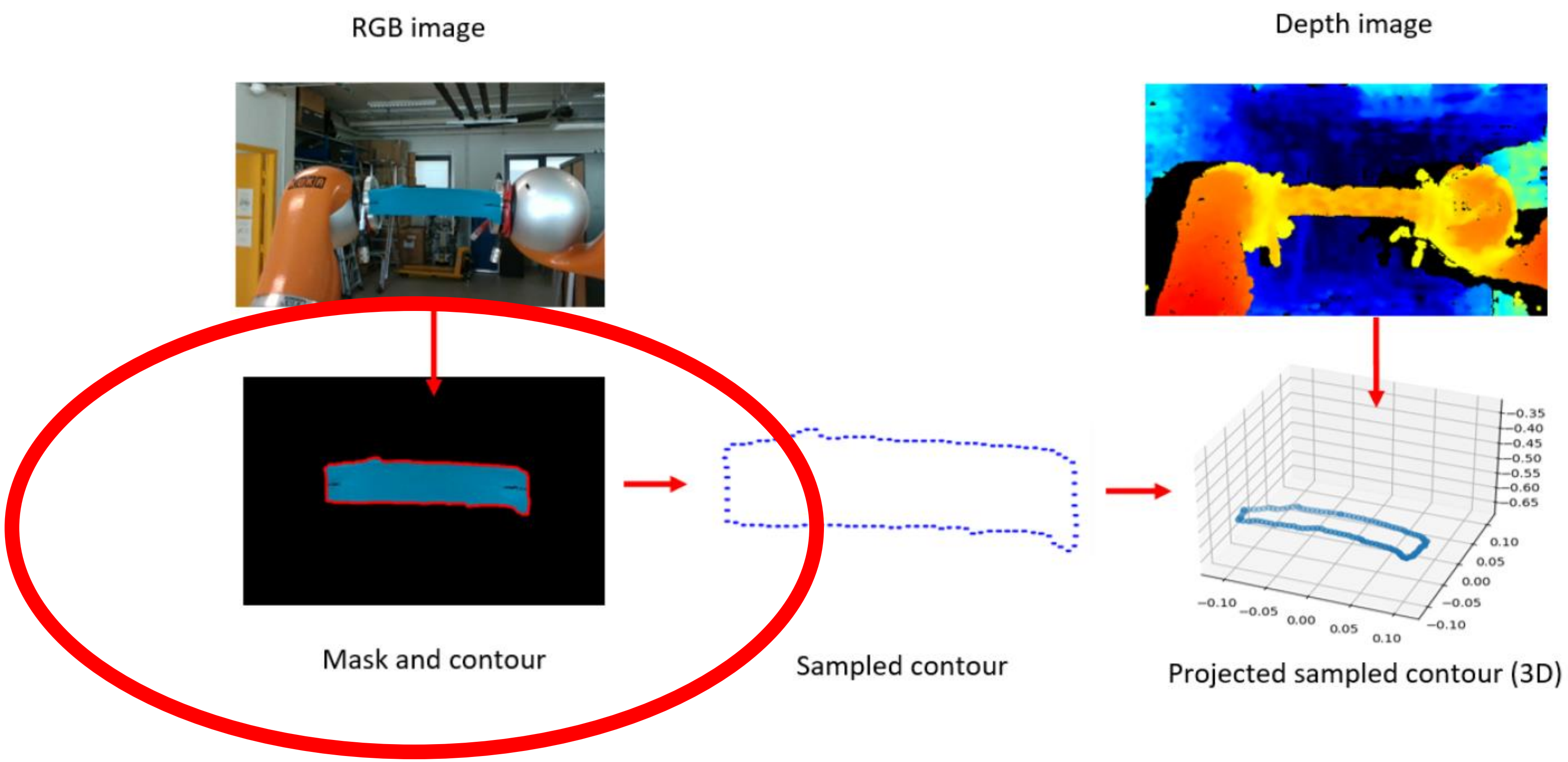
State of the Art

Ben Bartlett

Problem: 2 parts



1 Segment Image on Object

Problem: 2 parts

2 mapping movement between frames

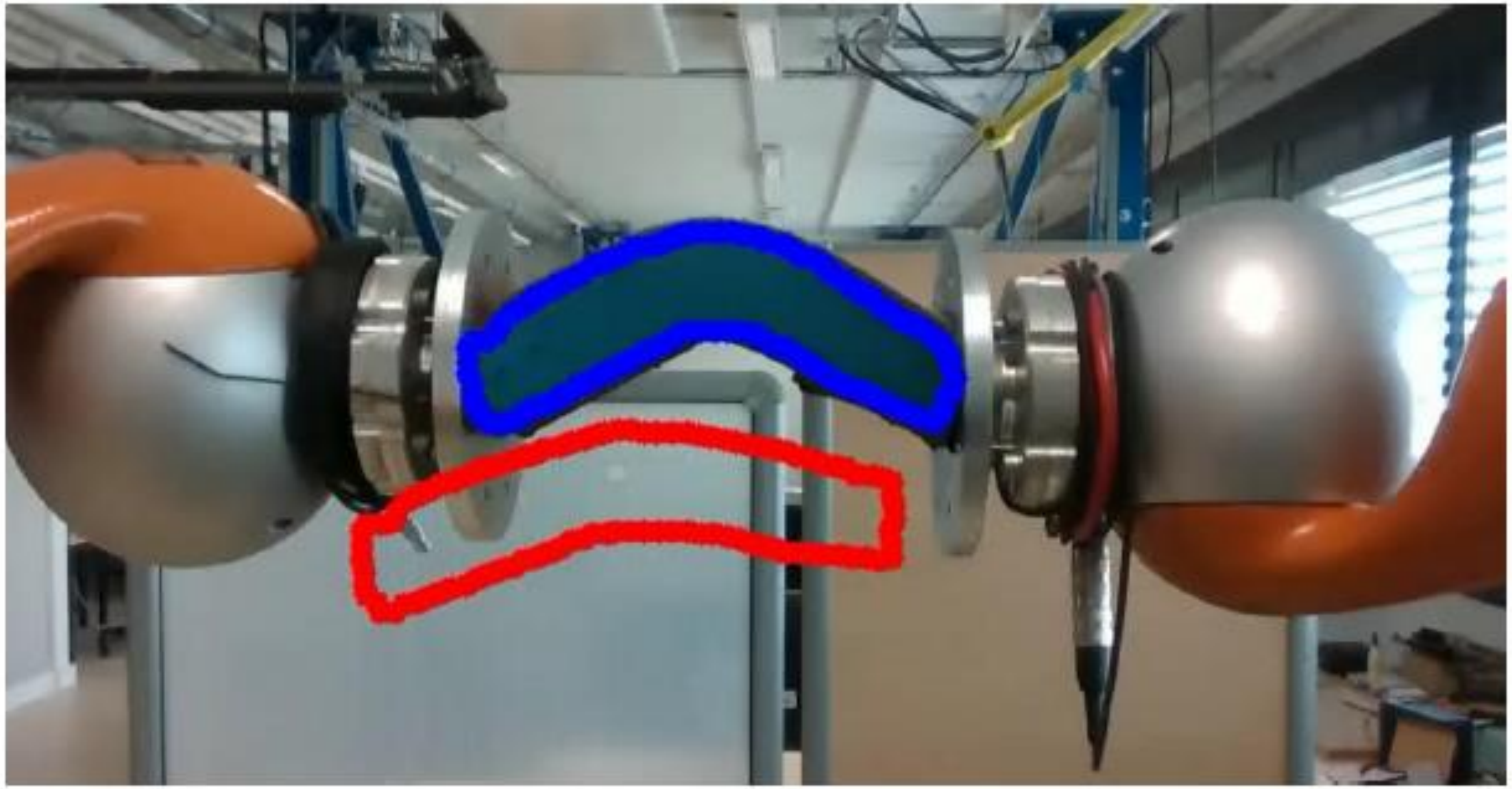


Image Segmentation

Overview of techniques (typically RGB)

- Classical Computer Vision Approaches
- AI Based Techniques

Image Segmentation

- Classical Computer Vision Approaches

1. Simple RGB -> Greyscale, thresholding
    1. Global (Single threshold for the entire image)
    2. Adaptive(Different regions calculated, good for varied lighting, Otsu method)

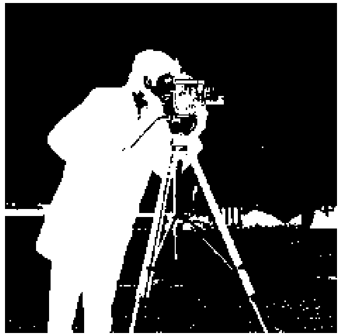
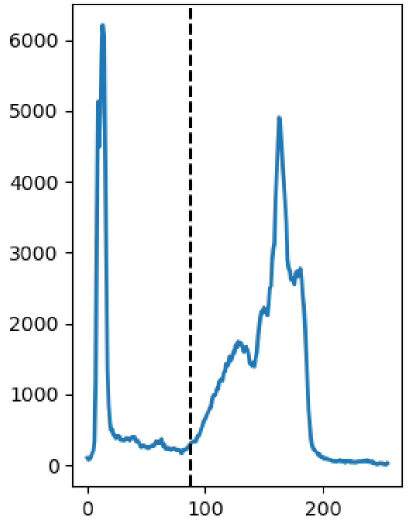




Image Segmentation

- Classical Computer Vision Approaches

2. Clustering, grouping based on groups of pixels, colours, intensity or texture
   1. K means (minimize variance within clusters, can be good when we have dominant colours)
   2. Mean shift Clustering (non-parametric, shifts points to average of the neighbourhood)
