



# 三维大模型 年度进展报告

 陈颖聪 助理教授

 人工智能学域

 香港科技大学 (广州)





# 大模型时代



视频大模型  
(Sora)



三维大模型





香港科技大学(广州)  
THE HONG KONG  
UNIVERSITY OF SCIENCE AND  
TECHNOLOGY (GUANGZHOU)

# 关键技术进展

Google

Google



Dreamfusion

Magic3D

Ponder



Zero123



UNIVERSITY OF  
CAMBRIDGE

ProlificDreamer



One-2-3-45

2022/9/29

2022/11/18

2022/12/31

2023/3/20

2023/5/25

2023/6/29



RichDreamer



LucidDreamer



One-2-3-45++



Zero123++



Ponder v2



LEAP

2023/11/28

2023/11/19

2023/11/14

2023/10/23

2023/10/12

2023/10/2



TriplaneGaussian



Point Transformer v3



Depth Anything



TripoSR



SV3D

2023/12/14

2023/12/15

2024/1/19

2024/3/4

2024/3/18



# 关键技术进展



香港科技大学(广州)  
THE HONG KONG  
UNIVERSITY OF SCIENCE AND  
TECHNOLOGY (GUANGZHOU)

Google

Google

Dreamfusion

Magic3D



THE UNIVERSITY OF  
SYDNEY

Ponder



Zero123



UNIVERSITY OF  
CAMBRIDGE

ProlificDreamer



One-2-3-45

2022/9/29

2022/11/18

2022/12/31

2023/3/20

2023/5/25

2023/6/29



RichDreamer



LucidDreamer



One-2-3-45++



Zero123++



Ponder v2



LEAP

2023/11/28

2023/11/19

2023/11/14

2023/10/23

2023/10/12

2023/10/2



TriplaneGaussian



Point Transformer v3



Depth Anything



TripoSR



SV3D

2023/12/14

2023/12/15

2024/1/19

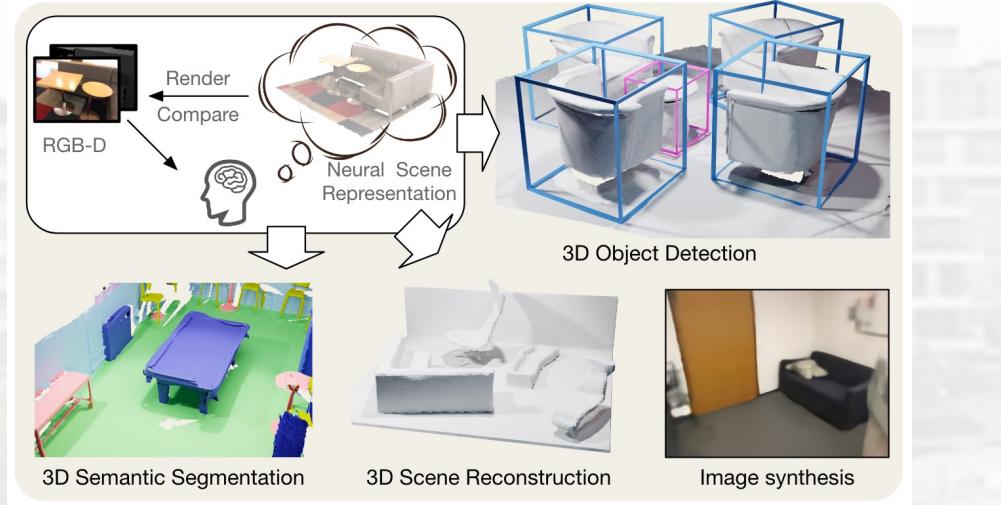
2024/3/4

2024/3/18

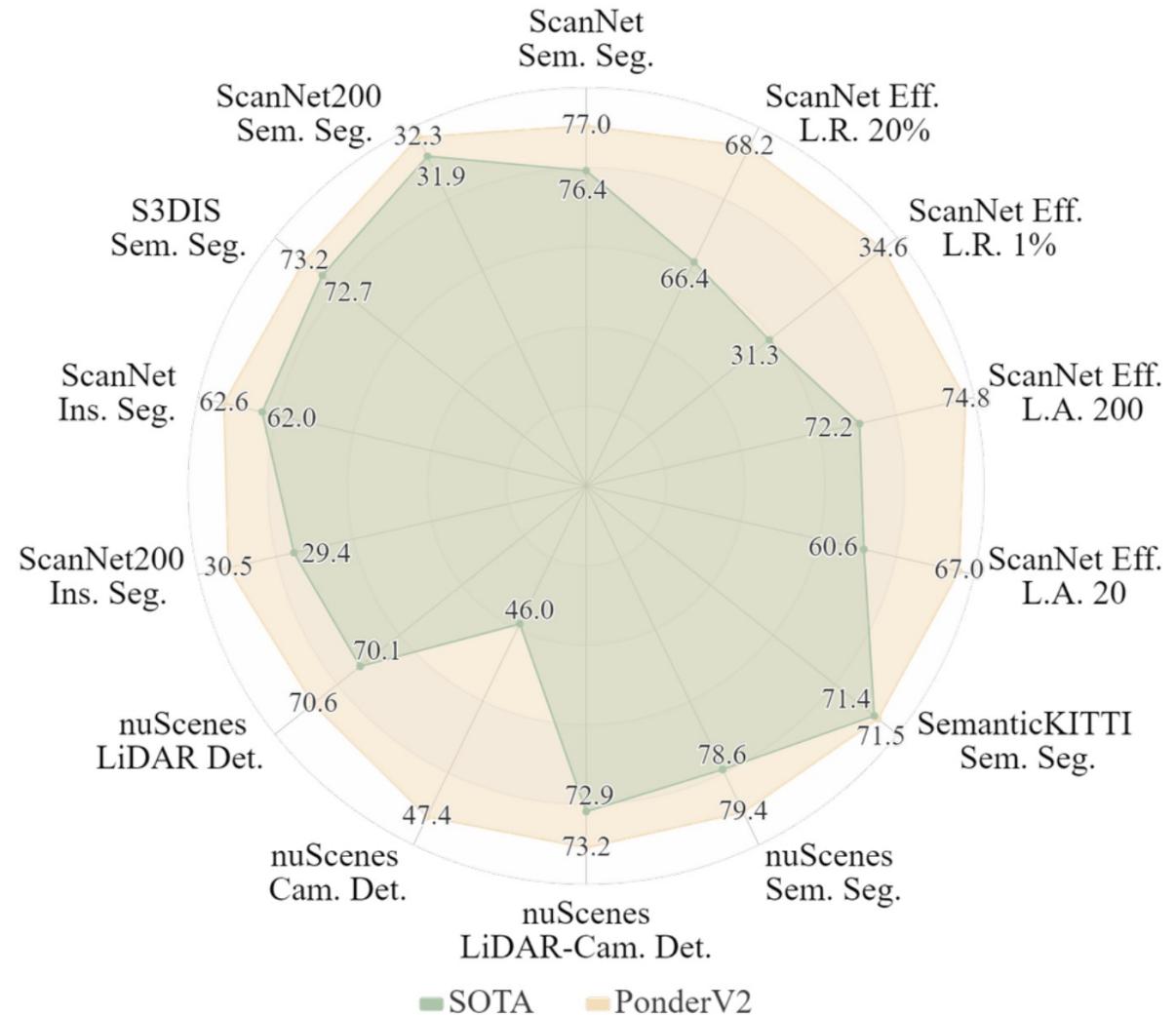


# 关键技术进展

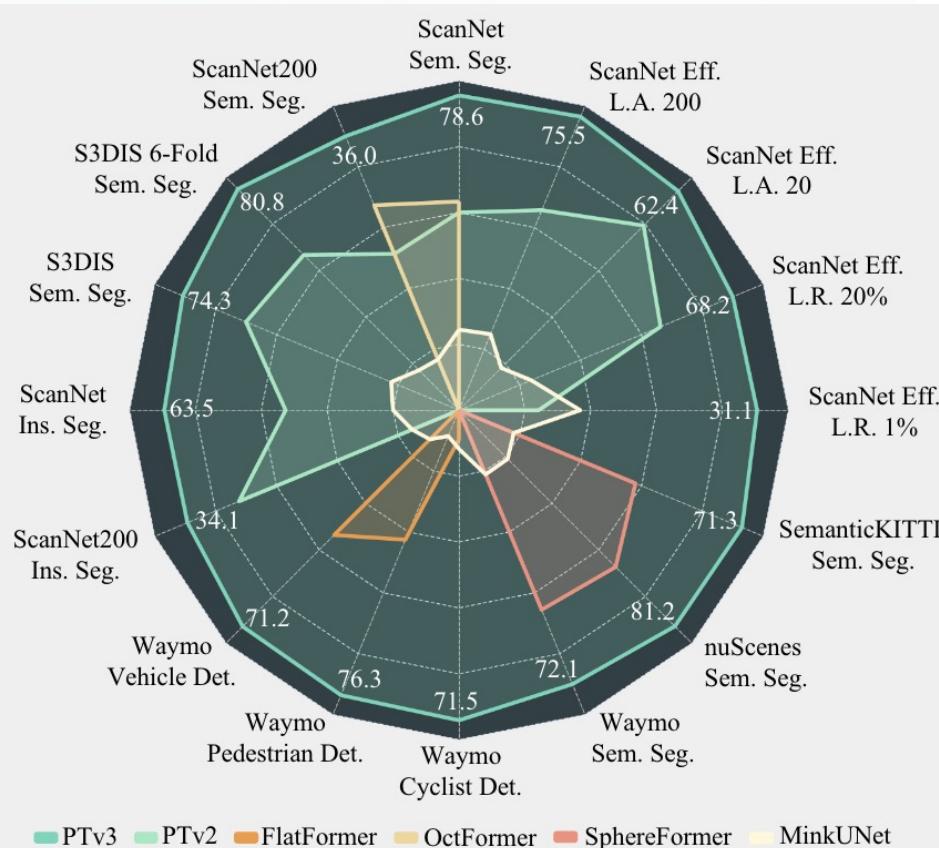
## Ponder v2



- 三维大模型预训练：Neural Rendering
- 其三维表示适配多种下游任务，包括3D检测、3D分割，3D重建、3D生成等
- 在11个涵盖室内外的基准测试中取得最佳效果



