PaintMe

Neural Texture Synthesis via Joint Multi-View Stochastic Diffusion Process

Seungwoo Yoo

TA: Juil Koo

Advisor: Minhyuk Sung

Geometric AI Lab, School of Computing

Submitted to

SIGGRAPH ASIA 2022

PaintMe: Neural Texture Synthesis via Joint Multi-View Stochastic Diffusion Process

Seungwoo Yoo
dreamy1534@kaist.ac.kr
KAIST
Daejeon, Korea, Republic of

Minhyuk Sung
mhsung@kaist.ac.kr
KAIST
Daejeon, Korea, Republic of

Figure 1: Our PAINTME generates realistic textures of 3D models via projective texture mapping. Starting from images rendered from the 3D model (the leftmost), our method leveraging a diffusion-based image generative model converts the images into plausible projective textures over the iterations of the diffusion process. Our joint process for multiple views also enables producing consistent images across the views. See the supplementary webpage for more results.

ABSTRACT

We propose PAINTME, a framework for creating realistic textures of 3D models using a diffusion-based image generative model and projective texture mapping. High-quality texture images play a crucial role in giving a photorealistic appearance to a 3D model. Previous work proposed to leverage SotA image GANs to produce high-quality textures, although it leads to several issues, such as 1) requiring architecture change and retraining due to the conditional setup, and 2) enforcing the silhouette constraint in the projective mapping while coping with such a *hard constraint* is not trivial in a conditional GAN. As a solution to avoiding these issues, we propose to leverage the recent diffusion-based generative model, which is not only known for outperforming SotA GANs but also has significant advantages. We observe that a diffusion-based model

challenge is the consistency across the views in the projective mapping. Our key contribution is a joint reverse process combining intermediate images from different views and guiding to align the per-view images. Our experimental results using a CompCars and ShapeNet dataset demonstrate significant outperformance of our method compared with the SotA GAN-based method: 10.45, 1.45, 50.4% gaps in FID, GIQA, and the user preference, respectively.

CCS CONCEPTS

• **Computing methodologies → Texturing; Neural networks.**

KEYWORDS

texturing, stochastic diffusion process

Motivation

Demands on 3D content are rapidly increasing.

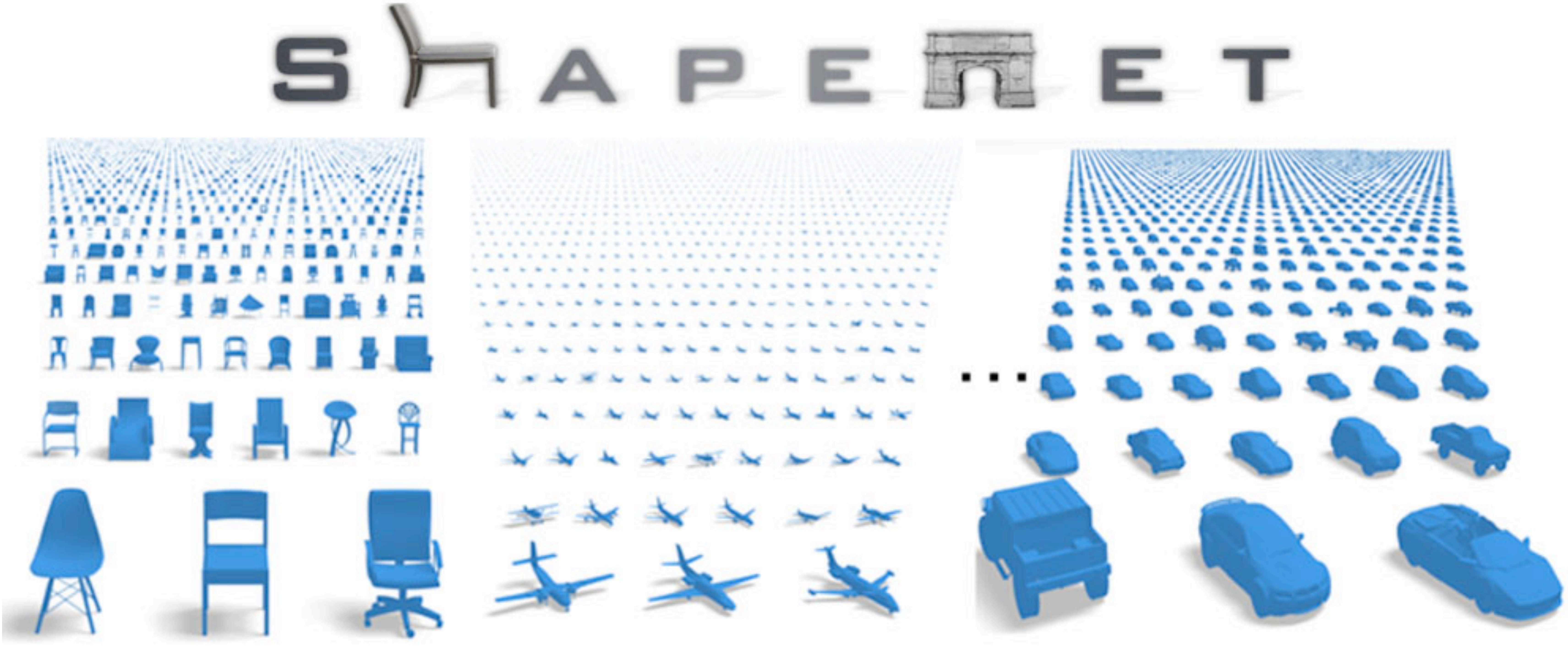
[1] https://blog.google/products/maps/three-maps-updates-io-2022/
[2] https://www.apple.com/kr/augmented-reality/

Motivation

Tools are hard to learn; Annotation is laborious.

[1] https://store.substance3d.com/blog/substance-painter-22-here

Motivation



3D objects in online repositories generally lack color information.

