9th Robotic Grasping & Manipulation Competition Summary

ICRA2024
YOKOHAMA | JAPAN

https://cse.usf.edu/~yusun/rgmc/2024.html

Five Tracks

- Cloth Manipulation Track
- In-Hand Manipulation Track
- Pick in Clutter Track
- Hand-Over Track
- Manufacturing Track

Schedule

<table>
<tr><th>Time</th><th colspan="2">May 14th</th><th colspan="2">May 15th</th><th>May 16th</th></tr>
<tr><td>9:00-10:00</td><td rowspan="9">Cloth Manipulation Track (9:00-18:00)</td><td rowspan="5">Manufacturing Track (9:00-14:00)</td><td rowspan="9">Cloth Manipulation Track (9:00-18:00)</td><td rowspan="4"></td><td rowspan="3">Handover Track (9:00-12:00)</td></tr>
<tr><td>10:00-11:00</td></tr>
<tr><td>11:00-12:00</td></tr>
<tr><td>12:00-13:00</td><td rowspan="2">ICRA Awards Lunch</td></tr>
<tr><td>13:00-14:00</td><td rowspan="4">Picking in Clutter (13:00-17:00)</td></tr>
<tr><td>14:00-15:00</td><td rowspan="3">In-hand Manipulation T. (14:00-17:00)</td><td rowspan="3">RGMC Winners' Presentation</td></tr>
<tr><td>15:00-16:00</td></tr>
<tr><td>16:00-17:00</td></tr>
<tr><td>17:00-18:00</td><td></td><td></td><td></td></tr>
</table>

Sponsors

MathWorks

ASTM INTERNATIONAL

IEEE Robotics & Automation Society

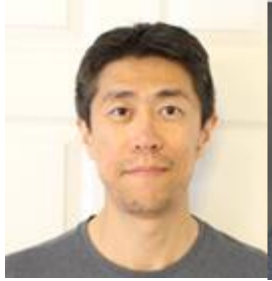
UNIVERSAL ROBOTS

amazon robotics

UMASS LOWELL NERVE CENTER

ICRA2024
YOKOHAMA | JAPAN

Organizers


Yu Sun


Berk Calli


Maximo A. Roa


Jose Nicolas Avendaño


Anastasia Mavrommati

Manufacturing Track


Kenny Kimble

Cloth Manipulation Sub-track


Francis wyffels


Victor-Louis De Gusseme

In-Hand Manipulation Sub-track


Kaiyu Hang


Podshara Chanrungmaneekul


Joshua Grace

Picking in Clutter Sub-Track


Salvatore D'Avella

Human-to-Robot Handovers Sub-track


Andrea Cavallaro


Alessio Xompero

Manufacturing Track

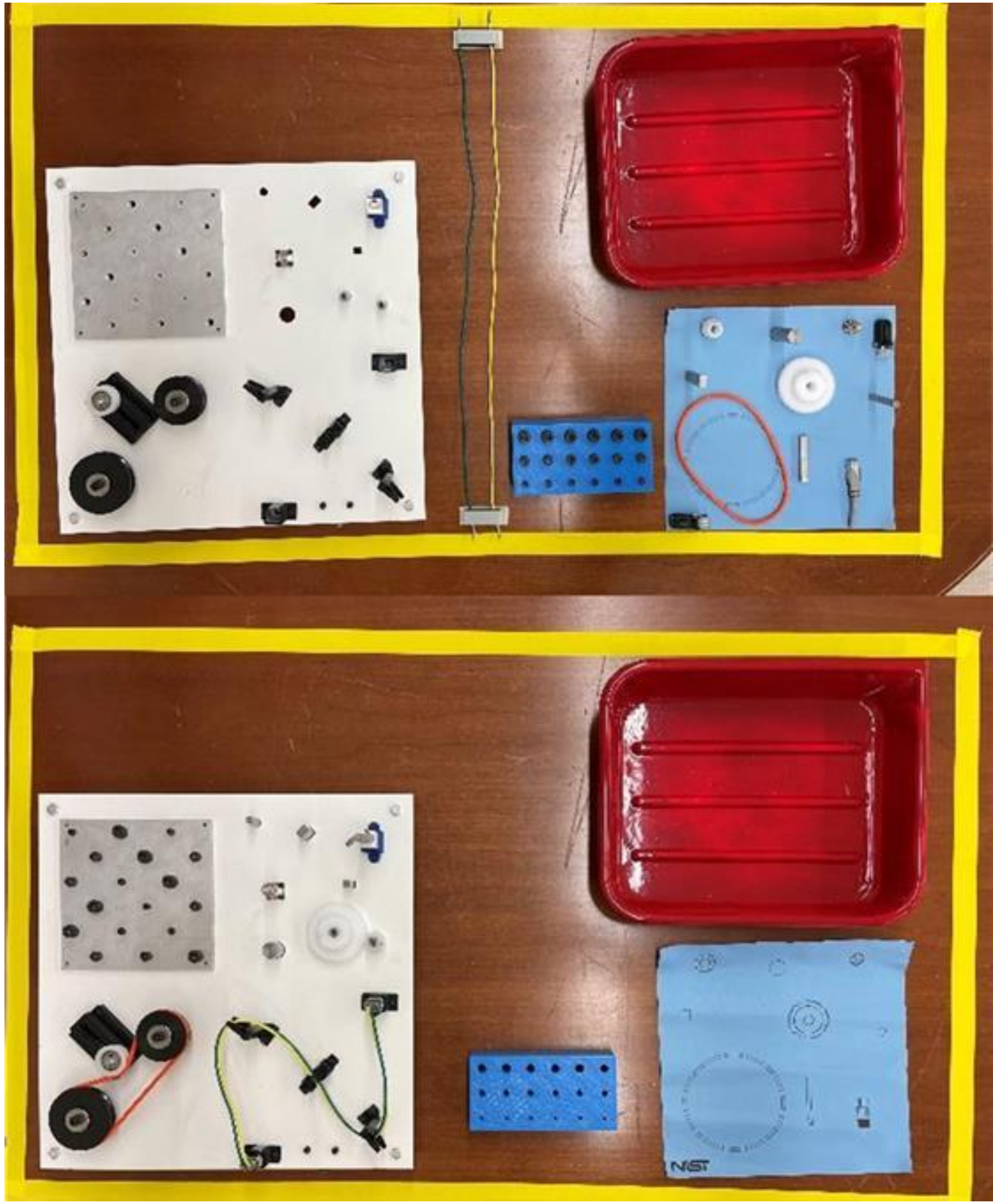
There will be two manufacturing sub-tasks in the competition, disassembly and assembly using a NIST Task Board (NTB). Due to the competition format, up to ten teams will be selected to compete. A pre-competition NTB design will be made available for selected teams to continue development of their systems leading up to the competition. A new NTB with modified assembly positions (very similar to the pre-competition task board), and new parts (identical to all parts used in practice), will be supplied prior to a team's scheduled competition run.

Rules

1. Time bonus points are only available if all parts are successfully assembled onto the applicable quadrant of the task board (maximum points achieved) as shown in Figure 1 and Figure 4(b).
2. No points earned for assembling something in the wrong quadrant other than the 1 point for getting the part to the board (see scoring section for complete point descriptions).
3. Points are only awarded for assembly of a part into its designated location. (e.g. no points for inserting 8mm peg into 16 mm hole)
4. No manual or teleoperated intervention by human operator (e.g., no manual tool changes)
5. No restriction on number of arms, grippers, sensors used
6. Use of hand tools (e.g., wrenches, electric drivers) is allowed provided the robot acquires these tools without human assistance.
7. Perception system markers (e.g., reflectors, AR tags, QR codes) may not be placed on the individual parts to be assembled but may be placed on the board, kit mat, or table.
8. Working area is the area within which the end-effector of the robot can move.
9. A reset is allowed in order to make program changes or repair/secure a task board. During a reset, teams must disassemble all parts from the current quadrant of the task board and reset in the kit area. All accumulated points are reset to zero. The clock continues to run throughout the reset.
10. Scoring is done after the completion of a quadrant.
11. While performing one quadrant of the task board, the other quadrants may be disassembled.

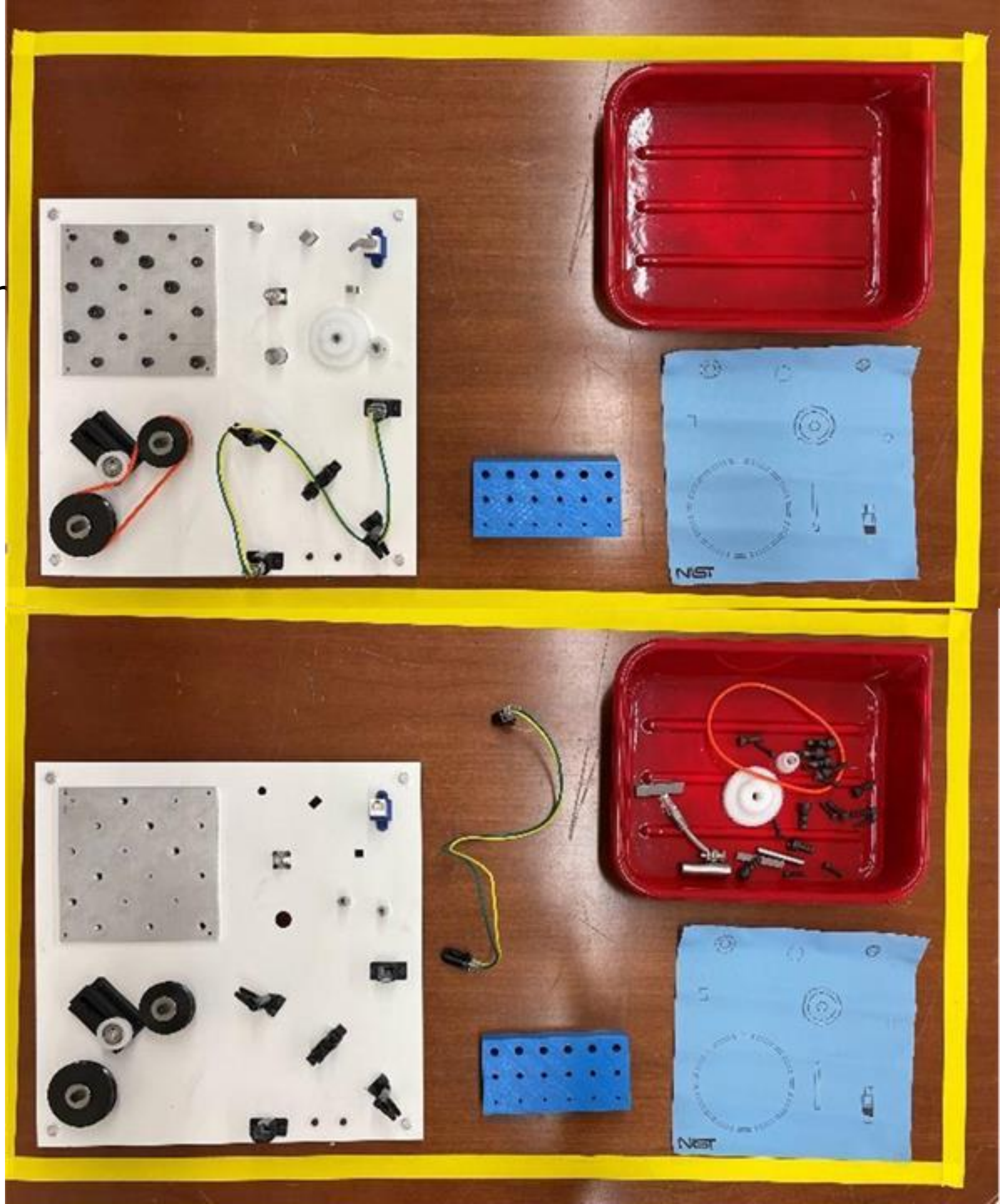
ASSEMBLY: 2hr 10min

- Time breakdown:
- 30 minutes for general programming setup/time
- 5 minutes to set up for quadrant 1 (threading)
- 20 minutes for assembly of quadrant 1
- 5 minutes to set up for quadrant 2 (insertion)
- 20 minutes for assembly of quadrant 2
- 5 minutes to set up for quadrant 3 (belt drive)
- 20 minutes for assembly of quadrant 3
- 5 minutes to set up for quadrant 4 (cable)
- 20 minutes for assembly of quadrant 4

