chist-era

CORSMAL
Collaborative object recognition, shared manipulation and learning

UKRI Engineering and Physical Sciences Research Council

2021 Intelligent Sensing Winter School

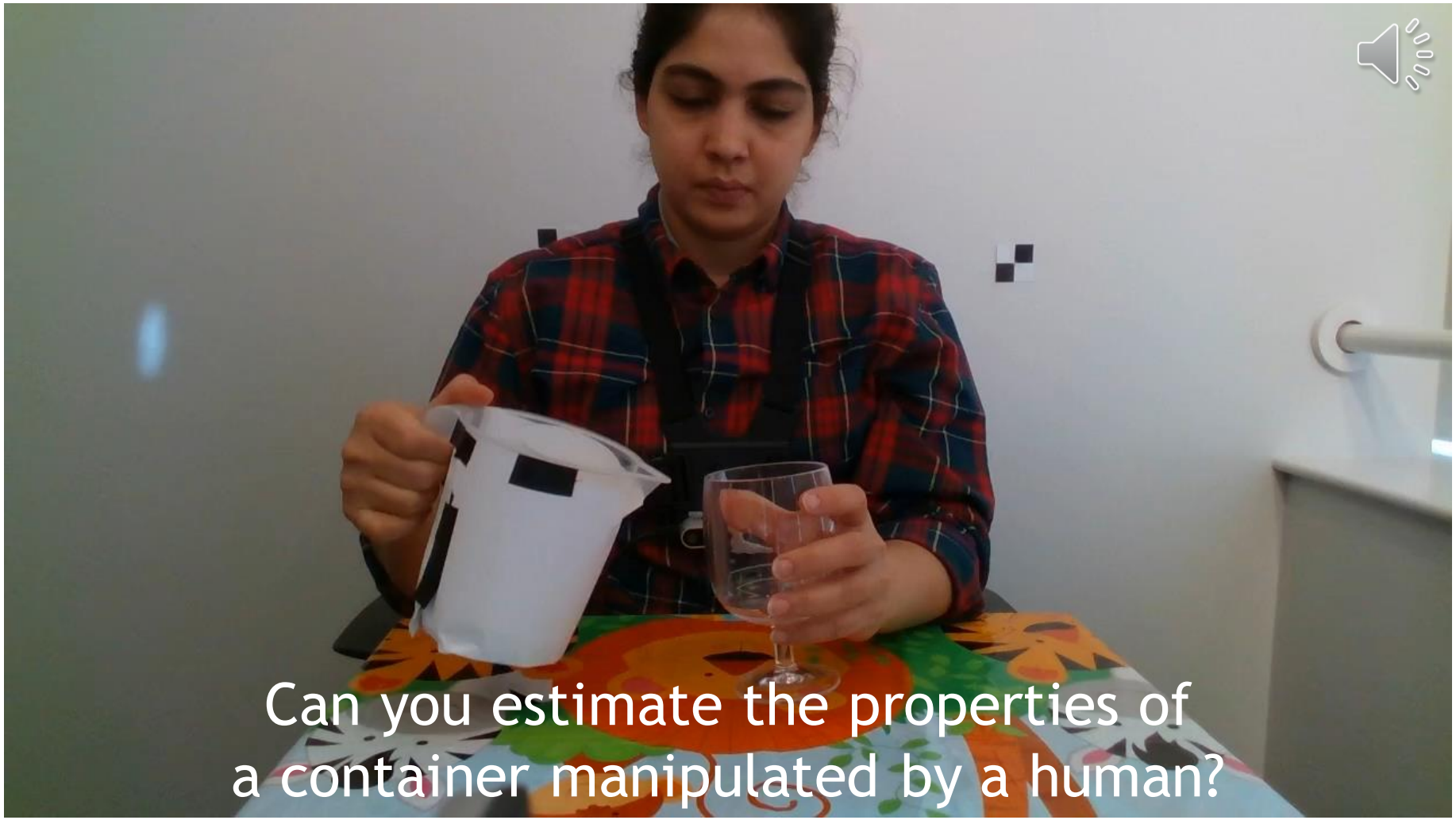
The CORSMAL challenge

10 December 2021

Queen Mary University of London

CIS centre for intelligent sensing

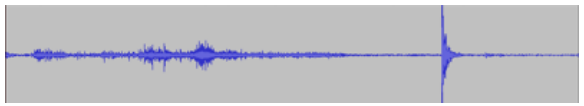
Introduction to the challenge



CORSMAL
Collaborative object recognition,
shared manipulation and learning

Physical properties estimation

Task for the Winter School



Filling type classification

Filling level classification

Container capacity estimation

Container mass estimation

Container dimensions estimation

Filling mass estimation

+

Object mass

CORSMAL
Collaborative object recognition, shared manipulation and learning

Containers and data splits for Winter School challenge

Training set
(684 audio-visual recordings)

S0
S1
S2

S0
S1
S2

Test set

84 audio-visual recordings/glass
