NEURAL INFORMATION PROCESSING SYSTEMS

PAD: A Dataset and Benchmark for Pose-agnostic Anomaly Detection

NeurIPS 2023 Datasets and Benchmarks Track

Qiang Zhou[1*], **Weize Li[1*](Presentor)** ,Lihan Jiang[2], Guoliang Wang[1], Guyue Zhou[1], Shanghang Zhang[3], Hao Zhao[1]

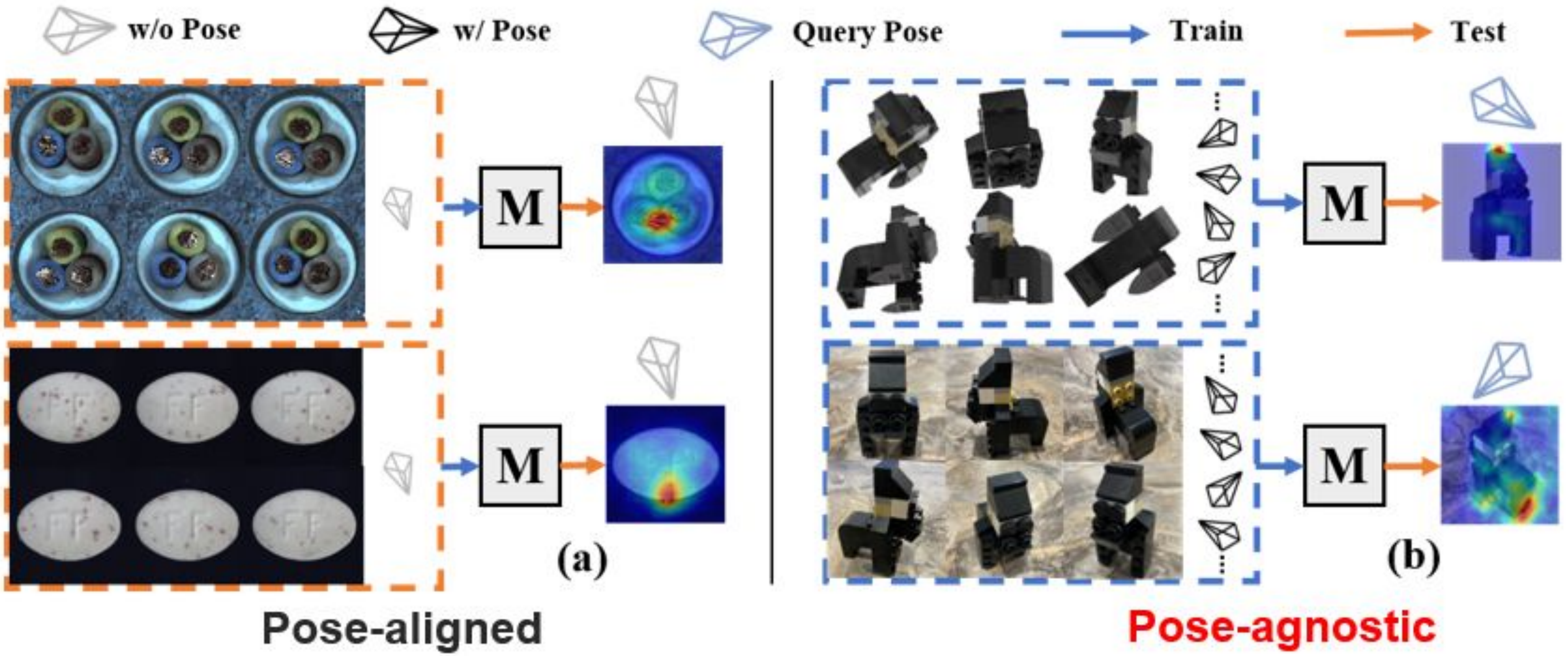
[1]AIR, Tsinghua University, [2]Wuhan University, [3] Peking University

清華大学 智能产业研究院
Institute for AI Industry Research, Tsinghua University

武漢大學 WUHAN UNIVERSITY

北京大學 PEKING UNIVERSITY

Motivation


w/o Pose
w/ Pose
Query Pose
Train
Test
M
M
(a)
M
M
(b)
Pose-aligned
Pose-agnostic

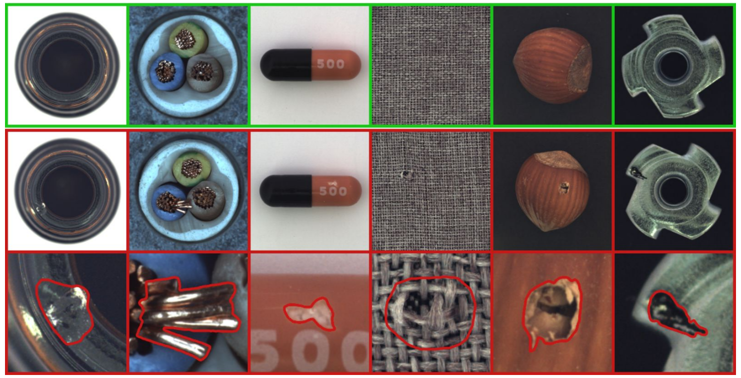
Contribution #1: Pose-agnostic Anomaly Detection Setting

