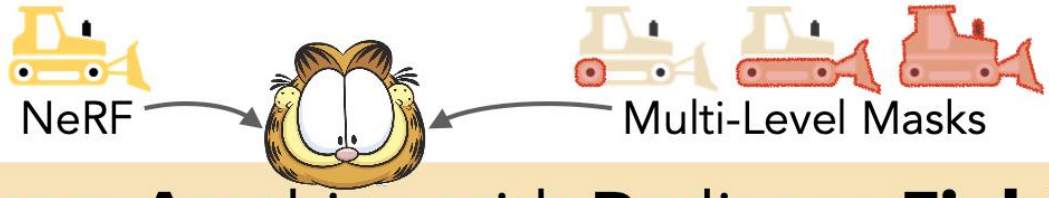
GARField: Group Anything with Radiance Fields

Chung Min Kim*[1] Mingxuan Wu*[1] Justin Kerr*[1] Ken Goldberg[1]
Matthew Tancik[2] Angjoo Kanazawa[1]
* Denotes equal contribution
[1]UC Berkeley [2] Luma AI

CVPR 2024

Citations: 2

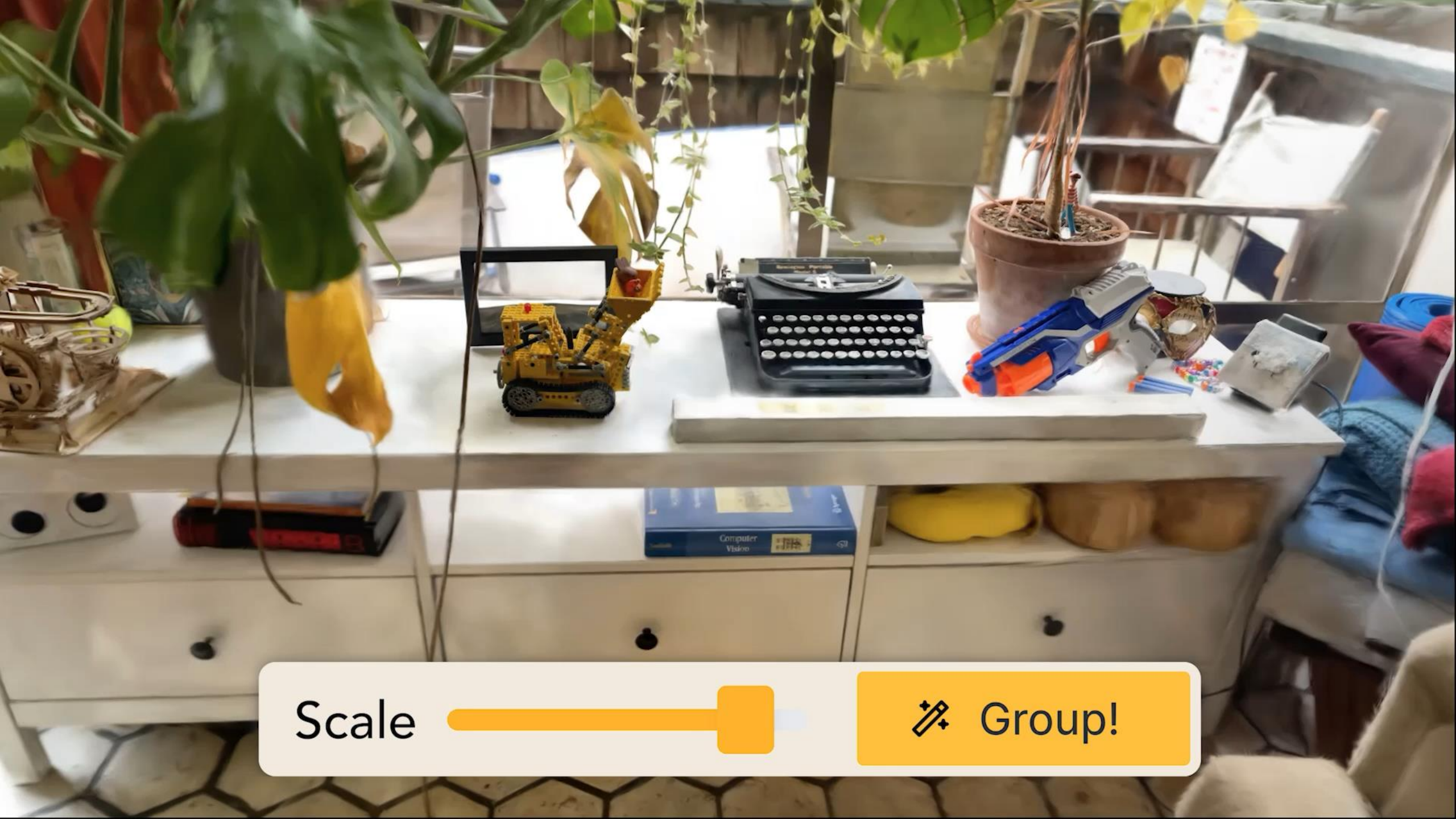
Source: https://www.garfield.studio/

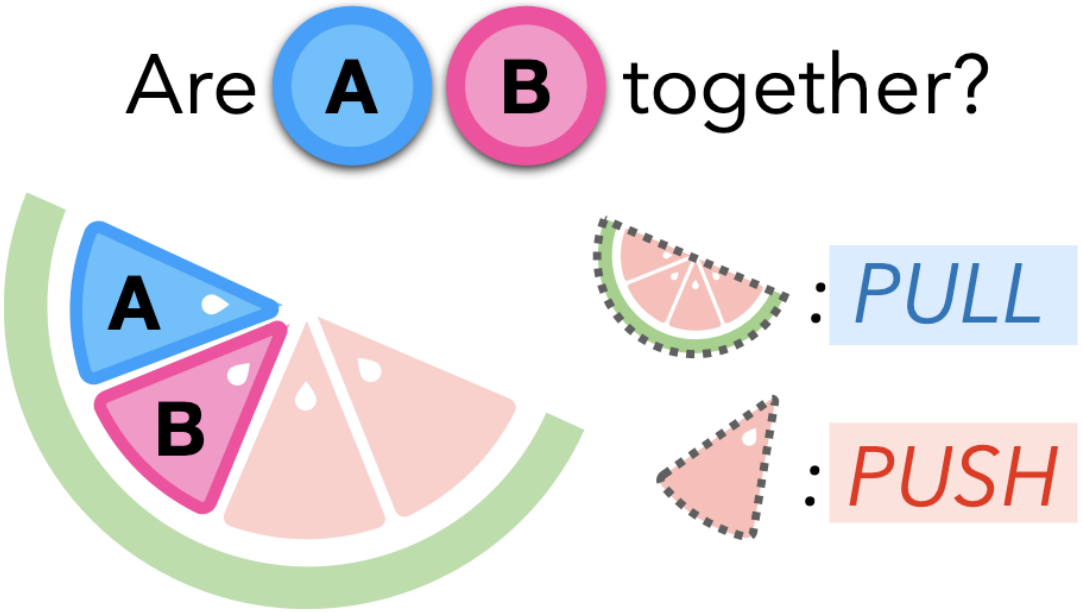


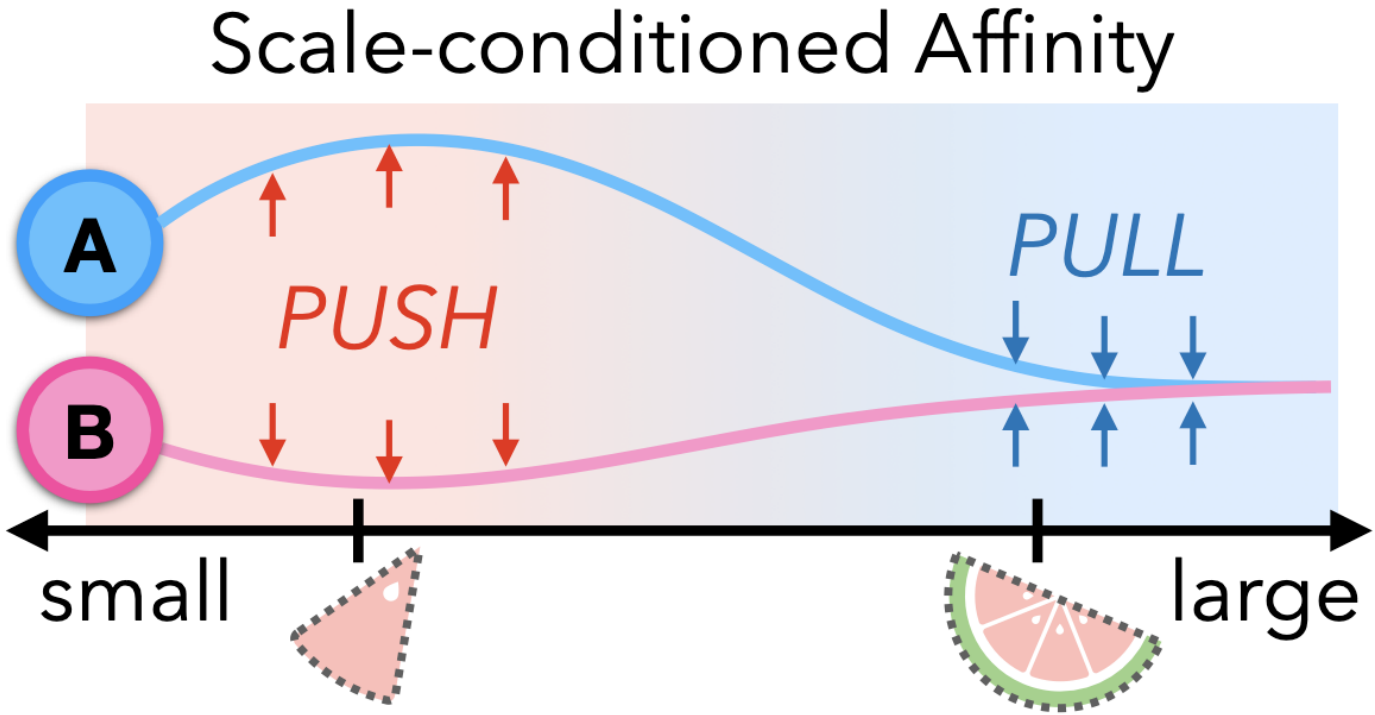
Group Anything with Radiance Fields

Hierarchical Asset Extraction

Multi-Scale Clustering


Computer
Vision
Scale
Group!