[1] https://cseweb.ucsd.edu/~shz338/research.html

Goal

A textureless 3D object

Deep-learning-based pipeline →

A 3D object with texture

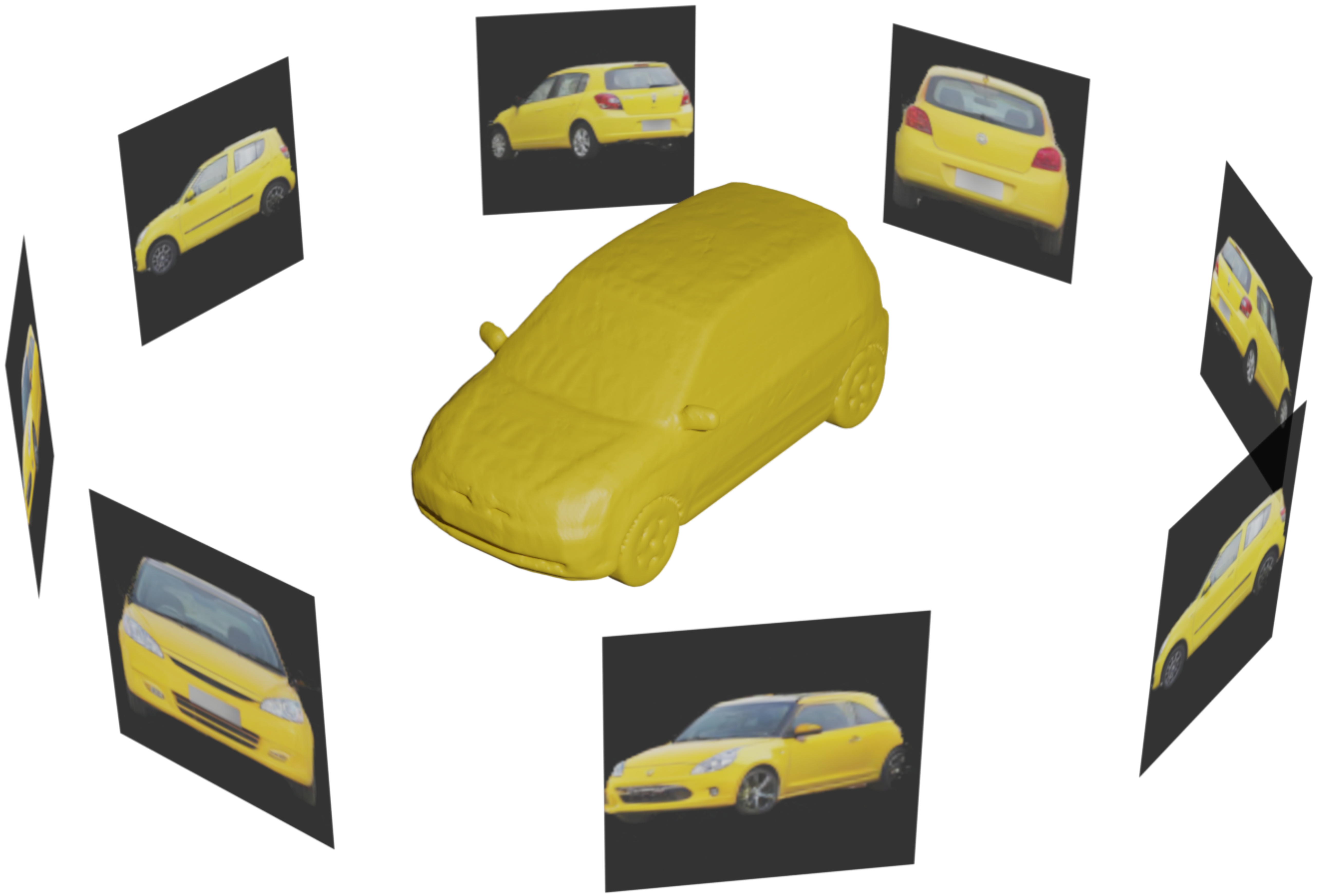
Challenge

Easy to collect 2D photos.

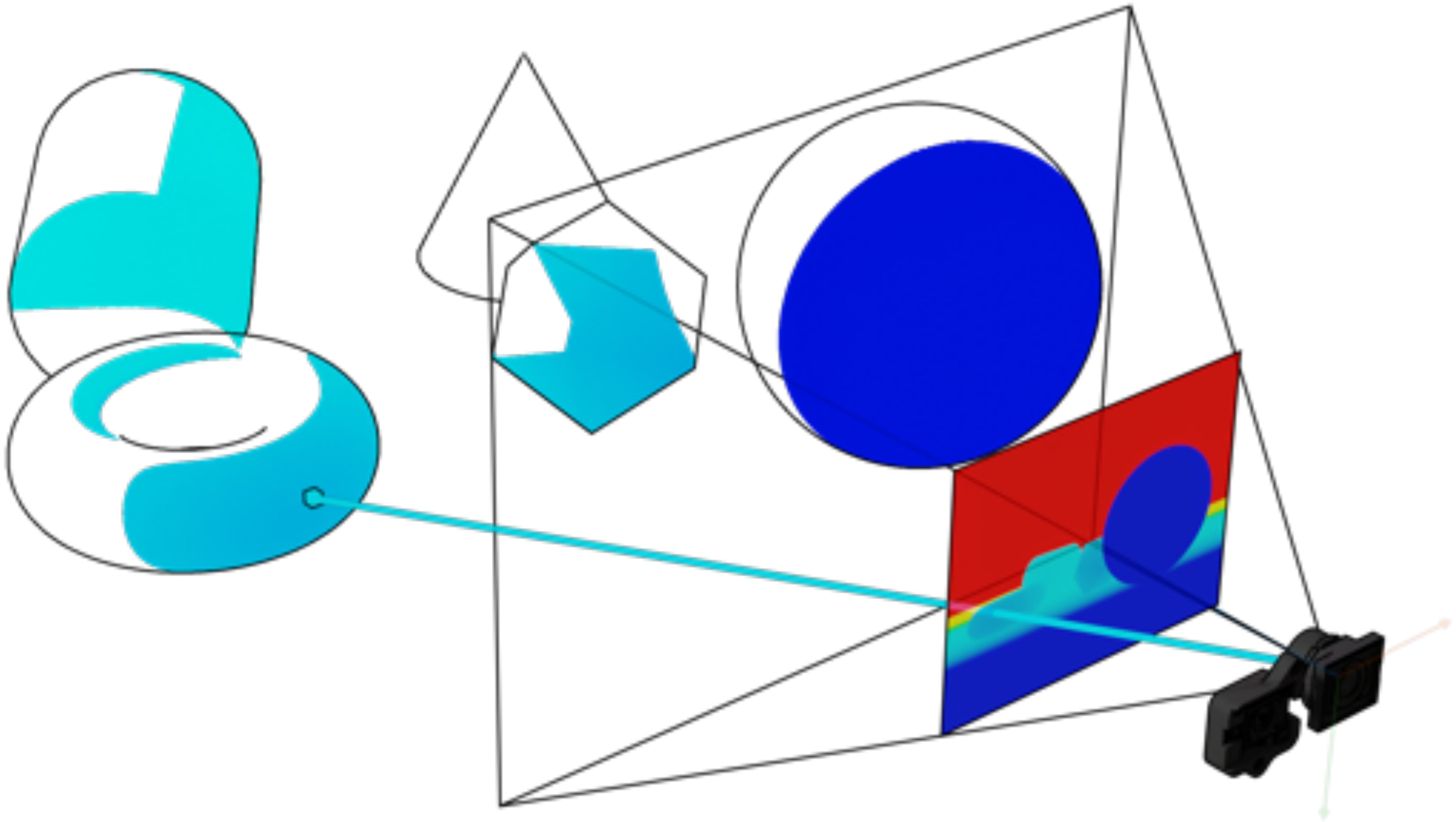
[1] https://cseweb.ucsd.edu/~shz338/research.html
[2] https://www.kaggle.com/code/mitchellodili/stanford-cars

