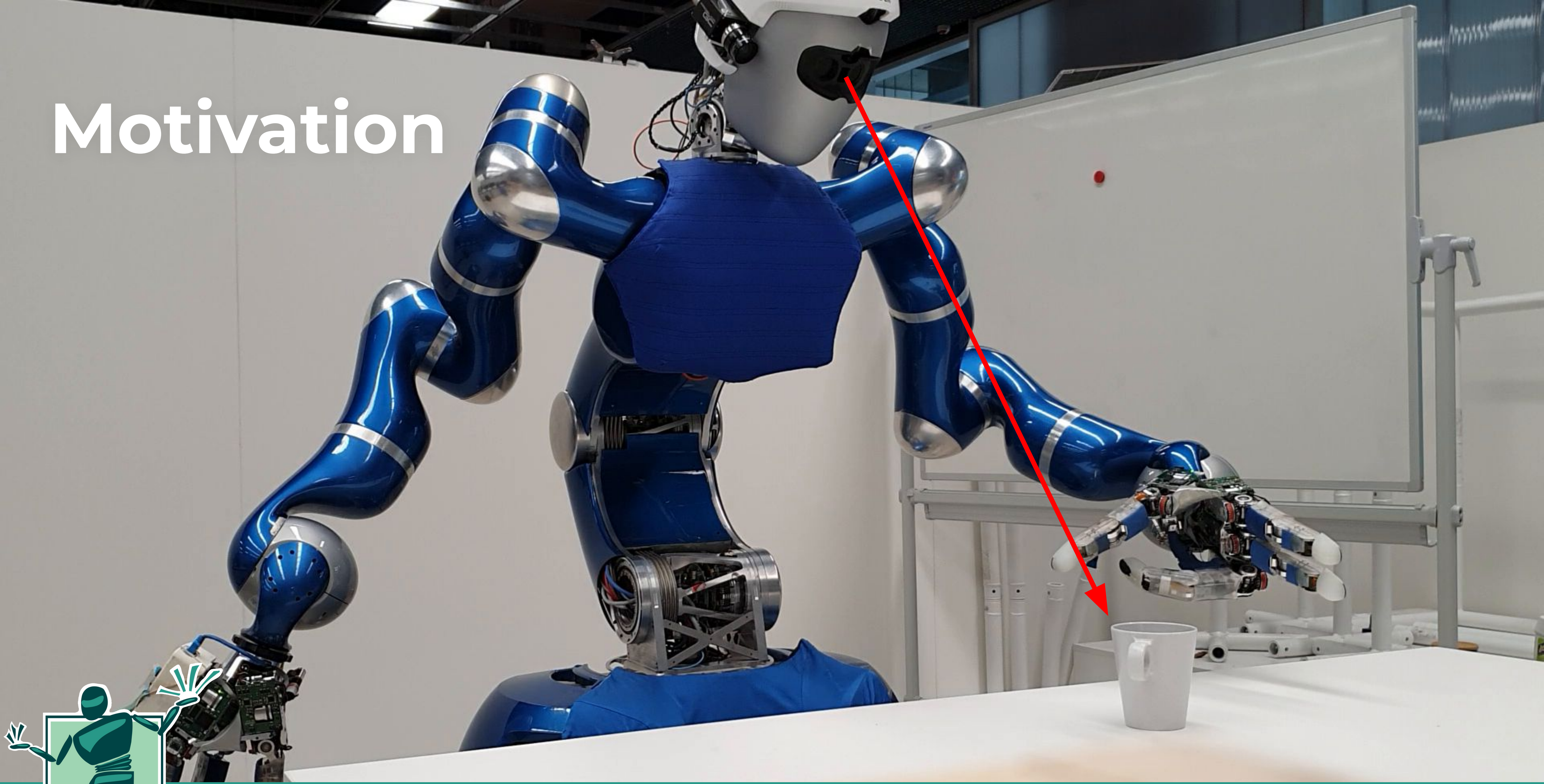
Shape Completion with Prediction of Uncertain Regions

Matthias Humt, Dominik Winkelbauer and Ulrich Hillenbrand

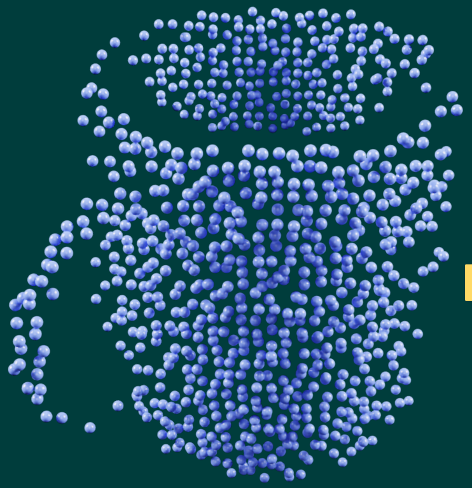
DLR Deutsches Zentrum für Luft- und Raumfahrt
German Aerospace Center

2023 IROS

Motivation



Shape Completion



1. point cloud (simulation)