84 audio-visual recordings/cup
60 audio-visual recordings/box

CORSMAL
Collaborative object recognition, shared manipulation and learning

Presentations by the teams and Q&A with the panel

Dr. Lin Wang

Audio signal processing

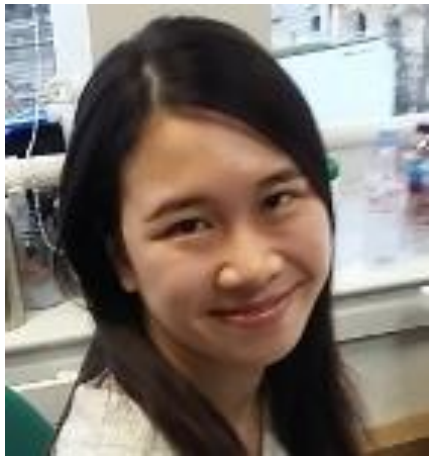
Vandana Rajan

Audio-visual fusion

Yik Lung Pang

CORSMAL Challenge co-organisers (vision, robotics)

Chau Yi Li



Vision, image processing

Dr. Riccardo Mazzon

Audio-visual, 3D geometry

If you need a certificate of attendance, email cis-web@eecs.qmul.ac.uk

CIS centre for intelligent sensing

CORSMAL
Collaborative object recognition, shared manipulation and learning

The Challenge is not over!

The CORSMAL challenge:
Audio-visual object classification for human-robot collaboration

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

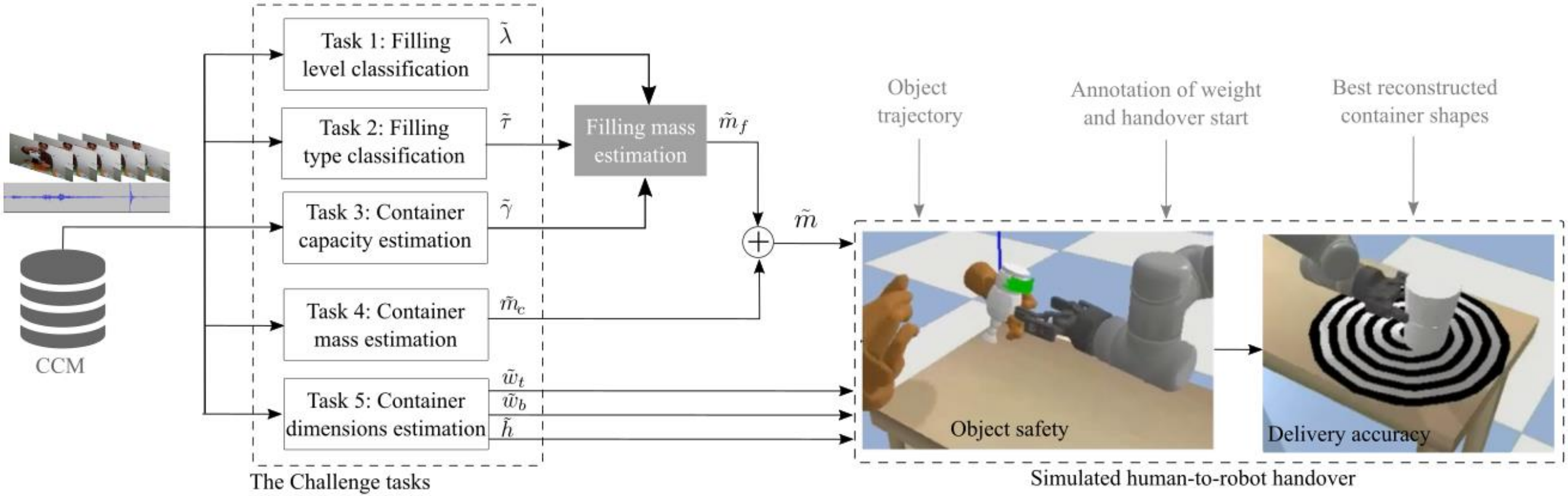
in conjunction with the 2022 IEEE International Conference on Acoustic, Speech and Signal Processing

