Self Introduction

- <u>**Education**</u>
  - [MS] UNIST, Graduate School of Artificial Intelligence [2021.09 ~ 2023.08]
  - [BS] Seoultech University, Electricity and Information Engineering [2015.03 ~ 2021.08]

- <u>**Publication**</u>
  1. “**DITTO: Dual and Integrated Latent Topologies for Implicit 3D Reconstruction**”, **under review**.
     <u>**Jaehyeok Shim**</u>, Kyungdon Joo.
  2. “**ContactGen: Contact-Guided Interactive 3D Human Generation for Partners**”, **AAAI 2024**.
     Dongjun Gu, <u>**Jaehyeok Shim**</u>, Jaehoon Jang, Changwoo Kang, Kyungdon Joo.
  3. “**Diffusion-Based Signed Distance Fields for 3D Shape Generation**”, **CVPR 2023**.
     <u>**Jaehyeok Shim**</u>, Changwoo Kang, Kyungdon Joo.

Github: https://github.com/Kitsunetic

Homepage: https://kitsunetic.github.io

Index

Research Project 2023.09 ~ 2023.12

1. DITTO:

Dual and Integrated Latent Topologies for Implicit 3D Reconstruction

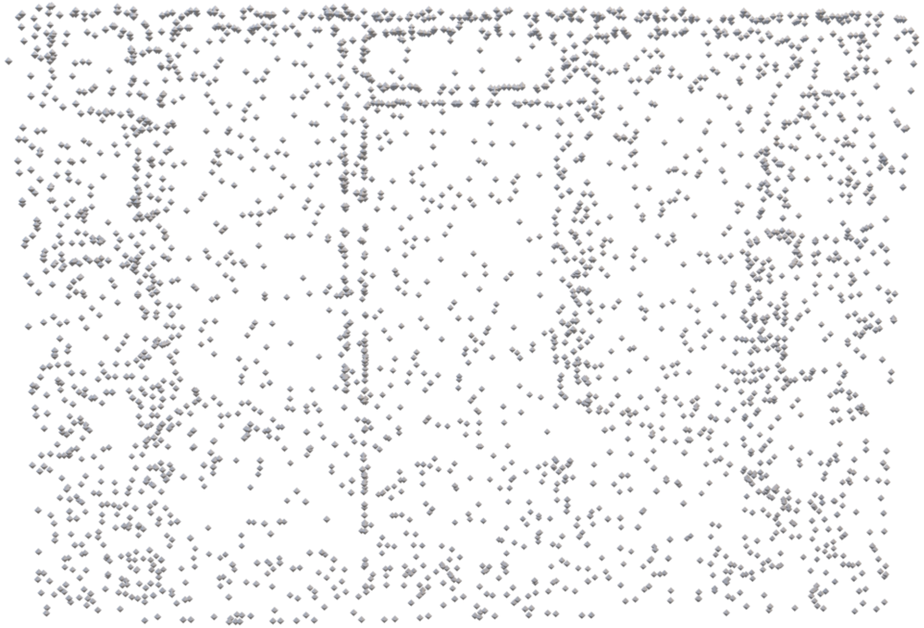
Jaehyeok Shim, Kyungdon Joo

under review

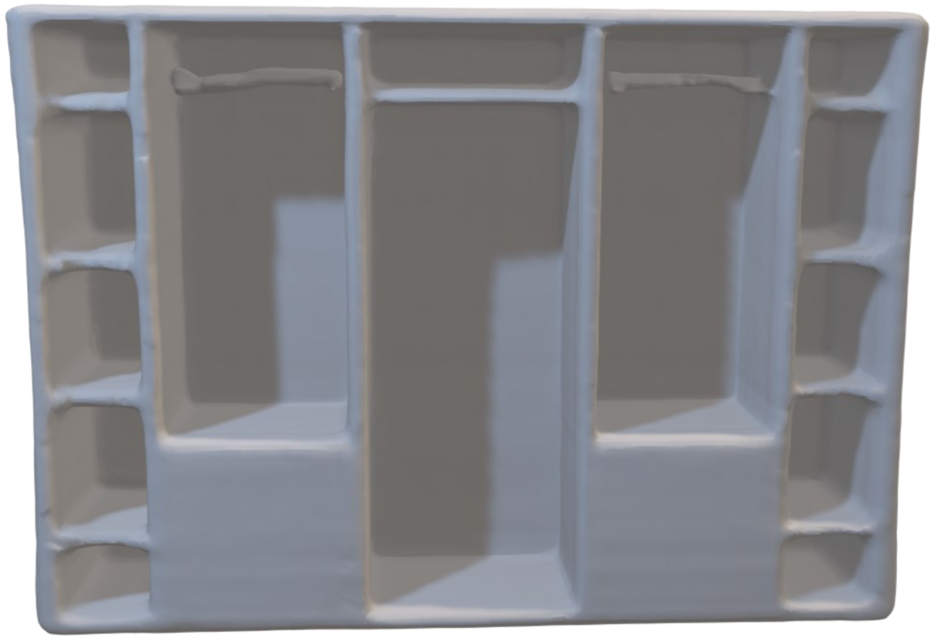
Introduction

Problem Statement

- Reconstructing meshes from input point clouds.



Input: Point cloud
($\mathbb{R}^{N\times3}$)

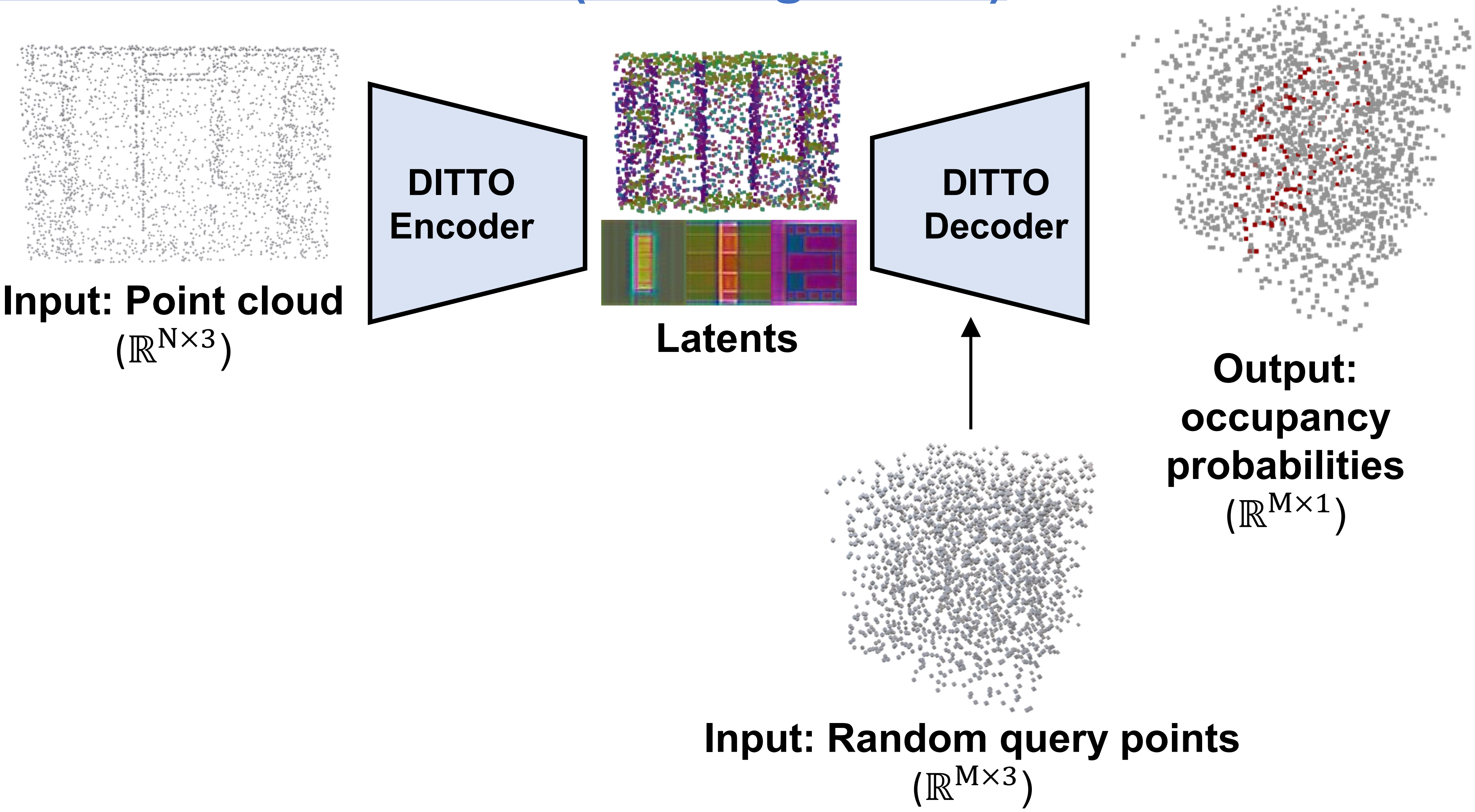


Output: Mesh
(vertices: $\mathbb{R}^{M\times3}$, faces: $\mathbb{R}^{L\times3}$)

- **Fundamental ill-posed problem** in 3D vision similar to super-resolution of image domain.

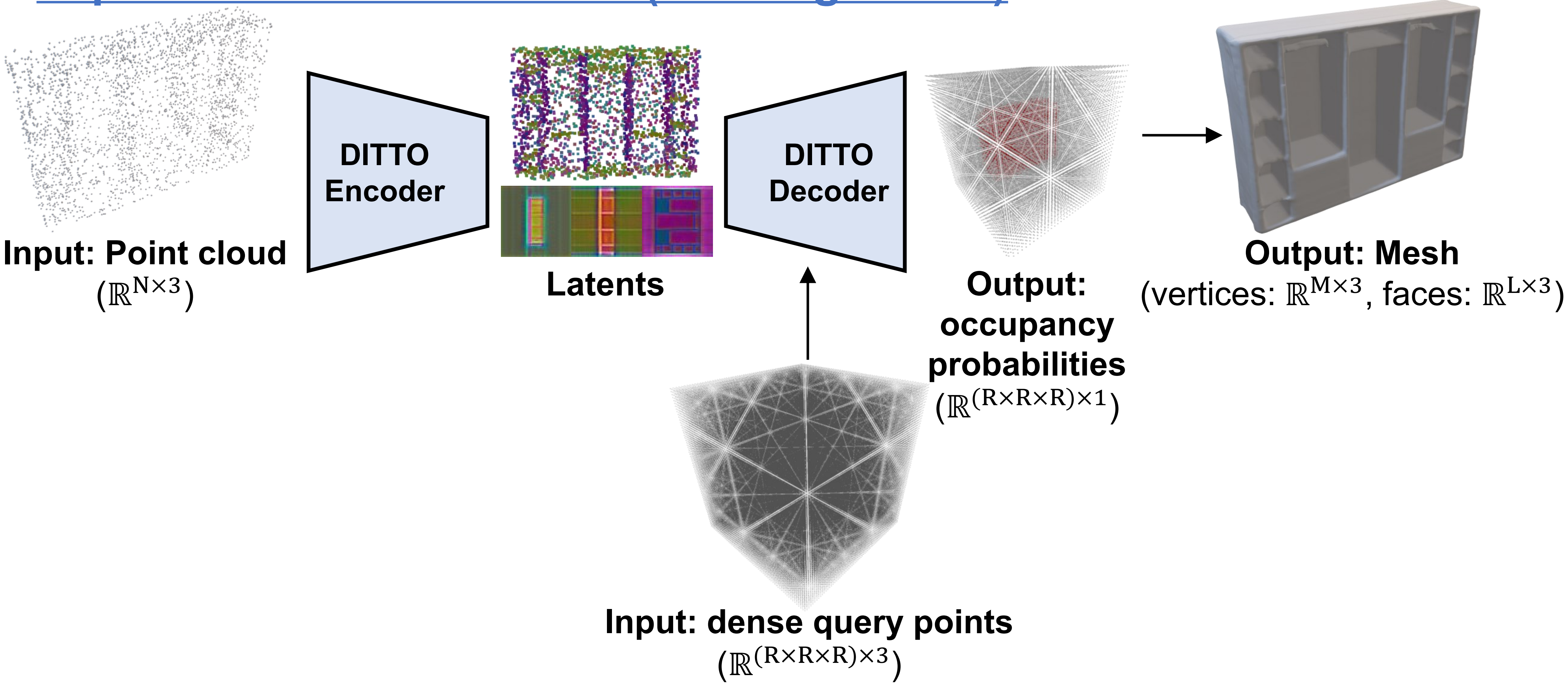
Introduction

Implicit 3D Reconstruction (Training Phase)



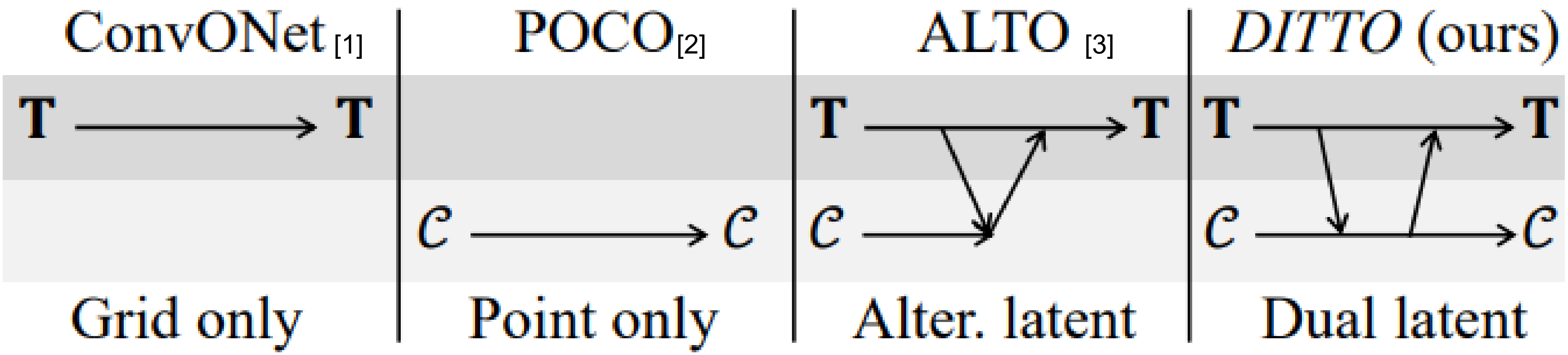
Introduction

Implicit 3D Reconstruction (Training Phase)



Methods

Concept of our DITTO Encoder



Conceptual Comparison of Encoder Architectures

$\mathbf{T} \in \mathbb{R}^{3\times d\times H\times W}$: grid latent

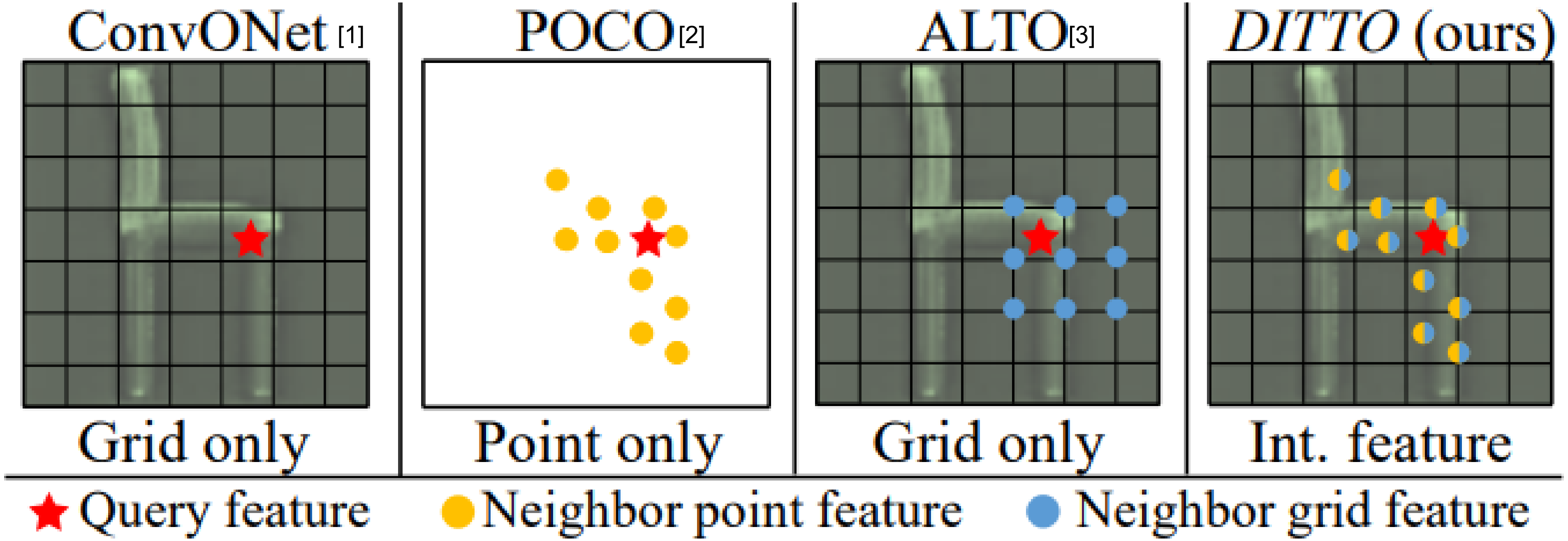
$\mathcal{C} \in \mathbb{R}^{N\times d}$: point latent

- ***DITTO*** encoder uses both grid $\mathbf{T}$ and point $\mathcal{C}$ latents at once.
- Previous methods mainly depend on a single latent type.
- Utilizing complementary strengths of each latent is important.

[1] Peng, Songyou, et al. "Convolutional occupancy networks." ECCV 2020.
[2] Boulch, Alexandre, and Renaud Marlet. "Poco: Point convolution for surface reconstruction." CVPR 2022.
[3] Wang, Zhen, et al. "Alto: Alternating latent topologies for implicit 3d reconstruction." CVPR 2023.

