"Event Horizon, Generation 28"

Exploring Ambiguity in Human and AI Art

Alex Amari[1]
Yale University, New Haven, CT 06520
alex.amari@yale.edu

1 Role of AI and ML

Beginning with the image "Galaxy M82" from the Event Horizon Telescope Collaboration, "Event Horizon, Generation 28" visualizes a feedback loop involving several collaborative machine learning models working with intermittent human feedback. The pipeline of models creatively transforms and interprets a starting image, ultimately producing an output that becomes the input of the subsequent "generation." Each generation consists of the following steps:

1. **Encoding by a Convolutional Autoencoder**: A starting image is processed by a convolutional autoencoder (**Figure 1**). The autoencoder was trained on 10,000 images generated by OpenAI's DALL-E 3 displaying people, places, and things. Training on the diffusion-generated dataset causes the autoencoder to introduce subtle visual artifacts and qualities characteristic of DALL-E 3 images during reconstruction. The 10-layer architecture includes an initial bilinear downsampling phase followed by 5 symmetrical encoding and decoding layers. Periodic dropout regularization during training introduces random changes to the latent space and helps enhance reconstruction quality.
2. **Prompt Generation with GPT-4 Vision**: The reconstructed image is passed to GPT-4 Vision via API. The purpose of this step is to generate a necessary prompt for the subsequent creative upscaling phase while also introducing an additional layer of creative interpretation by a separate AI model. GPT-4 Vision is prompted with: "Provide a vivid and creative description of this image in less than 150 tokens."
3. **Creative Upscaling with Stable Diffusion 3 and Human Feedback** The reconstructed image and generated caption are passed to Stable Diffusion 3's creative upscaler via API. The upscaled image creatively blends the target image with qualities from the caption (**Figure 2**). This process takes as an optional argument a negative prompt ("...what you do not wish to see in the output.") To increase the pace of creative drift across generations, short, negative prompts are added periodically by the human artist. In this work, the negative prompt "A solar eclipse" was added after the fifth generation, followed by "A black background" (10), "A sunflower" (15), "Flower petals" (20), and "An eyeball" (25) (**Figure 3**).

The pipeline was executed over 28 generations. The final visualization aims to capture the essence of the constituent machine learning processes while remaining interpretable for human viewers. Each generation is visualized with 14 frames scaled to dimensions corresponding with the layers of the encoder, decoder, and final upscaling. Ordered frames from all 28 generations were scaled to a 3152x3152 canvas before commencing at 6 frames per second.

2 Discussion of Ambiguity

This work aims to challenge the boundaries between "real" and "non-real," as well as "human" and "AI," in both visual quality and authorship. It begins with clear human origins: a scientific and artistic image applied to a pipeline conceived and operated by a human creator. But as model-driven interpretations accumulate, the work drifts artistically into realms that feel distinctly "AI." The subject matter extends quickly into the realm of the "non-real", while never fully abandoning compositional elements of its "real" origin. Such changes produce a dynamic blend of authorship that is never fully "human" or "AI." By generation 28, the lines between these notional categories converge on a new "event horizon"—a glimpse of the artistic and material ambiguity that so often characterizes human-AI art.

[1]Alex Amari is a researcher, data scientist, and artist pursuing an MBA at Yale University. His work focuses on human-AI collaboration in professional and creative industries. Alex also serves as a teaching fellow in Yale's Computer Science Department, where he enjoys introducing generative AI techniques to students from diverse backgrounds. He holds a BA from Rice University (2018) and an MSc from Oxford University (2019).

