

3D Generative Models

Mid-term Project Presentation

2025.11.10

Minseo Park, Jewoo Shin, Sangmin Lee

Team 2

Table of Contents

- **Introduction**
- **Related Work**
- **Limitations**
- **Proposal**
- **Teammate Roles**

체크리스트

(나중에 지우기) ⚠ (not finished) ✓ (finished) ✗ (empty)

Mid-term project presentation Expectations

- Introduce problems and explain why it is important -> 3D generation을 problem이자 task로 소개, 왜 중요한가? 3D asset의 사용처 (디자인, 가구배치, 게임 캐릭터, 3D 애니메이션 등등)들과 3D interaction이 필요한 분야들 (자율주행, motion transfer, motion 추출 등등) - ✓
- Give an overall idea on the related work - ✓
- Explain what problems those existing techniques have - ✓
- (Optional) explain how you can address those problems - ⚠
- Explain roles of each member - ✓

Introduction

Problem Definition: 3D Generation

- **Goal:** From diverse inputs (text, single/few images, video, sketches), automatically **create or reconstruct editable 3D assets/scenes** with coherent geometry and materials.
- **Scope:** Covers **single feed-forward models** (e.g., Trellis, SPARC3D), **optimization/distillation** pipelines (e.g., SDS/RF families), and **3D Representation models** (triplanes, 3DGS, meshes, SDFs).
- **Key quality criteria:** Multi-view consistency, geometric detail, material/lighting realism, controllability (pose/camera/style), diversity, speed/throughput, and animation/rigging readiness.
- **Core challenges:** 2D→3D ambiguity, scarce 3D data, high compute, representation trade-offs (NeRF/3DGS/mesh/SDF), temporal/pose consistency

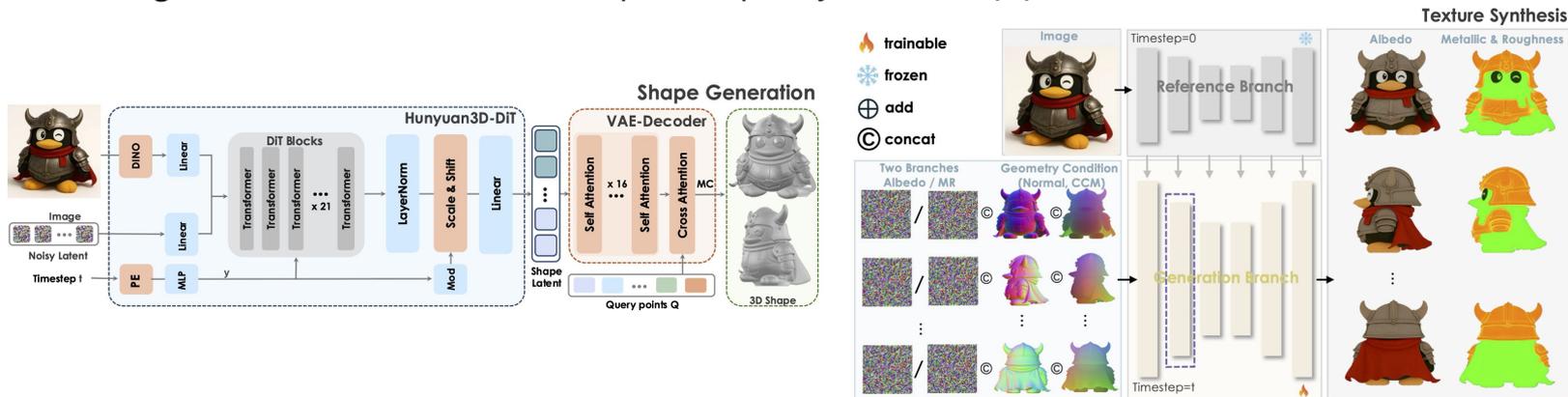
Why 3D Generation is Important?

- **Content creation:** Product & industrial design, interior/furniture layout, games (characters/environments), 3D animation/VFX, AR/VR/MR, digital twins, e-commerce visualization.
- **3D interaction domains:** Autonomous driving & robotics simulation, motion capture/extraction, motion transfer/retargeting, telepresence, medical planning, embodied AI.
- **Impact:** Dramatically lowers cost and turnaround time, speeds iteration and exploration, enables large-scale customization, and improves quality across 3D pipelines.



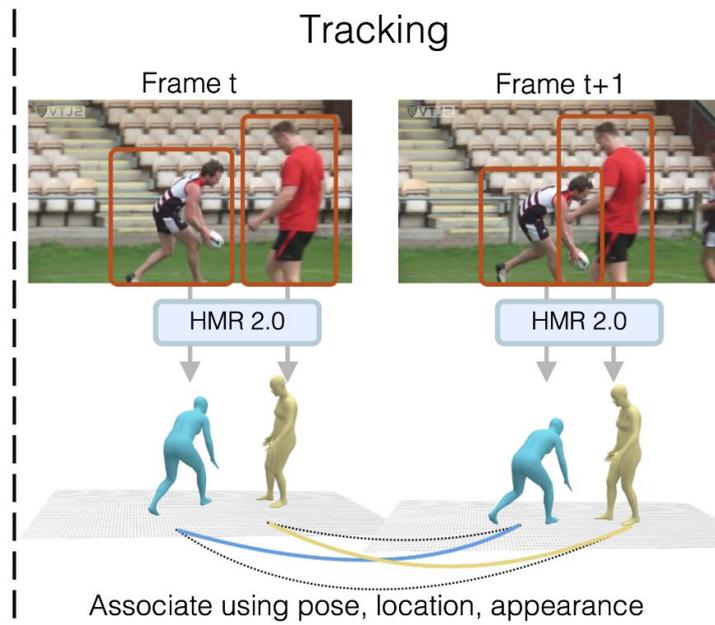
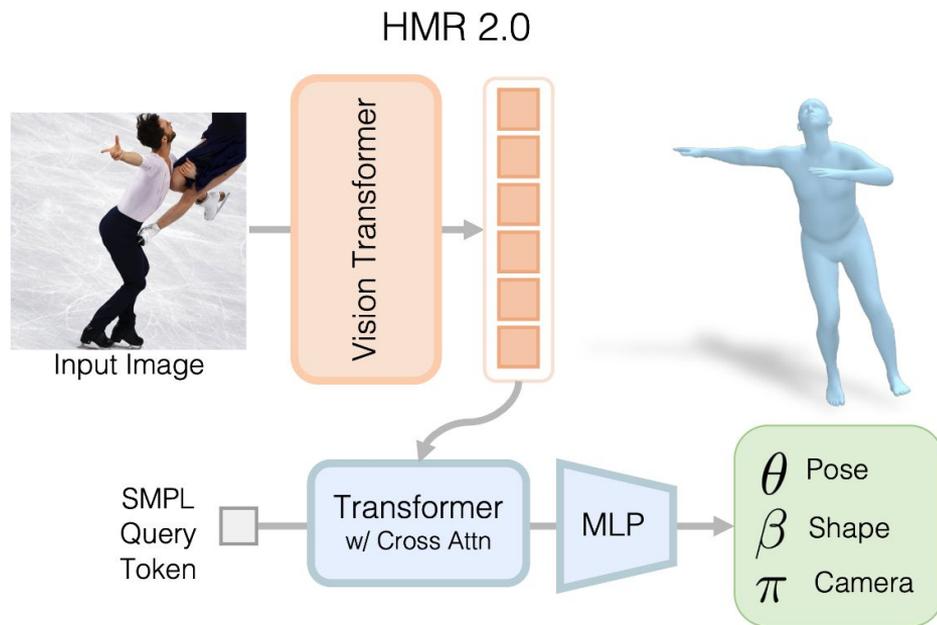
Why 3D Generation is Important?