Methods

Concept of our DITTO Decoder



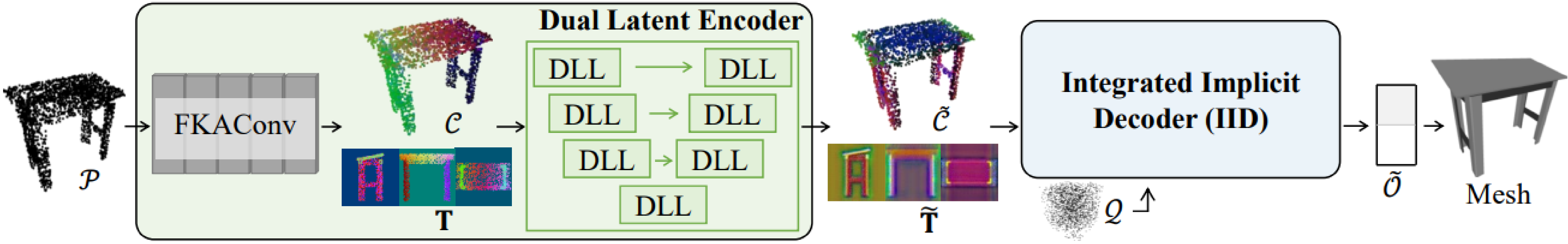
Conceptual Comparison of Decoder Architectures

- ***DITTO*** decoder integrates both grid and point latents to estimate occupancy.
- All previous methods use depend on a single latent.

[1] Peng, Songyou, et al. "Convolutional occupancy networks." ECCV 2020.
[2] Boulch, Alexandre, and Renaud Marlet. "Poco: Point convolution for surface reconstruction." CVPR 2022.
[3] Wang, Zhen, et al. "Alto: Alternating latent topologies for implicit 3d reconstruction." CVPR 2023.

Methods

DITTO Overview



FKAConv converts point $\mathcal{P}$ to latents $\mathcal{C}$ and $\mathbf{T}$

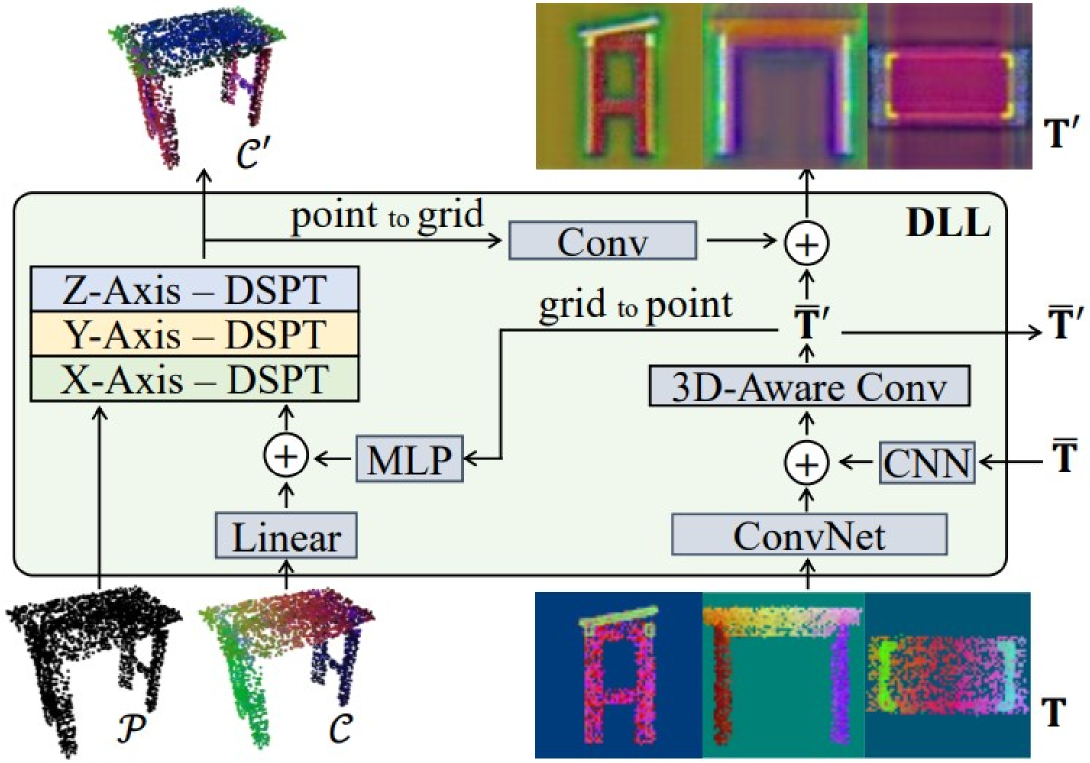
UNet with ***DLL***s refines $\mathcal{C}$ and $\mathbf{T}$ into $\tilde{\mathcal{C}}$ and $\tilde{\mathbf{T}}$

IID extracts occupancy from given query $\mathcal{Q}$

- **Notation**
- $\mathcal{P} \in \mathbb{R}^{\mathrm{N}\times 3}$: point cloud
- $\mathcal{C}$ or $\tilde{\mathcal{C}} \in \mathbb{R}^{\mathrm{N}\times \mathrm{d}}$: point latent
- $\mathbf{T}$ or $\tilde{\mathbf{T}} \in \mathbb{R}^{3\times \mathrm{d}\times \mathrm{H}\times \mathrm{W}}$: grid latent
- $\mathcal{Q} \in \mathbb{R}^{\mathrm{M}\times 3}$: query point
- $\mathcal{O}$ or $\tilde{\mathcal{O}} \in \mathbb{R}^{\mathrm{M}\times 1}$: occupancy

Methods

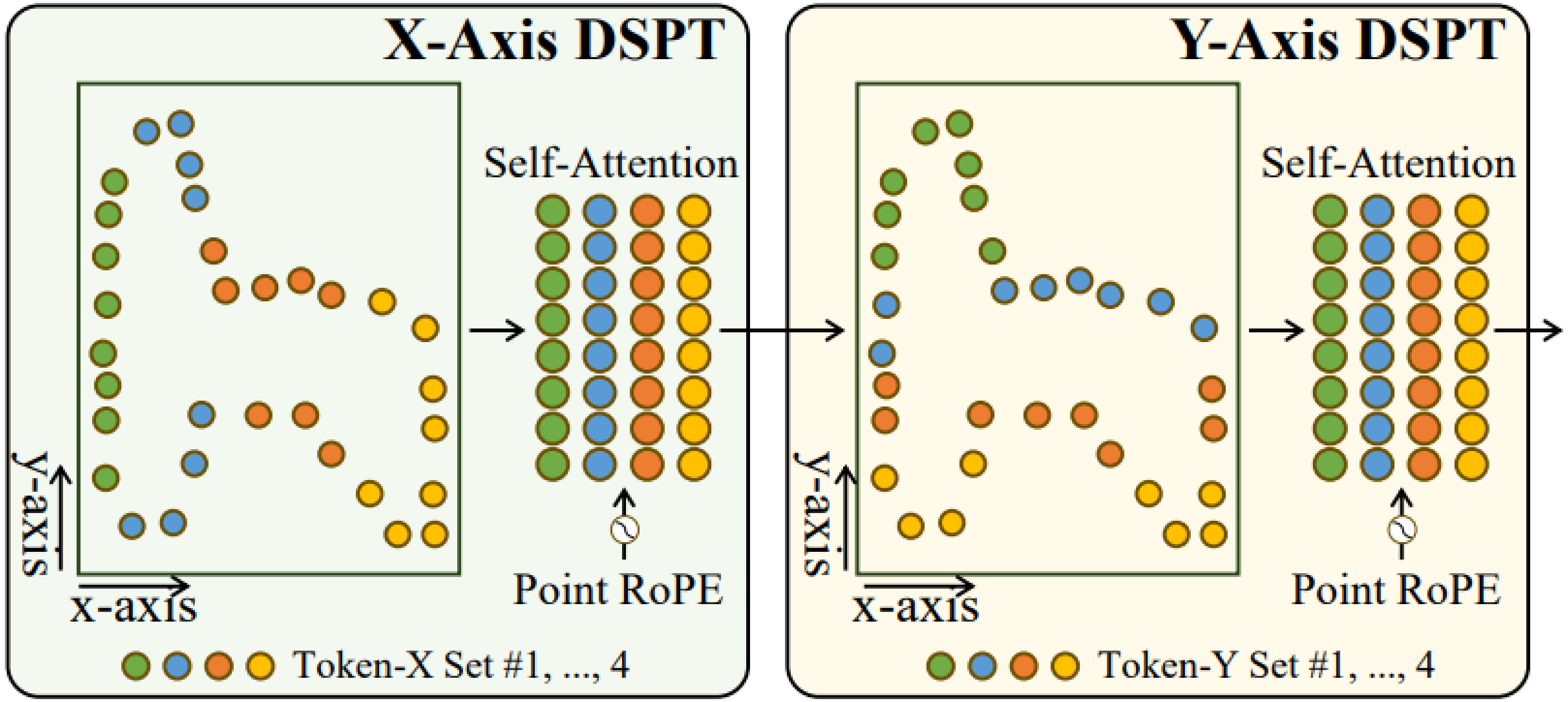
Dual Latent Layer (DLL)



- Enhances both grid $\mathbf{T}$ and point $\mathcal{C}$ latents to produce $\mathbf{T}'$ and $\mathcal{C}'$.
- ***DSPT*** especially enhances point latent $\mathcal{C}$.

Methods

Dynamic Sparse Point Transformer (DSPT)

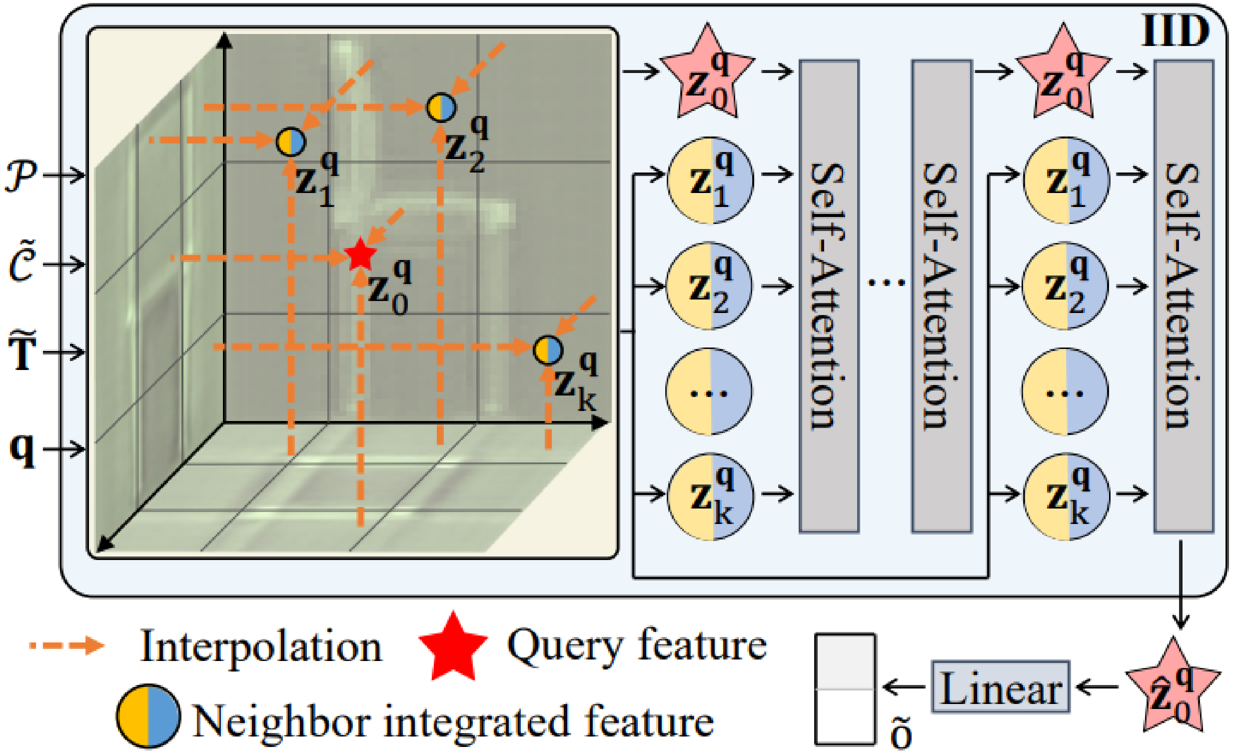


- Similar to SwinTransformer[4]. The windows are split based on point coordinates.

[4] Liu, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows." ICCV 2021.

Methods

Integrated Implicit Decoder (IID)



1. Find neighbor points $\tilde{\mathcal{P}}^{\mathbf{q}}$ and their latents $\tilde{\mathcal{C}}^{\mathbf{q}}$ of a query $\mathbf{q}$.
2. Combine $\tilde{\mathcal{C}}^{\mathbf{q}}$ with grid latent to make $\mathbf{Z}^{\mathbf{q}} = \{\mathbf{z}_1^{\mathbf{q}}, \dots, \mathbf{z}_k^{\mathbf{q}}\}$.
3. Interpolate grid feature at $\mathbf{q}$ to make $\mathbf{z}_0^{\mathbf{q}}$.
4. Apply self-attention multiple times to a sequence $\{\mathbf{z}_0^{\mathbf{q}}, \dots, \mathbf{z}_k^{\mathbf{q}}\}$ to make final occupancy.

Experiments

Object-Level 3D Surface Reconstruction

Method	normal (3K points & noise level 0.005) IoU ↑	Chamfer-L_1 ↓	NC ↑	F-score ↑	sparse (1K points & noise level 0.005) IoU ↑	Chamfer-L_1 ↓	NC ↑	F-score ↑	sparse (300 points & noise level 0.005) IoU ↑	Chamfer-L_1 ↓	NC ↑	F-score ↑
ONet [25]	0.761	0.87	0.891	0.785	0.772	0.81	0.894	0.801	0.778	0.80	0.895	0.806
ConvONet [30]	0.884	0.44	0.938	0.942	0.859	0.50	0.929	0.918	0.821	0.59	0.907	0.883
POCO [1]	0.926	0.30	0.950	0.984	0.884	0.40	0.928	0.950	0.808	0.61	0.892	0.869
ALTO [45]	0.930	0.30	0.952	0.980	0.905	0.35	0.940	0.964	0.863	0.47	0.922	0.924
DITTO (ours)	**0.949**	**0.27**	**0.957**	**0.988**	**0.926**	**0.32**	**0.949**	**0.975**	**0.882**	**0.43**	**0.931**	**0.940**

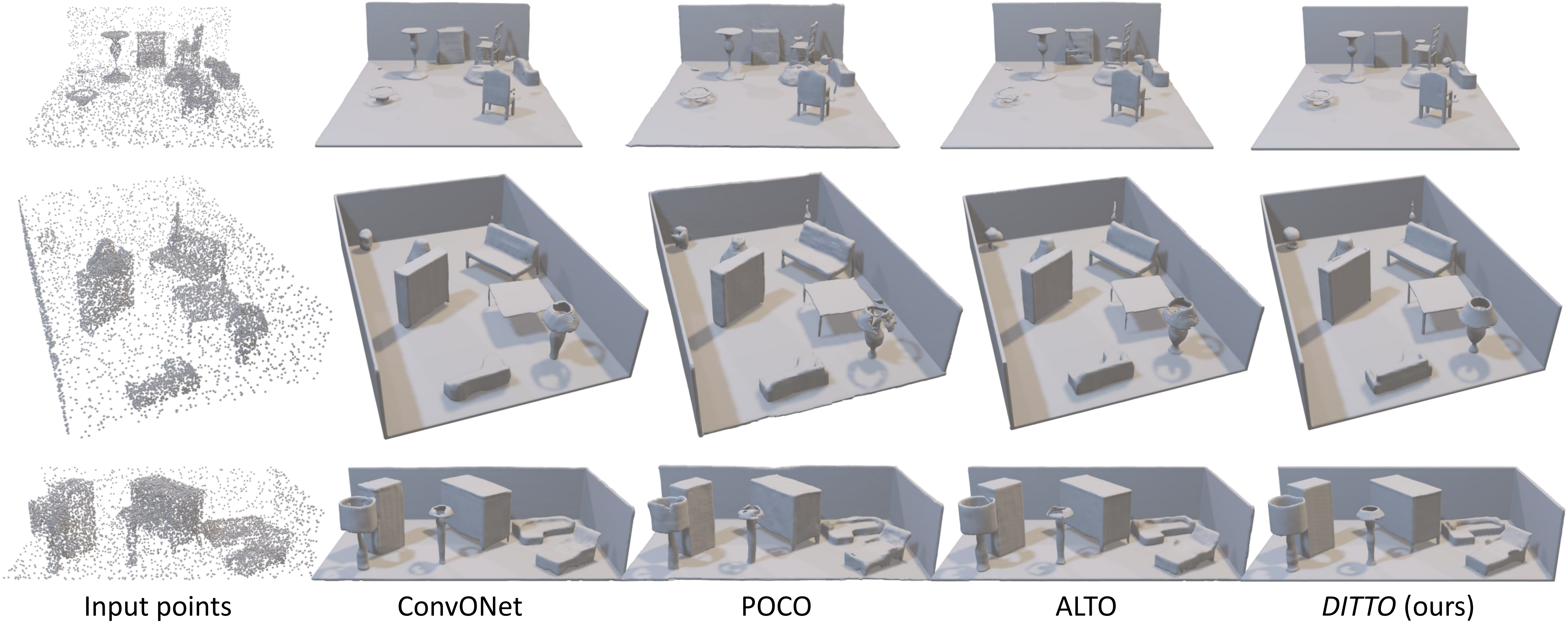
GT mesh   Input points   ConvONet   POCO   ALTO   *DITTO* (ours)

- SoTA in 3D surface reconstruction task in ShapeNet[5] dataset.

[5] Chang, Angel X., et al. "Shapenet: An information-rich 3d model repository." arXiv 2015.

Experiments

Scene-Level 3D Surface Reconstruction



- SoTA in 3D surface reconstruction task in SyntheticRooms dataset.

Peng, Songyou, et al. "Convolutional occupancy networks." ECCV 2020.

Conclusion

- New method how to handle data types including both grid and point cloud.
- New model architecture ***DITTO***, based on proposed ***DLL, DSPT, IID*** modules.
- Achieve SoTA in implicit 3D reconstruction tasks.

