



Understanding Latent Semantics in GANs

Bolei Zhou

The Chinese University of Hong Kong (CUHK)

GANs for Synthesizing Images

2014



GANs [Goodfellow et al.]

2015



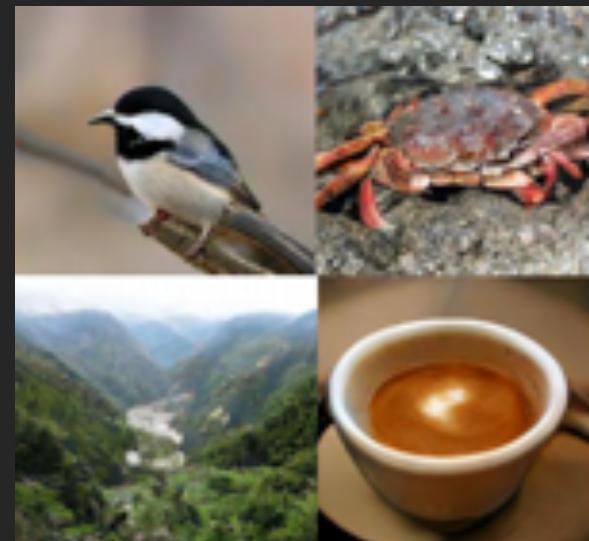
DCGAN [Radford et al.]

2017



PG-GAN [Karras et al.]

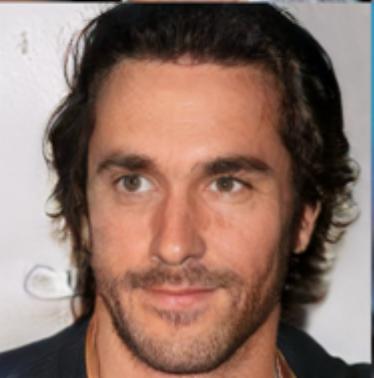
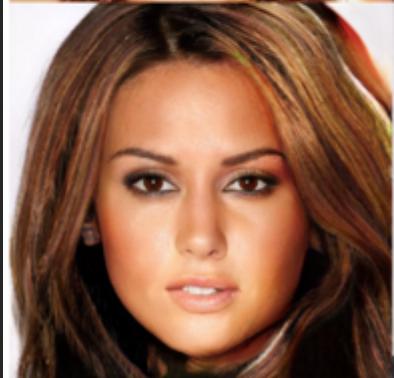
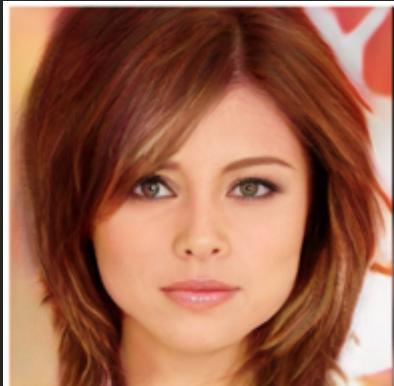
2018



BigGAN [Brock et al.]

...

GANs for Synthesizing Images

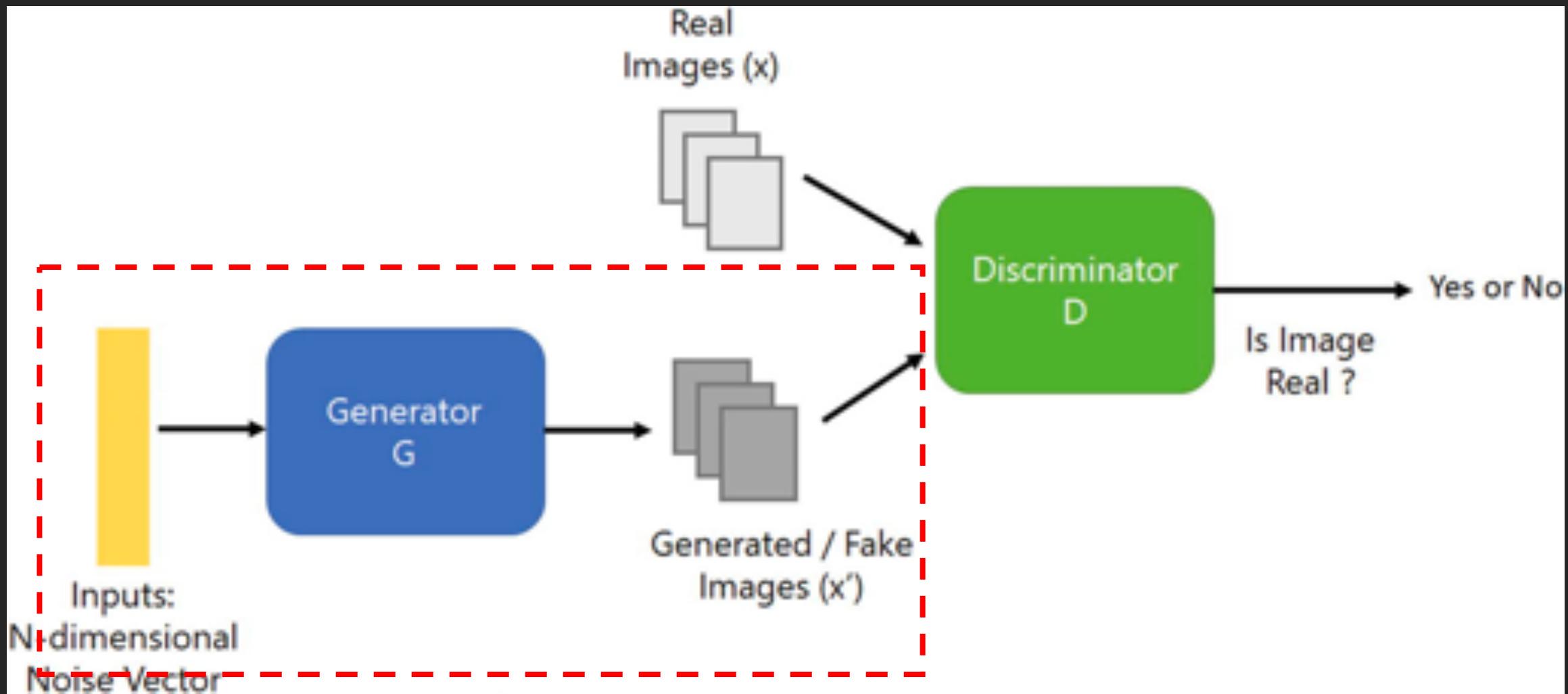


PG-GAN

StyleGAN

BigGAN

Training GANs



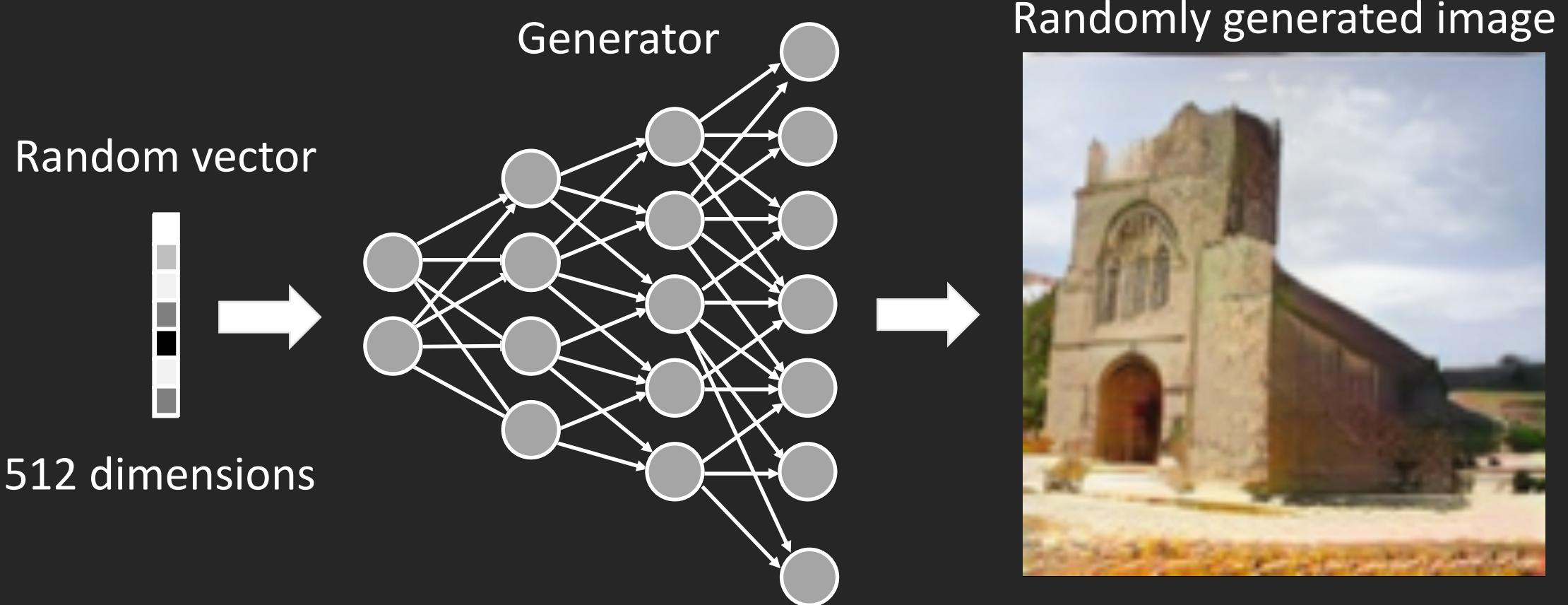
Tutorial Outline

- Interpretable units in intermediate layers
- Semantics in the latent space
- Inversion of real images
- Future directions

Tutorial Outline

- Interpretable units in intermediate layers
- Semantics in the latent space
- Inversion of real images
- Future directions

Deep Generative Representation

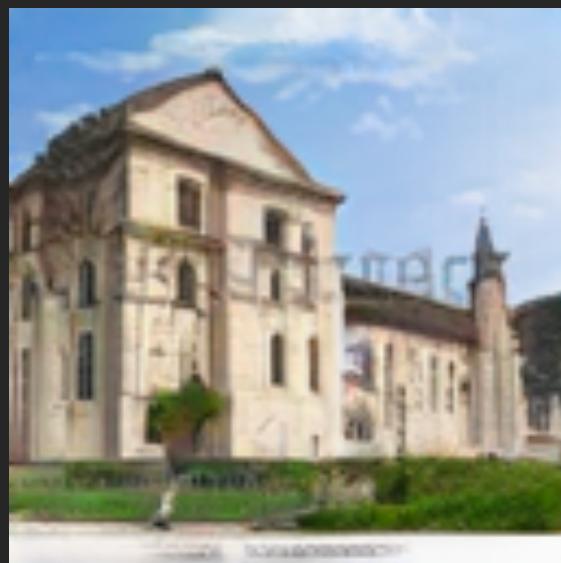




Progressive GANs [Karras et al., ICLR 2018]



Progressive GANs [Karras et al., ICLR 2018]

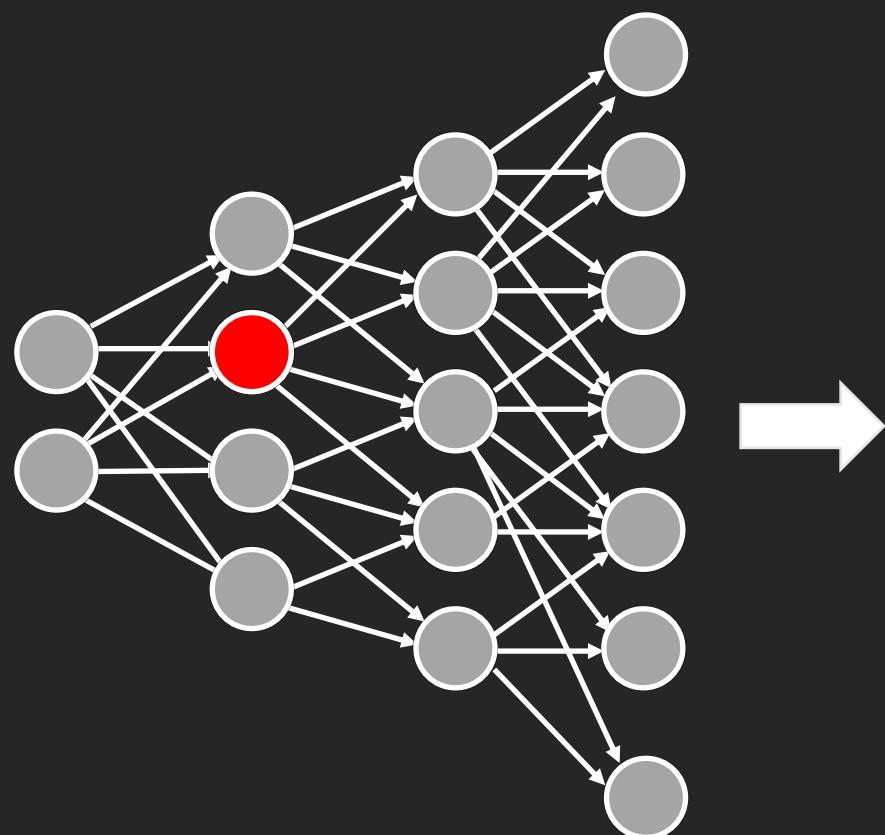


Are there watermark units?

