Annotation-Free and One-Shot Learning for Instance Segmentation of Homogeneous Object Clusters

Zheng Wu†, Ruiheng Chang†, Jiaxu Ma†, Cewu Lu†, Chi-Keung Tang‡

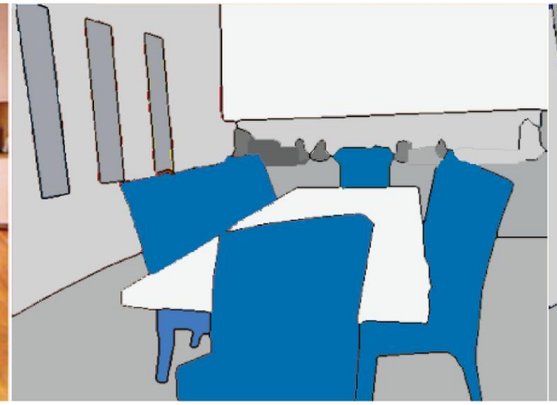
†Shanghai Jiao Tong University ‡HKUST

Email: 14wuzheng@sjtu.edu.cn

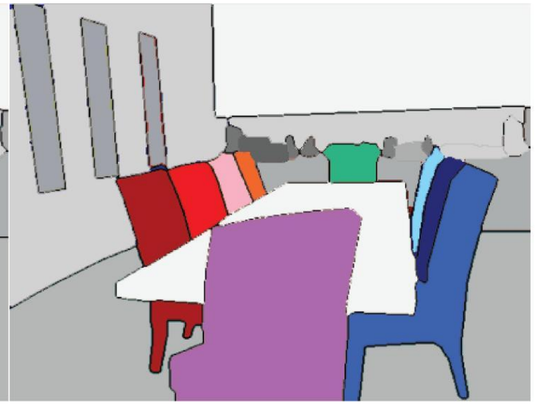
Jul.18th 2018 @ IJCAI

Background

Input Image



Semantic Segmentation



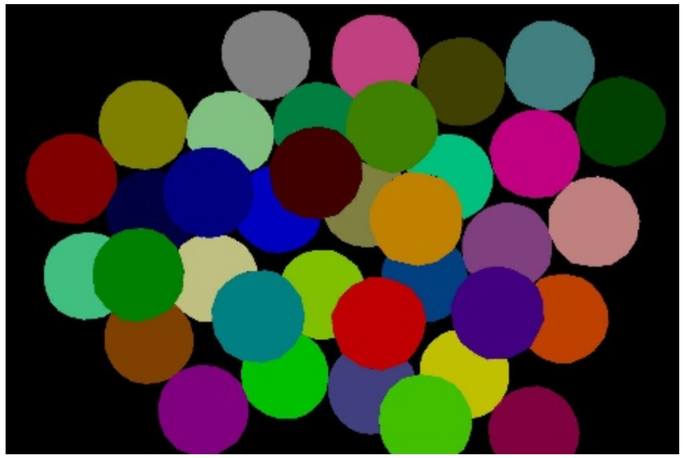
Semantic Instance Segmentation

Background

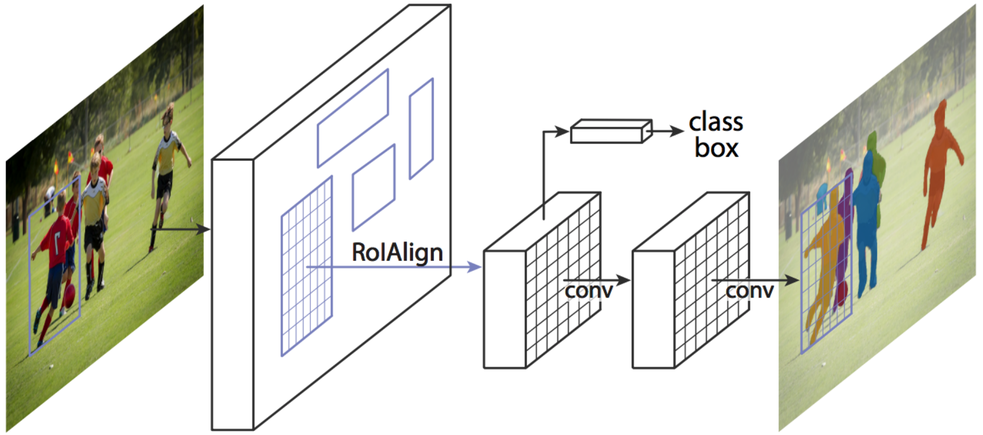
homogeneous object clusters (HOC)