   3. Fuzzy C-means Clustering, (can belong to multiple clusters with varying degrees of membership

Image Segmentation

- Classical Computer Vision Approaches

3. Edge Detection
    1. Canny (gradient based, noise reduction, non max suppression and hystersis)
    2. Sobel (Convolution, horizontal and vertical directions)
    3. LoG (Guassian smoothing with Laplacian edge detection)

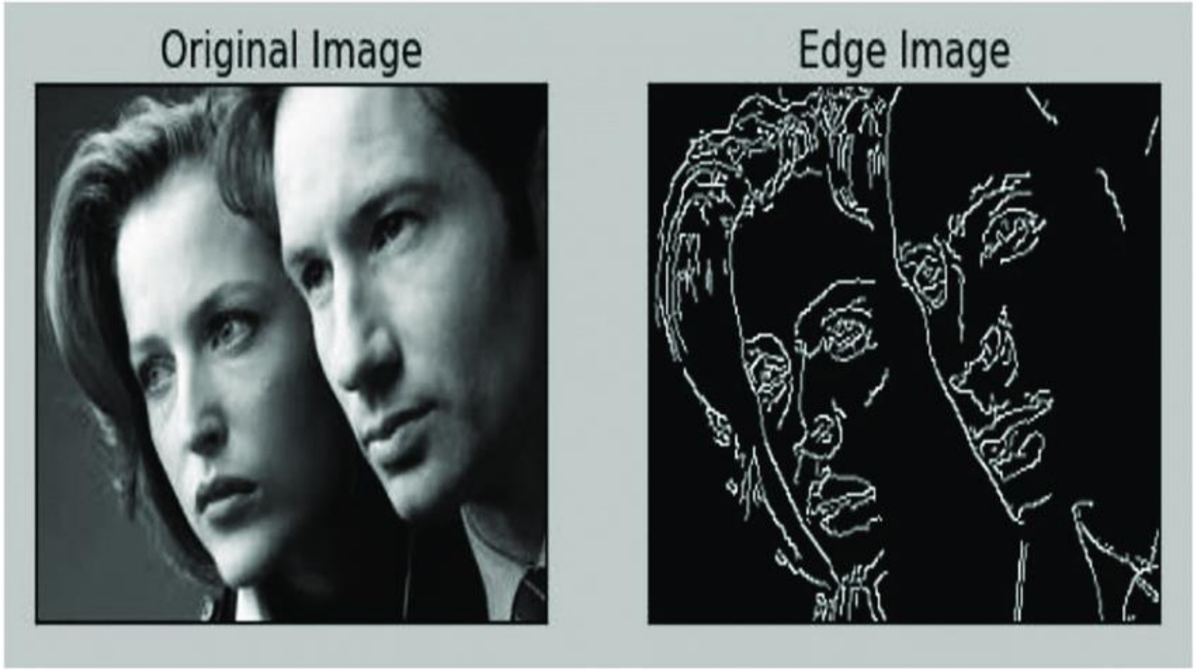


Image Segmentation

- Classical Computer Vision Approaches

4. Region Based partition to regions based on predefined criteria
    1. Region Growing, (add pixels with similar properties)
    2. Watershed (greyscale, treated as topographic image)
    3. Region splitting and merging (divide up image and merge)

Image Segmentation

- Classical Computer Vision Approaches

5. Graph Based Segmentation, represent image as a graph, pixels are nodes and edges represent the relationship between them

    1. Normalized cuts (minimizing cut cost, balancing of segments)
    2. Minimum spanning tree (find tree, cut edges to form segments)

