Seizing the Means of Production: Exploring the Landscape of Crafting, Adapting and Navigating Generative AI Models

AHMED M. ABUZURAIQ, Simon Fraser University, Canada
PHILIPPE PASQUIER, Simon Fraser University, Canada

In this paper, we map out the landscape of options available to visual artists for creating personal artworks, including crafting, adapting and navigating deep generative models. Following that, we argue for revisiting model crafting, defined as the design and manipulation of generative models for creative goals, and motivate studying and designing for model crafting as a creative activity in its own right.

Additional Key Words and Phrases: generative AI, generative models, model crafting, visual arts, personalization

ACM Reference Format:
Ahmed M. Abuzuraiq and Philippe Pasquier. 2024. Seizing the Means of Production: Exploring the Landscape of Crafting, Adapting and Navigating Generative AI Models. 1, 1 (March 2024), 7 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Large-scale Text-to-image Generation Models (LTGMs), including Dalle-E [37] and Stable Diffusion [41], are trained on large datasets with many high-end GPUs. Users of these models can produce diverse and high-quality visuals through meticulously written text prompts. These models mark a significant shift from the era when artists used personally-trainable generative models like Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs). LTGMs have made the generation of visuals accessible to everyone, but this shift in attention has overshadowed the practice of "model crafting", whereas artists personalize their work by experimenting with training sets, model architectures, and hyperparameters in addition to combining, adapting and manipulating pre-trained models [13]. Model crafting offered artists a sense of craftsmanship and ownership over the creative process and its outcomes. However, the high resource requirements and costs associated with training LTGMs have made them practically impossible for individual artists to train[1], and replaced process ownership with a contested product ownership [24]. In this paper, we map out the landscape of options available to artists for creating personal artworks, on LTGMs or personally-trainable models, and argue for revisiting model crafting as a significant venue for artistic personalization.

[1]For example, a recent work by Chen et al. [17] propose a model that can be trained with a fraction of the cost it takes to train a large model such as Stable Diffusion V1.5. However this fraction amounts to $26,000.

Authors' addresses: Ahmed M. Abuzuraiq, aabuzura@sfu.ca, Simon Fraser University, Canada; Philippe Pasquier, pasquier@sfu.ca, Simon Fraser University, Canada.

Table 1. Finer controls for personalizing generative models increase as we move (down) between navigating, adapting or crafting models, it also increases as we move (right) from training on existing, curated or manually created datasets.

	Personalization Method	Data			Parts Updated		Model Training		Systems	
		Existing	Curated	Created	Weights	Network	From Scratch	Pre-trained	Example Methods Used	No-Code
Navigate	Prompting, sampling, interpolating, ..etc	moving down (↓) or right (→) here increases creative control and needed effort.			✗	✗	✗	✓	Img2img [1] or Latent Projection [33]	✓
Adapt	Few-shot model adaptation				✓	✗	✗	✓	Textual Inversion [1, 2, 6]	✓
	Model fine-tuning				✓	✗	✗	✓	Train from a pre-trained model [8, 33]	✓
	Train an off-the-shelf model				✓	✗	✓	✗	Picking a GAN variation [8, 33]	✓
Craft	Active Divergence				✓	✓	✓	✓	Network Blending [33]	✓
	Train on a modified model				✓	✓	✓	✗	Customize standardized models [38, 45]	✗
	Train on own model				✓	✓	✓	✗	Deep Learning frameworks [9, 39]	✗

2 THE LANDSCAPE OF PERSONALIZATION

In a 2018 interview [57], computational artist Memo Akten outlined a spectrum of approaches for artistic creation with deep generative models along two dimensions: data (creating one's own, curating, or reusing existing sets) and models (designing original algorithms, modifying existing ones, or using pre-trained models). Having tried every combination, Akten argued that as artists move from using a custom to existing models/data in their work, it becomes "harder to give it a unique spin and make it your own". Inspired by this analysis, we explore the current landscape of options that artists have for creating personal artworks, with findings detailed in Table 1 and discussed later.

