

***Voyce Sabrina Durling-
Jones ,
Aigars Ceplitis***
***A Use Case for
Diffusion Models
in the Generation
of Hybrid AI,
Multi-Modal Live
Performances***

Abstract

In September of 2022, one month after Open AI's Stable Diffusion was released to the public, the authors of this text presented a Hybrid AI Multi-Modal Live Performances (AI MMLP) at the 30th Annual CEEMAN Conference in Bled, Slovenia, where a sequence of animations based on keynote addresses and generated using Stable Diffusion were projected on two large screens. The animations were experienced in conjunction with a musical score and an interpretive ballet solo performance, all designed to enhance the hybrid inter-medial nature of the piece. While now common in the mainstream, using text-to-image and image-to-image machine learning models at the time were just beginning to gain momentum among some tech savvy visual artists. This article offers insight into the importance of experimentation by artists as new AI approaches become accessible in the public sphere and provides an example of how once experimental techniques are now deployed across disciplines to produce novel and impactful approaches to generating moving image visualizations through human-computer creative collaboration.

Reflecting on the 22nd of September 2022 performance and from the viewpoint of practice-led researchers interested in experimenting with humanistic applications for AI, this article presents a use case for OpenAI Stable Diffusion in hybrid AI performances and offers commentary on how an audience of Business Education rectors, deans and administrators perceived the experience as viewers, and in the case of keynote speakers, as contributors.

Keywords

text-to-image, image-to-image, Stable Diffusion, hybrid AI performance, practice-led research, 30th Annual CEEMAN Conference.

...

A post-humanist art history would see instead all art works, from cave paintings through to the works of so-called Great Masters and contemporary experiments with all kinds of technologies, as having been produced by human artists in an assembly with a plethora of nonhuman agents: drives, impulses, viruses, drugs, various organic and nonorganic substances and devices, as well as all sorts of networks – from mycelium through to the Internet.

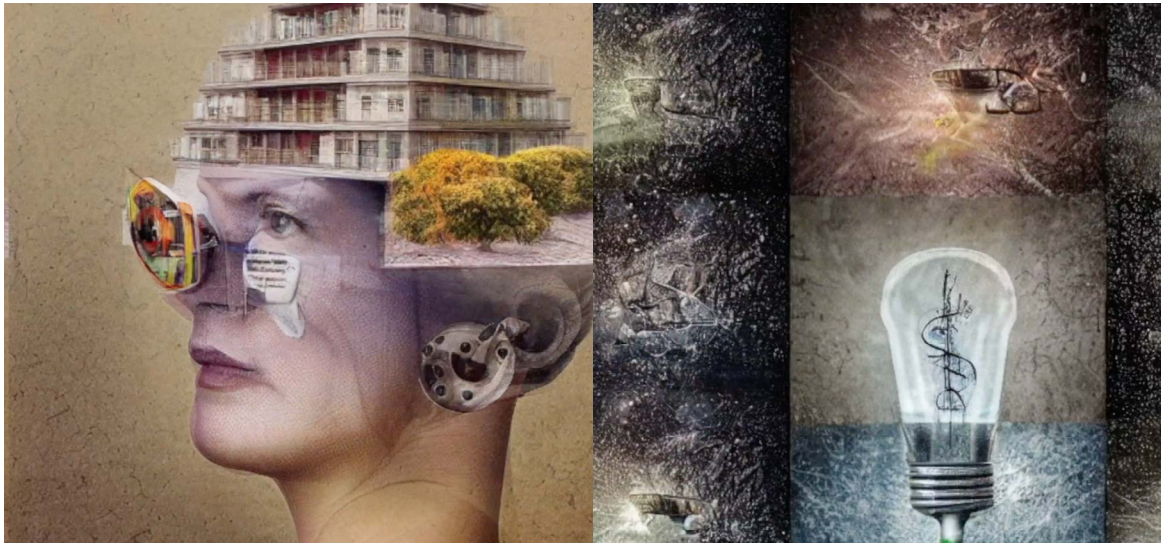
- Joanna Zalynska, *AI Art* (2020)

Introduction

Creative applications for machine learning (ML), or AI, have grown exponentially since OpenAI's StableDiffusion text-to-image visualization model was made available to the general public in August of 2022. Now, roughly ten months after the first StableDiffusion collab notebook was released, Digital Art, Generative Art and Creative Technologies communities have produced a plethora of innovative use cases that incorporate StableDiffusion text-to-image and image-to-image visualisations in the creative process, across a wide-ranging selection of HCI projects. In September of 2022 the authors of this article presented a Hybrid AI Multi-Modal Live Performances (AI MMLP) using the elements of dance, projections, and the StableDiffusion colab notebook generated art as part of the 30th Annual CEEMAN Conference hosted by IEDC-Bled School of Management in Slovenia. The text that follows is an explanation of the processes the authors used to create the AI visualizations projected during an interpretive ballet performance by Slovenian ballerina Tijuana Križman Hudernik.

