

# Learning Multi-Task Robot Trajectory Segmentation from Visual and Kinematic Streams

## Supplementary Material

### 8. Detailed Evaluation Breakdown

This appendix provides detailed per-task evaluation results, including representative trajectory visualizations and ablation study performance metrics.

#### 8.1. Task 327: Pickup items in the supermarket

This task is performed in Supermarket scenario using grippers, with typical duration of 30s to 48s and 1254 total action slices.

##### Task phases:

1. retrieve cucumber from the shelf.
2. place the held cucumber into the plastic bag in the shopping cart.
3. retrieve tomato from the shelf.
4. place the held tomato into the plastic bag in the shopping cart.
5. retrieve corn from the shelf.
6. place the held corn into the shopping carts plastic bag.

**Skill list:** Pick, Place

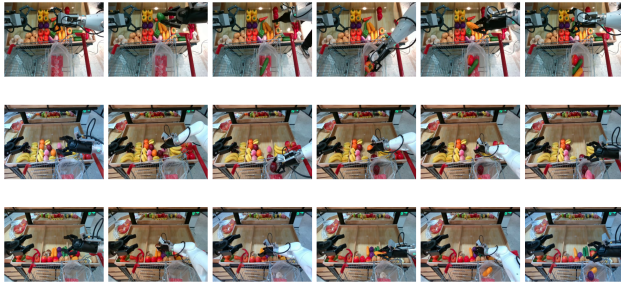


Figure 4. Three representative trajectory examples for Task 327: Pickup items in the supermarket.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	23.55	28.79	31.47	53.10	51.92	26.79
No Kinematics	21.40	24.53	24.53	41.61	33.71	14.47
MLP + MSE	21.28	23.49	23.62	40.43	33.88	12.67
RoboSegNet	<b>53.18</b>	<b>68.29</b>	<b>72.46</b>	<b>87.22</b>	<b>87.10</b>	<b>83.65</b>
<i>Multi-Task Training</i>						
No Visual	29.75	41.10	46.91	69.49	69.17	54.73
No Kinematics	21.58	24.53	24.53	41.61	31.63	14.67
RoboSegNet	21.10	39.95	47.84	68.21	68.19	50.58

#### 8.2. Task 354: Pickup items in the supermarket

This task is performed in Supermarket scenario using grippers, with typical duration of 7s to 25s and 1032 total action slices.

##### Task phases:

1. retrieve from the shelf.
2. place the held red battery into the shopping cart.

**Skill list:** Pick, Place



Figure 5. Three representative trajectory examples for Task 354: Pickup items in the supermarket.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	15.18	19.10	20.74	45.81	44.60	18.44
No Kinematics	14.56	16.64	16.64	36.15	24.83	8.67
MLP + MSE	14.64	15.86	15.94	34.98	24.92	7.79
RoboSegNet	<b>36.41</b>	<b>55.86</b>	<b>65.47</b>	<b>88.64</b>	<b>88.24</b>	<b>85.43</b>
<i>Multi-Task Training</i>						
No Visual	19.31	26.53	31.10	61.04	60.68	44.39
No Kinematics	14.71	16.64	16.64	36.15	23.20	9.63
RoboSegNet	15.87	29.45	35.52	65.59	65.42	48.15

#### 8.3. Task 356: Packing in the supermarket

This task is performed in Supermarket scenario using grippers, with typical duration of 20s to 35s and 2375 total action slices.

##### Task phases:

1. grasp the left arm with white plastic bag.
2. grab the blue cleansing milk on the table with the right arm.
3. place the held blue cleansing milk into white plastic bag.
4. grab the hawthorn tea on the table with the right arm.
5. place the held hawthorn tea into white plastic bag.