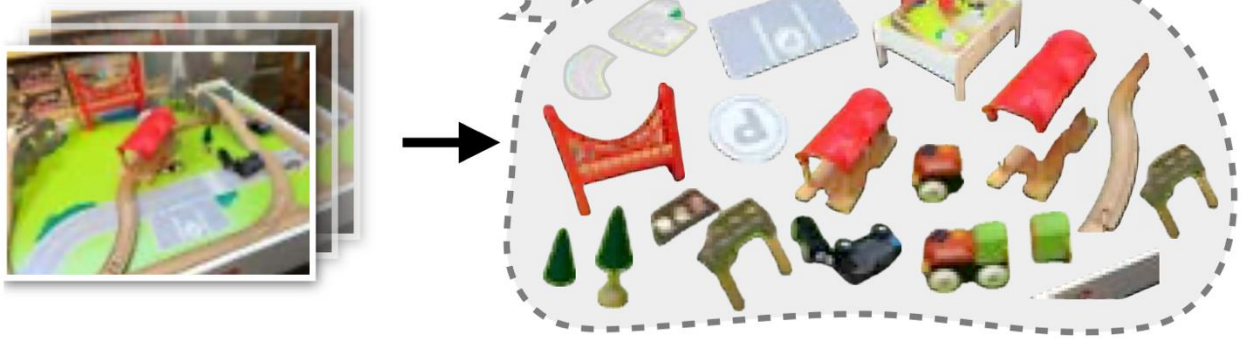
higher affinity @ large scale but low affinity @ a smaller scale

