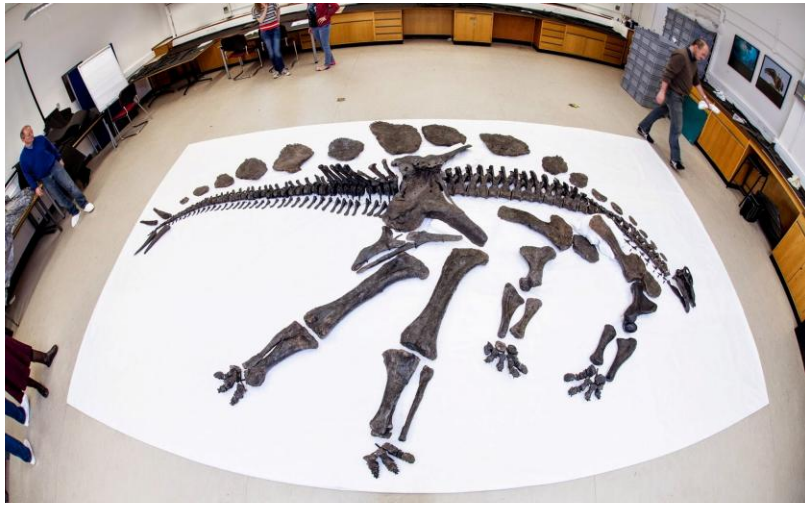
Can 3D Generative Models Help 3D Assembly?

ICML 2026

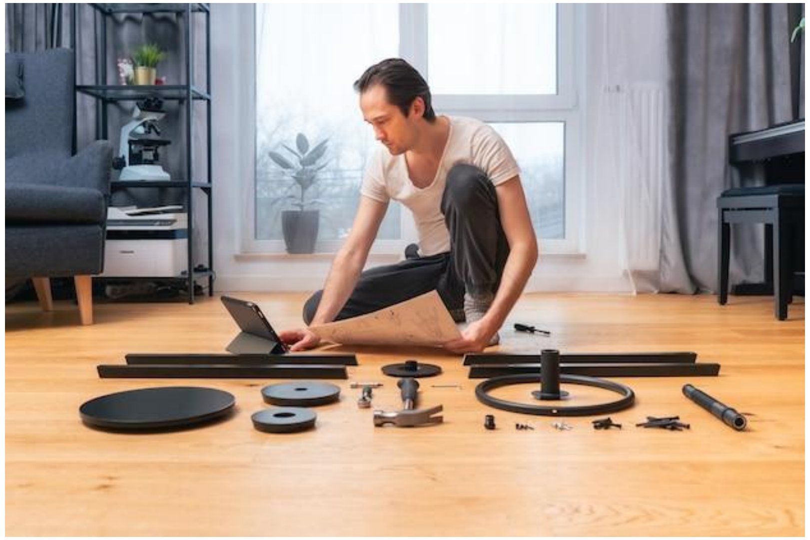
Zeyu Jiang*, Sihang Li*, Siqi Tan†, Chenyang Xu†, Juexiao Zhang,
Julia Galway-Witham, Xue Wang, Scott A. Williams, Radu Iovita,
Chen Feng§, Jing Zhang§

3D Assembly



Fossil

LEGO Toy



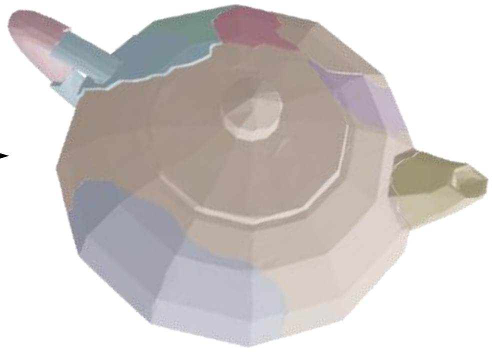
IKEA Furniture

Broken Item

3D Assembly

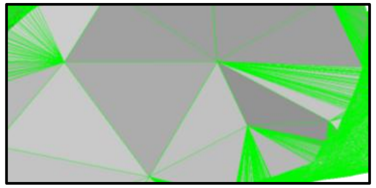
Rotation & Translation of Each Fragments

$$[\boldsymbol{R}, \boldsymbol{T}] = \mathcal{F}_\theta(\boldsymbol{P})$$

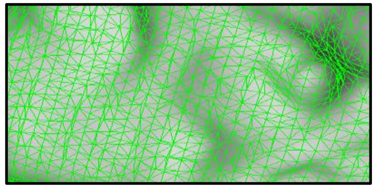


3D Assembly: Path to Physical Spatial Intelligence

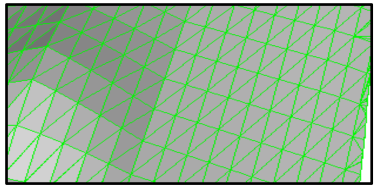
Diverse Local Feature



Physical Simulated



Real World Scan



Geometric Simulation

Diverse Object

PartNext[1]

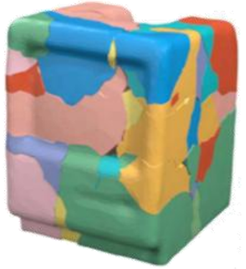
Bone, Ceramics, Lithics

Diverse Scale

N = 69

N = 41

N = 49



N = 45

1. Wang, Penghao, et al. "Partnext: A next-generation dataset for fine-grained and hierarchical 3d part understanding." NeurIPS(2025).