**Skill list:** Grasp, Pick, Place

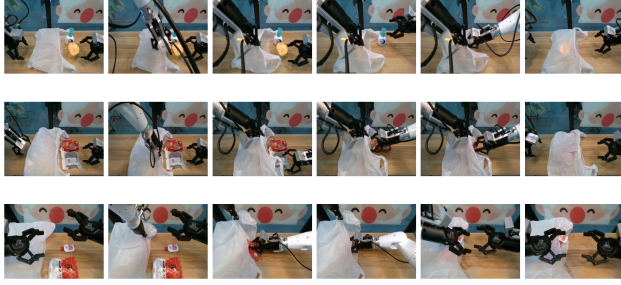


Figure 6. Three representative trajectory examples for Task 356: Packing in the supermarket.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	26.45	31.96	34.25	53.90	51.63	26.64
No Kinematics	24.40	27.74	27.74	43.86	36.90	16.52
MLP + MSE	24.34	27.84	28.26	44.73	39.56	17.83
RoboSegNet	<b>51.35</b>	<b>68.67</b>	<b>73.25</b>	<b>86.03</b>	<b>85.72</b>	<b>80.82</b>
<i>Multi-Task Training</i>						
No Visual	27.99	37.22	42.17	63.42	62.12	45.24
No Kinematics	24.28	27.74	27.74	43.86	34.61	16.66
RoboSegNet	23.19	43.16	53.38	72.78	72.74	53.38

#### 8.4. Task 357: Wash dishes with dishwasher

This task is performed in Home scenario using grippers, with typical duration of 81s to 149s and 6848 total action slices.

##### Task phases:

1. lift the dishwasher lid with both arms.
2. grasp the chopsticks with the right arm.
3. pass the chopsticks with both hands.
4. place the chopsticks held in the left arm into the dishwasher.
5. grasp the spoon with the right arm.
6. pass the spoon with both hands.
7. place the spoon in the dishwasher with the left arm.
8. lift the bowl with the right arm.
9. pass the bowl with both hands.
10. place the bowl in the dishwasher with the left arm.
11. lift the plate with the right arm.
12. pass the plate with both hands.
13. place the plate in the dishwasher with the left arm.
14. pull down the dishwasher lid with both arms.
15. press the dishwasher lid with the left arm.
16. press the button with the left arm to start the dishwasher.

**Skill list:** HandOver, Pick, Place, Press, Pull, Tap

#### 8.5. Task 358: Toast bread

This task is performed in Home scenario using grippers, with typical duration of 32s to 62s and 2350 total action slices.

##### Task phases:

1. lift a piece of whole wheat toast from the basket with the left arm.



Figure 7. Three representative trajectory examples for Task 357: Wash dishes with dishwasher.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	20.76	23.59	23.99	42.46	33.46	14.74
No Kinematics	23.36	26.54	26.57	45.68	35.83	17.34
MLP + MSE	36.27	41.80	43.40	65.74	65.10	53.94
RoboSegNet	<b>52.25</b>	<b>69.70</b>	<b>73.93</b>	<b>88.19</b>	<b>88.04</b>	<b>84.66</b>
<i>Multi-Task Training</i>						
No Visual	21.59	28.57	32.40	56.10	54.99	33.61
No Kinematics	19.60	22.40	22.40	40.13	28.42	13.31
RoboSegNet	21.00	40.40	50.27	72.41	72.33	52.41

2. insert the whole wheat toast held by the left arm into the empty slot behind white toaster.
3. lift a piece of whole wheat toast from the basket with the left arm.
4. place the whole wheat toast held in the left arm into the empty slot in front of white toaster.
5. press the white toaster start button with both arms.

**Skill list:** Pick, Place, PressButton

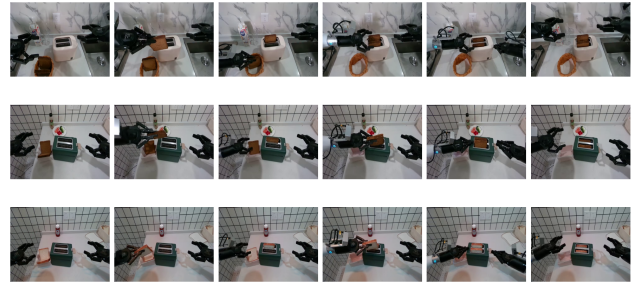


Figure 8. Three representative trajectory examples for Task 358: Toast bread.

#### 8.6. Task 359: Sort in the warehouse

This task is performed in Industry scenario using grippers, with typical duration of 40s to 50s and 3220 total action slices.

