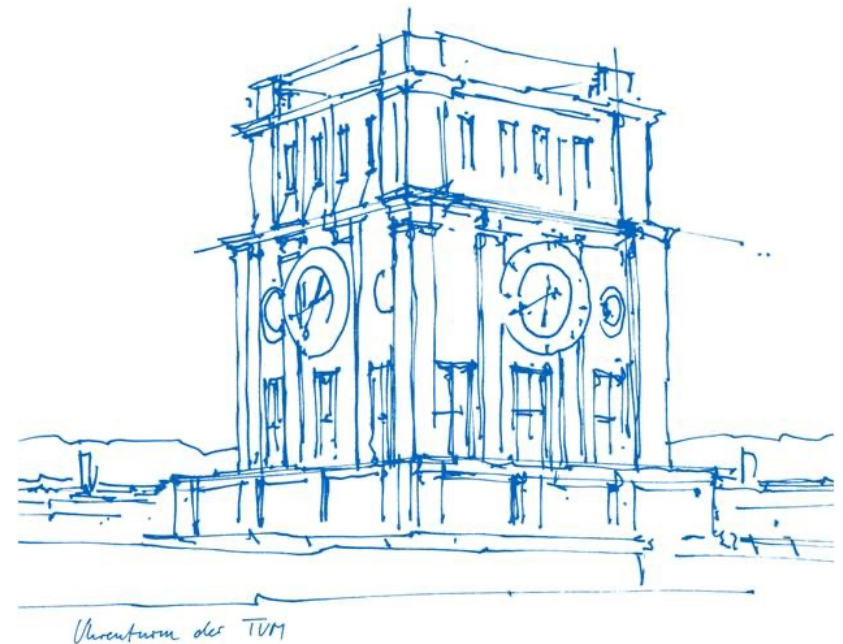


Seminar on Neural Radiance Fields and Gaussian Splatting – Gaussian Dreamer

Osama Salem

Technische Universität München

Munich, 27. May 2025



Motivation

- Democratizing content creation
 - Help amateur developers create 3D assets
- Help Guide Professionals
 - Reduce amount of tedious work
- Preceive 3D structure from partial observations



Motivation

3D Native

Speed: 🐇 (Seconds)

Data Requirements: ★★★★★

Consistency: ★★★★★



Shape-E



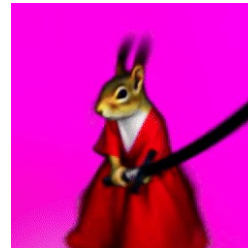
Point-E

2D Lifting

Speed: 🐢 (Hours)

Data Requirements: ★

Consistency: ★



DreamFusion

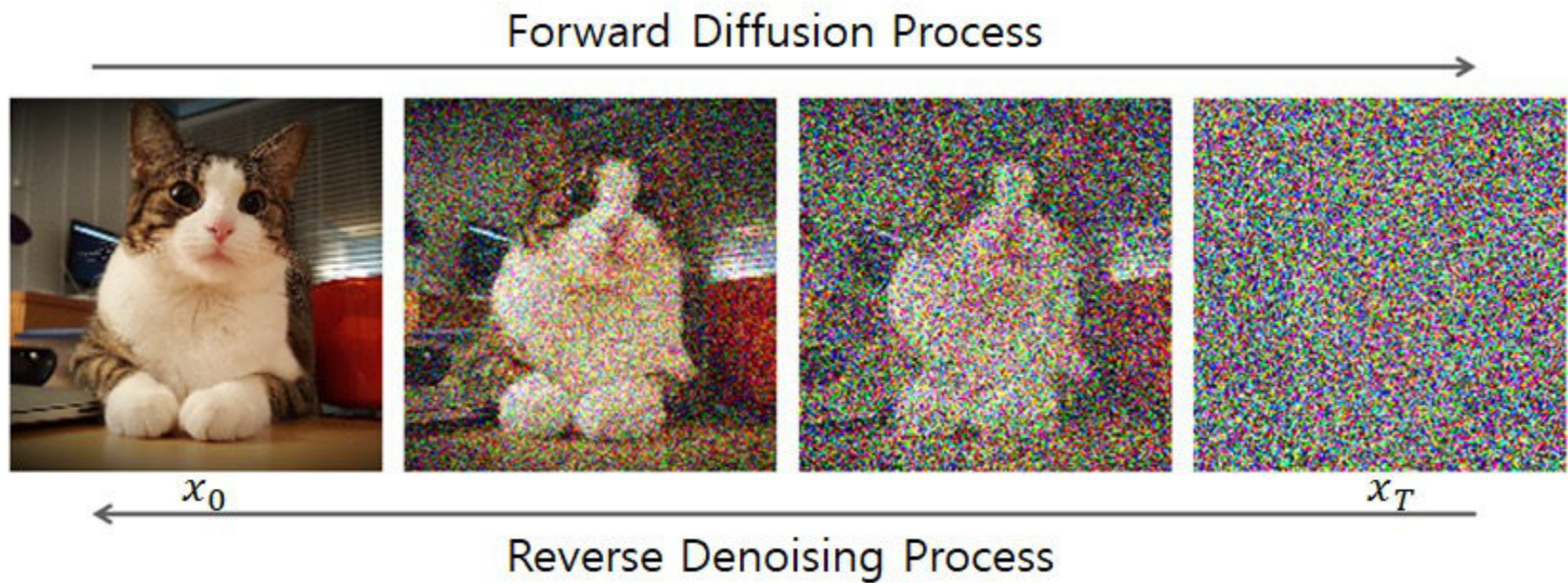


Fantasia3D



Gaussian Dreamer

Diffusion Models



Diffusion Models

Denoising diffusion models consist of two processes:

- Forward diffusion process that gradually adds noise to input – **Fixed**

$$x_t = \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} \varepsilon$$

Noise Scheduler
 $\mathcal{N}(0, I)$

- Reverse denoising process that learns to generate data by denoising – **Learnt**

