Artificial Intelligence and Robotics for Intelligent Agents — with emphasis on social robotics —

RoboCanes Lab

Ubbo Visser

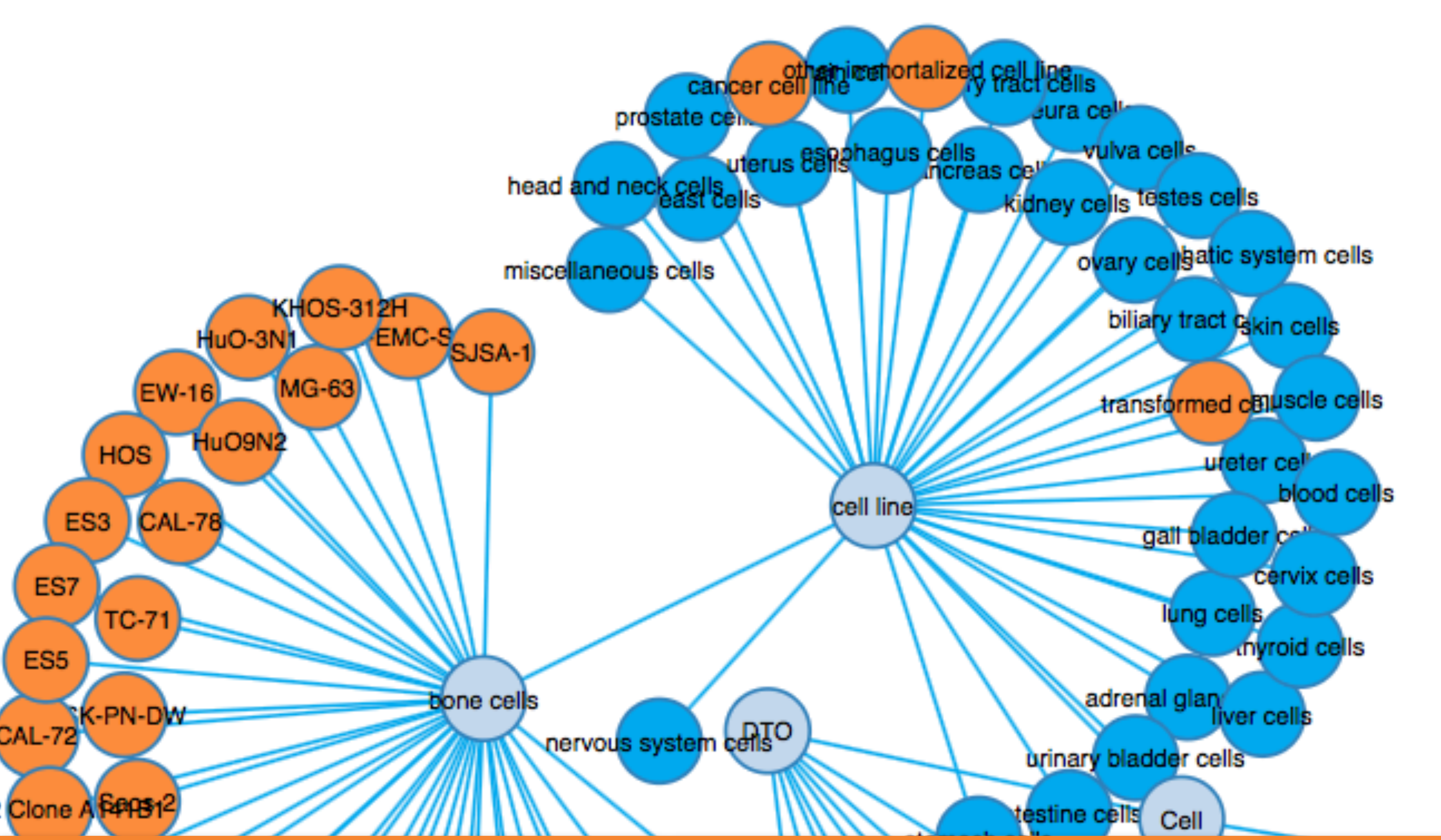
Associate Professor of Computer Science

President of the RoboCup Federation
Distinguished Faculty for Cognitive Science and Aging
Graduate Director of the Department of Computer Science
College of Arts and Sciences
University of Miami

2025

UNIVERSITY OF MIAMI
ROBOCANES

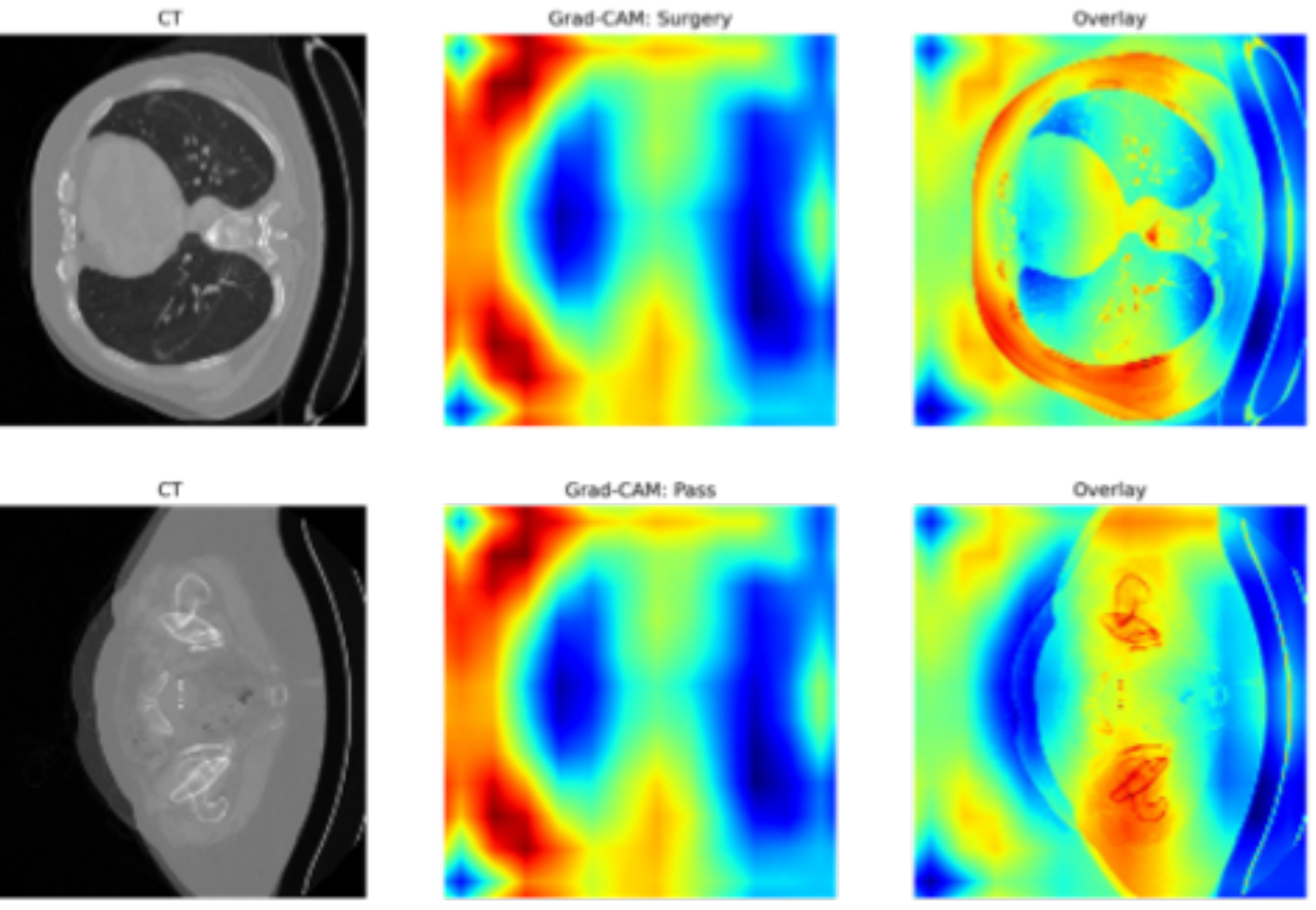

CYRUS 0:0 HELIOS2024 play_on 3466
FCPortugal
ASML
Ball Placement (23s)
ZJUNlict 0 3:45 0 TIGERs Mannheim
RoboCup 2024 @Home DSPL HMA Digest
Storing Groeries / 食料雑貨の片付け (Stage I)
x8.0
Navigating to the table 15
Opening the cabinet door (bonus) 200
Perceiving and picking up an object 65 x1
Placing the object next to similar objects 80 x1
Total 570

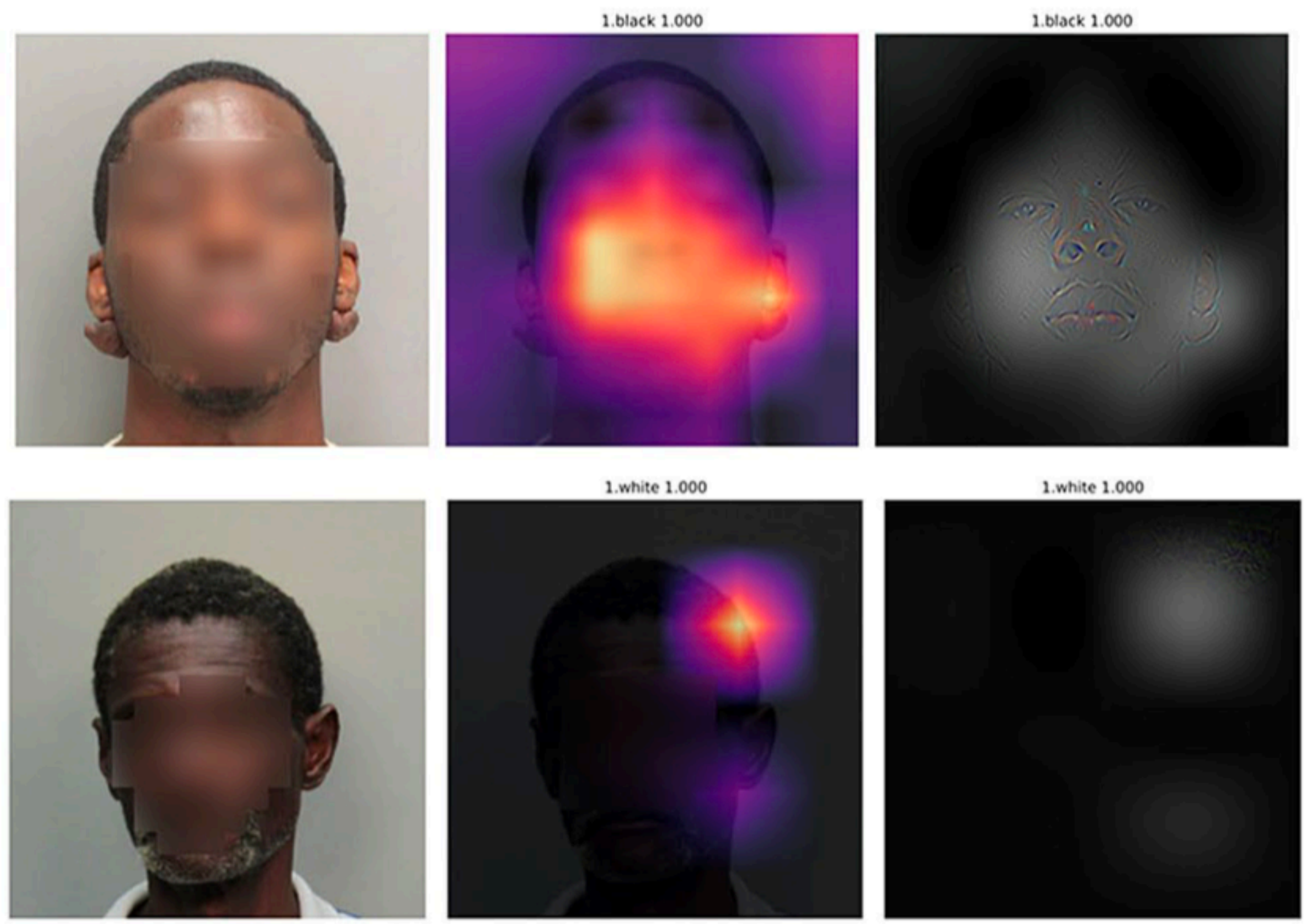


RegenBase
Knowledge Base

Ontologies, Knowledge Engineering, Semantic Web, Spatio-temporal Reasoning

Machine Learning, Medical Imaging, Digital Health, Therapy, FRT





Drive interdisciplinary AI & Robotics research to enable the next generation of robots that work and interact safely alongside humans.

RESEARCH CHALLENGES

- Autonomous robots shall act appropriately in dynamic, real time, and adversarial environments
- One of the biggest challenges in AI and robotics
- Numerous examples: rescue scenarios, home assistance, device assistance (e.g. cars, planes)
- Playing soccer with biped robots as a testbed for development in perception, multi-agent cooperation, complex motions, …
- Service robots for domestic environments
- RoboCup is a landmark project as well as a standard problem

RoboCanes Soccer, 5 NAO robots

RoboCanes@Home, HSR

WIDE RANGE OF RESEARCH CHALLENGES

real time sensor fusion

grasping and manipulation

context recognition

learning

real time planning

multi-agent systems

decision making

human-robot interaction

opponent/user modeling

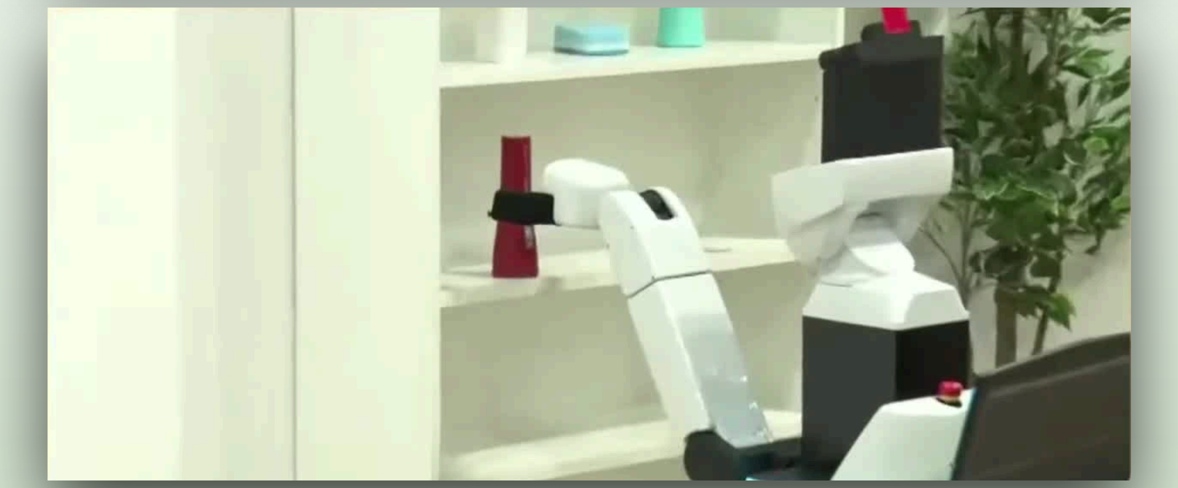
motor control

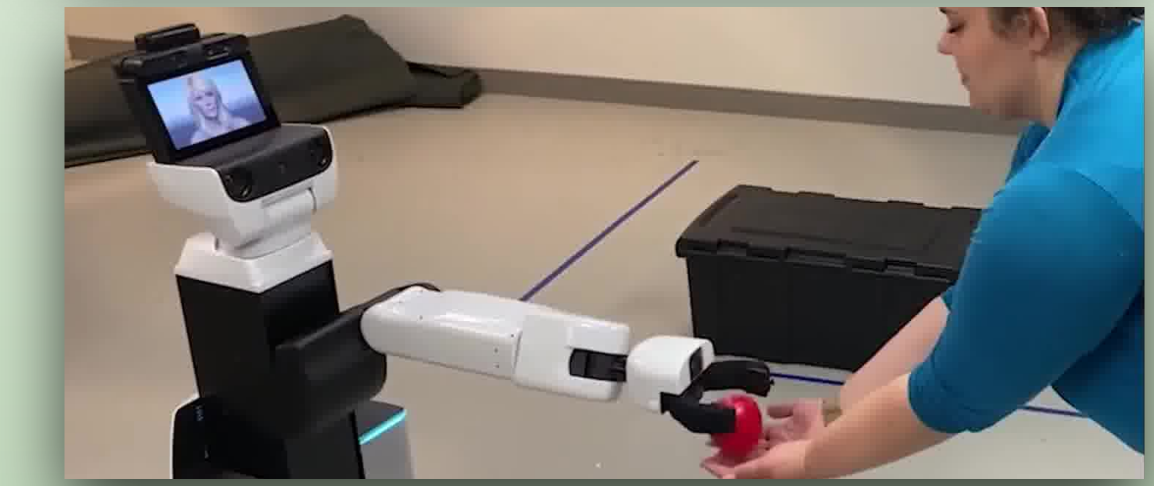
reactive/proactive behavior

and many more...

OUTLINE

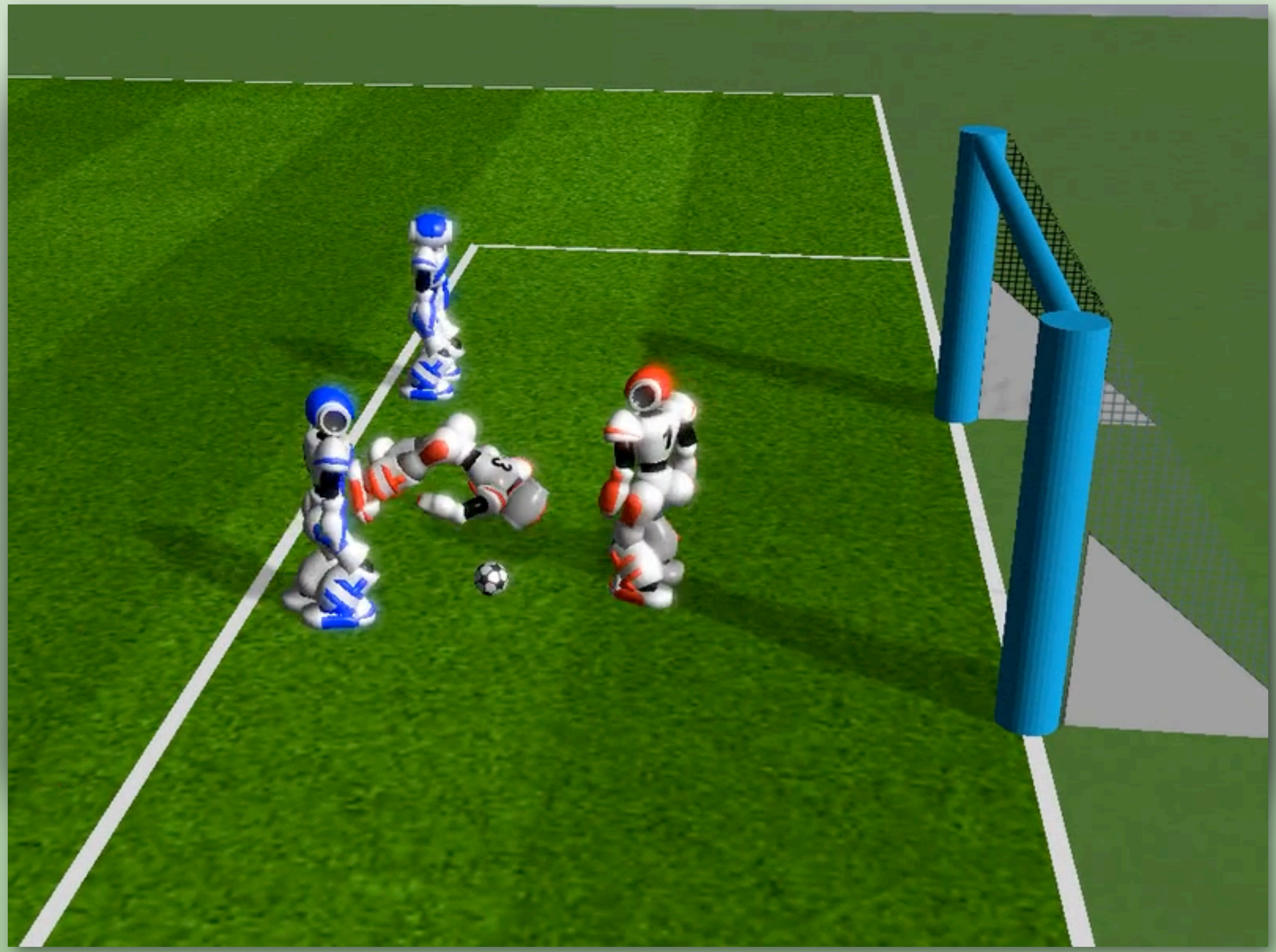
1. Multi-robot cooperation and communication
2. Manipulation systems
3. Human-Robot Interaction
4. Interfaces: VR meets AI & Robotics
5. Facial recognition on MD criminal justice data
6. Robot localization and path planning



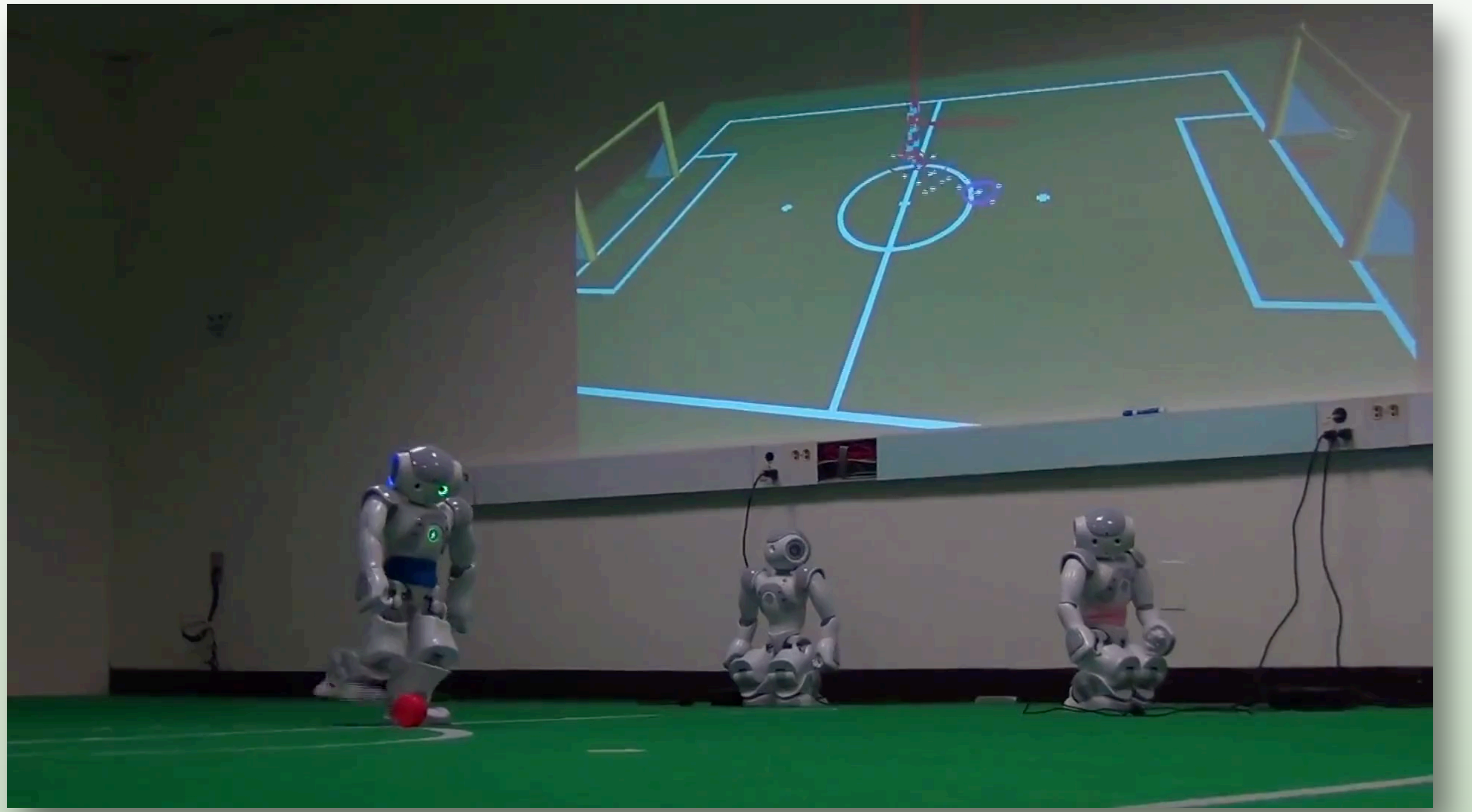


1. MULTI-ROBOT COOPERATION AND COMMUNICATION

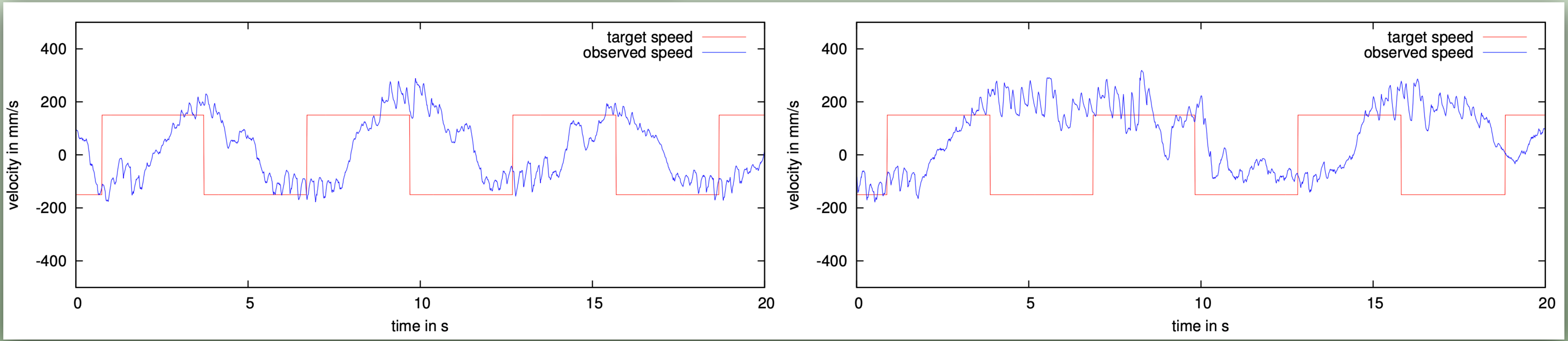
- Task and motion planning
  - Creation of high-level commands and collision-free trajectories to achieve goal
- State estimation and perception
  - Infer relevant quantities from sensor data (field objects, opponents, team mates, contacts/collisions, …)
- Communication
  - Communication with team mates, determine global geometry (e.g. ball position), ad-hoc sub-team building (e.g. offside trap, double-pass)
- Object manipulation
  - Determine good kick positions given relevant constraints (global geometry, local geometry, placement of ball)
- Trajectory generation and control
  - Real-time, reactive generation of control commands to move bipedal robot safely toward goal



CONTROL (1): DYNAMIC ADAPTIVE WALKING ENGINE



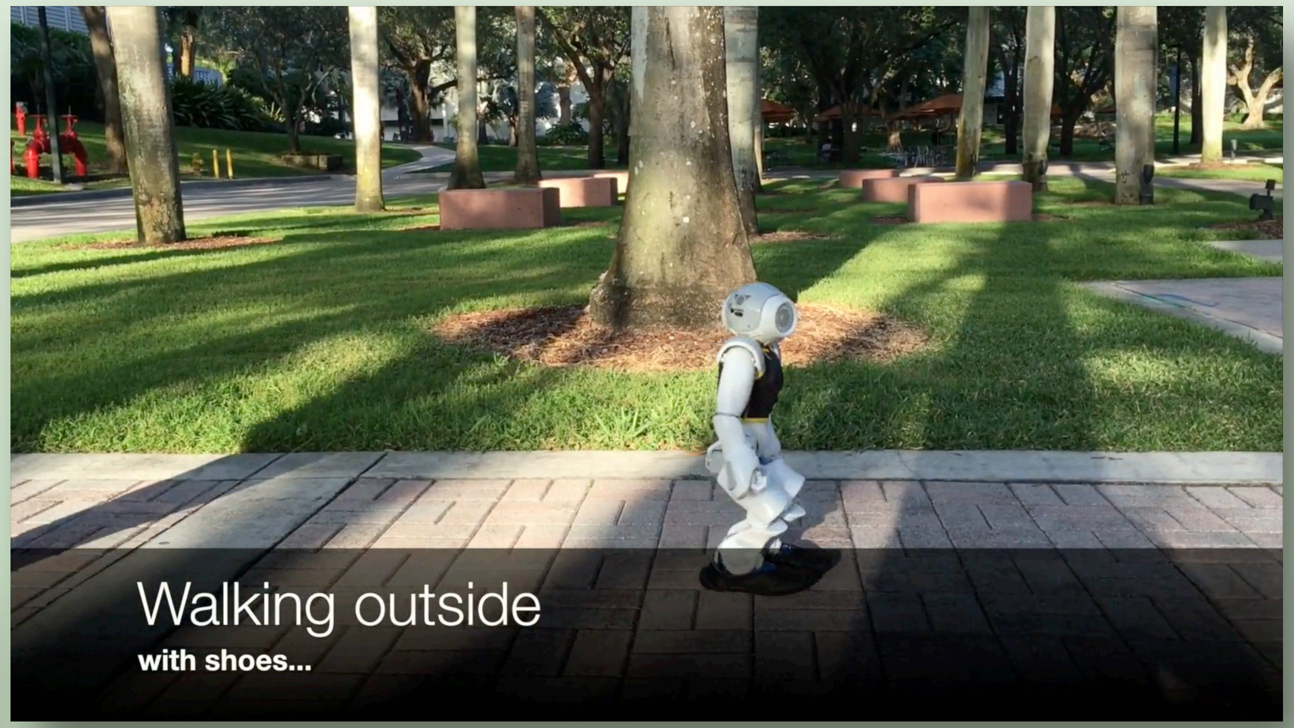
- First bipedal walk, open-loop, parameter prior optimized, no changes during runtime
- Becomes unstable due to changes (e.g. motor temperature)

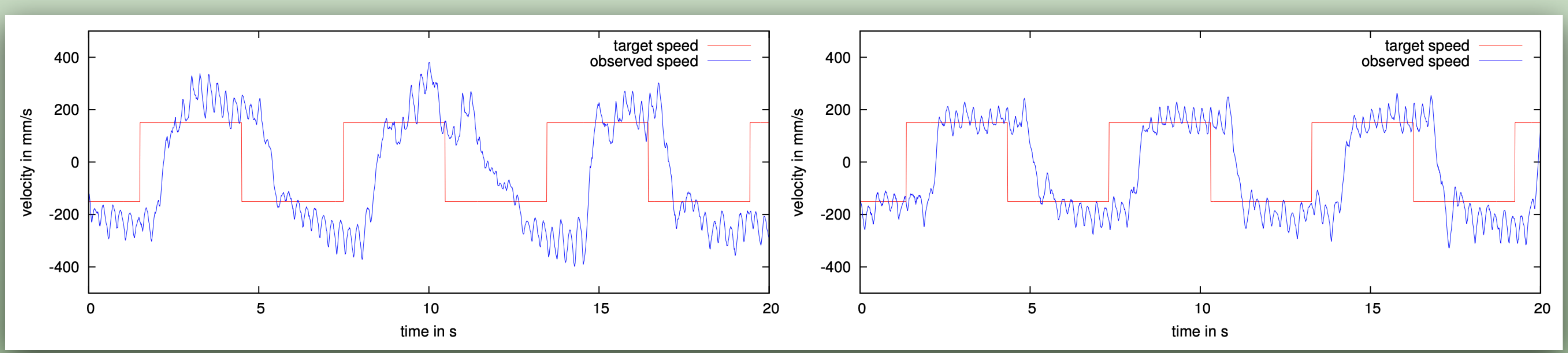


Forward/backward walk. Slow acceleration for stability (left), higher acceleration limit less stable (right)

CONTROL (1): DYNAMIC ADAPTIVE WALKING ENGINE

- LIPM-based closed-loop walk
- Adaptation by optimizing parameter of model in real-time onboard
- Result: new, stable, fast, and energy-efficient walk



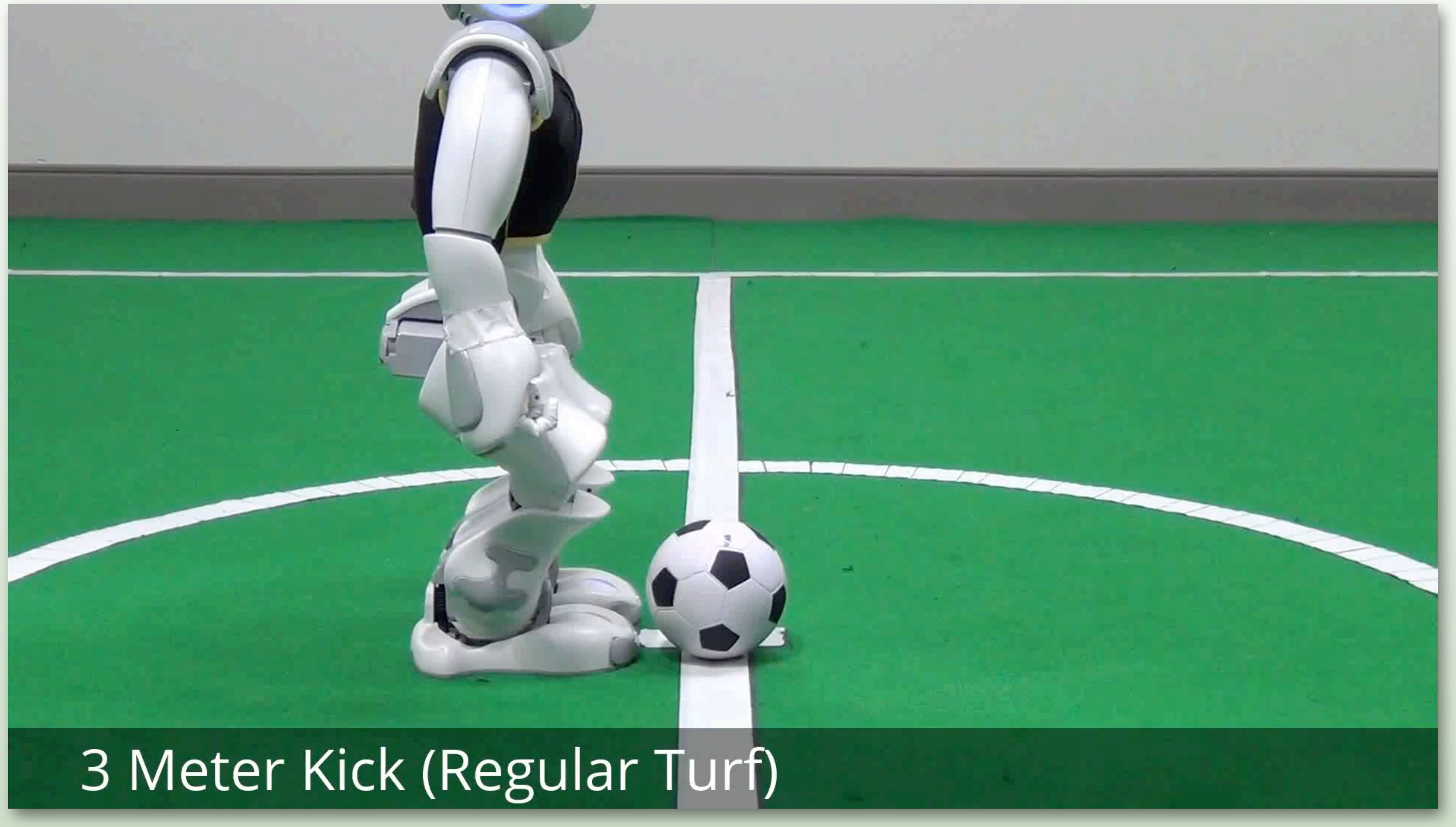


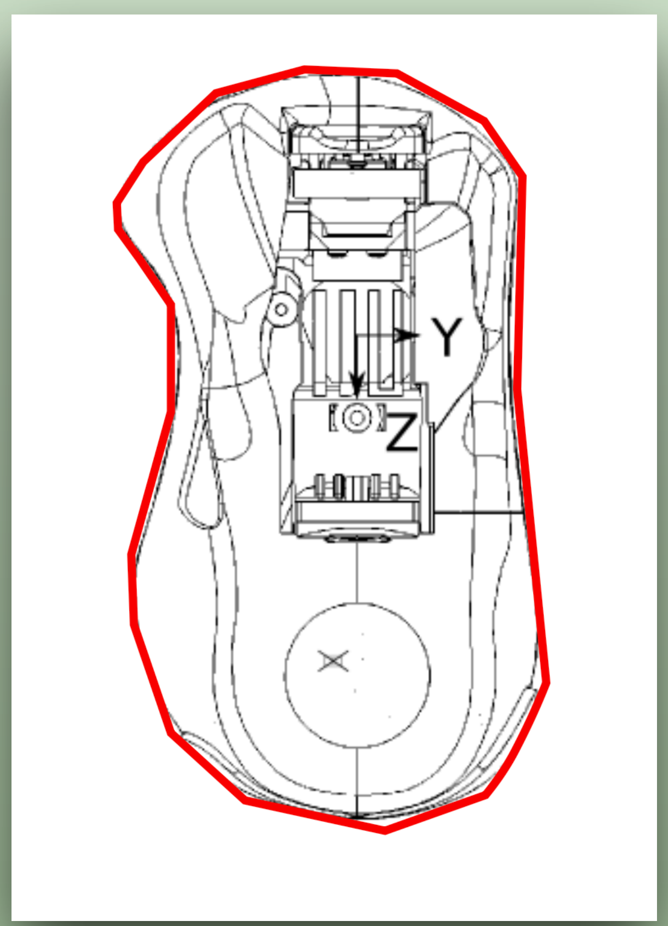
Forward/backward walk with 15cm/s with new walk. No optimization (left), with optimization (right)

Seekircher & Visser (*KI: German AI Journal, 2016*)

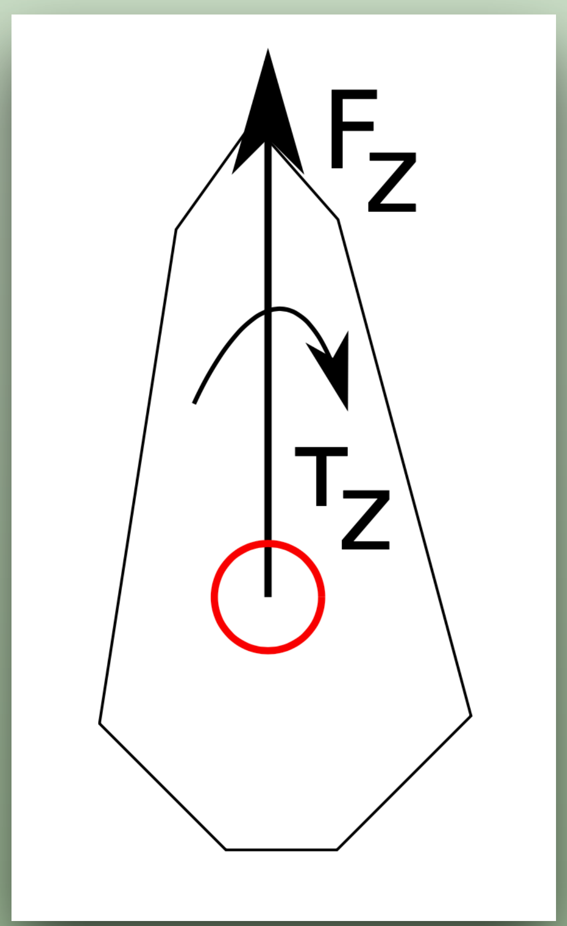
CONTROL (2): OMINDIRECTIONAL KICK

- Kick trajectory generation, arbitrary direction, no prior input or knowledge of the kick parameters
- Key idea: Zero Moment Point (ZMP) based preview controller that minimizes the ZMP error
- Covariance Matrix Adaptation Evolution Strategy (CMA-ES) for model optimization.