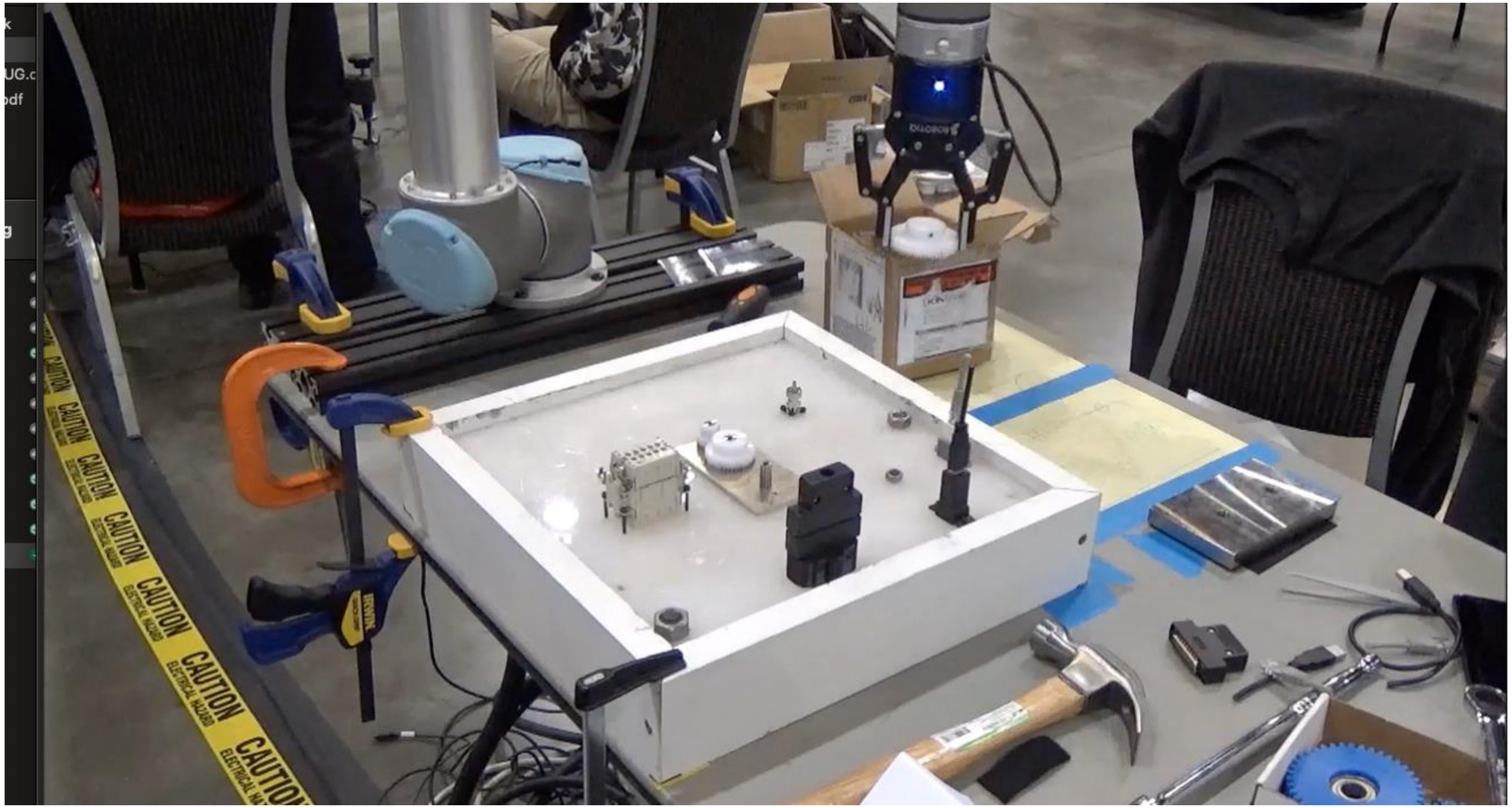


DISASSEMBLY: 1hr 30min

- Time breakdown:
- 30 minutes for general programming setup/tim
- 5 minutes to set up for quadrant 1 (threading)
- 10 minutes for assembly of quadrant 1
- 5 minutes to set up for quadrant 2 (insertion)
- 10 minutes for assembly of quadrant 2
- 5 minutes to set up for quadrant 3 (belt drive)
- 10 minutes for assembly of quadrant 3
- 5 minutes to set up for quadrant 4 (cable)
- 10 minutes for assembly of quadrant 4



Setup



First Place

- Team Name: Tsinghua 3C united; Leader: Sun Fuchun
- ASSEMBLY Scores
  - Threading: 0
  - Insertions: 6
  - Belt Tensioning: 25
  - Wire Harness: 0
  - Total: 31
- DISASSEMBLY Scores
  - Threading: 2
  - Insertions: 6
  - Belt tensioning: 15
  - Wire Harness: 25
  - Total: 48
- Total Score for assembly and disassembly: 79

Second Place

- Team Name: JAKS, Leader: Tokuo Tsuji
- ASSEMBLY Score
  - Threading: 7
  - Insertions: 7
  - Belt Tensioning: 13
  - Wire Harness: 3
  - Total: 30
- DISASSEMBLY Score
  - Threading: 2
  - Insertions: 9
  - Belt tensioning: 2
  - Wire Harness: 22
  - Total: 35
- Total Score for assembly and disassembly: 65

Third Place

- Team Name: AI & Robot; Leader: Haotian Tang
- ASSEMBLY
  - Threading: 21
  - Insertions: 0
  - Belt Tensioning: 1
  - Wire Harness: 0
  - Total: 22
- DISASSEMBLY
  - Threading: 18
  - Insertions: 15
  - Belt tensioning: 6
  - Wire Harness: 0
  - Total: 39
- Total Score for assembly and disassembly: 61