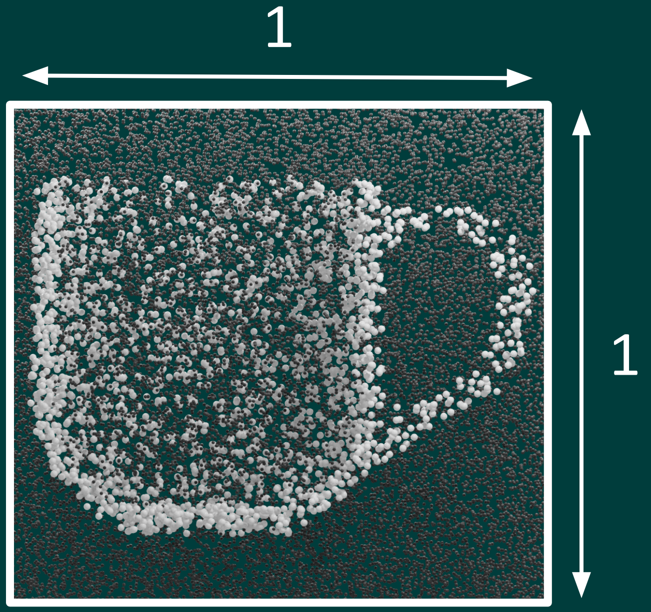
continuous query points

$\downarrow \mathbb{R}^3$

encoder → feature $\mathcal{X}$ → decoder → τ

$$f_\theta : \mathcal{X} \times \mathbb{R}^3 \rightarrow [0, 1]$$

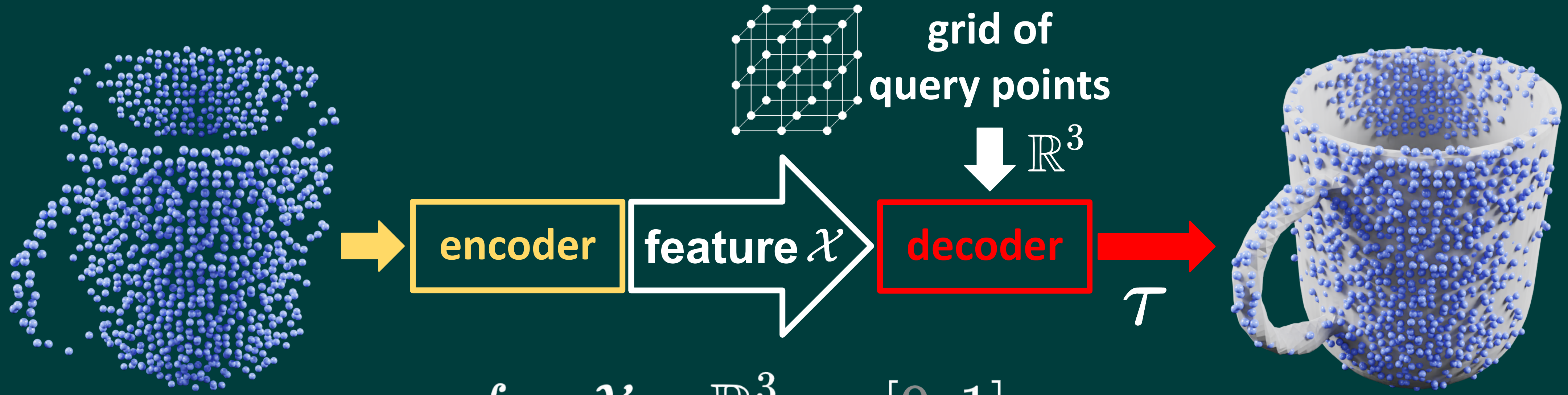
2. convolutional occupancy network



3. predicted occupancy probabilities

[1] Mescheder et al.: Occupancy Networks: Learning 3D Reconstruction in Function Space
[2] Peng et al.: Convolutional Occupancy Networks

Shape Completion



$$f_\theta : \mathcal{X} \times \mathbb{R}^3 \rightarrow [0, 1]$$

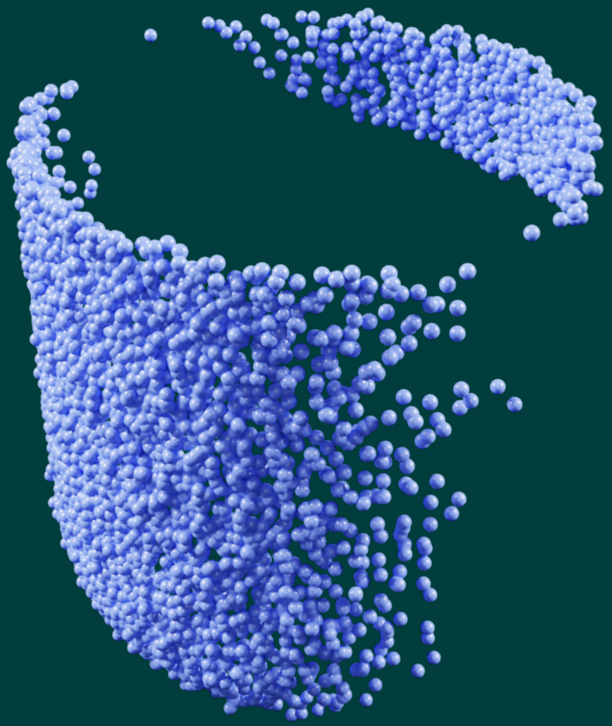
1. point cloud (simulation)

2. convolutional occupancy network

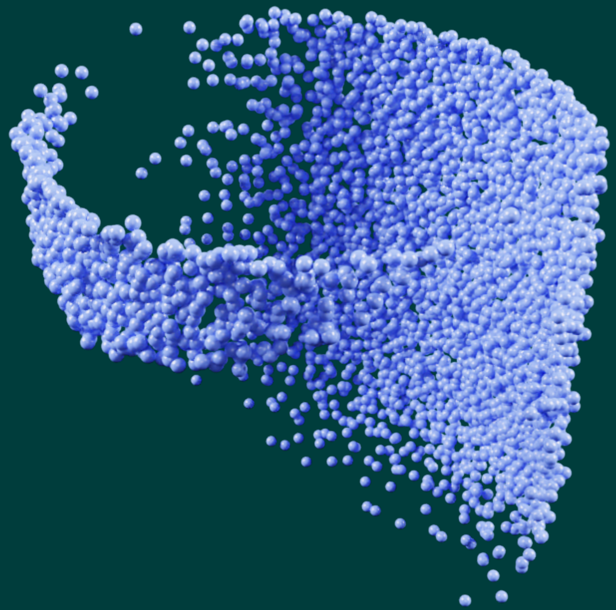
3. shape completion (marching cubes)

[1] Mescheder et al.: Occupancy Networks: Learning 3D Reconstruction in Function Space
[2] Peng et al.: Convolutional Occupancy Networks

