Neuro-Symbolic Generative Art: A Preliminary Study

Gunjan Aggarwal¹ and Devi Parikh²

¹Adobe
²Georgia Tech and Facebook AI Research

guaggarw@adobe.com parikh@gatech.edu

Abstract

There are two classes of generative art approaches: neural, where a deep model is trained to generate samples from a data distribution, and “symbolic” or algorithmic, where an artist designs the primary parameters and an autonomous system generates samples within these constraints. In this work, we propose a new hybrid genre: neuro-symbolic generative art. As a preliminary study, we train a generative deep neural network on samples from the symbolic approach. We demonstrate through human studies that subjects find the final artifacts and the creation process using our neuro-symbolic approach to be more creative than the symbolic approach 61% and 82% of the time respectively.

Introduction

Generative art refers to art generated using code, and typically includes an element of chance. Interactive versions of these systems can be viewed as Casual Creators (Compton and Mateas 2015). There are two dominant approaches for generative visual art. The first uses deep neural networks to generate images from a distribution that mimics training data. Indeed, generative artists train models on specific photographs they take or collect (e.g., Helena Sarin, Robbie Barrat) or perturb the weights of models to create artistic “glitches” in the generated art (e.g., Mario Klingemann). Another source of control is the random noise input to the model. Interpolations of two noise vectors smoothly control the generation in a local neighborhood. In the second approach an artist defines an algorithm to generate art. An autonomous system generates random samples using this algorithm. Early algorithmic artists include Georg Nees and Vera Molnar. Algorithmic art is often abstract, with geometric structures, or repeating or recursive patterns. These “symbolic” approaches typically have explicit parameters to control the generated art.

To the best of our knowledge, these two approaches to generative visual art – neural and symbolic – have been largely distinct. This work is a preliminary study in exploring their intersection: neuro-symbolic generative art. Specifically, we train a Generative Adversarial Network (GAN) on samples generated using a symbolic approach. We hypothesize that the organic, unpredictable aesthetic associated with neural approaches complements the crisp, designed aesthetic of symbolic approaches. Moreover, compatible with data-hungry deep models, symbolic approaches support generation of large amounts of training samples. Example generated art samples from our approach are shown in Figure 1. Our human studies show that subjects find the artifacts and the interactive creation process using the neuro-symbolic approach to be more creative 61% and 82% of the time respectively compared to the symbolic approach.

Related Work

Neural generative models. These include Generative Adversarial Networks (GANs) (Goodfellow et al. 2014), Autoregressive Models (Salimans et al. 2017), Latent Variable Models (Kingma and Welling 2014), etc. Recent progress in GANs enables realistic natural (Brock, Donahue, and Simonyan 2019) and high resolution human face (Karras, Laine, and Aila 2019) image generation. We limit our study to GANs. GANs to generate video game levels (Giacomello, Lanzi, and Loiacono 2018) are particularly relevant as neu-
A user can control the samples via the input noise vectors, and interpolations of two noise vectors. For interpolated generation, we sample two noise vectors and then create arbitrary linear interpolations between the two vectors. Neuro-Symbolic Interpolations (NSI) are then generated by feeding the model each of the interpolated latent vectors. Specifically, NSI $x$ is generated from the generative model $G$ as

$$
x = G(z); z = z_1 + \alpha \ast (z_2 - z_1)
$$

$z_1 \sim N(0, 1); z_2 \sim N(0, 1); \alpha \in (0, 1)$ set by the user.

Figures 3 and 9 show example interpolated samples.

**Human Evaluation**

We perform evaluation of both the artifact and the user-driven interactive creation process via human studies on Amazon Mechanical Turk (AMT). Subjects were from the US, with an AMT approval rating of 95% or higher, and having completed at least 5000 tasks on AMT in the past. They were paid above federal minimum wage.

**Artifact Evaluation**

We compare three artifacts – Symbolic, Neuro-Symbolic Generation (NSG), and Neuro-Symbolic Interpolated generation (NSI). These replicate the different artifacts a user might create when using the symbolic or neuro-symbolic interactive generative art tools. Human subjects were shown a pair of art pieces, one each from random two of the three types. They were asked which piece 1) seems more different from art you’ve seen in your life? 2) looks better? 3) is more creative? 4) is more artistic? 5) seems more likely to be hand-made? Subjects were also asked to optionally state why they felt one art was more likely to be hand-made than the other. The study consisted of 60 pairs, equally distributed across the three pairs of approaches. The study was completed by 20 unique subjects, resulting in a total of 1200 pairwise assessments.