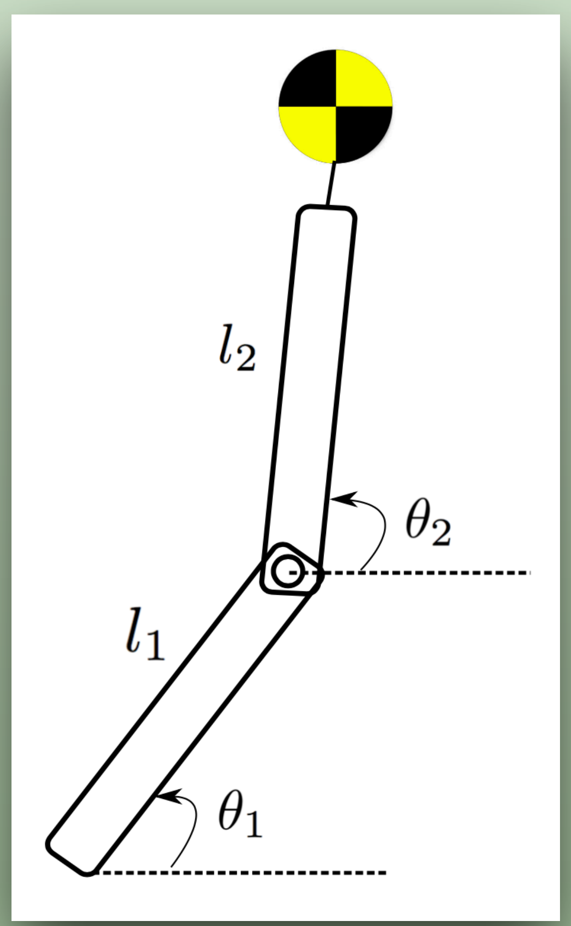




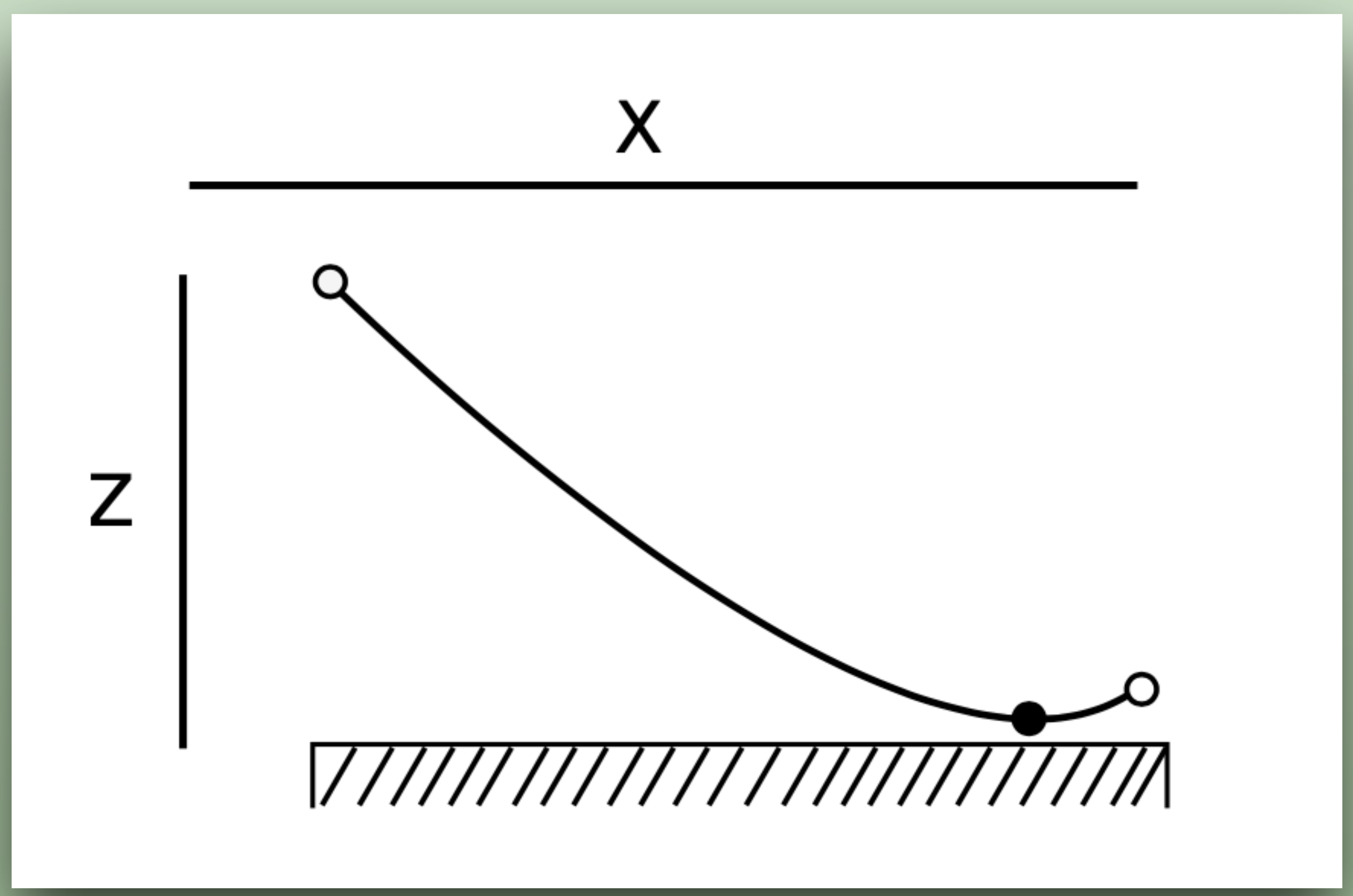
Support polygon (convex hull)



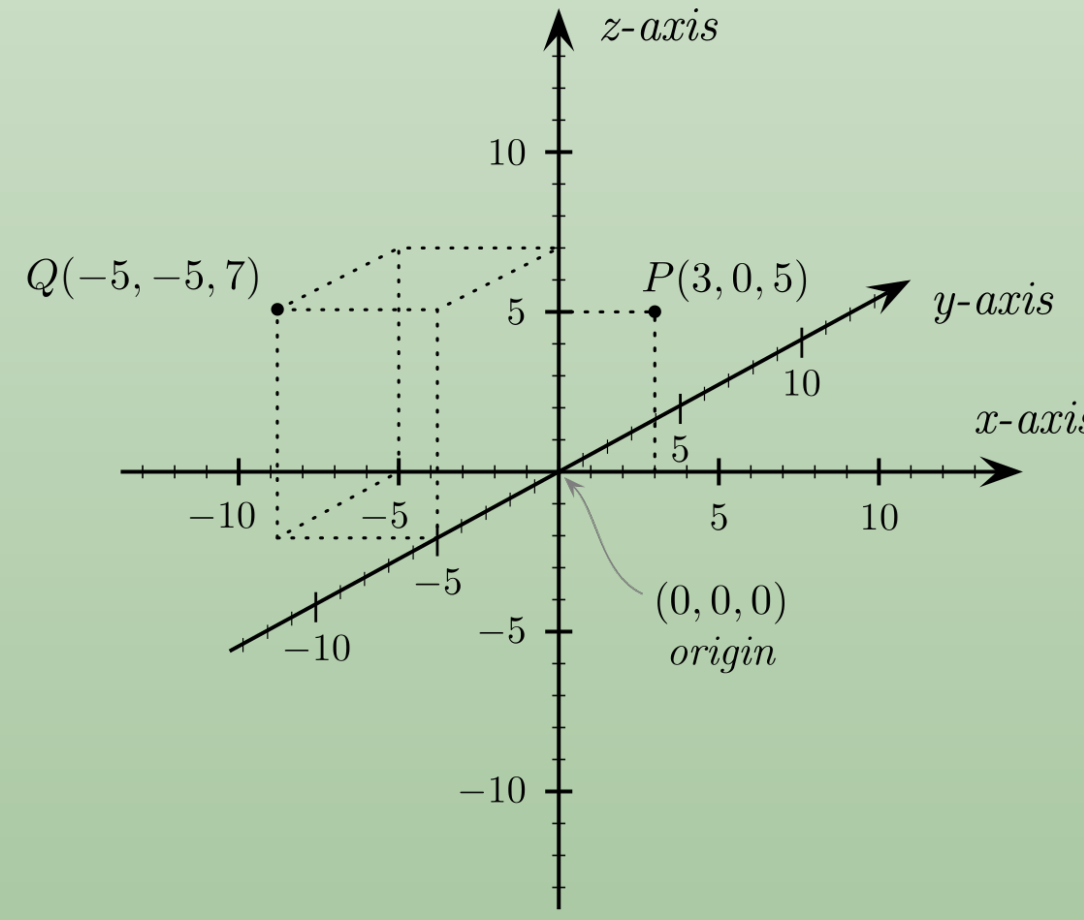
ZMP location



Approximation of inverse kinematic model



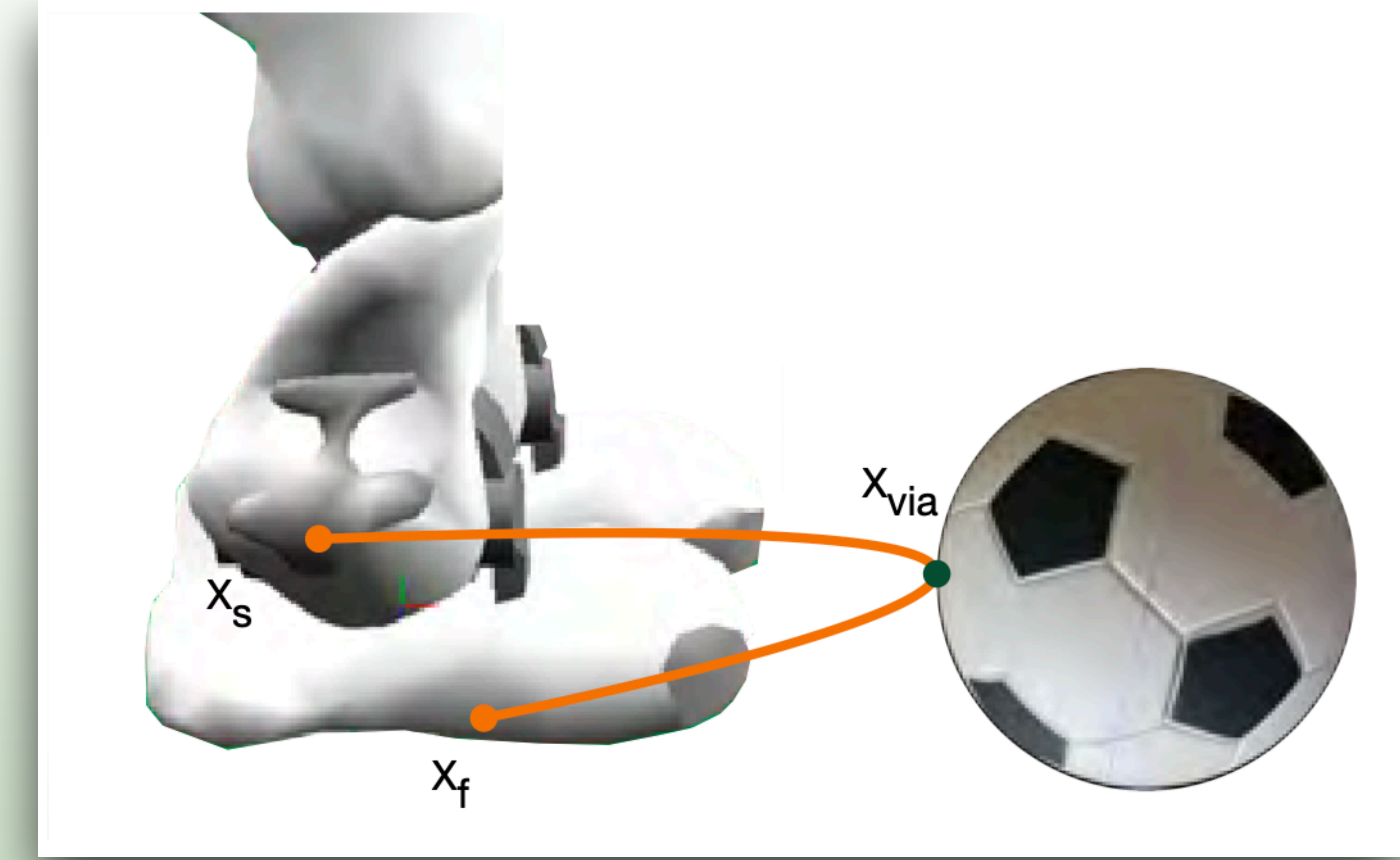
Two polynomials for front kick



Peña, Masterjohn & Visser (*RoboCup Symposium 2017*)

CONTROL (3): WALK-KICK ENGINE

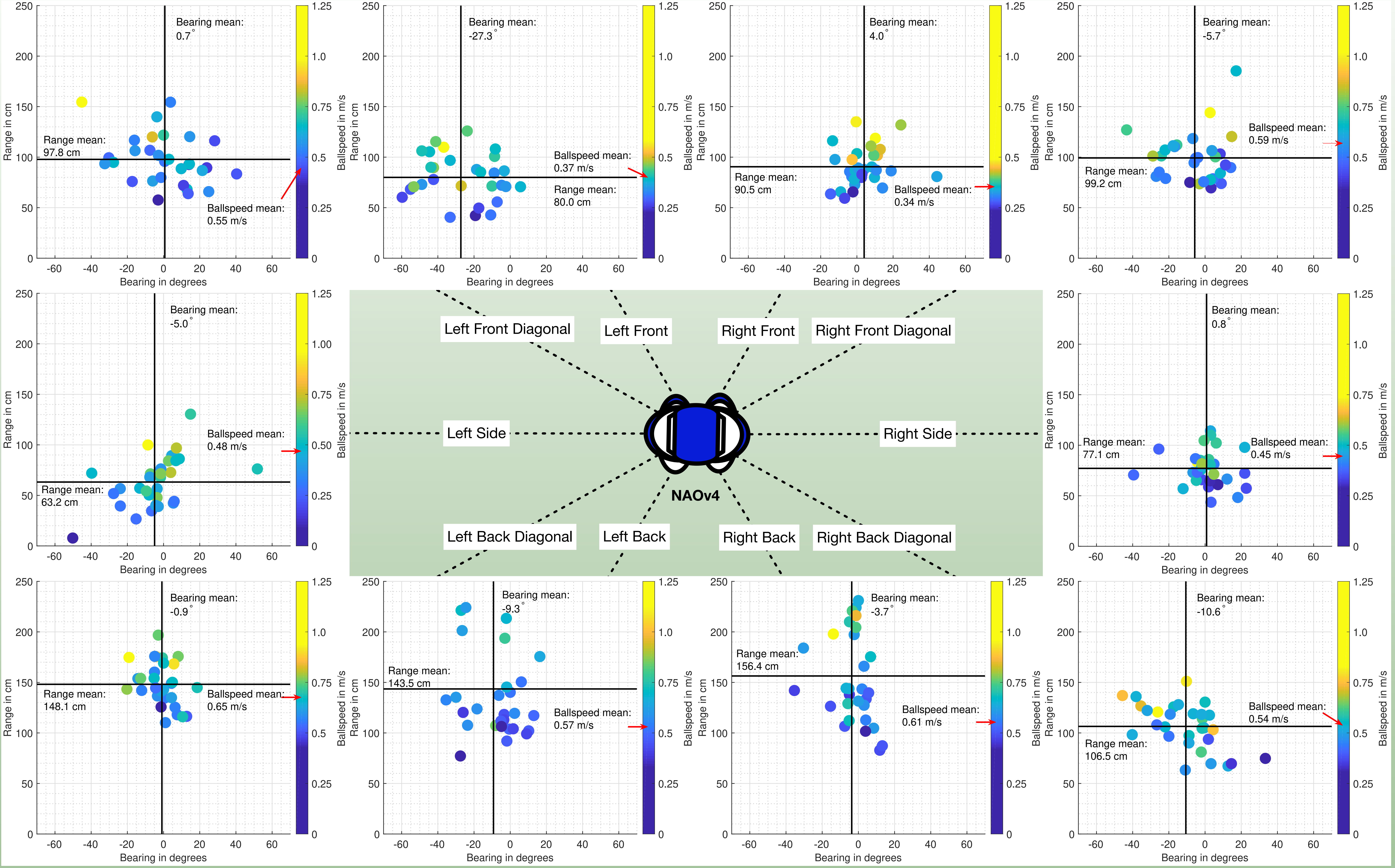
- Walk-kick frame-work generating a kick trajectory
  - any direction, no prior input, while walking
  - guaranteeing reaching a reference trajectory
- Uses kick interpolators from **dynamic kick engine** and walk trajectories generated from **adaptive walking engine** to generate motions
- Reliable in terms of the kick directions and stability of the robot overall (< 1% falling rate)
- Experiments verify that the walk-kick trajectories were consistent with an average absolute bearing of < 6° within any given direction.



Walk-kick motion where the via point is the point of contact with the ball.

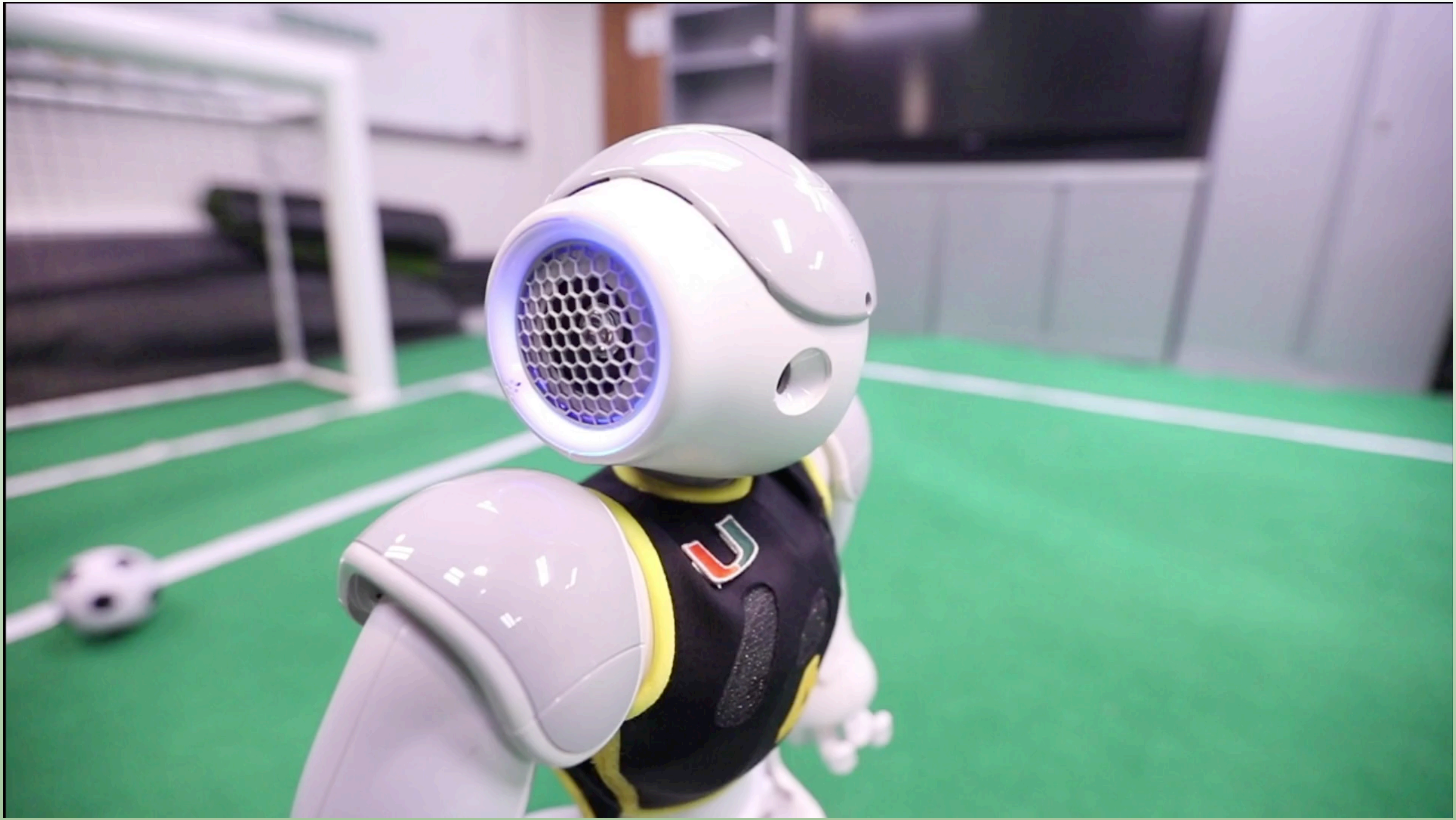
Peña & Visser (*RoboCup Symposium 2018)*

CONTROL (3): WALK-KICK ENGINE



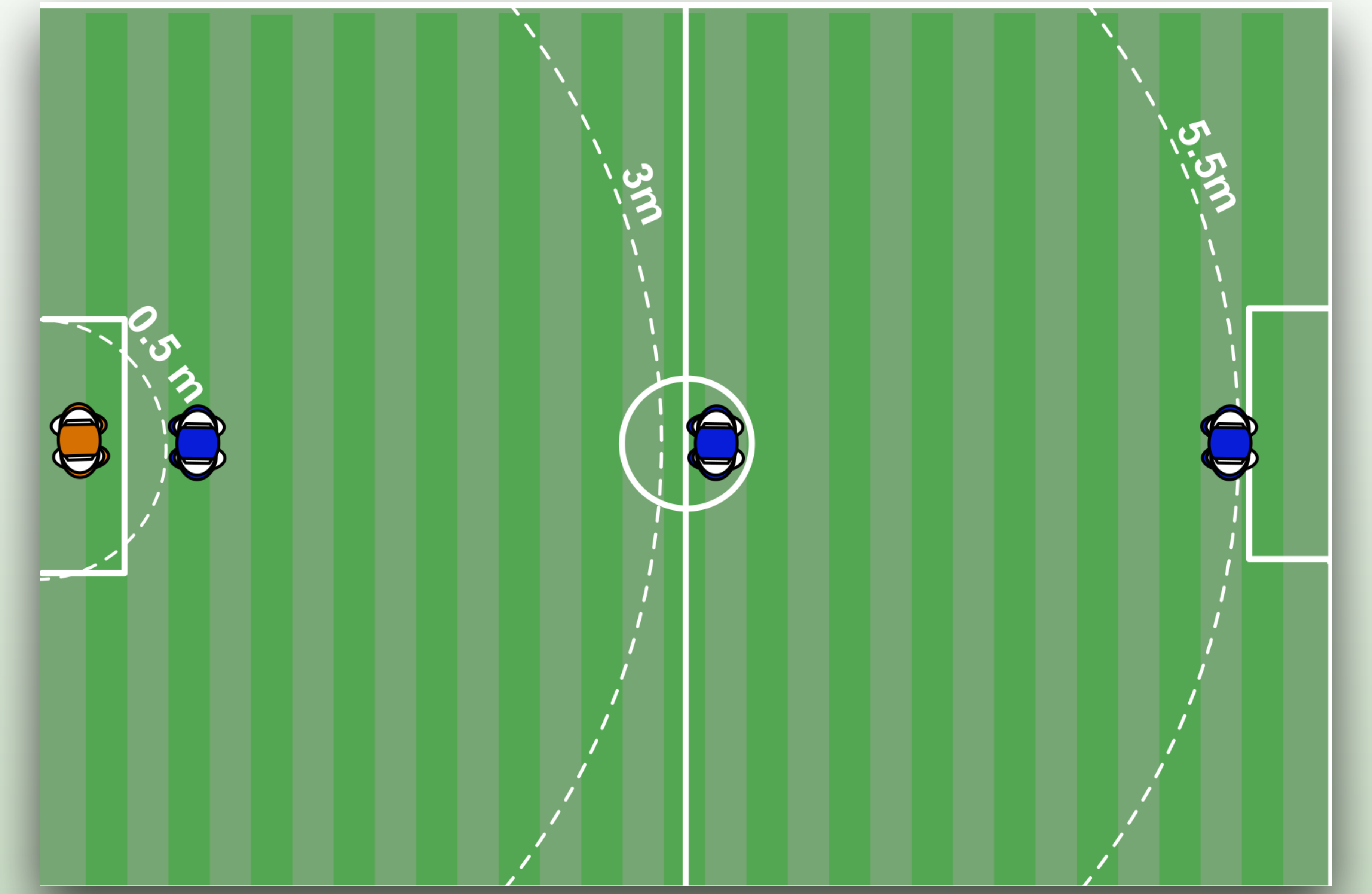
Peña & Visser (*RoboCup Symposium 2018*)

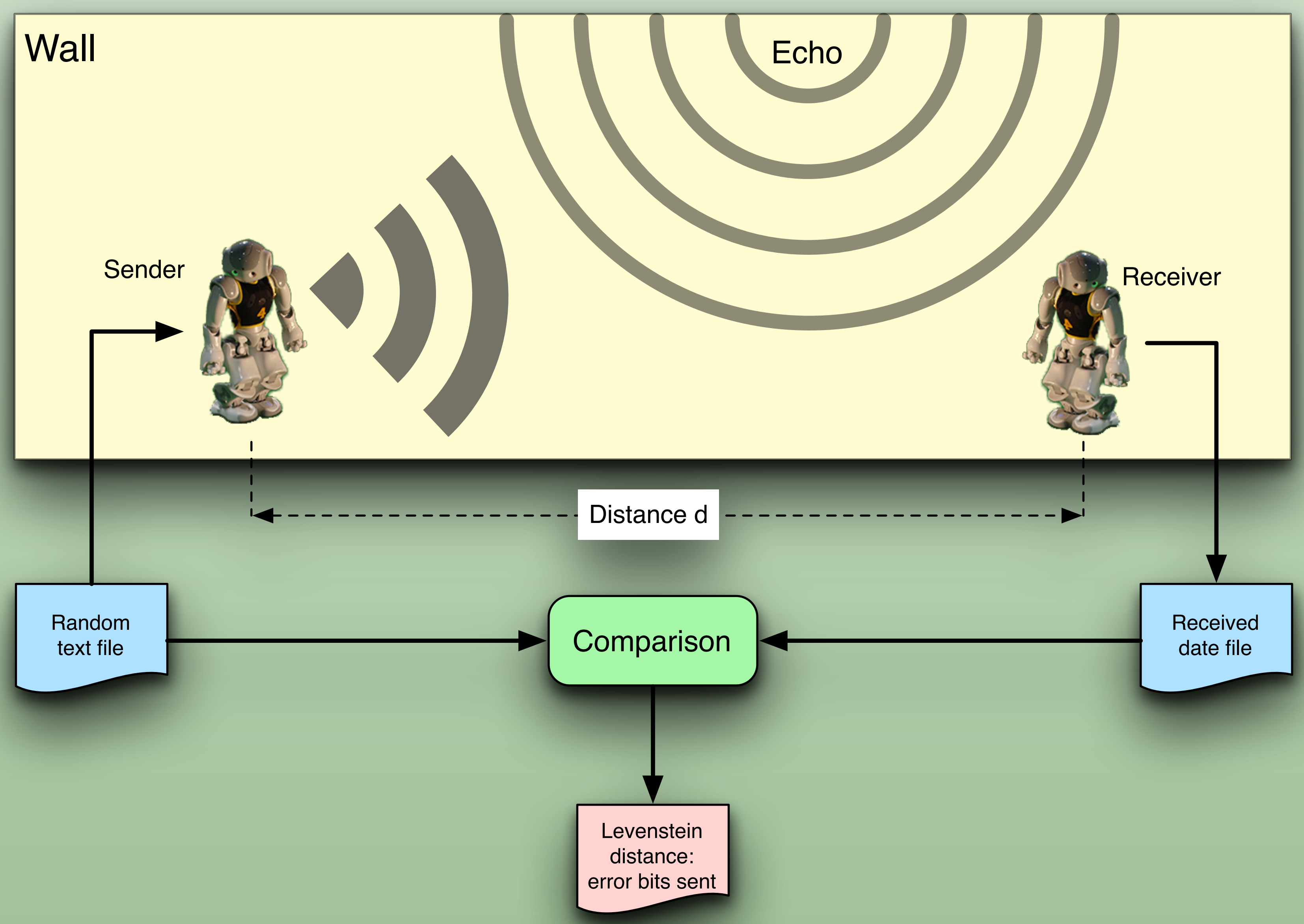
CONTROL (3): WALK-KICK ENGINE



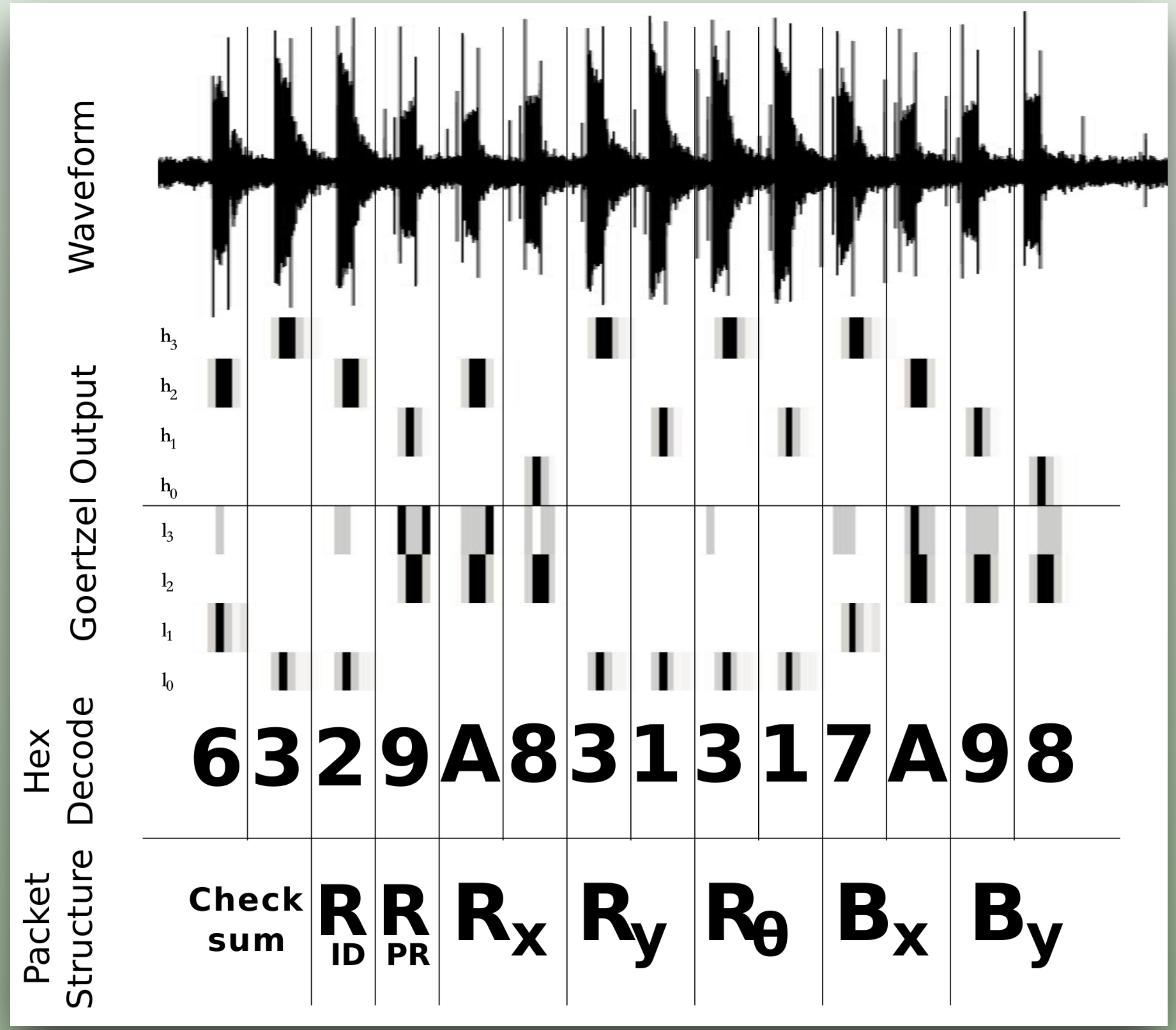
Peña & Visser (*RoboCup Symposium 2018*)

COMMUNICATION (1): DTMF BROADCASTING





- Multi-agent broadcasting based on fixed length DTMF messages



Poore et al. 2015 (*FLAIRS*)

