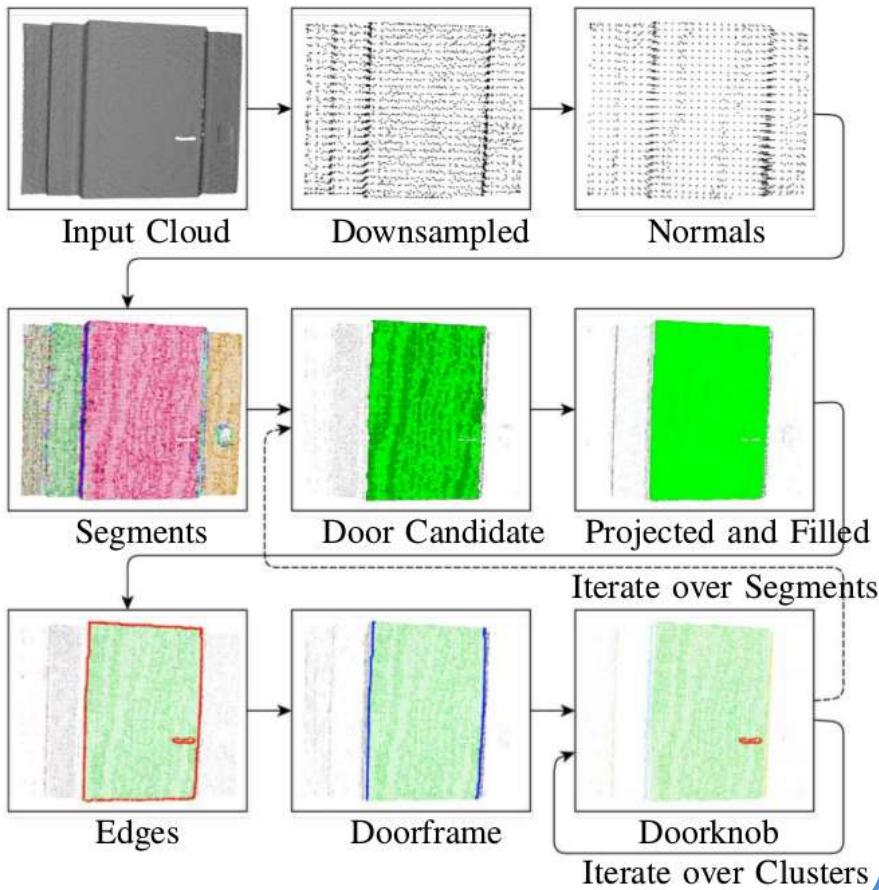


Door and Cabinet Recognition Using Convolutional Neural Nets and Real-time Method Fusion for Handle Detection and Grasping

