WORKSHOP #58:

Perceiving Unknown and Deformable Objects in Logistics and Service Robotics

Agenda

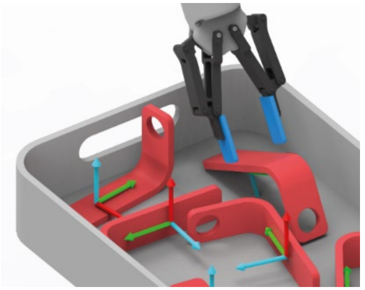
PERCEIVING UNKNOWN AND DEFORMABLE OBJECTS IN LOGISTICS AND SERVICE ROBOTICS

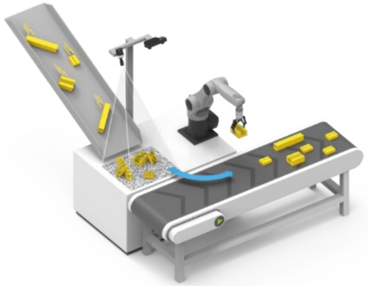
08:30	**Introduction and Definition of Key Statements / Questions** Dr. Michael Suppa, Roboception GmbH & University of Bremen, Germany
08:35	**Perception Challenges for Kitting in Automotive Assembly Lines** Florian Töper, Mercedes-Benz AG, Germany
08:45	**Insights on Pose Estimation and Grasp Prediction of Unknown Objects** Rudolph Triebel, Maximilian Durner, DLR, Germany
08:55	**Perceiving Deformable Linear Objects in Real-World Scenarios** Alessio Caporali, University of Bologna/Robosect srl, Italy
09:05	**AI-based Perception of Seen and Unseen flexible Objects** Dr. Michael Suppa, Roboception GmbH, Germany
09:15	**Interactive Poll Session / Round Table Discussion with the Audience**
09:45	**Closing Remarks and Take Home Messages**

Enabling AI Robotics

Perception is the Key Technology for Flexible Automation

INTRODUCTION





In flexible automation, robots must be able to reliably detect and locate work pieces and human collaborators - with varying illumination, work piece types and locations

- In **industrial automation**, accurate placement is usually the key challenge, rigid model data and process descriptions are available. The majority of objects is rigid.
- In **logistics**, manual work is still pre-dominant due to the complexity of tasks and the variation of objects. Process descriptions are available, and many objects are flexible i.e. bags
- In **service robotics,** process descriptions and process data is not available, many objects are non-rigid.

How to train, represent, and locate flexible and deformable objects is a cross domain key challenge for perception.

Key Questions

How are you dealing with flexible and deformable objects in your use cases? Please describe use cases such that we can foster the discussion!

Data sets and representation of flexible and deformable objects are difficult. How define a grasp point on a deformable object?

Synthetic training data generation helps to close the training data and ground truth gap. Do you have data for synthetic data generation of flexible objects available? Where does it come from and how do you assess the level of correctness?

INSIGHTS ON POSE ESTIMATION AND GRASP PREDICTION OF UNKNOWN OBJECTS

Perceiving Unknown and Deformable Objects in Logistics and Service Robotics

Rudolph Triebel / Maximilian Durner

ERF2025
STUTTGART | GERMANY
25–27 MARCH

DLR

Motivation

- Various robotic applications

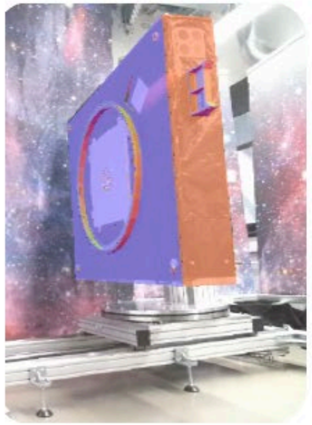
  > changing requirements

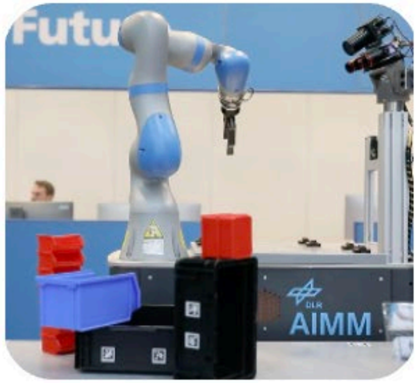
- General solutions:

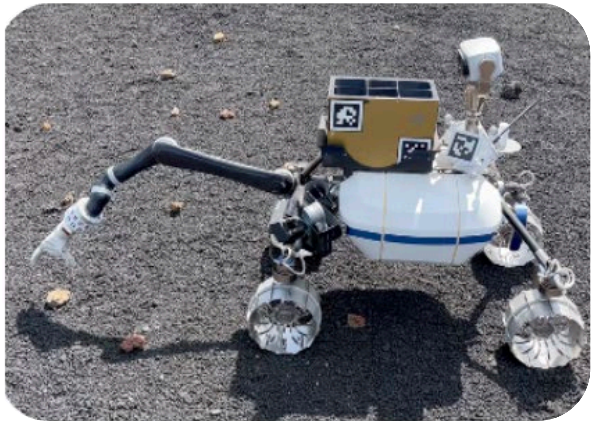
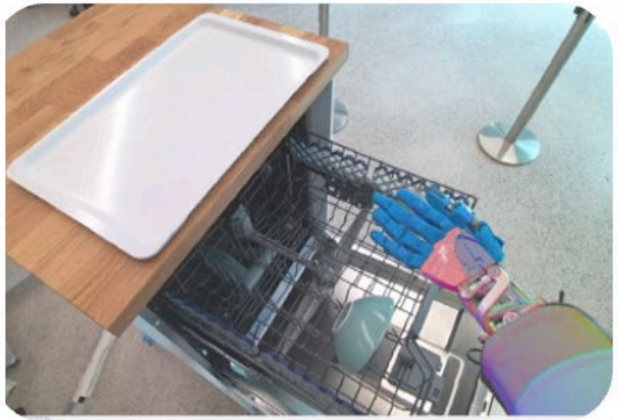
  > robust, precise, fast

- Attributes:
  - Object Information / Semantics
  - Training Effort / Adaptability
  - Rigidity / Flexibility







Synthetic training data

Indoor Simulation

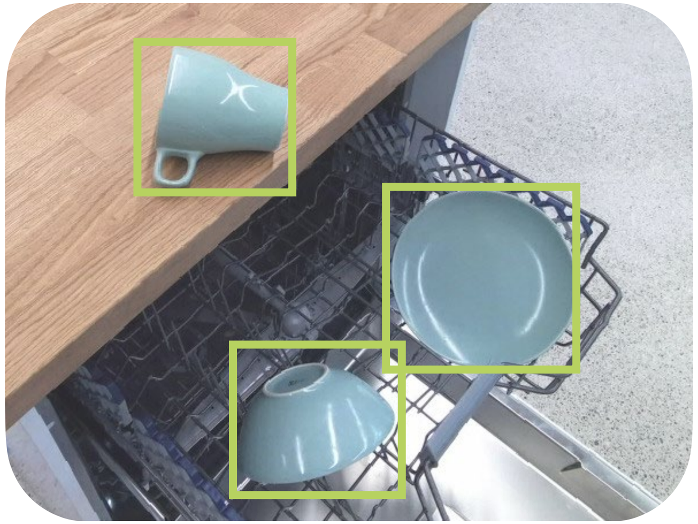
[Denninger et al., arXiv 2019]

Outdoor Simulation

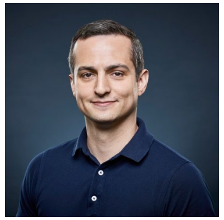
[Müller et al., IROS 2021]

Perception system

object
detector

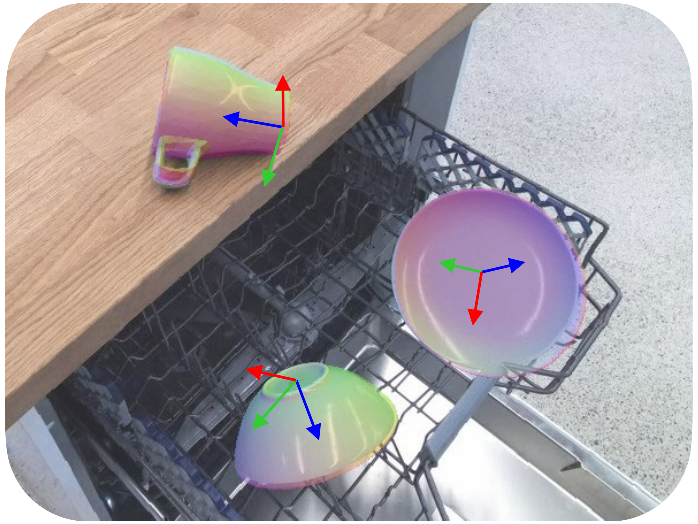


Perception system



DLR

object detector → 6D pose estimation

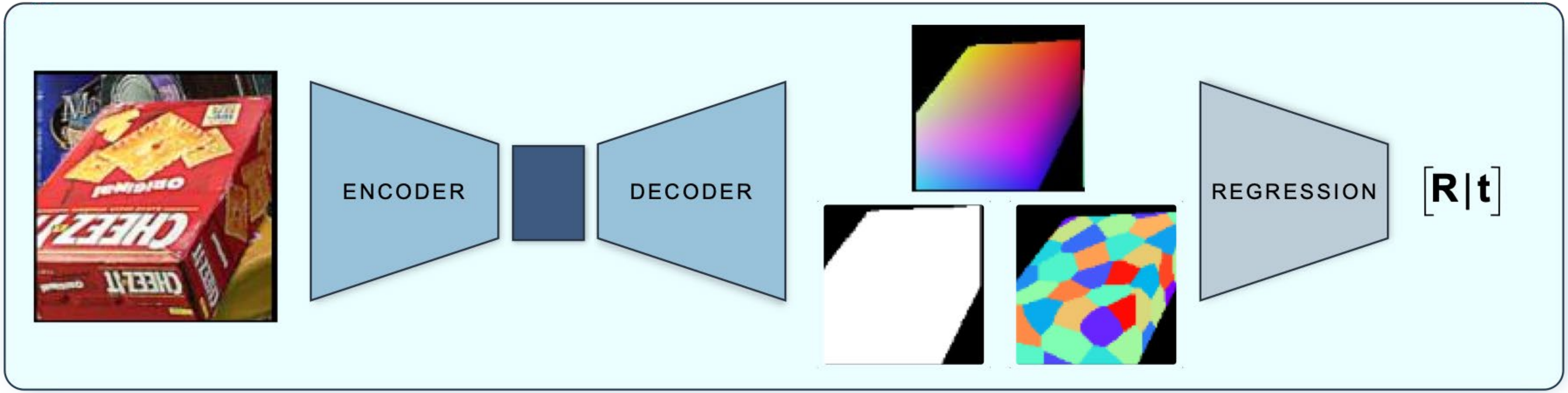


[Brucker et al., ISER 2019]

[Sundermeyer et al., ECCV 2018 (**Best Paper Award**)]

[Ulmer et al., IROS 2023 (**SPEED+ PostMortem Leader**)]

6D Pose estimation: Dense Correspondences

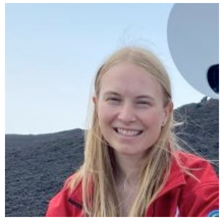


- Training purely on synthetic images
- RGB only approach
- Trained end-to-end

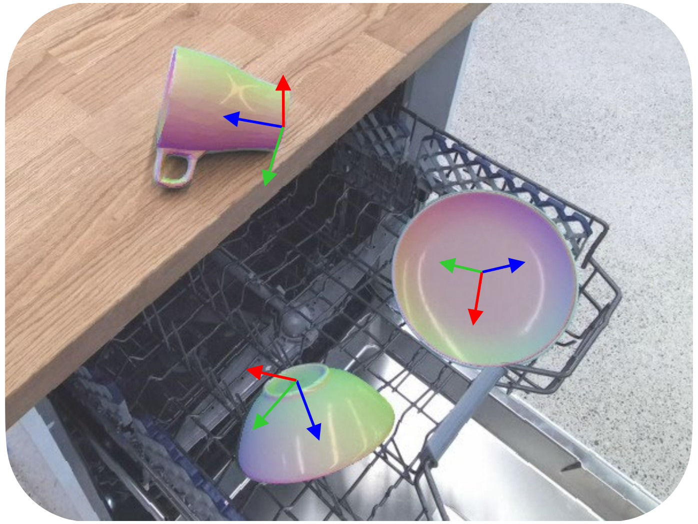
[Ulmer et al., IROS 2023 (**SPEED+ PostMortem Leader**)]

6D Pose estimation: Dense Correspondences

Perception system



object detector → 6D pose estimation → local refinement



[Stoiber et al., ACCV 2020 (**Best Paper Award**)]

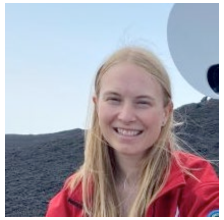
[Stoiber et al., CVPR 2022]

Local Refinement / 6D Tracking

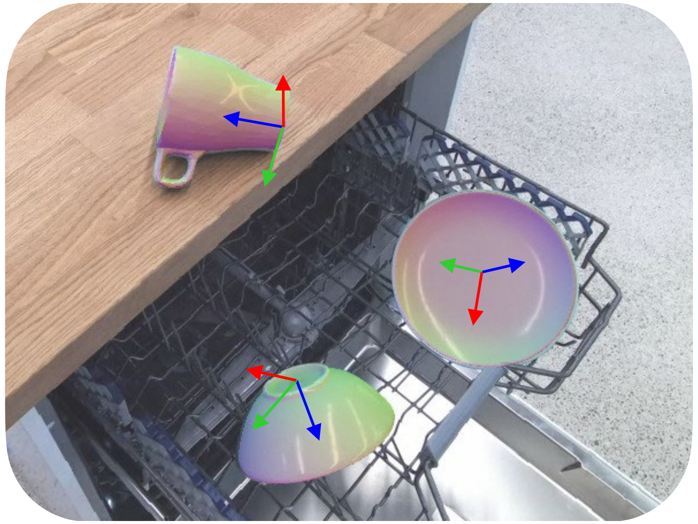
[Stoiber et al., ACCV 2020 (**Best Paper Award**)]

[Stoiber et al., CVPR 2022]