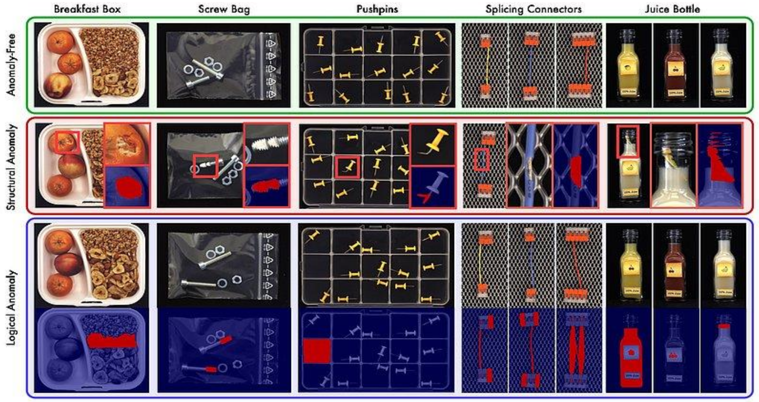
- For object anomaly detection from diverse viewpoints or poses.
- Breaking free from the constraints of stereotypical pose-aligned setting.
- Taking a step forward to practical anomaly detection and localization tasks.
- PAD setting:

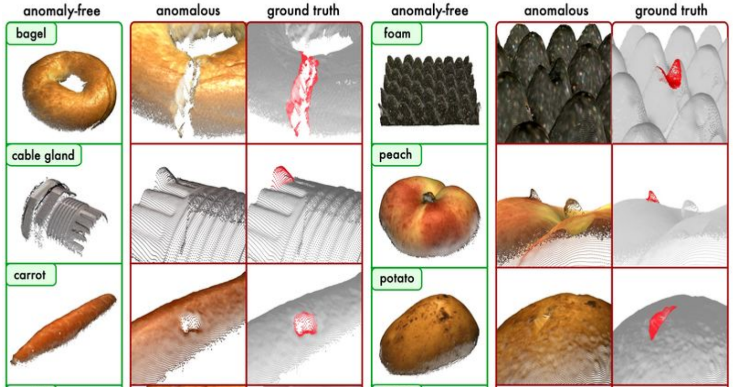
Given a training set $\mathcal{T} = \{t_i\}_{i=1}^{N}$, in which $\{t_1, t_2, \cdots, t_N\}$ are the anomaly-free samples from object's multi pose view and each sample t consists of a RGB image I_{rgb} w/ pose information θ_{pose}. In addition, $\mathcal{T}$ belongs to certain object o_j, $o_j \in \mathcal{O}$, where $\mathcal{O}$ denotes the set of all objects categories. During testing, given a query (normal or abnormal) image $\mathcal{Q}$ from object o_j w/o pose information θ_{pose}, the pre-trained AD model M should discriminate whether or not the query image $\mathcal{Q}$ is anomalous and localize the pixel-wise anomaly region if the anomaly is detected.

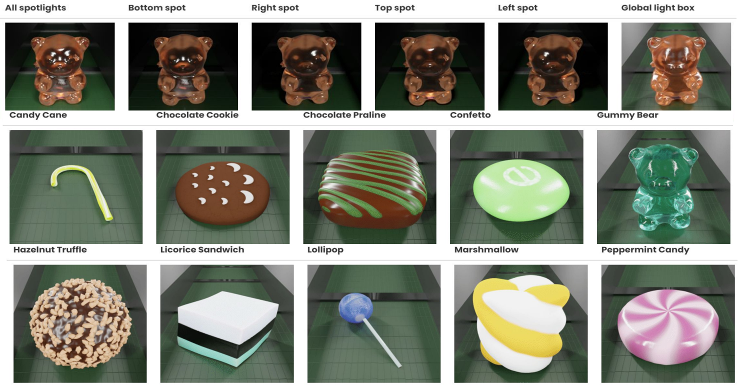
Challenges:

1) We lack anomaly detection datasets from multiple pose views of an object.







