

Ethical implications of AI-generated images

Manuel Muehlberger ¹ Sofia Kaltwasser ²

¹TU Munich

²University of Potsdam

Ethik 4 Nerds

08.03.2023

Presentation Overview

- 1 Introduction
- 2 technical
- 3 ethical
- 4 Conclusion

AI-generated images

- Significant advancements in recent years
- Realistic and high-quality images → almost indistinguishable from human-created images
- example Dall-E 2 by OpenAI [1]

Dall-E 2 generated images examples



(a) A tapir with the texture of an accordion [2]

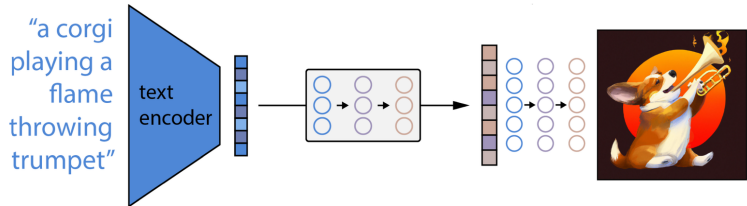


(b) Illustration of a baby hedgehog in a christmas sweater walking a dog [2]

Potential for misuse

- Highly realistic images of individuals without their consent
- Portrayal of socially unacceptable actions:
 - Violence
 - Child pornography
 - Racism
 - ...

High level abstraction of image generation process of DALL-E 2



High-level overview of operating principle of DalleE-2 [modified from [3]]

Contrastive Language-Image Pre-training (CLIP)

link between textual semantics and their visual representations
good explanation:
<https://www.assemblyai.com/blog/how-dall-e-2-actually-works/>

Possible solutions

- Clear labeling of AI-generated images
- Obtaining explicit consent from individuals depicted in AI-generated images
- Agreement on ethical framework for AI
- Establishing legal liability for misuse of such images

References



OpenAI.

Dall-e 2 openai, 2022.



Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever.

Zero-shot text-to-image generation.

CoRR, abs/2102.12092, 2021.



Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen.

Hierarchical text-conditional image generation with clip latents, 2022.