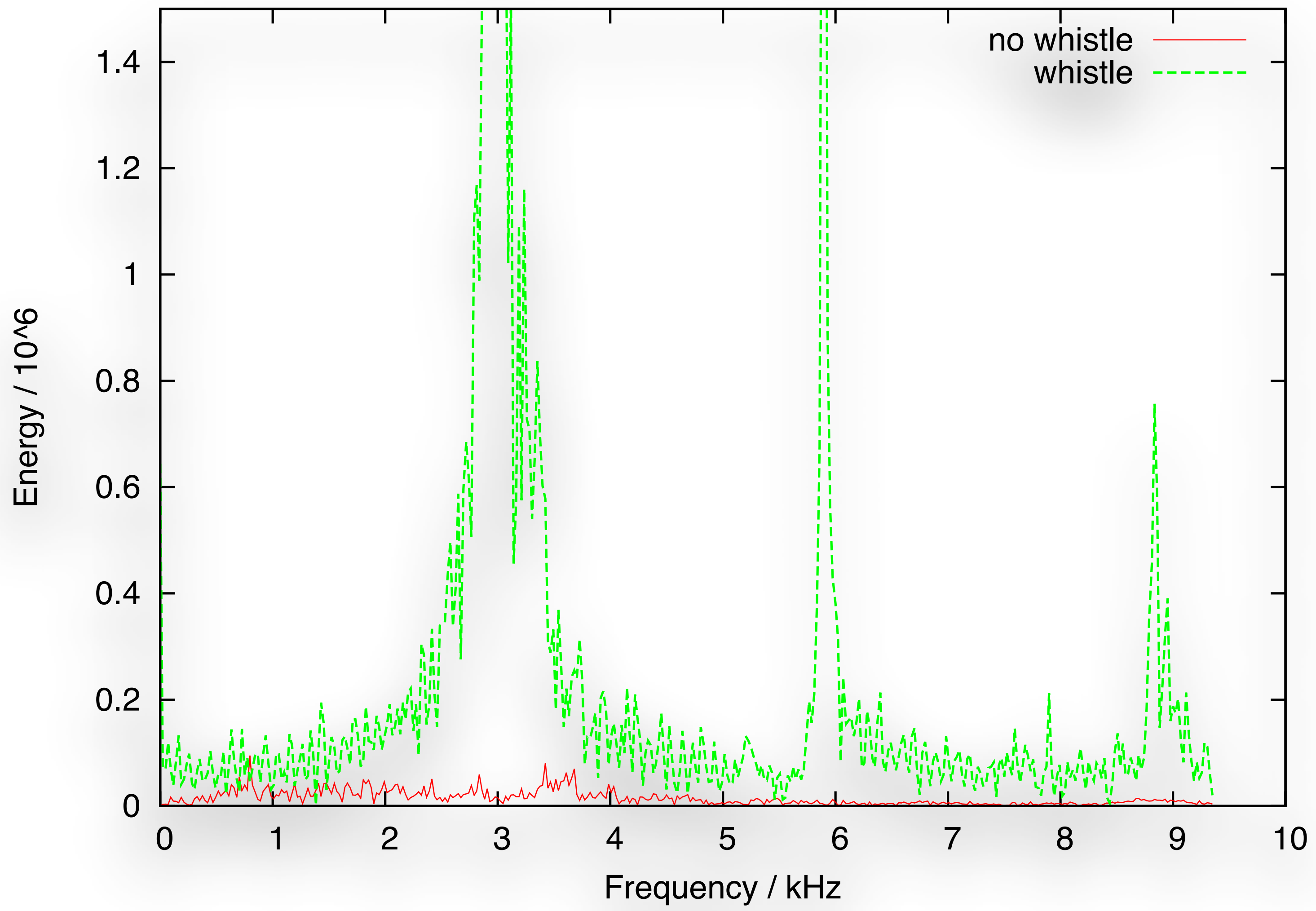
COMMUNICATION (1): DTMF BROADCASTING

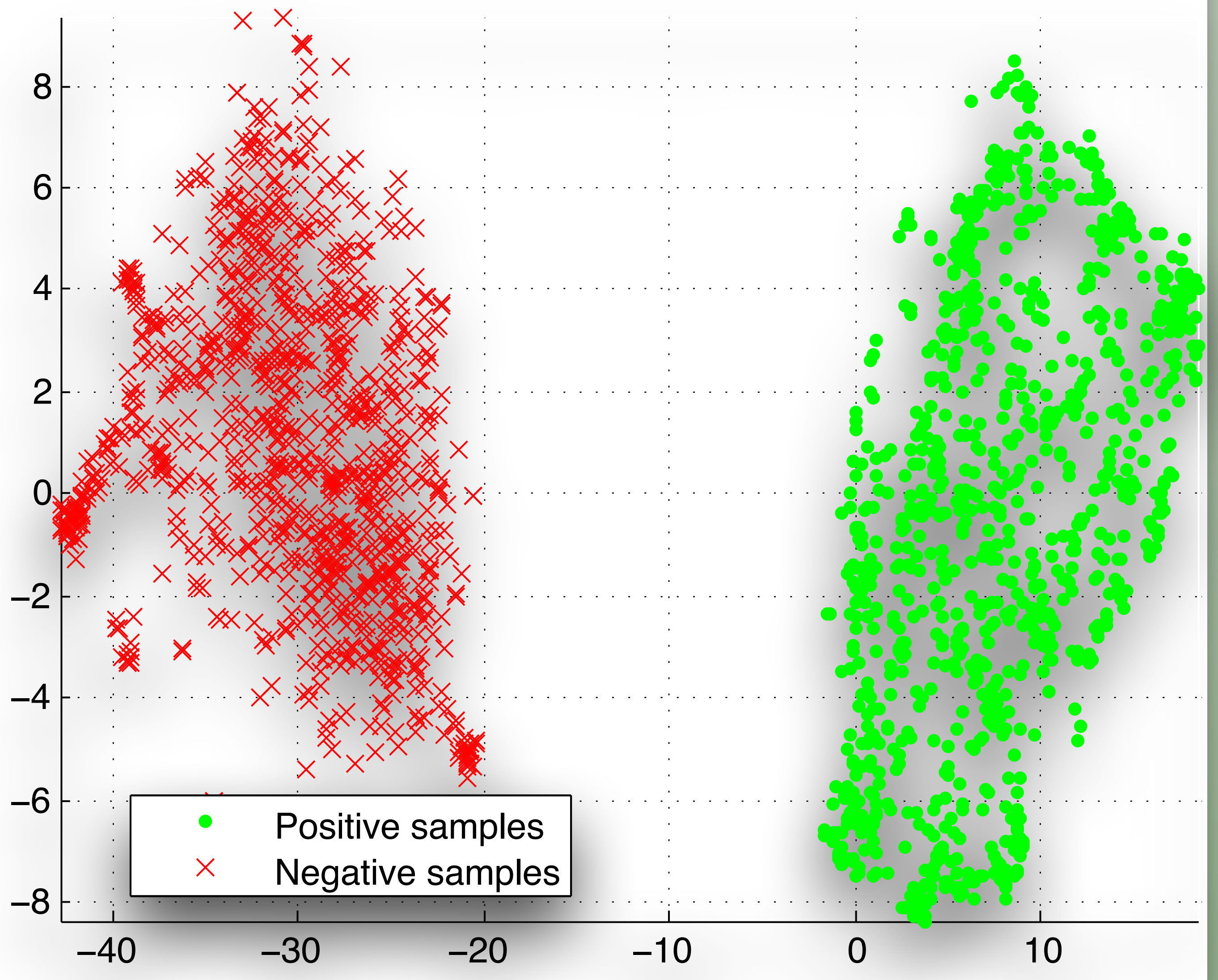
- Multi-agent broadcasting based on fixed length DTMF messages

Poore et al. 2015 (*FLAIRS*)

COMMUNICATION (2): WHISTLE DETECTION

- Based on Logistic Regression with L2-norm Regularization

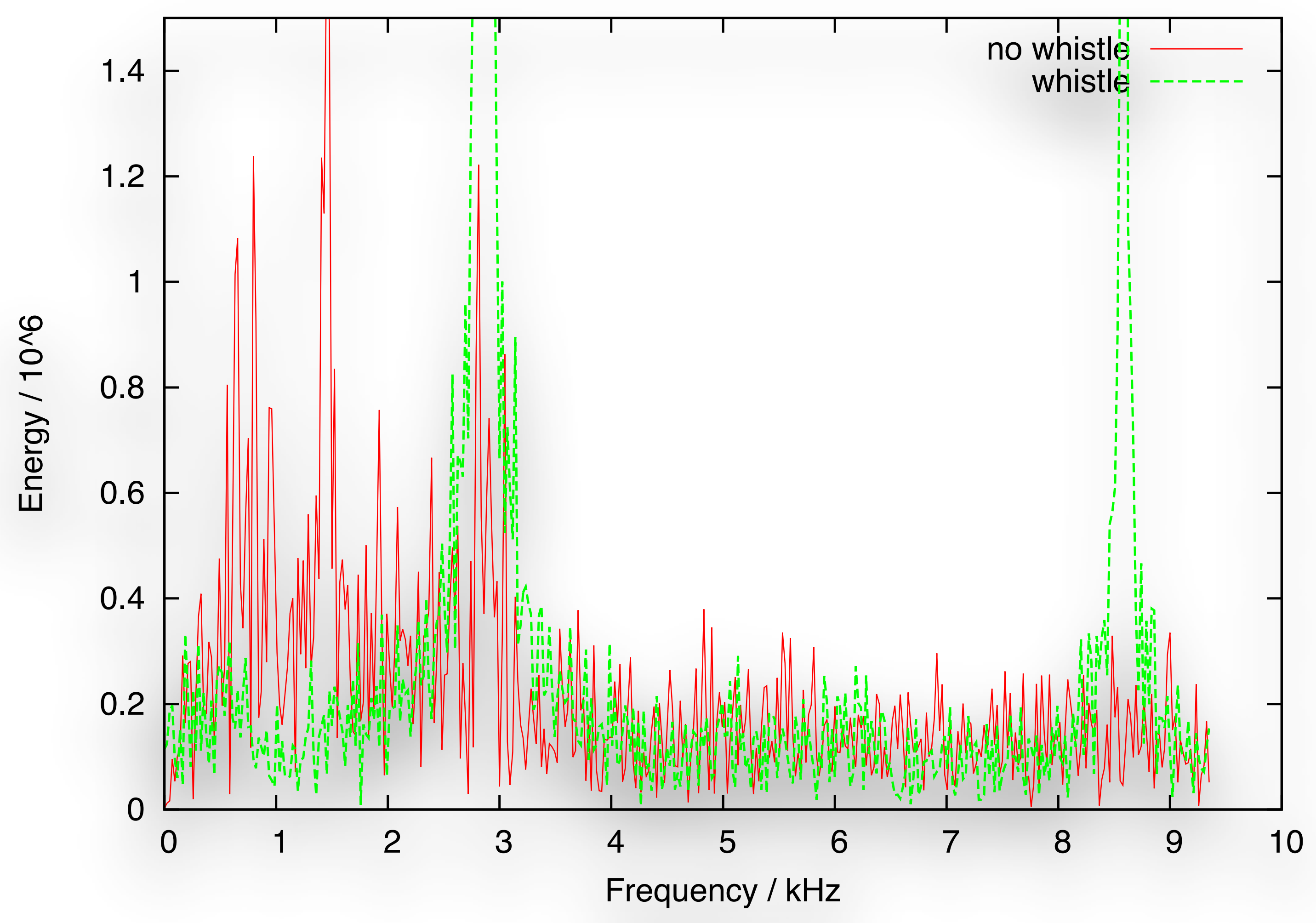


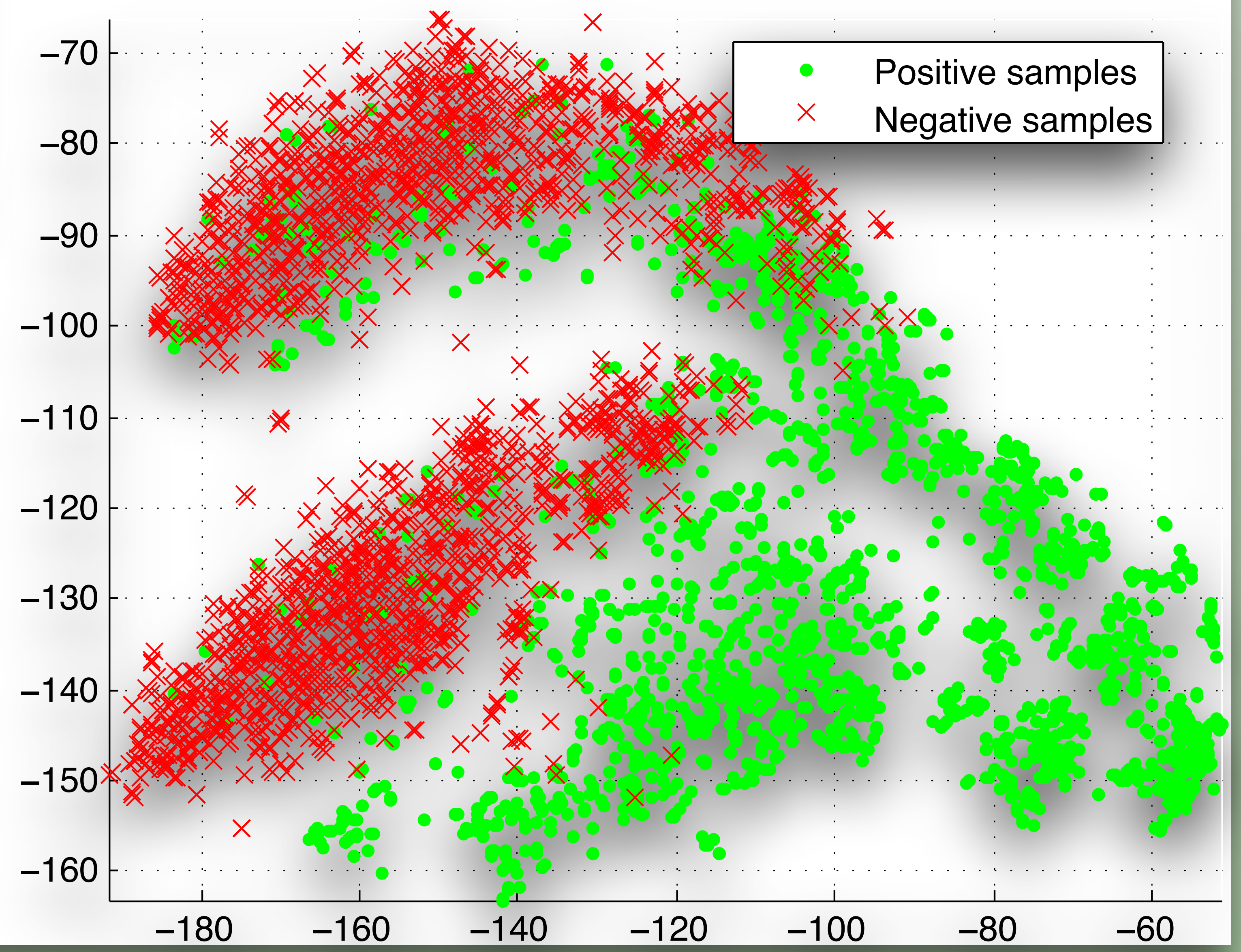


Poore et al. 2014 (*RoboCup Symposium*)

COMMUNICATION (2): WHISTLE DETECTION

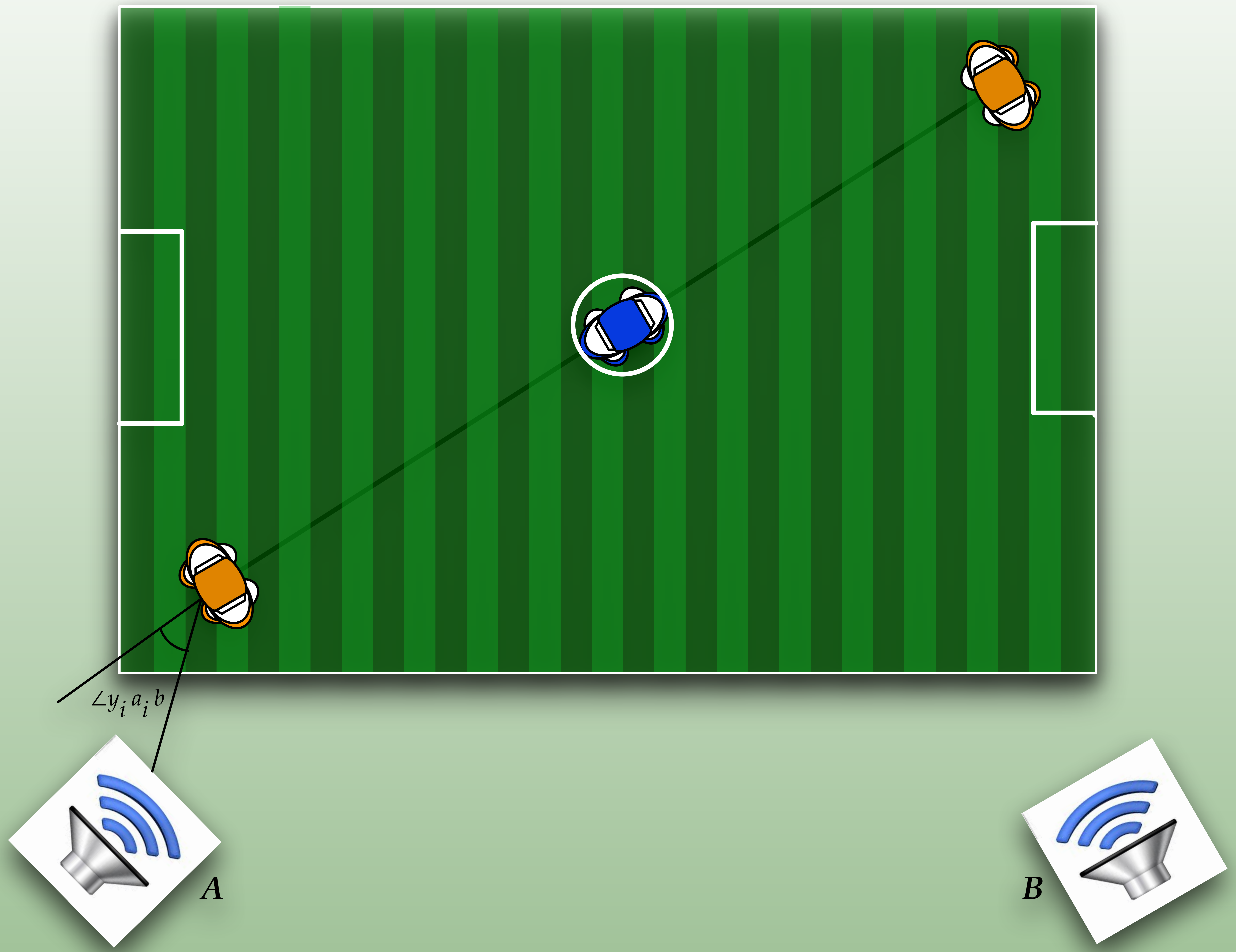
- Based on Logistic Regression with L2-norm Regularization





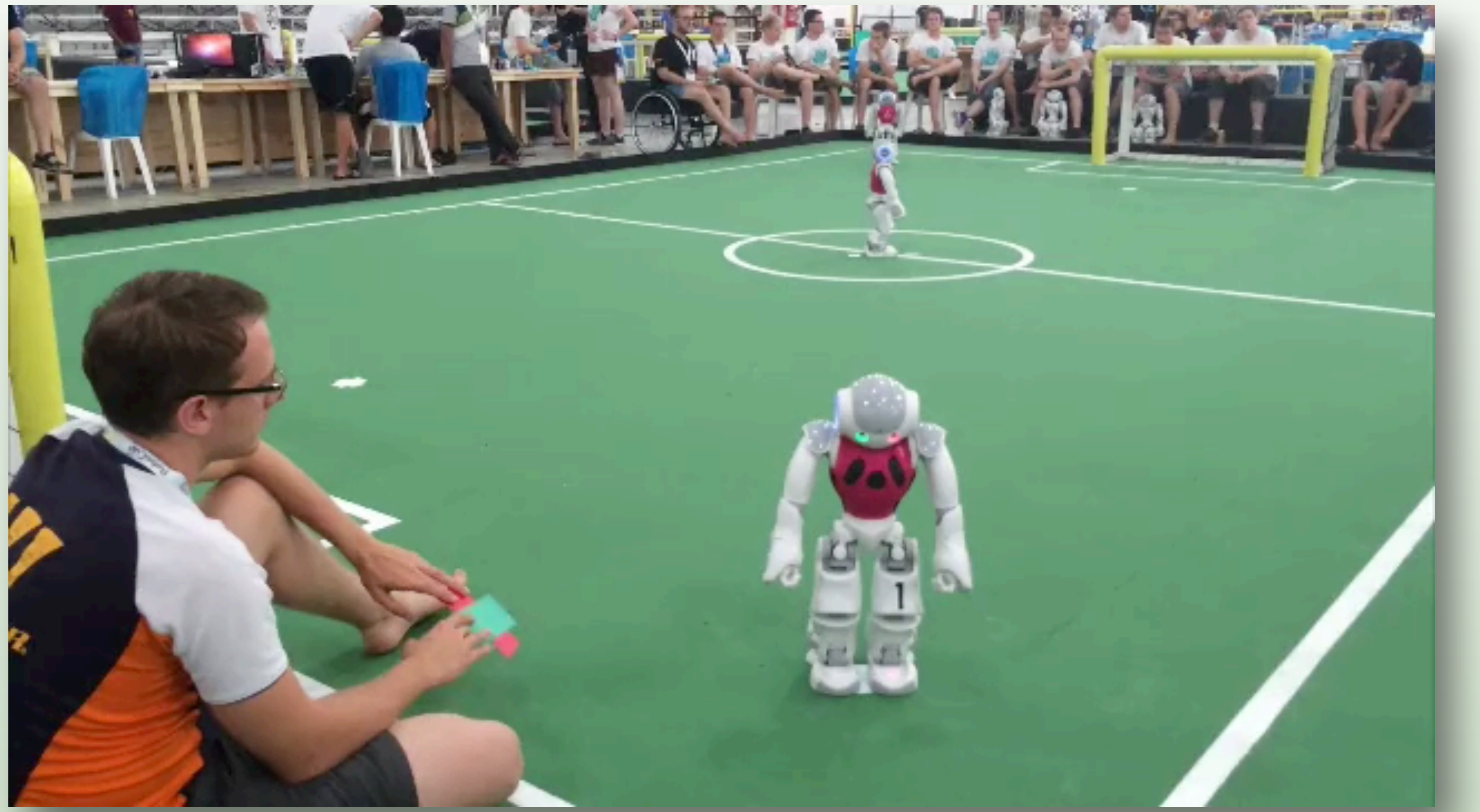
Poore et al. 2014 (*RoboCup Symposium)*

COMMUNICATION (2): WHISTLE DETECTION



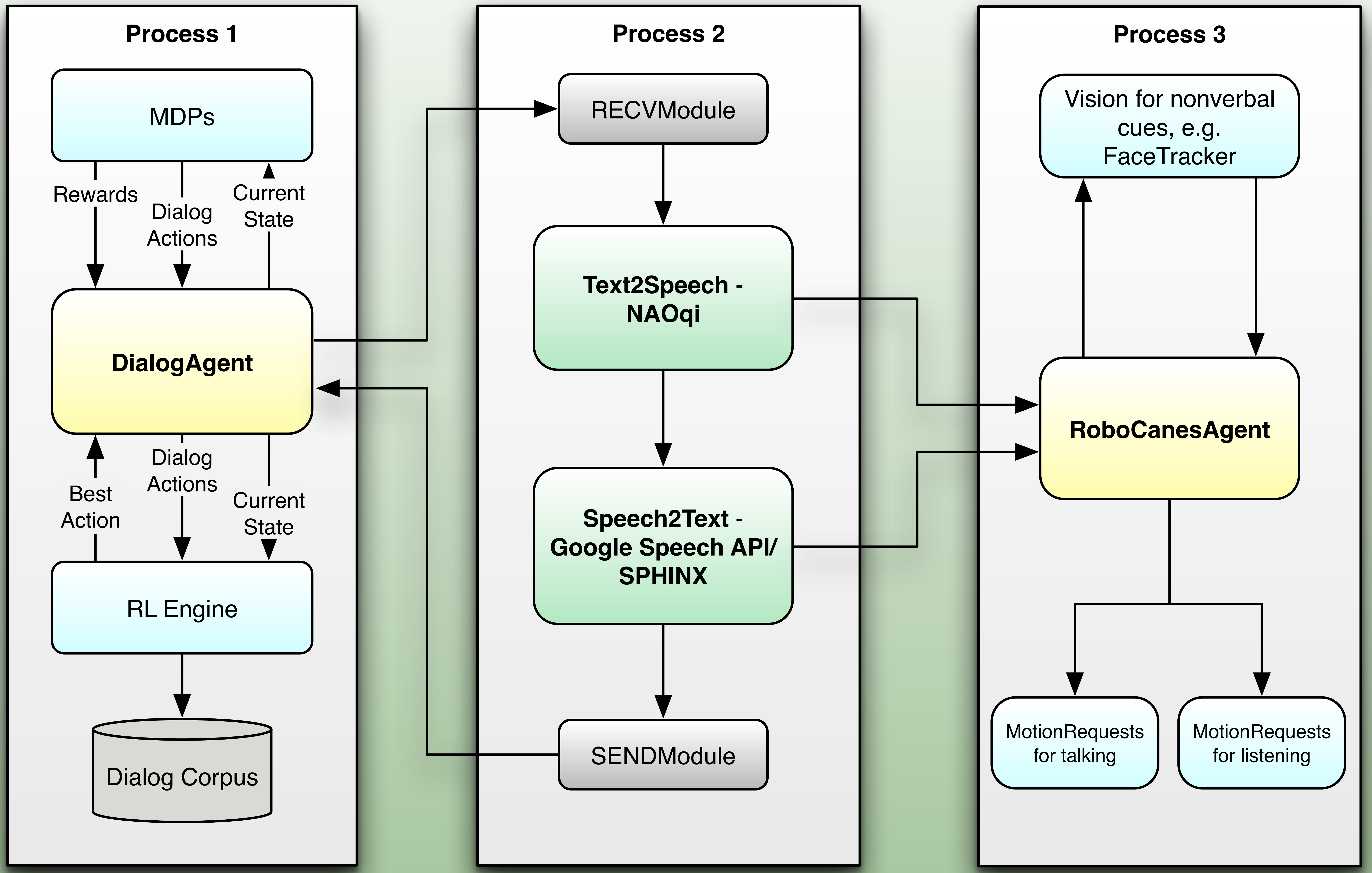
Poore et al. 2014 (*RoboCup Symposium*)

COMMUNICATION (2): WHISTLE DETECTION



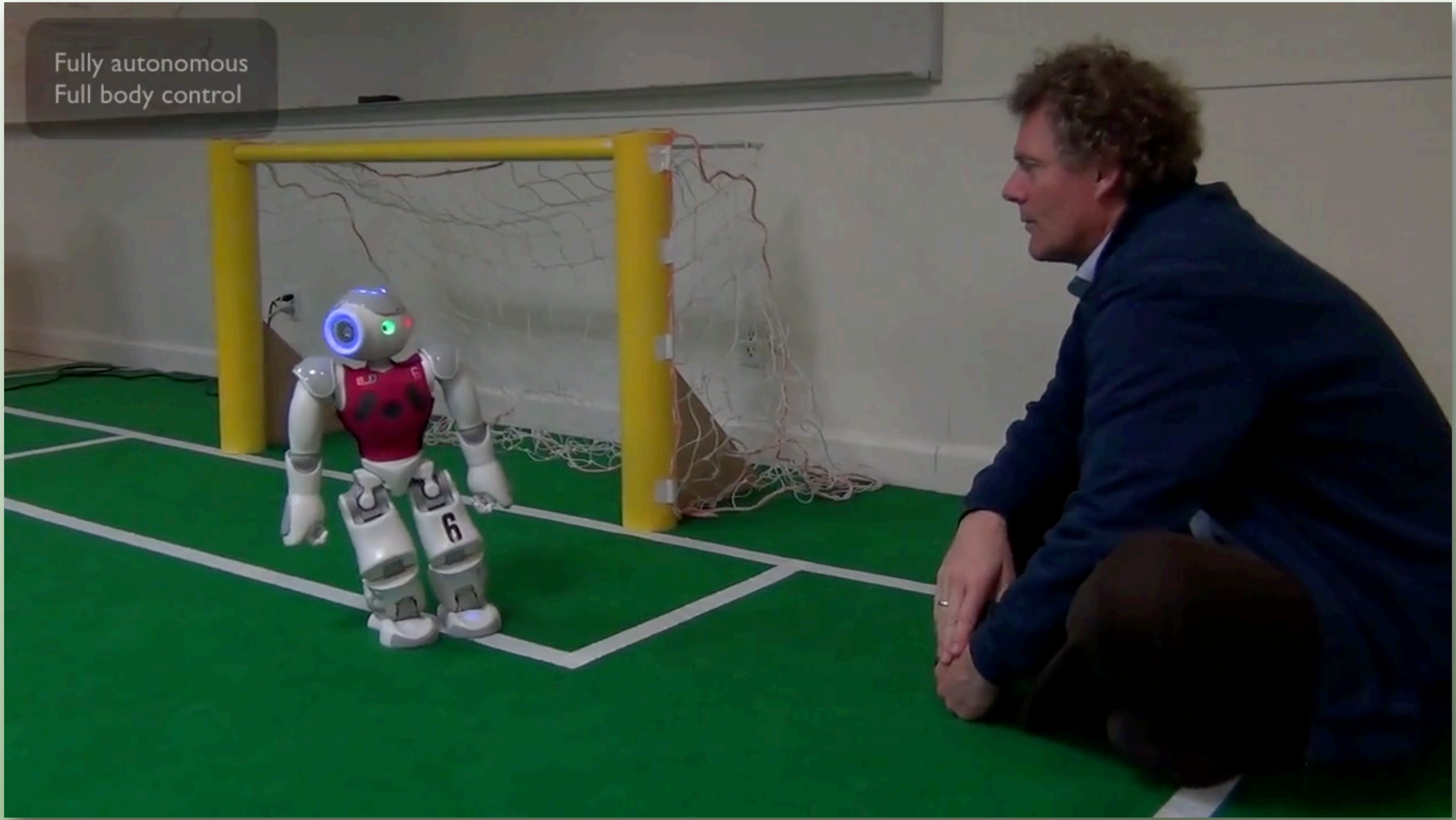
Poore et al. 2014 (*RoboCup Symposium*)

COMMUNICATION/CONTROL (3): SPOKEN DIALOG/HRI



Abeyruwan et al. 2014 (*AAAI Symposium*)

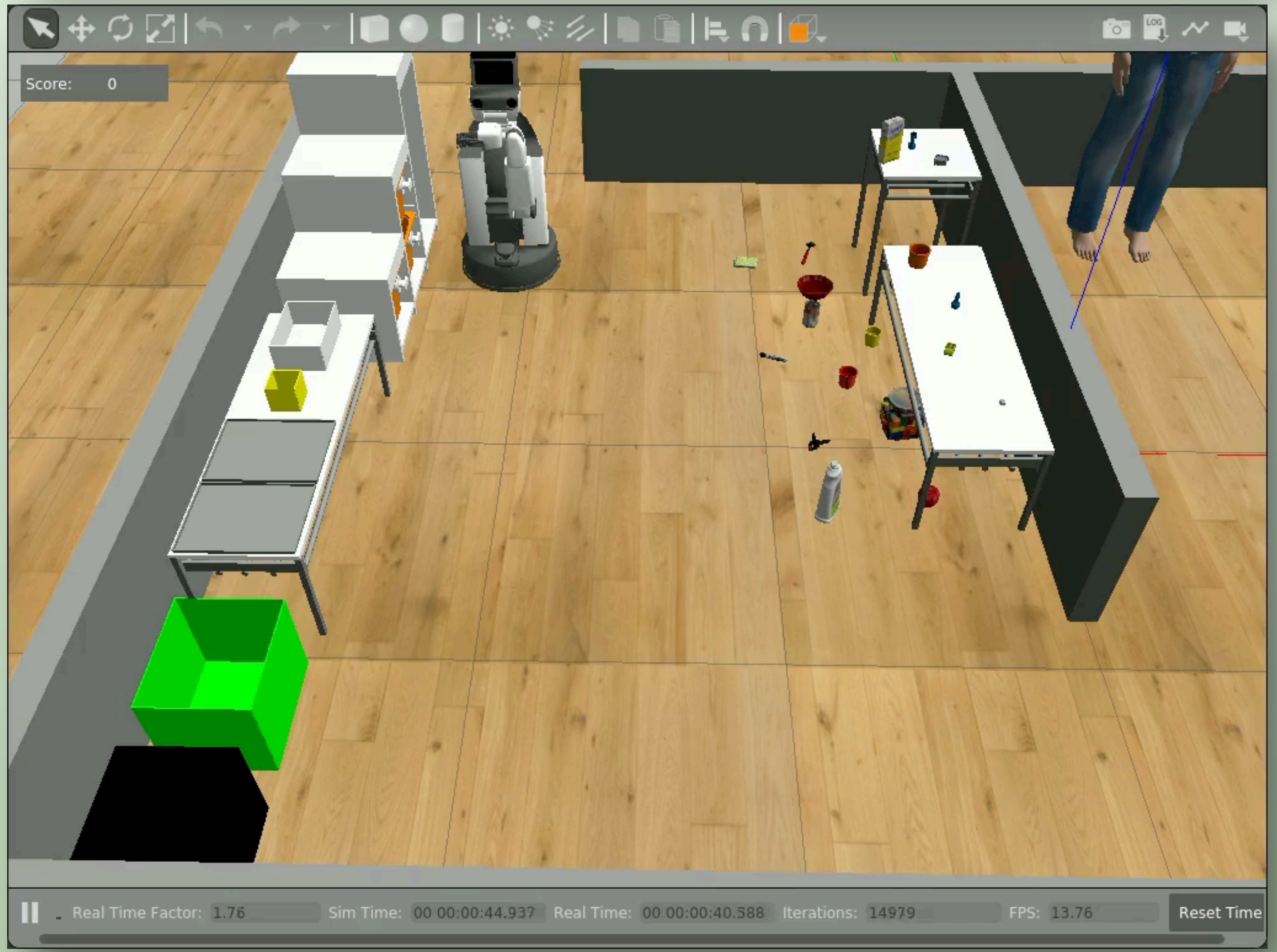
COMMUNICATION/CONTROL (3): SPOKEN DIALOG/HRI



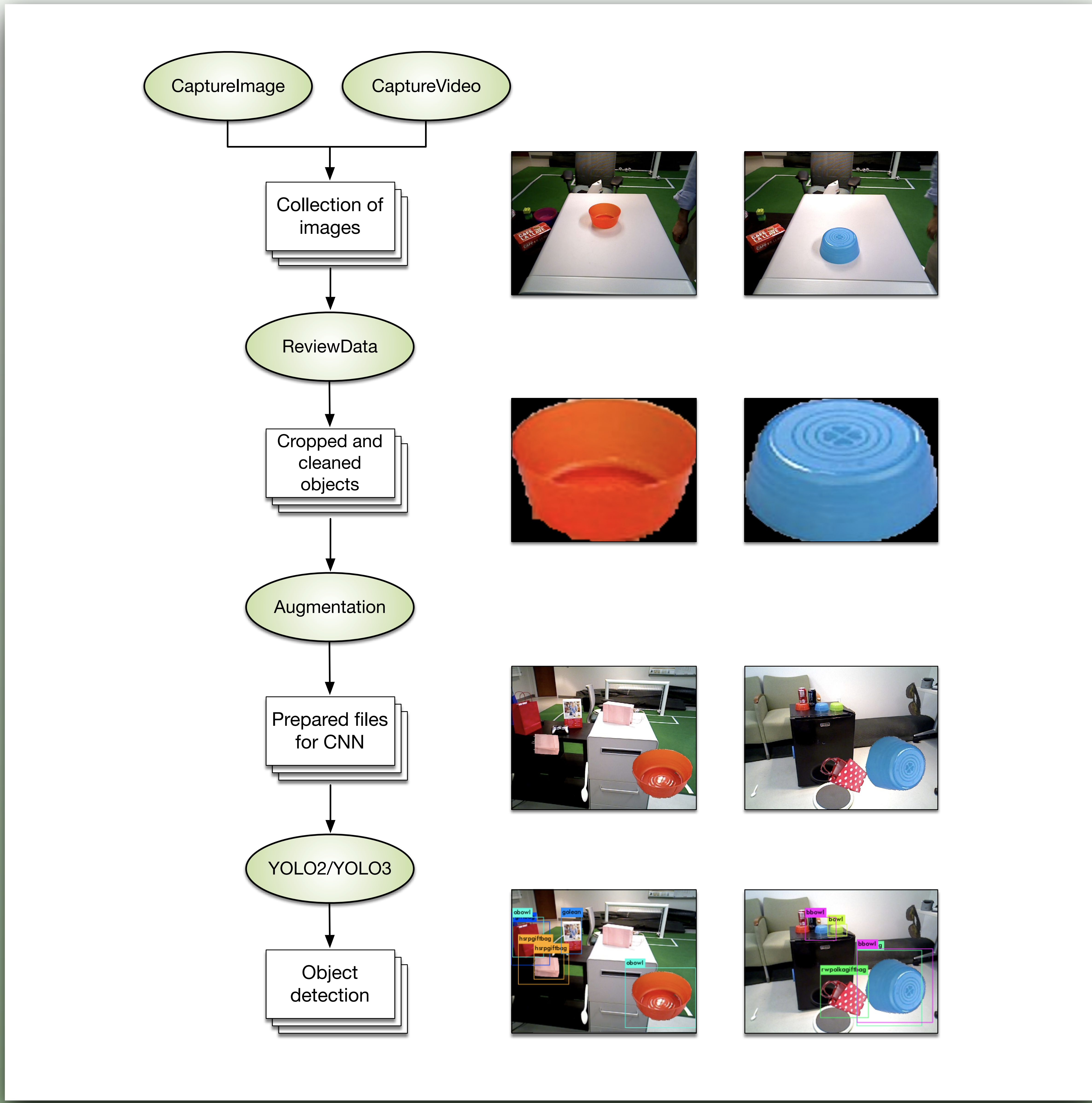
Abeyruwan et al. 2014 (*AAAI Symposium)*

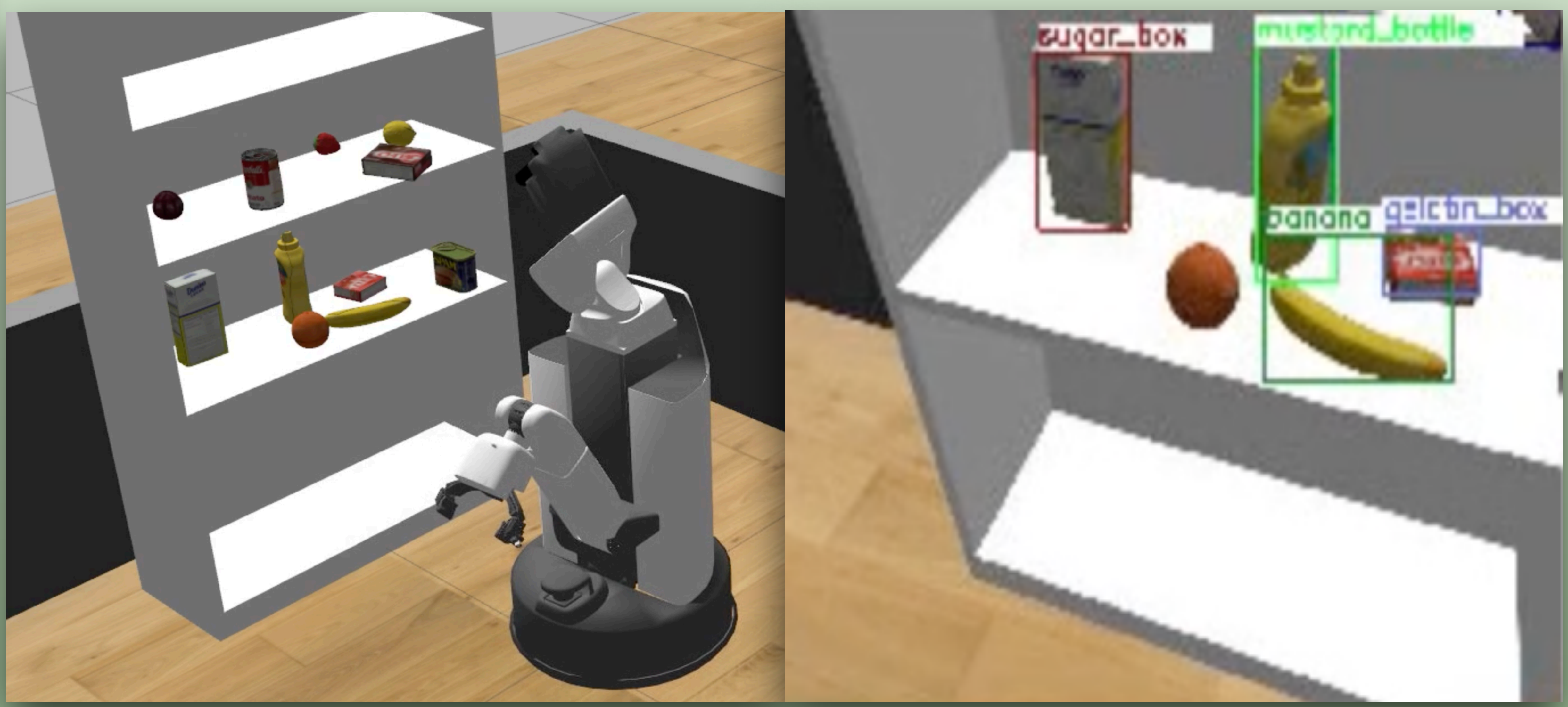
2. MANIPULATION SYSTEM

- Task and motion planning
  - Creation of high-level commands and collision-free trajectories to achieve goal
- State estimation and perception
  - Infer relevant quantities from sensor data (objects, drawers, manipulators, contacts/collisions, …)
- Object grasping and placement (pick-and-place)
  - Determine good grasps for objects given relevant constraints (gripper opening, local geometry, placement)
- Trajectory generation and control
  - Real-time, reactive generation of control commands to move robot (or parts) safely toward goal

