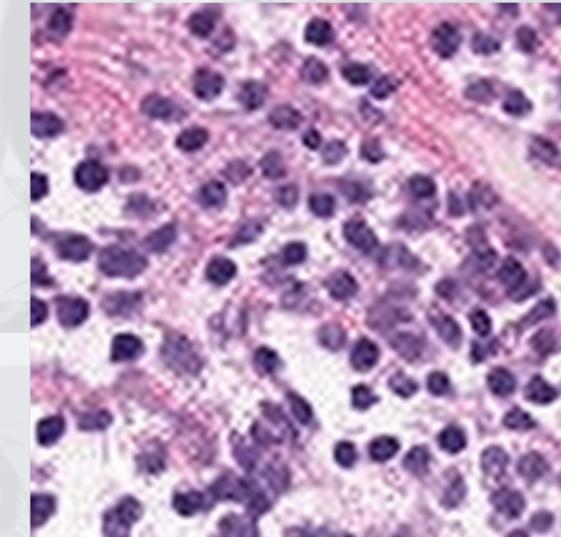
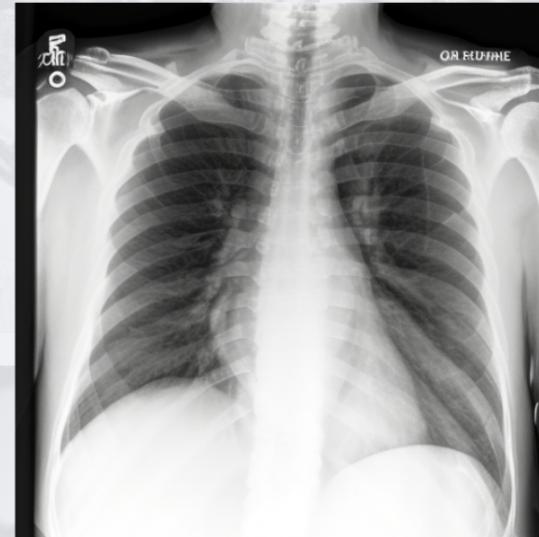
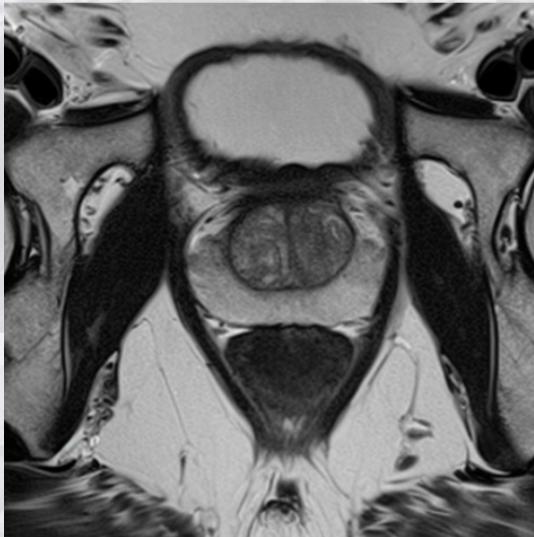


Medical diffusion on a budget

Textual Inversion for medical image generation



Bram de Wilde, Anindo Saha, Maarten de Rooij, Henkjan Huisman, Geert Litjens

Radboudumc
university medical center

Diffusion models

Popular for text-to-image modeling

Stable Diffusion: **open source** + inference on **single GPU**

Diffusion models

Popular for text-to-image modeling

Stable Diffusion: **open source + inference on single GPU**



Diffusion models

Popular for text-to-image modeling

Stable Diffusion: **open source + inference on single GPU**



✳️ Training requires a lot of **compute** and **data**

Diffusion models

Popular for text-to-image modeling

Stable Diffusion: **open source + inference on single GPU**



✳️ Training requires a lot of **compute and data**

🏥 Medical domain can have **rare diseases and local compute**

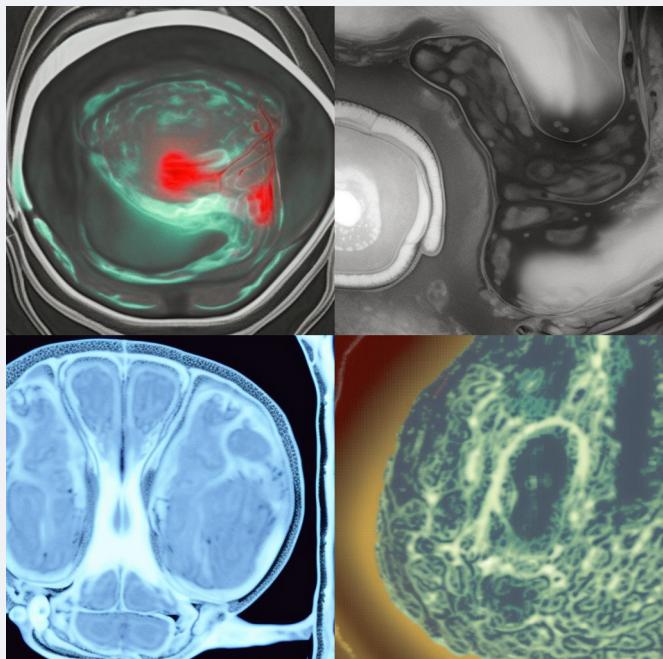
Prostate MRI?

 Does Stable Diffusion already know medical imaging?

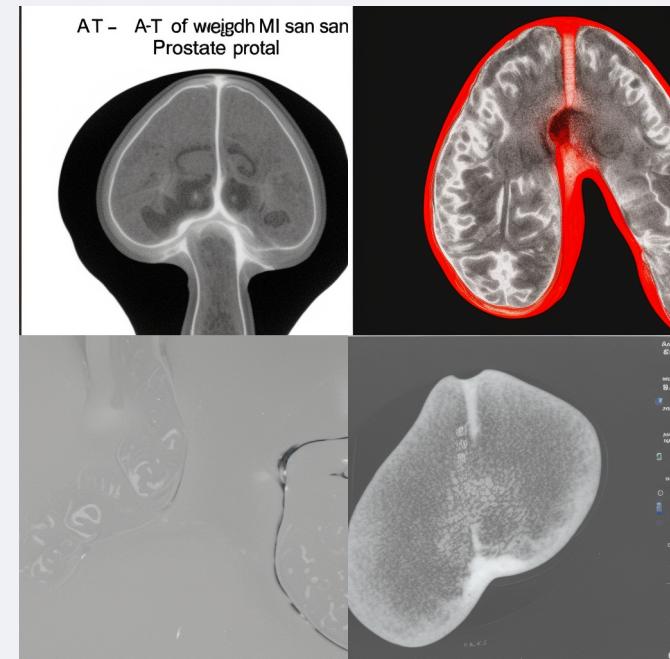
Prostate MRI?

🧐 Does Stable Diffusion already know medical imaging?