Motivation: Pose Ambiguity

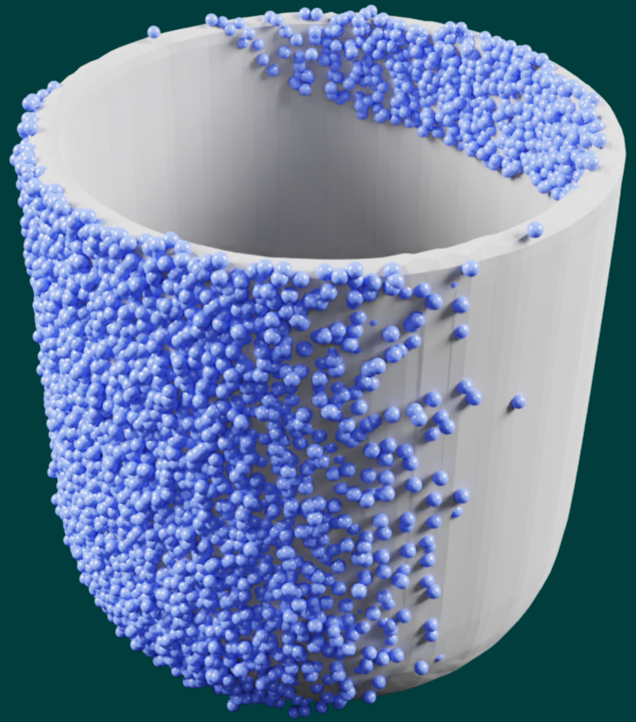


front view

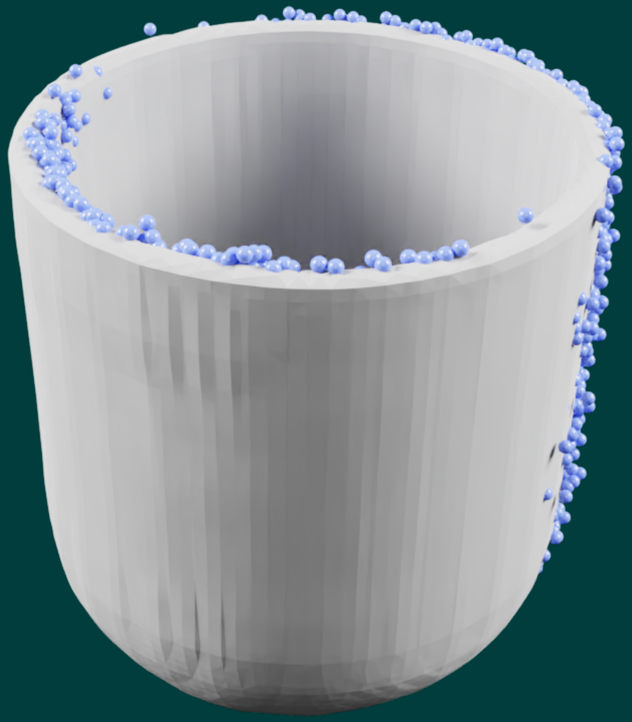


back view

Motivation: Pose Ambiguity

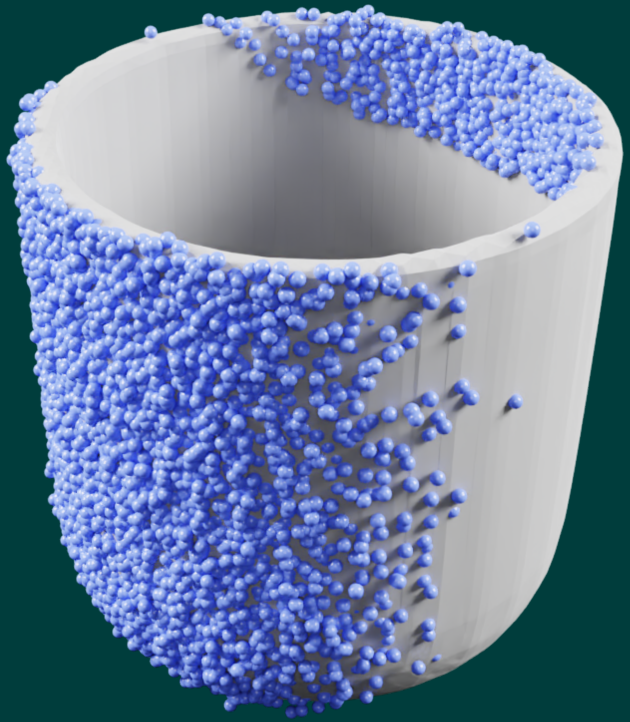


front view

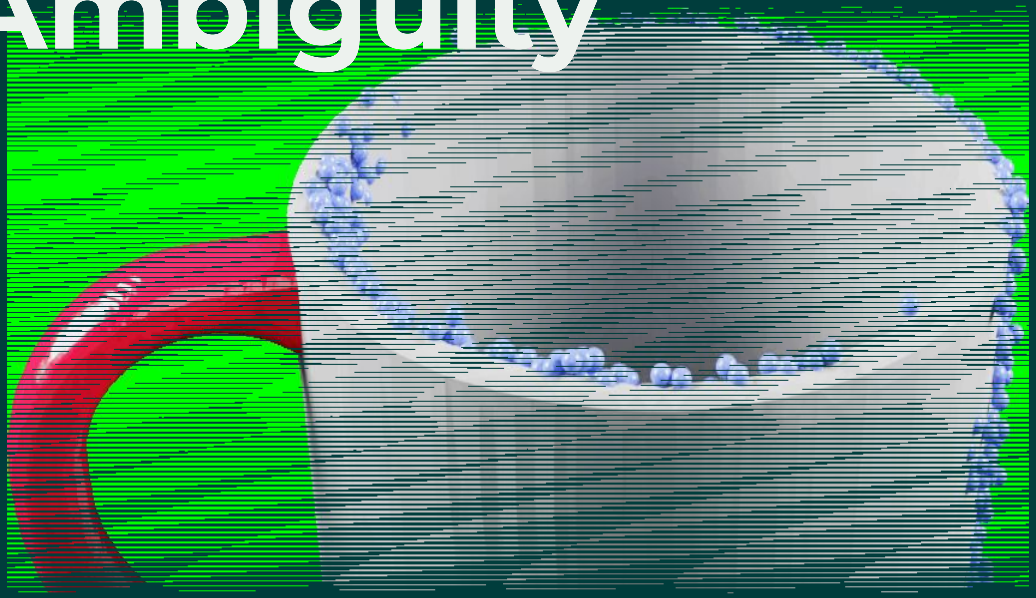


back view

Motivation: Pose Ambiguity



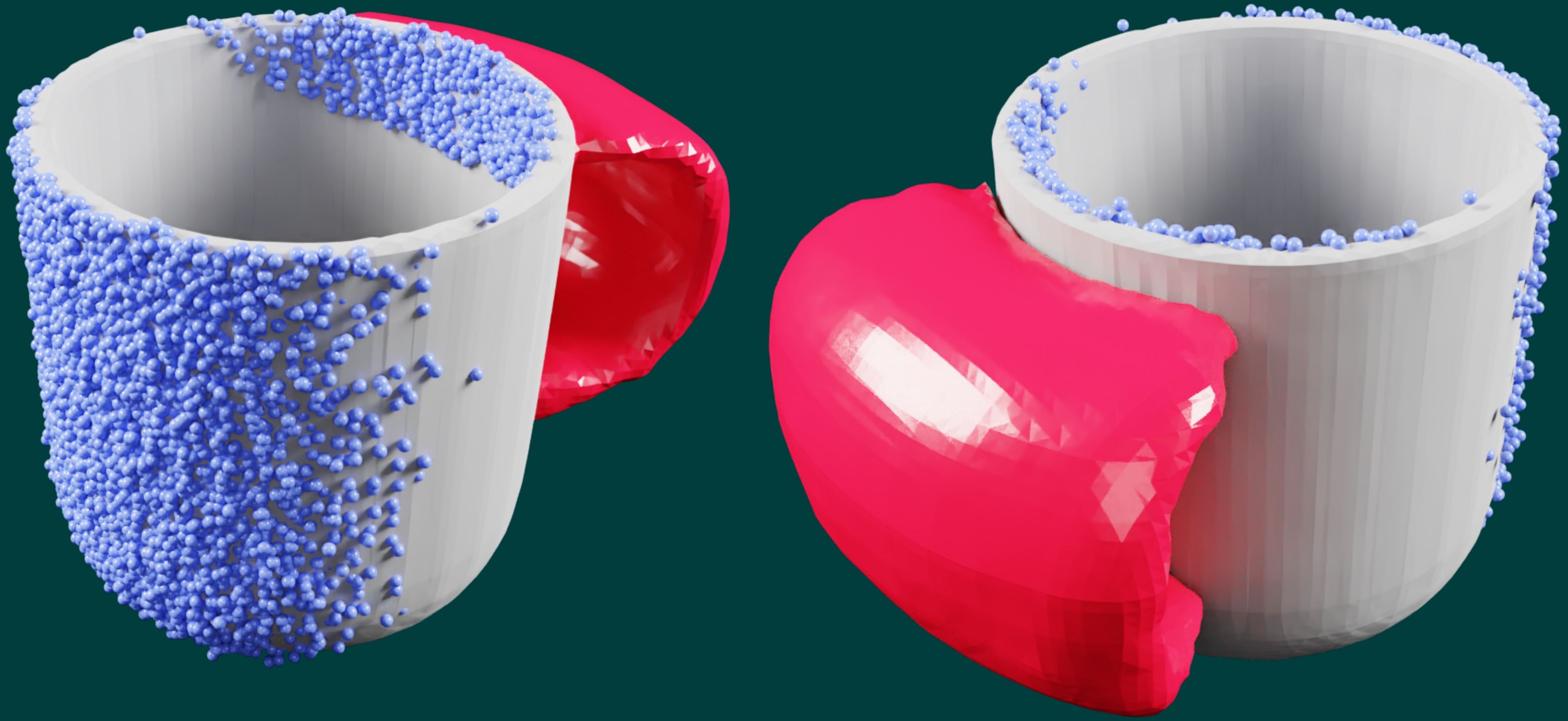
front view



back view

DLR

Motivation: Uncertain Regions

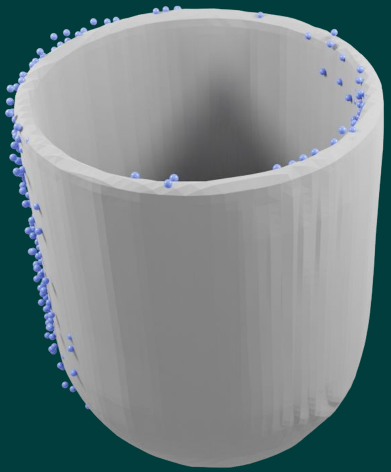


front view

back view

