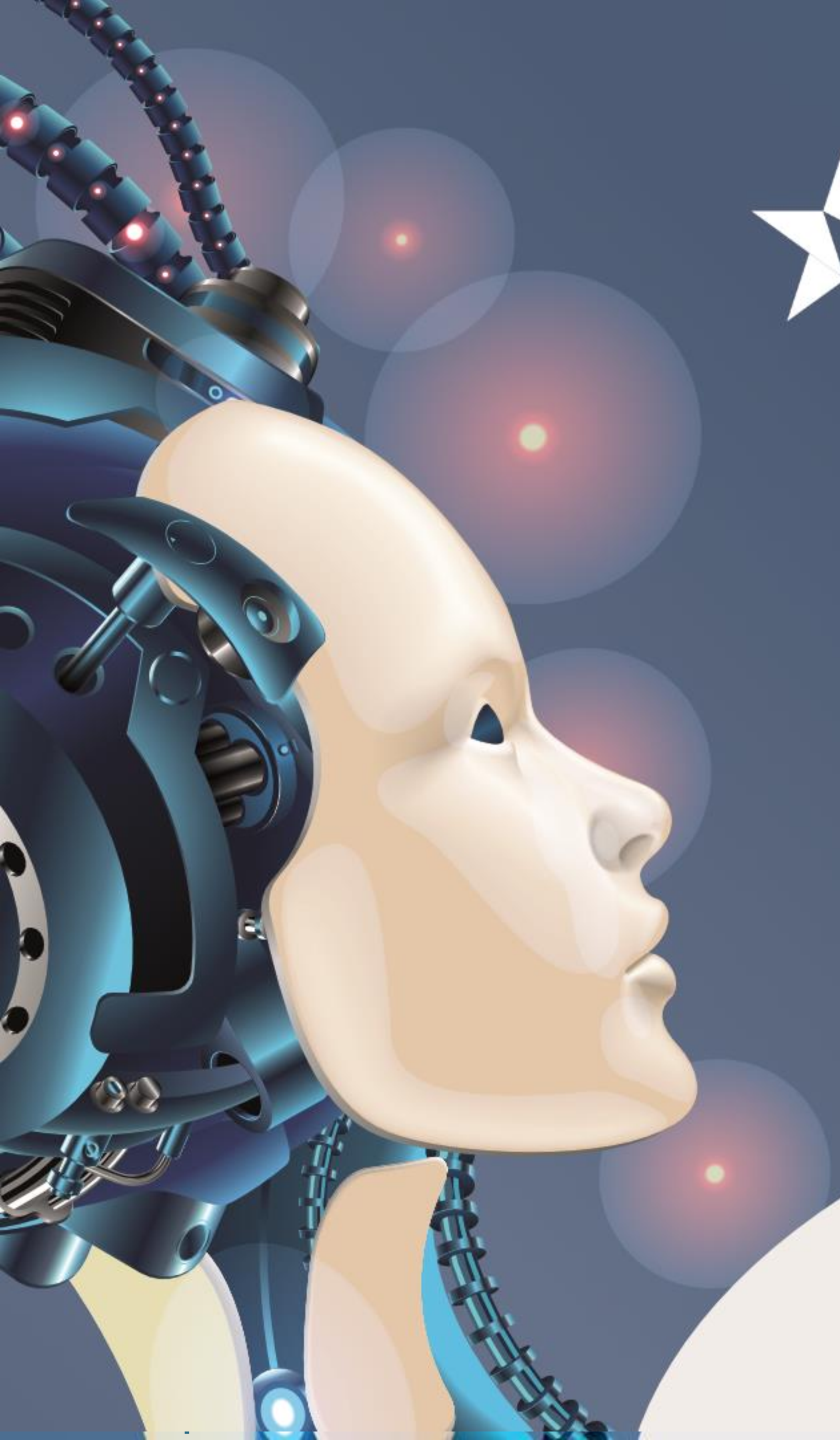


EUROPEAN ROBOTICS FORUM

eu ROBOTICS at the heart

13-15 March 2024
Rimini, Italy

ROBOTICS UNITES:
People, Countries, Disciplines

OWNER & CO-ORGANIZER

eu ROBOTICS

LOCAL CO-ORGANIZER

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

PLATINUM SPONSORS

COMAU LEONARDO SACMI SAXION

erf2024.eu

WORKSHOP:

