Exploring Latent Dimensions of Crowd-sourced Creativity

Umut Kocasarı, Alperen Bag, Efehan Atici, Pinar Yanardag

Bogazici University

Motivation

- Recently, discovering interpretable directions in the latent spaces of pre-trained GANs has become a popular topic.
- While existing works mostly consider directions for semantic image manipulations, we focus on an abstract property: Creativity. Can we manipulate an image to be more or less creative?
- We build our work on the largest AI-based creativity platform, Artbreeder.com, where users can generate unique images using pre-trained GAN models.
- We explore the latent dimensions of images generated on this platform and present a novel framework for manipulating images to make them creative.
- Our code and dataset are available at `http://github.com/catlab-team/latentcreative`.

Figure: Example images from the 'General' category of ArtBreeder. The creators of the images are annotated with labels.

References

[1] Andrew Brock, Jeff Donahue, and Karen Simonyan. Large scale GAN training for high fidelity natural image synthesis. CoRR, abs/1809.11096, 2018

[2] Lore Goetschalckx, Alex Andonian, Aude Oliva, and Phillip Isola. Ganalyze: Toward visual definitions of cognitive image properties. In Proceedings of ICCV, pages 5744–5753, 2019.

[3] David Hasler and Sabine E Suesstrunk. Measuring colorfulness in natural images. In Human vision and electronic imaging VIII, volume 5007, pages 87–95. International Society for Optics and Photonics, 2003.

[4] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In Proceedings of ICCV, pages 2961–2969, 2017.

Methodology

- Given a generator G, a class vector y, a noise vector z, and an assessor function A, the GANalyze [2] model solves the following problem: $\mathcal{L}(\theta) = \mathbb{E}_{\mathbf{z},\mathbf{y},\alpha}[A(G(T_\theta(\mathbf{z},\alpha),\mathbf{y})) - (A(G(\mathbf{z},\mathbf{y})) + \alpha))^2]$, where α is a scalar value representing the degree of manipulation, θ is the desired direction, and T is the transformation function defined as $T_\theta(\mathbf{z},\alpha) = \mathbf{z} + \alpha\theta$ that moves the input $\mathbf{z}$ along the direction θ.
- In this work, we extend the GANalyze framework where we use a neural network that uses noise drawn from a distribution $\mathcal{N}(0,1)$ which learns to map an input to different but functionally related (e.g., more creative) outputs:

$$\mathcal{L}(\theta) = \mathbb{E}_{\mathbf{z},\mathbf{y},\alpha}[(A(G(F_z(\mathbf{z},\alpha), F_y(\mathbf{y},\alpha)))) - (A(G(\mathbf{z},\mathbf{y})) + \alpha))^2]$$

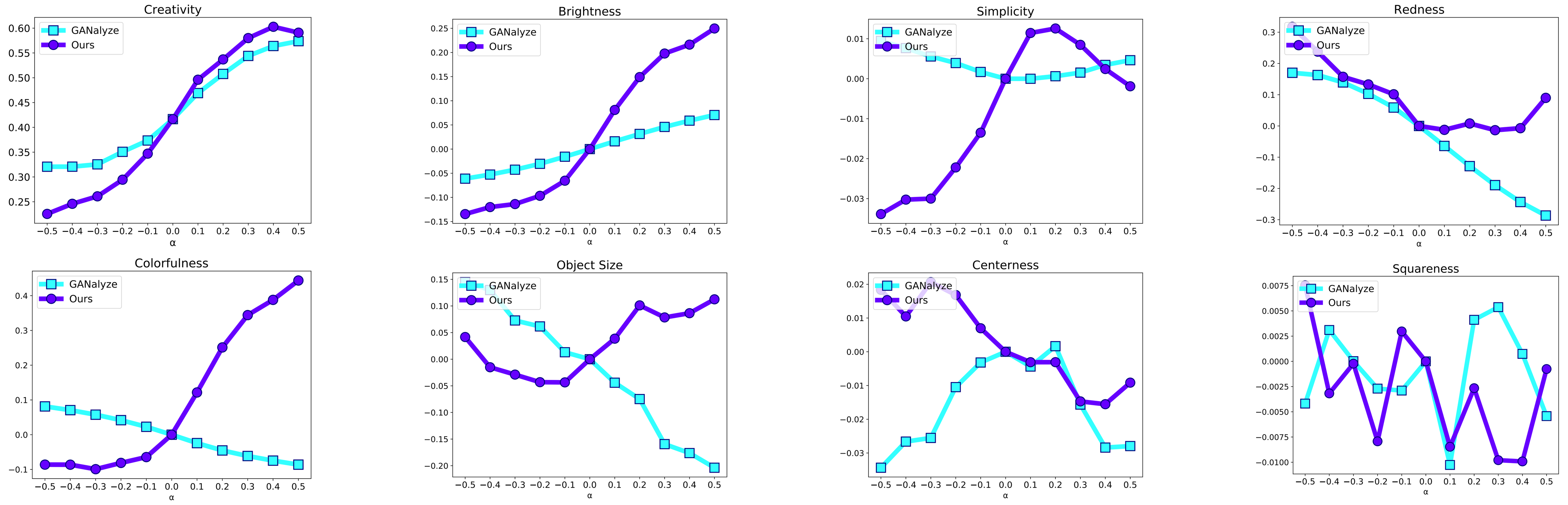
where the first term represents the score of the modified image after applying the function F with parameters $\mathbf{z}$, $\mathbf{y}$, and α, and the second term simply represents the score of the original image increased or decreased by α. F_z computes a diverse direction θ with a noise ϵ as $F_z(\mathbf{z},\alpha) = \mathbf{z} + \alpha \cdot \mathbf{NN}(\mathbf{z},\epsilon)$. We also learn a direction for class vectors as follows:

$$\mathcal{L}(\theta) = \mathbb{E}_{\mathbf{z},\mathbf{y},\alpha}[A(G(F_z(\mathbf{z},\alpha), F_y(\mathbf{y},\alpha))) - (A(G(\mathbf{z},\mathbf{y})) + \alpha))^2]$$

where F_y is calculated as $F_y(\mathbf{y},\alpha) = \mathbf{y} + \alpha \cdot \mathbf{NN}(\mathbf{y},\epsilon)$ where NN is a two layer neural network.

Results

- Our model achieves a much lower creativity score for low α values such as $\alpha = -0.5$, while simultaneously achieving higher creativity for $\alpha = 0.4$ and $\alpha = 0.5$.
- We investigate what kind of image factors are changed to achieve these improvements. Following [2], we examined redness, colorfulness, brightness, simplicity, squareness, centeredness and object size.



Results

Our method is capable of performing several different manipulations on the input images.

Conclusion and Limitations

Our work sheds light on the understanding of creativity and opens up possibilities to improve human-made art through GANs. We point out that using GAN-generated images as a proxy for creativity may bring its own limitations.