Approach 1: Gradient Threshold

ground truth



predicted occupied

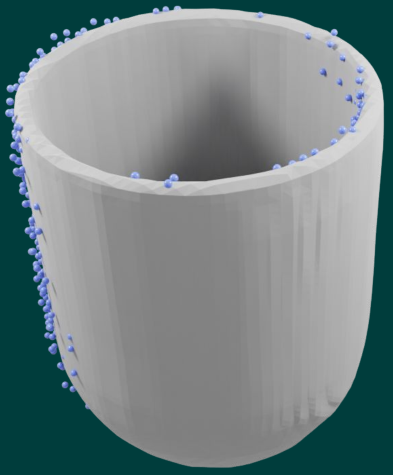
$$\tau_{\mathrm{occ}} = 0.5$$

predicted uncertain

$$\tau_{\mathrm{min}} < \tau_{\mathrm{max}} \leq \tau_{\mathrm{occ}}$$

Approach 1: Gradient Threshold

ground truth



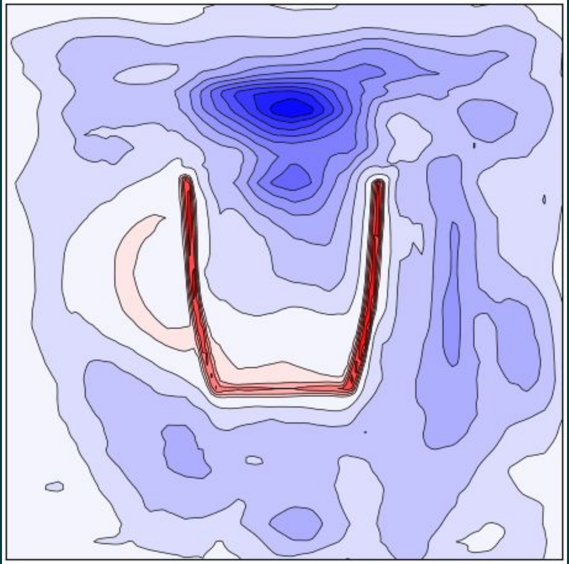
predicted occupied

$$\tau_{\text{occ}} = 0.5$$

predicted uncertain

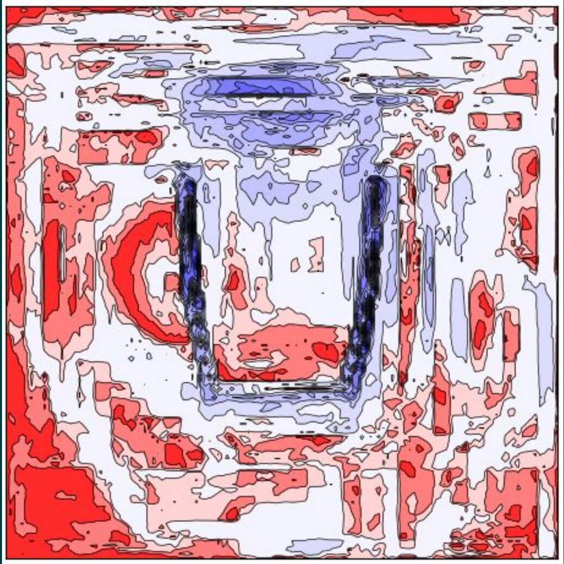
$$\tau_{\text{min}} < \tau_{\text{max}} \leq \tau_{\text{occ}}$$