We perform evaluation of both the artifact and the user-driven interactive creation process via human studies on Amazon Mechanical Turk (AMT). Subjects were from the US, with an AMT approval rating of 95% or higher, and having completed at least 5000 tasks on AMT in the past. They were paid above federal minimum wage.

**Approach**

We experimented with different GAN model architectures and found Progressive GAN (Karras et al. 2018) to work best. The final generated image is $512 \times 512$. The image starts at $4 \times 4$, and is doubled every $37k$ iterations. The input noise is $512d$. The learning rates for the generator and discriminator were $0.001$ for all image resolutions. We use different batch sizes for different image resolutions during training: $128$ till $16 \times 16$ and then halved after each subsequent increase in resolution. The training is run for $600k$ iterations with Adam optimizer. We refer to the samples generated from this model as Neuro-Symbolic Generations (NSG). Some examples of NSG can be found in Figure 8. Different NSG samples can be generated from the model by feeding it different input noise vectors.
The proportion of times users preferred one art form over another is shown in Figure 5. A one-sample proportion hypothesis test suggests that for our sample size, a “win ratio” over 0.54 (or below 0.46) is statistically significant at 95% confidence. These are shown as a horizontal lines in the figure. Novelty, unusualness: We find that the human evaluators rate NSG and NSI as being more “different” from art they’ve seen before than the Symbolic art about 66% of the time. Better quality, value: Subjects like NSG, NSI and Symbolic almost equally. Creativity: The third and fourth dimensions (“creative” and “artistic”) focus directly on the creative aspect of the artifacts. We see that human subjects find NSG and NSI to be more creative than Symbolic art about 61% and more artistic about 63% of the times. Note that NS(G/I) and Symbolic rated similarly for quality, but NS(G/I) were rated higher for novelty. We hypothesize that this results in NS(G/I) being rated higher in creativity overall (novelty + value, (Boden 2004)). Naturalness, hand-made: Subjects find NS(G/I) art to be more natural or more likely to be hand-made. Based on the comments shared, while certain subjects preferred Symbolic art as being more hand-made because of “perfect coloring”, about 59% of them chose the NS(G/I) art to be more likely to be hand-made because it “Looks like human error with paint dripping on to another color”, “The other piece of art has solid colors, where as the one I picked has various shades in spots.”, “mixture of color together”, “smudge”. Finally, we see that NSI is preferred over NSG for novelty and creativity.

Creation Process Evaluation Next, we evaluate live interactive generative art tools based on symbolic and our proposed neuro-symbolic approaches. Recall that the symbolic approach has 2 controllable parameters: color palette and the number of colors (maximum 5), as well as an option to generate a new variant of the art with the same parameters by changing the random seed. The neuro-symbolic tool has one controllable parameter $\alpha$ which generates an NSI between 2 NSG pieces, and an option to sample a new pair of NSG art by sampling new noise vectors. Human subjects were given links to both tools (Symbolic: http://genart.cloudcv.org/symbolic Neuro-symbolic: http://genart.cloudcv.org) and for each, were asked to: “Find an art piece that you like a lot and share it with us!” and describe “what characteristics of your favorite art made it stand out from others?” Additionally, subjects were asked which tool 1) generates better looking art? 2) generates more surprising / unusual / unpredictable art? 3) generates more creative art? 4) is more satisfying to work with? 50 unique subjects participated. Half were given the symbolic tool first, and the other half the neuro-symbolic tool. Both tools had an option to add up to 5 pieces to their “favorite” gallery so users can keep track of pieces they like as they encounter them. Users could delete pieces from the gallery to replace them with others. They were provided an easy way to copy the URL of their favorite piece and submit it.

The proportion of times subjects preferred the neuro-
symbolic (NS) tool over symbolic (S) is shown in Figure 6. Trends are similar to artifact evaluation. **Better quality, value:** They like art generated by both tools equally. **Novelty, unusualness:** Subjects rate NS to be more surprising and unusual than S 68% of the time. **Creativity:** Subjects find the NS tool to generate more creative art than S 82% of the time. **Satisfying:** Interestingly, while less creative, subjects find S to be more satisfying to work with (albeit, not with statistical significance). An indicative comment from a subject: “I liked the task. I found that in [NS] the colors felt like they mixed together more. I found that the art in [S] was more clean looking and that made it more satisfying to feel like they mixed together more. I found that the art in [S] was more clean looking and that made it more satisfying to work with (albeit, not with statistical significance).” Using S is perhaps more analogous to “zen” (relaxing) activities, while NS may be closer to cognitively taxing creative activities. Exploring this is future work. Other comments about the two tools: “I had a bit more creative control with [S], while [NS] did generate more interesting combinations, it was just harder to get there predictably.” “[NS] provided more creativity, versus taking out colors like in [S].” “I found that in [S] the ability to choose the number of colors. I felt that in [NS] it was a little harder using my mouse to get the form and shape of the circles I wanted.” “This was a very interesting experience, especially [NS]. I kind of felt like I didn’t know what to expect when I was trying to make hybrid art.”

