Autolume: A GAN-based No-Coding Small Data and Model Crafting Visual Synthesizer for Artistic Creation

Arshia Sobhan
Simon Fraser University
Vancouver, Canada
asobhans@sfu.ca

Ahmed M. Abuzuraiq
Simon Fraser University
Vancouver, Canada
aabuzura@sfu.ca

Philippe Pasquier
Simon Fraser University
Vancouver, Canada
pasquier@sfu.ca

Lionel Ringenbach
Ucodia
Vancouver, Canada
ucodia@live.com

Kris Krüg
TheUpgrade.ai
Vancouver, Canada
kk@theupgrade.ai

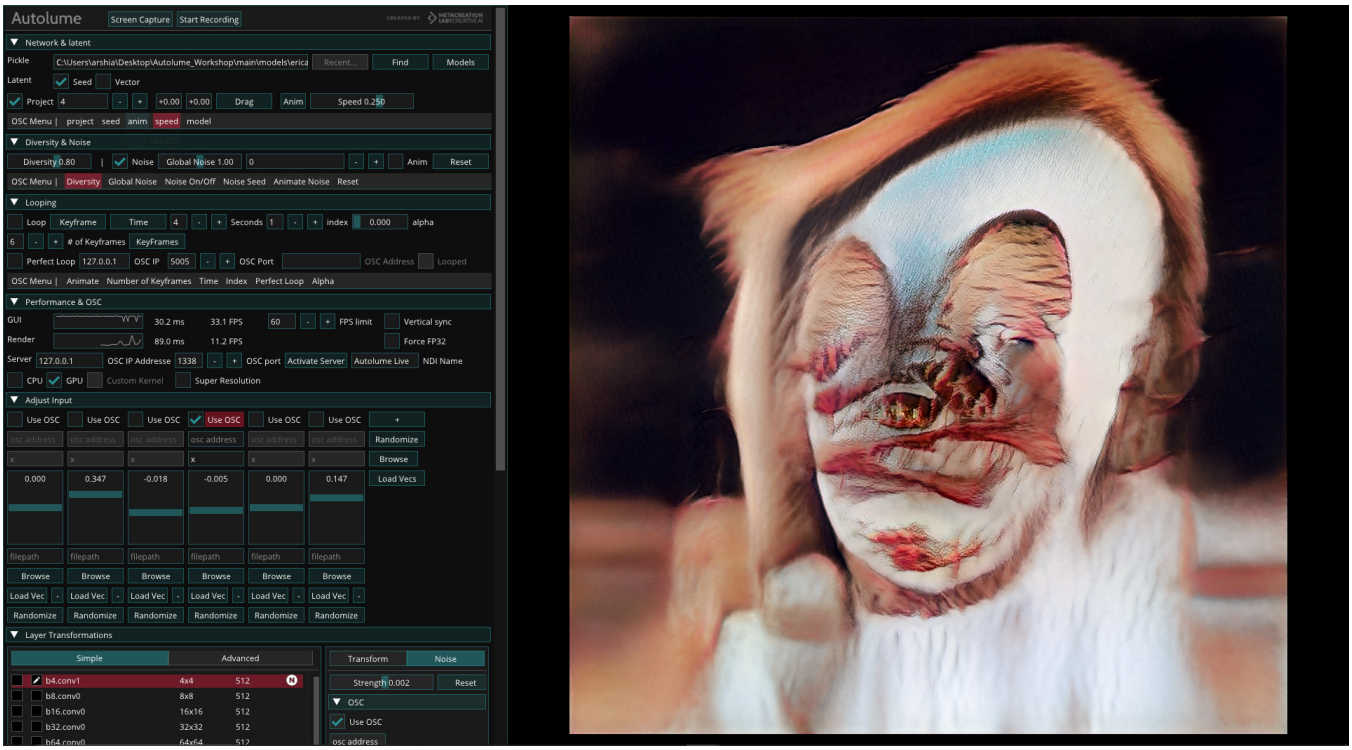

Figure 1: Screenshot of Autolume's real-time module displaying a model trained on the works of Erica Lapadat-Janzen.

Abstract

In this paper, we present a series of artworks created using Autolume, a no-coding, generative AI tool that leverages the artistic potential of Generative Adversarial Networks (GANs). Designed to empower artists without technical backgrounds, Autolume enables small-data model training, real-time latent space exploration, and interactive applications such as audio-reactive visuals. By preserving the fluidity and navigability of GAN latent spaces, Autolume offers unique possibilities for the creation of moving images and dynamic visual experiences. Through diverse case studies, including audiovisual performances, interactive installations, and small-dataset model crafting, we demonstrate how Autolume expands creative agency and supports new modes of artistic expression with generative AI.