Adrian Llopart, Nils A. Andersen and Ole Ravn

Other detection approaches

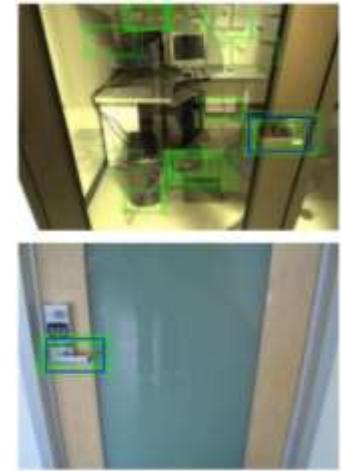
Depth



Labels/Signs



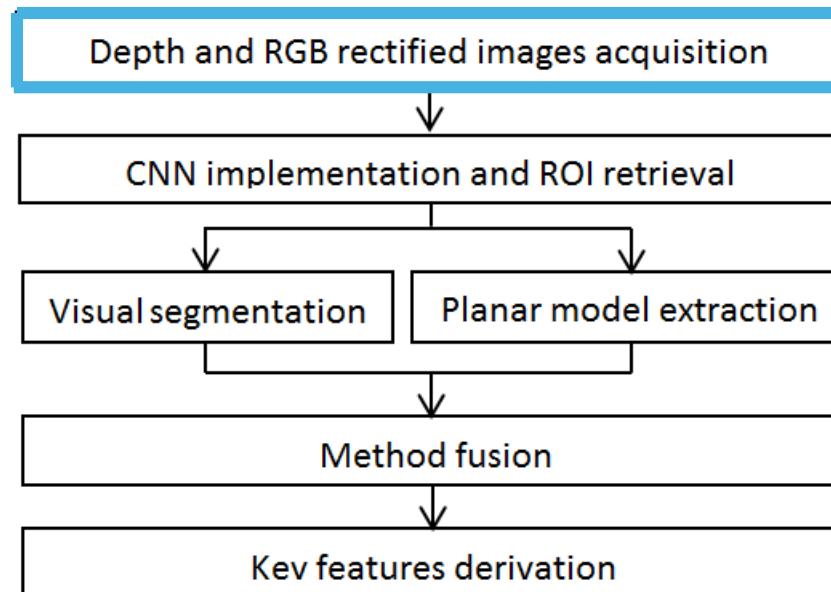
Handles



Visual geometric lines



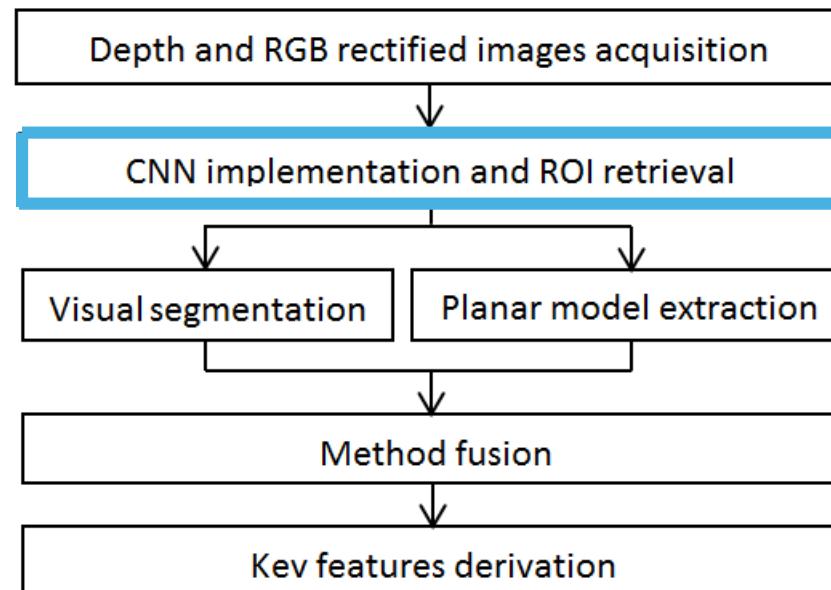
My proposed pipeline



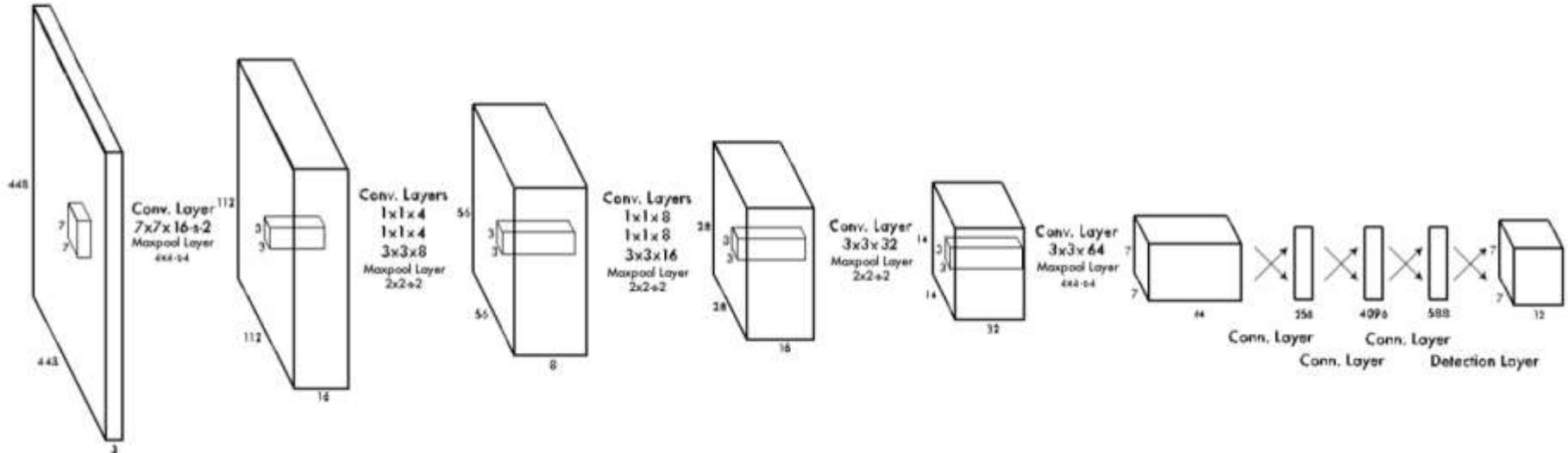
Robot description



My proposed pipeline



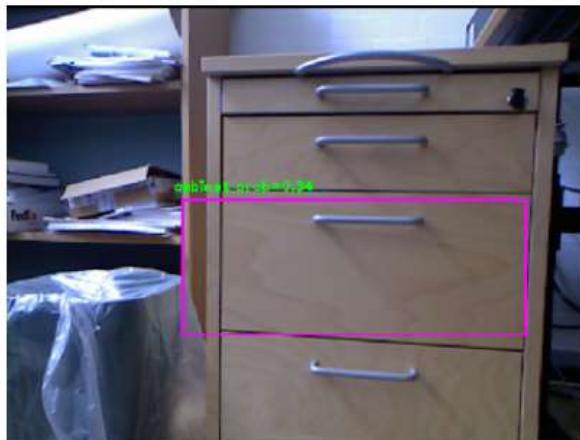
You Only Look Once (YOLO) detection system