How can we generate textures for 3D objects using only 2D images?

Our Approach



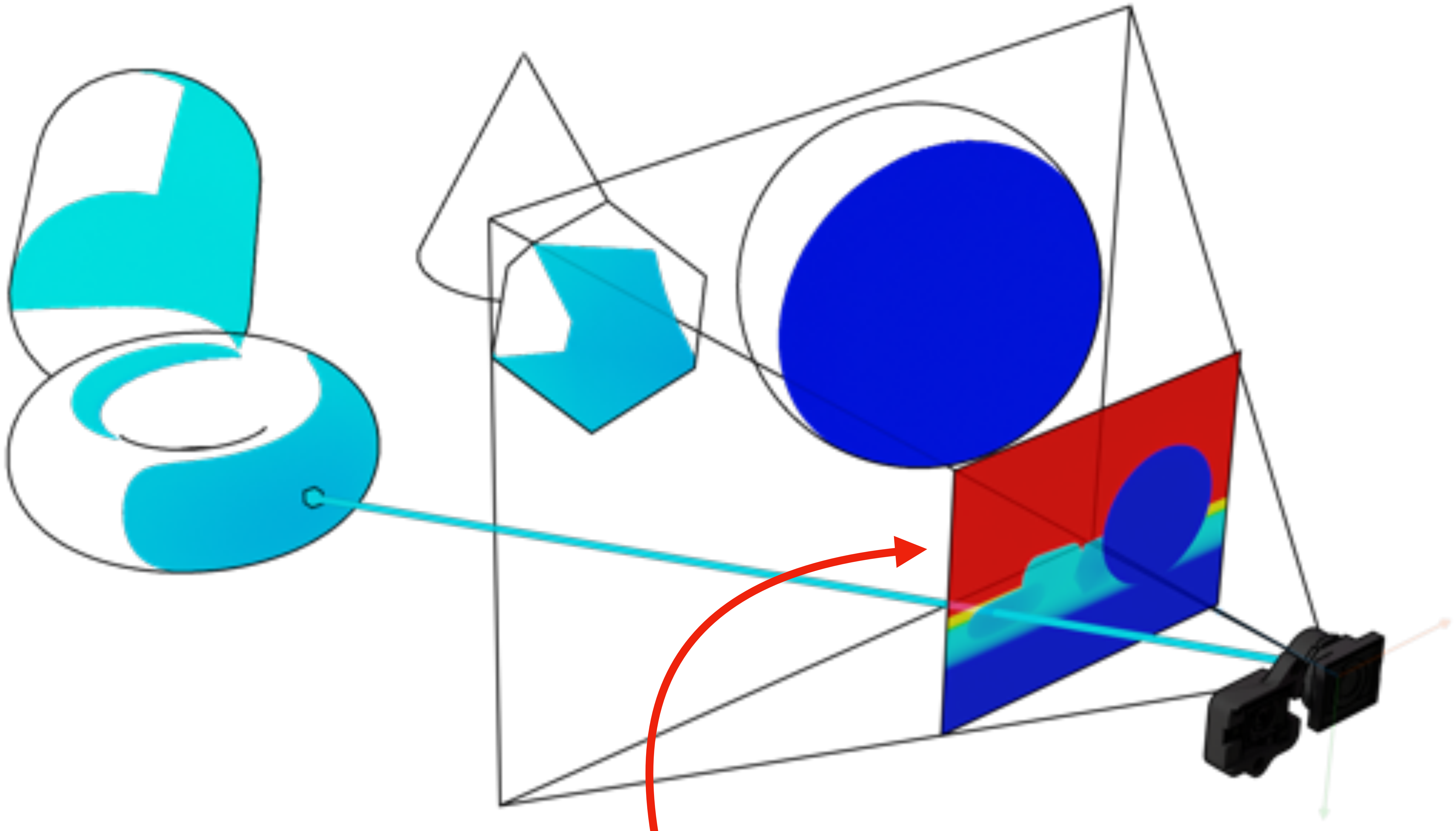
Project Mapping



Projective Texture Mapping (Williams, SIGGRAPH '78)

[1] https://dev.intelrealsense.com/docs/projection-texture-mapping-and-occlusion-with-intel-realsense-depth-cameras

Project Mapping



How to generate this image?

[1] https://dev.intelrealsense.com/docs/projection-texture-mapping-and-occlusion-with-intel-realsense-depth-cameras

Generative Adversarial Networks (GANs)

Image GANs can produce realistic images.

