Modern Evolution Strategies for Creativity: Fitting Concrete Images and Abstract Concepts

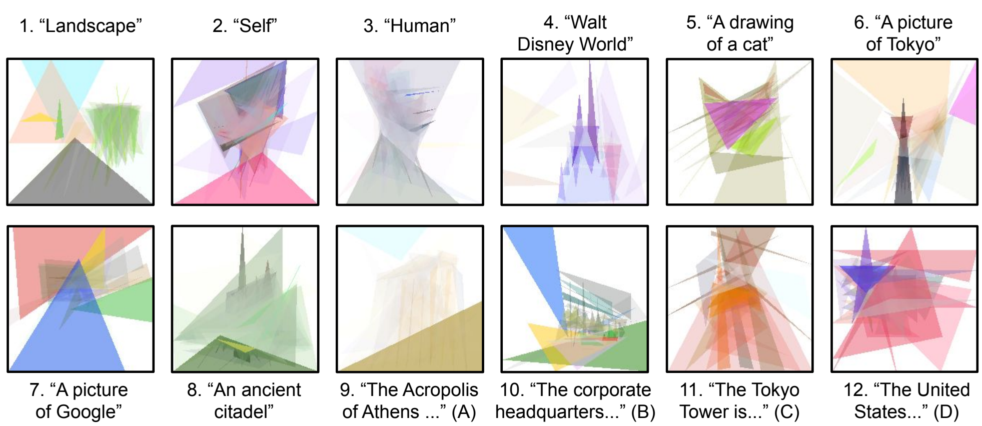
Yingtao Tian
Google Brain
alantian@google.com

David Ha
Google Brain
hadavid@google.com

Abstract

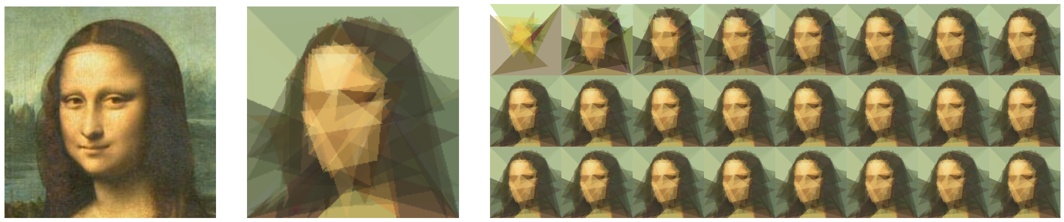
Evolutionary algorithms have been used in the digital art scene since the 1970s. A popular application of genetic algorithms is to optimize the procedural placement of vector graphic primitives to resemble a given painting. In recent years, deep learning-based approaches have also been proposed to generate procedural drawings, which can be optimized using gradient descent. In this work, we revisit the use of evolutionary algorithms for computational creativity. We find that modern evolution strategies (ES) algorithms, when tasked with the placement of shapes, offer large improvements in both quality and efficiency compared to traditional genetic algorithms, and even comparable to gradient-based methods. We demonstrate that ES is also well suited at optimizing the placement of shapes to fit the CLIP model, and can produce diverse, distinct geometric abstractions that are aligned with human interpretation of language. Videos and demo: https://es-clip.github.io/.

Staring from early 20th-century in the wider context of Modernism [29], a series of avant-garde art abandoned the depiction of objects from tradition rules of perspective and instead picking revolutionary, abstract point of views. Along this line were the Cubism art movement [41], where objects are analyzed by the artist, broken up, and reassembled in an abstract form consisting of geometric representations, and its successor, the geometric abstraction [8], which composed primitives that are either purely geometric and elementary. The impact is far-reaching: The use of simple geometry can be seen as one of styles found in the abstract expressionism [39] where artists expressed their subconscious or impulsive feelings. It also helped shaped the minimalist art [37] and architecture [42] movements, in which everything is stripped down to its essential quality to achieve simplicity [3].