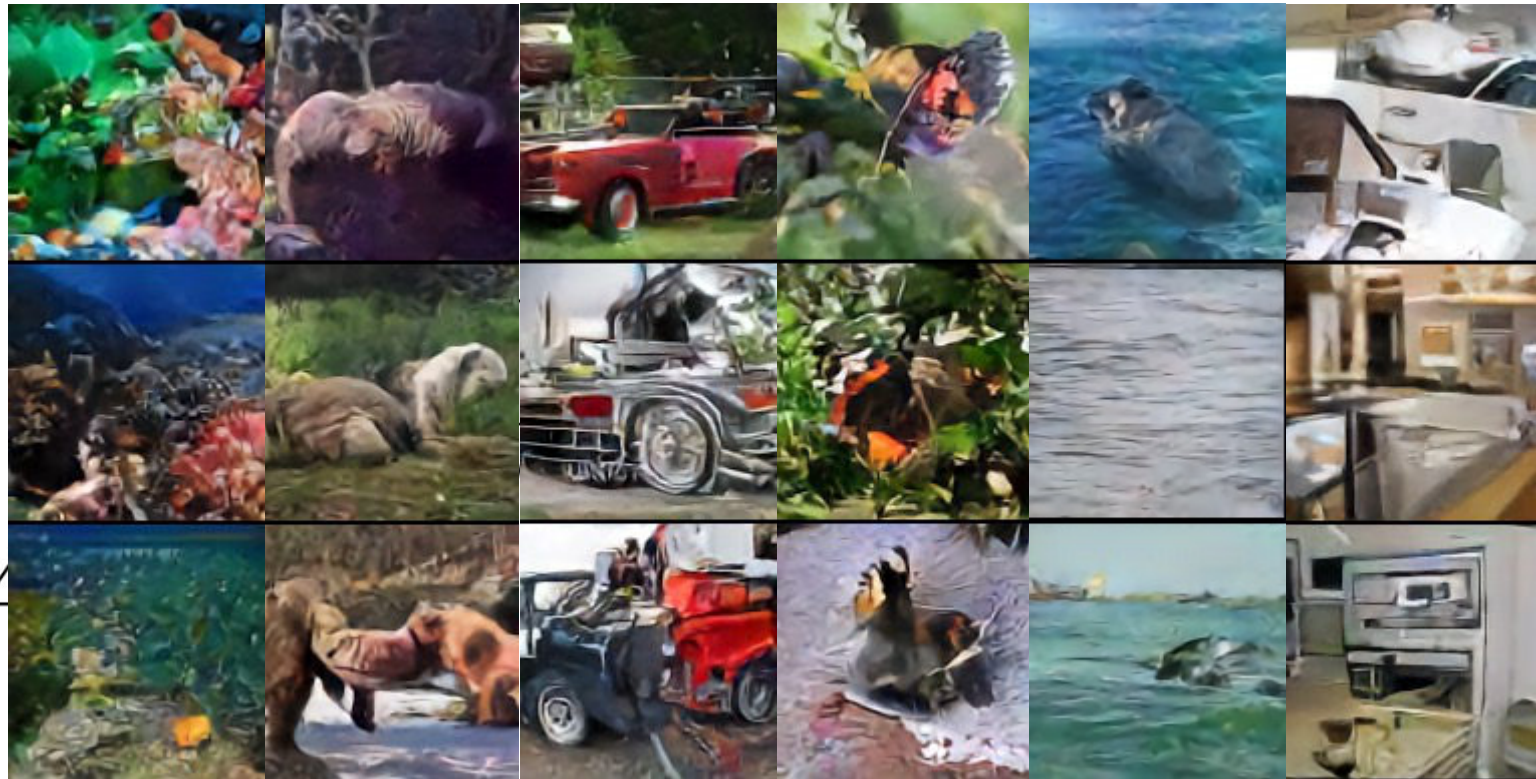
$$L = E_{t, x_0, \varepsilon} [\|\varepsilon - \varepsilon_\theta(x_t, t, y)\|_2^2]$$

Expectation
What the model predicts

Problem: Generated images are low resolution and bad – GANs are way ahead!

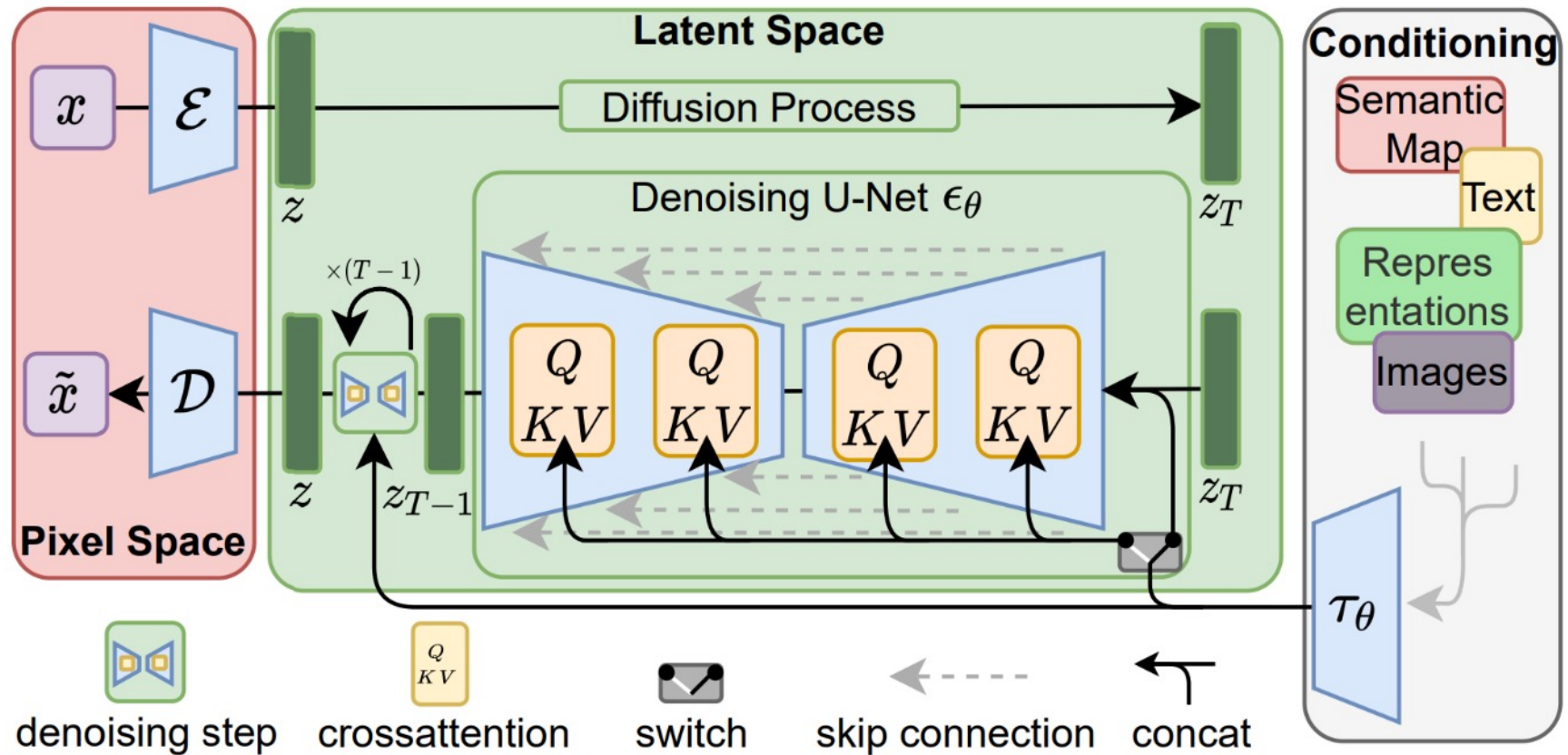
Latent Diffusion Models

VQ-VAE is great for reconstruction but not generation.

