Learning to Singulate Objects using a Push Proposal Network

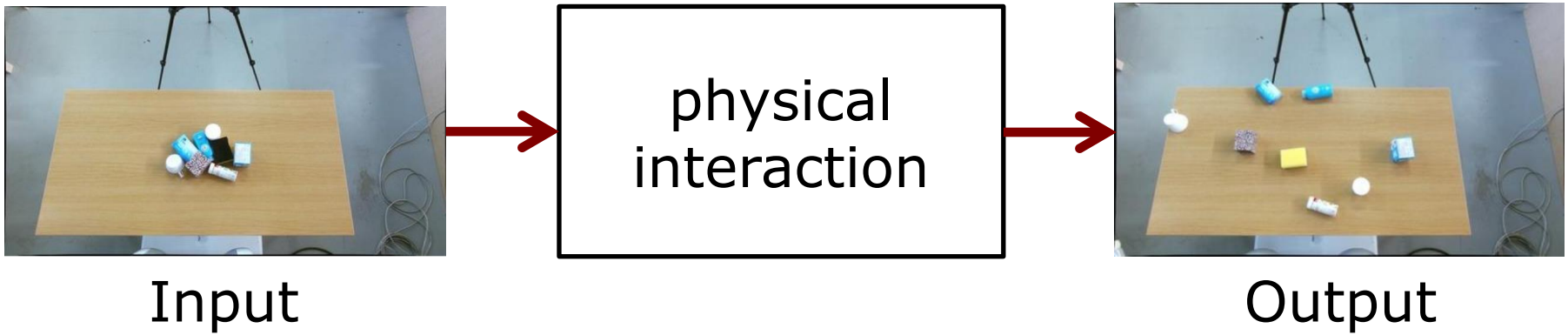
Andreas Eitel, Nico Hauff, Wolfram Burgard

AiS Autonomous Intelligent Systems

Removing Clutter is Hard

Manipulation of objects in unstructured scenes is challenging due to uncertainty from perception

Motivation



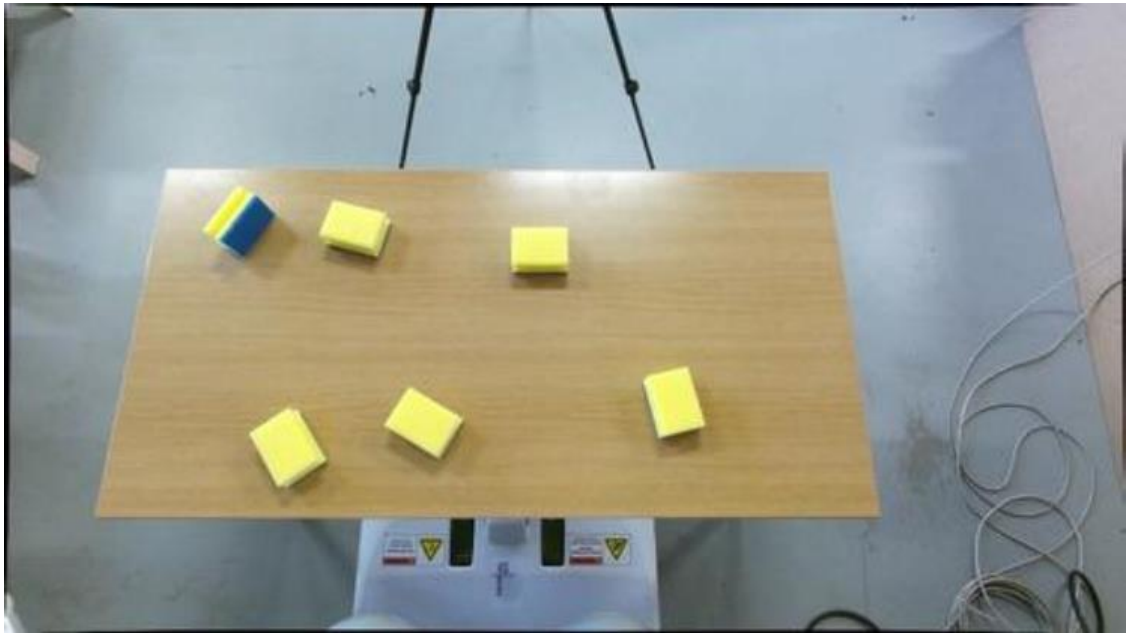
Object singulation = physically separating objects in cluttered tabletop scenes

Singulation and Interactive Object Perception

How many objects are in the image?

