# DeepRob

**Seminar 3**
**Object Pose, Geometry, SDF, Implicit Surfaces**
**University of Michigan and University of Minnesota**

# This Week: Rigid Body Objects

- Seminar 3: Object Pose, Geometry, SDF, Implicit Surfaces
    1. SUM: Sequential scene understanding and manipulation, Sui et al., 2017
    2. iSDF: Real-Time Neural Signed Distance Fields for Robot Perception, Oriz et al., 2022

- Seminar 4: Dense Descriptors, Category-level Representations
    1. Dense Object Nets: Learning Dense Visual Object Descriptors By and For Robotic Manipulation, Florence et al., 2018
    2. Normalized Object Coordinate Space for Category-Level 6D Object Pose and Size Estimation, Wang et al., 2019
    3. kPAM: KeyPoint Affordances for Category-Level Robotic Manipulation, Manuelli et al., 2019
    4. Single-Stage Keypoint-Based Category-Level Object Pose Estimation from an RGB Image, Lin et al., 2022

# Today: Object Pose, Geometry, SDF, Implicit Surfaces

- Seminar 3: Object Pose, Geometry, SDF, Implicit Surfaces
  1. SUM: Sequential scene understanding and manipulation, Sui et al., 2017
  2. iSDF: Real-Time Neural Signed Distance Fields for Robot Perception, Oriz et al., 2022

- Seminar 4: Dense Descriptors, Category-level Representations
  1. Dense Object Nets: Learning Dense Visual Object Descriptors By and For Robotic Manipulation, Florence et al., 2018
  2. Normalized Object Coordinate Space for Category-Level 6D Object Pose and Size Estimation, Wang et al., 2019
  3. kPAM: KeyPoint Affordances for Category-Level Robotic Manipulation, Manuelli et al., 2019
  4. Single-Stage Keypoint-Based Category-Level Object Pose Estimation from an RGB Image, Lin et al., 2022

# SUM

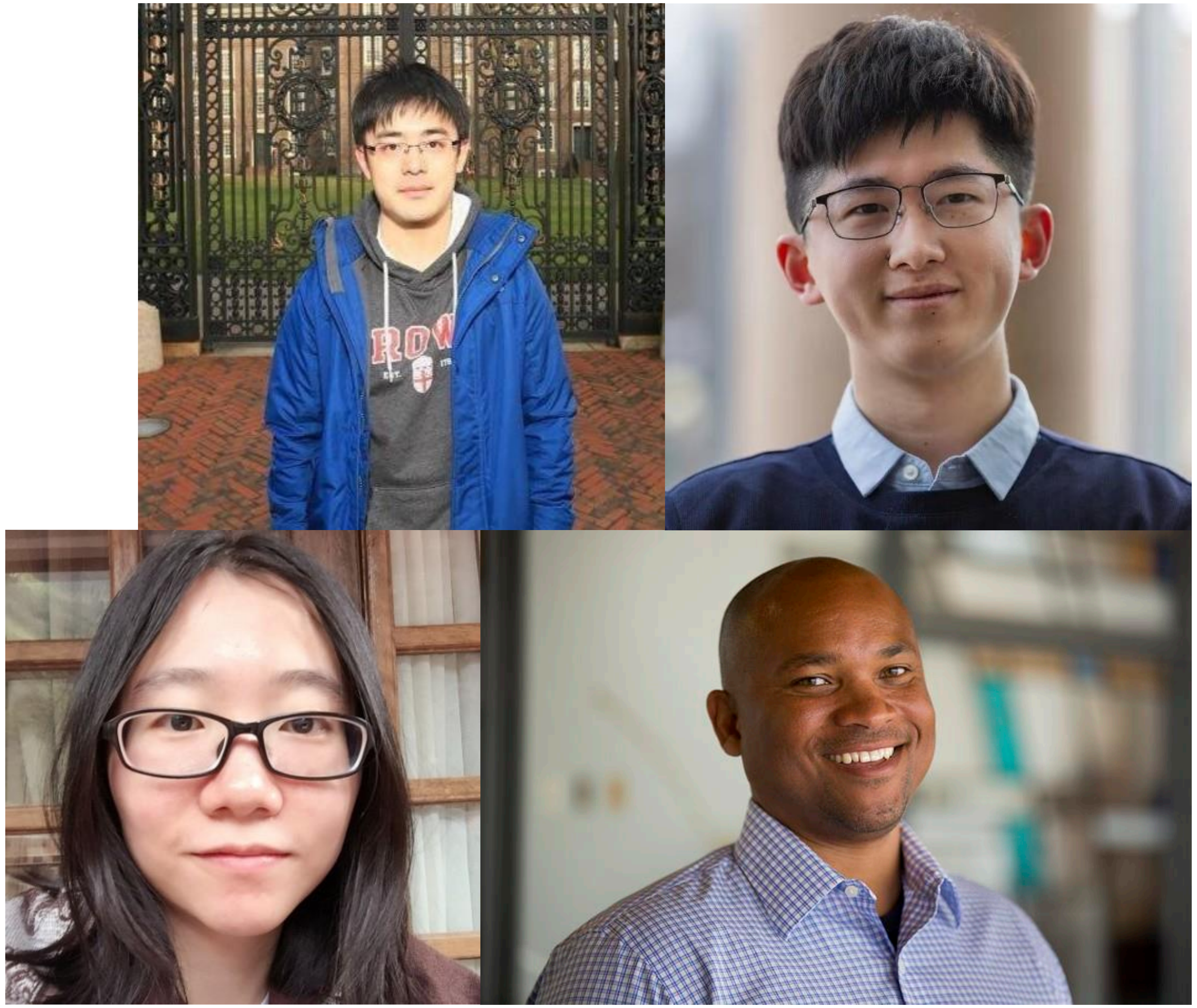
## Sequential Scene Understanding and Manipulation