Expanded 2025, Linz, Austria

ACM ISBN 979-8-4007-1532-7/25/09
https://doi.org/10.1145/3749893.3749968

CCS Concepts

• **Applied computing** → **Media arts**; • **Human-centered computing** → *Graphical user interfaces.*

Keywords

Generative AI, No-coding AI, Creative AI, Accessible AI, Small Data, Model Crafting

ACM Reference Format:

Arshia Sobhan, Ahmed M. Abuzuraiq, Philippe Pasquier, Lionel Ringenbach, and Kris Krüg. 2025. Autolume: A GAN-based No-Coding Small Data and Model Crafting Visual Synthesizer for Artistic Creation. In *Conference on Animation and Interactive Art (Expanded 2025), September 03–07, 2025, Linz, Austria.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3749893.3749968

1 Introduction

Recent advances in large-scale text-to-image diffusion models, such as DALL·E 3 by OpenAI [19] and Stable Diffusion by Stability AI [26], have significantly enhanced the ability to generate visuals from textual descriptions and led to the development of user-friendly tools that further democratize access to AI-driven art creation. Prior to that, small and personally-trainable models such as Generative Adversarial Networks (GANs) were more commonly used by AI artists with coding proficiency, such as Helena Sarin, Memo Akten, Mario Klingemann, and Anna Ridler, among others. [1]

Personal small models (or human-scale models as referred to by Tecks et al. [29]) offer a distinctive value for artists over large (diffusion) models, which warrant continued support and attention from the research community. Firstly, smaller models can be trained on smaller datasets, utilizing computing resources available to a wider range of artists. Working with smaller, personal or curated, datasets in tandem with personally-trainable models such as GANs offers a highly controllable and intimate artistic workflow that fosters creative ownership and provides a sense of craftsmanship [1, 29, 31]. Secondly, they avoid perpetuating biases associated with using large-scale models [31], which may be trained on datasets of artworks lacking proper attribution to creators. Thirdly, from an aesthetics point of view, models such as GANs provide a smoothly navigable and serendipitous latent space [6], offering artists opportunities to create expressive moving images through continuous interpolations and dynamic visual transitions, as well as a unique aesthetic of visual indeterminacy [9].

Fourthly, unlike large-scale text-based diffusion models, models such as GANs do not require the verbal articulation of tacit artistic intent [6, 14], which is important for some art forms (e.g. Arabic calligraphy) where there is an inherent lack of vocabulary for describing the pieces, whether due to the nature of the art form or because they are not present in training datasets. Similarly, the under-representation of some art forms in the training of large generative AI models may lead to inaccurate or culturally insensitive results. Lastly, AI artists work with algorithms and data as materials for creation [5], and so by virtue of being faster to train or fine-tune, variations of smaller models can become easier to explore[29]. Smaller models also offer high levels of creative control as they are amenable to hacking [2], adapting, and crafting, e.g. through different active divergence techniques [3].

In this paper, we present a series of artworks created with Autolume, a GAN-based art creation tool that embodies the ideas argued for earlier, including small-data and model crafting. To provide context for the artworks, we first introduce Autolume and its key features. The subsequent sections present a series of case studies that demonstrate how artists have engaged with Autolume across various creative contexts. We also briefly discuss workshops conducted to explore broader engagement with the artist community and to facilitate their use of Autolume.

2 Related Work

2.1 Text-to-Image Generation Tools

There are numerous text-to-image generative AI tools and cloud platforms. The open-source and free tools include Automatic1111 WebUI [2], ComfyAI [3], and InvokeAI [4]. The paid and freemium platforms, to name a few, include Midjourney [5], RunwayML [6], Firefly [7], Ideaogram [8], and Leonardo.ai [9], and image generators integrated into chat interfaces like DALL-E [10], Imagen [11], and Grok [12]. Most offer similar interfaces where users specify prompts, adjust the diffusion process parameters and generate. The interface itself may take the shape of a GUI, a node-based environment or an infinite canvas. Most also support artistic workflows and community-created extensions, but they do not offer the option to train from scratch, as it is costly to do so. Instead, artists choose between multiple pre-set models (such as from Stable Diffusion), models fine-tuned by others found on public model repositories (like HuggingFace [13], CivitAI [14], or Tensor.art [15]), or fine-tune/personalize base models on a few images using techniques like Textual Inversion [7], DreamBooth [23], and Low-Rank Adaptations (LoRAs) [10]. The majority of these text-to-image systems do not run in real-time unless specific workflows, e.g. StreamDiffusion [15], are deployed.

2.2 GAN-based Image Generation Tools

The options for working with GAN models are limited compared to diffusion models. Systems like Playform [22], and Autolume [18] are GAN-based systems aimed at artists with no coding experience. These systems offer options for training or fine-tuning GAN models in addition to some process-support features that are often found in digital arts workflows, such as image pre-processing in preparation for training (e.g., cropping, resizing, normalization), visualizing the training progress with charts and snapshots of the results, or post-processing the results (e.g. up-scaling). Other low-code options for working with GANs include Google Collab notebooks that facilitate training and generating images/videos with minor modifications to the notebooks [16]. Furthermore, systems such as as TorchGAN [20], StudioGAN [11] and the GANToolkit [24] enable users to train their

[1] To learn more about these artists and others, please visit https://aiartists.org/

[2] https://github.com/AUTOMATIC1111/stable-diffusion-webui
[3] https://comfy.org/
[4] https://github.com/invoke-ai/InvokeAI
[5] https://www.midjourney.com/
[6] https://runwayml.com/
[7] https://www.adobe.com/products/firefly/features/text-to-image.html
[8] https://ideogram.ai/
[9] https://leonardo.ai/
[10] https://openai.com/index/dall-e-3/
[11] https://deepmind.google/technologies/imagen-3/
[12] https://grok.com/
[13] https://huggingface.co/models
[14] https://civitai.green/
[15] https://tensor.art/
[16] A list of GAN-based Collab notebooks can be found here https://pharmapsychotic.com/tools.html

own GAN models from scratch by modifying high-level configuration files (e.g. JSON files). However, they require working with the command line or require coding skills to operate, and they do not generate in real-time.