- objects of the same class
- densely distributed, highly occluded

Background

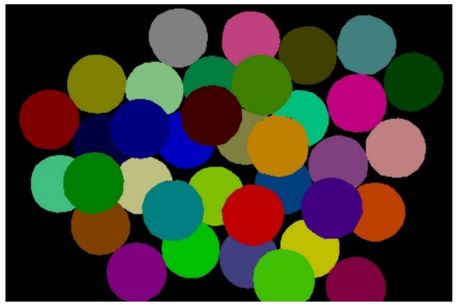


Mask R-CNN

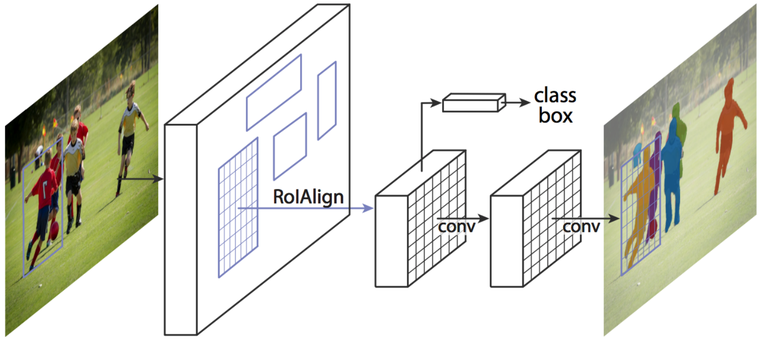


COCO: ~ **900k** instances

