AI Frontiers:
Augmenting Human Creativity

Creativity Applications Focus

1. Music

2. Visual Arts

3. Impact of AR/VR/XR

AI in Music Creation

➢ Sony (but all majors) apply AI to artists' music.
- Sometimes still left to artist choice. Will probably see an increment of applications of ML to commercial music to increase sales and audience development.

➢ Music Generation/Completion from an already existing melody
- Muse GAN (Generative Adversarial Networks) [1]

➢ Melody from text (custom-length)
- Stable Audio, Stability AI (Mixed Latent diffusion Model)[2]
- Prompt = "Trance, Ibiza, Beach, Sun, 4 AM, Progressive, Synthesizer, 909, Dramatic Chords, Choir, Euphoric, Nostalgic, Dynamic, Flowing" [3]

➢ Lyria Model, DeepMind & Youtube[4]
- **Dream Track** – Experiment to deepen connections between artists, creators, and fans through music creation.
- **Music AI tools** – Set of tools designed with artists, songwriters, and producers to help the creative processes.

AI in Visual Arts

➢ Generate new images from input images (Neural Style Transfer - CycleGAN)
- Over-imposing a specific style to an image
- Generation of scenery landscape photographs from paintings

➢ Image generation from text prompts
- Midjourney, Stability AI (Stable Diffusion)
- DALL-E3, OpenAI (Meta-Transformer)

➢ Robotic arms painting (Robohood & Pindar Van Arman)
- Quantum Computing + AI + Robotic arm painting

➢ Moving & Animating Images
- Runway ML: text-to-video (0:38)

➢ Eventually, Movies & Real-Time AI entertainment
- Initially blend of human-generated content with AI
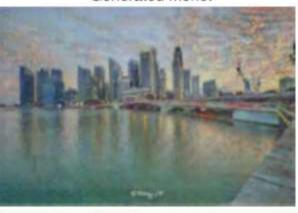- Eventually, 100% AI generated content

Cycle GAN

Content target + Style reference = Combination image

| Input | Monet | Van Gogh | Cezanne | Ukiyo-e |

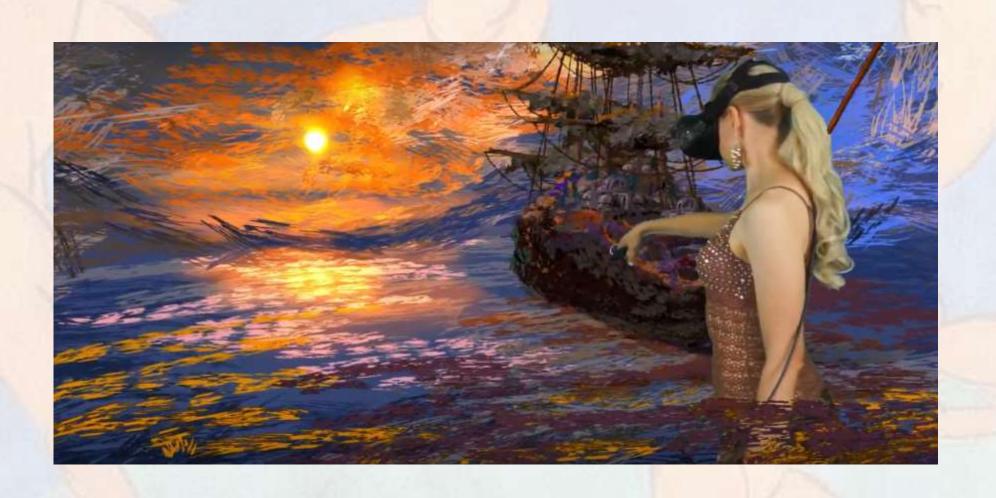# Cycle GAN 2

Midjourney vs. DALL-E3

Runway ML Example

➢ Script by GPT 4

➢ Images by Midjourney

➢ Animated with Runway ML

➢ Text-to-Speech by ElevenLabs

➢ Audio by Stable Audio

Multi-Modality, RL, AR/VR/XR

➢ Increasingly seamless multimodality interactions will aid creators

- A glance at early possibilities (text (or speech) to sound): Gemini Ultra

➢ As Reinforcement Learning is introduced into foundational models:

- We expect "planning capabilities over time steps" to get better. This will lead to the ability to create so called "semi-autonomous agents" (not AGI) and move away the time and effort of creators from medial tasks and outsource to multi-modal "semi-autonomous agents" to amplify the potential scope of creative projects.

➢ VR, XR, AR are just some examples of future trends. We expect these new forms of creativity to become more immersive as we unlock creativity to take advantage of new media of expression

ART in VR

Conclusion

➢ AI and technology in general will extend the medium of expression.
  - AI Kill creativity? Might spur an evolution of forms of creativity

➢ Technology Advancements will transform art and empower artists

References

1. Example: Doe, J. (2022). An Innovative Approach to Machine Learning. arXiv. https://arxiv.org/abs/1234.56789

2. Stability AI. (Year). Stable Audio: Efficient Timing Latent Diffusion. Retrieved from https://stability.ai/research/stable-audio-efficient-timing-latent-diffusion

3. Stable Audio. (Year). [Home page]. Retrieved from https://www.stableaudio.com/

4. Author(s). (Year). Transforming the Future of Music Creation. DeepMind. Retrieved from https://deepmind.google/discover/blog/transforming-the-future-of-music-creation/

5. Example: AI Innovations. (2023, March 5). Understanding Machine Learning. YouTube. https://youtu.be/UIZAiXYceBI?t=295

6. RunwayML. (Year). Generation 2: Text-to-Video. Retrieved from https://research.runwayml.com/gen2

7. Example: Foster, D. (2019). Generative Deep Learning. O'Reilly Media.

8. Example: VR Art World. (2021, June 10). Hopes: Virtual Reality Painting. YouTube. https://www.youtube.com/watch?v=xxxxx