Out of the GAN systems above, only Autolume [18] is currently free, open-source, operates locally, and offers real-time generation, among other features. Furthermore, Autolume is distinguished from text-to-image systems in that it focuses on GAN models, which offer real-time inference and smooth latent space interpolation, and Autolume's interface also offers options for model crafting [1], such as model bending (applying transformation to internal model layers) and mixing (picking layers from multiple models), which help push the aesthetics of the model towards experimental or abstract visuals and beyond the distribution it was trained on [3].

2.3 Moving Images

AI-based video generation encompasses a wide variety of tasks including creating videos from textual prompts (e.g., Deforum [17], Sora [18], Veo 2 [19], RunwayML Gen-4 [20]), video editing (e.g., VEED [21]), and animating stills (e.g., AnimateDiff [22] and KlingAI [23]). The survey by Sun et al. [27] summarizes the recent advances in the field of video generation. The breakdown in the survey reflects the current goals and aspirations of the field, including (1) *Excellent Pursuit*, which encompasses generating videos with extended duration, superior resolution, and seamless quality, and (2) *Realistic Panorama* which encompasses generating videos with dynamic motions, complex scenes, multiple objects, and rational layouts.

The unique affordances of Autolume focus on generating videos, such as audio-reactive visuals and video art, by smoothly moving through points in personal, artistic, or experimental latent spaces. As it pertains to *Excellent Pursuit*, the length of videos in Autolume depends on the number of points the artist chooses, without any inherent limitation. The resolution and quality depend on the underlying model being explored, rather than the tool itself, although upscaling models are available within Autolume and can be run in real-time. In relation to creating *Realistic Panoramas*, Autolume's emphasis is less on composing scenes with multiple interacting objects or characters, and more on generating moving visuals for aesthetic rather than functional purposes. Several artists we worked with mentioned using Autolume to create videos and experiment with new expressive modes that reinterpret their body of work, whether photographs or paintings. Finally, compared to current video generation tools, Autolume is distinctive in that it runs in real-time, enabling faster iteration.

3 Autolume

We present Autolume to make the artistic potential of personally-trainable generative AI models, particularly GANs, accessible to artists without technical expertise, while also expanding the possibilities for using these models in an interactive, real-time setting. Initially developed as a live VJing tool [16], Autolume is a no-coding, user-friendly visual synthesizer leveraging the artistic potential of GANs. It simplifies the entire GAN workflow, from dataset preprocessing and model training to real-time latent space exploration, feature extraction, and output upscaling, making it accessible to non-technical creatives. Moreover, Autolume introduces interactive generation via Open Sound Control (OSC) messaging, enabling audio-reactive and other forms of interactive art, thus broadening its creative possibilities. Finally, it offers options for adapting and crafting generative AI models [1]. Autolume offers both offline (pre-processing and post-processing) and real-time (latent space navigation, OSC messaging, semantic manipulation) modules. Autolume can be accessed on https://www.metacreation.net/autolume, and examples of artworks produced with Autolume are discussed in section 4.

3.1 Offline System Features

Autolume offers five offline modules, shown in Figure 2, for different processing features.

3.1.1 Model Training. The training module enables users to train StyleGAN2 [13] models from scratch or fine-tune pre-tained models, using either image datasets or frames extracted from uploaded videos. It provides data pre-processing, including resizing and framing options, and supports both square and non-square datasets. To enhance training with smaller datasets, it offers Adaptive Discriminator Augmentation (ADA) [12] and Differentiable Augmentation (DiffAug) [33] methods, each with various augmentation pipelines. These features facilitate training with smaller datasets and allow artists to use datasets as small as 100 images [24]. Users can adjust training hyperparameters like batch size, learning rates, and gamma to achieve desired results, and the training runs on the user's local GPU. While achieving optimal results may require multiple iterations, a robust GPU, such as NVIDIA GeForce RTX 3090, can yield meaningful outcomes in about 12 hours. Real-time progress monitoring is available via a grid display of randomly generated images to help visually evaluate the results during training.

3.1.2 Latent Space Projection. The projection module in Autolume allows users to enter a reference image, a text prompt, or both to find the closest latent vector in a trained model that resembles them, as well as generating a video of the search process. The resulting latent vector can be saved and loaded into the real-time module for further exploration.