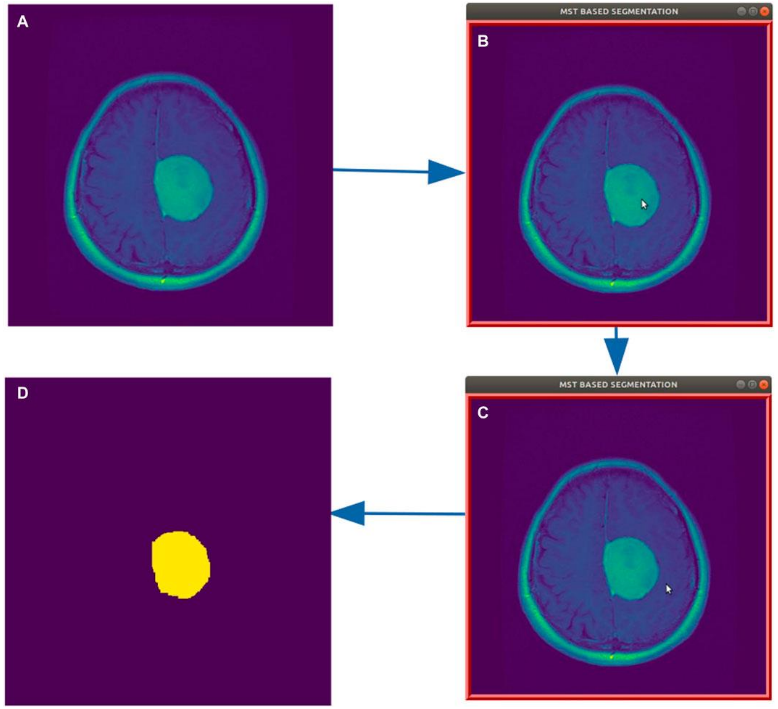
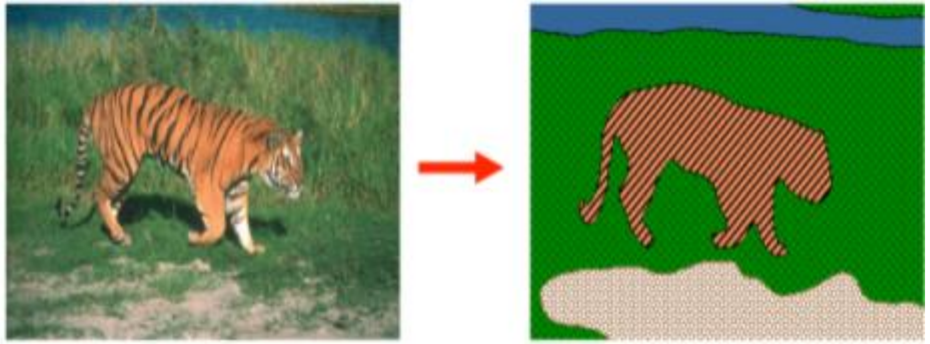


Image Segmentation



- Classical Computer Vision Approaches

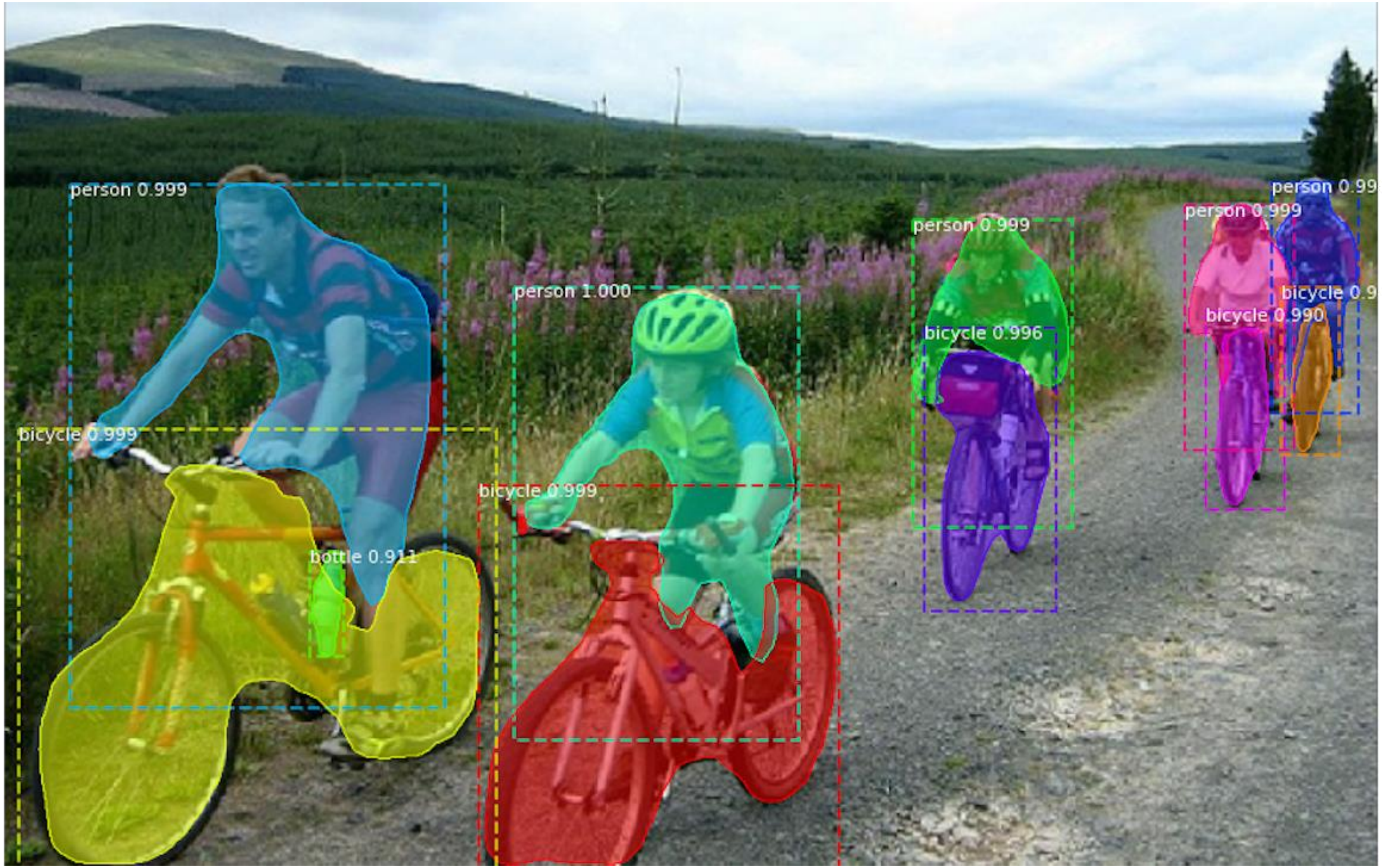
6. Model Based Segmentation, prior knowledge about shape, appearance or other properties
    1. Active Contours (evolves a contour to fit boundaries, minimizing energy function)
    2. Level Sets (Contour represented as zero level of higher dimensional function)

Image Segmentation

- A.I Based

1. Deep Learning Based Segmentation
    1. FCN (Replace fully connected layers with convolutional layers to produce pixel-wise segmentation maps.)
    2. U-net (Replace fully connected layers with convolutional layers to produce pixel-wise segmentation maps.)
    3. Mask R-CNN (Extends Faster R-CNN to generate high-quality segmentation masks along with object detection.)