✍ “a prostate MRI scan”



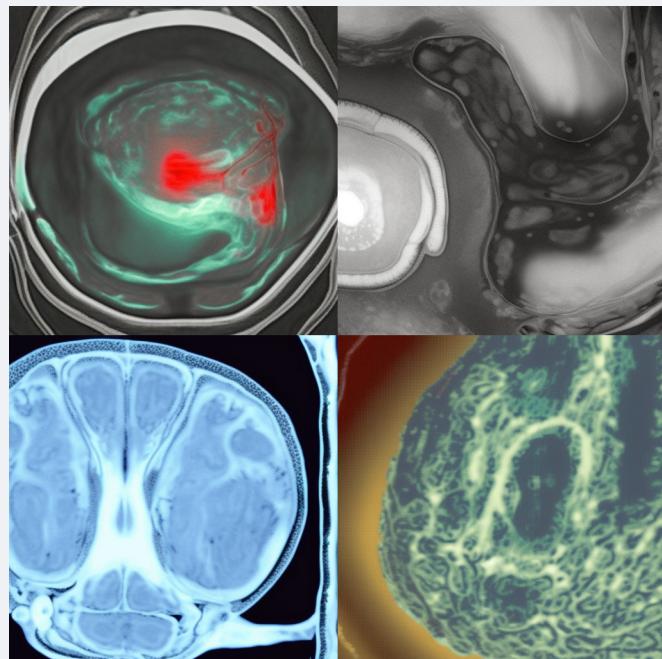
✍ “a T2-weighted MRI scan of a prostate”



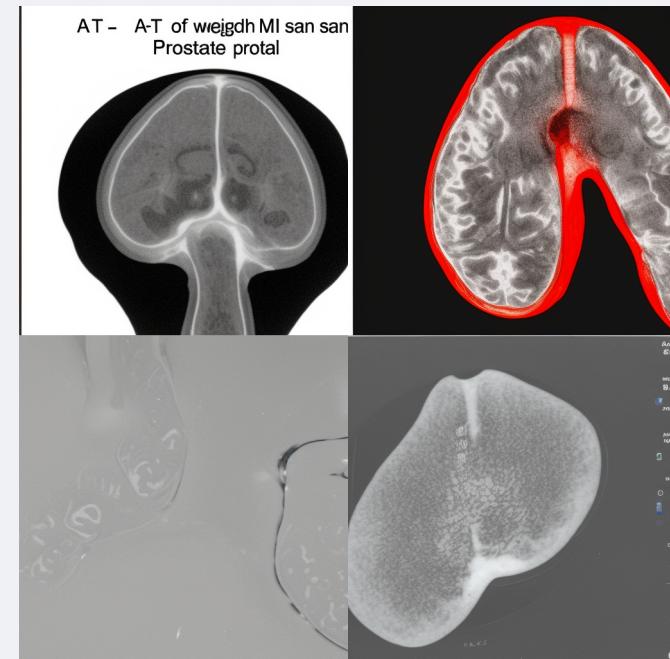
Prostate MRI?

💡 Does Stable Diffusion already know medical imaging?

💡 “a prostate MRI scan”



💡 “a T2-weighted MRI scan of a prostate”



💡 Fine-tuning towards medical domain necessary!

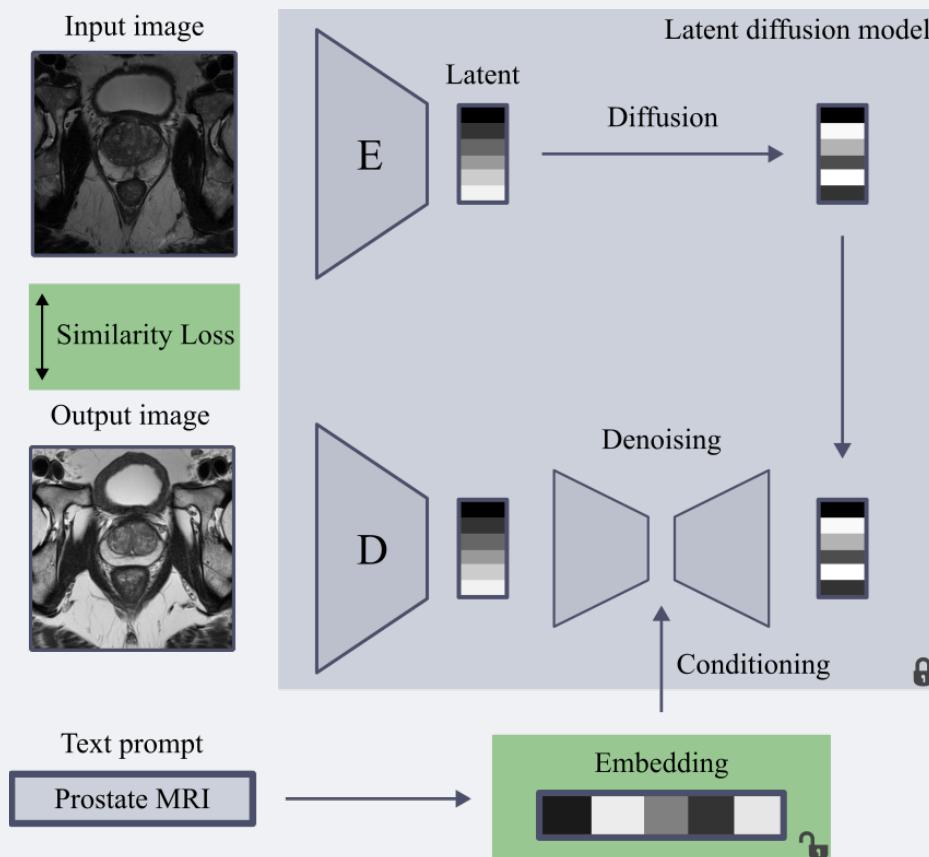
Fine-tuning diffusion models

Various methods fine-tune sub-parts of diffusion model

Fine-tuning diffusion models

Various methods fine-tune sub-parts of diffusion model

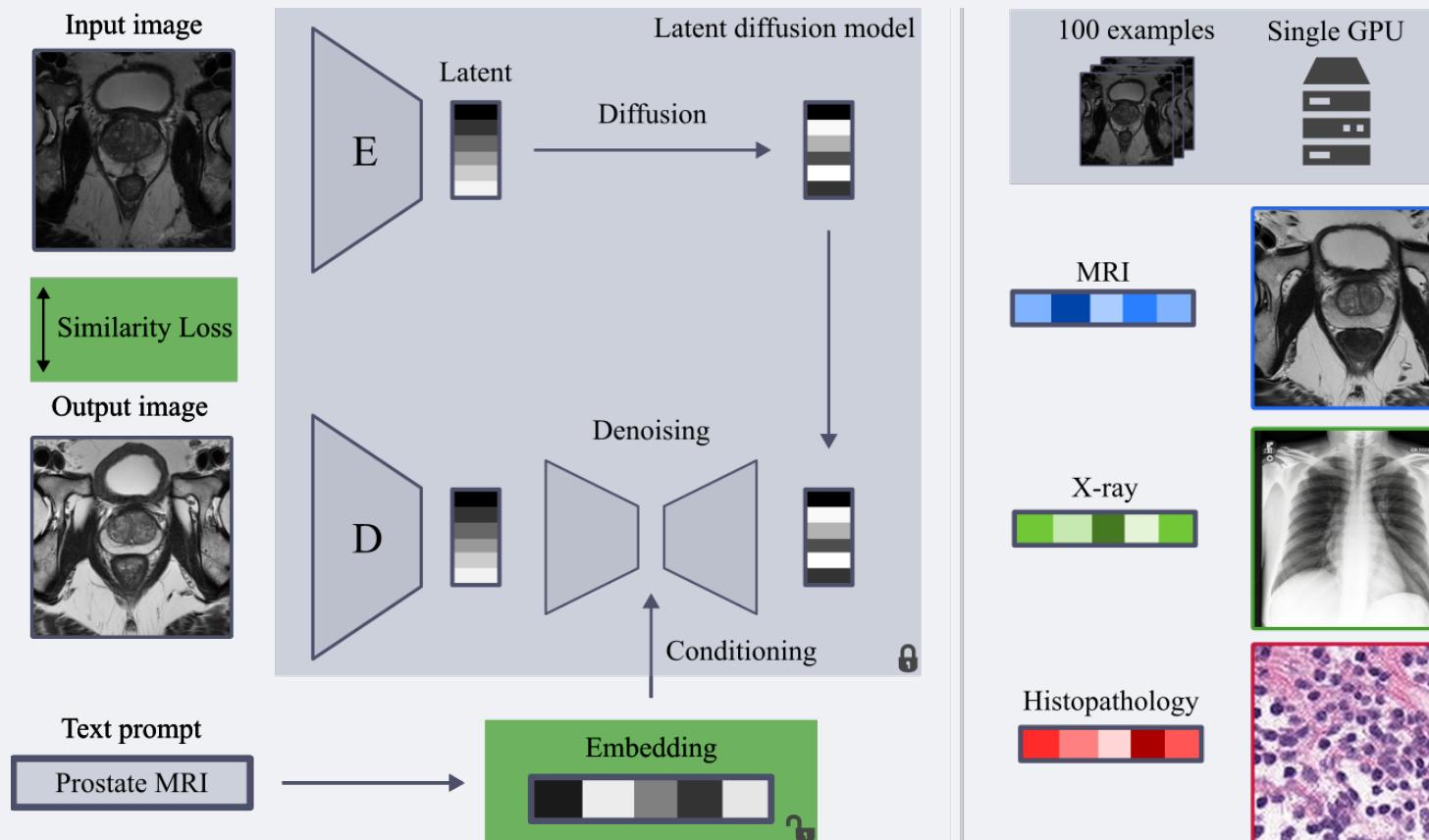
💡 Textual Inversion only trains token embedding



