

Beautification of Images by Generative Adversarial Networks (GANs)

Amar Music

Master of Psychology: Theory and Research

KU LEUVEN

Faculty of Psychology and Educational Sciences

Supervisor: Prof. Johan Wagemens

Co-supervisor: Anne-Sofie Maerten

May 28, 2022

Contents

1	Abstract	2
2	Beautification of images by Generative Adversarial Networks (GANs)	3
2.1	Philosophy	3
2.2	Psychology	3
2.3	Machine Learning	3
2.4	Solving problems of defining beauty with ML	3
3	Methods	4
3.1	Questionnaire	5
3.2	ss2	5
3.3	ss3	5
4	Results	6
4.1	Pilot	6
4.1.1	Participants	6
4.1.2	Data Analysis	6
5	Discussion	7
5.1	ss1	7
5.2	ss2	7
6	Appendix	9
6.1	Questionnaire	9
6.2	Instructions	9

1 Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

2 Beautification of images by Generative Adversarial Networks (GANs)

2.1 Philosophy

- What is beauty?
- Philosophy of aesthetics
- Transition to psychology to find answers

2.2 Psychology

- What is beauty
 - Studies on aesthetics
 - Evolutionary psychology
 - Experimental aesthetics
 - **Supernormal stimuli??** (probably discussion, not introduction)
 - **Hyperreality**
 - **Peak shift principle**
 - Refer to AestheticsNet

2.3 Machine Learning

- Introduction to GANs
- Use of GANs in cognitive science

2.4 Solving problems of defining beauty with ML

- GANalyze to obtain a visual definition
- AestheticsNet & BigGAN

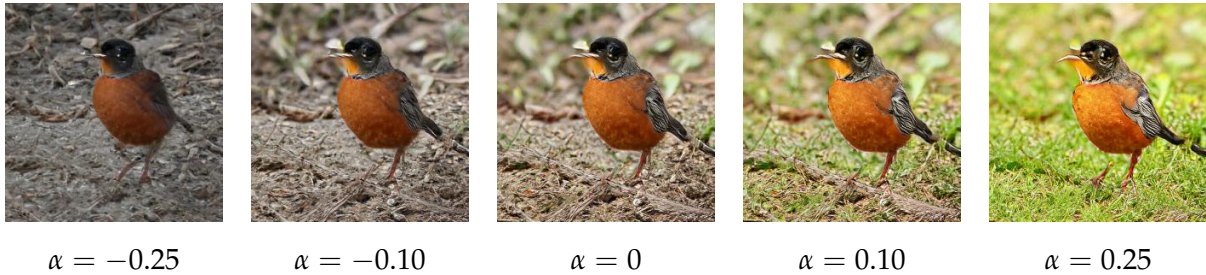
3 Methods

To determine whether GANalyze (Goetschalckx et al., 2019) can be used to study what it means for an image to be aesthetic, I conducted a study to test whether human participants indeed preferred the images that GANalyze generated to be more aesthetic over a corresponding base image.

I used GANalyze with AestheticsNet (**kongPhotoAestheticsRanking2016**) as the assessor and BigGAN-256 (**brockLargeScaleGAN2019**) pretrained on ImageNet (Rusakovsky et al., 2015) as the generator to train the GANalyze model for 400,000 iterations. The training resulted in GANalyze being able to produce image sequences based on 1,000 ImageNet categories. For each ImageNet category, I generated three different seeds (i.e., three different images of Siamese cats) to test for an effect of category. Within each seed, GANalyze produced 21 images with a supposedly increasing amount of aesthetic value (Figure 1). I picked a broad range of α -values with increasing density close to zero to obtain stimuli with obvious changes (e.g., $\alpha = 0.25$), but also subtle changes (e.g., $\alpha = 0.0025$). I chose these values because I hypothesized the participants to agree with GANalyze 100 of the time for trials with a large α -value, and 50 of the time (picking at random) for very small α -values, resulting in a psychometric function. The specific α -values were chosen based on the results of a pilot study with a very broad range.

Figure 1

Truncated Sample of an Image Sequence from One Seed Produced by GANalyze



489 ImageNet categories were manually excluded from the stimulus set because they either contained human faces (which are deliberately warped by BigGAN) or insects such as spiders which might make some participants feel uncomfortable. In addition, stimuli that were unrecognizable were also excluded. This resulted in 9xx categories, each with three seeds and 21 α -values, equaling a total of 30.807 stimuli.

The study was hosted on Prolific.co to obtain a large number of data points in a simple online task, ensuring each of the 30.807 stimuli was evaluated 2-3 times on average. Individuals younger than 17 or those without normal or corrected-to-normal vision were not allowed to participate. At the start of the study, participants were asked to fill in demographic details such as their age, gender, and country. They were also polled on their previous experience and familiarity with visual arts. Afterwards, they had to read and accept an informed consent document after which they were provided with instructions for the task. The participants carried out a spatial two-alternatives forced choice (2AFC) task in which they had to indicate which of the two presented GAN-generated images they considered to be most aesthetically pleasing. They did this by pressing "F" for the left, and "J" for the right image (Figure 2).

Figure 2

Example of a Sequence of Trials in the Study



Each trial consisted of two images generated from the same seed within an ImageNet category. One of the two images was always the base image with $\alpha = 0$ from that trial's seed while the comparison image was, according to GANalyze at least, either more aesthetic ($\alpha > 0$) or less aesthetic ($\alpha < 0$). The α -value for the comparison image differed for each trial, leading to differences in similarity between the base and comparison images for each trial. An α -value close to zero indicates high similarity and, consequently, difficulty to discriminate between the images. An α -value much larger or smaller than zero indicates low similarity and therefore easy discrimination. The position of the base and comparison image was randomized for each trial to prevent possible confounding factors. Each 10-minute block was created to contain a roughly uniform distribution of chosen α -values and ImageNet categories to ensure that each participant is presented with a comparable stimulus set. Furthermore, the same category never appeared more than one time in a block.

After 10 minutes, the task automatically turned to a screen thanking the participants for their assistance, after which they were given monetary compensation of £1.38 converted to their local currency.

3.1 Questionnaire

I used a shortened version of the Aesthetic Experience Questionnaire (AEQ; Wanzer et al., 2020) to measure experience and ... with art/aesthetics. I picked only the highest loading question for each of the six factors in order to save on time.

3.2 ss2

st2

3.3 ss3

Check out Goetschalckx et al. (2019).

4 Results

Results here please

4.1 Pilot

A pilot study was conducted to find which alpha values would work best and possibly to reduce the total number of required trials.

4.1.1 Participants

x participants have been excluded due for various reasons. Etc etc

4.1.2 Data Analysis

Statistics here. $p < 0.05$ etc very good

5 Discussion

Discussion here please

5.1 ss1

st1

5.2 ss2

st2

References

- Goetschalckx, L., Andonian, A., Oliva, A., & Isola, P. (2019). GANalyze: Toward Visual Definitions of Cognitive Image Properties. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 10. <https://doi.org/10.1109/ICCV.2019.00584>
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115(3), 211–252. <https://doi.org/10.1007/s11263-015-0816-y>
- Wanzer, D. L., Finley, K. P., Zarian, S., & Cortez, N. (2020). Experiencing flow while viewing art: Development of the Aesthetic Experience Questionnaire. *Psychology of Aesthetics, Creativity, and the Arts*, 14(1), 113–124. <https://doi.org/10.1037/aca0000203>

hello test2

6 Appendix

6.1 Questionnaire

In general when I view art...

Question 1:

Question 2:

Question 3:

Question 4:

Question 5:

Question 6:

6.2 Instructions