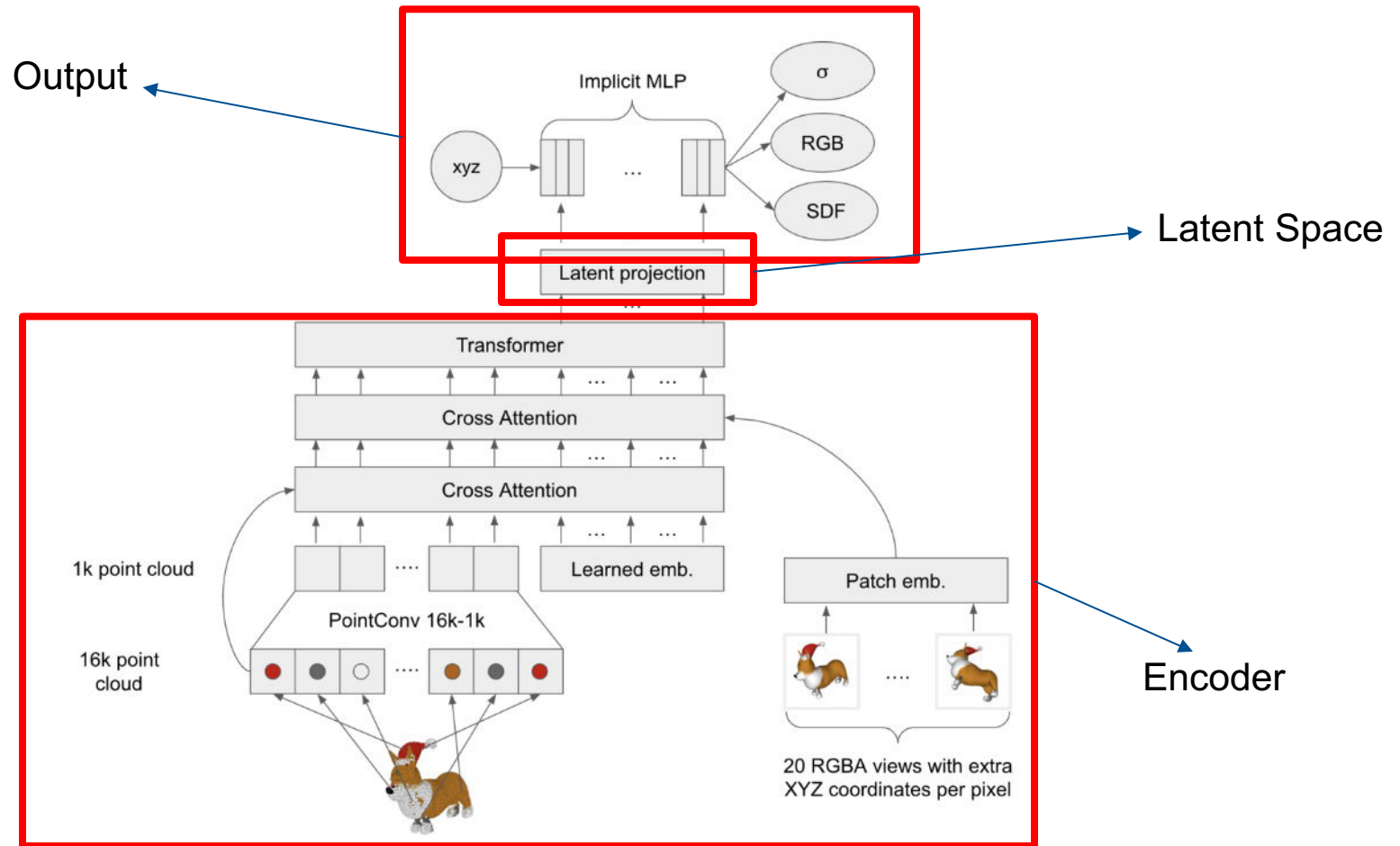


GANs are still better!

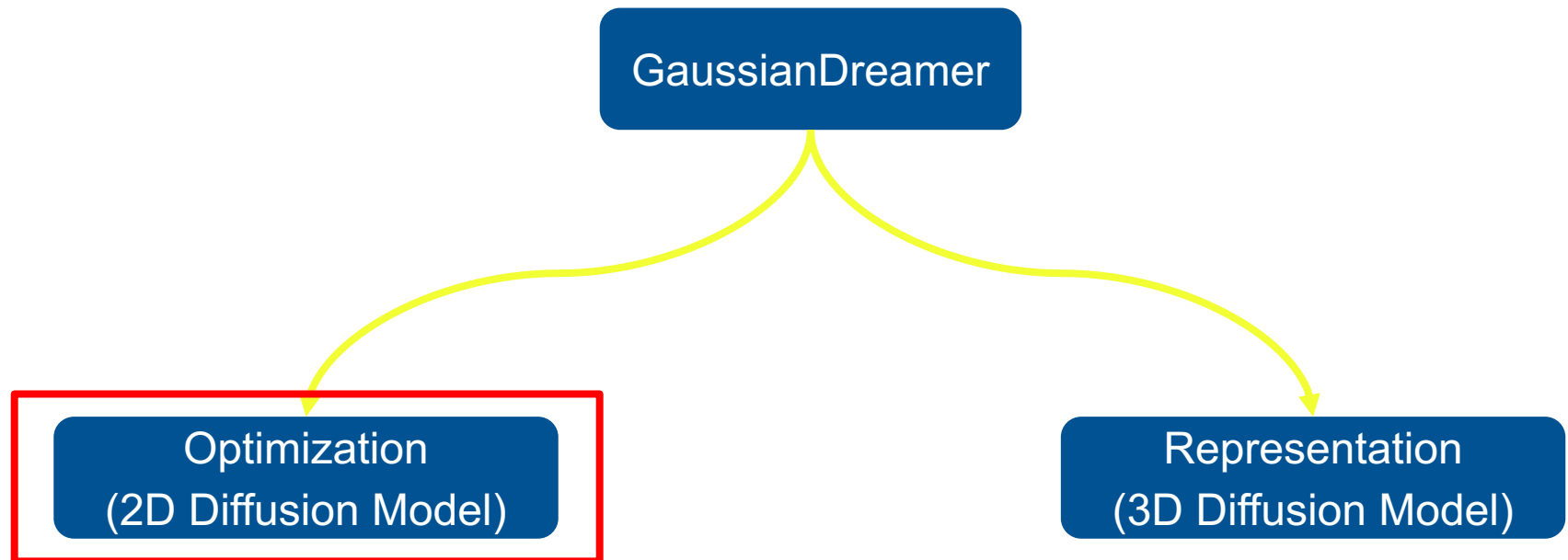
2D Latent Diffusion Models



3D Latent Diffusion Models



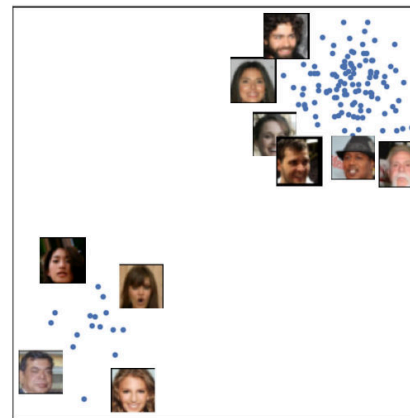
Method Overview



What is actually Diffusion?