Singulation and Interactive Object Perception

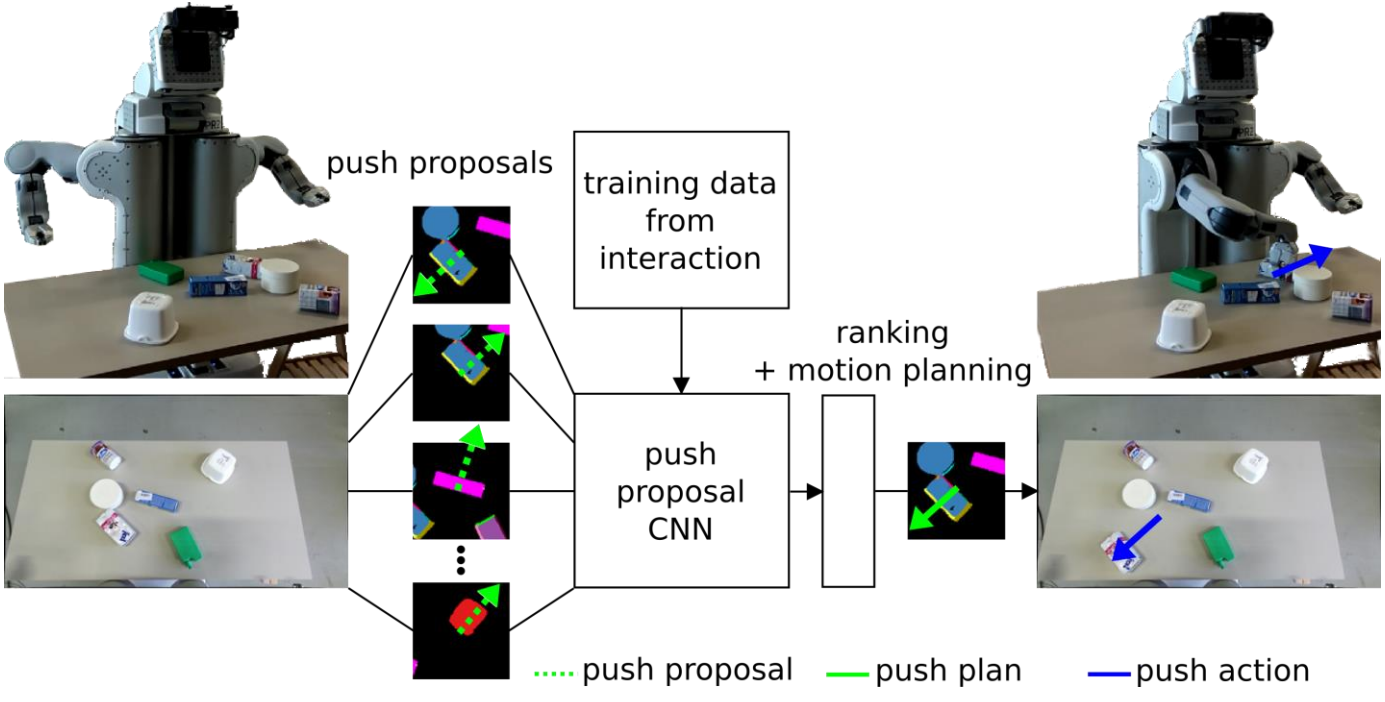
Singulation facilitates perception



Contributions

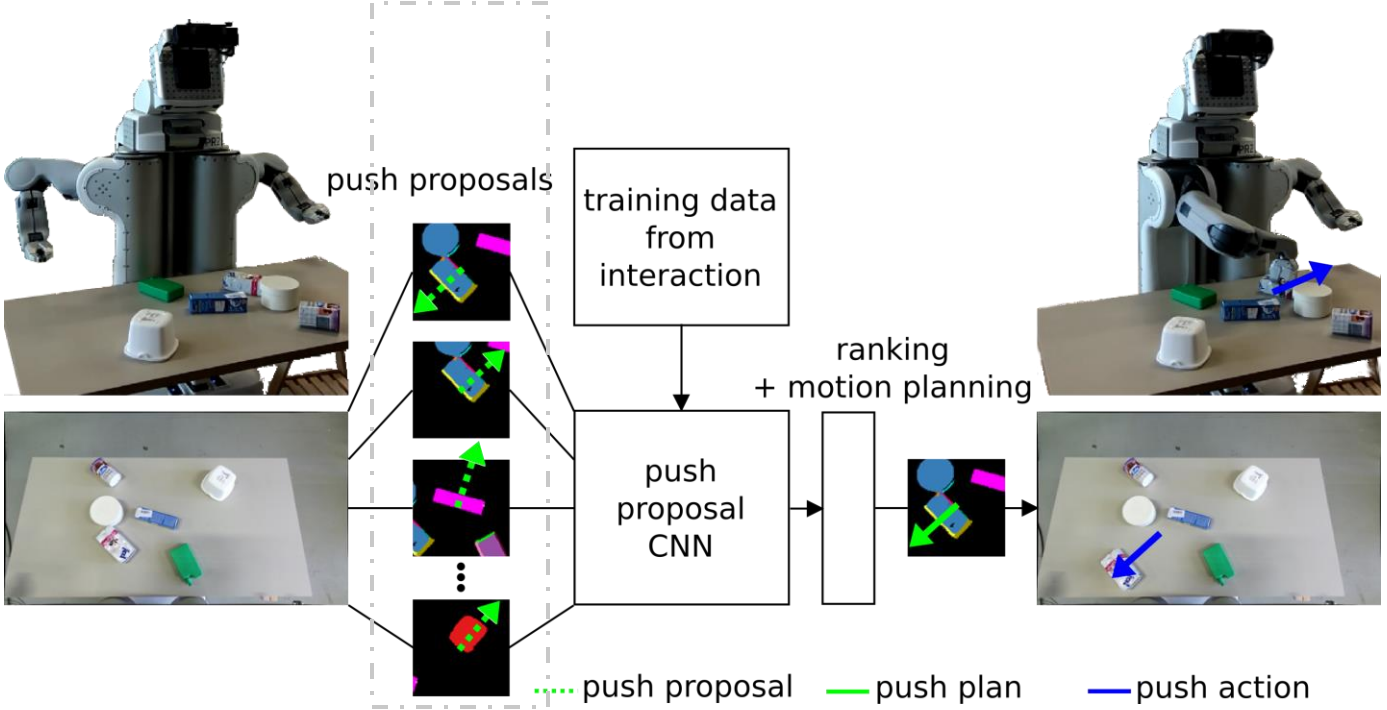
- We train a CNN to detect favourable push actions from over-segmented images in order to clear clutter
- In comparison to previous work
  1. Model-free approach, no physics simulator, no object knowledge
  2. Learn features automatically

Approach


push proposals
training data from interaction
ranking + motion planning
push proposal CNN
push proposal
push plan
push action