Random vector



512 dimensions



Randomly generated image





shutterstock - 1000x600 - 15551642



alexander alishev - 1000x600 - 15551643



alexander alishev - 1000x600 - 15551644



shutterstock - 1000x600 - 15551645



alexander alishev - 1000x600 - 15551646



alexander alishev - 1000x600 - 15551647



alexander alishev - 1000x600 - 15551648



alexander alishev - 1000x600 - 15551649

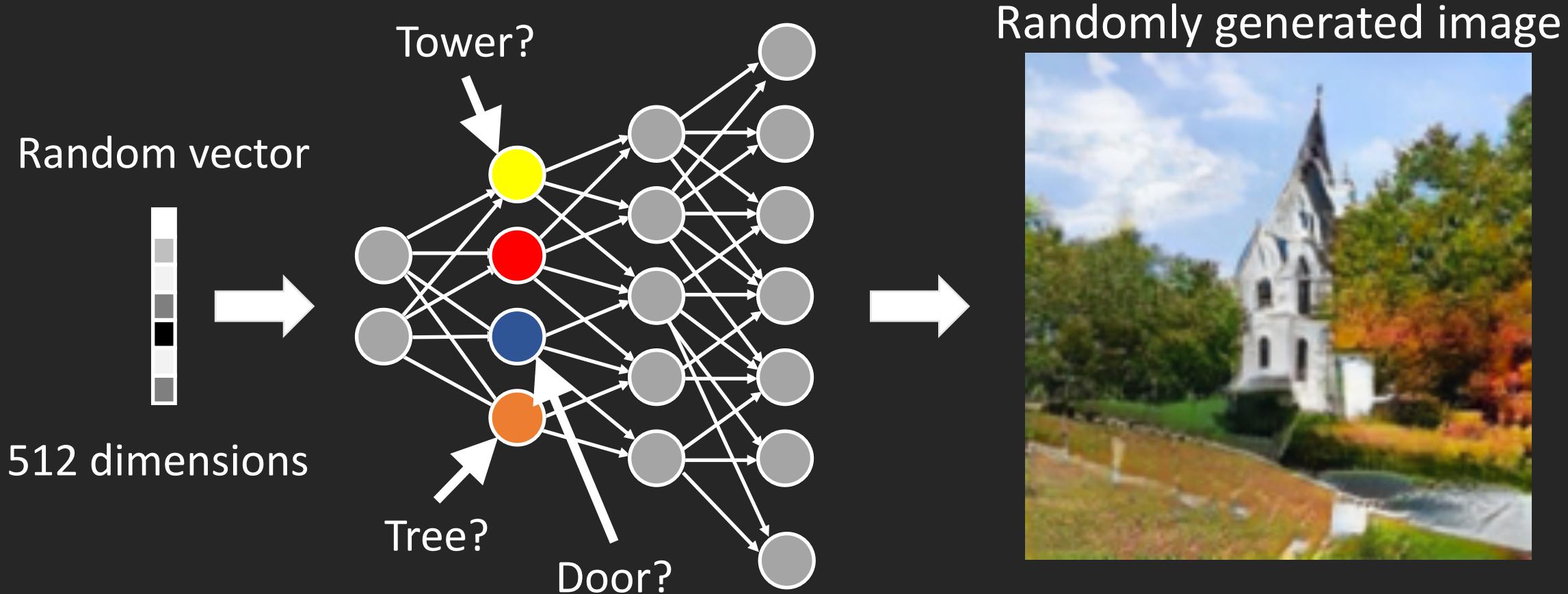
Deactivating banner units in layer 4



Deactivating watermark units in layer 4

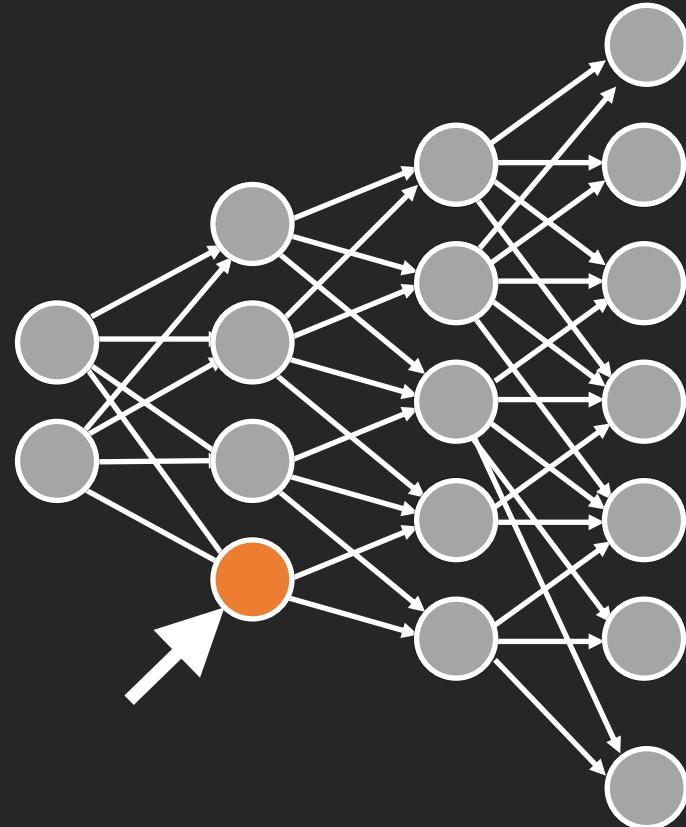


Are there other objects?



Are there other objects?

Random vector



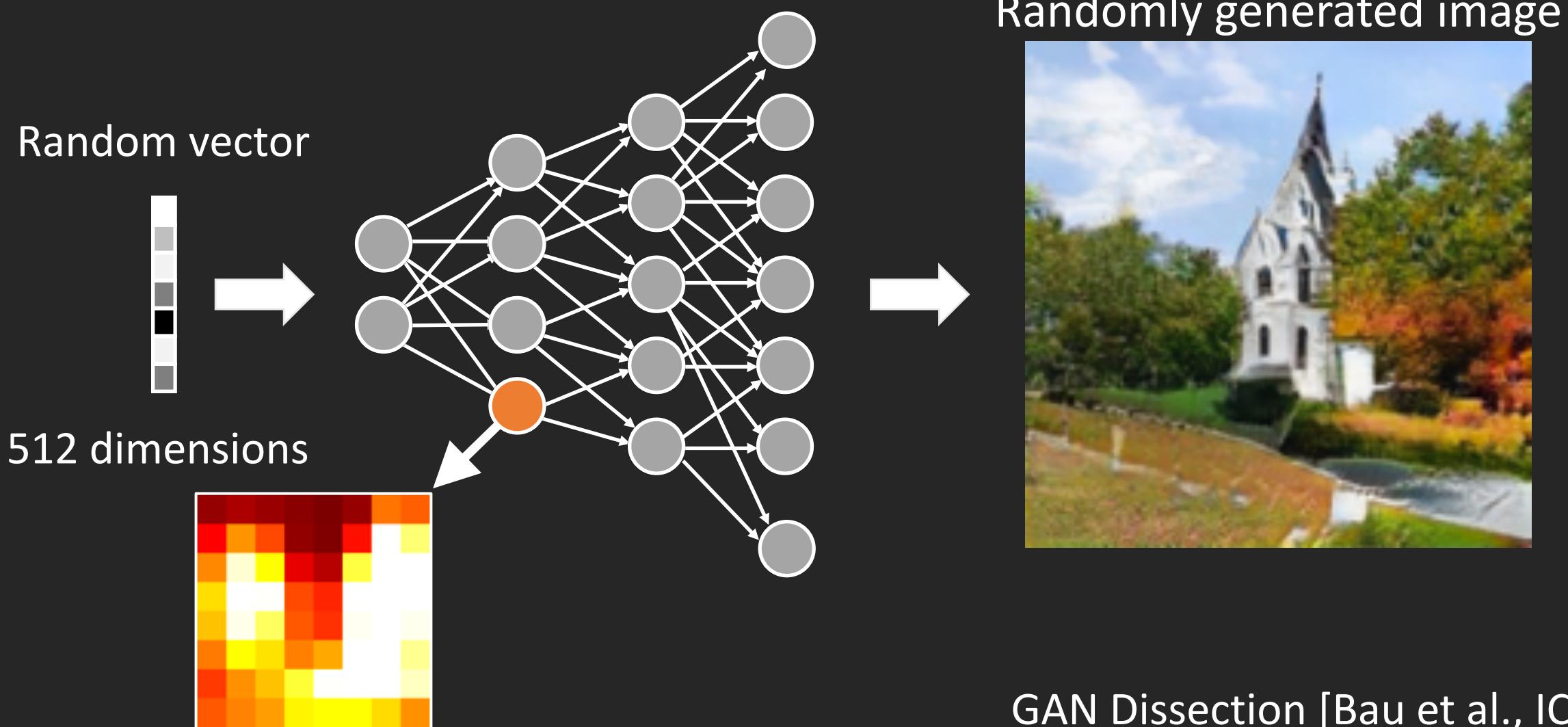
512 dimensions



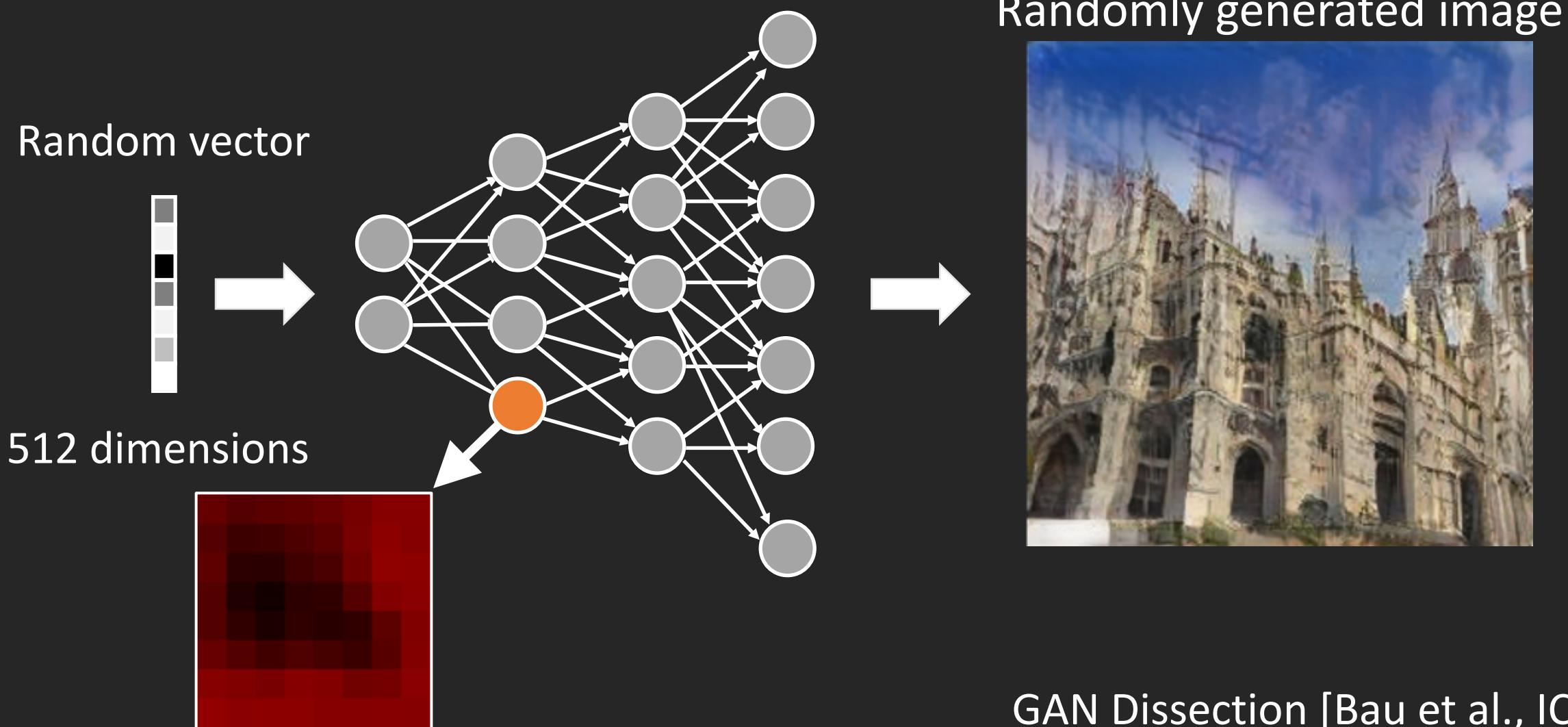
Randomly generated image



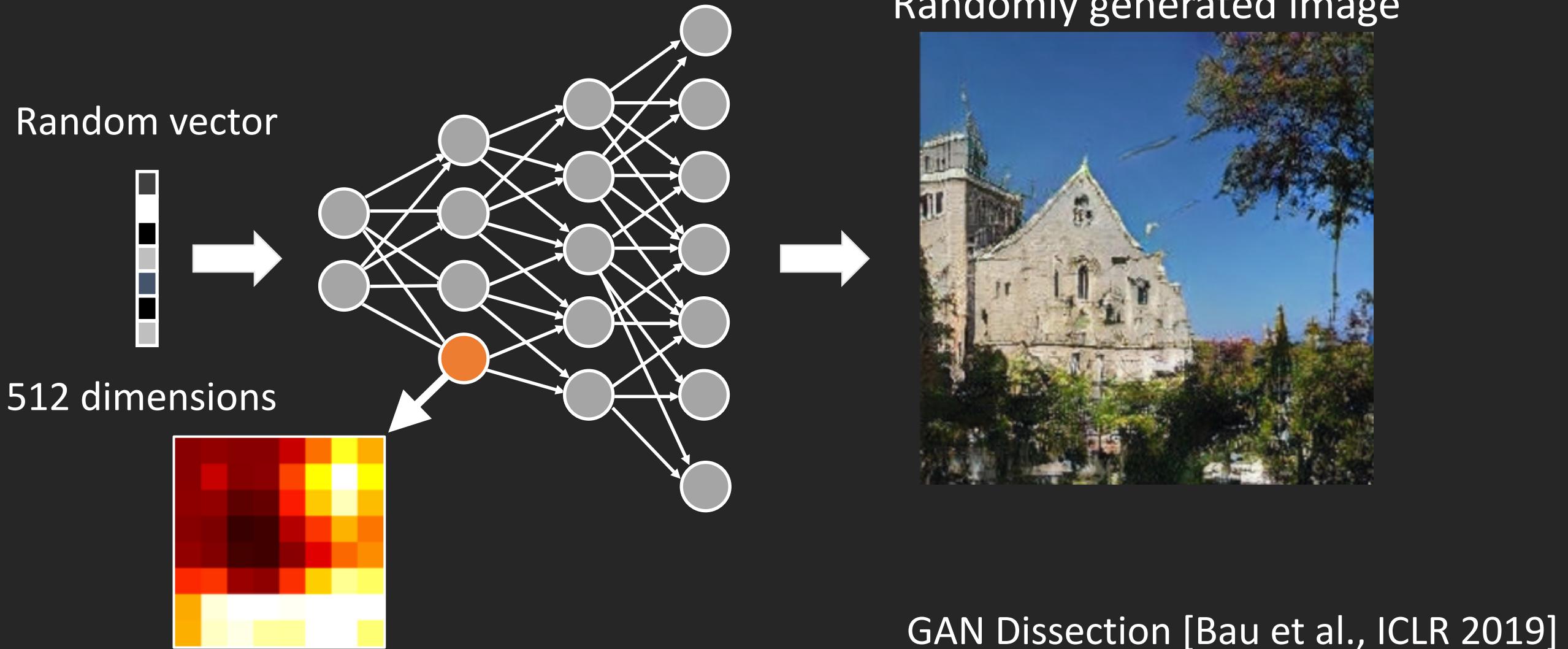
Are there other objects?



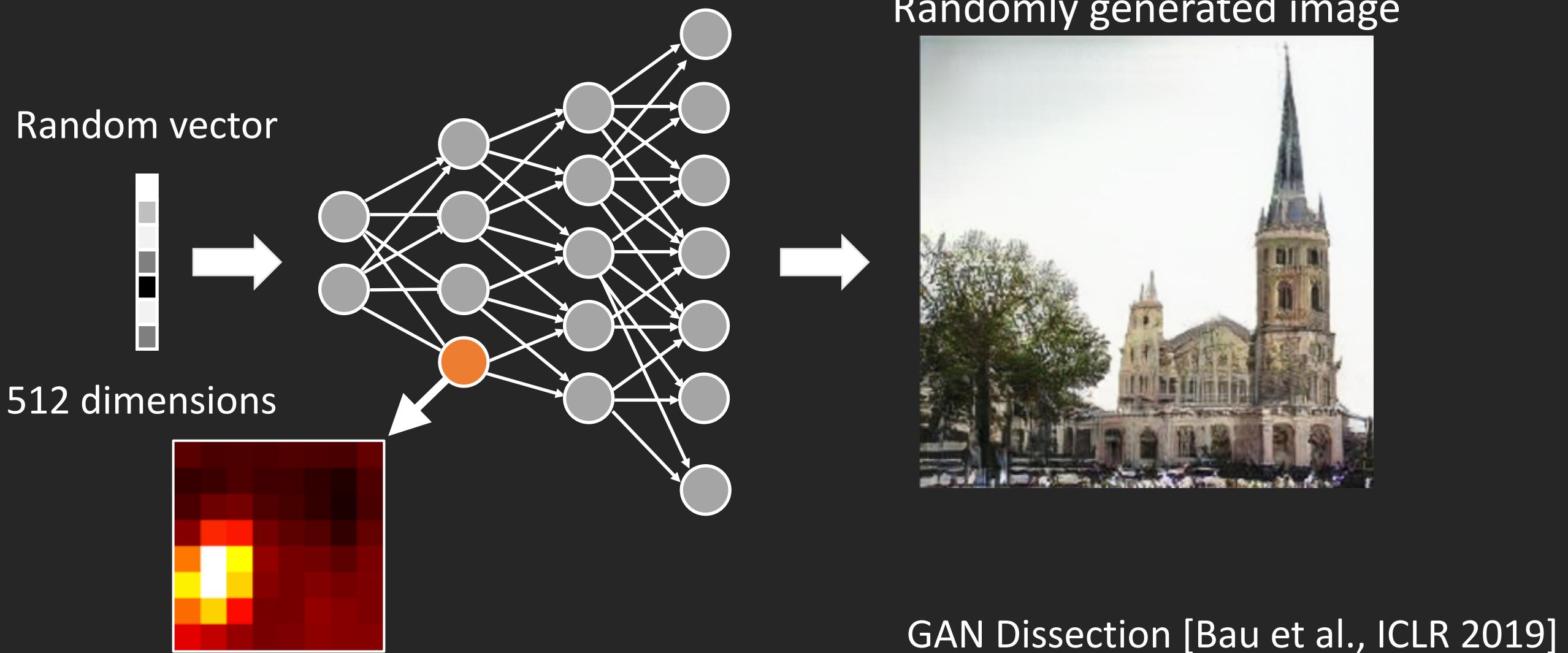
Are there other objects?