By: Zhiqiang Sui, Zheming Zhou, Zhen Zeng, Odest Chadwicke Jenkins

Presented by: Daniel Simmons

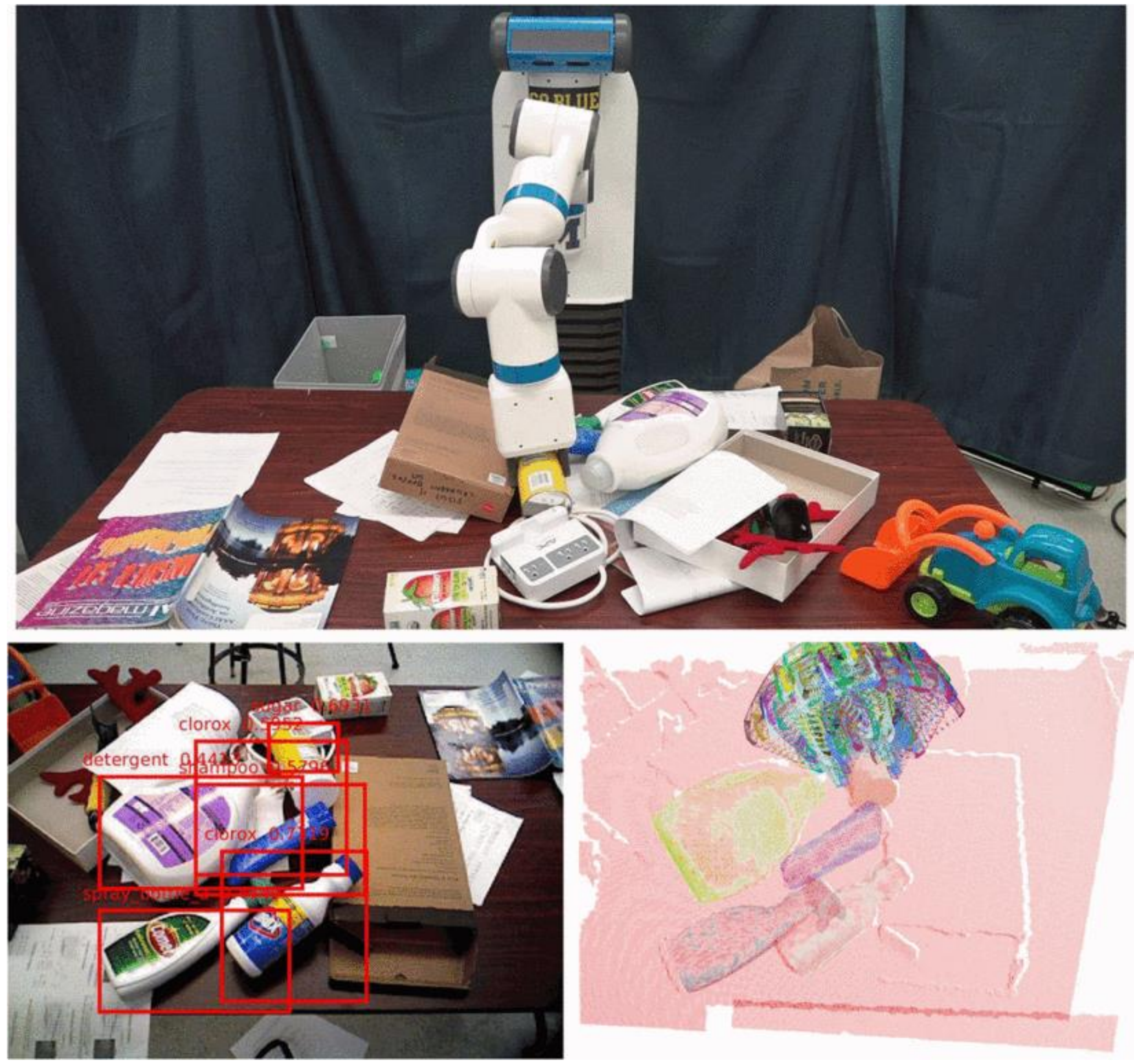
# The Handsome Authors

- Zhiqiang Sui
  - Former PhD Student at UMich
- Zheming Zhou
  - Former PhD Student at UMich
- Zhen Zeng
  - Former PhD Student at UMich
- Odest Chadwicke Jenkins
  - Professor at UMich