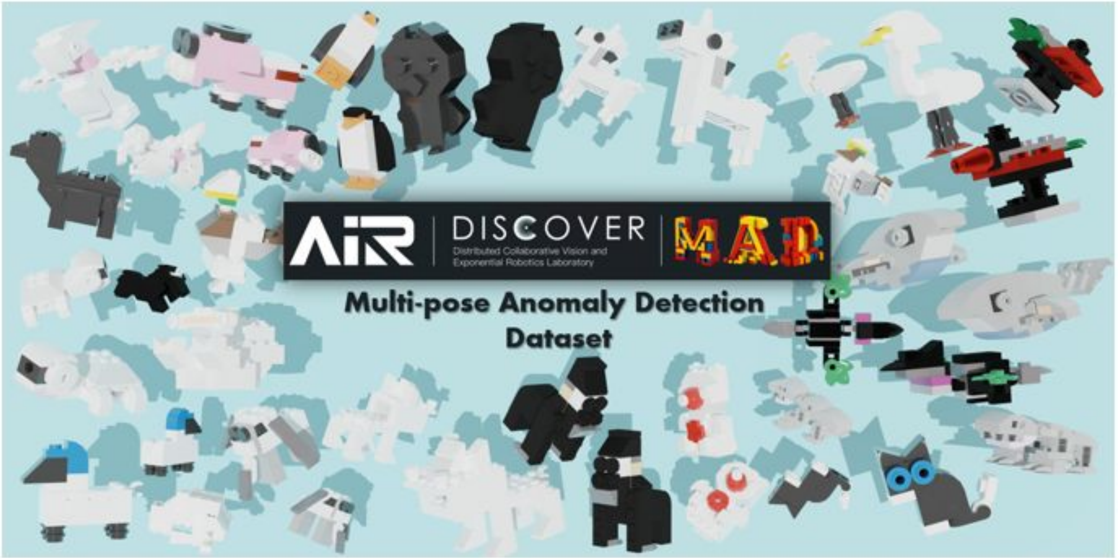


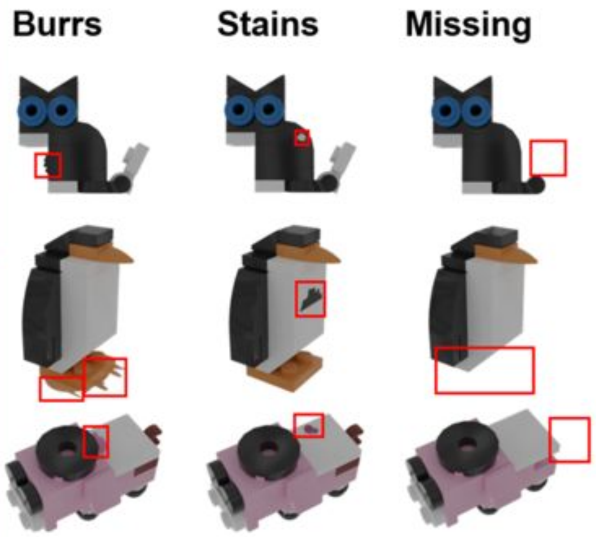
- w/o Multi-pose views
- w/o Pose annotations
- Do not want to use point cloud data that is difficult and expensive to collect.

……

Contribution #2: Multi-pose Anomaly Detection Dataset

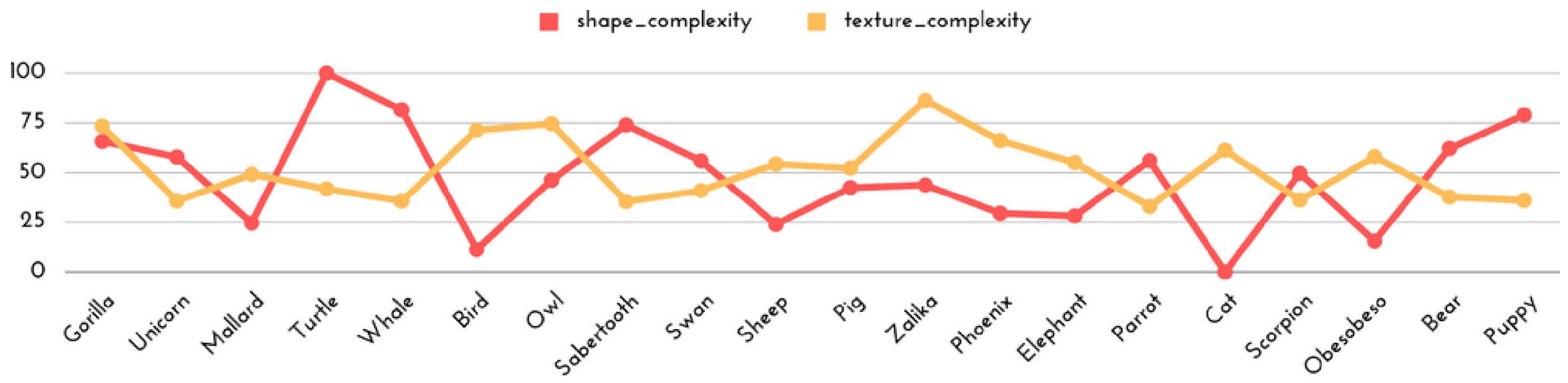
- Selected **20 Lego toys** with diverse shapes and colors.
- **Provided 11,000+** high-resolution RGB images from multi-pose view.
- Divided into **MAD-Simulated Set + MAD-Real Set**.
- Supported by **Pixel-precise** Ground Truth for **3 types of defects**.