2.1 Navigating Generative Spaces

By interacting with a generative model, artists navigate the generative space of the model, i.e. its space of possibilities, to find aesthetics that interest them. Navigation can take place through prompting (e.g., [5, 37]), latent projection and semantic sliders (e.g., on Autolume [33]), canvas-like [43] and paining-like [18] interactions, or by embedding generative models into bespoke workflows [2, 22], among others [46]. Steering interfaces that provide semantically meaningful controls improve artists' control, self-efficacy, and creative ownership [31]. We contend that navigation becomes more effective when it takes place in the generative space of expressive [32] and personalized models [11].

2.2 Adapting Generative Spaces

When navigating the model's generative space proves cumbersome or if it consistently fails to produce personal results, artists can adapt generative models, which helps in shaping and focusing the scope of navigation. Few-shot model adaptation relies on a handful of example images and can introduce novel concepts into the generative space of the model with minimal changes. Full fine-tuning (and training from scratch) requires more example images but it can produce visual results for domains that are difficult to model with a few images such as Arabic Calligraphy [49]. Numerous techniques for adapting and personalizing generative models exist [25, 27, 42, 44]. Furthermore, visual arts creation

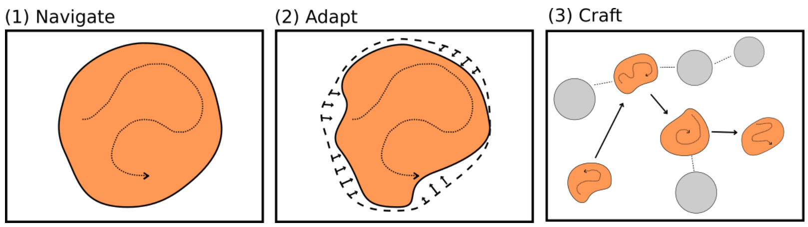

Fig. 1. Three modes for creating personal works using generative models: (1) model navigation, (2) model adaption, and (3) model crafting. In the first mode, artists navigate within a fixed generative space to find the aesthetics they like, e.g. by prompting. In the second, navigation takes places within an adapted generative space. In the third, artists (co-)explore the space of generative models (straight arrows) as well as navigating within the generative space of each model (curved arrows).

tools (e.g., [1–4, 7, 8, 33]) often integrate adaption into their interfaces, whether they operate on personally-trainable or large generative models, such that the adapted models become available for downstream navigation [11].

2.3 Crafting Generative Spaces

During model navigation a generative model is not modified, instead artists only sample from its generative space (Figure 1 - 1). During model adaption, the weights of the model are changed which results in a different generative space to navigate in (Figure 1 - 2). On the other hand, model crafting involves manipulating the model's architecture including designing new architectures, or editing and mixing existing models (c.f. in Active Divergence), which can produce qualitatively different and more personal generative spaces than is possible with model adaption alone, by virtue of affording finer control. Conceptually, model crafting can be defined as a creative exploration in the design space of generative models (Figure 1 - 3). Given the high computational cost of training large models, model crafting for artists is limited to small models in practice, noting that "small" depends on the artist's access to computational resources.

2.3.1 Active Divergence. Artists often seek novelty and surprise from generative models. However, generative deep learning models are, by definition, trained to model a target distribution. Active divergence methods [13], as the name indicates, are a collection of methods that diverge generative models from a target distribution to produce novel and out-of-distribution samples. Active divergence methods, such as network rewriting [12], bending [14], and blending [40], involve re-purposing and manipulating pre-trained models for art creation, with or without data.

