Supervision

Pre-examination artistic work
Bruno Caldas Vianna
PhD candidate at the Academy of Fine Arts, Uniarts Helsinki
Supervisor: Mika Elo

This work is a two-way mirror mounted on a panoramic window, overlooking the digital archive of the Finnish Museum of Photography. It informs my doctoral research on visual arts and artificial intelligence in the sense that it is the result of the use of computational tools to understand and organize the collection of the museum.

These tools are not built simply on the visual features of the photographs. They rely on artificial intelligence models which were trained on a huge number of images retrieved from the Internet in order to create a machinic method for classification. Embedded in the organization depicted in the piece, there are efforts to create an automated visualization of the world, based not only on the work of mathematicians and computer scientists, but also on uncredited digital workers which had to manually classify the objects in found in millions of images to build the model.

What does it mean to use machine learning to look and distribute 12.684 images from the museum collection? One hand, this would be a work almost impossible to perform by hand. There is clearly a machine mediation, an automated process through which we rely on the algorithm to find similarities and repetitions within the database. At the same time, this is not a purely mechanical process. On the surface, it is easy to see that there were artistic interventions to decide on the size, format, placement of the piece as a whole and on individual frames. Personally, I should add that I had to go through several iterations and attempts until I found a result which pleased me aesthetically, and that could convey my goals within the exhibition and also within my own artistic research project.

But most importantly, the mechanisms under the hood inform us essentially about ourselves. The mirror lets us gaze through at the pictures of the collection, but it also reflects the human imprint on the algorithms. From the selection of categories, the availability of images to the construction of the AI models, to the subjectivity of Mechanical Turk workers, all the biases present in creating these tools will also appear on the exhibited result, no matter if evident or subtly scattered. Finally, it also reflects the will for control that is brought by automation: the tools used for expressive visualization are the same used for surveillance. The title of the piece is meant to comment on this dangerous usage, as well as the superhuman yet mundane capacities of artificial intelligence.

The piece in the context of the doctoral investigation

This piece addresses a field within artificial intelligence which is commonly referred as *unsupervised learning.* In such processes, the algorithm will try to infer knowledge on existing data using solely machinic resources, that is, without human support in the form of of training or evaluation. Such a support would bring the analysis to the context of *supervised learning*, which is common is object recognition and image generation. The pieces belonging to my previous pre-examination dealt with this latter type of analysis.

The potential for the use of unsupervised learning in studies of media and culture is recognized by scholars like Lev Manovich (Manovich 2021), specially in the field of Digital Humanities (Pääkkönen 2021). However, as the models become more complex, it also becomes hard to affirm that the processes are exclusively machine-based. The image created for this work relied on pre-trained models, as discussed above.

Technical background

There were several steps involved in the creation of the tableau in the exhibition. The first was to use the ResNet50 framework (He et al. 2015) in order to provide a numerical translation of the content of images. One of the options of this framework is to use the Imagenet classification database (Deng et al. 2009), which contains one thousand different categories, from elephants to houses, cars or hammers. Imagenet was made possible by the appearance of Amazon's Mechanical Turk platform, used to outsource the work of training the model to thousands of distributed human workers. ("ImageNet: The Data That Spawned the Current AI Boom — Quartz" n.d.) This dataset was used in the Imagenet competition (Russakovsky et al. 2015), the event that sparked some of the most important developments within machine learning, such as the Alexnet neural network implementation. (Alom et al. 2018). Different frameworks were tested, like VGG16 and VGG19, with Resnet yielding better results.

In my implementation, the classification resulted in a vector of 2048 values containing classification information, which I wanted to reduce to only two: x and y, the position of the images in the tableau. For that I used standard techniques to reduce the dimensionality of the data, first to one hundred dimensions (PCA) and finally to two (T-SNE). These two vectors defined the final position of each photo in the frame.

Another feature was added from a suggestion of the curator: as the images formed clusters of categories, the representativity of each one within their cluster was indicated by their size. In other words, the biggest one image is the most representative of its group. If an image is small, it indicates that it couldn't be properly categorized and sits in between two groups.

Bibliography

Alom, Md Zahangir, Tarek M. Taha, Christopher Yakopcic, Stefan Westberg, Paheding Sidike, Mst Shamima Nasrin, Brian C. Van Esesn, Abdul A. S. Awwal, and Vijayan K. Asari. 2018. "The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches." *ArXiv:1803.01164 [Cs]*, March. http://arxiv.org/abs/1803.01164.

Deng, J., W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. 2009. "ImageNet: A Large-Scale Hierarchical Image Database." In *CVPR09*. https://image-net.org/static_files/papers/imagenet_cvpr09.pdf.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2015. "Deep Residual Learning for Image Recognition." arXiv. http://arxiv.org/abs/1512.03385.

"ImageNet: The Data That Spawned the Current AI Boom — Quartz." n.d. Accessed September 22, 2021. https://web.archive.org/web/20210826054758/https://qz.com/1034972/the-data-that-changed-the-direction-of-ai-research-and-possibly-the-world/.

Manovich, Lev. 2021. "Computer Vision, Human Senses, and Language of Art." AI & SOCIETY 36 (4): 1145–52. https://doi.org/10.1007/s00146-020-01094-9.

Pääkkönen, Juho. 2021. "Data Do Not Speak for Themselves: Interpretation and Model Selection in Unsupervised Automated Content Analysis." *Composition and Big Data*.

Russakovsky, Olga, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, et al. 2015. "ImageNet Large Scale Visual Recognition Challenge." *International Journal of Computer Vision (IJCV)* 115 (3): 211–52. https://doi.org/10.1007/s11263-015-0816-y.