##### Task phases:

1. retrieve trash bag from left material frame with the left arm.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	21.89	24.21	25.80	46.66	44.28	18.44
No Kinematics	21.69	24.50	26.07	48.50	46.74	20.28
MLP + MSE	18.16	21.39	22.69	40.58	39.45	13.60
RoboSegNet	<b>61.03</b>	<b>76.14</b>	<b>78.89</b>	<b>90.30</b>	<b>90.15</b>	<b>88.76</b>
<i>Multi-Task Training</i>						
No Visual	18.95	24.82	28.73	49.95	49.62	28.10
No Kinematics	22.36	26.73	27.11	47.15	40.58	15.82
RoboSegNet	23.30	41.47	48.02	69.24	69.24	51.80

- pass the trash bag with both hands.
- place the held trash bag into the right blue target box.
- retrieve toothpick with the right arm from right material frame.
- place the toothpick held in the right arm into the right blue target box.

**Skill list:** HandOver, Pick, Place

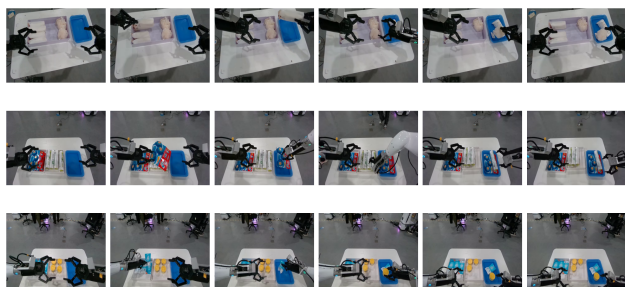


Figure 9. Three representative trajectory examples for Task 359: Sort in the warehouse.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	18.47	21.28	23.04	46.20	44.20	18.11
No Kinematics	17.03	19.39	19.39	38.04	27.43	10.35
MLP + MSE	16.14	18.00	18.16	36.31	27.74	8.86
RoboSegNet	<b>61.27</b>	<b>74.06</b>	<b>75.88</b>	<b>89.31</b>	<b>89.27</b>	<b>88.33</b>
<i>Multi-Task Training</i>						
No Visual	18.70	24.89	28.48	52.94	52.51	32.05
No Kinematics	16.96	19.39	19.39	38.04	25.43	11.22
RoboSegNet	19.94	36.54	43.61	68.11	68.09	50.56

### 8.7. Task 360: Packing in the supermarket

This task is performed in Supermarket scenario using grippers, with typical duration of 28s to 49s and 3619 total action slices.

**Task phases:**

- grasp the left arm with green felt bag.
- grab the tomato sauce on the table with the right arm.
- place the grasped tomato sauce into green felt bag with the right arm.

- grab the steel wool on the table with the right arm.
- place the grasped steel wool into green felt bag with the right arm.
- grasp the wash and care set on the table with the right arm.
- place the grasped wash and care set into green felt bag with the right arm.

**Skill list:** Grasp, Pick, Place

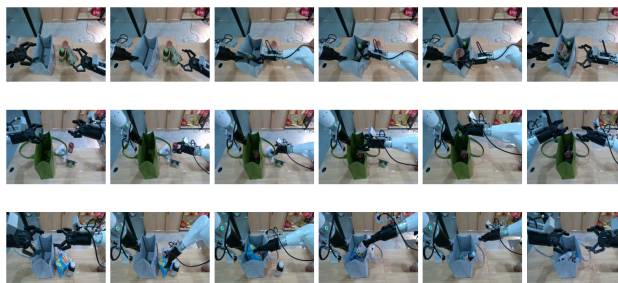


Figure 10. Three representative trajectory examples for Task 360: Packing in the supermarket.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	27.69	33.88	36.38	55.12	53.26	27.46
No Kinematics	25.62	36.22	41.26	57.96	57.91	34.04
MLP + MSE	26.13	30.60	31.59	48.19	45.12	21.28
RoboSegNet	<b>56.58</b>	<b>71.95</b>	<b>74.82</b>	<b>86.19</b>	<b>86.05</b>	<b>82.46</b>
<i>Multi-Task Training</i>						
No Visual	29.56	41.17	47.43	67.89	66.90	49.17
No Kinematics	25.99	29.57	29.57	45.15	37.48	17.61
RoboSegNet	25.06	44.57	53.37	70.90	70.88	49.22

### 8.8. Task 372: Packing in the supermarket

This task is performed in Supermarket scenario using grippers, with typical duration of 45s to 72s and 4833 total action slices.

