Beyond Prompts: Installation Project Dossier

by Garin Curtis

Contents

1 Introduction

This work examines the creative potential of generative AI by exploring its capabilities beyond conventional applications. While most image generative models are trained to replicate patterns from vast datasets using text inputs, this installation invites participants to engage with the system in a more experimental and interactive way. Through a touch interface, visitors can directly explore the layers of a generative neural network, revealing unexpected behaviours and outcomes.

By making AI's underlying processes accessible, the installation aims to demystify how these systems operate and generate their outputs. It offers a hands-on way to think critically about the mechanics of AI, moving beyond its role as a tool for replication to highlight its potential as a medium for exploration and discovery.

Using network-bending techniques, which modify and manipulate the internal representation of a stable diffusion model during inference, this work introduces novel interaction with the internal architecture of a deep generative model. The experience prompts visitors to reflect on the assumptions and possibilities of AI, fostering a deeper understanding of its potential role within the creative process.

2 Background

2.1 Current Limitations in Generative AI Tools

Diffusion-based generative tools for image creation primarily focus on text-to-image functionality, offering limited creative control beyond the initial input prompt. As [4] highlights, the reliance on textual prompts inherently restricts user control over outputs. This often results in generated images that closely resemble training data, minimising novelty and creative divergence [1].

For digital artists, particularly those working in live performance or iterative creative processes, the lack of real-time interaction capabilities presents significant limitations. Current tools emphasise accuracy in replication but lack flexibility for improvisation or experimentation. While techniques such as those by [3] introduce active divergence in diffusion-based models, they often require expert knowledge and intricate sys-

tem architecture manipulation, making them inaccessible to most artists.

2.2 Opportunities with Network Bending

Accessible, user-friendly interfaces that allow real-time manipulation of neural networks are scarce for diffusion models like Stable Diffusion. Although tools for Generative Adversarial Networks (GANs), such as StyleGAN, have been integrated into user-friendly systems like AutoLume [6], no equivalent exists for diffusion models. Furthermore, network bending, a technique that manipulates a model's internal representations, remains largely unexplored in diffusion models.

My thesis leverages network-bending techniques [2] within the StreamDiffusion pipeline [5] to address these gaps. By developing a prototype plugin in TouchDesigner, this project provides an intuitive, accessible interface for real-time interaction and creative exploration see Figure 1.

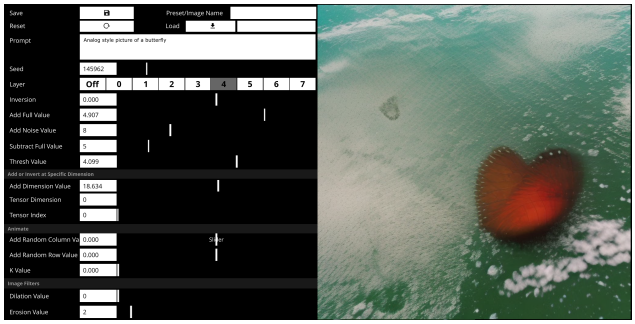

Figure 1: TouchDesigner Plugin allowing full control over network bending parameters.

See Plugin: `https://vimeo.com/1045850504?share=copy`

3 Motivations

This installation was inspired by the limitations of traditional generative AI tools and my master's thesis research. The primary goal is to empower artists and participants with tools for open-ended exploration of the latent space within generative AI, enabling greater artistic and expressive control over outputs.

By bridging the gap between technical complexity and creative accessibility, this work seeks to:

- Foster critical engagement with the mechanics of generative AI.
- Encourage experimentation and discovery through interactive exploration.
- Highlight AI's potential as a collaborative partner in the creative process.

4 Installation: Interaction and Experience

The installation was developed as a continuation of my master's thesis which aimed to establish a foundation for advancing real-time, artist-driven interaction with generative AI models. Using a tablet-based touch interface, users can interactively manipulate the model's internal architecture and observe the dynamic evolution of generated outputs.