- **Content creation:** Product & industrial design, interior/furniture layout, games (characters/environments), 3D animation/VFX, AR/VR/MR, digital twins, e-commerce visualization.
- **3D interaction domains:** Autonomous driving & robotics simulation, motion capture/extraction, motion transfer/retargeting, telepresence, medical planning, embodied AI.
- **Impact:** Dramatically lowers cost and turnaround time, speeds iteration and exploration, enables large-scale customization, and improves quality across 3D pipelines.



Shape Generation + Texture Synthesis

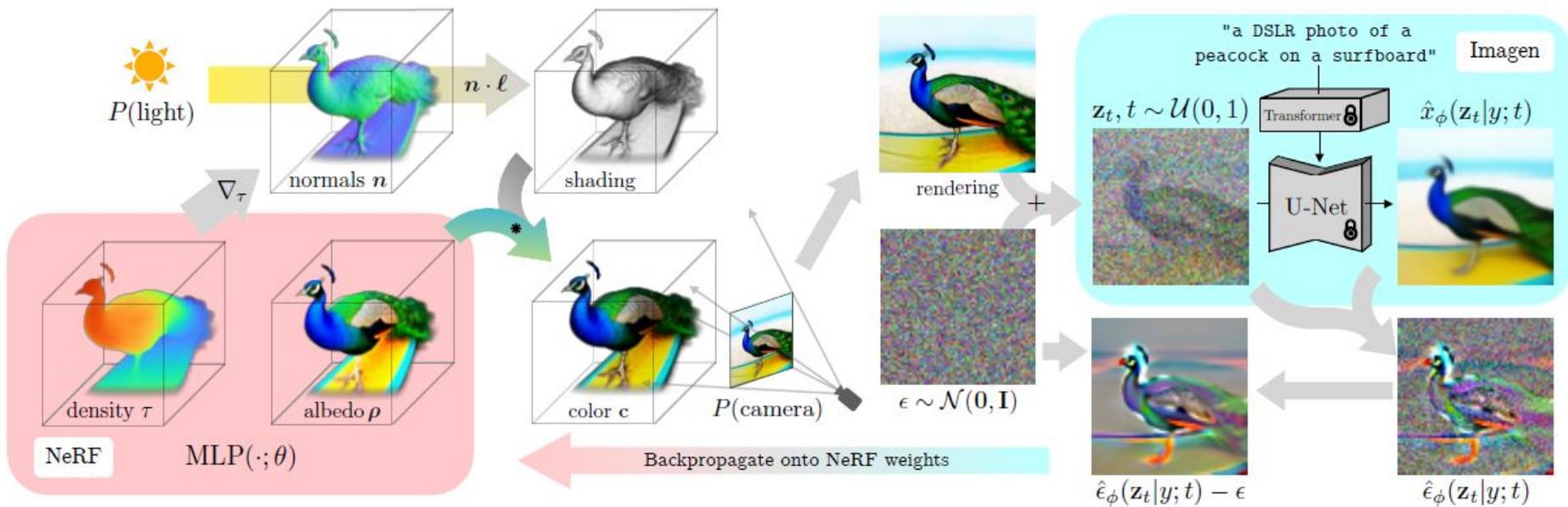
3D Generation



Human Mesh Recovery

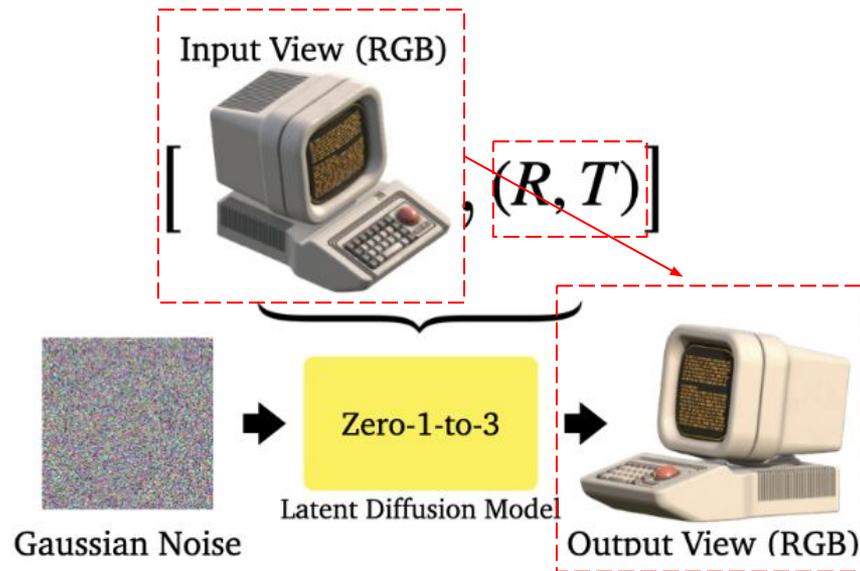
Related Work

DreamFusion



Single-view Image Generation Model based Distillation

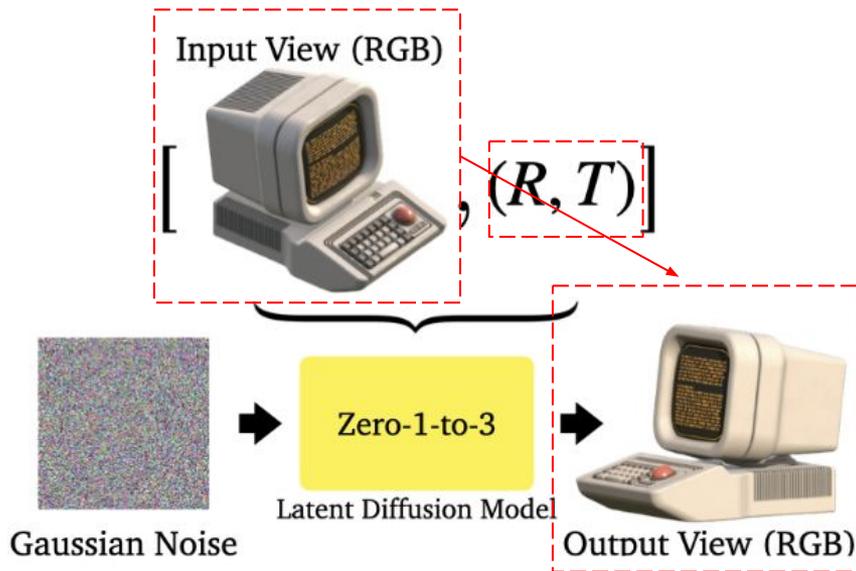
Zero-1-to-3: Zero-shot One Image to 3D Object



Novel View Synthesis

Multi-view Image Generation Model based Distillation

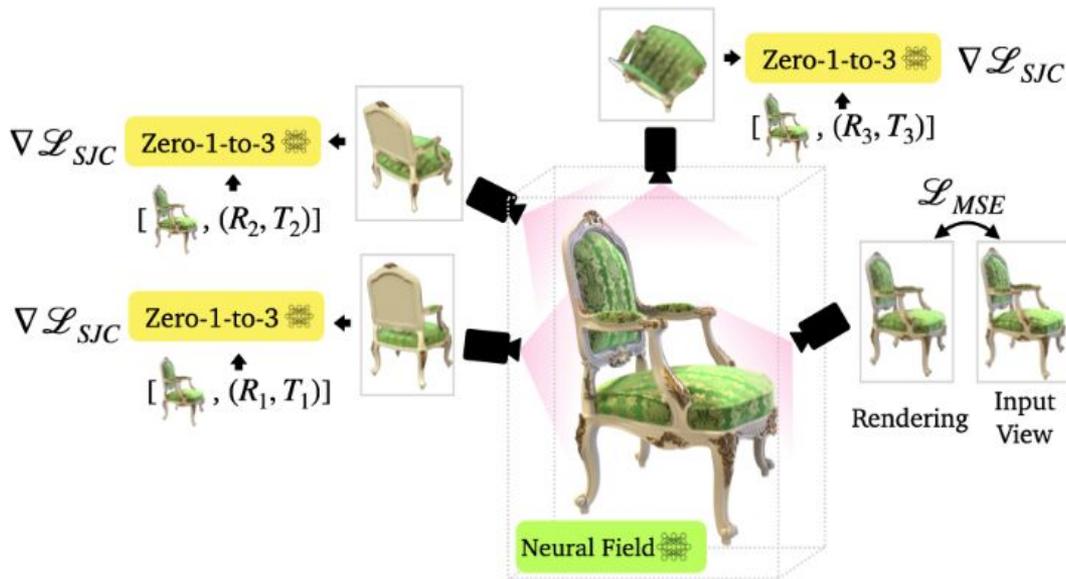
Zero-1-to-3: Zero-shot One Image to 3D Object