Obtaining Good Data for Agile Production, Logistics and Lab Automation

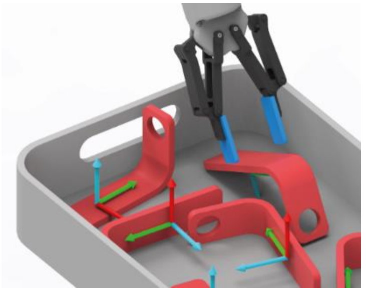
Agenda

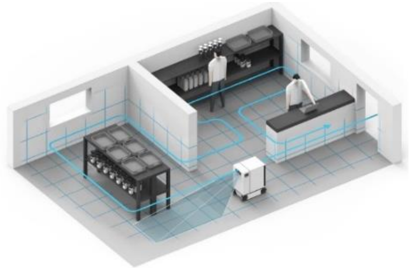
OBTAINING GOOD DATA FOR AGILE PRODUCTION, LOGISTICS, LAB AUTOMATION

14:40	**Introduction and Definition of Statements/ Key Questions** Dr. Michael Suppa, Roboception GmbH
14:45	**Towards Tricky (Transparent, Reflective, …) Object Models and Detection** Prof. Markus Vincze, TU Vienna, Austria
15:00	**Ocado Technology: On Grid Robotic Pick (OGRP)** Dr. Radhika Gudipati, Ocado Technologies, UK
15:15	**The power of synthetic data in Agile Production** Dr. Michael Suppa, Roboception GmbH, Germany
15:25	**Interactive Session/ Round Table Discussion**
15:55	**Conclusion and Take Home Messages**

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 **logistics**, manual work is still pre-dominant due to the complexity of tasks and the variation of objects.
- In **industrial automation**, accurate placement is usually the key challenge
- In **lab automation,** usually fragile and transparent objects must be handled in the processes including human interaction

Individual engineering of solutions is costly and does not scale

Key Questions

- Which level of expertise regarding 3D vision and machine learning is available in your area? (none, beginner, moderate, expert)
- Model data of the final product is usually available in agile production. How do you deal with the potential lack of knowledge during the production process when object handling is required at intermediate steps?
- Synthetic training data generation requires model data for simulation. Do you have this data available, where does it come from and how do you asses the level of correctness?

AUTOMATION & CONTROL INSTITUTE
INSTITUT FÜR AUTOMATISIERUNGS- & REGELUNGSTECHNIK

TECHNISCHE UNIVERSITÄT WIEN

Towards Tricky (Transparent, Reflective, …) Object Models and Detection

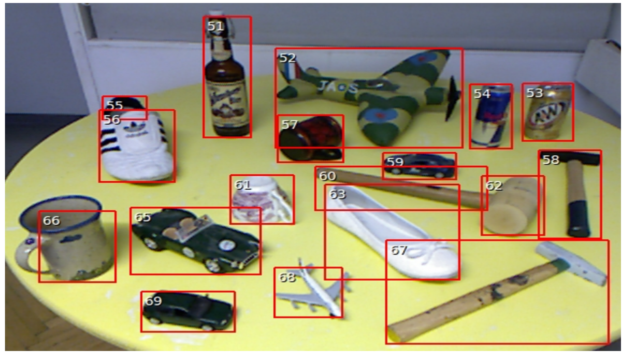
Markus Vincze
Philipp Ausserlechner, Dominik Bauer, Hrishikesh Gupta, Bernhard Neuberger, Tessa Pulli, Paolo Sebeto, Markus Suchi, Stefan Thalhammer, Jean-Baptiste Weibel

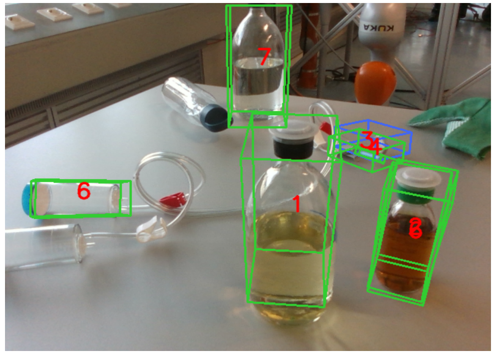
ERF 13.3.2024, WS „Good data for agile production, logistics, and lab automation"