Research Project 2023.09 ~ 2023.12

2. SDF-Diffusion:

Diffusion-Based Signed Distance Fields for 3D Shape Generation

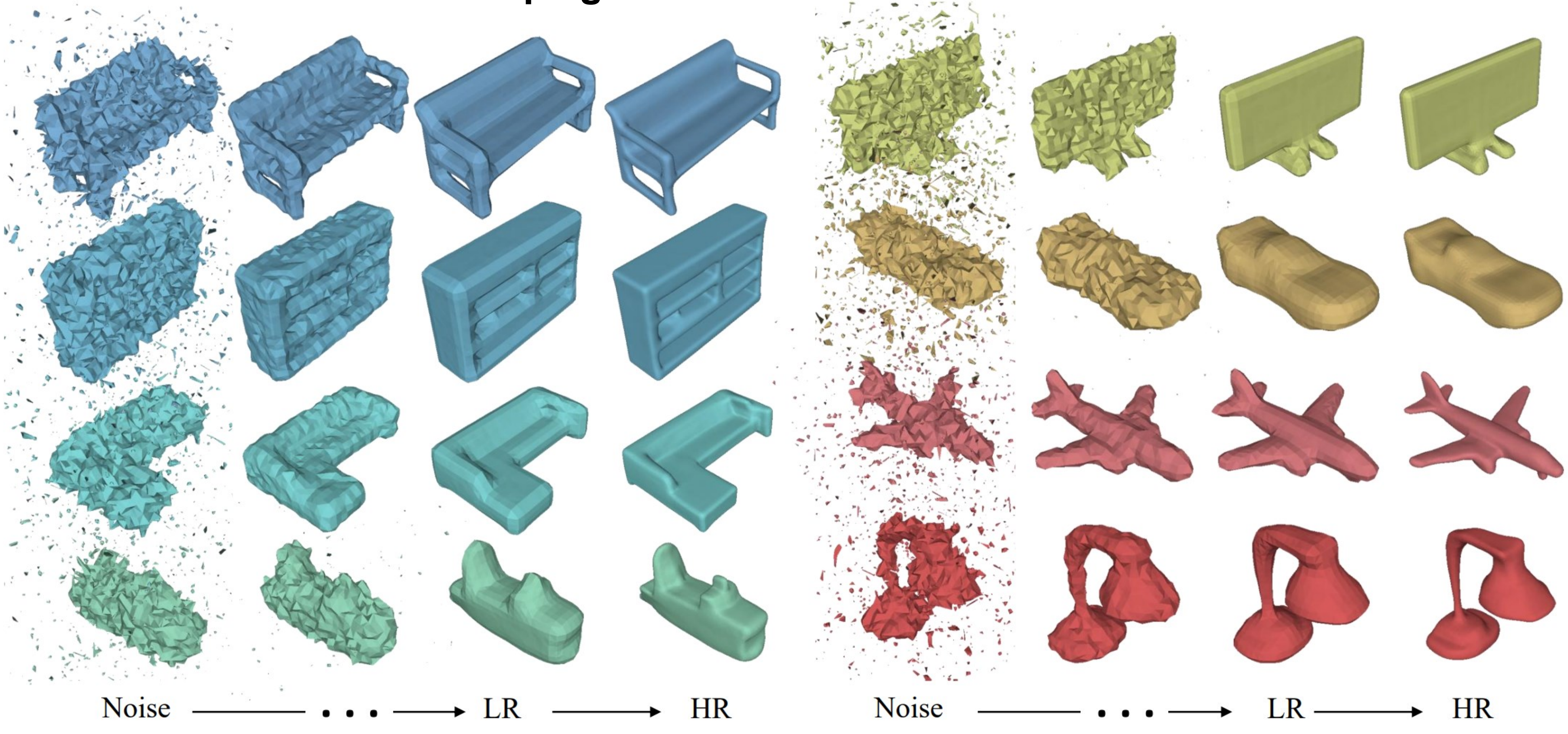
CVPR 2023

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Introduction

Problem Statement

- Make a **diffusion-based 3D shape generative** model in the form of **meshes**.

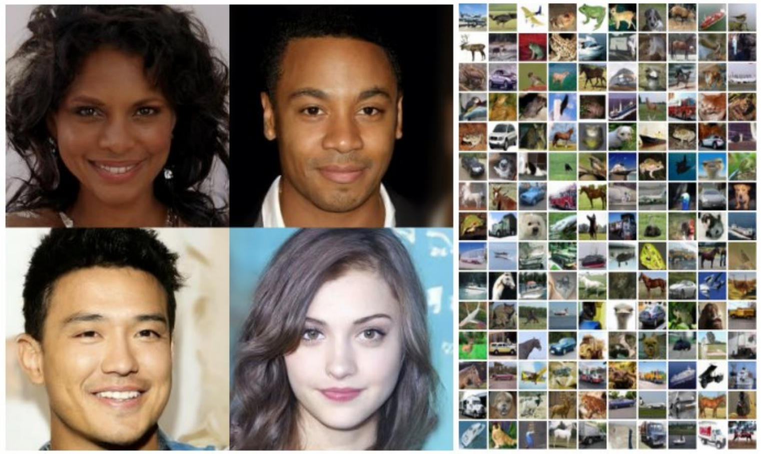


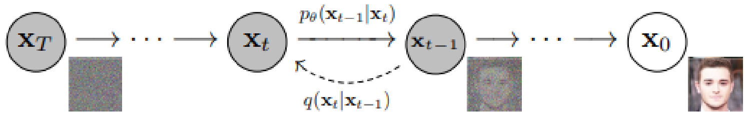
- **Input**: **TSDF Grid** (from mesh), **Output**: New **TSDF Grid** (can be converted into mesh)

Introduction

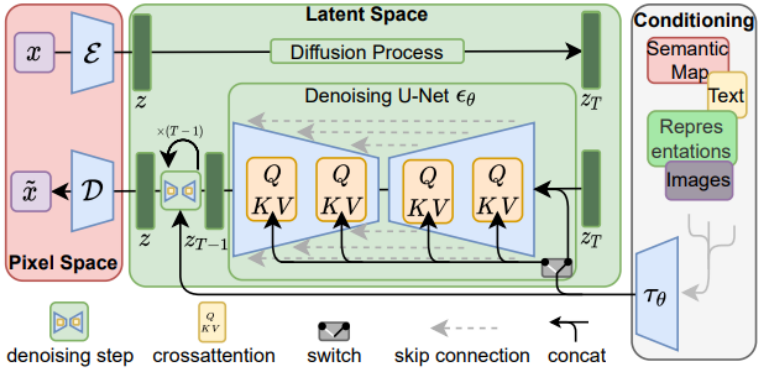
Motivation: Emergence of Denoising Diffusion Models for Image

DDPM (NIPS 2020) [1]

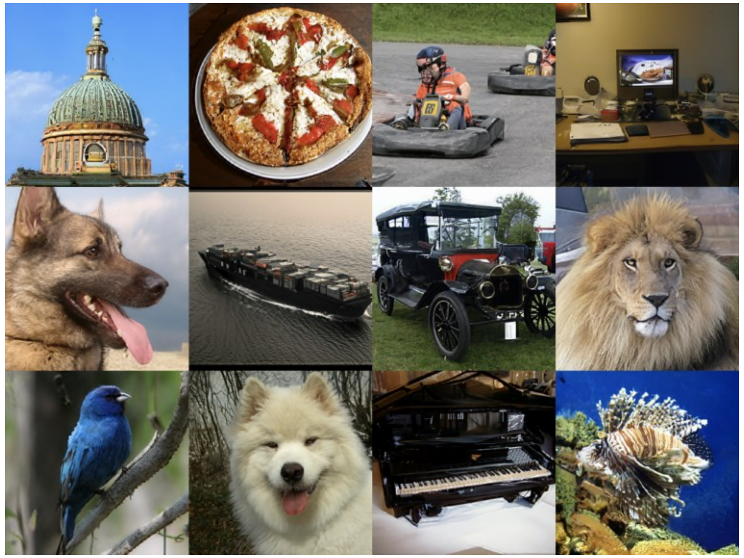


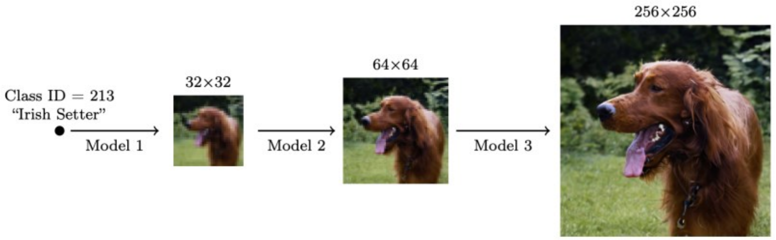


LDM (CVPR 2022) [2]



CDM (JMLR 2022) [3]



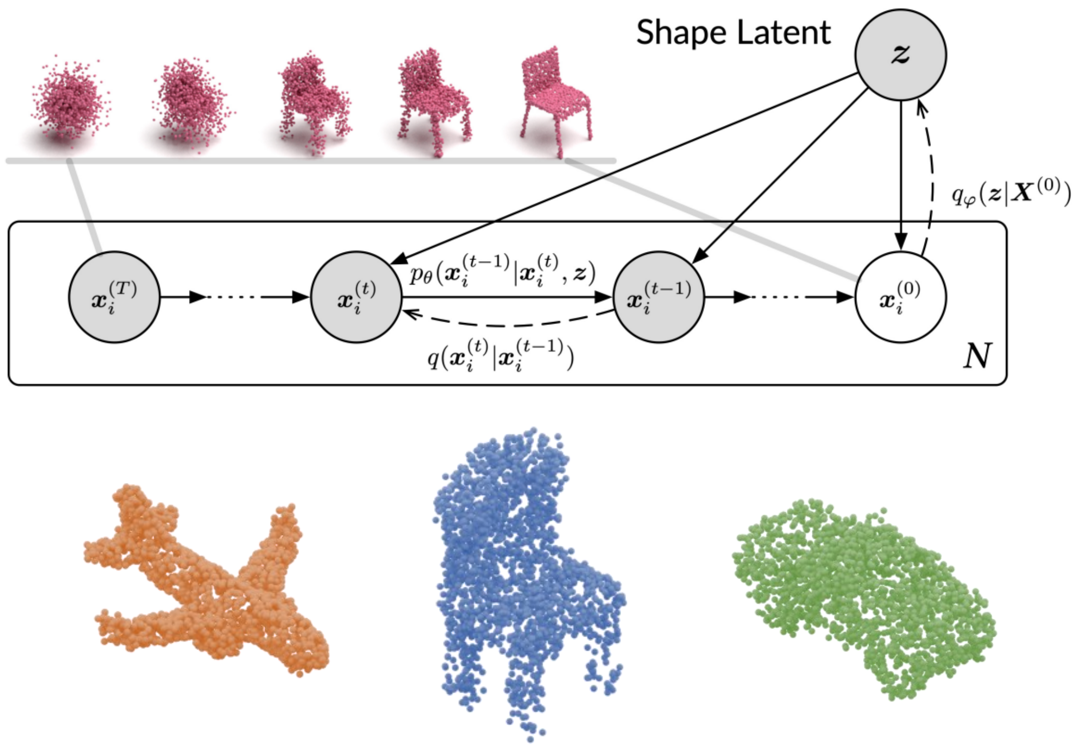


[1] Ho et al. "Denoising diffusion probabilistic models." NeurIPS 2020.
[2] Rombach et al. "High-resolution image synthesis with latent diffusion models." CVPR 2022.
[3] Ho et al. "Cascaded Diffusion Models for High Fidelity Image Generation." JMLR 2022.

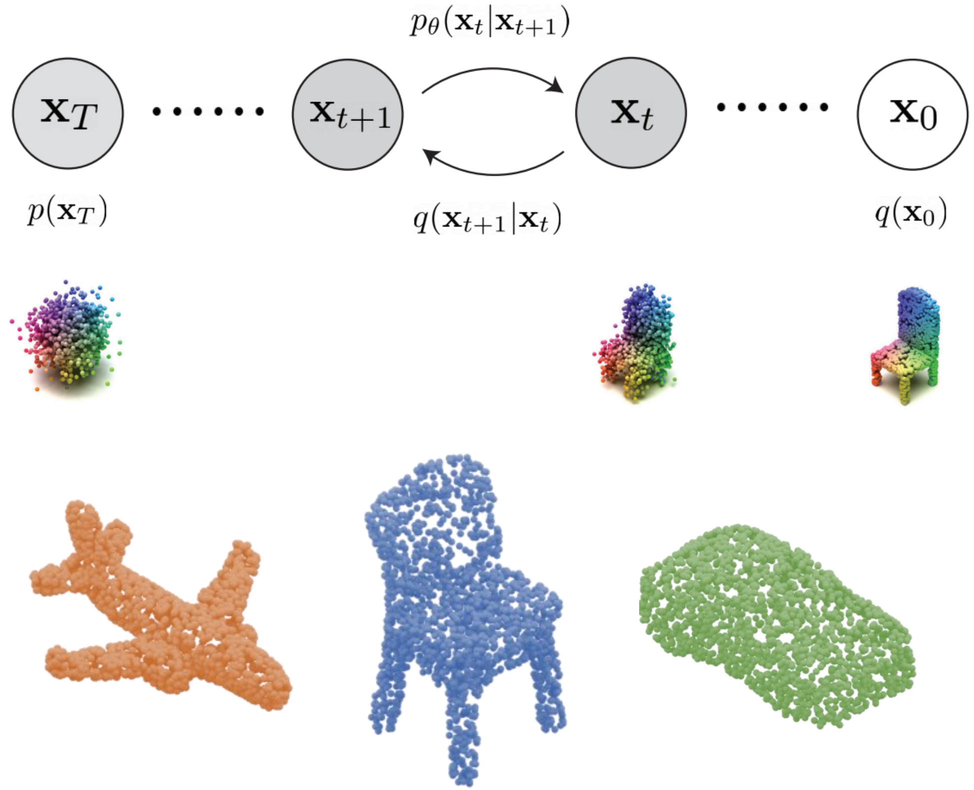
Introduction

Observation: Limitation of Previous Methods

DPM (CVPR 2021) [4]



PVD (ICCV 2021) [5]



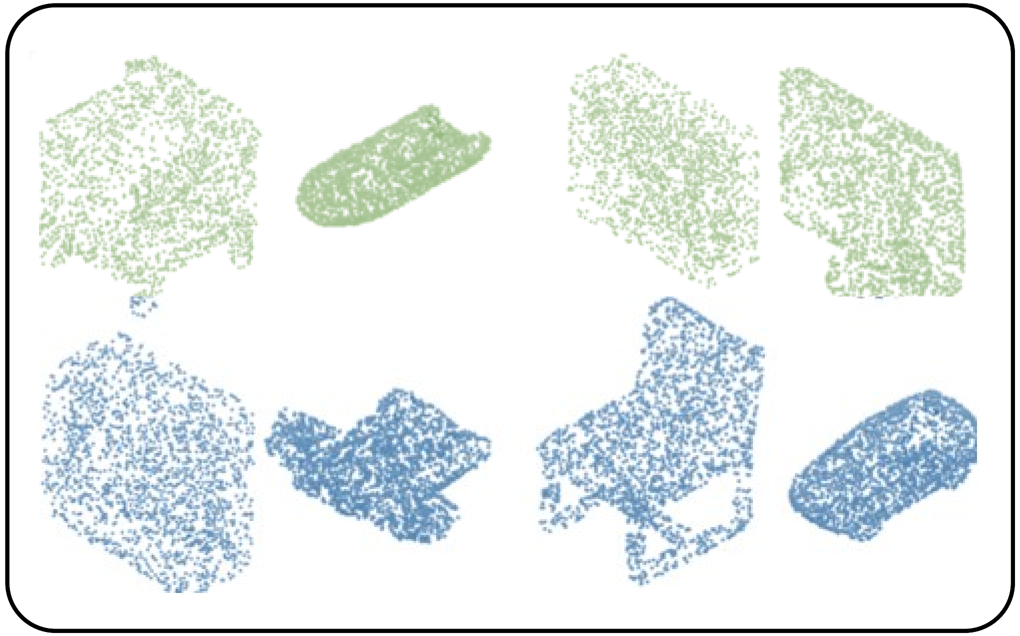
- Previous diffusion-based 3D shape generative models are based on point cloud representation.

[4] Luo and Hu. "Diffusion probabilistic models for 3d point cloud generation." CVPR 2021.
[5] Zhou et al. "3d shape generation and completion through point-voxel diffusion." ICCV 2021.

Introduction

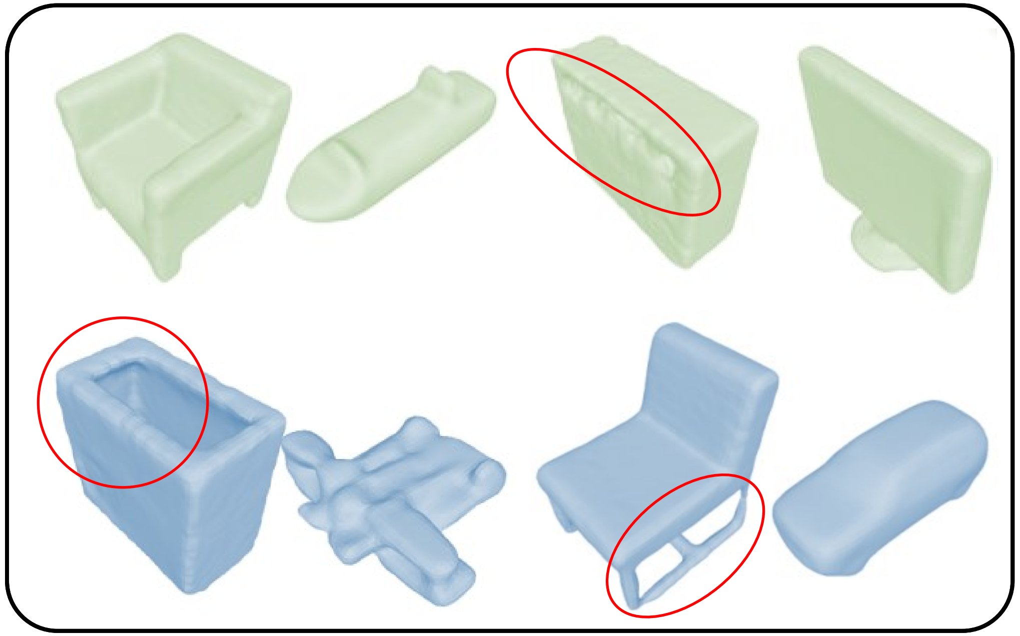
Observation : Limitation of Point Cloud

Point Cloud (from PVD [6])



Remeshing (pretrained SAP [7])

Reconstructed Mesh



- Point cloud requires **complex processes** to reconstruct mesh.
- Reconstructed mesh has **imperfect surface** due to partial characteristic of point cloud.