Are there other objects?

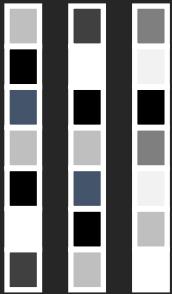


Are there other objects?

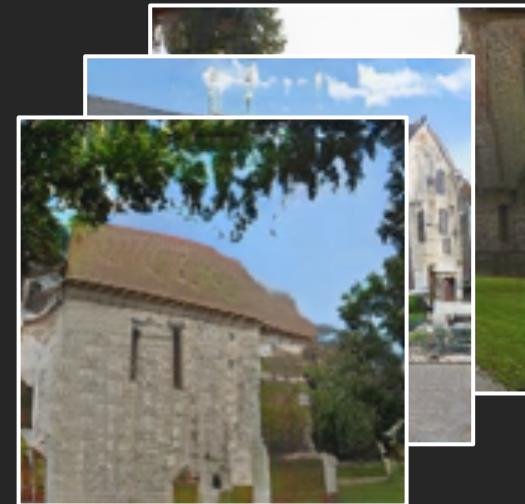


Dissecting a GAN

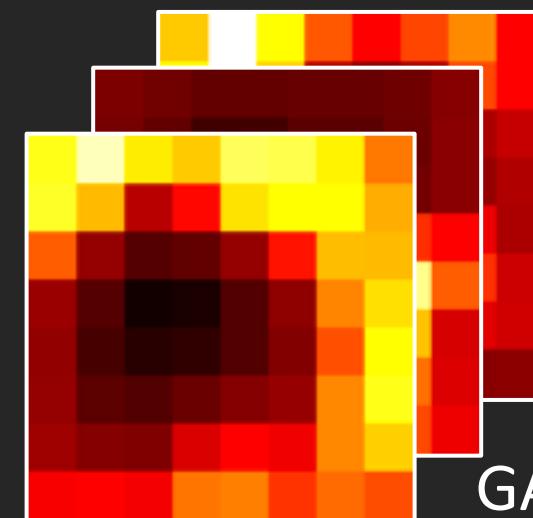
Random vectors



Generate lots of images



Semantic segmentation



Agreement

GAN Dissection [Bau et al., ICLR 2019]

Units Emerge as Drawing Objects

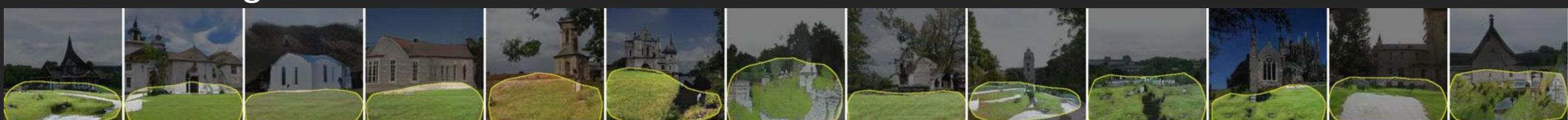
Unit 365 draws trees



Unit 43 draws domes



Unit 14 draws grass



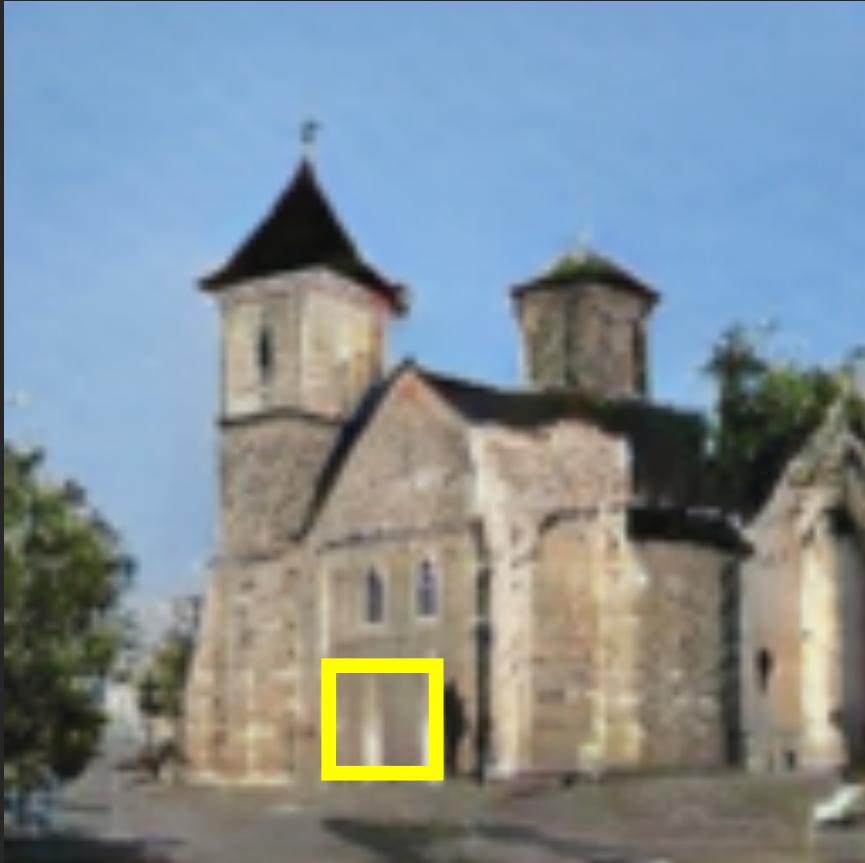
Unit 276 draws towers



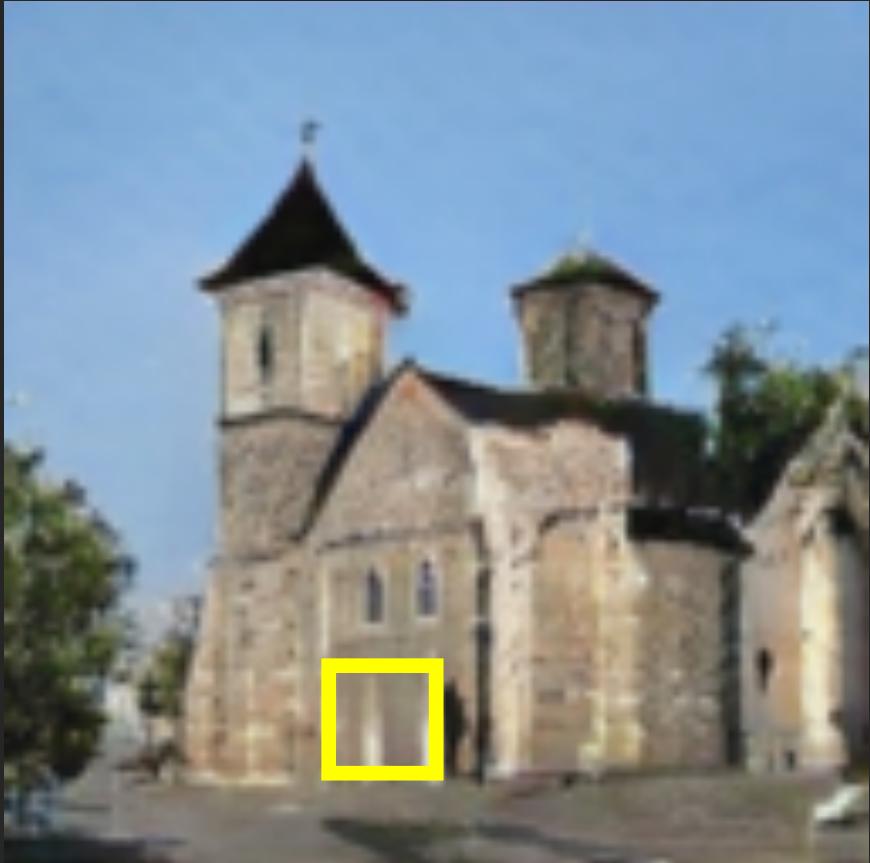
Turning off window units



Turning on door units



Turning on door units



Turning on door units



Turning on door units



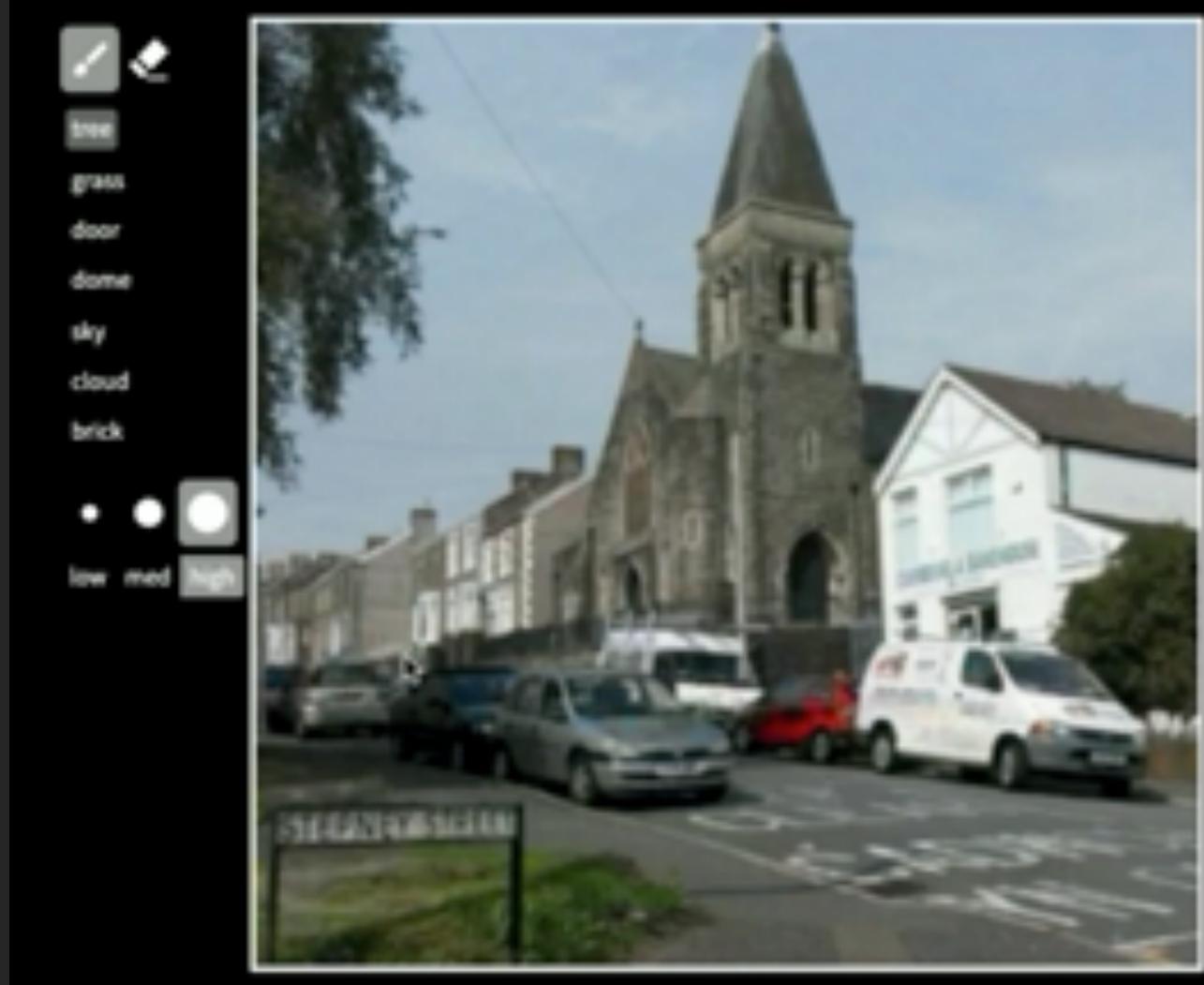
Turning on door units



Turning on door units



Interactive Object Editing



Online demo:
ganpaint.io

Editing Objects vs. Global Attributes



Manipulating Global Attributes

Indoor lighting



Wooden style



Different layout



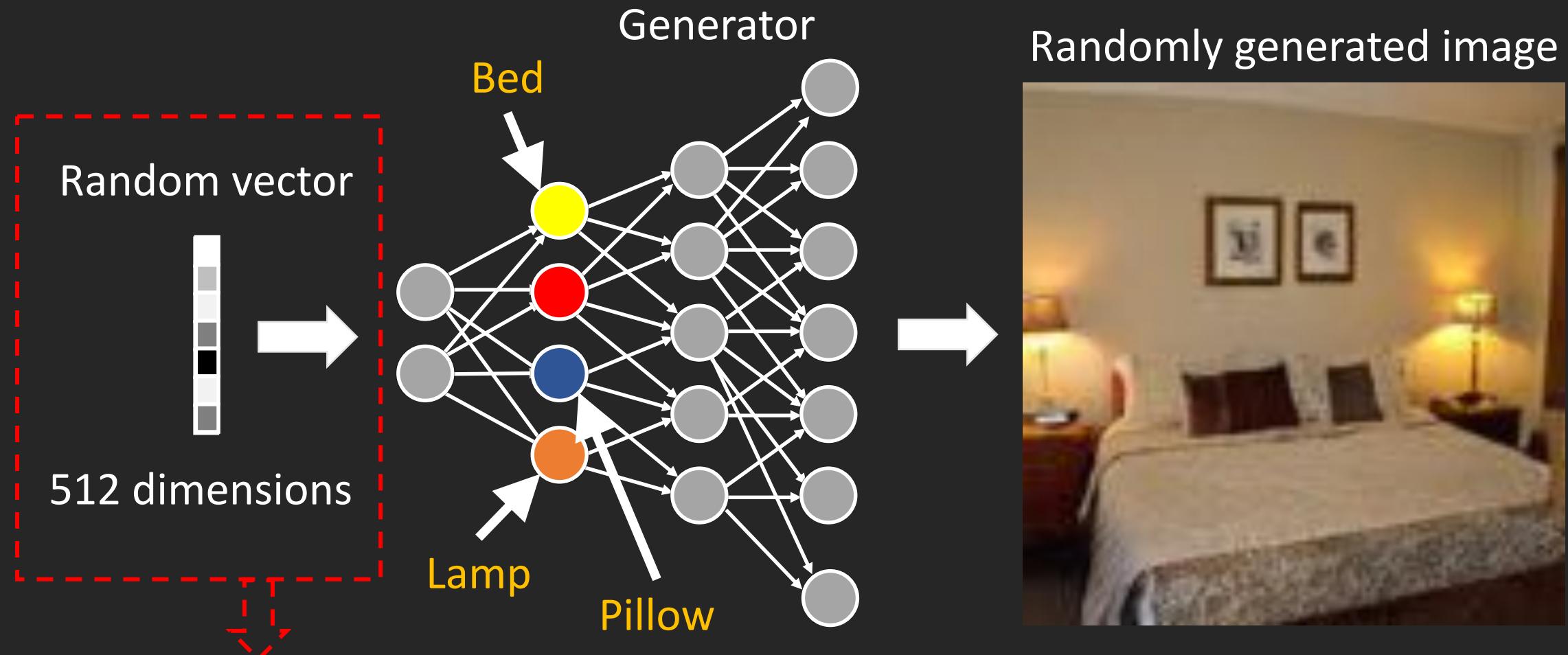
Living room



Tutorial Outline

- Interpretable units in intermediate layers
- Semantics in the latent space
- Inversion of real images
- Future directions

Deep Generative Representation



How latent code affects the output?