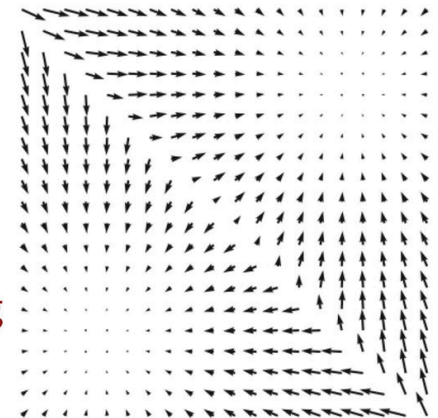
the **score function** $\nabla_x \log p(x)$ points in the direction of highest data density. It's a gradient field guiding samples towards data.

$$\nabla_{x_t} \log p_t(x_t | x_0) = -\frac{(x_t - x_0)}{1 - \bar{\alpha}_t}$$



Data samples

score
matching



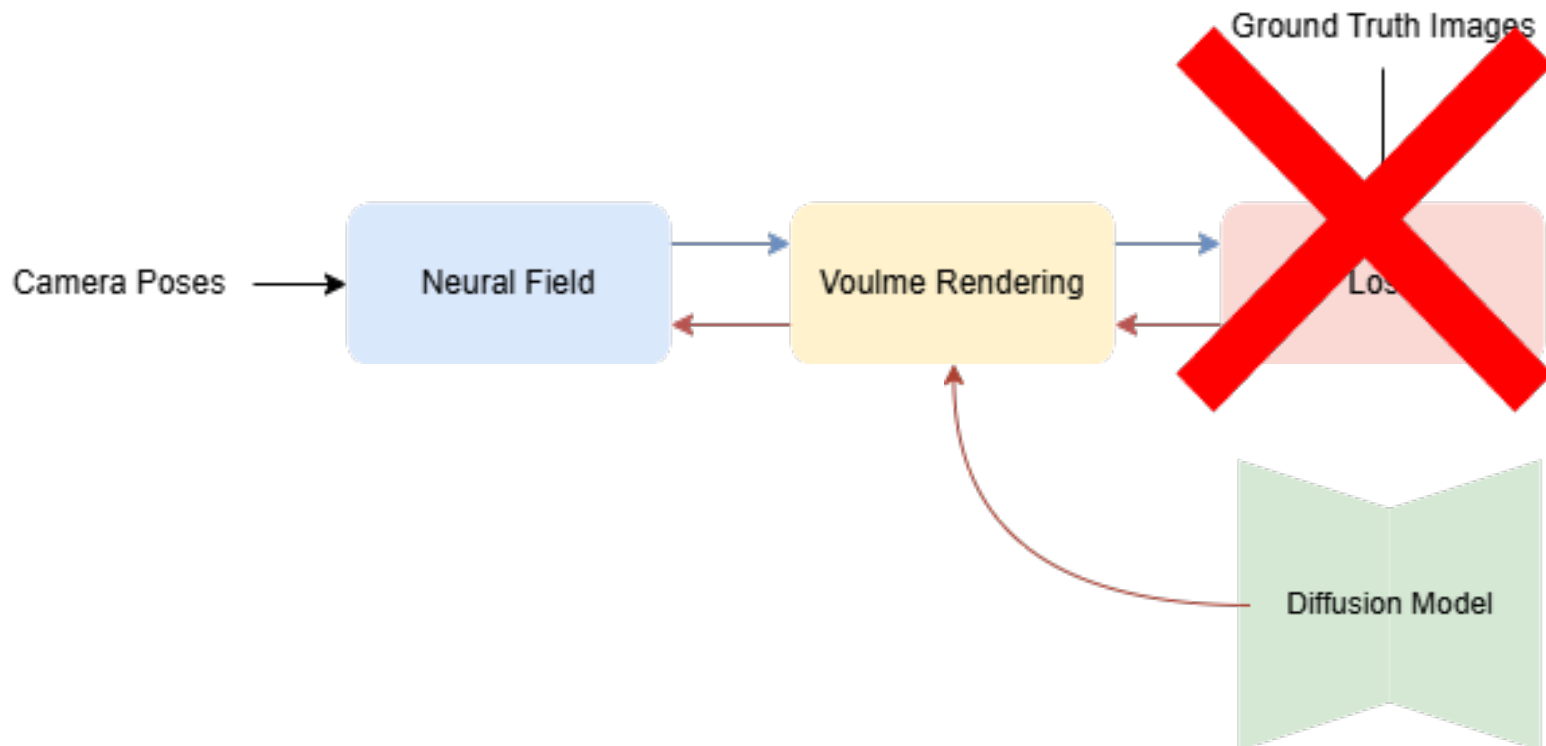
Scores

Recall Forward Process:

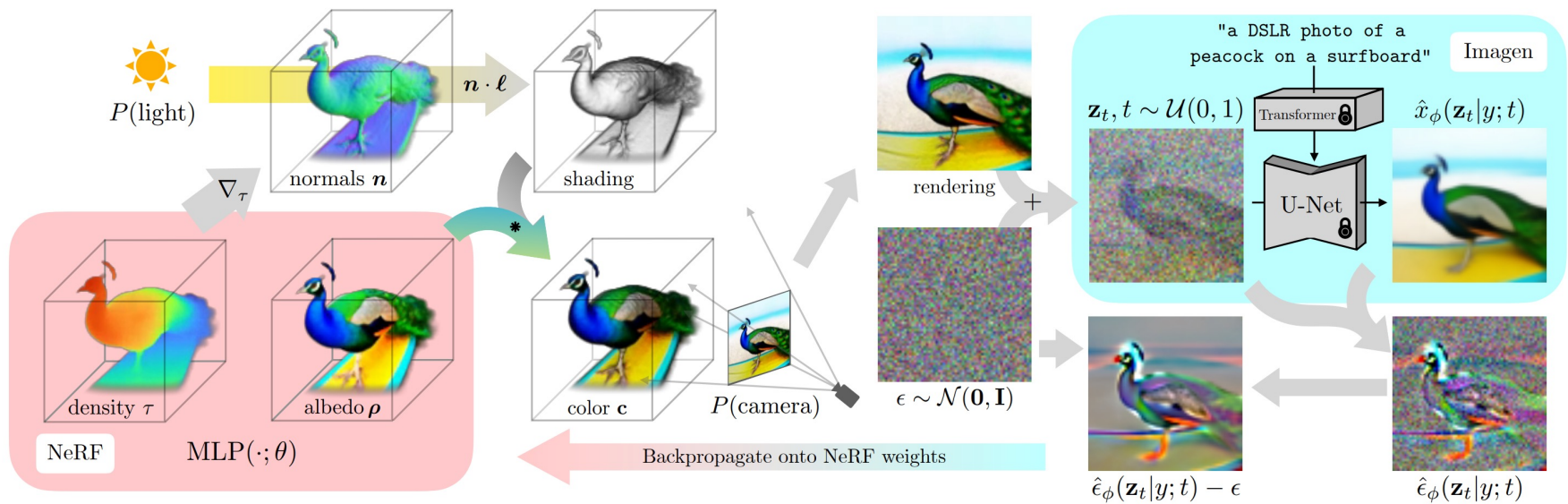
$$x_t = \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} \varepsilon \quad \longrightarrow \quad \varepsilon = \frac{(x_t - x_0)}{\sqrt{1 - \bar{\alpha}_t}}$$