[6] Liu et al. "Point-voxel cnn for efficient 3d deep learning." NeurIPS 2019.
[7] Peng et al. "Shape as points: A differentiable poisson solver.“ NeurIPS 2021.

Introduction

Challenge: Processing High-Resolution Voxel

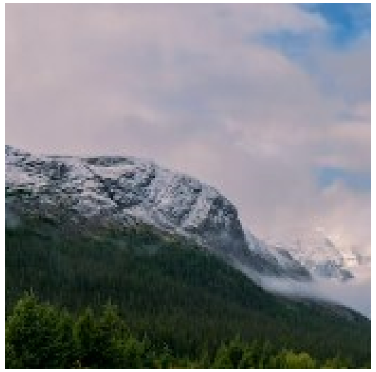
- Training deep neural network requires thousands of features maps.

Data Type	Spatial Complexity	Capacity
Image (2D)	128 x 128	**2 Mbytes**
Voxel (3D)	128 x 128 x 128	**256 Mbytes**

➔ **Out of memory!**

(for 32 channels of float32)

- Processing voxel like 2D image will exceeds **GPU memory and computation limitation**.



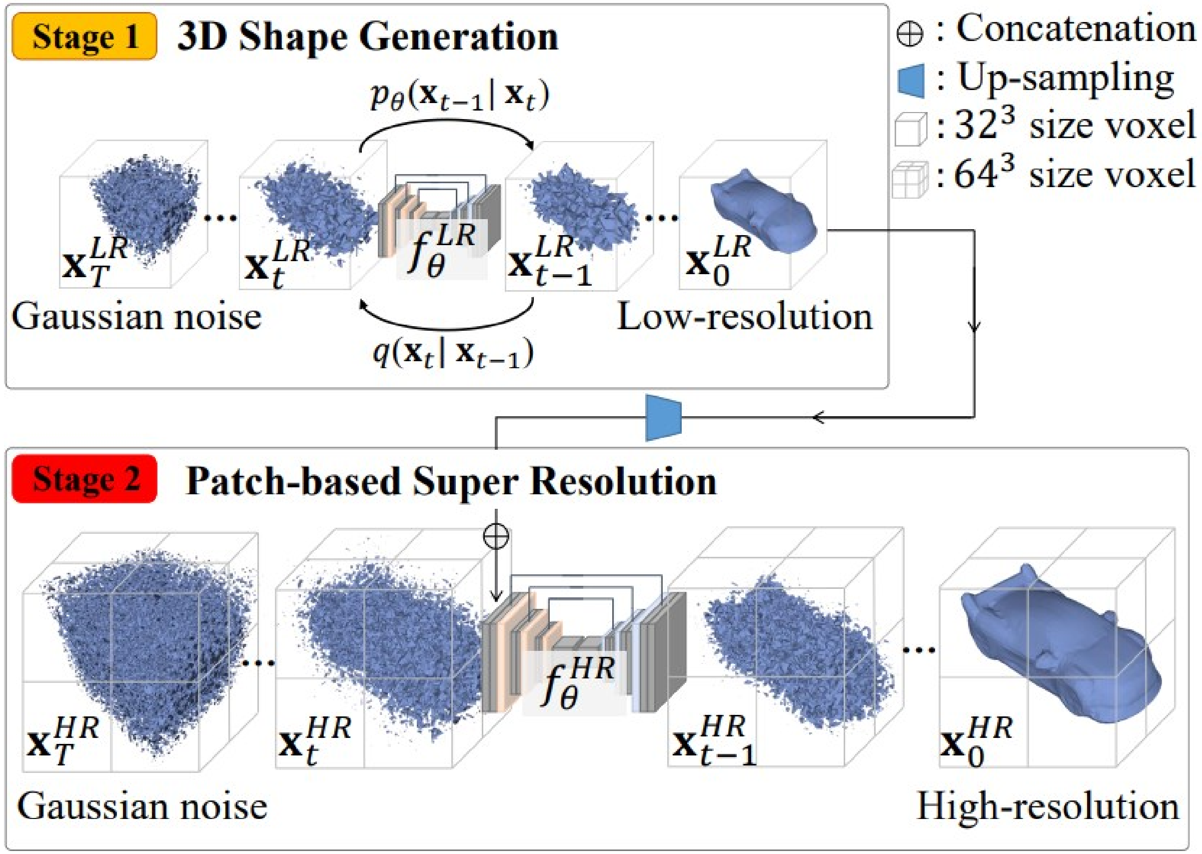
2D $128^2 = 2^{14}$

3D $128^3 = 2^{21}$

Methods

Two-Stage Diffusion Framework for TSDF Voxel

- Stage 1: Generates **low-resolution TSDF voxel** (at 32^3 resolution).
- Stage 2: **Upsamples TSDF voxel** into high-resolution ($32^3 \rightarrow 64^3$ or $64^3 \rightarrow 128^3$)

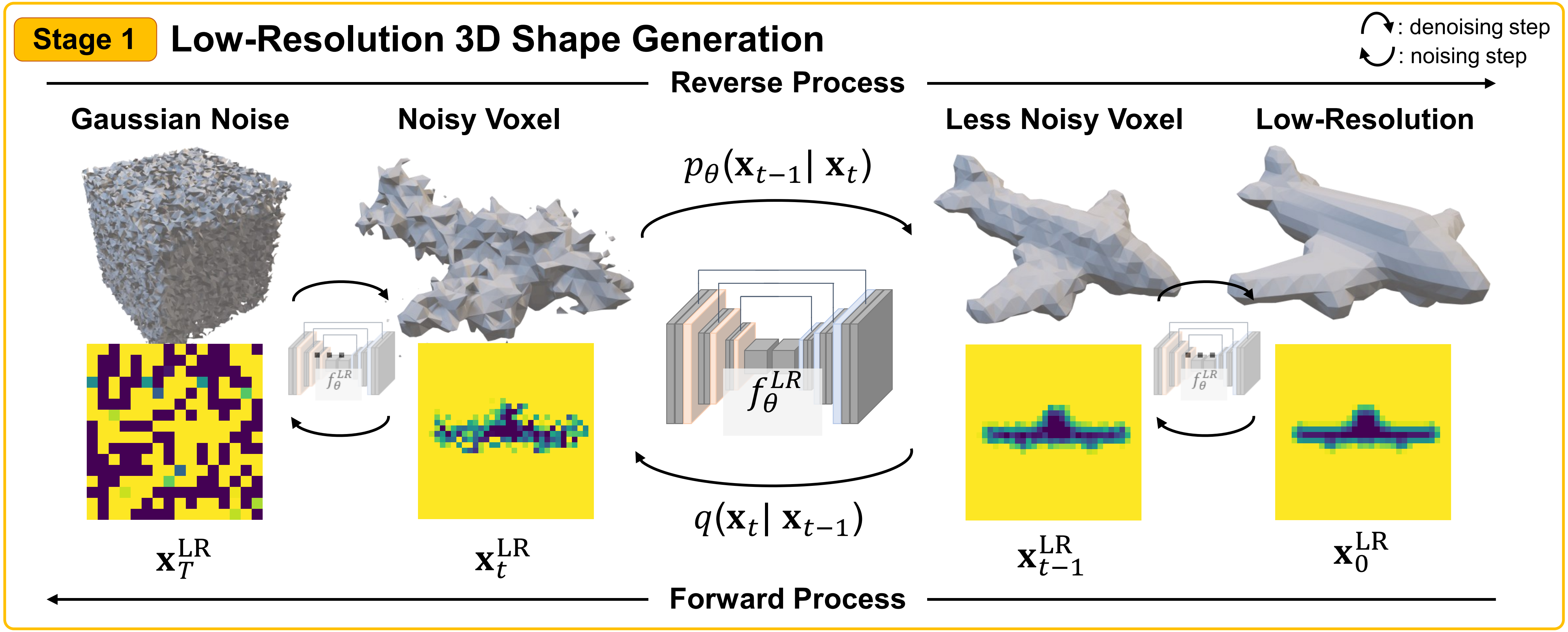


- $\mathbf{x}_t^{\mathrm{LR}}$: low-resolution TSDF voxel at timestep t.
- $\mathbf{x}_t^{\mathrm{HR}}$: high-resolution TSDF voxel at timestep t.
- f_θ^{LR}: first-stage generation model.
- f_θ^{HR}: second-stage super-resolution model.
- $q(\mathbf{x}_t|\mathbf{x}_{t-1})$: diffusion forward process.
- $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)$: diffusion reverse process.

Methods

Stage 1: Diffusion Models for Low-Resolution Generation

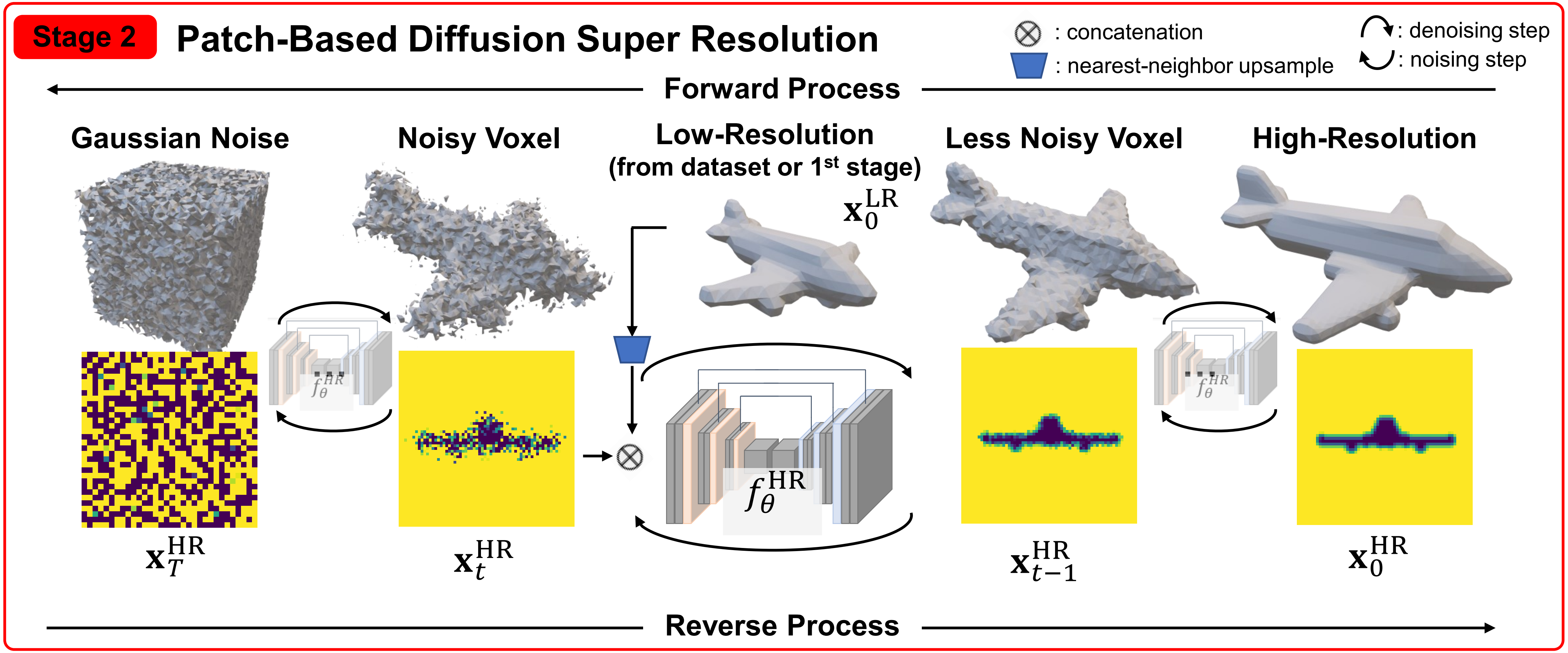
- **Stage 1:** generate **low-resolution** TSDF voxel using **diffusion framework**.



Methods

Stage 2: Diffusion Models for High-Resolution Super-Resolution

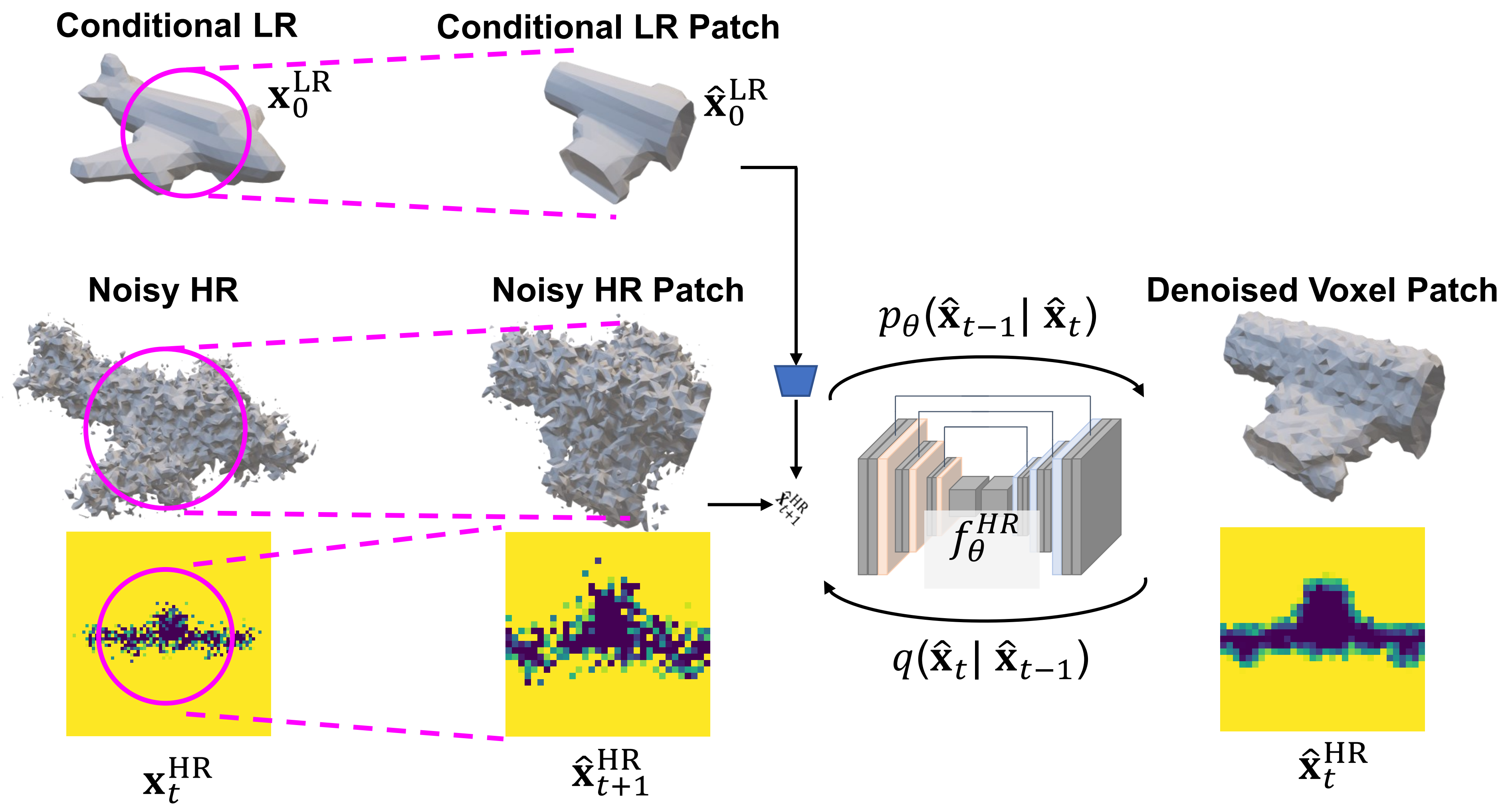
- **Stage 2:** generate **high-resolution** TSDF voxel using **diffusion-based super-resolution**.



Methods

Stage 2: Patch-Based Training Scheme

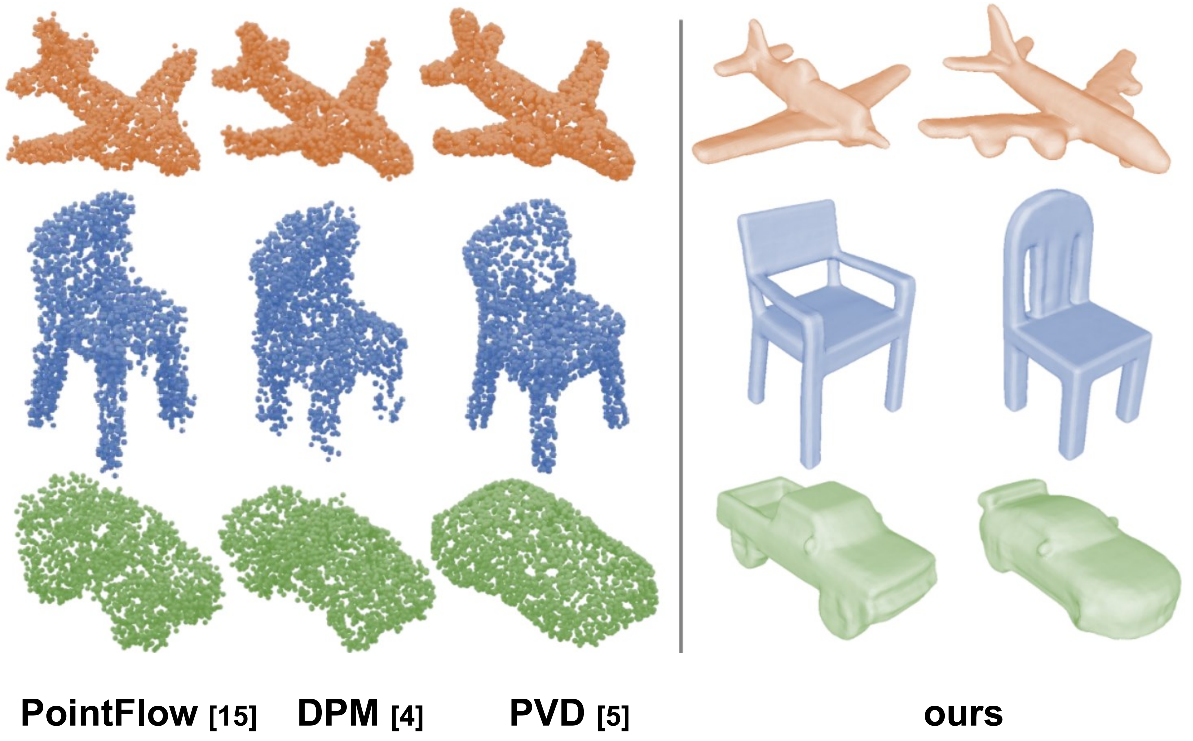
- Stage 2 is trained in **patch-by-patch** to alleviate **memory and computation**.