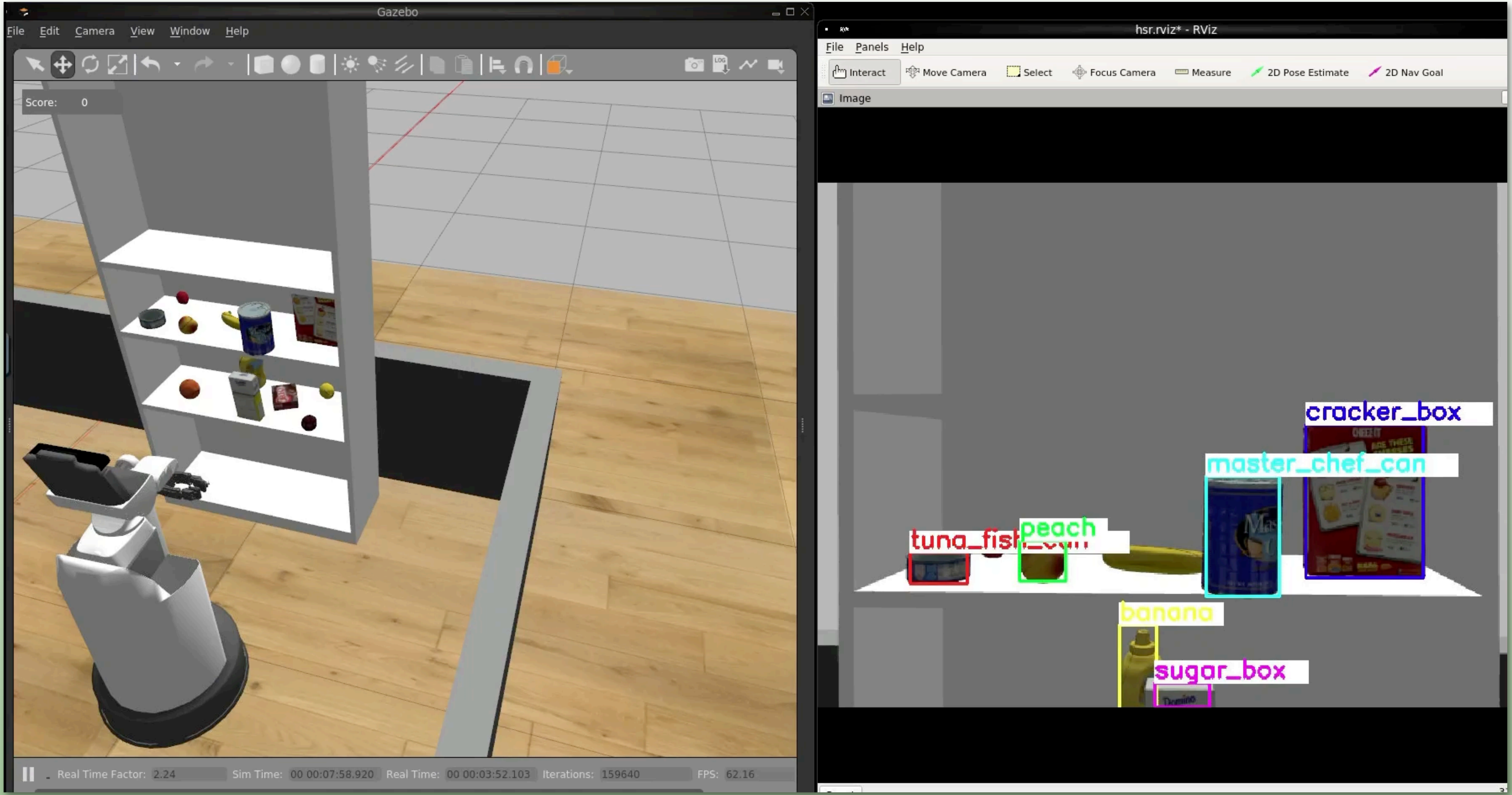


STATE ESTIMATION AND PERCEPTION

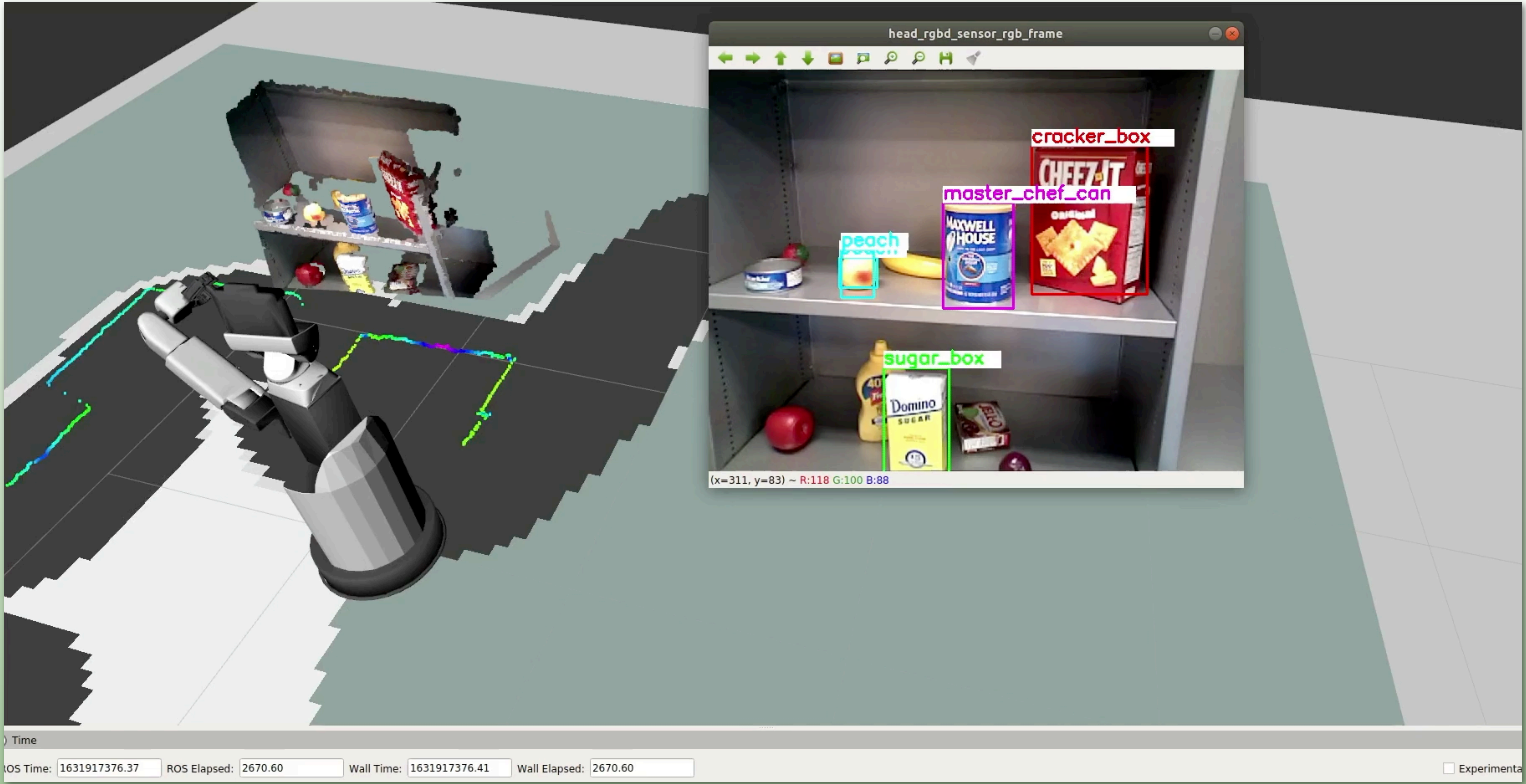




STATE ESTIMATION AND PERCEPTION

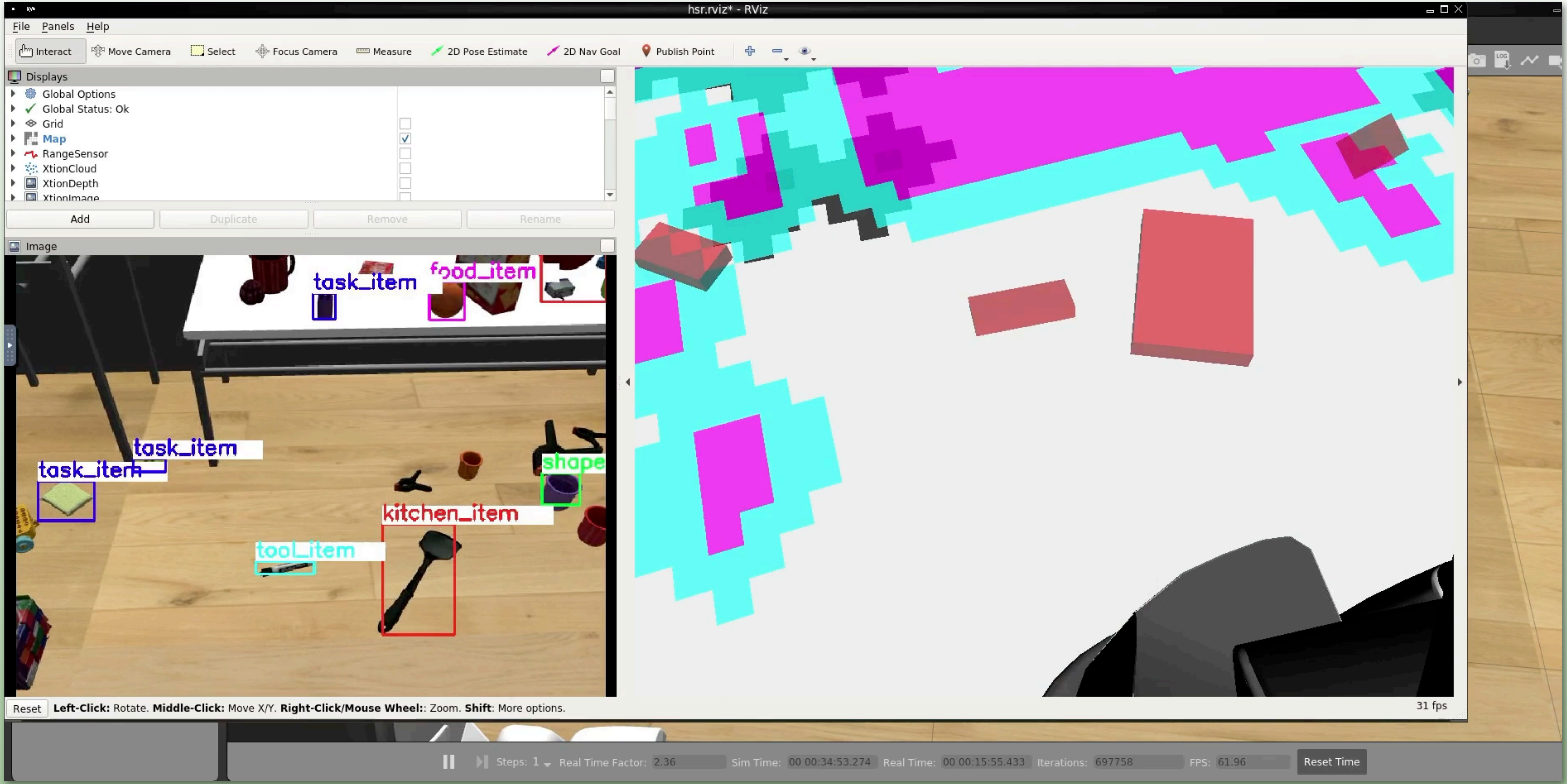


STATE ESTIMATION AND PERCEPTION



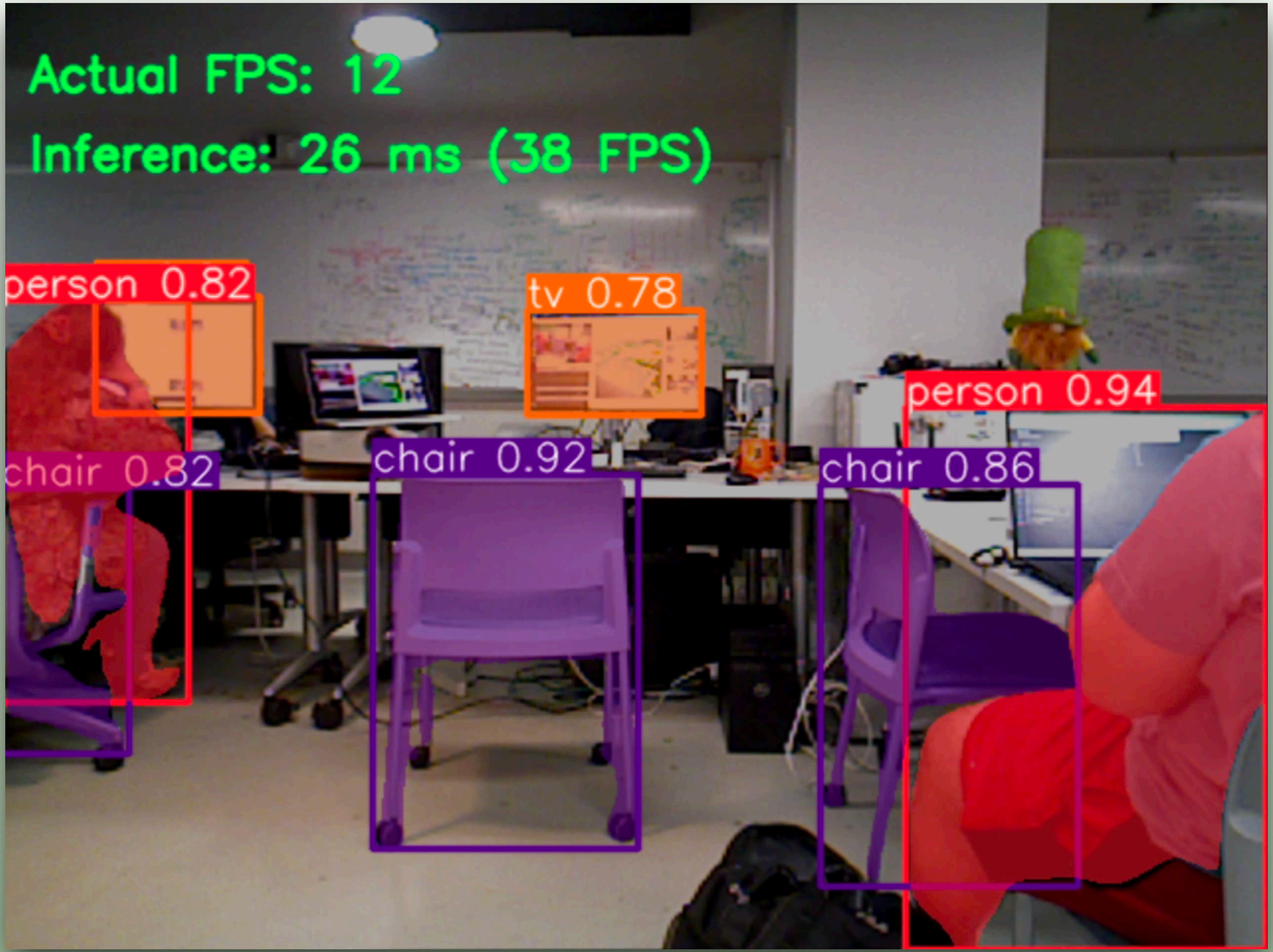
STATE ESTIMATION AND PERCEPTION

STATE ESTIMATION AND PERCEPTION

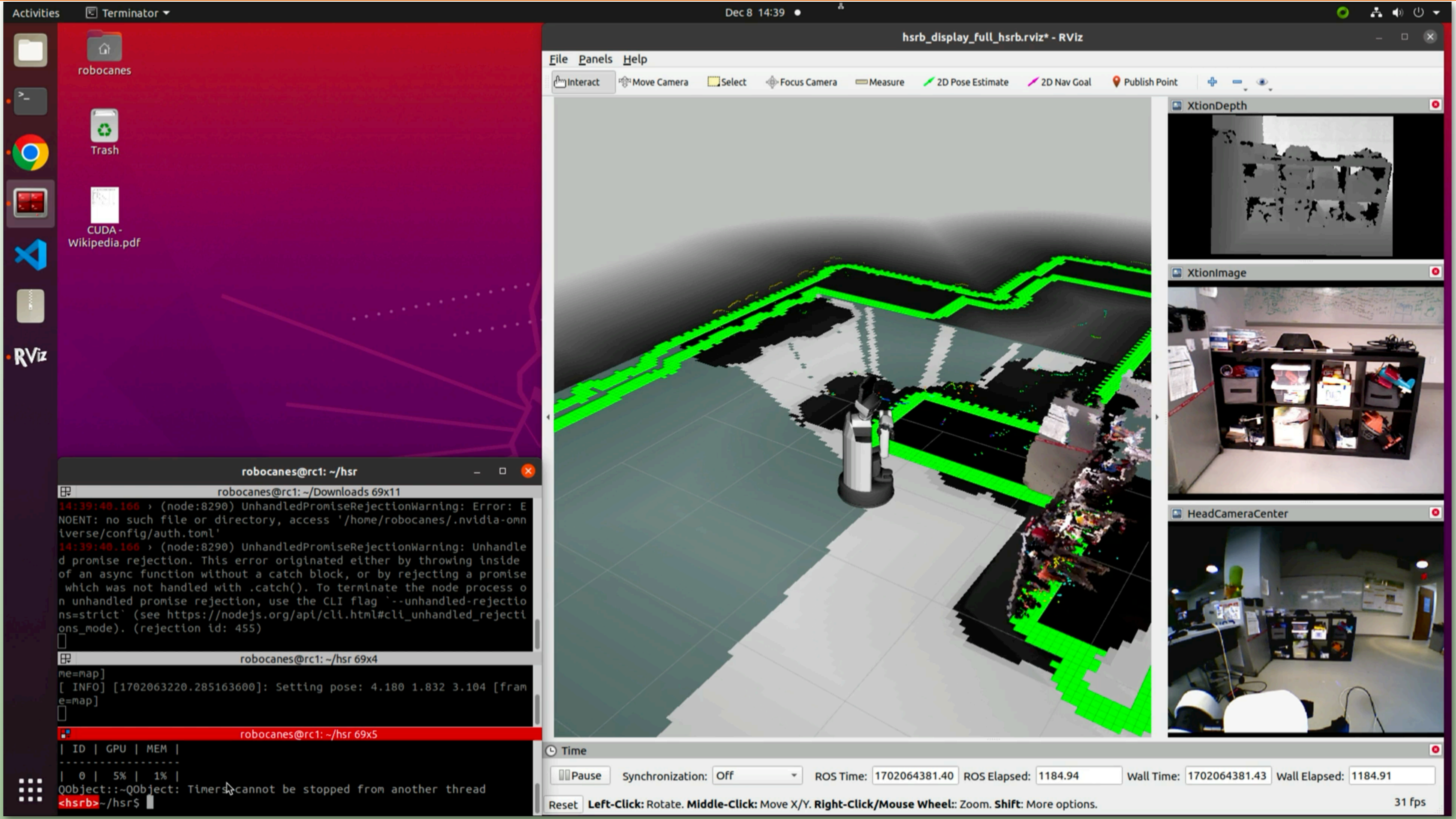


STATE ESTIMATION AND PERCEPTION (2) SEGMENTATION

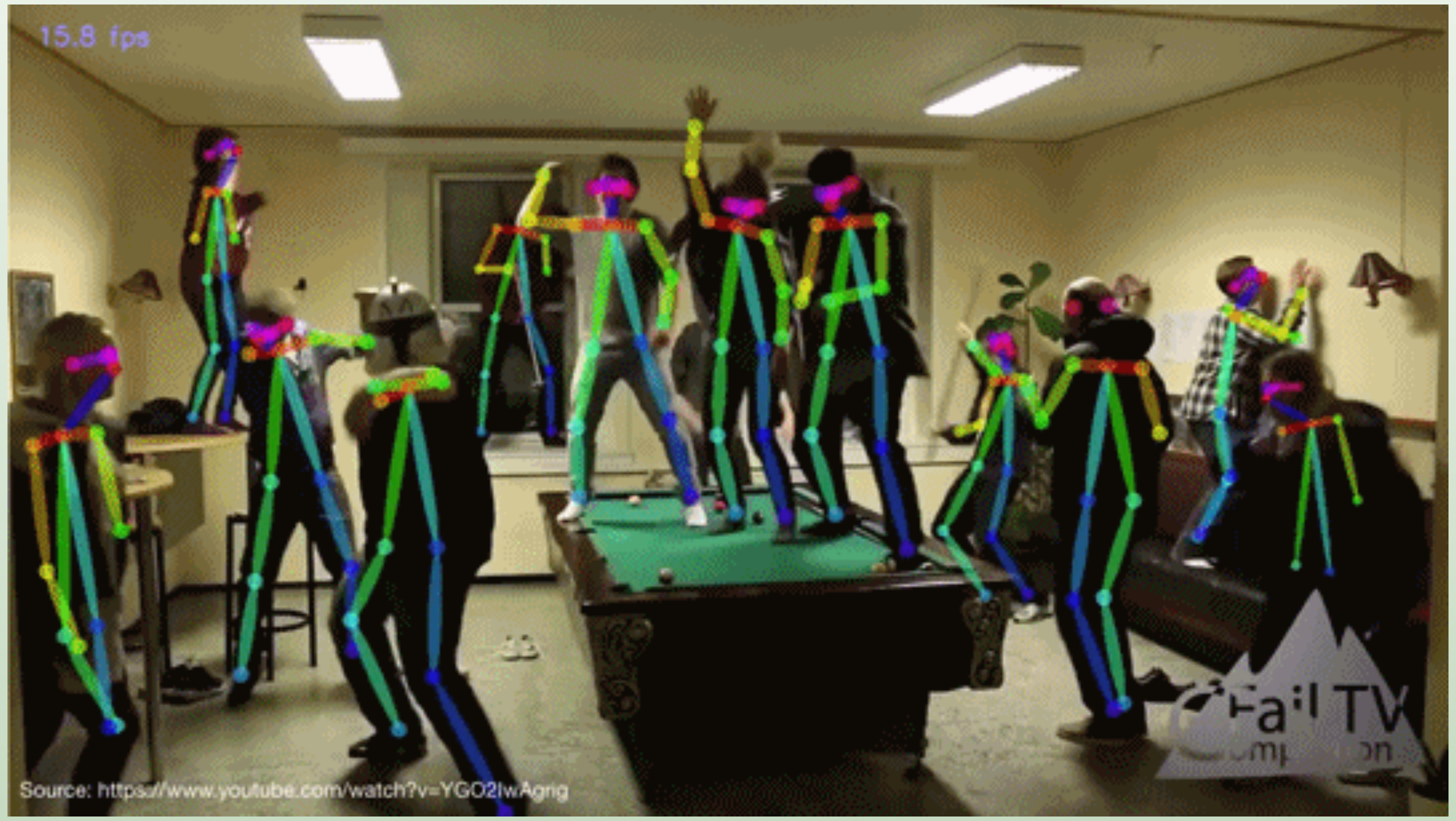
- Segmentation
  - Yolov8
    - Object detection
    - Object classification
    - Object segmentation
  - Here: yolov8x-seg standard model
    - 71.8M params
    - Avg inference 26 ms on RTX3070
    - Suitable for real-time



STATE ESTIMATION AND PERCEPTION (2) SEGMENTATION

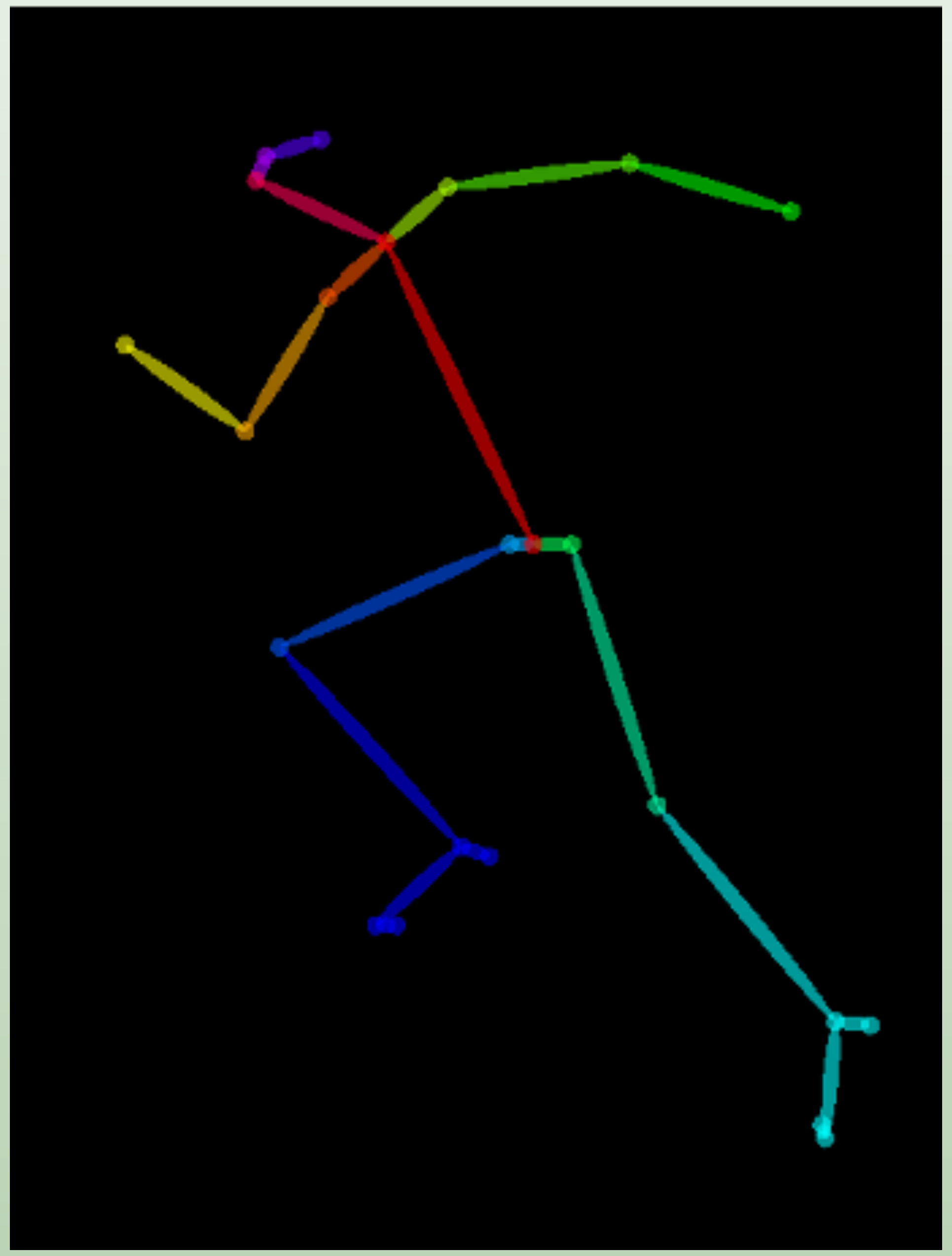


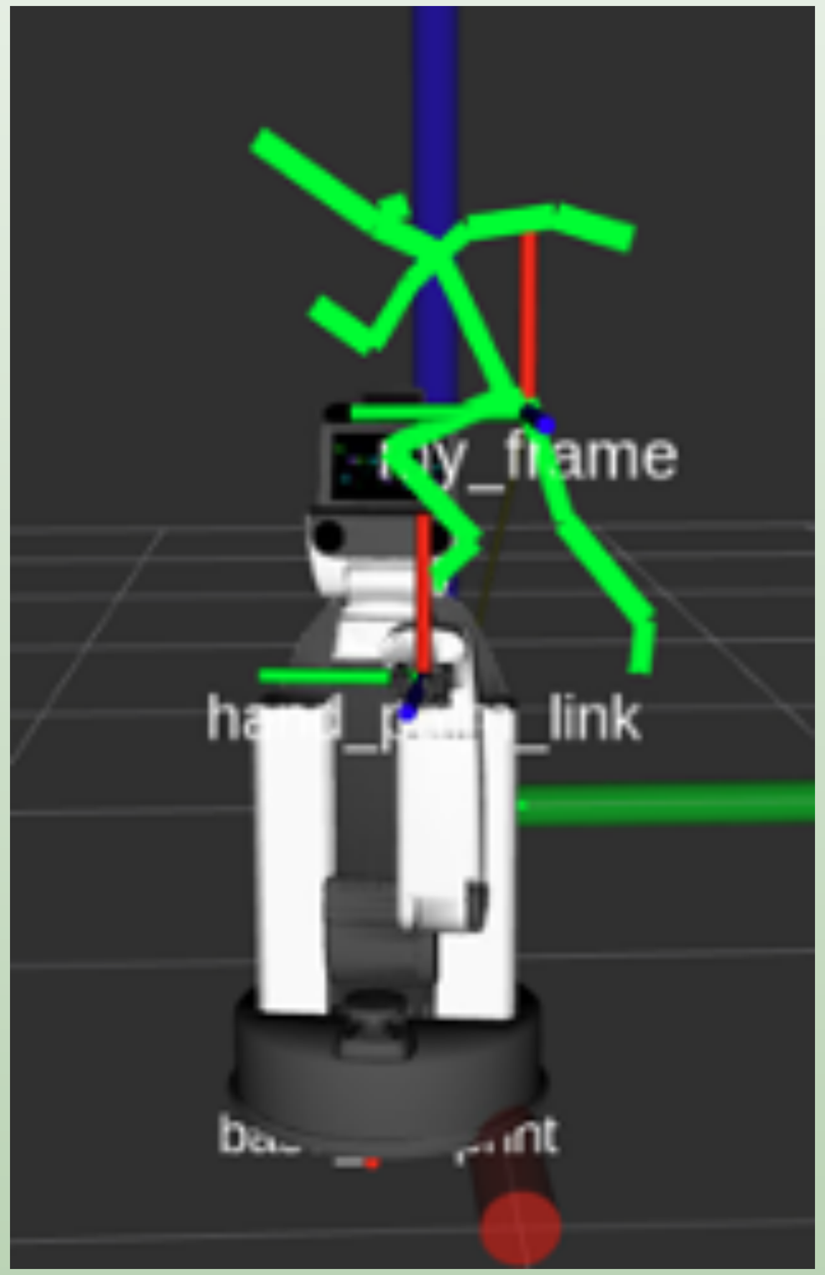
- OpenPose
  - Library containing a set of algorithms & neural networks that automatically detect humans and their current pose in a picture, video, or live image (Wei et al., 2016)
  - Goal: HSR retrieves images from its camera and draws the stick figure (from the augmented image) on a piece of paper (or a whiteboard).



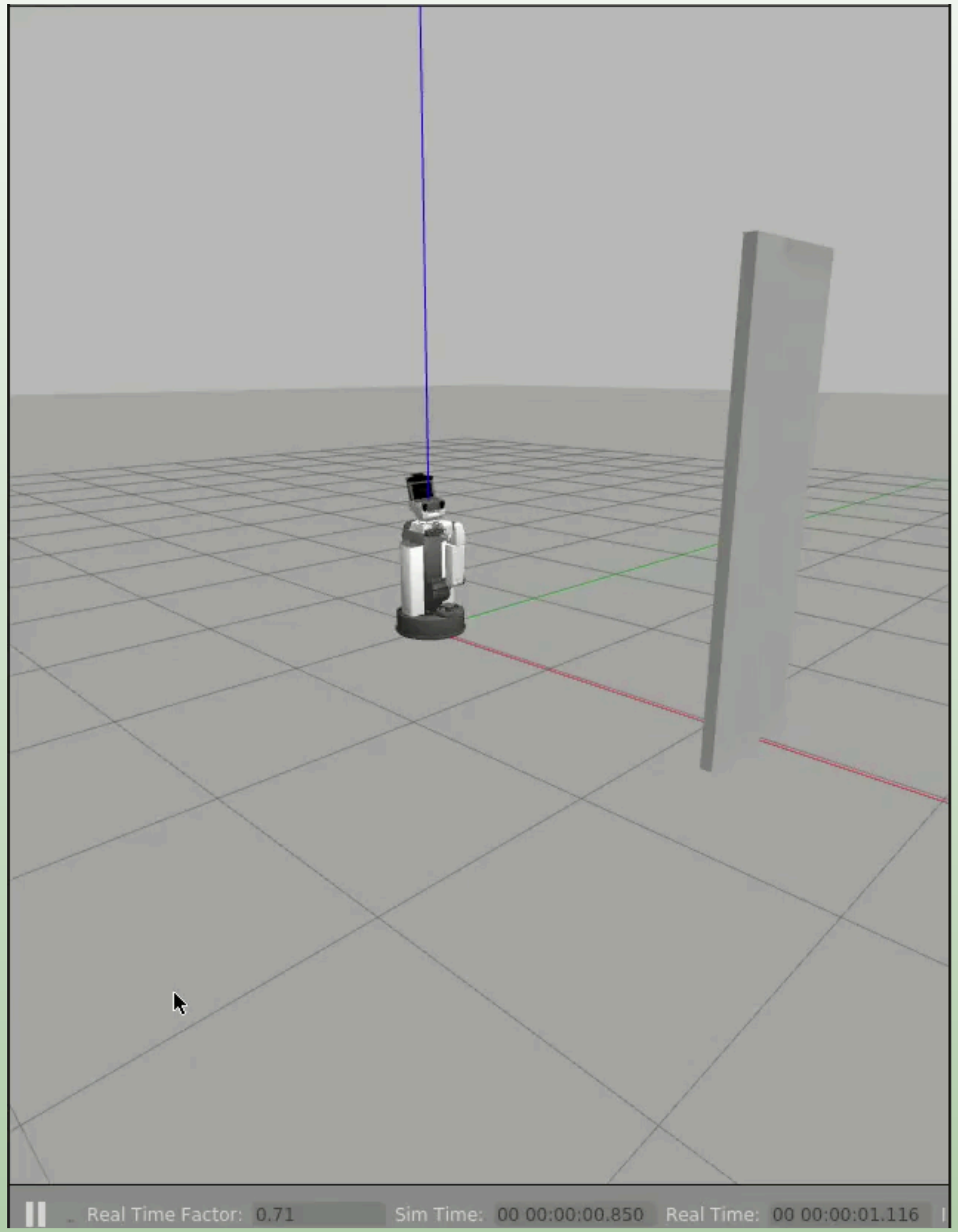
Source: https://cmu-perceptual-computing-lab.github.io/