**Task phases:**

- grasp the left arm with blue nylon bag.
- grab the iced tea on the table with the right arm.
- place the grasped iced tea into blue nylon bag.
- grasp the yogurt bowl on the table with the right arm.
- place the grasped yogurt bowl into blue nylon bag.
- retrieve the pork jerky from the table with the right arm.
- place the grasped pork jerky into blue nylon bag.
- grab the mouthwash on the table with the right arm.
- place the grasped mouthwash into blue nylon bag with the right arm.

**Skill list:** Grasp, Pick, Place

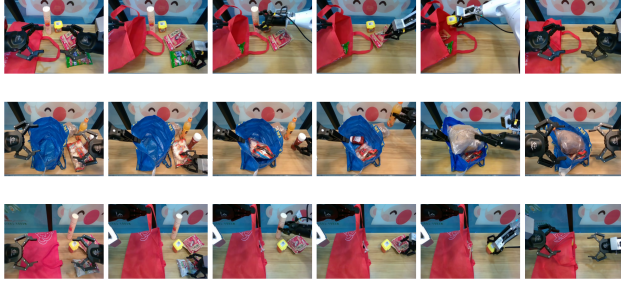


Figure 11. Three representative trajectory examples for Task 372: Packing in the supermarket.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	26.33	32.71	35.21	54.28	52.23	27.49
No Kinematics	24.89	34.52	37.00	56.50	53.97	30.53
MLP + MSE	25.38	28.96	29.50	45.91	41.41	18.54
RoboSegNet	<b>57.79</b>	<b>75.38</b>	<b>78.78</b>	<b>89.01</b>	<b>88.82</b>	<b>85.99</b>
<i>Multi-Task Training</i>						
No Visual	29.14	39.22	44.93	66.50	65.36	47.92
No Kinematics	24.85	28.38	28.38	44.31	35.54	17.01
RoboSegNet	26.95	47.97	56.54	74.11	74.08	56.93

## 8.9. Task 374: Sort laundry and personal care products

This task is performed in Industry scenario using grippers, with typical duration of 12s to 24s and 182 total action slices.

### Task phases:

1. lift the wash and care set in the top-left corner of the material frame with the left arm.
2. place the held wash and care set in the left blue target box.

**Skill list:** Pick, Place

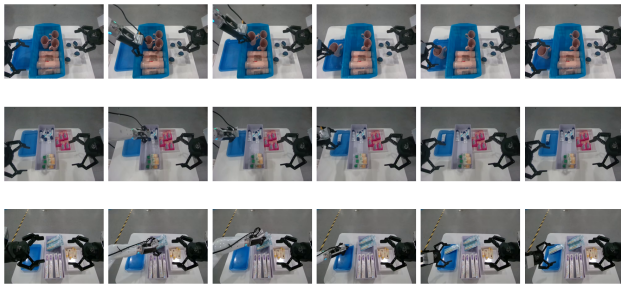


Figure 12. Three representative trajectory examples for Task 374: Sort laundry and personal care products.

## 8.10. Task 375: Brew tea

This task is performed in Home scenario using grippers, with typical duration of 55s to 110s and 3376 total action slices.