Experiments

Generation Performance

- Quantitative / qualitative evaluation of **single-category generation**.



		Trainset	PointFlow [15]	DPM [4]	PVD [5]	ours
Airplane	MMD (CD)	1.51	2.43	**2.24**	2.46	2.37
	COV (CD)	56.93	**50.50**	50.00	45.30	50.25
	1-NNA (CD)	45.92	72.28	67.45	62.62	**56.56**
	MMD (EMD)	2.33	1.67	1.64	1.55	**1.49**
	COV (EMD)	57.43	52.97	52.23	53.96	**55.20**
	1-NNA (EMD)	47.40	62.50	63.37	52.72	**48.14**
Car	MMD (CD)	2.44	2.61	2.57	**2.48**	**2.48**
	COV (CD)	54.07	41.92	44.19	44.33	**47.26**
	1-NNA (CD)	49.47	74.50	75.57	58.48	**58.28**
	MMD (EMD)	1.28	1.39	1.39	1.30	**1.28**
	COV (EMD)	52.34	41.92	41.92	48.33	**52.47**
	1-NNA (EMD)	50.53	71.90	71.90	**51.13**	53.20
Chair	MMD (CD)	8.05	8.27	**7.65**	7.87	8.00
	COV (CD)	54.95	46.53	47.86	48.89	**49.78**
	1-NNA (CD)	52.88	70.83	66.40	55.61	**53.69**
	MMD (EMD)	3.57	4.21	4.08	**3.56**	3.61
	COV (EMD)	53.47	49.63	41.65	**50.37**	49.31
	1-NNA (EMD)	48.74	74.74	76.66	53.03	**51.77**

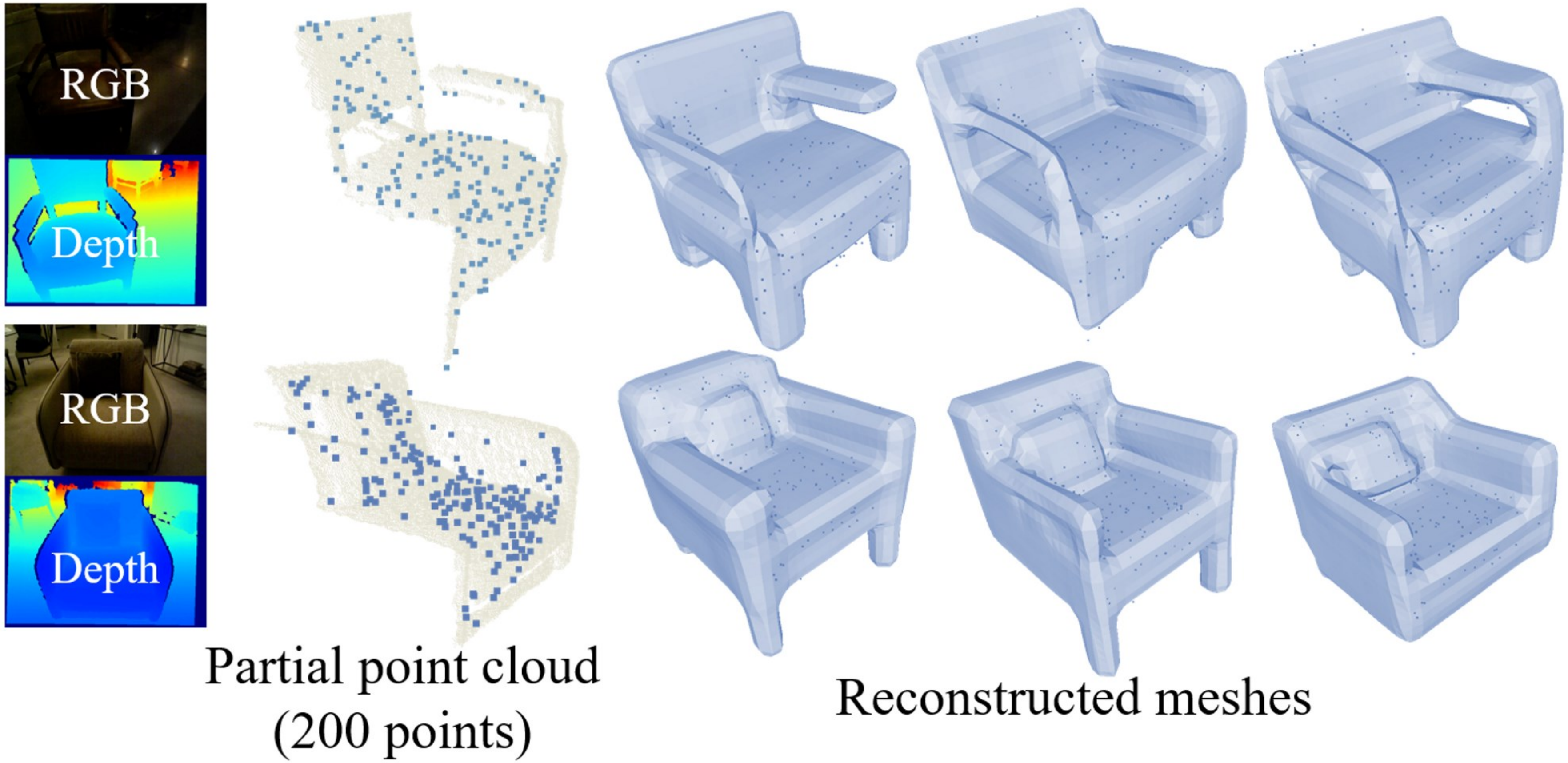
- Our framework **outperforms in most metrics** even though our **output is mesh**, **not point cloud**.

[15] Yang, et al. "Pointflow: 3d point cloud generation with continuous normalizing flows." ICCV 2019.
[4] Luo and Hu. "Diffusion probabilistic models for 3d point cloud generation." CVPR 2021.
[5] Zhou et al. "3d shape generation and completion through point-voxel diffusion." ICCV 2021.

Experiments

Application: Out-of-Distribution Shape Completion

- We test our shape completion method trained with **point cloud condition** about **unseen dataset** [16].



- Our model can generate various 3D shapes even from unseen dataset.

[16] Choi et al. "A large dataset of object scans." arXiv preprint arXiv:1602.02481 (2016).

Conclusion

- New 3D mesh generative model based on diffusion.
- Comparable results than SoTA with more complicated representation (point v.s. mesh)

Research Project 2023.09 ~ 2023.12

3. ContactGen:

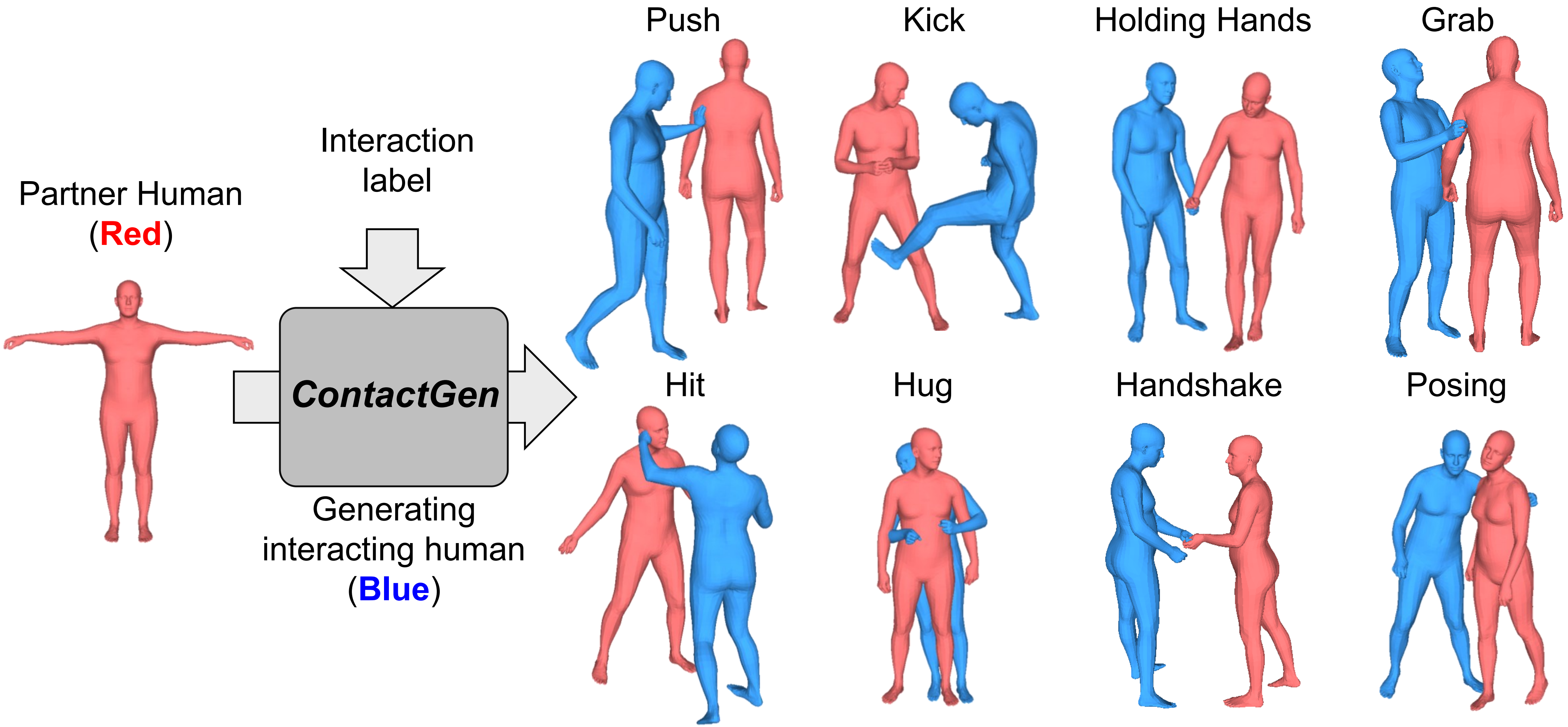
Contact-Guided Interactive 3D Human Generation for Partners

Dongjun Gu, Jaehyeok Shim, Jaehoon Jang, Changwoo Kang, Kyungdon Joo

AAAI 2024

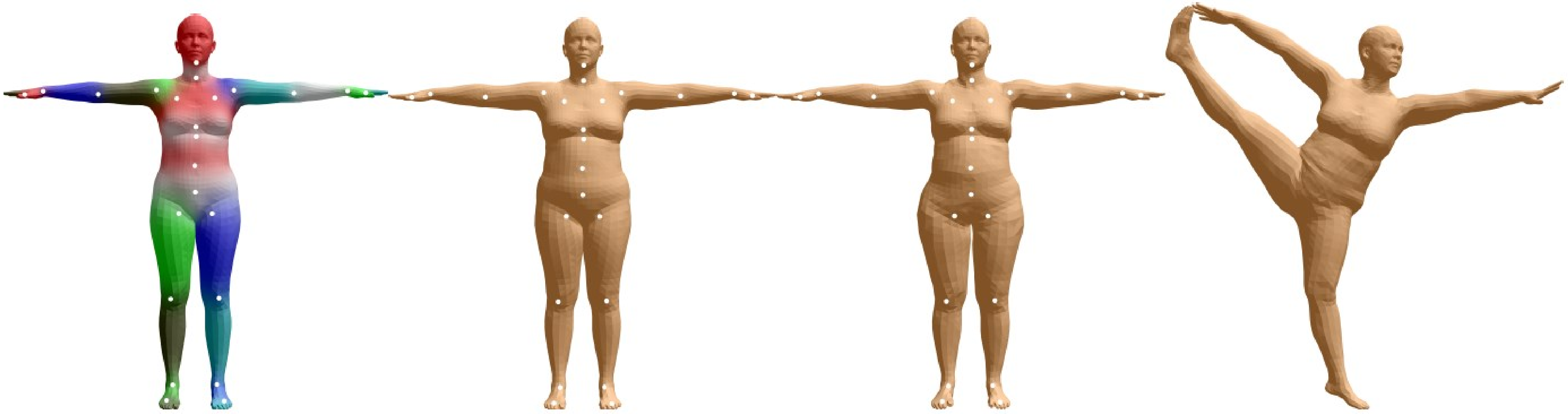
Introduction

Problem Statement



Background

SMPL-X[1] Parameters



(a) $\bar{\mathrm{T}}, \mathcal{W}$ (b) $\bar{\mathrm{T}} + B_S(\vec{\beta}), J(\vec{\beta})$ (c) $T_P(\vec{\beta}, \vec{\theta}) = \bar{\mathrm{T}} + B_S(\vec{\beta}) + B_P(\vec{\theta})$ (d) $W(T_P(\vec{\beta}, \vec{\theta}), J(\vec{\beta}), \vec{\theta}, \mathcal{W})$

Body parameters:
translation
global rotation
joint rotation
body shape
…

→ SMPL →

SMPL mesh
(vertices & faces)

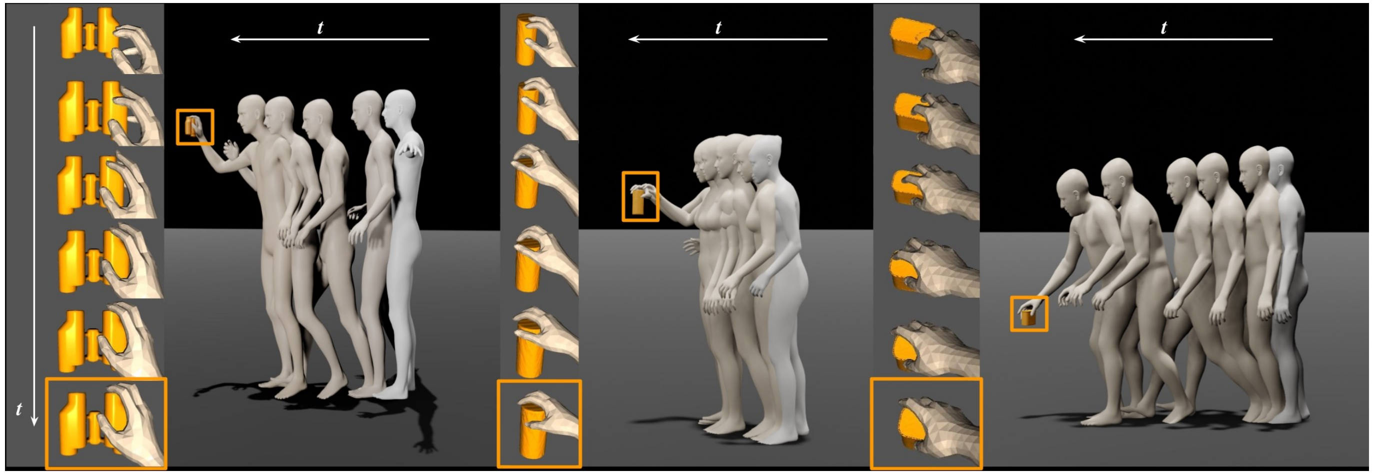
[1] Pavlakos, Georgios, et al. "Expressive body capture: 3d hands, face, and body from a single image." CVPR 2019.

Related Works

Previous Works: Human-Object Interaction

- **SAGA[2] (ECCV 2022):**
  - Generates human to grasp a given object.
- **SceneDiffuser[3] (CVPR 2023):**
  - Generates human conditioned to scene.

Biased to contact with simple and static object.



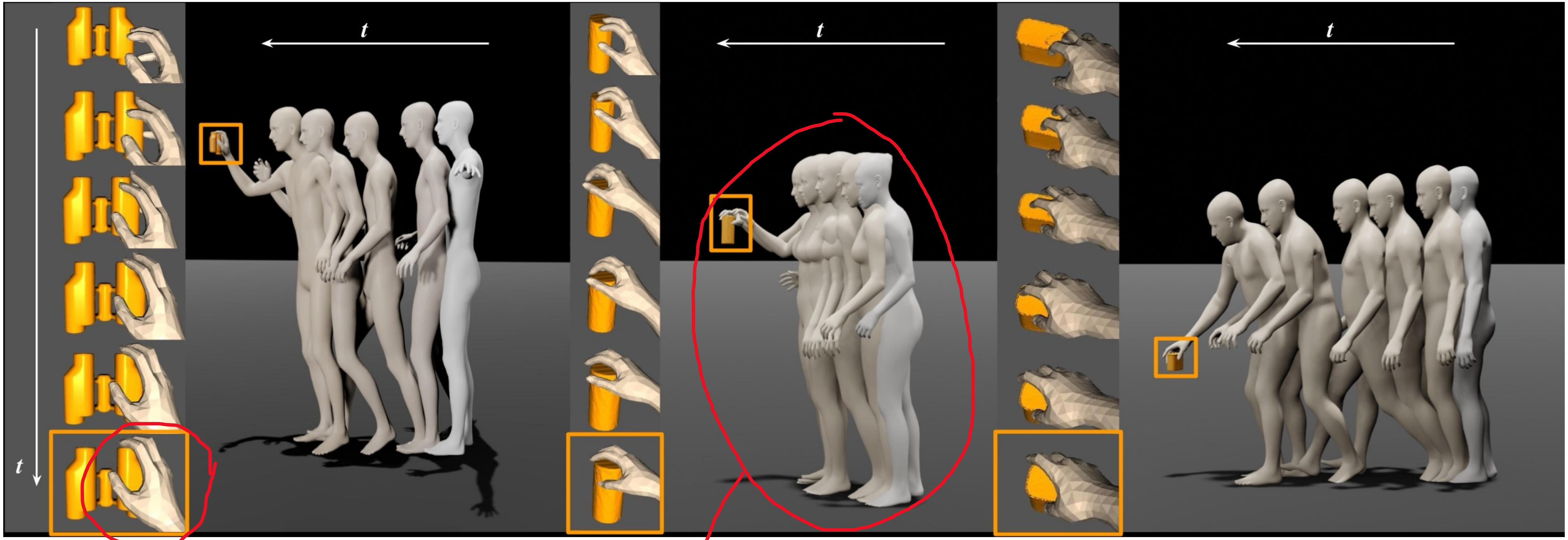
SAGA

SceneDiffuser

[2] Wu, Yan, et al. "Saga: Stochastic whole-body grasping with contact." ECCV 2022.
[3] Huang, Siyuan, et al. "Diffusion-based generation, optimization, and planning in 3d scenes." CVPR 2023.