Free registration: online form (or send this form to corsmal-challenge@qmul.ac.uk)

5 tasks, 10 performance scores, 1 real-to-simulation framework
10 leaderboards, 2 winners

CORSMAL
Collaborative object recognition, shared manipulation and learning

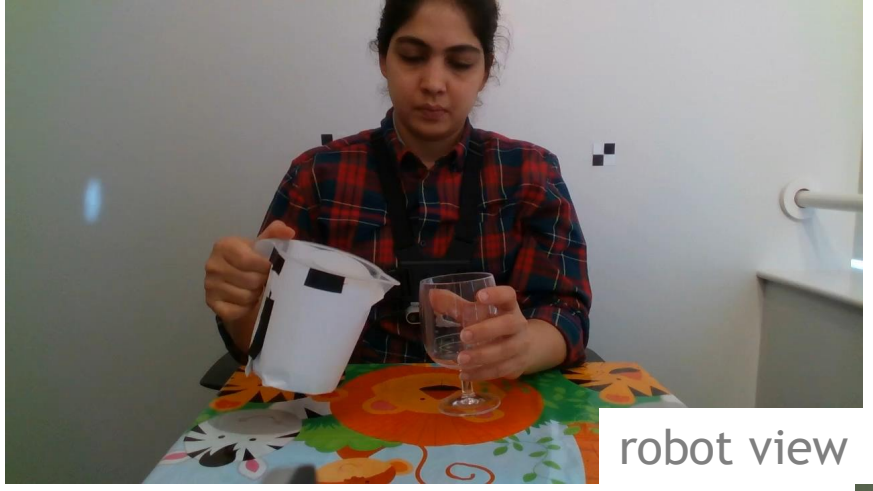
Physical properties estimation of unknown containers



Code of the real-to-simulation framework:
https://github.com/CORSMAL/safe_handover/

Towards safe human-to-robot handovers of unknown containers
Y. Pang, A. Xompero, C. Oh, and A. Cavallaro
IEEE Int. Conf. Robot and Human Interactive Comm. (RO-MAN), 8-12 August 2021.

15 containers
(5 cups, 5 drinking glasses, 5 food boxes)
3 filling types
3 filling levels



robot view

multi-view + multi-channel audio
4 views: 2 first-person + stereo pair
8 microphones: circular array

left view

CORSMAL Containers Manipulation dataset

1,140 audio-visual recordings
drinking glass: **84**
cup: **84**
food box: **60**