[1] https://github.com/rosinality/stylegan2-pytorch

Generative Adversarial Networks (GANs)

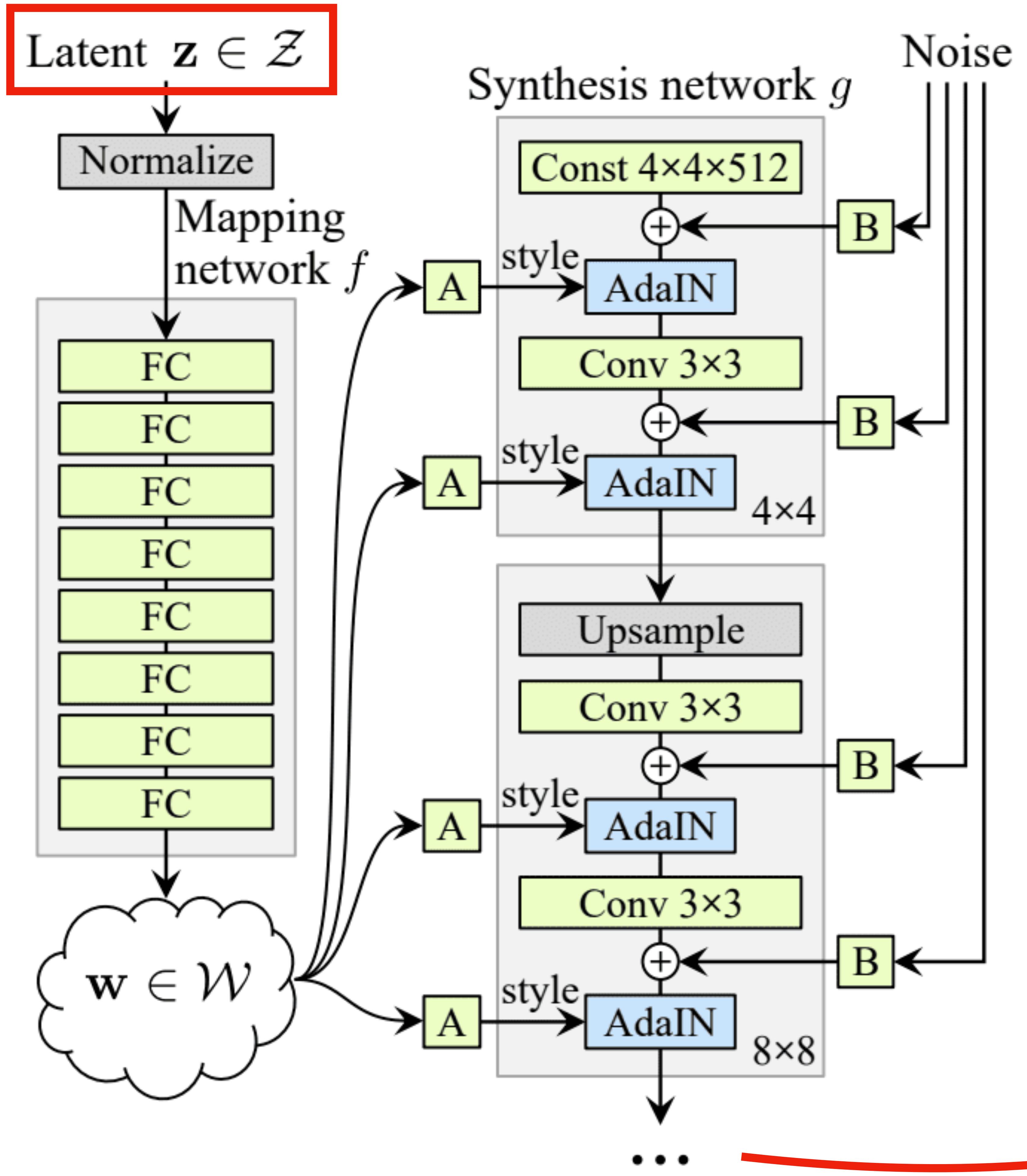
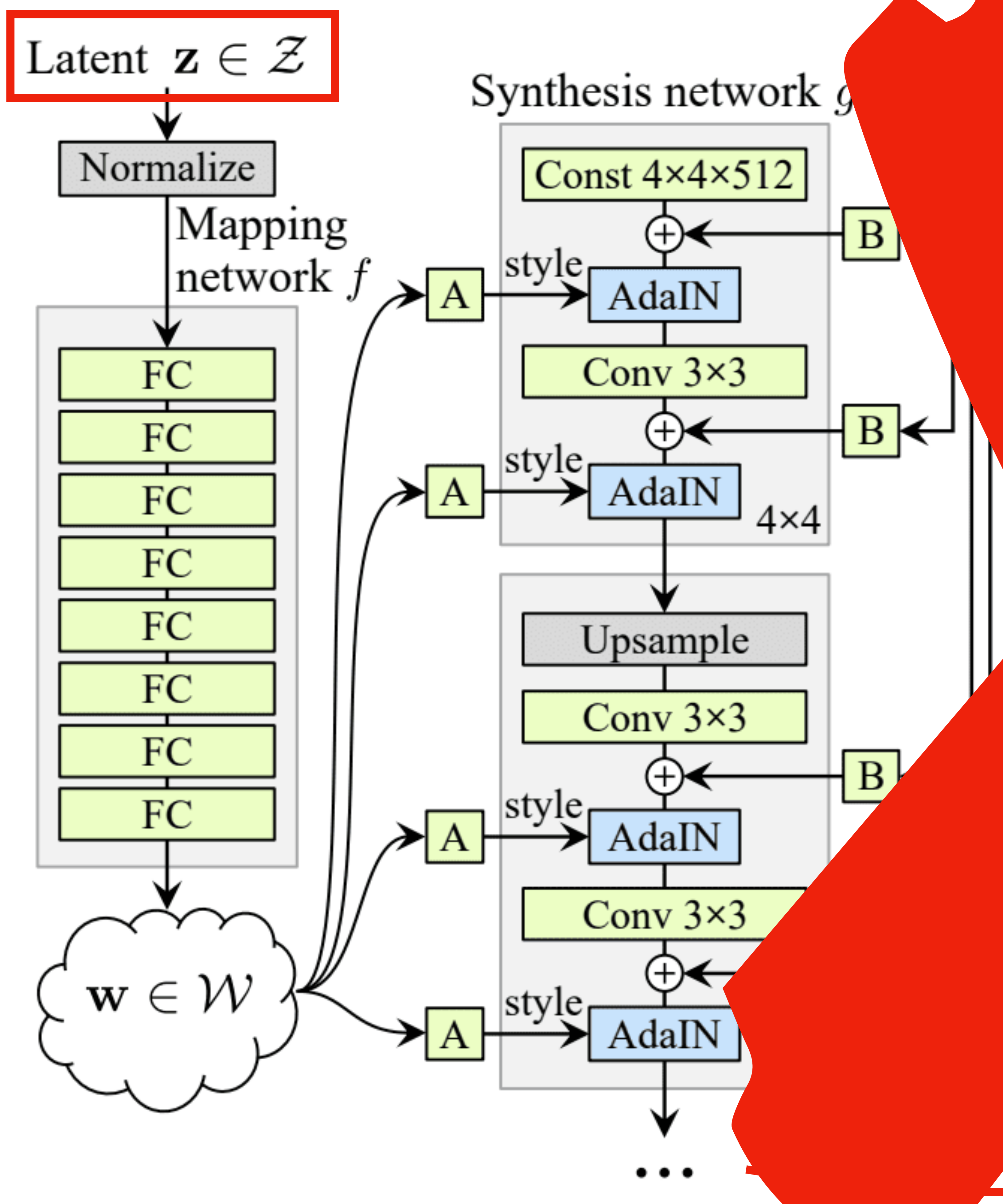