2D Mask Generation

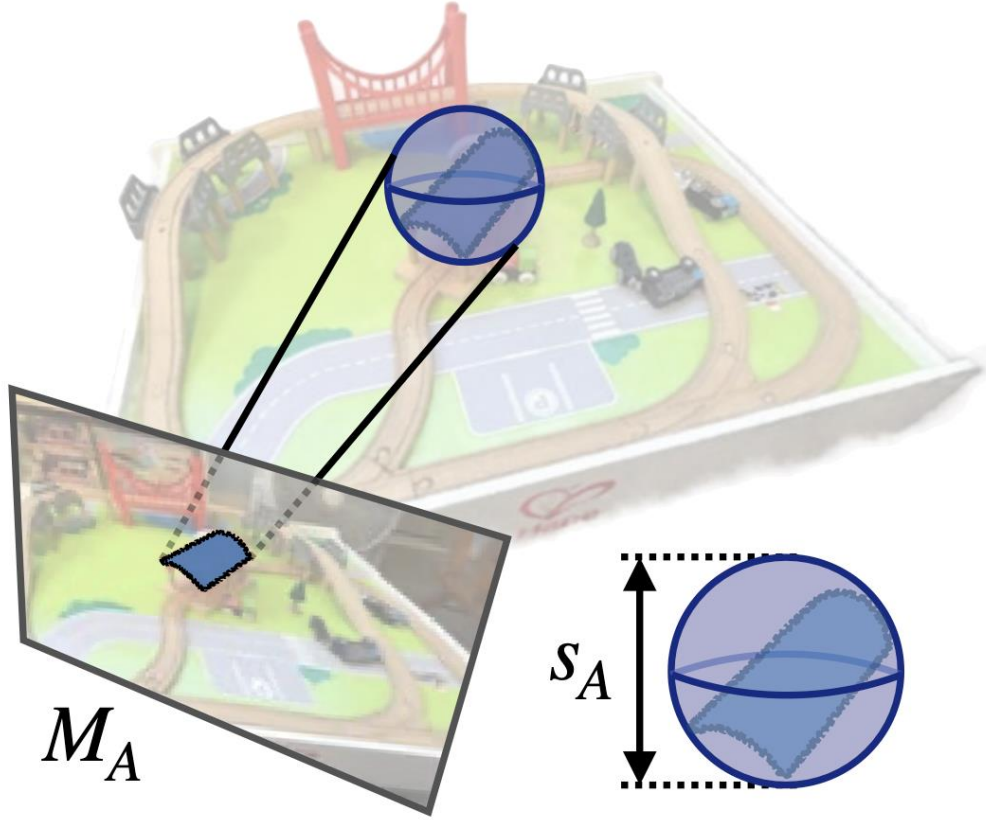
Group Preprocessing

Images

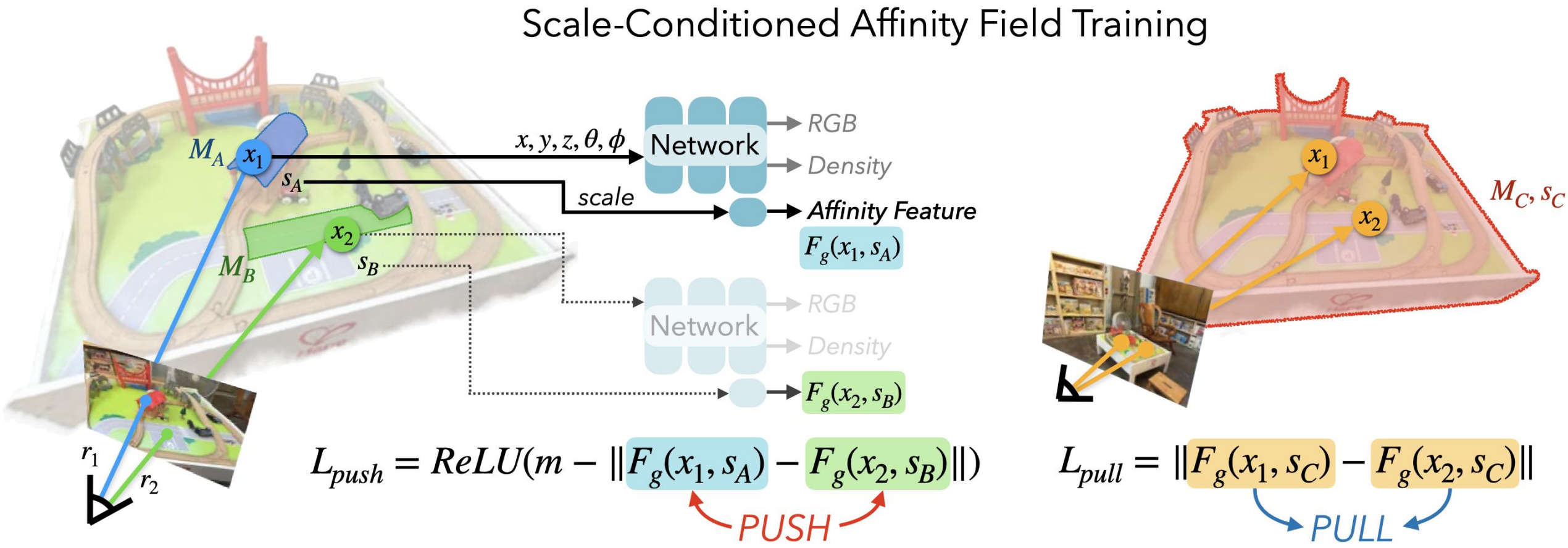
Set of SAM Masks

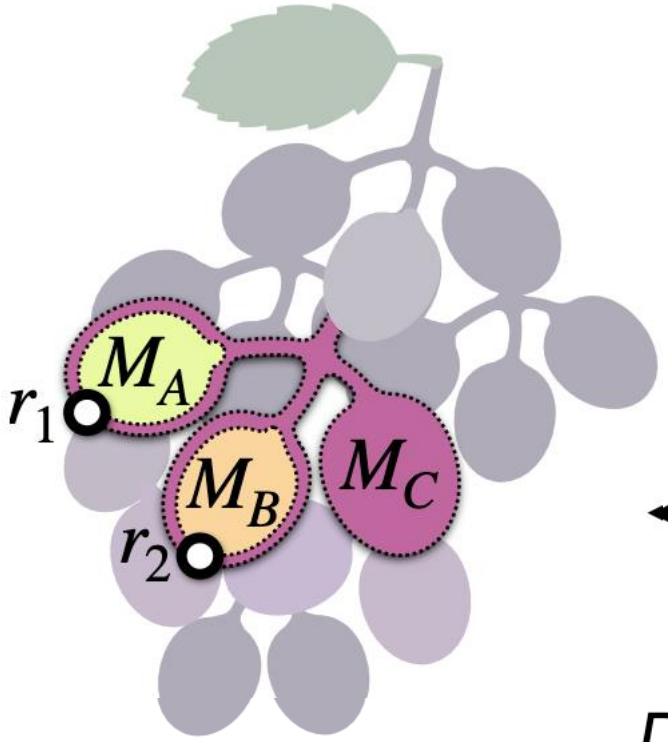


3D Mask Scale Generation

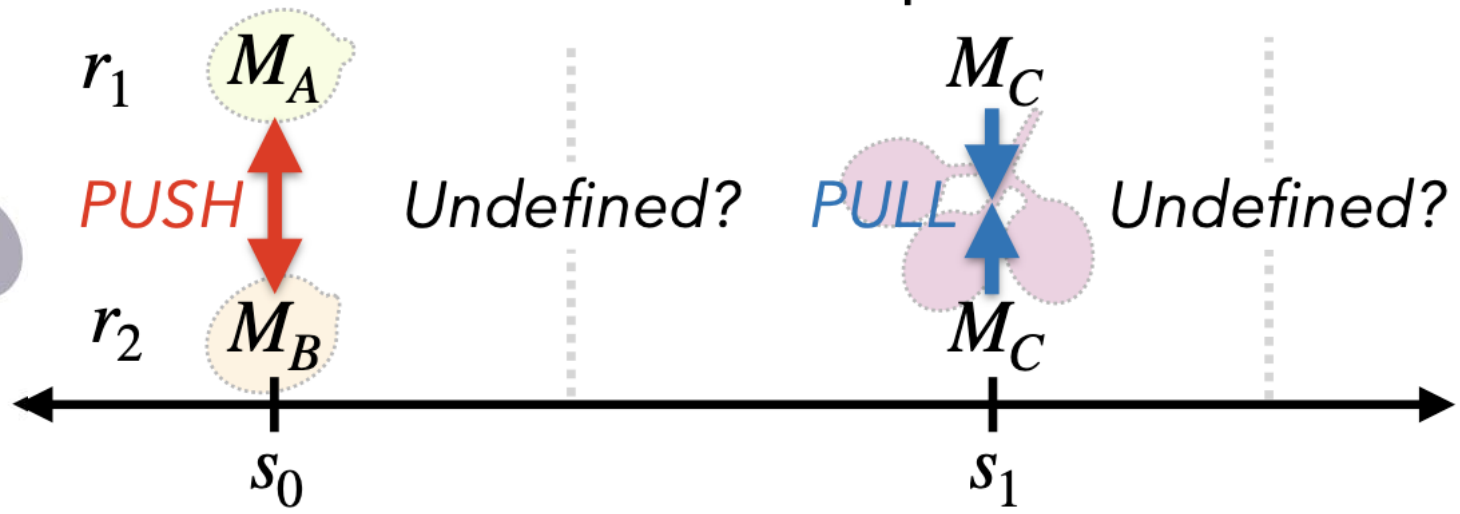


Scale-Conditioned Affinity Field

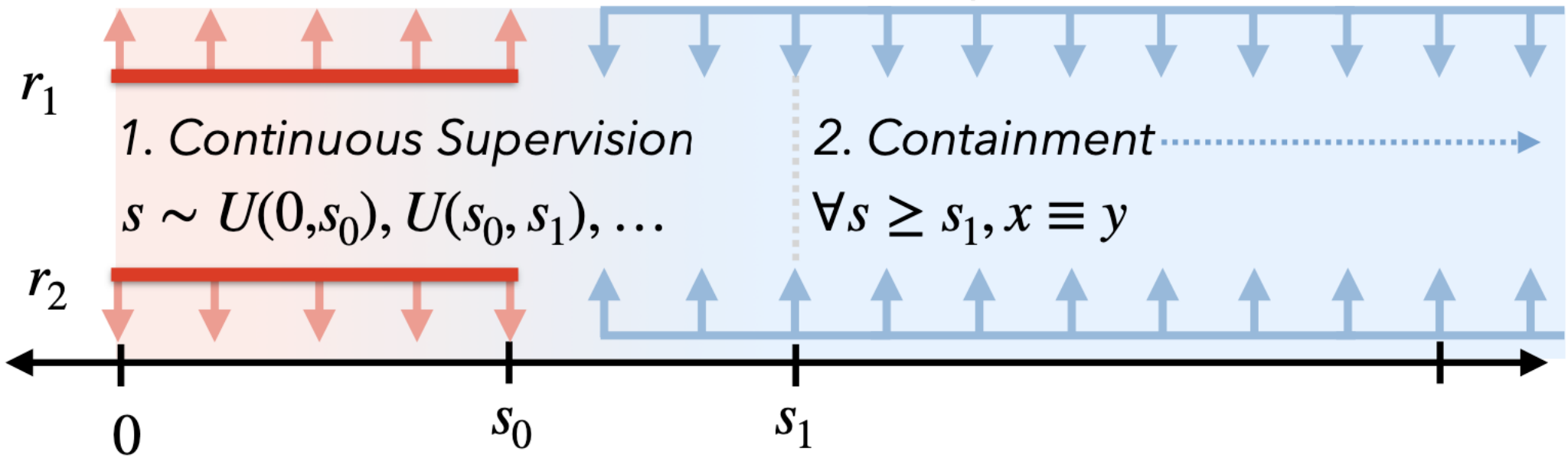




Naive Scale Supervision



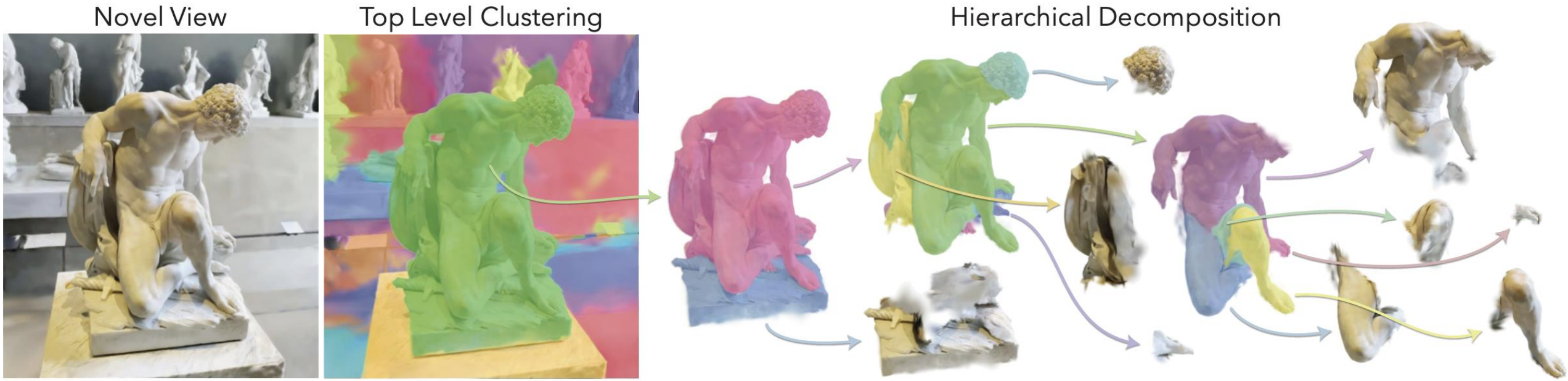
Densified Scale Supervision



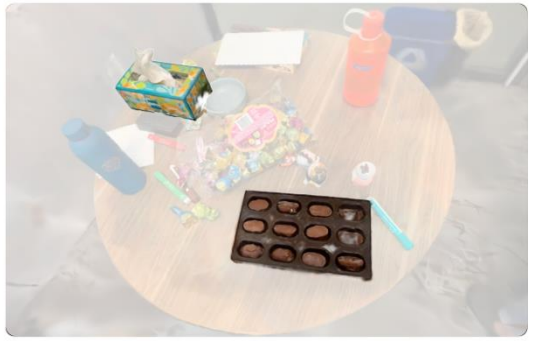
Interactive Selection

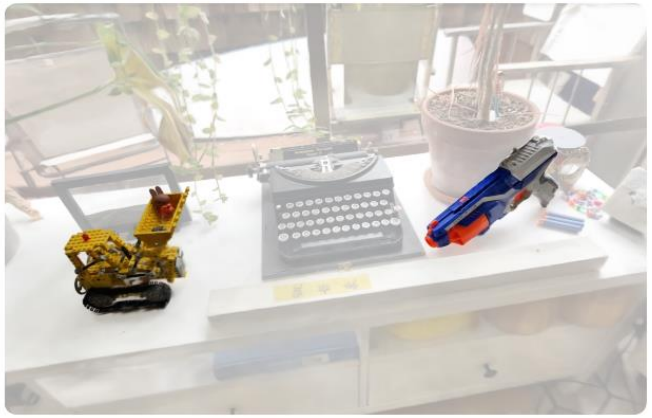
3D Consistent Asset Extraction

Hierarchical Decomposition







RGB SAM (2D) GARField

Affinity w/

GARField (-dense hierarchy)