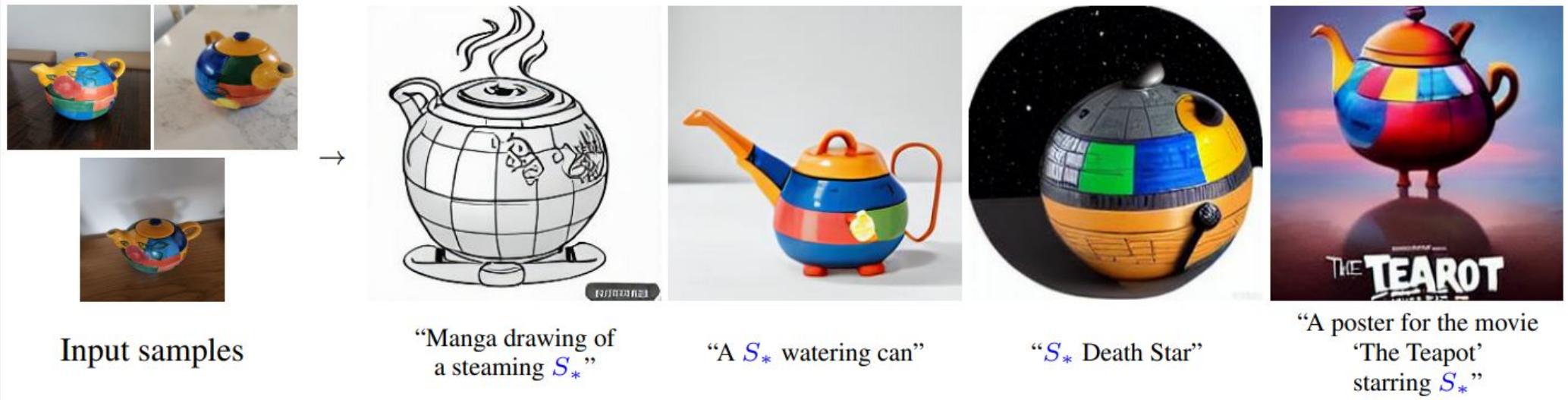
Fine-tuning diffusion models

Various methods fine-tune sub-parts of diffusion model

💡 Textual Inversion only trains token embedding

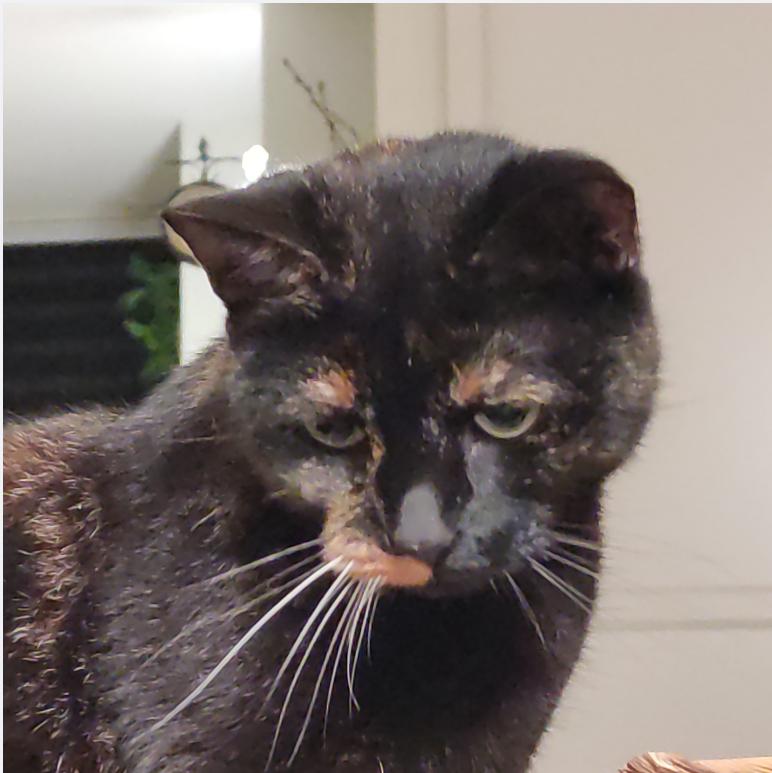


Textual Inversion



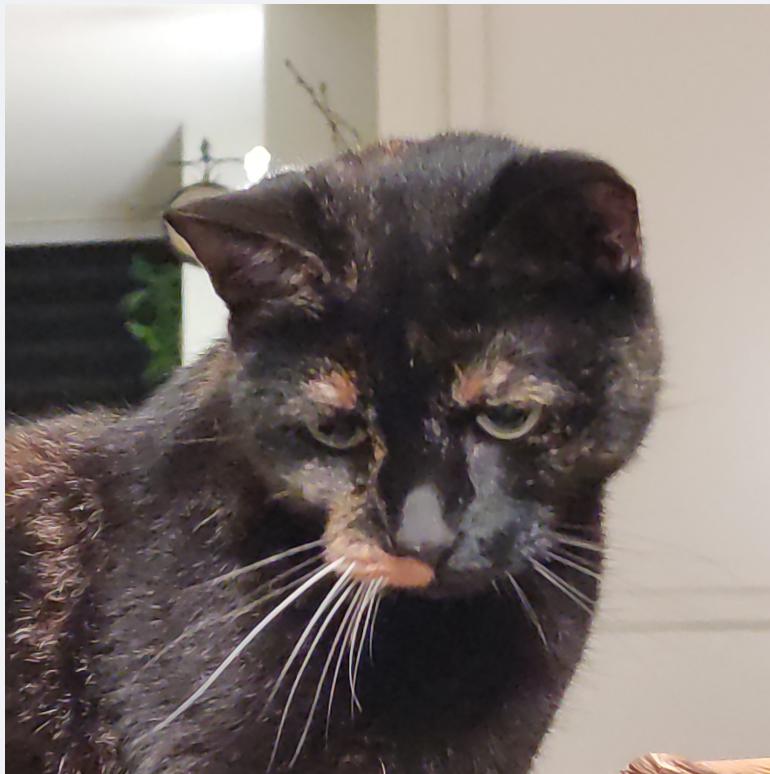
📋 An Image is Worth One Word: Personalizing Text-to-Image Generation using Textual Inversion, Gal et al., 2022