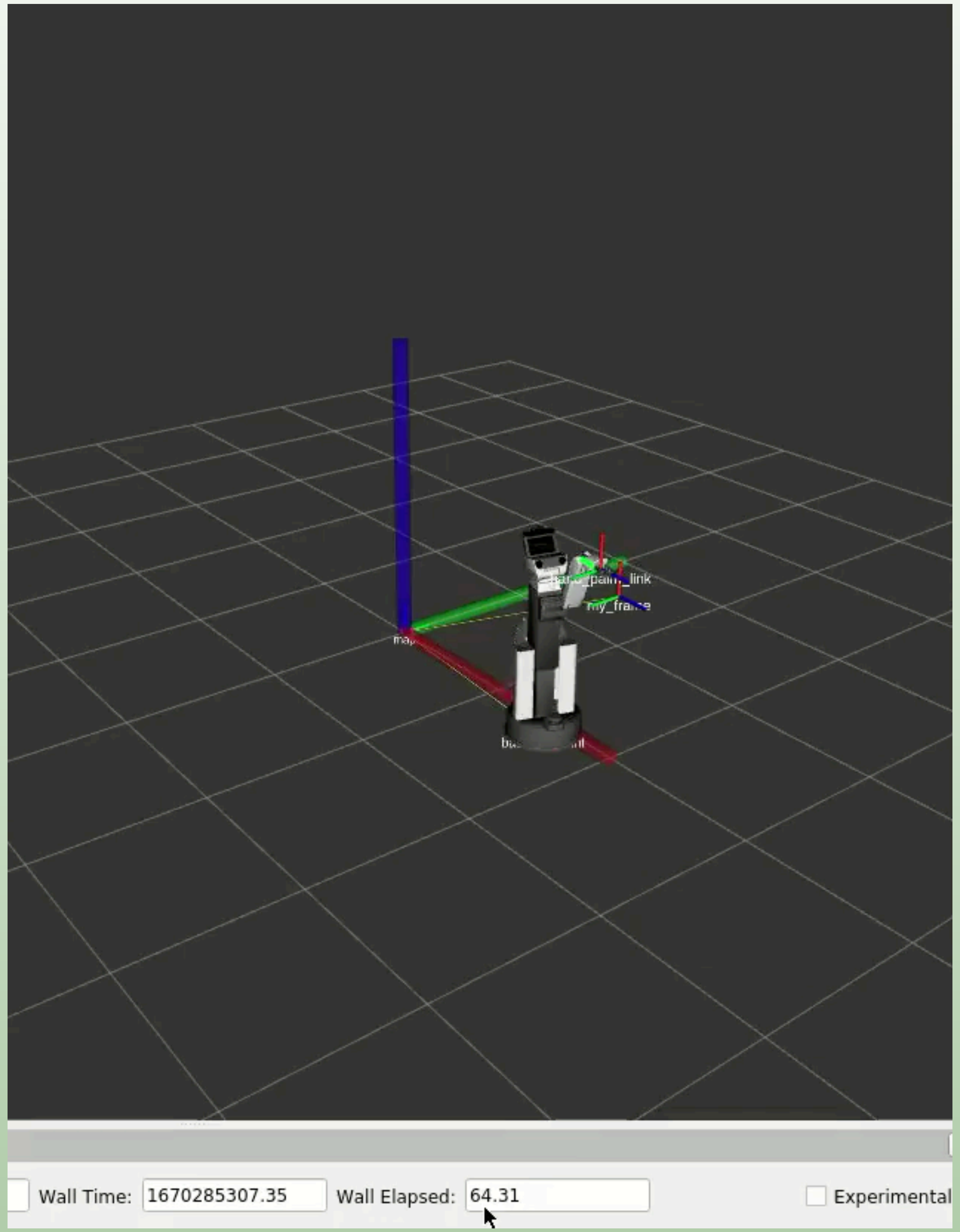
TRAJECTORY GENERATION AND CONTROL

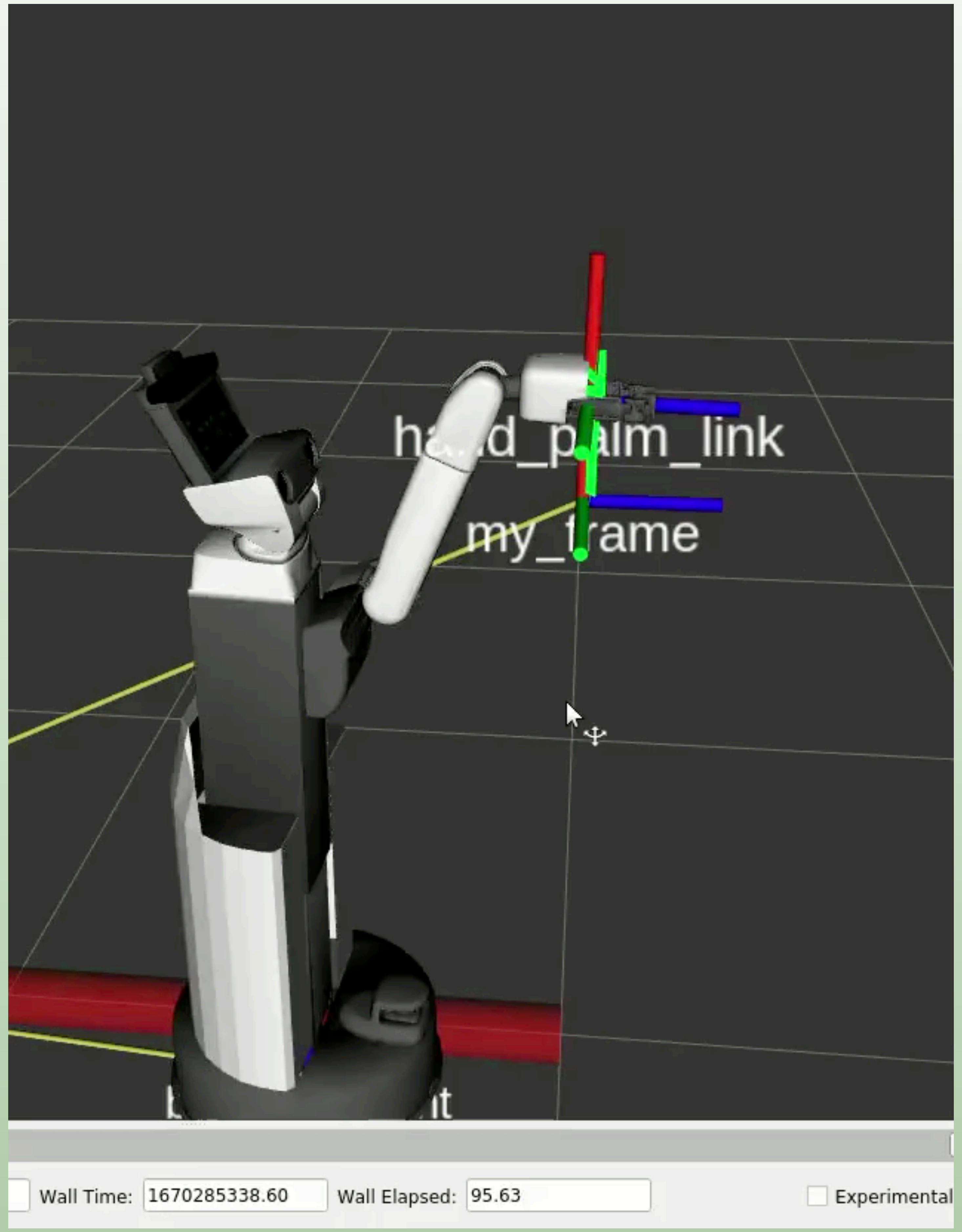




TRAJECTORY GENERATION AND CONTROL







3. HUMAN ROBOT INTERACTION - KEY FEATURES OF DIFFERENT ROBOCUP LEAGUES

- Abbreviations:
  - MSL, Middle Size League
  - RCLL, RoboCup Logistics League
  - SPL, Standard Platform League
  - SSL, Small Size League

League	Number of Robots	External Sensors	Objects to be Recognized	Research Topics	Motion Mechanism
RoboCupSoccer SSL	6–11	Global and fixed cameras shared by both teams	Ball, goals, field lines, teammates, and opponents	Real-time detection and tracking of objects, motion control, and teamwork	Wheel-based, omnidirectional locomotion with kicking device
RoboCupSoccer MSL	5	Onboard individual cameras (mainly horizontal motions) and omnidirectional vision systems	Ball, goals, field lines, teammates, and opponents	Real-time detection and tracking of objects, integration of individual robots' visual information into a global map, motion control, and teamwork	Wheel-based, omnidirectional locomotion with kicking device
RoboCupSoccer SPL and HL	7 SPL, 5 HL KidSize, 2 HL TeenSize and AdultSize	Onboard individual cameras (active vision through 3-D motion) with field of view limited to 180°	Ball, goals, field lines, teammates, and opponents	Real-time detection and tracking of objects, integration of visual information into a global map, localization with a limited field of view, motion control, and team coordination	Posturally stable human-like bipedal locomotion with various types of whole-body motion
RoboCup@Home	1	Onboard multiple cameras, common 2-D and 3-D sensors (e.g., lidar and RGB-D), and microphones	Humans, including their intentions and status; many daily-life objects; and the environment and its current status	Recognition of many objects and subjects, localization, and manipulation, communication, understanding and fulfilling tasks specified by humans on a human-like level	Wheel-based locomotion with arm and hand
RoboCupRescue	1–2	Onboard multiple color, 3-D, and thermal cameras; lidar; and microphones	Status of environment, including victims and damage found autonomously or with support from human operators	Moving from teleoperation to more autonomy in tasks such as mobility, navigation, and manipulation	Mainly tracked wheels for locomotion with arm and hand
RoboCupIndustrial	1 (RoboCup@Work), 3 (RCLL)	Onboard color and 3-D cameras and lidar (RoboCup@Work); camera, lidar, and multiple infrared distance sensors (RCLL)	Industrial parts (RoboCup@Work); status of the environment in exploration and production phases (to discover and utilize production machines) and status of other robots in the same environment (RCLL)	Recognition and autonomous manipulation of industrial parts, transportation, and navigation (RoboCup@Work); production planning and execution with a mobile multiagent team in a smart-factory environment (RCLL)	Wheel-based locomotion with arm and hand

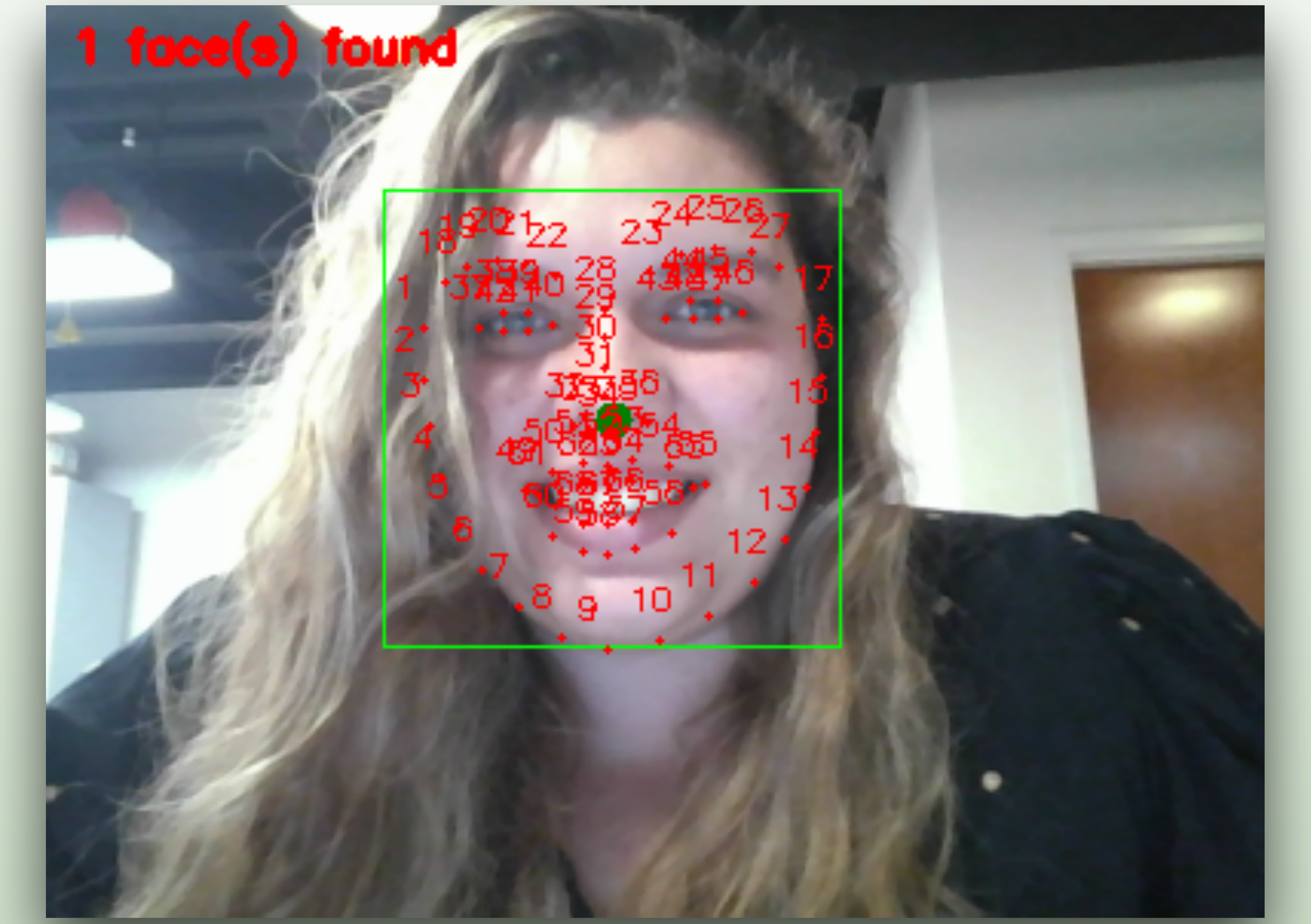
3. HUMAN ROBOT INTERACTION

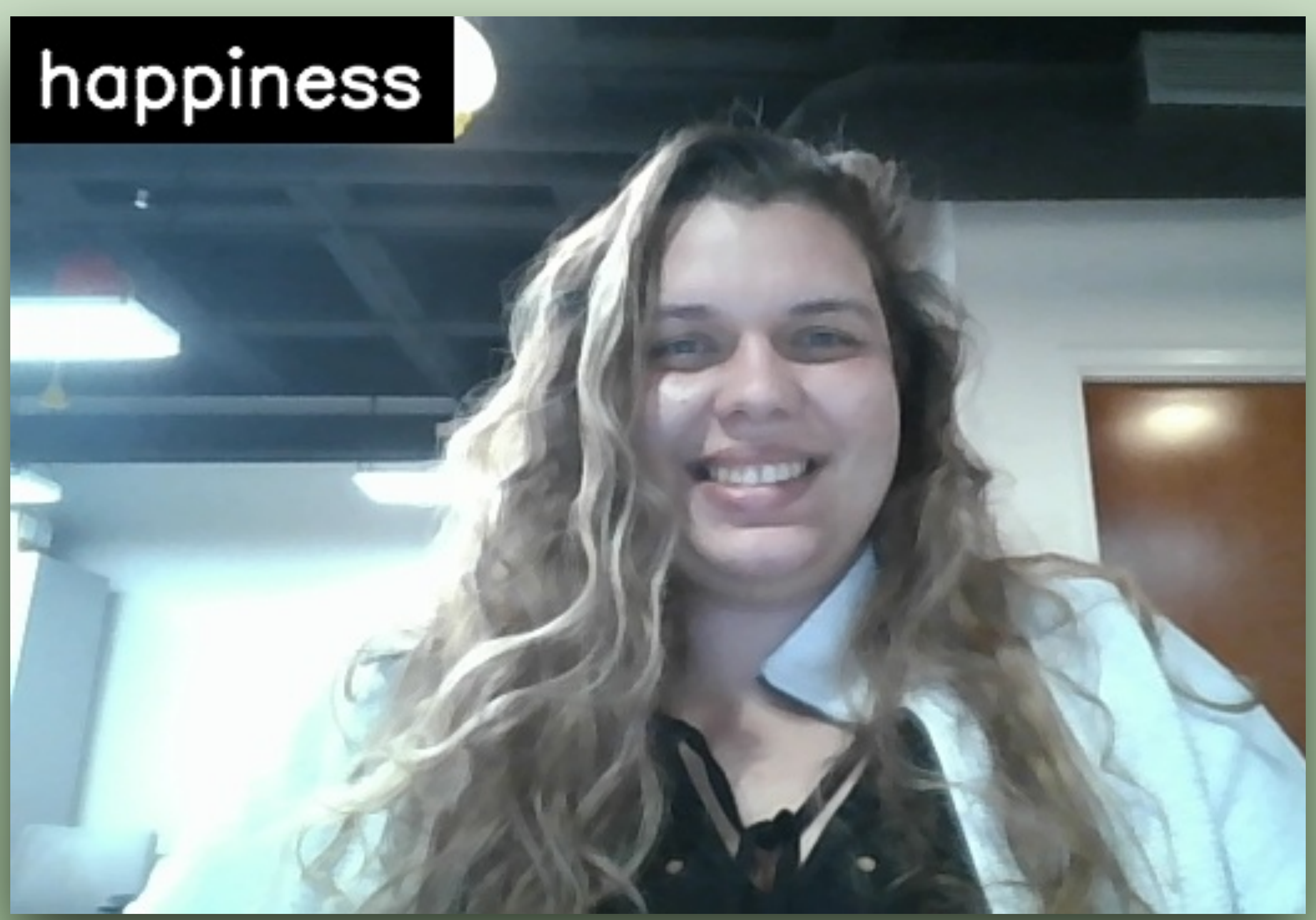
3. HUMAN ROBOT INTERACTION

- Task and motion planning
  - Creation of high-level commands and collision-free trajectories to achieve goal
- State estimation and perception
  - Infer relevant quantities from sensor data (human faces, human poses, emotions, NLP, race, ethnicity, …)
- Communication
  - Building rapport, interact socially, show affect, infer cultural differences, speech recognition, text2speech, people recognition
- Trajectory generation and control
  - Real-time, reactive generation of control commands to move robot (or parts) safely toward goal (people tracking, cleanup, go-get)
- Application in healthcare
  - Pediatric asthma education, SHARE
  - 2nd pilot, caregivers, healthcare professionals

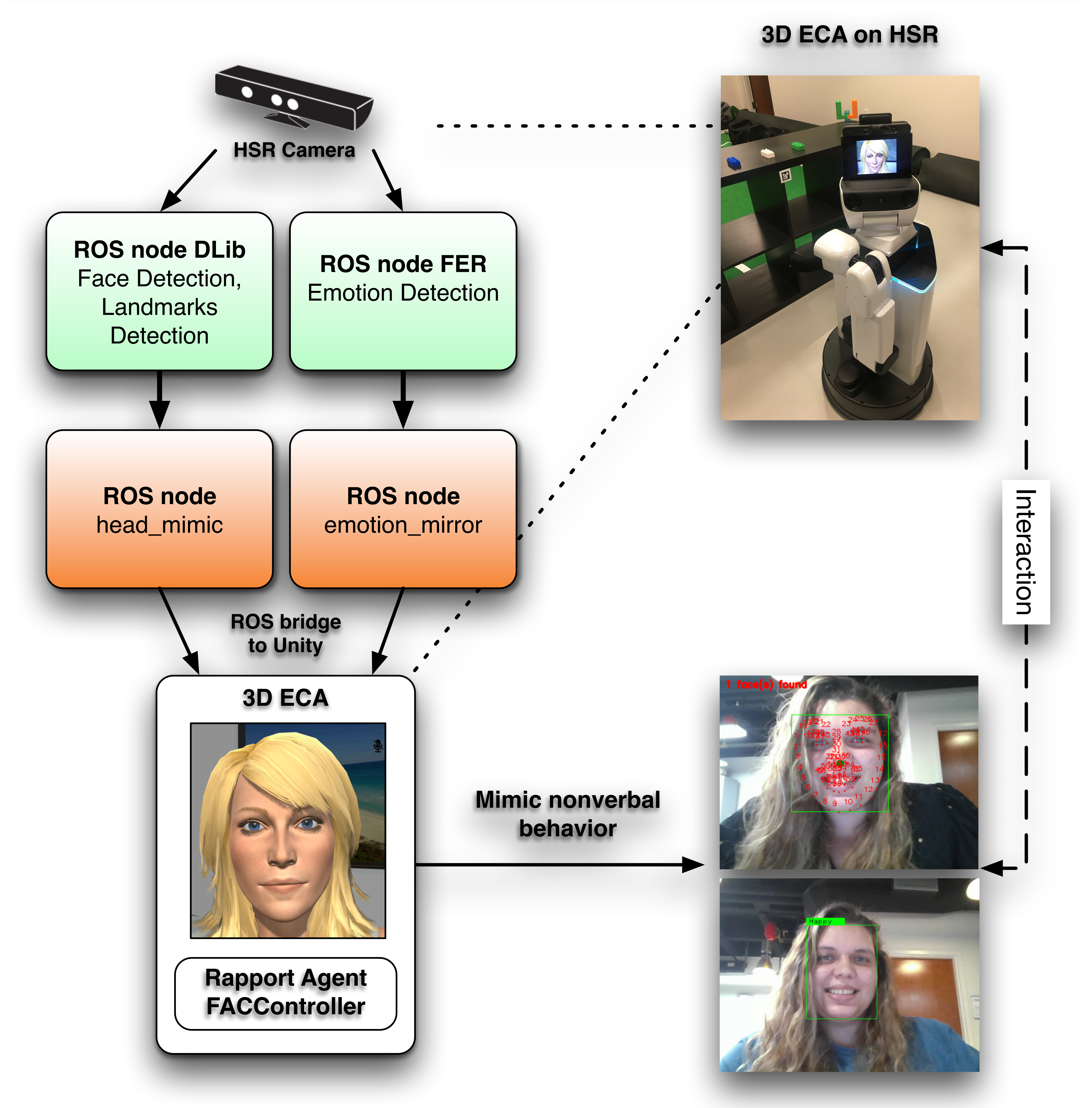
COMMUNICATION: BUILDING RAPPORT

- Rapport as result of a combination of socio-cultural-emotional complex processes, e.g. unconscious:
  - mutual attentiveness (mutual gaze, mutual interest, focus during interaction)
  - positivity (e.g., head nods, smiles, friendliness, and warmth)
  - unconscious coordination (e.g., postural mirroring, synchronized movements, balance, and harmony)
- Focus here on coordination/mirroring of
  - head movements and
  - facial emotions





COMMUNICATION: BUILDING RAPPORT



COMMUNICATION: BUILDING RAPPORT

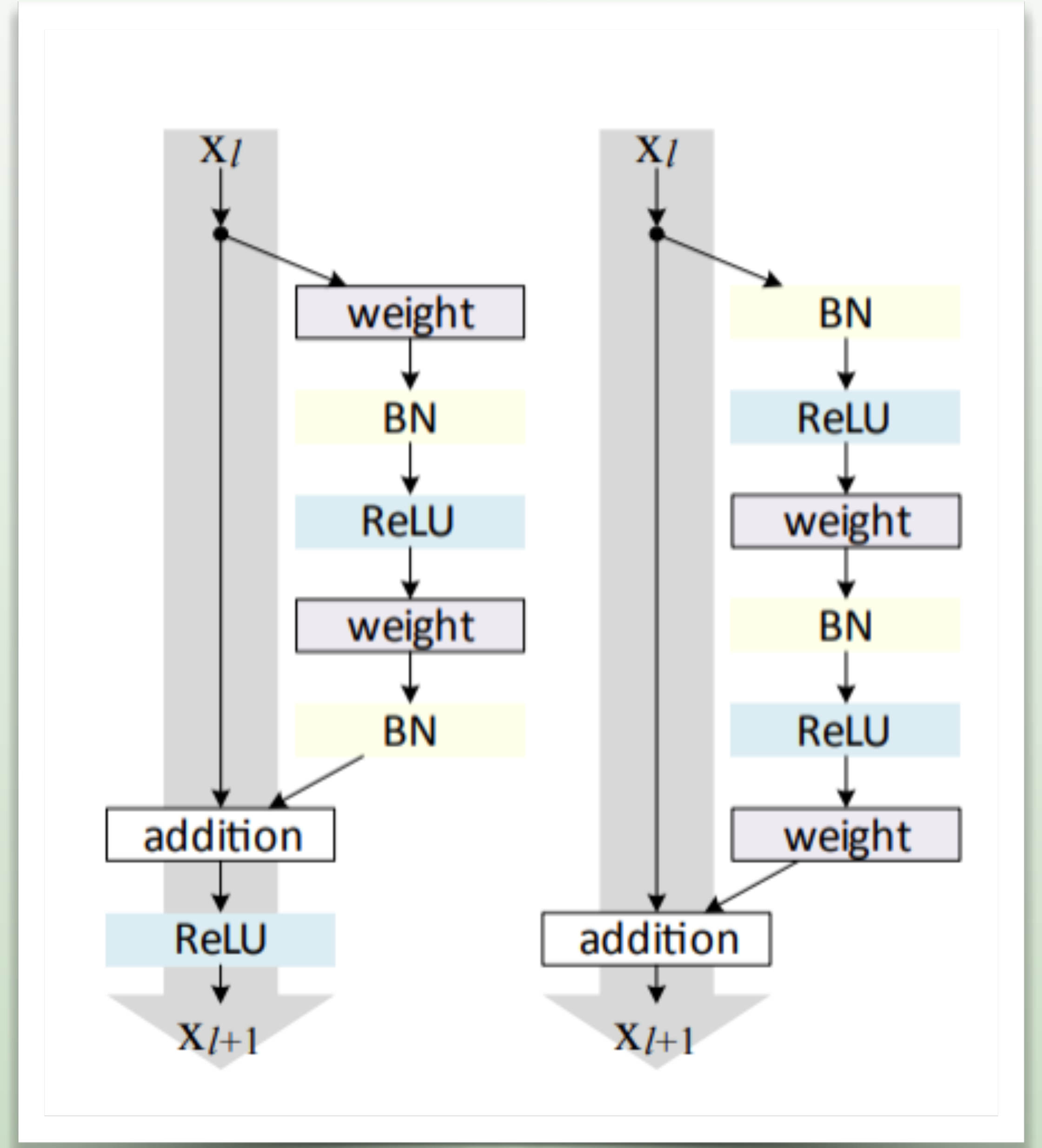
COMMUNICATION: BUILDING RAPPORT (2)

- Rapport as a result of a combination of socio-cultural-emotional complex processes
- Leveraging Deep Generative Modeling
- The metric model is a simple classifier predicting emotion categories given facial landmark inputs
- Classifier trained on AffectNet dataset
- Focus here on motion synthesis
  - head movements and
  - facial expressions

Mollahosseini et al., 2017

COMMUNICATION: BUILDING RAPPORT (2)

- Pre-Activation residual blocks
- Layers are resequenced to improve information flow



He et al., 2016

- Instance normalization to strip intro, sample, contrastive, information, making the generation process, simplified, and controllable.
- Swish activation, replacing Relu for better representation estimation of non-linearity

$$BN(x) = \gamma \frac{x - \mu_B(x)}{\sigma_B(x)} + \beta$$

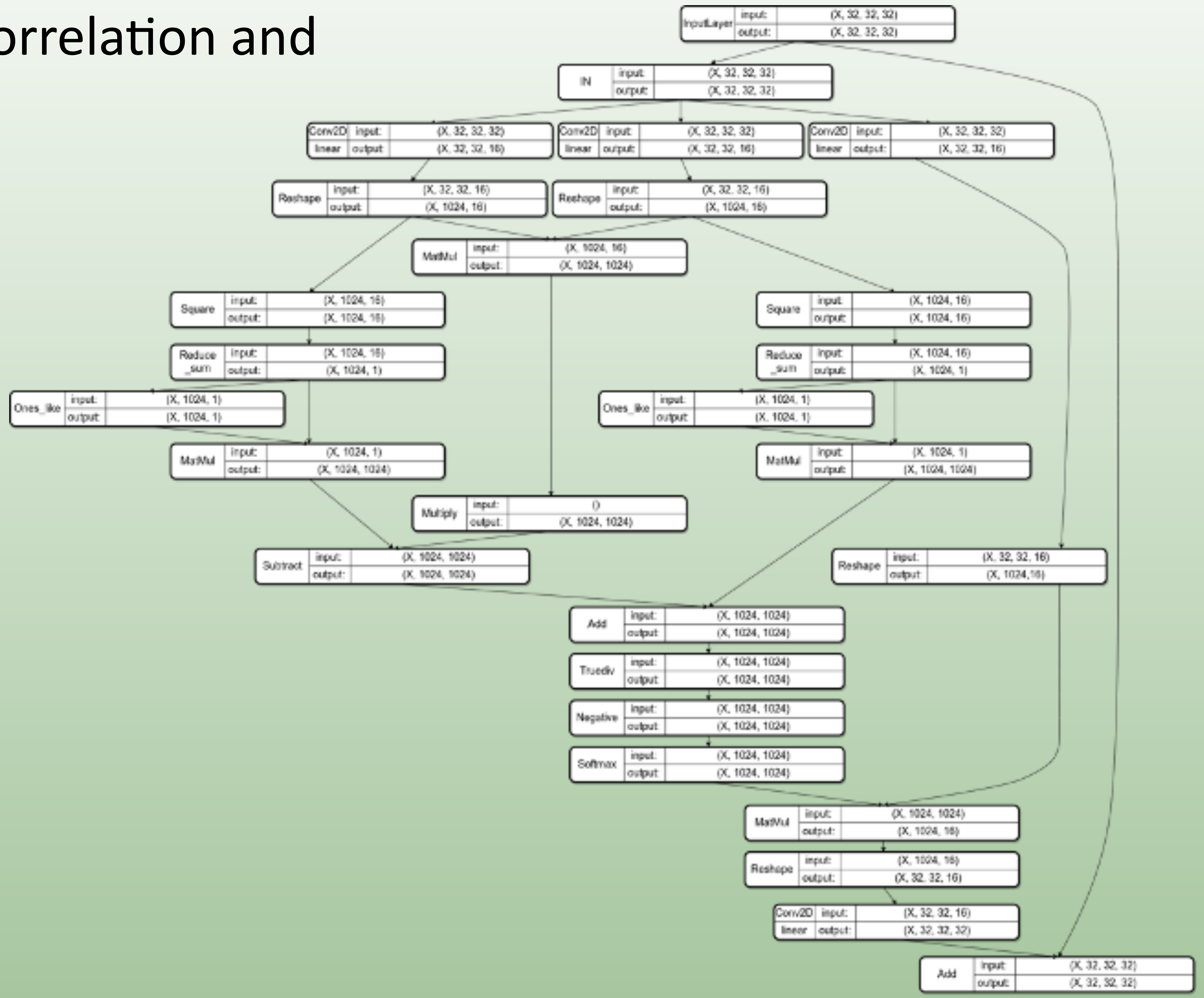
vs

$$IN(x) = \gamma \frac{x - \mu_{NC}(x)}{\sigma_{NC}(x)} + \beta$$

COMMUNICATION: BUILDING RAPPORT (2)

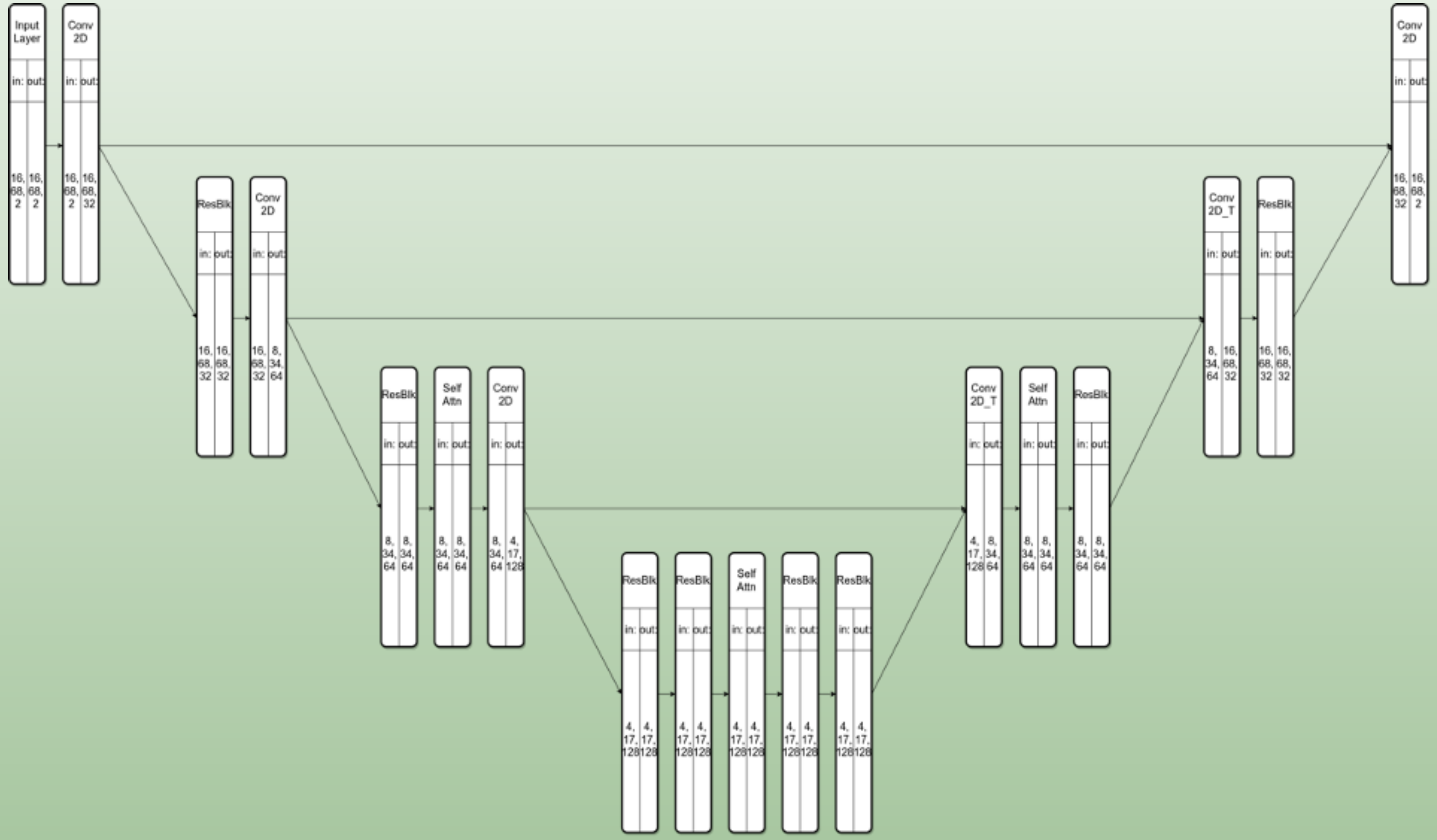
- L2 self-attention to capture global correlation and avoid Lipschitz unboundedness

$$L2Attention(Q, K, V) := exp\left(-\frac{\|x_i^T W^Q - x_j^T W^K\|_2^2}{\sqrt{d_k}}\right) V$$



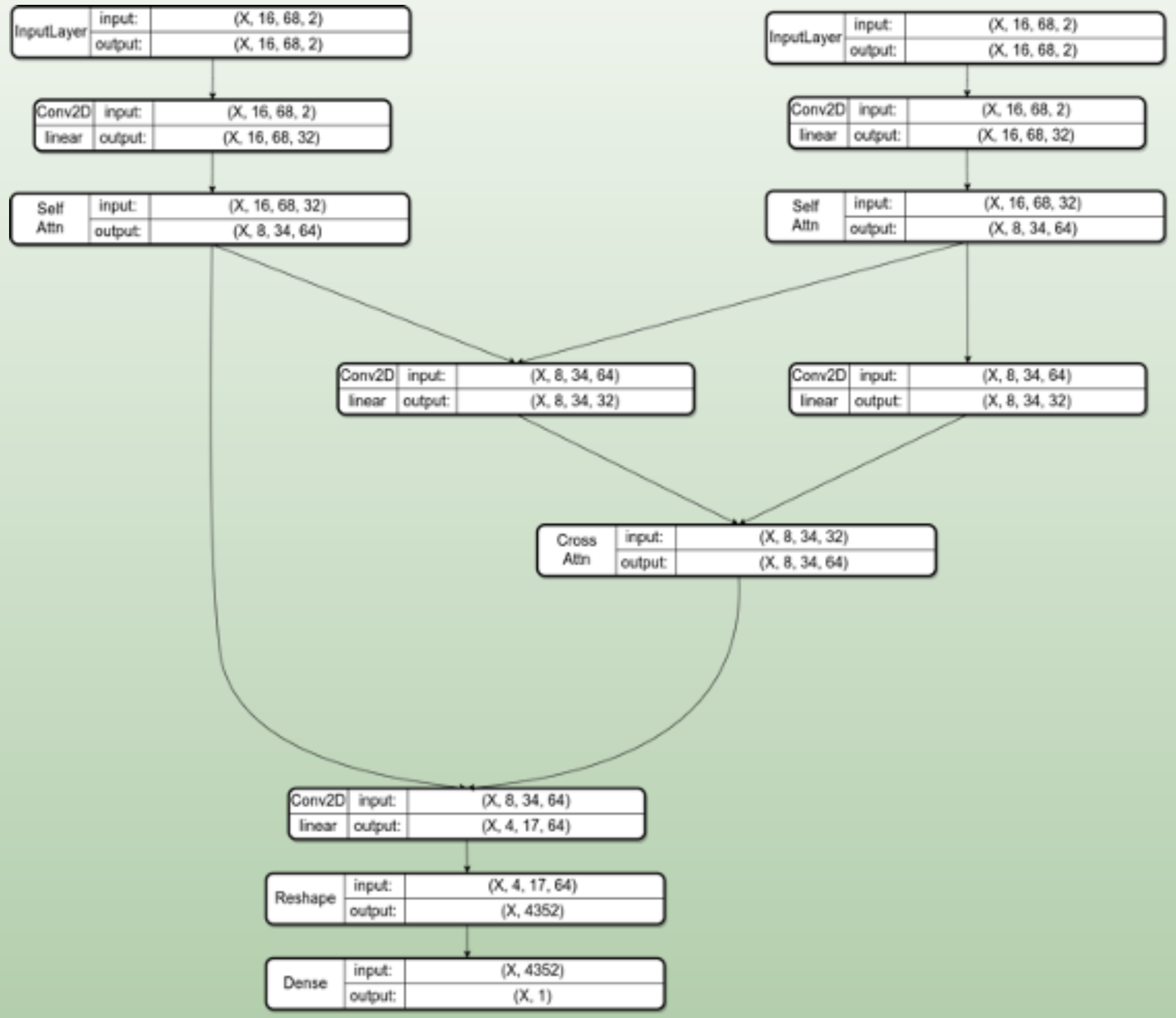
COMMUNICATION: BUILDING RAPPORT (2)

- U-Net architecture for symmetric skip connection



COMMUNICATION: BUILDING RAPPORT (2)

- The discriminator design follows mostly the generator. It additionally uses a cross-attention module to consider the corresponding client landmarks for accurate rapport modeling.