Novel View Synthesis

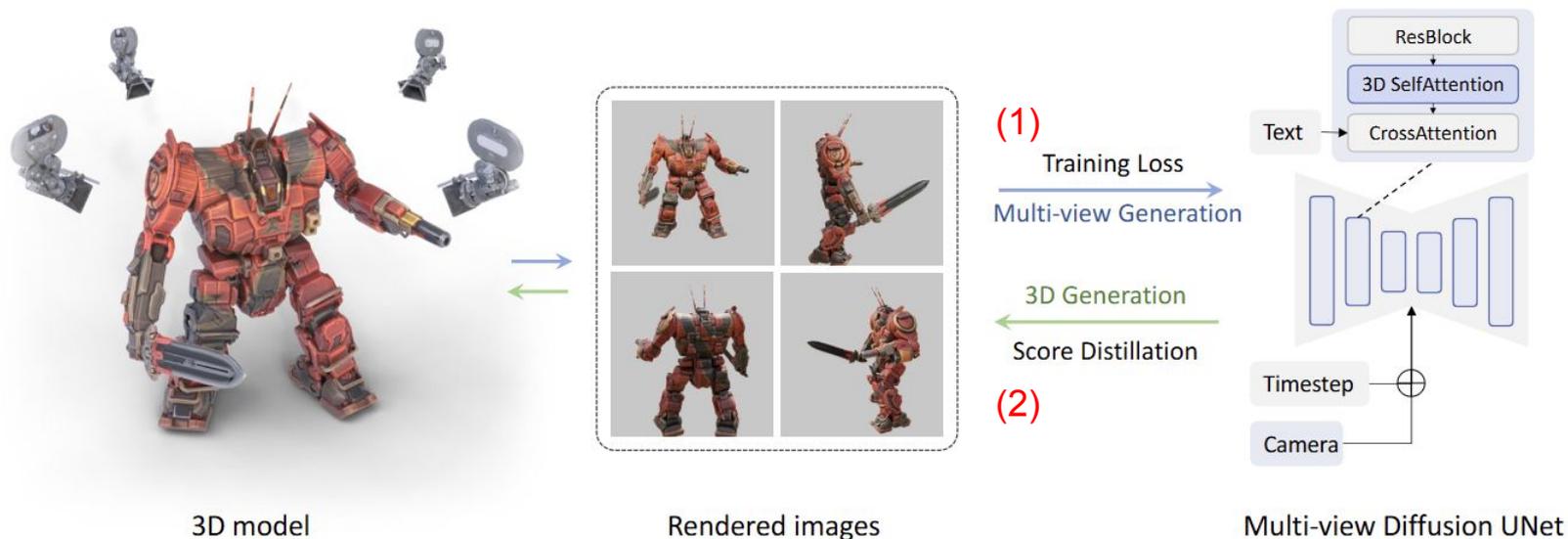
Train view-conditioned diffusion model

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



3D Reconstruction

MVDream: Multi-view Diffusion for 3D Generation



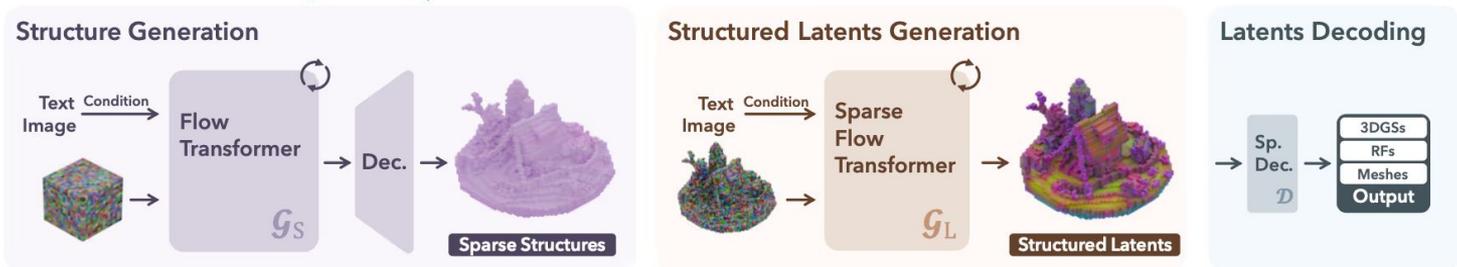
Multi-view Image Generation Model based Distillation

3D feed-forward models

