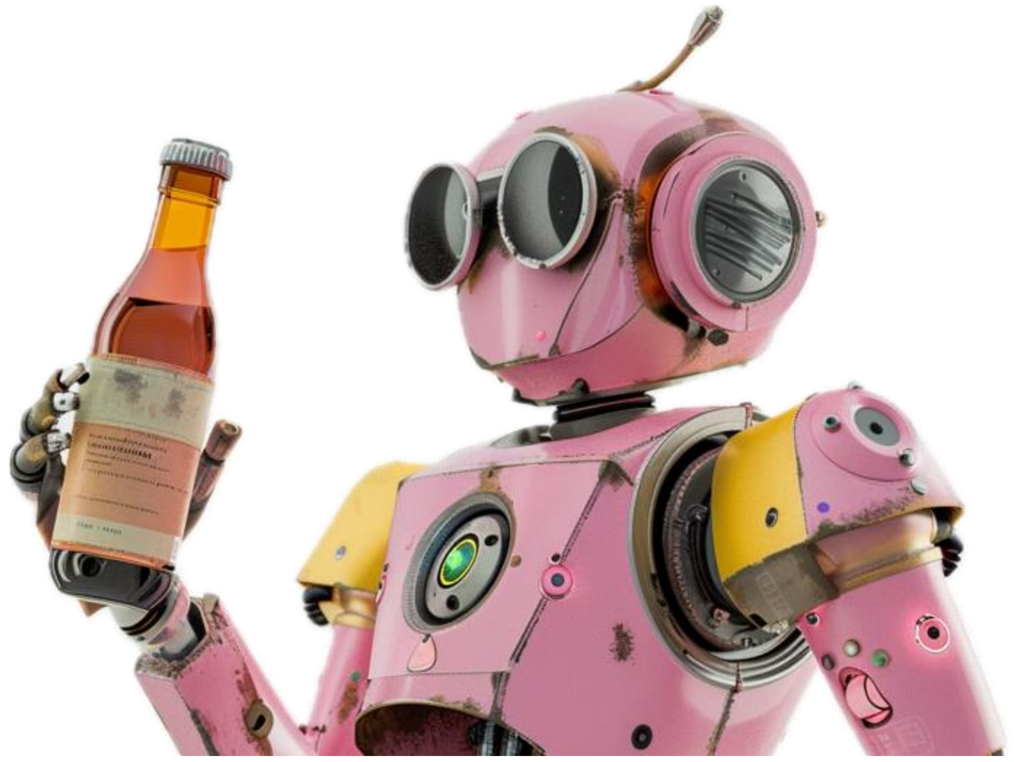
Tactile Exploration of Objects

J. Rüggemeier, D. Strauß, D. Winkelbauer, B. Bäuml



Problem

- Robot grasping relies on accurate 3D models, part of object often occluded
- Approach: use RL **to find optimal choice of grasping points for tactile exploration**
- Simplification: 2D models

Data Generation

- 3D models of cups and bottles from ShapeNet
- Converted to **2D** and **randomly rotated** 5 times each
- Sampling of tactile points of order 1 to 10
- Train/test split without overlap of labels

Reconstruction Network

- **U-Net** with **Bayesian optimizer** and **HyperBand**
- 100 samples for 10 epochs each