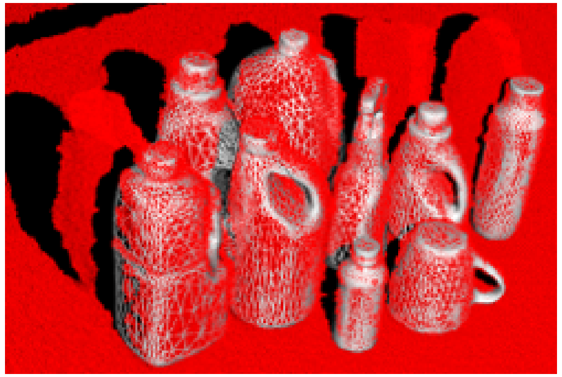
V4R – Vision for Robotics

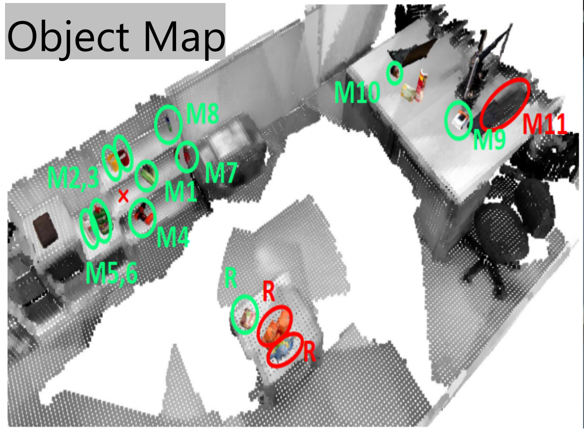
„We make robots see."

- Object X
  - Modelling & tracking
  - Recognition & classification
  - Function & affordances
  - Manipulation
- Rapid handling of novel objects
- RGB-D cameras