# 关键技术进展



## Point Transformer V3: Simpler, Faster, Stronger

- 在大模型时代，三维骨干网络的精度受规模化程度而非模型结构主导
- 其设计应以简单与效率为先
- 相比上一版，速度提升3.3倍，内存节省10.2倍
- 在超过20个下游任务中取得最先进的性能

	Inference Latency
MinkUNet	48ms
PTv2	146ms
PTv3	44ms ← 3.3× faster

Faster Speed

	Memory Consumption
MinkUNet	1.7G
PTv2	12.3G
PTv3	1.2G ← 10.2× lower

Lower Memory Consumption



香港科技大学(广州)  
THE HONG KONG  
UNIVERSITY OF SCIENCE AND  
TECHNOLOGY (GUANGZHOU)

# 关键技术进展

Google  
Dreamfusion

Google  
Magic3D



THE UNIVERSITY OF  
SYDNEY

Ponder



Zero123



UNIVERSITY OF  
CAMBRIDGE

ProlificDreamer



One-2-3-45

2022/9/29

2022/11/18

2022/12/31

2023/3/20

2023/5/25

2023/6/29



RichDreamer



香港科技大学(广州)  
THE HONG KONG  
UNIVERSITY OF SCIENCE AND  
TECHNOLOGY (GUANGZHOU)