Image Segmentation



RGB-D Specifically

- Application of classic methods with Depth weighting
- DCNN (depth aware Convolutional Neural Networks, both depth and RGB fed to model)
- Multi Modal Networks (i.e separate processing and later fusion)
- 3D CNN (application of voxelization such as done with LiDAR data)

Literature

- Detection of Cables
  - Transmission line detection in aerial images: An instance segmentation approach based on multitask neural networks

  Typical to see algorithms aimed at the detection of overhead powerlines in aerial surveys

  VGG16 backend followed by Binary segmentation and pixel embedding

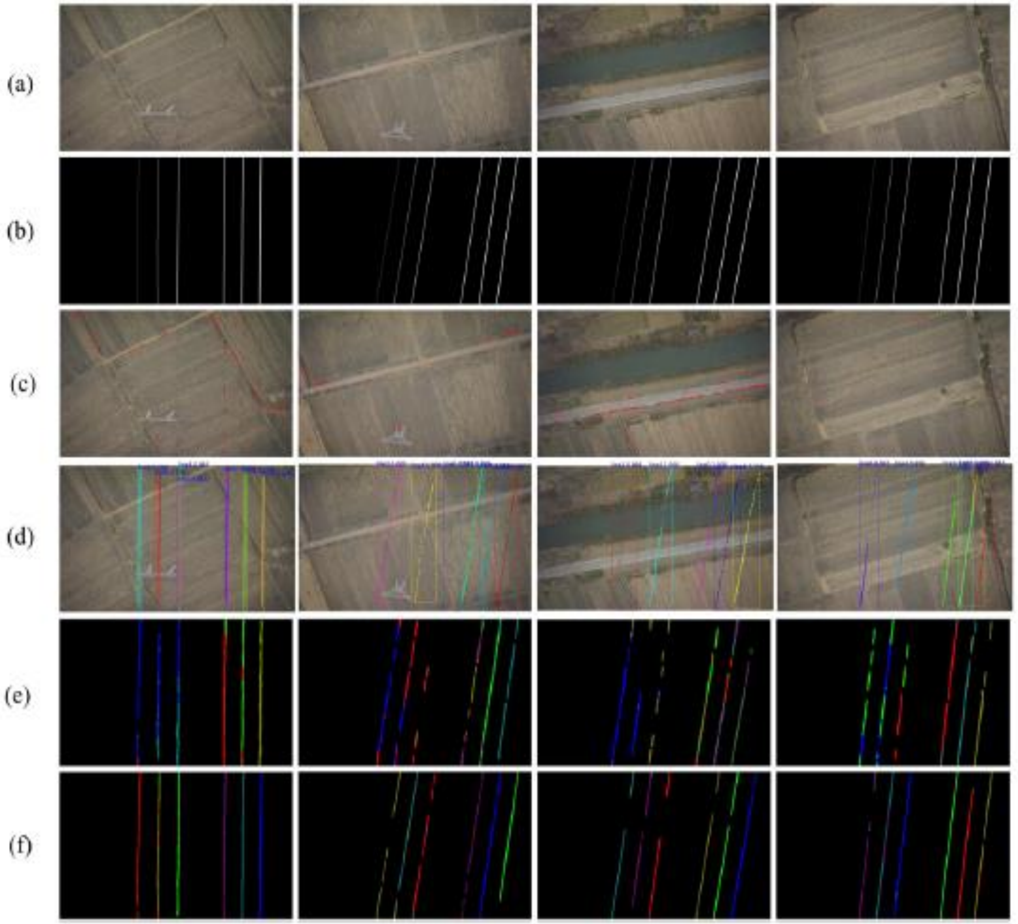


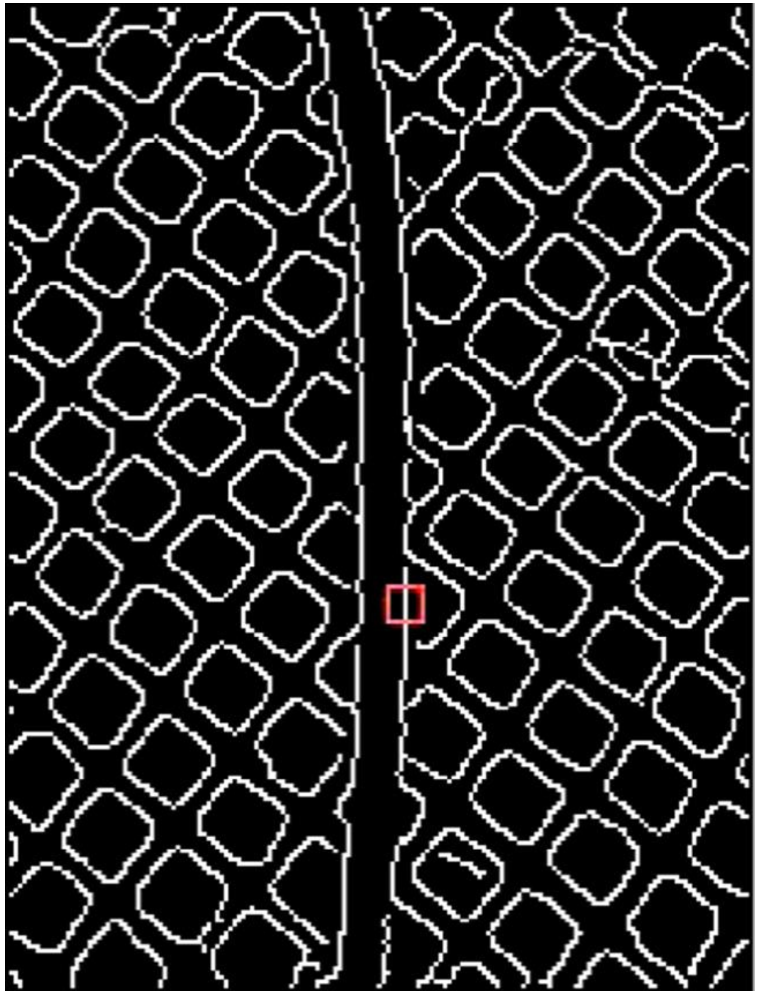
Fig. 6. Examples of the experimental results for the proposed algorithm and comparison methods in dataset II. (a) Original images. (b) Ground-truth images. (c)–(f) are the detection results of the LSD-based method, Mask R-CNN method, VGG-based method, and the proposed method, respectively. The detection results in (c) are denoted by red segments. The detection results in (d)–(f) are illustrated by segmented pixels, where colors indicate the different cable instances to which they belong. Note that the bounding boxes of detected cables are also represented in (d).