Perception system



object detector → 6D pose estimation → local refinement

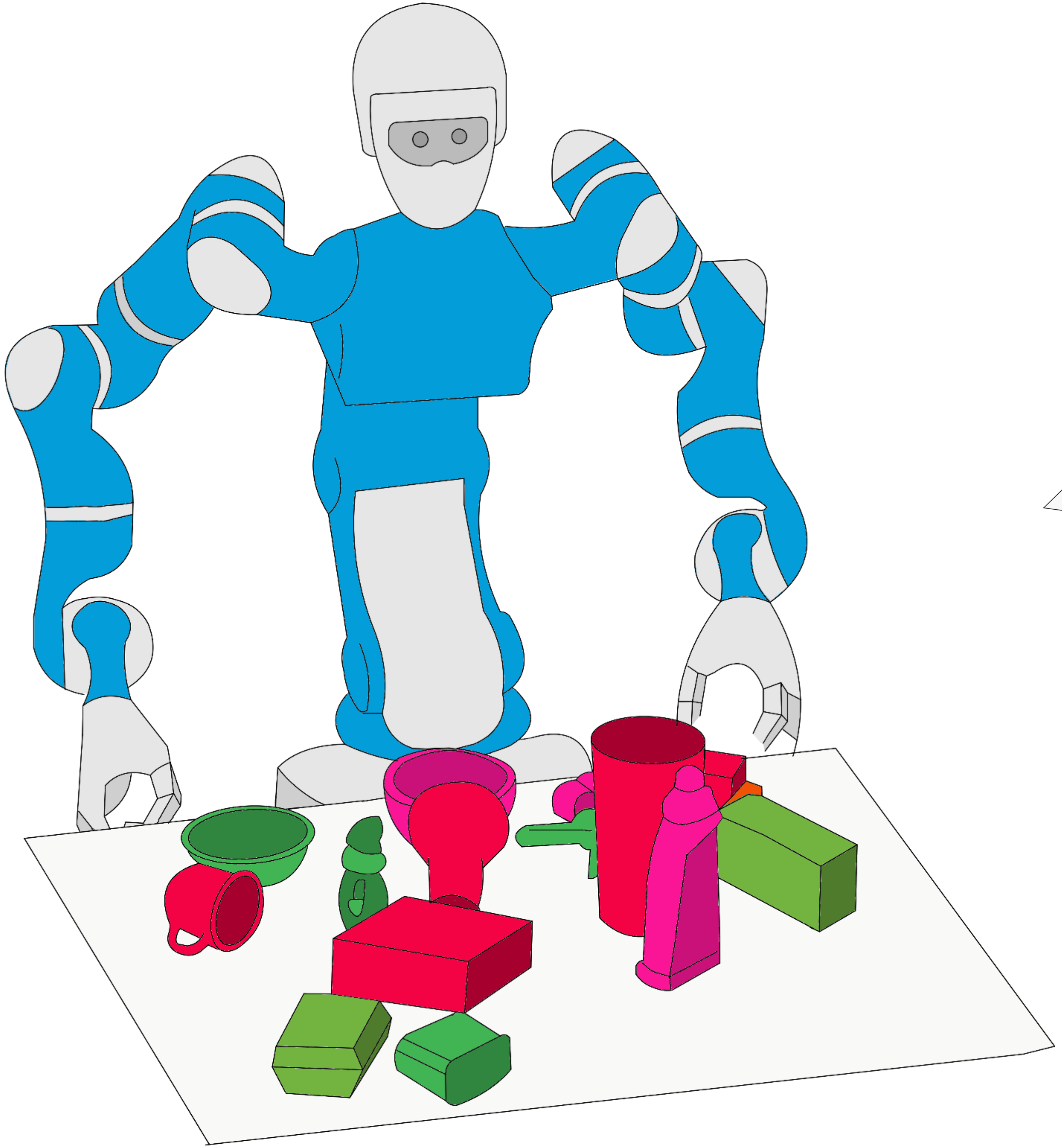


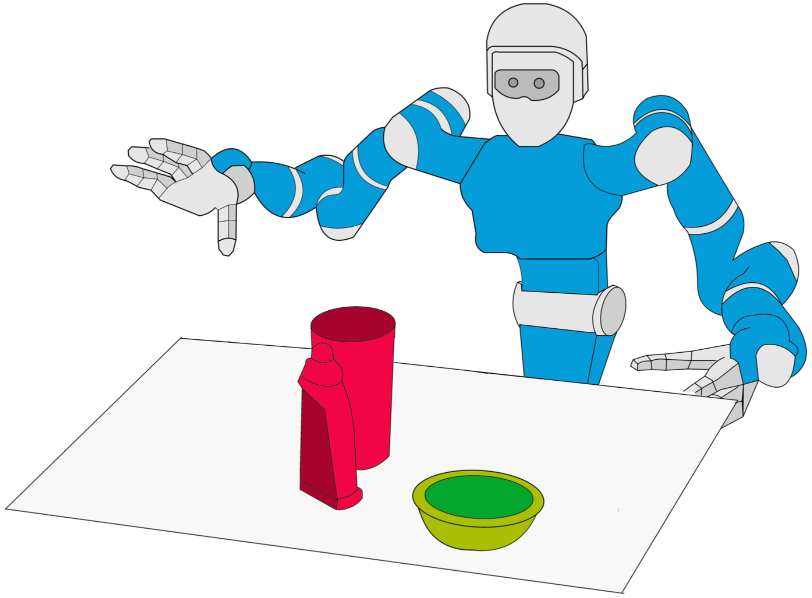
Object Information / Semantics

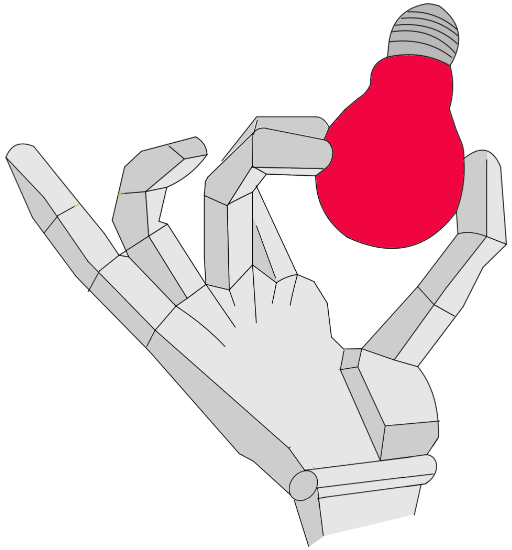
Set-Up Velocity

[Stoiber et al., ACCV 2020 (**Best Paper Award**)]

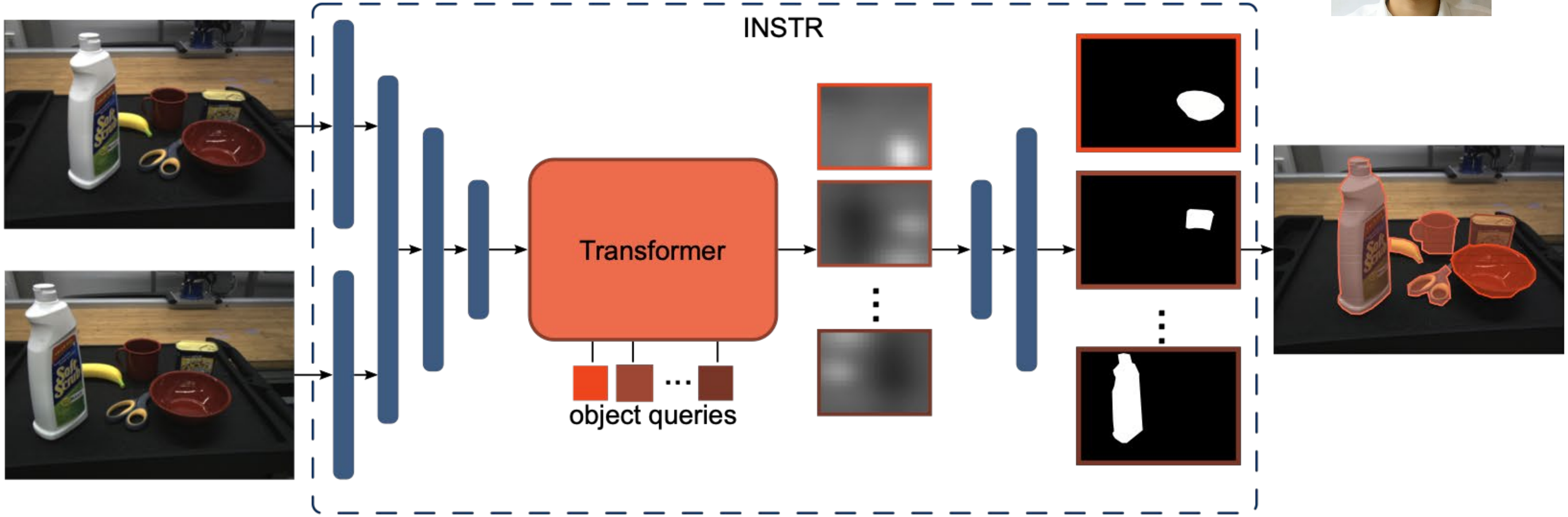
[Stoiber et al., CVPR 2022]







Class-agnostic Instance Segmentation



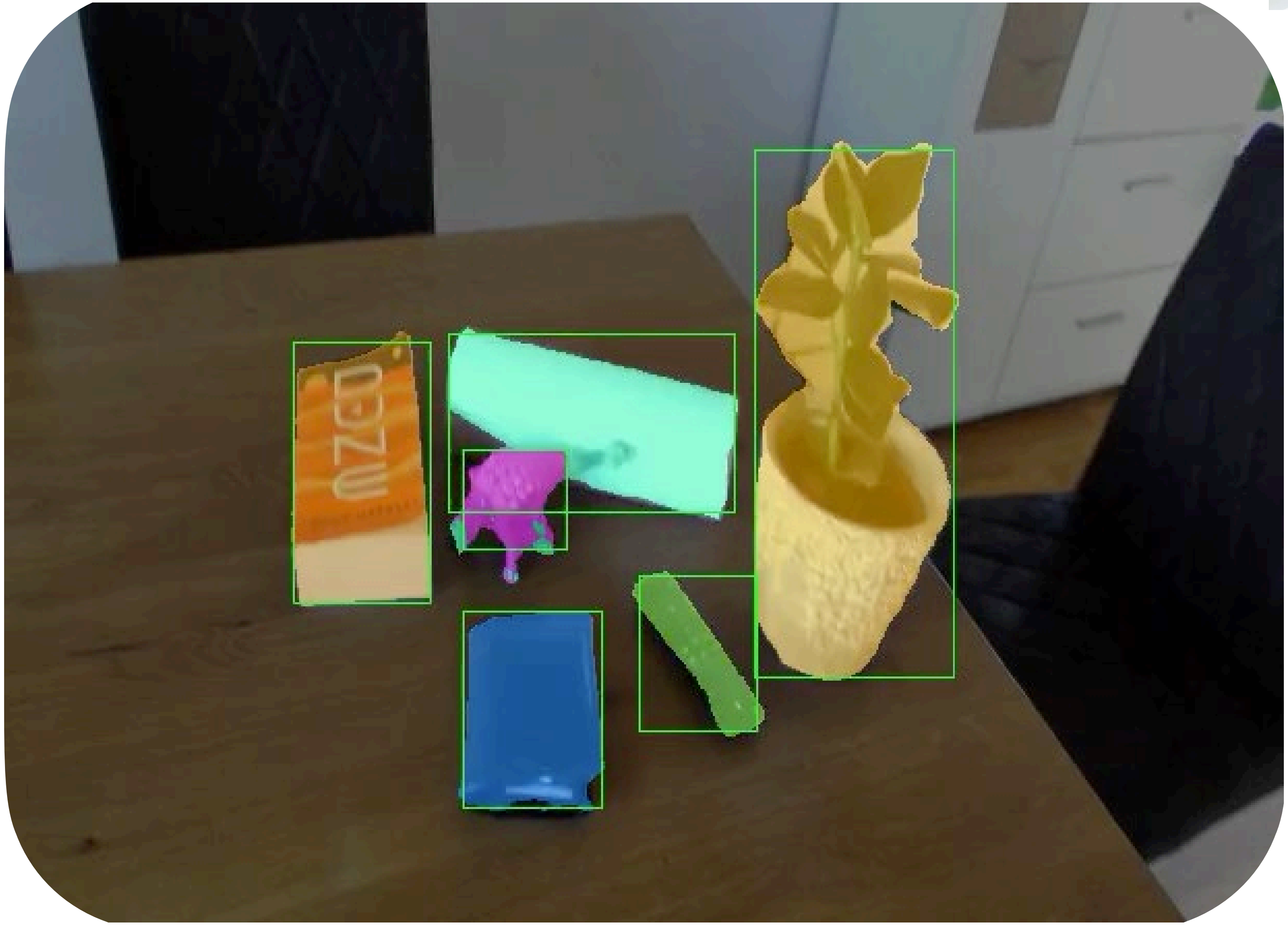
- Stereo-images as input
- Learns concept of objects
- No further training necessary

Object Information / Semantics

Set-Up Velocity

[Durner et al., IROS 2021]

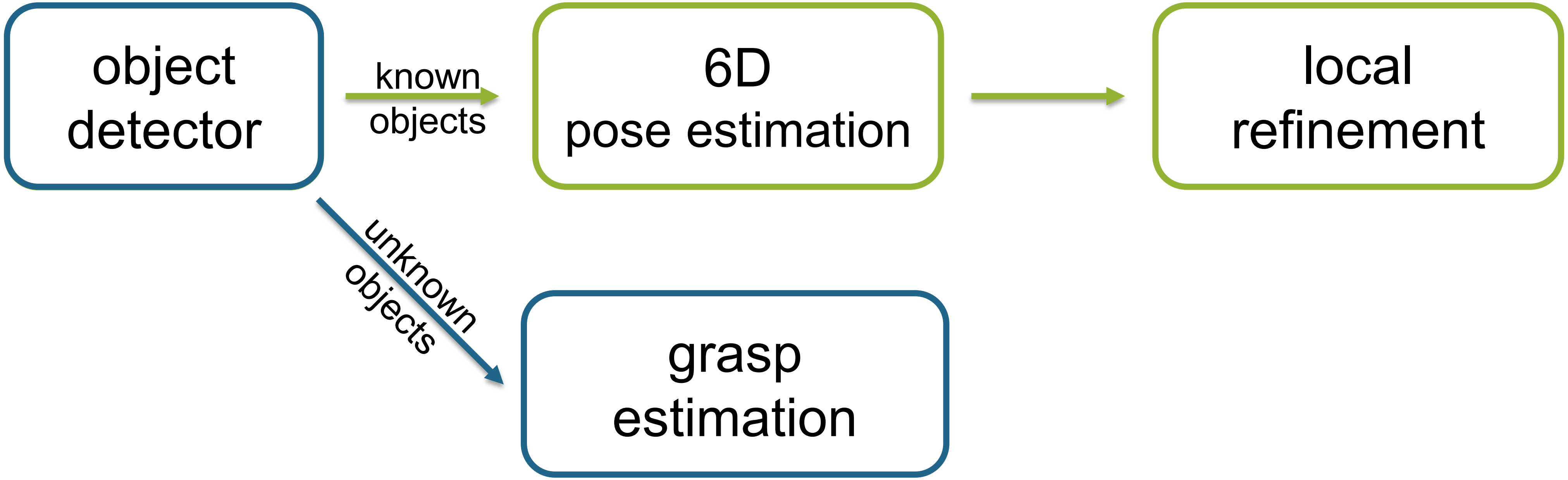
Class-agnostic Instance Segmentation



[Durner et al., IROS 2021]

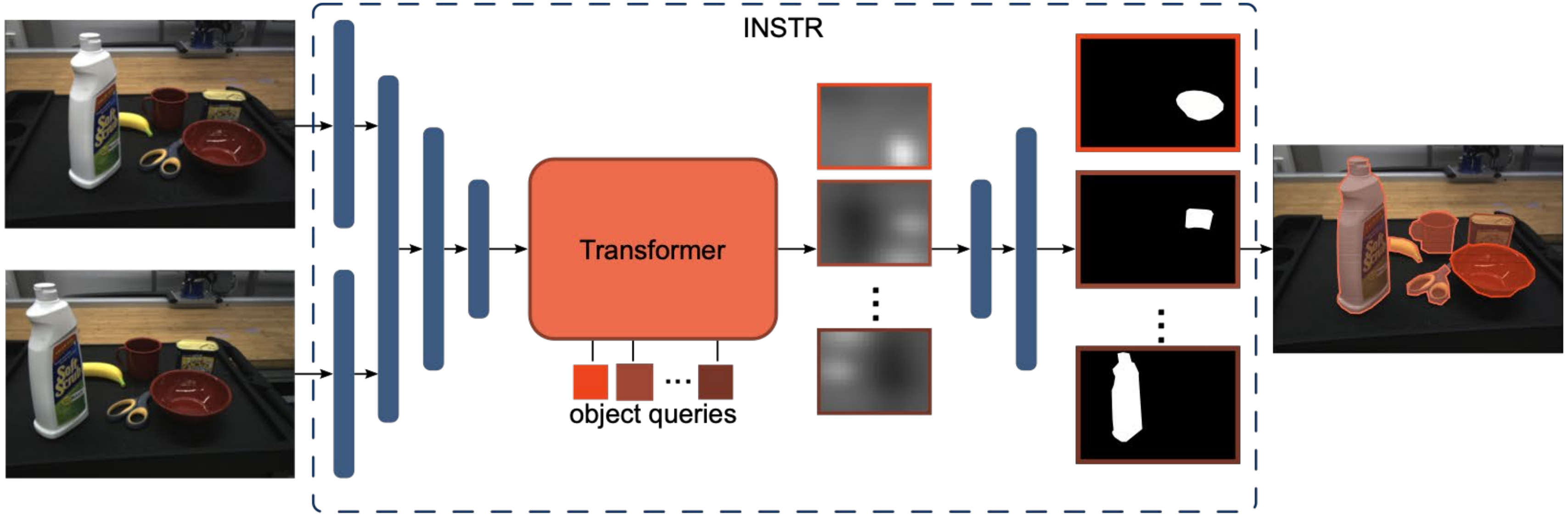
Perception system

Perception system



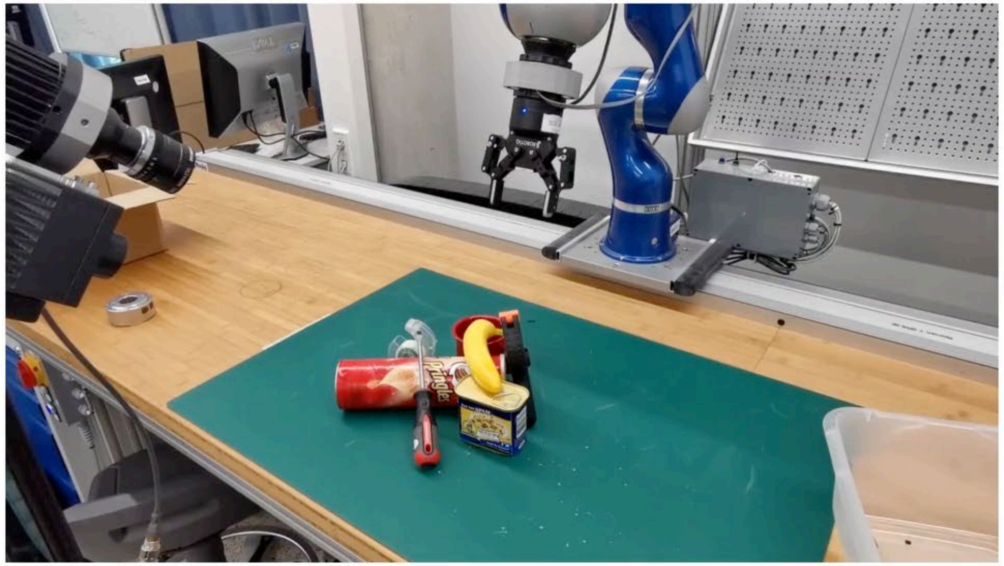
[Sundermeyer et al., ICRA 2021]
[Winkelbauer et al., IROS 2023]