Random Walk in Latent Space of Bedroom



Multiple Levels of Abstractions for Scenes

Scene category:

bedroom

Scene attributes:

nature lighting

wood

foliage

...

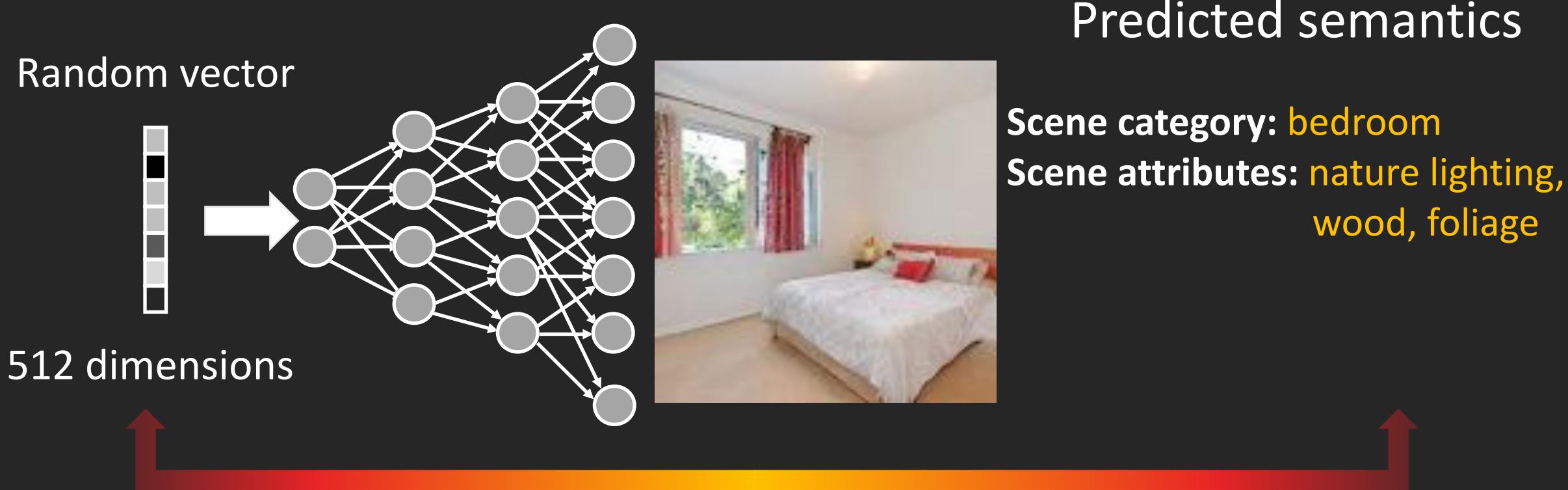
Layout



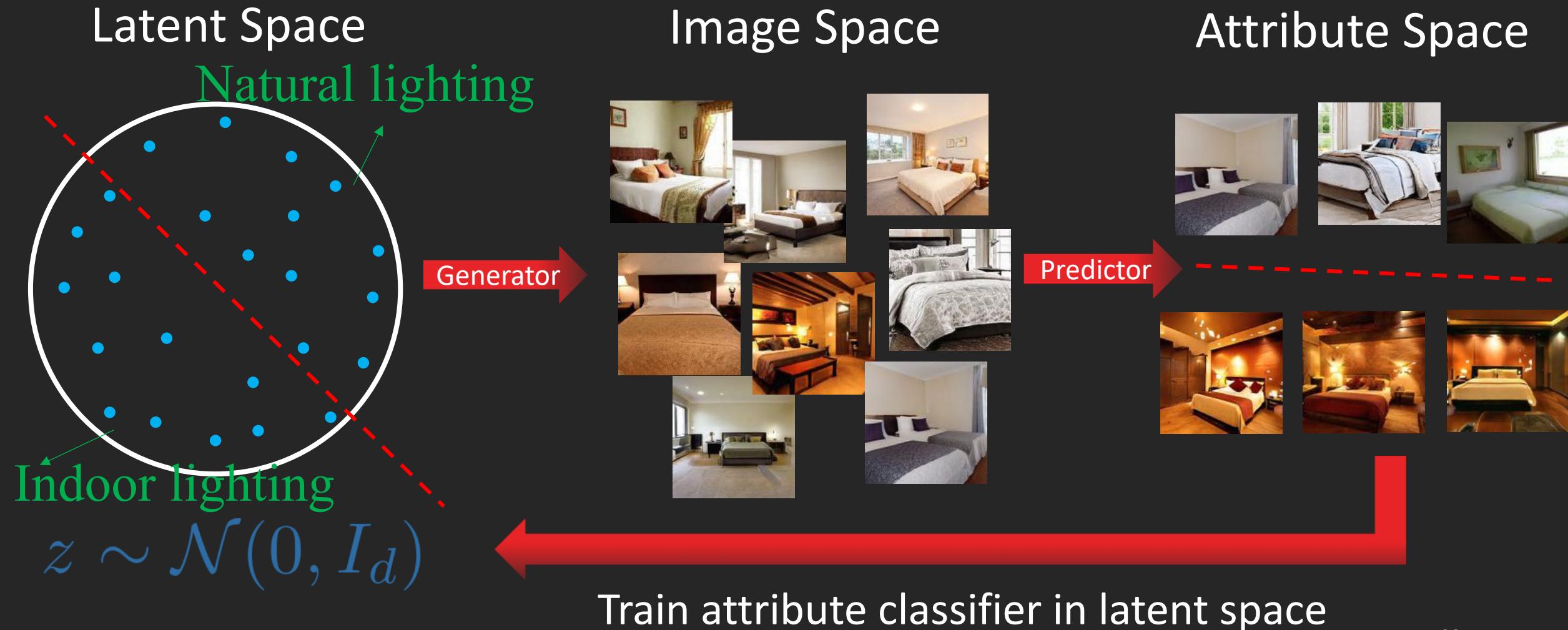
Segmentation



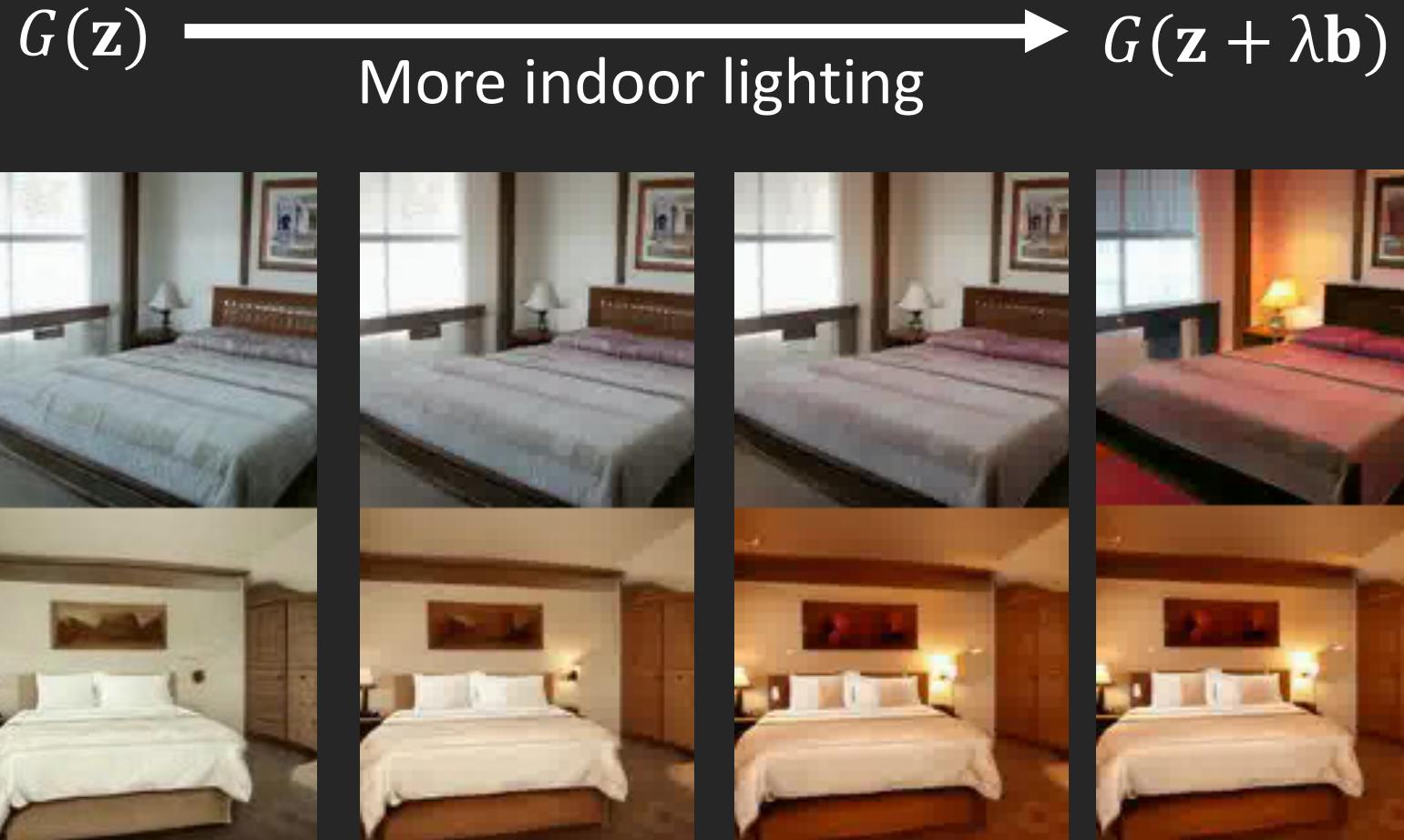
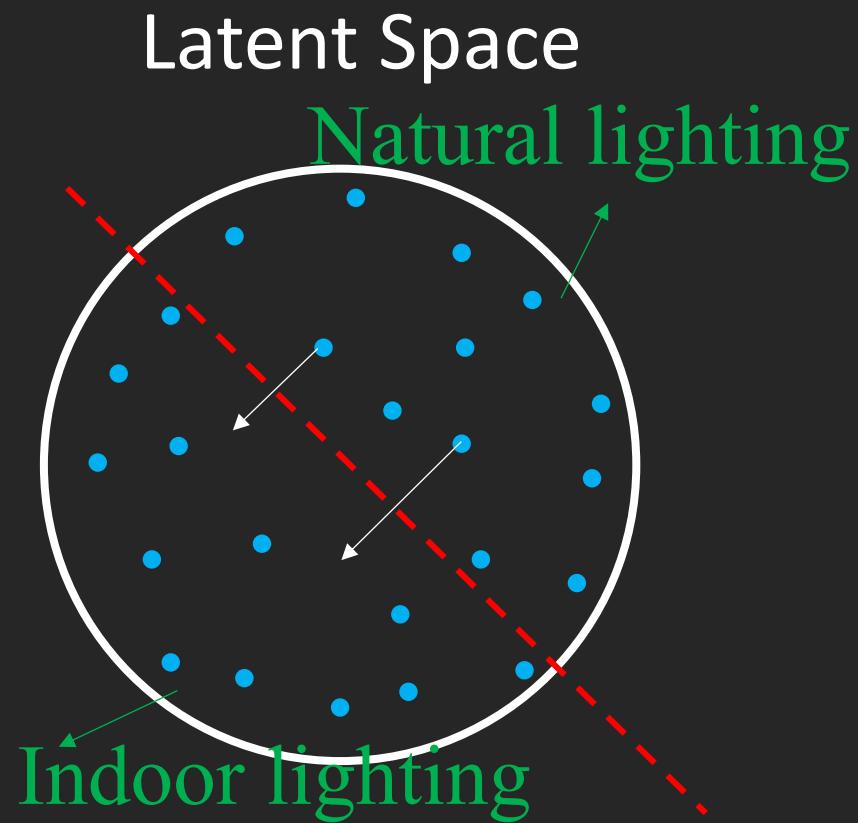
Causal Relations between Latent Code and Semantics



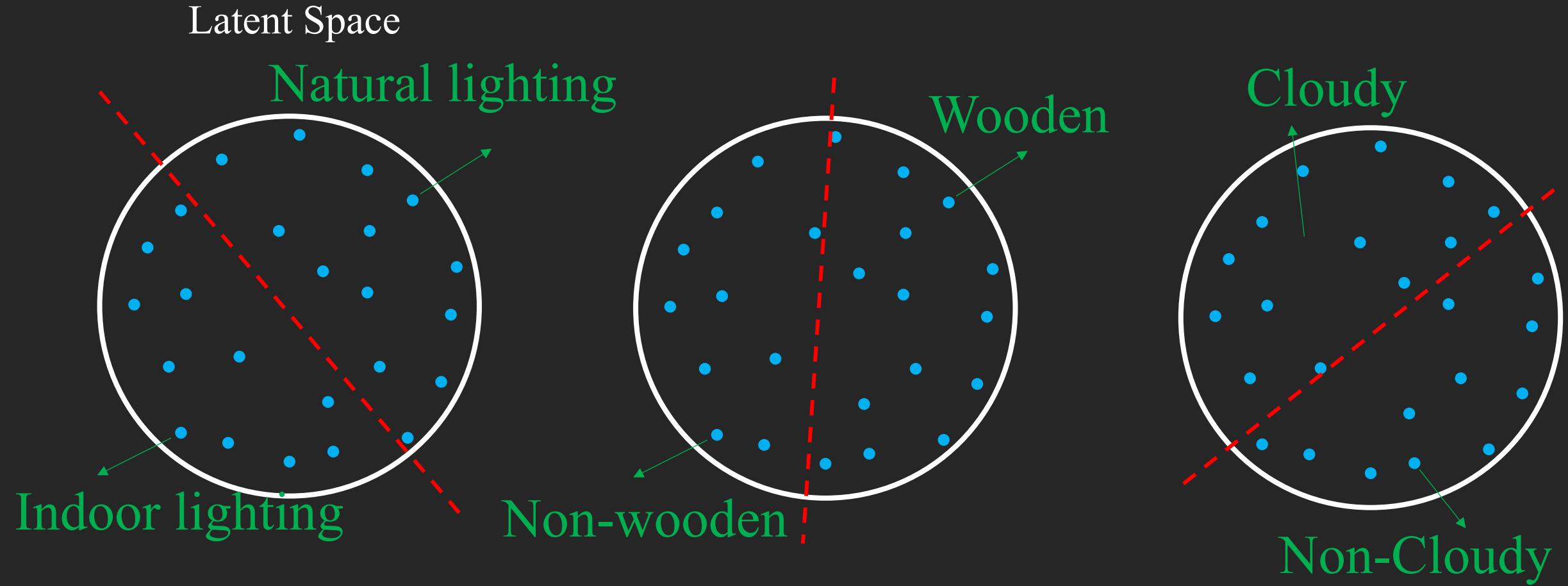
Aligning Latent Space with Attribute Space



Pushing Latent Code through Boundary



Various Attribute Boundaries in Latent Space



Turning up the light



How to edit my own photo?