3.1.3 Feature Extraction. The semantic feature extraction module uses GANSpace [8] to identify and extract interpretable feature directions in the latent space which provides an alternate approach for navigating the latent space. This feature includes options for feature sparsity and the number of directions to be calculated, and the calculated feature directions can be loaded and explored in the real-time module.

3.1.4 Super Resolution. Due to computational limitations, artists may opt to train their models on smaller resolutions (e.g., 256x256 or 512x512) and then upscale the results to achieve results with higher

[17] https://deforum.art/
[18] https://sora.com/
[19] https://deepmind.google/technologies/veo/veo-2/
[20] https://runwayml.com/research/introducing-runway-gen-4
[21] https://www.veed.io/
[22] https://animatediff.github.io/
[23] https://www.klingai.com/

[24] The quality of training depends on various factors such as the visual coherency of the dataset and its size.

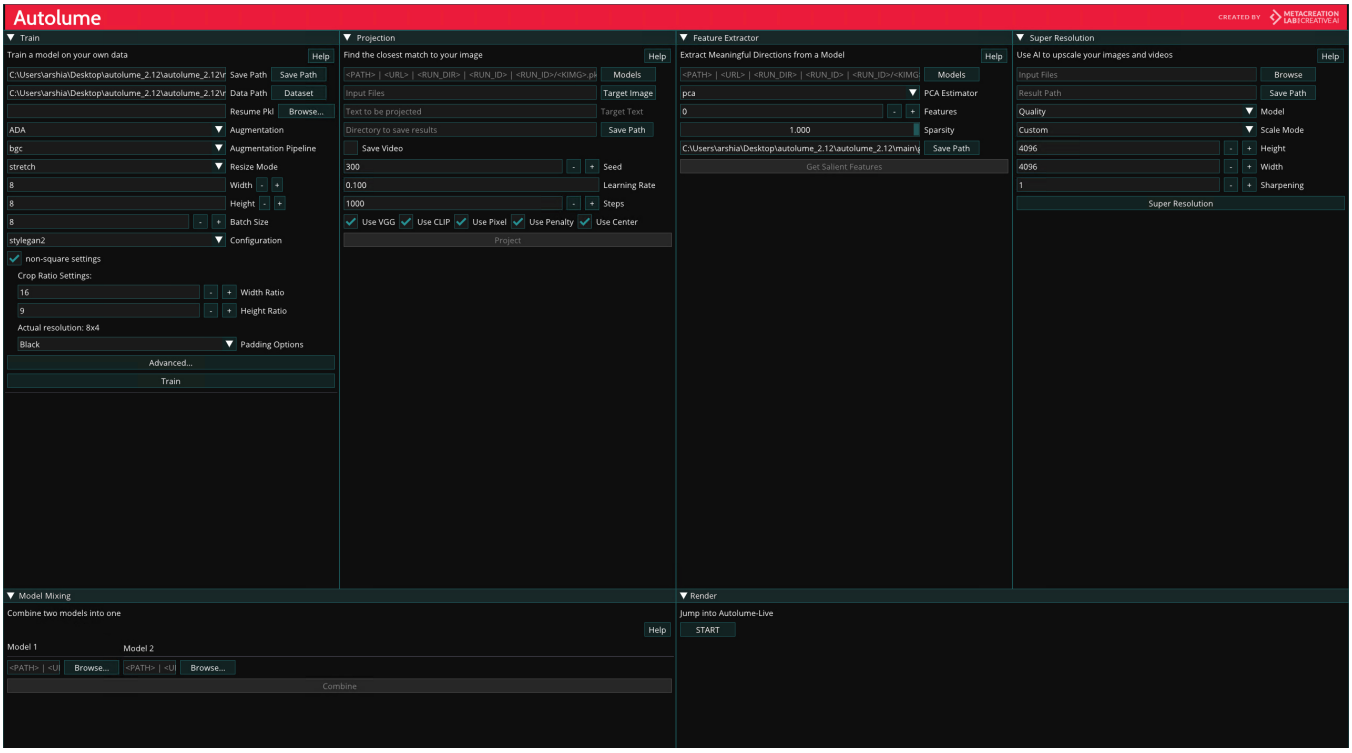

Figure 2: Screenshot of the Autolume offline modules interface, showing 4 modules on the top sections, and two modules on the bottom.

fidelity. The super-resolution module enables high-quality upscaling of images and videos using the Real-ESRGAN algorithm [32]. Users can select from three modes, which are Quality, Balance, and Speed, tailored to their desired fidelity and processing time. The module also allows customization of the upscaling factor or specific output dimensions, and also supports batch processing. The Speed mode in Autolume-live enables real-time upscaling during generation.

3.1.5 Model Mixing. The model mixing module enables users to blend two trained models by combining their layers [21], creating and saving a new model for future use. This allows for the fusion of aesthetic features from different models, generating unique visual styles and new aesthetic spaces. This feature is also available in the real-time module for a live preview of the results. The resulting hybrid models can be saved locally for future use.

3.2 Real-time System Features: Autolume-live

The Autolume-live module allows artists to load trained models and interact with them in real-time, offering a suite of features for exploring latent space and experimenting with network parameters to achieve unique aesthetics.