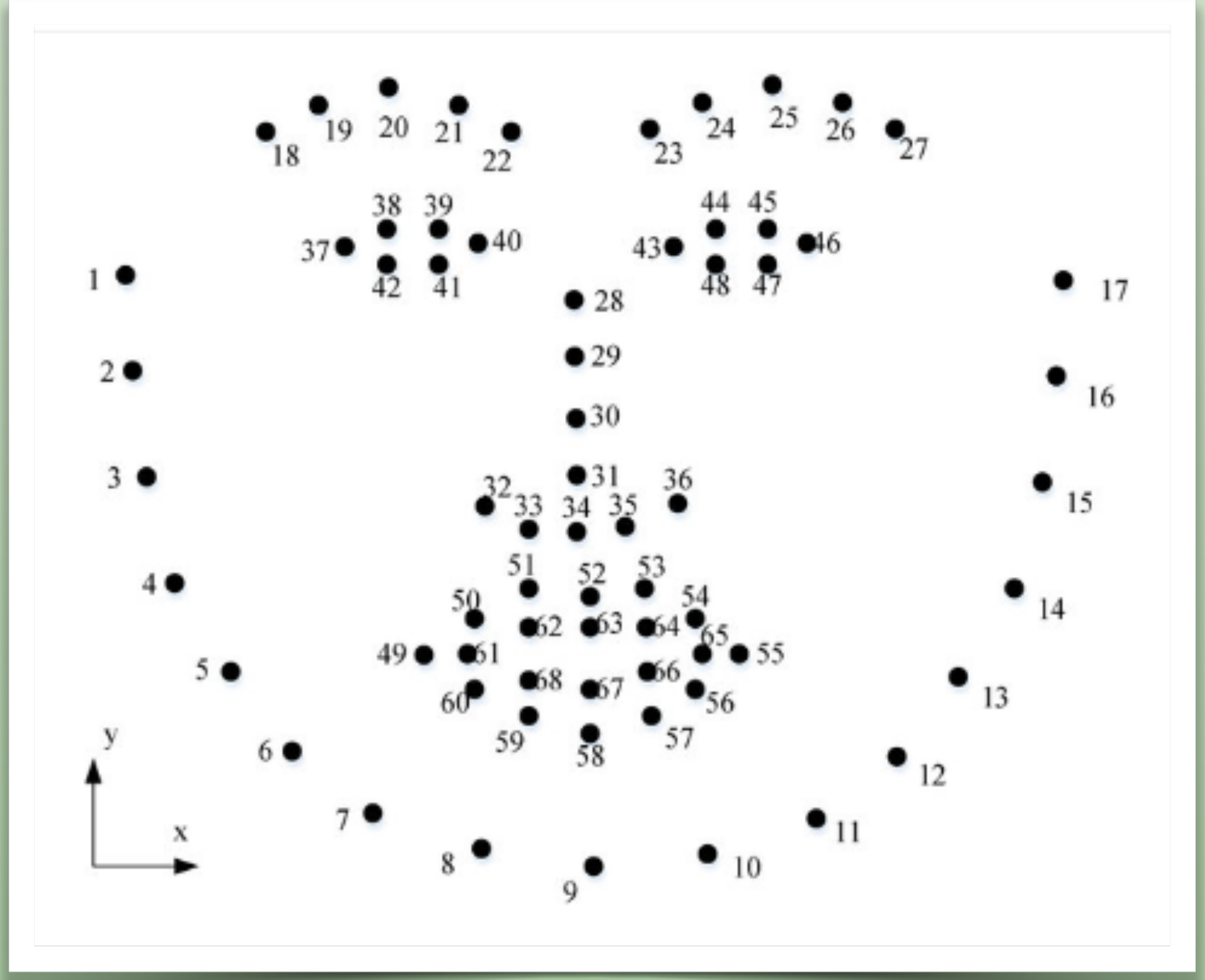
COMMUNICATION: BUILDING RAPPORT (2)

- GAN model, adversarial training:
  - Adversarial loss: $$\min_G \max_D \mathbf{E}_{x\sim p_{data}(x)}[log(D(x))] + \mathbf{E}_{z\sim p_z(z)}[log(1 - D(G(z)))]$$
  - Gradient penalty loss: $$L_{gp}(x') = \lambda_{gp}(\|\nabla_{x'} D_x(x')\|_2 - 1)^2$$ $$x' = \epsilon x + (1-\epsilon)F(y), \epsilon \in Uniform(0,1)$$
  - Signal magnitude regularization: $$L_{reg}(D) = \lambda_{reg} D^2$$
  - Additional supervised model: $$\mathbf{E}_{x\sim p_X(x), y\sim p_Y(y)}[\|G(x) - y\|_2]$$

COMMUNICATION: BUILDING RAPPORT (2)

- Rendering
  - Landmark translation to AU are not well defined.
  - A preliminary mapping is proposed here as a foundation

AU index	AU name	Landmark indices pair
1	Inner brow raiser	(21,22),(23,24)
2	Outer brow raiser	(18,19),(26,27)
5	Upper lid raiser	(34,52)
12	Lip corner puller	(32,49),(36,55)
13	Sharp lip puller	(32,49),(36,55)
15	Lip corner depressor	(58,63),(63,34)
20,23	Lip stretcher,Lip tightener	(49,55)
28	Lip suck	(52,58)



COMMUNICATION: BUILDING RAPPORT (2)

- Results
  - Metric model performance:

Variants	Accuracy	F1	R2
Linear	50.33	32.28	-4.83
1-D Self-Attention	51.01	33.32	-4.55
SVM	50.45	33.01	-3.96

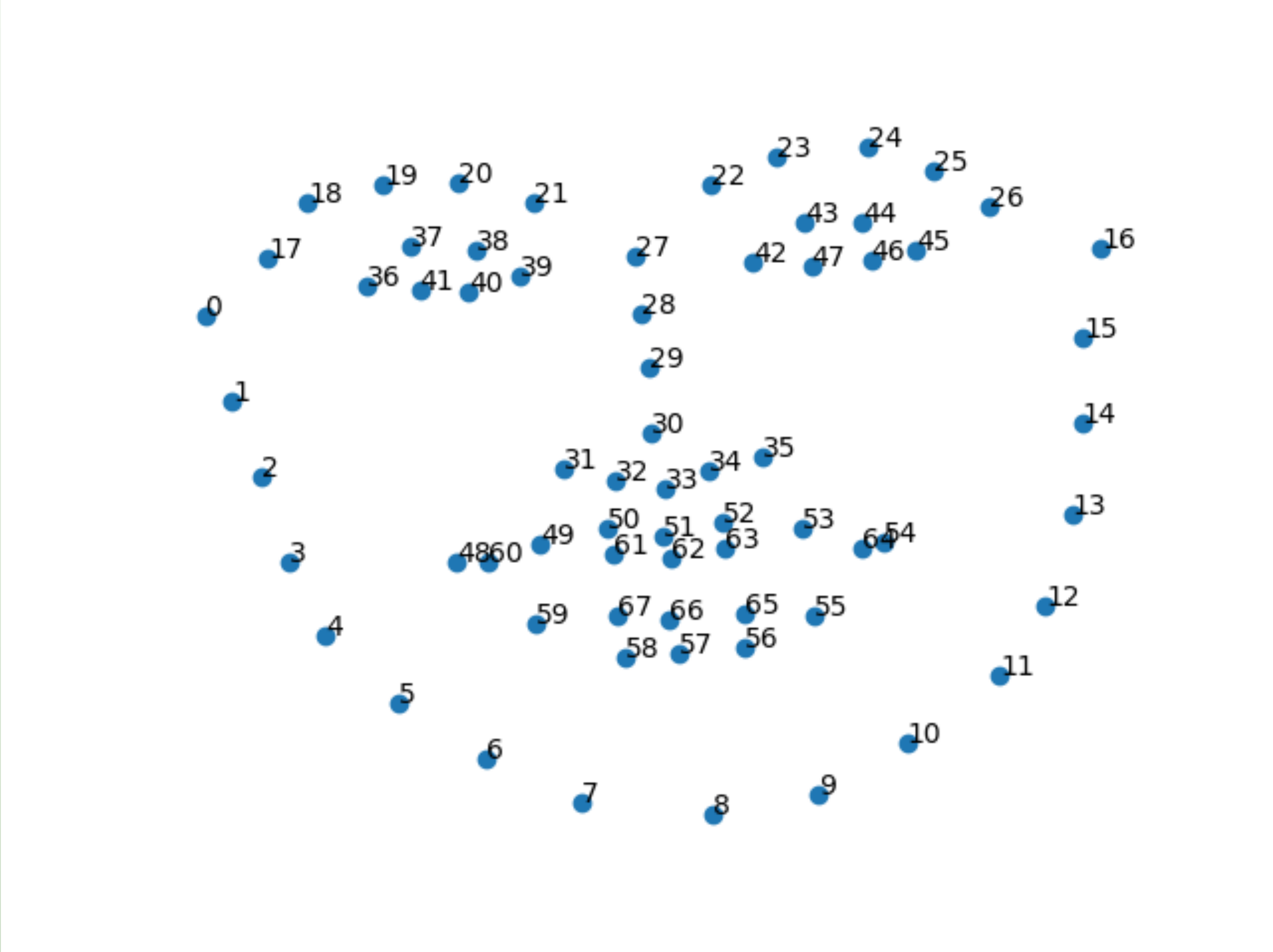
  - Generative model performance:

Variants	Emotion accuracy	Emotion F1	Emotion R2
Supervised	35.21/32.34/29.04	23.78/21.76/20.14	-10.45/-10.64/-10.22
GAN variant 1	35.98/34.71/32.50	24.03/24.15/23.29	-10.69/-10.40/-10.71
GAN variant 2	39.12/39.32/38.58	**28.56**/**28.31**/**27.99**	**-8.25**/**-9.16**/**-9.02**
GAN variant 3	36.35/33.41/25.12	23.97/20.13/13.33	-12.43/-15.61/-50.81
GAN variant 4	41.01/**43.65**/36.32	27.17/25.61/25.08	-10.08/-11.13/-11.02
GAN variant 5	36.22/34.40/29.12	24.56/24.72/22.01	-12.17/-12.33/-12.19
GAN variant 6	**42.76**/41.30/**39.07**	28.02/27.67/22.09	-10.83/-10.79/-10.91

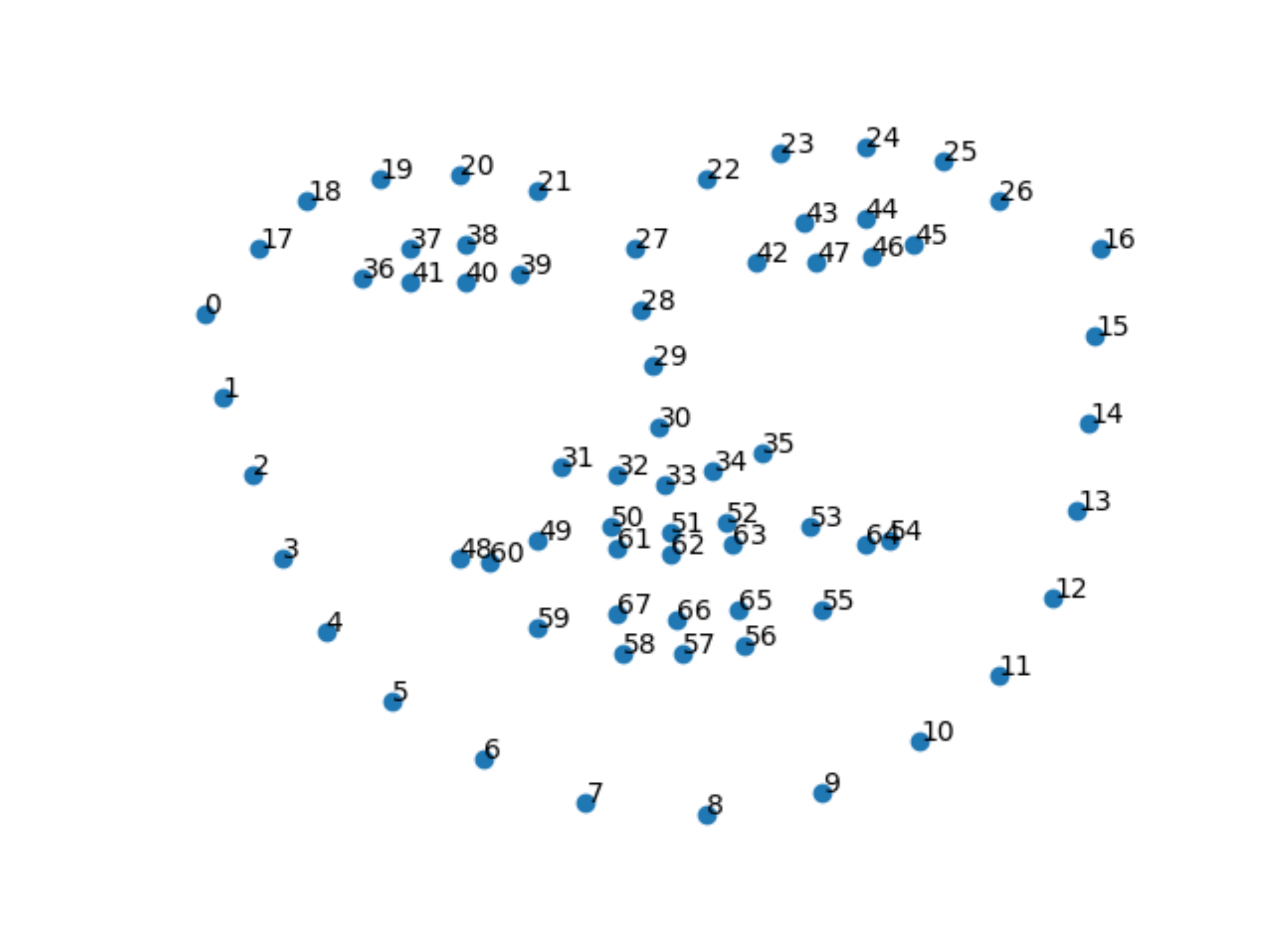
Variants	Landmark L2	Convergence rate
Supervised	**12.38**	100%
GAN variant 1	25.11	85%
GAN variant 2	23.09	100%
GAN variant 3	32.26	45%
GAN variant 4	29.87	90%
GAN variant 5	29.60	100%
GAN variant 6	26.57	100%

- Variant 1 - non saturating loss with dot product attention
- Variant 2 - non saturating loss with L2 norm attention
- Variant 3 - Wasserstein loss without regularization with dot product attention
- Variant 4 - Wasserstein loss without regularization with L2 norm attention
- Variant 5 - Wasserstein loss with regularization with dot product attention
- Variant 6 - Wasserstein loss with regularization with L2 norm attention

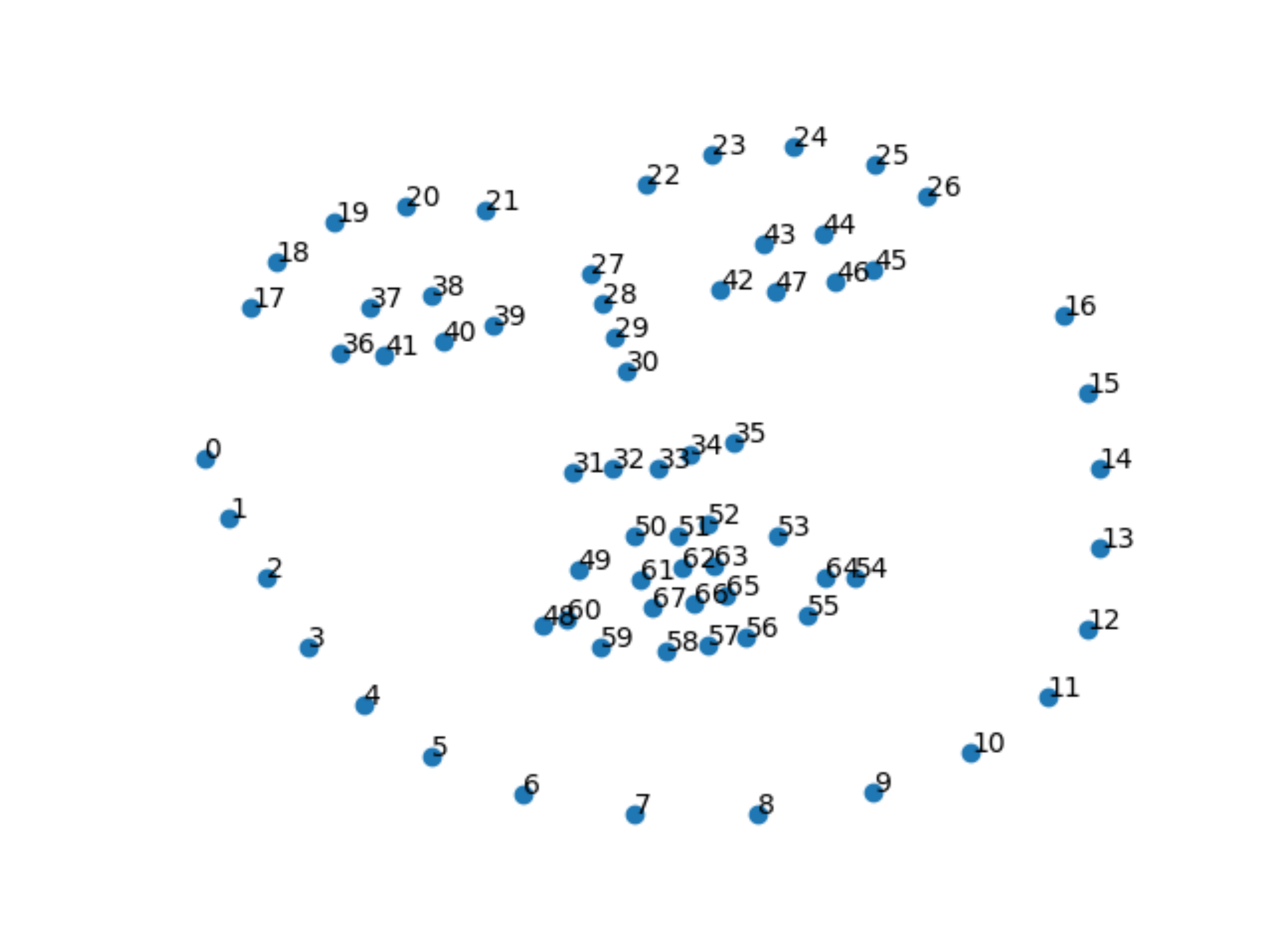
COMMUNICATION: BUILDING RAPPORT (2)



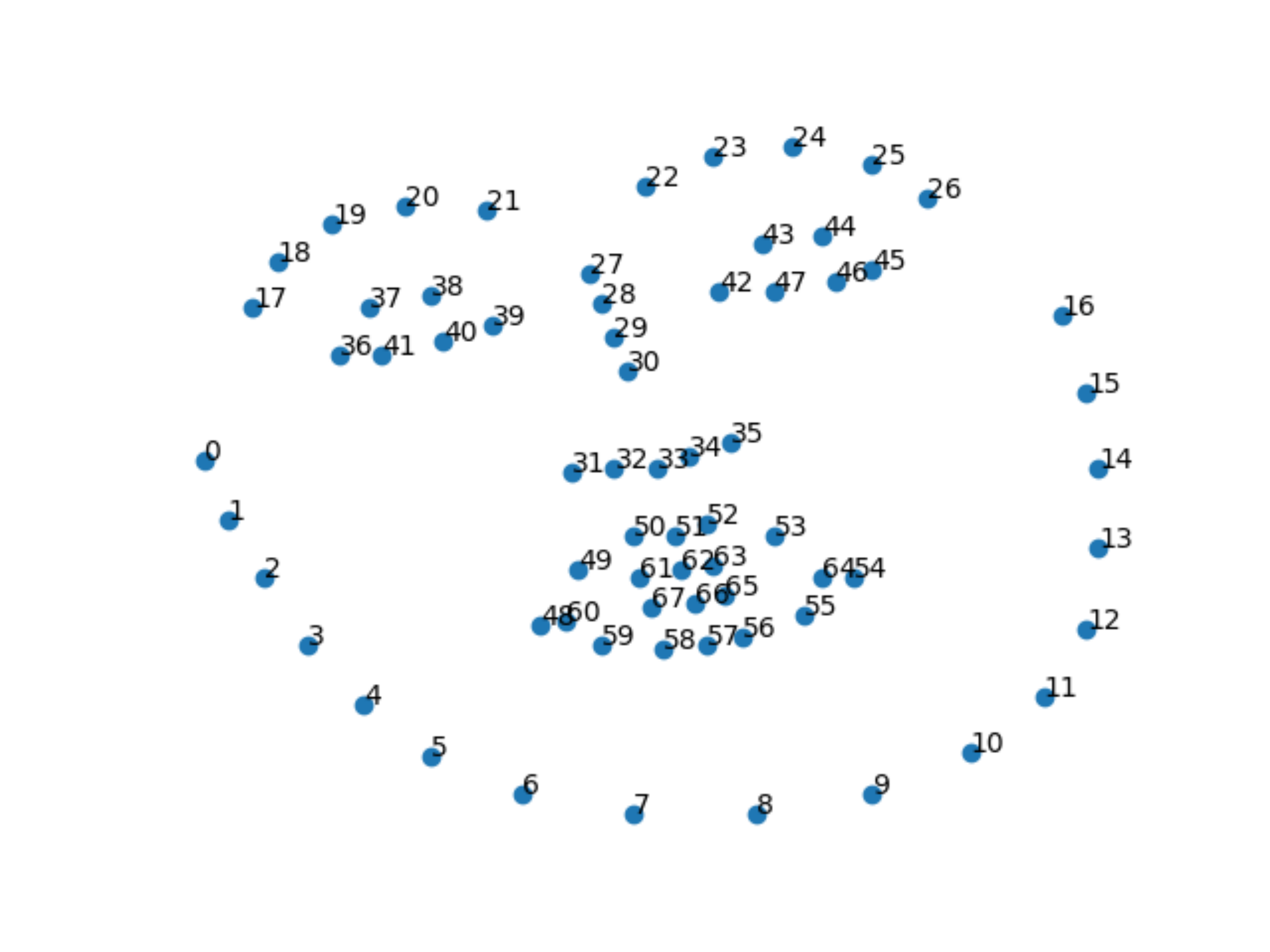
Testing set generated



Testing set ground truth



Training set generated



Training set ground truth

COMMUNICATION: BUILDING RAPPORT (2)




PEDRIATIC ASTHMA EDUCATION - SHARE

- Pediatric asthma affects 6.3 million children in the U.S., making it a leading cause of hospitalizations.
- Discharge education is crucial but often rushed.
- Two completed stages of research:
  - 1.Pilot study – simulated discharge with student participants.
  - Clinical study – real-world outpatient clinic with caregivers and healthcare professionals

Goal: Develop and evaluate a socially assistive robot (SAR) to deliver structured, engaging, and repeatable asthma education.

PEDRIATIC ASTHMA EDUCATION - SHARE

- Within-subjects pilot study: 11 students simulating discharge interactions.
- Content developed with input from:
  - National Heart, Lung, and Blood Institute (NHLBI) guidelines
  - Asthma Academy resources
  - Consultations with a pediatric pulmonologists and clinical visits.
- Sessions held at UM's Simulation Hospital (SHARE).
- Session has five components: asthma pathophysiology, medication overview, inhaler technique, asthma action plan, and discharge reminders.
- Delivered through HSR using a multimodal approach of speech and gestures, expert videos, and printed hand out.

SHARE Simulation Hospital environment

Study setup in the space with caregivers and pediatric patients

PEDRIATIC ASTHMA EDUCATION - SHARE

CERA

Family Caregiver Education with Robotic Assistance

UNIVERSITY OF MIAMI
COLLEGE of ARTS & SCIENCES

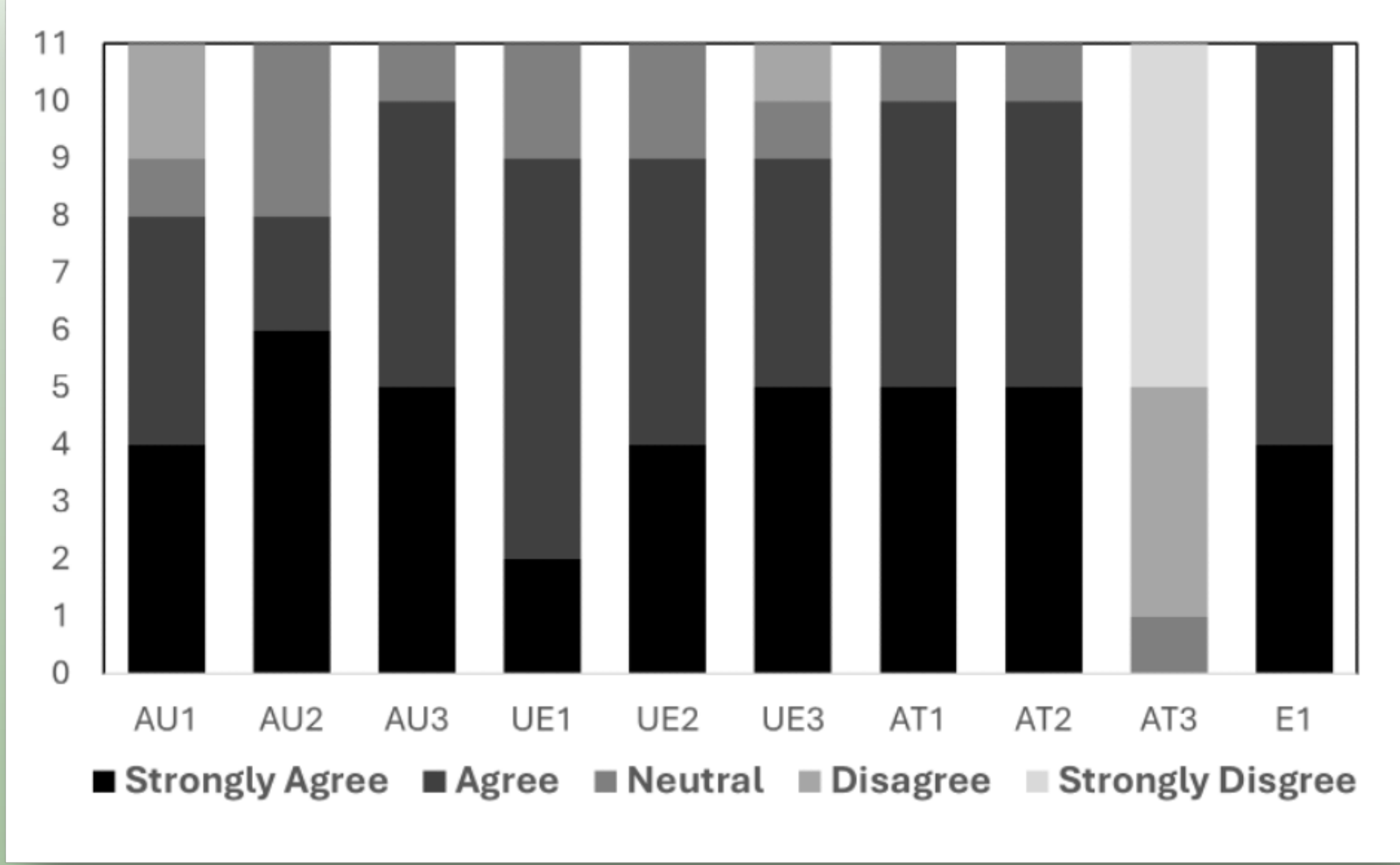
UNIVERSITY OF MIAMI
ROBOCANES

SHARE.

UNIVERSITY OF MIAMI
SCHOOL of NURSING & HEALTH STUDIES

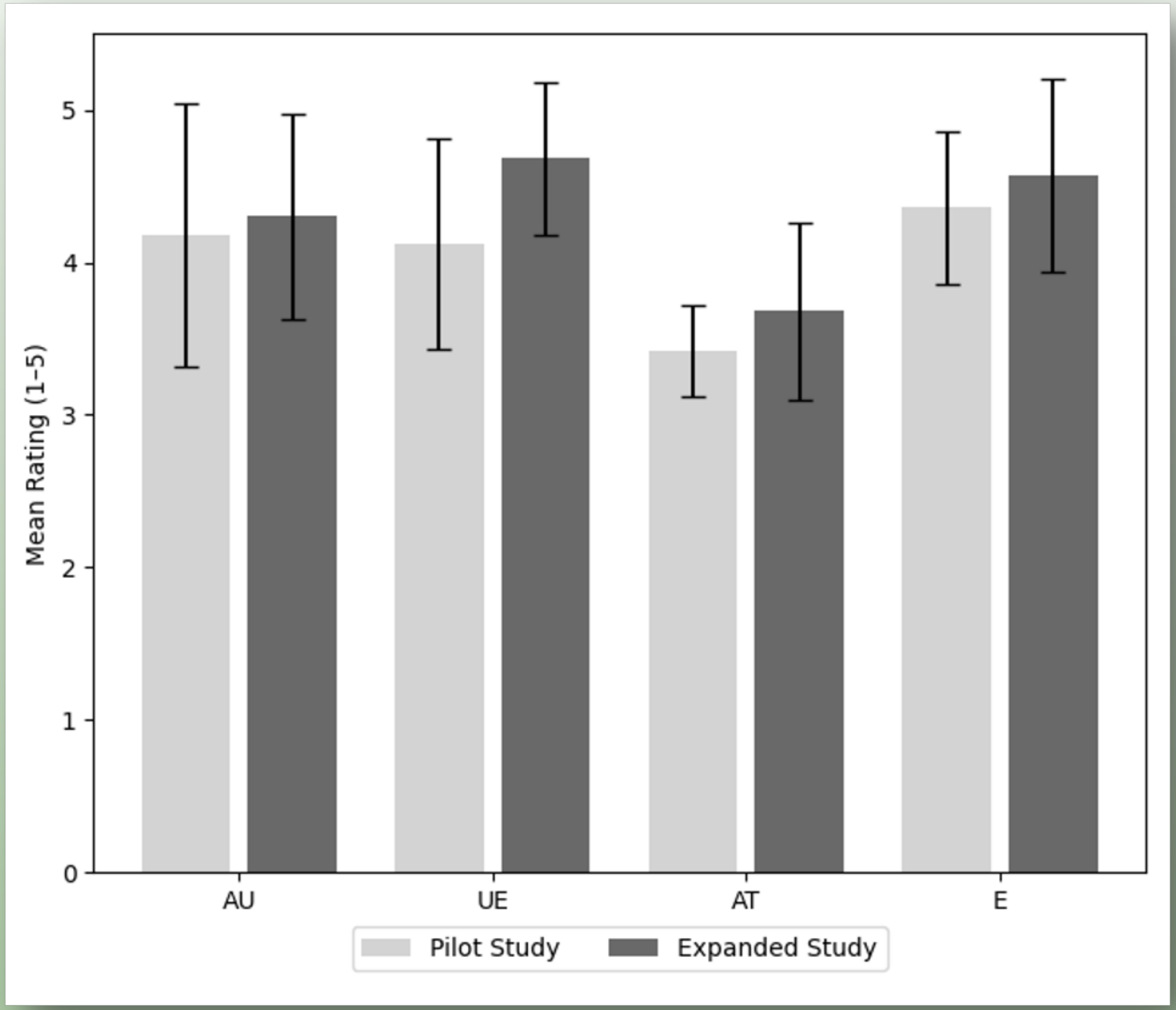
PEDRIATIC ASTHMA EDUCATION - SHARE - PILOT

- Results:
- 78% rated usability positively.
- 82% found interaction engaging.
- >90% reported positive attitude toward the robot.
- 100% satisfied with educational experience.
- Key Insight: Robot-led asthma education is feasible, easy to use, and engaging for first-time users.



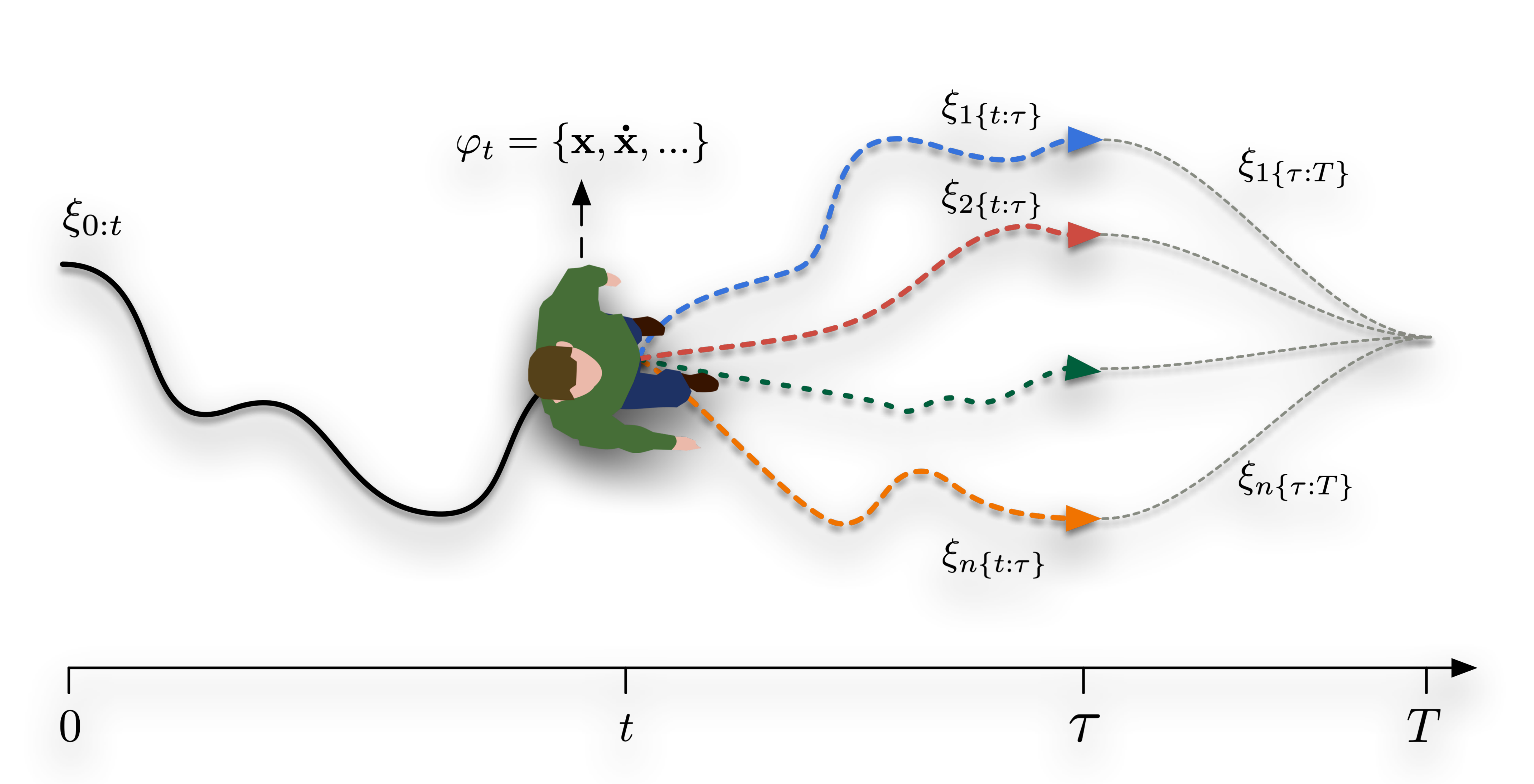
PEDRIATIC ASTHMA EDUCATION - CLINICAL STUDY WITH CAREGIVERS

- Conducted at UHealth Batchelor Children's Research Institute (BCRI) outpatient clinic.
- Participants: 30 total – 14 caregivers of children with asthma, 16 healthcare professionals.
- Findings:
  - High ratings across all constructs (means 4.3–4.7/5).
  - Engagement significantly higher than pilot ($p = 0.028$, Cohen's $d = 0.76$).
  - Caregivers and providers rated usability, engagement, and educational value highly.
  - Qualitative feedback: approachable design, clear instruction, reduced anxiety, reinforced educational learning.