Approach 1: Gradient Threshold



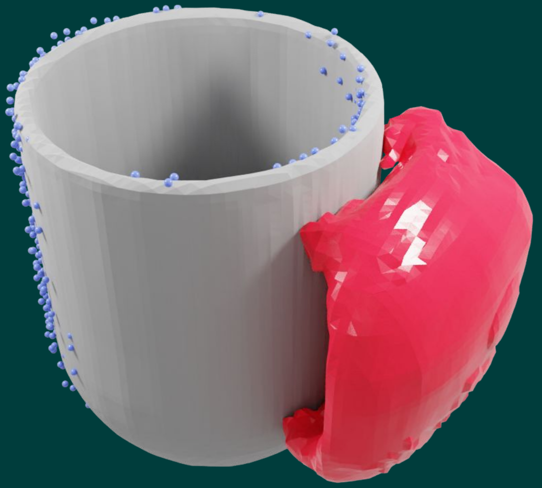
a: predicted occupancy probability

$\cap$



b: gradient magnitude of (a) w.r.t. the query

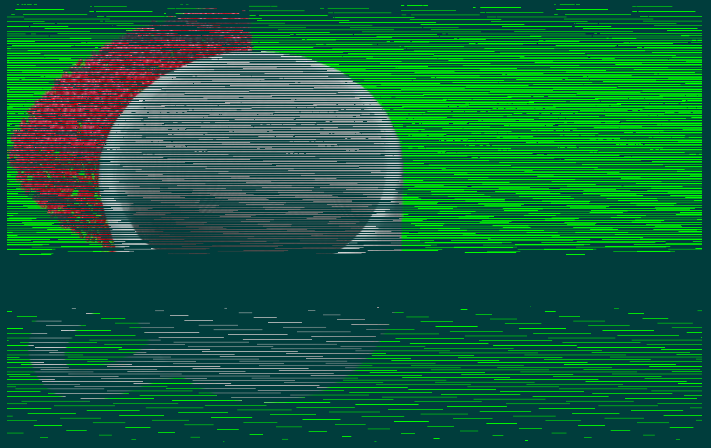
$=$



c: shape completion & uncertain region

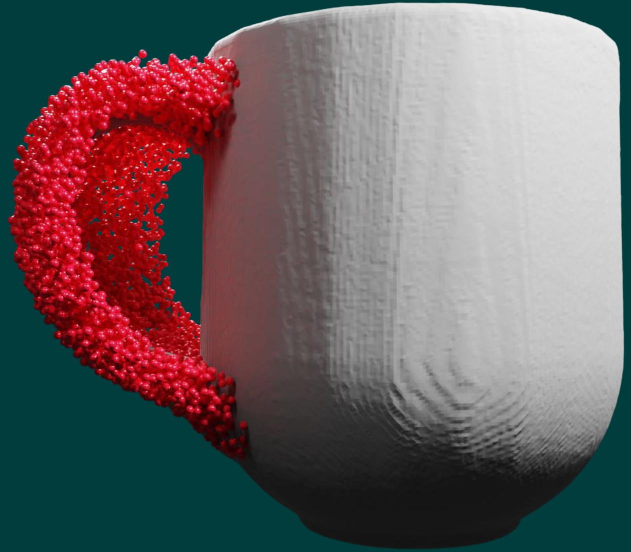
$$\forall q \in \mathcal{Q} : \quad \| \nabla_q \sum_{i=1}^{N} \hat{y}_i \| < \frac{1}{|\mathcal{Q}|} \sum_{q \in \mathcal{Q}} \| \nabla_q \sum_{i=1}^{N} \hat{y}_i \|$$

Approach 2: Additional Label



rotations

virtual camera



side view

top view

$$f_\theta : \mathcal{X} \times \mathbb{R}^3 \rightarrow [0, 1, 2]$$

Experiments

1. **Novel View:** Generalize to new viewpoints of known object instances
2. **Novel Instance:** Generalize to new instances of a known class
3. **Sim2Real:** Generalize from simulated to real data

Evaluation: Metrics

$$\mathrm{IoU} = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}}$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP+FP}}$$

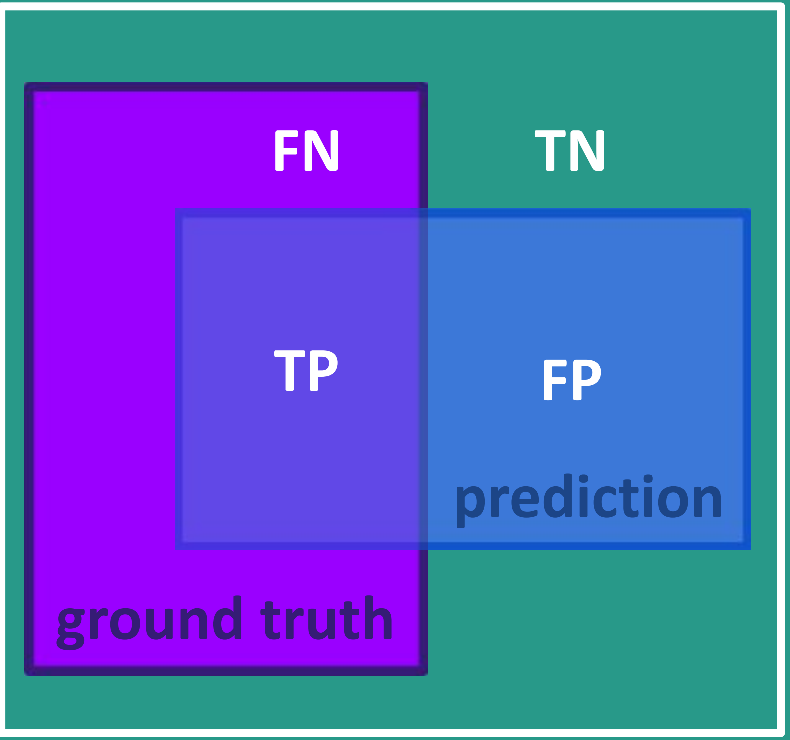
$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP+FN}}$$

$$\boxed{\mathrm{F1} = \frac{2 \times \mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision+Recall}}}$$

→ **harmonic mean between precision & recall**

- **FN: false negative**
- **FP: false positive**
- **TN: true negative**
- **TP: true positive**

confusion matrix



Evaluation: Grasping

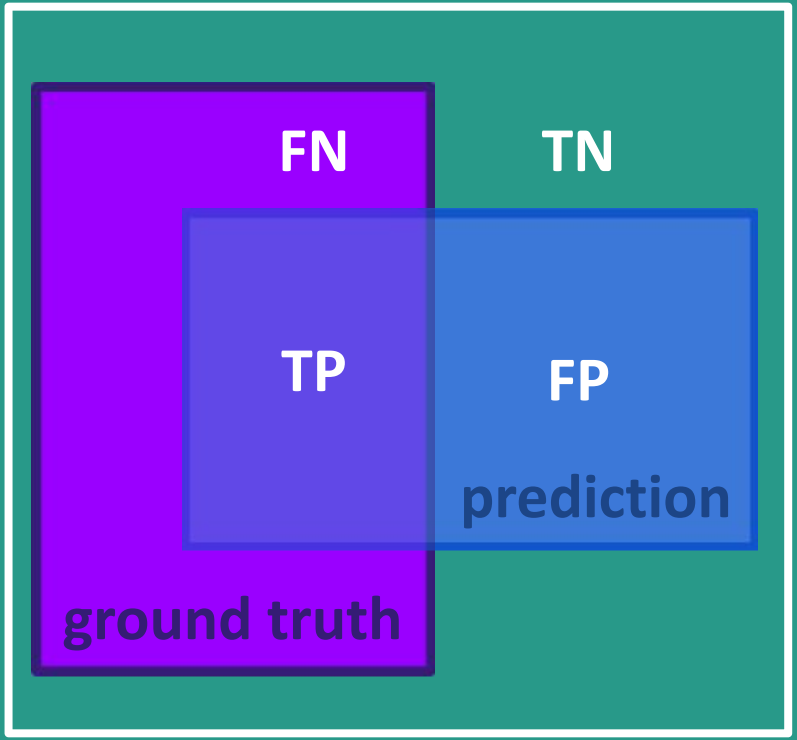
Grasp Collision Risk: $\mathrm{GCR} = \frac{\mathrm{FN}_{\mathrm{occ}} + \mathrm{FN}_{\mathrm{unc}}}{\mathrm{TP}_{\mathrm{occ}} + \mathrm{FN}_{\mathrm{occ}} + \mathrm{TP}_{\mathrm{unc}} + \mathrm{FN}_{\mathrm{unc}}}$