Previous

PuzzleFusion++ (ICLR25)

GARF (ICCV25)

RPF (NeurIPS25)

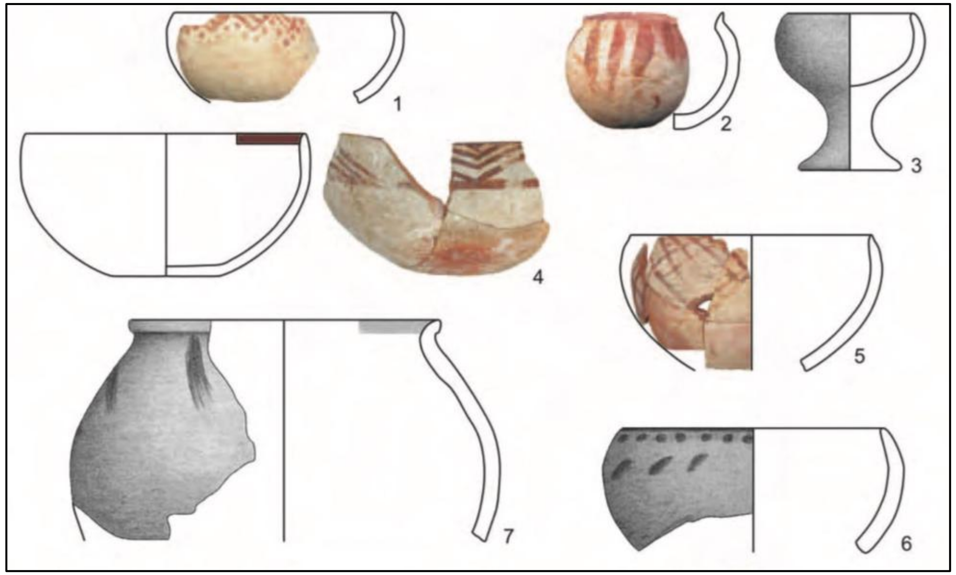
Previous

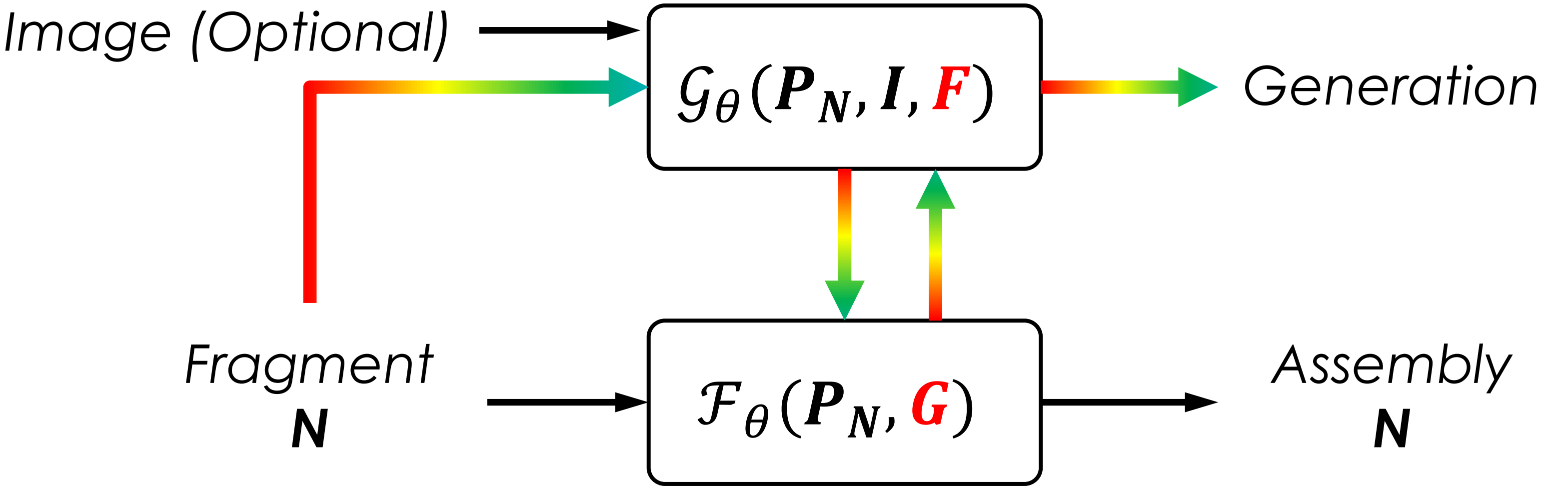
Fragment **N** → $\mathcal{F}_\theta(\boldsymbol{P}_N)$ → *Assembly* **N**

PuzzleFusion++ (ICLR25)	GARF (ICCV25)	RPF (NeurIPS25)
Noise ∈ $\mathcal{N}(\mu, \sigma^2)$	Noise ∈ $\mathcal{N}(\mu, \sigma^2)$	Noise ∈ $\mathcal{N}(\mu, \sigma^2)$
↓	↓	↓
Diffusion	Flow Matching	Rectified Flow
↓	↓	↓
Pose ∈ $\mathbb{R}^7$	Pose ∈ $\mathbb{R}^7$	Point Clouds ∈ $\mathbb{R}^{n \times 3}$
		↓ SVD
		Pose ∈ $\mathbb{R}^7$