3.2.1 Latent Space Navigation and Controls. Artists can control the visual features by adjusting parameters such as global noise and truncation. Various latent space exploration options are available, including setting specific seed numbers as starting points of the exploration, performing random walks with adjustable speeds, and using different looping techniques. For instance, artists can identify interesting points in the latent space and loop between them as keyframes, allowing them to create curated experiences of the latent space. Keyframes can also be saved and loaded into other parts of the interface. Moreover, pre-calculated directions from Feature Extraction, derived in the offline module, can be loaded and explored in real-time.

3.2.2 Model Bending and Blending. The Layer Transformations panel enables custom noise control and transformations (e.g. translation, scaling, and inversion) on individual model layers, allowing for highly detailed control over visual outputs and expanding creative possibilities [4]. The model mixing feature, available in the offline module, can also be used in real-time for a live preview of the results.

3.2.3 Interactivity. Autolume-live supports Open Sound Control (OSC) which is a protocol for networking different applications, such as sound and visual synthesizers or multimedia devices, enabling real-time data exchange and control. OSC enables artists to control the real-time module parameters for audio-reactive and other interactive applications. The OSC feature also offers fine-tuning capabilities, such as applying mathematical expressions to OSC signals, for precise mapping within Autolume. Lastly, the generated visuals can be sent through a Network Device Interface (NDI) video streaming protocol to other software for additional post-processing.

3.2.4 Presets. To save the progress made during navigation and to enable revisiting preferred settings, artists can save all their current settings, including OSC mappings or slider values, as presents that are linked to their specific models.

4 Artworks

Autolume has been implemented in diverse artistic applications, demonstrating its adaptability and creative potential across various contexts. This section highlights distinct approaches to using Autolume in artworks, particularly emphasizing its capacity to generate and manipulate moving images through real-time and small-data workflows. These applications include training a model from scratch using an artist's personal dataset, creating dynamic audio-visual pieces, supporting live performances as an autonomous agent, and enabling interactive installations with real-time generative features. Together, they showcase how Autolume empowers artists to explore new creative possibilities with generative AI by blending aesthetic exploration, movement, and interaction within a small-data, no-coding approach.

Aesthetically, Autolume embraces a distinct visual language shaped by the fluid geometry of GAN latent space, often resulting in imagery that evokes dreamlike abstraction, seamless transitions, and a sense of temporal ambiguity. The system's capacity for serendipitous interpolation and fine-grained manipulation fosters the emergence of organic, ambiguous, or uncanny forms that challenge conventional visual narratives. At the same time, Autolume supports artists pursuing deliberate and meticulously crafted aesthetics, offering precise control over generative processes. As such, it accommodates both open-ended exploration and clearly defined creative visions, amplifying a wide range of artistic intentions.

4.1 Dreamscape

In *Dreamscape*, Autolume was used to train a model from scratch using a small dataset of 302 artworks provided by Vancouver-based visual artist Erica Lapadat-Janzen. Some samples from the dataset are shown in Figure 3. Collaborating closely with the artist, we utilized Autolume to explore the latent space and fine-tune parameters, crafting visuals that resonated with the artist's unique aesthetic. The result was a collection of 12 stills and 9 video loops—conceptualized as slowly moving paintings—intended for public presentation. See Figure 4 for samples of generated stills. This process highlights Autolume's capacity to enable artists to train personalized models and achieve refined outputs that align with their creative vision, even with limited data. More samples from the collection, including video loops, can be accessed at https://www.metacreation.net/projects/dreamscape.

4.2 Autolume Mzton

Autolume Mzton, a collaboration between Jonas Kraasch and Philippe Pasquier, explores the notion of dawn using Autolume audio-reactive features (Figure 5). Driven by the piece Mzton, from the analog modular rhizome of the French band Robonom[25], the neural aesthetic of generative visuals unexpectedly evokes early experimental analog cinema. Themes of horizons, sunsets, and the vibrancy of the real-time audio-visual coupling between the soundtrack and

[25]https://soundcloud.com/monobor/sets

Figure 3: Samples from the *Dreamscape* training dataset. Images courtesy of the artist.

Figure 4: Three stills from *Dreamscape* at Ars Electronica 2024 in Linz, Austria.

Autolume visual content evoke utopian new beginnings. This piece can be viewed at https://vimeo.com/527564204.

Figure 5: *Autolume Mzton* at Distopya Sound Art Festival 2021 in Istanbul, Turkey.

4.3 Reprising Elements

Reprising Elements is an audio-visual performance, by Arshia Sobhan and Joshua Rodenberg (see Figure 6). This performance unfolds

as a co-creative exploration among three distinct agents: a calligrapher, Autolume operating as a real-time generative visual system, and a sound artist, all interacting within intricate audio-visual feedback loops.