LucidDreamer



One-2-3-45++



Zero123++



Ponder v2



LEAP

2023/11/28

2023/11/19

2023/11/14

2023/10/23

2023/10/12

2023/10/2



TriplaneGaussian



Point Transformer v3



Depth Anything



TripoSR



SV3D

2023/12/14

2023/12/15

2024/1/19

2024/3/4

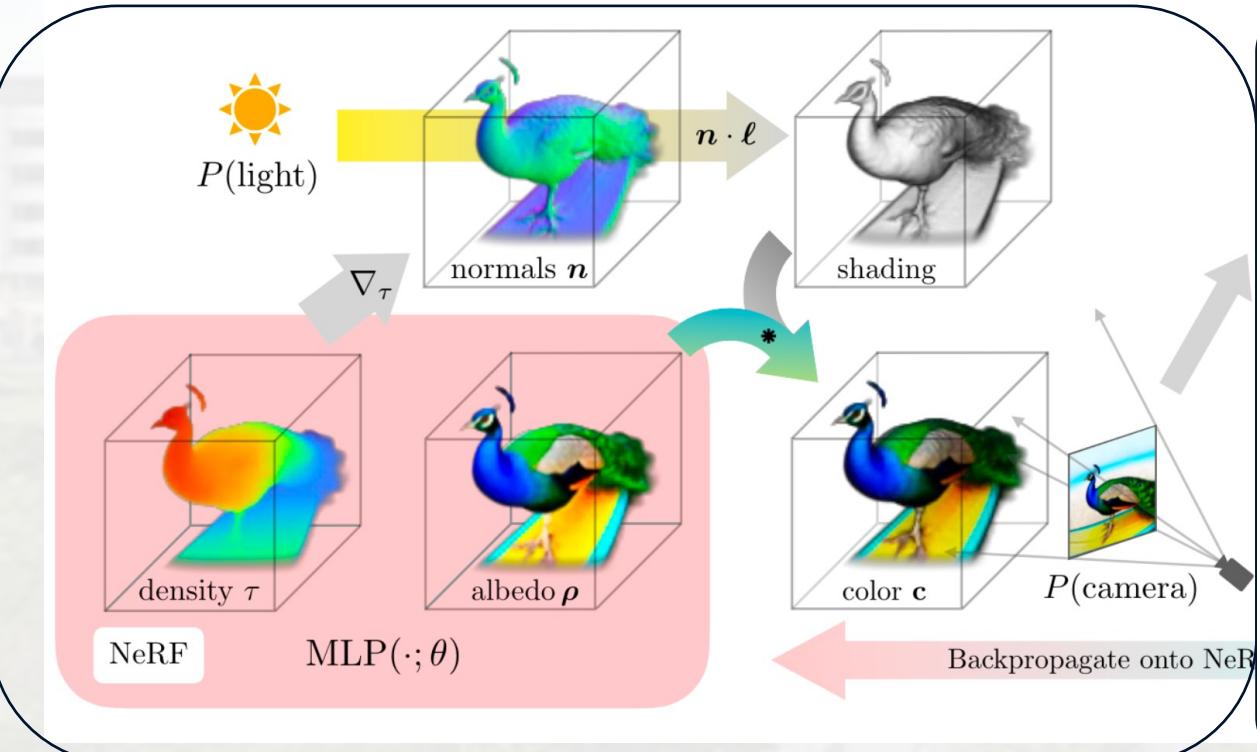
2024/3/18



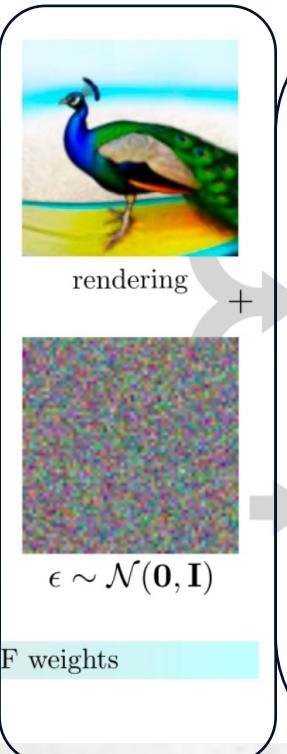
# 关键技术进展

## DreamFusion: Text-to-3D using 2D Diffusion

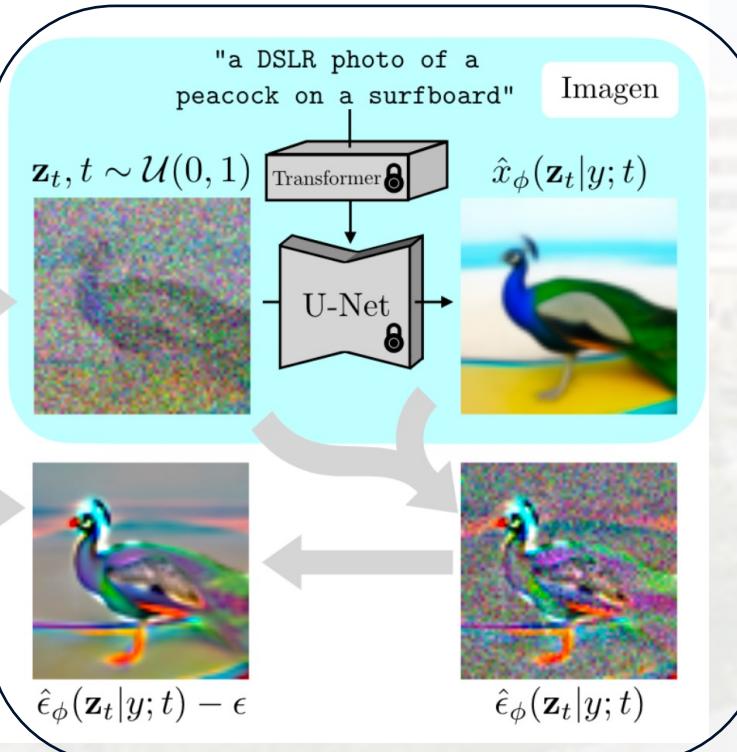
### 三维表示



### 蒸馏



### 2D Diffusion



将二维大模型的知识蒸馏到三维表示中，生成三维模型，点燃了社区对三维模型生成的兴趣。



香港科技大学(广州)  
THE HONG KONG  
UNIVERSITY OF SCIENCE AND  
TECHNOLOGY (GUANGZHOU)