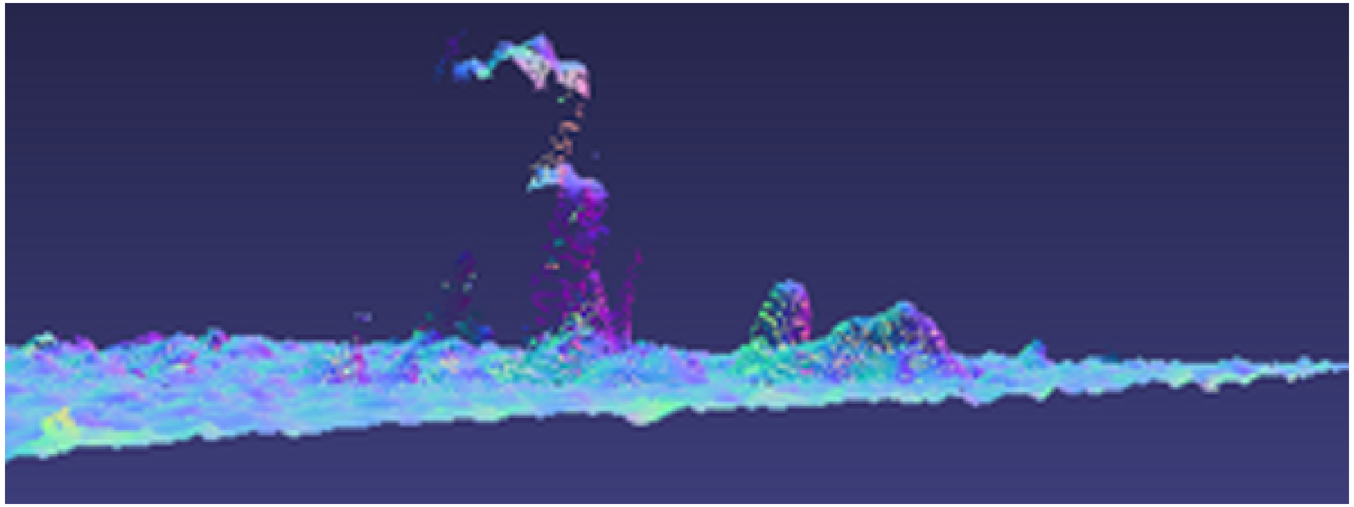






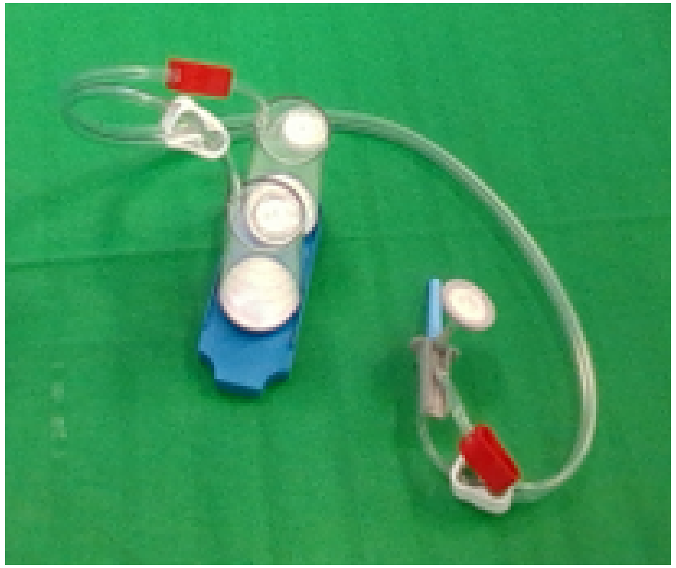
Tricky Objects and their Challenges

- Missing depth data
- Background mixed with object
- Need to work with RGB



Approach

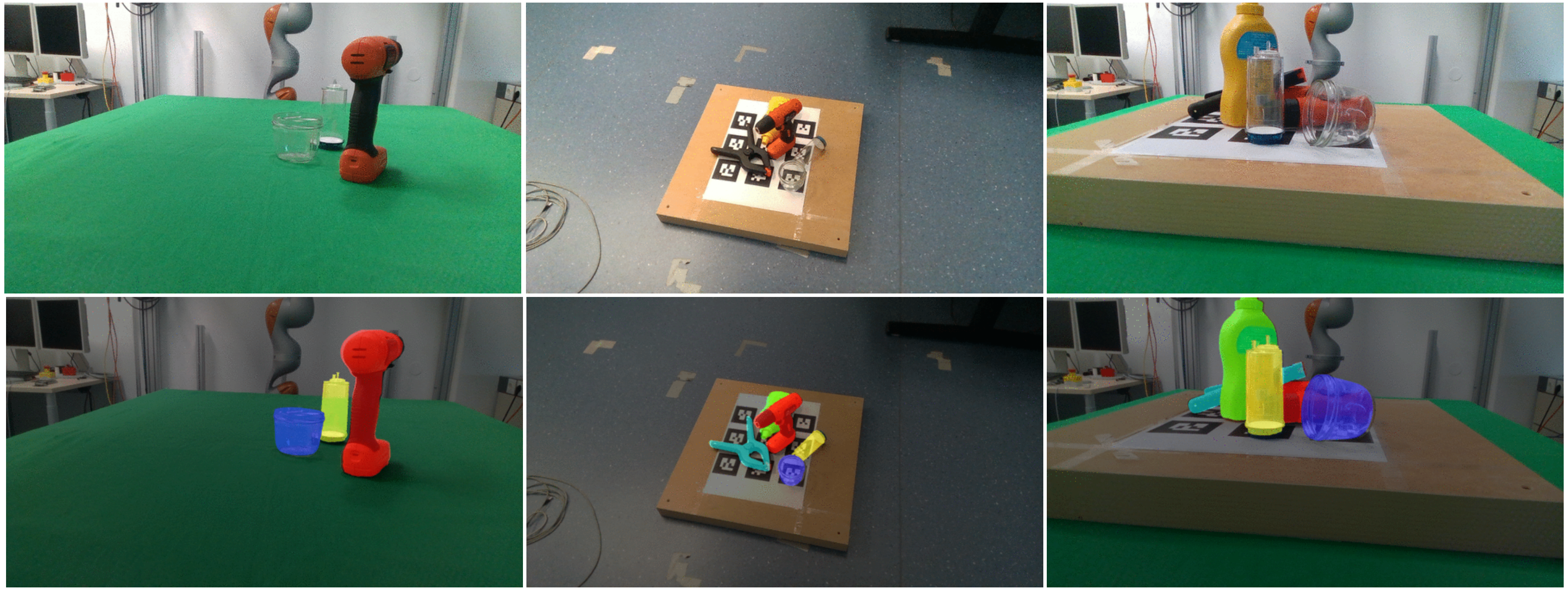