right view

12 people
3 scenarios
2 backgrounds
2 illuminations

body-cam view

Annotations
Filling type & level
Container volume
Container & filling masses

http://corsmal.eecs.qmul.ac.uk/containers_manip.html

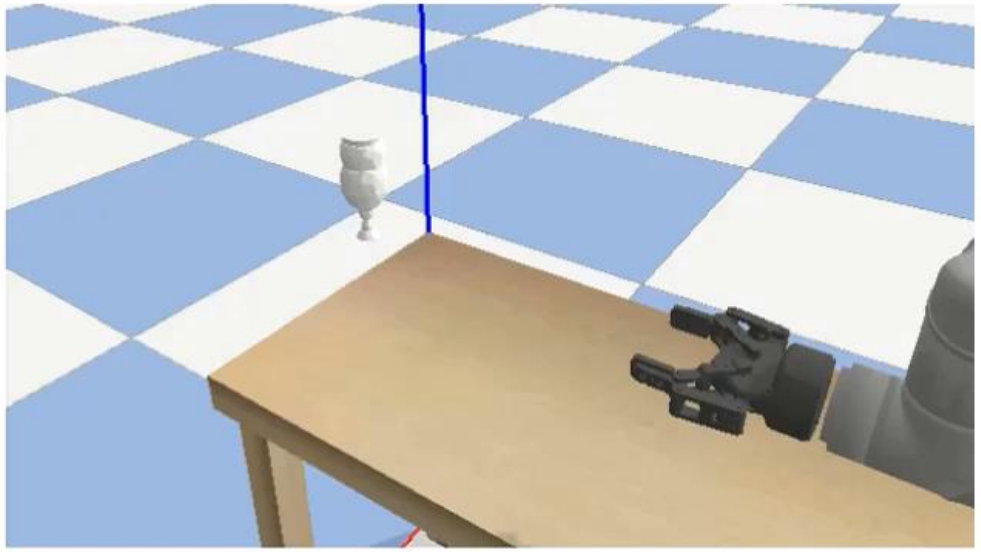
CORSMAL
Collaborative object recognition, shared manipulation and learning

Real-to-simulation framework

RGB video

Perceptual estimations + Annotations

Simulation



Object manipulation

Object manipulation

Evaluation

public testing set
(3 unseen containers,
228 configurations)

private testing set
(3 unseen containers,
228 configurations)

combined public & private testing sets

Performance scores

Description	Unit	Measure	Score
Filling level		λ^j	$s_1 = F_1(\lambda^1, \ldots, \lambda^J, \hat{\lambda}^1, \ldots, \hat{\lambda}^J)$
Filling type		τ^j	$s_2 = F_1(\tau^1, \ldots, \tau^J, \hat{\tau}^1, \ldots, \hat{\tau}^J)$
Capacity	mL	γ^j	$s_3 = \frac{1}{J}\sum_{j=1}^{J} \mathbb{1}e^{-\varepsilon^j(\gamma^j, \hat{\gamma}^j)}$
Width at top	mm	w_t^j	$s_4 = \frac{1}{J}\sum_{j=1}^{J} \mathbb{1}\sigma_1(w_t^j, \hat{w_t}^j)$
Width at bottom	mm	w_b^j	$s_5 = \frac{1}{J}\sum_{j=1}^{J} \mathbb{1}\sigma_1(w_b^j, \hat{w_b}^j)$
Height	mm	h^j	$s_6 = \frac{1}{J}\sum_{j=1}^{J} \mathbb{1}\sigma_1(h^j, \hat{h}^j)$
Container mass	g	m_c^j	$s_7 = \frac{1}{J}\sum_{j=1}^{J} \mathbb{1}e^{-\varepsilon^j(m_c^j, \hat{m}_c^j)}$
Filling mass	g	m_f^j	$s_8 = \frac{1}{J}\sum_{j=1}^{J} \mathbb{1}e^{-\epsilon^j(m_f^j, \hat{m}_f^j)}$
Object mass (container + filling)	g	m^j	$s_9 = \frac{1}{J}\sum_{j=1}^{J} \mathbb{1}\psi^j(m^j, \hat{F}^j)$
Container pose at delivery	(mm, °)	(α^j, β^j)	$s_{10} = \frac{1}{J}\sum_{j=1}^{J} \Delta_j(\alpha^j, \beta^j, \eta, \phi)$

Details: https://corsmal.eecs.qmul.ac.uk/resources/challenge/PerformanceScores.pdf

Evaluation toolkit: https://github.com/CORSMAL/CORSMALChallengeEvalToolkit

Ranking and rules

- Overall ranking:
  - aggregation of the performance scores
  - weighed based on the number of submitted tasks
  - filling mass & object mass:
    using random case results if one (or more) of the tasks is (are) not submitted by a team

- **Winners:**
  - team with the best-performing solution (highest score, above 60);
  - team with the most innovative solution (as judged by the organisers among the submissions with score above 60).