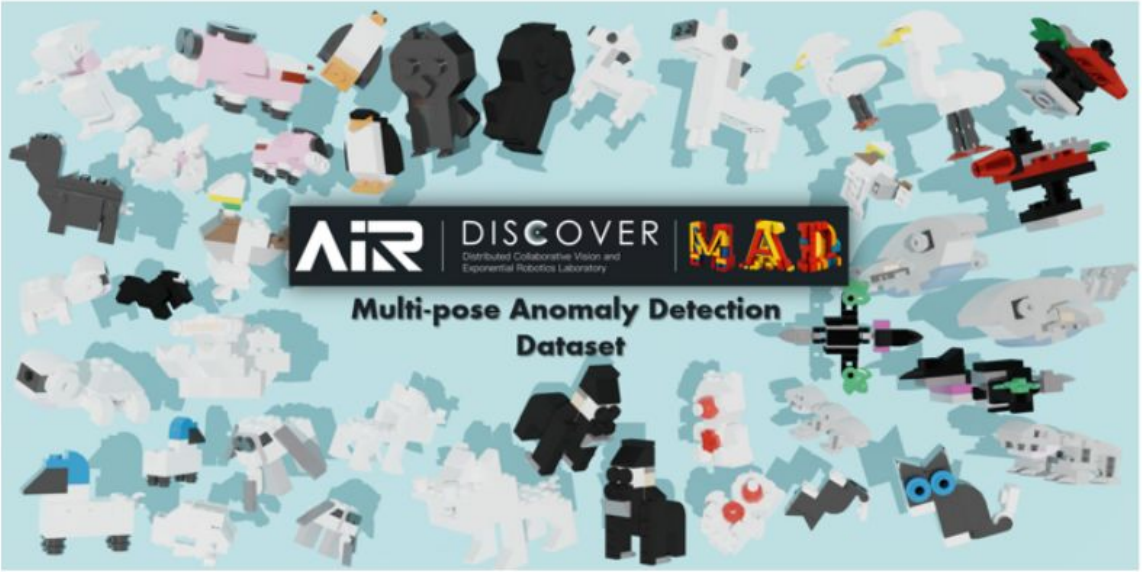
Contribution #2: Multi-pose Anomaly Detection Dataset

- Quantified the **attributes** of each Lego toy in **shape** complexity and **texture** complexity.



Challenges:

2) We lack criteria for evaluating the performance of pose-agnostic anomaly detection models.