# Robots can't deal with this


- Need to find  the cleaning items!




**Human Environment**

# Cleaning up the mess

- Recognizing objects in cluttered environments is a critical challenge for a variety of tasks


- Video can be found at:

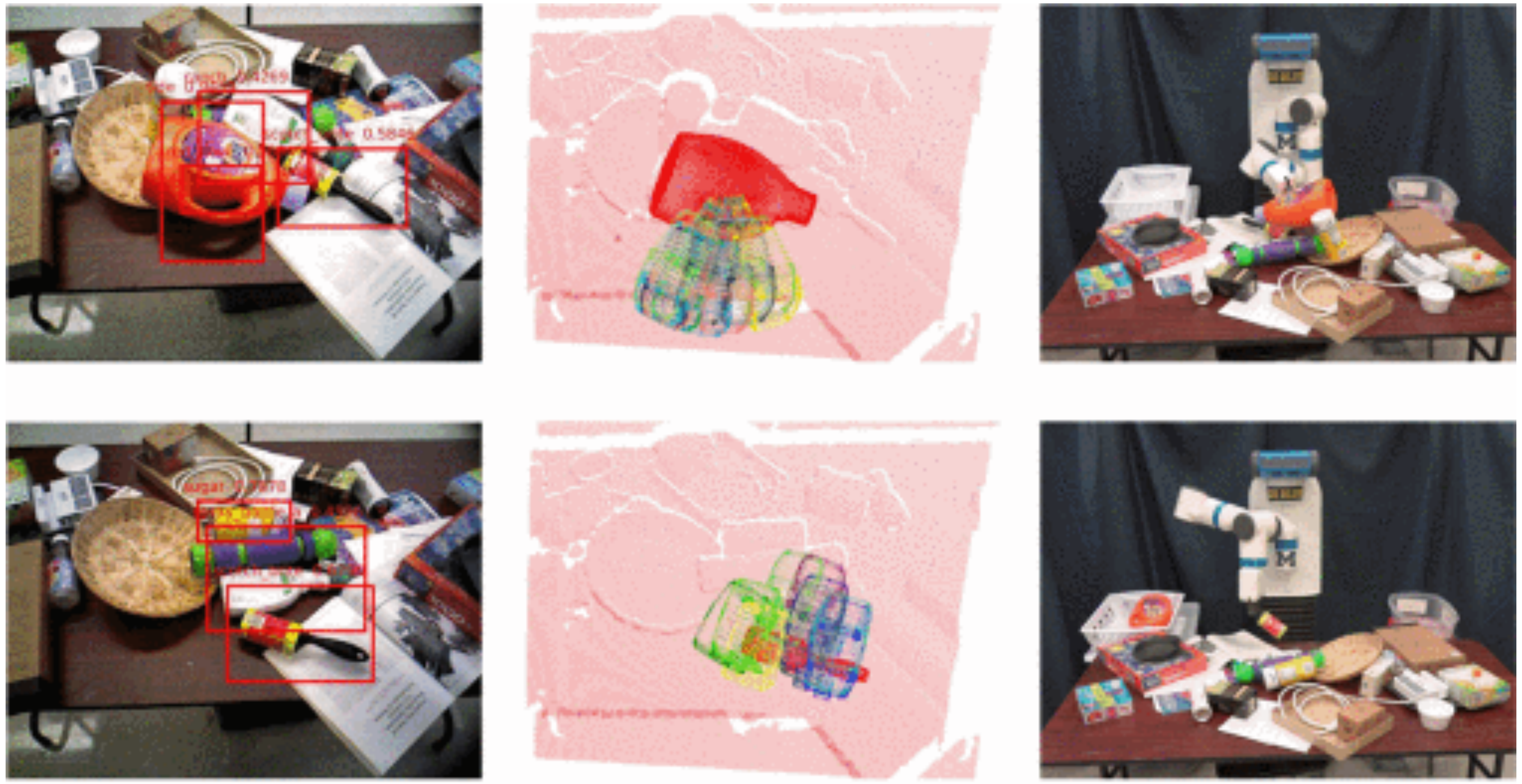
https://www.youtube.com/watch?v=ry0mqY5I-04