Grasp Miss Risk: $\mathrm{GMR} = \frac{\mathrm{FP}_{\mathrm{occ}}}{\mathrm{FP}_{\mathrm{occ}} + \mathrm{TP}_{\mathrm{occ}}}$

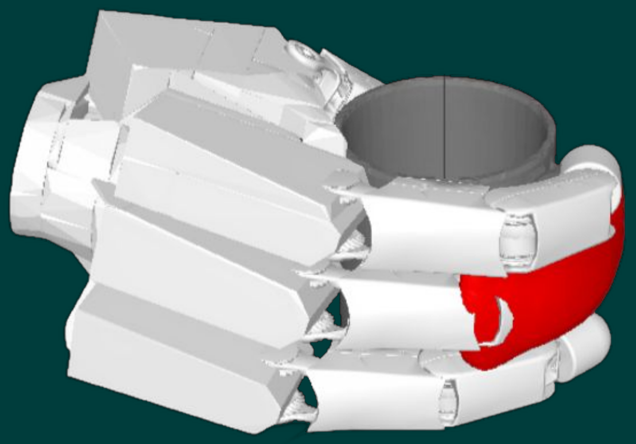
Grasp Exclusion Risk: $\mathrm{GER} = \frac{|\mathcal{FP}_{\mathrm{occ}} \cup \mathcal{FP}_{\mathrm{unc}}|}{\mathrm{FP}_{\mathrm{occ}} + \mathrm{TN}_{\mathrm{occ}}}$