Best parameters

- Batch size: 4
- LR: 2.074e-5
- Depth: 6
- Channels: 128

Dataset	Loss	Jacc
Train	0.0276	82.45
Eval	0.0350	77.37
Test	0.0350	77.48

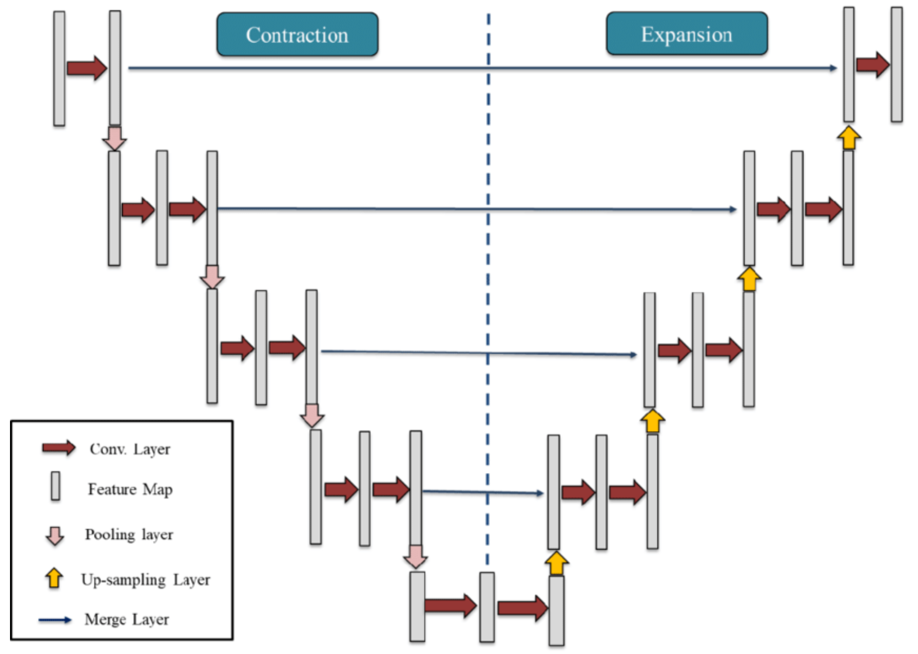
Metric: Jaccard Index

$$\frac{|O \cap L|}{|O \cup L|} * 100$$

Boolean threshold of 0.5 on reconstruction and label

O: Set of True pixels in output

L: Set of True pixels in label



Deep Learning Methods for Abnormality Detection and Segmentation in Computed Tomography and Magnetic Resonance Images - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/9-An-overview-of-a-typical-U-Net-with-five-stages_fig8_343774108 [accessed 5 Jun, 2024]

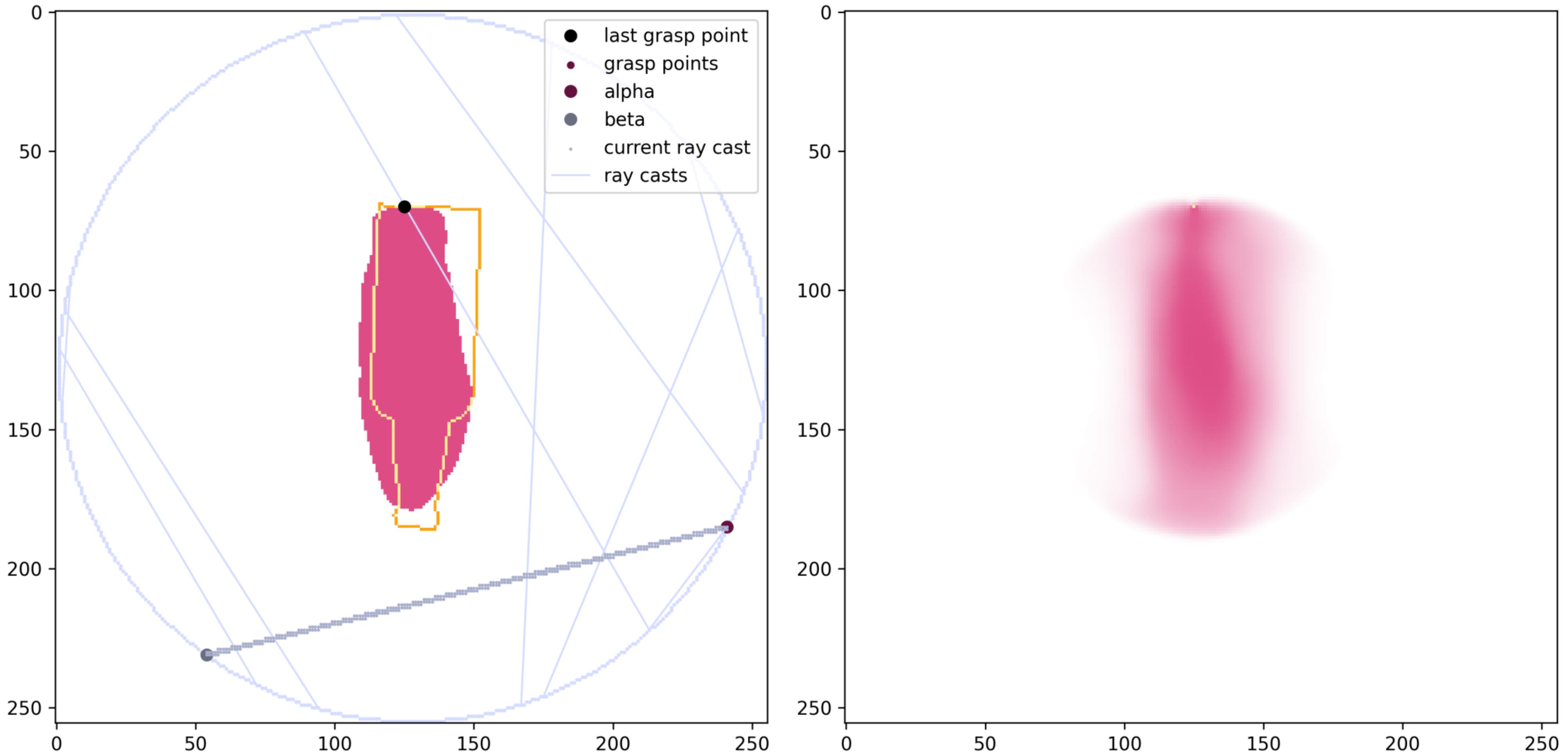
Reinforcement Learning

- Ray simulates grasp approach
- Based on two points on circle points
- Termination after 10 steps
- Penalty of -10 on misses and doubles

