

Poetics of Autonomy

Exhibition at Research Pavilion 2021, Helsinki

Bruno Caldas Vianna

Doctoral Programme in Fine Arts
Fine Arts Academy, University of the Arts Helsinki

Research theme

This doctoral project research investigates the relations between visual arts and artificial intelligence. A great weight is given to the concept of autonomy, as some kind of Holy Graal of cybernetics: the creation of entities that can fend off for themselves, like robots, autonomous cars or artificial software agents. As such, the project has the current working title of "Poetics of Autonomy". But this concept is also dear to the arts. Much has been written and discussed on the potential for autonomy within the art object, and how much modernism and conceptual art has changed it.

My concern at this exhibition was to understand the tools AI offers for visual arts and issues they might convey. I also address the concept of autonomy from a perspective of the post-human, or the cyborg. What art can be done from an apocalyptic perspective? When no humans are left on the planet, could there be something left that continues to create art? And what kind of serendipity can be created - or at least simulated - with the current state of AI?

The exhibition took place at the Hietsun Pavilion in Helsinki, from the 1st to the 30th July, 2021, in the context of the 2021 edition of Uniarts Helsinki's Research Pavilion.

Post-colonialism

This work is an effort to visualize biases in visual artificial intelligence tools. It is presented as a triptic, with each of the three pieces measuring 100x100cm. The central piece has a grid of 10.000 fake human faces generated by the StyleGAN2 algorithm.

This algorithm is of special interest to me not only because it is probably the most popular platform for AI-generated imagery, but also because it is one of the

offsprings of the original generative adversarial network (GAN), invented in 2014 by researcher Ian Goodfellow. This invention was a breakthrough in the field, as for the first time it allowed a computer program to be trained in almost any visual style and reproduce it, creating new works which look similar to the training images but are never equal. GANs are the base for a few landmark works, like the piece created by the French collective named Obvious, which sold for 432.000 dollars by the auction house Christie's in New York City. It is a tool used by many artists such as Helena Sarin, Anna Riedler or Gene Kogan. (Zeilinger 2021; Zylinska 2020)

In order to generate new images, this platform requires a great number of images to be trained on - in the orders of thousands (although artistic projects might sacrifice quality for creativity and use just a few hundred images to train). Training is a computing-expensive process, which is usually done in the research centers. The graphics cards company NVidia - where the current version of StyleGAN2 is being developed - offers a few pre-trained models for experimentation. Probably the most popular of them is based on the Flickr faces dataset.¹ This collection of face images was created by scraping the Flickr website. Images with permissive licences were downloaded and cleaned up to arrive at a total of 70.000 1024x1024 pictures of faces. Even though the creators of the site claim that it "contains considerable variation in terms of (..) ethnicity", its distribution is clearly biased towards caucasian people.

In order to make this bias visible, I ran another artificial intelligence algorithm - perhaps just as problematic as the dataset - that proposes to classify human faces according to ethnicity. Using it as a base, I wrote a software that erases all the caucasian faces from the same grid created before, creating empty spaces and therefore leaving only non-white people. This is the piece on the left side of the triptic. As one could expect the result of the effacement is also highly questionable. The program does a terrible job, leaving persons who are clearly white, while erasing faces of non-western people. It should be said, in any case, that the software used is not professional or peer-reviewed. In fact, most of the companies which offered such services stopped offering them due to the obvious controversy of the possible uses of such a tool.

Faced with those concerns, I decided to add the third element of the piece. In this last one, I did the selection of non-white faces manually. This allowed me to reflect on the subjectivity and criteria that such a task entails. Ethnicity is not just a visual feature: I weighed my decisions on factors I did not suspect, such as what I evaluated to be the wealth of the subject. Even though I might have corrected what I considered to be errors of the ethnicity algorithm, I certainly have brought my own flaws into play. It

¹ <https://github.com/NVLabs/ffhq-dataset>

also allowed me to compare manual labor against machine labor. The task took me about 16 hours to complete, considering I had to go through 10.000 faces, and that I might have spent a long time on a few ones which were more undefined for my own subjectivity. The automated process, in comparison, took just a couple of minutes.

No conclusion is wished or possible from this process. The results are the reflections that will be incorporated into the thesis documentation and overall artistic research process.