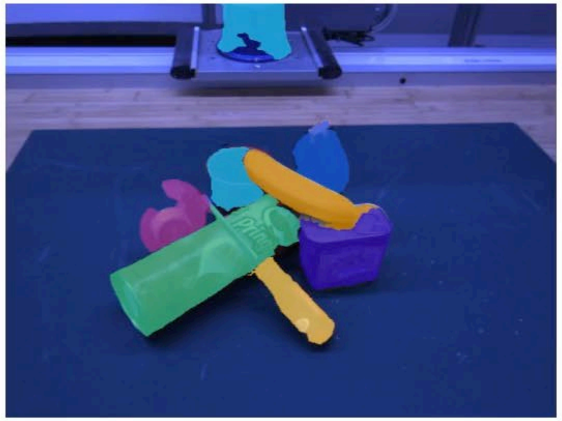
Class-agnostic Instance Grasping



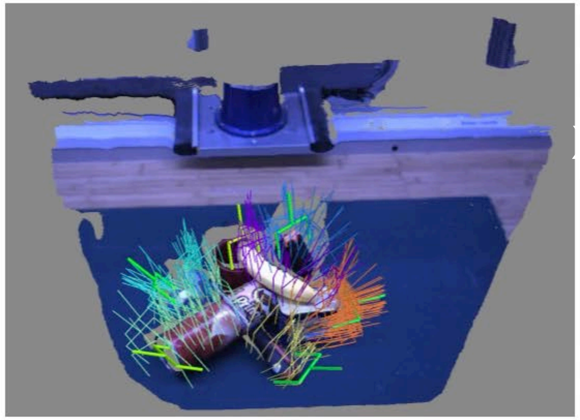
[Sundermeyer et al., ICRA 2021]

Perception pipeline





instance masks



6D grasp estimation

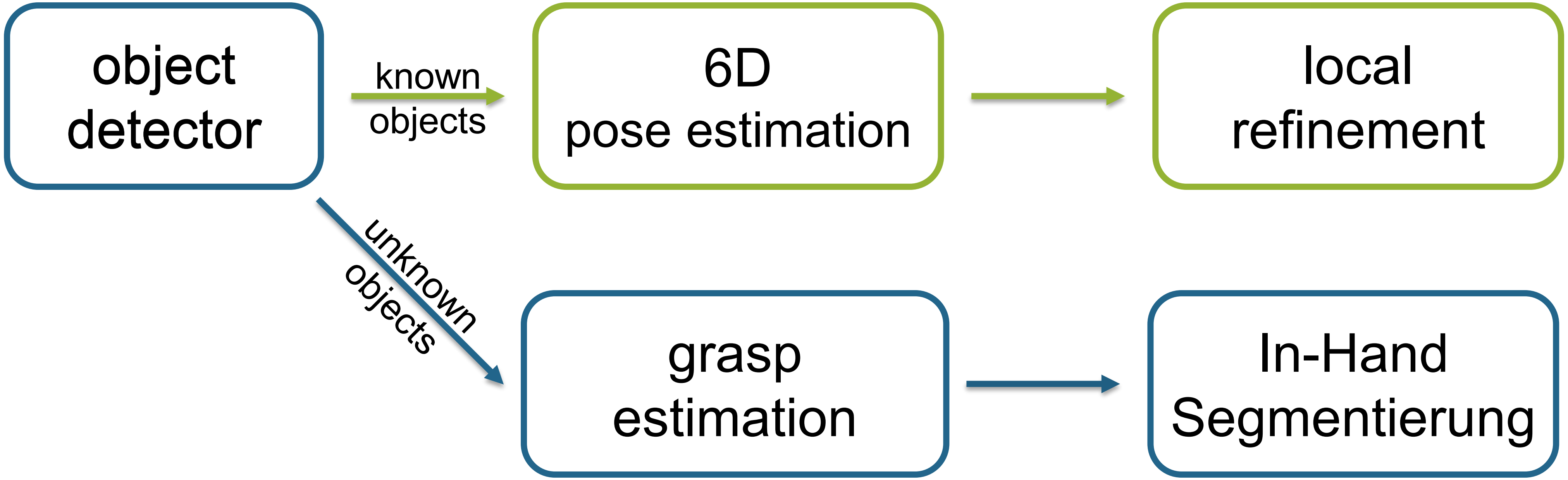
Perception system

[Winkelbauer et al., IROS 2022]

[Humt et al., IROS 2023]

[Miller et al., ICRA 2024]

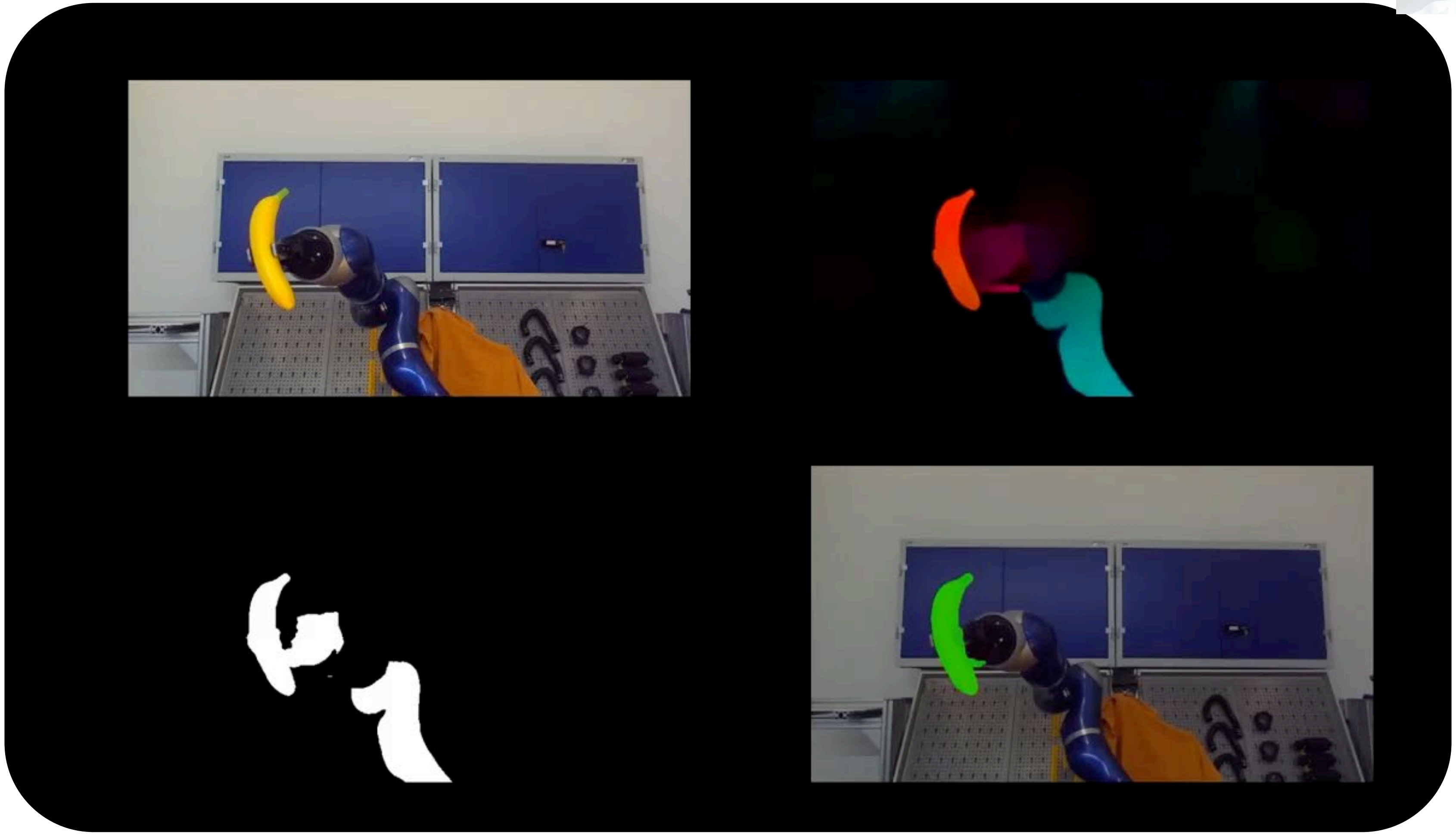
Perception system

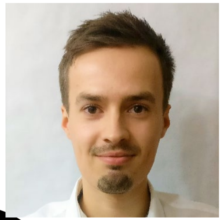


[Boerdijk et al., CORL 2020]

[Boerdijk et al., ICRA 2021]

Perception system

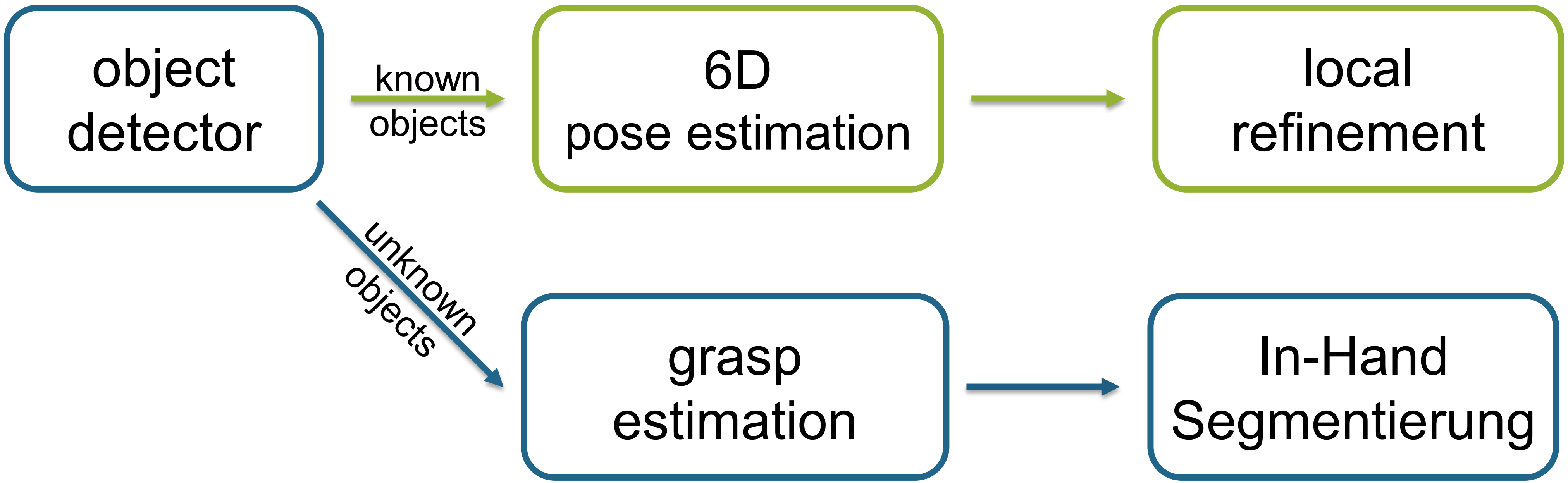




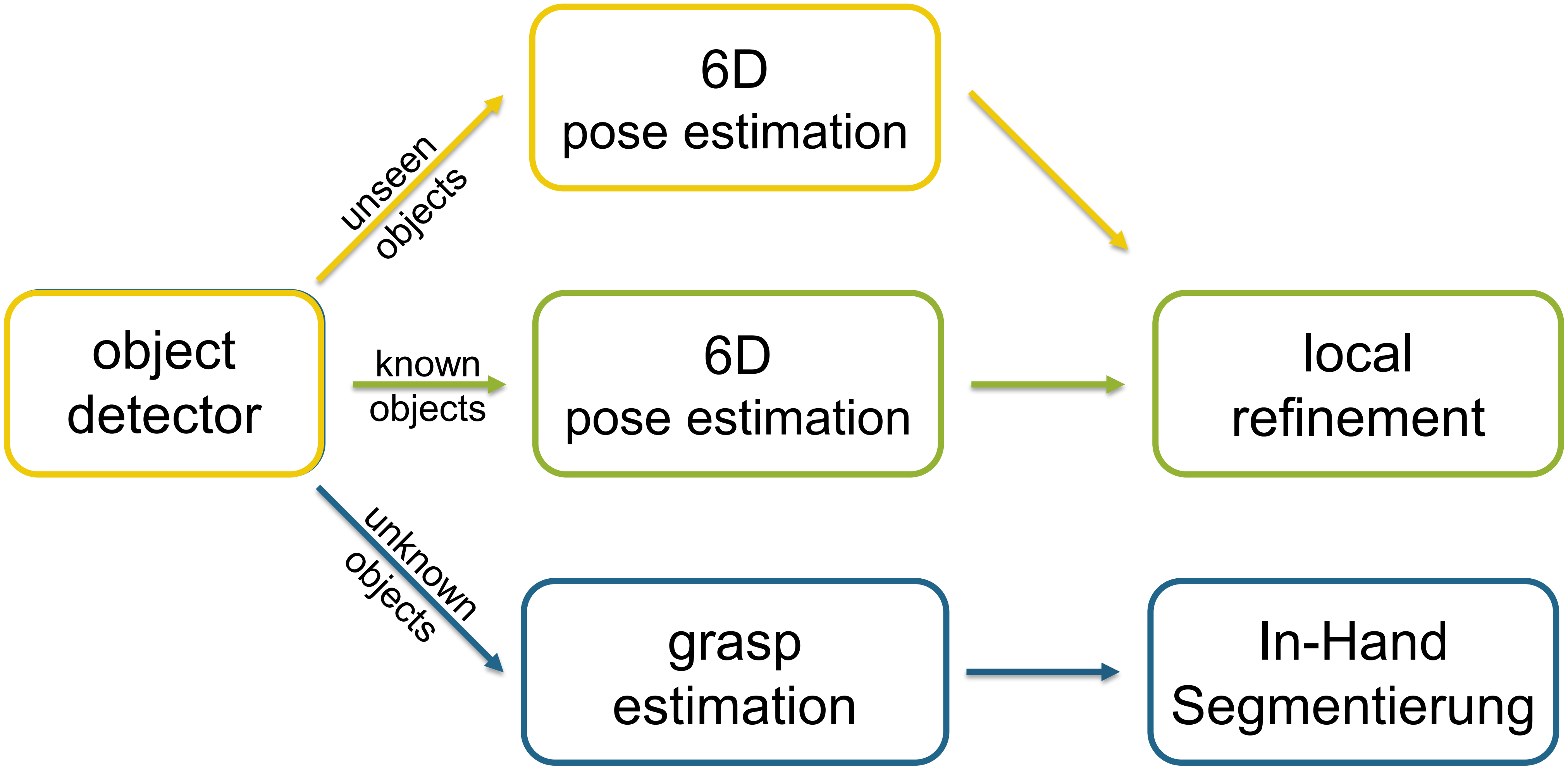
[Boerdijk et al., CORL 2020]

[Boerdijk et al., ICRA 2021]