### Task phases:

1. lift the right arm with white teapot.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	14.17	18.24	20.50	44.36	<b>44.36</b>	20.63
No Kinematics	<b>14.42</b>	16.11	16.11	35.79	24.06	10.78
MLP + MSE	14.32	16.25	16.25	36.02	24.22	7.99
RoboSegNet	2.97	4.82	7.07	33.27	31.74	<b>21.72</b>
<i>Multi-Task Training</i>						
No Visual	14.32	<b>18.83</b>	<b>20.60</b>	<b>46.45</b>	41.52	20.03
No Kinematics	14.20	16.11	16.11	35.79	22.42	9.14
RoboSegNet	13.60	27.73	34.93	63.68	63.52	48.69

2. pour the water from white teapot into the cup.
3. place the right arm back on the table.
4. lift the tea bag with the left arm.
5. pass the tea bag with both hands.
6. dip the tea bag into the cup and lift it out.
7. dip the tea bag into the cup and lift it.
8. place the tea bag in the cup with the right arm, leaving the tea bag tag outside the cup.

**Skill list:** HandOver, Pick, Place, Pour, dip

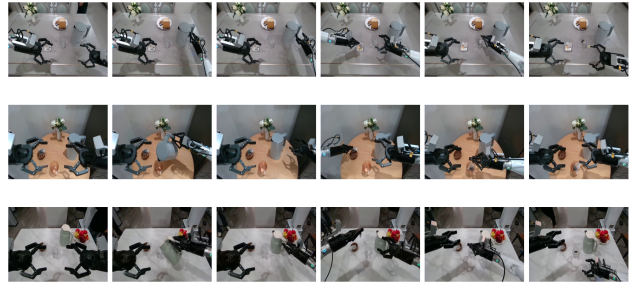


Figure 13. Three representative trajectory examples for Task 375: Brew tea.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	19.27	21.98	24.02	43.50	41.77	16.36
No Kinematics	18.32	20.92	20.93	39.12	30.10	11.34
MLP + MSE	25.81	34.47	38.58	63.96	62.54	47.73
RoboSegNet	<b>42.73</b>	<b>62.76</b>	<b>70.03</b>	<b>87.43</b>	<b>87.19</b>	<b>83.53</b>
<i>Multi-Task Training</i>						
No Visual	18.21	23.78	27.38	49.39	49.17	28.44
No Kinematics	18.36	20.92	20.92	39.11	28.13	12.21
RoboSegNet	16.02	31.63	40.03	65.18	65.08	40.70

## 8.11. Task 376: Sort electronic products

This task is performed in Industry scenario using grippers, with typical duration of 25s to 50s and 2015 total action slices.

### Task phases:

1. retrieve smart charger from bottom left of the material frame with the left arm.

2. pass the smart charger with both hands.
3. place the held smart charger into the right blue target box.
4. retrieve ethernet cable with the right arm from top right corner of the material frame.
5. place the ethernet cable held in the right arm into the blue target box on the right.

**Skill list:** HandOver, Pick, Place

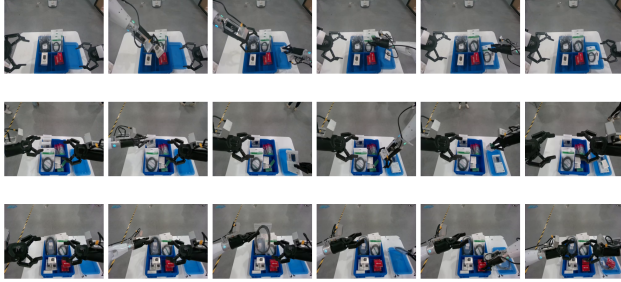


Figure 14. Three representative trajectory examples for Task 376: Sort electronic products.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	19.57	22.94	25.07	48.03	44.83	19.69
No Kinematics	21.39	24.05	24.72	44.75	38.16	15.45
MLP + MSE	18.62	21.36	21.59	39.99	33.20	12.45
RoboSegNet	<b>48.21</b>	<b>67.90</b>	<b>73.45</b>	<b>88.54</b>	<b>88.36</b>	<b>86.32</b>
<i>Multi-Task Training</i>						
No Visual	21.08	28.00	31.91	56.15	55.75	37.49
No Kinematics	18.51	21.12	21.12	39.23	28.45	12.20
RoboSegNet	20.63	38.01	43.61	68.41	68.33	52.18

### 8.12. Task 377: Packing in e-commerce

This task is performed in Industry scenario using grippers, with typical duration of 80s to 160s and 4764 total action slices.