Image GANs generate images with one forward pass.

[1] A Style-Based Generator Architecture for Generative Adversarial Networks, Karras et al., CVPR 2019

Generative Adversarial Networks (GANs)

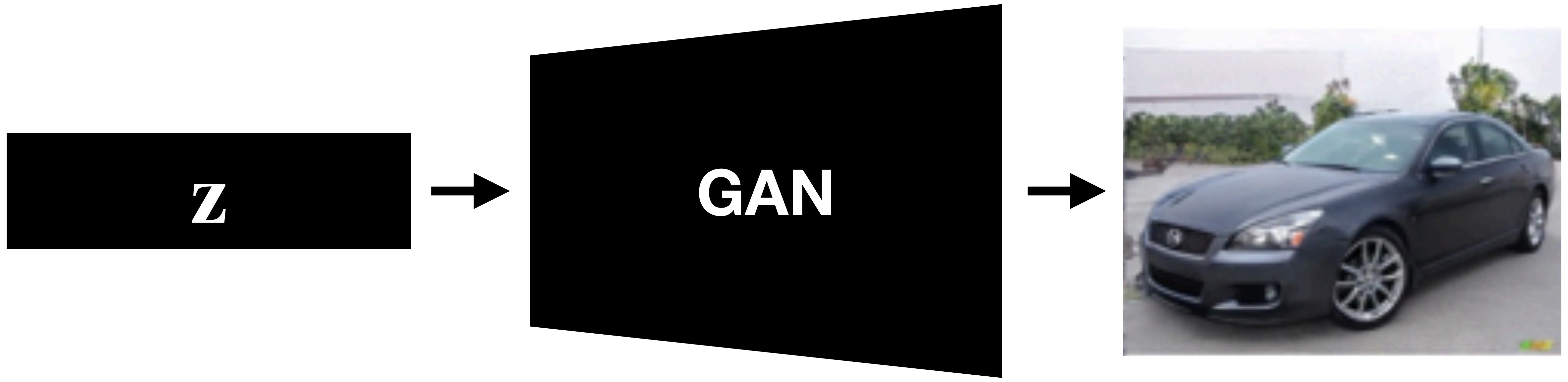


...ANs generate images

...th one forward pass.

[1] A Style-Based Generator Architecture for Generative Adversarial Networks, Karras et al., CVPR 2019

Generative Adversarial Networks (GANs)



Latent-code-to-Image

GANs

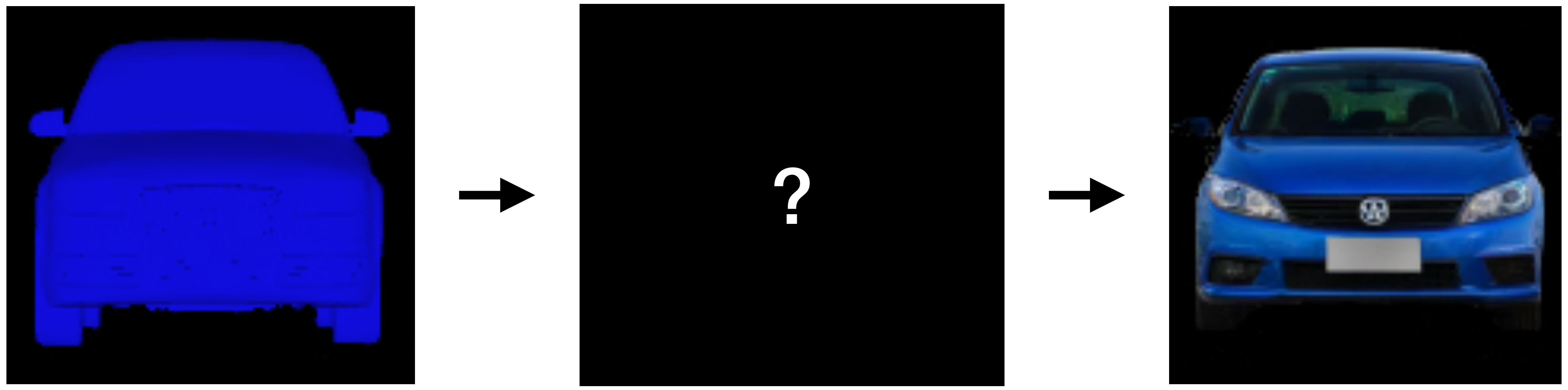


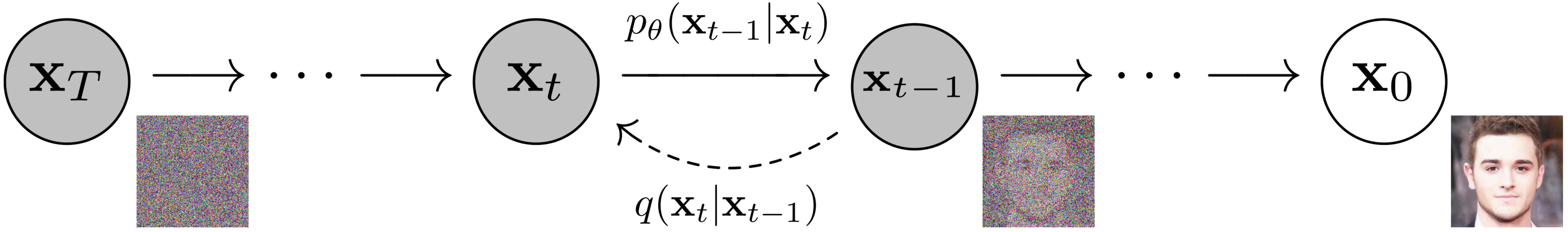
Image-to-Image

Ours

Diffusion Models

GANs can do this, but with heavy engineering.