However

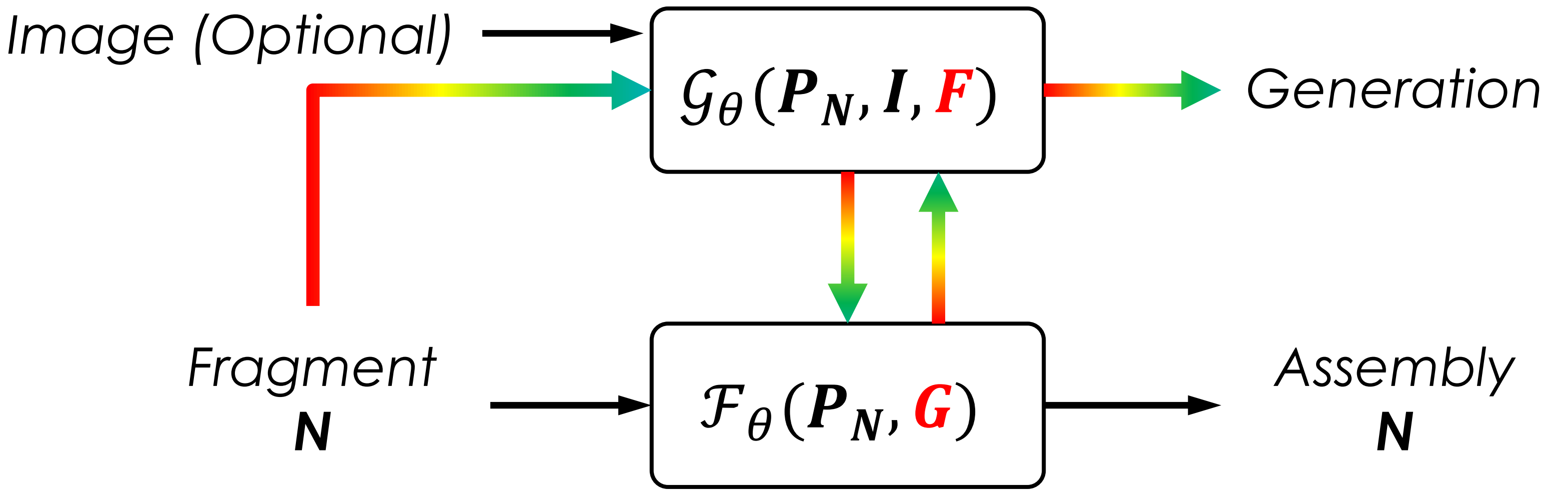
- *There are always* ***missing parts****: conservators often infer absent regions from the available fragments and perform gap-filling to restore a plausible, complete form.*



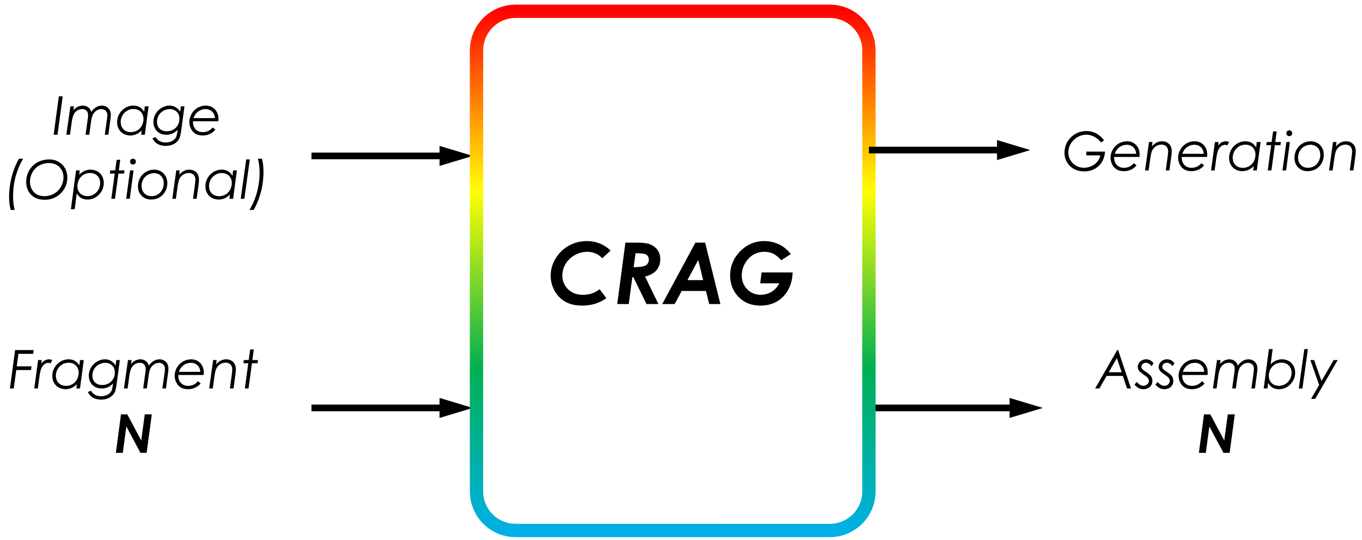


Could 3D assembly and generation be unified?

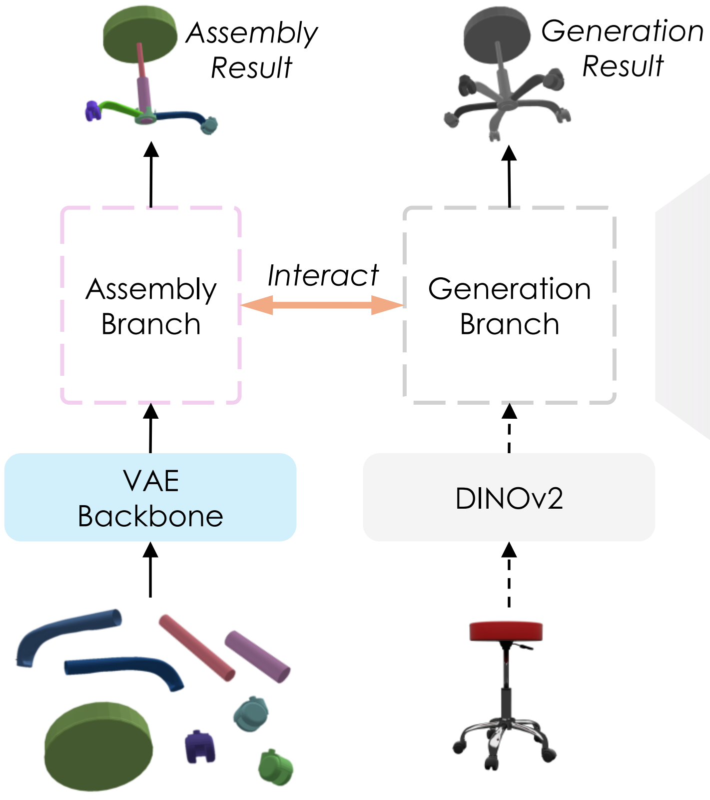
- Challenge 1: How to **infer the global shape** across diverse object categories.
- Challenge 2: a **shared latent space** that can embed an unordered set of variable-size, partial, and noisy fragments into a representation for both assembly and generation
- Challenge 3: effective **bidirectional information exchange**



Could 3D assembly and generation be unified?