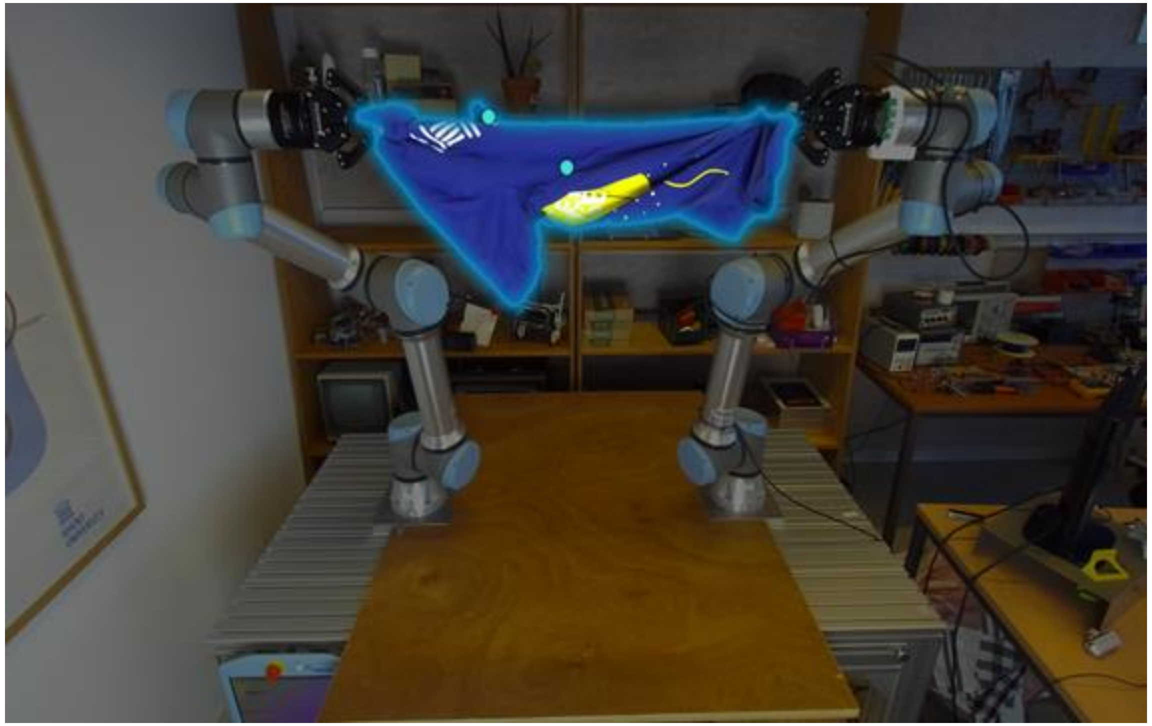
Winning Team Photo

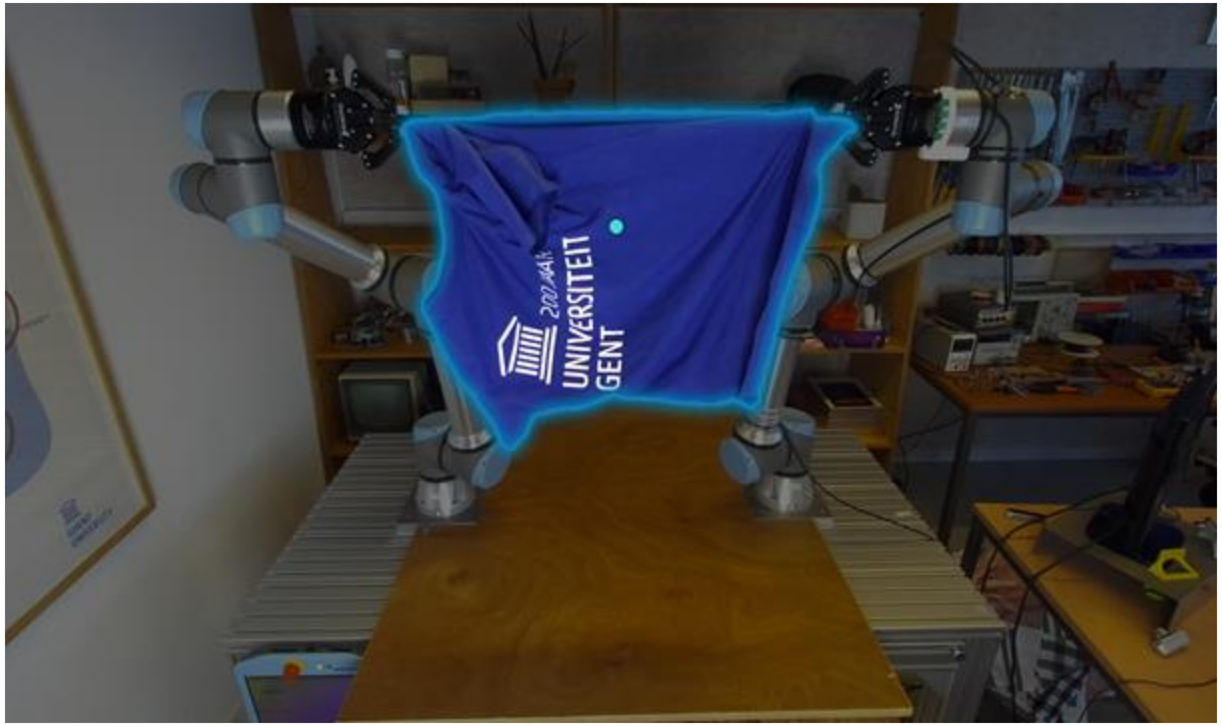
Cloth Manipulation

Essential Skill Track #1

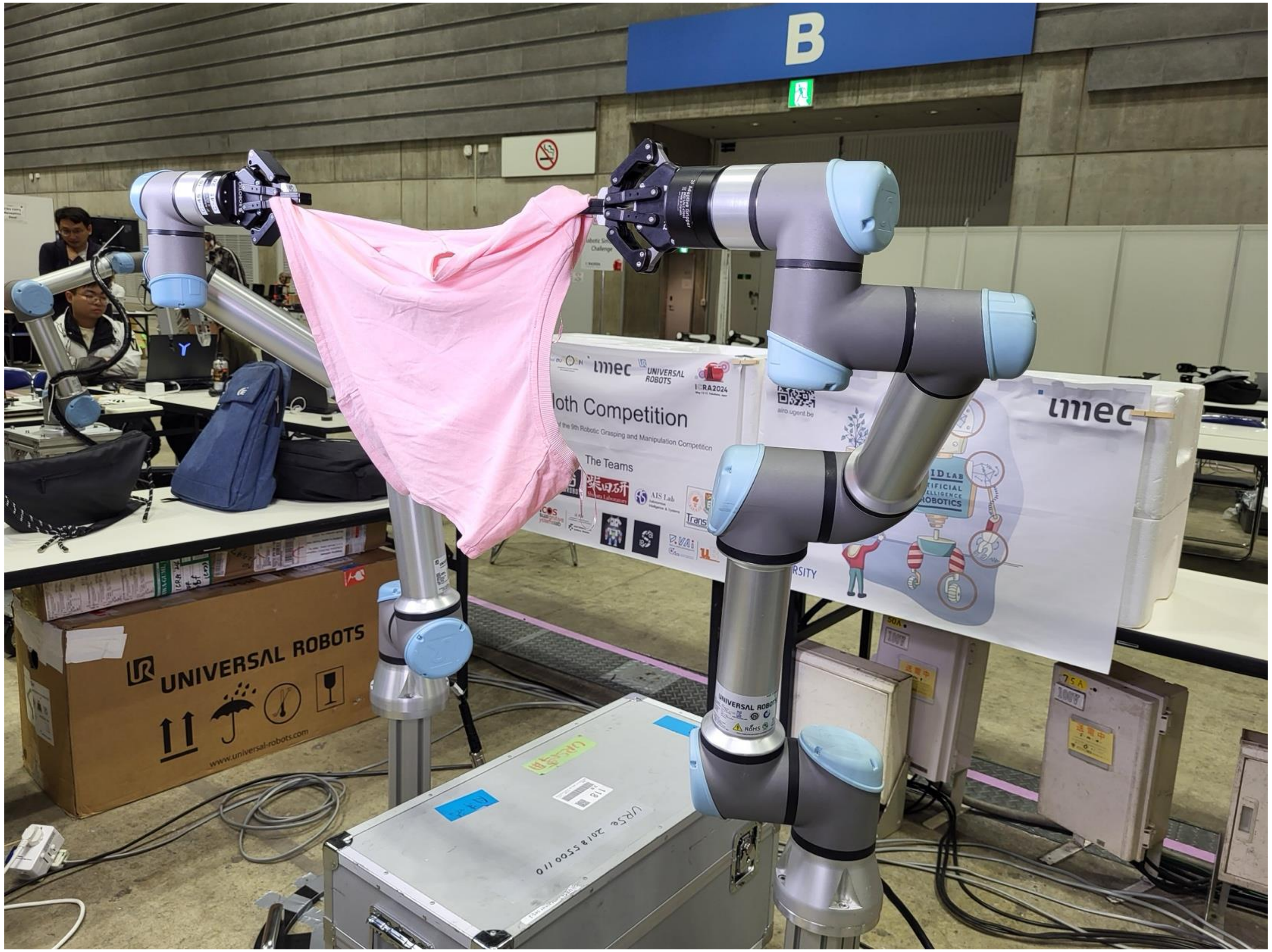
The Challenge: Grasp clothes in a way that leads to the best unfolding

Evaluation: Surface area after grasping and stretching





Setup


B
imec
UNIVERSAL ROBOTS
loth Competition
The Teams
imec
UNIVERSAL ROBOTS
www.universal-robots.com

Results

Rank	Team Name	Score (Average Coverage)
🥇 1st	AIR-JNU	0.60
🥈 2nd	Team Ljubljana	0.57
🥉 3rd	Ewha Glab	0.55
4th	SCUT-ROBOT	0.53
5th	Team Greater Bay	0.53
6th	Samsung Research China - Beijing	0.48
7th	Shibata Lab	0.46
8th	AI&ROBOT LAB	0.45
9th	UOS-Robotics	0.39
10th	AIS Shinshu	0.37
11th	3C1S	0.35

Winning Team Photo

In-Hand Manipulation

Essential Skill Track #2

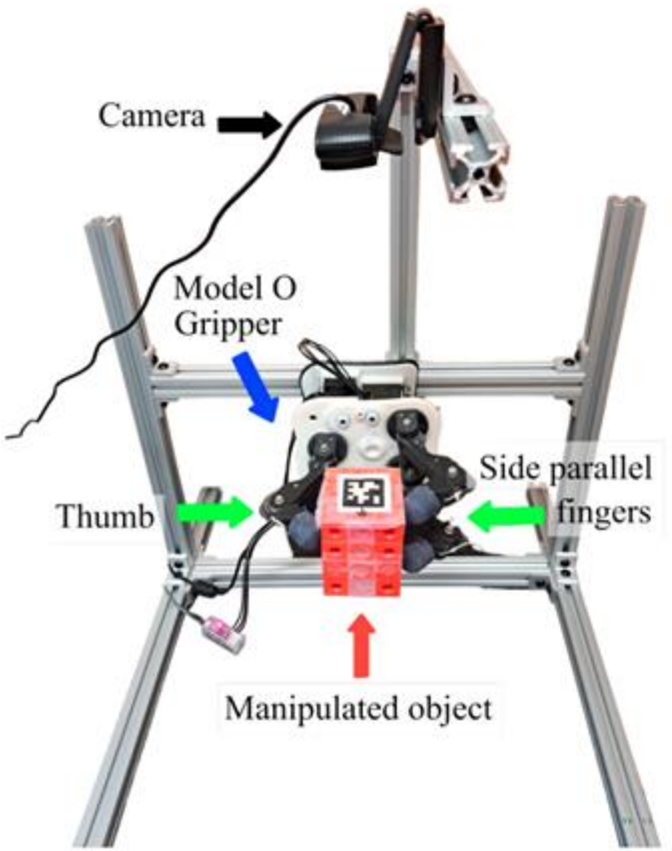
Goal

The goal is to provide an international venue for in-hand manipulation researchers to compete and gain more insights on the opportunities and challenges of:

1. Different robot hand designs
2. Manipulation planning algorithms
3. Perception strategies for in-hand manipulation
4. Robustness of various manipulation controllers
5. Data-driven manipulation frameworks

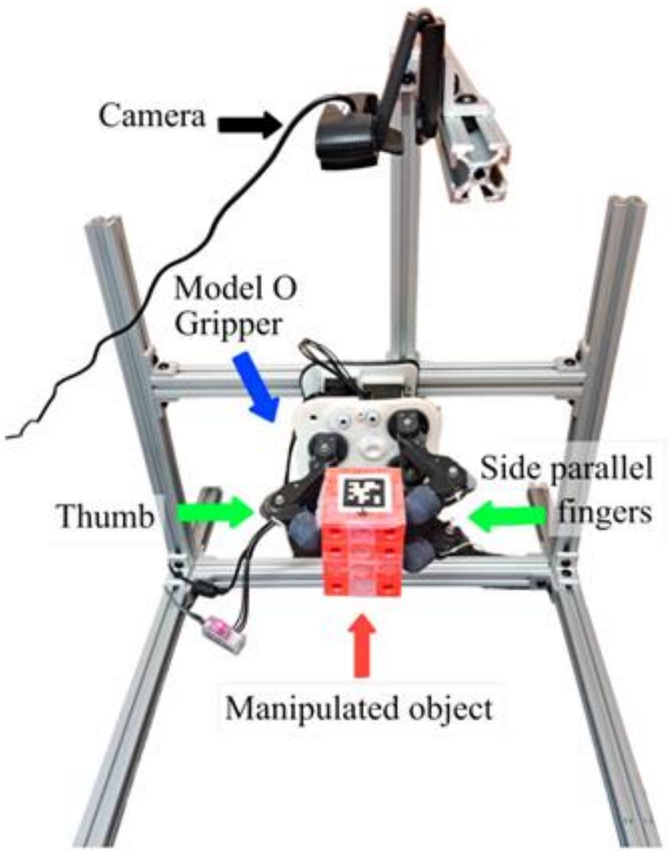
Workspace Setup

- Every team will bring their own robot hand for the competition
- The hand can be commercial models, open source models, or models designed the participating teams
- The robot hand will be installed on a stationary mount designed by the teams
- The teams can change the hand mounting poses as needed for different tasks
- Only hand actions are allowed, e.g., even if a team decides to mount their robot hand on a robot arm, the robot arm has to be stationary during the task executions.