STATE ESTIMATION AND PERCEPTION: HUMAN POSES

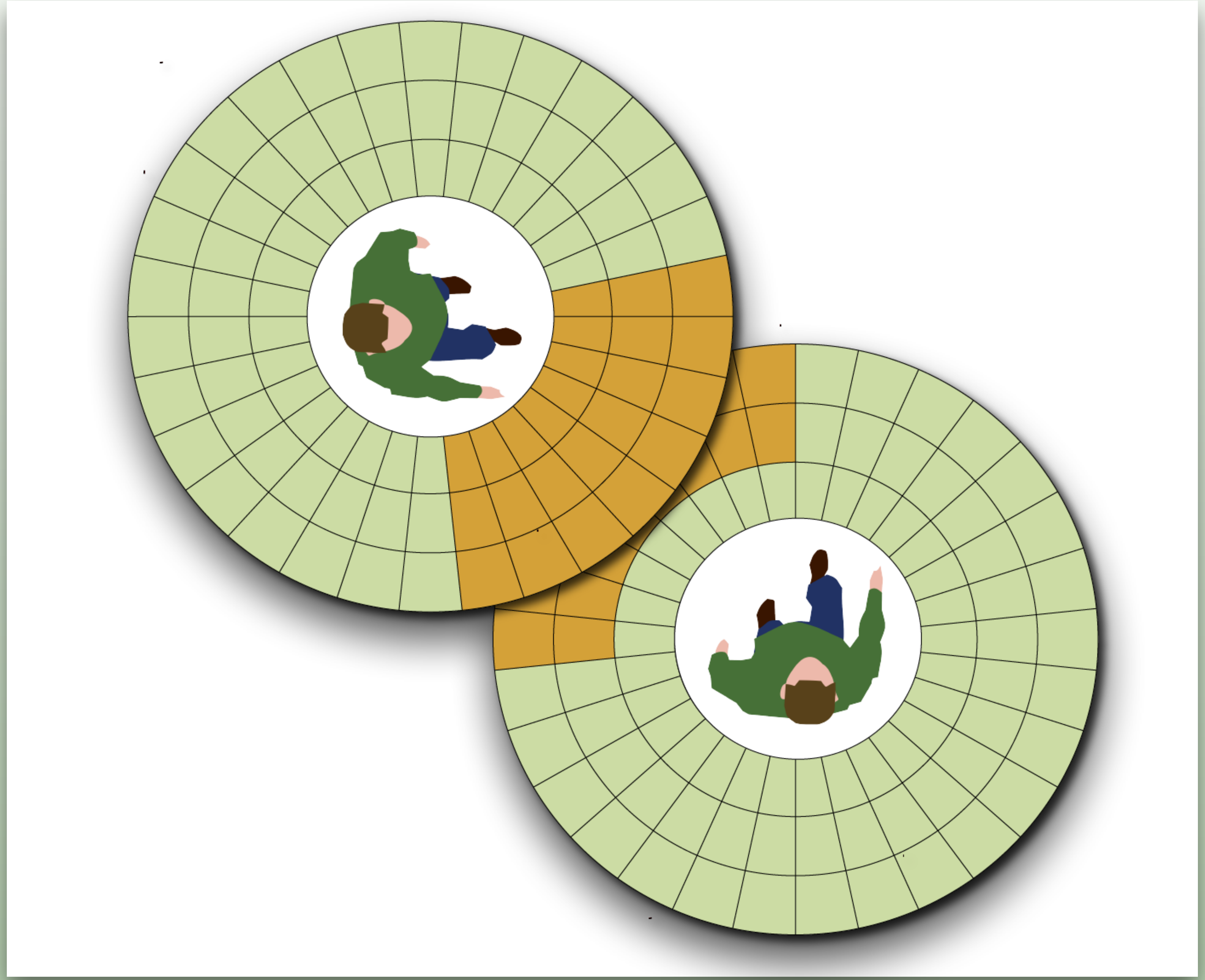
- Tracking and predicting humans in 3D space
- Novel probabilistic framework in which multiple models can be fused into a circular probability-map to forecast human poses
- ITP: Inverse Trajectory Planning



Peña & Visser (*KI: German AI Journal, 2020)*

STATE ESTIMATION AND PERCEPTION: HUMAN POSES

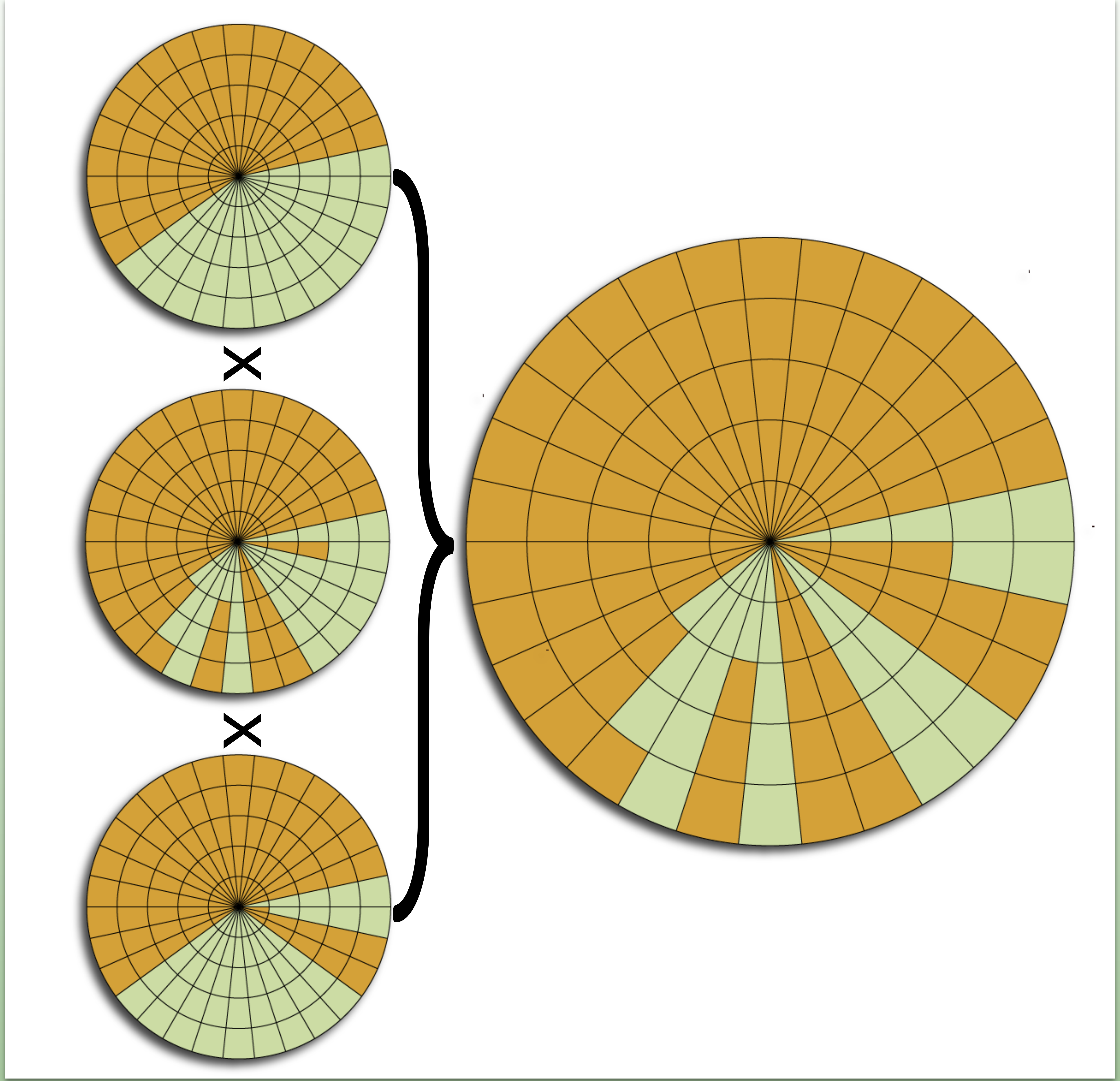
- Circular probability-maps for the social force model
- ITP allows interaction between two people or two probability-maps



Peña & Visser (*KI: German AI Journal, 2020*)

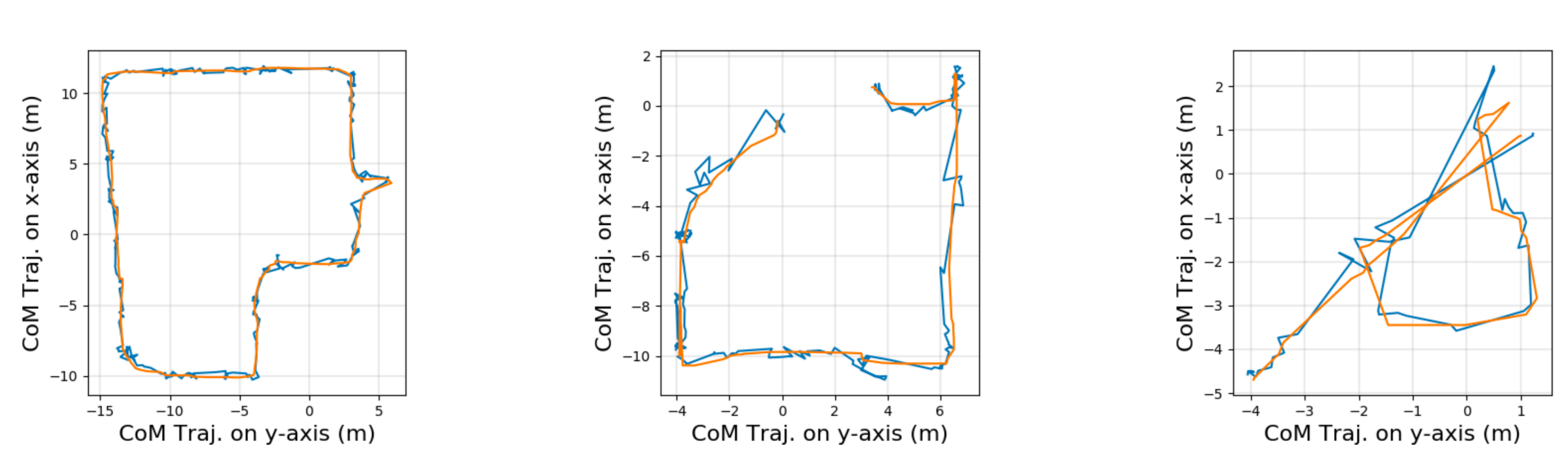
STATE ESTIMATION AND PERCEPTION: HUMAN POSES

- Lidar-based model for 2D obstacles
- OctoMap model for 3D obstacles
- Social force model for other humans
- Result: Heading model as prediction of future states



Peña & Visser (*KI: German AI Journal, 2020*)

STATE ESTIMATION AND PERCEPTION: HUMAN POSES



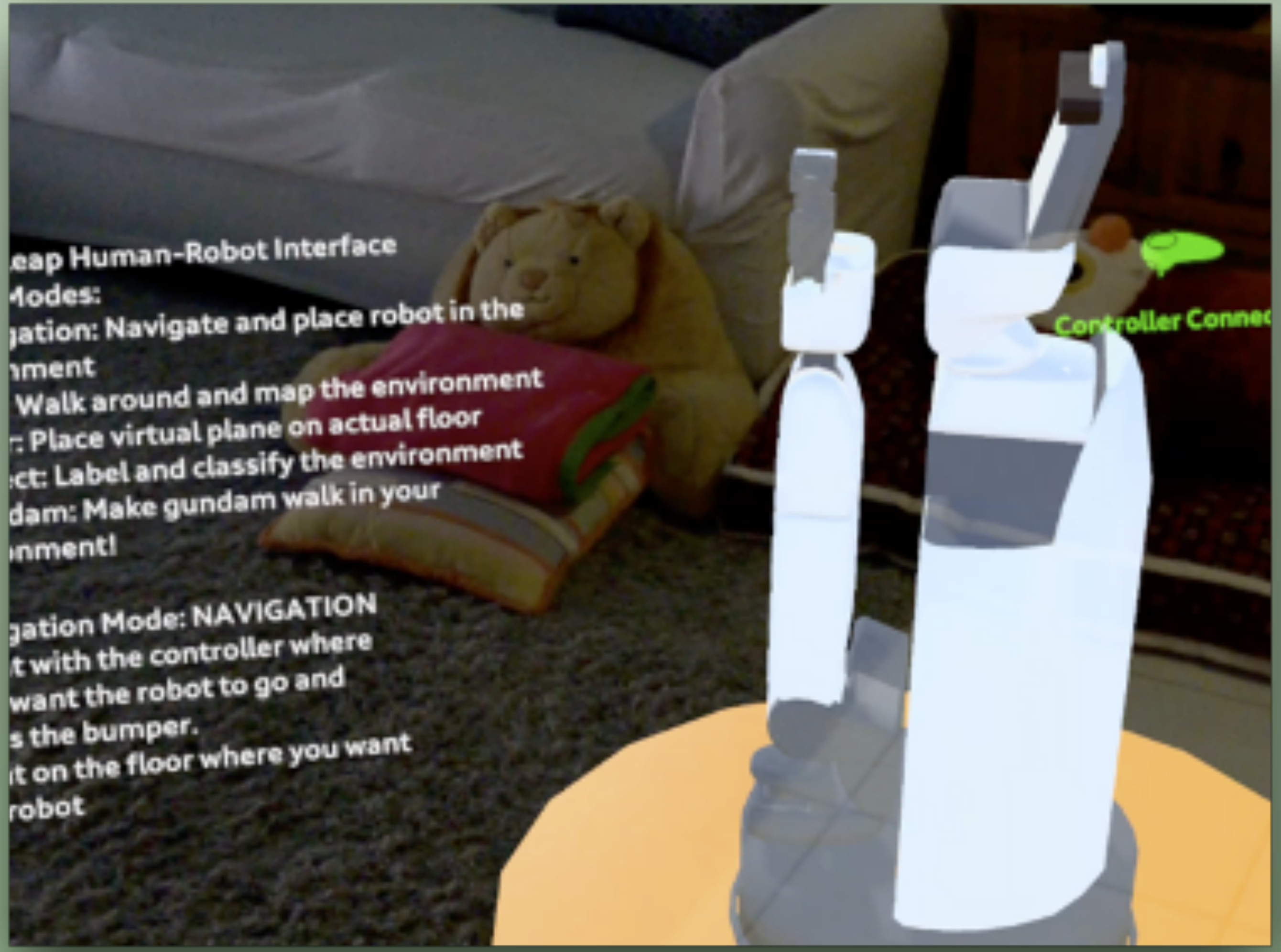
Peña & Visser *(KI: German AI Journal, 2020)*

STATE ESTIMATION AND PERCEPTION: HUMAN POSES

Peña & Visser (*KI: German AI Journal, 2020*)

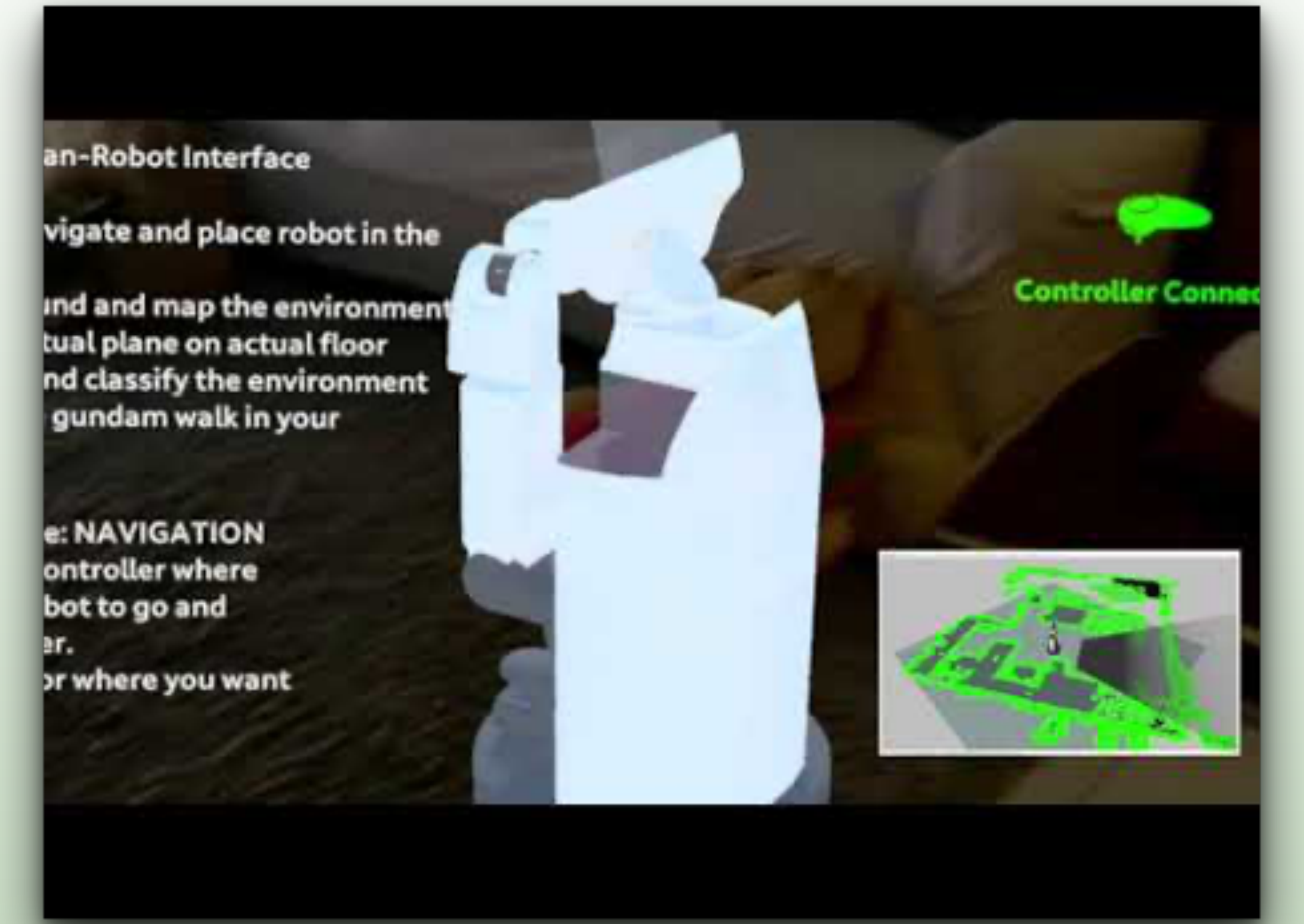
4. INTERFACES: VR MEETS AI AND ROBOTICS

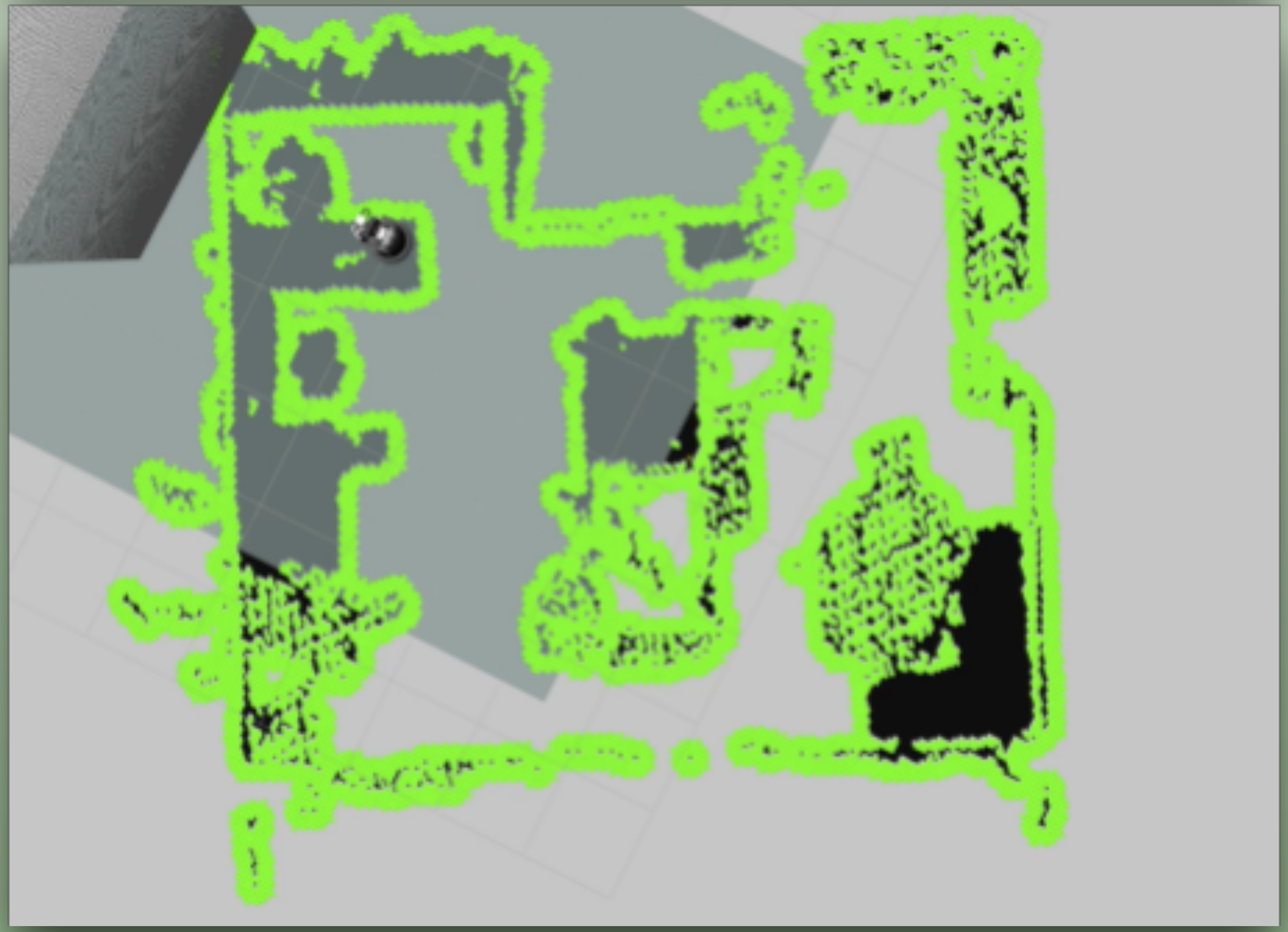
- Implement a human-robot interface that is intuitive and does not require a computer
- Geriatric (not robotics) expert uses magic leap to map the environment and label important features (bed, bathroom, sink, chair, TV, fridge, emergency path, etc.)
- Localization and navigation with no calibration process
- The interface is implemented in Unity Engine
- ROS# is a ROS (Robot Operating System) bridge for Unity
- Unity application communicates with ROS



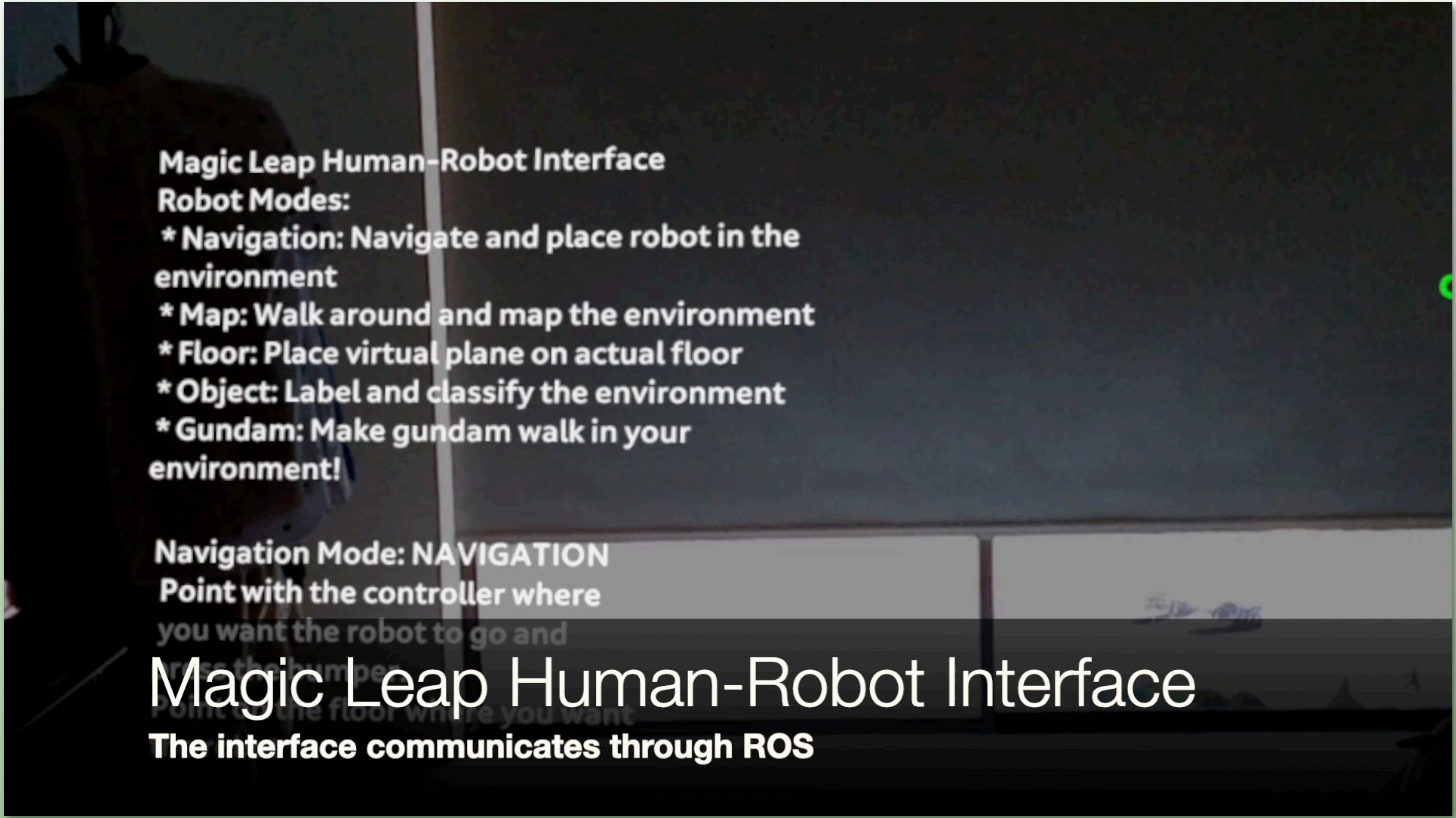
INTERFACES: IMPLEMENTATION (MAPPING AN ENVIRONMENT)

- The meshing feature of Magic Leap is used to create a 2D map
- A virtual camera in the Unity scene is placed on top of the scene generating a birds-eye view of the environment
- The image of the virtual camera is rendered in a texture
- The texture is used by the map server script in Unity to generate a occupancy grid map used by the navigation stack of the robot
- Along with the 2D costmap, the robot sends the transformation from Unity to ROS obtained from Magic Leap’s global frame





MAGIC LEAP INTERFACE FOR HSR



5. FACIAL RECOGNITION: DETECTING RACIAL INEQUALITIES IN CRIMINAL JUSTICE

- Recent events have highlighted large-scale systemic (race, gender, skin tone, etc.) disparities in U.S. criminal justice.
- Propose an experimental methodology based on ethical AI principles to generate binary racial categories using mugshots
- Data: ~200K defendants from Miami-Dade County Clerk of Records
- Ground truth: 2 sources
  - Dataset (N=14,177 random images) labeled by official court records single rater
  - Dataset (N=14,018 random images) is formed using consensus-driven racial categorization by multiple raters

Ousmane Bah (left) was arrested but let go after a detective realized he looking nothing like the suspect (right) caught stealing in a from an Apple store. (Photos courtesy of KTVU and Subhan Tariq)

N.Y. Teen Blames Apple's Facial Recognition for Wrongful Arrest, Files $1B Lawsuit

By Tanasia Kenney - April 29, 2019 4084 0

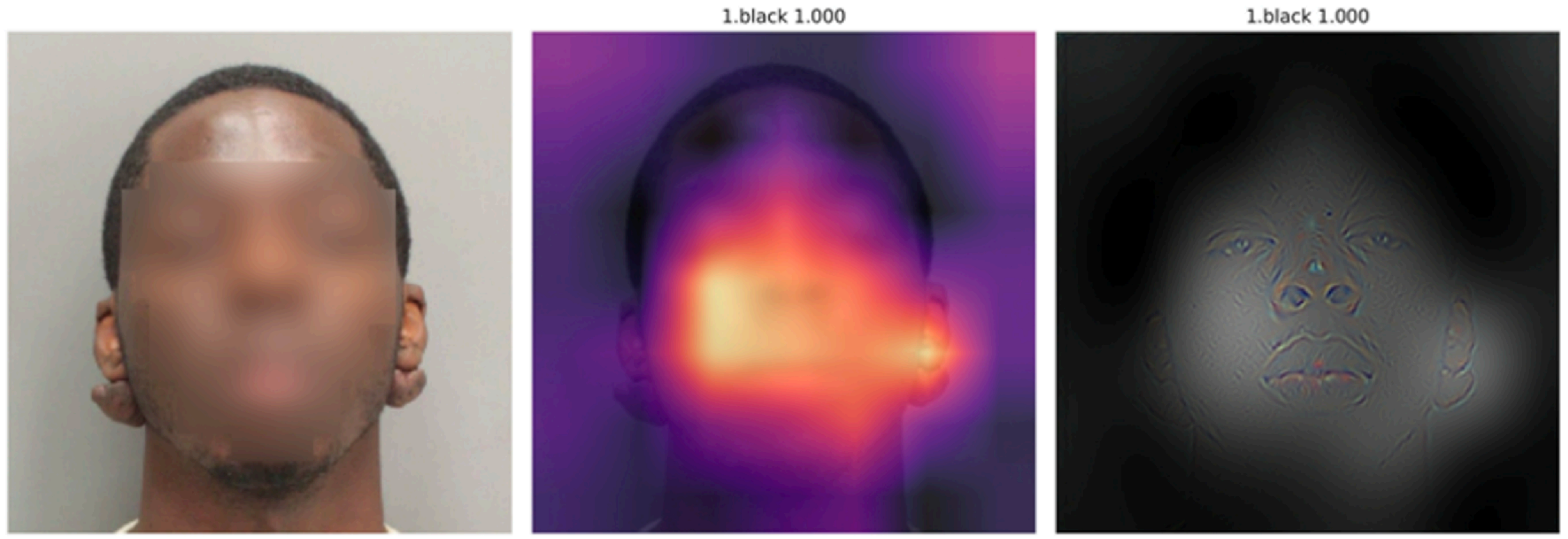
DETECTING RACIAL INEQUALITIES IN CRIMINAL JUSTICE

- First study showed 90%+ accuracy for race (black/white), < 75% accuracy for ethnicity (black/white/black hispanic/white hispanic) with SOA FRT (U-Link project)
- Results show that data preprocessing is crucial

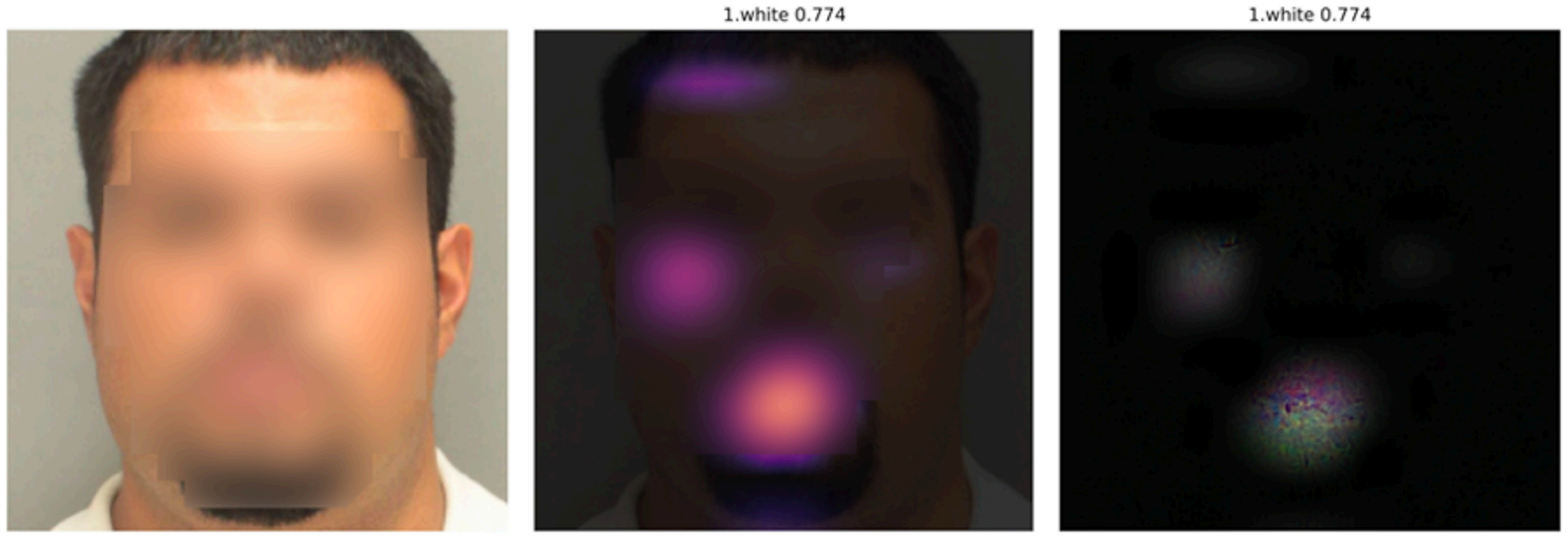
- DL library *fastai* using 7 state-of-the-art computer vision architectures (DenseNet161, ResNet50, InceptionV4, SE-ResNet50, SE-ResNeXt50_32x4d, AlexNet, and VGG19_bn)
- Trained with different preprocessing steps using our 2 sources, test on remaining images
- Highest accuracy is 97.75% (SE-ResNet50 with OpenFace preprocessing)
- Lowest accuracy is 1.28% (InceptionV4 with MTCNN preprocessing)

Dass et al. (*CRV, 2020*)

DETECTING RACIAL INEQUALITIES IN CRIMINAL JUSTICE



(c) “Best” Black mugshot by MTCNN preprocessed court trained SE-ResNet50 model.



(b) “Worst” Black mugshot by OpenFace preprocessed court trained DenseNet161 model.

(e) “Best” Black mugshot by OpenFace preprocessed court trained InceptionV4 model.

(f) “Worst” Black mugshot by OpenFace preprocessed court trained InceptionV4 model.

DETECTING RACIAL INEQUALITIES IN CRIMINAL JUSTICE

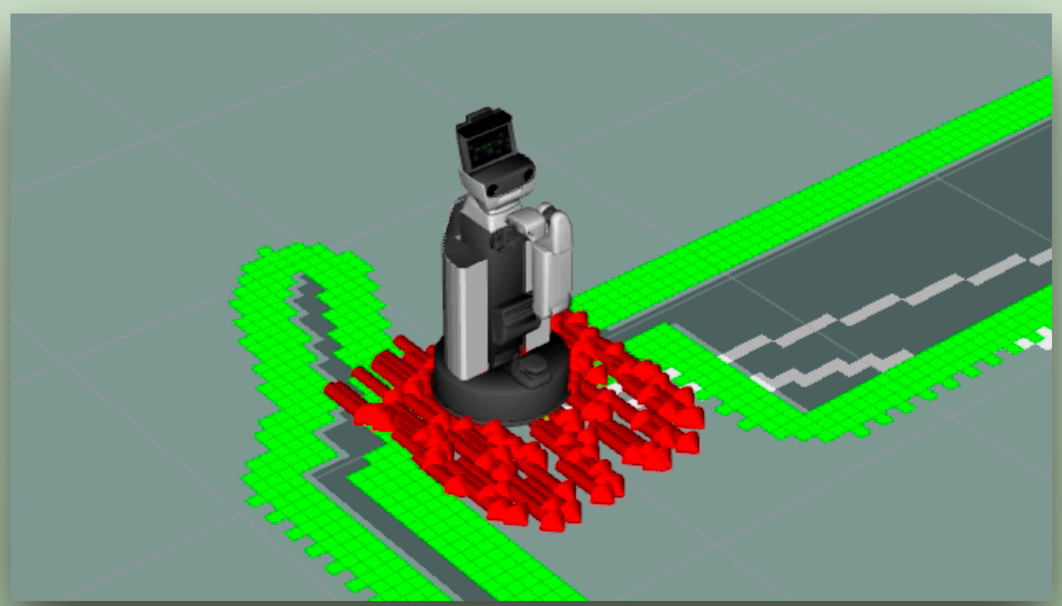
- Method
  - Data generation (multiple “ground truths” to tackle **labeling bias)**
  - Data preprocessing (face segmentation, aligning, cropping, pose, illumination, …)
  - DLM architectures (7 SOA methods with 6 experimental combinations) to tackle **algorithmic bias**
  - Training setup to avoid **representation bias**
  - Testing (model inference and interpretability) to tackle **historical** and **evaluation bias; self-auditing** (using 42 fine-tuned models for 12 test scenarios: 252 in total)

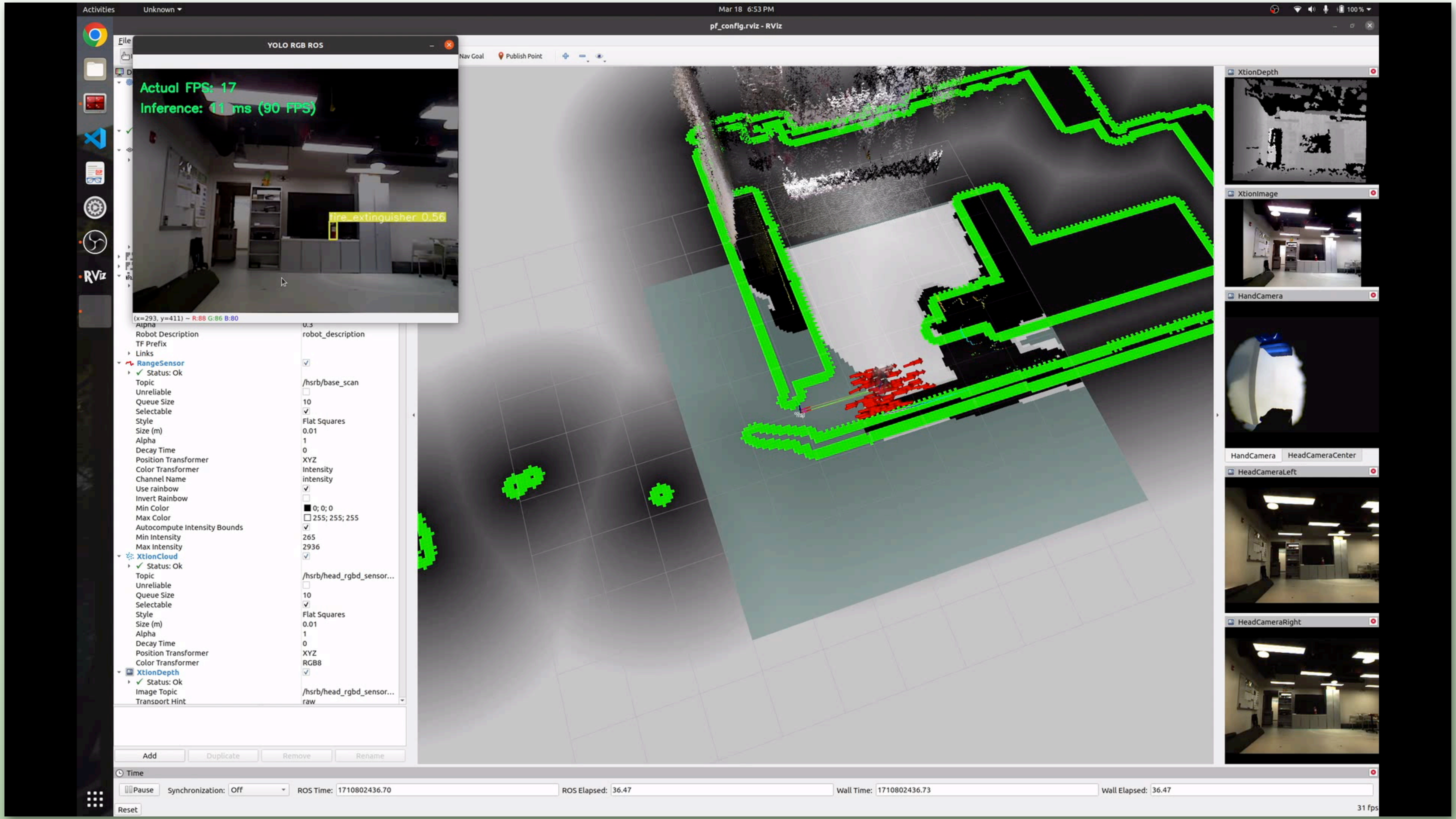
Table 4: A comparison of **validation accuracies** during the training process from seven deep learning-based vision models.