Darknet Reference Model

- 2 classes (510 door images and 420 cabinet images)
- 14 Hz
- GeForce GT 730

CNN: Example results

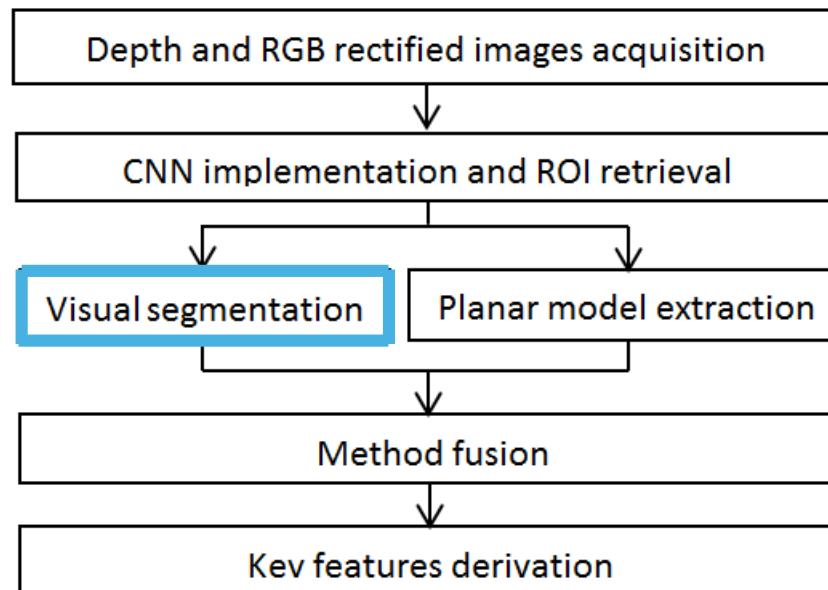


CNN: Comparison

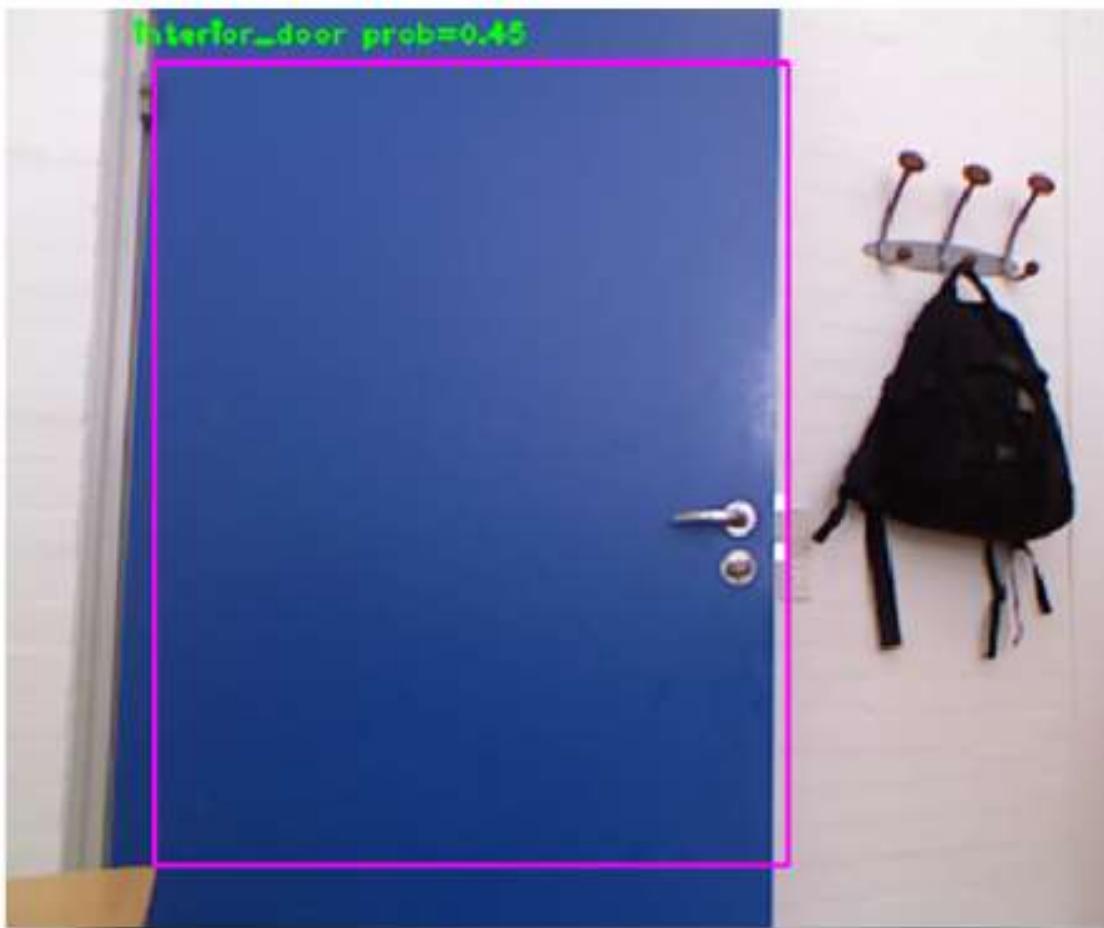
TABLE I. DOOR DETECTION

	Shalaby et. al 2014			Yang et. al 2010			Proposed CNN		
	N _d	TP	FP	N _d	TP	FP	N _d	TP	FP
Simple	55	100%	3.6%	55	98.2%	1.8%	63	98.4%	7.53%
Medium	94	82.7%	17.3%	93	91.4%	1.1%	107	95.3%	4.7%
Complex	63	68.2%	31.7%	56	85.7%	7.1%	71	74.6%	8.45%
Total	212	82.8%	26.4%	204	91.7%	2.9%	241	90.0%	6.64%

My proposed pipeline



Visual Segmentation



Assumptions:

- Correct ROI
- Color differentiation

Process:

1. ROI extraction
2. K-means color clusterization
3. Canny edge detection
4. Contour size, form and orientation evaluation
5. Binary mask creation
6. Disparity image masking
7. Handle PCL derivation
8. False positives elimination (height)

False positives:

- Lock, signs, posters

Visual Segmentation



Assumptions:

- Correct ROI
- Color differentiation

Process:

1. ROI extraction
2. K-means color clusterization
3. Canny edge detection
4. Contour size, form and orientation evaluation
5. Binary mask creation
6. Disparity image masking
7. Handle PCL derivation
8. False positives elimination (height)