But fair benchmarking？

Contribution #3: PAD Benchmark

Feature Embedding-based: Patchcore, STFPM, Fastflow, CFlow, Cutpaste

Reconstruction-based: DRAEM, FAVAE, OCR-GAN, UniAD

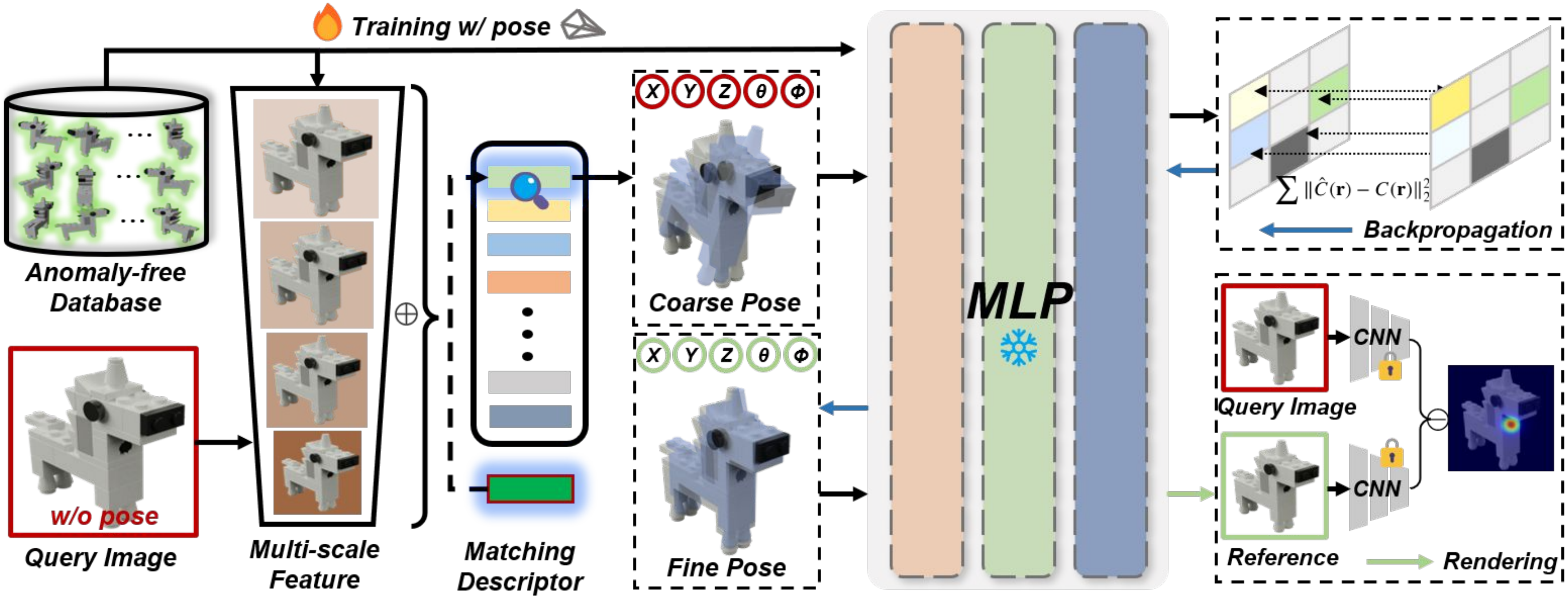
Pseudo-anomaly: Cutpaste, DRAEM

Category	Feature Embedding-based Patchcore[32]	STFPM[45]	Fastflow[59]	CFlow[18]	CFA[21]	Cutpaste[22]	Reconstruction-based DRAEM[46]	FAVAE[54]	OCRGAN[28]	UniAD[57]	Ours
Gorilla	88.4/66.8	93.8/65.3	91.4/51.1	94.7/69.2	91.4/41.8	36.1/-	77.7/58.9	92.1/46.8	94.2/-	93.4/56.6	**99.5/93.6**
Unicorn	58.9/92.4	89.3/79.6	77.9/45.0	89.9/82.3	85.2/85.6	69.6/-	26.0/70.4	88.0/68.3	86.7/-	86.8/73.0	**98.2/94.0**
Mallard	66.1/59.3	86.0/42.2	85.0/72.1	87.3/74.9	83.7/36.6	40.9/-	47.8/34.5	85.3/33.6	88.9/-	85.4/70.0	**97.4/84.7**
Turtle	77.5/87.0	91.0/64.4	83.9/67.7	90.2/51.0	88.7/58.3	77.2/-	45.3/18.4	89.9/82.8	76.7/-	88.9/50.2	**99.1/95.6**
Whale	60.9/**86.0**	88.6/64.1	86.5/53.2	89.2/57.0	87.9/77.7	66.8/-	55.9/65.8	90.1/62.5	89.4/-	90.7/75.5	**98.3**/82.5
Bird	88.6/82.9	90.6/52.4	90.4/76.5	91.8/75.6	92.2/78.4	71.7/-	60.3/69.1	91.6/73.3	**99.1**/-	91.1/74.7	95.7/**92.4**
Owl	86.3/72.9	91.8/72.7	90.7/58.2	94.6/76.5	93.9/74.0	51.9/-	78.9/67.2	96.7/62.5	90.1/-	92.8/65.3	**99.4/88.2**
Sabertooth	69.4/76.6	89.3/56.0	88.7/70.5	93.3/71.3	88.0/64.2	71.2/-	26.2/68.6	94.5/82.4	91.7/-	90.3/61.2	**98.5/95.7**
Swan	73.5/75.2	90.8/53.6	89.5/63.9	93.1/67.4	95.0/66.7	57.2/-	75.9/59.7	87.4/50.6	72.2/-	90.6/57.5	**98.8/86.5**
Sheep	79.9/89.4	93.2/56.5	91.0/71.4	94.3/80.9	94.1/86.5	67.2/-	70.5/59.5	94.3/74.9	**98.9**/-	92.9/70.4	97.7/**90.1**
Pig	83.5/85.7	94.2/50.6	93.6/59.6	97.1/72.1	95.6/66.7	52.3/-	65.6/64.4	92.2/52.5	93.6/-	94.8/54.6	**97.7/88.3**
Zalika	64.9/68.2	86.2/53.7	84.6/54.9	89.4/66.9	87.7/52.1	43.5/-	66.6/51.7	86.4/34.6	94.4/-	86.7/50.5	**99.1/88.2**
Phoenix	62.4/71.4	86.1/56.7	85.7/53.4	87.3/64.4	87.0/65.9	53.1/-	38.7/53.1	92.4/65.2	86.8/-	84.7/55.4	**99.4/82.3**
Elephant	56.2/78.6	76.8/61.7	76.8/61.6	72.4/70.1	77.8/71.7	56.9/-	55.9/62.5	72.0/49.1	91.7/-	70.7/59.3	**99.0/92.5**
Parrot	70.7/78.0	84.0/61.1	84.0/53.4	86.8/67.9	83.7/69.8	55.4/-	34.4/62.3	87.7/46.1	66.5/-	85.6/53.4	**99.5/97.0**
Cat	85.6/78.7	93.7/52.2	93.7/51.3	94.7/65.8	95.0/68.2	58.3/-	79.4/61.3	94.0/53.2	91.3/-	93.8/53.1	**97.7/84.9**
Scorpion	79.9/82.1	90.7/68.9	74.3/51.9	91.9/79.5	92.2/91.4	71.2/-	79.7/83.7	88.4/66.9	**97.6**/-	92.2/69.5	95.9/**91.5**
Obesobeso	91.9/89.5	94.2/60.8	92.9/67.6	95.8/80.0	96.2/80.6	73.3/-	89.2/73.9	92.7/58.2	**98.5**/-	93.6/67.7	98.0/**97.1**
Bear	79.5/84.2	90.6/60.7	85.0/72.9	92.2/81.4	90.7/78.7	68.8/-	39.2/76.1	90.1/52.8	83.1/-	90.9/65.1	**99.3/98.8**
Puppy	73.3/65.6	84.9/56.7	80.3/59.5	89.6/71.4	82.3/53.7	43.2/-	45.8/57.4	85.6/43.5	78.9/-	87.1/55.6	**98.8/93.5**
Mean	74.7/78.5	89.3/59.5	86.1/60.8	90.8/71.3	89.8/68.2	59.3/-	58.0/60.9	89.4/58.0	88.5/-	89.1/62.2	**97.8/90.9**