GAN-Synthesized Kitchen

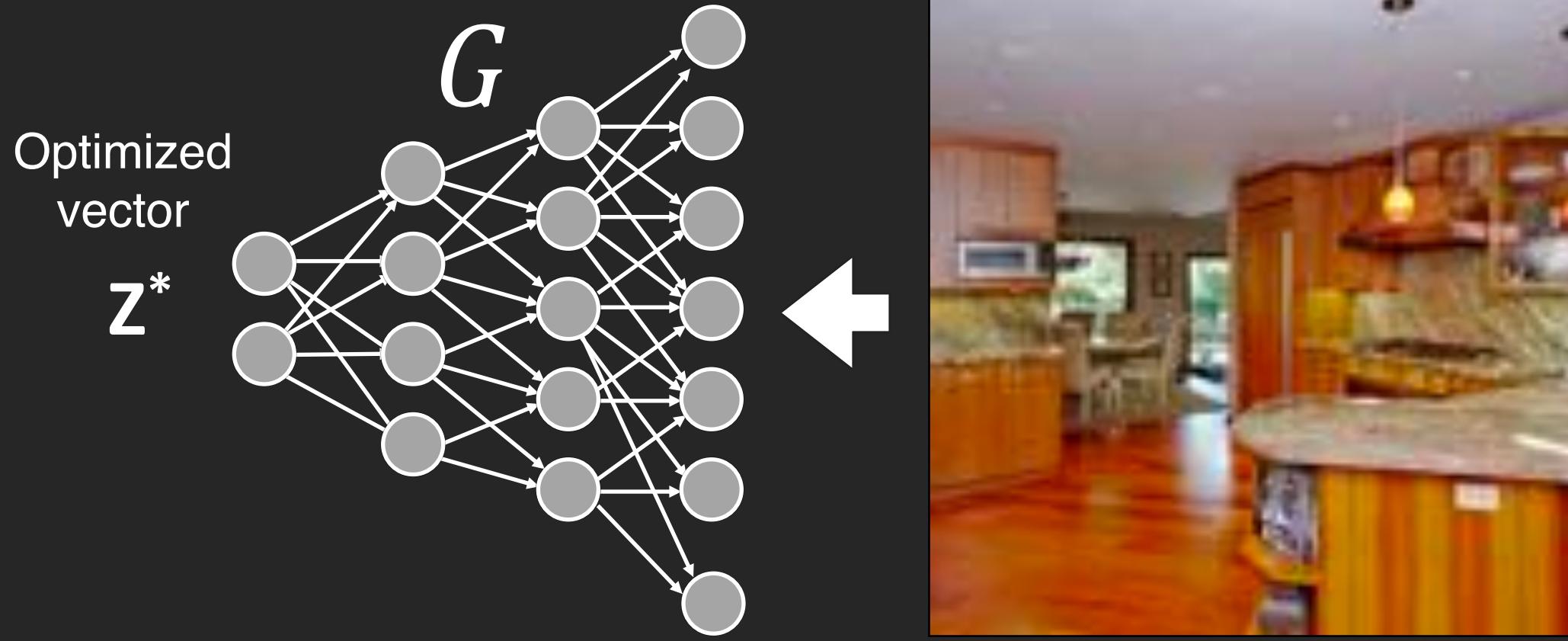


My Kitchen Photo

Tutorial Outline

- Interpretable units in intermediate layers
- Semantics in the latent space
- Inversion of real images
- Future directions

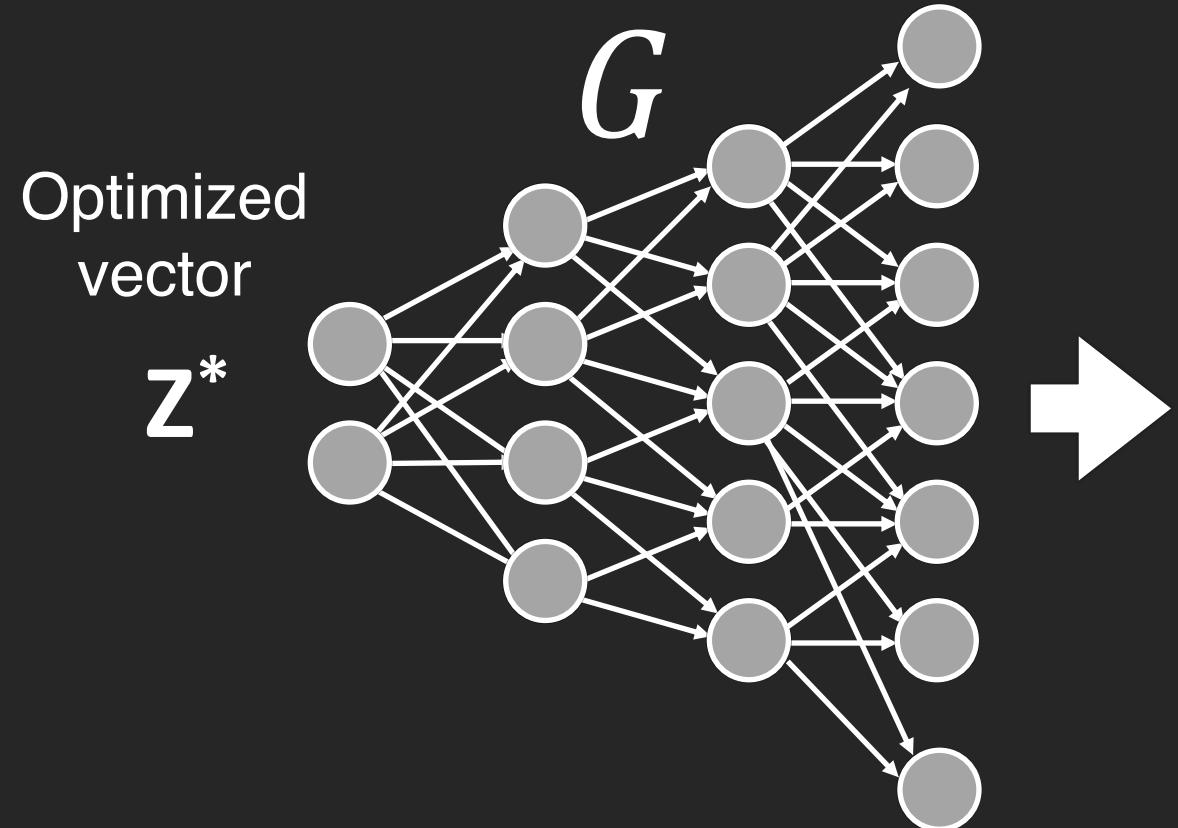
Inverting image to GAN latent code



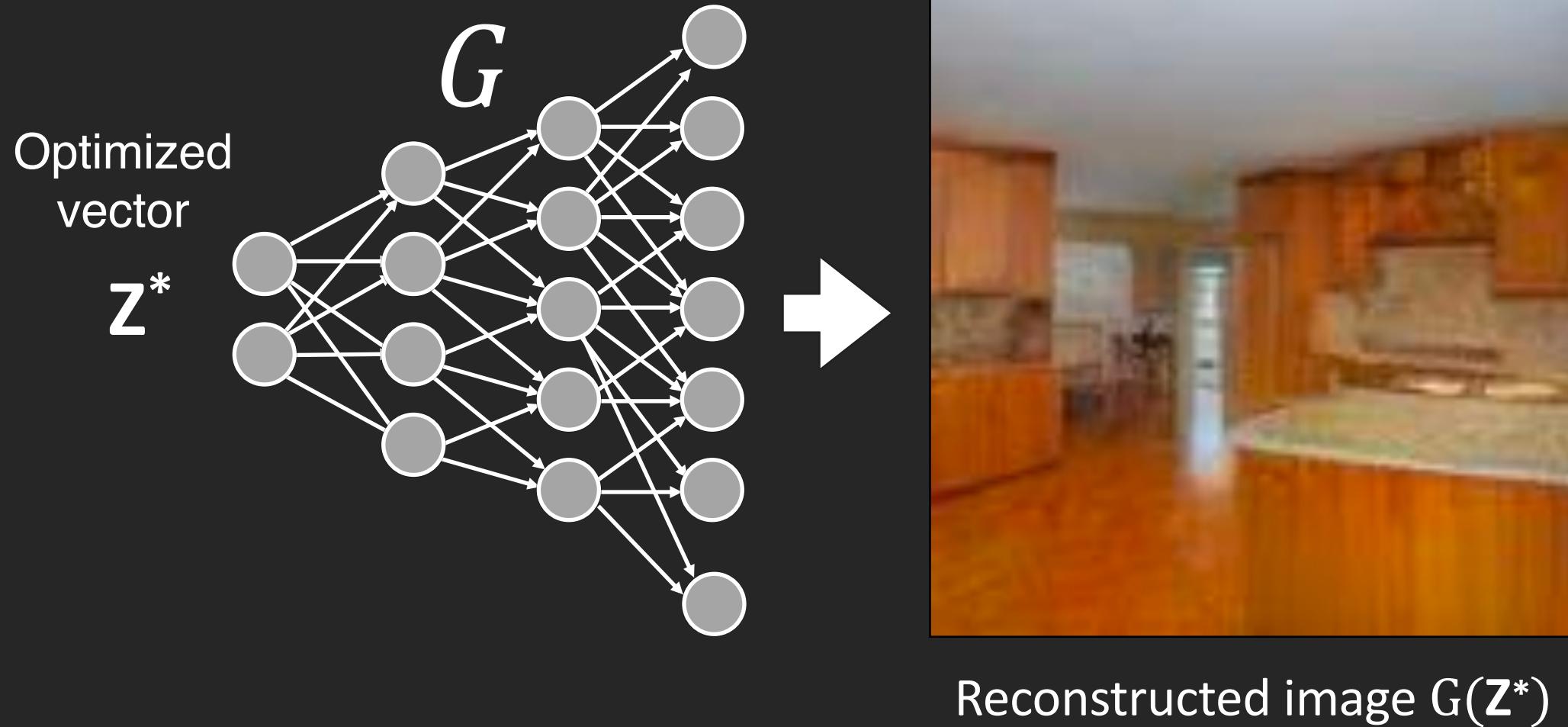
$$\mathbf{z}^* = \arg \min_{\mathbf{z}} \frac{1}{2} \|G(\mathbf{z}) - \mathbf{x}\|^2$$

Real image \mathbf{x}

Inverting image to GAN latent code



Inverting image to GAN latent code



Find the differences...



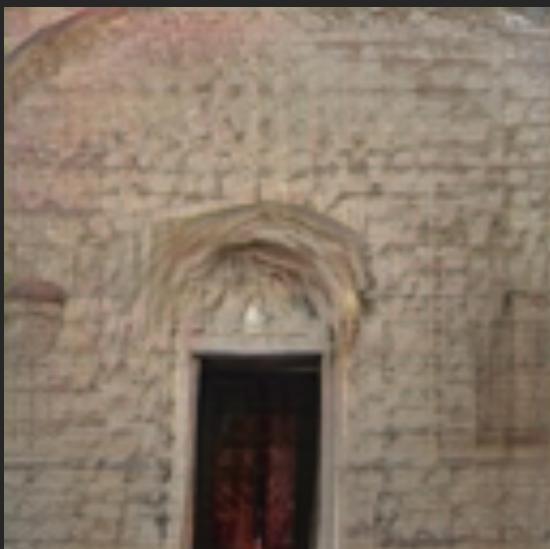
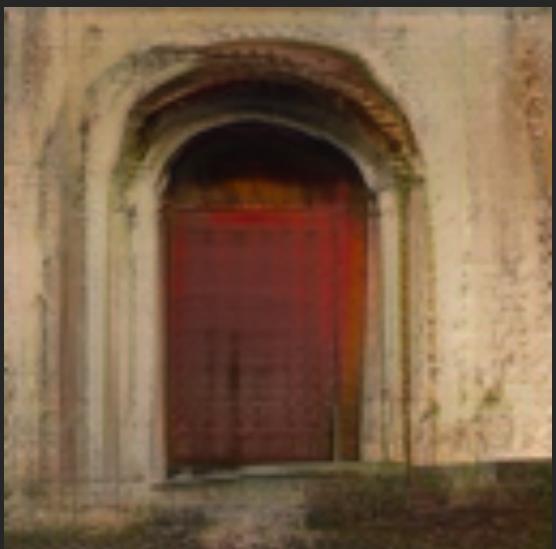
Original image



GAN reconstructed image

Fact:
GANs don't like a lot of things!

GANs don't like people

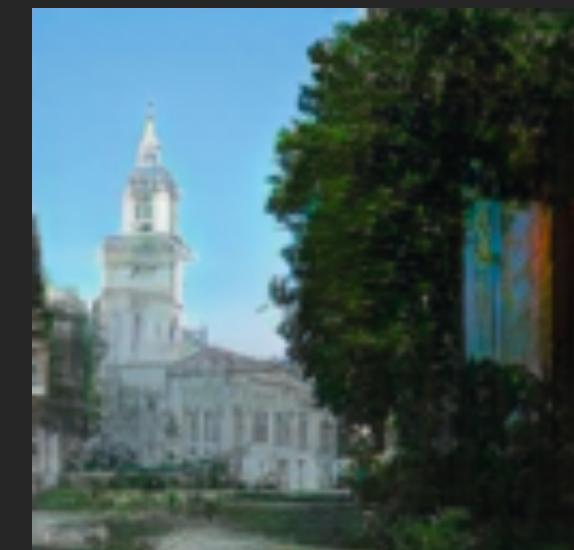


GANs don't like people

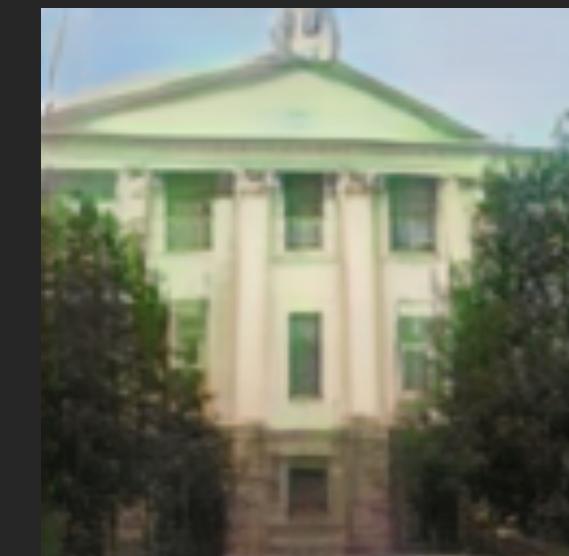
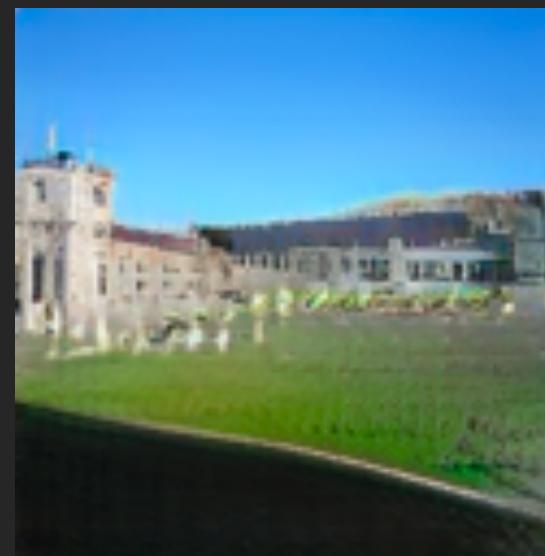
BigGAN's synthesis of Bar