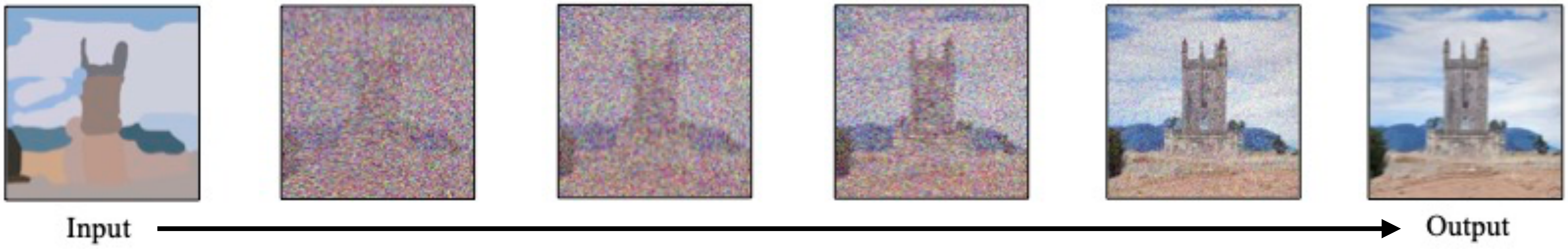
Diffusion Models



Diffusion models generate images in an iterative manner.

[1] Denoising Diffusion Probabilistic Models, Ho et al., NeurIPS 2020

Diffusion Models



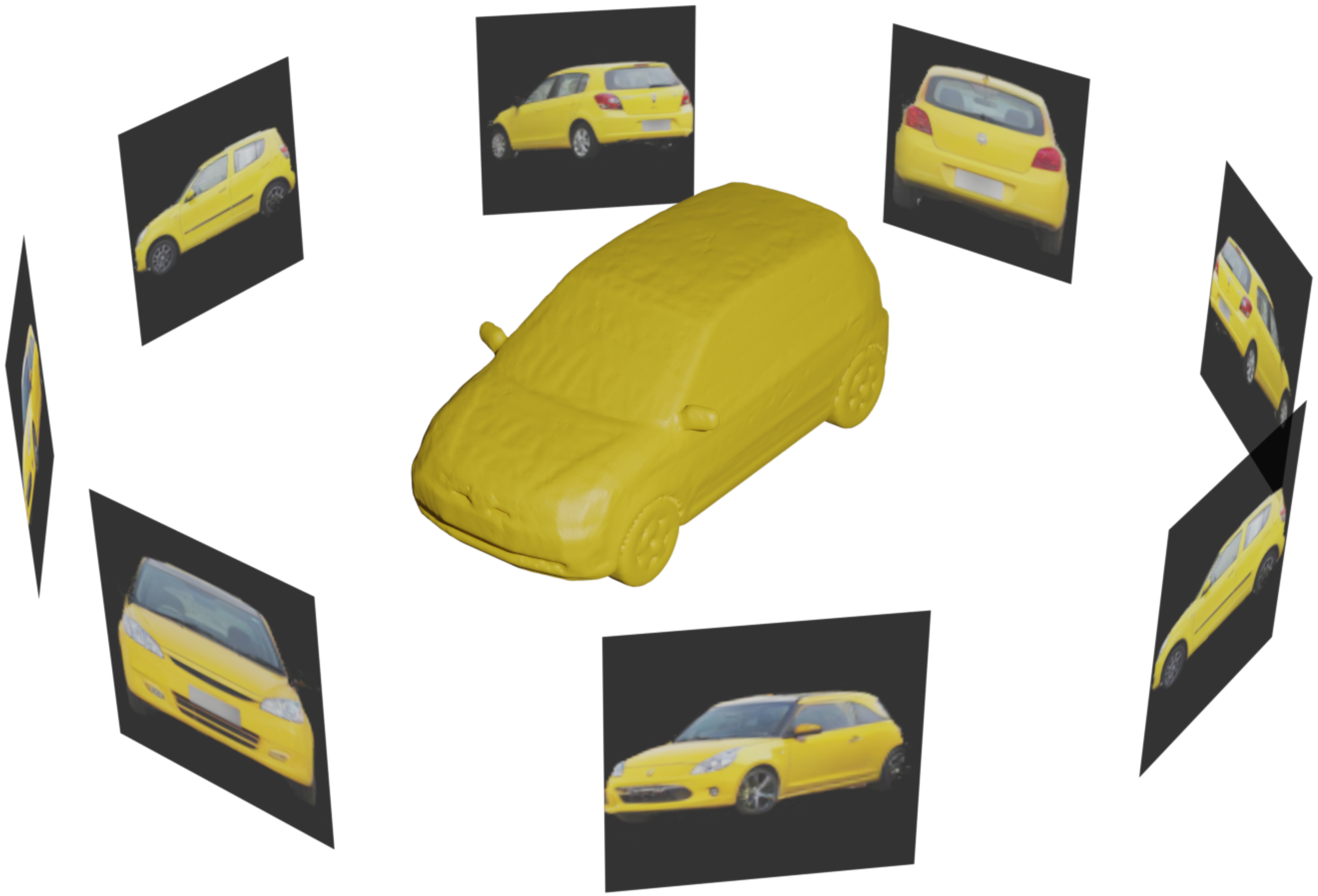
We can easily generate images that match silhouettes.