3D Assets Encoding & Decoding



3D Assets Generation

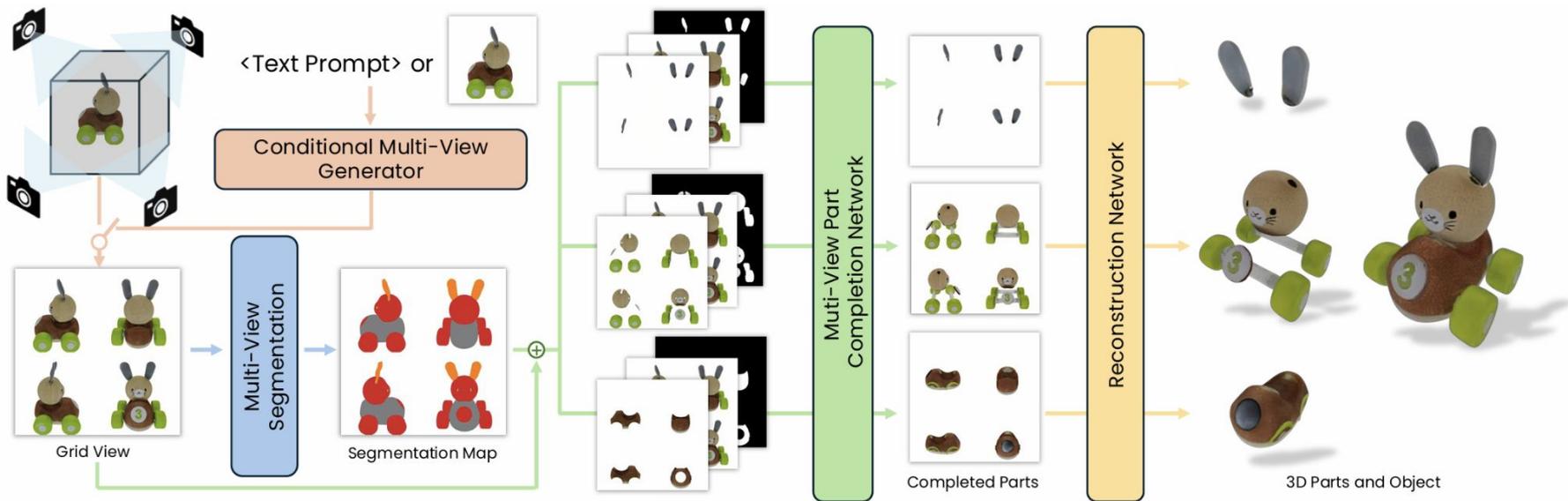


3D Neural Representation + Shape Generative model

Part-based approach

PartGen: Part-level 3D Generation and Reconstruction with Multi-View Diffusion Models (CVPR 2025)

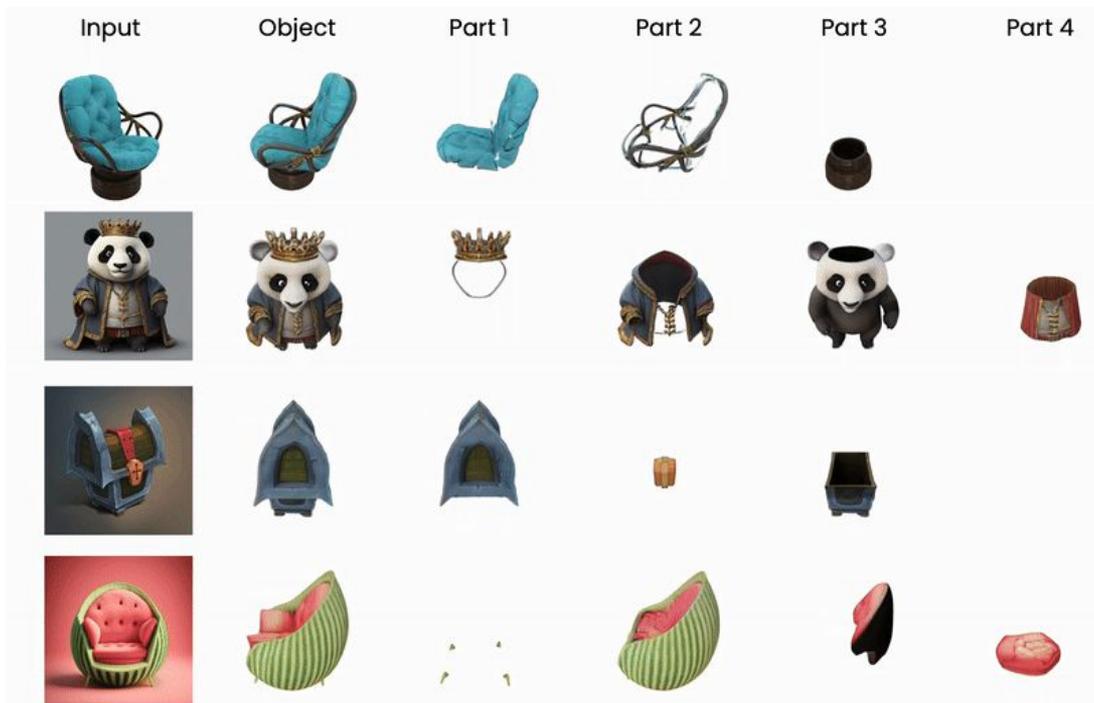
→ Part segmentation applied to multi-view diffusion models