Challenge 1: Infer the Global Shape



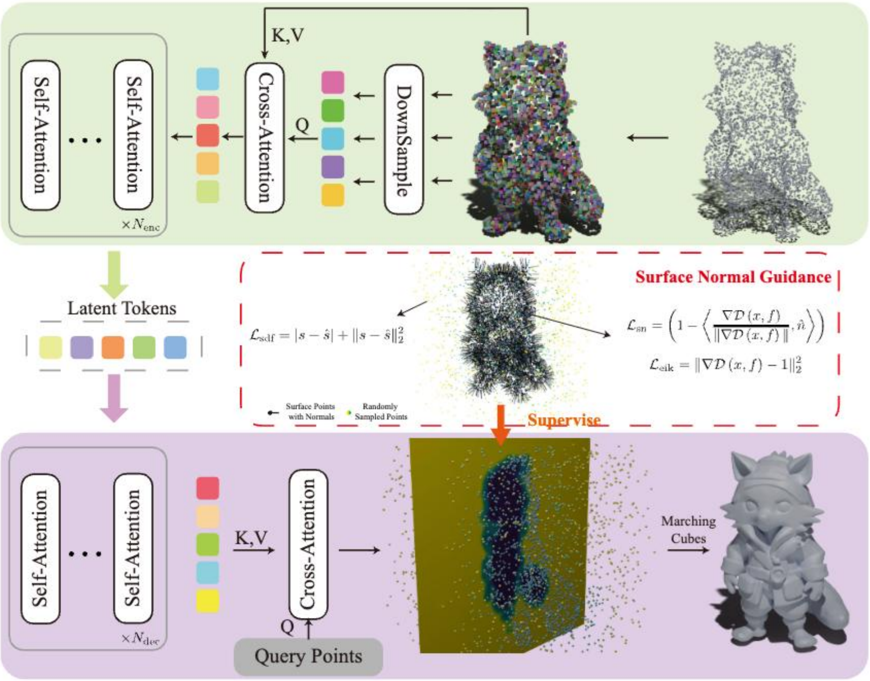
(a) Coupled Architecture

3M *Large-scale Object Dataset*

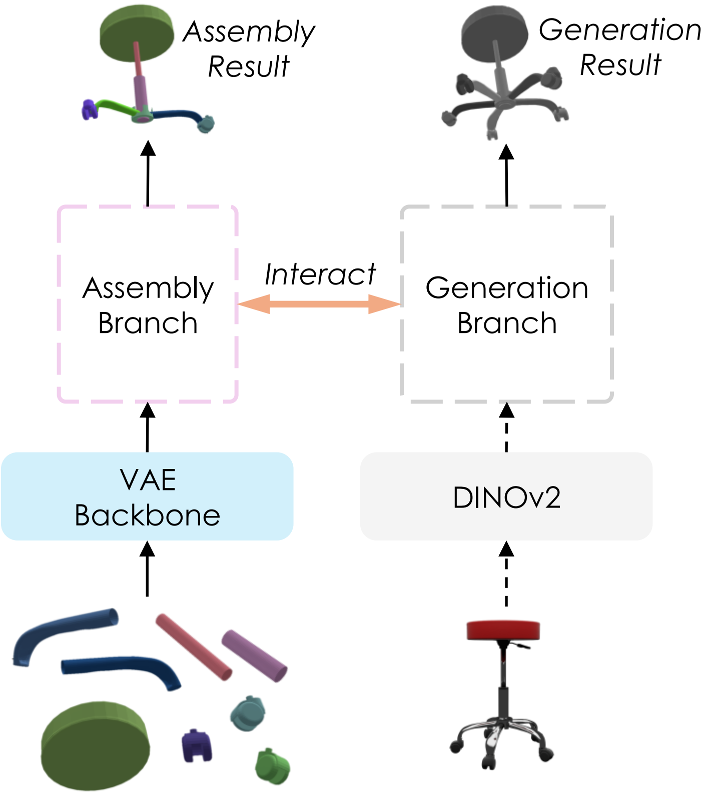
Pretrained TripoSG Model

Li, Yangguang, et al. "Triposg: High-fidelity 3d shape synthesis using large-scale rectified flow models." IEEE Transactions on Pattern Analysis and Machine Intelligence (2025).