**Task phases:**

1. pick up the air column film outside the carton.
2. place the air column film into the carton.
3. pick up trash bag with right arm.
4. place the held trash bag into the carton.
5. pick up the air column film outside the cardboard box.
6. place the air column film into the carton on trash bag.

**Skill list:** Pick, Place

### 8.13. Task 378: Clear table in the restaurant

This task is performed in Restaurant scenario using grippers, with typical duration of 55s to 111s and 8868 total action slices.

**Task phases:**

1. coordinate arms to place the tray on the counter.
2. pick up the fork on the table.

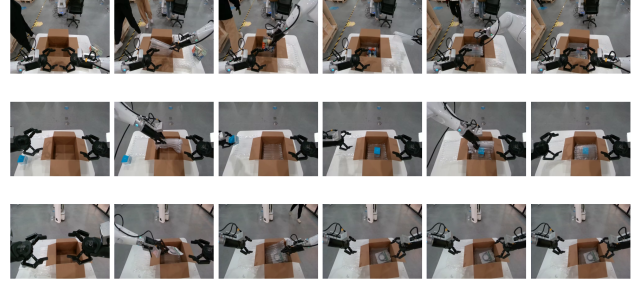


Figure 15. Three representative trajectory examples for Task 377: Packing in e-commerce.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
MLP + MSE	22.13	24.97	25.89	56.90	56.11	38.97
RoboSegNet	<b>59.39</b>	<b>71.62</b>	<b>75.50</b>	<b>92.71</b>	<b>92.58</b>	<b>90.87</b>
<i>Multi-Task Training</i>						
No Visual	10.67	14.30	16.23	43.77	43.45	19.57
No Kinematics	10.47	11.97	11.97	32.97	15.32	6.97
RoboSegNet	23.88	42.16	50.50	81.36	81.31	71.33

3. place the held fork onto the dining plate on the table.
4. pick up the bowl on the table.
5. place the bowl on the plate on the counter.
6. pick up the plate on the table.
7. stack the held plate onto the bowl in the dining tray.
8. pick up the cup on the table.
9. place the held cup onto the tray on the counter.
10. pick up the napkin on the table.
11. place the held napkin into the plate on the table.
12. coordinate arms to lift the tray.

**Skill list:** Pick, Place, Stack

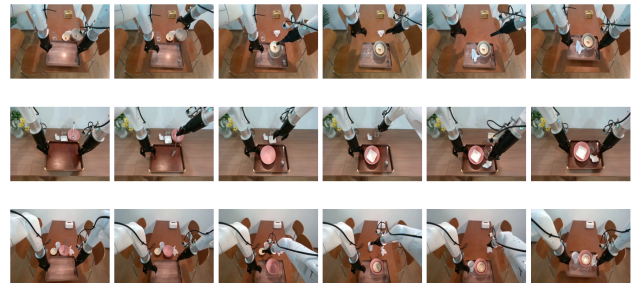


Figure 16. Three representative trajectory examples for Task 378: Clear table in the restaurant.

### 8.14. Task 384: Insert a book into the bookshelf

This task is performed in Home scenario using grippers, with typical duration of 38s to 76s and 3232 total action slices.

**Task phases:**

1. lift the horizontally placed book on the shelf with the right arm.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	30.46	36.51	39.13	56.23	54.64	31.84
No Kinematics	28.21	33.58	34.23	50.37	45.32	22.26
MLP + MSE	47.32	56.17	59.01	75.62	74.68	63.81
RoboSegNet	<b>55.23</b>	<b>72.27</b>	<b>75.47</b>	<b>86.60</b>	<b>86.42</b>	<b>82.51</b>
<i>Multi-Task Training</i>						
No Visual	30.89	41.92	47.66	67.99	67.29	51.78
No Kinematics	26.69	30.46	30.46	45.78	38.14	18.33
RoboSegNet	23.94	45.47	55.32	71.71	71.70	51.04

- push aside the books on the shelf with the left arm.
- insert the book held in the right arm into the gap cleared by the left arm on the shelf.
- push the misplaced book on the shelf to the end with your right arm.