Perception system



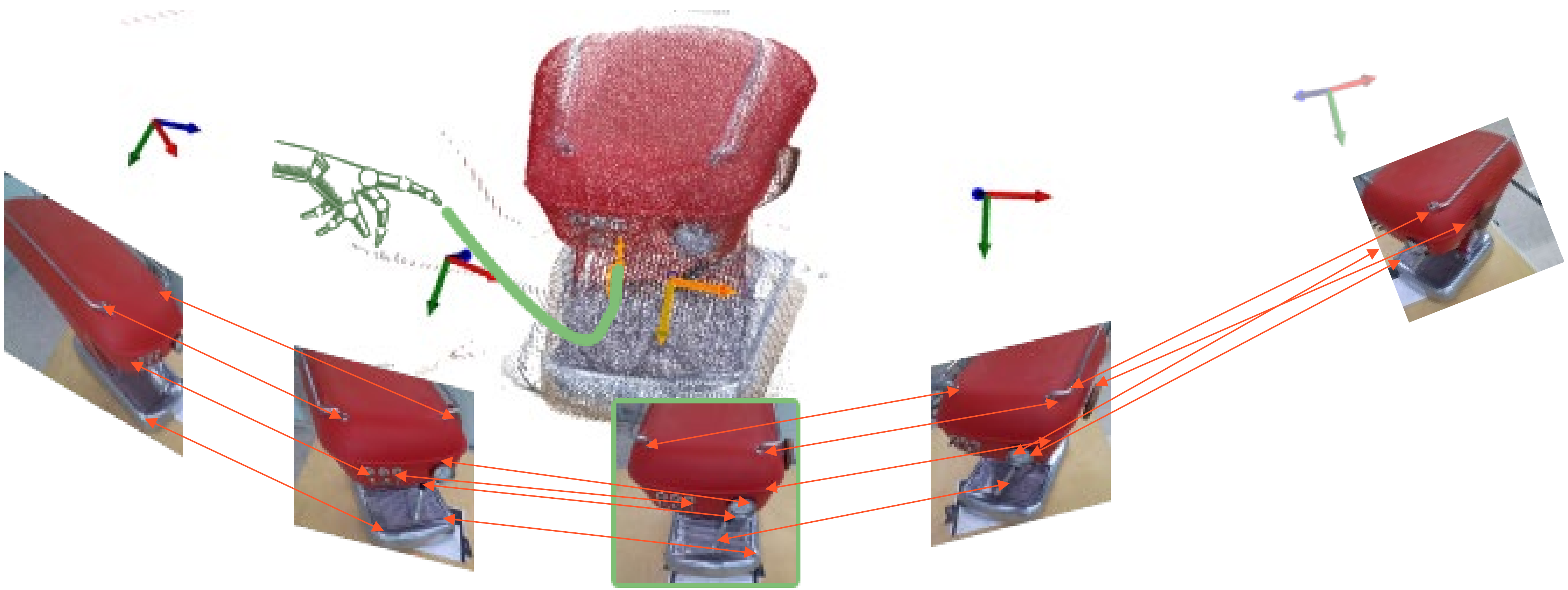
Perception system



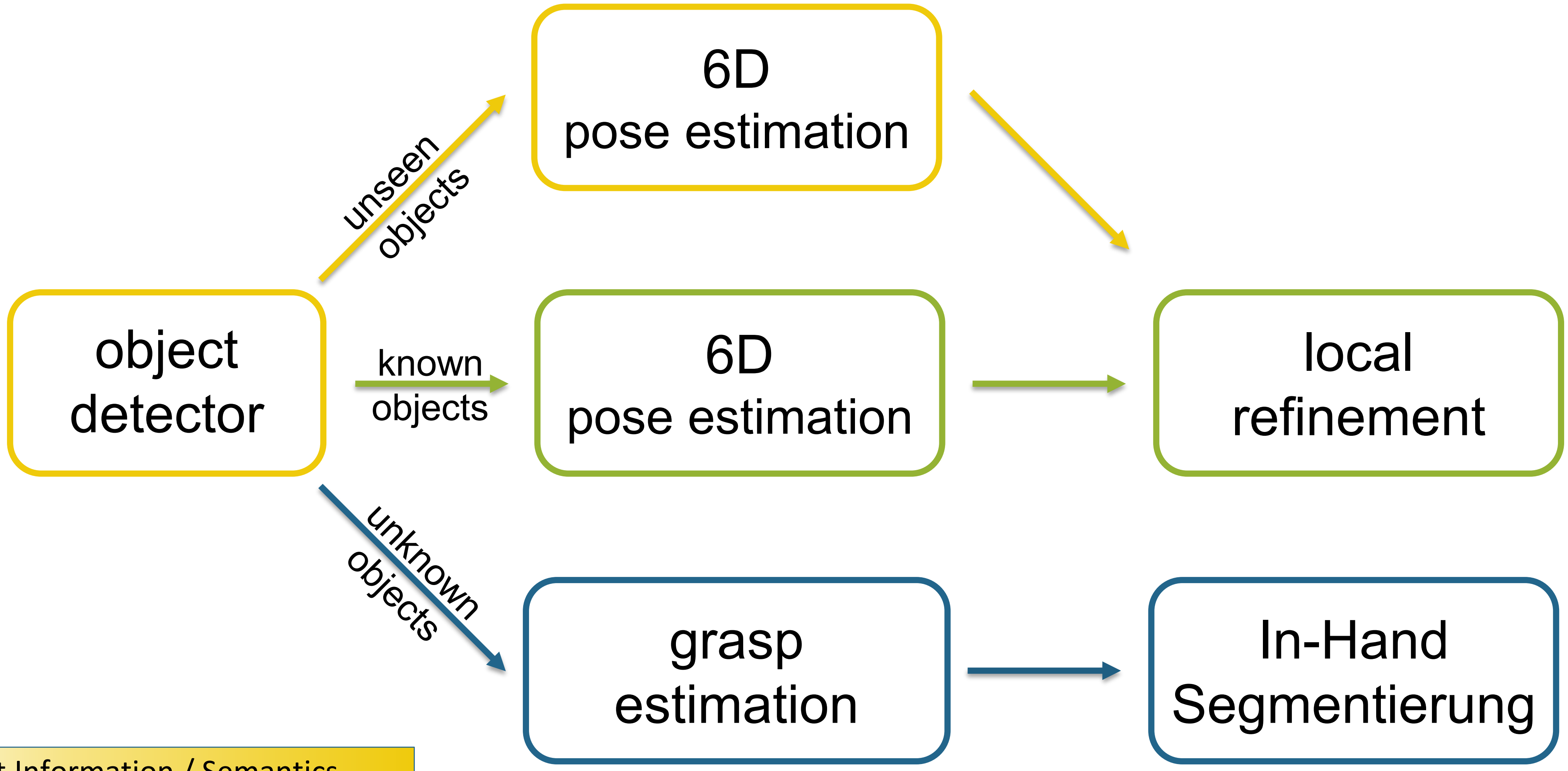
Unseen Object Perception

Set-Up

Inference



Perception system



Object Information / Semantics

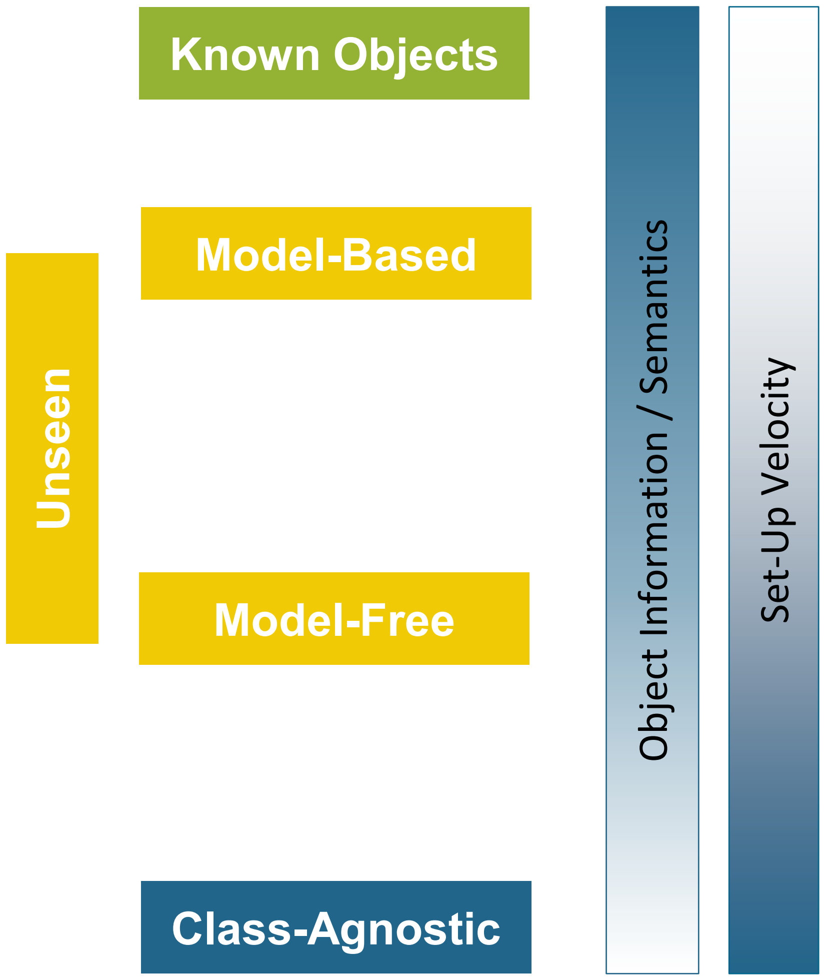
Set-Up Velocity

Take Aways

- Synthetic Data
- Large advances
- Task-dependent: Precision vs. Generalizability

Outlook

- Broader Objects Groups: articulated, deformable
- Category- / Semantic-Level
- Include additional
  - knowledge
  - modalities



THANK YOU FOR YOUR ATTENTION!

Referenzen

- **M Durner**, S Kriegel, S Riedel, M Brucker, ZC Marton, F Balint-Benczédi, R Triebel. “Experience-based optimization of robotic perception”, *ICAR 2017 (**Best Paper Finalist**)*
- M Sundermeyer, ZC Marton, **M Durner**, M Brucker, R Triebel. “Implicit 3D orientation learning for 6D object detection from rgb images”, *ECCV 2018 (**Best Paper Award**)*
- J Feng*, **M Durner***, ZC Márton, F Bálint-Benczédi, R Triebel. “Introspective robot perception using smoothed predictions from bayesian neural networks”, *ISRR 2019*
- M Brucker, **M Durner**, ZC Marton, F Balint-Benczédi, M Sundermeyer, R Triebel. “6dof pose estimation for industrial manipulation based on synthetic data", *ISER 2019*
- J Lee, M Humt, J Feng, R Triebel. “Estimating model uncertainty of neural networks in sparse information form”, *ICML 2020*
- M Sundermeyer, ZC Marton, **M Durner**, R Triebel. “Augmented Autoencoders: Implicit 3D orientation learning for 6D object detection”, *IJCV 2020*
- M Sundermeyer, **M Durner**, E Puang, ZC Marton, R Triebel. “Multi-path Learning for Object Pose Estimation Across Domains”, *CVPR 2020*
- W Boerdijk, M Sundermeyer, **M Durner**, R Triebel. “Self-Supervised Object-in-Gripper Segmentation from Robotic Motions”, *CORL 2020*
- M Stoiber, M Pfanne, KH Strobel, R Triebel, A Albu-Schäfer. “A sparse gaussian approach to region-based 6dof object tracking”, *ACCV 2020 (Best Paper Award)*
- J Lee, J Feng, M Humt, M G Müller, R Triebel. “Trust your robots! Predictive Uncertainty Estimation of Neural Networks with Sparse Gaussian Processes”, *CORL 2021*
- M Sundermeyer, A Mousavian, R Triebel, D Fox, “*Contact-GraspNet: Efficient 6-DoF Grasp Generation in Cluttered Scenes”*, *ICRA 2021*
- **M Durner***, W Boerdijk*, M Sundermeyer, W Friedl, ZC Marton, R Triebel. “Unknown Object Segmentation from Stereo Images”, *IROS 2021*
- W Boerdijk, M Sundermeyer, **M Durner**, R Triebel. “What’s This? - Learning to Segment Unknown Objects from Manipulation Sequences”, *ICRA 2021*
- J Feng, J Lee**, M Durner,** R Triebel. “Bayesian active learning for sim-to-real robotic perception“, IROS *2022*
- **M Durner***, W Boerdijk*, Y Fanger, R Sakagami, D Risch, R Triebel, A Wedler. “Autonomous Rock Instance Segmentation for Extra-Terrestrial Robotic Missions”, *Aero. Conf. 2023*
- M Ulmer, **M Durner**, M Sundermeyer, M Stoiber, R Triebel. “6D Object Pose Estimation from Approximate 3D Models for Orbital Robotics”, *IROS 2023 (**SPEED+ Post-Mortem Leader**)*

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ERF2025
STUTTGART | GERMANY
25–27 MARCH

Perceiving Deformable Linear Objects in Real-World Scenarios

Alessio Caporali

WS#58
Perceiving unknown and deformable objects in logistics and service robotics

COESIONE ITALIA 21-27 EMILIA-ROMAGNA
Co-funded by the European Union
Regione Emilia-Romagna

About Me…

📌 **Current Position**

🎓 *Junior Assistant Professor*, University of Bologna (Italy)

📌 **Education & Research**

🎓 *PhD Defense (04/2024)* – Thesis: *"Robotic Perception and Manipulation of Deformable Linear Objects"*

🎓 *BS & MS in Automation Engineering*

🔬 Research on **Deformable Objects** since **2020**

ROBOSECT

ROBOSECT is a spin-off project of the University of Bologna.

📌 ***Co-founder** of ROBOSECT SRL*

📌 *Technology-driven company focused on robotic solutions for the perception and manipulation of electrical cables*

What are Deformable Linear Objects?

Abbreviated as DLOs

- Flexible objects with an elongated shape
- Length significantly larger than diameter
- Examples: cables, wires, ropes, tubes…

Presence in Various Environments (domestic, industrial assembly…)

Perception Challenges

- Ambiguous and complex appearance
- Lack of distinct features
- Small size
- Deformability

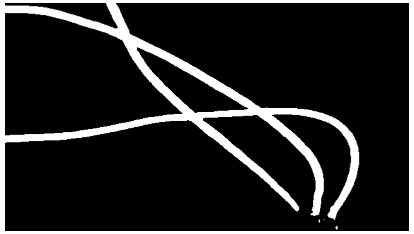
Main Perception “Components”…

Dataset Generation → 2D Perception → 3D Perception

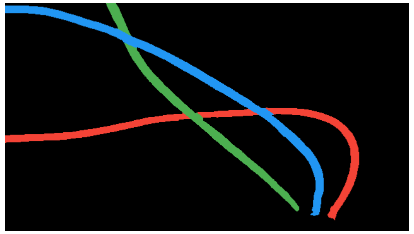
Why Segmentation?

Segmentation is the only vision-based task able to properly characterize the DLO shape.
Example: centerline from mask

Input Image



Semantic Seg.



Instance Seg.