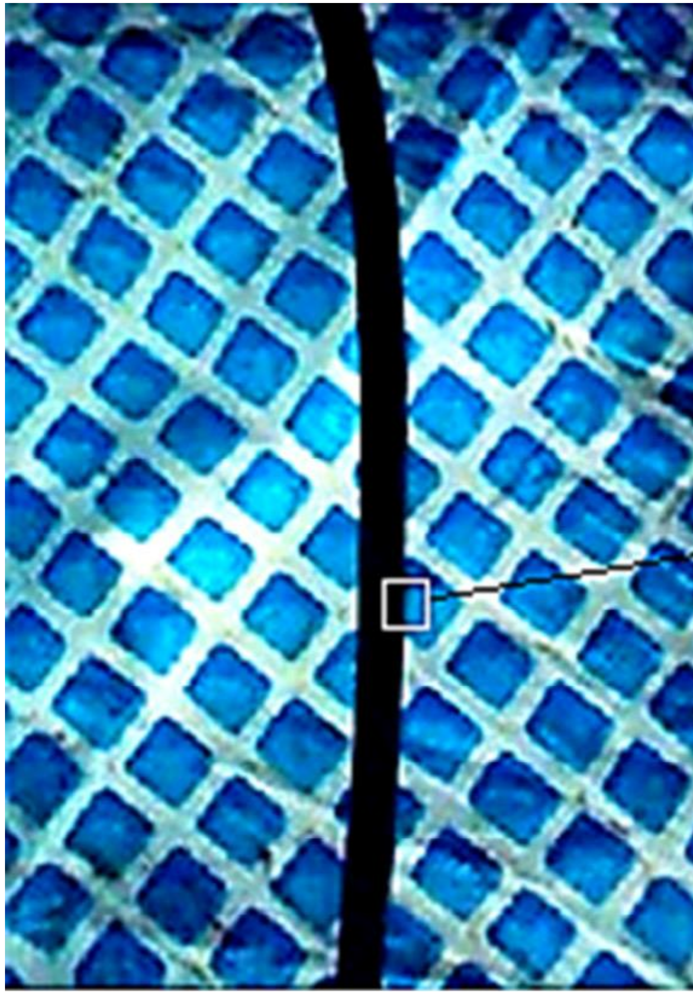
Literature

- Detection of Cables
  - Underwater cable detection in the images using edge classification based on texture information

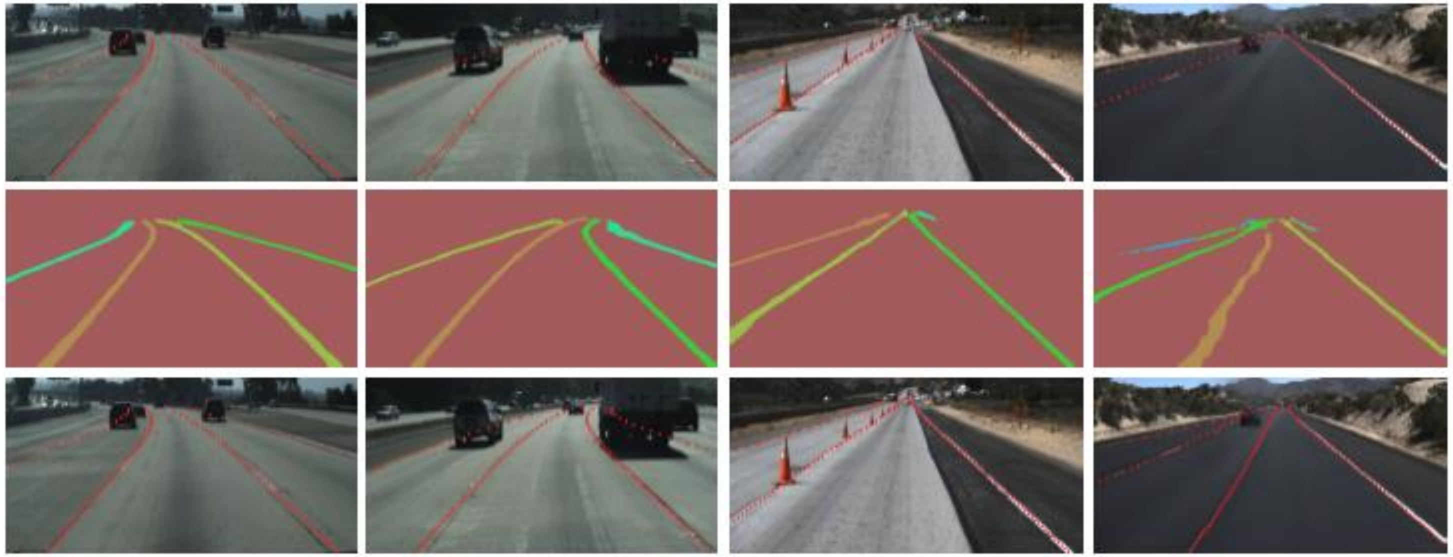
Typical to see algorithms aimed at the detection of cables underwater

Detection based on background texture





Literature



- Detection of Lanes
  - Towards End-to-End Lane Detection: an Instance Segmentation Approach

From the exploration of Overhead line detection, another well studied field is lane detection for autonomous vehicles

Another Multi Modal Model, 1 section binary mask output and 1 section outputting instance segmentation