Part-based approach

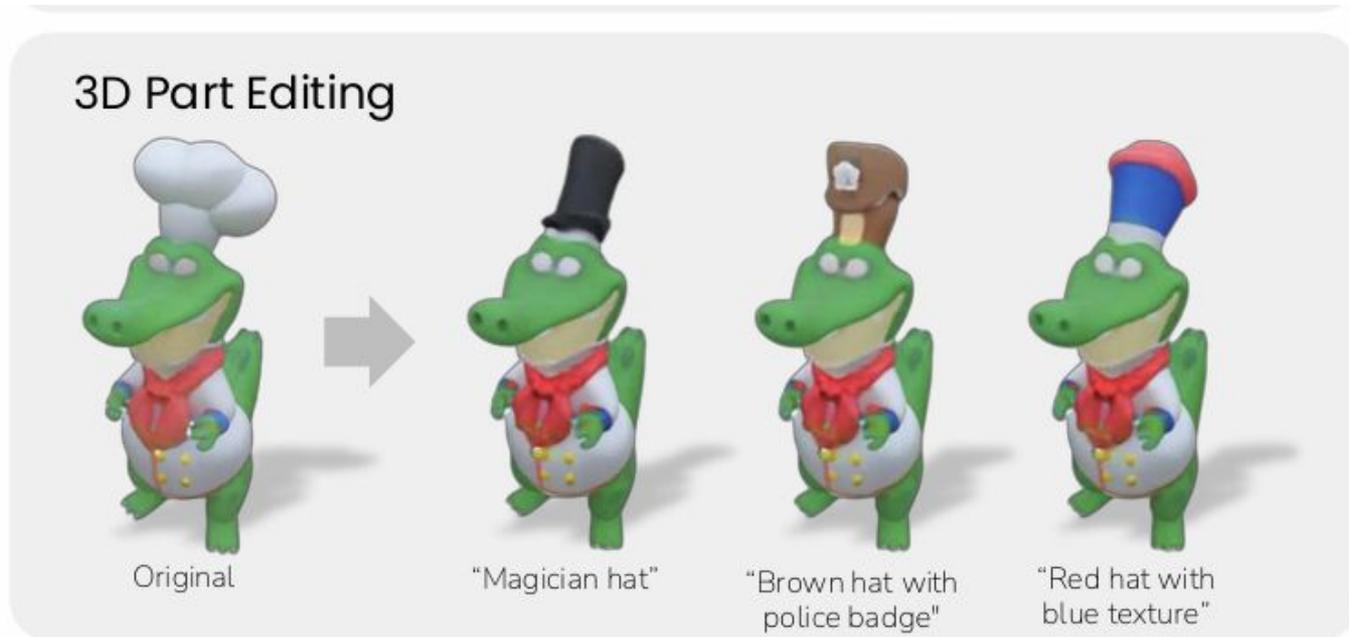
PartGen: Part-level 3D Generation and Reconstruction with Multi-View Diffusion Models (CVPR 2025)

→ Part segmentation applied to multi-view diffusion models



Part-based approach

PartGen: Part-level 3D Generation and Reconstruction with Multi-View Diffusion Models (CVPR 2025)
→ Part segmentation applied to multi-view diffusion models



Limitations of Previous Works

Limitations

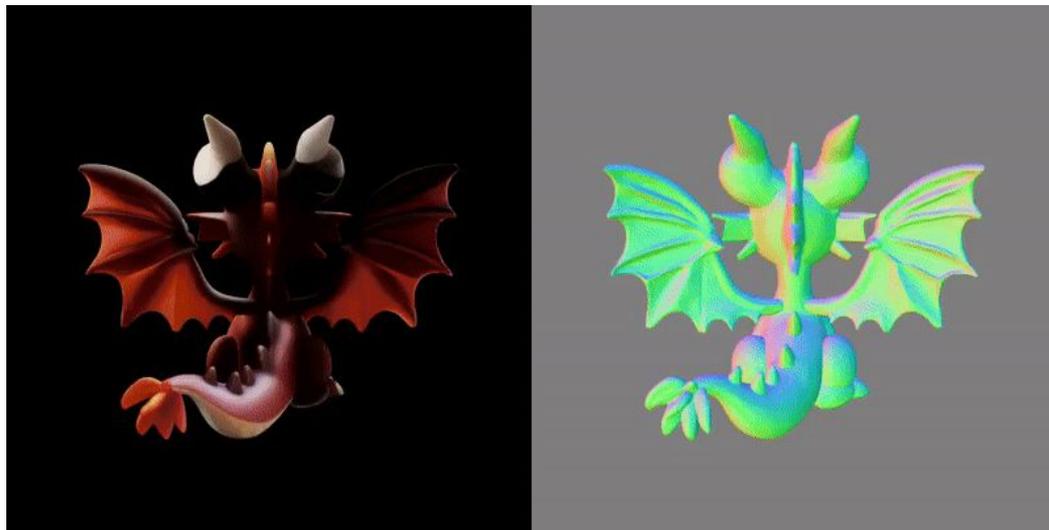
- **Problems of (Multi-view) 2D Diffusion-based models:**
 - Low 3D Geometric Consistency & “Janus Problem” - ⚠️
 - Slow generation speed & High computational cost - ⚠️
 - Lack of fine-grained details - ⚠️
- **3D feed-forward models:**
 - Low 3D Geometric Consistency & “Janus Problem” - ✅
 - Slow generation speed & High computational cost - ✅
 - Scarce High-quality 3D Datasets - ⚠️
 - Lack of fine-grained details - ⚠️

