



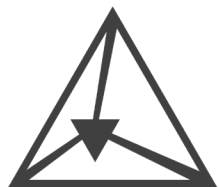
# Single Image to 3D using Cross-Domain Diffusion

## Wonder3D

Xiaoxiao Long

# 3D generation

- Lift 2D prior (text-to-image diffusion models) into 3D
  - text-to-3D generation
  - image-to-3D generation



## threestudio



ProlificDreamer

DreamFusion

Magic3D

SJC

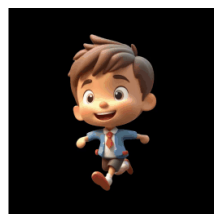
LatentNeRF

Fantasia3D

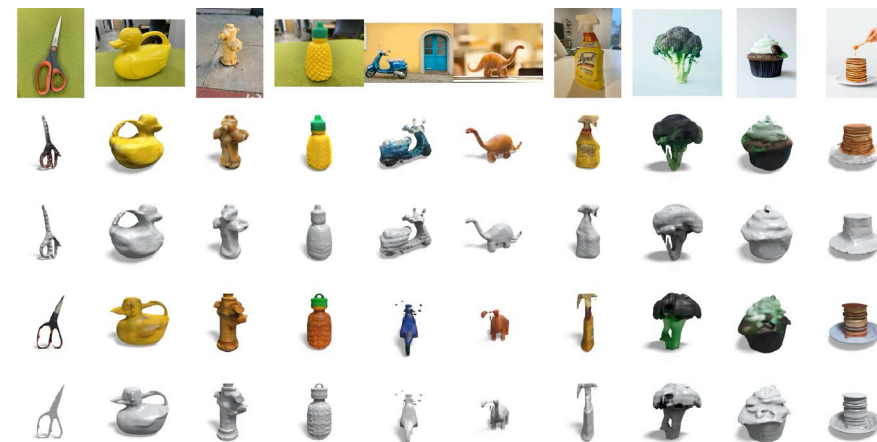
TextMesh



Zero-1-to-3



Magic123

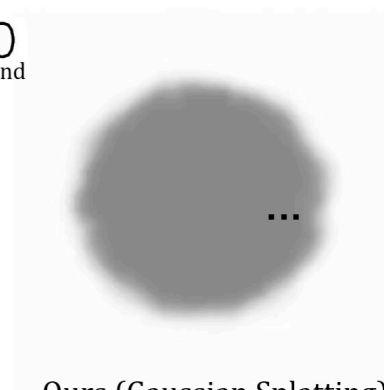


One-2345



Zero-1-to-3 (NeRF)

00:00  
Minute Second



Ours (Gaussian Splatting)

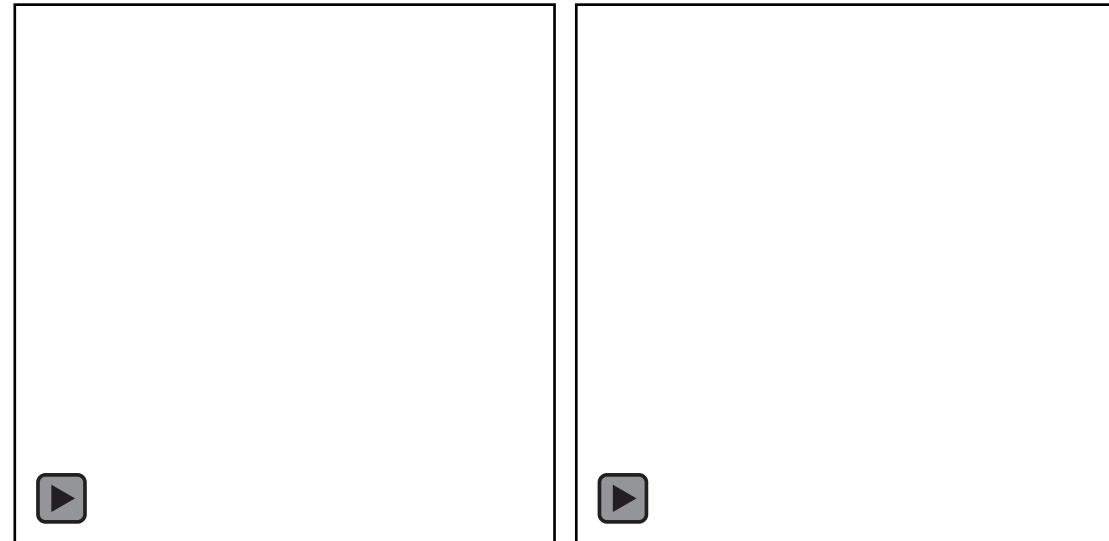


# Limitations

- Janus Problem / Multi-face problem
- Time-consuming optimization
- Handy parameter tuning