- Tools for creating data
- Modelling/rendering tricky objects
- Object pose estimation and verification
- Integration on robot for object grasping

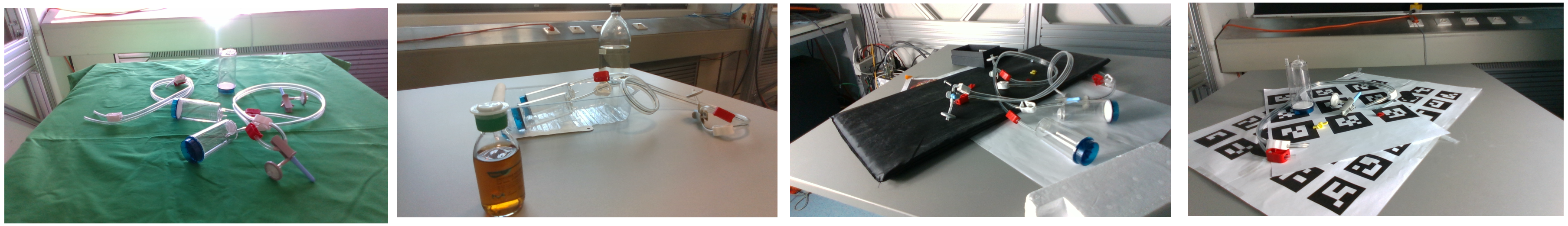


Good Data: Groundtruth?