GARField

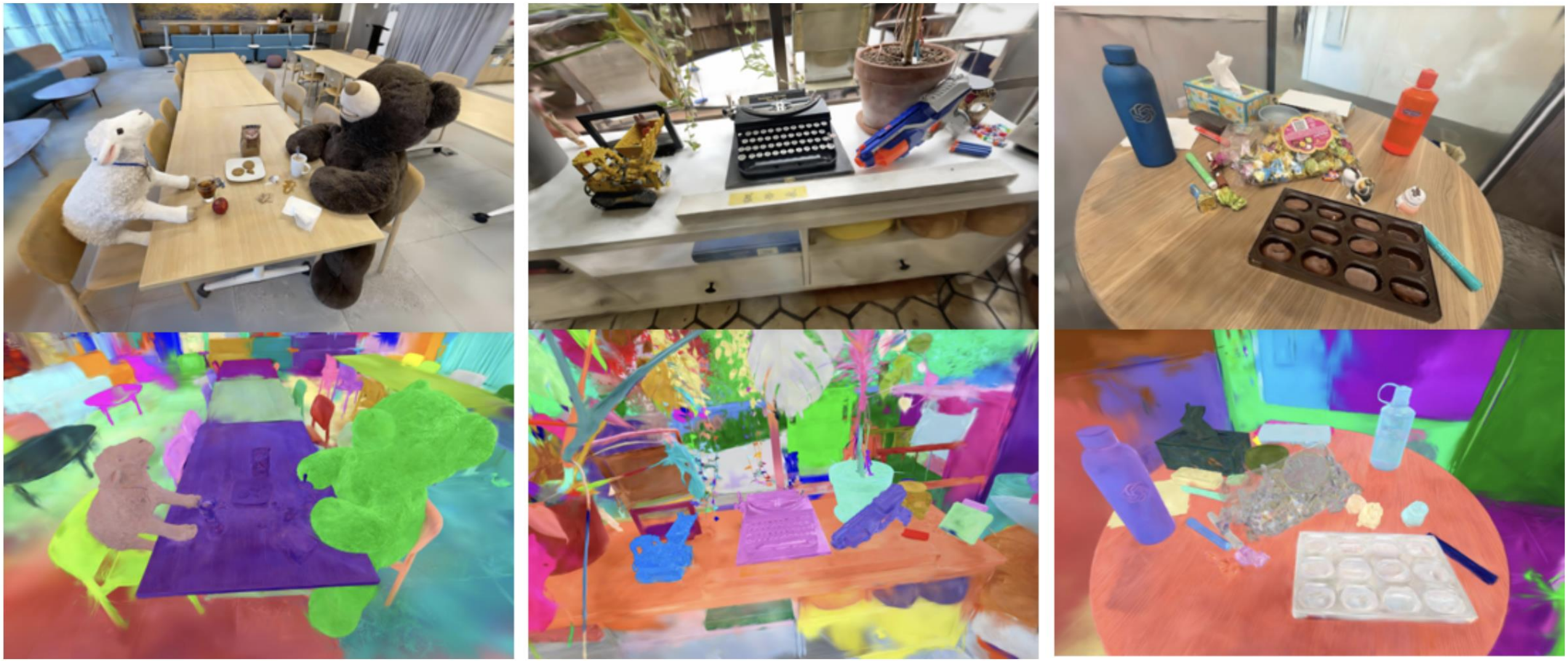
s 0.0 0.04 0.06 0.22 0.5

Scene	Fine		Medium		Coarse	
	SAM	Ours	SAM	Ours	SAM	Ours
teatime	81.6	**92.7**	97.3	**97.9**	-	-
bouquet	17.4	**76.0**	73.5	**81.6**	76.1	**85.4**
keyboard	65.3	**88.8**	73.6	**98.4**	-	-
ramen	53.3	**79.2**	74.7	**90.7**	92.6	**95.5**
living_room	85.3	**90.5**	74.2	**80.7**	88.6	**94.4**

Table 1. **3D Completeness.** We report mIOU of scene annotations for a single point with up to three levels of hierarchy. SAM struggles to produce view-consistent fine groups compared to GARField.

Scene	SAM [15]	Ours (-scale)	Ours (-dense)	Ours
ramen	74.9	64.1	74.1	**85.6**
teatime	64.9	67.7	66.1	**86.6**
keyboard	23.2	57.6	73.1	**77.9**
bouquet	34.4	49.8	72.9	**76.4**
living_room	59.6	49.7	62.1	**76.6**

Table 2. **Hierarchical Grouping Recall:** We report mIOU against human annotations of multi-scale groups of different objects.



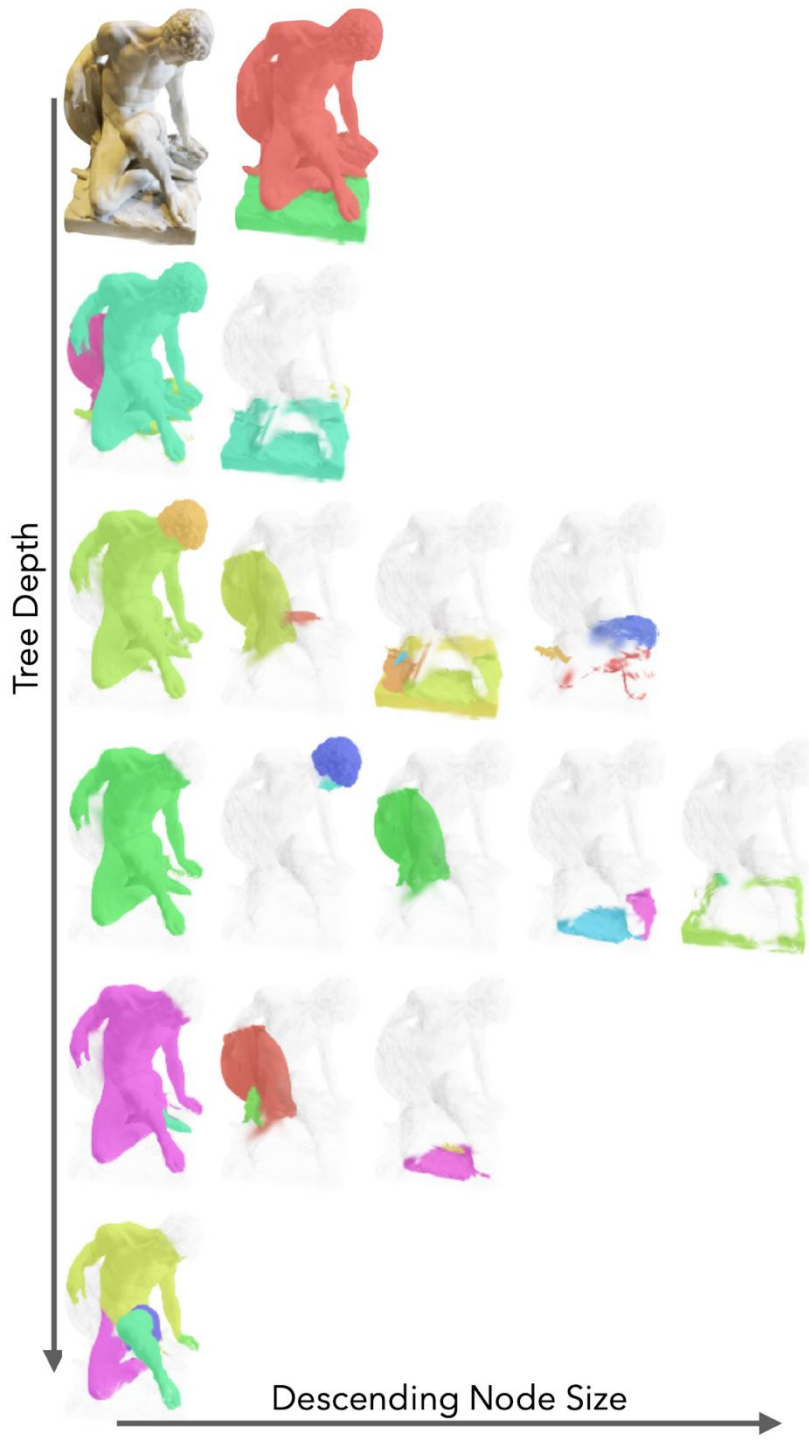
Limitations

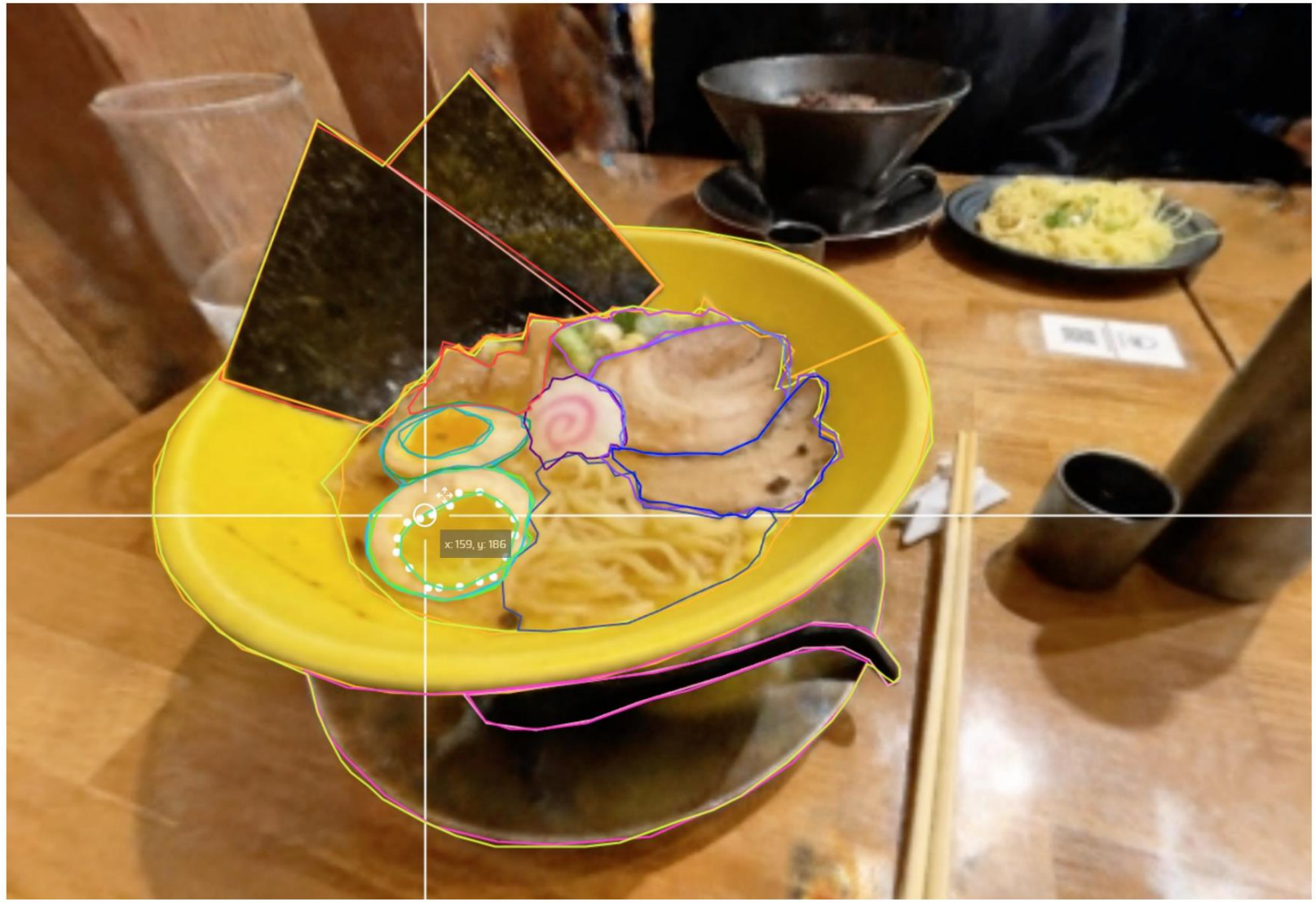
GARField → 2D Mask Generator → desired group not in mask → not in 3D