Related Works

Limitation of Human-Object Interaction



contact area is limited

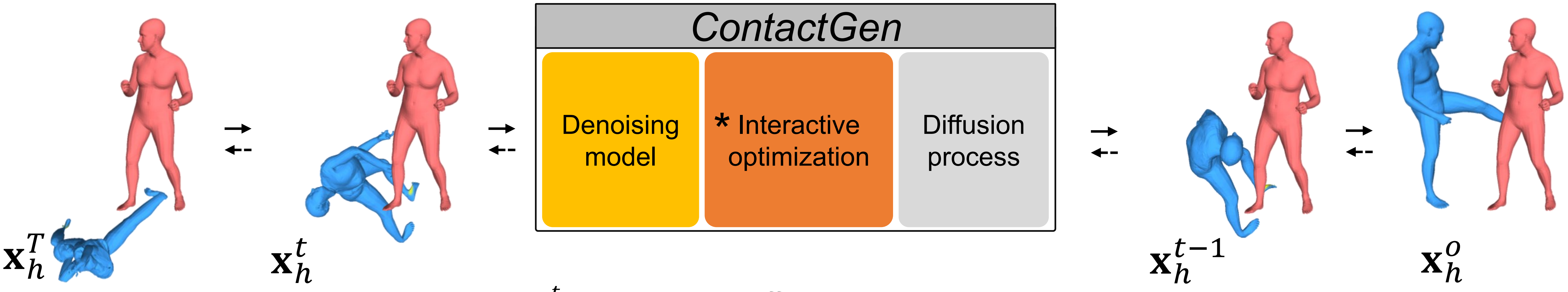
pose deformation is simple and small

[2] Wu, Yan, et al. "Saga: Stochastic whole-body grasping with contact." ECCV 2022.

Methods

ContactGen: Method Overview

- Generate new human pose based on given partner's pose with diffusion models.



$\mathbf{x}_h^t$: human at diffusion timestep t

$\mathbf{x}_p$: partner human

$\mathbf{l}$: interaction label

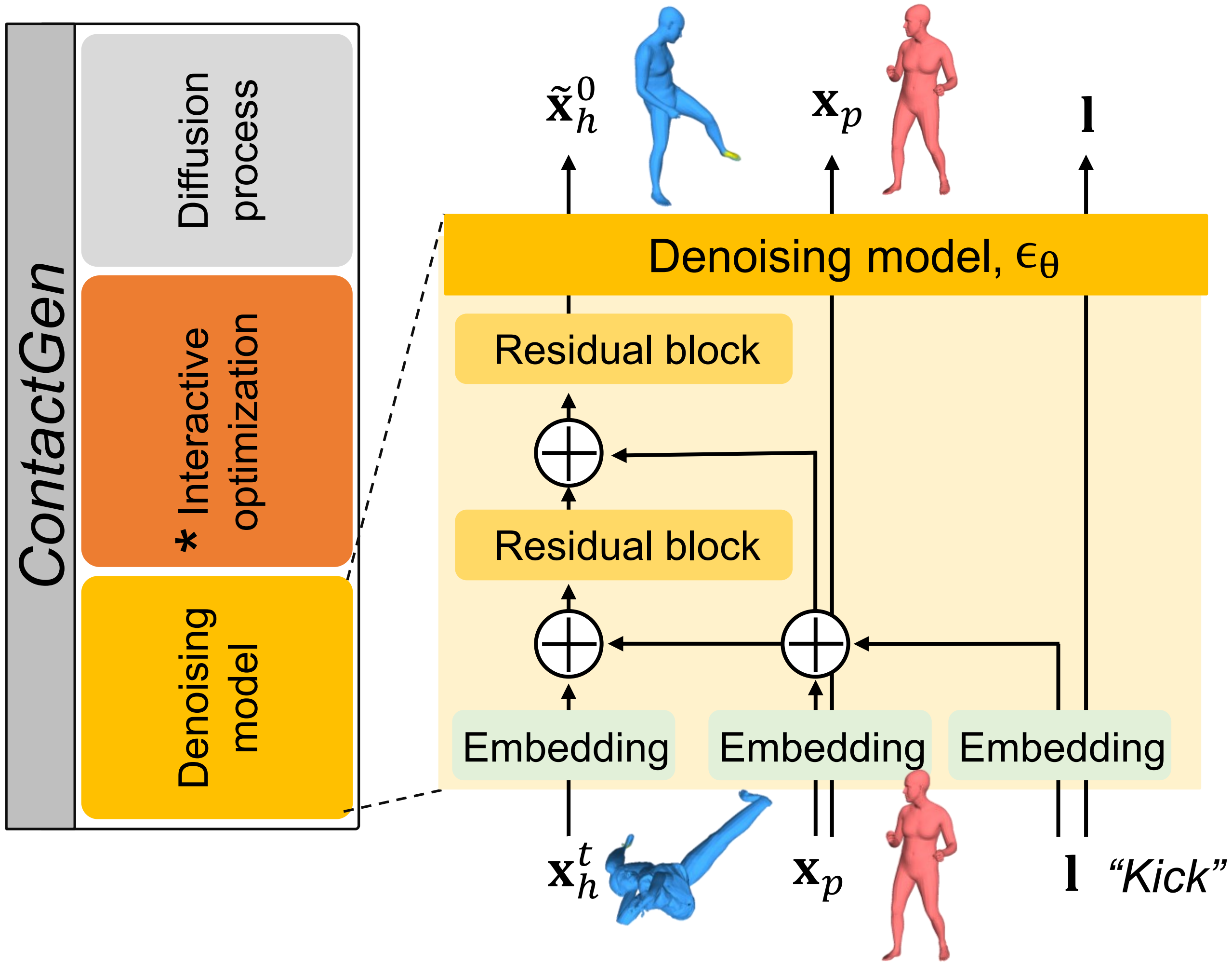
⋯→ : forward process

⟶ : reverse process

* : sampling only

Methods

ContactGen: Denoising Model



Training Phase

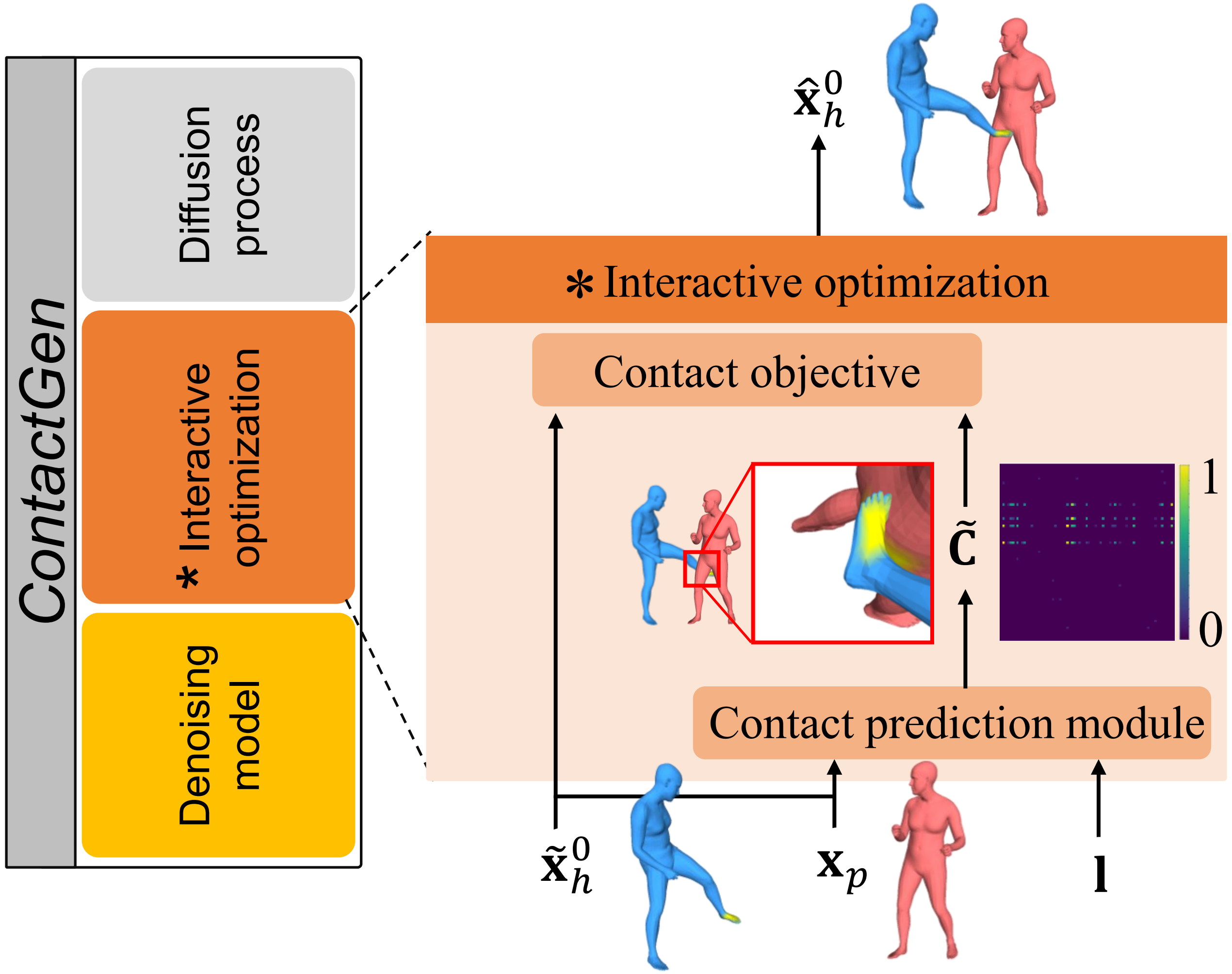
- Diffusion model ϵ_θ learns how to denoise SMPL.
- Predict SMPL noisy: $\tilde{\epsilon} = \epsilon_\theta\left(\mathbf{x}_h^t, t, \mathbf{x}_p, \mathbf{l}\right)$
- Calculate loss: $\mathcal{L} = \sum_i \frac{1}{\mathrm{N}_i} \|\tilde{\epsilon}_i - \epsilon_i\|_2^2$

Inference Phase

- Predict denoised SMPL:
  - $\tilde{\mathbf{x}}_h^0 = f_\theta\left(\mathbf{x}_h^t, t, \mathbf{x}_p, \mathbf{l}\right) = \frac{\mathbf{x}_h^t - \sqrt{1-\hat{\alpha}_t}\,\epsilon_\theta(\mathbf{x}_h^t, t, \mathbf{x}_p, \mathbf{l})}{\sqrt{\hat{\alpha}_t}}$

Methods

ContactGen: Interactive Optimization



Inference Phase

- Contact prediction module estimates contact $\tilde{\mathbb{C}}$.
- Denoised human $\tilde{\mathbf{x}}_h^0$ is updated contact objective $\mathcal{O}$.
- Guidance $\boldsymbol{g}$ is applied with as a gradient form following Guided Diffusion[5].

- $\mathcal{O}_{contact}(\mathbf{x}_h, \mathbf{x}_p, \mathbf{l}) = \sum_{(r_i, r_j) \in \mathbb{C}} d_{CD}(V_h^{r_i}, V_p^{r_j})$
- $\boldsymbol{g} = \nabla_{\mathbf{x}_h} \mathcal{O}(\mathbf{x}_h, \mathbf{x}_p, \mathbf{l})\big|_{\mathbf{x}_h = \tilde{\mathbf{x}}_h^0}$
- $\hat{\mathbf{x}}_h^0 = \tilde{\mathbf{x}}_h^0 - \lambda \odot \boldsymbol{g}$

[5] Dhariwal, Prafulla, and Alexander Nichol. "Diffusion models beat gans on image synthesis." NeurIPS 2021.

Results

Quantitative Comparison

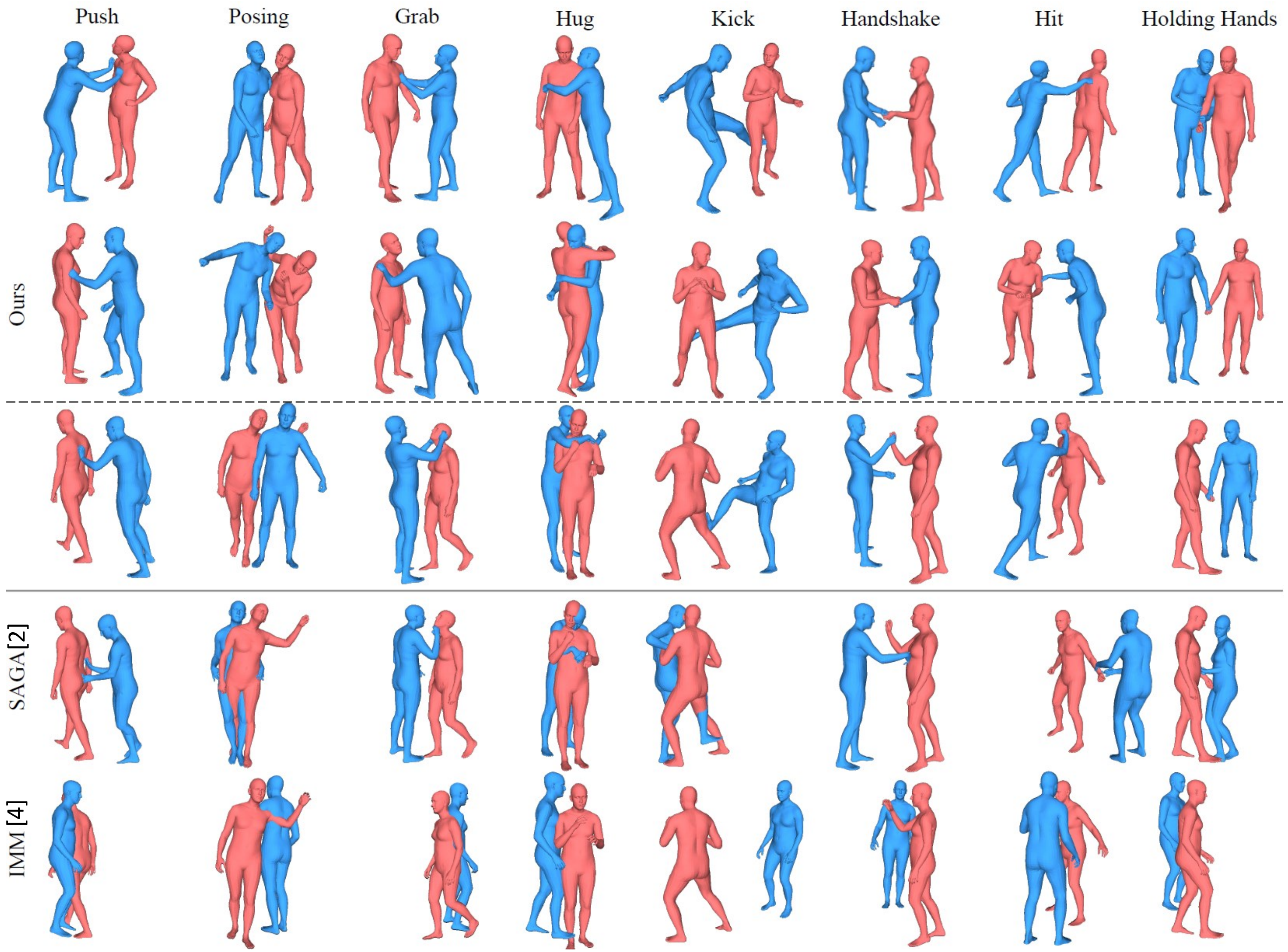
Interaction	IMM [4]					SAGA [2]					Ours				
	FHID↓	top-1↑	top-3↑	contact↓	non-coll↑	FHID↓	top-1↑	top-3↑	contact↓	non-coll↑	FHID↓	top-1↑	top-3↑	contact↓	non-coll↑
Push	51.00	0.00	82.89	17.52	94.08	**34.54**	18.43	72.74	7.44	**95.10**	56.78	**87.52**	**98.85**	**3.61**	94.77
Posing	174.78	11.90	23.13	14.81	**93.54**	**53.23**	62.50	94.51	**4.33**	92.15	71.70	**84.15**	**99.54**	8.58	92.72
Grab	59.23	**92.78**	**100.00**	13.26	96.98	46.70	84.41	94.06	5.95	**97.00**	**28.88**	91.32	96.92	**4.76**	96.50
Hug	235.84	18.72	34.76	12.45	**88.17**	57.16	93.38	99.24	**3.16**	76.21	**21.43**	**99.75**	**100.00**	5.96	78.71
Kick	29.89	66.67	92.59	10.01	**99.70**	**42.04**	30.17	64.22	**4.62**	97.41	46.55	**95.69**	**100.00**	11.98	98.49
Handshake	146.90	11.90	45.24	16.14	99.10	34.39	66.12	92.12	**4.14**	98.16	**12.97**	**92.12**	**99.63**	7.34	**99.13**
Hit	43.01	61.40	**100.00**	10.94	**98.52**	44.10	50.60	96.99	6.22	96.98	**20.38**	**67.47**	99.40	**5.16**	96.55
Holding Hands	499.84	0.00	0.00	15.17	96.80	54.07	73.93	92.62	**4.27**	96.41	**26.20**	**99.02**	**100.00**	6.42	**97.12**
All	92.25	34.49	58.93	14.24	94.41	24.26	67.47	90.76	**5.23**	94.62	**6.15**	**90.30**	**98.67**	6.07	**94.89**

- ContactGen shows better qualitative and quantitative results than baseline methods: IMM[6] and SAGA[2].