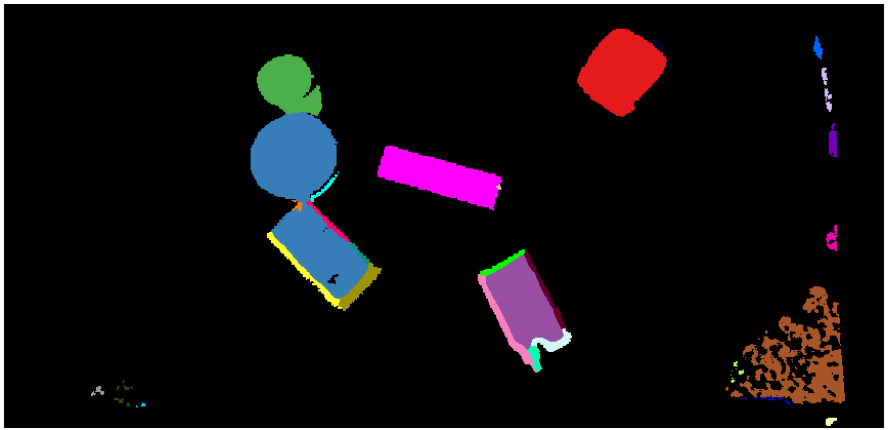
Approach

1. Sample push proposals from over-segmented RGB-D image



Over-segmentation

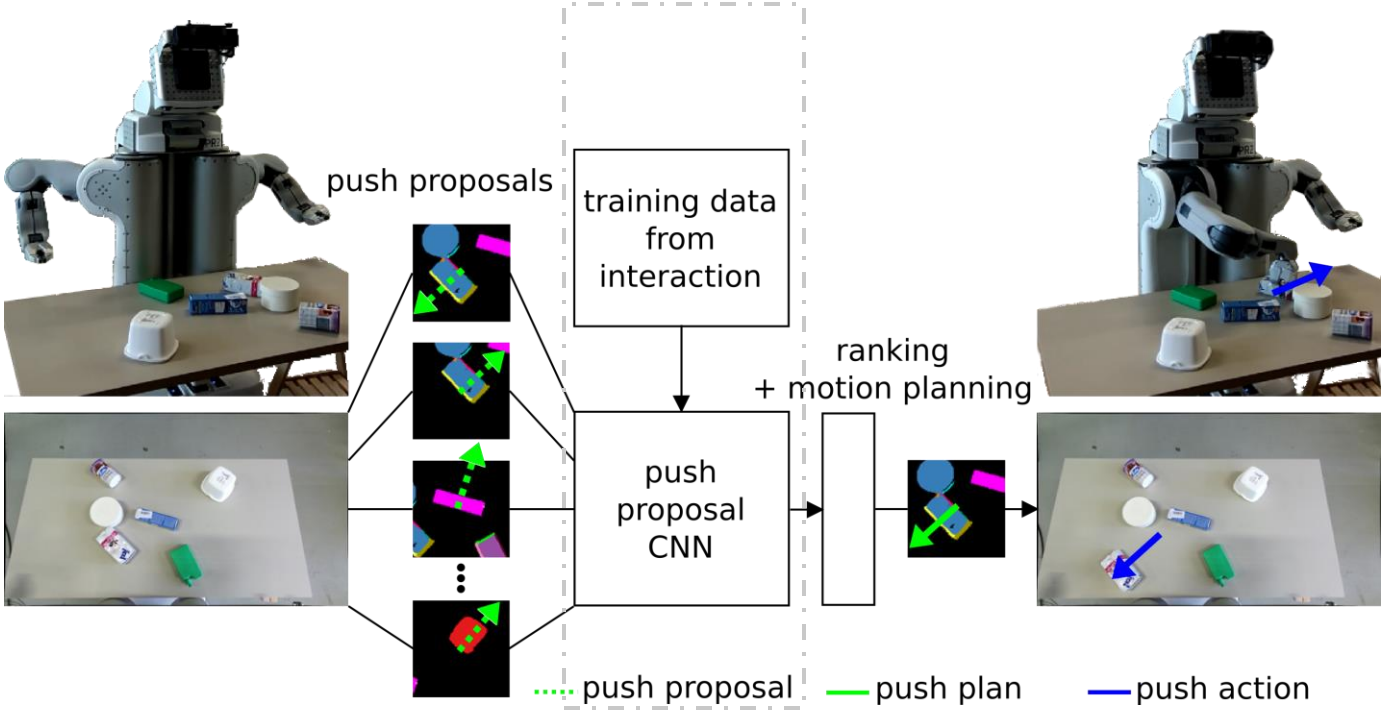
Objects get segmented into multiple facets using RGB-D Segmenter [1]