Challenges:

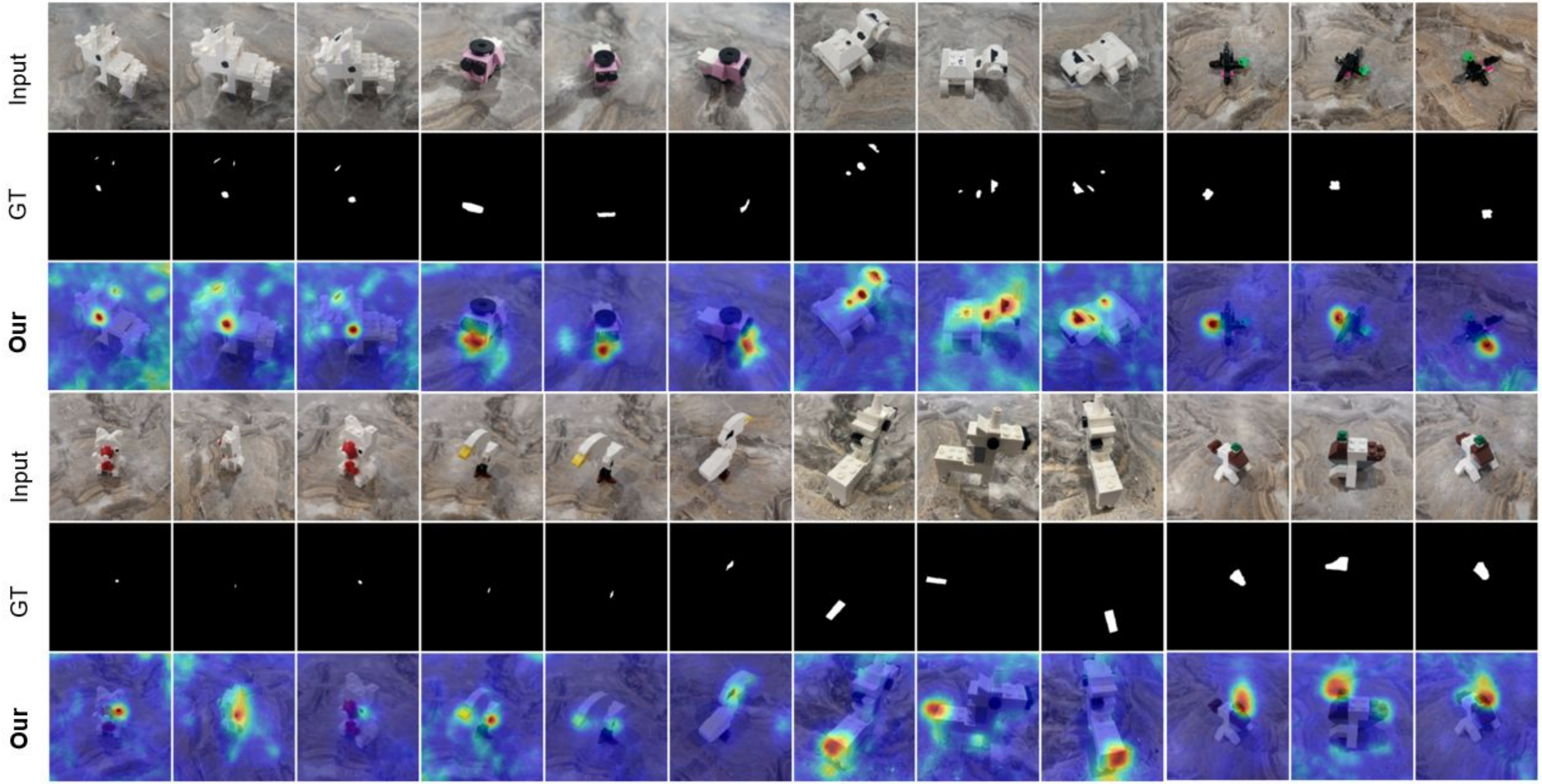
3) We lack a baseline method for pose-agnostic anomaly detection setting.

Contribution #4: OmniposeAD



Experiments Result

◆ *OmniposeAD* enables anomaly detection for arbitrary object poses.