Dataset Generation

- DLOs can feature a wide variety of shapes and colors.
- Lack of distinctive feature to exploits with standard CV approaches

solution →

Deep learning methods that can generalize well on the challenging class of DLOs.
Quality and size of the dataset is crucial!

but

Dataset Collection Issues

📌 **Manual Annotation is Slow & Tedious**

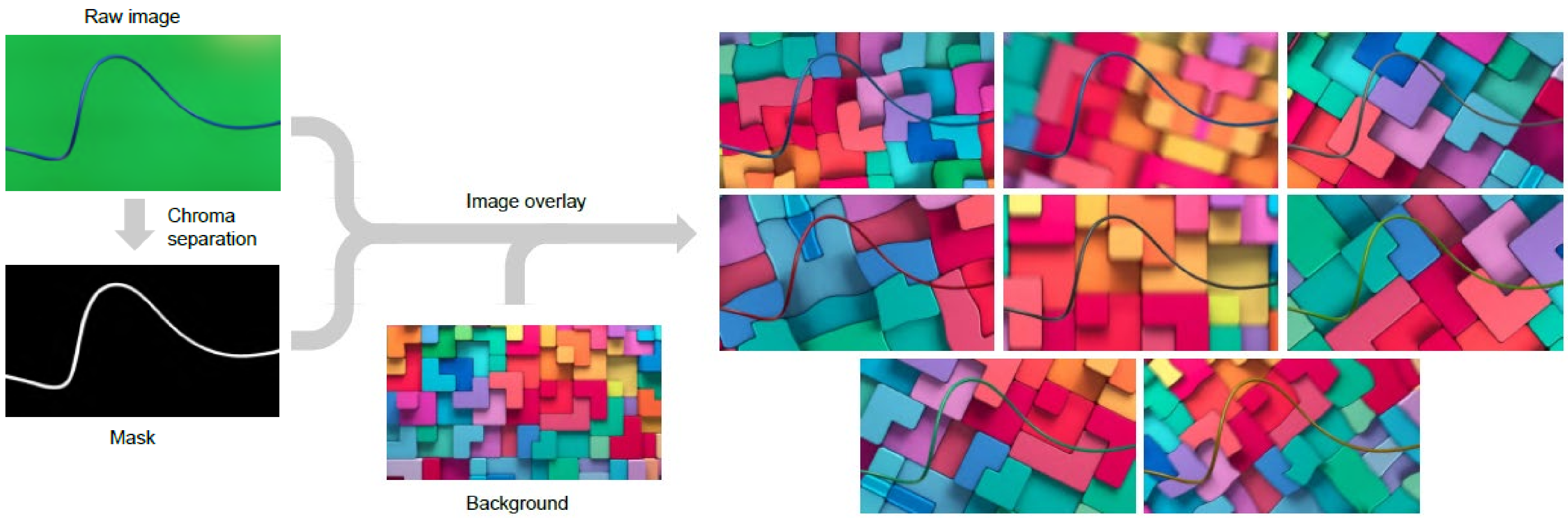
📌 Goal: **Automate Large-Scale Dataset Generation** with minimal human effort

🔍 **Proposed Methods**

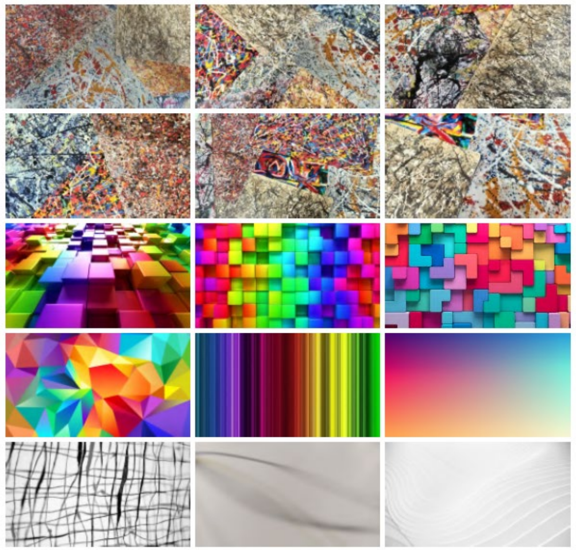
1. **Chroma Separation** – Leveraging color differences for mask separation
2. **Synthetic Image Rendering** – Generating diverse, annotated images in a simulation environment

[Alessio Caporali, et al. IEEE RA-L (2023)]

Chroma Separation



Real **foreground** image + (synthetic) **background** augmentation



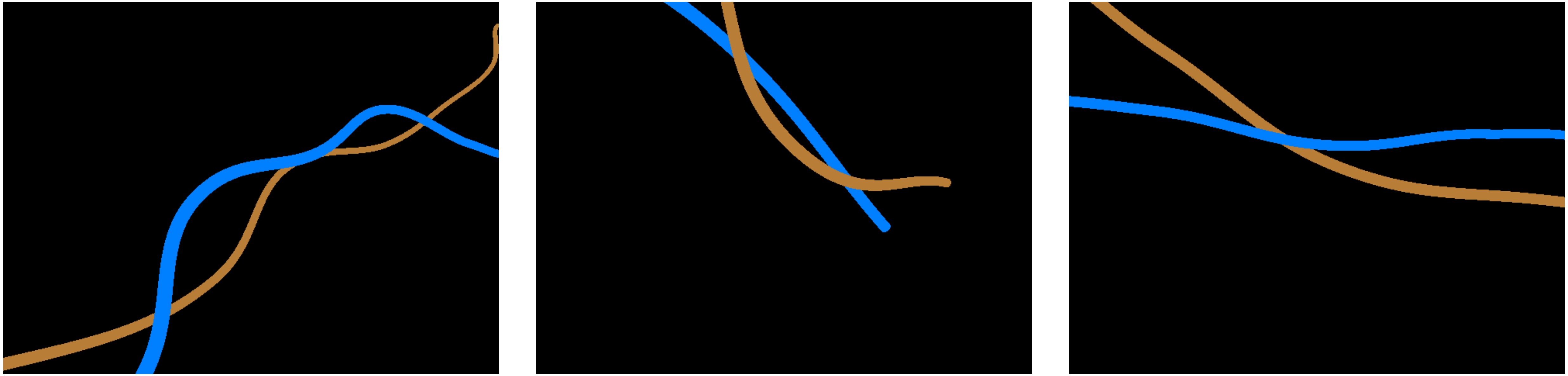
complex backgrounds

Human input still required!

[Riccardo Zanella, Alessio Caporali, et al. IEEE ICCCR (2021)]

Synthetic Images Rendering

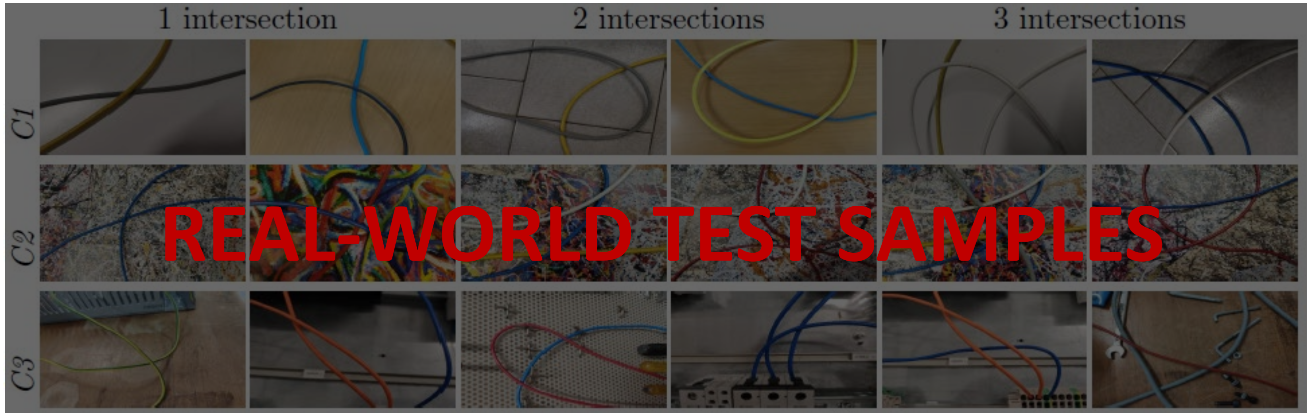
No need for manual image collection and labeling
Fully automated dataset generation

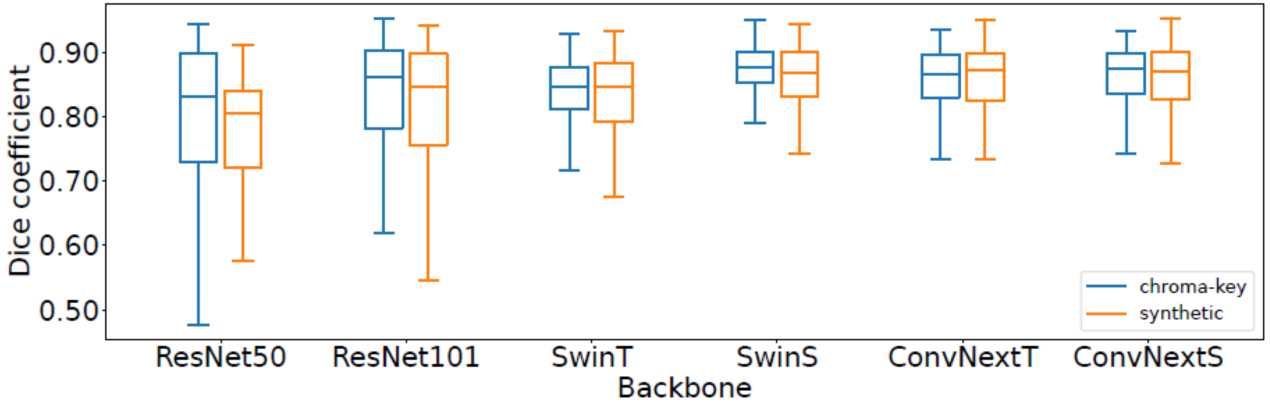


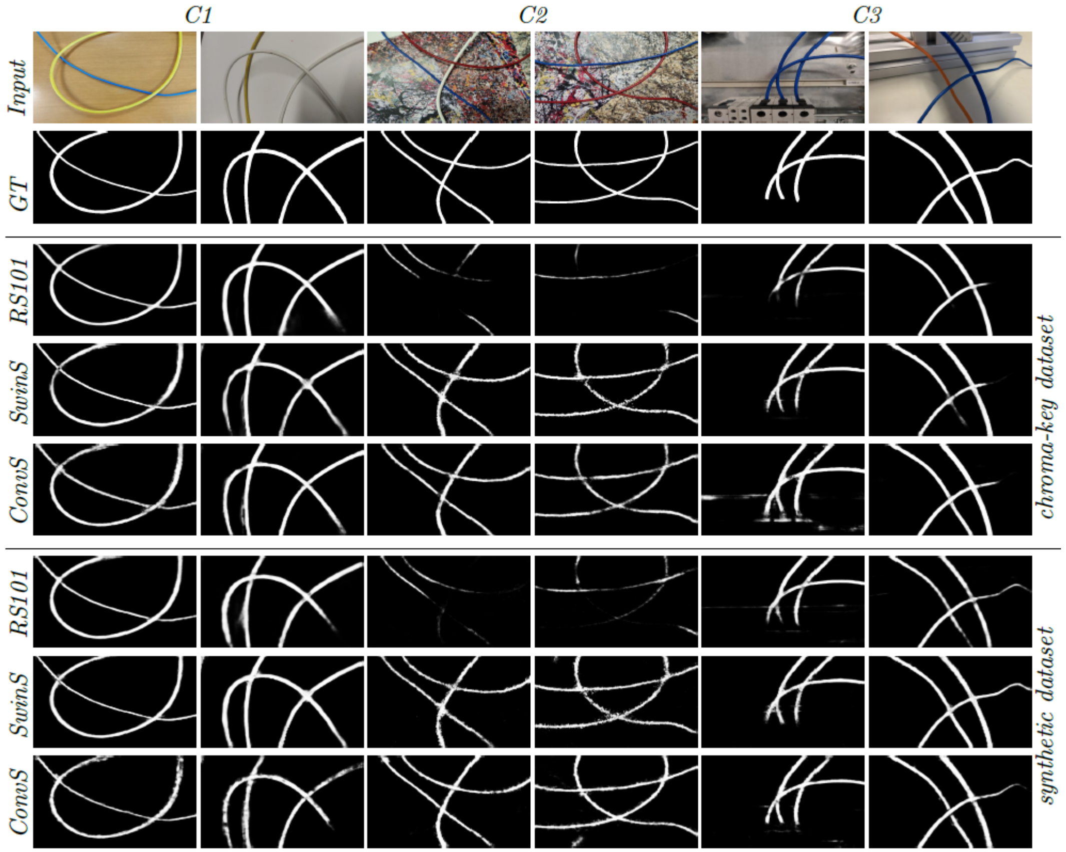
[Alessio Caporali, et al. IEEE RA-L (2023)]

Dataset Generation

Chroma-Separation vs Synthetic Datasets for Semantic Segmentation



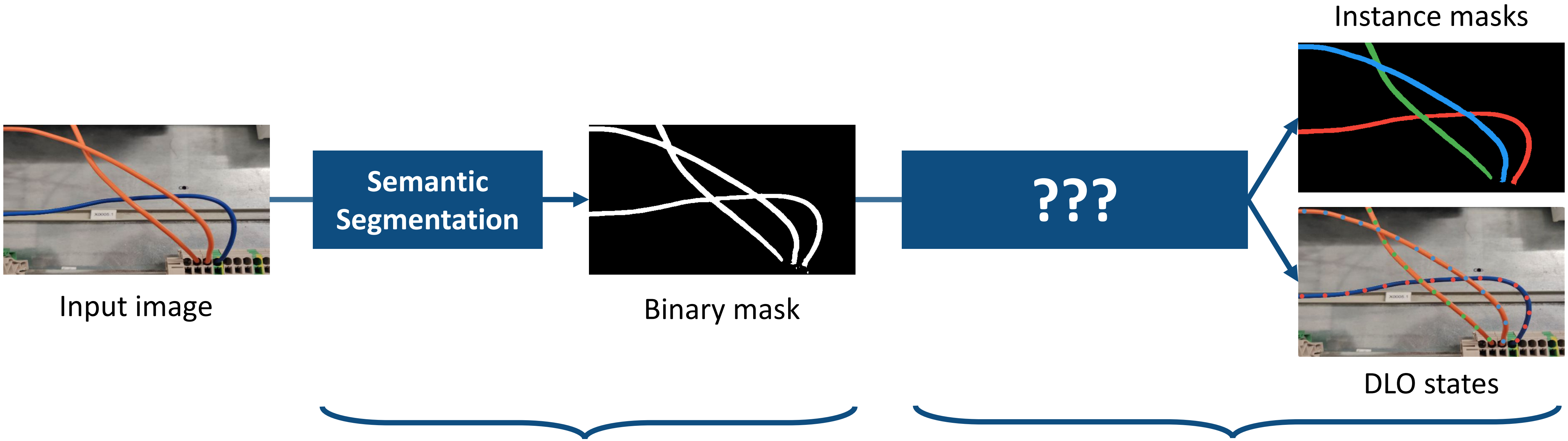




Synthetic Dataset viable approach!

2D Perception

General Structure (Ariadne+[1], FASTDLO[2], RT-DLO[3])

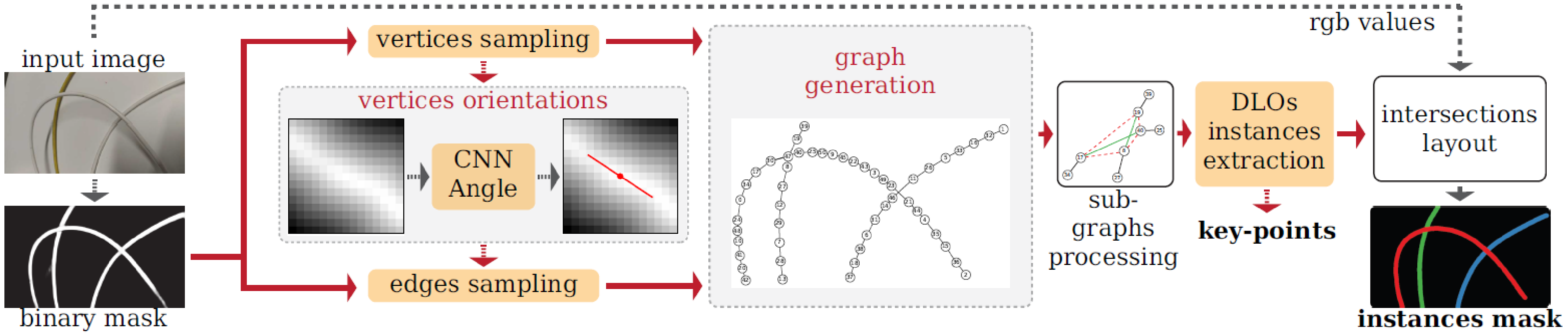


Previous Slides!