$$\varepsilon = -\sqrt{1 - \bar{\alpha}_t} \nabla_{x_t} \log p_t(x_t | x_0)$$

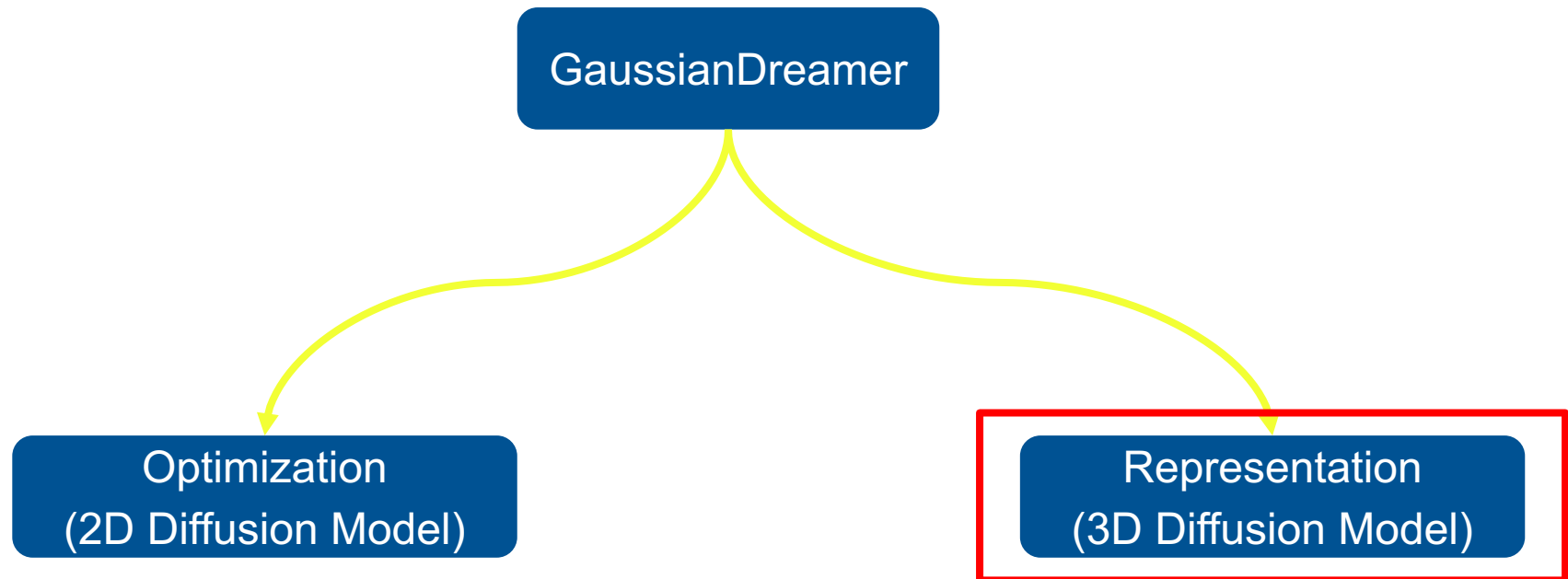
Neural Radiance Fields



DreamFusion3D



Method Overview



3D Gaussian Splatting



3D Gaussian Spaltting

imgflip.com



Point Clouds

Gaussian Splatting Attributes

Where? Position



How? Spherical Harmonics

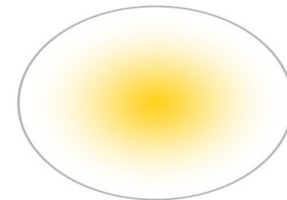


How Transparent? Opacity