[1] Richtsfeld et al. 2012

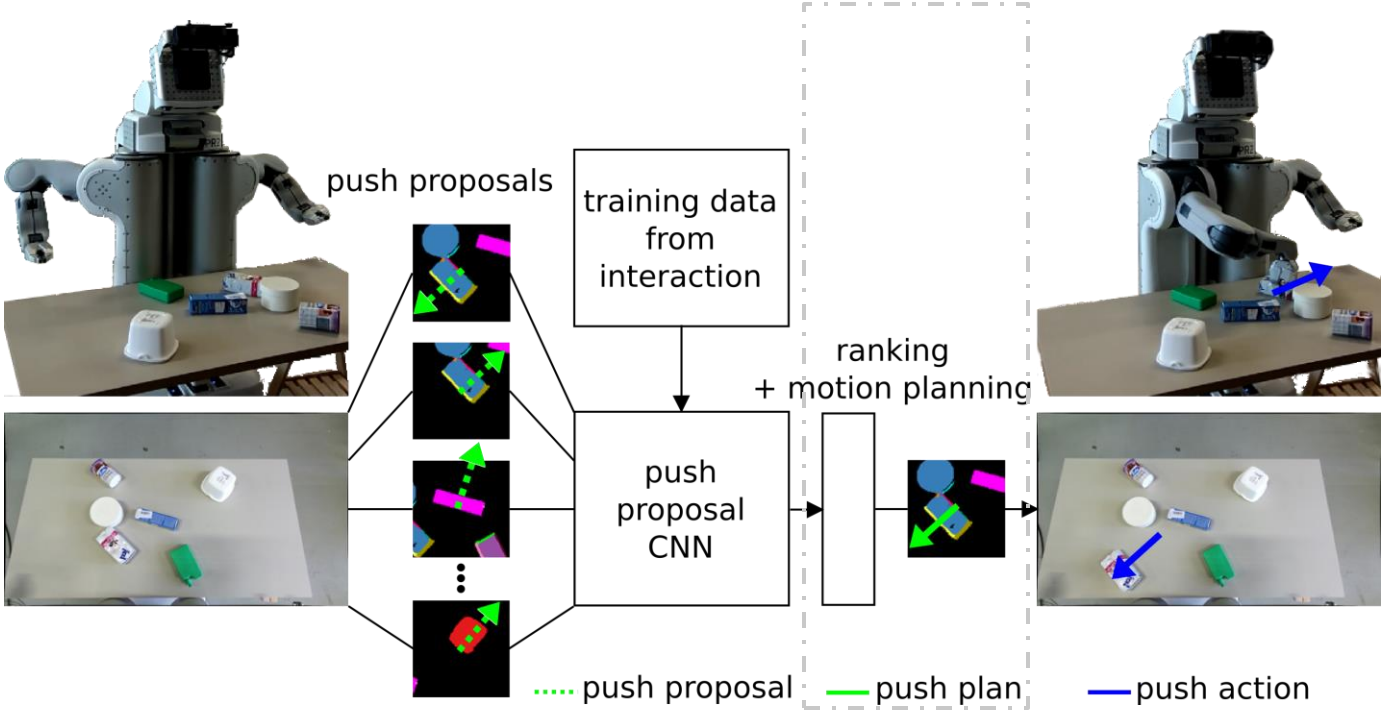
Approach

2. Classify set of sampled push proposals with push proposal network



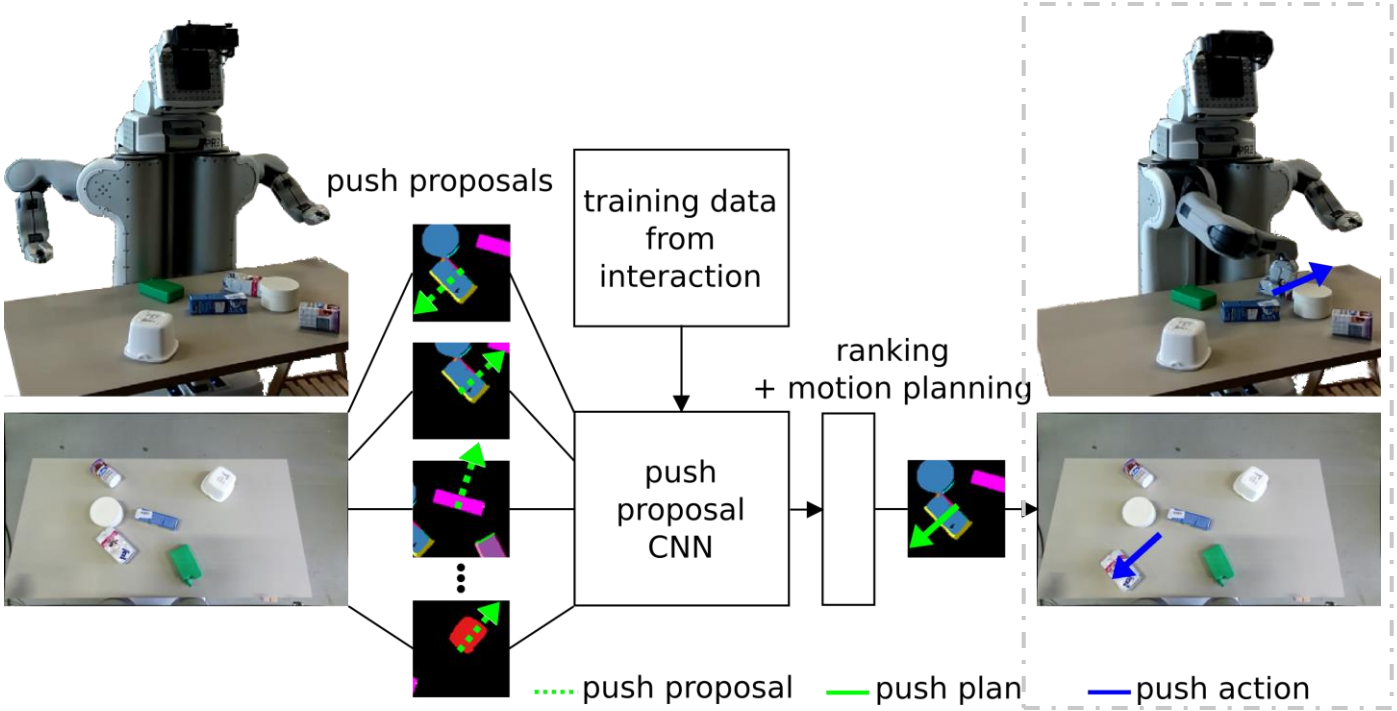
Approach

3. Perform motion planning



Approach

4. Execute first successful motion plan

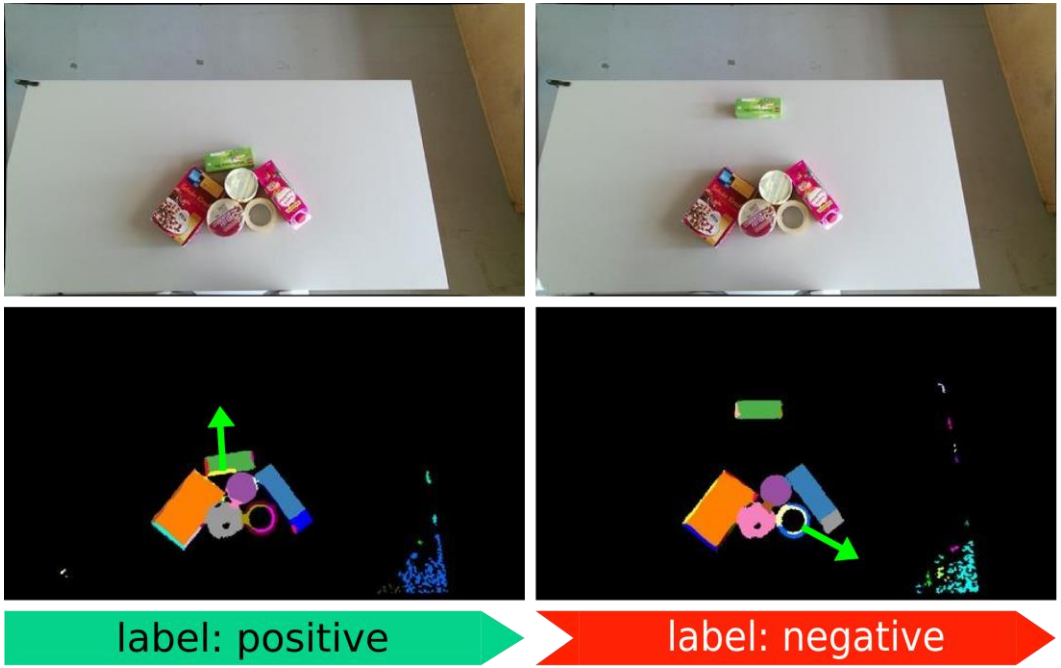


Definitions

- Neural network $F(\mathbf{o}, \mathbf{a}; \theta)$ with parameters θ