Limitations

2D image of Dragon:



3D geometry generated by TRELIS:

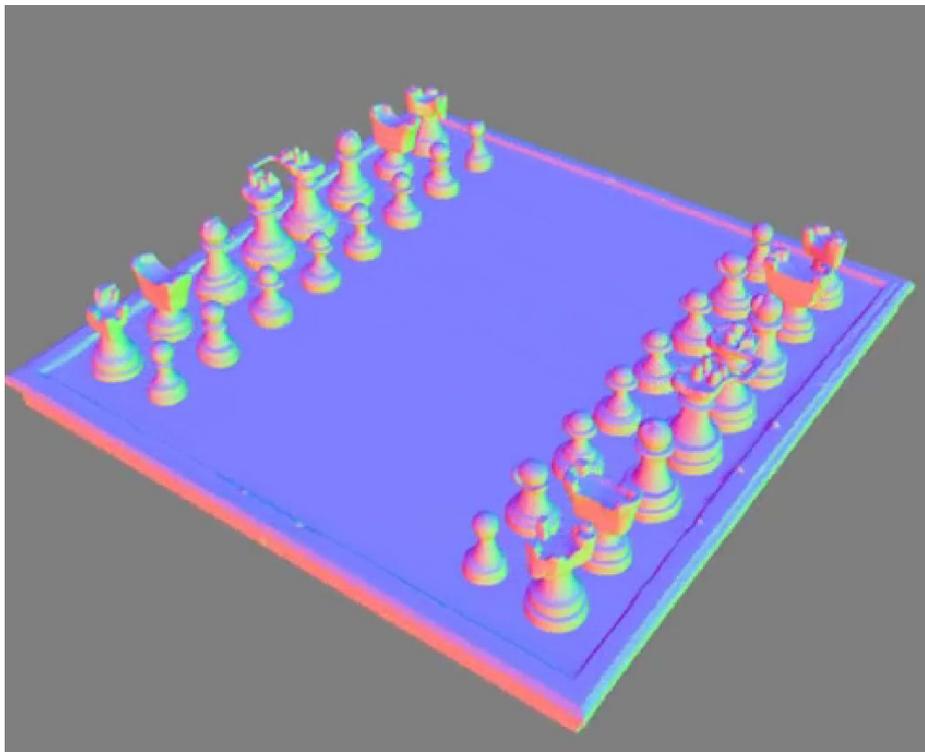


Limitations

2D image of chessboard:

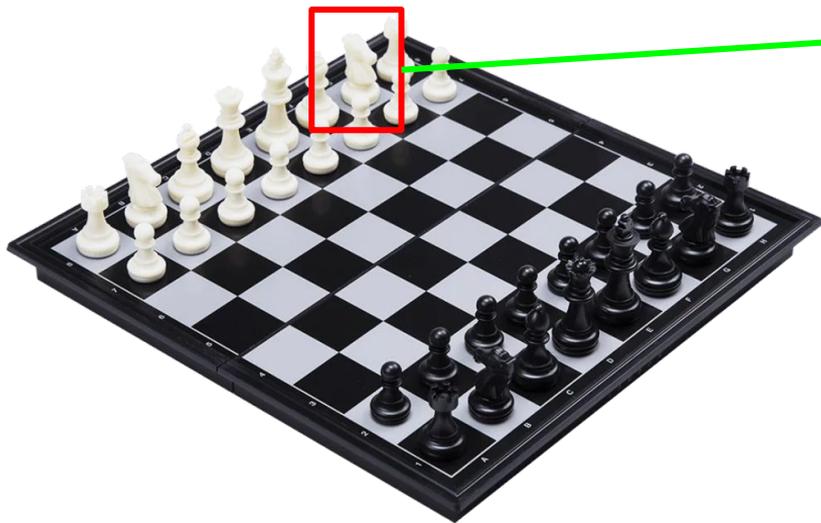


3D geometry generated by TRELIS: ⚠

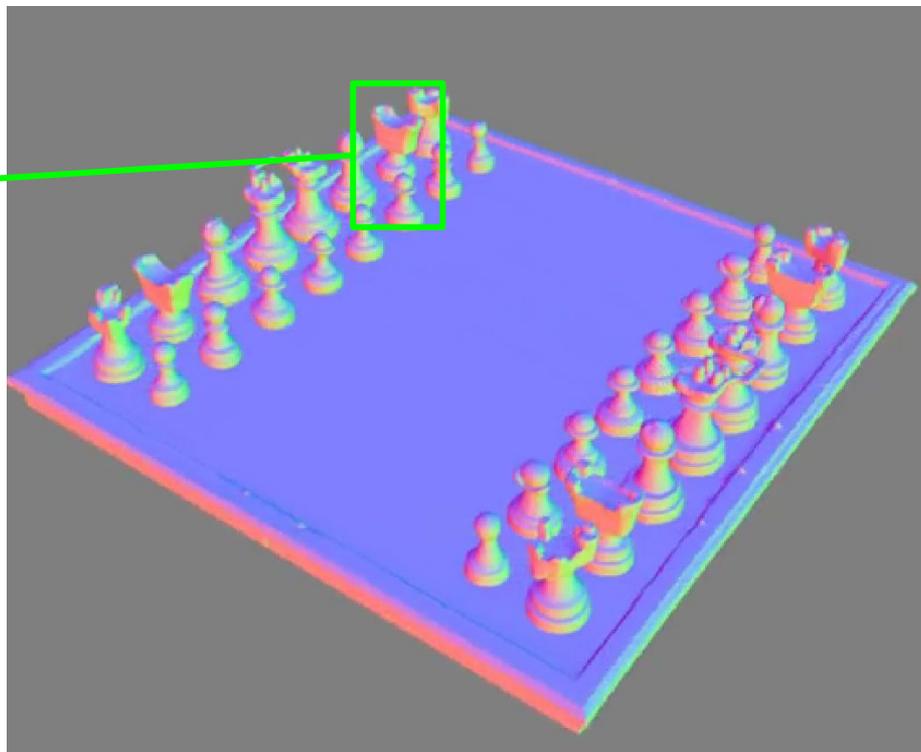


Limitations

2D image of chessboard:

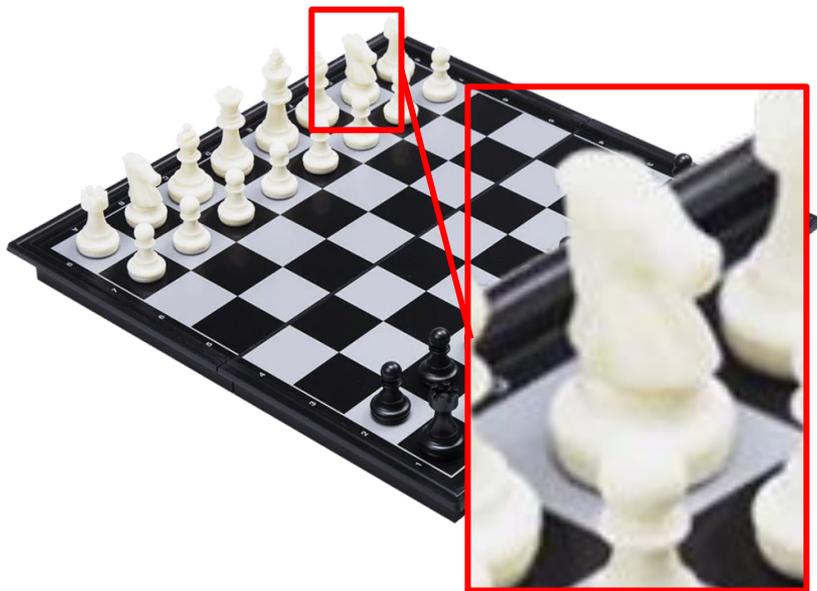


3D geometry generated by TRELIS: ⚠

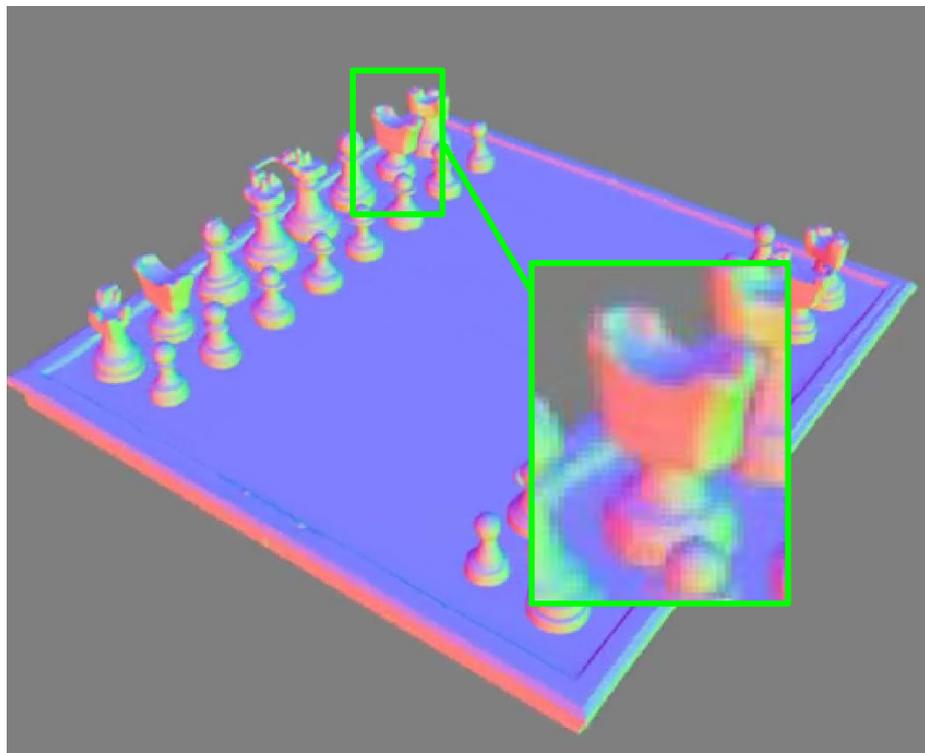


Limitations

2D image of chessboard:



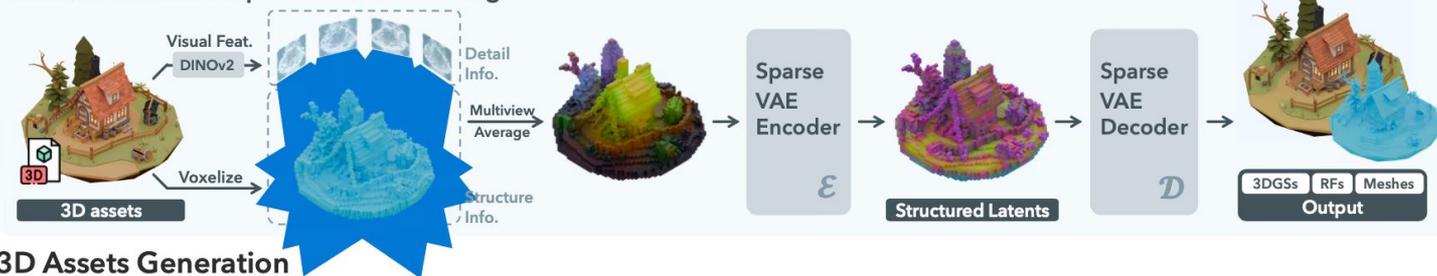
3D geometry generated by TRELIS: ⚠



Limitations

3D Assets Encoding & Decoding

Structured Latent Representation Learning



3D Assets Generation

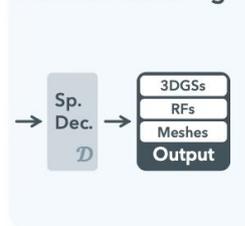
Structure Generation



Structured Latents Generation



Latents Decoding



Detail loss from Compression & Decoding latents

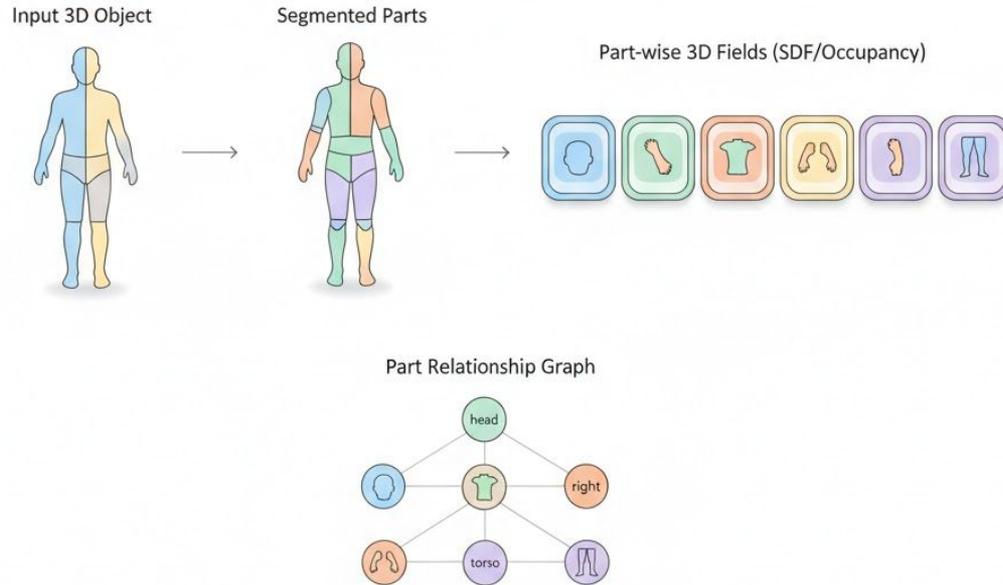
Proposal

Our Problem: Part-Level 3D Generation

- **Input:** Single Image or Few Images
- **Goal: 3D generation** with **Part-level** correlation, with single feedforward model
- **Dataset:** large-scale 3D part dataset named **Partverse** from Objaverse
- **Key quality criteria:** Multi-view consistency, geometric detail, material/lighting realism, controllability (pose/camera/style), diversity, speed/throughput, and Whole aware/Part aware quality
- **Why Part-Level 3D Generation Important?** Editability, High geometric details, etc



Main Idea



Segment 2D image into parts, generate each part with an independent 3D field, connect them via a part relationship graph, enabling part-consistent geometry & multi-part generation

Datasets

PartVerse Dataset

12K objects

91K parts

Part priority sort

Texture & geometry

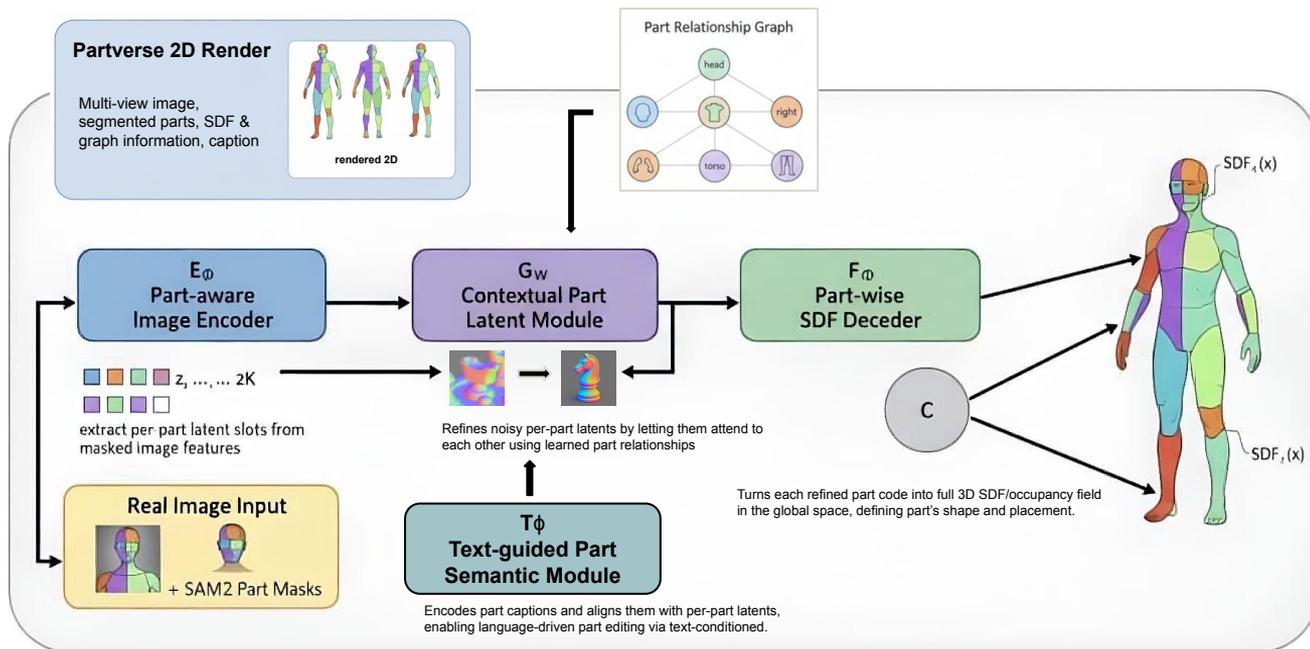
Part-level text caption

Brief: A camera lens from a Sony video camera.
Detailed: The camera lens is a circular component with a multi-faceted, faceted design, featuring shades of gray and black. It appears to be made of glass or a similar transparent material, with a metallic frame surrounding it.

Brief: A blue Sony logo panel from a camcorder.
Detailed: The blue Sony logo panel is a rectangular component with a smooth, glossy finish and golden-colored 'Sony' text embossed on it. It is part of the camcorder's exterior design, serving as a brand identifier.



Model



Teammate Roles

Roles of each member

- **Minseo Park: Prepare presentation, Investigate project topic(Read paper, Run codes for failure cases), Dataset Collection, Code Implementation**
- **Jewoo Shin: Prepare presentation, Investigate project topic(Read paper, Run codes for failure cases), Design Generative models, Code Implementation**
- **Sangmin Lee: Prepare presentation, Investigate project topic(Read paper, Run codes for failure cases), Result Analysis, Code Implementation**

Thank you