First steps

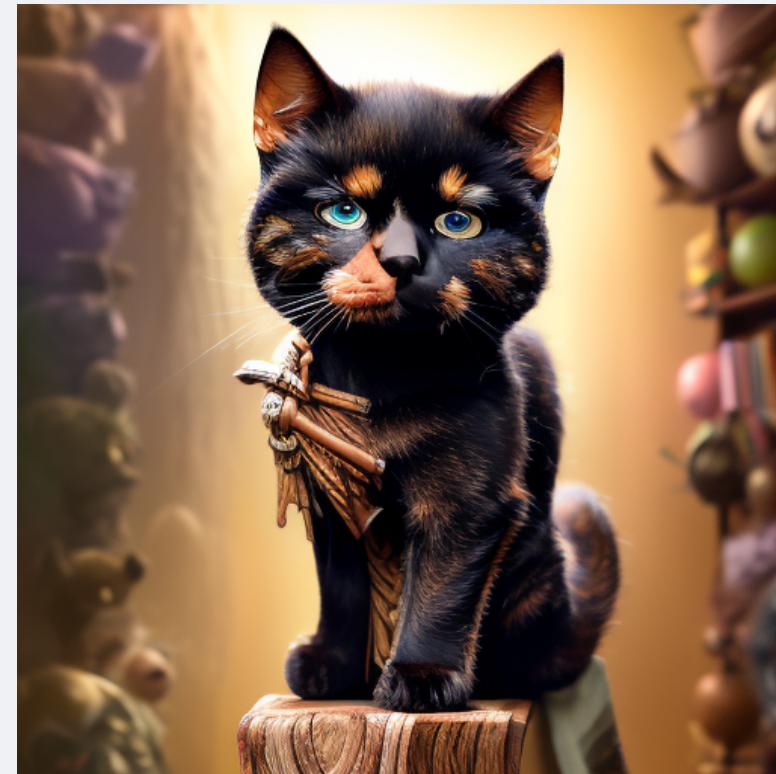


 Shelly

First steps

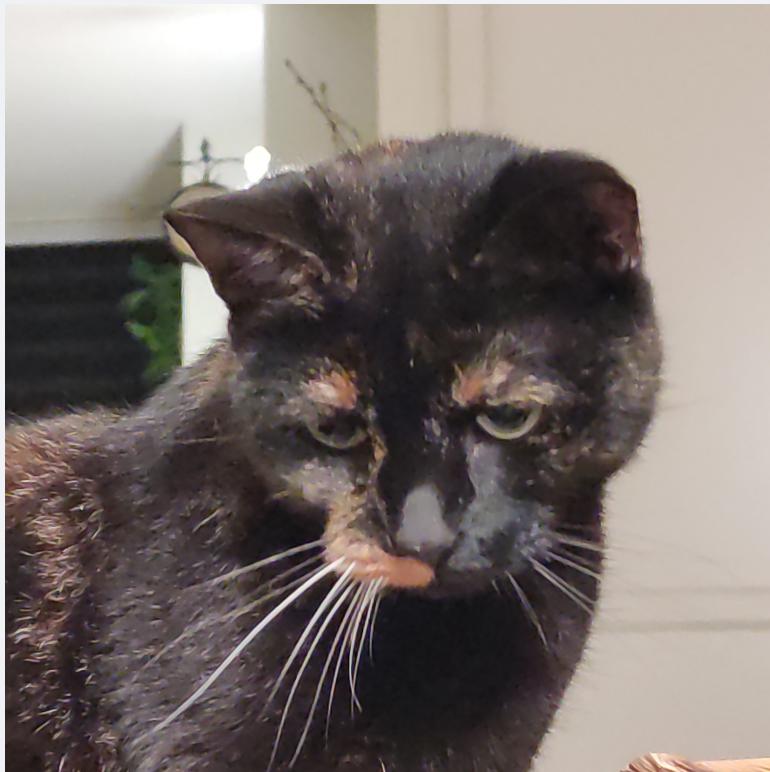


🐱 Shelly



🖌️ “a painting of a 🐱 as a knight”

First steps



😺 Shelly



🖌️ “a painting of a 🐱 as a knight”



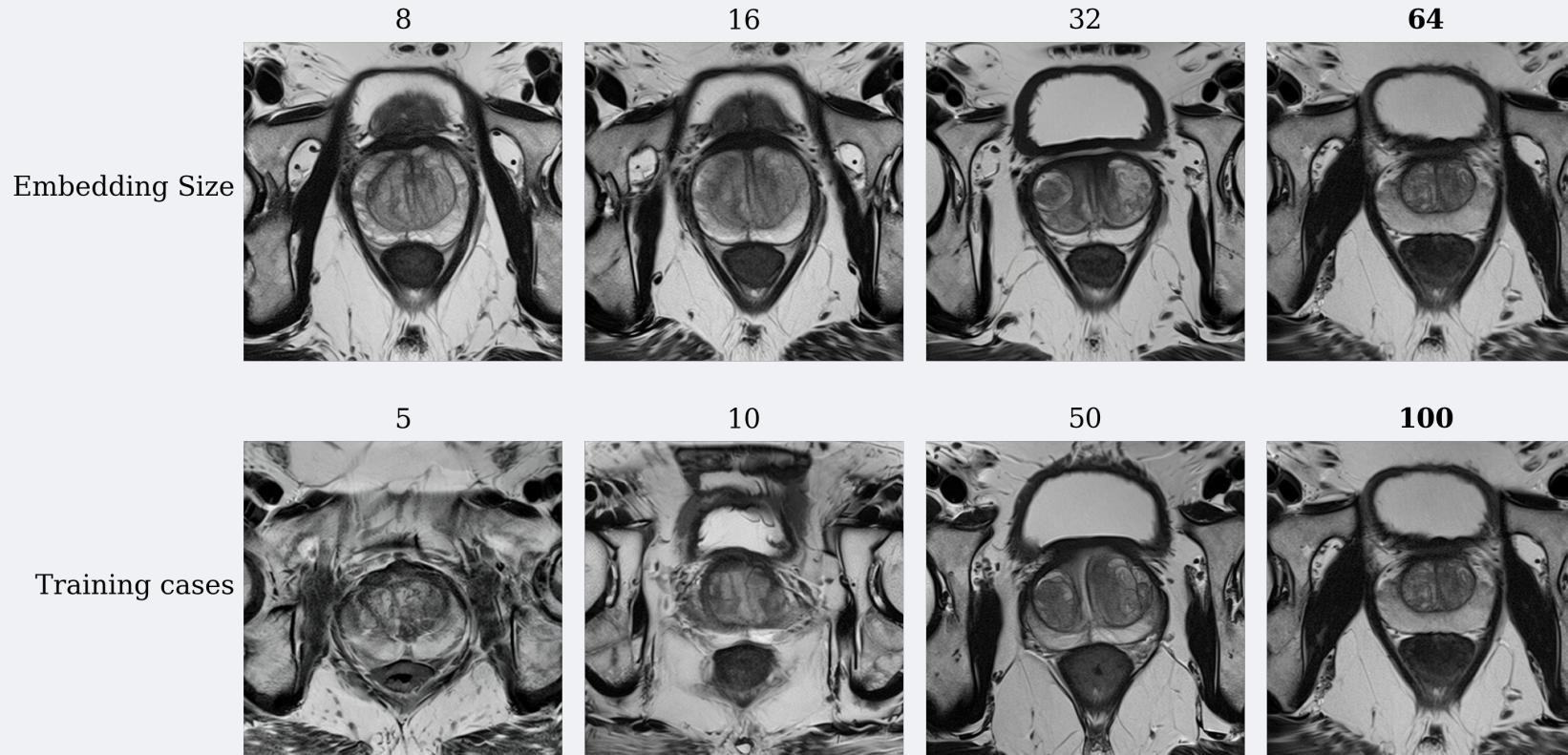
Explore Textual Inversion for medical images