# Addressing Limitations

- Previous models assumed object detection was always accurate
- Objects were assumed to be static
- Generative methods were often used to predict layout

# Approach

1. Probability-evaluated object detection
2. Physics and state models
3. Pose evaluation

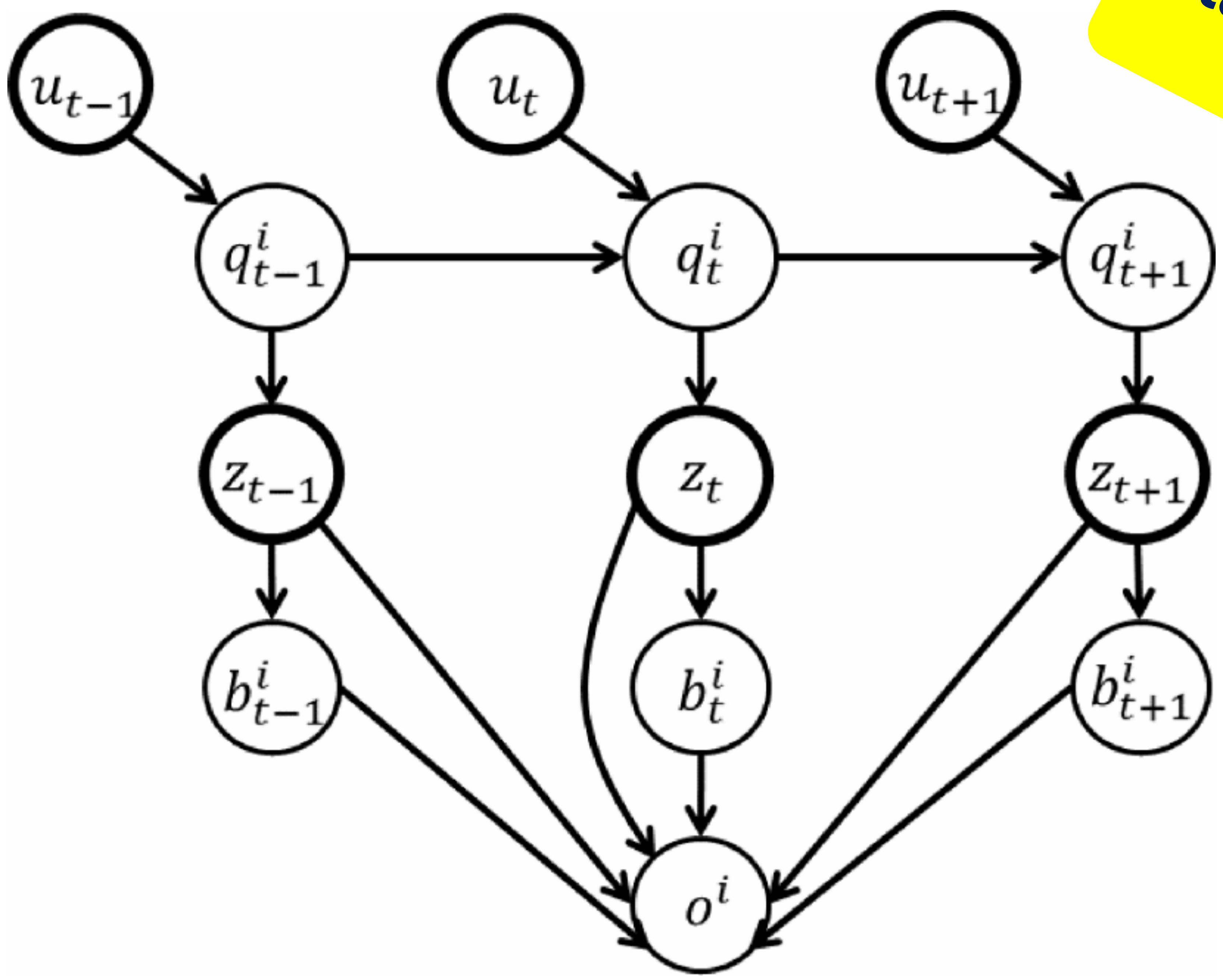


# State Estimation

- Given RGBD observations estimate as objects with labels and poses
- Iterated with a bayes filter to assess accuracy and particle Filter

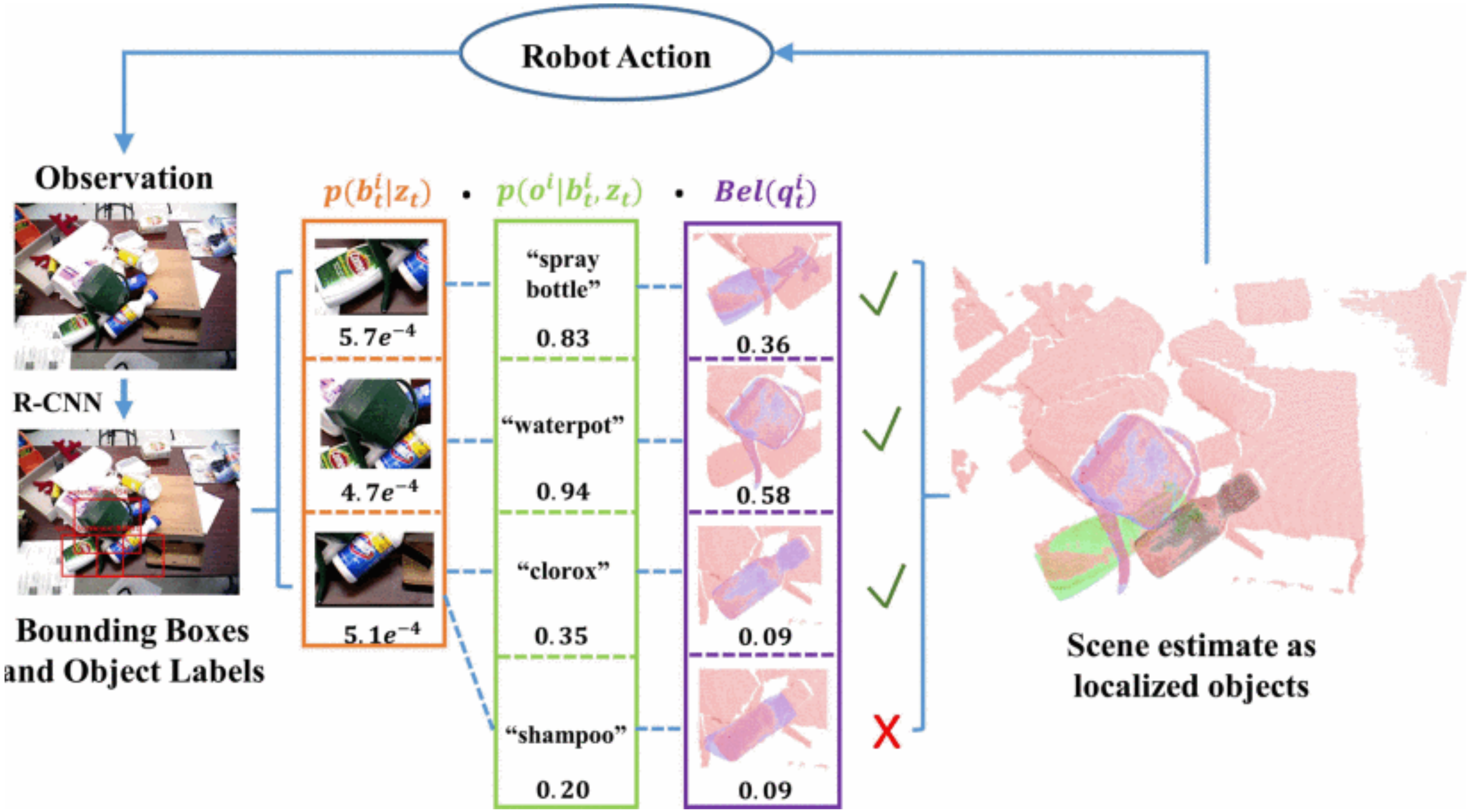
Probability!