At the heart of this performance is the traditional calligraphic practice of *siyah mashq* in Persian calligraphy, a meditative exercise where calligraphers repetitively inscribe letters and combinations, transforming paper into a richly layered black canvas. In *Reprising Elements*, this ritual is reimagined by using Autolume acting as a generative collaborator and a visual guide. Trained on a dataset of calligraphy samples [25], the model running in Autolume proposes novel calligraphy combinations and challenges the calligrapher to follow and replicate the generated samples. Beyond visual guidance, Autolume captures and integrates the distinct auditory nuances of the calligraphic process—the screeching of the reed pen on paper. These sounds are recorded and transformed by the sound artist into an integral part of the performance, adding auditory depth to the visual spectacle. Using Autolume's interactive features, the generative calligraphy responds to these auditory inputs in real time, completing a feedback loop where each element—the calligrapher, the AI, and the sound artist—influences and reshapes the others' contributions. This interaction not only challenges but also expands the boundaries of traditional calligraphic art, integrating contemporary technology while staying true to the essence of the practice. Excerpts from this performance can be viewed at https://youtu.be/q5oesRsqs7k.

Figure 6: *Reprising Elements* performance at VCUQ arts, 2023 in Doha, Qatar.

4.4 Alpha Prism

This installation rests on a model trained in Autolume on the 20-year archive of analog film portraits by photographer Kris Krüg. This collaborative project, conceptualized by Philippe Pasquier, with a model trained by Arshia Sobhan, and interactivity programmed by Lionel Ringenbach, consists of an ever-growing video loop between portraits of the audience. Photography of the audience is taken on-site and projected into Krüg's portrait model. These projections are added to the video loop slowly morphing from one visage to another. *Alpha Prism* is a reflection on deepfakes and identity through the aesthetic lens of Krüg and a unified latent space of human visages. Figure 7 shows the overall workflow of the Alpha Prism installation, illustrating how audience portraits are processed, projected into the latent space, and integrated into the evolving video loop using Autolume.

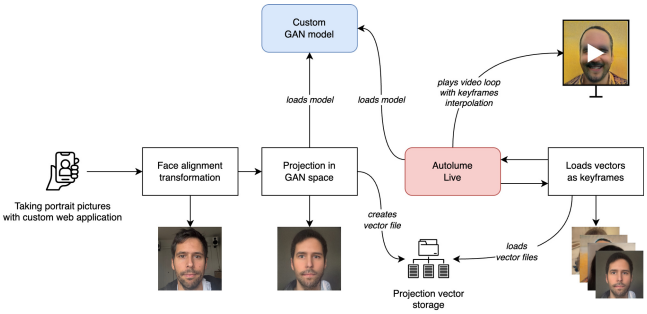

Figure 7: Workflow of the *Alpha Prism* installation.

Alpha Prism explores the porous boundaries between human memory, machine vision, and synthetic identity. Autolume's real-time small-data approach preserves the material textures and emotive grain of analog photography, allowing the installation to generate organic, fluid transitions between visages rather than sterile or hyperreal outputs. Conceptually, *Alpha Prism* reframes deepfakes not as deceptive artifacts but as portals to new forms of creative subjectivity, where identities are mutable, relational, and continuously co-constructed between humans and intelligent systems. Through this entangled process, *Alpha Prism* invites reflection on authorship, authenticity, and the evolving nature of human presence in a hybrid computational world.

Figure 8: *Alpha Prism.*

4.5 Revival

Revival is a live audiovisual performance by the artist collective K-Phi-A trio, consisting of electronic music performer Philippe Pasquier, percussionist Keon Ju Maverick Lee, and VJ Amagi (Jun Yuri). The performance is powered by award-winning musical AI agents MASOM [28], SpireMuse [30], MACAT [17] and Autolume as the visual synthesizer. Together, they produce electronic and electroacoustic music tightly integrated with audio-reactive visuals. The performance is centred on real-time co-creative improvisation between human performers and musical AI agents developed at the Metacreation Lab for Creative AI. These musical AI agents are

trained on the works of deceased composers and compositions created by our artist collective, hence the project title. Using Autolume and Touch Designer, VJ Amagi generates abstract audio-reactive visual artworks. An extended version of the piece is available at https://vimeo.com/1041308161.

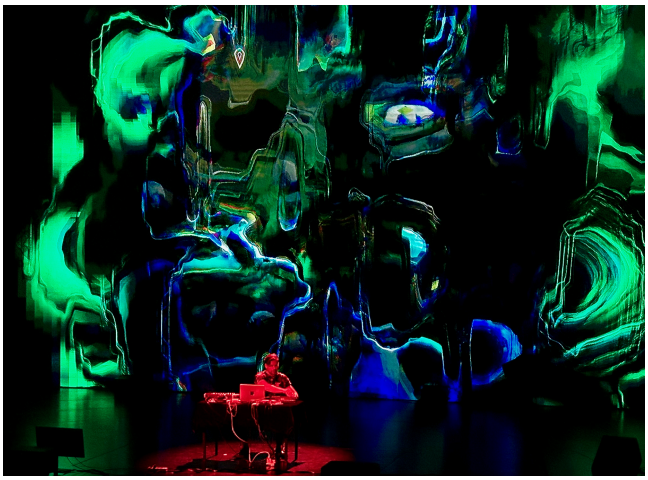

Figure 9: *Revival* at Musicacoustica 2024, Hangzhou, China.