**Vanilla 2D diffusion  
models are not 3D aware!**



# Motivation

- High-efficiency
- Feed-forward manner
- High-quality geometry
- Easy to export textured meshes
- Strong generalization

**Make 2D diffusion  
models 3D-aware!**

## Zero-1-to-3: Zero-shot One Image to 3D Object

Ruoshi Liu  
Columbia University

Rundi Wu  
Columbia University

Basile Van Hoorick  
Columbia University

Pavel Tokmakov  
Toyota Research  
Institute

Sergey Zakharov  
Toyota Research  
Institute

Carl Vondrick  
Columbia University

TL;DR: We learn to control the camera perspective in large-scale diffusion models, enabling zero-shot novel view synthesis and 3D reconstruction from a single image.



## SyncDreamer: Generating Multiview-consistent Images from a Single-view Image

Arxiv 2023

Yuan Liu<sup>1,2</sup>, Cheng Lin<sup>2</sup>, Zijiao Zeng<sup>2</sup>, Xiaoxiao Long<sup>1</sup>, Lingjie Liu<sup>3</sup>, Taku Komura<sup>1</sup>, Wenping Wang<sup>4</sup>

<sup>1</sup>The University of Hong Kong <sup>2</sup>Tencent Games <sup>3</sup>University of Pennsylvania <sup>4</sup>Texas A&M University