$$p(x_t^i|z_{0:t},u_{1:t}) = \underbrace{p(b_t^i|z_t)}_{detection}\underbrace{p(o^i|b_t^i,z_t)}_{recognition}\underbrace{p(q_t^i|b_t^i,o^i,z_{0:t},u_{1:t})}_{Bel(q_t^i)}$$




State Estimation!

# The Loop


Robot Action
Observation
R-CNN
Bounding Boxes
and Object Labels
$p(b_t^i|z_t)$ • $p(o^i|b_t^i, z_t)$ • $Bel(q_t^i)$
5.7e−4
4.7e−4
5.1e−4
"spray bottle"
0.83
"waterpot"
0.94
"clorox"
0.35
"shampoo"
0.20
0.36
0.58
0.09
0.09
Scene estimate as
localized objects

# Trials

- 8 experiments with a Fetch Mobile Manipulation Robot
  - 15 objects, 625 particles with 20 resampling iterations

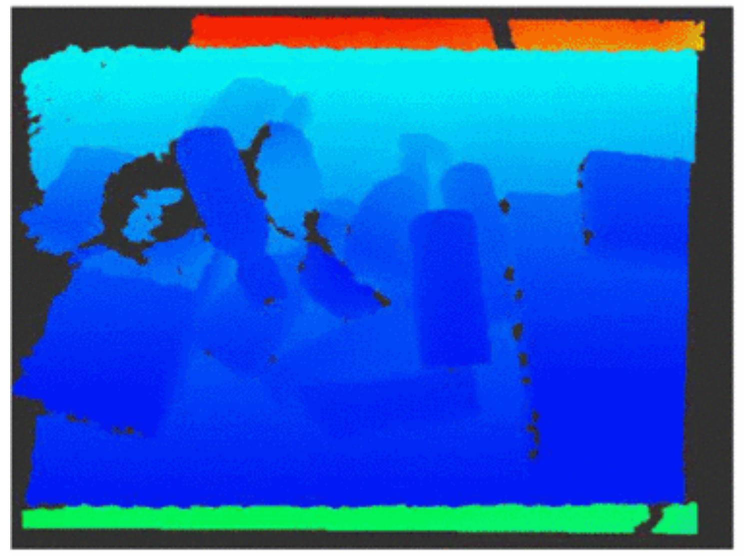
Dataset

(a)