Mask R-CNN



=

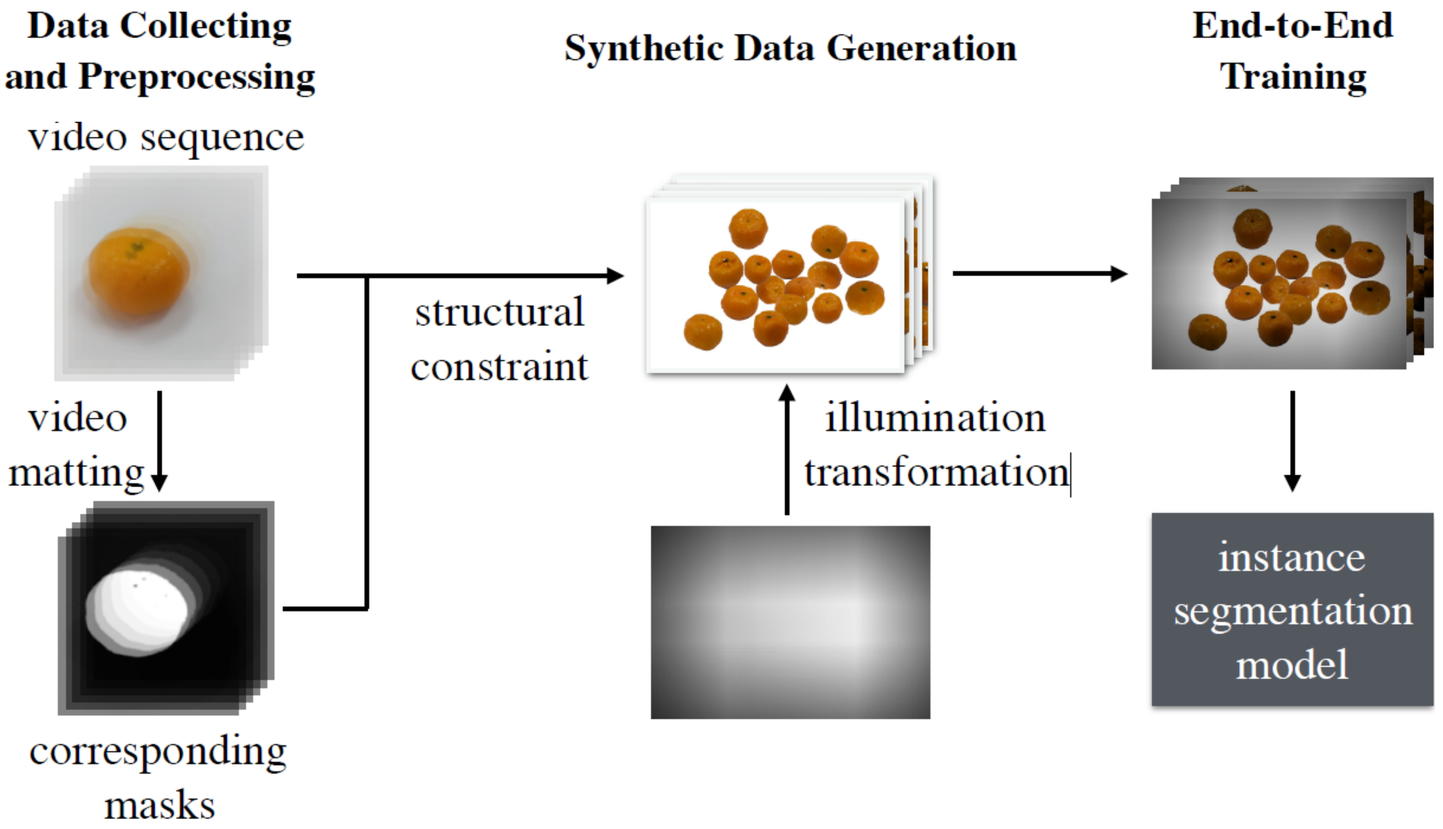


+

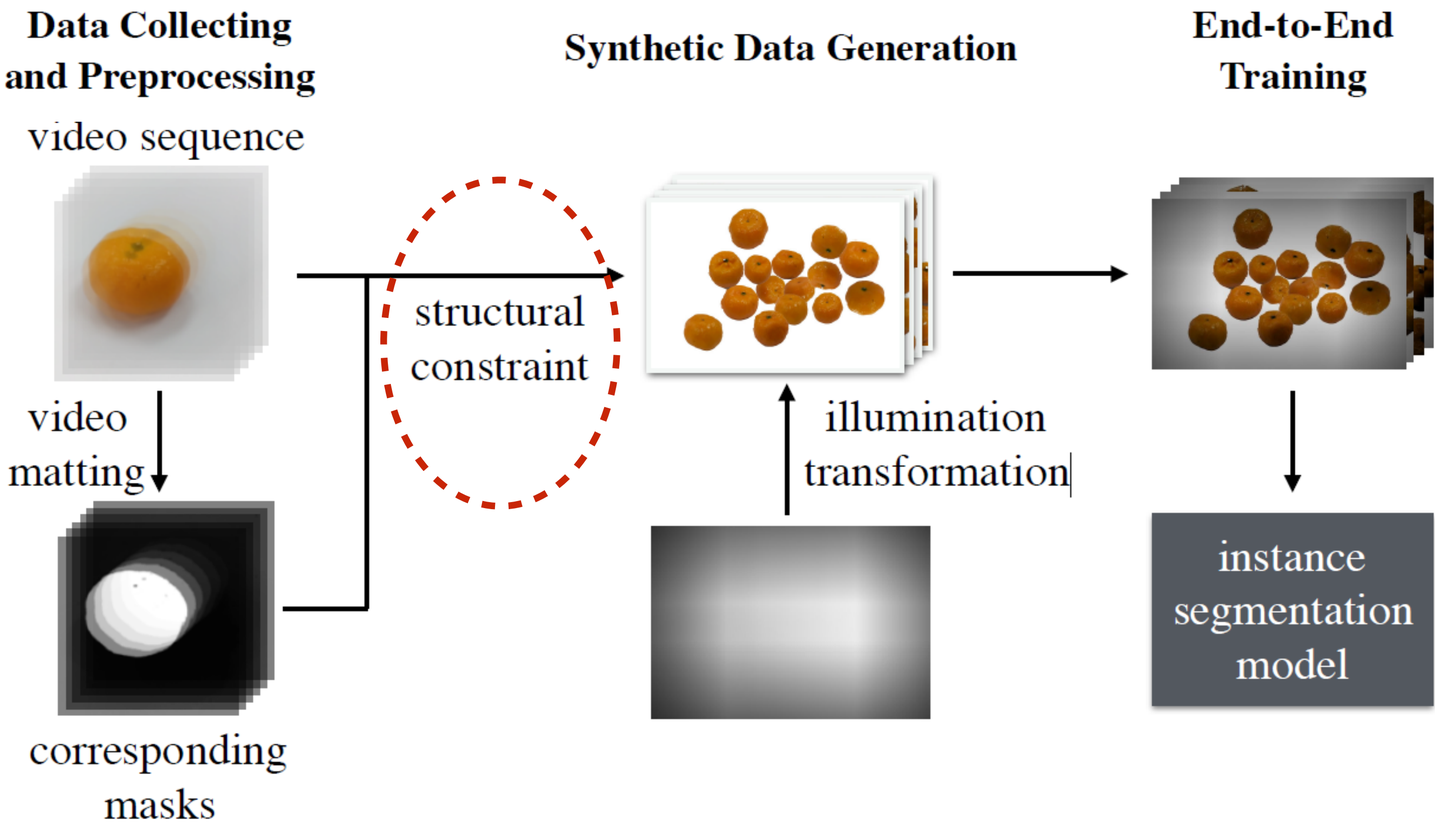
thousands of **annotated** images of pingpong balls

Can we do it cheaply?