HOW?
Graph-based approach

1. [Alessio Caporali, et al. IEEE TII (2022)]
2. [Alessio Caporali, et al. IEEE RA-L (2022)]
3. [Alessio Caporali, et al. IEEE TII (2023)]

2D Perception with RT-DLO



- **Semantic Segmentation:** synthetic images dataset
- **Main Processing: fully graph-based representation**
- **Intersection Solving**: cosine similarity
- **Intersection layouts:** std color values

Real-Time Instance Segmentation of Deformable Linear Objects

[Alessio Caporali, et al. IEEE TII (2023)]

2D Perception with RT-DLO

RT-DLO: Real-Time Deformable Linear Objects
Instance Segmentation

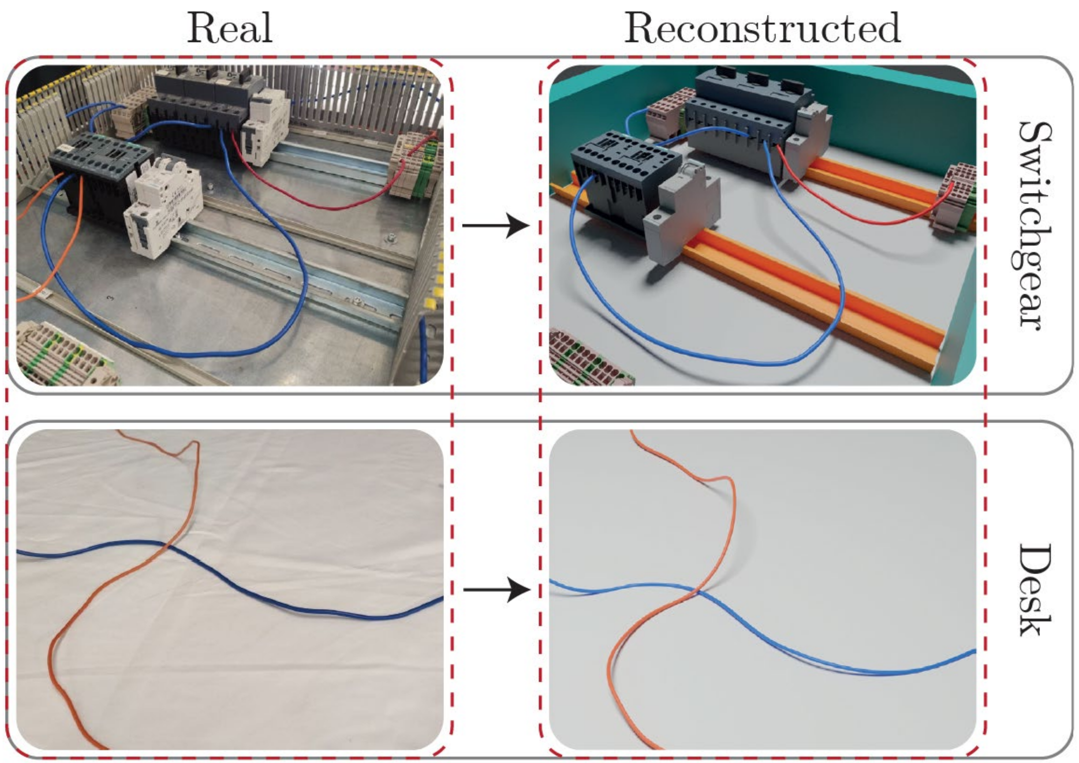
3D Perception of DLOs

Challenges in 3D Perception

- **Small DLOs (<10-15mm)** are **invisible** in common depth sensors (e.g., Intel RealSense, Kinect Azure)
- High-end cameras are **expensive, bulky, and slow**, making robotic integration impractical

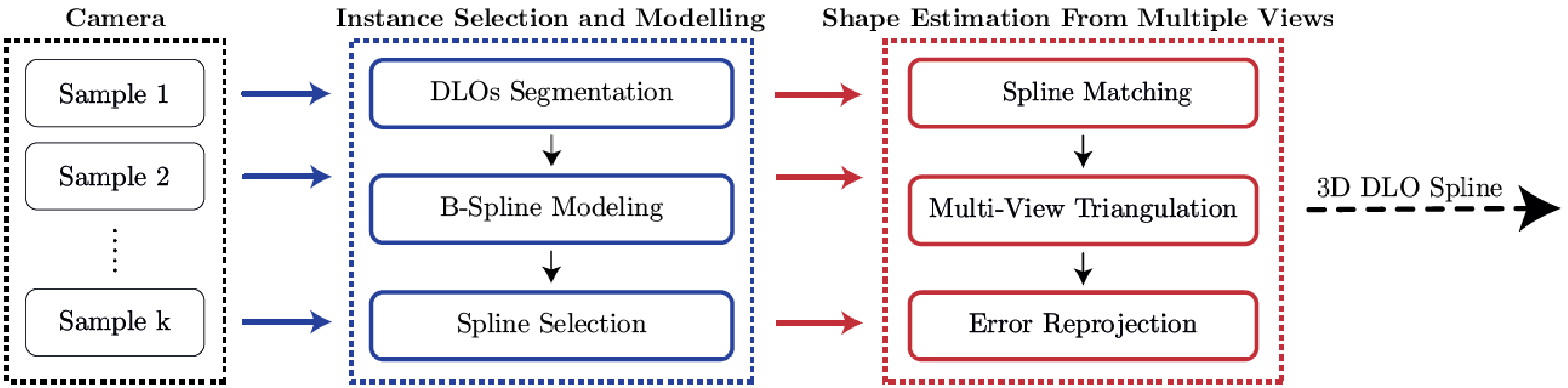
Proposed Solution: DLO3DS

- Uses a **simple 2D camera** on the robot end-effector
- **2D Spline-Based Representation** from processed images
- **Multi-view triangulation** reconstructs 3D information



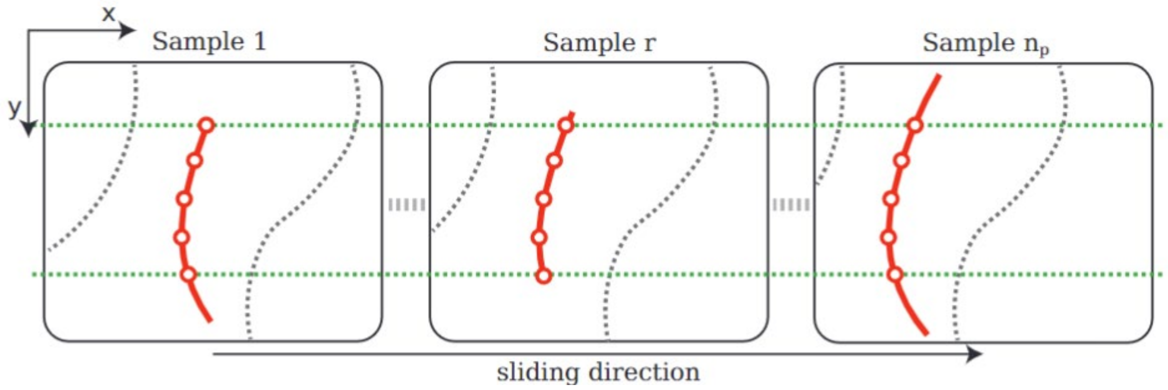
[Alessio Caporali, et al. IEEE RA-L (2023)]

3D Perception of DLOs



2D Perception tools!

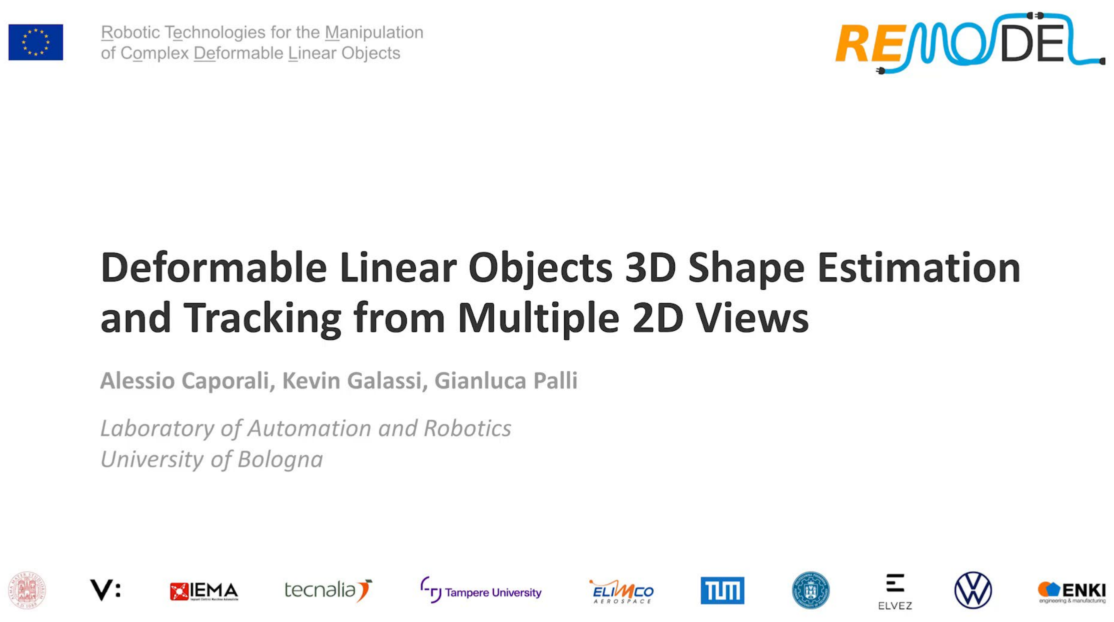
Possibility of optimizing baseline and distance from target online!



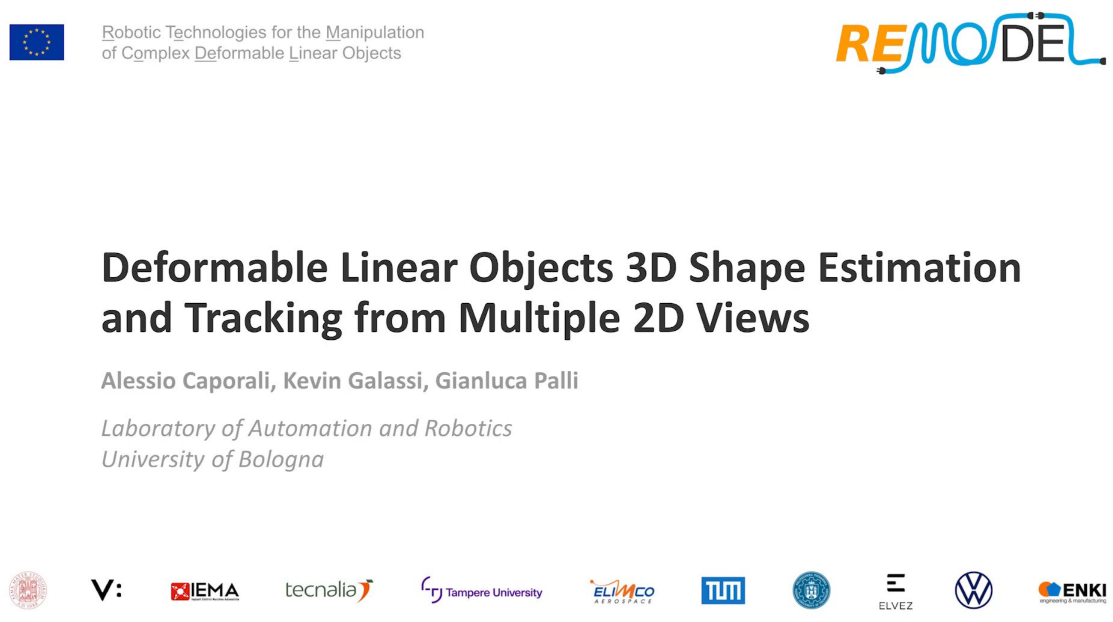
[Alessio Caporali, et al. IEEE RA-L (2023)]

3D Perception of DLOs

Shape Reconstruction (3 views)



Testing



[Alessio Caporali, et al. IEEE RA-L (2023)]

Conclusions

Perception of DLOs…

📌 **Synthetic Data for Deep Learning** – Efficient for dataset creation [2]

📌 **Semantic Segmentation** – Simplifies DLO perception [1,2]

📌 **Graph-Based Representation** – Enables **instance segmentation** [3,4,5]

📌 **3D Reconstruction** – Achieved via **instance segmentation + multi-view triangulation** [6]

Key References

[1] Zanella Riccardo, et al. "***Auto-generated Wires Dataset for Semantic Segmentation with Domain-independence***." 2021 IEEE International Conference on Computer, Control and Robotics.

[2] Caporali Alessio, et al. "***A Weakly Supervised Semi-automatic Image Labeling Approach for Deformable Linear Objects***." IEEE Robotics and Automation Letters (2023).

[3] Caporali Alessio, et al. "***Ariadne+: Deep Learning-Based Augmented Framework for the Instance Segmentation of Wires***." IEEE Transactions on Industrial Informatics (2022).

[4] Caporali Alessio, et al. "***FASTDLO: Fast Deformable Linear Objects Instance Segmentation.***" IEEE Robotics and Automation Letters (2022).

[5] Caporali Alessio, et al. "***RT-DLO: Real-Time Deformable Linear Objects Instance Segmentation***." IEEE Transactions on Industrial Informatics (2023).

[6] Caporali Alessio, et al. "***Deformable Linear Objects 3D Shape Estimation and Tracking From Multiple 2D Views***." IEEE Robotics and Automation Letters (2023).

Thank You for Your Attention!

Contacts

Alessio Caporali
alessio.caporali@unibo.it
www.alessiocaporali.github.io

ROBOSECT SRL
info@robosect.it
www.robosect.it

COESIONE ITALIA 21-27 EMILIA-ROMAGNA

Co-funded by the European Union

Regione Emilia-Romagna

ALESSIO CAPORALI IS FUNDED BY EUROPEAN FUNDS FROM THE EMILIA-ROMAGNA REGION - CUP J19J23000730002

AI-based Perception of Seen and Unseen flexible Objects

Dr. Michael Suppa

Many complex automation tasks rely on high-precision 3D robot vision to perform efficiently

Robots Need 3D Vision

WHY IMPLEMENT ROBOT VISION?

Today's robotic systems cannot handle unorganized items.

Pre-organization and standardization need (down)time, space, money.

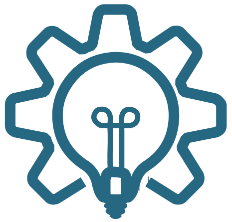
Many existing processes are labor-intensive.

Some tasks are impossible to automate.

Challenge: Automation of Complex Tasks

REQUIRES CONSIDERATION OF VARIATIONS

Environment



Lighting conditions

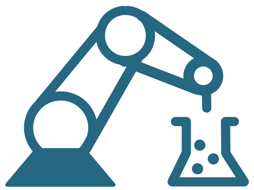
Perspective

Objects

Material varies from shiny and transparent to translucent and black

Different sizes and distances

Application



Parameterization requires expertise

Test & implementation time

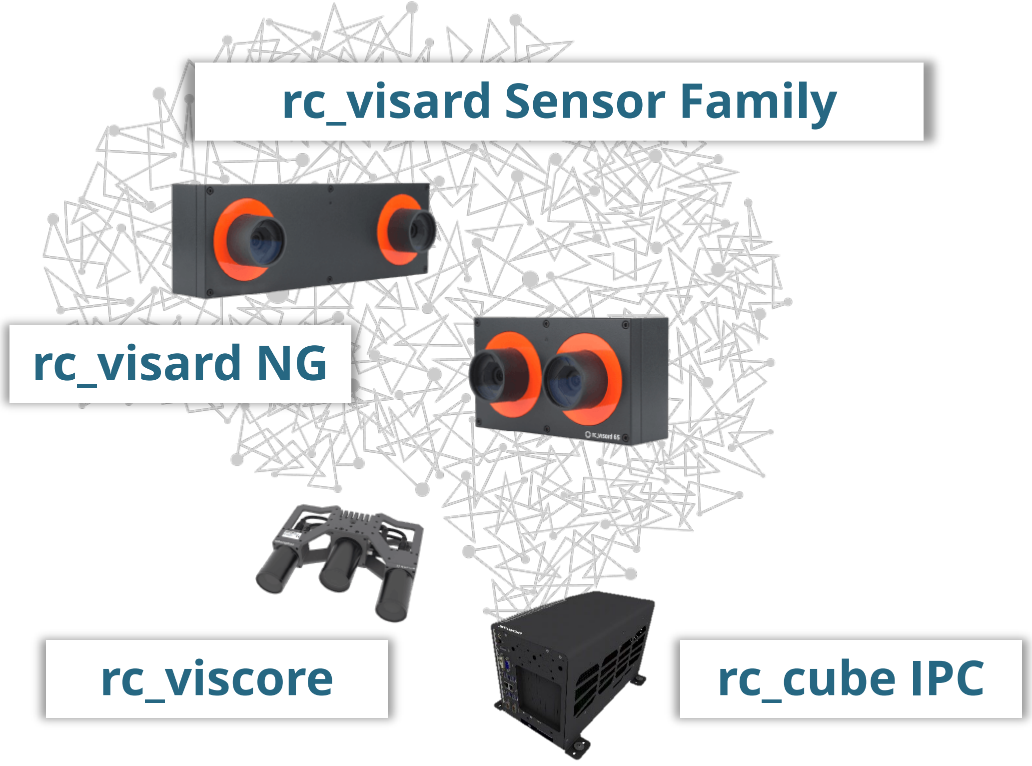
>97% availability (successful picks)

Data must cover all variations: Hard (impossible) to achieve with real data/ human input.