- Valid submissions:
  - must have the source code (evaluation on private test set)
  - not reproducible source codes: 0 score

Additional rankings
(recognition of best-performing solution)

joint filling type & level classification
container capacity & dimensions estimation
filling mass estimation
filling level estimation
container capacity estimation
container mass estimation

Additional info at:
https://corsmal.eecs.qmul.ac.uk/challenge.html

Leaderboards (1/2)

Combined CCM test sets

Team	T1	T2	T3	T4	T5	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	O...
BIT	✓	✓	✓			79.65	94.26	60.57	0.00	0.00	0.00	0.00	65.46	0.00	0.00	37.49
HVRL	✓	✓	✓			78.56	96.95	54.79	0.00	0.00	0.00	0.00	61.77	0.00	0.00	36.51
Average	✓	✓	✓	✓	✓	33.15	23.01	40.73	76.89	58.19	64.32	22.06	51.44	0.00	0.00	29.61
Concatenation	✓	✓	✓			43.53	41.83	62.57	0.00	0.00	0.00	0.00	63.54	0.00	0.00	26.43
NTNU		✓	✓			0.00	86.89	67.30	0.00	0.00	0.00	0.00	39.91	0.00	0.00	24.26
Random	✓	✓	✓	✓	✓	33.15	23.01	23.78	32.33	25.36	42.48	29.42	40.32	0.00	0.00	22.88
Challengers	✓	✓				48.71	75.24	0.00	0.00	0.00	0.00	0.00	26.50	0.00	0.00	18.81

CCM public test set

Team	T1	T2	T3	T4	T5	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	O...
HVRL	✓	✓	✓			82.63	97.83	57.19	0.00	0.00	0.00	0.00	63.52	0.00	0.00	37.65
BIT	✓	✓	✓			78.14	93.83	60.56	0.00	0.00	0.00	0.00	66.17	0.00	0.00	37.34
Average	✓	✓	✓	✓	✓	33.42	22.91	54.39	79.56	65.53	82.90	21.88	59.48	0.00	0.00	33.51
Concatenation	✓	✓	✓			44.31	41.77	63.00	0.00	0.00	0.00	0.00	62.32	0.00	0.00	26.42
Random	✓	✓	✓	✓	✓	33.42	22.91	31.63	26.82	22.55	50.34	30.59	44.04	0.00	0.00	24.48
NTNU		✓	✓			0.00	81.97	66.92	0.00	0.00	0.00	0.00	41.65	0.00	0.00	23.82
Challengers	✓	✓				50.73	78.58	0.00	0.00	0.00	0.00	0.00	29.36	0.00	0.00	19.83

Leaderboards (2/2)

Container capacity estimation

Team	Input modalities									Publ	Priv	Comb
		View 1		View 2		View 3		View 4				
	A	RGB	D	RGB	D	RGB	D	RGB	D			
NTNU							✓			66.92	67.67	67.30
Concatenation		✓		✓		✓		✓		63.00	62.14	62.57
BIT		✓		✓						60.56	60.58	60.57
HVRL		✓	✓							57.19	52.38	54.79
Average										54.39	27.08	40.73
Random										31.63	15.92	23.78

Container mass estimation

Team	Input modalities									Publ	Priv	Comb
		View 1		View 2		View 3		View 4				
	A	RGB	D	RGB	D	RGB	D	RGB	D			
Random										30.59	27.58	29.08
Average										21.88	21.88	21.88