5 Autolume Workshops

Figure 10: Autolume workshops at MANIFEST:IO Symposium in Berlin, February 2025 (top), in Montreal (bottom-right) and in Toronto (bottom-left), October 2024.

Three in-person workshops and an online workshop on Autolume have been conducted so far, targeting artists and creative professionals. These workshops aimed to facilitate participants' access to Autolume and demonstrate its features, while simultaneously gathering insights into how they incorporate the tool into their creative practice. Held in collaboration with OCAD University in Toronto, NAD (École des arts numériques, de l'animation et du design) and SAT (Society for Arts and Technology) in Montreal, and MANIFEST:IO Symposium in Berlin, our in-person workshops attracted a total of 40 participants. Additionally, our online workshop hosted more than 50 participants from multiple countries. Each workshop spanned two days. On the first day, participants were introduced to Autolume's features through a system demonstration. By the end of the day, participants who brought their own created or ethically curated datasets initiated overnight model training. On the second day, the trained models were reviewed and discussed, providing participants with an opportunity to explore the results and share their experiences. Despite time constraints, multiple participants managed to train their own custom models with small datasets.

The workshops also explored the challenges and benefits participants experienced when using Autolume through post-workshop surveys and open-ended discussions. Initial feedback from artists indicates that Autolume fosters a strong sense of authorship over their creations and is generally perceived as intuitive to use. It enables users to uncover new aspects of their datasets and can quickly generate diverse outputs. Participants also felt confident that, with more effort in navigating the latent space, they could arrive at desired or interesting aesthetics. Additionally, the tool has inspired many to consider integrating it into their future creations. However, participants also identified challenges, such as the need for tooltips to improve learnability, and progressive feature reveals to make the tool more accessible to novices. This research is ongoing, and we plan to publish the findings in a future paper, further contributing to the development and adoption of Autolume in artistic practices.

6 Conclusion & Future Work

We presented Autolume, a tool for art creation that adopts model crafting and small-data mindsets. The goal of this project is to provide artists with tools that maximize creative ownership and address ethical concerns associated with large-scale generative models. We presented multiple artworks that showcase the versatility of the tool, whether in creating real-time performances, interactive installations, or personalized visual synthesis. Ongoing research includes workshops conducted with artists and creative professionals, where we explore their experiences and gather insights on incorporating Autolume into their practice. These workshops have provided valuable feedback for refining the tool and understanding its role in diverse artistic workflows.

Looking ahead, we are pursuing several directions for further development. These include integrating small-data methods for diffusion models to broaden Autolume's capabilities and expanding the latent space navigation techniques to offer more intuitive and precise control over generative processes. Additionally, we are actively seeking artistic collaborations across varied creative approaches to deepen our understanding of the dynamic between artists and the system. By fostering such collaborations, we aim to ensure that Autolume continues to evolve as a versatile and inclusive tool for generative art.

References

[1] Ahmed M. Abuzuraiq and Philippe Pasquier. 2024. Seizing the Means of Production: Exploring the Landscape of Crafting, Adapting and Navigating Generative AI Models. In *the 3rd Generative AI and HCI Workshop* (Hawaii, USA).

[2] Terence Broad. 2024. Using Generative AI as an Artistic Material: A Hacker's Guide. In *In Proceedings of Explainable AI for the Arts Workshop 2024 (XAIxArts 2024)*.

[3] Terence Broad, Sebastian Berns, Simon Colton, and Mick Grierson. 2021. Active Divergence with Generative Deep Learning–A Survey and Taxonomy. In *Proceedings of the 12th International Conference on Computational Creativity (ICCC '21)*. doi:arXivpreprintarXiv:2107.05599

[4] Terence Broad, Frederic Fol Leymarie, and Mick Grierson. [n. d.]. 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, 20–36.

[5] Baptiste Caramiaux and Sarah Fdili Alaoui. 2022. "Explorers of Unknown Planets" Practices and Politics of Artificial Intelligence in Visual Arts. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–24.

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

[7] Rinon Gal, Yuval Alaluf, Yuval Atzmon, Or Patashnik, Amit H Bermano, Gal Chechik, and Daniel Cohen-Or. [n. d.]. An Image Is Worth One Word: Personalizing Text-to-Image Generation Using Textual Inversion. ([n. d.]). arXiv:2208.01618

[8] Erik Härkönen, Aaron Hertzmann, Jaakko Lehtinen, and Sylvain Paris. 2020. Ganspace: Discovering interpretable gan controls. *Advances in neural information processing systems* 33 (2020), 9841–9850.

[9] Aaron Hertzmann. 2020. Visual indeterminacy in GAN art. In *ACM SIGGRAPH 2020 Art Gallery*. 424–428.