Example generated art samples beyond those shown in Figures, screen captures of our interactive generative art tools, and “favorite” pieces created by subjects along with a description of why they like the pieces can be found here: [https://sites.google.com/view/neuro-symbolic-art-gen](https://sites.google.com/view/neuro-symbolic-art-gen)

**Conclusion**

We present a preliminary study on neuro-symbolic generative art. It combines what have typically been two distinct approaches to generative visual art: neural and algorithmic/symbolic. We trained GANs on data generated via a symbolic approach. We evaluate the generated art and build live interactive generative art tools using both approaches. Human studies show that subjects find the neuro-symbolic generated art and creation process to be more creative than symbolic counterparts 61% and 82% of the time respectively. Overall, we see promising indications that neuro-symbolic generative art may be a viable new genre.

**Future Work.** We will explore other symbolic art styles, and train a model over multiple styles to potentially discover entirely novel styles. We further plan to interpolate between two symbolic images instead of two neuro-symbolic images. For this, we will explore techniques that map real images to latent representations. A user can then first design the two ends points (symbolically), and then generate an intermediate piece (neurally). Neither symbolic nor neuro-symbolic approaches alone allow for this level of control. Finally, training GANs directly on symbolic representations is an interesting and open research question.

**Acknowledgment.** Abhishek Sinha for helpful discussions.

**References**

[Boden 2004] Boden, M. A. 2004. The Creative Mind: Myths And Mechanisms. Psychology Press.

[Brock, Donahue, and Simonyan 2019] Brock, A.; Donahue, J.; and Simonyan, K. 2019. Large Scale GAN Training For High Fidelity Natural Image Synthesis. In ICLR 2019.

[Compton and Mateas 2015] Compton, K., and Mateas, M. 2015. Casual Creators. In Proceedings of the Sixth International Conference on Computational Creativity June, 228.

[Giacomello, Lanzi, and Loiacono 2018] Giacomello, E.; Lanzi, P. L.; and Loiacono, D. 2018. DOOM Level Generation Using Generative Adversarial Networks. In NeurIPS.

[Goodfellow et al. 2014] Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative Adversarial Nets. In NeurIPS.

[Härkönen et al. 2020] Härkönen, E.; Hertzmann, A.; Lehtinen, J.; and Paris, S. 2020. GANSpace: Discovering Interpretable GAN Controls. arXiv:2004.02546.

[Karras et al. 2018] Karras, T.; Aila, T.; Laine, S.; and Lehtinen, J. 2018. Progressive Growing Of GANs For Improved Quality, Stability, And Variation. In ICLR 2018.

[Karras, Laine, and Aila 2019] Karras, T.; Laine, S.; and Aila, T. 2019. A Style-Based Generator Architecture For Generative Adversarial Networks. In CVPR.

[Kingma and Welling 2014] Kingma, D. P., and Welling, M. 2014. Auto-Encoding Variational Bayes. In Bengio, Y., and LeCun, Y., eds., In ICLR 2014.

[Radford, Metz, and Chintala 2016] Radford, A.; Metz, L.; and Chintala, S. 2016. Unsupervised Representation Learning With Deep Convolutional Generative Adversarial Networks. In Bengio, Y., and LeCun, Y., eds., In ICLR 2016.

[Salimans et al. 2017] Salimans, T.; Karpathy, A.; Chen, X.; and Kingma, D. P. 2017. PixelCNN++: Improving ThePixelCNN With Discretized Logistic Mixture Likelihood And Other Modifications. In In ICLR 2017.

[Shen et al. 2019] Shen, Y.; Gu, J.; Tang, X.; and Zhou, B. 2019. Interpreting The Latent Space Of Gans For Semantic Face Editing. arXiv:1907.10786.
Figure 7: Additional examples of symbolic generated art.
Figure 8: Additional examples of neuro-symbolic generated art.
Figure 9: Additional examples of neuro-symbolic interpolations.