Slaughterhouse

This piece is also based on the StyleGAN2 architecture, which offers a tool named “projector”. This allows the user to execute a search in the latent space - the set of possible images to be generated from the model. This search is made from any given image, belonging or not to the training dataset or even or the latent space. It will try to show the image most similar to the given one. For instance, if I use my passport photo as input, it will give me a face similar to me, from within its universe of possible generated images.

What happens, though, if the search image has nothing to do, visually, with the original training set? The system still tries to find something similar - but will certainly end up with something grotesque. This peculiar result may be seen as a blended image between the latent image space and the searched image. But in more exact terms, it is the image belonging to the latent space that looks most similar to the search one. If we think of GANs as a tool for visual artists, then this is one of the best ways to subvert its usage.

In the work created for the exhibition, I still worked with the Flickr Faces dataset model. In order to disrupt the tool, my technique was to feed the architecture with faces of animals. Since they bear features common to human faces (eye, mouth...), I supposed correctly that animal-like humans (or human looking animals) could be found. The search is a labor intensive task, since it can be run multiple times, with different random results for each input; the pieces are the result of my personal curatorship.

I have also been looking to find architectural features in latent spaces of natural landscapes, although I couldn't finish this experiment in time for the exhibition.

Photosynthesis

This is an exercise in autonomous generative art. The idea of a program as a set of rules to create visuals is very common. In some cases, like Sol Lewitts' projects, a system can be devised that does well without the need of computers; however, most algorithm art is based on digital tools. These tools are contingent on the infrastructure built around our society and will cease to work if not connected to a source of electricity. And most sources of all energy in our planet are somehow derived from the sun - fossil fuels, hydropower, wind and, naturally, solar power. In the end, humans are creatures who devised a plethora of new photosynthesis methods to fulfill our energy needs.

This work exposes this relation, which is usually hidden, by means of a combination of low-energy screen, computing chip and solar panel. The algorithm inscribed in the chip will generate new drawings of a tree whenever it is being lit by the sun. The e-paper display is of the same kind used in electronic book readers such as the Kindle. It has the characteristic of being very slow to update, but will retain the image even if the device is not being powered.

Crank of Life

This work references one of the most famous concepts in cybernetics: the game of life, designed by mathematician John Conway. It consists of a set of rules that will make an initial random pattern evolve over time, creating endless combinations of new patterns that can merge, die, or change shape. It spawned a new category of algorithmic simulations named cellular automata. The game has a hypnotic feature which allows it to be watched looking for chaos, stabilizations, repetitions.

My installment of the game of life also uses the low energy chips used in Photosynthesis, except with a faster display that needs constant energy. To put the system in motion, the visitor must activate a hand cranked dynamo. Life in the small screen will evolve as long as the crank is in movement. Every time it stops and starts again, a new random initial configuration is created, which in turn evolves into infinite possibilities.

Aiku

This work is designed around the prospect of serendipity, inspiration and imagination

in an autonomous system. It is also a note on the flaws of current commercial AI systems, and how limited can some of them be. In fact, the mistakes the software makes are likely more interesting than the intended results, which are on the verge of uselessness. The name is a wordplay on haikus, AI and the beginning of the Finnish word for adulthood (aikuisuus).

The piece runs on a single board computer connected to the Internet. This device is equipped with a Google Coral TPU, an accessory that allows cheap computers to run inferences, that is, to use AI models that make classifications and other processing intensive tasks. It is useful for machine vision tasks, such as video surveillance. It is used here not with live images, but by retrieving random picture files from wikipedia. The reason for using an online source instead of a camera is that the local solution would offer not much input variation - there are only a few different objects that can be found at a gallery.

The software downloads a picture and makes an inference on it that is most likely to be flawed: the Google dataset used has the capacity to recognize only 90 different things. The pictures and results are displayed on a small screen. These mistaken inferences are then fed into another AI system, which uses the words as a prompt to create haikus. The nonsensical poems are finally displayed in another micro display.

REFERENCES

- Zeilinger, Martin. 2021. *Tactical Entanglements: AI Art, Creative Agency, and the Limits of Intellectual Property*. Lüneburg: meson press.
- Zylinska, J. 2020. *AI Art: Machine Visions and Warped Dreams*. Media: Art: Write: Now. Open Humanities Press. <https://books.google.es/books?id=ZbaozQEACAAJ>.