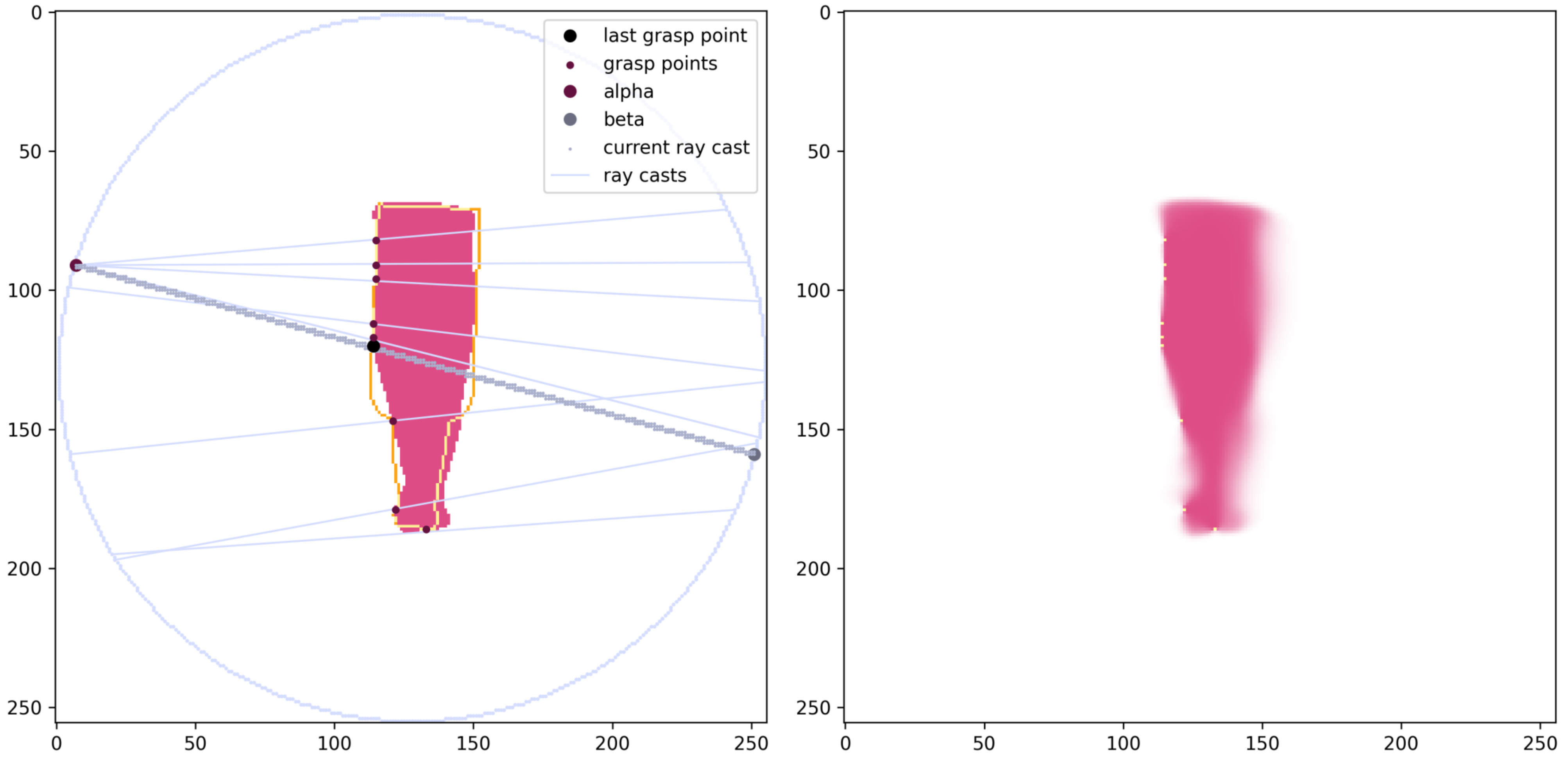
Training

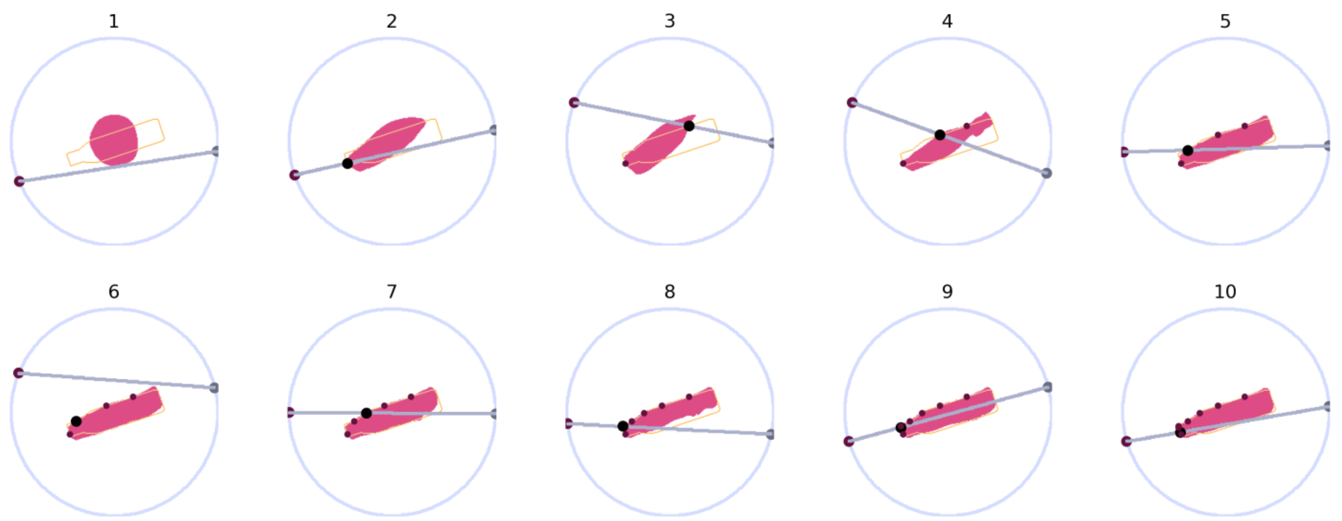
- **PPO** with default settings
- Converged to single action with bad reward -> **Entropy coefficient** of 0.01
- Training with different reward functions