Adapt to medical imaging

Original work uses small embeddings and ~5 examples

Adapt to medical imaging

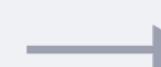
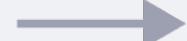
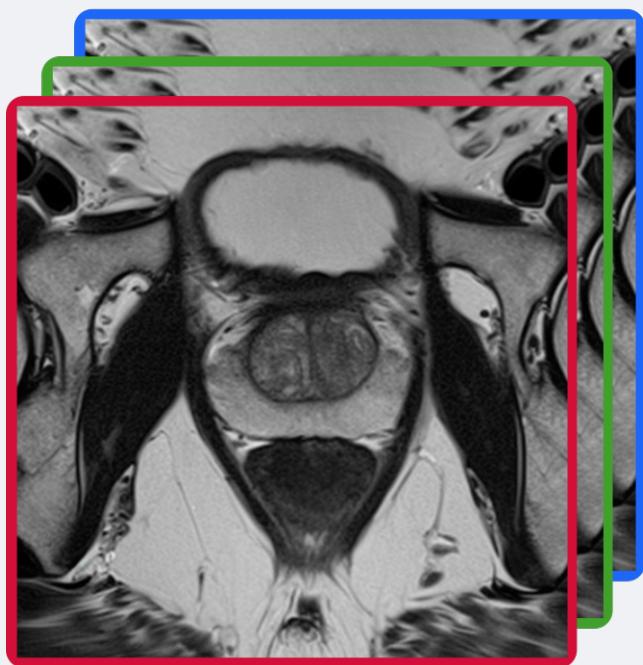
Original work uses small embeddings and ~5 examples



 Use larger embeddings and more examples

Classification

Prostate MRI

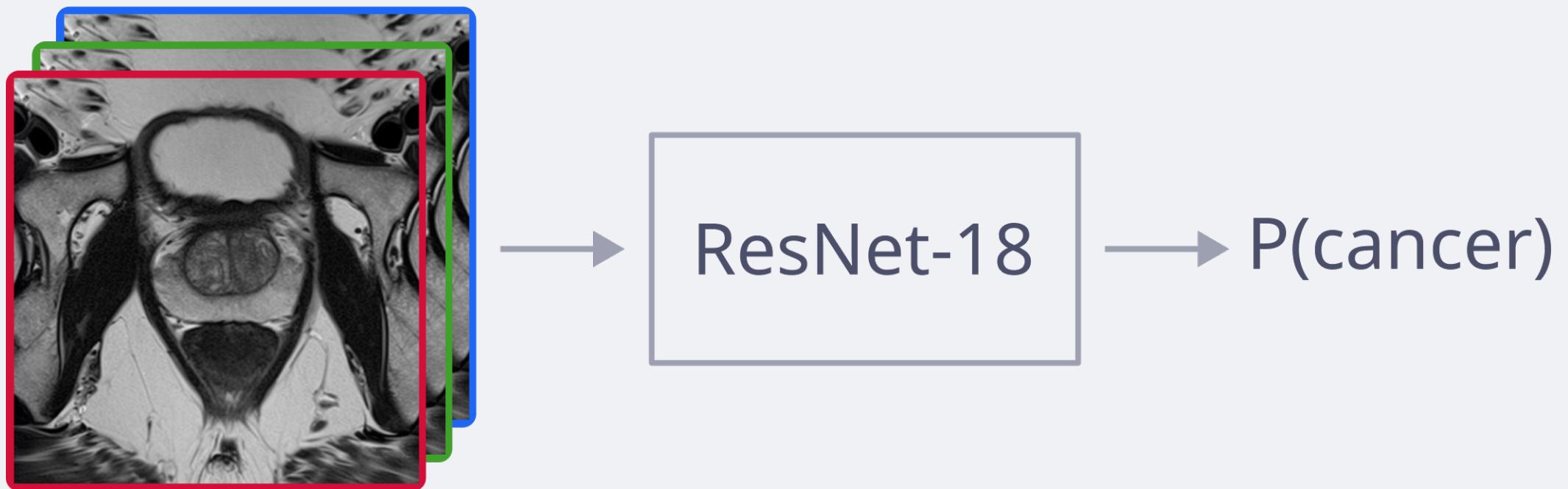


$P(\text{cancer})$

Classification

⚡ Train negative+positive embedding on 100 cases

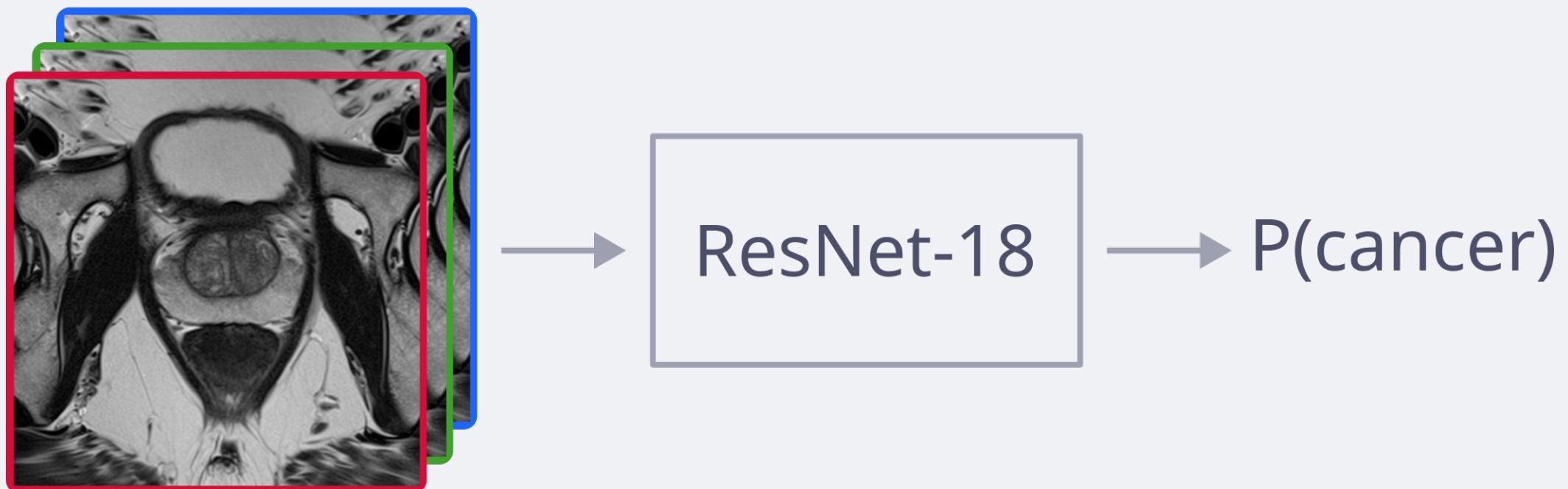
Prostate MRI



Classification

- ⚡ Train negative+positive embedding on 100 cases
- 💿 Encode modalities as RGB channels (T2W, ADC, DWI)