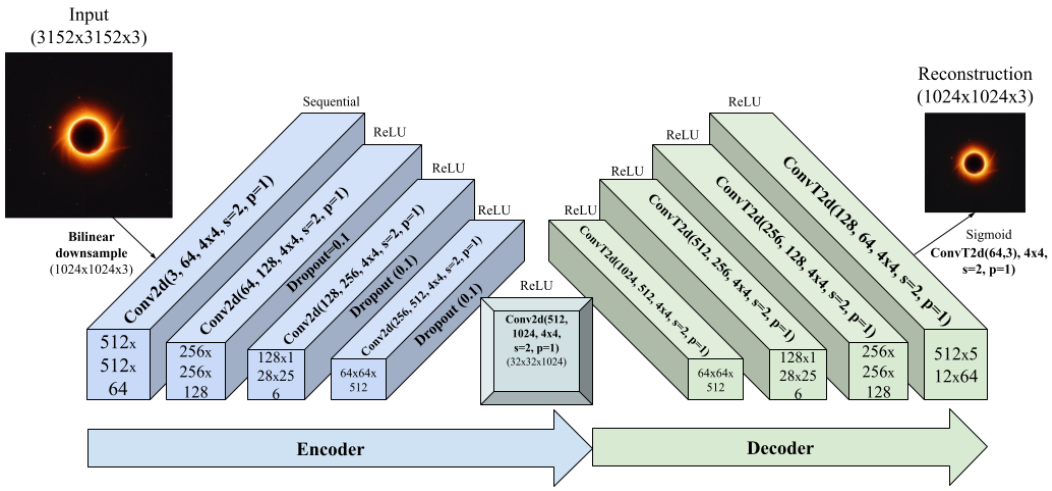

Figure 1: Convolutional Autoencoder Architecture, Samples from Generation 3

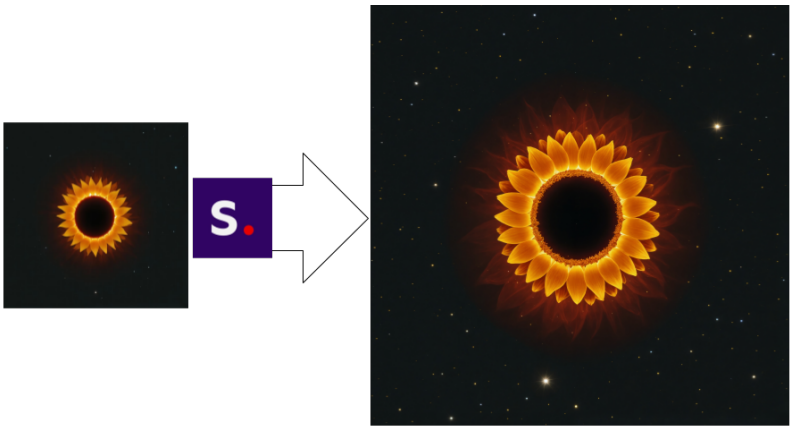

Figure 2: Creative Upscaling of Reconstructed Image, Generation 6

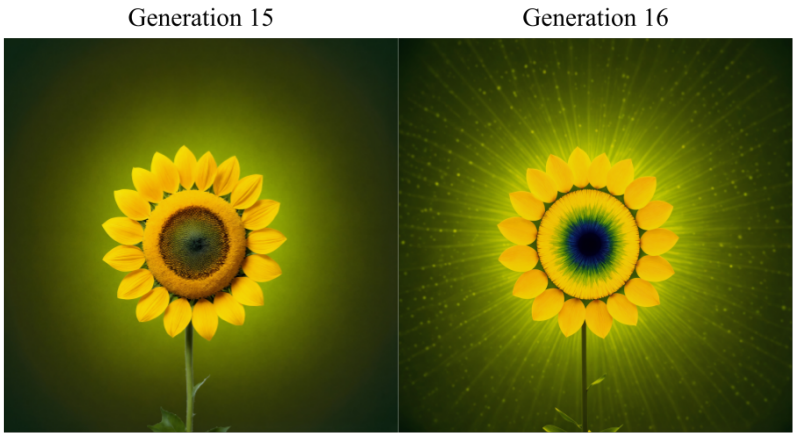

Figure 3: Introduction of Negative Prompt "A sunflower" at Generation 15