Literature

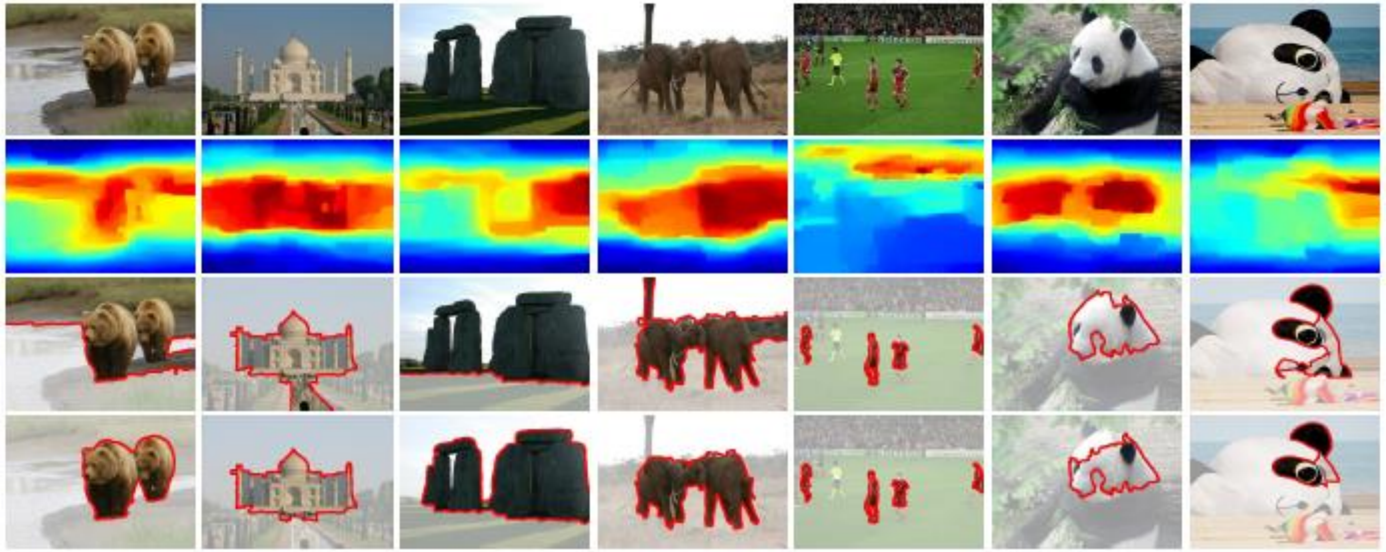


Figure 6. Our RGBD co-segmentation results on the iCoseg dataset. Top to bottom: one of the input images, its estimated depth map, and the results of our RGB and our RGBD co-segmentation. (Best viewed in color.)

- RGBD Segmentation
  - Object-based RGBD image co-segmentation with mutex constraint

Region Based Segmentation and Histogram matching, Particularly focus on foreground elimination

Literature

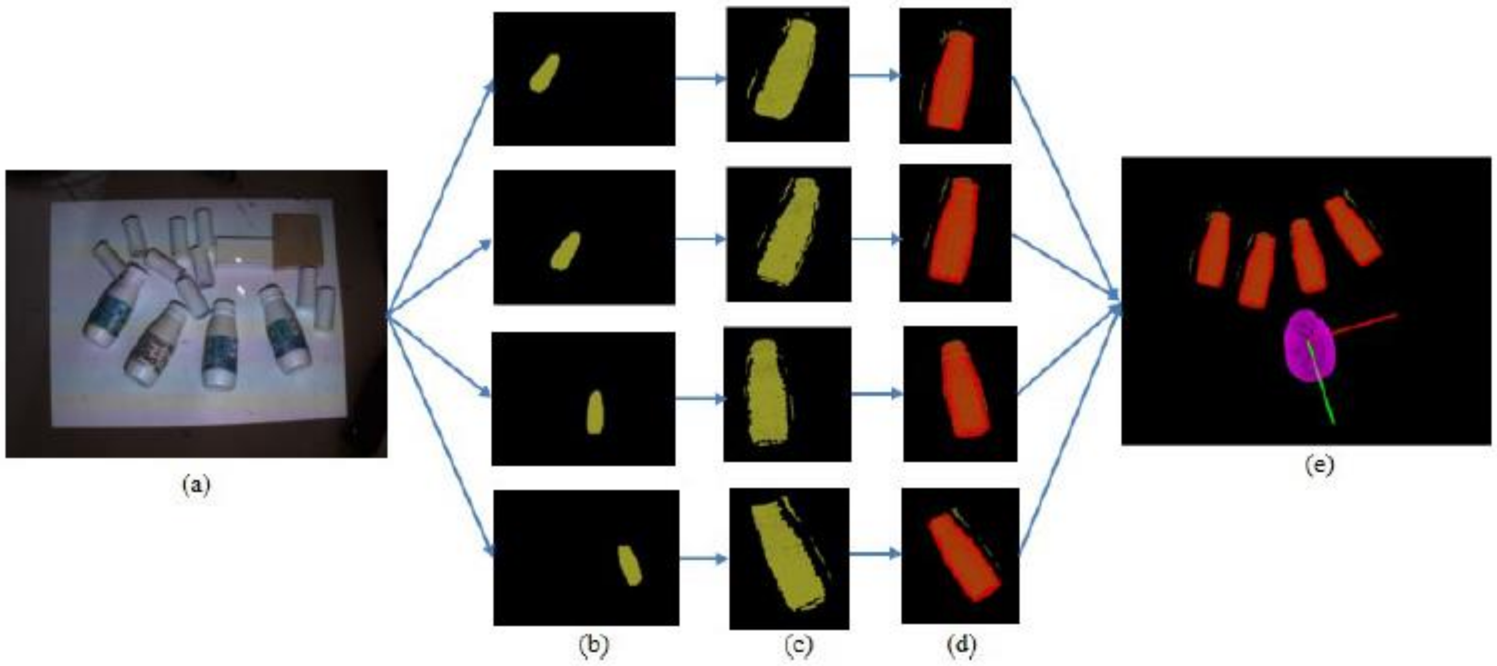


Fig. 4. The overview of our approach for 6D pose estimation: (a) one input RGB image from the RGB-D camera, (b) object detection and segmentation in multi-layer based on Mask R-CNN, (c) 3D detected objects in the multiple layers with the 3D information from the structured light system. (d) estimating the 6D object pose by the ICP algorithm in multi-layer, and (e) output results with the object pose and the 3D CAD model.

