OVERVIEW

- What is Stable Diffusion?
- Business Opportunity
- Research
- Coding Demonstration
- Ethical Concerns
- What's Next?

STABLE DIFFUSION 2.1

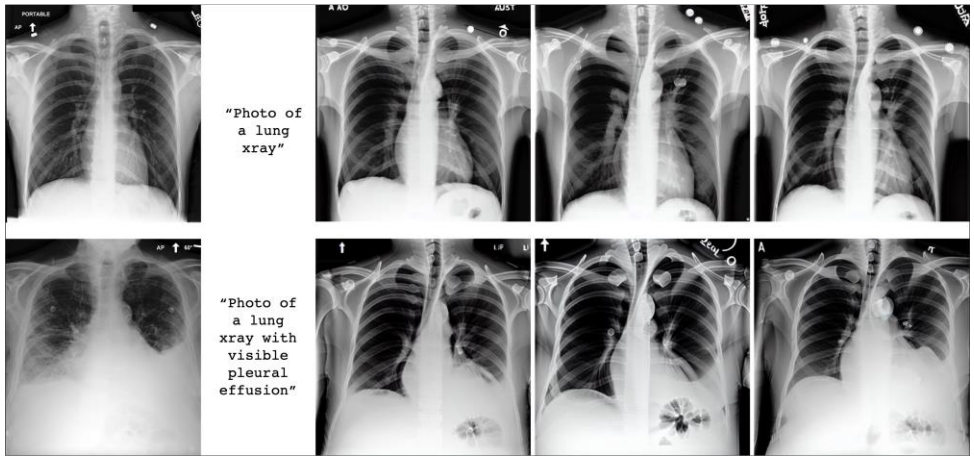
- Patrick Esser & Robin Rombach
- Released December 27, 2022
- Python
- General Disclaimer

WHAT IS STABLE DIFFUSION

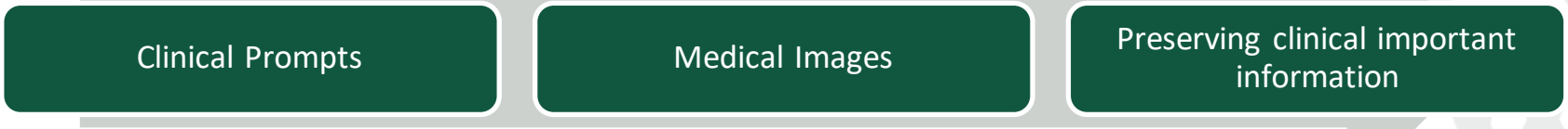
- High-resolution image synthesis
- Text to image
- Text to video
- Image to image
- Image to video

PROBLEM STATEMENT

- If Stable Diffusion models can accurately produce medically realistic images of observable abnormalities, rare diseases and clinically relevant features.



- First study to look at latent diffusion modeling for generating medical images.
- Chest X-Ray's (CXR's) most common imaging modality



- Researchers found the U-Net tuning to be the most effective.
- Overall, they found that medical diagnostic features were well preserved when generating realistic medical images.

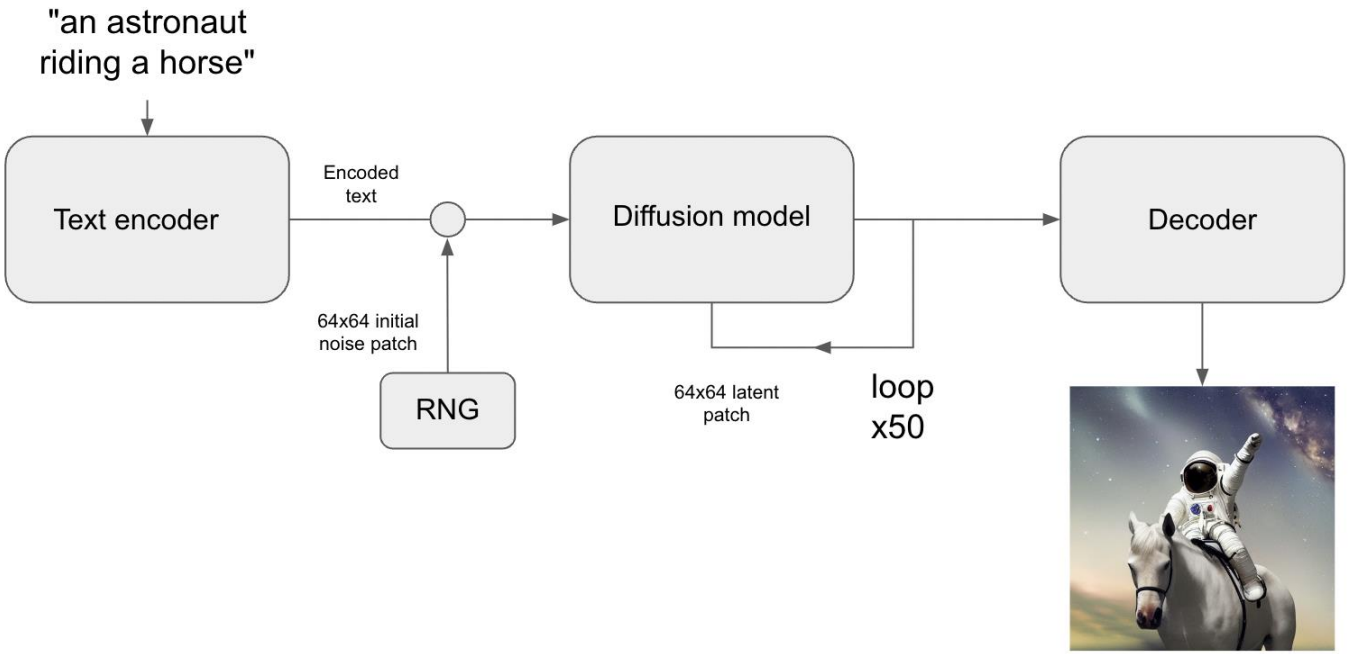
LIMITATIONS

- Questionable accuracy measures to assess clinical correctness
- Low diversity in images generated
- Text prompts do not fully correspond with words used in the clinical setting.

FUTURE DIRECTIONS

- Image generation can extend to a wide range of abnormalities.
- The ability to combine abnormalities (comorbidity).
- Models will be able to generate images for other body parts.

PROCESS


"an astronaut riding a horse"
Text encoder
Encoded text
Diffusion model
Decoder
64x64 initial noise patch
RNG
64x64 latent patch
loop x50

TECHNICAL DEMO

ETHICAL CONCERNS

Copyright / Ownership

Human job loss

Malicious use

WHAT'S NEXT?

Stated Goals

- Improved performance/quality
- Optimize for chipsets other than NVIDIA
- Reduce adverse outcomes

Current Limitations

- Perfect photorealism
- Legible text
- Compositionality
- People/Faces
- Trained in English

QUESTIONS?

VIDEO DEMONSTRATION

- https://twitter.com/i/status/1557517982095626241

WORKS CITED

Chambon, P., Bluethgen, C., Langlotz, C. P., & Chaudhari, A. (2022). Adapting pretrained vision-language foundational models to medical imaging domains.

Chollet, F., Wood, L., & Gupta, D. (2022, December 9). High-performance image generation using stable diffusion in KERASCV Tensorflow Core. TensorFlow. Retrieved April 12, 2023.

CompVis. (2022). Stable diffusion online. Stable Diffusion Online. Retrieved April 12, 2023.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.

Blattmann, A., Rombach, R., Oktay, K., Müller, J., & Ommer, B. (2022). Retrieval-augmented diffusion models. Advances in Neural Information Processing Systems.