Random actions on test sample

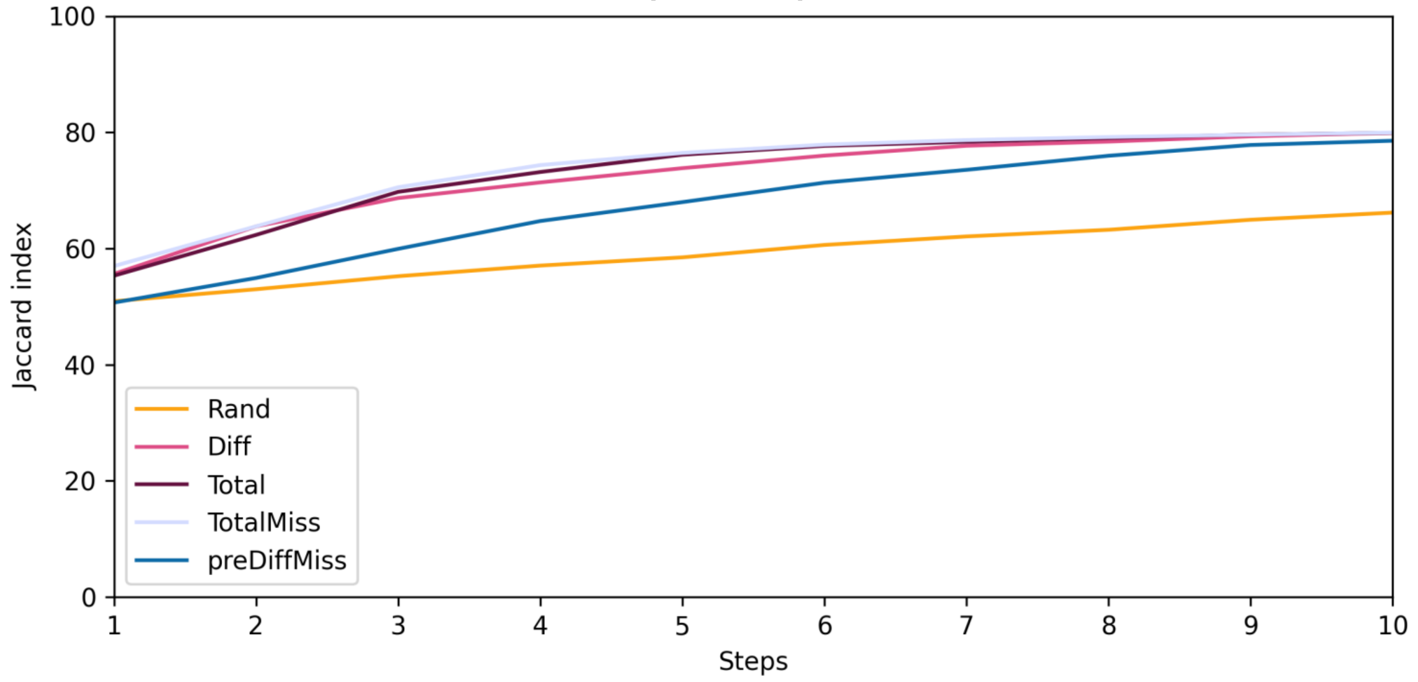


CNN-Policy with Total-reward function on test sample

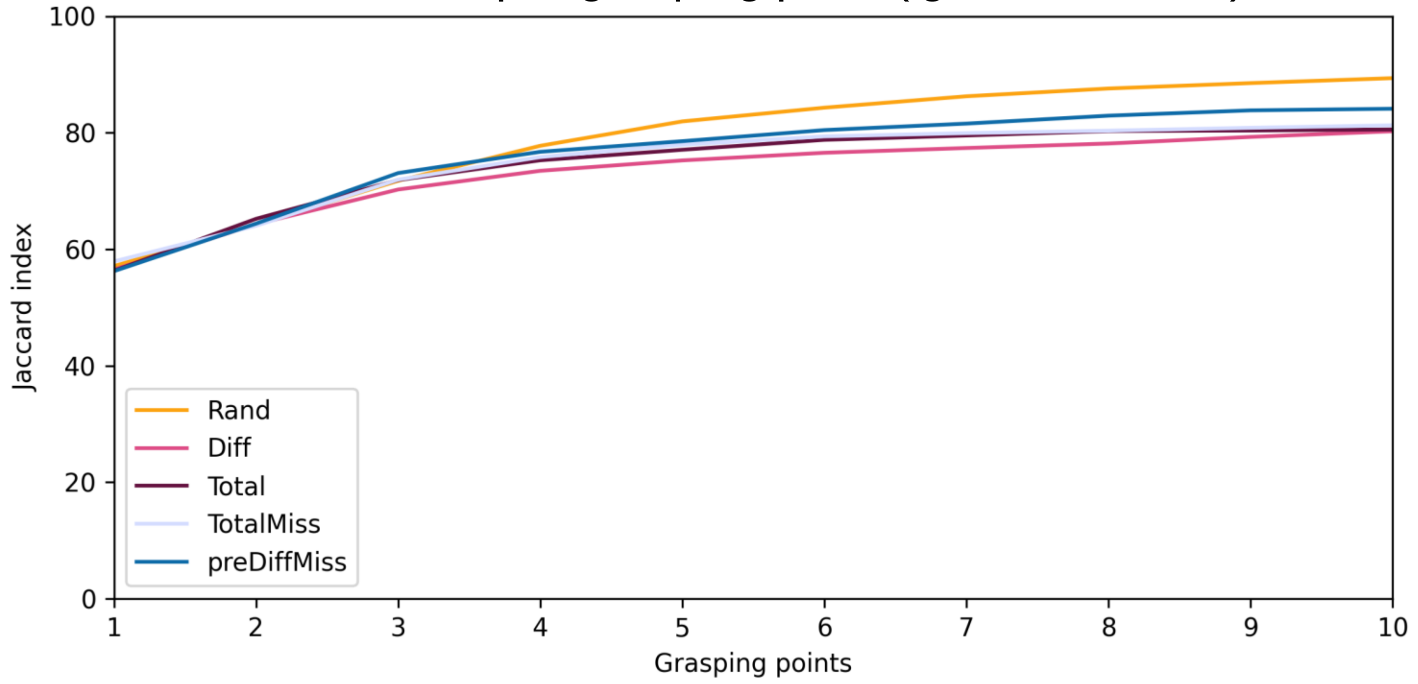




Jaccard index per steps on test set



Jaccard index per grasping point (ignores misses)



Conclusions

- Good generalization on unseen data
- Trained policies find better grasp points than random policy
- Mainly due to avoiding misses

Future

- Generalize for 3D datasets
- Provide more information for RL-policy
- Train on larger datasets