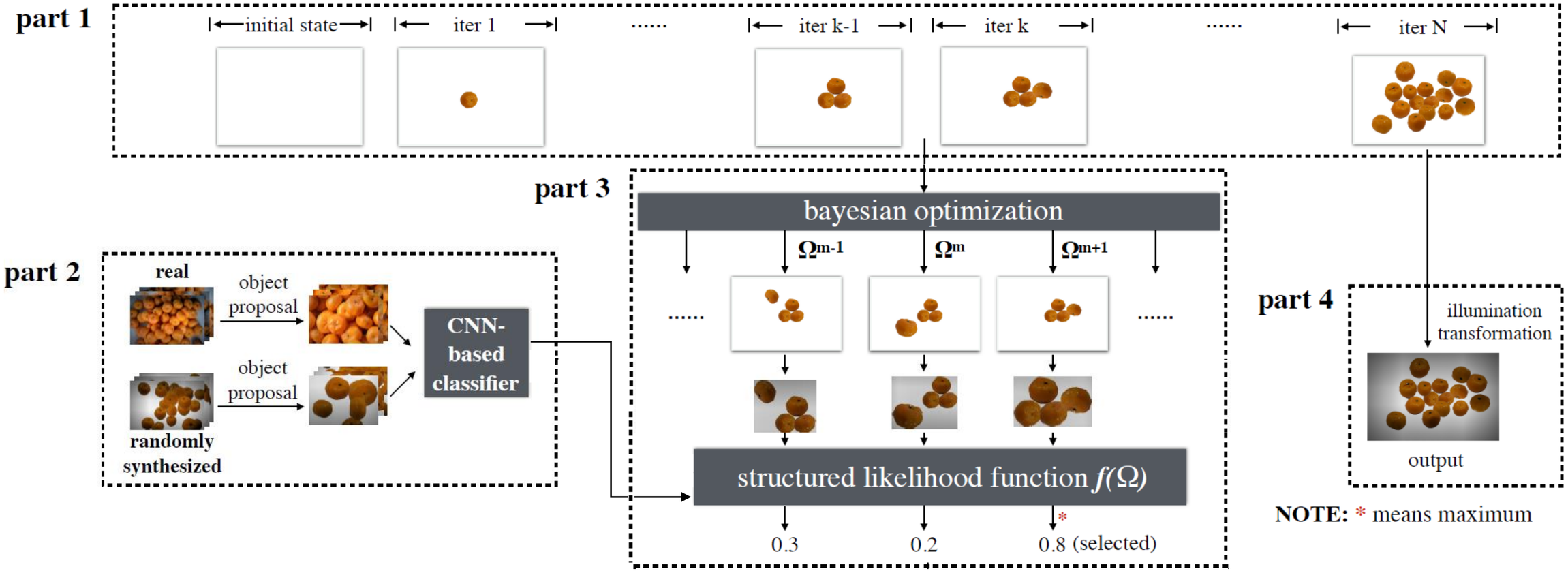
Our Pipeline



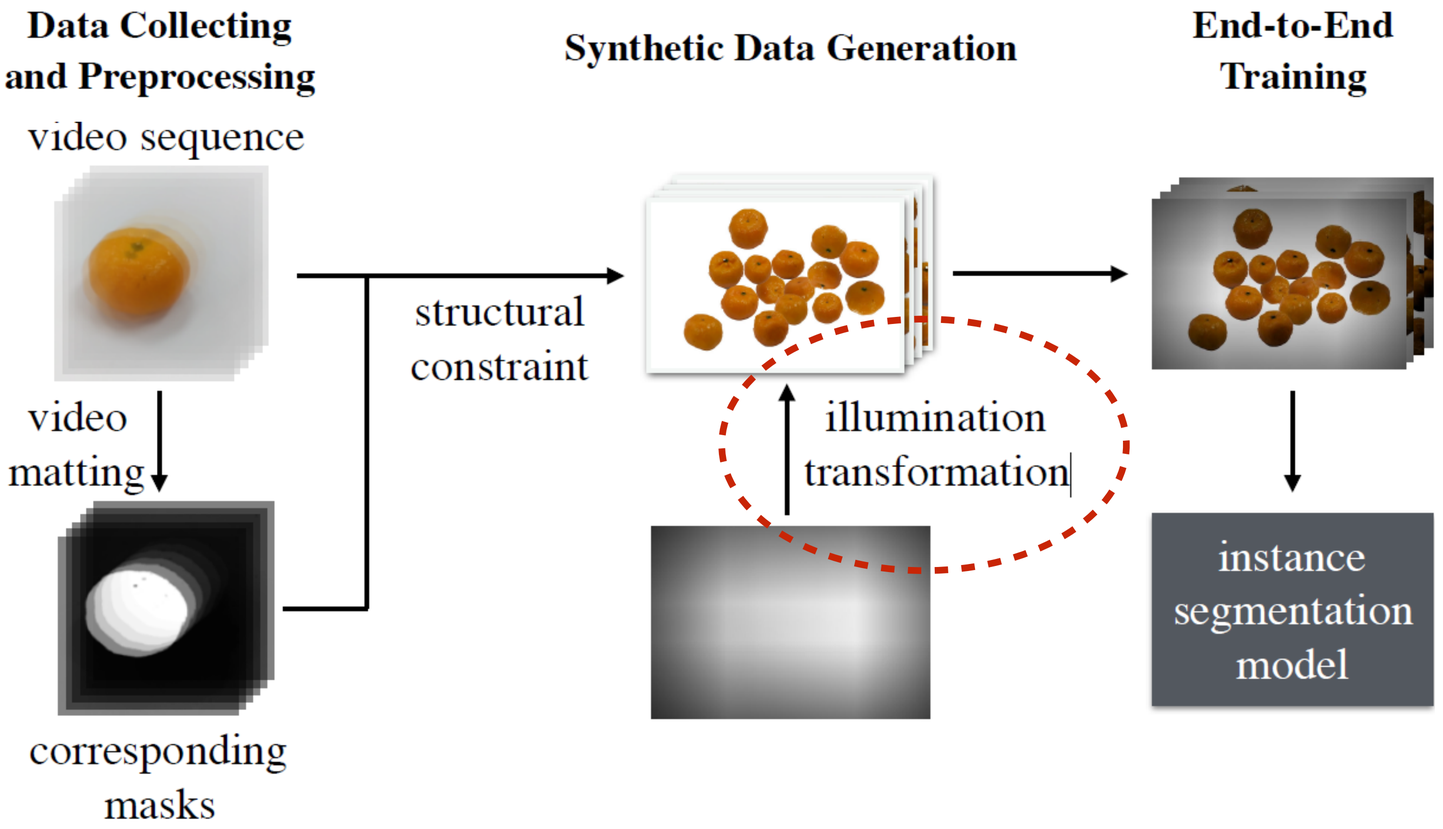