2.3.2 Model Crafting vs. Model Development. Developing generative models, and machine learning models in general, is a mature field with support tools that span the entire development lifecycle including data preparation and pre-processing, hyperparameters optimization, model version control, and training process tracking as well as tools for testing and deploying models [15, 53]. Artists can lean into this ecosystem in developing generative models for creative applications, but the emphasis on efficiency, generality, and scalability, around which those tools are built is not always relevant for art creation. In fact, when it comes to art creation, the bias and overfitting associated with small datasets or simple models can be desired [54]. Art creation involves experimentation, visual exploration, and reflection and is characterized by the lack of optimal solutions, which is why we prefer the term Model Crafting over Model Development. If we concede that crafting for creative endeavours is different from model development for common purposes, we can deduce that it also requires different tool support. As Table 1 and previous sections show, model navigation and adaption for the visual arts are well supported with no-code tools while crafting largely happens on coding editors and paper sketches [55]. A notable exception is Autolume [33] which is a no-code visual synthesis system that combines model navigation (via semantic sliders and latent projection), adaption (fine-tuning), and crafting (via network blending).

3 DESIGNING FOR MODEL CRAFTING

3.1. Challenges. Model crafters and researchers who design tools for them face multiple challenges including: (1) the prolonged training times for training generative models, and (2) the interdisciplinary and technical expertise required for model crafting. However, we join Abuzuraiq and Pasquier [11] in arguing that the proliferation of cloud computing and the advances on data-efficient, limited-data and few-shot generative models [10, 30, 35, 56] can lead high-quality models that are faster to train (technical-perspective). Furthermore, that the adoption of "small data" [54] and "slow technology" [26] mindsets can provide effective lenses for the design of model crafting tools (design-perspective).

3.2. Opportunities. Large generative AI systems impact the artist's ability to perceive generated works as their own (i.e. authorship), and the systems' black-boxed nature impacts the artist's sense of control over the results (i.e., agency). In this context, model crafting offers the following distinct advantages:

A. Owning both the Creative Process and its Products: Creative AI, epitomized by recent generative AI systems, emphasizes the products over processes [19]. When people are presented with a generative AI system that creates impressive results without offering them means to understand or contribute to its process, their agency is decreased and they lose an opportunity to reflect on their own creative process, in addition to understanding human and computational creativity at large. The advent of large text-to-image generative models raised questions over the attribution (authorship) of what is produced and challenged artist's sense of ownership (agency) when using those models, possibly due to the incurred loss of control [50]. To reclaim agency over the created products, artists personalize and adapt generative models. However, product ownership (and authorship) can also come as a by-product of process ownership, which can be achieved through model crafting along with carefully curated or created datasets.

B. Sense of Craftsmanship: Model crafting is challenging and requires weaving expertise that spans coding, deep learning, and specific domain expertise. But it is also a rewarding activity that can be a source of pride and craftsmanship, which is a motivation they share with makers in the Do-It-Yourself (DIY) community [34]. Crafters of generative systems also often report surprising [23] and serendipitous [29] encounters as they experiment with generative systems.

3.3. Creativity Support for Model Crafting: Given the advantages outlined above, we encourage researchers to study and design tools for the model crafters. The work in this direction is scarce, so we weave together some starting points. First, by analogy to sketching in design, Lam et al [28], propose a system for Model Sketching by which users explore high-level concepts (e.g. profanity or racism) based on which ML classifiers are subsequently created. In the same vein of simplifying ML model design for non-experts, multiple visual programming frameworks for deep learning are introduced [16, 51], as well as direct manipulation tools for visualizing and experimenting with deep learning models [48]. To reduce the barrier for non-expert users, Croisdale et al. [21] propose a data-flow system of cards for exploring multi-modal generative workflows. Similarly, Compton et al. [20] present physical cards for creating generative pipelines. Finally, crafting involves finer control including mixing and matching. Ulberg et al. [52] suggest visualizing and crafting model weights for better control over neural networks for the arts. Shimizu et al. [47] enable creating bespoke text-to-media mappings between generative models with a few examples.