False positives:

- Lock, signs, posters

Visual Segmentation



Assumptions:

- Correct ROI
- Color differentiation

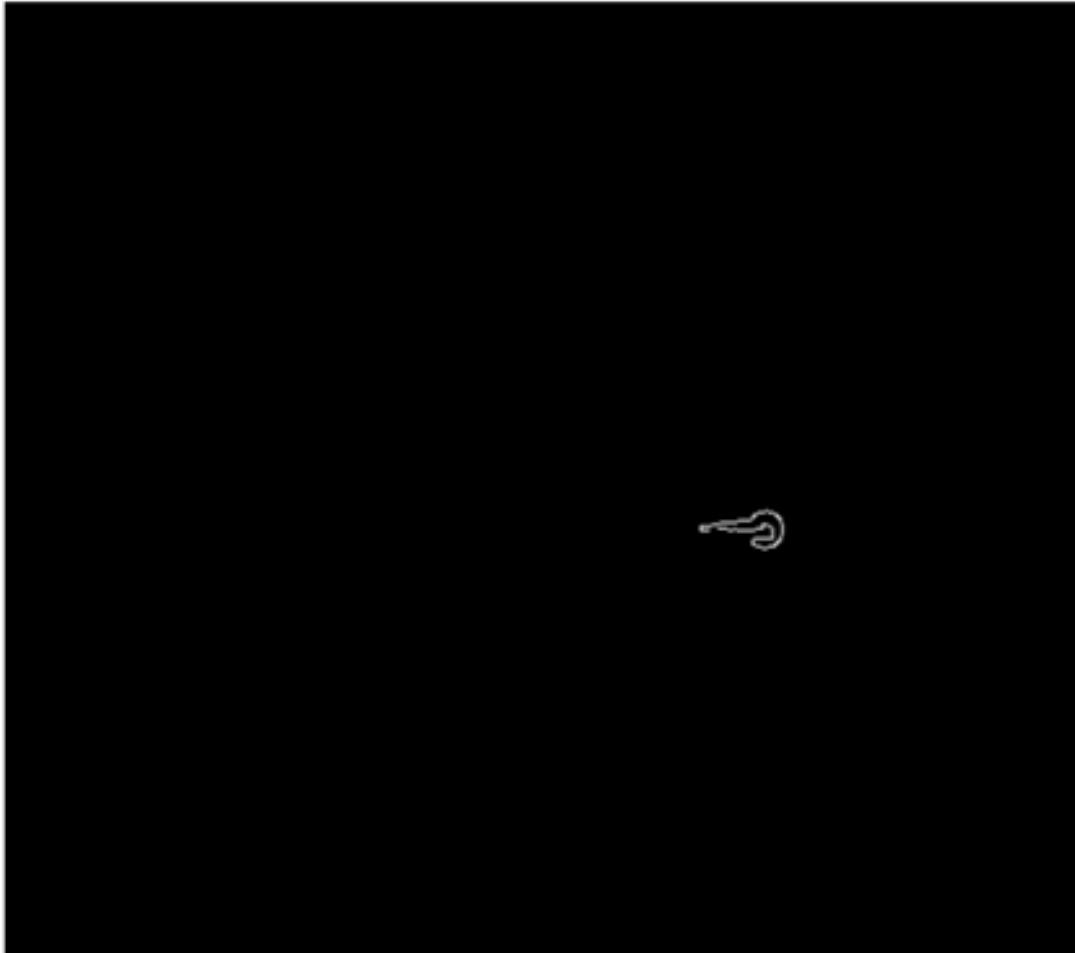
Process:

1. ROI extraction
2. K-means color clusterization
3. Canny edge detection
4. Contour size, form and orientation evaluation
5. Binary mask creation
6. Disparity image masking
7. Handle PCL derivation
8. False positives elimination (height)

False positives:

- Lock, signs, posters

Visual Segmentation



Assumptions:

- Correct ROI
- Color differentiation

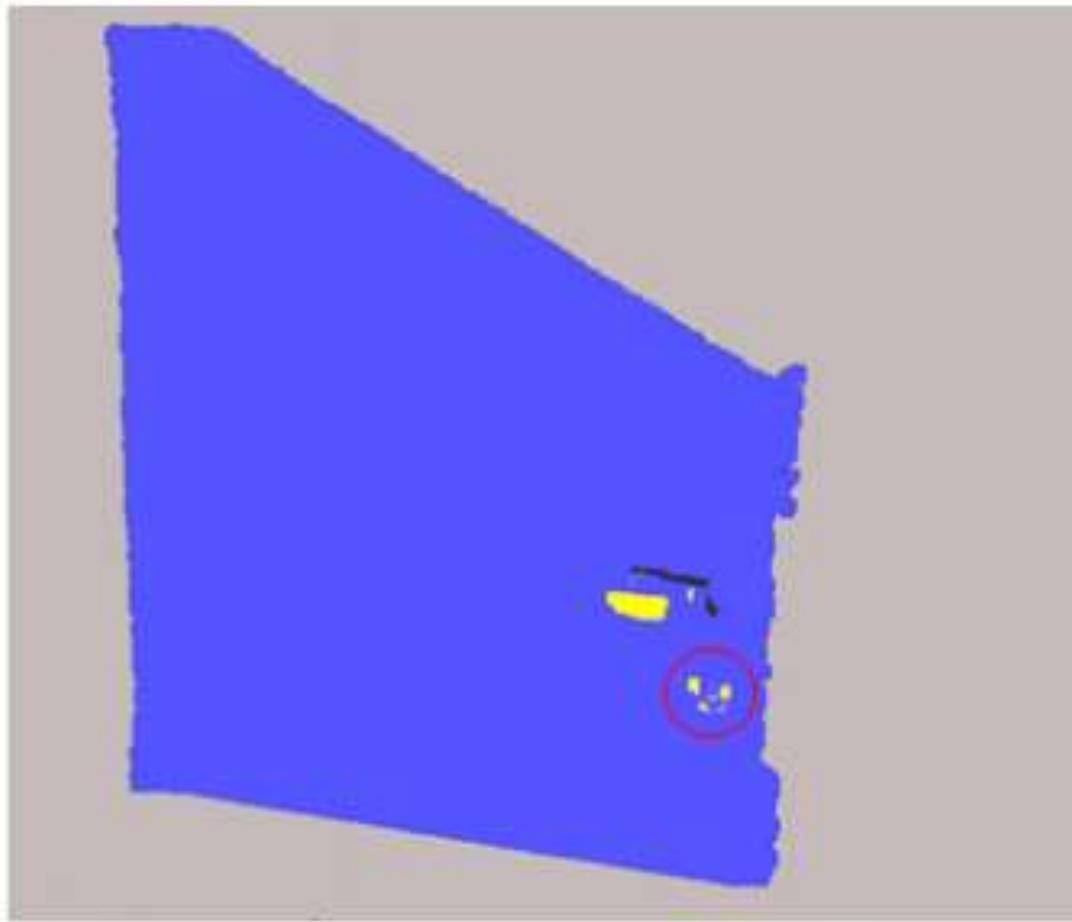
Process:

1. ROI extraction
2. K-means color clusterization
3. Canny edge detection
4. Contour size, form and orientation evaluation
5. Binary mask creation
6. Disparity image masking
7. Handle PCL derivation
8. False positives elimination (height)

False positives:

- Lock, signs, posters

Visual Segmentation



Assumptions:

- Correct ROI
- Color differentiation

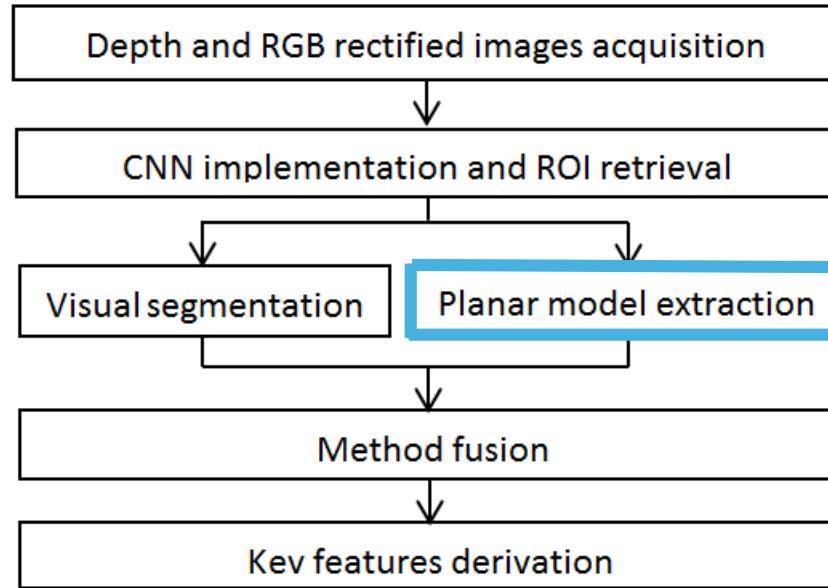
Process:

1. ROI extraction
2. K-means color clusterization
3. Canny edge detection
4. Contour size, form and orientation evaluation
5. Binary mask creation
6. Disparity image masking
7. Handle PCL derivation
8. False positives elimination (height)

False positives:

- Lock, signs, posters

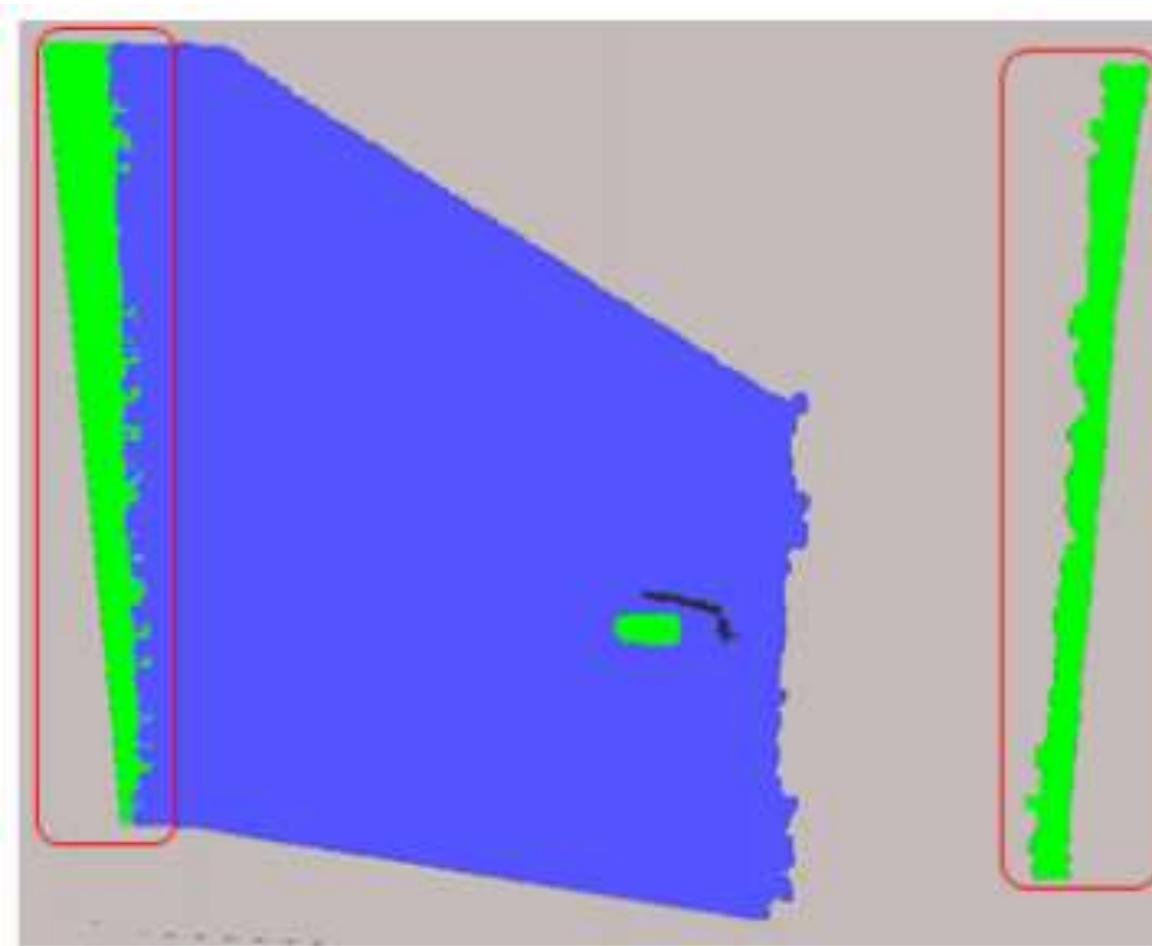
My proposed pipeline



Planar model extraction

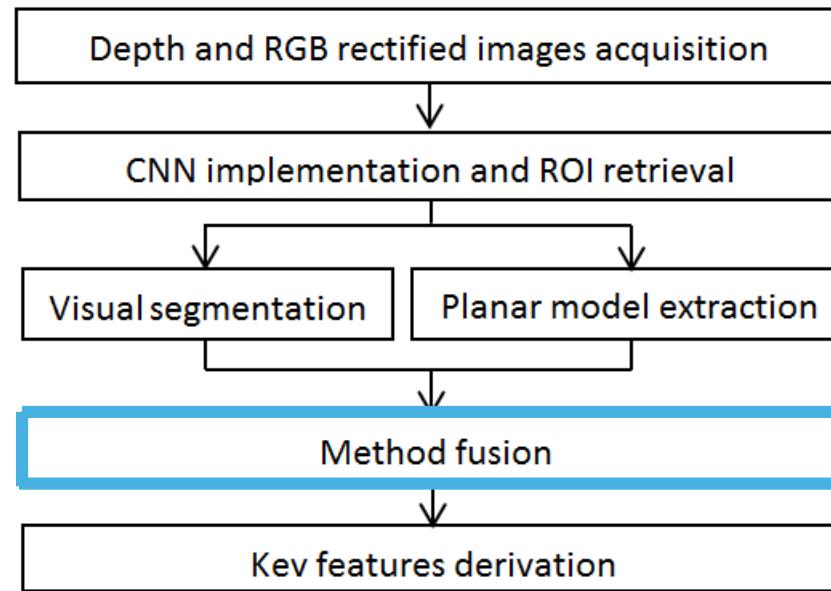
1.
 - Masking of ROI from CNN onto Disparity Image
 - Environment Point cloud generation
2.
 - Filtration of points outside reasonable height
3.
 - Planar segmentation (3-4 cm resolution)
 - Outliers represent handle's point cloud