Model	**original**		**OpenFace**		**MTCNN**	
	Courts	Students	Courts	Students	Courts	Students
AlexNet	92.50%	**94.00%**	97.25%	92.25%	92.75%	93.00%
DenseNet161	**97.00%**	93.00%	97.00%	92.25%	96.50%	94.00%
InceptionV4	93.25%	90.25%	90.75%	90.50%	90.75%	91.75%
ResNet50	96.50%	92.50%	97.00%	91.50%	95.25%	93.25%
SE-ResNet50	95.00%	92.75%	**97.75%**	**92.50%**	96.75%	93.25%
SE-ResNeXt50	96.75%	91.75%	97.75%	90.00%	96.75%	94.25%
VGG19	96.00%	89.75%	97.00%	92.25%	**96.75%**	**94.75%**

6. ROBOT LOCALIZATION AND PATH PLANNING

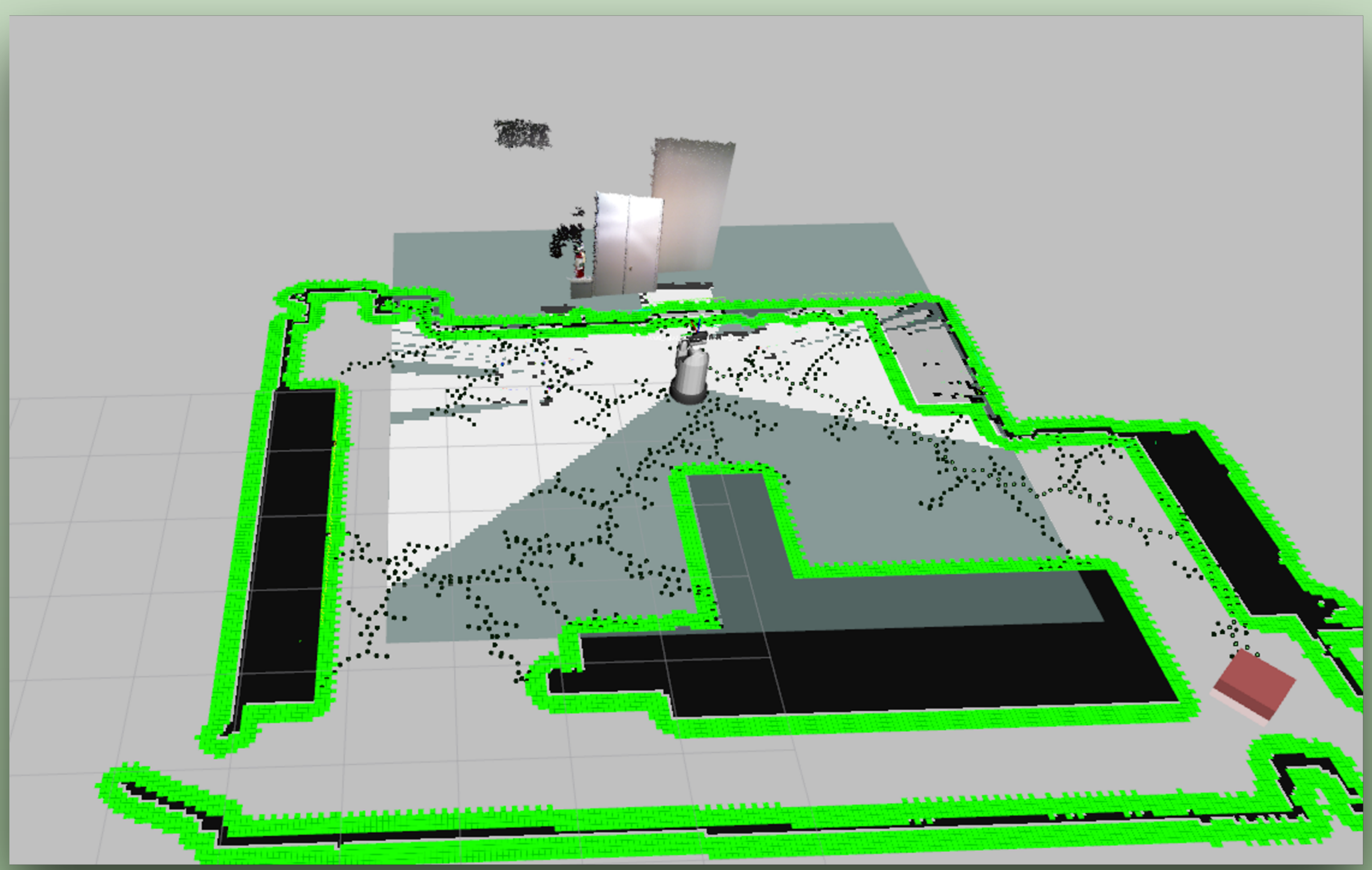
- Self-localization with Particle filter (MDP)
  - Based on data from LiDAR
  - Based on visual landmarks

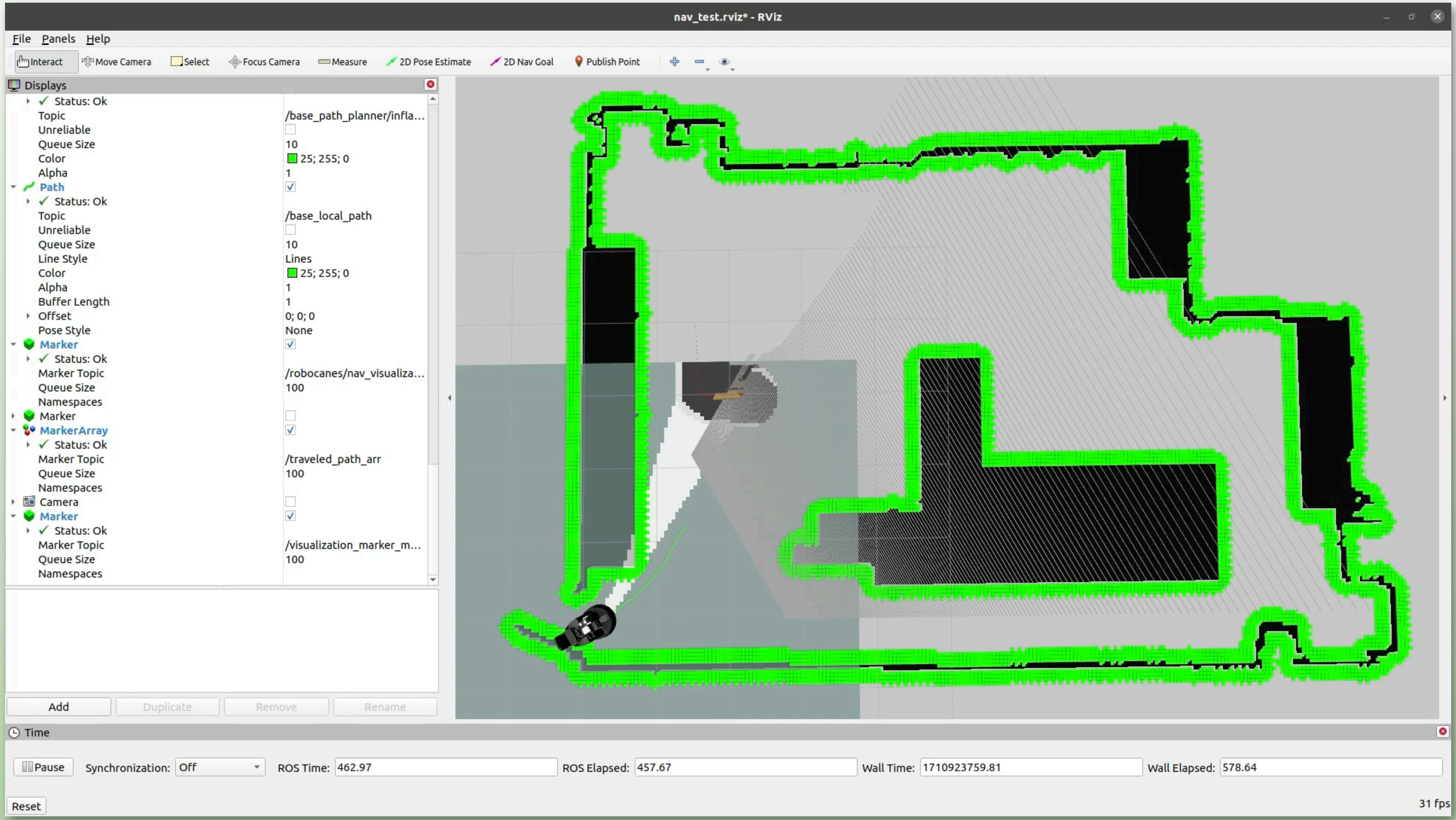




6. ROBOT LOCALIZATION AND PATH PLANNING

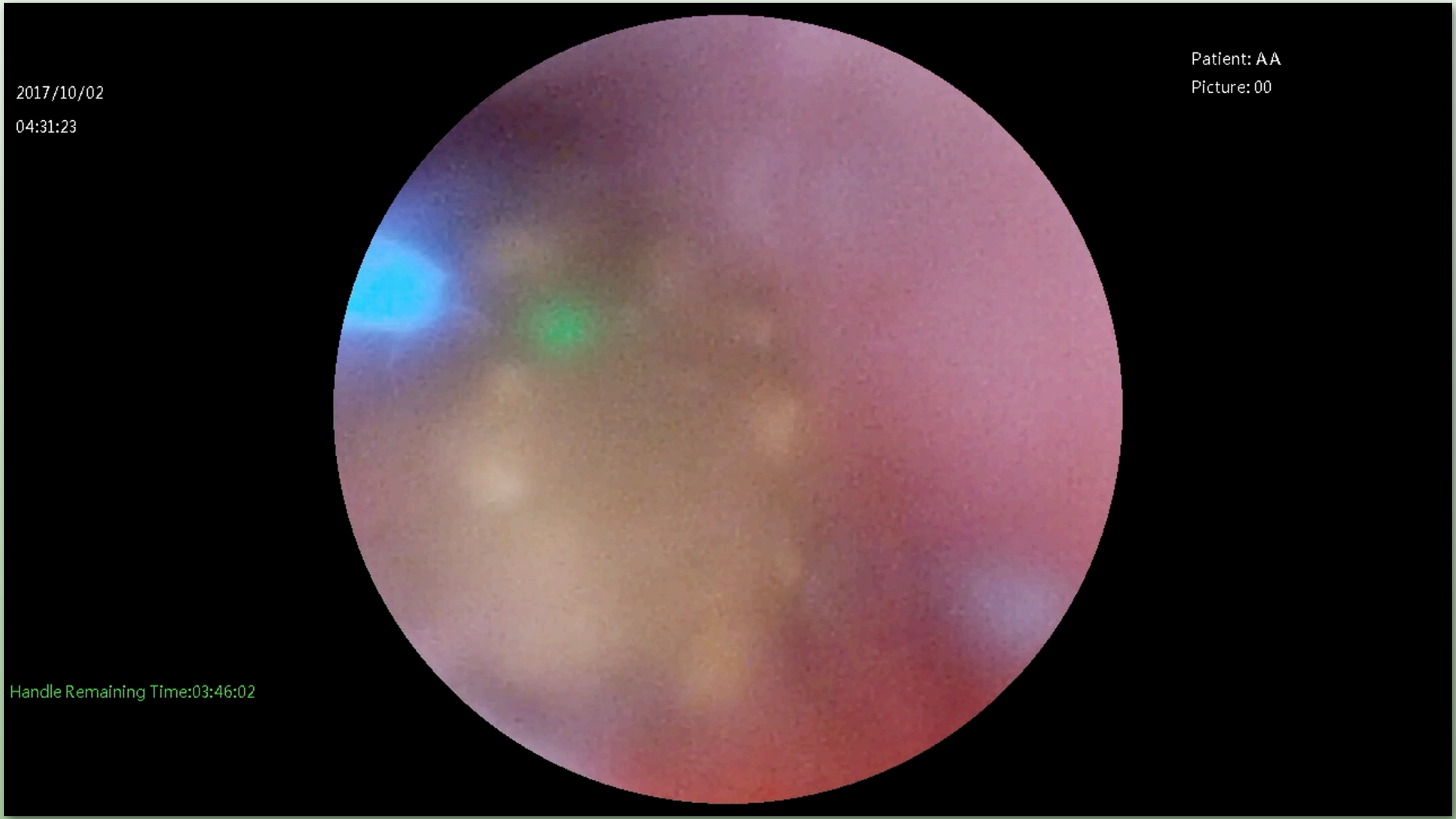
- Path planning with sample-based method: RRT
  - Rapidly exploring Random Trees
  - Smoothing for final path





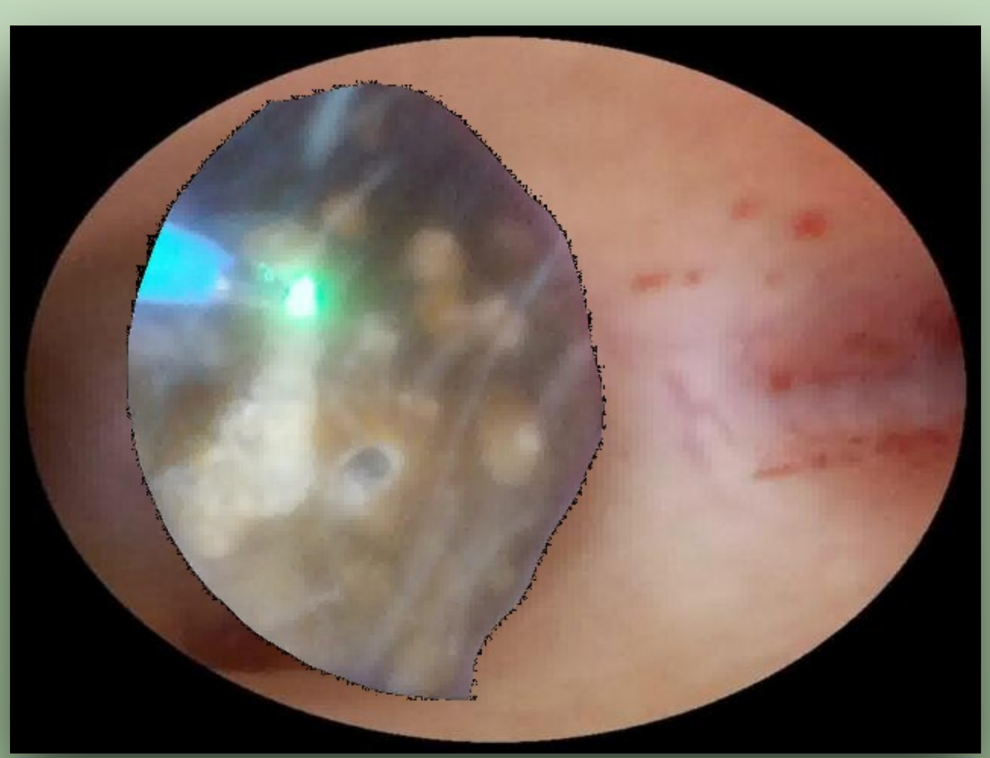
7. GLIDAR

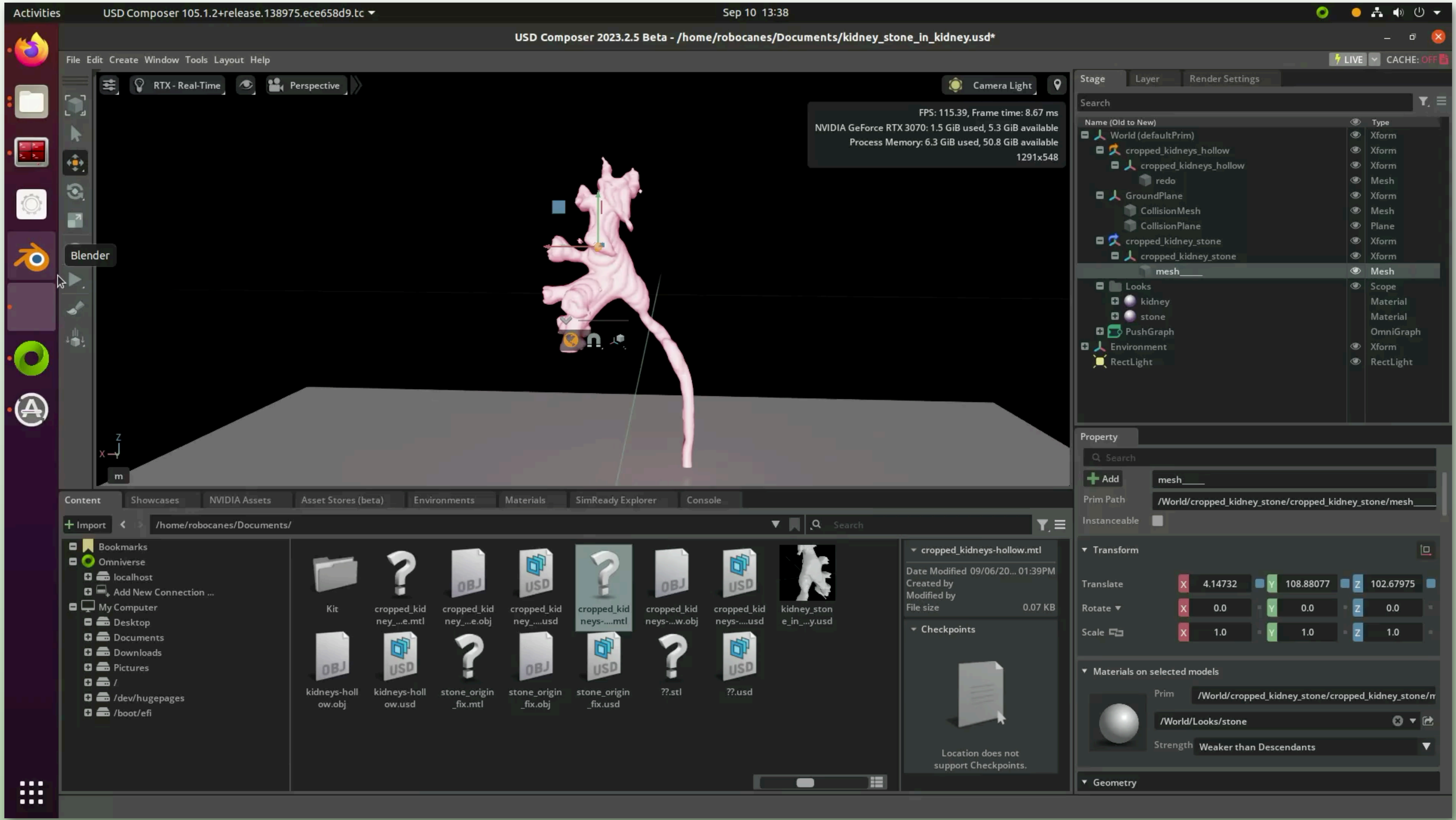
- Guided Lithotripsy with Dynamic Autonomous Robot



7. GLIDAR

- Simulation of renal collecting ducts
  - Renal calculus dynamics based on physical simulation
  - PhysX and Isaac Sim
  - Segmenting renal calculus



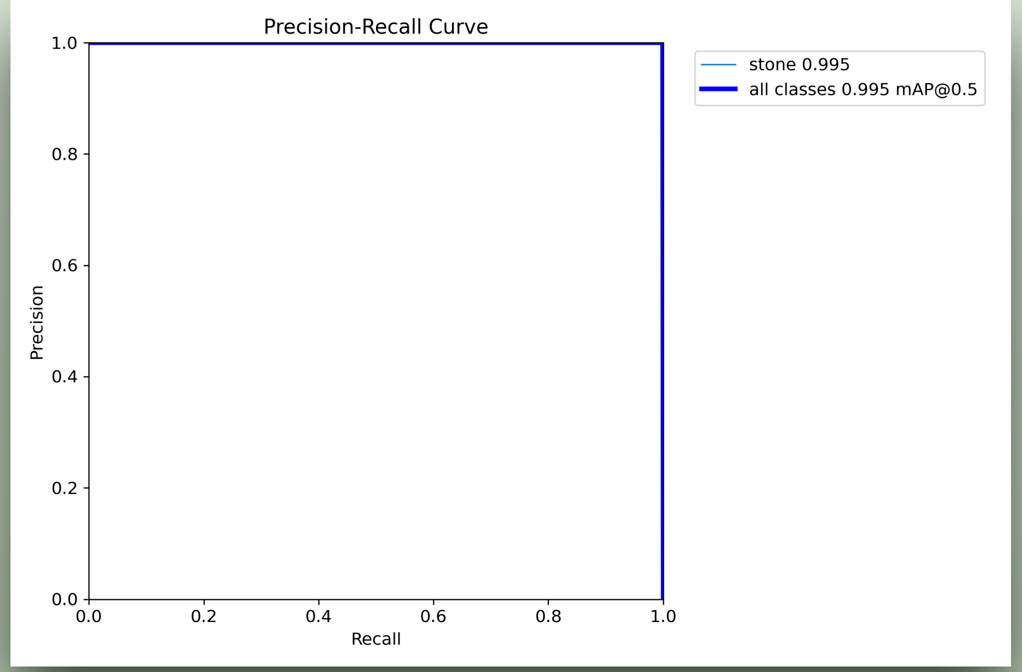


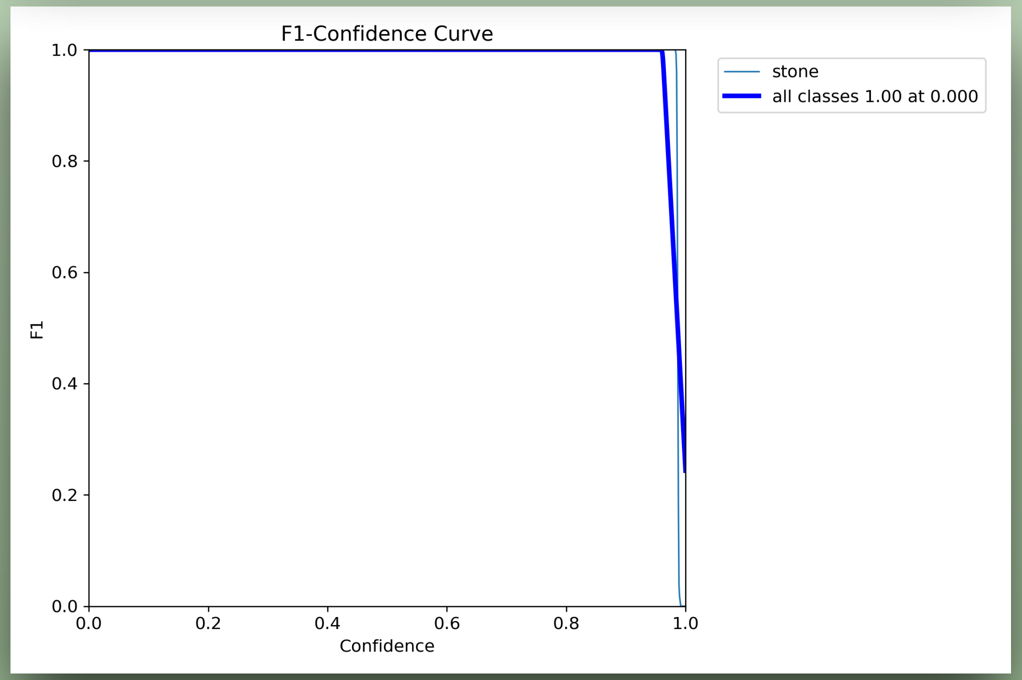
7. GLIDAR

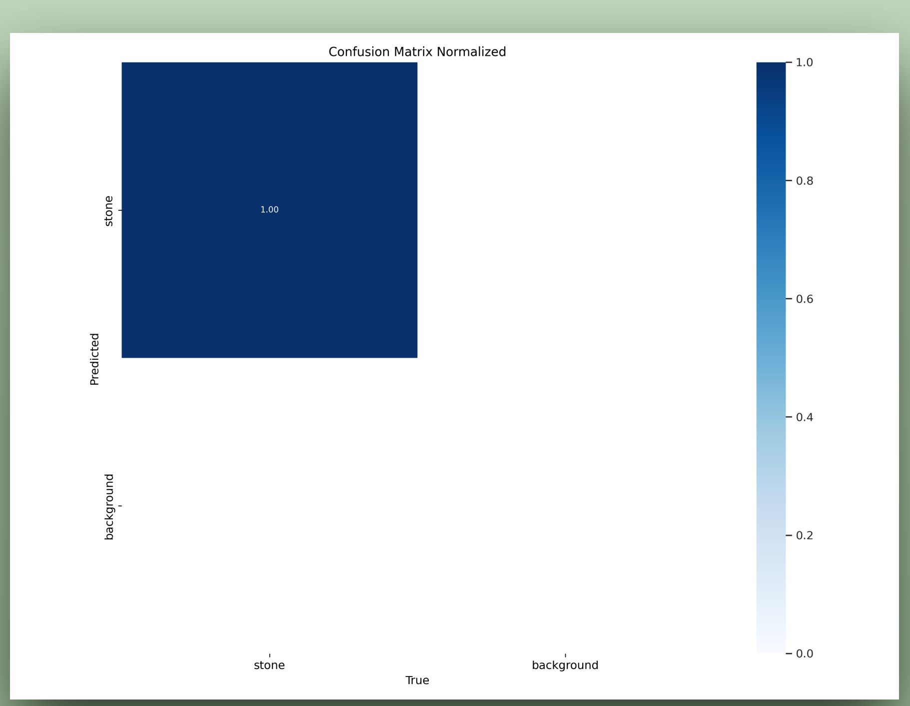
Autonomous Robot-Assisted Ureteroscopy (ARA-URS) for the Treatment of Kidney Stones

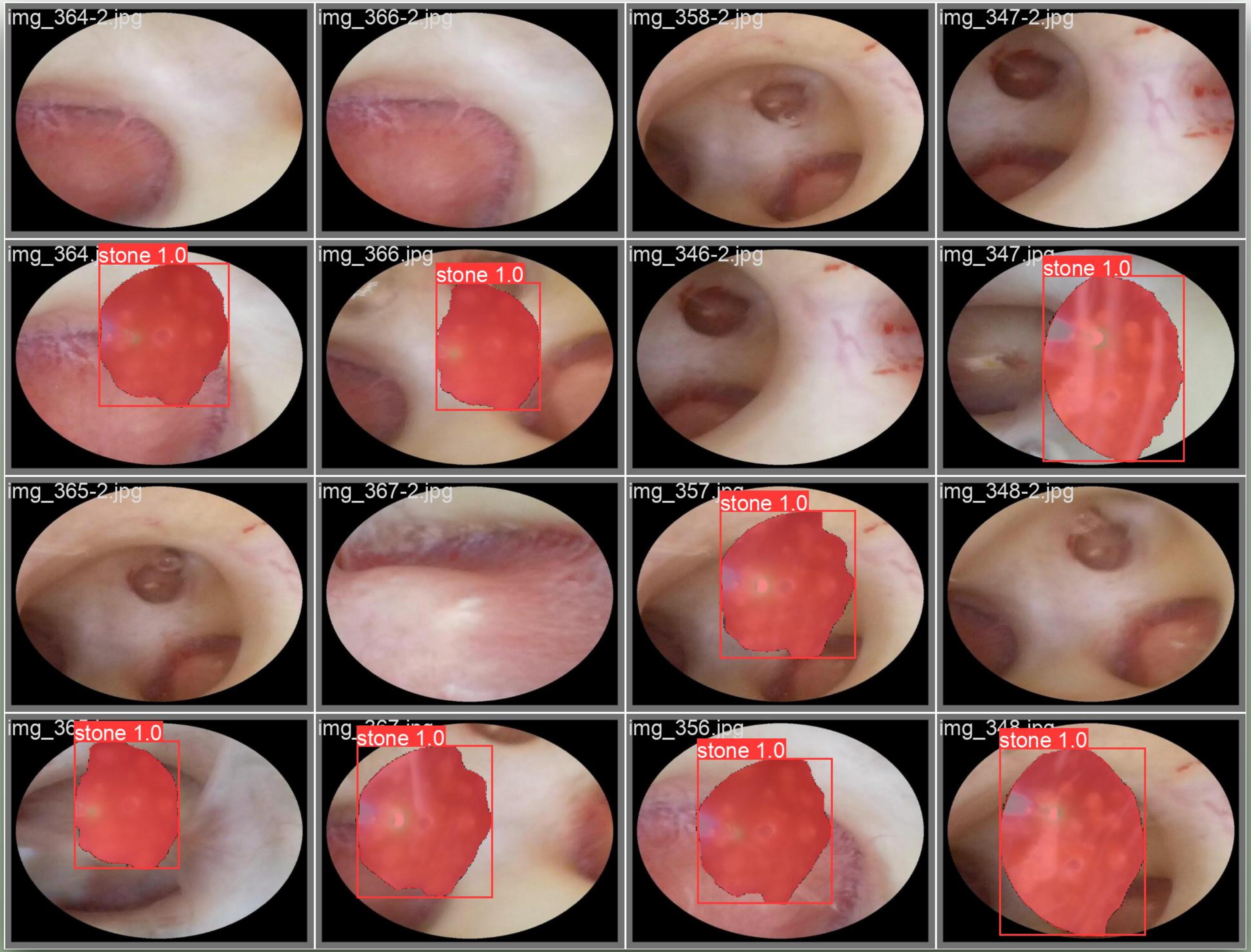
7. GLIDER

- Batch prediction
- Precision/recall
- F1
- Confusion matrix









CONCLUSIONS

- Task and motion planning
- State estimation and perception
- Communication
- Object grasping and placement
- Trajectory generation and control

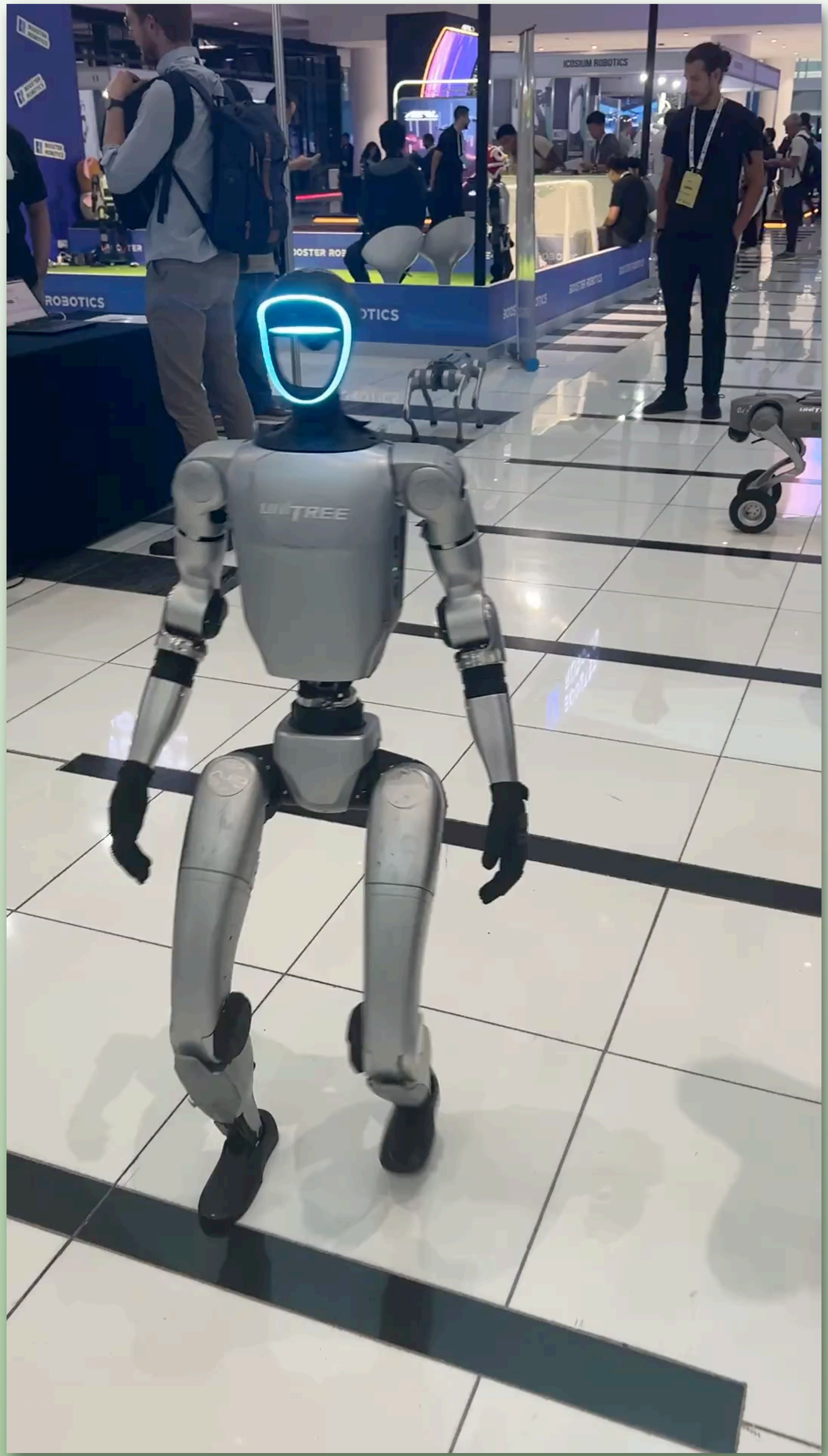
THOUGHTS

- Lidar (e.g. Strap)
- SLAM (e.g. real-time visual SLAM with RGB-D)
- Visual + Haptics (e.g. Maptic)
- Visual + Audio (e.g. Horus)
- Visual + VR (e.g. MyEye)

APPEARANCE

Toyota's HSR-B and HSR-C

HUMANOIDS

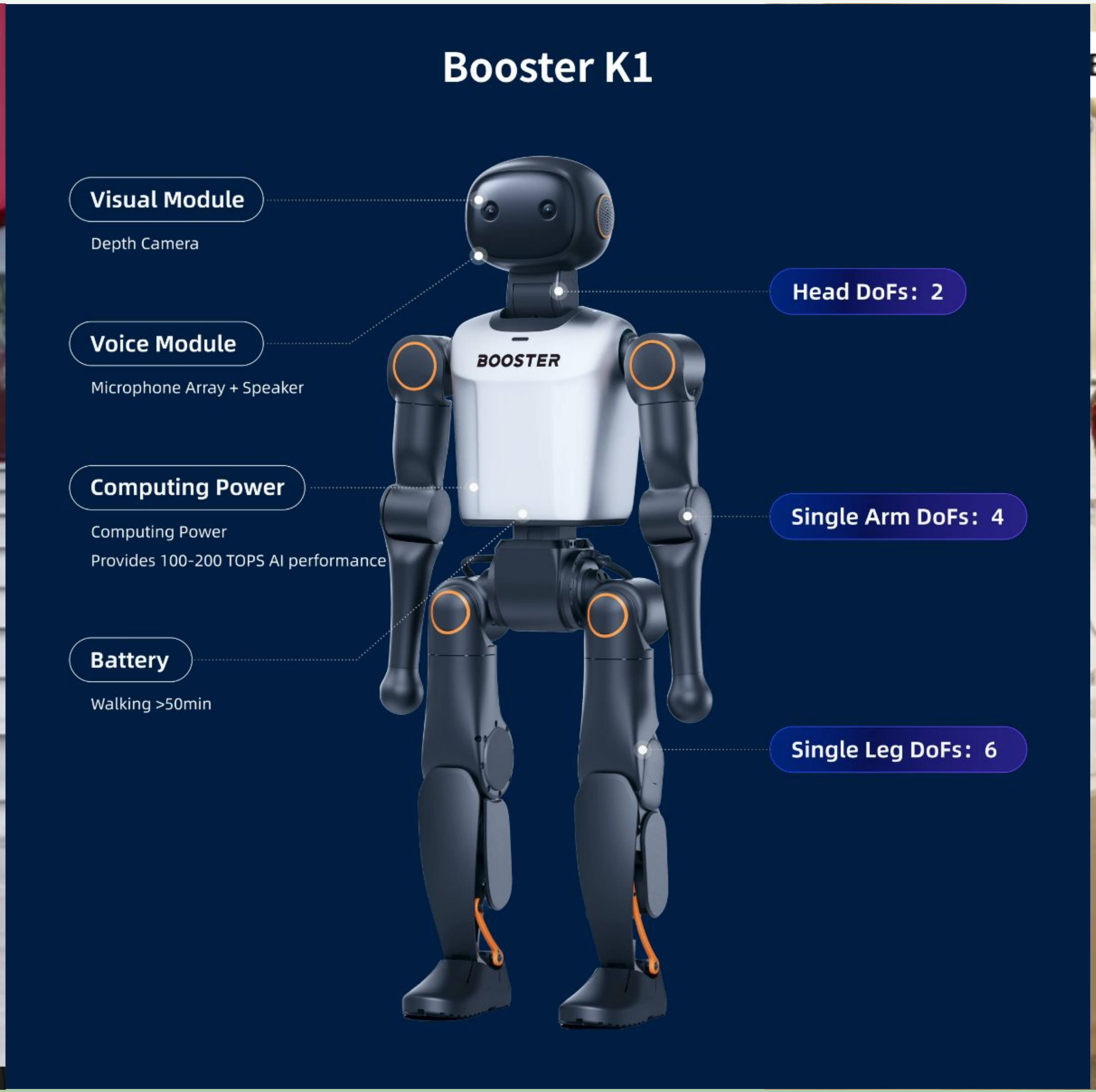


Unitree G1: https://www.unitree.com/

Booster T1: https://www.boosterobotics.com

Fourier Intelligence GR2, https://www.fftai.com/

HUMANOIDS



Unitree R1: https://www.unitree.com/

Booster K1: https://www.boosterobotics.com

Fourier Intelligence N1, https://www.fftai.com/

THANK YOU!


Saminda Abeyruwan


Chloe Arluck


Lloyd Beaufils


Michael Davis


Rahul Dass


Alexander Härtl


Nasir Laskar


Shengxin (Tony) Luo


Joe Masterjohn


Phil McKenna


Katarzyna (Kasia) Pasternak


Pedro Peña


Kyle Poore


Andreas Seekircher


Justin Stoecker


Zishi Wu


Chris Duarte

With help from the VISAGE group (FIU): Dr. Christine Lisetti, Dr. Mihai Polceanu, Stephanie Lunn; ENIB/FRANCE: Dr. Cédric Buche

- Contact
  - Web: https://www.cs.miami.edu/~visser
  - Email: visser@cs.miami.edu