**Skill list:** Insert, Pick, Push

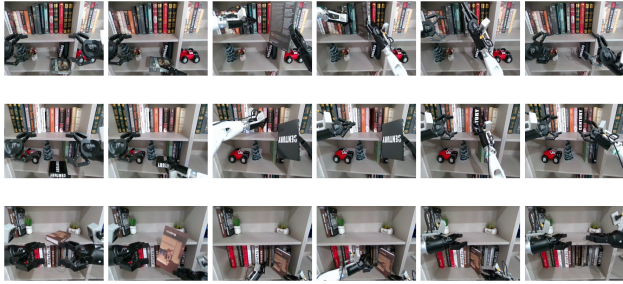


Figure 17. Three representative trajectory examples for Task 384: Insert a book into the bookshelf.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	13.68	16.41	17.81	42.82	39.28	13.84
MLP + MSE	12.63	14.81	15.29	37.00	31.11	9.55
RoboSegNet	<b>39.10</b>	<b>56.72</b>	<b>63.64</b>	<b>87.92</b>	<b>87.73</b>	<b>85.23</b>
<i>Multi-Task Training</i>						
No Visual	13.16	17.92	21.22	50.88	49.97	28.63
No Kinematics	12.50	14.25	14.25	34.52	19.57	8.18
RoboSegNet	15.22	28.66	34.46	66.50	66.43	49.01

### 8.15. Task 385: Pickup in the supermarket produce section

This task is performed in Supermarket scenario using grippers, with typical duration of 30s to 48s and 1284 total action slices.

**Task phases:**

- retrieve potato from the shelf.
- place the held potato into the plastic bag in the shopping cart.
- retrieve potato from the shelf.
- place the held potato into the plastic bag in the shopping cart.

5. retrieve potato from the shelf.

6. place the held potato into the shopping carts plastic bag.

**Skill list:** Pick, Place

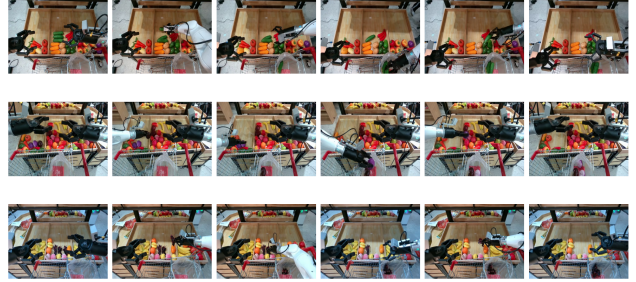


Figure 18. Three representative trajectory examples for Task 385: Pickup in the supermarket produce section.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	23.96	28.89	31.70	52.87	51.15	27.71
No Kinematics	22.04	25.21	25.21	42.09	35.36	13.97
MLP + MSE	21.73	24.42	24.70	41.43	36.32	13.71
RoboSegNet	<b>38.75</b>	<b>59.81</b>	<b>67.89</b>	<b>85.47</b>	<b>85.06</b>	<b>79.86</b>
<i>Multi-Task Training</i>						
No Visual	29.82	41.28	48.04	70.35	69.70	55.95
No Kinematics	22.17	25.21	25.21	42.09	32.75	14.91
RoboSegNet	21.33	40.38	49.06	68.41	68.40	49.60

### 8.16. Task 389: Pickup in the supermarket

This task is performed in Supermarket scenario using grippers, with typical duration of 26s to 52s and 1248 total action slices.

**Task phases:**

- retrieve the fruit and vegetable juice from the cold storage cabinet.
- place the held fruit and vegetable juice into the shopping cart.
- retrieve the packaged yogurt from the cold storage cabinet.
- place the held packaged yogurt into the shopping cart.

**Skill list:** Pick, Place

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	16.07	19.64	21.64	46.10	42.45	18.27
No Kinematics	15.15	17.31	17.31	36.61	25.20	9.12
MLP + MSE	15.10	17.53	17.75	37.35	29.75	9.63
RoboSegNet	<b>56.46</b>	<b>71.32</b>	<b>73.66</b>	<b>89.48</b>	<b>89.46</b>	<b>88.50</b>
<i>Multi-Task Training</i>						
No Visual	22.62	31.50	36.89	65.65	65.08	50.73
No Kinematics	15.17	17.31	17.31	36.61	23.49	9.97
RoboSegNet	16.76	32.02	38.67	67.74	67.47	49.72



Figure 19. Three representative trajectory examples for Task 389: Pickup in the supermarket.