[1] SDEdit: Guided Image Synthesis and Editing with Stochastic Differential Equations, Meng et al., ICLR 2022

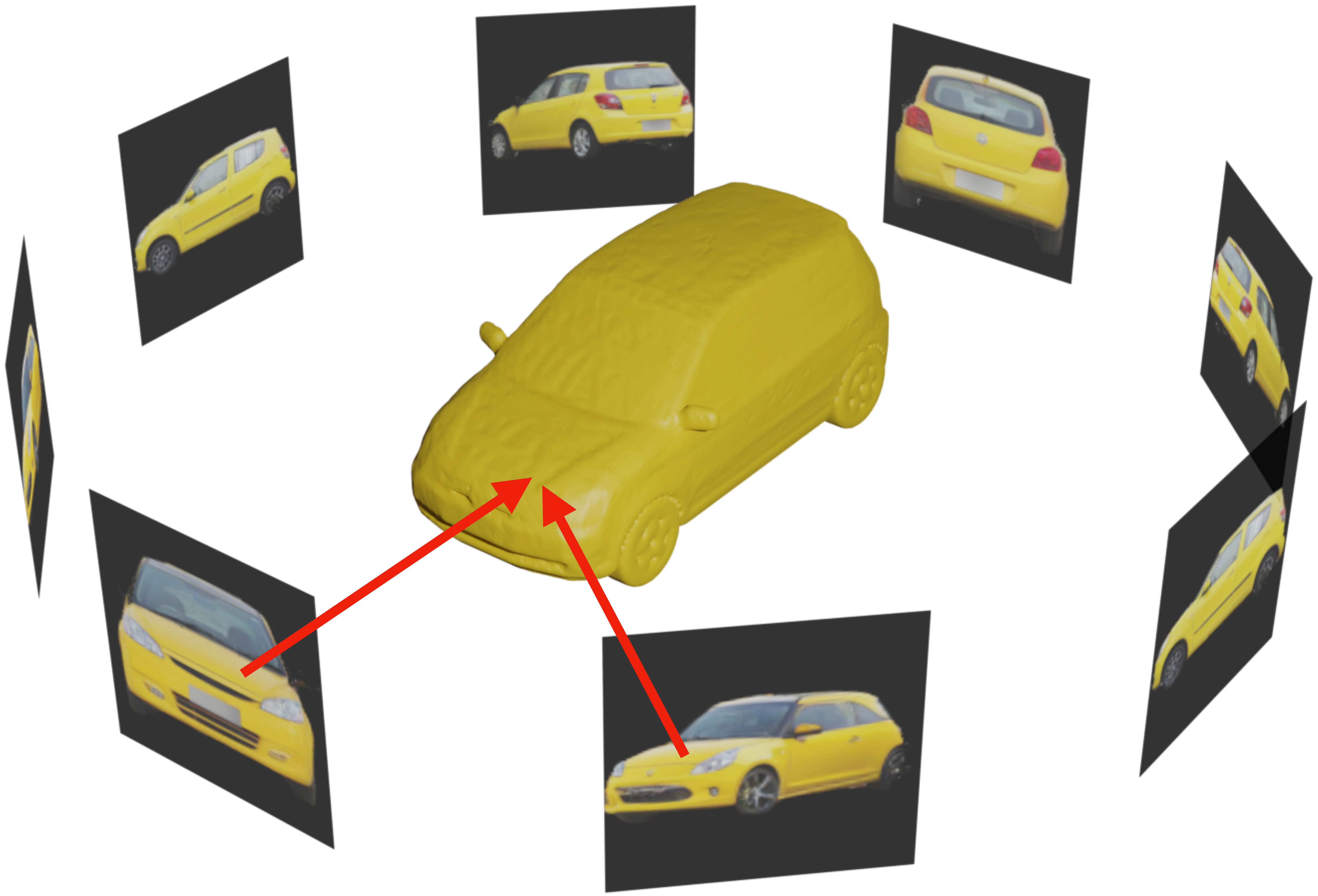
Diffusion Models

Failure case: Totally different images across the viewpoints.