How Big? Covariance

Same Everywhere? Gaussian Decay

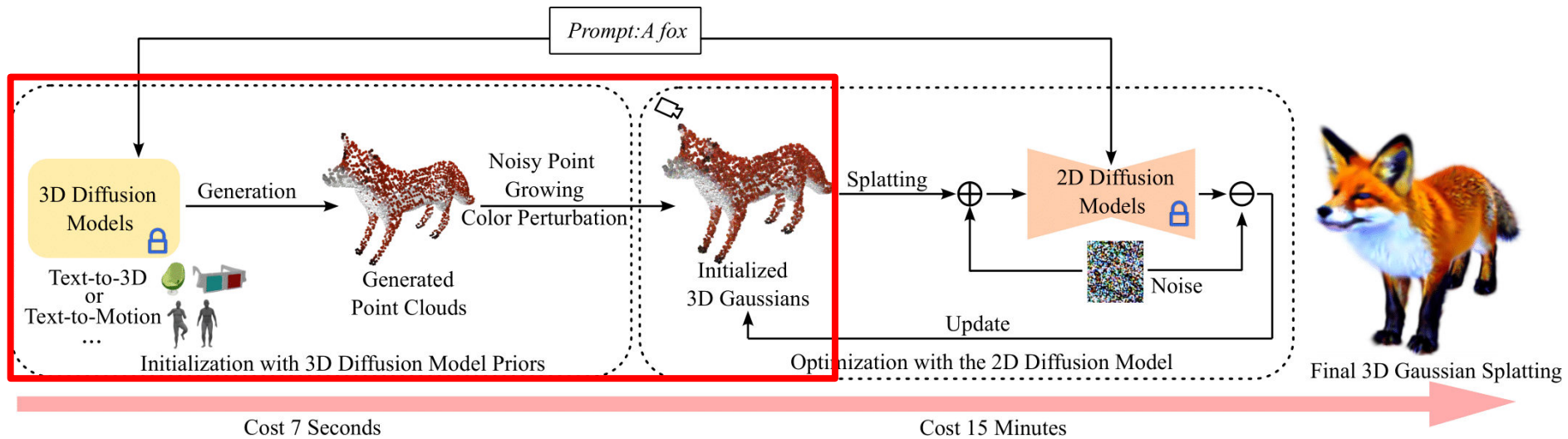


Alpha Blending

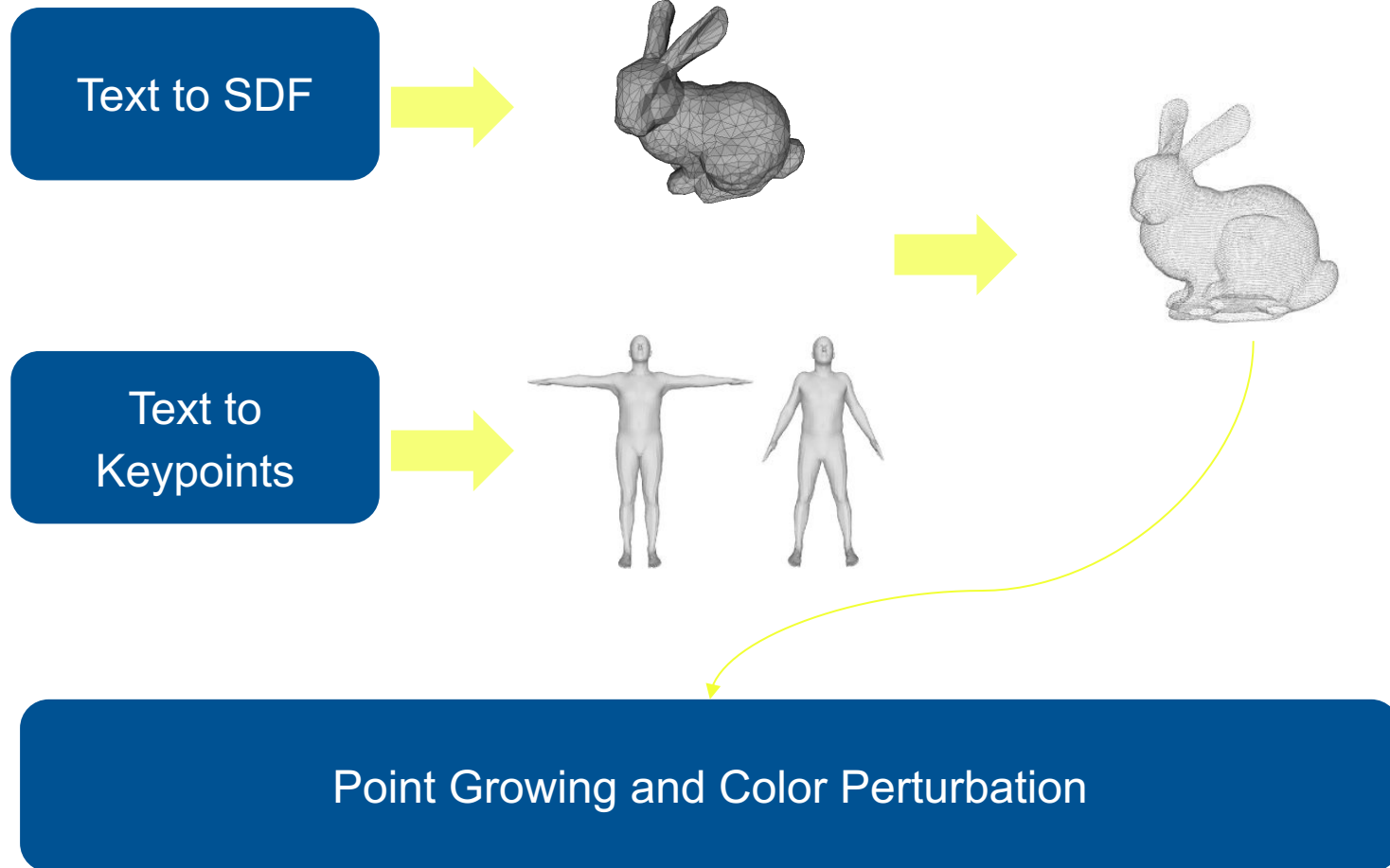




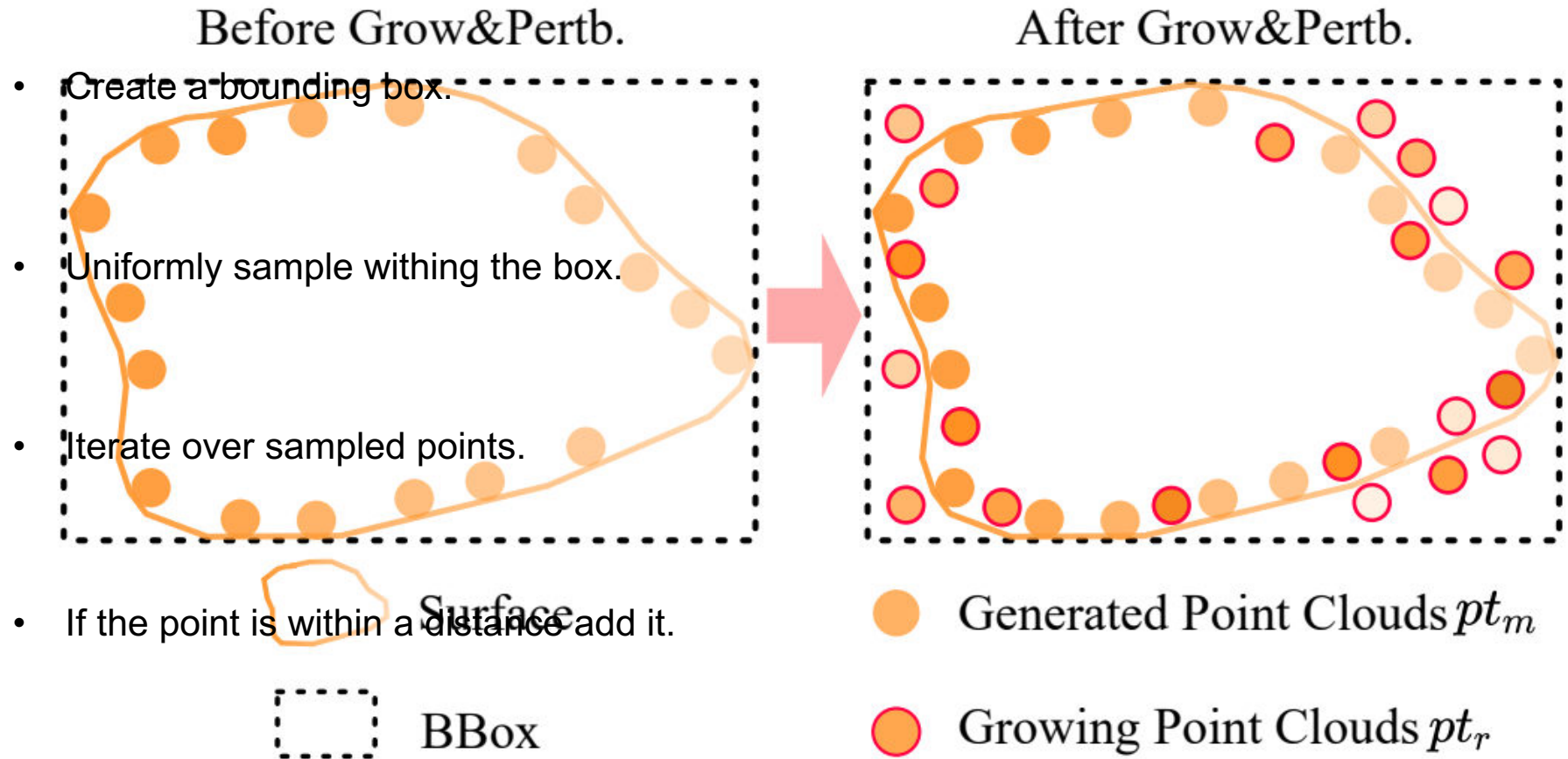
Gaussian Dreamer



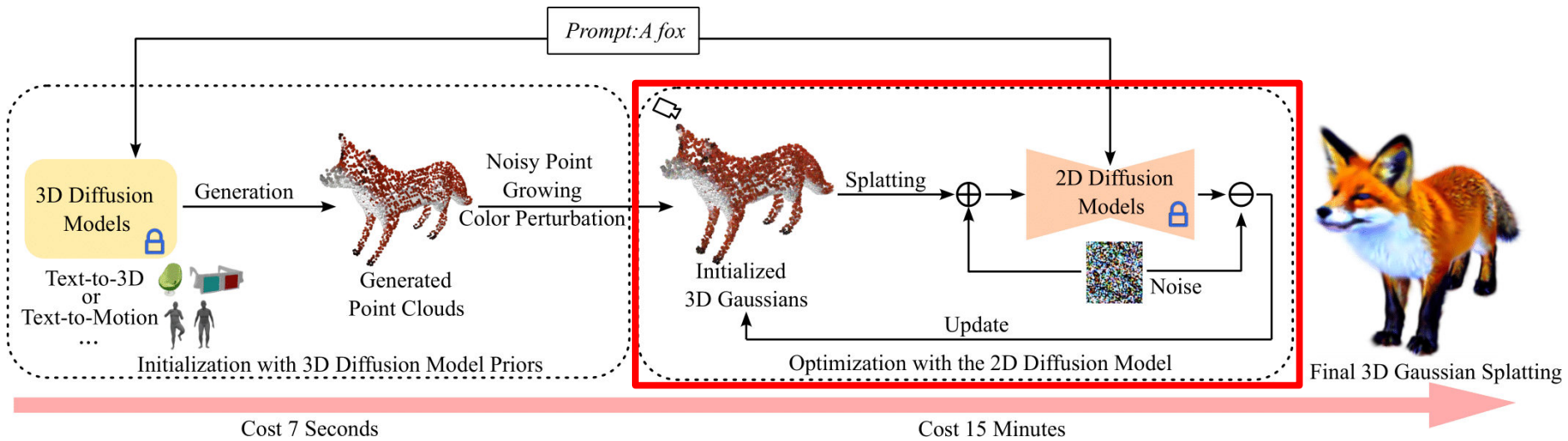
Initialization



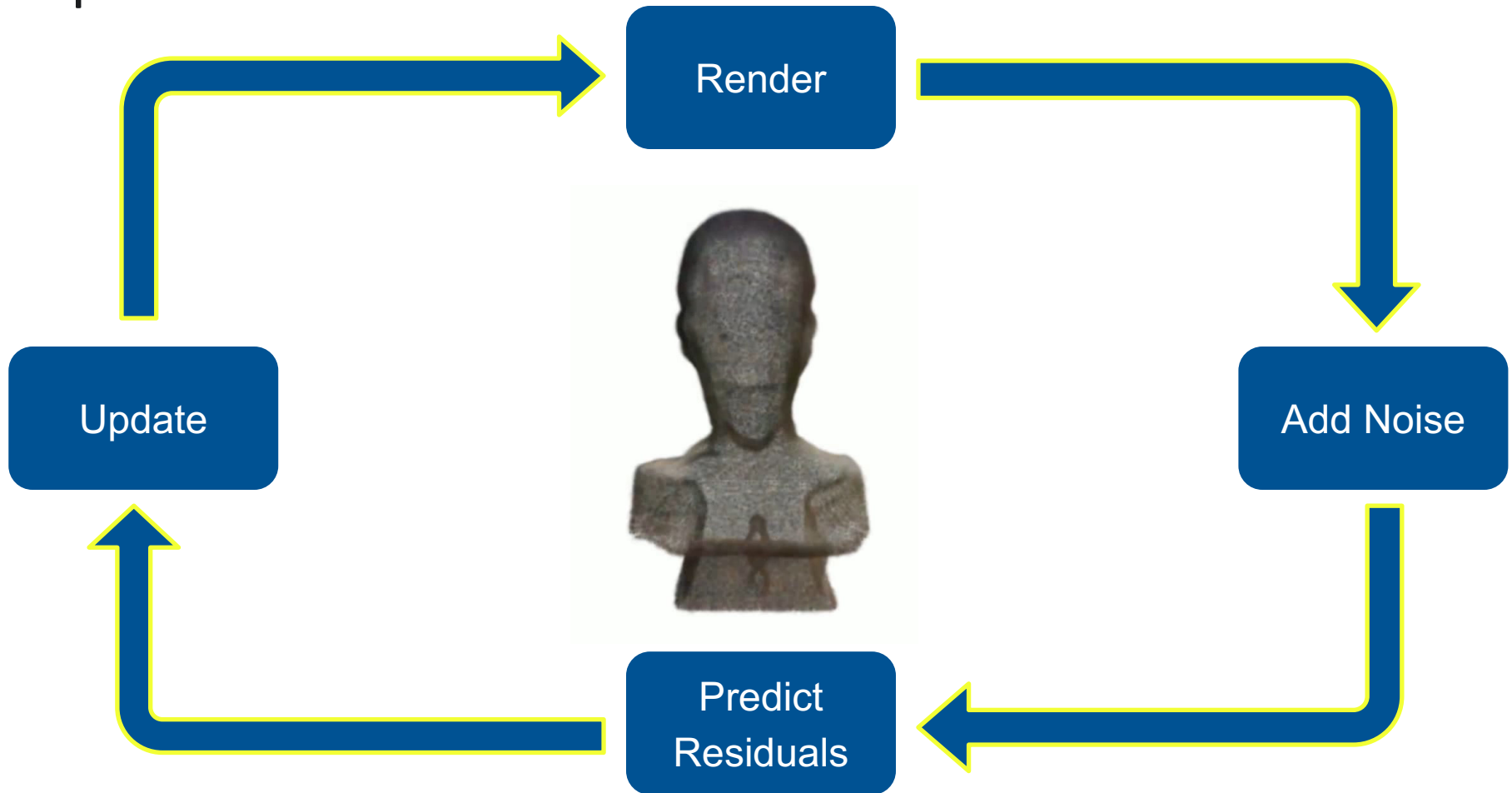
Point Growing & Color Perturbation



Optimization



Optimization



Quantitative Results – T³ Benchmark

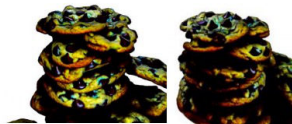
Method	Time	Single Obj.	Single w/ Surr.	Multi Obj.	Average
SJC	-	24.7	19.8	11.7	18.7
DreamFusion	6 Hours	24.4	24.6	16.1	21.7
Fantasia3D	6 Hours	26.4	27.0	18.5	24.0
LatentNeRF	15 Minutes	33.1	30.6	20.6	28.1
Magic3D	5.3 Hours	37.0	35.4	25.7	32.7
ProlificDreamer	10 Hours	49.4	44.8	35.8	43.3
Gaussian Dreamer	15 Minutes	54.0	48.6	34.5	45.7