Sensors

- Every team is required to use at least one camera that can support the use of apriltag_ros
- As exemplified in the figure, the object being manipulated will be tracked by this camera through an apriltag
- The scoring of the competition is heavily relying on the apriltags, and the manipulation goals are specified in the camera's view/frame based on the readings of apriltags. Therefore, **we do not require hand-camera calibration**
- There is no restriction on the sensors used by teams. Vision sensors, tactile sensors, force/torque sensors, etc., are all welcome to be included in the setups



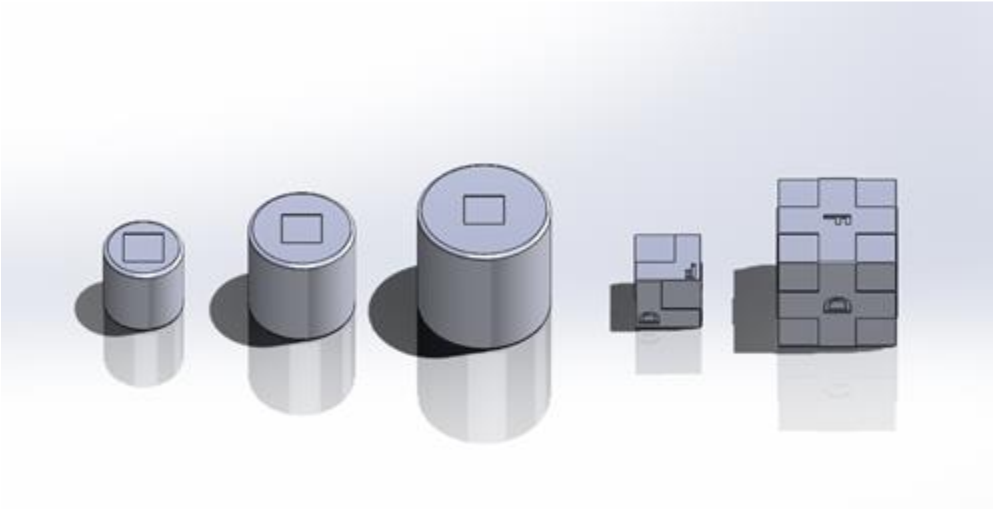
Define Robot Hands?

- There is no formal definition of robot hands in any literature or in our community

- We use "Common Sense" to evaluate whether or not a design or setup is considered a robot hand
  - Hardware qualification required 2 months before the competition
  - In-hand CNC machines, in-hand gantry robots, etc., are not qualified as valid solutions

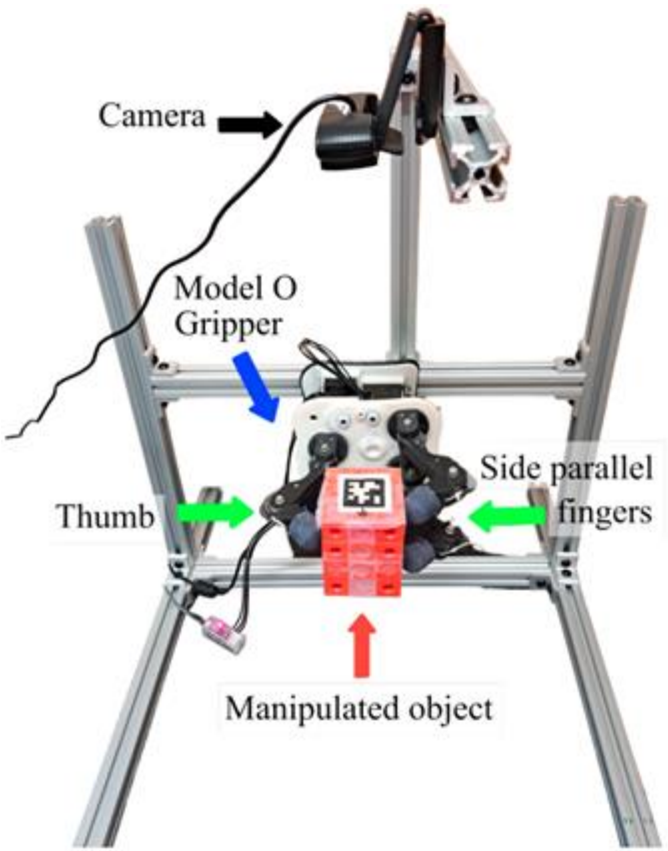
Competition Objects

- CAD model provided for 3D printing so that every team has access
- 3 objects used in the competition with size options
  - Cylinders (Task A): 60mm, 80mm, 100mm in diameters and 80mm tall
  - Cubes (Task B): 50*50*50 mm^3 and 90*90*90 mm^3
  - Novel object (Task A): unknown before the competition



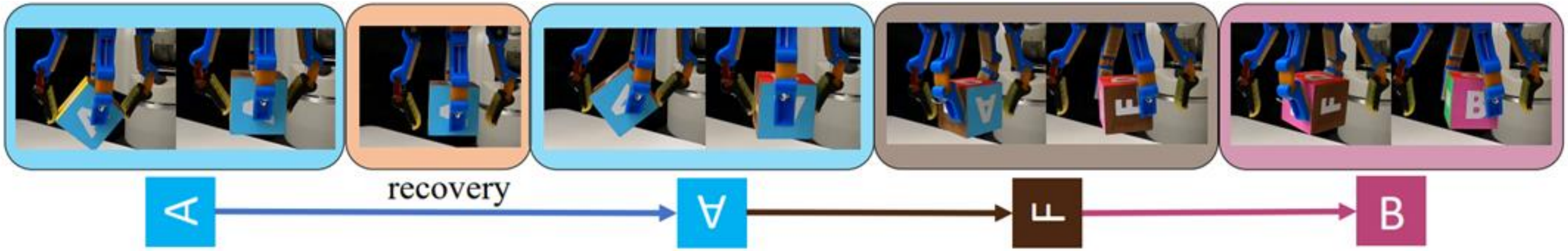
Task A: Object Position Control for Waypoints Tracking

- Object is grasped and the apriltag is tracked
- The goal positions for the object are specified in the camera's frame
- Once the grasp is initially stabilized, goals are given as a sequence of 10 waypoints within a workspace of size [-2.5cm, 2.5cm] * [-2.5cm, 2.5cm] * [-2.5cm, 2.5cm], centered at the object's initial position
- The goal positions should be reached one by one without skipping



Task B: Object Re-orientation

- The goal orientations for the object are specified in the camera's frame
- Once the grasp is initially stabilized, goals are given as a sequence of 10 waypoints of "A, B, C, D, E, F" letters, which mark the face of the object
- The goal positions should be reached one by one without skipping



Competition Scoring

- The scoring of the both Tasks A and B are fully automatic and focused on two metrics
  - Accuracy
  - Speed

- An open source evaluator is provided on github: https://github.com/Rice-RobotPI-Lab/RGMC_In-Hand_Manipulation_2024

- At the competition, the judge will bring a computer to join the team's ROS network and run the auto-evaluator

Teams

- 11 teams signed up and 6 teams are qualified
  - Tsinghua University
  - MIT
  - ETH
  - University College Dublin
  - Tsinghua Shenzhen International Graduate School
  - Chinese Academy of Sciences & Yantai Institute of Technology

- All qualified teams had good partial solutions to some tasks, and 2 Teams developed successful solutions to complete all tasks
  - Major challenge observed: sim-to-real transfer

Winner: XL Team, Tsinghua University

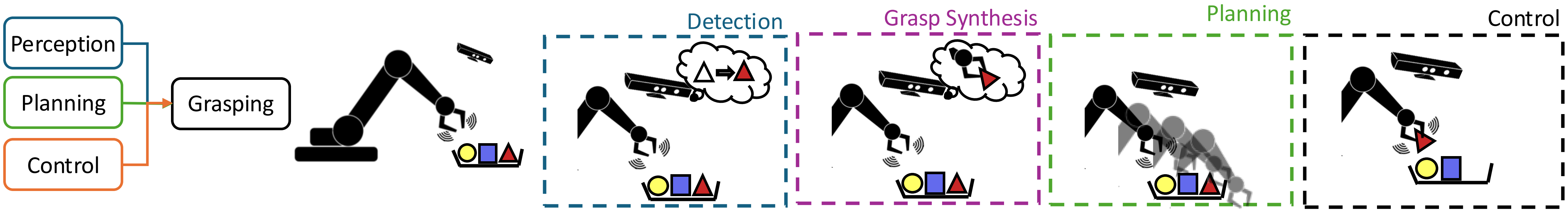
- Task A: Accuracy 0.54mm/0.62mm
- Task B: 81.18 seconds for all 10 target orientations

Picking in Clutter

Essential Skill Track #3

The Benchmark [1/4]

- Evaluation of *Perception, Planning, and Control* aspects of Robotic Grasping and Manipulation systems



- Modular design:

  organized in **Stages** with one or more *intermediate phases* and a *final phase test*.
  Stages are meant to represent an industrial-relevant task, which is identified by the final phase test, while the intermediate phases of each Stage aim at evaluating a specific sub-problem of the manipulation task
    - Stage 1: Pick and Place of **non sequential** objects in cluttered environments
    - Stage 2: Pick and Place of **non-sequential unknown** objects in cluttered environments
    - Stage 3: Pick and Place of **sequential** objects in cluttered environments

[Ref] D'Avella, S., Bianchi, M., Sundaram, A. M., Avizzano, C. A., Roa, M. A., & Tripicchio, P. (2023).
The Cluttered Environment Picking Benchmark (CEPB) for Advanced Warehouse Automation: Evaluating the Perception, Planning, Control, and Grasping of Manipulation Systems. IEEE Robotics & Automation Magazine.

The Benchmark [2/4]

- **40 objects**
  from YCB, ACRV-APC, and T-LESS
  - Standards rigid objects with square or cylindrical shapes
  - More articulated rigid objects
  - Soft, deformable, and fragile
  - Transparent
  - Objects with complex dynamics
  - Very small or very big
  - Heavy

object	original dataset	generalization	difficulty				
			parallel-jaw	suction	soft grippers (i.e., UJG)	vision	total
1 Cheez-it Cracker box	YCB obj#1		1	1	3	1	1,5
2 French's Mustard bottle	YCB obj#9		3	2	3	1	2,25
3 Tomato Soup can	YCB obj#10		2	1	3	1	1,75
4 Scissors	YCB obj#35		3	2	1	2	2
5 Foam Brick	YCB obj#57		1	3	1	1	1,5
6 Small Clamp	YCB obj#46		3	3	2	2	2,5
7 Starkist Tuna Fish can	YCB obj#7		1	1	1	2	1,25
8 Plastic banana	YCB obj#11		2	2	1	1	1,5
9 Meat can	YCB obj#5		1	2	3	1	1,75
10 Mug	YCB obj#31		2	2	1	1	1,5
11 Padlock	YCB obj#38	X	2	3	1	2	2
12 Baseball	YCB obj#51	X	3	2	3	1	2,25
13 Bowl	YCB obj#25	X	3	1	3	1	2
14 Sleeve	T-LESS obj#13	X	1	2	1	2	1,5
15 Coca-cola bottle half	CEPB	X	3	3	3	3	3
16 ICRA duckie	APC/ACRV obj#1	X	2	3	2	1	2
17 Elmers school glue	APC/ACRV obj#20	X	1	1	2	1	1,25
18 Dice	YCB obj#58	X	3	1	1	2	1,75
19 Eggs plush puppies	APC/ACRV obj#13	X	2	3	3	1	2,25
20 Scotch duct tape	APC/ACRV obj#16	X	1	2	2	1	1,5
21 Haribo golden bears	CEPB		3	1	3	1	2
22 Plint Board *	T-LESS obj#14		1	2	1	3	1,75
23 Flat Screwdriver	YCB obj#43		3	3	2	1	2,25
24 Clamping Plate *	T-LESS obj#15		1	2	1	3	1,75
25 Cocal-cola bottle full	CEPB		2	2	3	2	2,25
26 Kong duck dog toy	APC		2	3	3	1	2,25
27 Spoon	YCB obj#27		3	3	2	2	2,5
28 Plastic strawberry	YCB obj#12		3	3	2	1	2,25
29 Paper towels	CEPB		3	3	3	1	2,5
30 Stabilo OHPen	CEPB		2	1	3	2	2
31 Plastic white cup	APC/ACRV obj#30		3	3	3	2	2,75
32 Wine glass	YCB obj#30		3	3	2	3	2,75
33 Key	YCB obj#38		3	3	2	2	2,5
34 Nail	YCB obj#40		3	3	2	3	2,75
35 Adjustable Wrench	YCB obj#44		3	3	3	1	2,5
36 t-shirt	YCB obj#70		2	3	3	2	2,5
37 Rolodex jumbo pencil cup	APC/ACRV obj#3		2	2	2	3	2,25
38 Glove	APC/ACRV obj#22		2	3	3	2	2,25
39 Laugh Out Loud Joke Book	APC		3	1	2	1	1,75
40 Pringles Chips can	YCB obj#8		1	1	3	1	1,5
41 Timer	YCB obj#71						
42 Clear Box	YCB obj#63						
43 Clear Box	IKEA obj#SAMLA(301.029.74)						

- Difficulties have been assigned through a **consensus protocol** disseminating questionnaires among colleagues with different levels of experience. The total level of difficulty takes into account the **gripper typology** among parallel-jaw, suction, and soft grippers, and the **vision**.
- The objects are divided into **4 subsets** of 10 items each with an increasing level of difficulty.
  - It is worth noticing that given a subset, the mean difficulty is the same for every gripper typology in order to not privilege one type of gripper over another.