Uneven viewpoints → artificial group boundaries

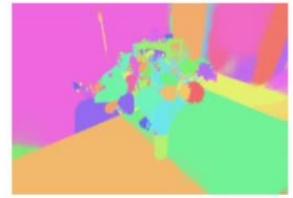
Multiple grouping within a single scale

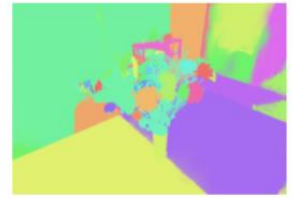
object parts of different sizes branch off the tree separately

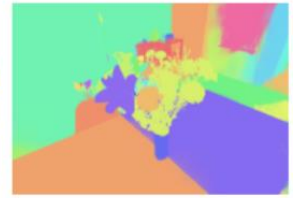

Tree Depth
Descending Node Size

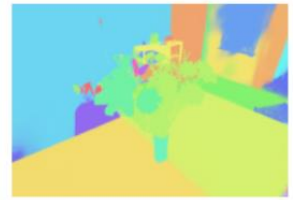

x: 159, y: 186

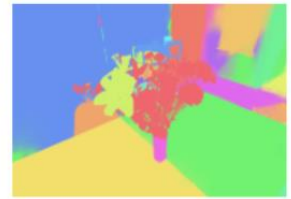
RGB (novel view)

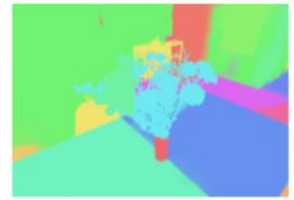


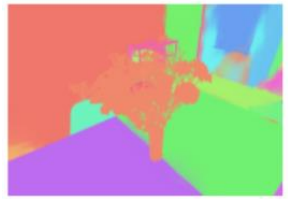


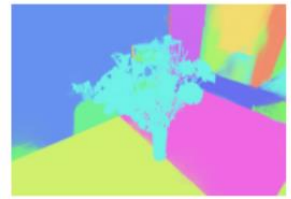


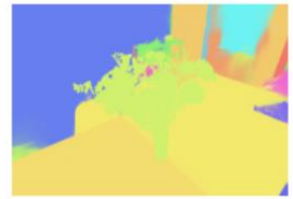


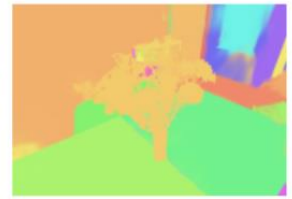


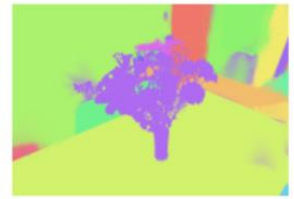


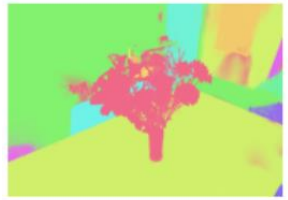


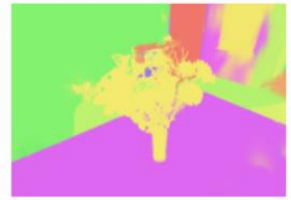


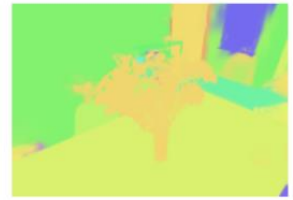


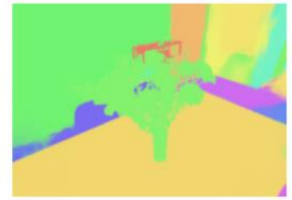


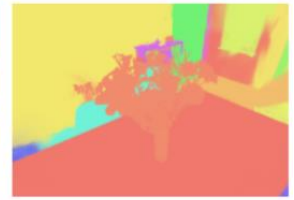


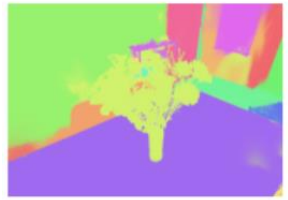


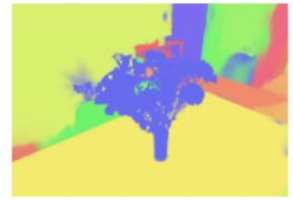


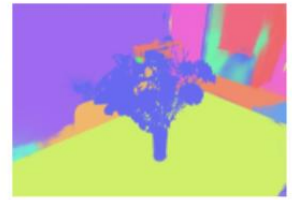


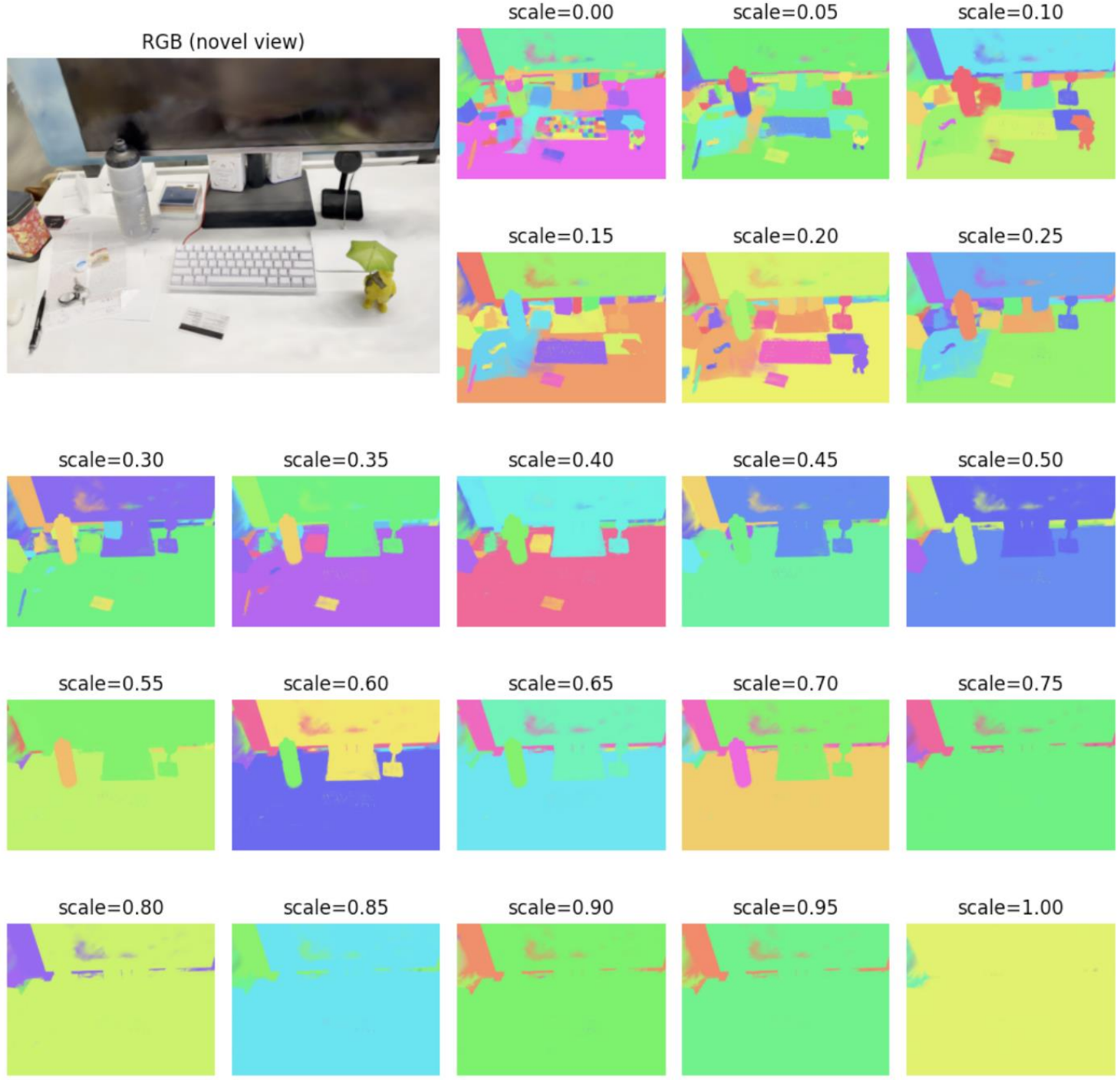











RGB (novel view)
scale=0.00
scale=0.05
scale=0.10
scale=0.15
scale=0.20
scale=0.25
scale=0.30
scale=0.35
scale=0.40
scale=0.45
scale=0.50
scale=0.55
scale=0.60
scale=0.65
scale=0.70
scale=0.75
scale=0.80
scale=0.85
scale=0.90
scale=0.95
scale=1.00

RGB (novel view)

scale=0.00

scale=0.05

scale=0.10

scale=0.15

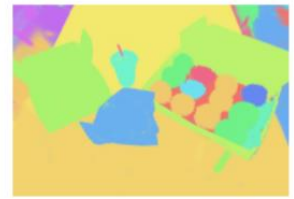
scale=0.20

scale=0.25

scale=0.30

scale=0.35

scale=0.40



scale=0.45

scale=0.50

scale=0.55

scale=0.60

scale=0.65

scale=0.70

scale=0.75

scale=0.80

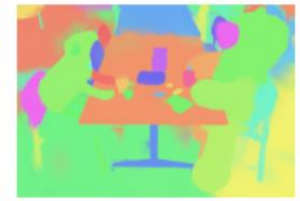
scale=0.85

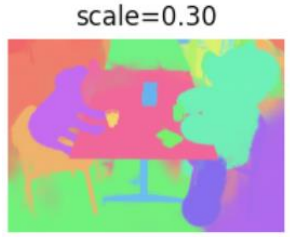
scale=0.90

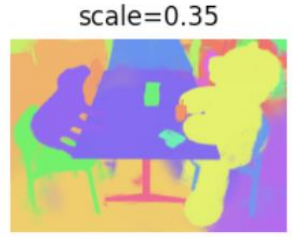
scale=0.95

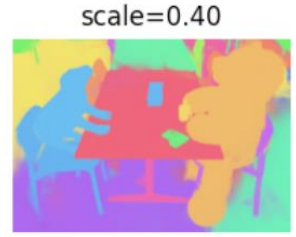
scale=1.00

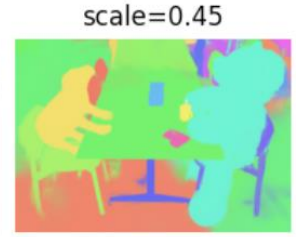
RGB (novel view)