A. “The Acropolis of Athens is an ancient citadel located on a rocky outcrop above the city of Athens and contains the remains of several ancient buildings including the Parthenon.”
B. “The corporate headquarters complex of Google located at 1600 Amphitheatre Parkway in Mountain View, California.The corporate headquarters complex of Google located at 1600 Amphitheatre Parkway in Mountain View, California.”
C. “The Tokyo Tower is a communications and observation tower in the Shiba-koen district of Minato Tokyo Japan.”
D. “The United States of America commonly known as the United States or America is a country primarily located in North America.”

Figure 1: Our proposed painting synthesization places transparent triangles using an Evolution Strategy (ES). Each concept represented as a text prompt is accompanied by its corresponding synthesized image. Here, the fitness is defined as the cosine distance between rendered canvas and text, both embedded by CLIP [40], and we optimize the position and the color of triangles using ES, which produces diverse and distinct geometric abstractions.



“Mona Lisa” Evolved Results Evolution Process

Figure 2: We use modern ES (PGPE with ClipUp), with 50 triangles and running evolution for $10,000$ steps to fit the target image “Mona Lisa”.

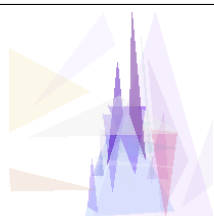
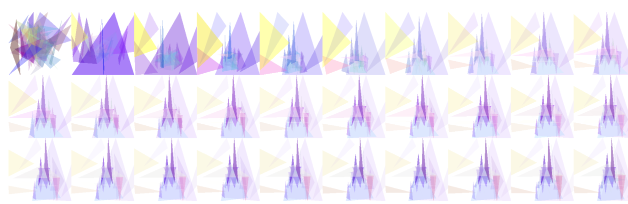
Prompt	**Evolved Results**	**Evolution Process**
“Self”		
“Walt Disney World”		

Figure 3: Our method fits the concept represented in text prompt, using 50 triangles and running evolution for $2,000$ steps.

The idea of minimalist art has also been explored in computer art with a root in mathematical art [35]. Schmidhuber [45] proposed low-complexity art, as a minimal art form that attempts to depict the essence of an object by making use of ideas from algorithmic complexity [28]. Similarly, algorithmic art [52], and in a broad sense genetic algorithm, proposed to generate arts using the algorithm and/or rules designed by the artist. As one example, a basic genetic algorithm using evolution has been proposed [2, 25] to represent a target image using semi-transparent, overlapping triangles and has gained popularity over the years with the creative coding community, resulting in a number of sophisticated extensions [6, 11, 47]. Interestingly, since these methods are iterative, they also echo Process art [49] which emphasizes the process of making. More discussion regarding the related works can be found in Appendix A.

With the recent resurgence of interest in evolution strategies (ES) in the machine learning community [17, 43], in this work, we revisit the use of ES for creativity applications as an alternative to gradient-based methods (details in Appendix B). For the image-approximation with shapes task, illustrated in Figure 2, we find that modern ES algorithms offer large improvements in both quality and efficiency when compared to traditional genetic algorithms, and as we will also demonstrate, even comparable to state-of-the-art differentiable rendering methods [30] (details in Appendix C). We show that ES is also well suited at optimizing the placement of shapes to fit the CLIP [40] model targeting concepts in the form of text, as illustrated in Figure 3, and can produce diverse, distinct geometric abstractions that are aligned with human interpretation of language (details in Appendix D). In both tasks, we study the effect of a wide range of target images and concepts to choose from, the number of triangles as an computational and artistic budget, and the behavior of other method compared to differentiable method optimized with gradient descent. We also explore differences from multiple runs which may be helpful for artists in the loop. Interestingly, the results produced by our method resemble Abstract expressionism [39] and Minimalist art [37, 42], and opens the door for the further studying into computer generated art (discussion in Appendix E). We provide a reference code implementation of our approach so that it can be a useful tool in the computational artist’s toolbox.

Ethical Implications. This project itself does not pose ethical concerns *per se*. Using the framework provided in this project, a user could let the method produce art works fitting on the provided target image or text prompt, both under the sole control of the user. However since we use CLIP to interpret the meaning of text provided by the user, the intrinsic bias in CLIP itself, as identified by its very developers [1], may impact the generated art. Although such impact is hard to anticipate, partially due to the abstract and indirect nature of the generated art, we still believe it is worthy of consideration.