[2] Wu, Yan, et al. "Saga: Stochastic whole-body grasping with contact." ECCV 2022.
[4] Goel, Aman, Qianhui Men, and Edmond SL Ho. "Interaction Mix and Match: Synthesizing Close Interaction using Conditional Hierarchical GAN with Multi-Hot Class Embedding." Computer Graphics Forum 2022.

Methods

Qualitative Comparison



[2] Wu, Yan, et al. "Saga: Stochastic whole-body grasping with contact." ECCV 2022.
[4] Goel, Aman, Qianhui Men, and Edmond SL Ho. "Interaction Mix and Match: Synthesizing Close Interaction using Conditional Hierarchical GAN with Multi-Hot Class Embedding." Computer Graphics Forum 2022.

Conclusion

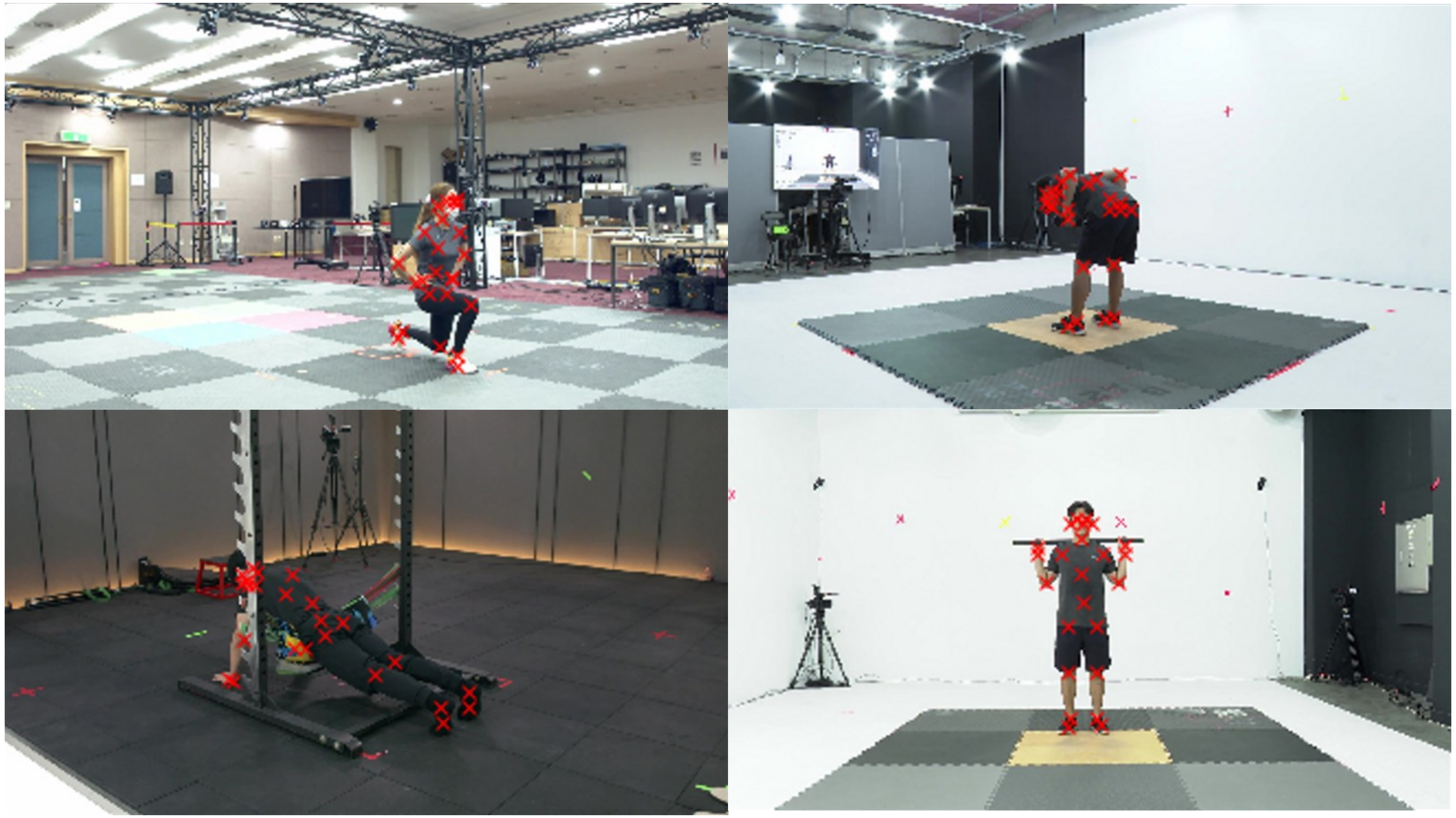
- New task suggested, human generation for partner.
- New human generative model based on diffusion.
- Outperforming baseline methods about human generation with contact.

4. Various AI Challenges

DACON: Human Pose Estimation Challenge

Problem Statement

- **Goal**: Find human keypoint from given image.
- **Input**: Image, **Output**: Human Keypoint

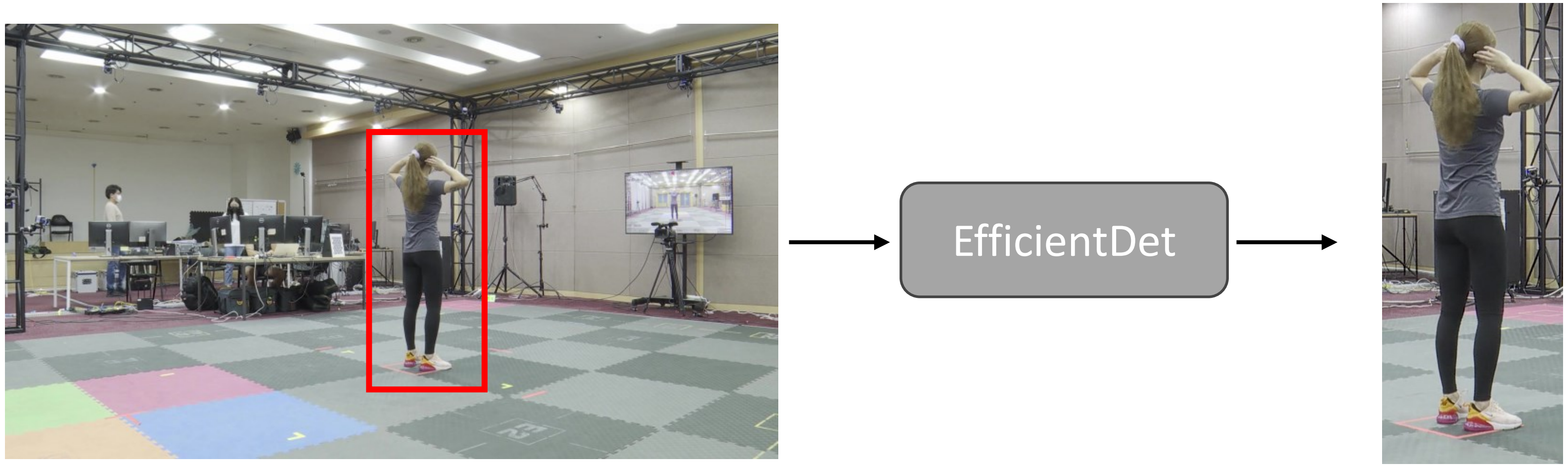


Notion: https://www.notion.so/AI-54065ab4da62480f9ea595373c6f6454
Github: https://github.com/Kitsunetic/motion-keypoint

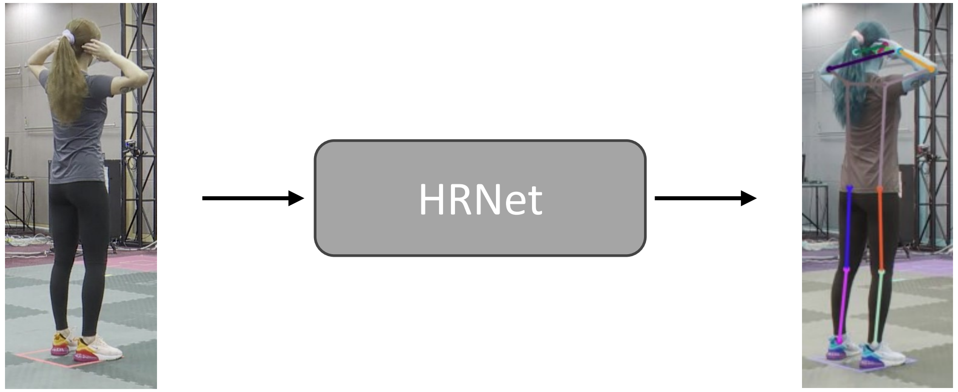
DACON: Human Pose Estimation Challenge

Approaches

1. Detect humans using pretrained EfficientDet[1], crop RoI.



2. Detect human keypoints by finetuning pretrained HRNet[2].

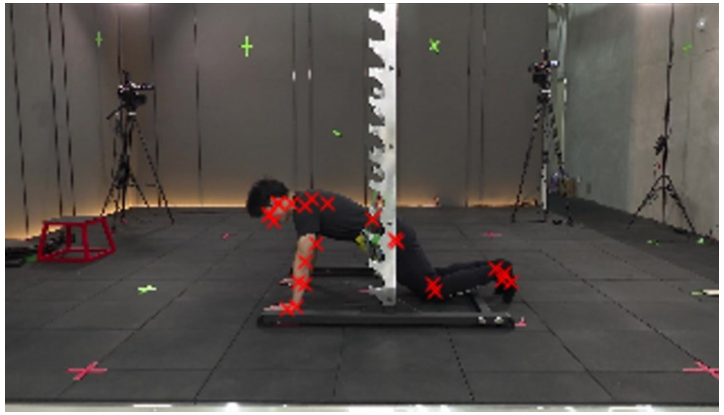


[1] Tan, Mingxing, Ruoming Pang, and Quoc V. Le. "Efficientdet: Scalable and efficient object detection." CVPR 2020.
[2] Sun, Ke, et al. "Deep high-resolution representation learning for human pose estimation." CVPR 2019.

DACON: Human Pose Estimation Challenge

Challenge

1. **Occlusion**: Occlusion bothers pose model detecting keypoints.



2. **Irregular bbox aspect ratio**: aspect ratio of person largely differs regarding to pose.

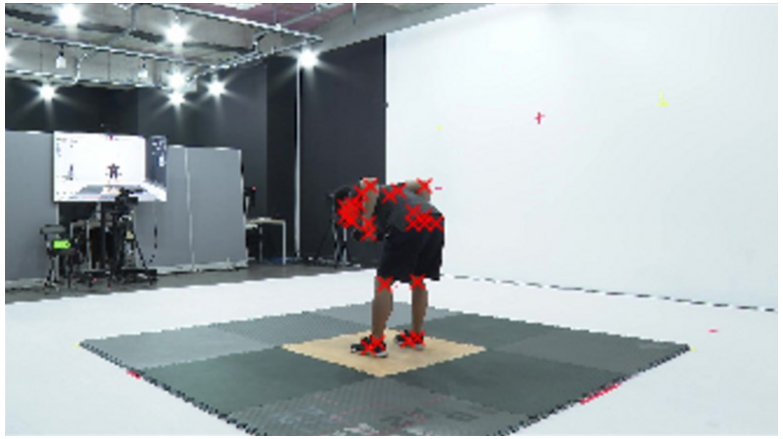
Outputs of detection model

Warped images has large distortion

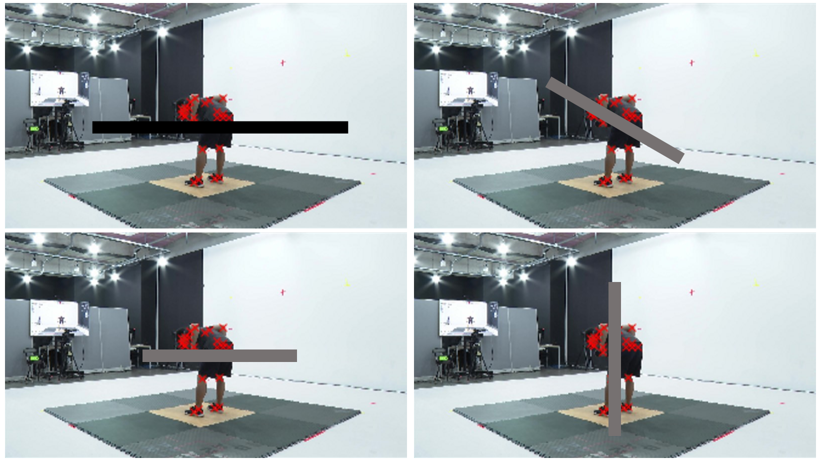
DACON: Human Pose Estimation Challenge

Approaches

1. **Occlusion**: Propose new augmentation imitating occlusions.



Augmentation



2. **Irregular bbox aspect ratio**: Padding and resizing instead of warping to reduce distortion.

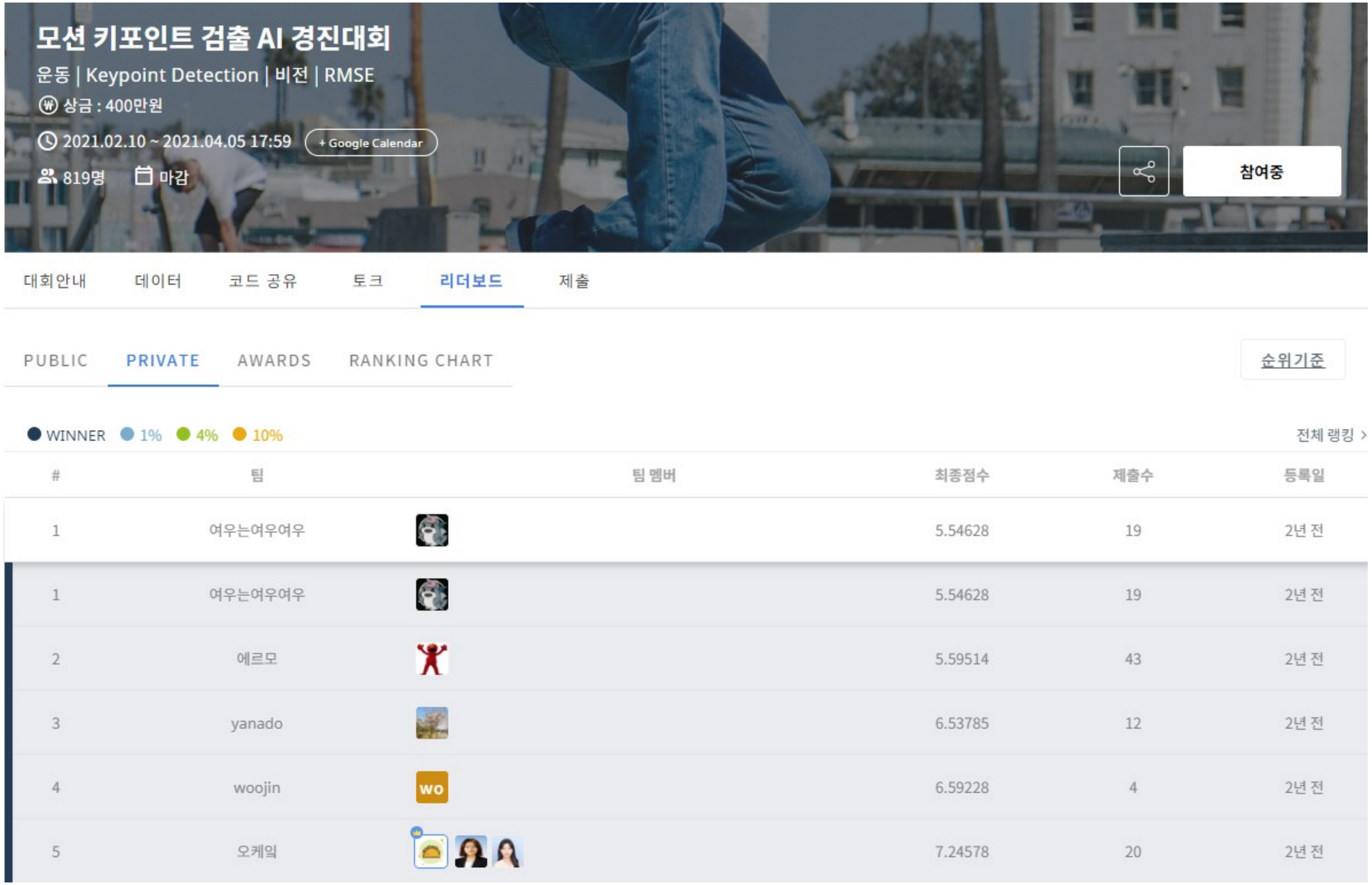
Padding/resize to make square

Outputs of detection model **Padded image has less distortion**

DACON: Human Pose Estimation Challenge

Result

1st place win against 819 participants.



#	팀	팀 멤버	최종점수	제출수	등록일
1	여우는여우여우		5.54628	19	2년 전
1	여우는여우여우		5.54628	19	2년 전
2	에르모		5.59514	43	2년 전
3	yanado		6.53785	12	2년 전
4	woojin		6.59228	4	2년 전
5	오케잌		7.24578	20	2년 전

Other Experiences

- Participated various AI challenges alone.
- Win various challenges especially in DACON.
- Total 4th place in DACON when highest rank.