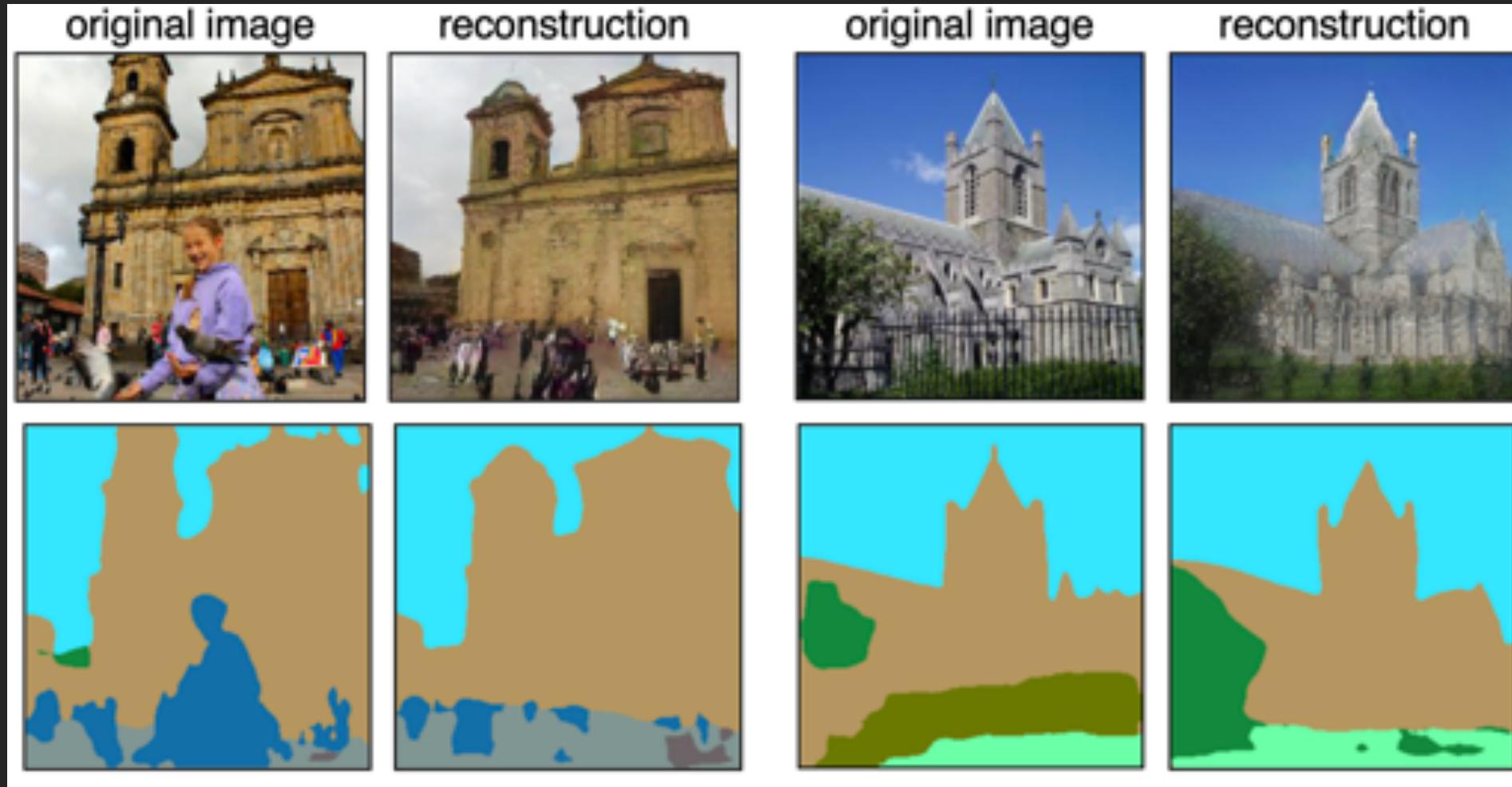
GANs don't like vehicles



GANs don't like signs

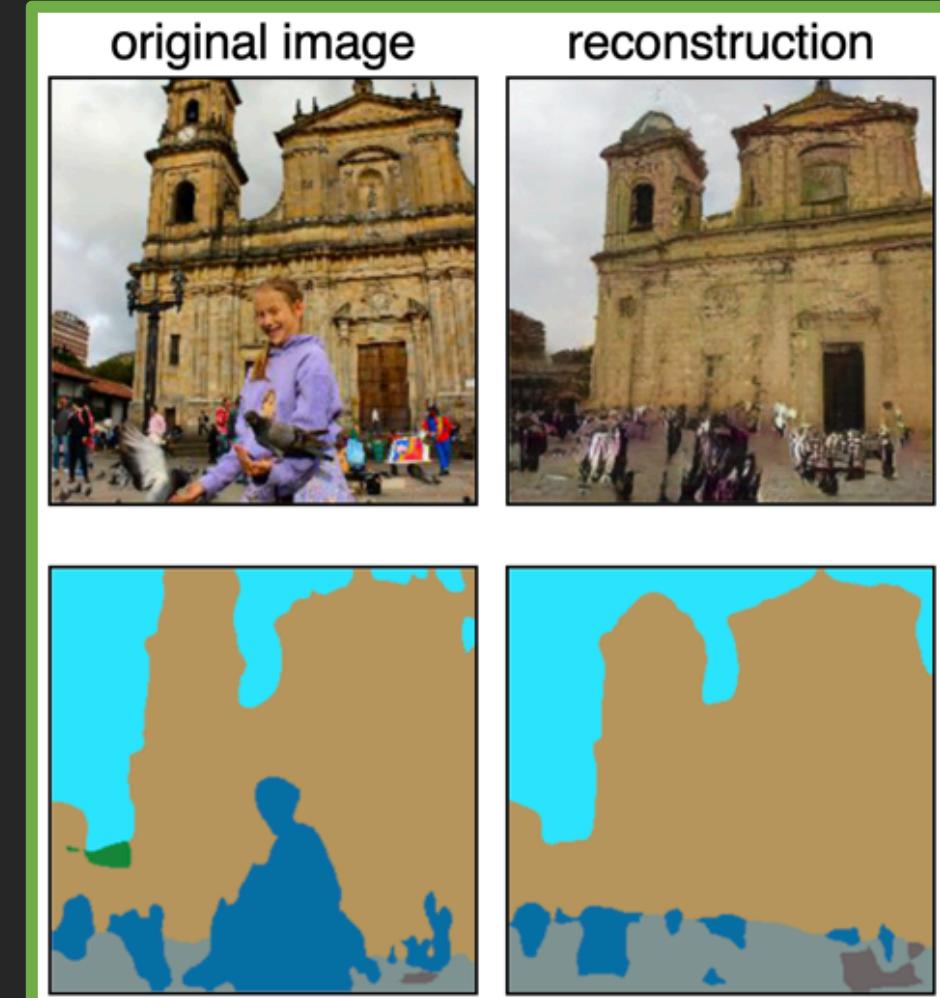
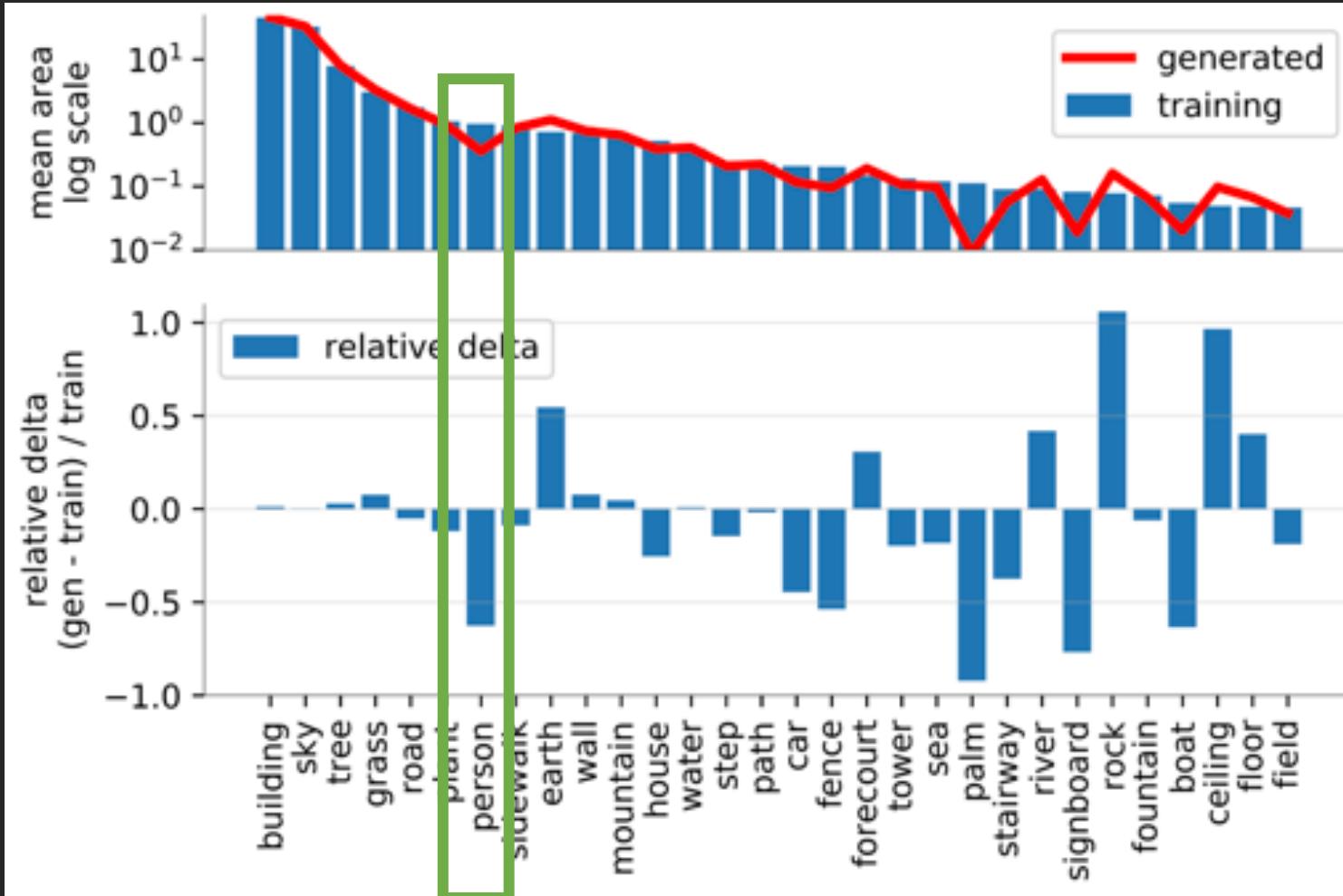


Identify what's missing by comparing masks



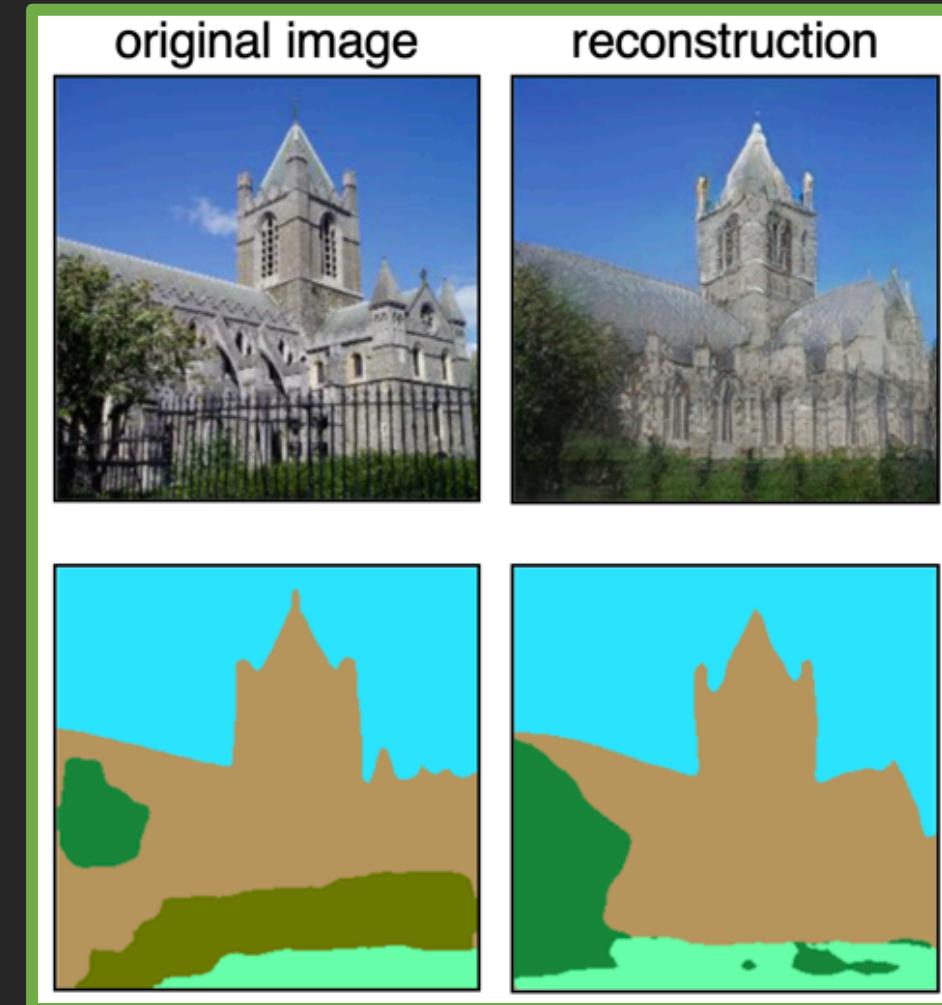
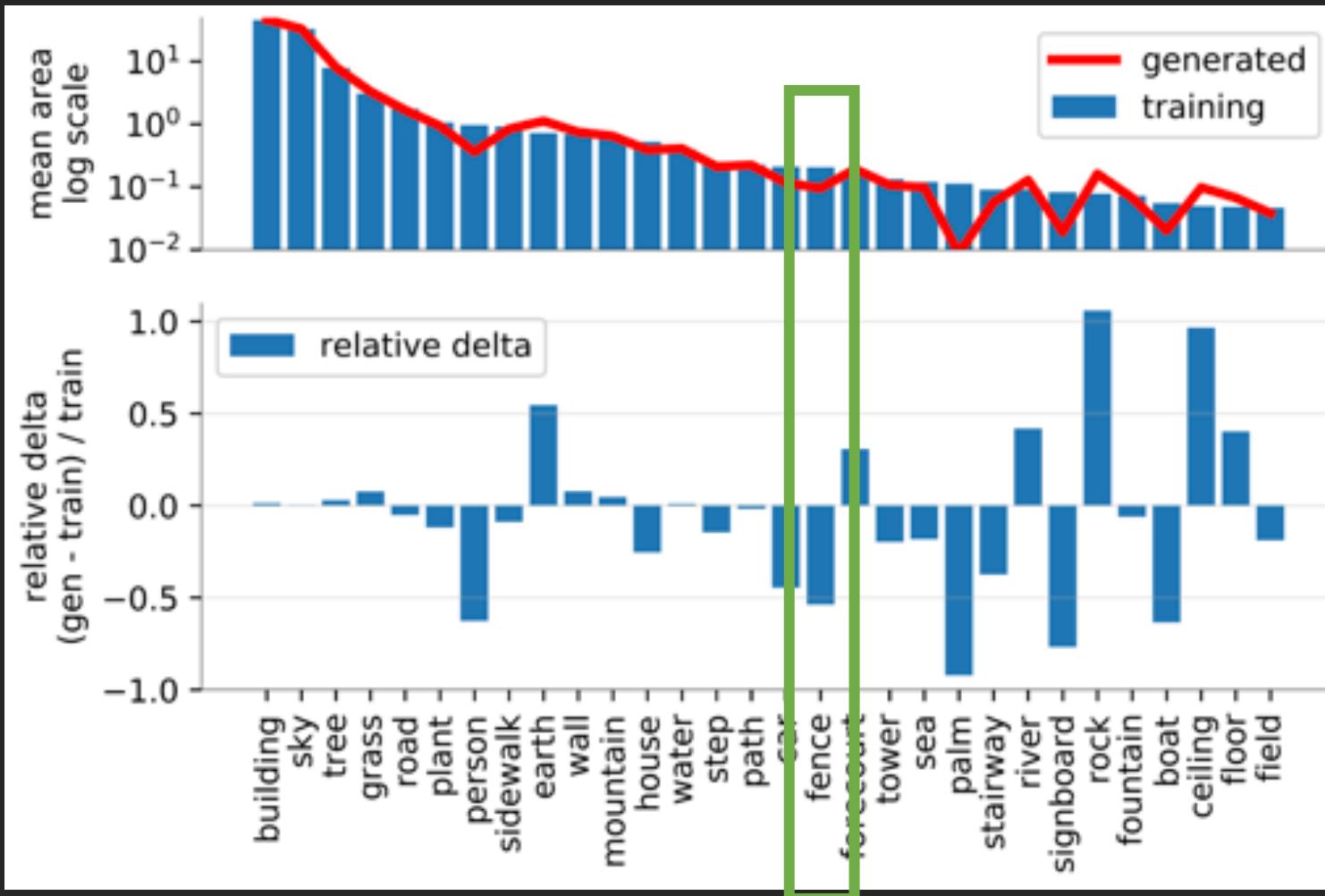
Thursday, October 31, 0900–1030 Oral 3.1A