Key Features:

- **Network Bending**
  - This technique applies mathematical transformations to the image tensor within the computational graph during inference.
  - Transformations create expressive and unpredictable modifications to outputs. For example, starting with the prompt "a portrait of a person," outputs dynamically evolve into abstract forms through real-time manipulation.
- **Interactive Interface**
  - Users control the system by moving their finger across a 2D space visualized on a tablet.
  - A machine-learning algorithm maps the XY coordinates to ten simultaneous transformations applied to the image tensor, ensuring smooth and intuitive interaction.
  - Participants can also select specific neural network layers to apply the transformations. Each layer corresponds to a distinct stage in the denoising process during image generation, with each stage uniquely shaping the final output and its visual characteristics.
- **Configuration**
  - The system is set up for image-to-image generation using animated 3D Perlin noise as the input. This approach ensures continuous visual evolution and interaction.
- **User Feedback**
  - The installation was exhibited at UAL's Creative Computing Institute Winter Festival and won the industry award for storytelling.

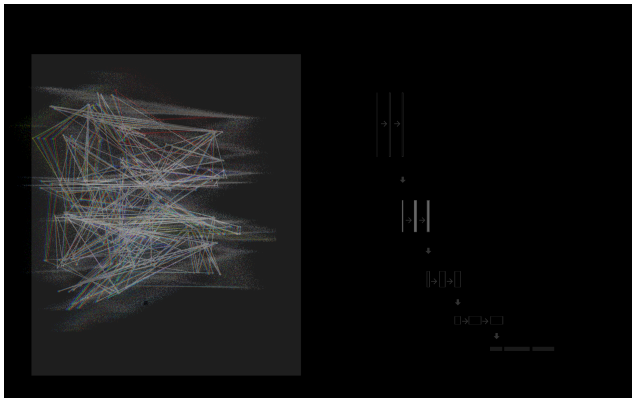

Figure 2: Touch screen interface with XY grid on the left and steps to select neural network layer on the right.

See https://vimeo.com/1045833997?share=copy

5 Technical Description and Requirements

5.1 Minimum requirement

The minimum version of this project would just include the interactive system of the installation as shown on the right-hand side of Figure 3. The installation has been well tested and successfully ran for 8 hours per day over 3 days at the UAL Creative Computing Institute Winter Festival. The installation requires the following components:

5.1.1 Technical Setup:

Samsung Galaxy A9+ Tablet Used for the touch interface (provided by the artist and pre-tested).

Razer Blade 16 RTX 4080 Laptop Runs the generative AI and TouchDesigner plugin (provided by the artist and pre-tested).

4K Monitor 27" Screen To display the work (Can be provided by the artist if necessary).

Power Requires two power outlets one for the laptop and one for the 4K monitor. The tablet runs power off of the laptop.

5.1.2 Installation Details:

Space Requirements 40 x 40 cm plinth space for the Tablet and Monitor).

Setup Time 1- 2 hours by one person (method pre-tested and optimised).

Figure 3: Beyond Prompts Installation. Audiovisual on the left and interactive system on the right.

See Audiovisual: https://vimeo.com/1044132057?share=copy

6 Artworks

A collection of Artworks made using just the tool and upscaled to 1024 x 1024.

7 Conclusion

This installation bridges the gap between generative AI's technical complexities and its creative possibilities. By introducing network-bending techniques and real-time interaction, it transforms AI from a tool of replication into a medium for exploration and artistic expression.

Through this hands-on, intuitive interface, participants are invited to engage critically and creatively with the mechanics of AI, discovering its potential as a partner in the creative process.

Reference List

[1] Sebastian Berns and Simon Colton. “Bridging Generative Deep Learning and Computational Creativity.” In: *ICCC*. 2020, pp. 406–409.

[2] Terence Broad, Frederic Fol Leymarie, and Mick Grierson. “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*. Springer. 2021, pp. 20–36.

[3] Luke Dzwonczyk, Carmine Emanuele Cella, and David Ban. “Network Bending of Diffusion Models for Audio-Visual Generation”. In: *arXiv preprint arXiv:2406.19589* (2024).

[4] Hyung-Kwon Ko et al. “Large-scale text-to-image generation models for visual artists’ creative works”. In: *Proceedings of the 28th international conference on intelligent user interfaces*. 2023, pp. 919–933.

[5] Akio Kodaira et al. “Streamdiffusion: A pipeline-level solution for real-time interactive generation”. In: *arXiv preprint arXiv:2312.12491* (2023).

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

Contact

Mobile Number: +44 7748655773

Contact: garincurtis@gmail.com