Structural Constraint



Structural Constraint



Illumination Transformation



Illumination Transformation

(a) (b) (c)

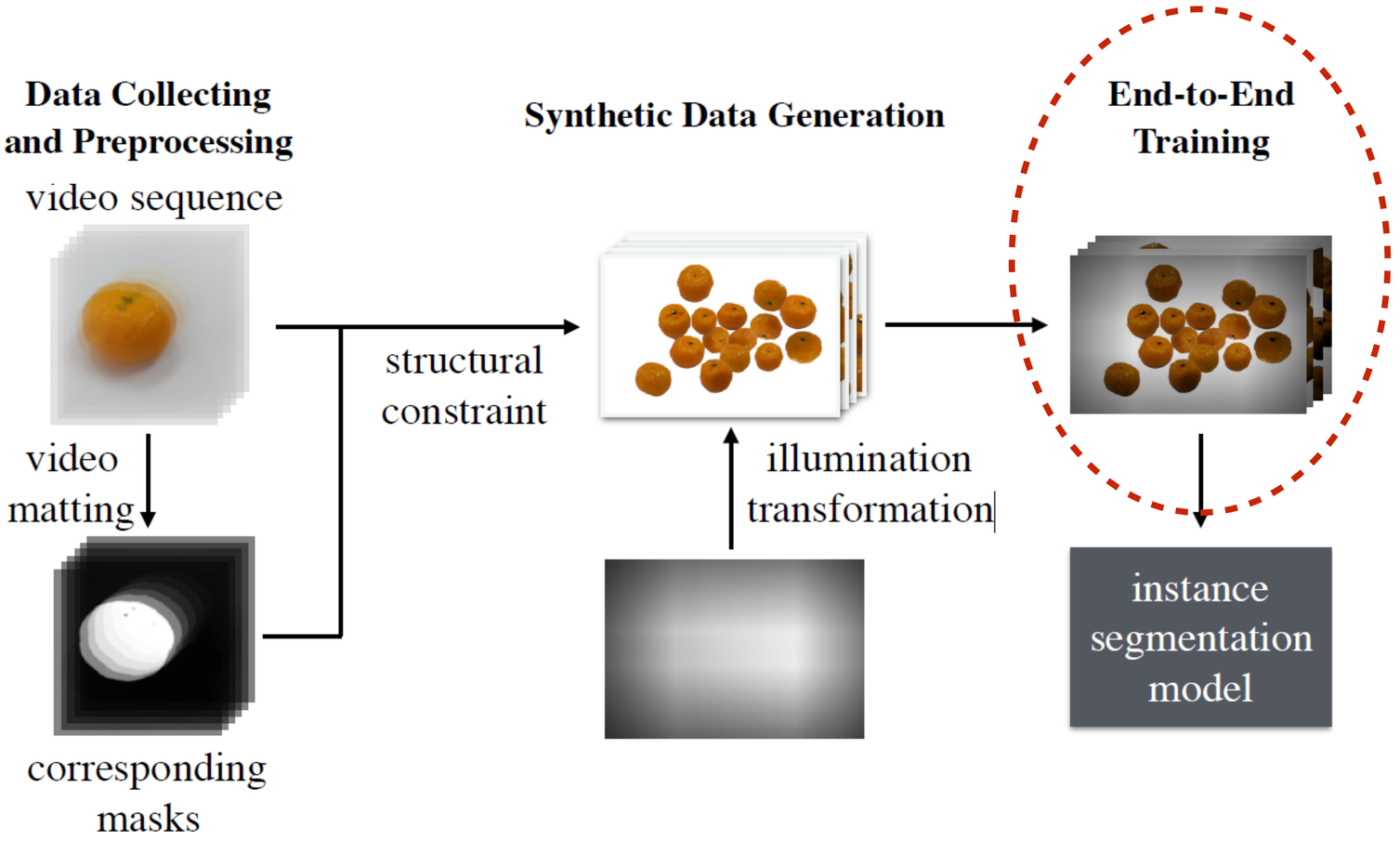
detail removing: apply large kernel Gaussian smoothing