[Ref] D'Avella, S., Bianchi, M., Sundaram, A. M., Avizzano, C. A., Roa, M. A., & Tripicchio, P. (2023).
The Cluttered Environment Picking Benchmark (CEPB) for Advanced Warehouse Automation: Evaluating the Perception, Planning, Control, and Grasping of Manipulation Systems. IEEE Robotics & Automation Magazine.

The Benchmark [3/4]

Photorealistic companion dataset:

- Realized in **Unity** with the **Flex physical simulator** that is able to reproduce soft and deformable objects or objects with complex dynamics like bottles filled with liquids
- **1.5 M images** with
  - three different camera positions
  - three light conditions
  - multiple High Dynamic Range Imaging (HDRI) maps for domain randomization purposes
  - high level of clutter
- YAML file for each scene containing
  - 2D and 3D bbox coordinate
  - translation and orientation
  - visibility percentage of each object.

[Ref] D'Avella, S., Bianchi, M., Sundaram, A. M., Avizzano, C. A., Roa, M. A., & Tripicchio, P. (2023). **The Cluttered Environment Picking Benchmark (CEPB) for Advanced Warehouse Automation: Evaluating the Perception, Planning, Control, and Grasping of Manipulation Systems**. IEEE Robotics & Automation Magazine.

The Benchmark [4/4]

Evaluation metric:

1. average of the complexity of the object successfully grasped
2. taking into account the difficulty of the clutter γ
3. normalized by the spent time

40s x 2N, where N = 20 -> ≈ 26 min

3) $$s_F(t) = \frac{s_F \times (max\_time - t)}{max\_time},$$

2) $$s_F = s_I \times (\sum_{i=1}^{N} d_i + \gamma)$$

1) $$s_I = \sum_{i=1}^{N} \frac{d_i^*}{d_i}$$

Algorithm 1: Algorithm used to compute the clutter percentage of the starting scene and the scene difficulty

Data: 3D bounding box for each object $cuboid$, objects difficulties, $difficulties$

Result: γ, scene_difficulty

```
objects = objects used in image I;
cuboids = 8 x, y, z coordinates of the objects in I;
polygons = 2D polygons derived from cuboids;
for each obj in objects do
    cuboid_obj = pose corresponding to obj;
    polygon_obj = 2D polygon corresponding to obj;
    max_depth_obj = max(cuboid_obj(z));
    compute the union union_other of the other
      polygons for which
      max_depth_obj ≥ max_depth_other;
    compute the intersection occlusion_obj between
      polygon_obj and union_other;
    occluded_area_obj = area(occlusion_obj) / area(polygon_obj);
    γ_obj = sigmoid(occluded_area_obj) × 0.5;
end
```

$$\gamma = \frac{\sum_{obj \in objects} \gamma_{obj}}{\#objects} \times 100;$$

$$scene\_difficulty = \sum_{obj \in objects} difficulty_{obj} \times (1 + \gamma);$$

[Ref] D'Avella, S., Bianchi, M., Sundaram, A. M., Avizzano, C. A., Roa, M. A., & Tripicchio, P. (2023).
The Cluttered Environment Picking Benchmark (CEPB) for Advanced Warehouse Automation: Evaluating the Perception, Planning, Control, and Grasping of Manipulation Systems. IEEE Robotics & Automation Magazine.

ICRA 2024 Adaptation

Objective:

Bin-picking of **known and novel objects** in cluttered scenes
≈ final phase of Stage 2 (of the Benchmarking paper - slide 3)
on **Subset 1 + Subset 4**
where the 4 non YCB objects are announced 1 hour before the start of the competition to test the generalization capabilities of the system

- The score is computed by applying the previous evaluation metric

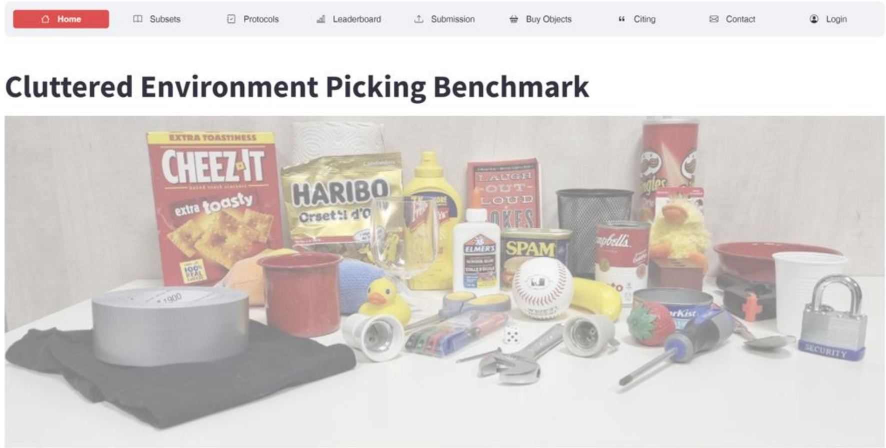
$$s_F(t) = \frac{s_F \times (max\_time - t)}{max\_time},$$

- Results are shown on the benchmark webpage

http://cepbbenchmark.eu/

	Subset 1	Subset 4
1	Cheez-it Cracker box	Plastic white cup
2	Tomato Soup can	Wine glass
3	Plastic banana	Key
4	Starkist Tuna Fish can	Adjustable Wrench
5	Scissors	Dice
6	Foam brick	Rolodex jumbo pencil cup
7	Meat can	Nail
8	Small Clamp	Gloves
9	Mug	Laugh Out Loud Joke Book
10	French's Mustard bottle	Pringles Chips can

⇒ non YCB objects ⇒ novel objects

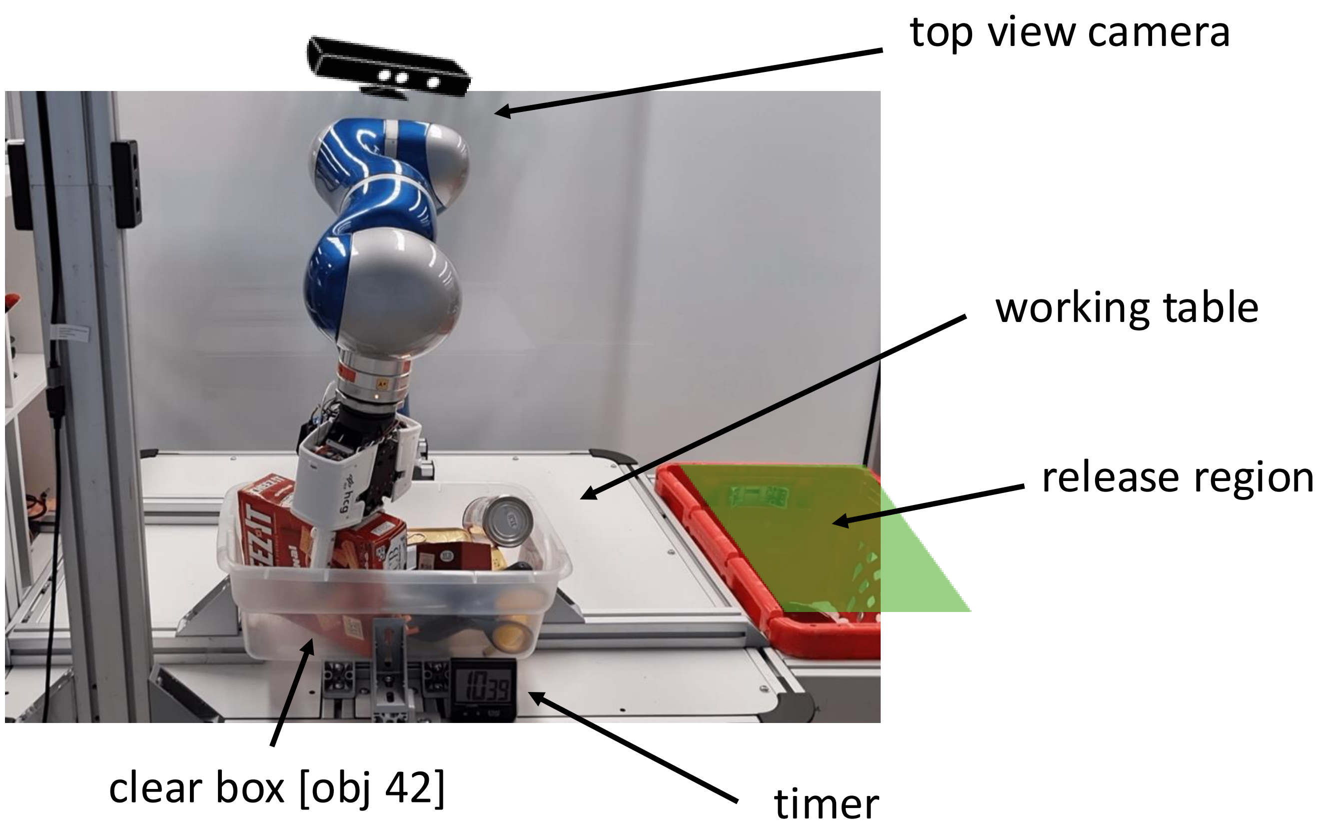


General Rules

- Maximum 5 members per team
- Teams have to bring and mount their own setup:
  - robotic arm & end effector
  - laptop/workstation
  - cameras
  - sensors
  - objects

    Those who did not have YCB objects could purchase their own local objects (validated by the committee).
- Teams have at least half a day for mounting and testing their setup on-site before the start of the competition

Setup



- Users can decide their own setup according to the properties of their grasping system.
- It is important to declare in advance the release region, which cannot be changed afterward.
- No human intervention or teleoperation is possible
- If an object falls out of the table, it cannot be reintroduced.