- Do you really trust GT?
- Nets learn what is annotated
  - Tune architecture & parameters?
- GT needs to capture the actual challenge!
  - Accurate data
  - Tricky objects: transparent, reflective, shiny, small, …

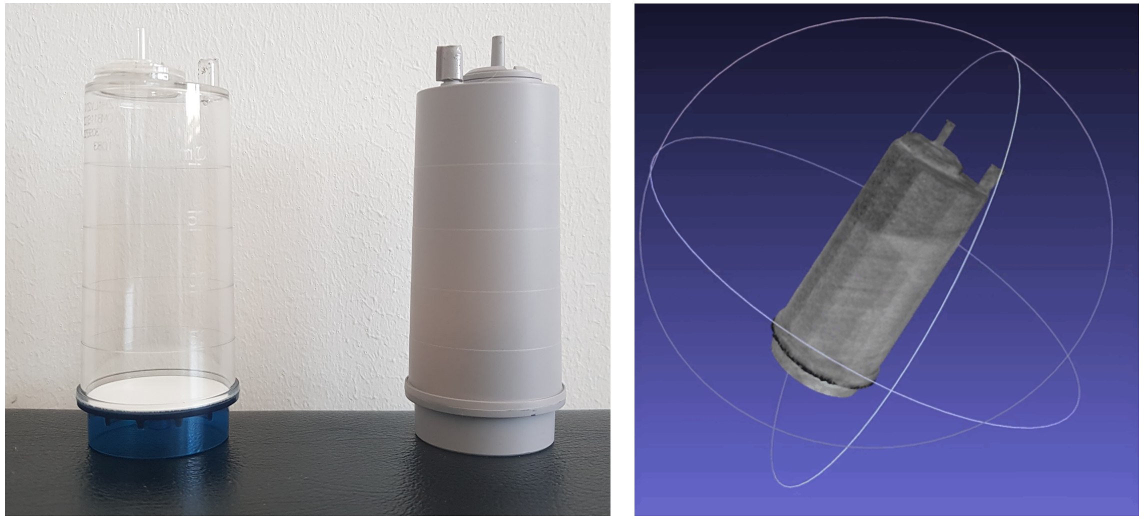
Example Scenes





Creating Models and Datasets

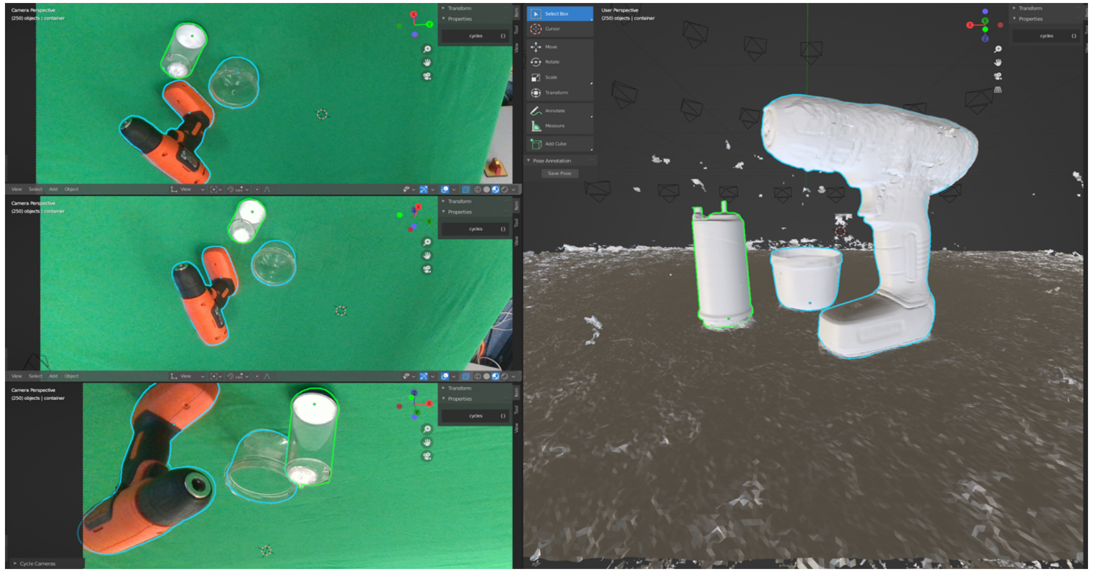
Recording sequence

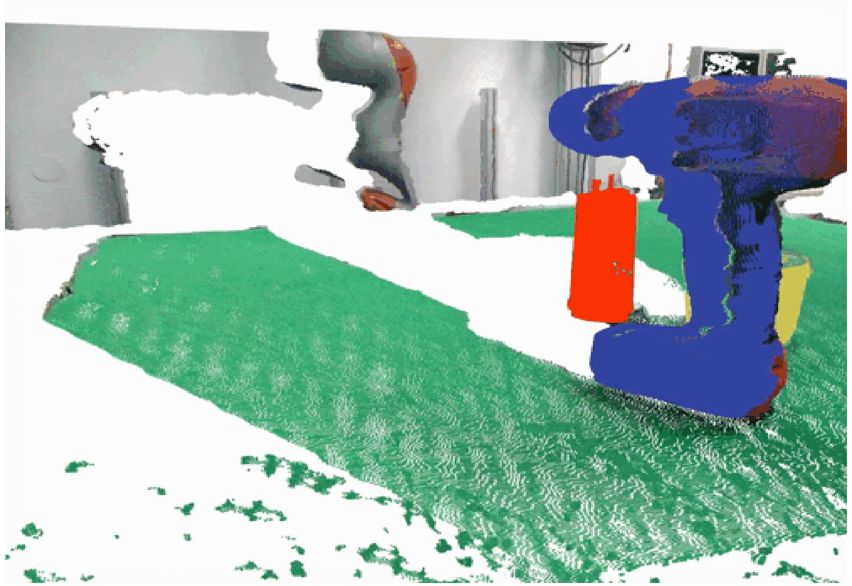


3D model from coated transparent objects

- Marker and calibrated KUKA arm for accurate camera pose
- RealSense D415, D435, automatic motion planning
- Accurate models from scans using Photoneo sensor
- Scan with up to 104 views per scene