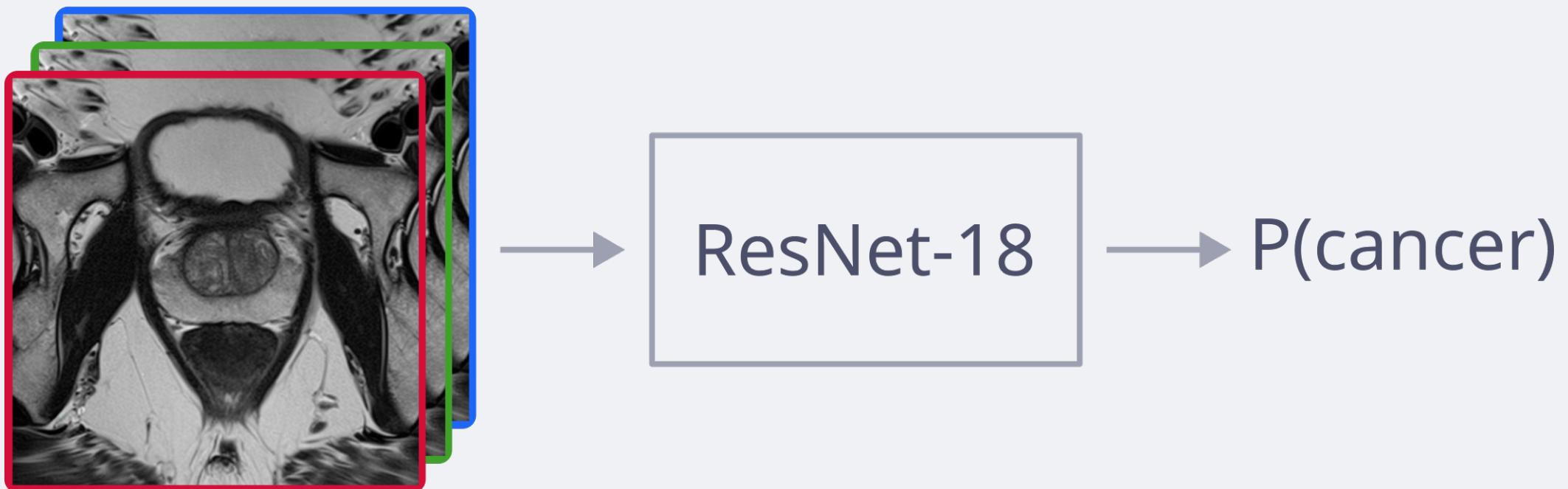
Prostate MRI



Classification

- ⚡ Train negative+positive embedding on 100 cases
- 💿 Encode modalities as RGB channels (T2W, ADC, DWI)
- 💾 200/200/200 train/val/test

Prostate MRI



Classification

| #Real | #Synthetic | AUC \pm std - Prostate MRI |
|-------|------------|------------------------------|
| 200 | 0 | 0.780 ± 0.017 |
| 200 | 2000 | 0.803 ± 0.009 |

 Adding synthetic cases **maintains or improves** performance

Classification

| #Real | #Synthetic | AUC ± std - Prostate MRI |
|-------|------------|--------------------------|
| 200 | 0 | 0.780 ± 0.017 |
| 200 | 2000 | 0.803 ± 0.009 |
| 0 | 2000 | 0.766 ± 0.020 |

- ✓ Adding synthetic cases maintains or improves performance
- ✗ Using only synthetic data gives small performance drop

Classification

| #Real | #Synthetic | AUC ± std - Prostate MRI |
|-------|-------------------|--------------------------|
| 200 | 0 | 0.780 ± 0.017 |
| 200 | 2000 | 0.803 ± 0.009 |
| 0 | 2000 | 0.766 ± 0.020 |
| 0 | 2000 ¹ | 0.562 ± 0.036 |

- ✓ Adding synthetic cases maintains or improves performance
- ✗ Using only synthetic data gives small performance drop
- 👎 Training on 10-case embeddings shows quality difference

Comparison to GAN baseline

 What about GANs?

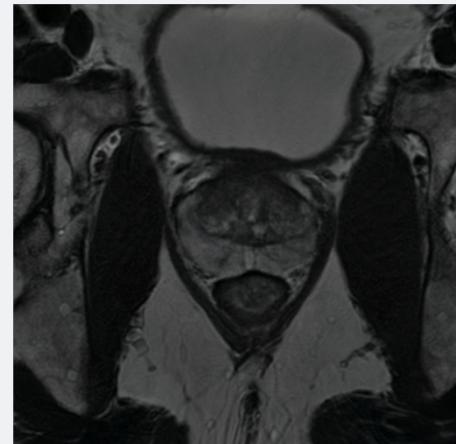
Comparison to GAN baseline

🔍 What about GANs?

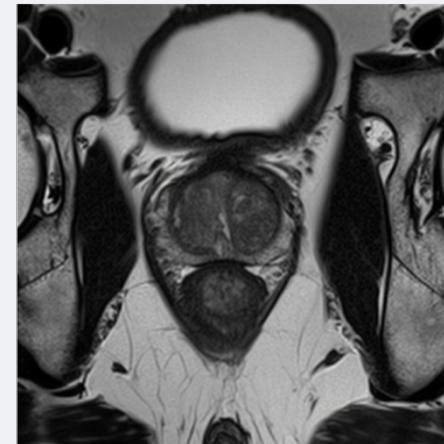
🔬 Fine-tune a pre-trained StyleGAN3 on 100 images

⚖️ Similar training time and compute

StyleGAN3



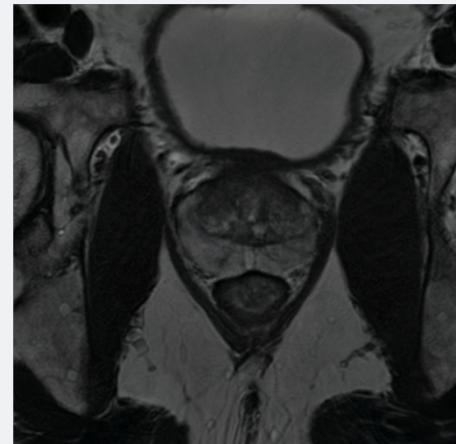
Stable Diffusion



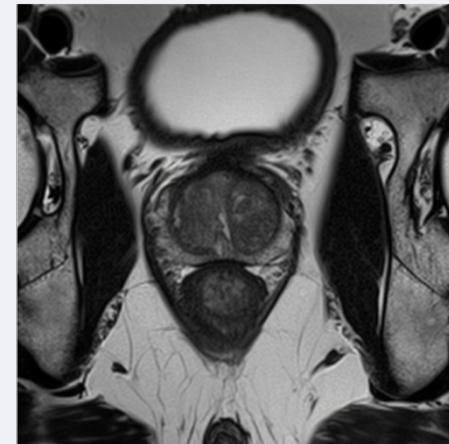
Comparison to GAN baseline

- 🔍 What about GANs?
- 🔬 Fine-tune a pre-trained StyleGAN3 on 100 images
- ⚖️ Similar training time and compute