Protocol [1/2]

1. A team member has to shake the objects in the smaller clear box per subset (Subset 1 and Subset 4 separately) and throw them in the bigger clear box, starting from Subset 1.
2. A team member has to put a clear box in the designed release region. It cannot be moved afterward.
3. The judge has to activate the timer announcing "3, 2, 1, go", so that a team member can activate the system. After that, human intervention is no longer possible! A top-view picture of the starting configuration of the objects must be taken for evaluation purposes.
4. The judge has to take note of the successfully grasped objects (possibly taking note of the order). An object is successfully grasped if it is completely placed in the release region. If the object has some parts outside the smaller clear box, please take note of this. For example, the T-shirt can have some parts outside the box.
5. Exit conditions that can trigger the end of the trial/stop of the timer:
    a) The bigger clear box has been emptied after the last pick and place movement.
    b) The system is unable to proceed further since it is stacked failing to grasp multiple times (i.e., 5 times) the exact same object, or in general, it does consecutive fail grasps of different objects (i.e., 15 times). This choice is up to the judge.
    c) Max time is reached = 26 minutes per trial.
    d) The system decides to quit earlier. To choose such an option, the team must communicate to the judge the quit pose of the robot before the start of the competition. The quit pose must be different from the home configuration. An example is the robot looking up to the ceiling.
6. If one of the previous conditions holds, the judge stops the clock, takes note of it, and double checks the grasped objects.
7. The loop repeats three times.

Protocol [2/2]

- Hardware faults (e.g., power drop, cables disconnections) that stop the system at the beginning of the trial or during the trial can be tolerated, resetting the timer and repeating the trial. Software faults should not be repeated, but the final decision is up to the judge depending on the situation.

- If the system blocks due to collision checking the team is allowed to re-activate it through the robot controller. The timer must not be paused during this recovery operation.

- **Judges**:
    - One main competition organizer (Berk)
    - Main sub-track organizer (Salvatore)
    - 3 people from Amazon Robotics (Kapil D. Katyal's group)

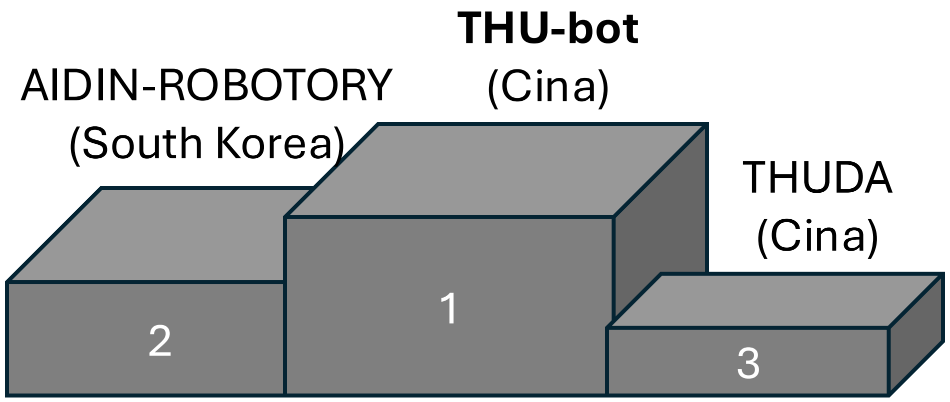
Teams

1. **ICR USTC**
   Zhen Kan, Zhangli Zhou, Xiangcheng Liu, Tangyu Qian , Yinxiao Tian
   from University of Science and Technology of China – Intelligent Control and Robotics
2. **THUDA Team**
   Shihefeng Wang, Xiang Li, Weiduo Gong,Yingyue Li, Junqi Ge
   from Tsinghua University, Department of Automation
3. **THU-bot Team**
   Xueqian Wang, Shoujie Li, Xiankun Zhu, Yucheng Xin, Xianru Tian
   from Tsinghua University, Shenzhen International Graduate School
4. **AIDIN-ROBOTURY Team**
   Yeong Gwang Son, Seunghwan Um, Bui Tat Hieu, Ho Sang Jung, Hyouk Ryeol Choi
   from Sungkyunkwan University and AIDIN ROBOTICS Inc., Anayang, South Korea
5. **TCS Smart Machines Team**
   Chandan Kumar Singh, Devesh Kumar, Vipul Sanap, Mayank Khandelwal, Aman Srivastava
   from Smart Machines – TCS Research Noida

Two other teams withdrawn last minute.

Winners

Team	Final Score	Trial Scores 1st - 2nd - 3rd	Picked Objects 1st - 2nd - 3rd (trial)	Elapsed Time 1st - 2nd - 3rd (trial)
THU-bot	1,17	1,03 - 1,11 - 1,37	16 - 19 - 20	10:00 - 11:03 - 8:35
AIDIN-ROBOTURY	0,94	1,57 - 0,38 - 0,85	20 - 8 - 20	6:25 - 18:31 - 15:58
THUDA	0,78	0,73 - 0,93 - 0,70	14 - 17 - 10	12:46 - 12:12 - 6:45
TCS Smart Machines	0,36	0,54 - 0,30 - 0,26	2 - 9 - 7	11:00 - 26:00 - 26:00
ICR USTC	0,053	0,14 - 0,02 - 0,02	11 - 6 - 6	14:14 - 11:02 - 14:01



Comments:

- TCS Smart Machines and ICR USTC arrived late for the dry run and had less time to prepare and test their setup on-site.
- AIDIN-ROBOTURY had hardware problems with the gripper: a small bolt fixing the linear actuator and the slider was loosened, causing the suction cup to not work in the second and third trials, wasting some time and stopping the second trial

Challenges of on-site competition

- Bringing the system and setup up it on-site has been not easy for some teams.
- Hardware problems to custom realized hardware occurred during the competition.
- Illumination of the environment was different from laboratory conditions for all the team and affected the perception pipeline. However, the improvisation and adaptability skills of the teams made things work.

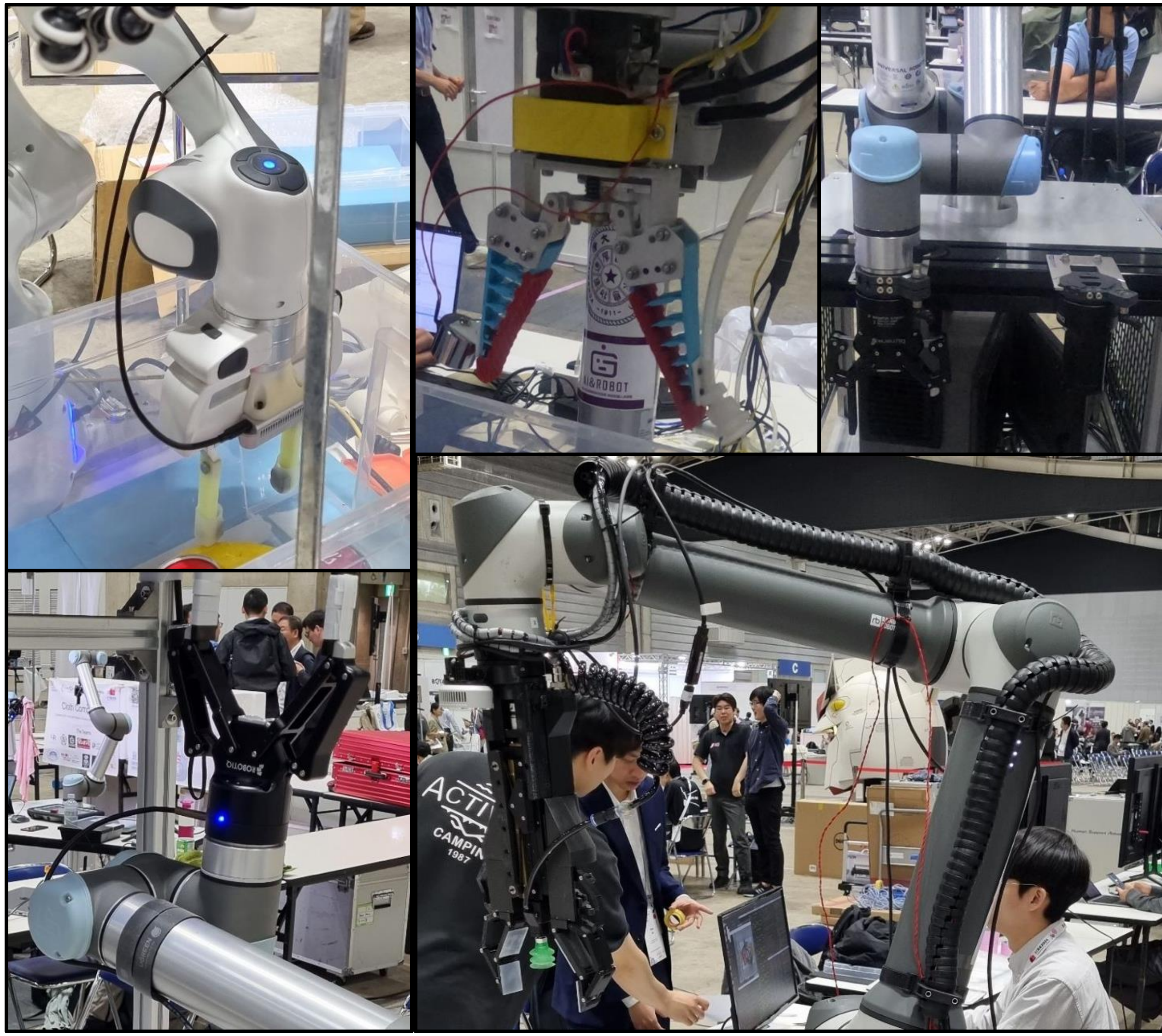
Some considerations and pictures [1/2]

- Teams who used a hybrid gripper, i.e., a gripper that is able to provide multiple grasping modalities, or a multimodal grasping solution, on average performed better as they were able to handle the different properties of the objects and the object configurations in the clutter. Indeed, teams with a traditional gripper (parallel jaw) had difficulties grasping objects in a specific pose and location in the bin.

- Most of the teams had problems grasping very small objects like nails, keys, or tricky objects like the magazines. In general, the introduction of the novel objects caused some troubles not for the generalization of the vision part but rather to the final grasp (gripper-object interaction).

- Another challenge for many teams was the grasping of the objects in the corner or near the wall of the bin.

- Indeed, a lot of hardware stops for collision detection were registered.



Some considerations and pictures [2/2]

- Some grippers were designed for power grasp that helped bring a successful grasp even with a non optimal grasp but damaged or broke some fragile objects like the plastic cup or the Cheez-it cracker box.

- A couple of teams implemented some decluttered actions that helped to disentangle some tricky configurations. Others, instead, developed a logic to understand if an object has been correctly grasped in order to not waste time performing the release movement uselessly.

- Several graspings of two objects at the same time, which was not allowed, were registered. In addition, sometimes happened that the grasped object was different from the target one.

See you next year. Be ready for more complex challenges!