Challenge 2: Shared Latent Space



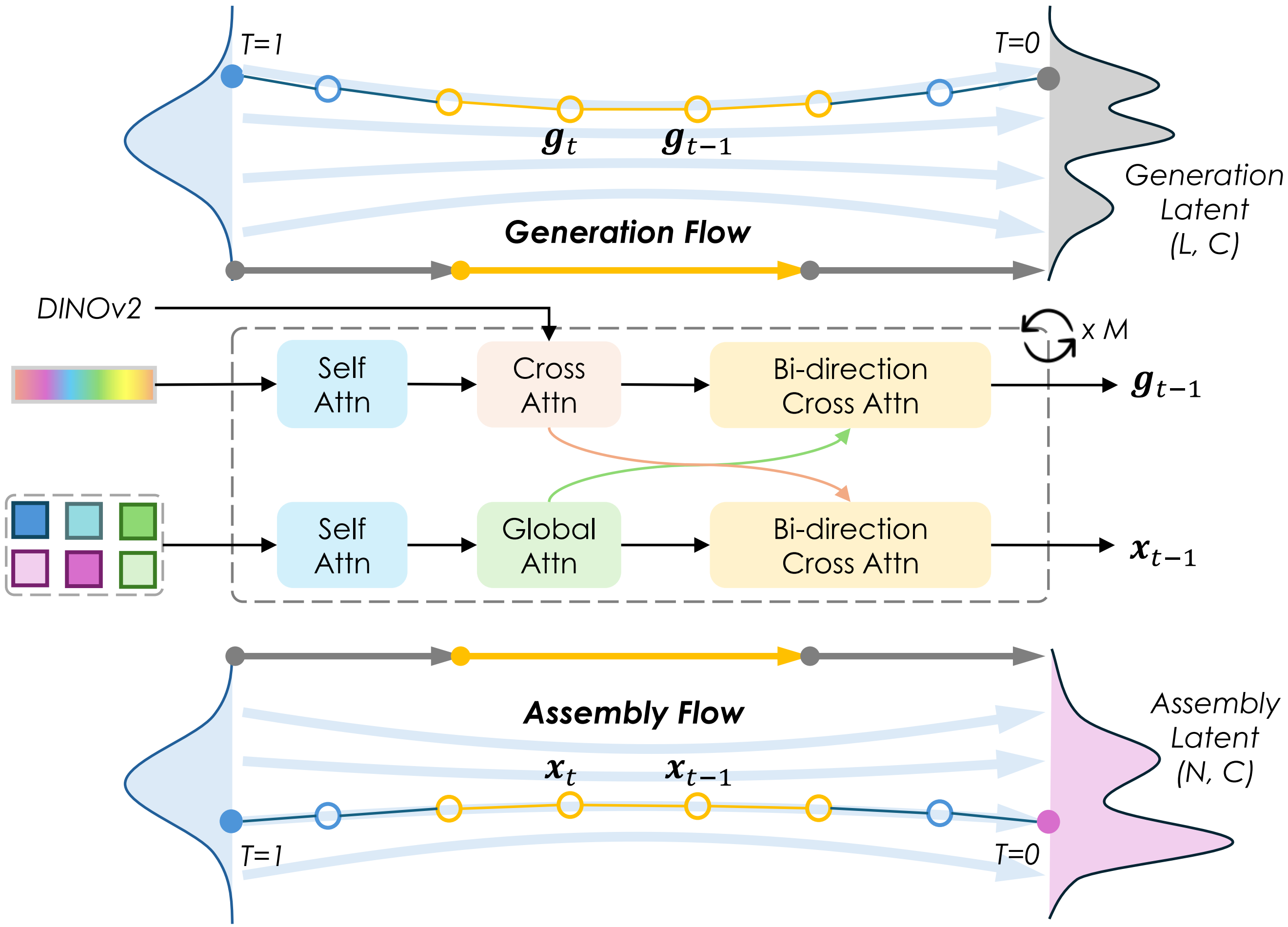
Shared VAE for Fragment Embedding



(a) Coupled Architecture

Li, Yangguang, et al. "Triposg: High-fidelity 3d shape synthesis using large-scale rectified flow models." IEEE Transactions on Pattern Analysis and Machine Intelligence (2025).

Challenge 3: Bidirectional Information Exchange



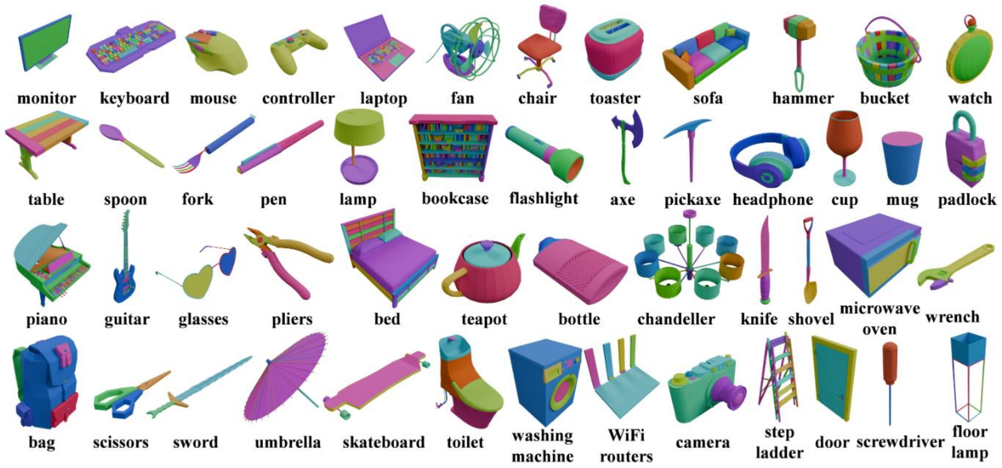
(b) Iteration with Joint Flow

Bi-direction Cross Attention

$$h_{\text{gen}}^{(\ell)} = h_{\text{gen}}^{(\ell)} + \mathrm{CrossAttn}\big(h_{\text{gen}}^{(\ell)}, h_{\text{asm}}^{(\ell)}\big),$$
$$h_{\text{asm}}^{(\ell)} = h_{\text{asm}}^{(\ell)} + \mathrm{CrossAttn}\big(h_{\text{asm}}^{(\ell)}, h_{\text{gen}}^{(\ell)}\big),$$

Dataset

Part-level Assembly
PartNext[1]



Training set: 15563
Evaluation set: 3903

Fracture Assembly
Breaking Bad[2]

Training set: 7872
Evaluation set: 3697

Fracture Assembly
Morphosourse[3]

Training set: 1744
Evaluation set: 437

1. Wang, Penghao, et al. "Partnext: A next-generation dataset for fine-grained and hierarchical 3d part understanding." NeurIPS(2025).
2. Sellán, Silvia, et al. "Breaking bad: A dataset for geometric fracture and reassembly." NeurIPS(2022).
3. Boyer, Doug M., et al. "Morphosource: archiving and sharing 3-D digital specimen data." The Paleontological Society Papers 22 (2016): 157-181.

Experiment

Does generation improve assembly by providing a holistic prior?

Can we still assemble the observed parts and synthesize a complete object from incomplete fragment sets?