### 8.17. Task 390: Checkout and scan barcode in the supermarket

This task is performed in Supermarket scenario using grippers, with typical duration of 71s to 140s and 8250 total action slices.

#### Task phases:

1. grab the barcode scanner on the counter with the right arm.
2. lift the round bottle yogurt on the cashier counter with the left arm.
3. aim the barcode scanner at the barcode of the held round bottle yogurt with the right arm.
4. place the held round bottle yogurt back on the counter.
5. lift the salty peach meat on the counter with the left arm.
6. scan the barcode of the held salty peach meat with the right arm using the scanner.
7. place the salty peach meat held in the left arm back onto the cashier counter.
8. lift the scouring pad on the counter with the left arm.
9. scan the barcode of the held scouring pad with the right arm using a barcode scanner.
10. place the scouring pad held in the left arm back onto the cashier counter.
11. place the barcode scanner back on the counter.

**Skill list:** Pick, Place, Scan



Figure 20. Three representative trajectory examples for Task 390: Checkout and scan barcode in the supermarket.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
MLP + MSE	26.84	36.48	40.22	60.74	59.90	38.44
RoboSegNet	<b>45.92</b>	<b>66.53</b>	<b>72.09</b>	<b>86.29</b>	<b>85.99</b>	<b>79.62</b>
<i>Multi-Task Training</i>						
No Visual	22.07	29.45	34.06	55.50	54.78	33.91
No Kinematics	23.16	26.44	26.44	42.94	32.57	15.97
RoboSegNet	23.11	44.21	55.16	73.82	73.71	51.18

### 8.18. Task 398: Sort clothes

This task is performed in Industry scenario using grippers, with typical duration of 25s to 50s and 865 total action slices.

#### Task phases:

1. retrieve blue long-sleeve shirt with the right arm from right material frame.
2. pass the blue long-sleeve shirt with both hands.
3. place the blue long-sleeve shirt held in the left arm into the left gray target box.
4. retrieve blue-black casual trousers from left material frame with the left arm.
5. place the blue-black casual trousers held in the left arm into the left gray target box.

**Skill list:** HandOver, Pick, Place

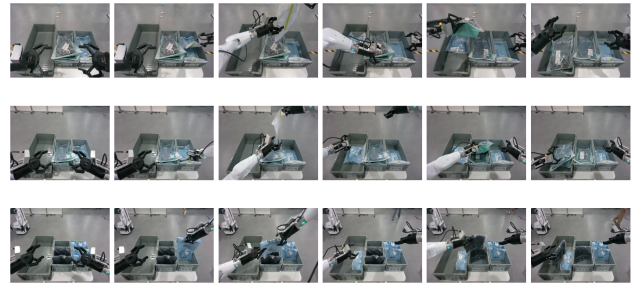


Figure 21. Three representative trajectory examples for Task 398: Sort clothes.

Method	F1@1s (%)	F1@2s (%)	F1@3s (%)	Seg-F1@10% (%)	Seg-F1@25% (%)	Seg-F1@50% (%)
<i>Single-Task Training</i>						
No Visual	17.88	21.13	22.97	47.53	45.85	20.23
No Kinematics	16.57	18.96	18.96	37.75	27.84	9.97
MLP + MSE	16.71	18.12	18.12	36.47	27.12	8.49
RoboSegNet	<b>39.75</b>	<b>60.38</b>	<b>67.00</b>	<b>86.41</b>	<b>86.20</b>	<b>83.53</b>
<i>Multi-Task Training</i>						
No Visual	17.30	23.98	28.45	56.47	55.22	34.20
No Kinematics	16.66	18.96	18.96	37.75	26.02	10.91
RoboSegNet	16.61	31.12	38.84	65.74	65.67	47.00