Statistics tells what GANs cannot generate



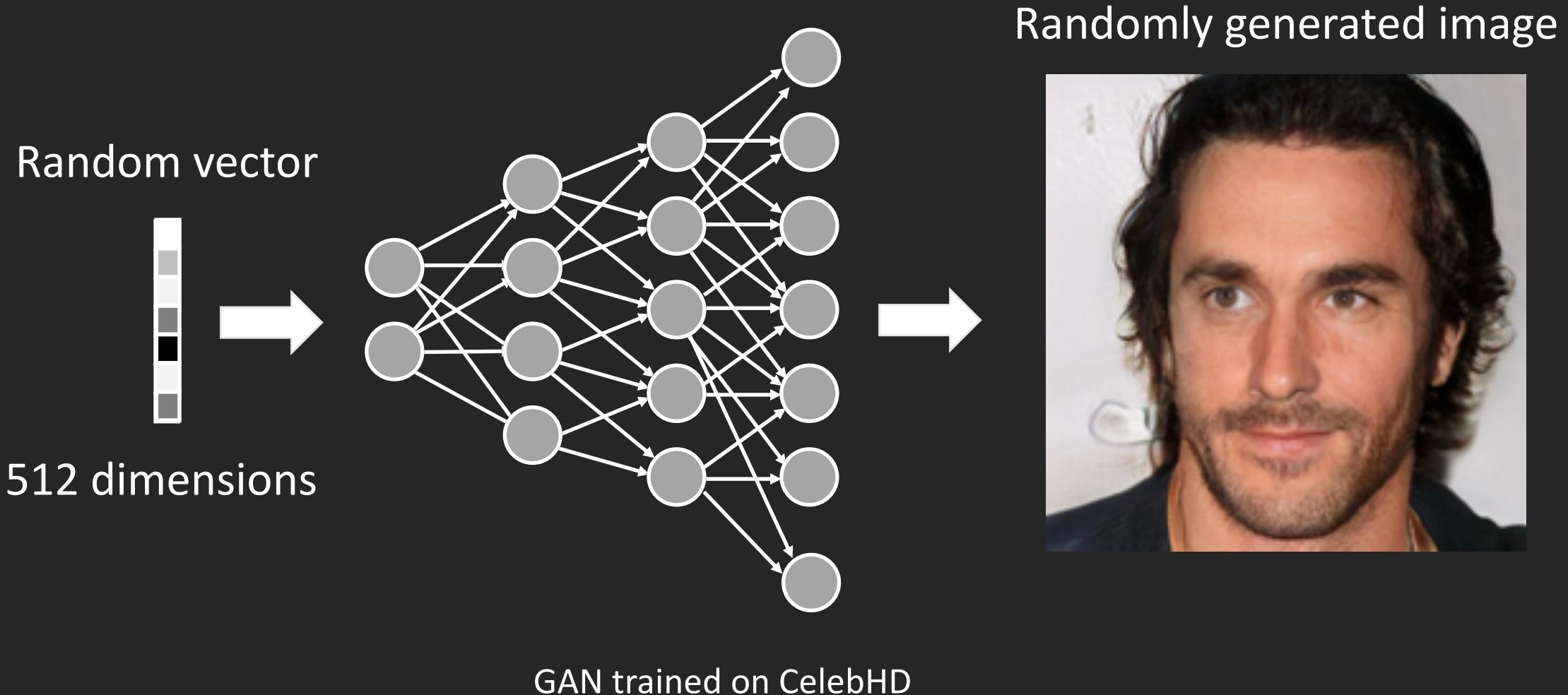
Thursday, October 31, 0900–1030 Oral 3.1A

Statistics tells what GANs cannot generate

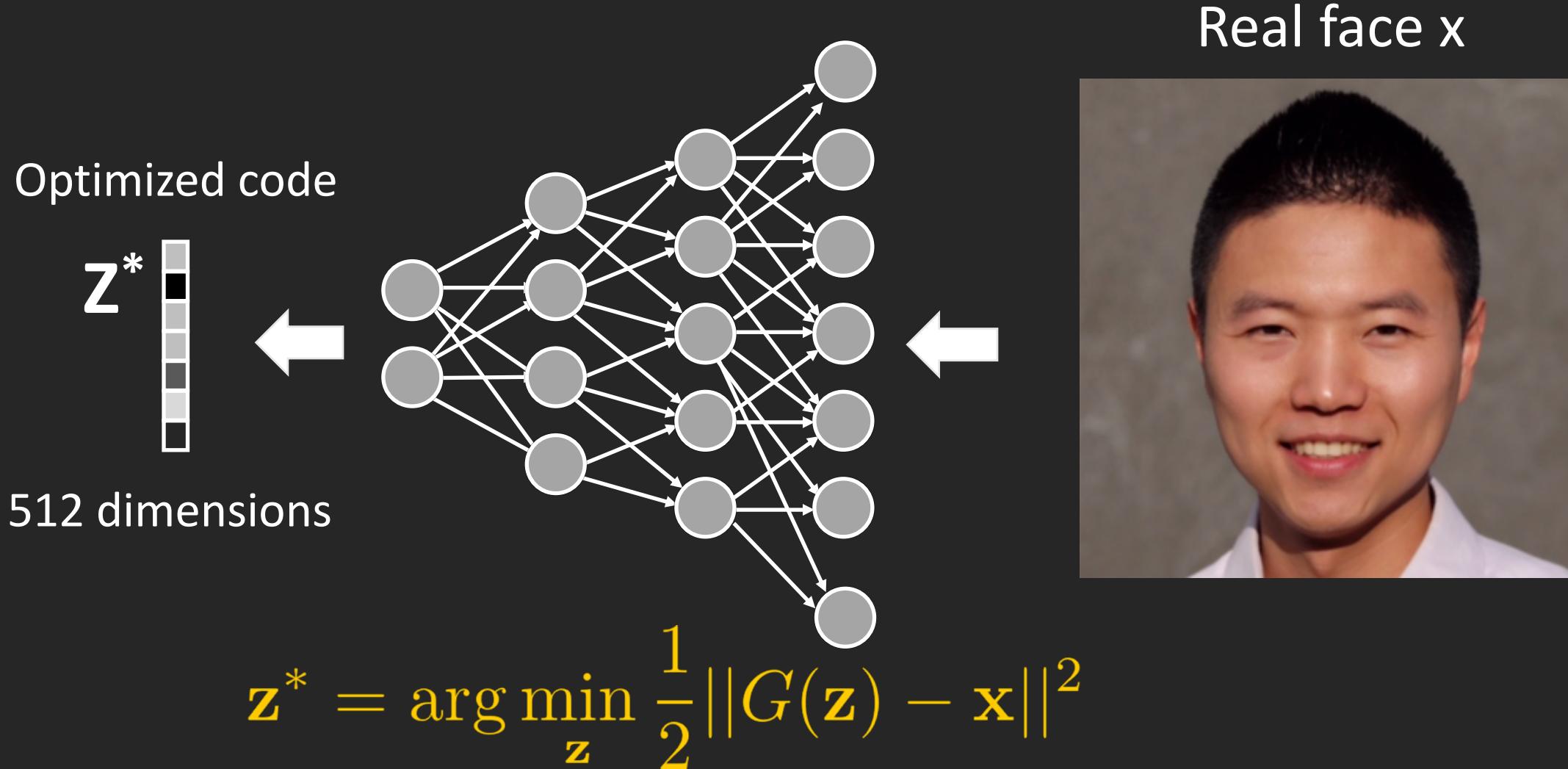


Thursday, October 31, 0900–1030 Oral 3.1A

GAN Inversion for Faces



GAN Inversion for Faces

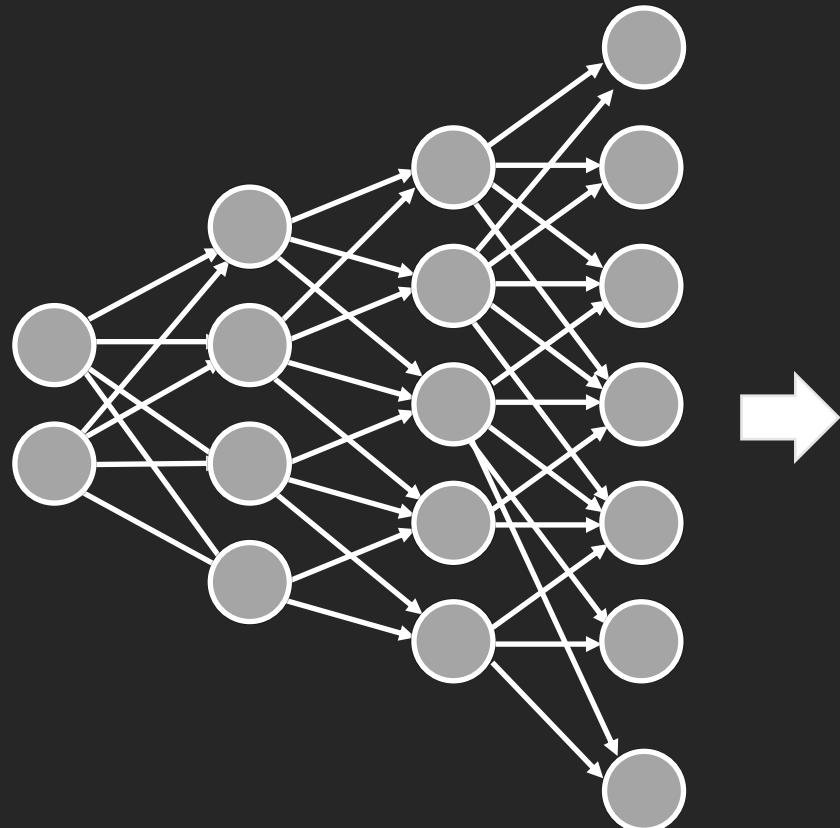


GAN Inversion for Faces

Optimized code



512 dimensions



Reconstructed face



Bias in GANs

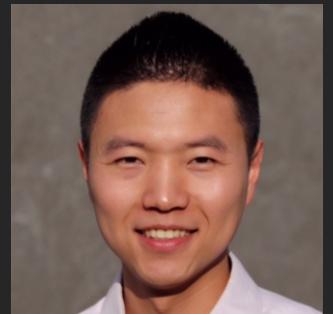
Real face



Reconstructed face

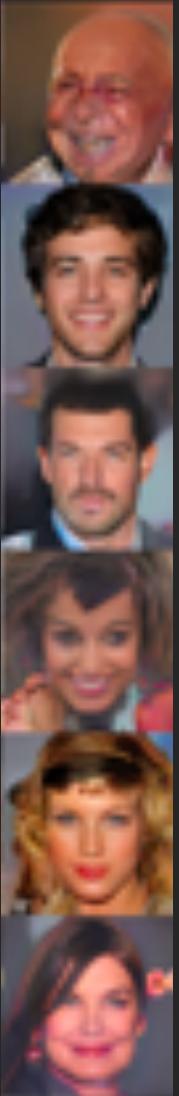


Different random initializations

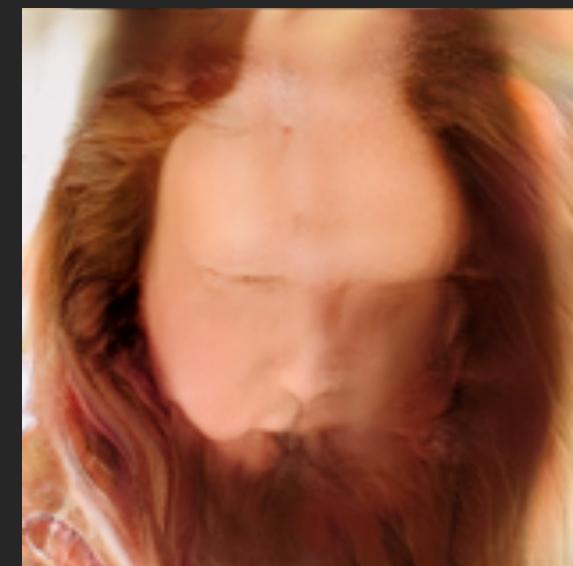
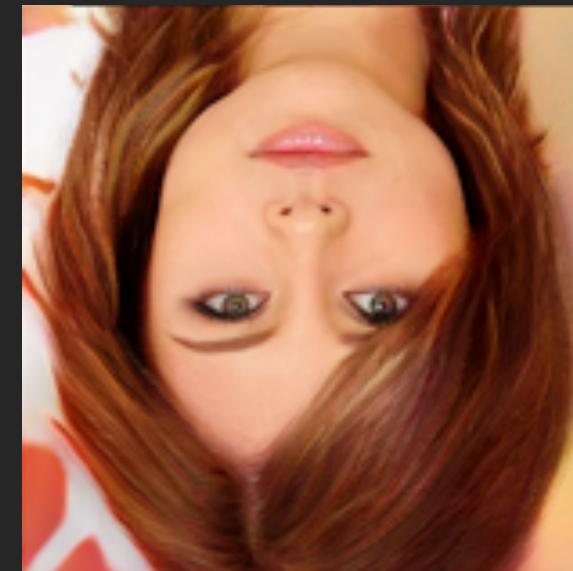