Comparison Results

DreamFusion
6.0 Hours



Magic3D
5.3 Hours



Fantasia3D
6.0 Hours



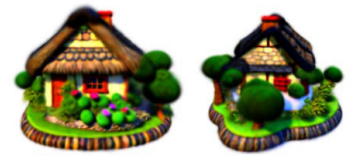
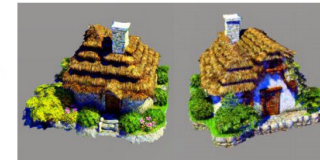
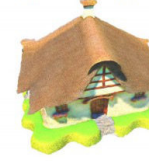
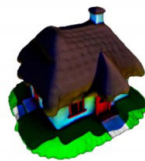
ProlificDreamer
~ Hours



GaussianDreamer (Ours)
15 Minutes

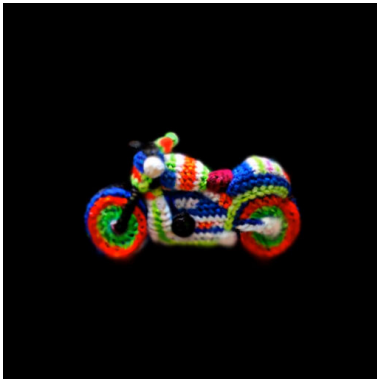


A plate piled high with chocolate chip cookies.

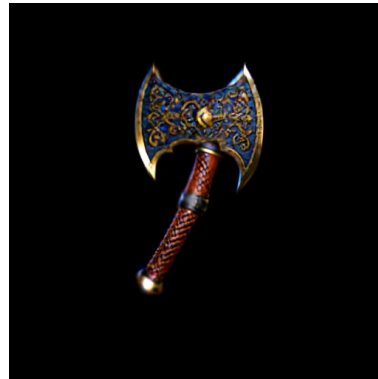


A 3D model of an adorable cottage with a thatched roof.

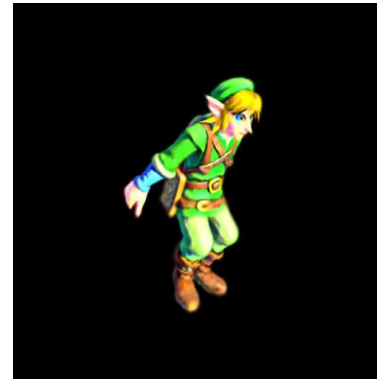
Further Results



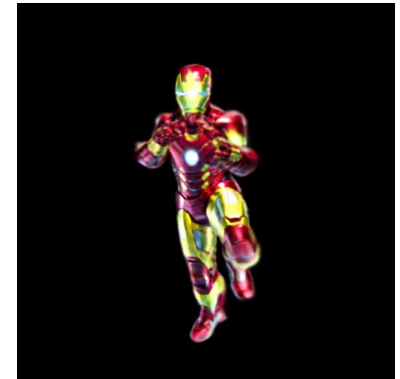
a zoomed out DSLR photo of
an amigurumi motorcycle



Viking axe, fantasy, weapon,
blender, 8k, HD



Link in Zelda jumped down
from the sky



Iron man kicks with his left leg

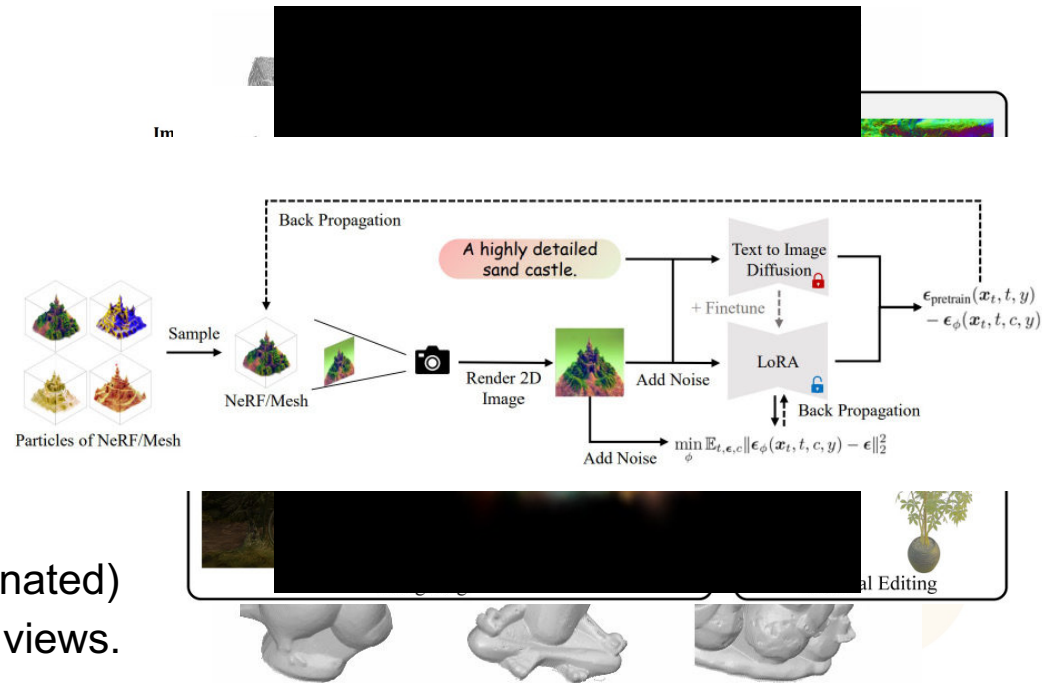
Discussion – Strengths & Innovations

- Unprecedented Speed for Quality
 - ~15 Minutes Optimization + ~7 Seconds Initialization
- Geometric Consistency via 3D Prior
 - Enforcing an inductive bias mitigates Janus Problem
- High-Fidelity Explicit Representation
 - Gaussian Splatting enables detail and fast rendering ~200 FPS



Discussion – Limitations & Challenges

- Single Object Optimization
 - Dealing with object optimization.
- Gaussian Splatting as representation
 - Unbounded to surfaces.
 - Hard to edit and re-light.
- View Consistency (Mitigated, Not Eliminated)
 - 2D Diffusion model biased to certain views.



Thanks for your attention

Osama Salem

Garching, 27. May 2025