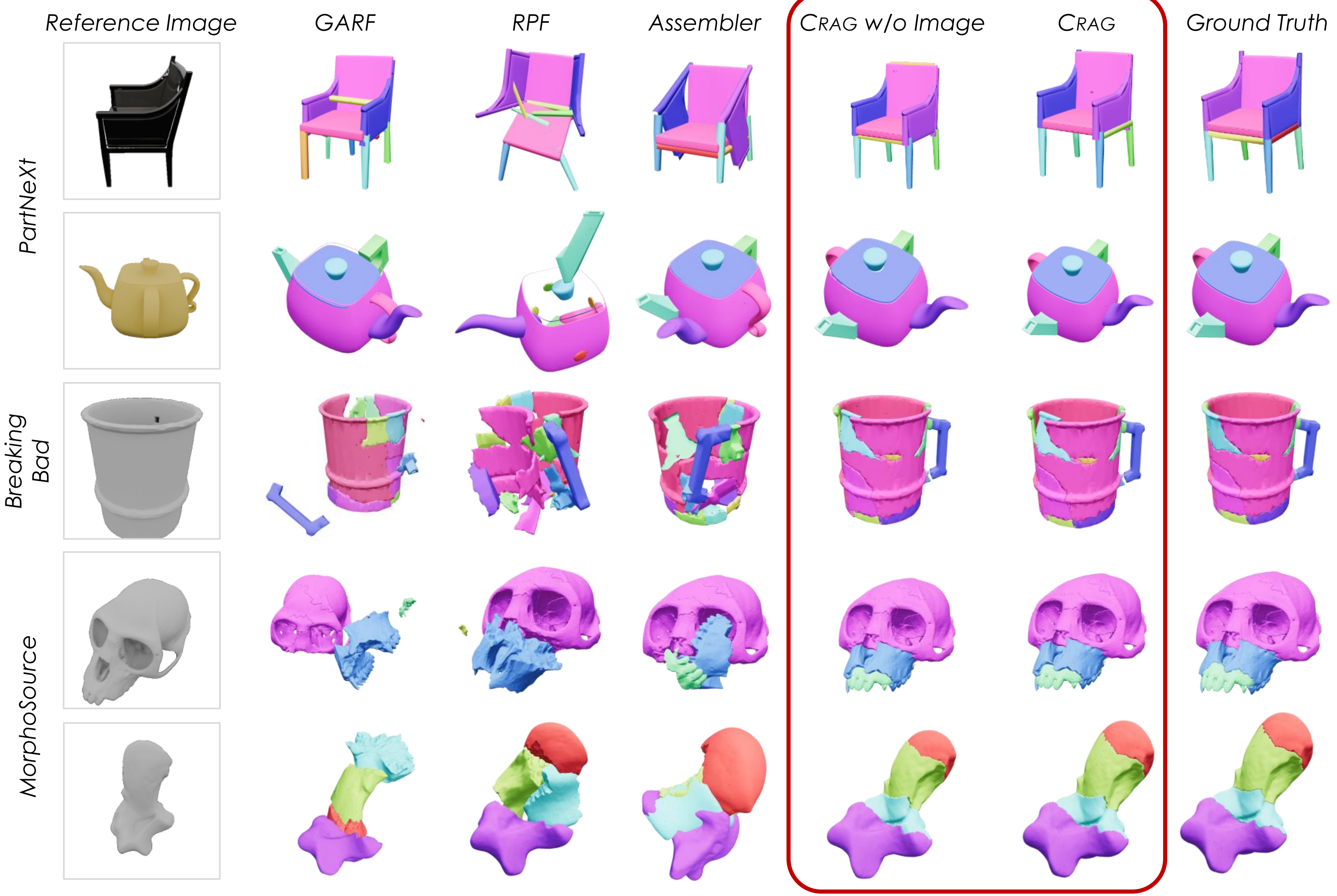
Does part-level evidence help disambiguate image conditioned 3D generation?

Does generation improve assembly by providing a holistic prior?