(a) (b) (c) (d)

$$V_{syn} = V_{syn} - \text{mean}(V_{syn}) + blur(V_{real})$$

$$V_{real} = V_{real} - \text{mean}(V_{real}) + blur(V_{real})$$

End-to-End Training



Dataset

- 10 classes
- 200 images in total
- 18.3 instances per image
- different backgrounds and lighting condition

Baselines

Single: images containing single object as training data

Random: randomly synthesized images as training data

Random+illumination: applies illumination transformation method to transform the synthetic data generated in **Random**

Random+structure: uses the method proposed in “Structural Constraint” section to generate structurally realistic training data

Experiments

Table 1: Results on mAP^r@0.5 on our dataset. All numbers are percentages %.

	badminton	battery	clothespin	grape	milk	hexagon nut	orange	ping pong	tissue	wing nut	mAP
Single	12.1	23.2	4.4	19.3	8.2	17.5	17.6	14.2	13.1	21.1	15.1
Random	40.6	50.9	38.4	50.8	26.7	52.9	63.8	67.2	83.9	32.9	50.8
Random+illumination	44.6	48.7	34.3	41.6	26.0	46.3	54.9	64.2	68.9	39.4	46.9
Random+structure	34.2	39.3	52.6	**72.7**	31.3	62.8	**90.3**	**81.7**	**90.7**	23.7	57.9
Ours	**53.0**	**69.5**	**67.7**	72.5	**52.6**	**73.6**	90.0	81.2	90.4	**48.4**	**69.9**

Table 2: Results on mAP^r@[0.5:0.95] on our dataset. All numbers are percentages %.

	badminton	battery	clothespin	grape	milk	hexagon nut	orange	ping pong	tissue	wing nut	mAP
Single	8.7	21.2	2.0	16.8	5.0	15.2	16.0	11.8	9.9	18.8	12.5
Random	33.1	44.7	29.3	44.8	20.9	43.4	56.3	59.5	68.4	27.7	42.8
Random+illumination	36.5	43.2	28.4	37.2	20.8	38.3	50.4	56.4	56.8	35.5	40.4
Random+structure	29.8	36.5	36.7	**66.5**	22.0	45.2	83.8	**76.4**	**80.4**	20.7	49.8
Ours	**44.2**	**60.3**	**46.2**	65.4	**41.3**	**54.8**	**84.0**	75.6	**80.4**	**39.3**	**59.2**