Initializations

Iteration 1 -> 500



GAN Inversion of Non-face Images

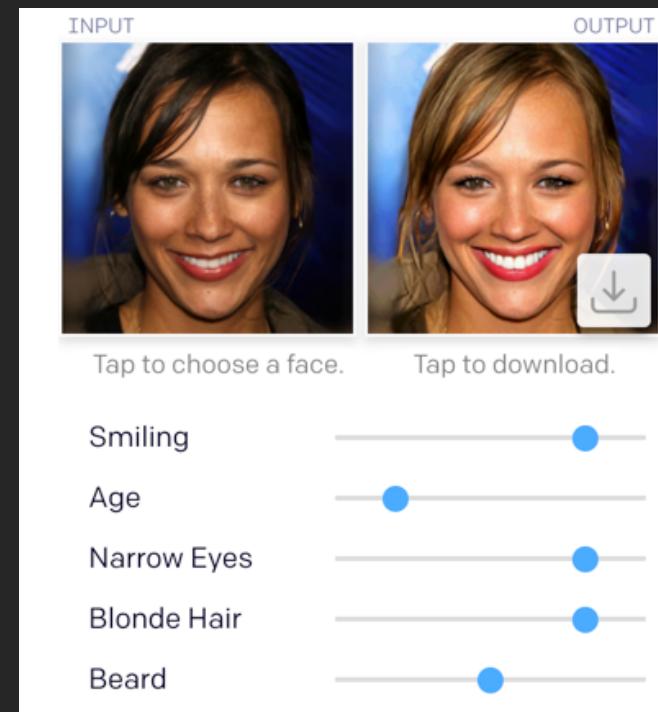
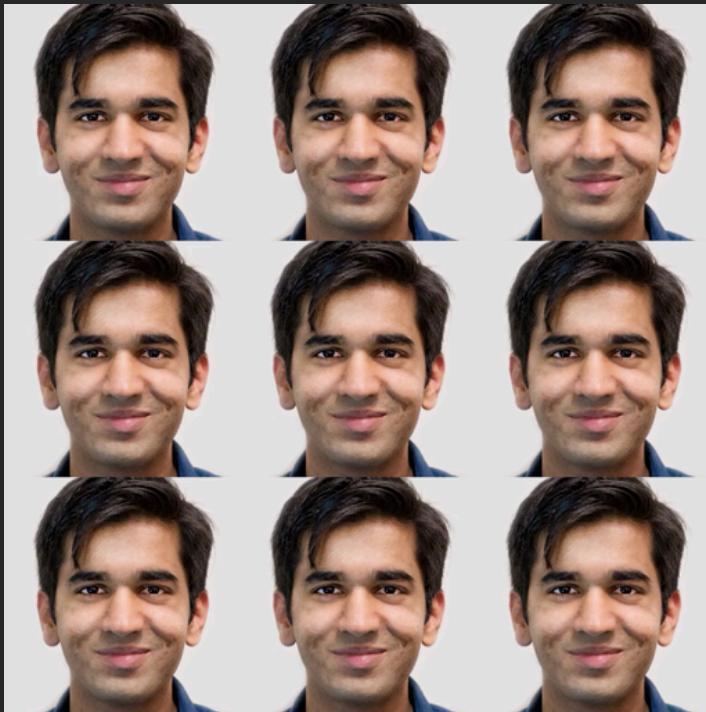


Tutorial Outline

- Interpretable units in intermediate layers
- Semantics in the latent space
- Inversion of real images
- Future directions

Invertible Generative Models

- Explicitly design reversible generative models
- Nice (Dinh et al, ICLR'15), RealNVP (Dinh et al, ICLR'17), Glow (Kingma et al, NIPS'18)

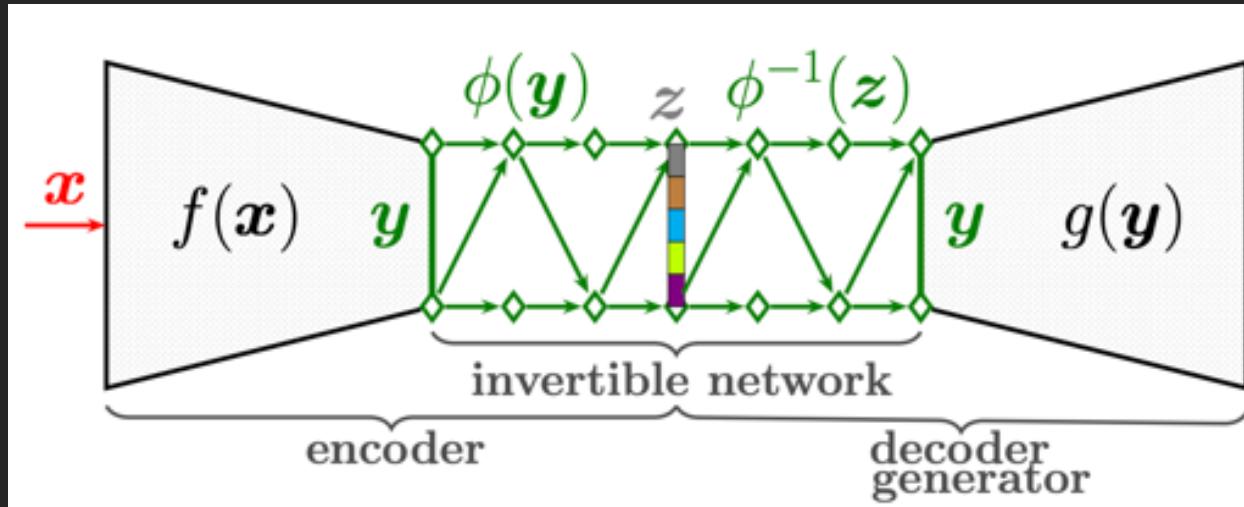


<https://arxiv.org/pdf/1410.8516.pdf>

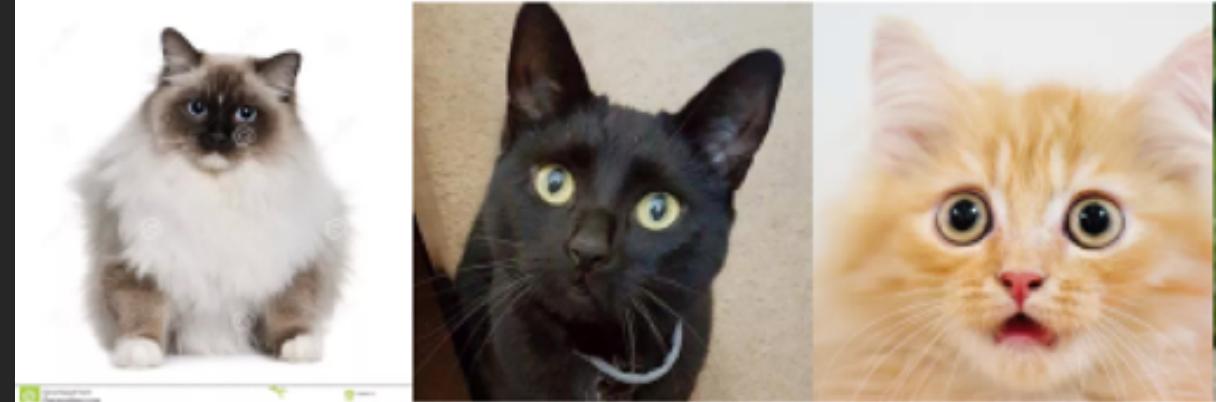
<https://arxiv.org/pdf/1807.03039.pdf>

<https://openai.com/blog/glow/>

Invertible Network + VAEs



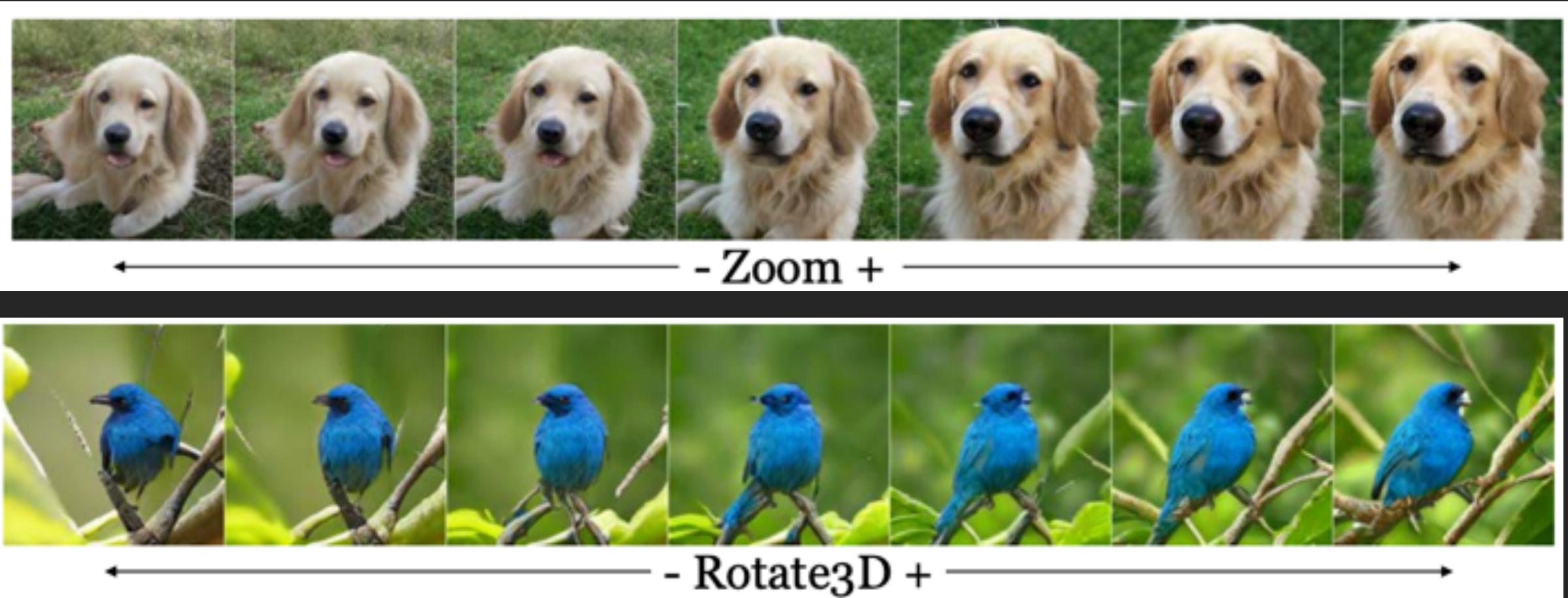
Input



Reconstruction



Steerability of GANs



Facial Attributes Disentangled in Latent Space

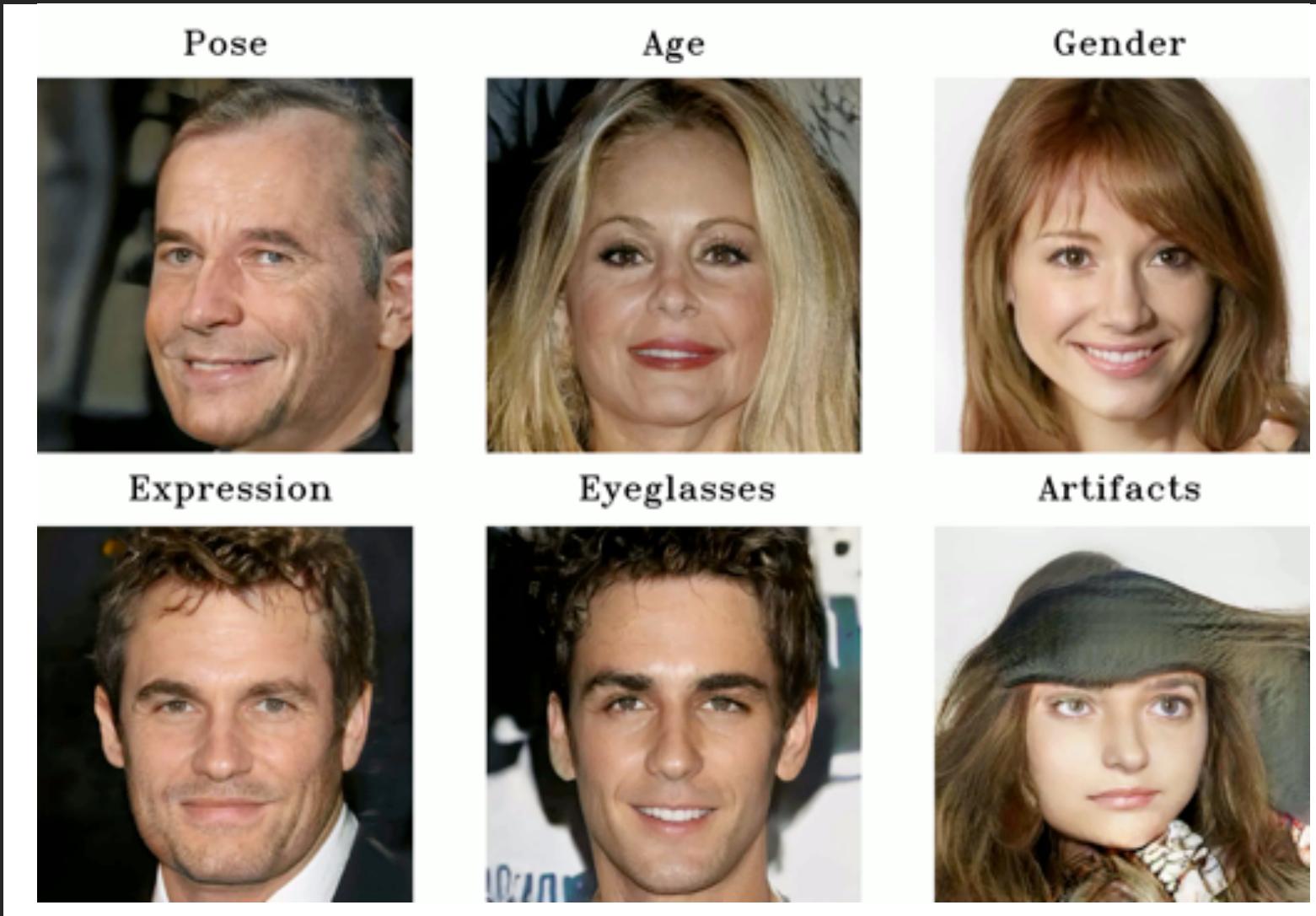


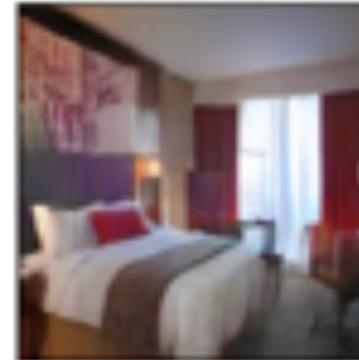
Image Processing with GAN Prior



(a) GAN inversion result



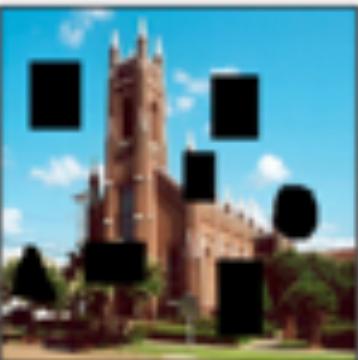
(b) Colorization



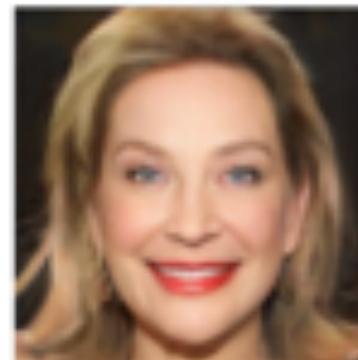
(c) Super-Resolution



(d) Image reconstruction



(e) Image inpainting



(f) Semantic image manipulation

Acknowledgement

MIT team

Key player

David Bau



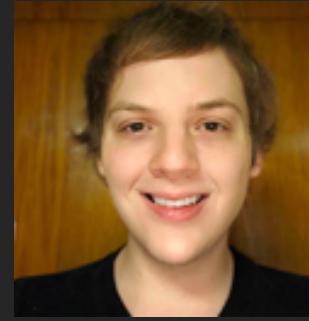
Jun-Yan Zhu



Hendrik Strobelt



Bill Peebles

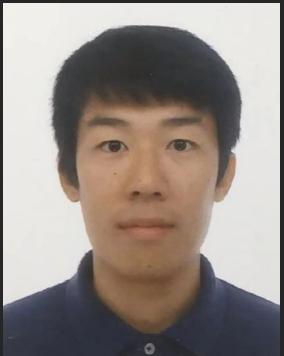


Antonio Torralba



CUHK team

Yujun Shen



Ceyuan Yang



Jinjin Gu



Conclusion

- Latent semantics emerge in GANs
- Code and papers at <http://bzhou.ie.cuhk.edu.hk>

