creature appears to be of human-like form in a crucifix position, of sorts. The image appears more and more chaotic the further down one looks.</p>
B- Grad student, Stanford CS	<p>Antisymmetric: left side is very artistic, rich in flavor and shades; right is more monotonic but has more texture.</p> <p>Reminds me of the two different sides of the brain through the anti-symmetry</p>
C- Data Scientist, Facebook Inc	Futurism
D- CS undergrad, U-Michigan	<p>It's visually confusing in the sense that I couldn't tell if I was looking at a 3D object with a colorful background or a painting.</p> <p>It's not just abstract, but also mysteriously detailed in areas to the point that I doubt that a human created these</p>
E - Senior software engineer, Mt View	<p>The symmetry implies a sort of intentionally.</p> <p>I get a sense of Picasso mixed with Frieda Callo here.</p>
F- Data Scientist, SF	<p>Reminds me of a bee and very colorful flowers, but with some nightmarish masks hidden in some places. Very tropical</p>

Table 2: Responses received for the neural art image in Fig 1

References

- [1] A. Brock, J. Donahue, and K. Simonyan, "Large scale gan training for high fidelity natural image synthesis," *arXiv preprint arXiv:1809.11096*, 2018.
- [2] "Collaborative art tool for discovering images." <https://ganbreeder.app>, 2019. [Online; accessed 9-8-2019].
- [3] M. A. Boden and E. A. Edmonds, "What is generative art?," *Digital Creativity*, vol. 20, no. 1-2, pp. 21–46, 2009.
- [4] A. Hertzmann, "Aesthetics of neural network art," *arXiv preprint arXiv:1903.05696*, 2019.
- [5] "Is artificial intelligence set to become art's next medium?." <https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx>, 2019. [Online; accessed 9-8-2019].
- [6] J. L. Fishman, "Is diamond smuggling forever-the kimberley process certification scheme: The first step down the long road to solving the blood diamond trade problem," *U. Miami Bus. L. Rev.*, vol. 13, p. 217, 2004.
- [7] S. Baran, "Visual patriarchy: Peta advertising and the commodification of sexualized bodies," in *Women and Nature?*, pp. 43–56, Routledge, 2017.
- [8] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255, Ieee, 2009.
- [9] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.
- [10] K. Fort, G. Adda, and K. B. Cohen, "Amazon mechanical turk: Gold mine or coal mine?," *Computational Linguistics*, vol. 37, no. 2, pp. 413–420, 2011.
- [11] V. Williamson, "On the ethics of crowdsourced research," *PS: Political Science & Politics*, vol. 49, no. 1, pp. 77–81, 2016.