A Very Brief History of Computer Art

Computer Art, or Generative Art, has been around since the 1950's, and "from the theoretical point of view, this new art originated in cybernetics and general systems theory" (Boden 2019, 1). It



was around the same time that mathematicians and computer scientists at Dartmouth University began using games like chess to measure artificial intelligence and the ability of programmers to train machines from various types of data sets derived from board games. Since that time, however, there has been a great deal of debate regarding machines and creativity. Are computers – and more recently, AI, capable of creating Art? While a viable topic worthy of debate, this article is not framed from the perspective of whether machine learning models (AI) can produce art. Rather, it is an exploration of how it is possible for artists to collaborate with machines to generate creative outputs that explore and draw comparisons between professional fields and academic disciplines that result in new observations that might not have been otherwise possible to see.

Concept

Keynote speakers at the CEEMAN conference, held by The International Association for Management Development in Dynamic Societies, were asked by event organizers at IEDC-Bled

Figure 1. “I think this performance was food for our souls and a great way of demonstrating how technology, art and the idea of education come together.”

School of Management to imagine what the future of management education might look like. The speakers recorded videos two to three minutes in length delivering an imagined inspirational address to their employees in the future (2027). Each speaker hypothetically looked back at the past five years (2022-2027) to synthesize how they envisioned business education and professions evolving between now and then. RISEBA University of Applied Sciences Media Art and Creative Technologies (MACT) PhD candidate Sabrina Durling-Jones used key phrases and images in each speaker’s video to generate AI visualizations with Stable Diffusion text-to-image and image-to-image processing. Assistant Professor Aigars Cepītis, M.F.A., and the Dean of RISEBA’s Faculty of Media and Creative Technologies, the originator of the concept and created the score for the hybrid performance piece, which combined AI moving image projections

as the backdrop for a soloist ballet dancer who interpreted the computer-processed imagery in the collaborative piece through movement. The core issue of the concept revolves around the impact the business community has on the development of historic events while often being oblivious to the processes it has initiated. Notwithstanding the precarious economic conditions at present, business disciplines continue to direct their gaze forward, focusing on digitization and sustainability and creating a growing number of collaborative clusters comprised of business professionals and interdisciplinary artists working in innovative ways with immersive technologies, visual data, and Artificial Intelligence in order to expand into markets of the future. AI MMLP, thus, explored the potential of collaborations between business and the arts and to reflect on the electrifying possibilities such endeavors may offer, allowing the participants to come visually face to face with the residuals of the future projections they themselves have set off.

Process

There is currently a great deal of ire surrounding computer visualizations derived from the vast data set of the Internet. Certainly, from the artist's perspective, creating visual output using AI (essentially machine learning model training) is, in many ways, unattractive. If one is an artist whose work is among those used by online image libraries for model training, questions of ownership, privacy and compensation are rightfully very important. Also significant is the acknowledgment that images generated when ML models process the data they are trained with, whether textual or visual, artists who do experiment or work with AI likely will not see their personal vision or artistic perspectives reflected in what AI generates from images not of the artists' own making. Meaning, computer vision is not often a reflection of an artist's personal creative vision. In the case of the CEEMAN conference, a

convoluted approach to creating AI animations was taken to see if it was possible to inject more of the artists' and conference speakers' perspective into the visuals generated by AI. It was important to have the animations reflect the themes discussed by the conference speakers so the audience at the conference would be able to identify on some level with the subject matter being projected. As a result, when generating visualizations using diffusion models, it was necessary to continuously nudge the models and make slight adjustments throughout collab notebook runs so the image interpretations produced by the machine reflected the messages conference speakers wanted to emphasize. Some of the approaches taken included processing each of the videos the keynote speakers provided with a modified Disco Diffusion collab notebook, inserting key phrases from each speaker to generate a new AI vision version of their videos. A strong, representative image was then extracted from each of the videos and used as the initial image when generating visualizations using the Stable Diffusion collab notebook. Further excerpts from the presenters' speeches, along with the starting image for each, were then used to run 1200 visualizations in Stable Diffusion. Once the 1200 images for each speaker were generated, they were upscaled to 4K and edited into animated video clips, adjusting frame rate to reduce flicker and to create as smooth of a video interpolation as possible. Once the video projections were completed, the composer was then able to design a musical score and work with the ballet dancer to make sure all three elements (video, audio, performance) were effective in creating a live, immersive and multi-sensory hybrid AI experience for the audience.