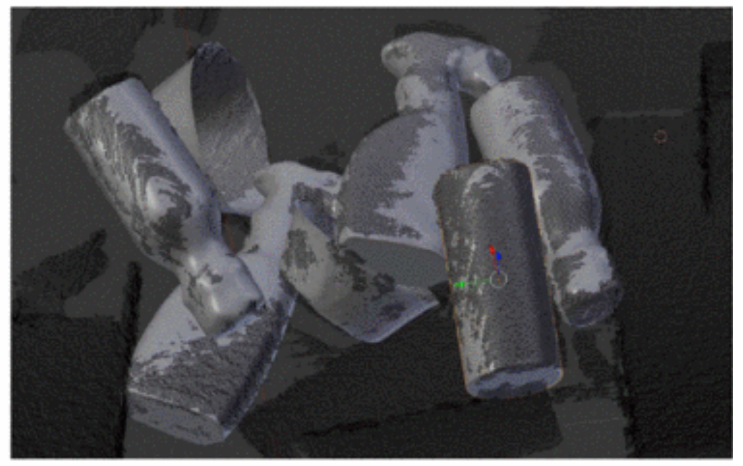
(b)

Clutter

Depth Map



(c)



(d)

Ground Truth

# Sequence Table

| | Sequence (a) | Sequence (b) | Sequence (c) | Sequence (d) | Sequence (e) | Sequence (f) | Sequence (g) | Sequence (h) |
|---|---|---|---|---|---|---|---|---|
| Number of total objects | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| Number of Manipulation Errors | 1 | 1 | 2 | 0 | 0 | 1 | 1 | 0 |
| Number of Manipulation Trials | 4 | 6 | 7 | 5 | 5 | 5 | 6 | 5 |
| Completion Ratio | 0.80 | 1.0 | 1.0 | 1.0 | 1.0 | 0.8 | 1.0 | 1.0 |

# Sequence Table

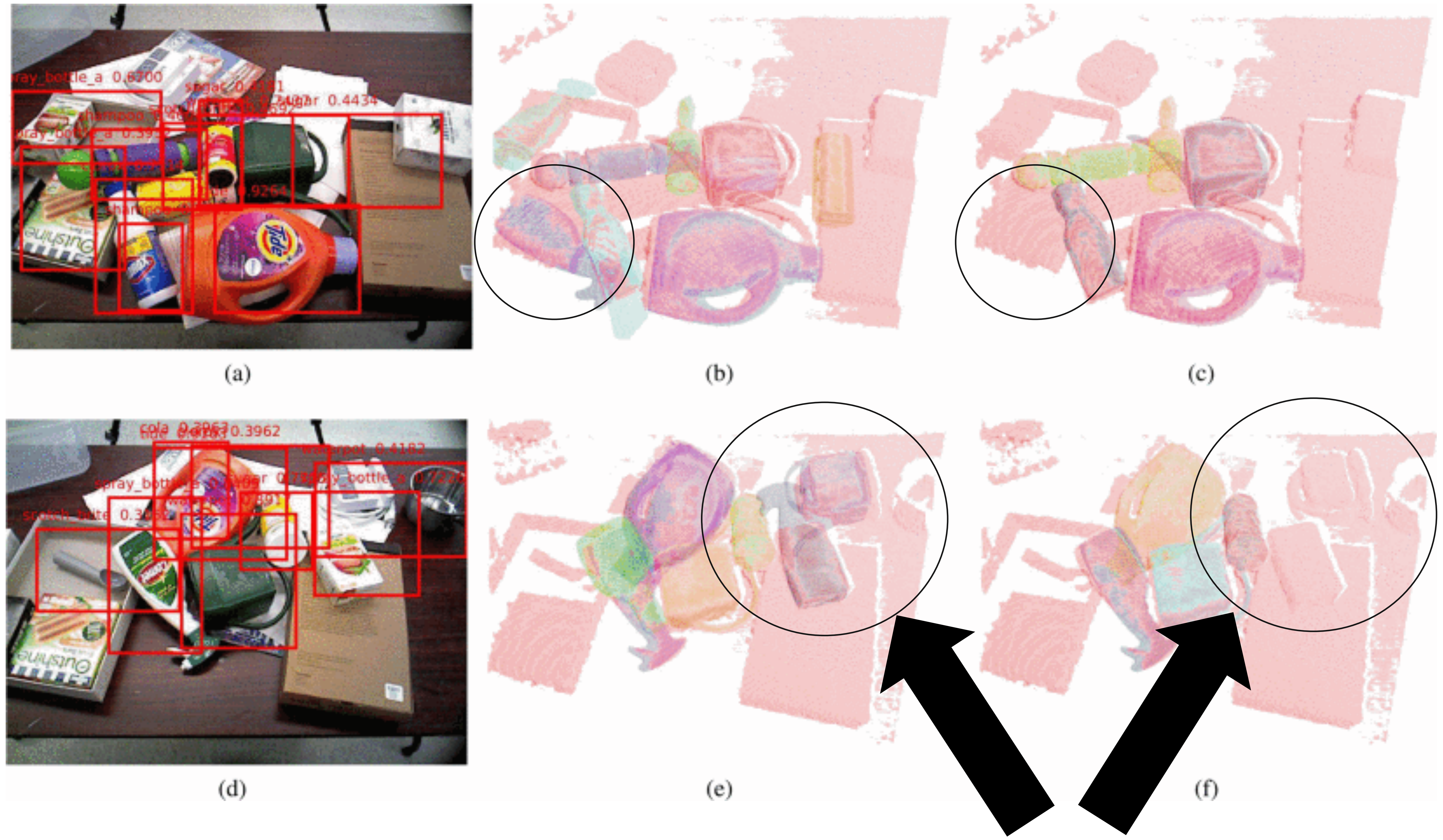
| | Sequence (a) | Sequence (b) | Sequence (c) | Sequence (d) | Sequence (e) | Sequence (f) | Sequence (g) | Sequence (h) |
|---|---|---|---|---|---|---|---|---|
| Number of total objects | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| Number of Manipulation Errors | 1 | 1 | 2 | 0 | 0 | 1 | 1 | 0 |
| Number of Manipulation Trials | 4 | 6 | 7 | 5 | 5 | 5 | 6 | 5 |
| Completion Ratio | 0.80 | 1.0 | 1.0 | 1.0 | 1.0 | 0.8 | 1.0 | 1.0 |