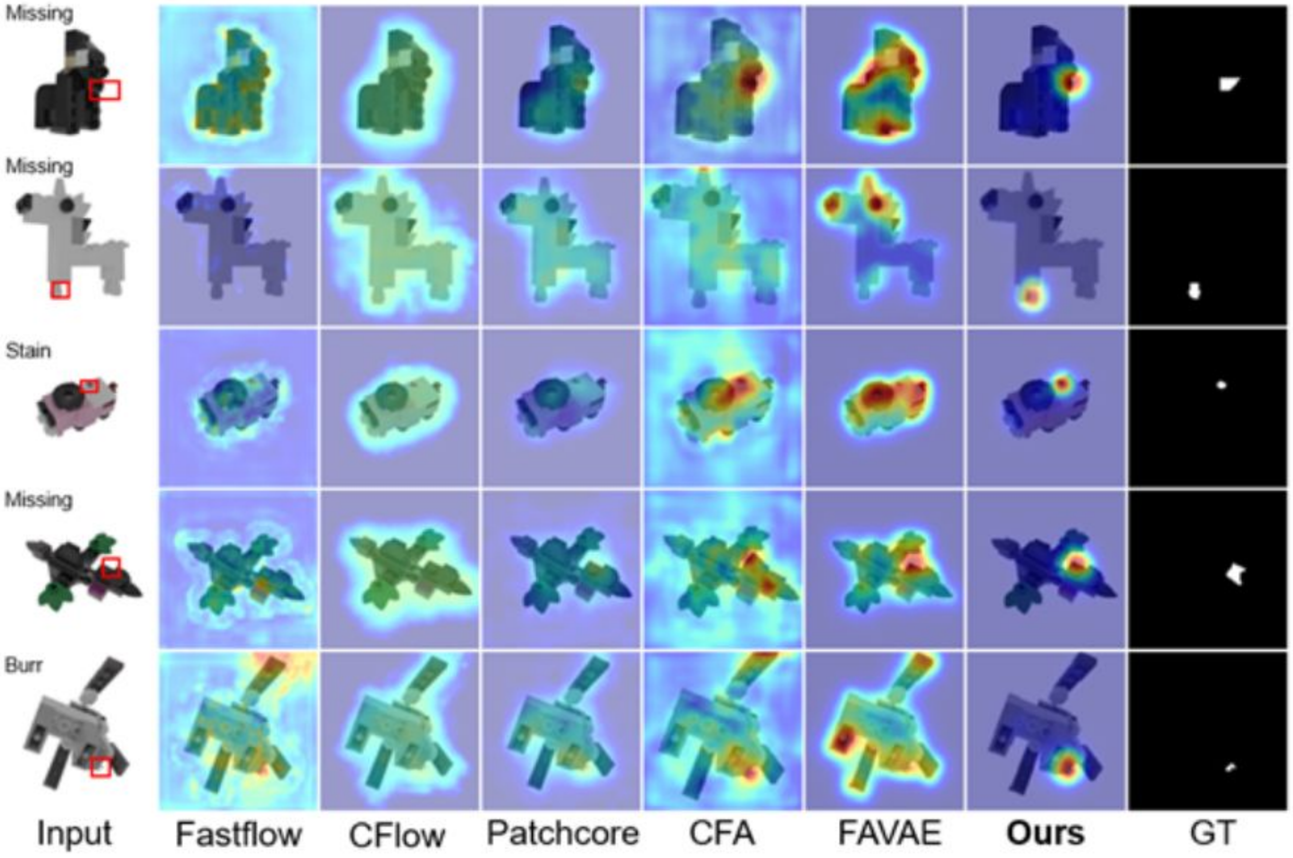
Experiments Result

◆ OmniposeAD's Dense to sparse view evaluation on MAD-Real.

Sample Size / Object Class	100 Image	100 Pixel	70 Image	70 Pixel	50 Image	50 Pixel
Gorilla	**93.6**	**99.5**	85.4	97.5	91.0	97.4
Unicorn	**86.2**	**96.5**	84.4	96.1	80.2	94.9
Mallard	**87.3**	**96.3**	80.3	94.4	74.6	95.1
Turtle	**99.1**	**95.4**	95.1	92.3	83.4	90.4
Whale	**82.1**	**98.5**	82.5	97.0	76.0	93.1
Bird	**92.0**	**95.1**	91.6	91.7	89.1	93.4
Owl	**89.1**	**99.3**	78.8	98.1	89.1	98.8
Sabertooth	**93.8**	**97.8**	84.9	95.1	81.5	94.4
Swan	**80.1**	**98.0**	77.0	97.3	71.3	96.7
Sheep	**84.3**	**97.6**	83.2	97.6	82.3	92.8