Joint filling type and level classification

Team	Publ	Priv	Comb
HVRL	82.14	73.40	77.81
BIT	75.00	77.86	76.45
Challengers	40.61	38.34	39.55
VA2M	24.59	23.98	24.32
Random	9.73	7.87	8.88

Container dimensions and capacity estimation

Team	Publ	Priv	Comb
Average	65.19	42.01	53.60
NTNU	33.46	33.84	33.65
VA2M	31.50	31.07	31.28
BIT	30.28	30.29	30.28
Random	32.43	24.73	28.58
HVRL	28.59	26.19	27.39

Filling level classification

Team	Input modalities									Publ	Priv	Comb
		View 1		View 2		View 3		View 4				
	A	RGB	D	RGB	D	RGB	D	RGB	D			
ACC	✓									80.22	81.46	80.84
BIT	✓	✓		✓		✓		✓		78.14	81.16	79.65
HVRL	✓									82.63	74.43	78.56
ZCR+MFCC+RF	✓									70.04	63.11	66.80
A5F+RF	✓									64.18	63.94	64.74
ZCR+MFCC+SVM	✓									66.27	57.19	61.87
A5F+SVM	✓									60.77	58.57	60.09
ZCR+MFCC+kNN	✓									63.63	54.97	59.35
Spectrogram+kNN	✓									59.15	53.47	56.38
A5F+kNN	✓									55.49	53.22	54.47
Spectrogram+SVM	✓									47.66	51.54	49.67
Challengers	✓									50.73	47.08	48.71
Mask R-CNN + RN18		✓								58.51	32.93	47.00
Spectrogram+RF	✓									45.43	45.59	45.49
Spectrogram+PCA+RF	✓									46.79	42.46	44.66
Concatenation	✓	✓		✓		✓		✓		44.31	42.70	43.53
Mask R-CNN + RN18				✓						48.90	26.73	39.00
Spectrogram+PCA+k...	✓									39.03	37.16	38.31
Random										33.35	41.86	37.62
Spectrogram+PCA+S...	✓									30.08	31.99	31.64
Mask R-CNN + RN18						✓				36.52	25.52	31.46
Mask R-CNN + RN18								✓		25.12	21.99	23.68

Filling mass estimation

Team	T1	T2	T3	Publ	Priv	Comb
BIT	✓	✓	✓	64.98	65.15	65.06
HVRL	✓	✓	✓	63.32	61.01	62.16
Concatenation	✓	✓	✓	52.80	54.14	53.47
NTNU-ERC		✓	✓	38.56	39.80	39.18
Random	✓	✓	✓	43.61	31.65	35.06
Average	✓	✓	✓	59.05	59.05	35.06
Challengers	✓	✓		29.25	23.21	26.23
ACC	✓	✓		28.25	21.89	25.07
Mask R-CNN + R...	✓	✓		19.46	9.59	14.53
Mask R-CNN + R...	✓	✓		15.15	9.96	12.56
Mask R-CNN + R...	✓	✓		17.28	6.99	12.14
Mask R-CNN + R...	✓	✓		12.95	10.25	11.60

Schedule

CORSMAL
Collaborative object recognition, shared manipulation and learning

Public training set available for download	November 10, 2021
Release of the password for the public test set	January 10, 2022
Submission of **papers**, estimation results on the public test set, & source code	**January 24, 2022**
Paper acceptance notification	February 10, 2022
Release of the results on the leaderboards	February 10, 2022
Camera-ready papers for ICASSP 2022 Proceedings due	February 16, 2022

All deadlines are **11.59pm PDT**

You may complete, extend, and submit your solution
Opportunity to submit a scientific paper about your solution

IEEE

2022 IEEE International Conference on
Acoustic, Speech and Signal Processing
22-27 May 2022, Singapore

icassp 2022 Singapore

chist-era

CORSMAL
Collaborative object recognition, shared manipulation and learning

UKRI Engineering and Physical Sciences Research Council

2021 Intelligent Sensing Winter School

Thank you!

Queen Mary University of London

CIS centre for intelligent sensing