- Input is an over-segmented image $\mathbf{o}$ and a push proposal action $\mathbf{a} = (\mathbf{c}, \alpha)$
- The push proposal consists of a start position pixel $\mathbf{c} = (x, y)$ and a push angle α both defined in the image plane

Supervised Learning

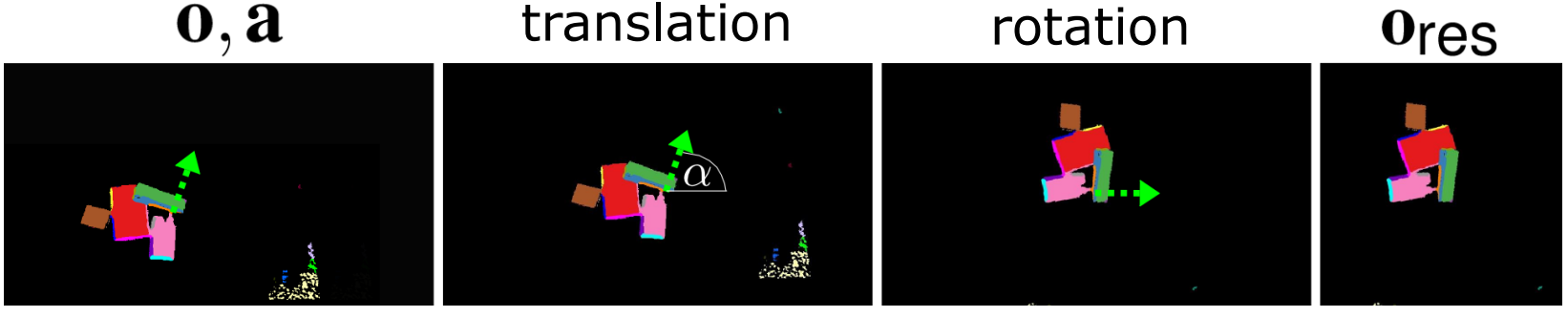
Training data is labeled by an expert user who gives a binary label for successful or unsuccessful push action