Qualitative Results

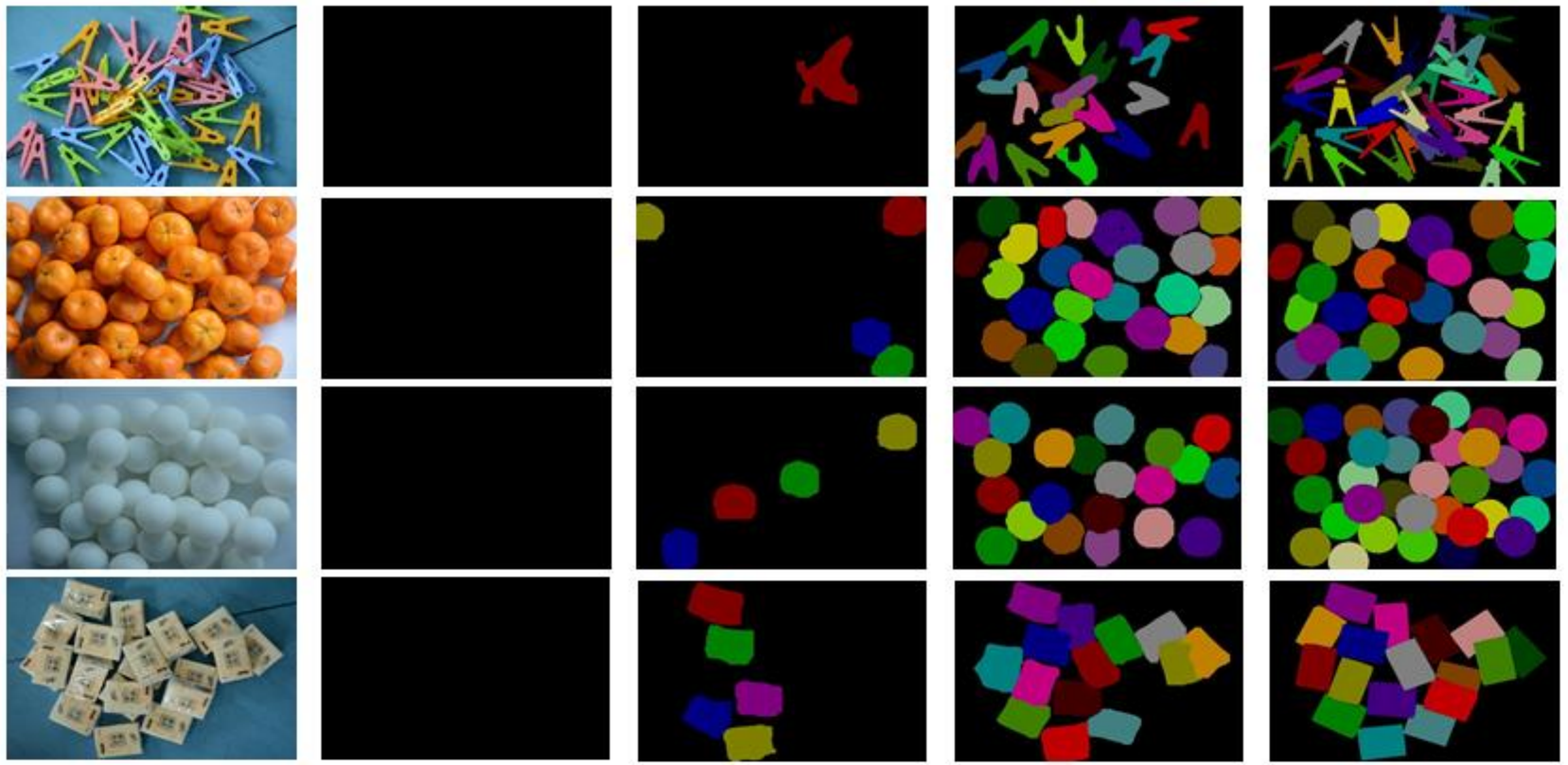


Image Single Random Ours Groundtruth

Thanks

Q & A