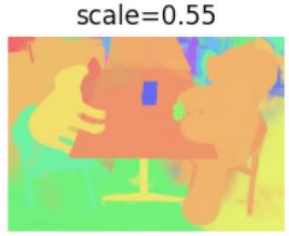


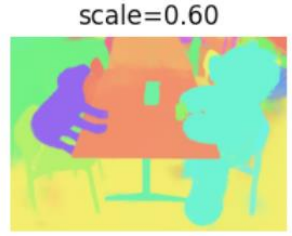


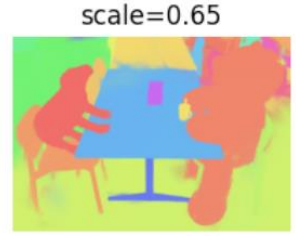


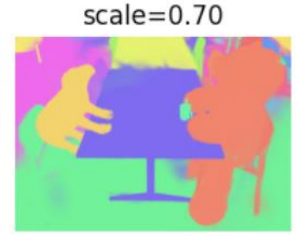


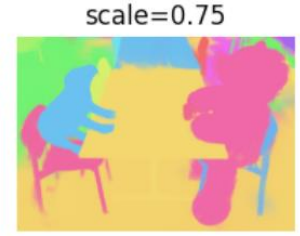


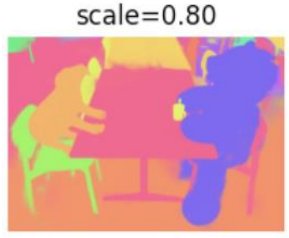


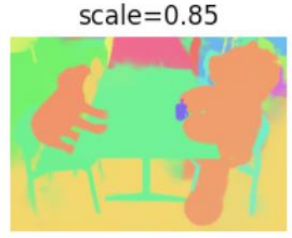


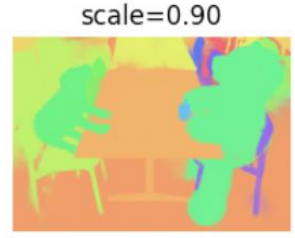


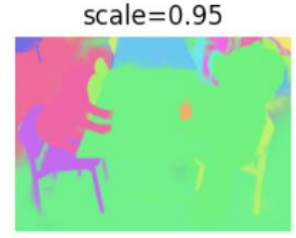


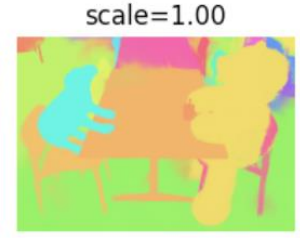






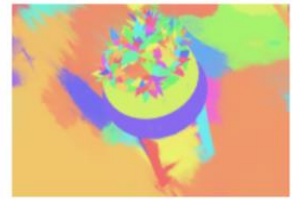






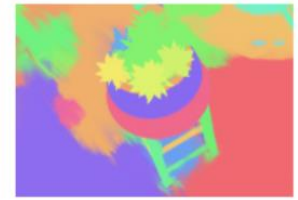
RGB (novel view)

scale=0.00



scale=0.05

scale=0.10



scale=0.15

scale=0.20

scale=0.25

scale=0.30

scale=0.35

scale=0.40

scale=0.45

scale=0.50

scale=0.55

scale=0.60

scale=0.65

scale=0.70

scale=0.75

scale=0.80

scale=0.85

scale=0.90

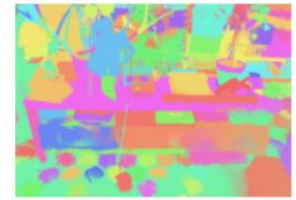
scale=0.95

scale=1.00

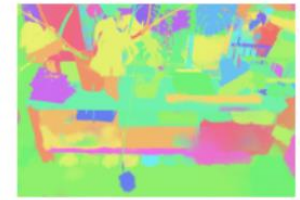
RGB (novel view)