Competition Gold

34
of 86,676
현재 랭킹

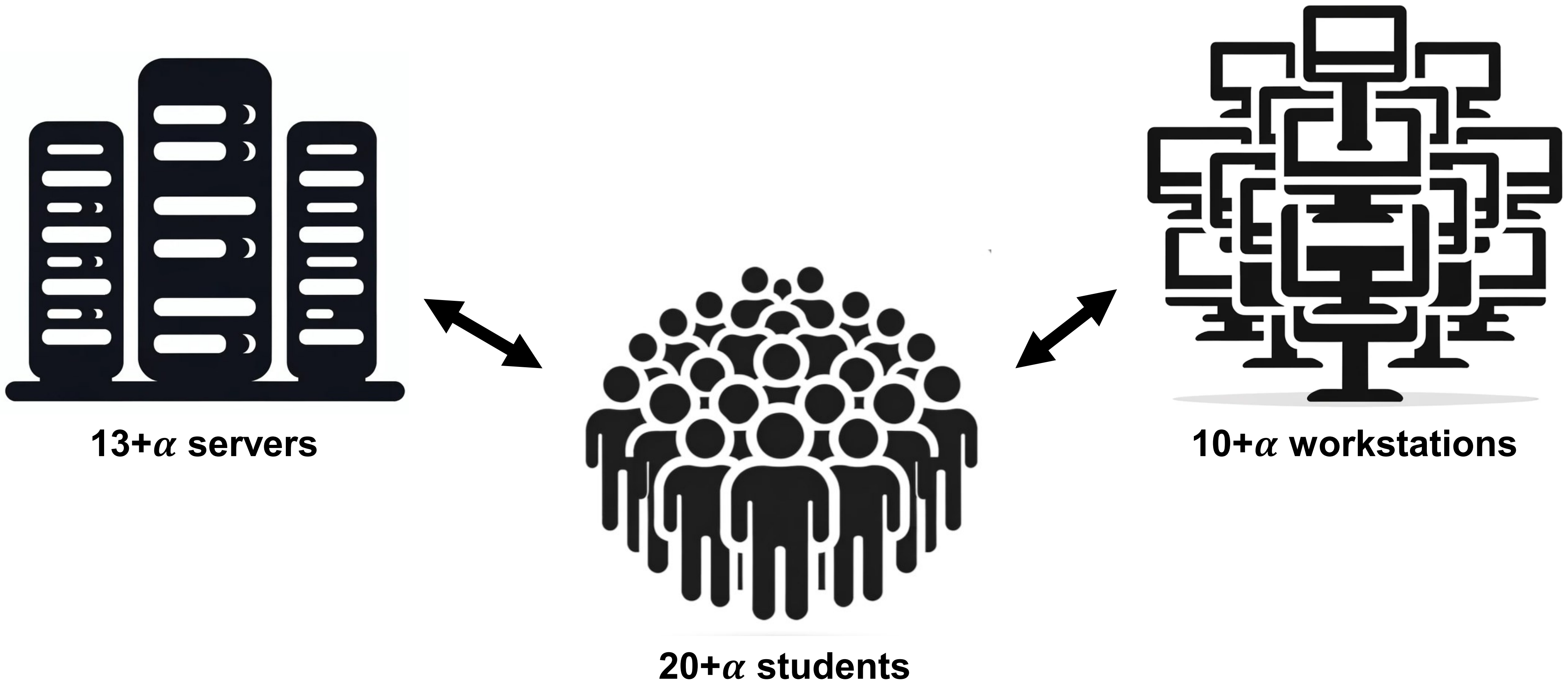
4
최고 랭킹

- OCR Challenge , from KYOWON Group with DACON, 2022.12, 7th place win from 430 players (top 2%). OCR task of Korean language. I improved accuracy with transfer learning of ConvNeXT by proposing language specific loss.
- NAVER CLOVA AI-RUSH 2022 Round 2 , from NAVER CLOVA, 2022.08, 7th place from 15 players. Regress specific scores from given images. I improved accuracy with transfer learning of CoAT with various augmentation.
- NAVER CLOVA AI-RUSH 2022 Round 1 , from NAVER CLOVA, 2022.07, 15th place win from 27 players. Classifies given images. I improved accuracy through transfer learning of EfficientNetV2 with various augmentations.
- Ego-Vision Hand Gesture Recognition AI Contest , from DACON, 2021.06, 3th place win from 290 players (top 1%). Classifies hand gesture from given images. I achieved high-accuracy with transfer learning of EfficientNetV2 with cross validation.
- News Topic Classification AI Contest , from DACON, 2021.05, 3 th place win from 256 players (top 1%). This competition is to classify topics of given text articles. I improved accuracy with noisy student training strategy about BeRT-based model.
- NAVER CLOVA AI-RUSH 2021 Round2 , from NAVER CLOVA, 2021.05, 6th place from 13 players. Clustering of given text dataset. I improved model performance with self-supervised learning.
- NAVER CLOVA AI-RUSH 2021 Round1 , from NAVER CLOVA, 2021.04, 4th place win from 35 players. Classification of given image dataset with limited model capacity. I achieved high-accuracy with transfer learning of EfficientNetV2 with careful hyper-parameter tuning.
- Predicting Danger of System Log Messages , from DACON, 2021.04, 2th place win from 152 players (top 1%). Finding out-of-distribution data not appeared in training dataset. I achieved high-accuracy with DistilBeRT-based anomaly detection.
- Finding Human Key-Points from Motion Images , from DACON, 2021.02, 1th place win from 156 players (top 0.6%). Estimating human key-points from given image dataset. I achieved high-accuracy with fine-tuning HRNet for pose estimator and EfficientDet for human detector, and with data-driven augmentations.
- The 2th Computer Vision AI Contest , from DACON, 2021.02, 30th place from 216 players (top 13%).
- Motion Classification AI Contest , from DACON, 2021.01, 21th place from 336 players (top 6%). Classifies human motions from acceleration of human body parts captured with various sensors.
- 정보통신대학 ICT AI 및 알고리즘 경진대회 , from SNUT, 2021.01, 1th place win.
- 태양광 발전량 예측 AI 경진대회 , from DACON, 2020.11, 94 place from 461 players (top 20%).
- 공공데이터 활용 수력 댐 강우예측 AI 경진대회 , from DACON, 2020.10, 39 place from 132 players (top 29%).
- 컴퓨터 비전 학습 경진대회 , from DACON, 2020.08, 11th place from 356 players (top 3%).
- NAVER CLOVA AI-RUSH 2020 Round1 , from NAVER CLOVA, 2020.06, 37th place.
- 위성관측 데이터 활용 강수량 산출 AI 경진대회 , from DACON, 2020.05, 72th place from 213 players (top 33%).

5. Server Management Experience

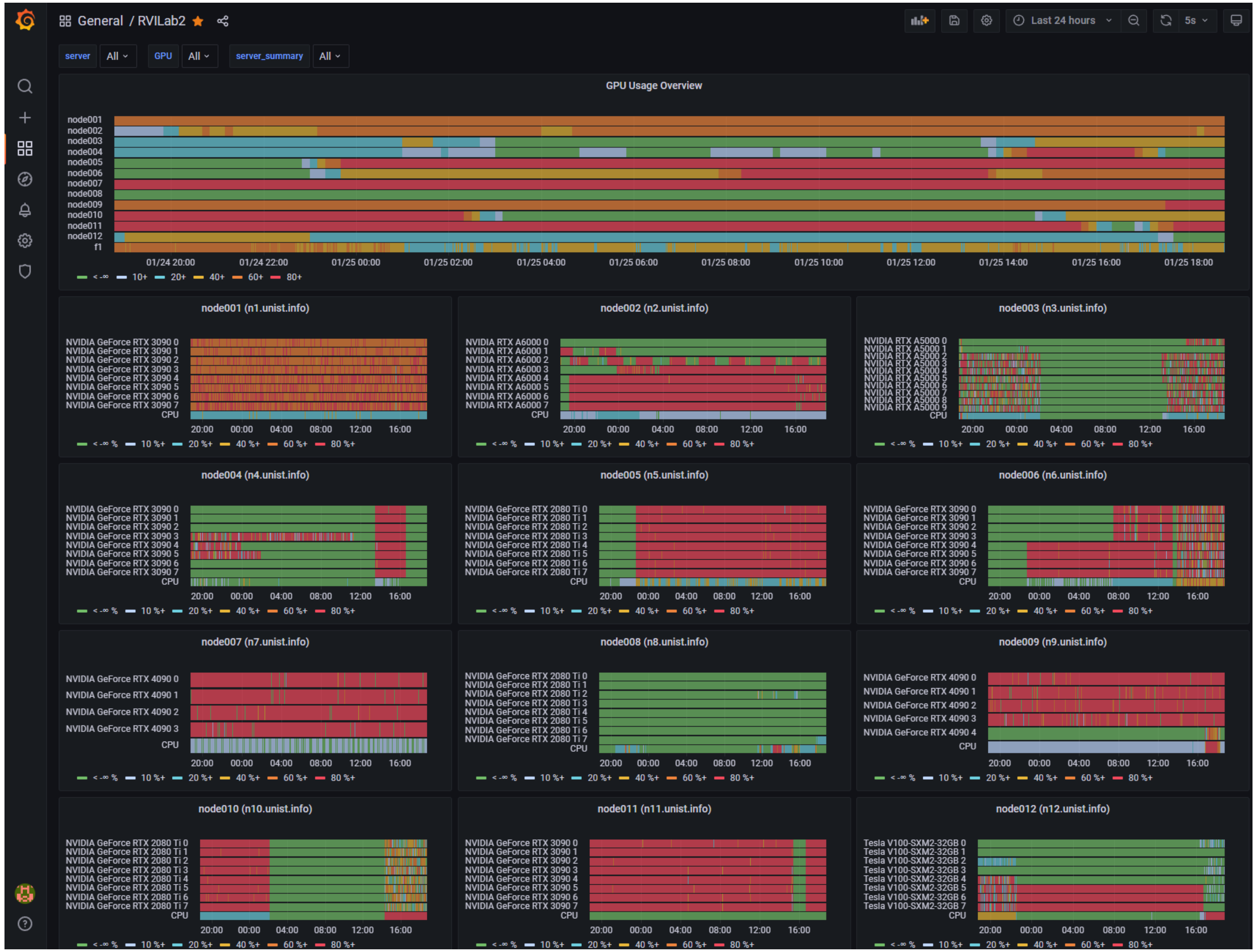
Lab Server Managing

- For more than 2 years, manager rule for lab servers.
- There are more than 20 students, 13 servers, 10 workstations.



Integrated Monitoring System

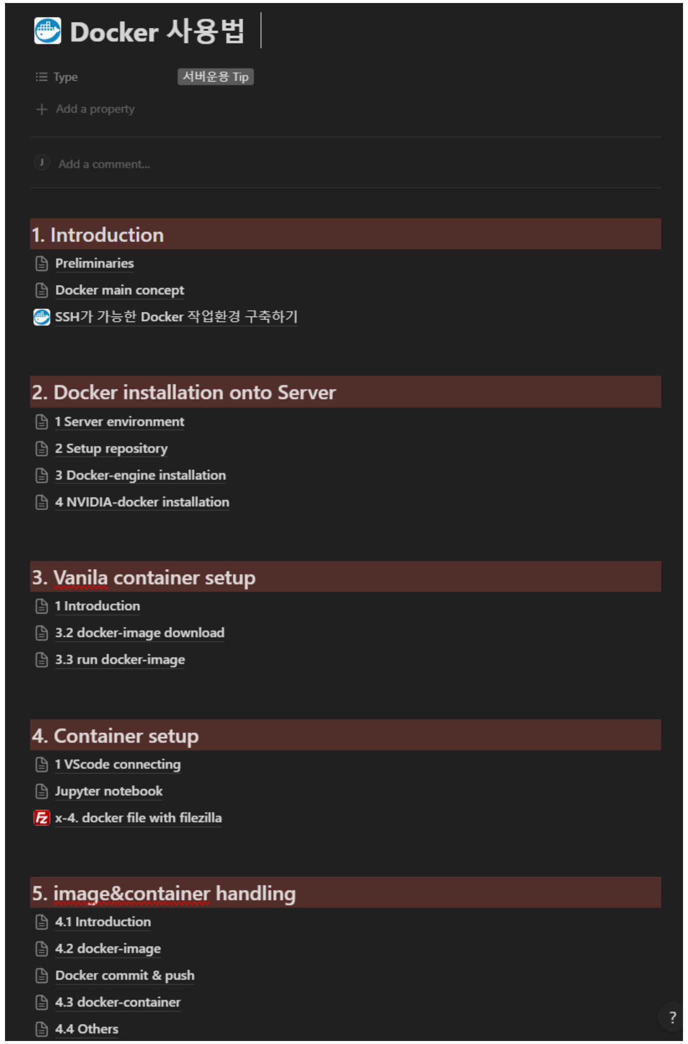
- For convenience of managing, have built monitoring system using Prometheus and Grafana.



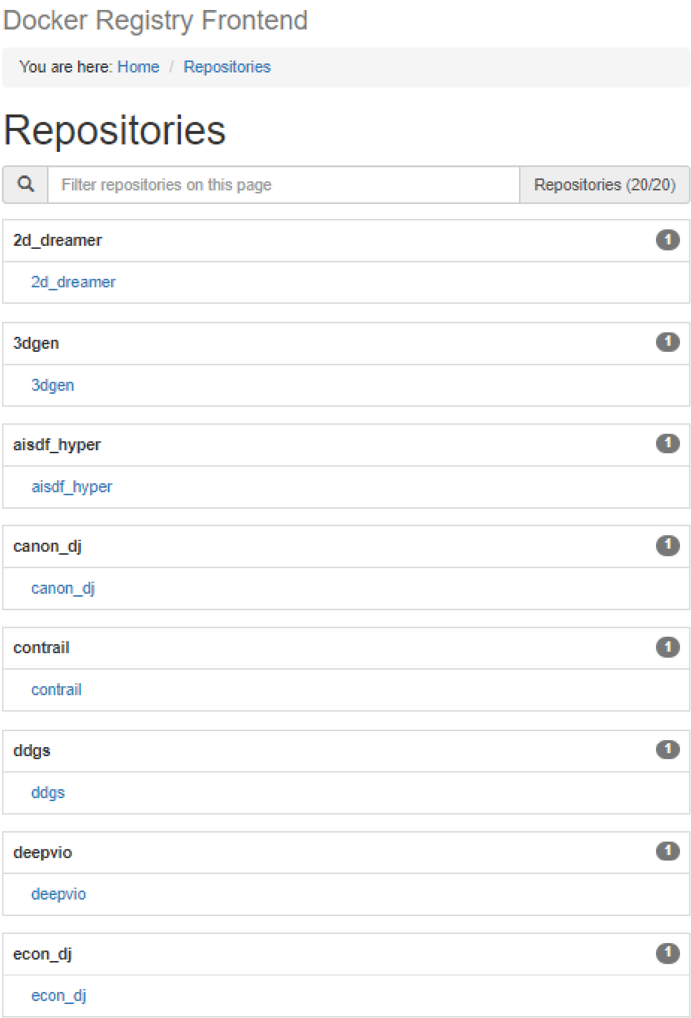
Grafana page

Integrated Monitoring System

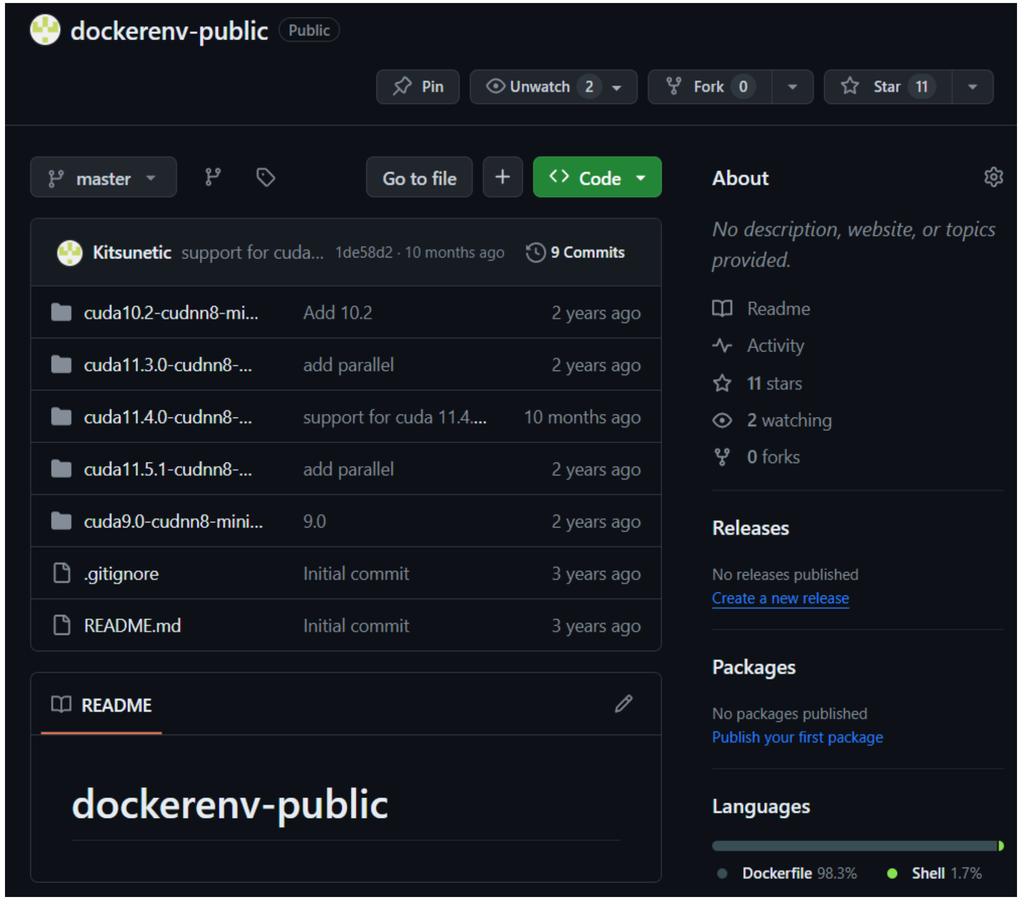
- For students ease of use, have built guide, custom docker image and docker hub.



Server usage guide



Custom docker hub



Custom docker image base

Thank You