Planar model extraction



False positives:
- Walls

My proposed pipeline

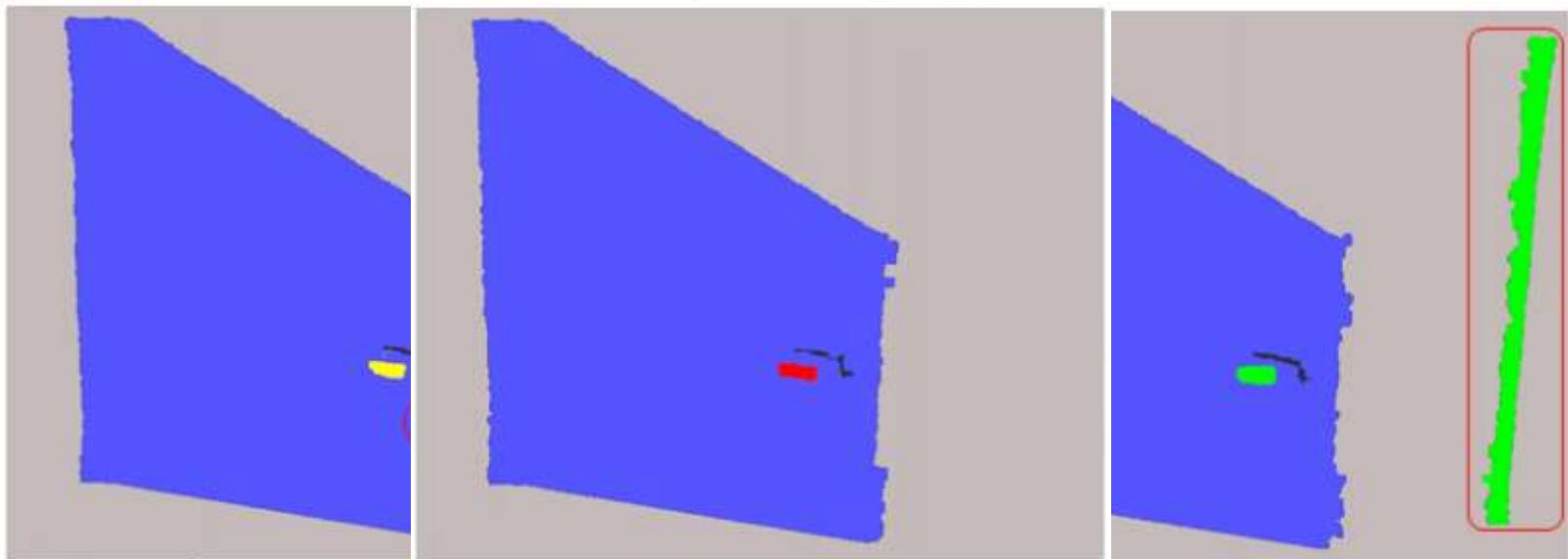


Method fusion

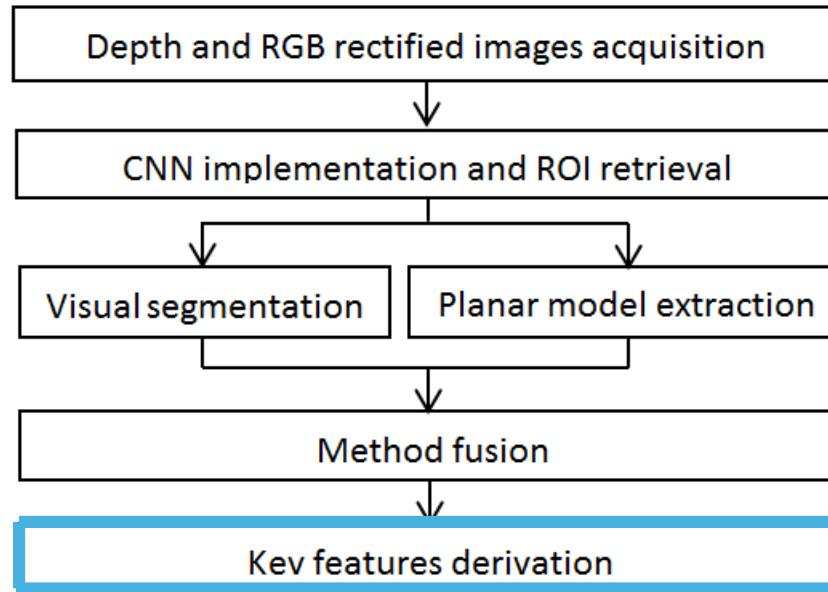
Visual segmentation



Point cloud processing



My proposed pipeline



Key features derivation

Orthogonal vector
to the door plane

- Know which direction to push/pull.
- Derived from the planar segmentation coefficients.

Turning direction
of handle

- Obtained by knowing to which side of the door is the handle closer to.

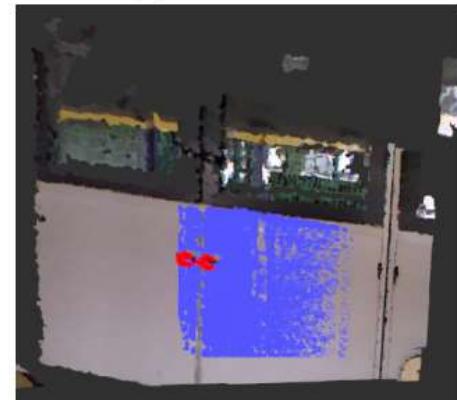
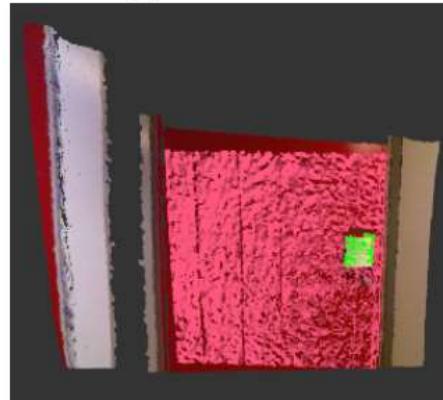
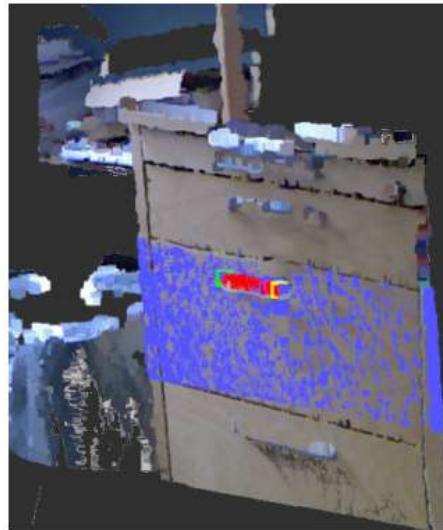
Turning point

- Closest point to door side depending on handles turning direction.