StyleGAN3



Stable Diffusion



- 👨‍⚕️ Prostate radiologist preferred diffusion model (36/50)

Composing embeddings

↔ Interpolate between healthy and diseased state

Composing embeddings

↔ Interpolate between healthy and diseased state

⚡ Train two embeddings: **healthy**, **diseased**

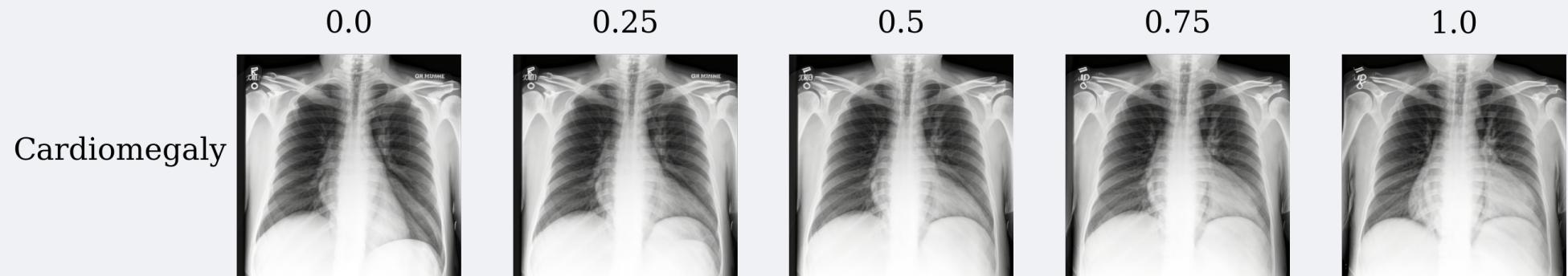
✍ “**healthy:30% AND diseased:70%**”

Composing embeddings

↔ Interpolate between healthy and diseased state

⚡ Train two embeddings: **healthy**, **diseased**

✍ “**healthy:30% AND diseased:70%**”



Composing embeddings

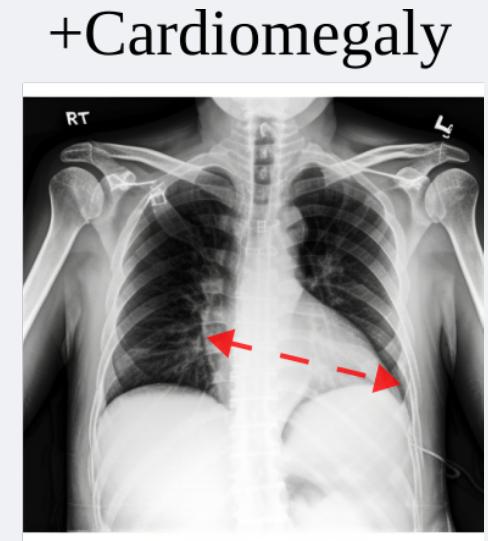
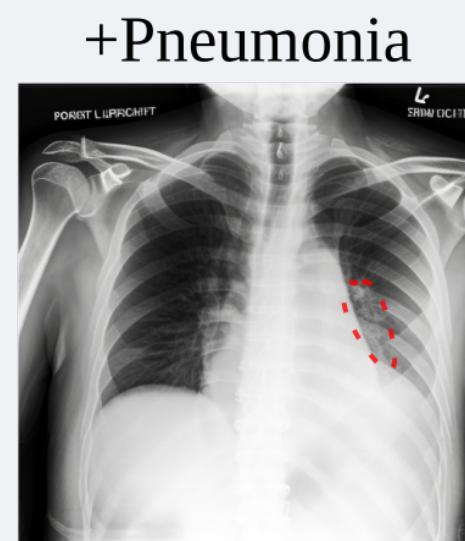
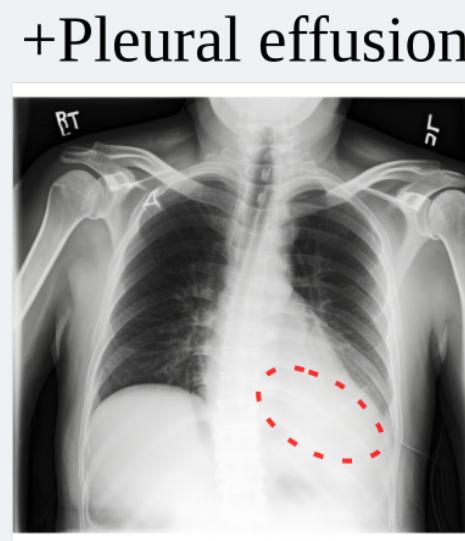
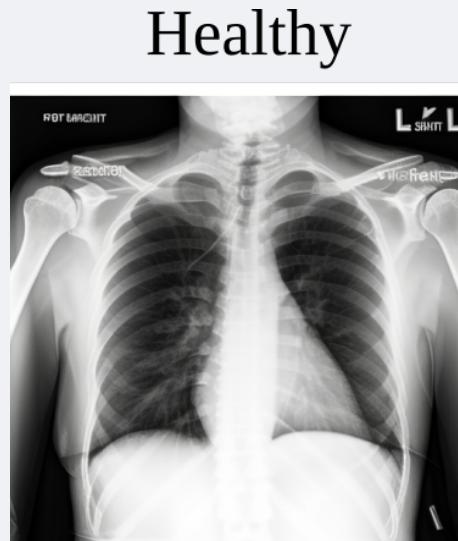
- + Combine multiple embeddings to show multiple diseases

Composing embeddings

- + Combine multiple embeddings to show multiple diseases
- ⚡ Train embeddings per disease
- ✍ “pleural_effusion AND pneumonia”

Composing embeddings

- + Combine multiple embeddings to show multiple diseases
- ⚡ Train embeddings per disease
- 📝 “pleural_effusion AND pneumonia”