Key Idea: Multi-View Diffusion

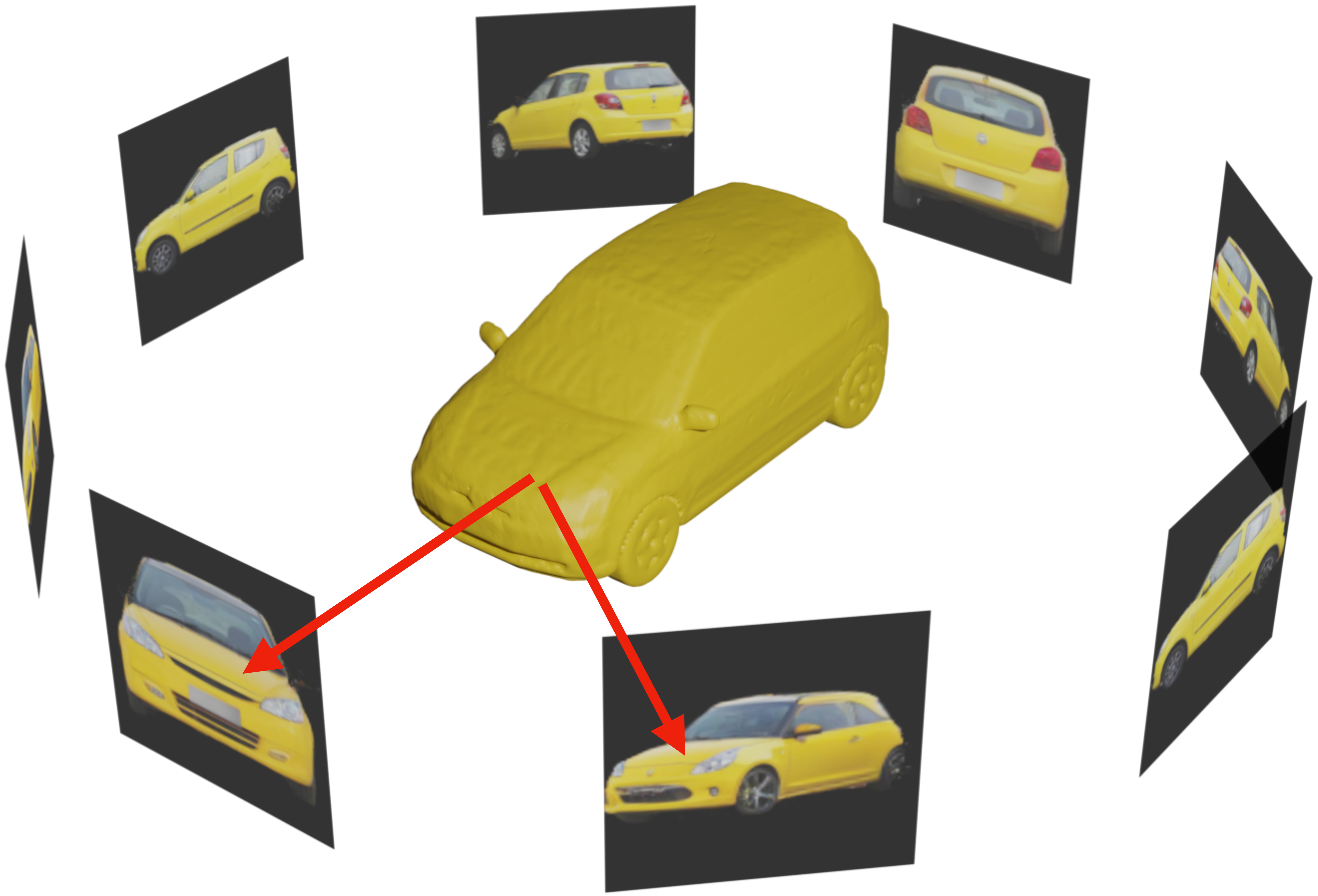


Key Idea: Multi-View Diffusion



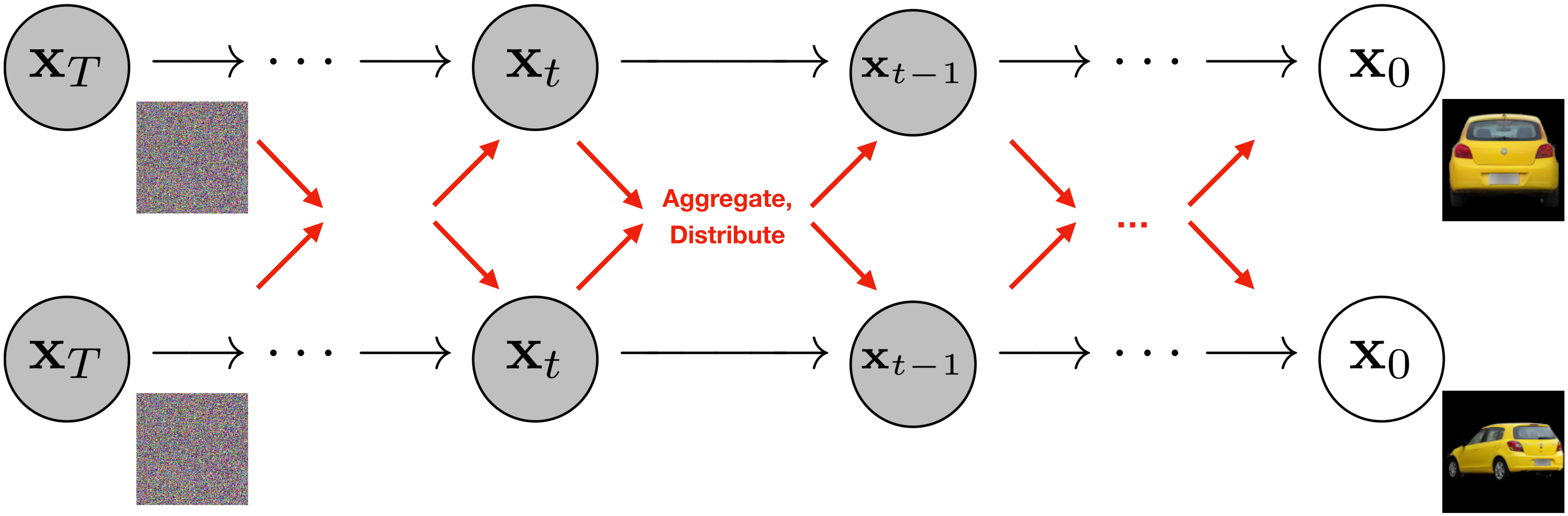
Use 2D-to-3D mapping to aggregate pixel colors.

Key Idea: Multi-View Diffusion



Distribute the aggregated colors back to the images.

Key Idea: Multi-View Diffusion



***Synchronize* image generation at different viewpoints.**

[1] Denoising Diffusion Probabilistic Models, Ho et al., NeurIPS 2020

Results

Setup

Data

- 184 3D car objects from the *ShapeNet* dataset
- 16,000 real-world car images from the *Comprehensive Cars* dataset
- 8 diffusion models each generating images for each viewpoint

Baseline

- Learning Texture Generators for 3D Shape Collections from Internet Photo Sets, BMVC 2021

Qualitative Evaluation

(GAN-Based)

Yu et al., BMVC 2021

Ours

Qualitative Evaluation

(GAN-Based)

Yu et al., BMVC 2021

Ours

Qualitative Evaluation

(w/o multi-view diffusion)

(w/ multi-view diffusion)

Quantitative Evaluation

These metrics tell us "how realistic the generated images are".

Method	FID (↓)	GIQA ($\times 10^2$, ↑)
Yu et al., BMVC 2021	44.84	9.79
Ours	**34.39**	**11.24**

User Study

“Here are two images depicting the same type of car.
Which of these two cars do you think is more realistic?"

	Yu et al., BMVC 2021	Ours
Preference (%)	24.8 -+ 3.8	**75.2 +- 3.8**

Conclusion

- Propose a novel framework for 3D object texture generation based on 2D image generation and projective texture mapping.

Conclusion

- Propose a novel framework for 3D object texture generation based on 2D image generation and projective texture mapping.
- Introduce a multi-view diffusion framework for consistent image generation and silhouette alignment.

Conclusion

- Propose a novel framework for 3D object texture generation based on 2D image generation and projective texture mapping.
- Introduce a multi-view diffusion framework for consistent image generation and silhouette alignment.
- Demonstrates outperformance compared with previous GAN-based state-of-the-art method.

PaintMe

Neural Texture Synthesis via Joint Multi-View Stochastic Diffusion Process

Seungwoo Yoo

dreamy1534@kaist.ac.kr

TA: Juil Koo
Advisor: Minhyuk Sung

Geometric AI Lab, School of Computing