4 CONCLUSION & FUTURE WORK

Crafting, adapting and using generative systems has a long and diverse lineage in many academic fields and communities of practice, even if not under the term Generative AI or if not based on machine learning [2]. Other names included procedural content generation in games, generative design in architectural or industrial design, generative music, and computational arts to name a few. With roots in extensive bodies of work on generative and procedural systems, model crafting can also extend its reach to other fields. Artists have always pioneered advances in the AI field (e.g. DeepDream [36] as a predecessor for visualizing neural networks activations), and work on creative model crafting for small models can inform the design and development of large generative models as well. In the future, we will continue to design for and study model crafting as a creative endeavour in its own right.

REFERENCES

[1] Automatic1111 WebUI. https://github.com/AUTOMATIC1111/stable-diffusion-webui, 2024.
[2] ComfyAI. https://github.com/comfyanonymous/ComfyUI, 2024.
[3] InvokeAI (community-version). https://github.com/invoke-ai/InvokeAI, 2024.
[4] InvokeAI (industry-version). https://invoke.ai/, 2024.
[5] MidJourney. https://www.midjourney.com/, 2024.
[6] Photobooth. https://openart.ai/photobooth, 2024.
[7] Playform: No-Code AI for Creative People. https://www.playform.io/train, 2024.
[8] RunwayML. https://runwayml.com/ai-magic-tools/ai-training/, 2024.
[9] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015.
[10] Abdollahzadeh, M., Malekzadeh, T., Teo, C. T. H., Chandrasegaran, K., Liu, G., and Cheung, N.-M. A survey on generative modeling with limited data, few shots, and zero shot. *ArXiv abs/2307.14397* (2023).
[11] Abuzuraiq, A. M., and Pasquier, P. Towards Personalizing Generative AI with Small Data for Co-Creation in the Visual Arts. In *HAI-GEN 2024: 4th Workshop on Human-AI Co-Creation* (South Carolina, USA, 2024).
[12] Bau, D., Liu, S., Wang, T., Zhu, J.-Y., and Torralba, A. Rewriting a deep generative model. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16* (2020), Springer, pp. 351–369.
[13] Broad, T., Berns, S., Colton, S., and Grierson, M. Active divergence with generative deep Learning–A survey and taxonomy. In *Proceedings of the 12th International Conference on Computational Creativity (ICCC '21)* (2021).
[14] Broad, T., Leymarie, F. F., and Grierson, M. Network bending: Expressive manipulation of deep generative models. In *Artificial Intelligence in Music, Sound, Art and Design: 10th International Conference, EvoMUSART 2021, Held as Part of EvoStar 2021, Virtual Event, April 7–9, 2021, Proceedings 10* (2021), Springer, pp. 20–36.
[15] Budras, T., Blanck, M., Berger, T., and Schmidt, A. An in-depth comparison of experiment tracking tools for machine learning applications. *International Journal on Advances in Software Volume 15, Number 3 & 4, 2022* (2022).
[16] Calò, T., and De Russis, L. Towards a visual programming tool to create deep learning models. In *Companion Proceedings of the 2023 ACM SIGCHI Symposium on Engineering Interactive Computing Systems* (New York, NY, USA, 2023), EICS '23 Companion, Association for Computing Machinery.
[17] Chen, J., Yu, J., Ge, C., Yao, L., Xie, E., Wang, Z., Kwok, J., Luo, P., Lu, H., and Li, Z. PixArt-α: Fast training of diffusion transformer for photorealistic text-to-image synthesis. In *The Twelfth International Conference on Learning Representations* (2024).
[18] Chung, J. J. Y., and Adar, E. PromptPaint: Steering text-to-image generation through paint medium-like interactions. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (2023), pp. 1–17.
[19] Colton, S. From computational creativity to creative AI and back again. *Interalia Magazine* (2019).
[20] Compton, K., and Mateas, M. A generative framework of generativity. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* (2017), vol. 13, pp. 41–48.
[21] Croisdale, G. T., Chung, J. J. Y., Huang, E., Birchmeier, G., Wang, X., and Guo, A. DeckFlow: A Card Game Interface for Exploring Generative Model Flows. In *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco CA USA, Oct. 2023), ACM, pp. 1–3.
[22] Derivative. TouchDesigner: A node-based visual programming language for real-time interactive multimedia content. https://derivative.ca/, 2024.