Visualizations and Audience Perception

The performance, which was commissioned by IEDC-Bled School of Management, provided a unique opportunity to blend human-AI creative collaboration experiments with Business Education



Figure 2. “I really like the idea of having a dancer here because I think she anchored us back to humanity. This performance was about technology and the mental shift, but at the end of the day machines will be machines and humans will be humans. I think the way she performed helped us remember that.”

themes and concepts and gain some insight into how computers can process human language to produce creative outputs. It also provided the unique opportunity to gauge audience reactions to experiencing humanistic applications of mechanical processes. Because it is not possible to include footage of the live performance, provided here instead are key still images from the piece, accompanied by audience member observations regarding the performance that unfolded before them ¹.

Conclusions

While computers have been successfully deployed in music, film, video gaming and many other creative sectors for decades, “AI in the creative industries have dramatically increased in the last five years” (Anantrasirichai & Bull 2020, 591). Although there is a great deal of debate about computers and AI having any capacity at all to be creative, the approach taken in this article is to demonstrate how a collaboration between carefully curated human inputs and diffusion models can

help an audience synthesize concepts across disciplines. Using text and image contributions from the creators of those concepts alongside computer generated visual outputs influenced by the human contributors, allow for a re-imagined, or re-conceptualised approach to the synthesis of ideas across disciplines. The residual effect of for those working in the field of economics are not only to witness the electrifying possibilities the collaborative efforts may offer, but, more importantly, to raise an astute awareness of the impact some business decisions might have on local communities and the world at large, whenever the obliviousness to the processes the business people have initiated have not faced up to the ugly and the redemptive at an appropriate time.

...



Figure 3. “The artificial-intelligence-generated art provided us with the imagery of both hope and anxiety about the future. The music was atonal and gave us anxiety, but there were also moments of calm. We discussed the surrealism in the imagery and how everything blended together. My colleague also mentioned that she felt a lack of connection with nature – and is that not just an organic occurrence since it was interpreted by artificial intelligence?”

¹ Audience reactions presented here are quoted from the published Proceedings of the 30th Annual CEEMAN Conference, 21-23 September 2022.



Figure 4. “One of our key impressions was that we saw and felt a world in motion; a world in permanent change, a flow without finality and in progress. We also saw a mix of anxiety and serenity on one hand, and a mix of distance and proximity with this very reality. On the other hand, we could also identify key topics such as gender, sustainability and war. There was a shift from despair to hope, a shift from darkness to light.



References

Anantrasirichai, N., & Bull, D., 2020. "Artificial intelligence in the creative industries: a review." *Artificial Intelligence Review*, 55, 589-656.

Audry, S., 2021. *Art in the Age of Machine Learning*. MIT Press.

Benjamin, W., 1969. *The Work of Art in the Age of Mechanical Reproduction*. Penguin Books.

Boden, M.A., 2019. *A Taxonomy of Computer Art. From Fingers to Digits*. MIT Press.

Chefer, H., Lang, O., Geva, M., Polosukhin, V., Shocher, A., Irani, M., Mosseri, I., & Wolf, L., 2023 . "The Hidden Language of Diffusion Models." *arXiv*, abs/2306.00966v3, 1 June 2023.

Lee, S., Hoover, B., Strobel, H., Wang, Z.J., Peng, S., Wright, A.P., Li, K., Park, H., Yang, H., & Chau, D. "Diffusion Explainer: Visual Explanation for Text-to-image Stable Diffusion." *arXiv*, abs/2305.03509, 4 May 2023.

Purg, D. et. al., 2022. Conference Proceedings of the 30th CEEMAN Annual Conference. *The International Association for Management Development in Dynamic Societies*. Retrieved from <https://www.ceeman.org/publications/30th-ceeman-annual-conference-proceedings>.

Sautoy, M.D., 2019. "The Creativity Code." *Science*, 364, 842 - 842.

Zylinska, J., 2020. *AI Art: Machine Visions and Warped Dreams*. Open Humanities Press.

...