3D-DAT – Annotation Tool

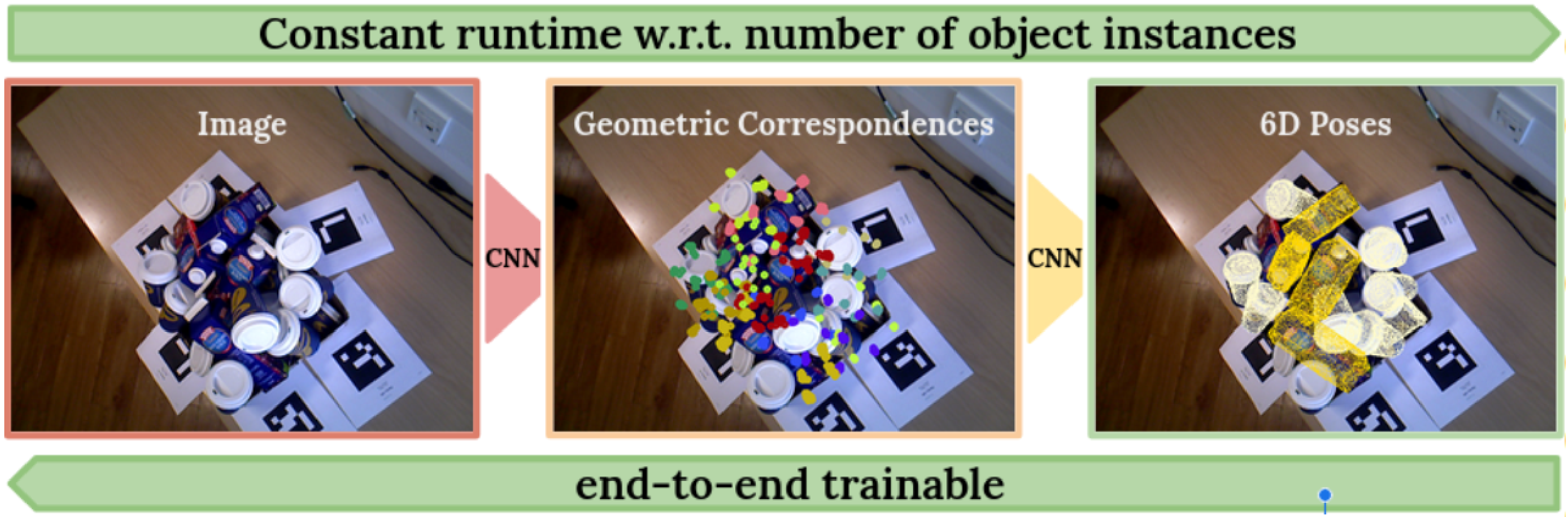




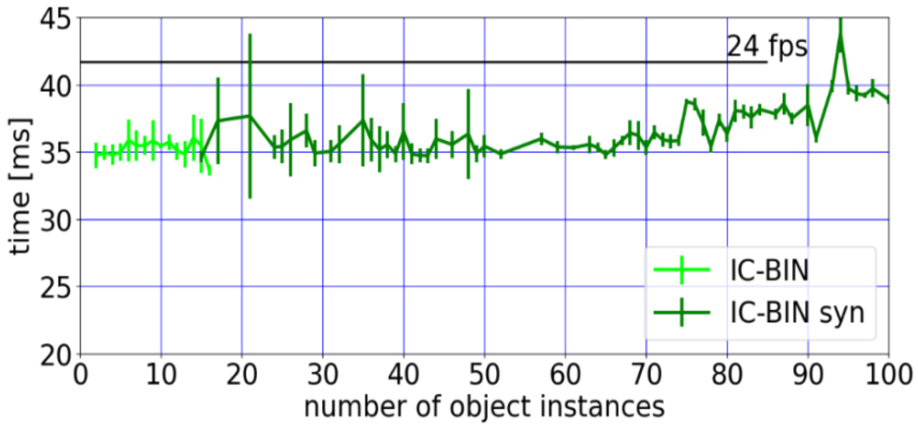
Pose annotation of tabletop scene [Suchi et al., RA-L 2021, ICRA 2023]

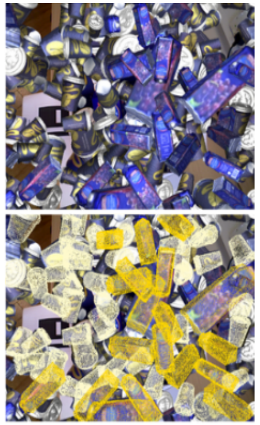
- Annotate one model with multiple views → transfer to other 103 views
- Automated NeRF reconstruction, modelling, and fit to data
  - NeRF - Neural Rediance field – volume modelling
- Allows to model objects with poor depth data (tricky objects)

COPE - Constant Runtime Object Pose Estimation



- **Learn from synthetic data only → SotA**
  Simulate sensor and target scenes
- Calculating mutual IoU in RGB image:
  Pose hypotheses are clustered →
  constant runtime wrt. number of object instances
- 24 fps with 90 objects;
  LineMod 74%; LM-O: 35%