- *occ*: occupied space
- *unc*: uncertain space

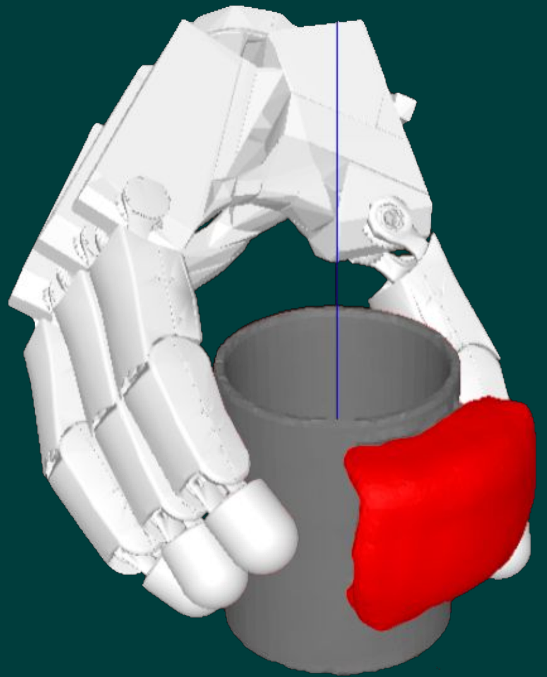
confusion matrix



Evaluation: Grasping



ignoring uncertain region

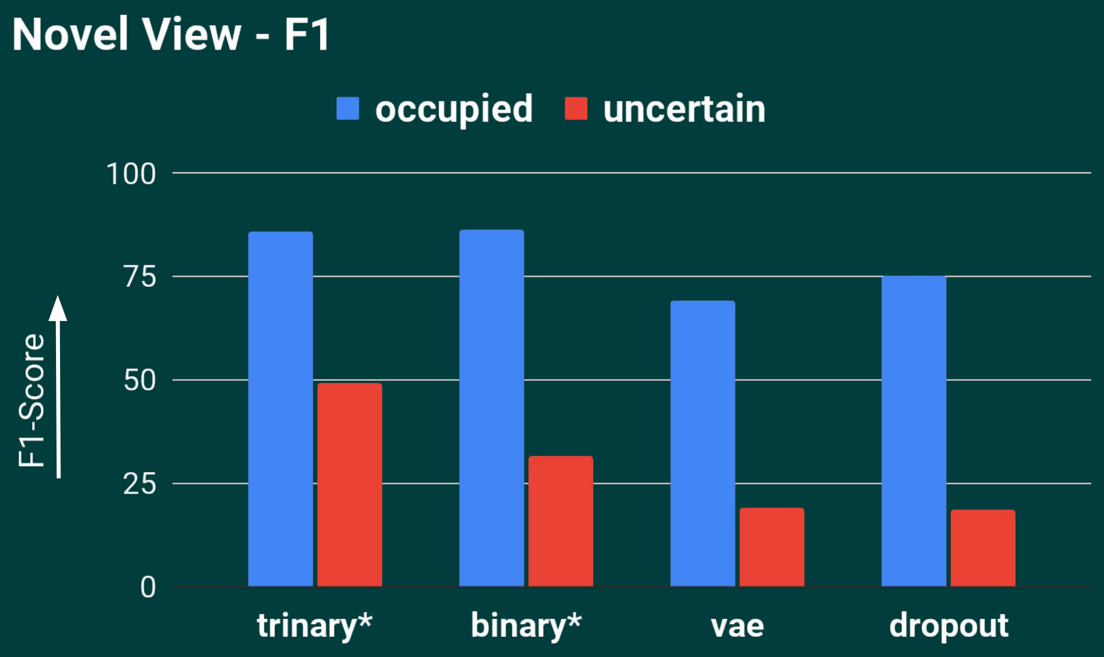


considering uncertain region

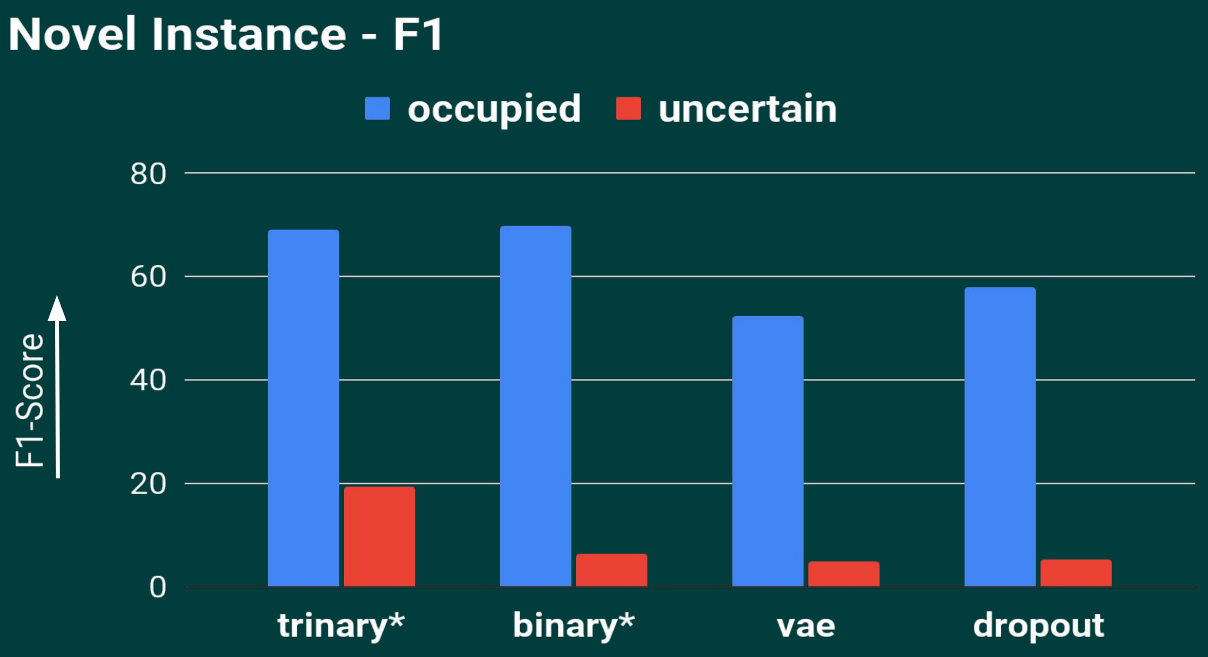
<ins>Improved Epsilon Quality</ins> metric (IEQ): ***"The minimal external force applied to the ground truth object mesh that would break the grasp."***

[3] Winkelbauer et al.: A two-stage learning architecture that generates high-quality grasps for a multi-fingered hand