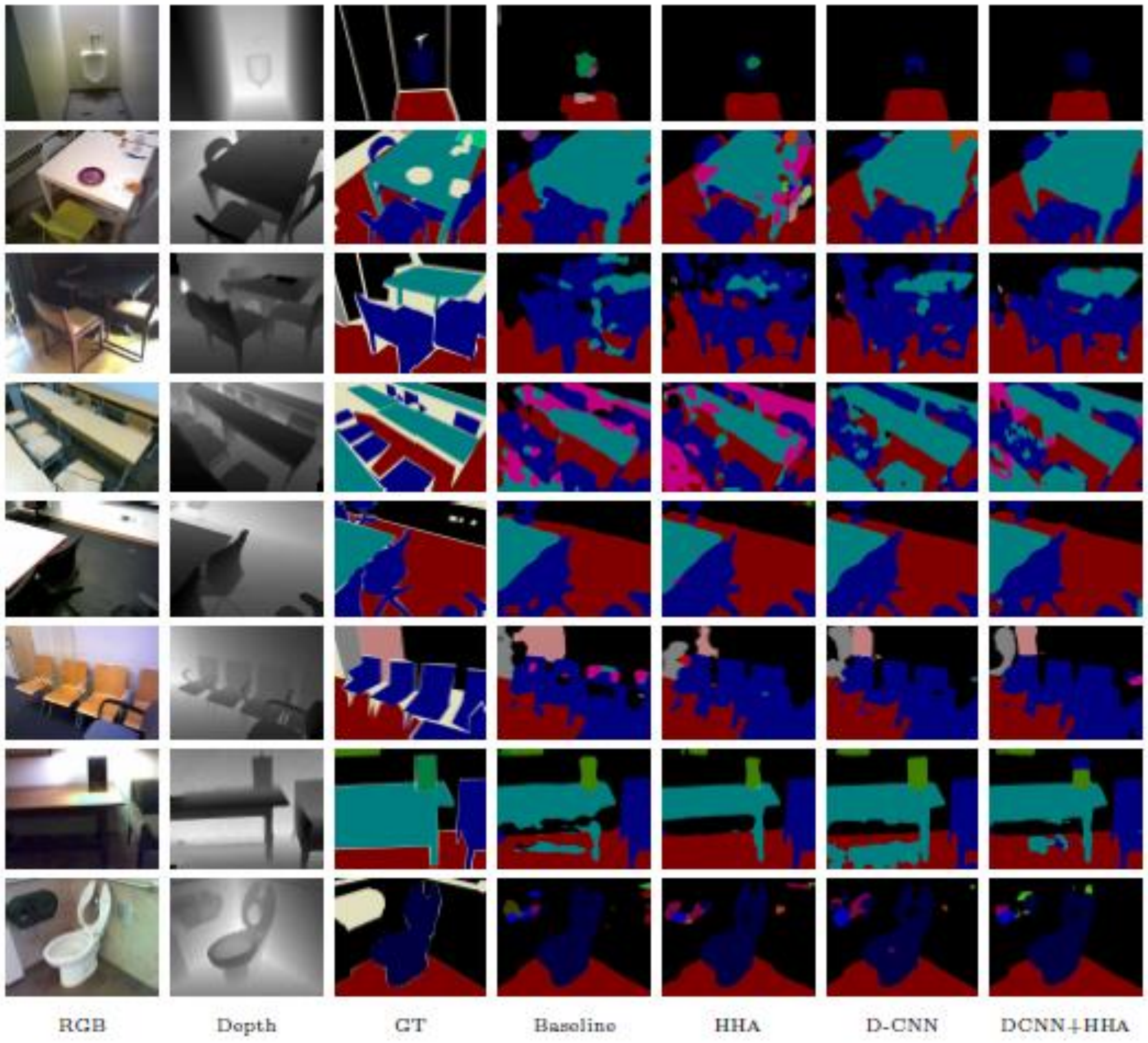
- RGBD Segmentation
  - 3D Object Detection and 6D Pose Estimation Using RGB-D Images and Mask R-CNN

Mask R-CNN detection with 6D pose estimation, optimization with CAD ICP

Literature

- RGBD Segmentation
  - Depth-aware CNN for RGB-D Segmentation

Depth aware convolution , depth similarity term for average pooling



Literature

- RGBD Segmentation
  - Graph-Based Segmentation for RGB-D Data Using 3-D Geometry Enhanced Superpixels

Classical approach

K means clustering, “superpixel” generation utilizing depth info, graph-based energy optimization.

Fig. 5. Segmentation results for *Art*, *Books*, *Moebius*, *Laundry*, *Baby*, and *Plastic* from top to bottom. Each row from left to right presents the ground-truth segmentation, and the segmentation results produced by EG [59], MGD [60], CDHS [48], KGC [61], Ours_LM [30], and our method, respectively.

Literature

Figure 7. Sequence 1, (A) : RGB image, (B): Results from [24], (C): Results from [7] (N/A is used when the frame is beyond the length it can run), (D): Our Results

- RGBD Segmentation
  - Efficient Hierarchical Graph-Based Segmentation of RGBD Videos

Classical approach

Hierarchical tree combines depth and RGB related segments with 3d optical flow across frames to aid

Literature

- Potential Optimizations
  - Image Segmentation Based on RGBD Data
  - Fast Depth Edge Detection and Edge Based RGB-D SLAM

Frame to frame prioritise areas where edges were detected previously

Utilization of expected distance to create mask from depth image

Fig. 2: Occluding depth edges (red) and the associated image patches which were flagged for edge detection (green).

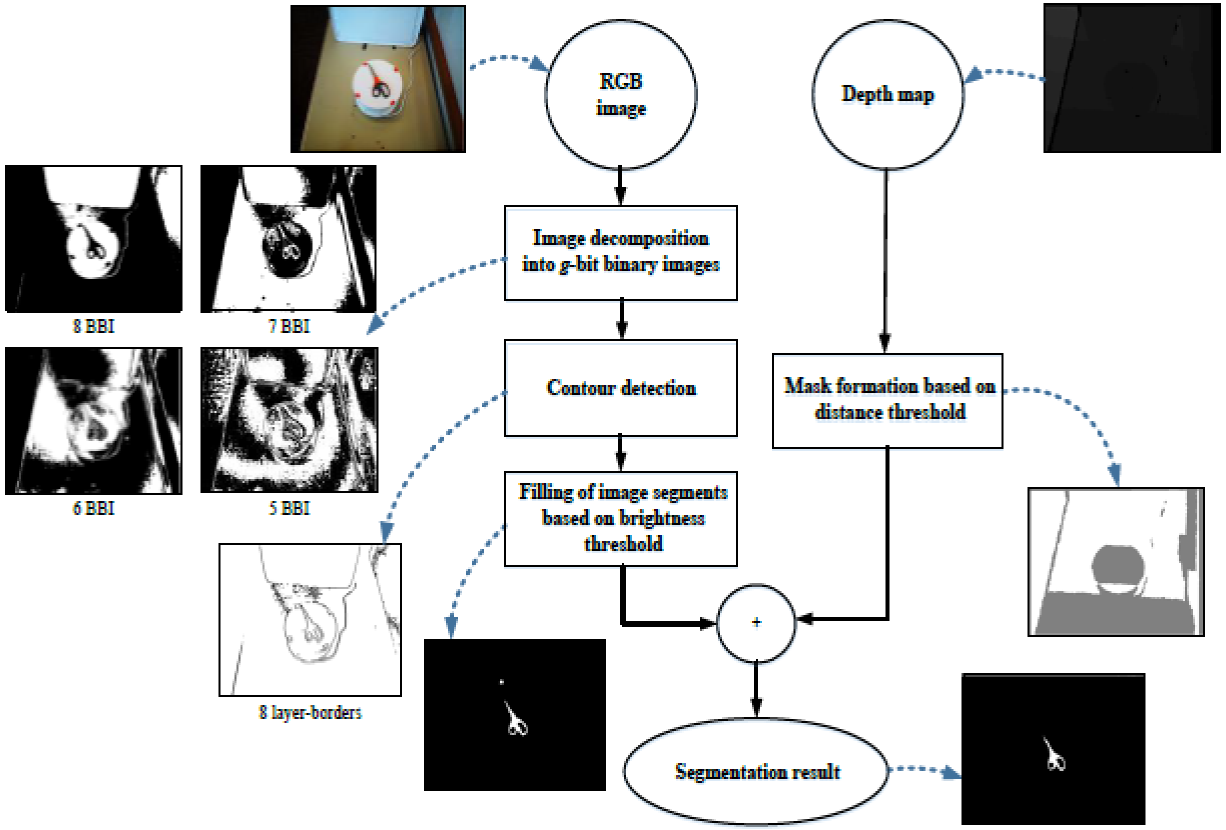


Fig. 2. Flowchart of RGBD image segmentation algorithm.