[10] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. [n. d.]. Lora: Low-rank Adaptation of Large Language Models. ([n. d.]). arXiv:2106.09685

[11] Minguk Kang, Joonghyuk Shin, and Jaesik Park. [n. d.]. StudioGAN: A Taxonomy and Benchmark of GANs for Image Synthesis. 45 ([n. d.]), 15725–15742. https://api.semanticscholar.org/CorpusID:249889489

[12] Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila. 2020. Training generative adversarial networks with limited data. *Advances in neural information processing systems* 33 (2020), 12104–12114.

[13] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2020. Analyzing and improving the image quality of stylegan. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 8110–8119.

[14] Hyung-Kwon Ko, Gwanmo Park, Hyeon Jeon, Jaemin Jo, Juho Kim, and Jinwook Seo. [n. d.]. Large-Scale Text-to-Image Generation Models for Visual Artists' Creative Works. In *Proceedings of the 28th International Conference on Intelligent User Interfaces* (New York, NY, USA, 2023) *(IUI '23)*. Association for Computing Machinery, 919–933. doi:10.1145/3581641.3584078

[15] Akio Kodaira, Chenfeng Xu, Toshiki Hazama, Takanori Yoshimoto, Kohei Ohno, Shogo Mitsuhori, Soichi Sugano, Hanying Cho, Zhijian Liu, and Kurt Keutzer. 2023. Streamdiffusion: A pipeline-level solution for real-time interactive generation. *arXiv preprint arXiv:2312.12491* (2023).

[16] Jonas Kraasch and Philippe Pasquier. 2022. Autolume-Live: Turning GANs into a Live VJing Tool. In *Proceedings of the 10th Conference on Computation, Communication, Aesthetics & X* (Coimbra, Portugal). 152–169.

[17] Keon Ju M. Lee and Philippe Pasquier. 2024. Musical Agent Systems: MACAT and MACataRT. In *Proceedings of the Creativity and Generative AI NIPS (Neural Information Processing Systems) Workshop*. https://creativity-ai.github.io/assets/papers/64.pdf.

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

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

[20] Avik Pal and Aniket Das. [n. d.]. TorchGAN: A Flexible Framework for GAN Training and Evaluation. 6 ([n. d.]), 2606. https://api.semanticscholar.org/CorpusID:202540048

[21] Justin N. M. Pinkney and Doron Adler. 2020. 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).). doi:arXivpreprintarXiv:2010.05334

[22] Playform. 2024. Playform: No-Code AI for Creative People. https://www.playform.io/train.

[23] Nataniel Ruiz, Yuanzhen Li, Varun Jampani, Yael Pritch, Michael Rubinstein, and Kfir Aberman. [n. d.]. 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). 22500–22510.

[24] Anush Sankaran, Raunak Sinha, and Naveen Panwar. [n. d.]. *IBM GAN Toolkit*. https://https://github.com/IBM/gan-toolki

[25] Arshia Sobhan, Philippe Pasquier, and Adam Tindale. 2024. Unveiling New Artistic Dimensions in Calligraphic Arabic Script with Generative Adversarial Networks. *Proc. ACM Comput. Graph. Interact. Tech.* 7, 4, Article 61 (July 2024), 12 pages. doi:10.1145/3664210

[26] StabilityAI. 2024. Stable Diffusion. https://stability.ai/stable-image.

[27] Rui Sun, Yumin Zhang, Tejal Shah, Jiahao Sun, Shuoying Zhang, Wenqi Li, Haoran Duan, Bo Wei, and Rajiv Ranjan. 2024. From sora what we can see: A survey of text-to-video generation. *arXiv preprint arXiv:2405.10674* (2024).

[28] Kıvanç Tatar and Philippe Pasquier. 2017. MASOM: A musical agent architecture based on self-organizing maps, affective computing, and variable Markov models. In *Proceedings of the 5th International Workshop on Musical Metacreation (MUME 2017). Atlanta, Georgia, USA.*

[29] Austin Tecks, Thomas Peschlow, and Gabriel Vigliensoni. [n. d.]. Explainability Paths for Sustained Artistic Practice with AI. In *In Proceedings of Explainable AI for the Arts Workshop 2024 (XAIxArts 2024)* (2024-07-21). arXiv. doi:10.48550/arXiv.2407.15216 arXiv:2407.15216 [cs]

[30] Notto J. W. Thelle and Philippe Pasquier. 2021. Spire Muse: A Virtual Musical Partner for Creative Brainstorming. In *NIME 2021*. https://nime.pubpub.org/pub/wcj8sjee.

[31] Gabriel Vigliensoni, Phoenix Perry, and Rebecca Fiebrink. 2022. A Small-Data Mindset for Generative AI Creative Work. In *In Generative AI in HCI Workshop, CHI '22* (New York, NY, USA). 5.

[32] Xintao Wang, Liangbin Xie, Chao Dong, and Ying Shan. 2021. Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In *Proceedings of the IEEE/CVF international conference on computer vision*. 1905–1914.

[33] Shengyu Zhao, Zhijian Liu, Ji Lin, Jun-Yan Zhu, and Song Han. 2020. Differentiable Augmentation for Data-Efficient GAN Training. In *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (Eds.), Vol. 33. Curran Associates, Inc., 7559–7570.