scale=0.00

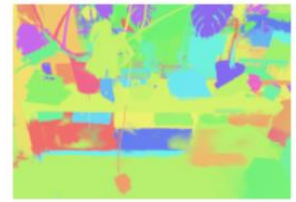
scale=0.05



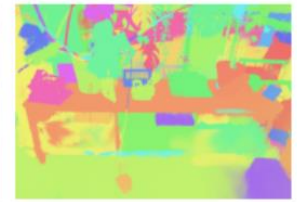
scale=0.10



scale=0.15

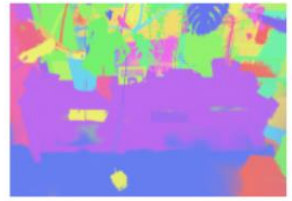


scale=0.20

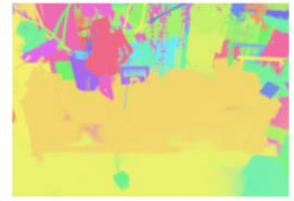


scale=0.25

scale=0.30



scale=0.35

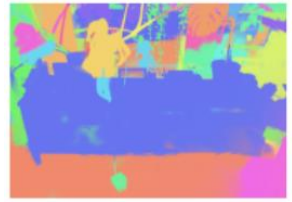


scale=0.40

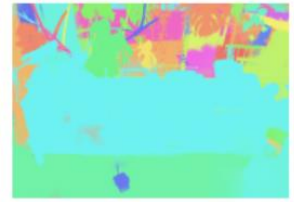
scale=0.45

scale=0.50

scale=0.55



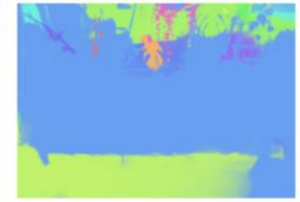
scale=0.60



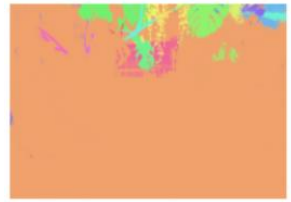
scale=0.65

scale=0.70

scale=0.75



scale=0.80

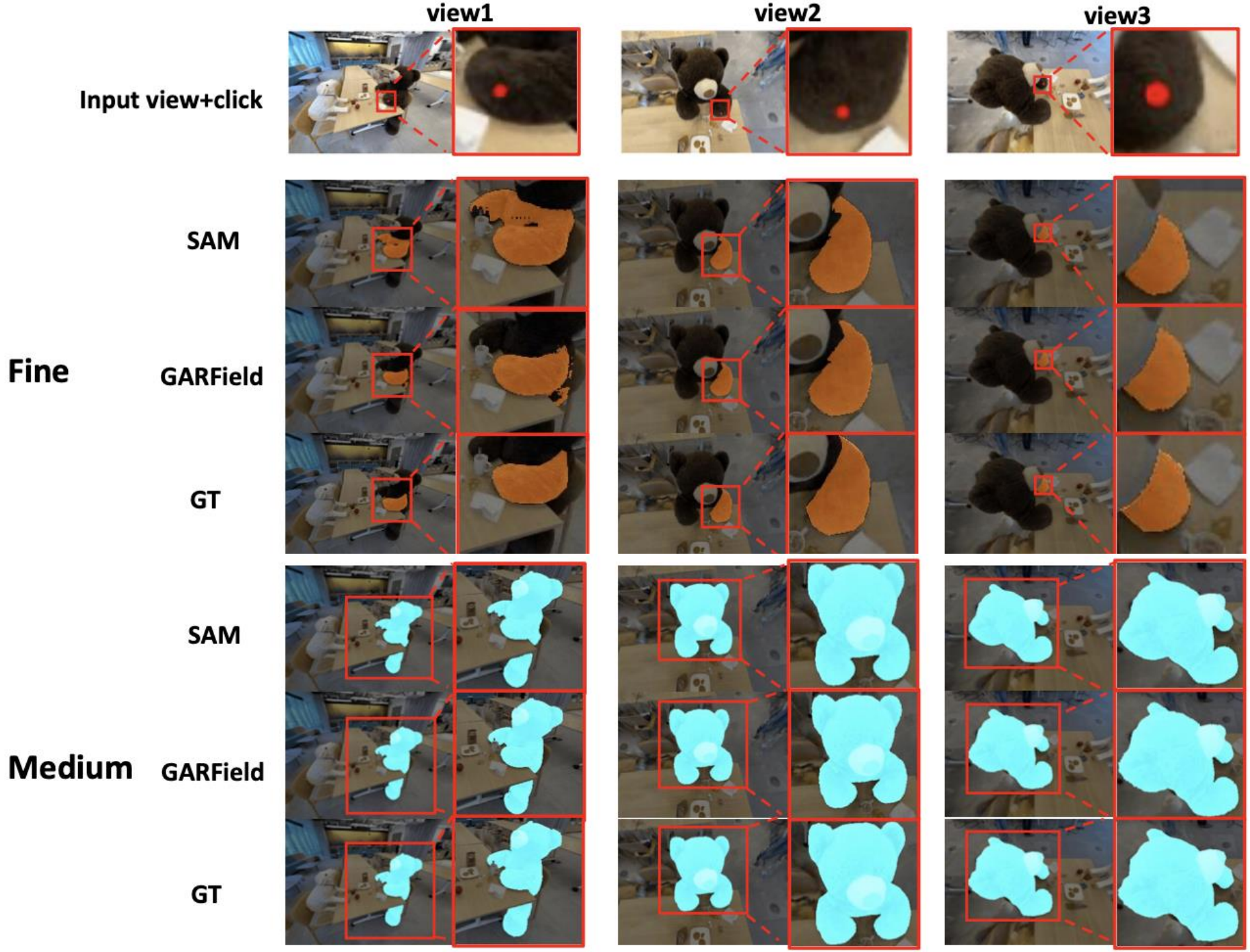


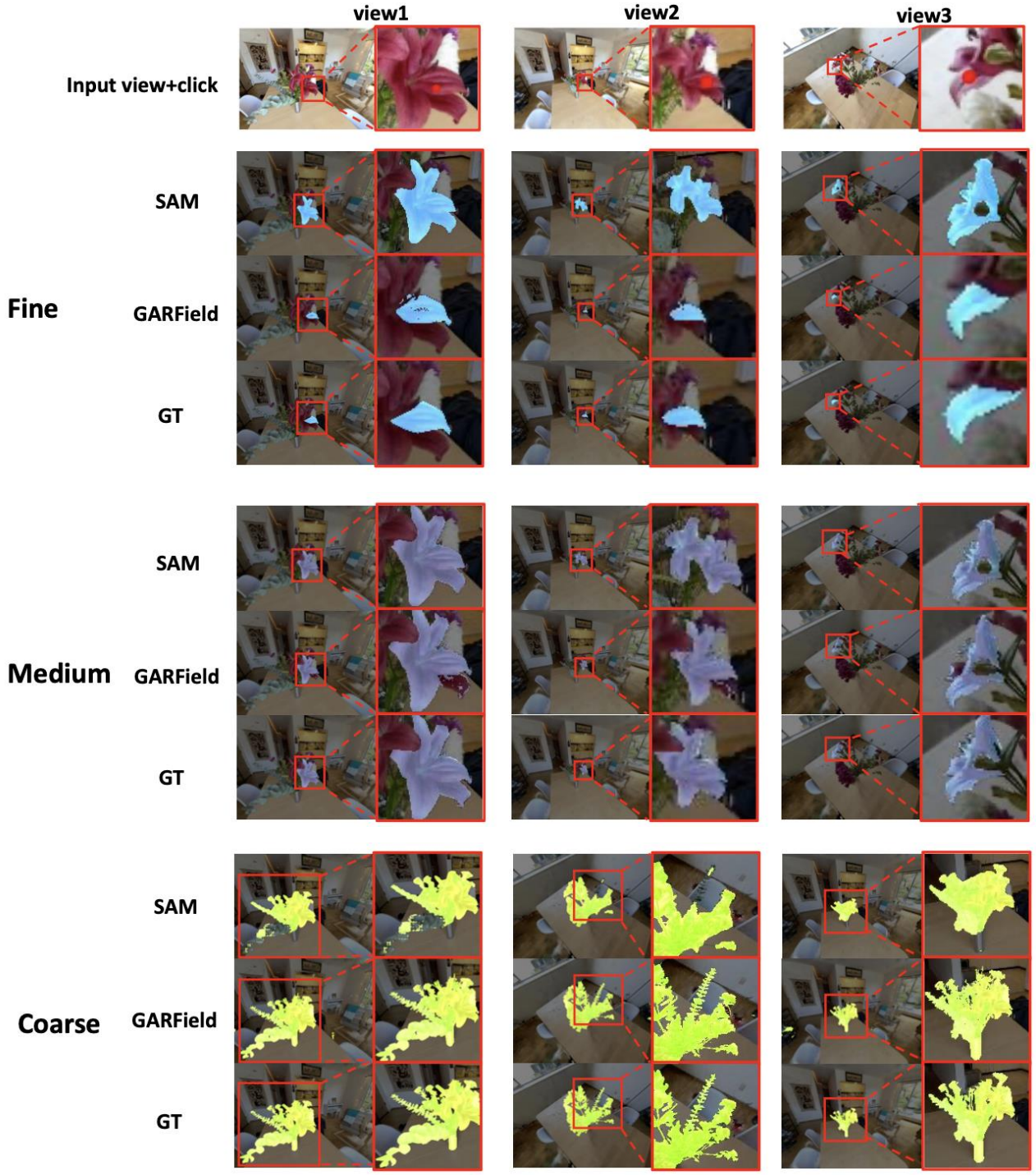
scale=0.85

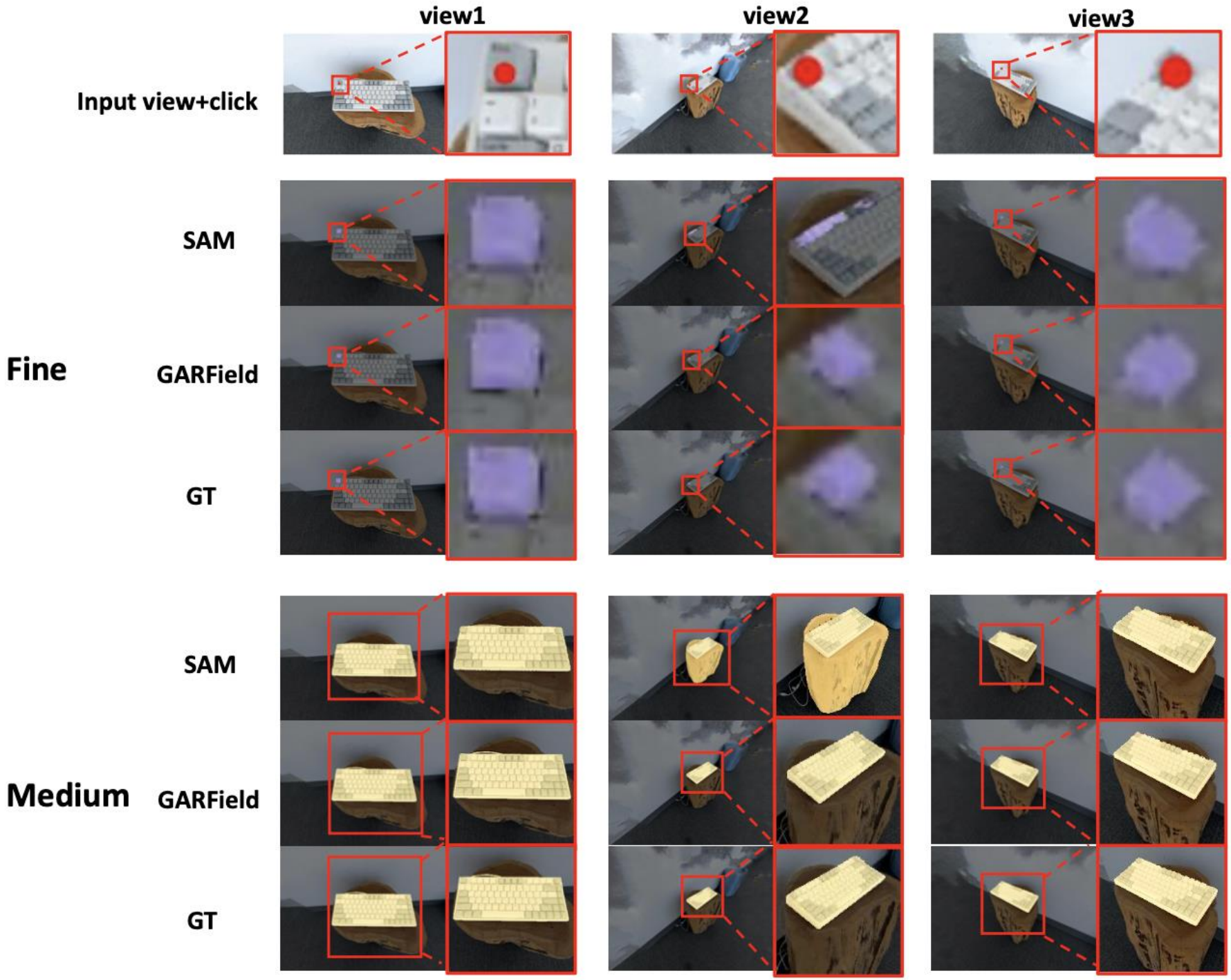
scale=0.90