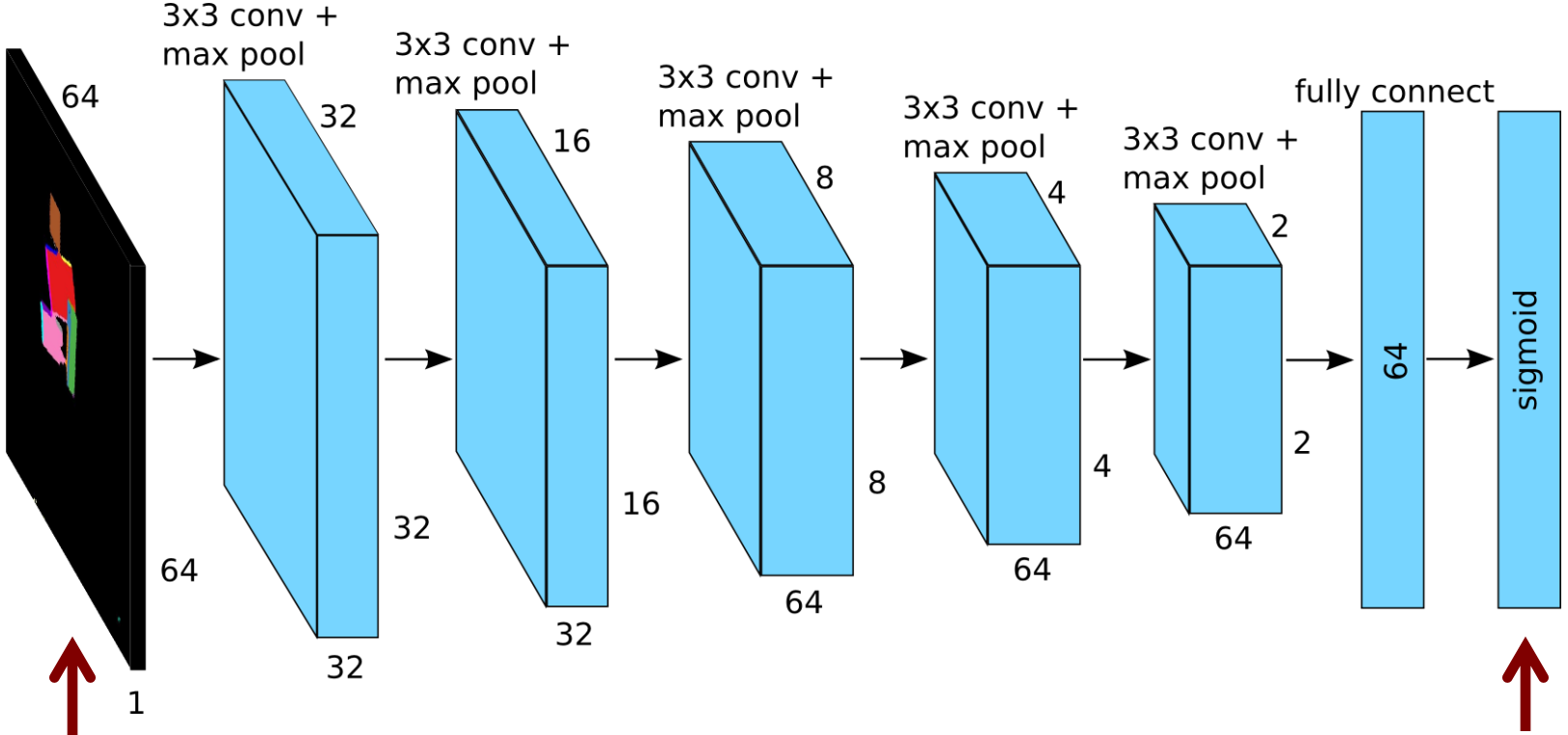
Iterative Training

How to Fuse Image and Push Proposal Action

- Key idea: only need to capture local context between objects, not global
- Fuse image $\mathbf{o}$ and push proposal action $\mathbf{a}$ using rigid image transformations
- Result is a local push-centric image $\mathbf{o}_{\text{res}}$