References

[1] Sandhini Agarwal, Gretchen Krueger, Jack Clark, Alec Radford, Jong Wook Kim, and Miles Brundage. Evaluating clip: Towards characterization of broader capabilities and downstream implications. *arXiv preprint arXiv:2108.02818*, 2021.

[2] Alteredqualia. Evolution of mona lisa in javascript and canvas, 2008. URL https://alteredqualia.com/visualization/evolve/.

[3] Franco Bertoni. *Minimalist architecture*. Birkhauser, 2002. URL https://www.worldcat.org/title/minimalist-architecture/oclc/469368801?loc=.

[4] Hans-Georg Beyer. *The theory of evolution strategies*. Springer Science & Business Media, 2001. URL https://www.springer.com/gp/book/9783540672975.

[5] Hans-Georg Beyer and Hans-Paul Schwefel. Evolution strategies–a comprehensive introduction. *Natural computing*, 1(1):3–52, 2002.

[6] Kenny Cason. Genetic draw, 2016. URL https://github.com/kennycason/genetic_draw.

[7] Wikimedia Commons. File:040 okertalsperre.jpg — wikimedia commons the free media repository, 2020. URL https://commons.wikimedia.org/w/index.php?title=File:040_Okertalsperre.jpg&oldid=496749636. [Online; accessed 25-August-2021].

[8] Magdalena Dabrowski. Geometric abstraction, 2004. URL https://www.metmuseum.org/toah/hd/geab/hd_geab.htm.

[9] Chrisantha Fernando. Royal academy summer exhibition 2021 submission, 2021. URL https://www.chrisantha.co.uk/post/royal-academy-summer-exhibition-2021-submission.

[10] Chrisantha Fernando, SM Eslami, Jean-Baptiste Alayrac, Piotr Mirowski, Dylan Banarse, and Simon Osindero. Generative art using neural visual grammars and dual encoders. *arXiv preprint arXiv:2105.00162*, 2021. URL https://arxiv.org/abs/2105.00162.

[11] Michael Fogleman. Primitive pictures, 2016. URL https://github.com/fogleman/primitive.

[12] Kevin Frans, LB Soros, and Olaf Witkowski. Clipdraw: Exploring text-to-drawing synthesis through language-image encoders. *arXiv preprint arXiv:2106.14843*, 2021. URL https://arxiv.org/abs/2106.14843.

[13] Federico A Galatolo, Mario GCA Cimino, and Gigliola Vaglini. Generating images from caption and vice versa via clip-guided generative latent space search. *arXiv preprint arXiv:2102.01645*, 2021. URL https://arxiv.org/abs/2102.01645.

[14] Yaroslav Ganin, Tejas Kulkarni, Igor Babuschkin, SM Ali Eslami, and Oriol Vinyals. Synthesizing programs for images using reinforced adversarial learning. In *International Conference on Machine Learning*, pages 1666–1675. PMLR, 2018. URL http://proceedings.mlr.press/v80/ganin18a.html.

[15] Robert A. Gonsalves. Ganscapes: Using ai to create new impressionist paintings, 2021. URL https://towardsdatascience.com/ganscapes-using-ai-to-create-new-impressionist-paintings-d6af1cf94c56.

[16] Karol Gregor, Ivo Danihelka, Alex Graves, Danilo Rezende, and Daan Wierstra. Draw: A recurrent neural network for image generation. In *International Conference on Machine Learning*, pages 1462–1471. PMLR, 2015. URL http://proceedings.mlr.press/v37/gregor15.html.

[17] David Ha. Evolving stable strategies. *blog.otoro.net*, 2017. URL http://blog.otoro.net/2017/11/12/evolving-stable-strategies/.

[18] David Ha and Douglas Eck. A neural representation of sketch drawings. *arXiv preprint arXiv:1704.03477*, 2017. URL https://arxiv.org/abs/1704.03477.

[19] Inman Harvey. The microbial genetic algorithm. In *European conference on artificial life*, pages 126–133. Springer, 2009. URL http://users.sussex.ac.uk/~inmanh/MicrobialGA_ECAL2009.pdf.