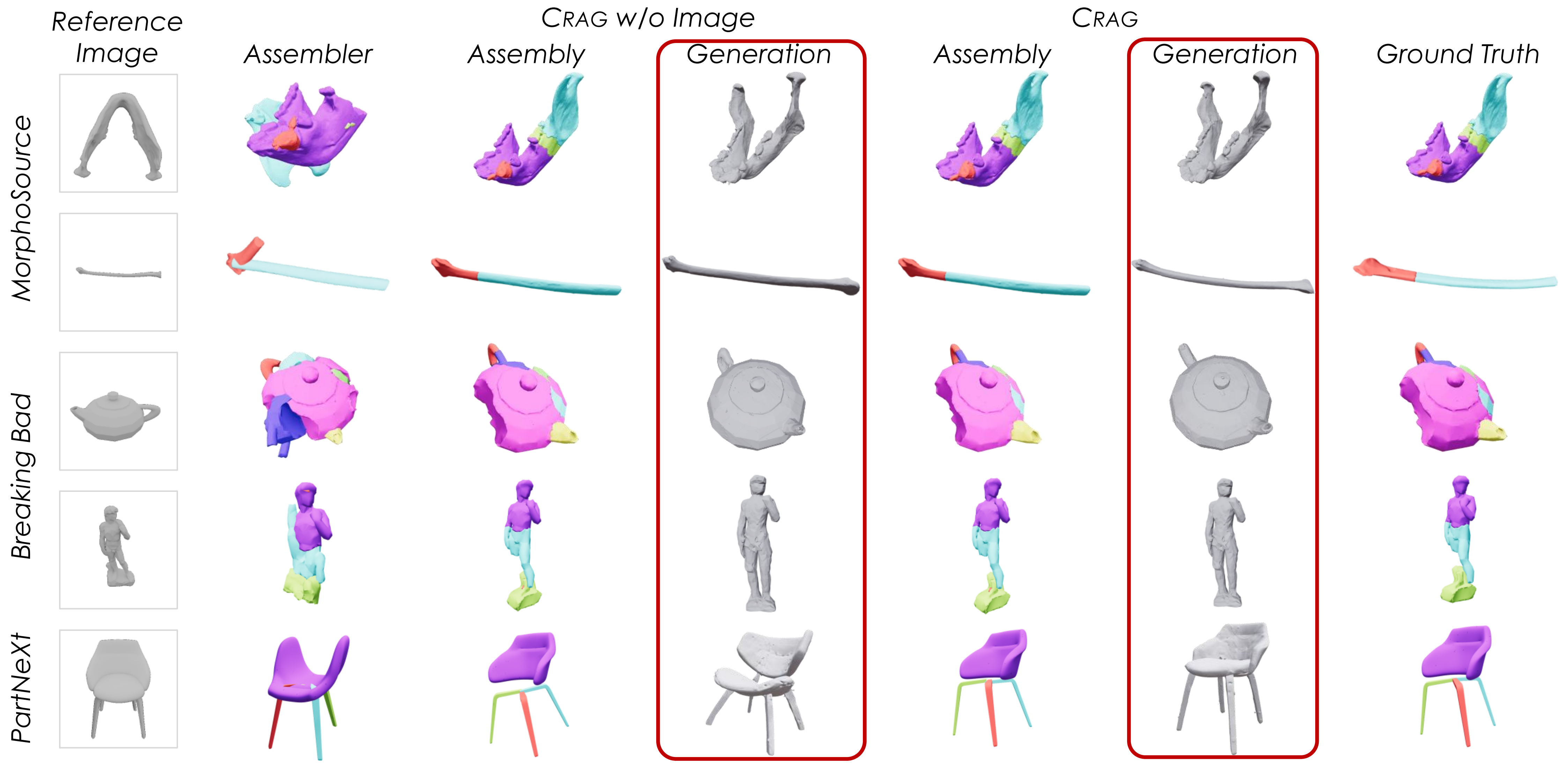
RE: Rotation Error *TE: Translation Error* *PA: Part Accuracy* *CD: Chamfer Distance*

	PartNeXt (Wang et al., 2025a)								Breaking Bad (Sellán et al., 2022)							
Part Status	Complete				Missing				Complete				Missing			
Method	RE ↓	TE ↓	PA ↑	CD ↓	RE ↓	TE ↓	PA ↑	CD ↓	RE ↓	TE ↓	PA ↑	CD ↓	RE ↓	TE ↓	PA ↑	CD ↓
GARF (Li et al., 2025a)	52.52	10.68	58.19	<u>3.10</u>	49.40	9.73	60.71	5.78	<u>8.75</u>	1.72	93.56	0.42	14.59	3.01	85.55	3.43
RPF (Sun et al., 2025)	54.99	29.54	46.17	10.46	42.49	22.95	57.20	8.21	30.59	13.67	80.23	1.01	31.04	14.96	77.05	1.20
CRAG w/o img (Ours)	<u>45.45</u>	<u>9.82</u>	<u>61.67</u>	3.31	<u>42.46</u>	<u>8.70</u>	<u>66.74</u>	<u>5.17</u>	**8.00**	<u>1.37</u>	<u>94.64</u>	<u>0.22</u>	**11.43**	<u>1.79</u>	<u>92.03</u>	<u>0.52</u>
Assembler (Zhao et al., 2025a)	85.82	17.14	44.18	27.93	83.80	15.00	49.86	24.71	74.92	13.22	48.38	7.46	74.45	13.43	45.93	7.78
CRAG (Ours)	**45.12**	**9.13**	**65.89**	**2.40**	**42.33**	**7.86**	**71.81**	**4.21**	**8.00**	**1.36**	**94.68**	**0.21**	<u>11.44</u>	**1.77**	**92.07**	**0.50**

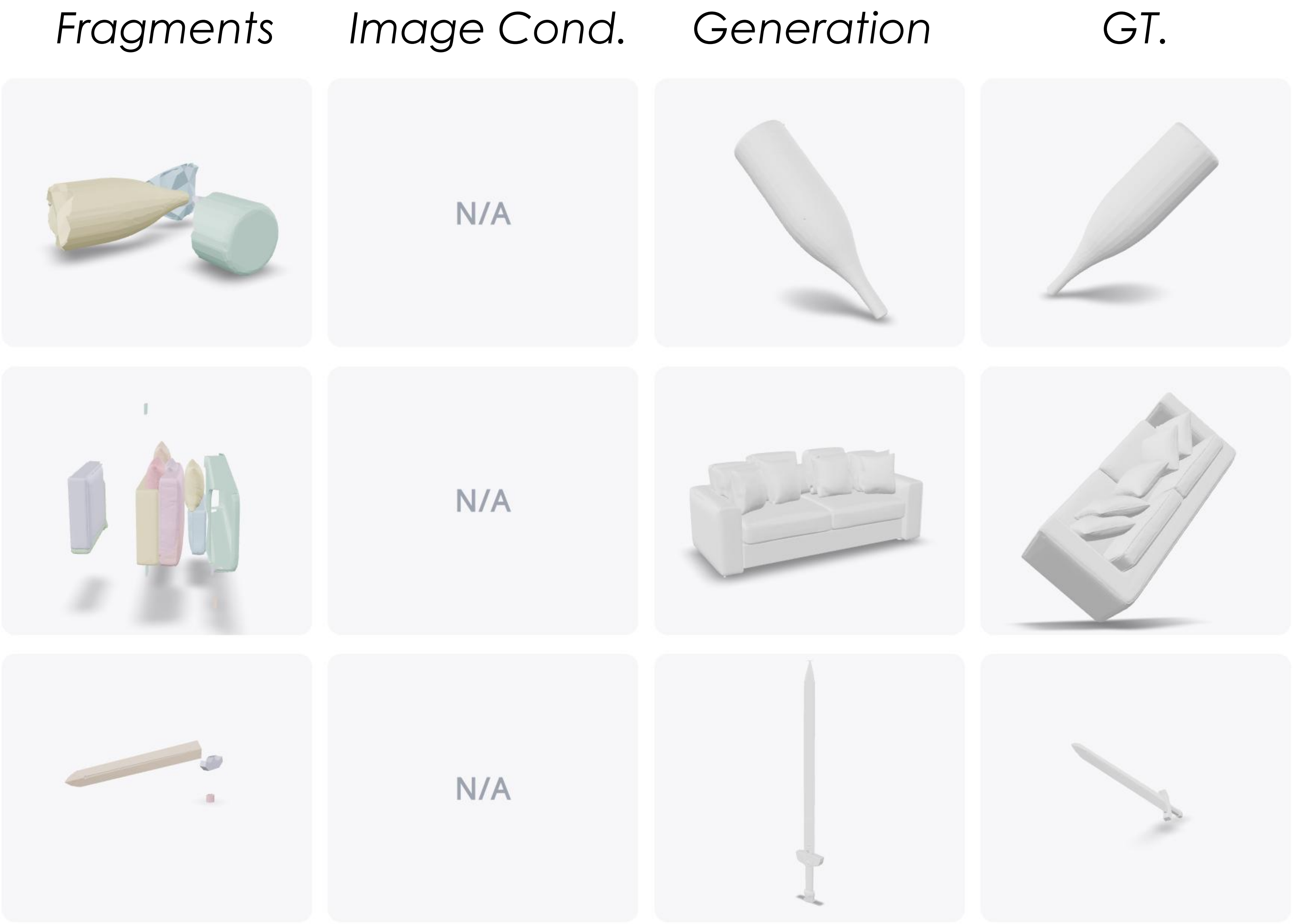
Does generation improve assembly by providing a holistic prior?