# 关键技术进展

Google

Dreamfusion

Google

Magic3D



THE UNIVERSITY OF  
SYDNEY



Zero123



UNIVERSITY OF  
CAMBRIDGE

ProlificDreamer



One-2-3-45

2022/9/29



RichDreamer

2022/11/18



LucidDreamer

2022/12/31



One-2-3-45++

2023/3/20



Zero123++

2023/5/25



Ponder v2

2023/6/29



LEAP

2023/11/28



TriplaneGaussian

2023/11/19



Point Transformer v3

2023/11/14



Depth Anything

2023/10/23



TripoSR

2023/12/14

2023/12/15

2024/1/19

2024/3/4

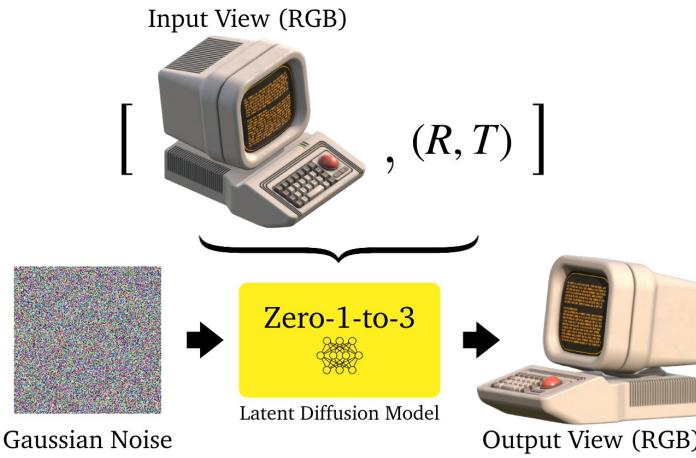
2024/3/18



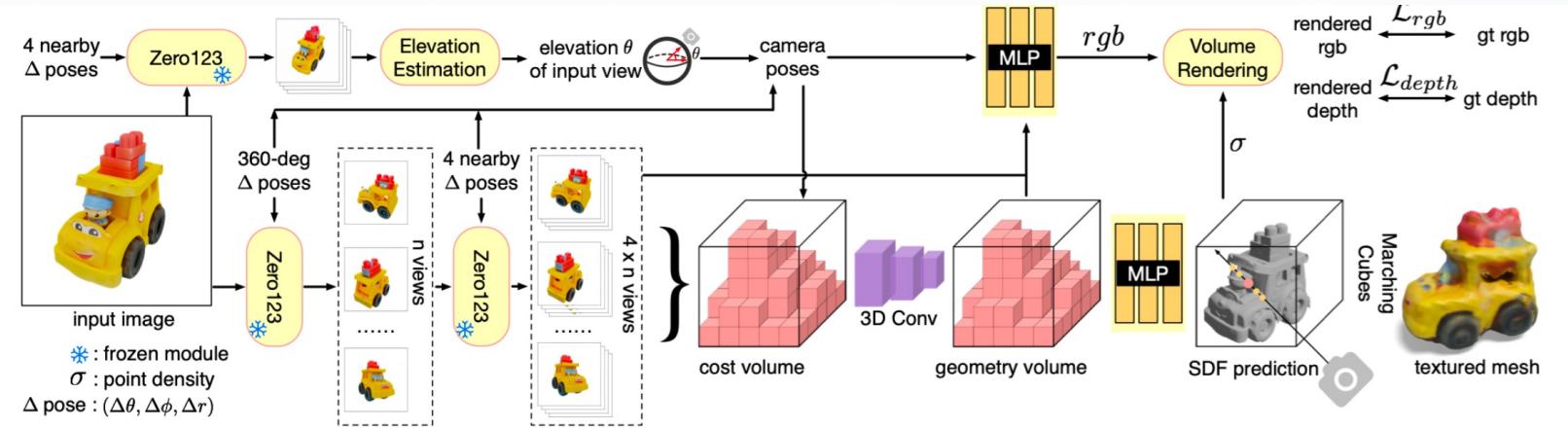
SV3D

# 关键技术进展

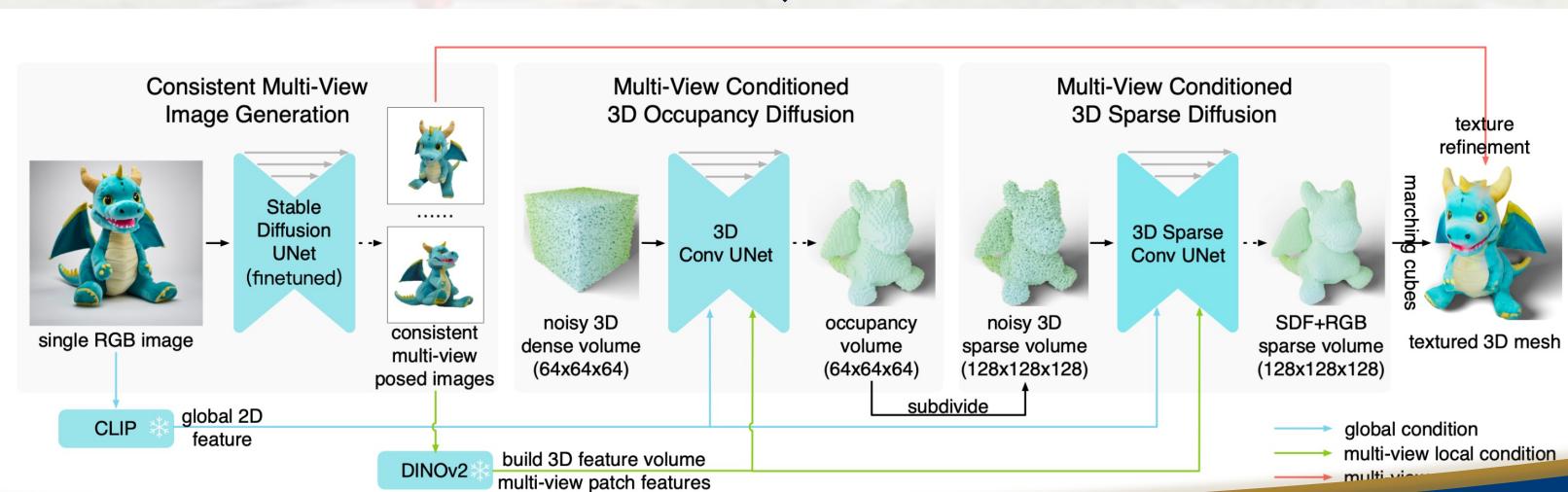
## Zero-1-to-3



## One-2-3-45



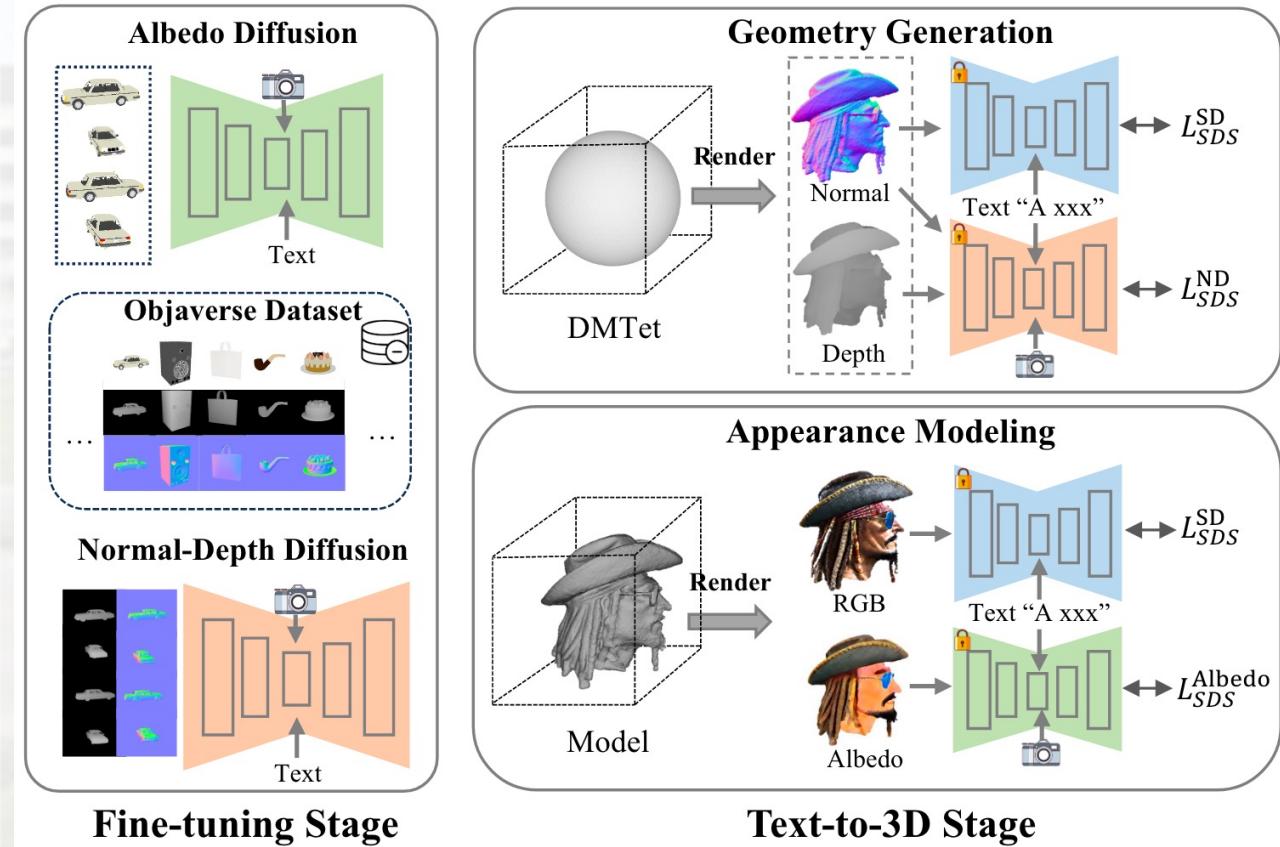
## One-2-3-45 ++



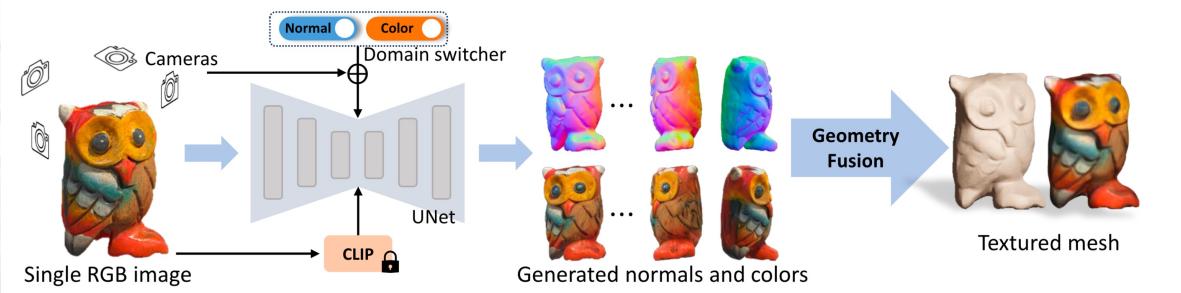
- Zero-1-to-3：扩散模型可直接输出不同视图图像
- One-2-3-45 (++)
  - 提升多视角一致性（性能）
  - 专用可泛化重建模型（速度）