[20] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9 (8):1735–1780, 1997. URL https://direct.mit.edu/neco/article/9/8/1735/6109/Long-Short-Term-Memory.

[21] Forrest Huang and John F Canny. Sketchforme: Composing sketched scenes from text descriptions for interactive applications. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pages 209–220, 2019. URL https://dl.acm.org/doi/10.1145/3332165.3347878.

[22] Forrest Huang, Eldon Schoop, David Ha, and John Canny. Scones: towards conversational authoring of sketches. In *Proceedings of the 25th International Conference on Intelligent User Interfaces*, pages 313–323, 2020. URL https://dl.acm.org/doi/10.1145/3377325.3377485.

[23] Zhewei Huang, Wen Heng, and Shuchang Zhou. Learning to paint with model-based deep reinforcement learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 8709–8718, 2019. URL https://openaccess.thecvf.com/content_ICCV_2019/html/Huang_Learning_to_Paint_With_Model-Based_Deep_Reinforcement_Learning_ICCV_2019_paper.html.

[24] Chao Jia, Yinfei Yang, Ye Xia, Yi-Ting Chen, Zarana Parekh, Hieu Pham, Quoc V Le, Yunhsuan Sung, Zhen Li, and Tom Duerig. Scaling up visual and vision-language representation learning with noisy text supervision. *arXiv preprint arXiv:2102.05918*, 2021. URL https://arxiv.org/abs/2102.05918.

[25] Roger Johansson. Genetic programming: Evolution of mona lisa, 2008. URL https://rogerjohansson.blog/2008/12/07/genetic-programming-evolution-of-mona-lisa/.

[26] Hiroharu Kato, Deniz Beker, Mihai Morariu, Takahiro Ando, Toru Matsuoka, Wadim Kehl, and Adrien Gaidon. Differentiable rendering: A survey. *arXiv preprint arXiv:2006.12057*, 2020. URL https://arxiv.org/abs/2006.12057.

[27] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. URL https://arxiv.org/abs/1412.6980.

[28] Andrei Kolmogorov. Three approaches to the quantitative definition of information. *Problems of information transmission*, 1(1):1–7, 1965. URL https://www.tandfonline.com/doi/abs/10.1080/00207166808803030?journalCode=gcom20.

[29] Kathleen Kuiper. Modernism, 2021. URL https://www.britannica.com/art/Modernism-art.

[30] Samuli Laine, Janne Hellsten, Tero Karras, Yeongho Seol, Jaakko Lehtinen, and Timo Aila. Modular primitives for high-performance differentiable rendering. *ACM Transactions on Graphics (TOG)*, 39(6):1–14, 2020. URL https://arxiv.org/abs/2011.03277.

[31] Tzu-Mao Li, Michal Lukáč, Michaël Gharbi, and Jonathan Ragan-Kelley. Differentiable vector graphics rasterization for editing and learning. *ACM Transactions on Graphics (TOG)*, 39(6): 1–15, 2020. URL https://people.csail.mit.edu/tzumao/diffvg/.

[32] Aristid Lindenmayer. Mathematical models for cellular interactions in development i. filaments with one-sided inputs. *Journal of theoretical biology*, 18(3):280–299, 1968. URL https://www.sciencedirect.com/science/article/abs/pii/0022519368900799.

[33] Songhua Liu, Tianwei Lin, Dongliang He, Fu Li, Ruifeng Deng, Xin Li, Errui Ding, and Hao Wang. Paint transformer: Feed forward neural painting with stroke prediction. *arXiv preprint arXiv:2108.03798*, 2021. URL https://arxiv.org/abs/2108.03798.

[34] Raphael Gontijo Lopes, David Ha, Douglas Eck, and Jonathon Shlens. A learned representation for scalable vector graphics. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 7930–7939, 2019. URL

https://openaccess.thecvf.com/content_ICCV_2019/html/Lopes_A_Learned_Representation_for_Scalable_Vector_Graphics_ICCV_2019_paper.html.

[35] Joseph Malkevitchn. Mathematics and art, 2003. URL https://www.ams.org/publicoutreach/feature-column/fcarc-art1.