Can we still assemble the observed parts and synthesize a complete object from incomplete fragment sets?



Can we still assemble the observed parts and synthesize a complete object from incomplete fragment sets?



Can we still assemble the observed parts and synthesize a complete object from incomplete fragment sets?

RE: Rotation Error *TE: Translation Error* *PA: Part Accuracy* *CD: Chamfer Distance*

	PartNeXt (Wang et al., 2025a)								Breaking Bad (Sellán et al., 2022)							
Part Status	Complete				Missing				Complete				Missing			
Method	RE ↓	TE ↓	PA ↑	CD ↓	RE ↓	TE ↓	PA ↑	CD ↓	RE ↓	TE ↓	PA ↑	CD ↓	RE ↓	TE ↓	PA ↑	CD ↓
GARF (Li et al., 2025a)	52.52	10.68	58.19	3.10	49.40	9.73	60.71	5.78	8.75	1.72	93.56	0.42	14.59	3.01	85.55	3.43
RPF (Sun et al., 2025)	54.99	29.54	46.17	10.46	42.49	22.95	57.20	8.21	30.59	13.67	80.23	1.01	31.04	14.96	77.05	1.20
CRAG w/o img (Ours)	45.45	9.82	61.67	3.31	42.46	8.70	66.74	5.17	**8.00**	1.37	94.64	0.22	**11.43**	1.79	92.03	0.52
Assembler (Zhao et al., 2025a)	85.82	17.14	44.18	27.93	83.80	15.00	49.86	24.71	74.92	13.22	48.38	7.46	74.45	13.43	45.93	7.78
CRAG (Ours)	**45.12**	**9.13**	**65.89**	**2.40**	**42.33**	**7.86**	**71.81**	**4.21**	**8.00**	**1.36**	**94.68**	**0.21**	11.44	**1.77**	**92.07**	**0.50**

20% Missing Part

Does part-level evidence help disambiguate image conditioned 3D generation?

Does part-level evidence help disambiguate image conditioned 3D generation?

CRAG vs 3D Generation Model

Method	CD-L1 ($\times10^{-2}$) ↓	CD-L2 ($\times10^{-3}$) ↓	F-Score@1% (%) ↑	F-Score@5% (%) ↑	EMD ($\times10^{-3}$) ↓	Failure Rate ↓
TripoSG (Li et al., 2025b)	9.12	11.23	14.63	73.18	16.67	25.0% (1157/4626)
CRAG w/o img	9.26	11.46	15.02	72.61	16.96	**0%**
CRAG	**6.53**	**5.83**	**20.49**	**83.87**	**12.92**	**0%**

CRAG vs 3D Completion Model

Method	CD-L1 ($\times10^{-2}$) ↓	CD-L2 ($\times10^{-3}$) ↓	F-Score@1% (%) ↑	F-Score@5% (%) ↑	EMD ($\times10^{-3}$) ↓
SDFusion (Cheng et al., 2023)	22.51	23.93	1.90	28.58	28.60
MGPC (Liu et al., 2026)	12.59	11.54	12.75	70.54	16.99
CRAG w/o img	12.31	10.32	13.84	74.11	16.12
CRAG	**9.39**	**5.76**	**18.23**	**82.82**	**13.09**

Takeaways

- *Imagining the final global shape really helps 3D assembly.*
- *3D assembly can act as auxiliary task to promote 3D generation model.*
    - *You can try joint training (if you have enough computation resources)*
- *Data / Data Distribution matters in domain-specific 3D assembly, like in archaeology or medicine.*
    - *We observe some overfittings in MorphoSource.*

Future Works

- *Data!*
  - *How to synthesize diverse fragments with controllable fracture surface / simulate the real fracture for different domains.*
- *Agentic 3D assembly:*
  - *When you try to assemble object with more than **1000** pieces*
  - *Extract subgroup*
  - *Verifier and rewind*
  - *Combination of end-to-end and optimization-based method*
- *Language-driven arbitrary-shape assembly*
- *Learn from human video*
  - *Physical constraints for assembly*
  - *Novel object*

Thanks!











Project Page

Code Repo