RGB-D Benchmark Datasets

- NYUv2
- SUN-RGBD
- SceneNet RGB-D
- Stanford 2D-3D-S
- Matterport3D
- Cityscapes
- ScanNet

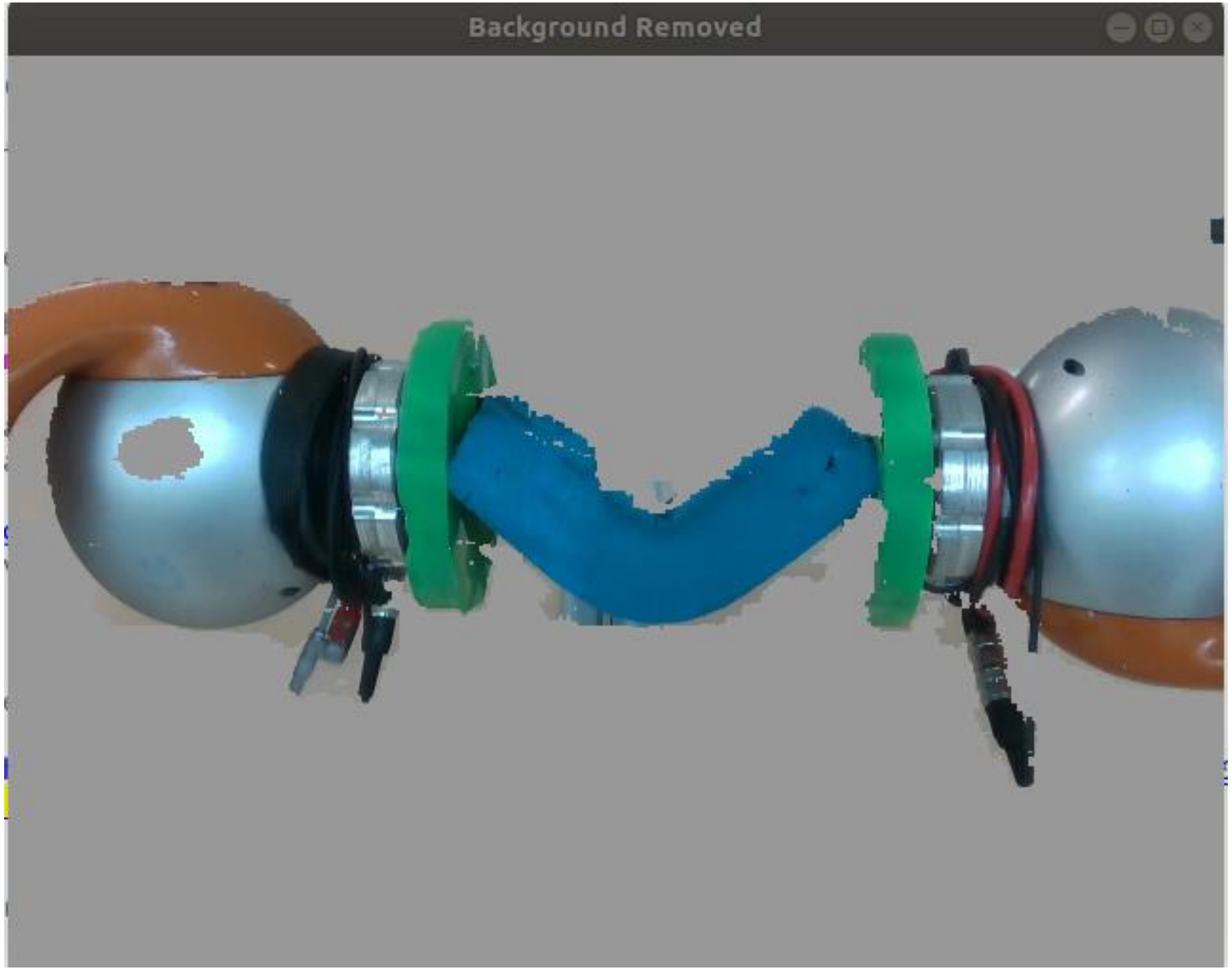
Model Benchmarks on different sets

- https://github.com/Yangzhangcst/RGBD-semantic-segmentation

RGB-D Benchmark Datasets

NYUDv2

Method	PixAcc	mAcc	mIoU	f.w.IOU	Input	Ref. from	Published	Year
POR	59.1	28.4	29.1		RGBD		CVPR	2013
RGBD R-CNN	60.3	35.1	31.3	47(in LSD-GF)	RGBD		ECCV	2014
DeconvNet	69.9	56.4	42.7	56	RGB	**LSD-GF**	ICCV	2015
DeepLab	68.7	46.9	36.8	52.5	RGBD	**STD2P**	ICLR	2015
CRF-RNN	66.3	48.9	35.4	51	RGBD	**STD2P**	ICCV	2015
Multi-Scale CNN	65.6	45.1	34.1	51.4	RGB	**LCSF-Deconv**	ICCV	2015
FCN	65.4	46.1	34	49.5	RGBD	**LCSF-Deconv**	CVPR	2015
Mutex Constraints	63.8	31.5		48.5 (in LSD-GF)	RGBD		ICCV	2015
E2S2	58.1	52.9	31	44.2	RGBD	**STD2P**	ECCV	2016
BI-3000	58.9	39.3	27.7	43	RGBD	**STD2P**	ECCV	2016
BI-1000	57.7	37.8	27.1	41.9	RGBD	**STD2P**	ECCV	2016
LCSF-Deconv		47.3			RGBD		ECCV	2016
LSTM-CF		49.4			RGBD		ECCV	2016
CRF+RF+RFS	73.8				RGBD		PRL	2016
RDFNet-152	76	62.8	50.1		RGBD		ICCV	2017
SCN-ResNet152			49.6		RGBD		ICCV	2017
RDFNet-50	74.8	60.4	47.7		RGBD		ICCV	2017


Background Removed