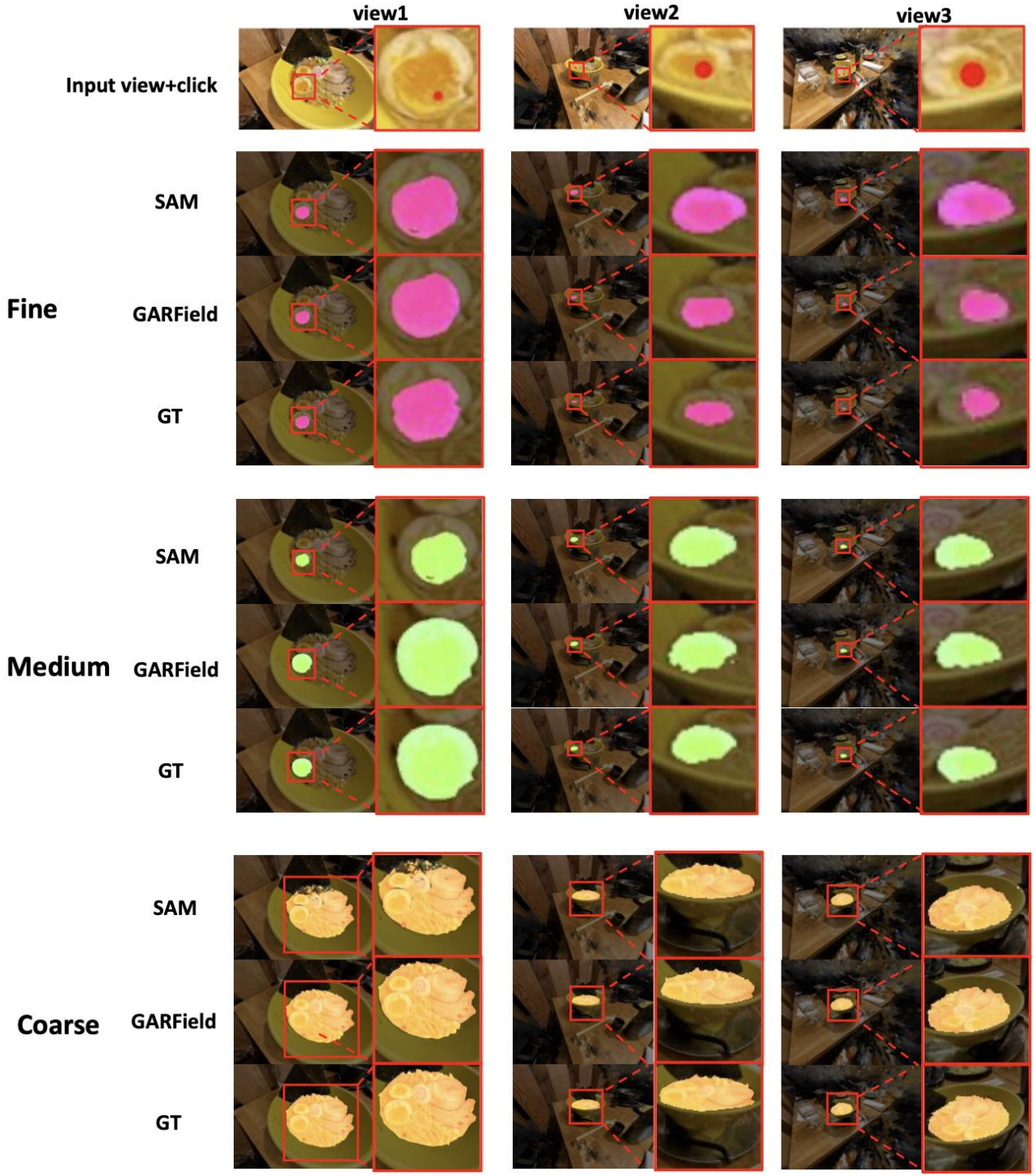
scale=0.95

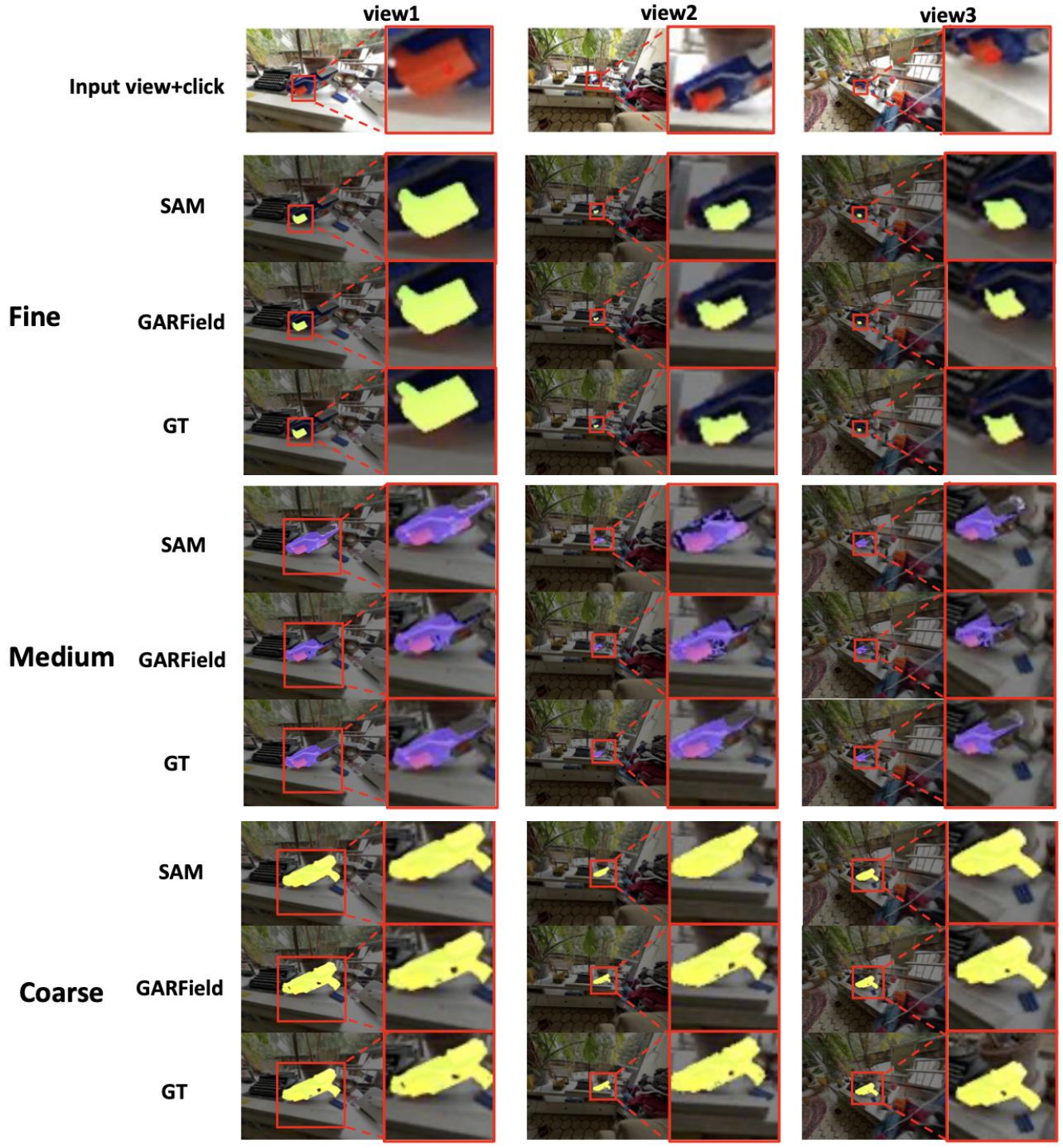
scale=1.00

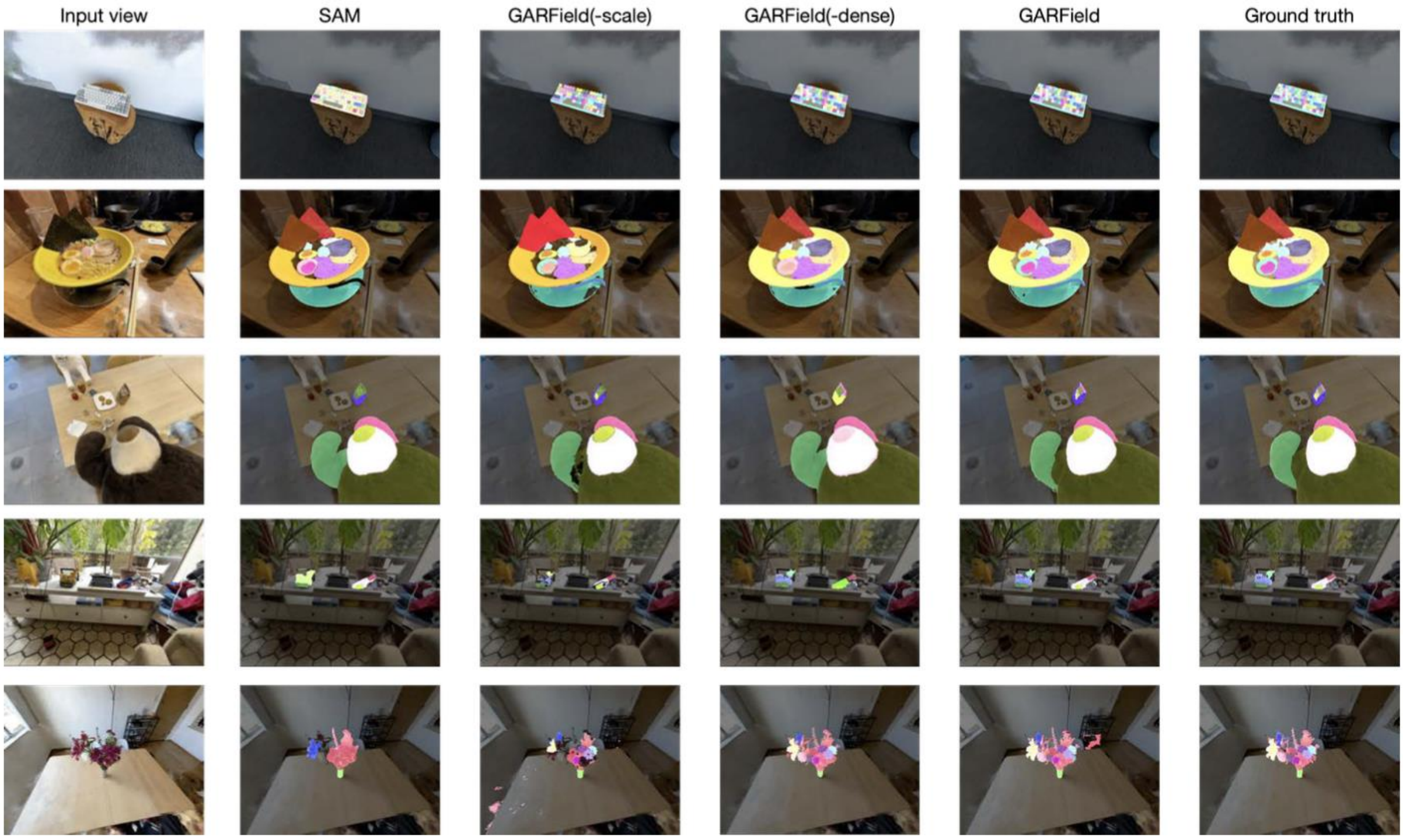

view1
view2
view3
Input view+click
Fine
SAM
GARField
GT
Medium
SAM
GARField
GT


view1
view2
view3
Input view+click
Fine
SAM
GARField
GT
Medium
SAM
GARField
GT
Coarse
SAM
GARField
GT


view1
view2
view3
Input view+click
Fine
SAM
GARField
GT
Medium
SAM
GARField
GT


view1
view2
view3
Input view+click
Fine
SAM
GARField
GT
Medium
SAM
GARField
GT
Coarse
SAM
GARField
GT


view1
view2
view3
Input view+click
Fine
SAM
GARField
GT
Medium
SAM
GARField
GT
Coarse
SAM
GARField
GT


Input view
SAM
GARField(-scale)
GARField(-dense)
GARField
Ground truth

Q&A

Source: https://www.garfield.studio/