# Sequence Table

| | Sequence (a) | Sequence (b) | Sequence (c) | Sequence (d) | Sequence (e) | Sequence (f) | Sequence (g) | Sequence (h) |
|---|---|---|---|---|---|---|---|---|
| Number of total objects | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| Number of Manipulation Errors | 1 | 1 | 2 | 0 | 0 | 1 | 1 | 0 |
| Number of Manipulation Trials | 4 | 6 | 7 | 5 | 5 | 5 | 6 | 5 |
| Completion Ratio | 0.80 | 1.0 | 1.0 | 1.0 | 1.0 | 0.8 | 1.0 | 1.0 |

# False Positives

- Able to overcome mistakes



- Object detection made mistakes
- Scene estimation recognized incorrect objects

# Conclusions

- SUM is a generative and discriminative approach
  - Maintains belief over a sequence of actions
- Provides robust estimation and manipulation
  - Can parse and sort cluttered environments

# Limitations and Directions for Future Work

- Limitations
  - Can solve for impossible joint positions
  - Frequent manipulation erros
- Future directions
  - Test with data other than RGBD
  - Apply motion to objects

# Thank you

# Next Time: Dense Descriptors, Category-level Representations

- Seminar 3: Object Pose, Geometry, SDF, Implicit Surfaces
    1. SUM: Sequential scene understanding and manipulation, Sui et al., 2017
    2. iSDF: Real-Time Neural Signed Distance Fields for Robot Perception, Oriz et al., 2022

- Seminar 4: Dense Descriptors, Category-level Representations
    1. Dense Object Nets: Learning Dense Visual Object Descriptors By and For Robotic Manipulation, Florence et al., 2018
    2. Normalized Object Coordinate Space for Category-Level 6D Object Pose and Size Estimation, Wang et al., 2019
    3. kPAM: KeyPoint Affordances for Category-Level Robotic Manipulation, Manuelli et al., 2019
    4. Single-Stage Keypoint-Based Category-Level Object Pose Estimation from an RGB Image, Lin et al., 2022

# DeepRob

**Seminar 3**
**Object Pose, Geometry, SDF, Implicit Surfaces**
**University of Michigan and University of Minnesota**