Paper

Code

Model

Live Demo

## MVDream: Multi-view Diffusion for 3D Generation

Yichun Shi<sup>1</sup>

Peng Wang<sup>1</sup>

Jianglong Ye<sup>2</sup>

Long Mai<sup>1</sup>

Kejie Li<sup>1</sup>

Xiao Yang<sup>1</sup>

<sup>1</sup>ByteDance

<sup>2</sup>University of California San Diego

Paper

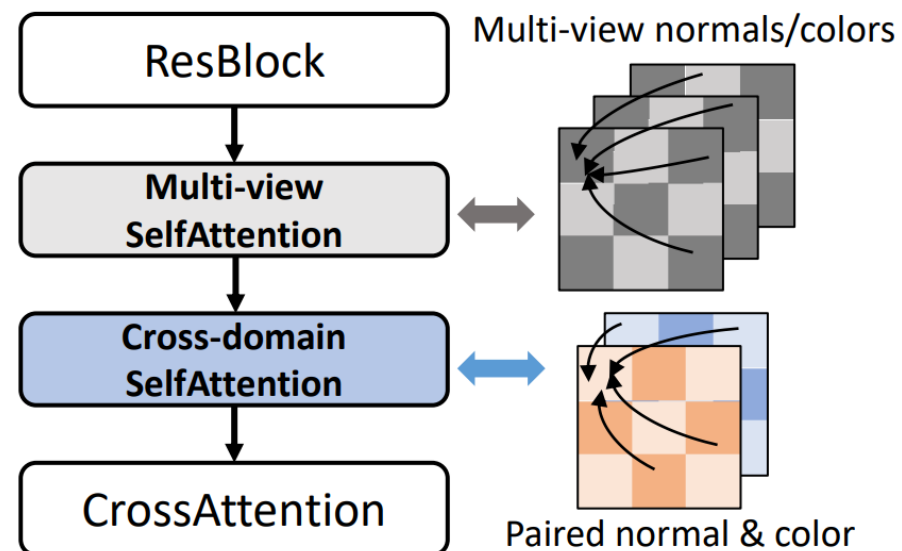
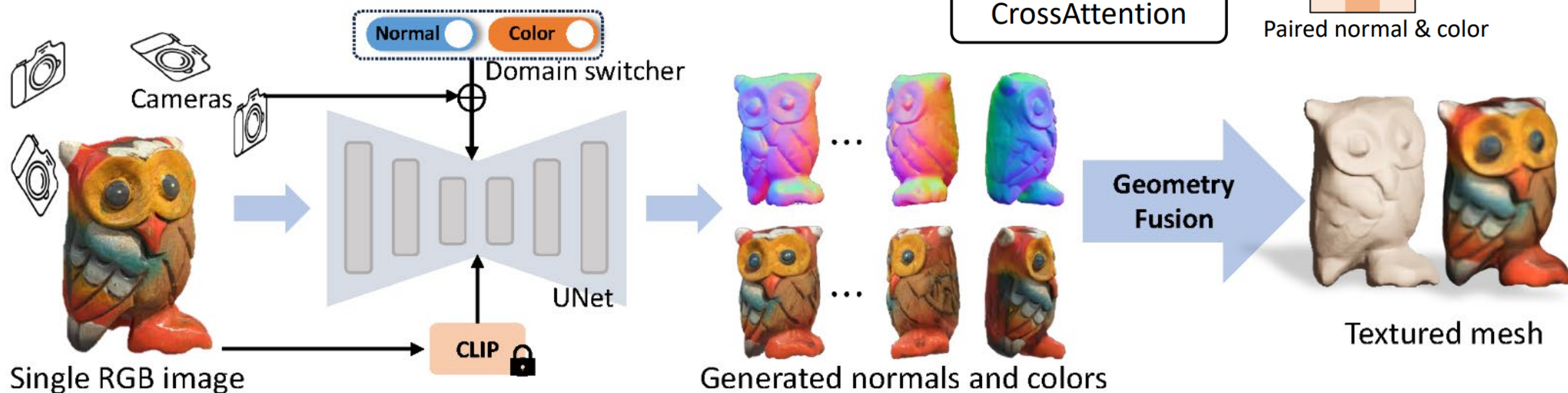
Project

Code

Gallery (New)

# Pipeline

## Cross-domain diffusion model Normal & RGB domains



Overview of Wonder3D.