[2] Generative systems can be rule-based, probabilistic or constraint satisfaction systems, or they can use genetic search algorithms, and machine learning models (specifically generative deep learning)

[23] Elgammal, A. Text-to-Image Generators Have Altered the Digital Art Landscape—But Killed Creativity. Here's Why an Era of A.I. Art Is Over. https://news.artnet.com/art-world-archives/ahmed-elgammal-op-ed-ai-art-is-over-2304028, June 2023.

[24] Epstein, Z., Hertzmann, A., Creativity, t. I. o. H., Akten, M., Farid, H., Fjeld, J., Frank, M. R., Groh, M., Herman, L., Leach, N., Mahari, R., Pentland, A. S., Russakovsky, O., Schroeder, H., and Amy Smith. Art and the Science of Generative AI. *Science (New York, N.Y.) 380*, 6650 (2023), 1110–1111.

[25] Gal, R., Alaluf, Y., Atzmon, Y., Patashnik, O., Bermano, A. H., Chechik, G., and Cohen-Or, D. An image is worth one word: Personalizing text-to-image generation using textual inversion. *arXiv preprint arXiv:2208.01618* (2022).

[26] Hallnäs, L., and Redström, J. Slow technology – designing for reflection. *Personal Ubiquitous Comput. 5*, 3 (Jan. 2001), 201–212.

[27] Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., and Chen, W. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685* (2021).

[28] Lam, M. S., Ma, Z., Li, A., Freitas, I., Wang, D., Landay, J. A., and Bernstein, M. S. Model Sketching: Centering Concepts in Early-Stage Machine Learning Model Design. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2023), CHI '23, Association for Computing Machinery, pp. 1–24.

[29] Lehman, J., Clune, J., Misevic, D., Adami, C., Altenberg, L., Beaulieu, J., Bentley, P. J., Bernard, S., Beslon, G., Bryson, D. M., et al. The surprising creativity of digital evolution: A collection of anecdotes from the evolutionary computation and artificial life research communities. *Artificial life 26*, 2 (2020), 274–306.

[30] Li, Z., Wu, X., Xia, B., Zhang, J., Wang, C., and Li, B. A comprehensive survey on data-efficient GANs in image generation. *ArXiv abs/2204.08329* (2022).

[31] Louie, R., Coenen, A., Huang, C. Z., Terry, M., and Cai, C. J. Novice-AI Music Co-Creation via AI-Steering Tools for Deep Generative Models. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2020), CHI '20, Association for Computing Machinery, pp. 1–13.

[32] Louie, R., Engel, J., and Huang, C.-Z. A. Expressive communication: Evaluating developments in generative models and steering interfaces for music creation. In *27th International Conference on Intelligent User Interfaces* (2022), pp. 405–417.

[33] Metacreation Lab. Autolume: A Neural-network based Visual Synthesizer. https://www.metacreation.net/autolume, 2024.

[34] Milne, A. What makes a maker: Common attitudes, habits and skills from the Do-It-Yourself (DIY) community. Master's thesis, Simon Fraser University, 2014.

[35] Moon, T., Choi, M., Lee, G., Ha, J.-W., and Lee, J. Fine-tuning diffusion models with limited data. In *NeurIPS Workshop on Score-Based Methods* (2022).

[36] Mordvintsev, A., Olah, C., and Tyka, M. Inceptionism: Going deeper into neural networks. https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html, 2015.

[37] OpenAI. DALL-E 3. https://openai.com/dall-e-3, 2024.

[38] Pal, A., and Das, A. TorchGAN: A flexible framework for GAN training and evaluation. *J. Open Source Softw. 6* (2019), 2606.

[39] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S. PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019, pp. 8024–8035.