Grasping points

- Inferred from the point cloud itself.
- The centroid is highly recommended

Door perception examples



Handle grasping examples



Experimental results

Range (m)	Angle (°)	HANDLE DETECTION											
		CLOSED DOOR (0°)				SEMI-OPEN DOOR (35°)				OPEN DOOR (70°)			
		Good (%)	Unknown(%)	Bad (%)	Attempts	Good (%)	Unknown (%)	Bad (%)	Attempts	Good (%)	Unknown (%)	Bad (%)	Attempts
0.5 m	-30°	70.6	23.5	5.9	17	66.7	28.6	4.8	21	* ¹	* ¹	* ¹	-
	0°	90.5	9.5	0.0	21	100.0	0.0	0.0	16	100.0	0.0	0.0	16
	30°	81.0	19.0	0.0	21	69.2	30.8	0.0	26	68.2	31.8	0.0	22
1 m	-30°	88.9	11.1	0.0	18	93.8	6.2	0.0	16	* ¹	* ¹	* ¹	-
	0°	57.9	36.8	5.3	19	87.5	12.5	0.0	16	61.9	23.8	14.3	21
	30°	55.6	44.4	0.0	18	84.2	10.5	5.3	19	80.0	20.0	0.0	20
1.5 m	-30°	18.8	81.2	0.0	16	18.8	62.4	18.8	16	* ¹	* ¹	* ¹	-
	0°	33.3	60.0	6.7	15	41.2	47.1	11.7	17	14.3	71.4	14.3	14
	30°	11.8	76.4	11.8	17	30.0	60.0	10.0	20	22.2	66.7	11.1	18

*¹ In this position, the robot would be looking almost perpendicular to the door plane so the handle extraction was not possible.

Calibration errors

- Rectification inconsistencies.
- Main reason for *unknowns* after 1m mark.

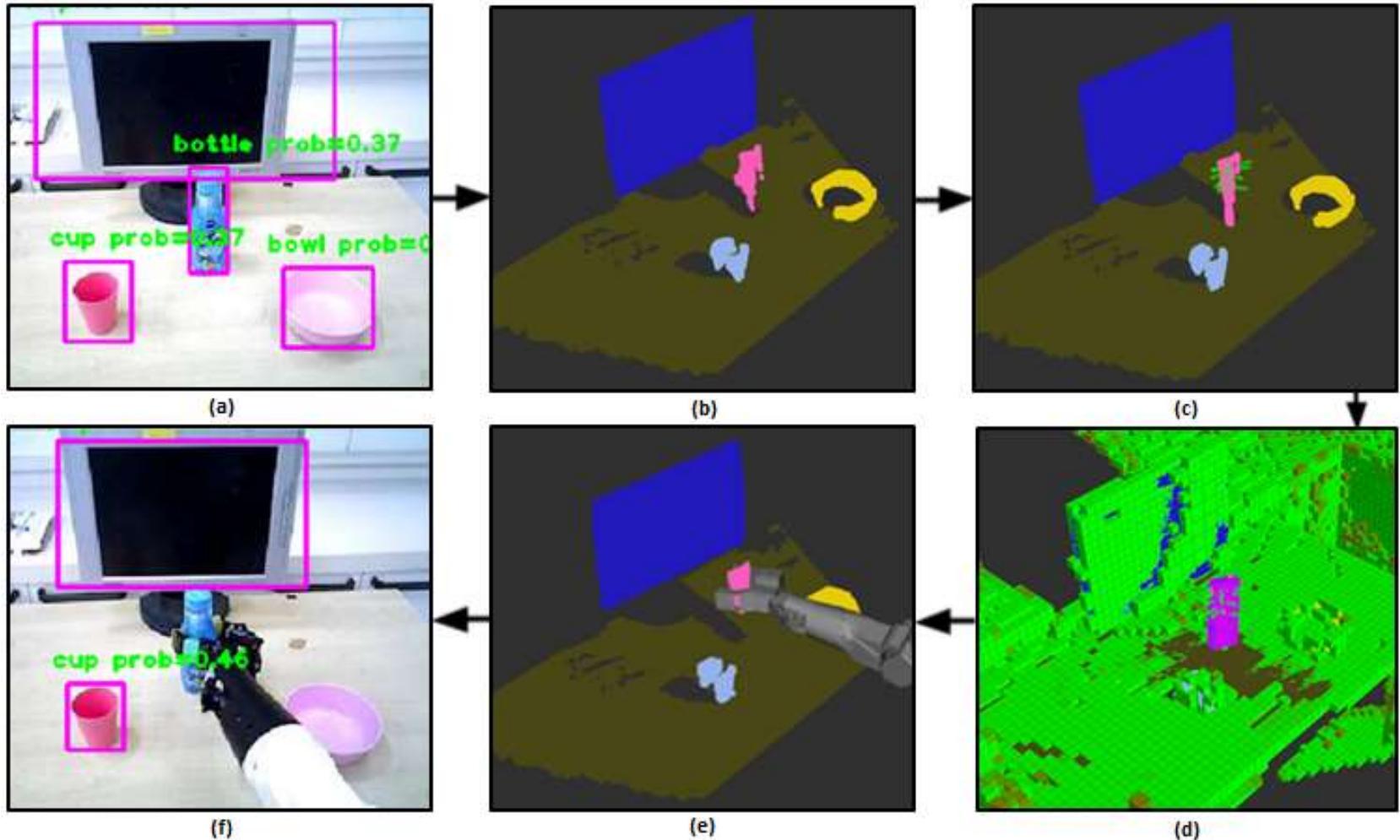
Tightness of ROI

- Walls appear as part of handles point cloud.
- Effects become present after 1.5 meters.

Depth resolution

- After 1.5 m mark, the depth sensor starts failing to differentiate door and handle.

Future work



Thank you