Results: Novel View & Instance



F1-Score

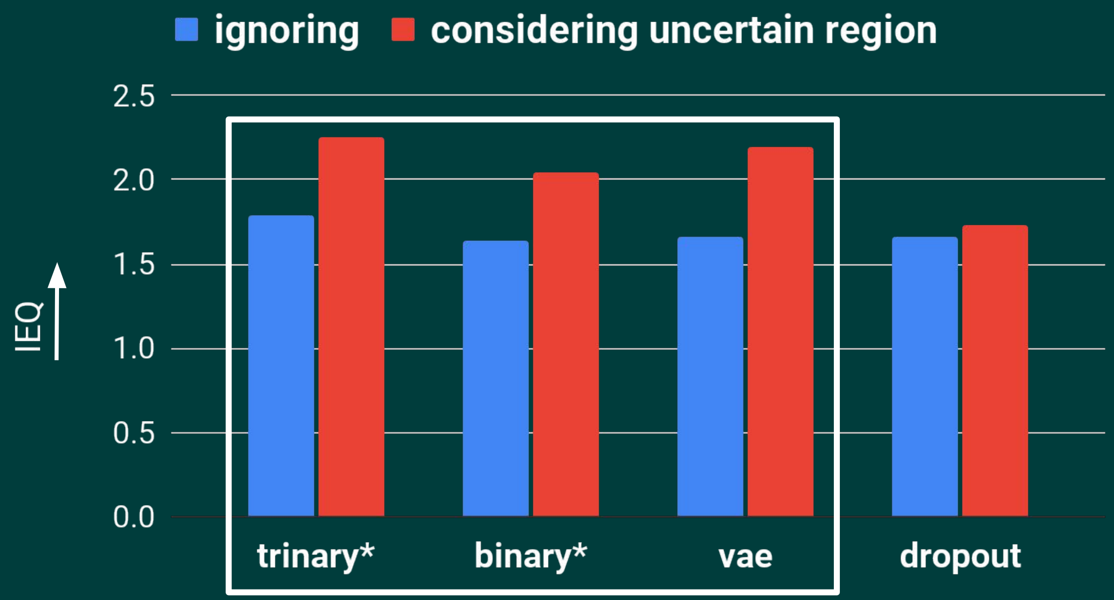


F1-Score

*ours

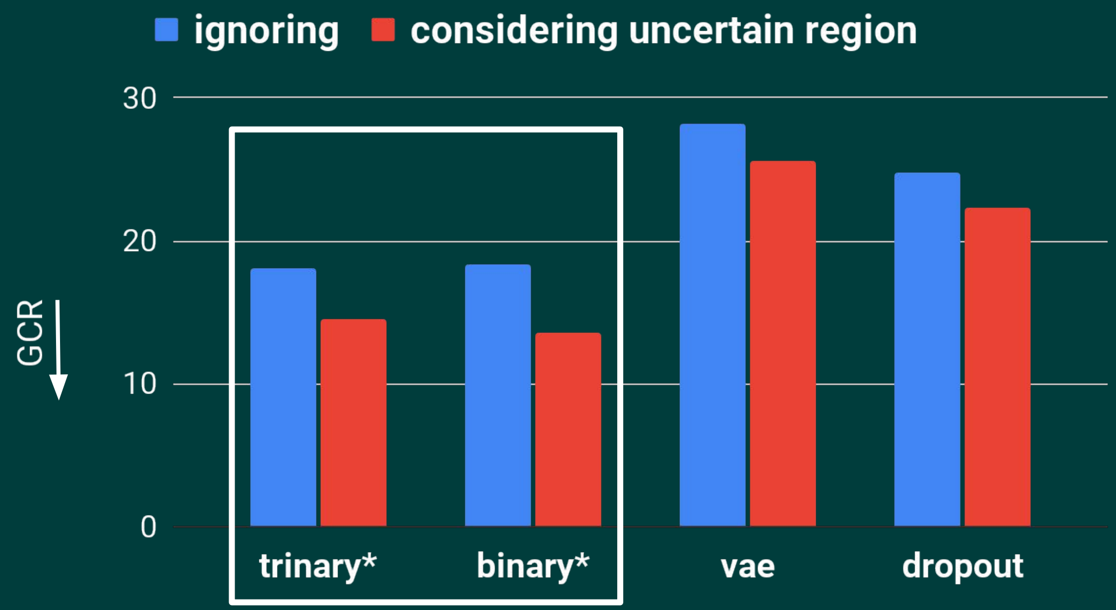
Results: Novel View Grasping

Novel View - IEQ



Improved Ɛ-Quality

Novel View - GCR

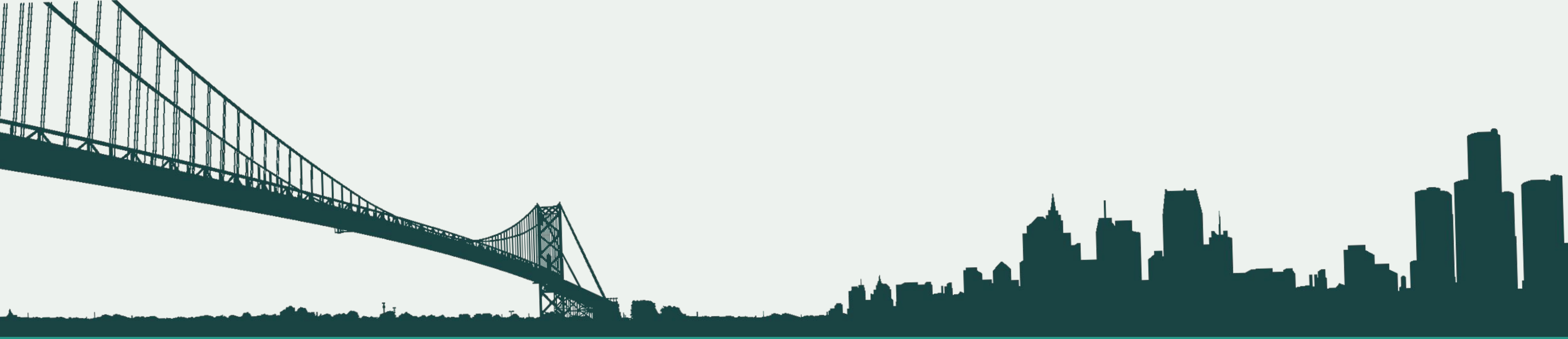


Grasp Collision Risk

*ours

Qualitative Sim2Real Results: Trinary

Qualitative Sim2Real Results: Binary



Thank you for your attention!

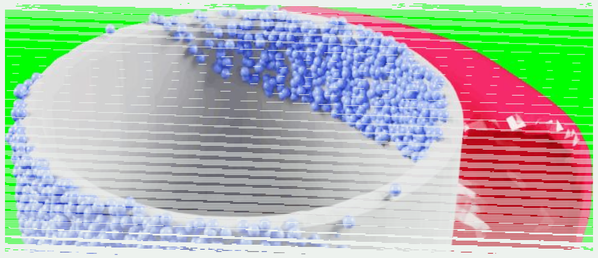
Paper

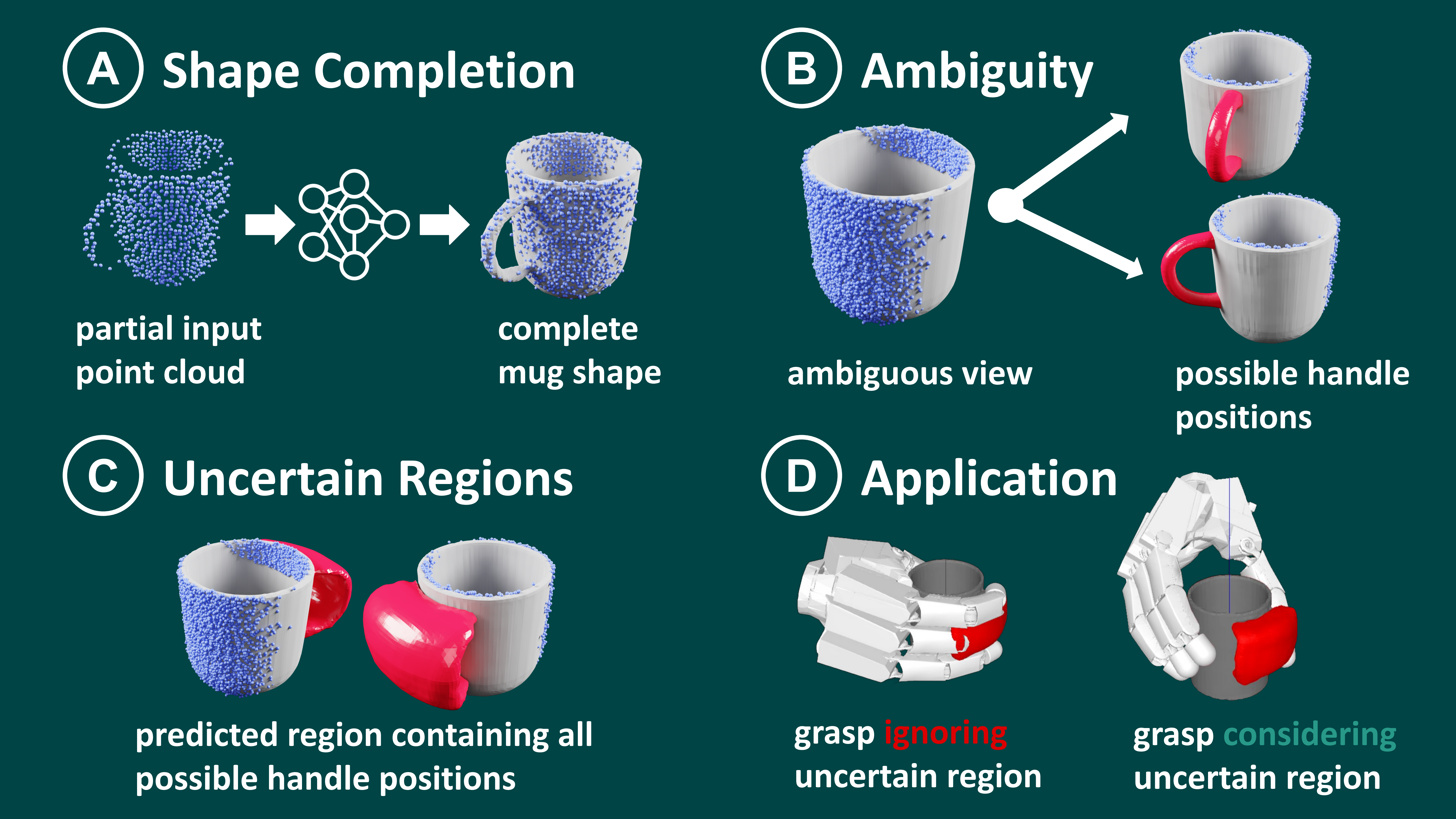
https://arxiv.org/abs/2308.00377

Code & Dataset

https://github.com/DLR-RM/shape-completion

Poster: MoAIP-12.9




A Shape Completion
partial input
point cloud
complete
mug shape
B Ambiguity
ambiguous view
possible handle
positions
C Uncertain Regions
predicted region containing all
possible handle positions
D Application
grasp ignoring
uncertain region
grasp considering
uncertain region