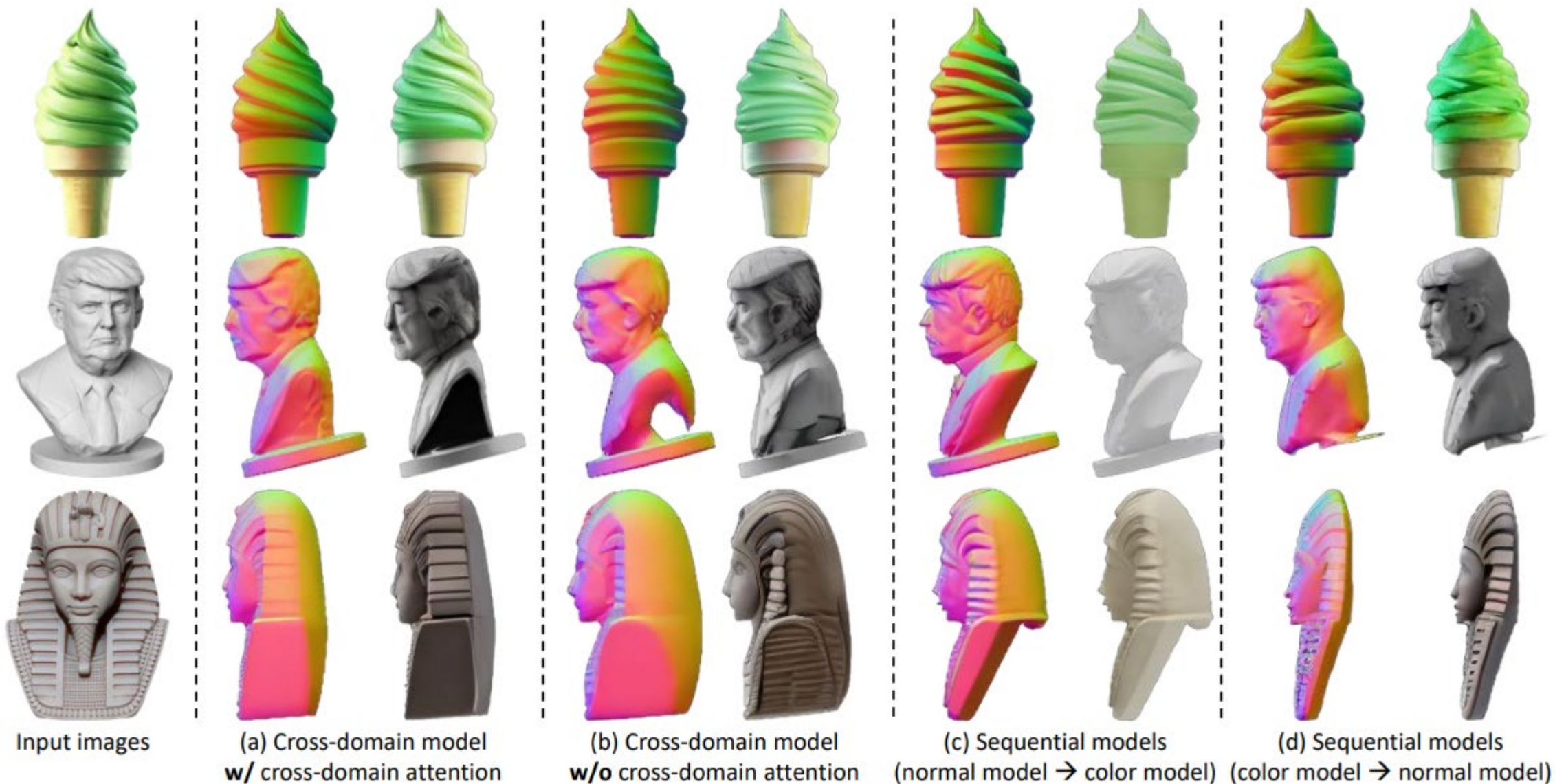


# Training details

- Objaverse LVIS subset: **30k+ objects**
- Finetuning on stable diffusion image variants for **30k steps**

**Stable diffusion model has huge  
potentials to understand the 3D world !**

# Ablation on Cross-domain models

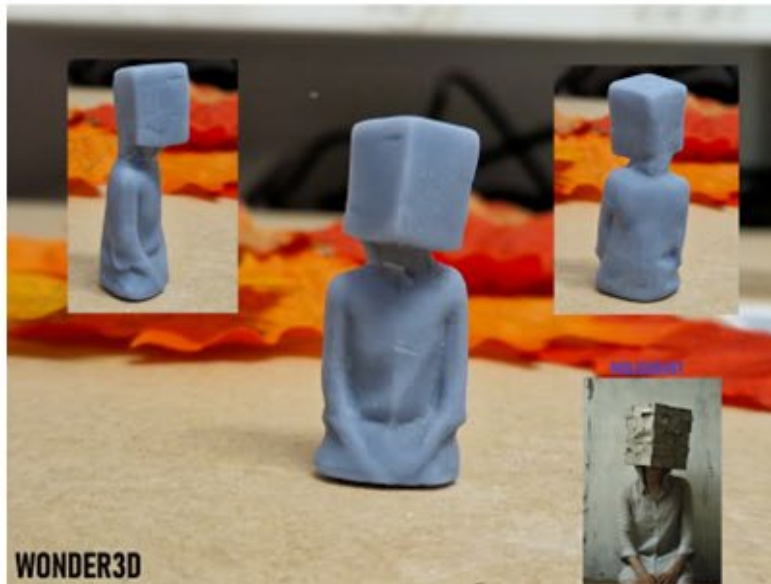








# 3D printing



# Easy to create thousands of 3D assets





# Create lego-style objects with Houdini



# Conclusions

- **Normal domain does matter.**
  - Encodes high-frequency geometries
  - Boosts the novel view synthesis of color domain
- **Fixed viewpoints leads to better view synthesis.**
  - Difficult to train a model with total random viewpoints.
- **Don't significantly modify stable diffusion model.**
  - Avoid model forgetting.
  - Keep strong generalization from SD model.

# Improvements

- Higher resolution.
- More views.
- Replace the fusion stage by Large Reconstruction Model (LRM).





龙霄潇

中国香港



扫一扫上面的二维码图案，加我为朋友。