# 关键技术进展

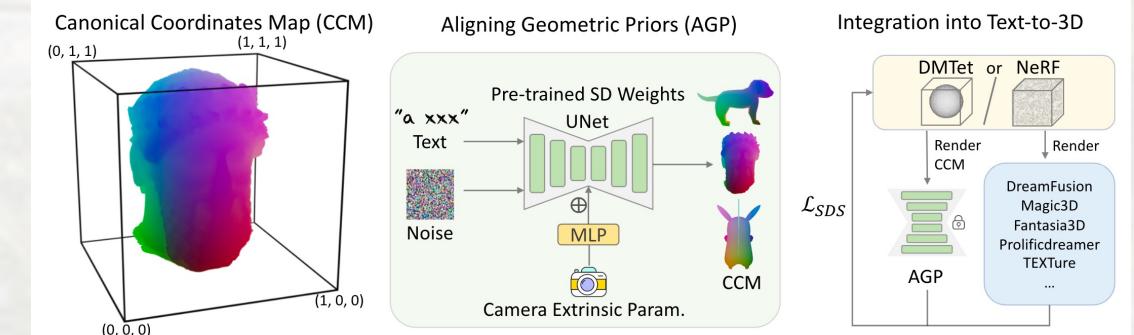
## RichDreamer



## Wonder3D



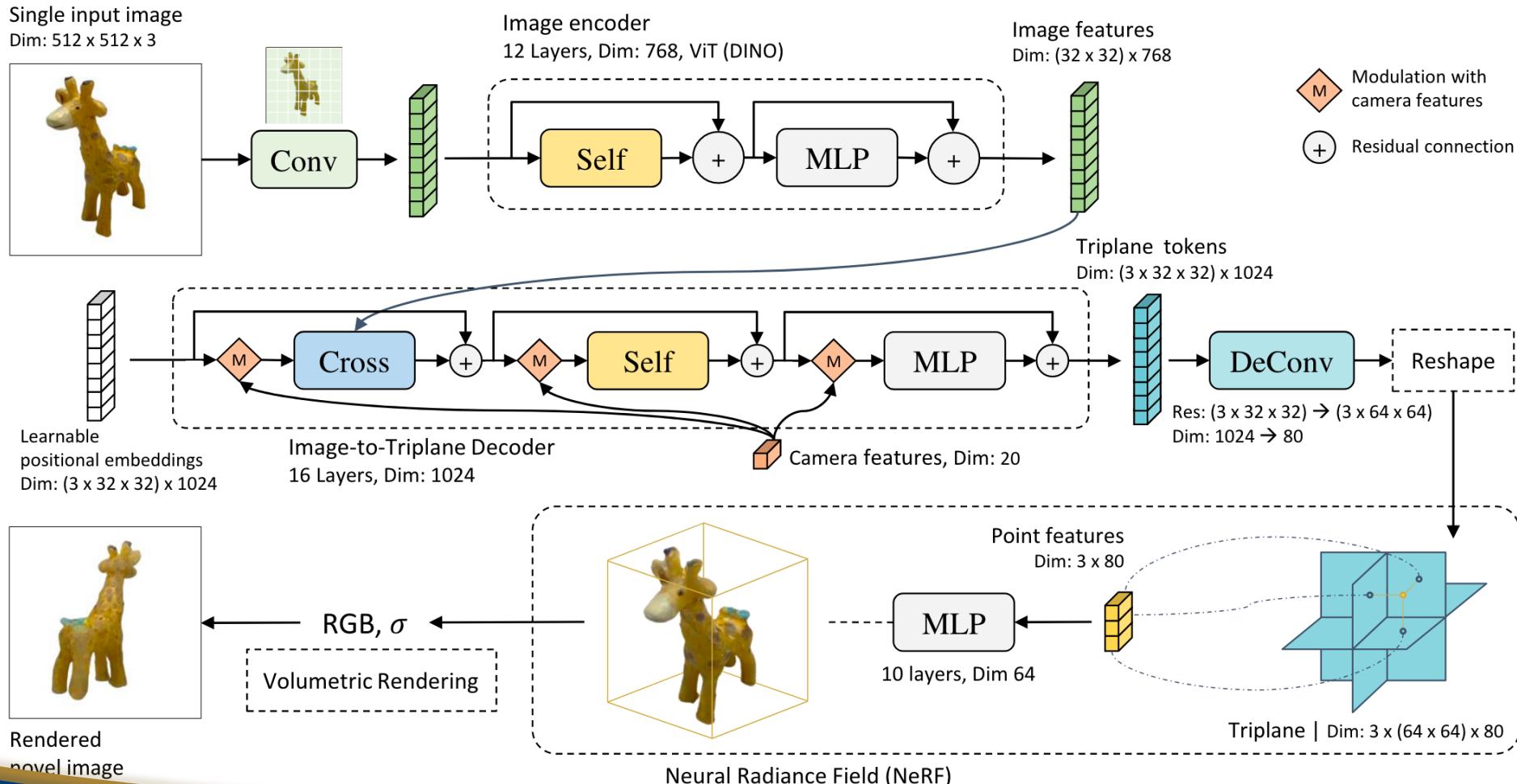
## SweetDreamer



在扩散模型中嵌入三维先验信息（如normal, depth, CCM等）有助于提升几何一致性。

# 关键技术进展

## LRM: Large Reconstruction Model for Single Image to 3D

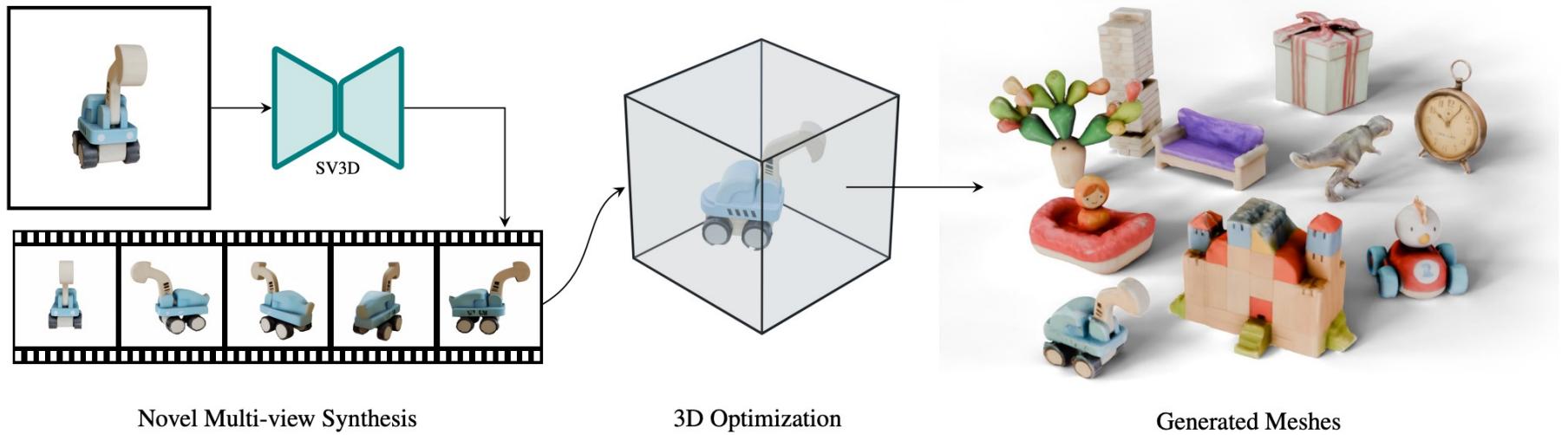


文字大模型的成功  
取决于三大要素：