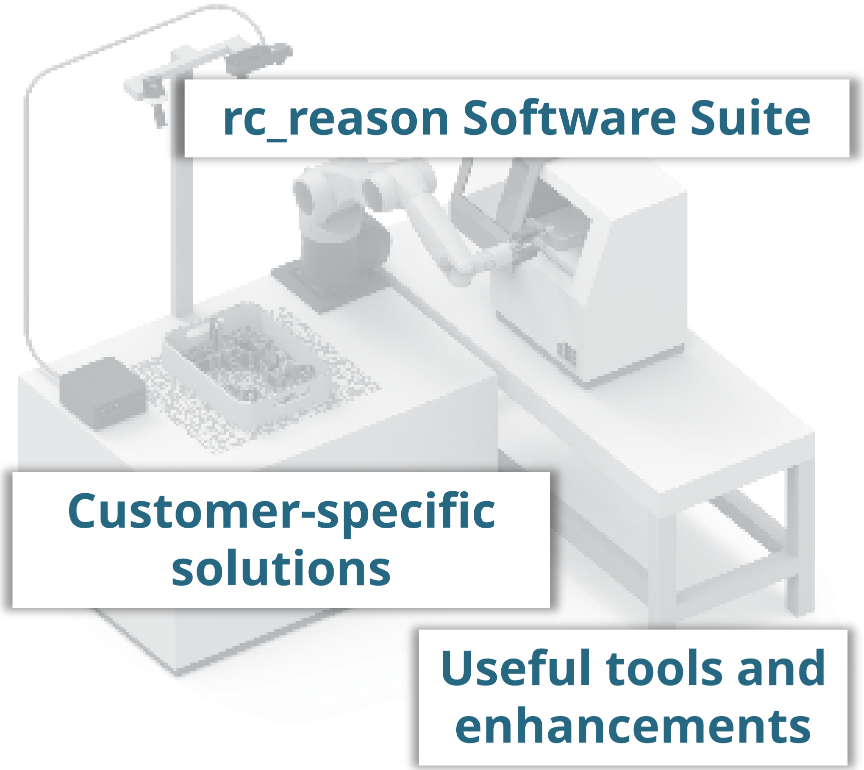
Product Portfolio

VERSATILE SENSORS AND INTUITIVE ROBOTICS SOFTWARE

3D Stereo Vision for Your Robot

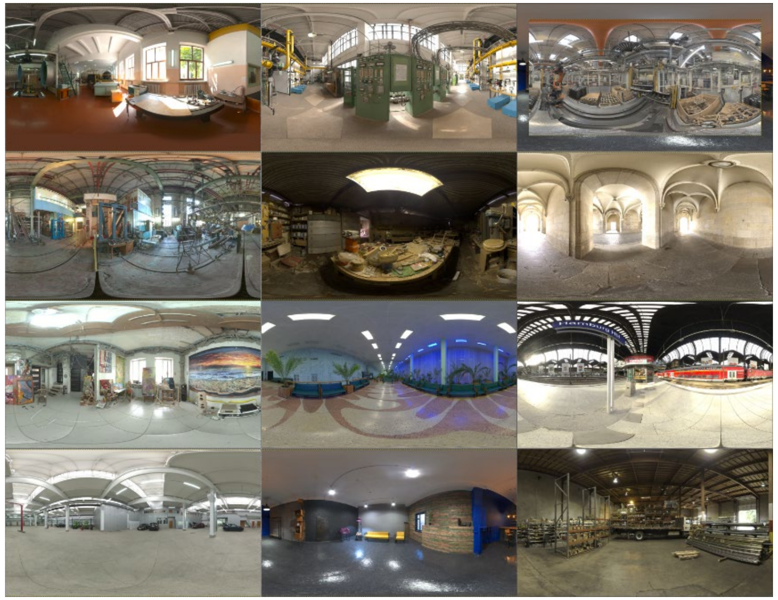


Advanced AppliedAI Software

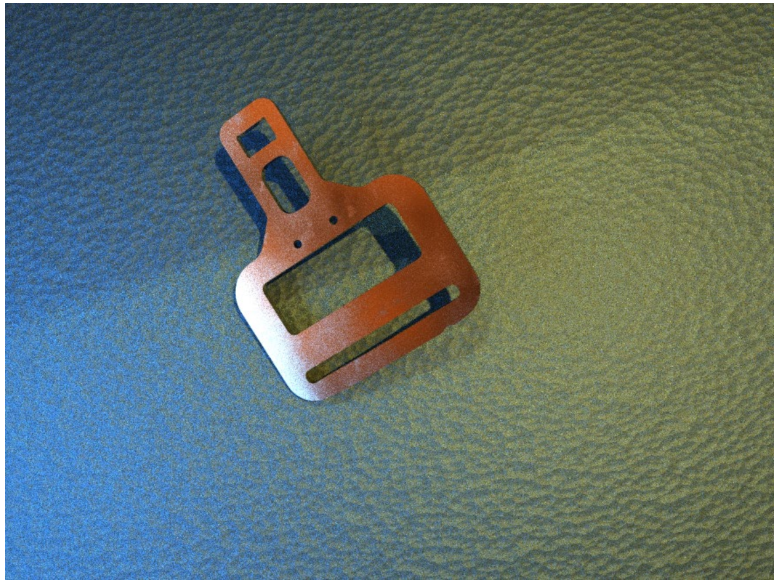


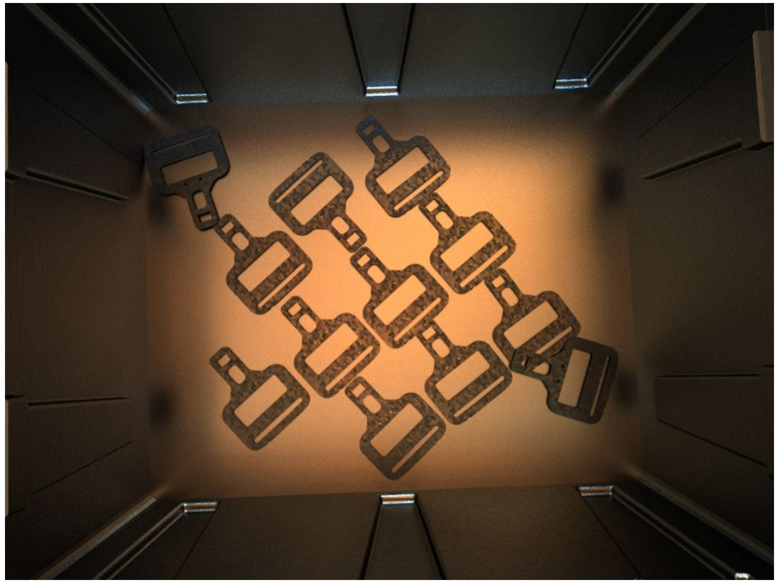
AI-Based Detection of Known Items

TRAINING WITH SYNTHETIC DATA

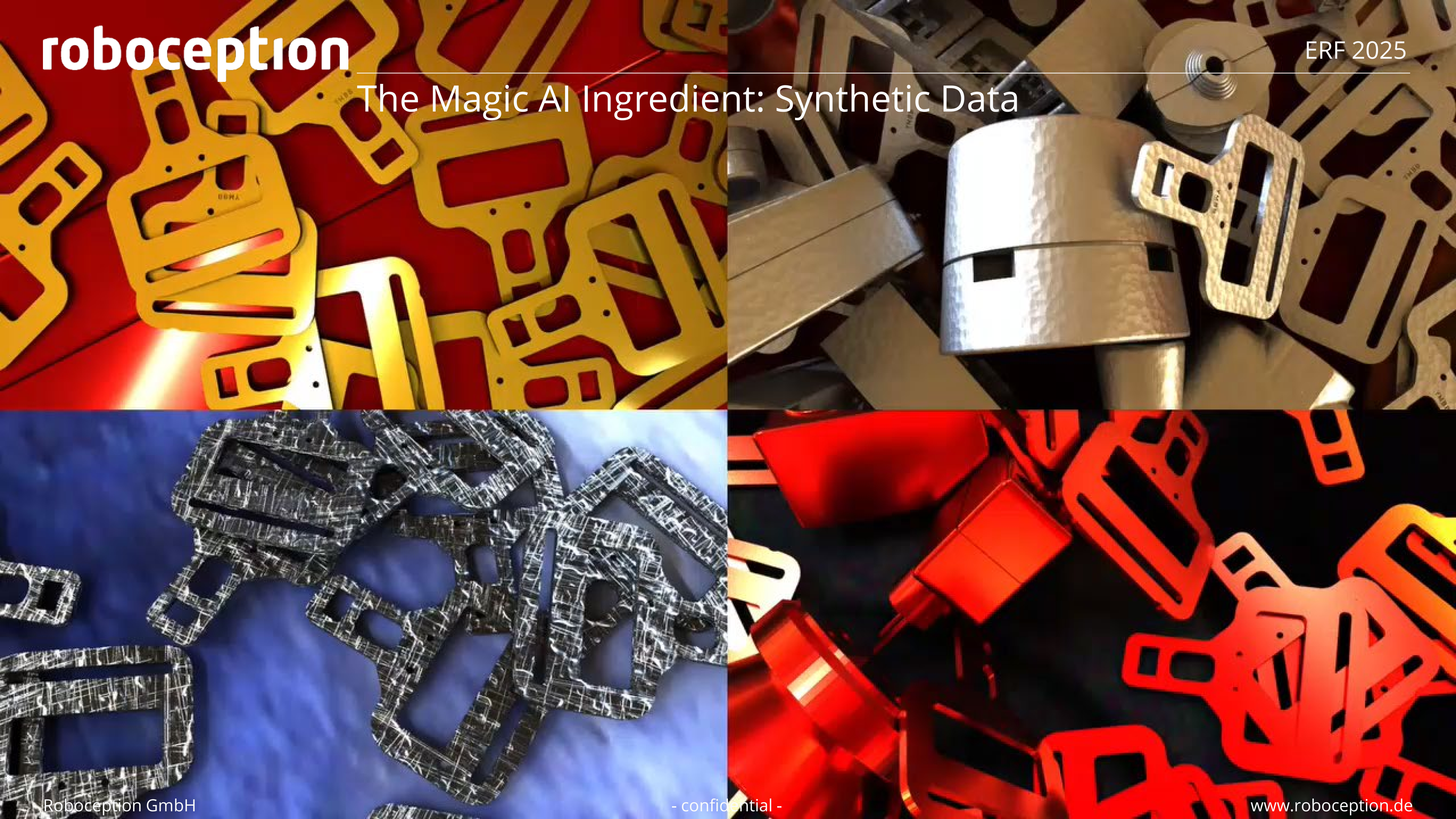


- Training data from a photorealistic simulation environment
- Material library, lighting simulation, CAD-models
- Support for various applications and parts
- Automated creation of templates as a cloud service





The Magic AI Ingredient: Synthetic Data



Robot Interfaces

Standard Interface:
- REST-API is a generic modern interface
- Most industrial robot interfaces are very proprietary
- Robot controllers/PLCs have reduced computing performance

EXPERT ROBOT PROGRAMMERS

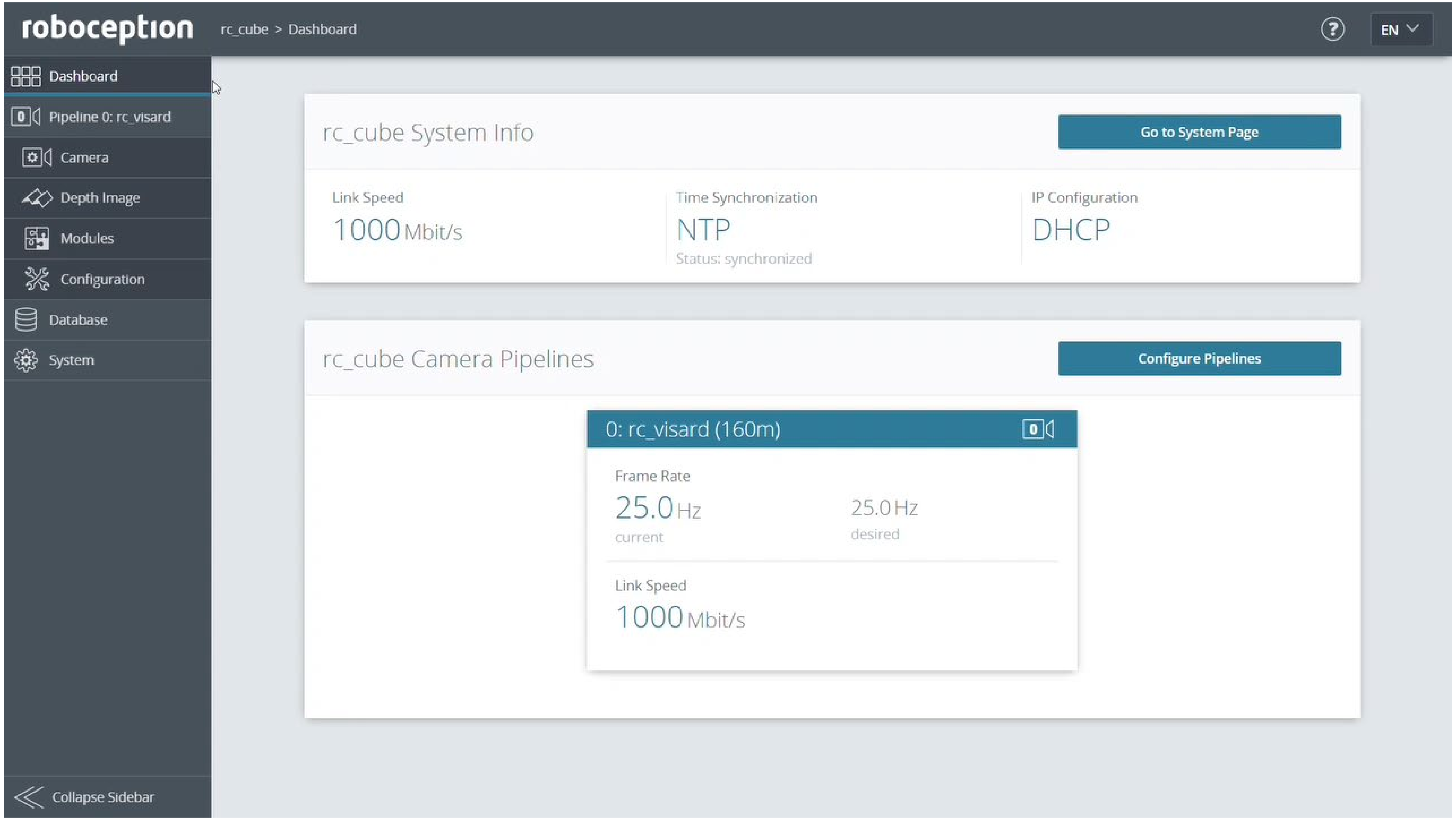
HIGH-LEVEL USERS

Perception skills for various platforms

ctrlX AUTOMATION   intrinsic   vorausrobotik

Ease of Use for Non-Vision Expert

INTUITIVE WEB INTERFACE ENABLES NON-EXPERT USE



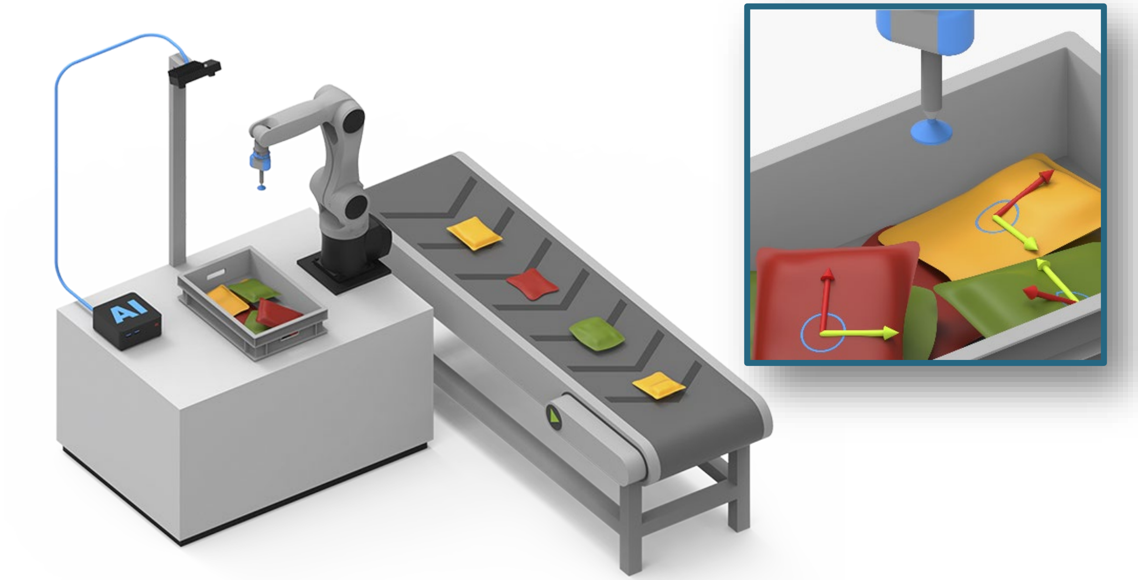
Three Major Trends in Robotics

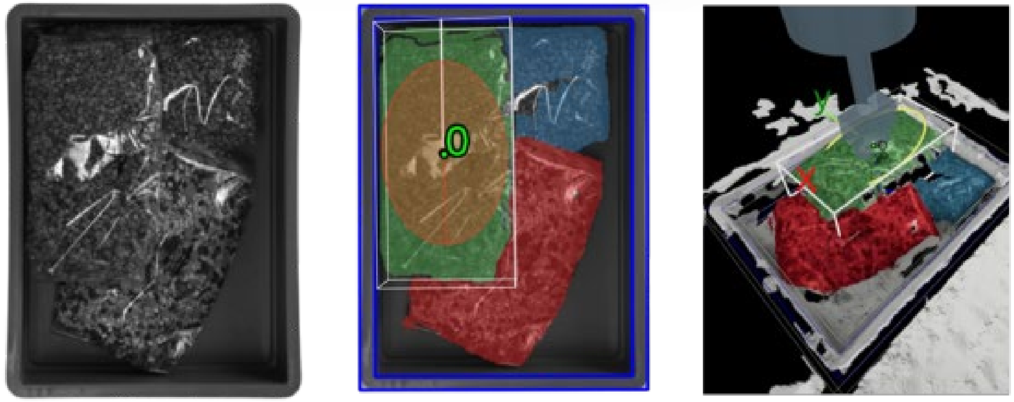
#1 GOOD DATA INSTEAD OF BIG DATA	• Generation of detection templates based on CAD data • Simulations create realistic training data using model-knowledge	IMPLEMENTATION WITH MINIMUM EFFORT
#2 PLUG-AND-PRODUCE	• Integrators and end users can add modules on the same platform • Smart sensors enable distribution of computing resources	SCALABLE MACHINE LEARNING PLATFORM
#3 EASE-OF-USE	• AI reduces parameterization effort for the user • Web interfaces with wizards enable non-expert use	EASE-OF-USE FOR VISION-NEWBIES