Salvatore D'Avella
Assistant Professor
Sant'Anna School of Advanced Studies
Mechanical Intelligence Institute
Department of Excellence in Robotics and AI
Pisa, Italy

Human-to-Robot Handovers

Essential skills sub-track 4

Prof. Andrea Cavallaro
Ecole Polytechnique Fédérale de Lausanne
Switzerland

Dr. Alessio Xompero
Queen Mary University of London
United Kingdom

https://corsmal.eecs.qmul.ac.uk/rgm24icra/index.html

CORSMAL
Collaborative object recognition, shared manipulation and learning

The setup

CORSMAL
Collaborative object recognition, shared manipulation and learning

Delivery area

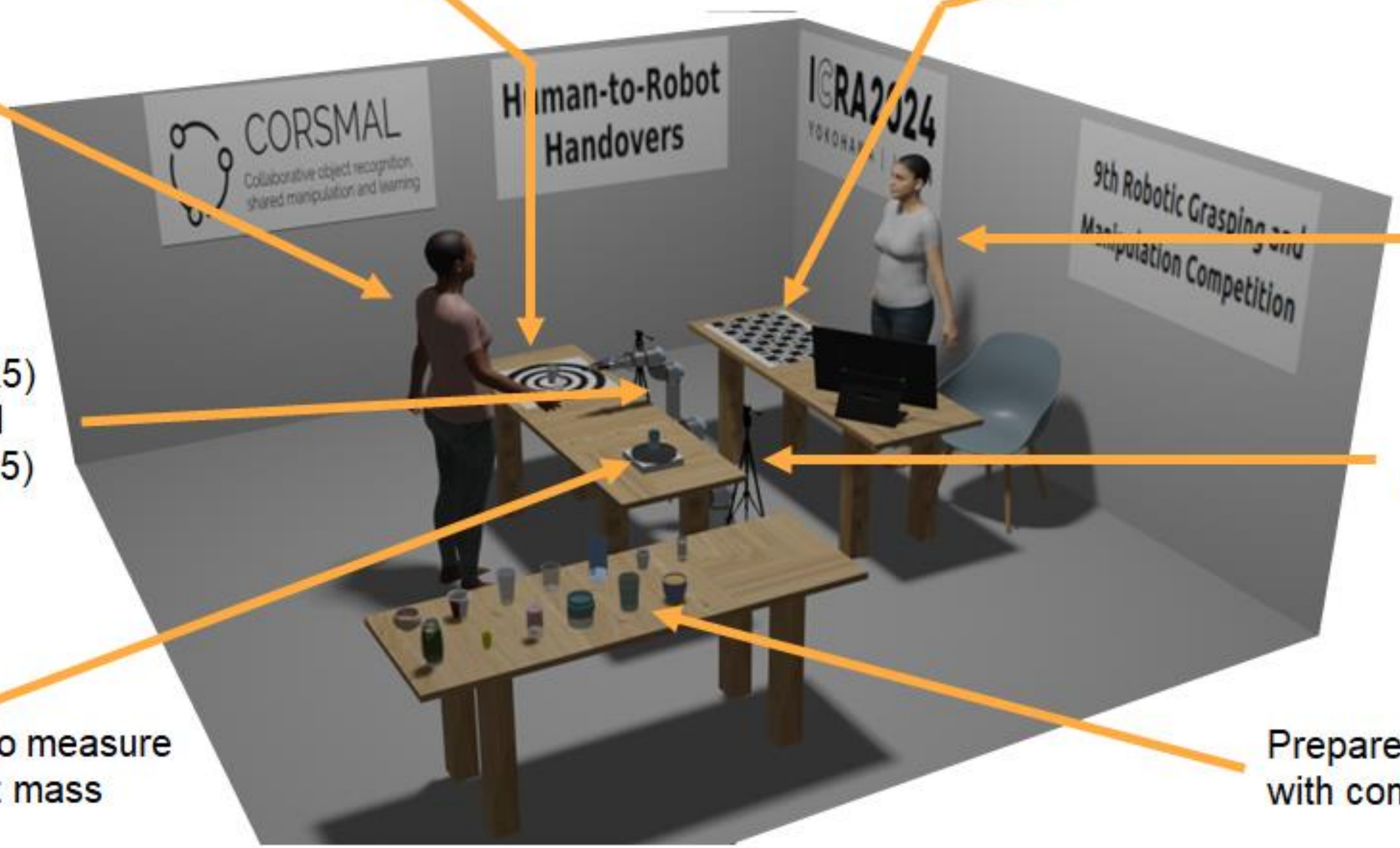
Calibration board

Volunteer (from another team)

Operator

Robotic arm (e.g., UR5) and 2-finger parallel gripper (Robotiq 2F-85)

2 cameras mounted on tripods (e.g., IntelRealSense 465i)



Digital scale to measure the object mass

Prepared objects with content (if any)

Safety: second person ready to stop the robot in case of an anomalous behavior

Image made in Blender 3D by Tommaso Apicella

Objects and fillings

CORSMAL
Collaborative object recognition, shared manipulation and learning

2 containers previously used for the preparation phase (known)

Unknown containers: 4 new containers made available on the competition day

Known filling: Rice
content previously used for the preparation phase

Unknown fillings: **Pasta** & **beans**
contents made available on the competition day

Procedure

For each configuration:

1. Prepare the container either empty or filled with its predefined content type and level
2. Weight the (filled) container before the execution of the task
3. Place the container at the centre of the table, at a distance not reachable by the robotic arm (safety)
4. The volunteer grasps the container from its location with a natural grasp
5. The volunteer carries the container with the intention of handing it over to the robot
6. The robot should track and predict the pose of the container to move the arm towards the handover area
7. The volunteer hands the container over to the robot
8. The robot closes the end effector and grasps the container
9. The robot delivers the container upright within the predefined area.
10. Measure the distance between the initial and the delivery location of the container (if not failed)
11. Weight the (filled) container after the execution of the task

This procedure has been revised from the CORSMAL Human-to-Robot Handover Protocol document

Handover configurations

CORSMAL
Collaborative object recognition, shared manipulation and learning

Handover **configuration**: predefined combination of a container with a content type and amount

Easy (5 points)

Red cup	None	Empty
Wine glass	None	Empty
Clear jar	None	Empty
Wine glass	Rice	90%
Red cup	Rice	90%
Champagne flute	None	Empty

Medium (10 points)

Red cup	Pasta	90%
Wine glass	Pasta	90%
Wine glass	Pasta	50%
Wine glass	Rice	50%
Wine glass	Beans	90%
Clear jar	Pasta	90%

Difficult (15 points)

Clear jar	Rice	50%
Clear jar	Rice	90%
Champagne flute	Pasta	50%
Champagne flute	Pasta	90%
Champagne flute	Rice	50%
Clear jar	Beans	90%

Hard (20 points)

Biscuit box	None	Empty
Tea box	None	Empty
Biscuit box	Pasta	50%
Tea box	Rice	50%
Tea box	Pasta	90%
Red cup	Beans	50%

24 configurations with increasing difficulty: 300 total points

Unknown object, content, or level

Unknown object, content, or level

Content non-visible

Rules: teams and their duties on-site

- Maximum 5 members

- to bring and mount their own setup:
  - robotic arm & end effector (e.g., 2-finger gripper, such as Robotiq)
  - laptop/workstation (tower PC with monitor)
  - cameras (with tripods, USB-C cables)
  - digital scale
  - calibration pattern
  - sockets extension lead & socket adapter
  - ethernet cable connecting PCs and robotic arm

- To prepare the sensing setup on-site:
  camera synchronization, camera calibration, camera-robot calibration

- To verify the behavior of the robotic arm prior to the execution of the task
  (e.g., end-effector, speed, kinematics, etc.)

- To weigh the mass of the container and content, if any, for each configuration before and after executing the handover to the robot, using a digital weight scale

Rules: handover configurations

CORSMAL
Collaborative object recognition, shared manipulation and learning

- All configurations executed by the same volunteer (same hand, natural grasp) from another team
- Natural and dynamic handovers:
  - no intention to help the robot (assistive behavior)
  - no intention to make it difficult for the robot (adversarial behavior)
- Each handover configuration executed only once
- Pre-defined initial location of the container on the table (not reachable by the robotic arm)
- Each configuration executed within 5 seconds
- Failed handover configuration: 0 points
  - robot cannot grasp and/or hold the container during the delivery phase
  - object falls after the robot places it on the table

Rules: designing the solution for the task

CORSMAL
Collaborative object recognition, shared manipulation and learning

Allowed

- Calls to existing Large Language Models
- Use of tactile sensors
- Both RGB and RGB-D inputs
- Any initial robot pose can be chosen with respect to the environment setup

  however, the volunteer expected to stand on the opposite side of the table w.r.t. the robot

Not allowed

- Prior knowledge of the objects (e.g., prior 3D object model)
- Learning across executions of configurations

Evaluation with a 100-point based scoring system

Accuracy of the delivery location

$$\delta(d_i, \rho) = \begin{cases} 1 - \frac{d_i}{\rho} & \text{if} \quad d_i < \rho \\ 0 & \text{otherwise,} \end{cases}$$

Maximum allowed distance from the target location: 500 mm

Euclidean distance from the target location

For each configuration, teams to report x and y coordinates in mm on the table of

1. the target location
2. the location where the container is delivered

CORSMAL
Collaborative object recognition, shared manipulation and learning

Evaluation with a 100-point based scoring system

Accuracy of
the delivery location

Efficiency
(Total execution time)

$$\gamma(t_i, \tau, \eta) = \begin{cases} 1 - \dfrac{\max(t_i, \eta) - \eta}{\tau - \eta} & \text{if} \quad t_i < \tau \\ 0 & \text{otherwise,} \end{cases}$$

Maximum allowed execution time to perform a handover: 5000 ms

Minimum expected time to perform a handover: 1000 ms

For each configuration, teams to report the total execution time in *milliseconds* (manually or automatically measured)

1. *Start time*: person enters in contact with the container (grasp by human)
2. *End time*: moment the robot releases the gripper at the delivery location

CORSMAL
Collaborative object recognition, shared manipulation and learning

Evaluation with a 100-point based scoring system

Accuracy of the delivery location

Efficiency (Total execution time)

Final mass of the delivered container

$$\mu(m_i, \hat{m}_i) = \begin{cases} 1 - \frac{|m_i - \hat{m}_i|}{\hat{m}_i} & \text{if} \quad |m_i - \hat{m}_i| < \hat{m}_i \\ 0 & \text{otherwise,} \end{cases}$$

m_i: Measured mass of the filled container (if not empty) *after* the handover execution

$\hat{m}_i$: Measured mass of the filled container (if not empty) *before* the handover execution

For each configuration, teams to report the measured mass in *grams* of the container + content, if any, (using the digital scale)

1. Before the execution of the handover configuration
2. After the container is delivered by the robot (<u>0 g if failing to deliver</u>)

CORSMAL
Collaborative object recognition, shared manipulation and learning

Evaluation with a 100-point based scoring system

Accuracy of the delivery location

Efficiency (Total execution time)

Final mass of the delivered container

Benchmark score:

$$S = \frac{1}{3}\sum_i \left[\omega_i \Psi_i \frac{\delta(d_i, \rho) + \gamma(t_i, \tau, \eta) + \mu(m_i, \hat{m}_i)}{3}\right]$$