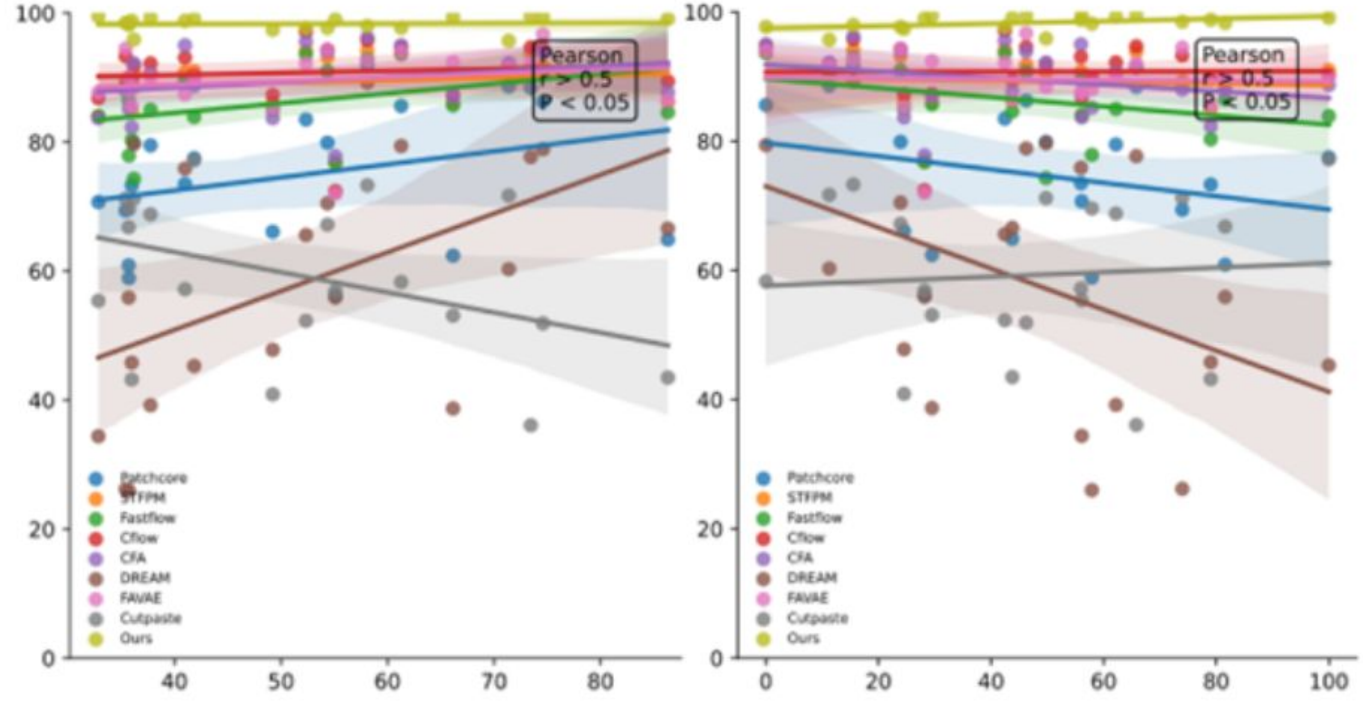
Sample Size / Object Class	100 Image	100 Pixel	70 Image	70 Pixel	50 Image	50 Pixel
Pig	**75.7**	**96.4**	73.9	95.7	72.8	95.4
Zalika	**88.2**	**99.1**	78.8	96.8	72.8	96.0
Phoenix	**86.0**	**99.4**	84.6	99.2	82.6	98.8
Elephant	**91.0**	**98.1**	82.5	95.9	85.1	96.5
Parrot	**91.5**	**99.1**	85.5	98.2	87.9	96.9
Cat	**78.9**	**97.6**	73.5	97.6	73.7	97.1
Scorpion	**87.4**	**94.2**	77.0	91.8	74.3	91.5
Obesobeso	**88.5**	**96.1**	86.4	97.9	81.8	97.8
Bear	**97.8**	**99.3**	92.0	98.0	87.8	97.9
Puppy	**93.1**	**98.5**	86.5	95.8	80.3	95.3

◆ Qualitative results on MAD-Sim.

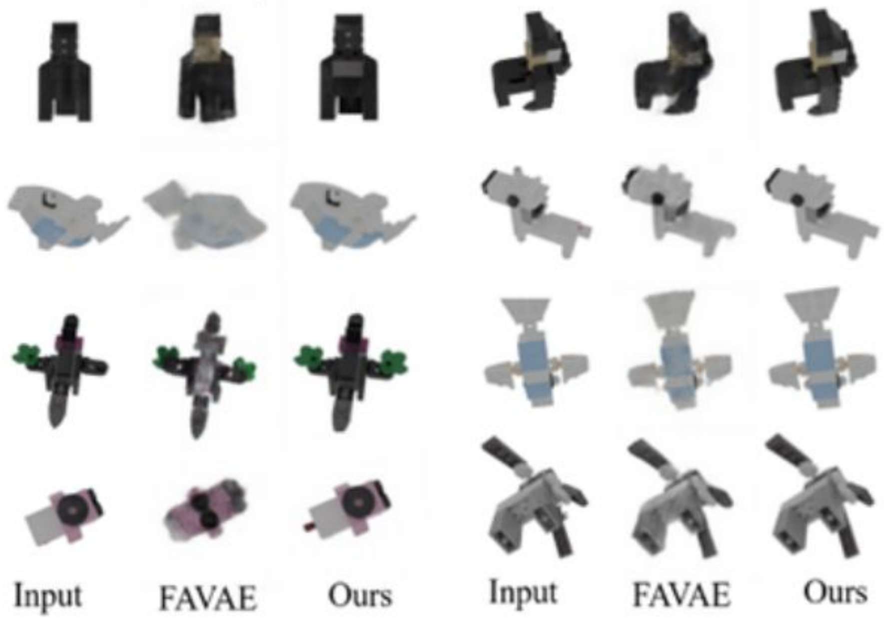


Experiments Result

◆ Correlation of performance with object attributes: *Texture & Shape*



◆ Reference Reconstruction: *OmniposeAD* vs. *FAVAE*



(a) train and test with different poses (b) train and test with similar poses

Thanks!

Check out our paper/dataset on arXiv/Github :

Paper

Github
Repo