Disease inpainting

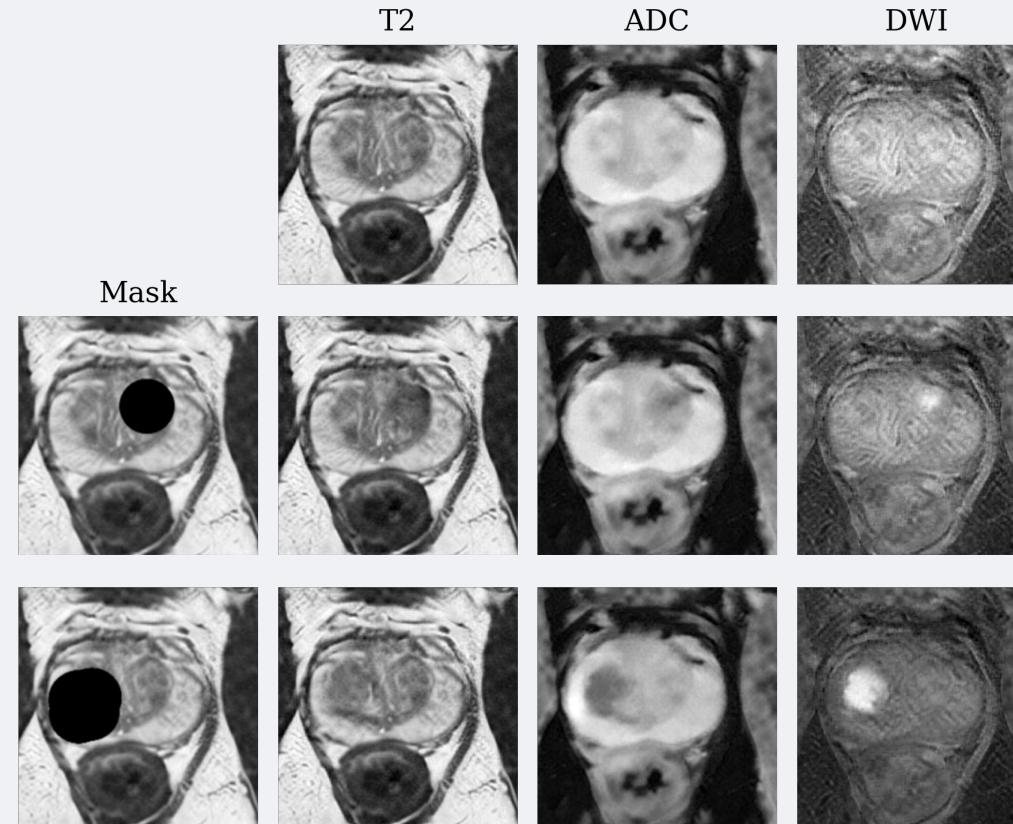
 Mask part of image and denoise

 Precise control over disease location

Disease inpainting

🖌️ Mask part of image and denoise

🔍 Precise control over disease location



Conclusion

⚡ Fine-tune Stable Diffusion with 100 images

Conclusion

- ⚡ Fine-tune Stable Diffusion with 100 images
- 🔍 Images are diagnostically accurate

Conclusion

- ⚡ Fine-tune Stable Diffusion with 100 images
- 🔍 Images are diagnostically accurate
- ↔ Compose multiple embeddings

Conclusion

- ⚡ Fine-tune Stable Diffusion with 100 images
- 🔍 Images are diagnostically accurate
- ↔ Compose multiple embeddings
- 🖌️ Inpainting for disease appearance

Conclusion

- ⚡ Fine-tune Stable Diffusion with 100 images
- 🔍 Images are diagnostically accurate
- ↔ Compose multiple embeddings
- 🖌️ Inpainting for disease appearance

Future directions

- 👉 Fine-tune medical diffusion model

Conclusion

- ⚡ Fine-tune Stable Diffusion with 100 images
- 🔍 Images are diagnostically accurate
- ↔ Compose multiple embeddings
- 🖌️ Inpainting for disease appearance

Future directions

- 👉 Fine-tune medical diffusion model
- 👤 Applicable to 3D text-to-image models

Conclusion

- ⚡ Fine-tune Stable Diffusion with 100 images
- 🔍 Images are diagnostically accurate
- ↔ Compose multiple embeddings
- 🖌️ Inpainting for disease appearance

Future directions

- 👉 Fine-tune medical diffusion model
- 👤 Applicable to 3D text-to-image models
- 💊 Study utility for rare diseases

Conclusion

- ⚡ Fine-tune Stable Diffusion with 100 images
- 🔍 Images are diagnostically accurate
- ↔ Compose multiple embeddings
- 🖌️ Inpainting for disease appearance

Future directions

- 👉 Fine-tune medical diffusion model
- 👤 Applicable to 3D text-to-image models
- 💊 Study utility for rare diseases
- 🔧 Controlled synthesis with **composing** and **inpainting**



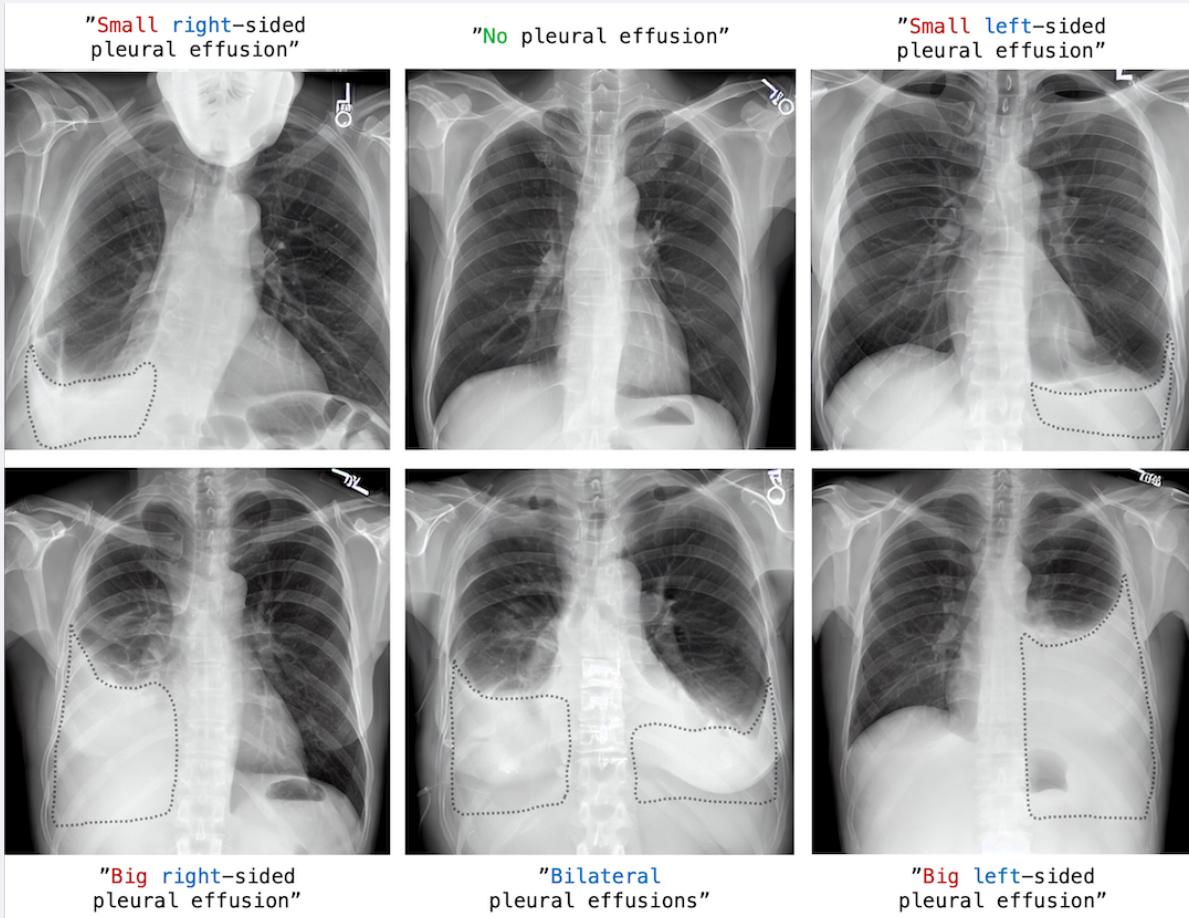
Try it out!



Scan for    & more!

Radboudumc
university medical center

Chambon et al. 2022



⚡ Train diffusion model on $O(10^5)$ chest x-ray + report

🔥 64 A100 / 1 day