Push Proposal CNN



Gets push-centric image $\mathbf{o}_{res}$ as input

Predicts probability of singulation success for one proposal

Experimental Setup

- PR2 robot with Kinect 2
- All experiments with unknown objects in cluttered initial configurations
- Increasing difficulty level 4-8 objects
- Singulation trial is successful if all objects are separated by at least 3cm

Results with 6 and 8 Objects

success

fail, two objects at top still touching

success

fail, no motion plan found objects too close to robot

Quantitative Results

- Success rate of our best network
  - 6 objects 70%, 25 trials
  - 8 objects 40%, 10 trials
- Improvement with respect to manual baseline method is 30%

Video

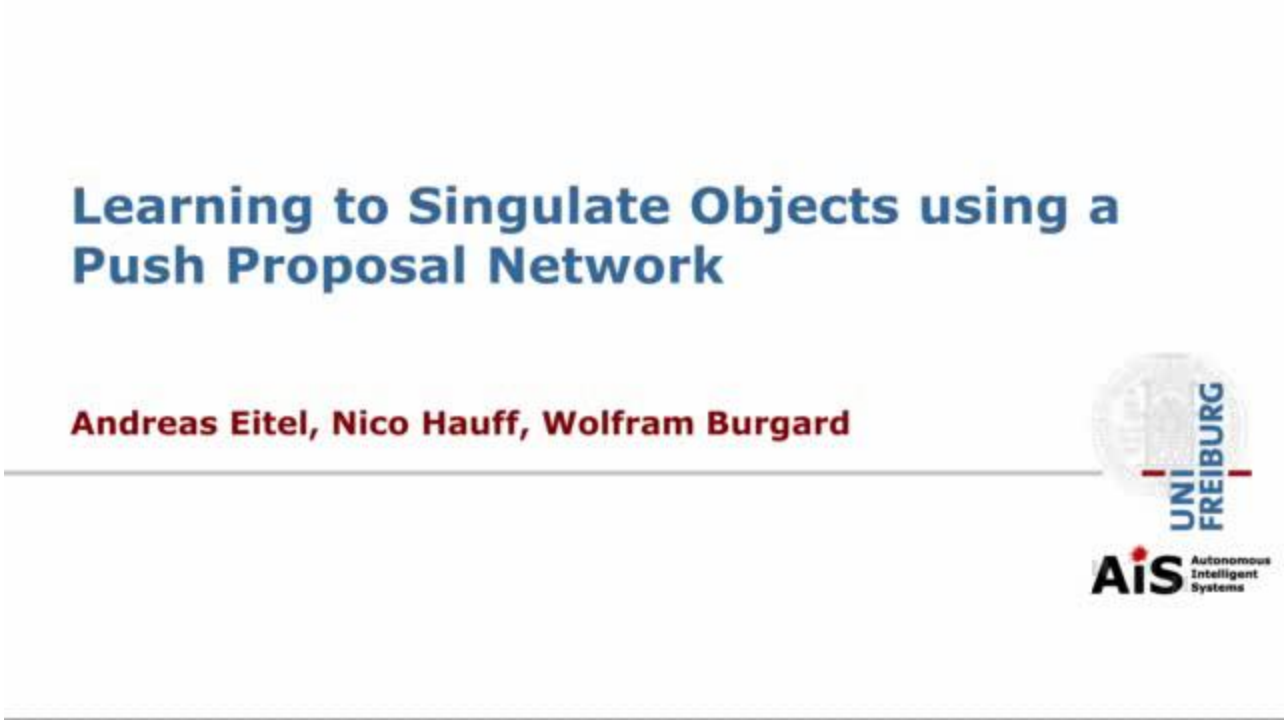
Conclusions

- Novel learning-based approach for clearing object clutter based on CNN
- Neural network generalizes well to novel objects and cluttered object configurations
- Novel method for fusing image and action representation into network
- Successful singulation experiments with up to 8 cluttered objects

Future Work

- Move from supervised to semi- and self-supervised learning
- Extension of network with multi-class output for prediction of varying push lengths

Thank you for your attention!



Learning to Singulate Objects using a Push Proposal Network

http://robotpush.cs.uni-freiburg.de