- Transformer
- 大数据集
- 自监督

这三大要素能否复  
制到三维大模型中？

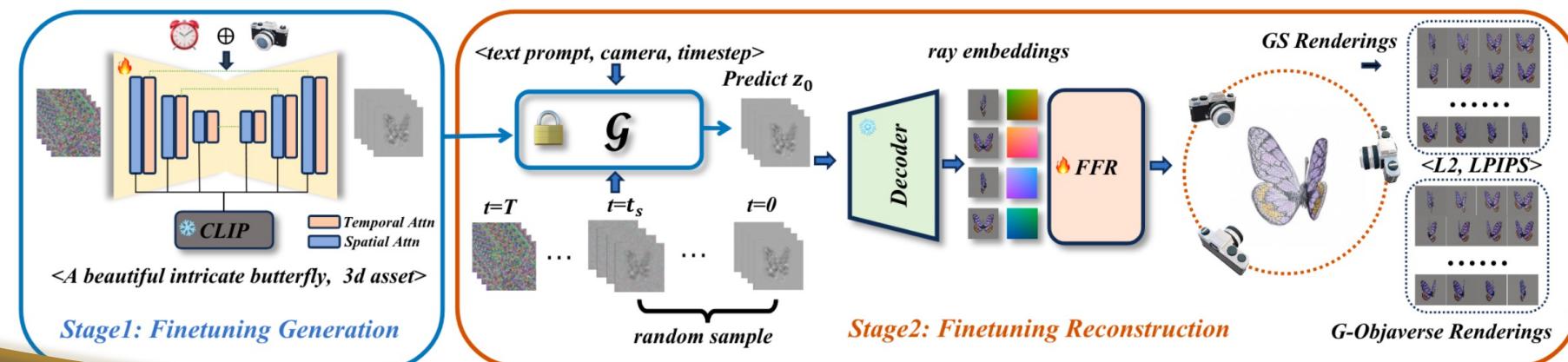
# 关键技术进展



**SV3D:** Novel Multi-view Synthesis and 3D Generation from a Single Image using Latent Video Diffusion

基于视频大模型的3D生成

帧间一致性≈多视图一致性



**VideoMV:** Consistent Multi-View Generation Based on Large Video Generative Model

# 商业化实践