[40] Pinkney, J. N. M., and Adler, D. Resolution dependent gan interpolation for controllable image synthesis between domains. In *Machine Learning for Creativity and Design Workshop* (34th Conference on Neural Information Processing Systems (NeurIPS 2020)., 2020).

[41] Podell, D., English, Z., Lacey, K., Blattmann, A., Dockhorn, T., Müller, J., Penna, J., and Rombach, R. SDXL: Improving Latent Diffusion Models for High-Resolution Image Synthesis, July 2023.

[42] Robb, E., Chu, W.-S., Kumar, A., and Huang, J.-B. Few-shot adaptation of generative adversarial networks. *arXiv preprint arXiv:2010.11943* (2020).

[43] Rost, M., and Andreasson, S. Stable walk: An interactive environment for exploring stable diffusion outputs. In *HAI-GEN 2023: 4th Workshop on Human-AI Co-Creation* (Sydney, Australia, 2023), Joint Proceedings of the ACM IUI Workshops 2023, pp. 89–97.

[44] Ruiz, N., Li, Y., Jampani, V., Pritch, Y., Rubinstein, M., and Aberman, K. Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023), pp. 22500–22510.

[45] Sankaran, A., Sinha, R., and Panwar, N. IBM GAN Toolkit. https://https://github. com/IBM/gan-toolki, 2018.

[46] Shi, J., Jain, R., Doh, H., Suzuki, R., and Ramani, K. An HCI-Centric survey and taxonomy of human-generative-AI interactions. *arXiv preprint arXiv:2310.07127* (2023).

[47] Shimizu, J., Olowe, I., Broad, T., Vigliensoni, G., Thattai Ravikumar, P., and Fiebrink, R. Interactive machine learning for generative models. In *Machine Learning for Creativity and Design Workshop* (2023).

[48] Smilkov, D., Carter, S., Sculley, D., Viégas, F. B., and Wattenberg, M. Direct-manipulation Visualization of Deep Networks. In *Proceedings of the 33rd International Conference on Machine Learning*. (New York, NY, USA, 2016).

[49] Sobhan Sarbandi, A. *Navigating the Latent: Exploring the Potentials of Islamic Calligraphy with Generative Adversarial Networks*. PhD thesis, OCAD University, 2021.

[50] Steinbrück, A., and Stankowski, A. Creative ownership and control for generative AI in art and design. In *In Generative AI in HCI Workshop, CHI '23* (2023).

[51] Tamilselvam, S. G., Panwar, N., Khare, S., Aralikatte, R., Sankaran, A., and Mani, S. A visual programming paradigm for abstract deep learning model development. In *Proceedings of the 10th Indian Conference on Human-Computer Interaction* (2019), pp. 1–11.

[52] Ulberg, E., Llach, D. C., and Byrne, D. Hand-crafting neural networks for art-making. In *Proceedings of the 11th International Conference on Computational Creativity (ICCC'20)* (2020).

[53] Vidal Domínguez, M. Tools for Deep Learning Experimentation. Master's thesis, Universidad Politécnica de Madrid, Madrid, 2023.

[54] Vigliensoni, G., Perry, P., and Fiebrink, R. A small-data mindset for generative AI creative work. In *In Generative AI in HCI Workshop, CHI '22* (New York, NY, USA, 2022), p. 5.

[55] Wongsuphasawat, K., Smilkov, D., Wexler, J., Wilson, J., Mane, D., Fritz, D., Krishnan, D., Viégas, F. B., and Wattenberg, M. Visualizing dataflow graphs of deep learning models in tensorflow. *IEEE transactions on visualization and computer graphics 24*, 1 (2017), 1–12.

[56] Yang, M., and Wang, Z. Image synthesis under limited data: A survey and taxonomy. *ArXiv abs/2307.16879* (2023).

[57] Zachariou, R. Machine Learning Art: An Interview With Memo Akten. https://www.artnome.com/news/2018/12/13/machine-learning-art-an-interview-with-memo-akten, Dec. 2018.

Received 26 February 2024