[36] John FJ Mellor, Eunbyung Park, Yaroslav Ganin, Igor Babuschkin, Tejas Kulkarni, Dan Rosenbaum, Andy Ballard, Theophane Weber, Oriol Vinyals, and SM Eslami. Unsupervised doodling and painting with improved spiral. *arXiv preprint arXiv:1910.01007*, 2019. URL https://arxiv.org/abs/1910.01007.

[37] Tate Modern. Minimalism, 2018. URL https://www.tate.org.uk/art/art-terms/m/minimalism.

[38] Reiichiro Nakano. Neural painters: A learned differentiable constraint for generating brushstroke paintings. *arXiv preprint arXiv:1904.08410*, 2019. URL https://arxiv.org/abs/1904.08410.

[39] Stella Paul. Abstract expressionism, 2004. URL https://www.metmuseum.org/toah/hd/abex/hd_abex.htm.

[40] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. *arXiv preprint arXiv:2103.00020*, 2021. URL https://arxiv.org/abs/2103.00020.

[41] Sabine Rewald. Heilbrunn timeline of art history: Cubism. *The Metropolitan Museum of Art*, 2014. URL https://www.metmuseum.org/toah/hd/cube/hd_cube.htm.

[42] Barbara Rose. Abc art. *Art in America*, 53(5):57–69, 1965. URL https://www.artnews.com/art-in-america/features/abc-art-barbara-rose-1234580665/.

[43] Tim Salimans, Jonathan Ho, Xi Chen, Szymon Sidor, and Ilya Sutskever. Evolution strategies as a scalable alternative to reinforcement learning. *arXiv preprint arXiv:1703.03864*, 2017. URL https://arxiv.org/abs/1703.03864.

[44] @samburtonking. Introduction to vqgan+clip, 2021. URL https://docs.google.com/document/d/1Lu7XPRKlNhBQjcKr8k8qRzUzbBW7kzxb5Vu72GMRn2E/edit.

[45] Jürgen Schmidhuber. Low-complexity art. *Leonardo*, 30(2):97–103, 1997. URL https://www.idsia.ch/~juergen/locoart/locoart.html.

[46] Frank Sehnke, Christian Osendorfer, Thomas Rückstieß, Alex Graves, Jan Peters, and Jürgen Schmidhuber. Parameter-exploring policy gradients. *Neural Networks*, 23(4):551–559, 2010. URL https://people.idsia.ch//~juergen/nn2010.pdf.

[47] Shahriar Shahrabi. Procedural paintings with genetic evolution algorithm, 2020. URL https://github.com/IRCSS/Procedural-painting.

[48] "Sizigi Studios". Waifu labs, 2019. URL https://waifulabs.com/.

[49] Tate. Process art, 2021. URL https://www.tate.org.uk/art/art-terms/p/process-art.

[50] Nihat Engin Toklu. Pgpelib. https://github.com/nnaisense/pgpelib, 2020. URL https://github.com/nnaisense/pgpelib.

[51] Nihat Engin Toklu, Paweł Liskowski, and Rupesh Kumar Srivastava. Clipup: A simple and powerful optimizer for distribution-based policy evolution. In *International Conference on Parallel Problem Solving from Nature*, pages 515–527. Springer, 2020. URL https://arxiv.org/abs/2008.02387.

[52] Roman Verostko. Algorithmic art, 1994. URL http://www.verostko.com/algorithm.html.

[53] Phil Wang. Big sleep: A simple command line tool for text to image generation, using openai's clip and a biggan, 2021. URL https://github.com/lucidrains/big-sleep.

[54] Phil Wang. Deep daze: A simple command line tool for text to image generation using openai's clip and siren (implicit neural representation network), 2021. URL https://github.com/lucidrains/deep-daze.

[55] Tom White. Shared visual abstractions. *arXiv preprint arXiv:1912.04217*, 2019. URL https://arxiv.org/abs/1912.04217.

[56] Ningyuan Zheng, Yifan Jiang, and Dingjiang Huang. Strokenet: A neural painting environment. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=HJxwDiActX.