24 configurations: 300 total points

Configuration points

Indicator function
0: maximum time exceed
1: otherwise

Temporarily discarded to avoid over penalizing teams

Python-based evaluation

CORSMAL Benchmark → revised → ICRA2024 Competition

CORSMAL Benchmark	ICRA2024 Competition
• 4 rotationally symmetric cups (known)	• 2 known cups + 4 new containers (not only rotationally symmetric objects)
• Only rice as content • cups either empty or full	• Rice + pasta + beans • Containers either empty, half full or full
• 288 configurations: 4 volunteers x 4 cups x 2 content level x 3 grasp types x 3 handover locations	• 24 configurations: 1 volunteer, 6 containers, 3 contents, 3 content levels, natural grasp, 1 location
• 13 granular performances scores (5 x vision, 3 x robotic, 5x task)	• Simplified to 3 task scores: delivery, mass, handover time

https://corsmal.eecs.qmul.ac.uk/benchmark.html

CORSMAL
Collaborative object recognition, shared manipulation and learning

Starting toolkit and documentation

Benchmark for human-to-robot handovers of unseen containers with unknown filling
R. Sanchez-Matilla, K. Chatzilygeroudis, K., A. Modas, N.F. Duarte, A., Xompero, A., P. Frossard, A. Billard, A. Cavallaro
IEEE Robotics and Automation Letters, 5(2), pp.1642-1649, 2020
[Open Access] [code] [webpage]

Towards safe human-to-robot handovers of unknown containers
Y. L. Pang, A. Xompero, C. Oh, A. Cavallaro
IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), Virtual, 8-12 Aug 2021
[Open Access] [code] [webpage]

LoDE
Method that jointly localises container-like objects and estimates their dimensions with a generative 3D sampling model and a multi-view 3D-2D iterative shape fitting, using two wide-baseline, calibrated RGB cameras.
[paper] [code] [webpage]

CORSMAL Challenge
Perception solutions for the estimation of the physical properties of manipulated objects prior to a handover to a robot arm.
[challenge] [paper 1] [paper 2]

CORSMAL Containers Manipulation [Data set]
A. Xompero, R. Sanchez-Matilla, R. Mazzon, A. Cavallaro
Queen Mary University of London [https://doi.org/10.17636/101CORSMAL1]

Additional References

CORSMAL
Collaborative object recognition, shared manipulation and learning

Human-to-robot handovers: statistics

- 9 teams registered

  in brackets the number of teams per country

- 4 teams showed up
  at the competition event (13-16 May 2024)

- 3 teams completed the competition

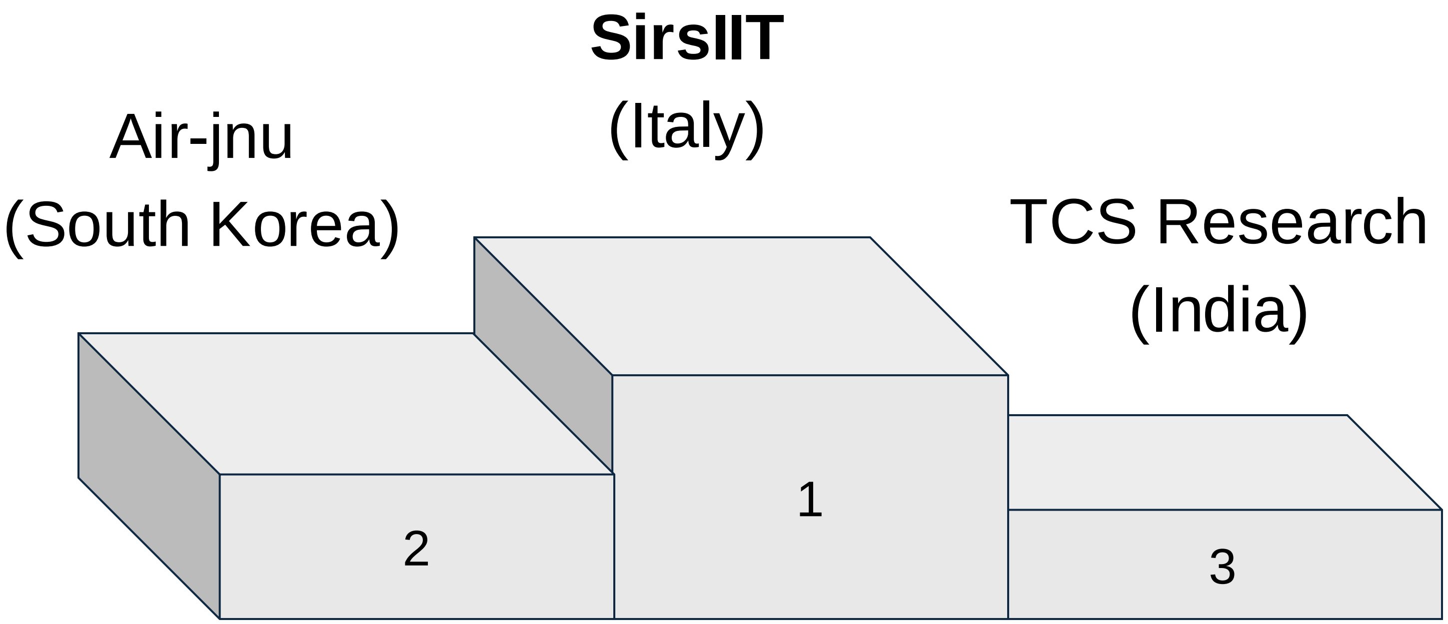
  (15 people, 5 people/team):
  - SirsIIT (Italy)
  - Air-jnu (South Korea)
  - Smart Machines TCS Research (India)

Legend

- At least 1 team competed
- At least 1 team showed up
- At least 1 team registered

CORSMAL
Collaborative object recognition, shared manipulation and learning

Human-to-robot handovers: winners



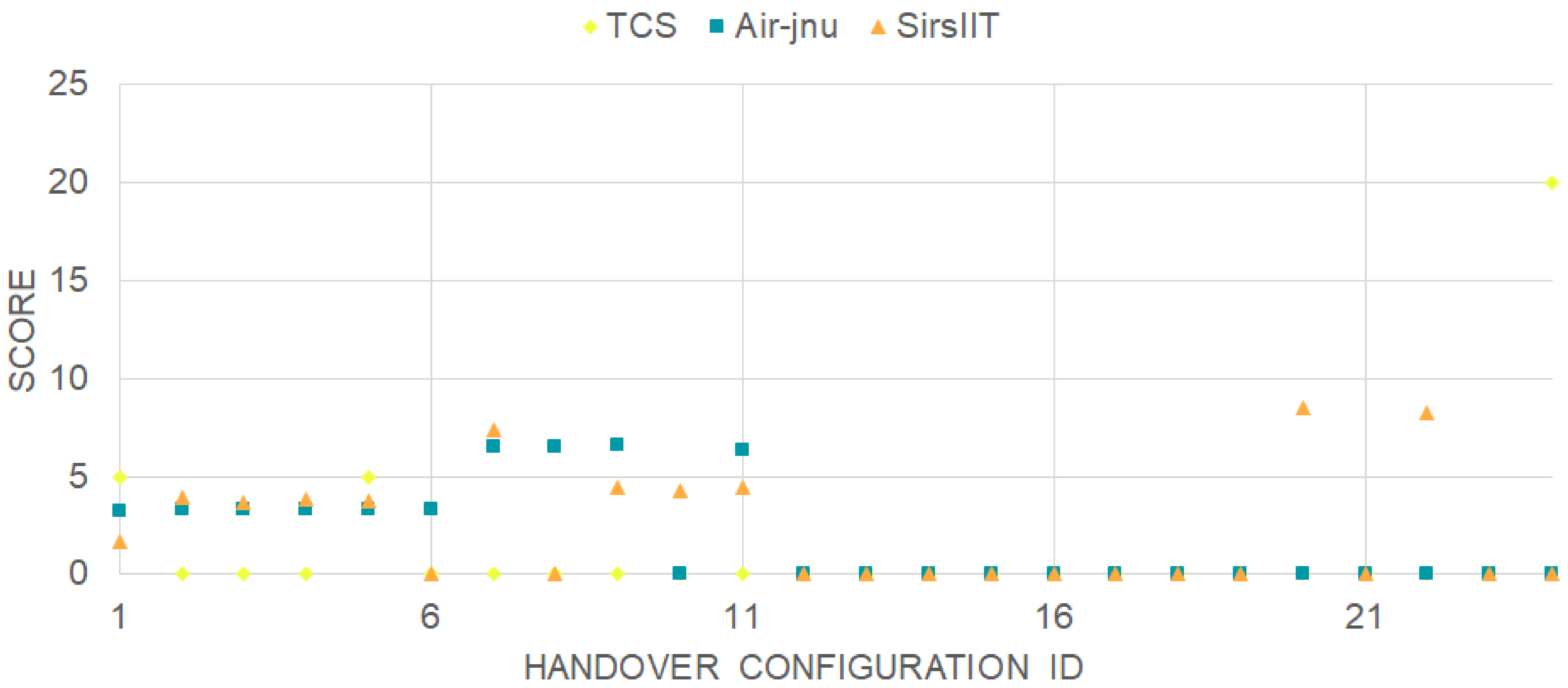
Leaderboard	
Team	**Score**
SirsIIT	18/100
Air-jnu	15/100
TCS Research	10/100

ICRA2024 YOKOHAMA | JAPAN

9th Robotics Grasping and Manipulation Competition

CORSMAL
Collaborative object recognition, shared manipulation and learning

Results per configuration

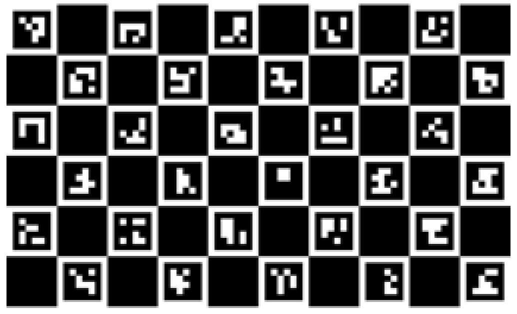


Teams did not have enough time to perform all configurations

Challenges of on-site competition

- Setting up the provided and available robot without any remote trial before the competition (version compatibility, software and libraries used)

- Calibration of the overall system, especially calibration of the cameras



- Illumination of the environment affecting the perception system (recognition of the object)

Smart Machines - TCS Research team (India)

Smart Machines – TCS Research Noida

Chandan Kumar Singh

Devesh Kumar

Vipul Sanap

Mayank Khandelwal

Ajin J

Air-jnu team (South Korea)

Chonnam National University

Giwan Lee, Jiyoung Choi, Jeongil Choi, Geon Kim, Phayuth Yonrith

(MSc students)

SirsIIT team (Italy)

Enrico Turco (PostDoc), ***Istituto Italiano di Tecnologia***

Valerio Bo (PostDoc), ***Istituto Italiano di Tecnologia***

Chiara Castellani (PhD student), ***Istituto Italiano di Tecnologia***

Maria Pozzi (Researcher), ***University of Siena***

Gionata Salvietti (Ass. Prof.), ***University of Siena***

Human-to-robot handovers of unknown containers with unknown content


ICRA2024
YOKOHAMA | JAPAN
ICRA EXPO
Prize of **$1,000**
sponsored by
amazon robotics
IEEE Robotics & Automation Society
IEEE
IEEE International Conference on
Robotics and Automation
13-17 May | Pacifico Yokohama
9th Robotics Grasping and
Manipulation Competition
CORSMAL
Collaborative object recognition,
shared manipulation and learning
SirsIIT team (Italy)
ICRA2024
YOKOHAMA | JAPAN
IEEE
Photo credit: Gionata Salvietti ©

Most Elegant Solution Winner

- XL Team, Tsinghua University: Xiang Li, Yongpeng Jiang, Mingrui Yu, Yongyi Jia, and Chen Chen
- Selected by competition organizers and judges