rc_reason ItemPickAI

FOR ROBOTIC PICK-AND-PLACE APPLICATIONS USING SUCTION GRIPPERS

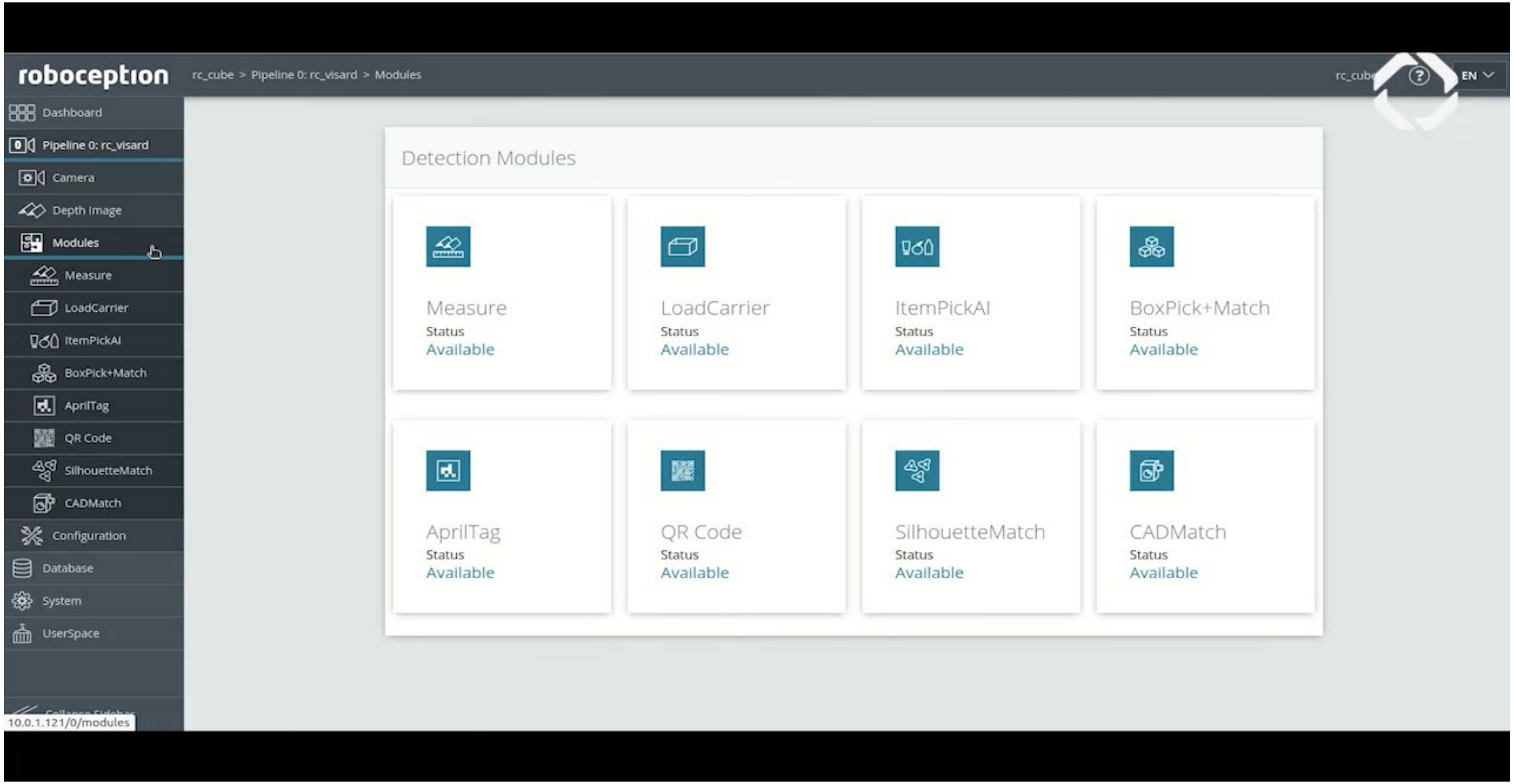
- Segmentation of deformable objects in mixed and unmixed scenes
- Delivers oriented grasp poses for an oriented placement
- Object category 'bag' with various volumes and fill levels





rc_reason ItemPickAI

FOR ROBOTIC PICK-AND-PLACE APPLICATIONS USING SUCTION GRIPPERS



Contact

Dr. Michael Suppa

CEO & Co-Founder

phone: +49 89 8895079-11
cell: +49 172 4195266
email: michael.suppa@roboception.de

Round Table Discussion

roboceptıon

How are you dealing with flexible and deformable objects in your use cases?
Please describe use cases such that we can foster the discussion!

18 Participants

- Soft grippers
- Highly cluttered environments with deformable object, fruit and vegetables, also different shapes. Pick and pack densely without damaging other objects
- What about combining force/tactile sensing with vision? For improving handling
- Branched DLO/ wire harness
- With a success rate of 97% at a rate of 600 picks per hour, the robot would fail 18 times in 1 hour of work....how do you handle these failures?
- Do you know approaches that Consider the Center of mass e. g to minimize swinging of flexible bags
- Any interest in using 3d mechanical models (mass point masses with springs... or more complex) to assess the state of the deformable object?
- How well does the DLO perception models react to different types of DLO e.g. different thickness of cables?
- Waste manipulation
- In handling of DLIs typically specialized machines are used, how do you see robots can be EFFECITEVELY applied
- stretchy and sticky deformable objects
- Are those models about 6d pose estimation suitable for running conveyors?
- Removal/pulling of cables during dismantling in recycling
- Big range of model-free objects
- How reproducible are results with deformable objects? What parameters are most problematic?
- Food manipulation
- Cables manipulation
- Textile manipulation
- Cabling of electronics
- Clothes manipulation

Data sets and representation of flexible and deformable objects are difficult. How define a grasp point on a deformable object?

16 Participants

- Requires realistic simulation of deformable objects to predict the objects future state and enable dataset generation. Do you know / work on any 'general' deformable dataset generation tools?
- grasping points are tightly linked to the expected deformation process, predicting the deformation is essential for this problem, how can we have accurate predictions of such processes?
- The deformation of a bag highly depend on the content. What about using such info?
- Identify stable geometric features or topology-preserving regions that remain consistent despite deformation, combined with real-time force feedback during execution.
- Why not use the robot to collect data or understand properties of deformable objects by planned disturbances. This seem to be not mostly used - any disadvantages?
- Multiple? To manage cable while plugging something in would help a lot
- Requires a combination of geometry, accurate prediction of how the object deforms (mechanical properties), and the interaction if the gripper and the object
- We need to be able to predict the deformation and compare it with the task objective to optimize the grasping point and manipulation maneuver
- Based on the post grasping task
- how stable is the grasp point for pick and place?
- Any experience in scaling the synthetic data from vision only to synthetic vision/action datasets for joint training?
- Depends on gripper capabilities
- Have you encountered problems where the object deformation modifies the grasp point too much , which were not in training datasets?
- Estimation of physics model based on perception then grasp estimation based on physics and gravitybestimation
- How applicable are grasp planning solutions trained on rigid objects to deformable objects?
- Rigid features

roboception

Synthetic training data generation helps to close the training data and ground truth gap. Do you have data for synthetic data generation of flexible objects available? Where does it come from and how do you assess the level of correctness?

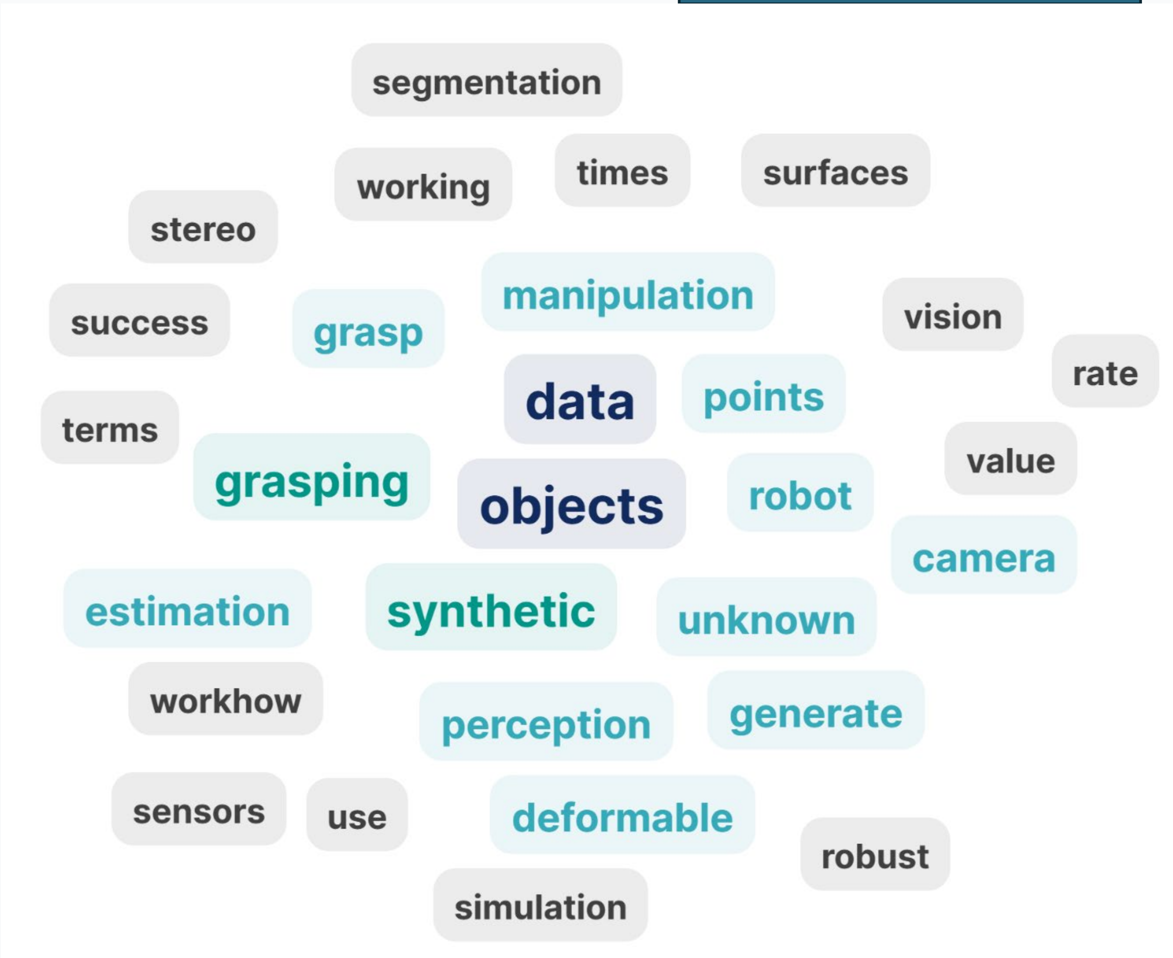
8 Participants

- How would you generate a representative dataset of naturally distinct objects? (Natural objects, anatomical parts, defects…)
- What incentives could the community create to accelerate open dataset generation?
- NVIDIA Isaac Sim and Unity maybe?
- Potential to interface with industrial cable design software like EPLAN Harness
- Any issues caused by the models available in the physics engine(s)?
- Any hope that pure learned generative models will surpass "Blender-like" generation any time soon? What are the main challenges?
- Are there any tools for synthetic data generation? Can the end-users be assured their data will stay confidential while using these tools?
- Modelling the deformation capability is the bottleneck

Audience Q&A

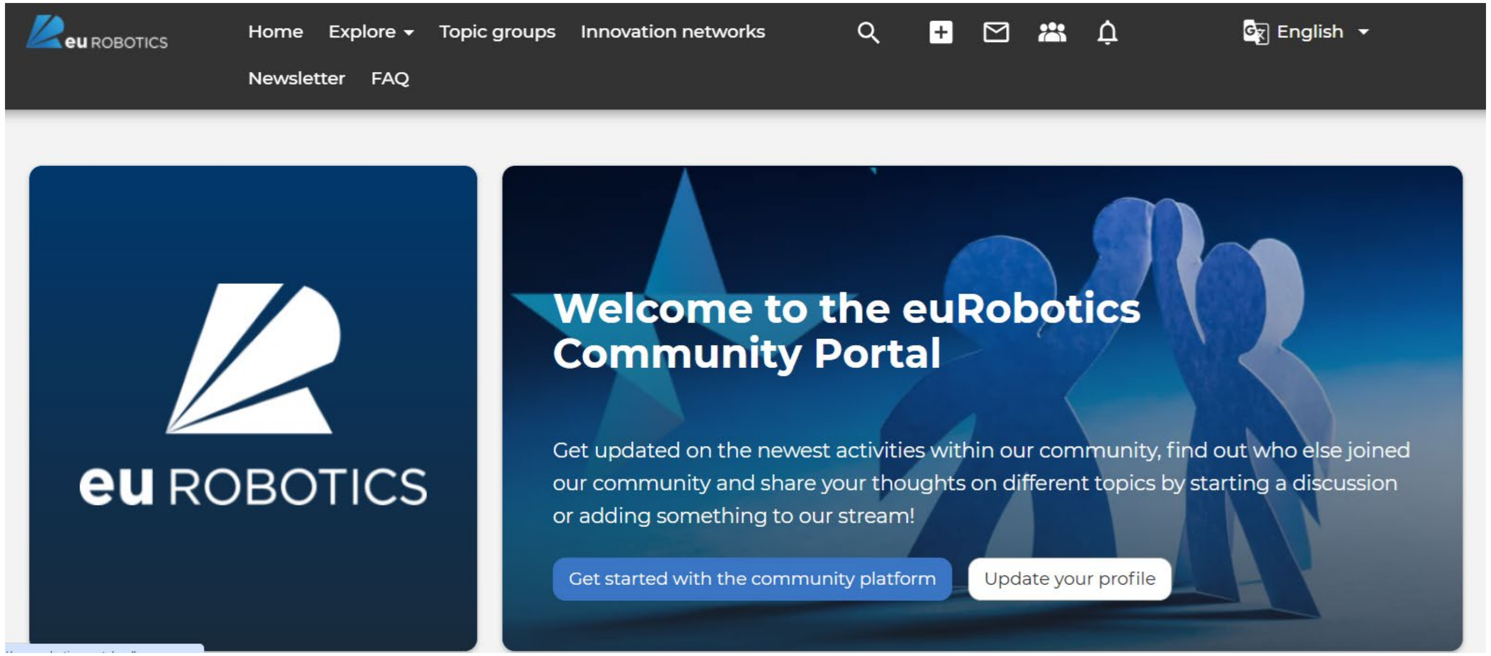
13 Participants

- Do you include contact information in the synthetic data, for grasping and packing when in clutter?
- What are we still missing in terms of synthetic data generation. For example, can we realistically generate data for deformable or articulated objects.
- Why not use the robot to collect data with hand mounted camera to collect data from different perspectives. This seem to be not mostly used - any disadvantages?
- With a success rate of 97% at a rate of 600 picks per hour, the robot would fail 18 times in 1 hour of work....how do you handle these failures?
- Can you provide some examples of unknown objects in automotive industry? Why they appear and what you are doing with them?
- You mentioned working with stereo vision data to assess added value in addition to RGB. Any conclusion(s) in your case(s)?



Closing

Slides will be published on the website:
https://roboception.com/workshop-at-erf-2025/



Join the TG Perception: https://www.robotics-portal.eu