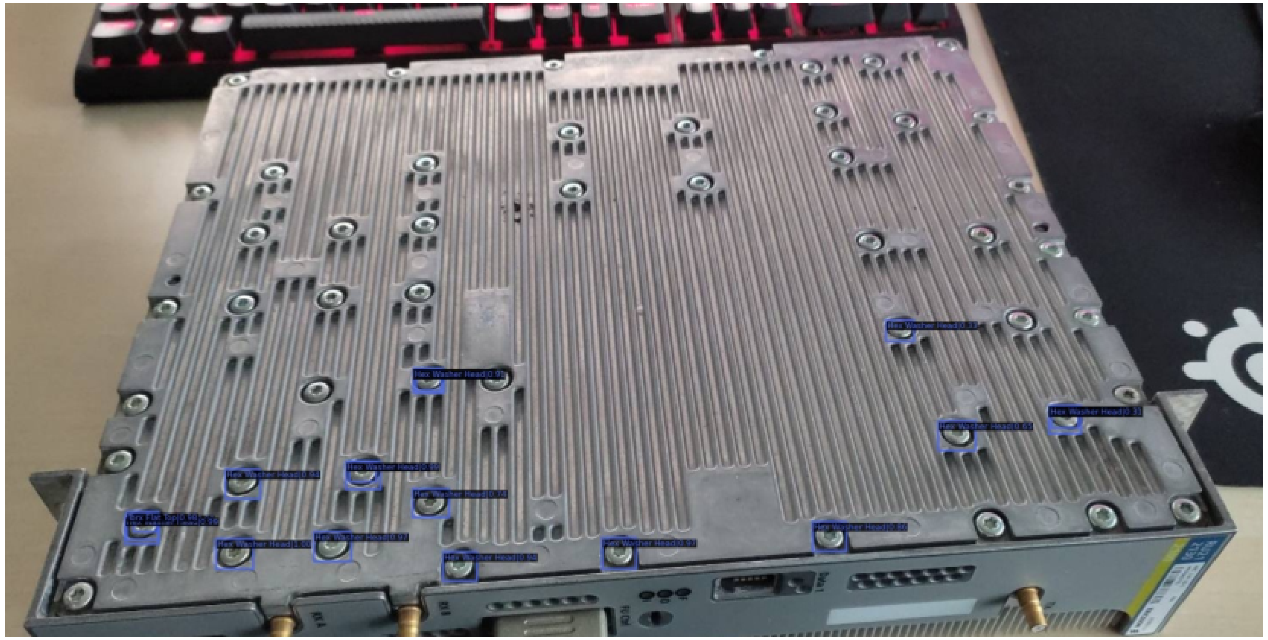


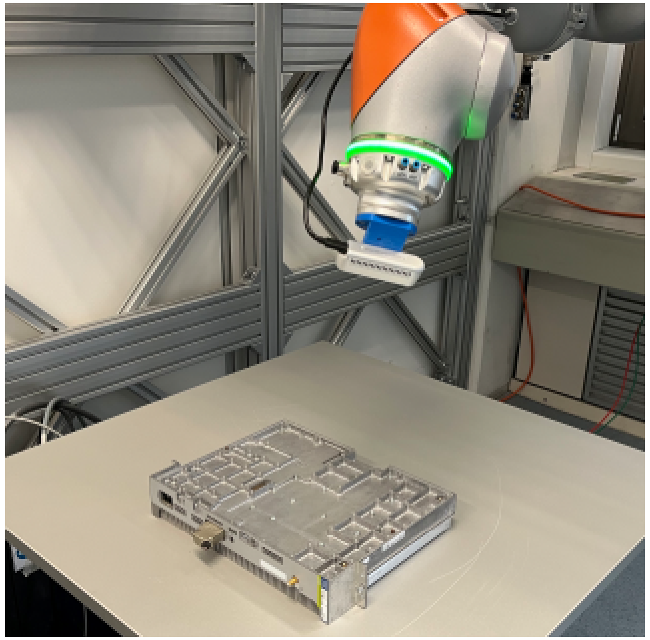
[Thalhammer et al.: COPE, WACV 2023]

COPE for Recycling

SmartDis – Smart Disassembly with a Knowledge-based Automation System

- Task: unscrew plates, > 50 screws, several layers
- Learning a few types of screws



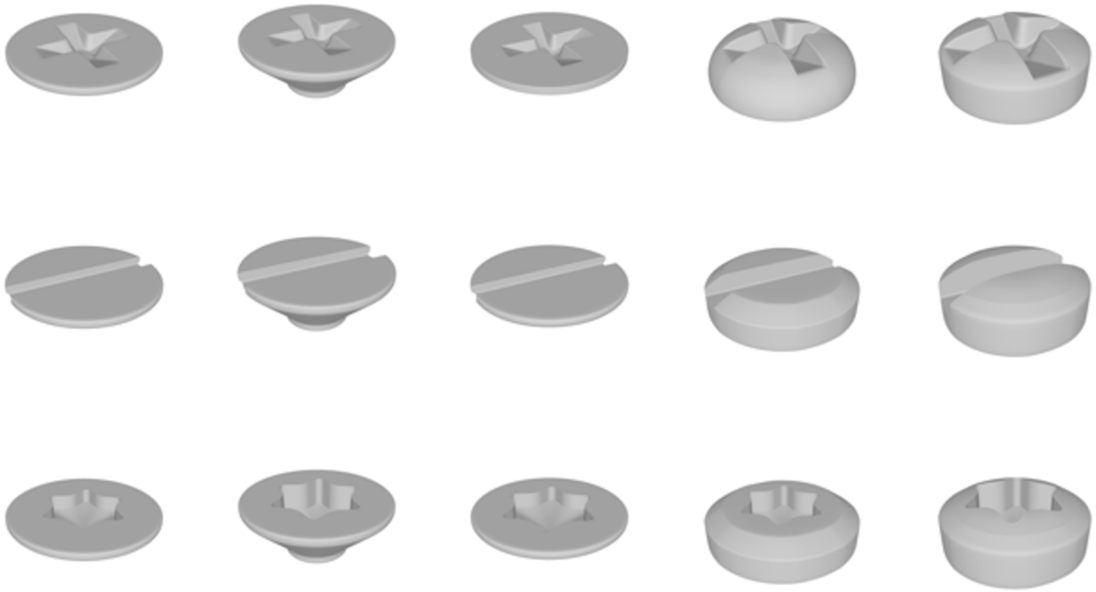


Partners:
Augusta,
SPS,
U Kiel,
Dokulil,
PRIA

[Höbert, ARW '21]

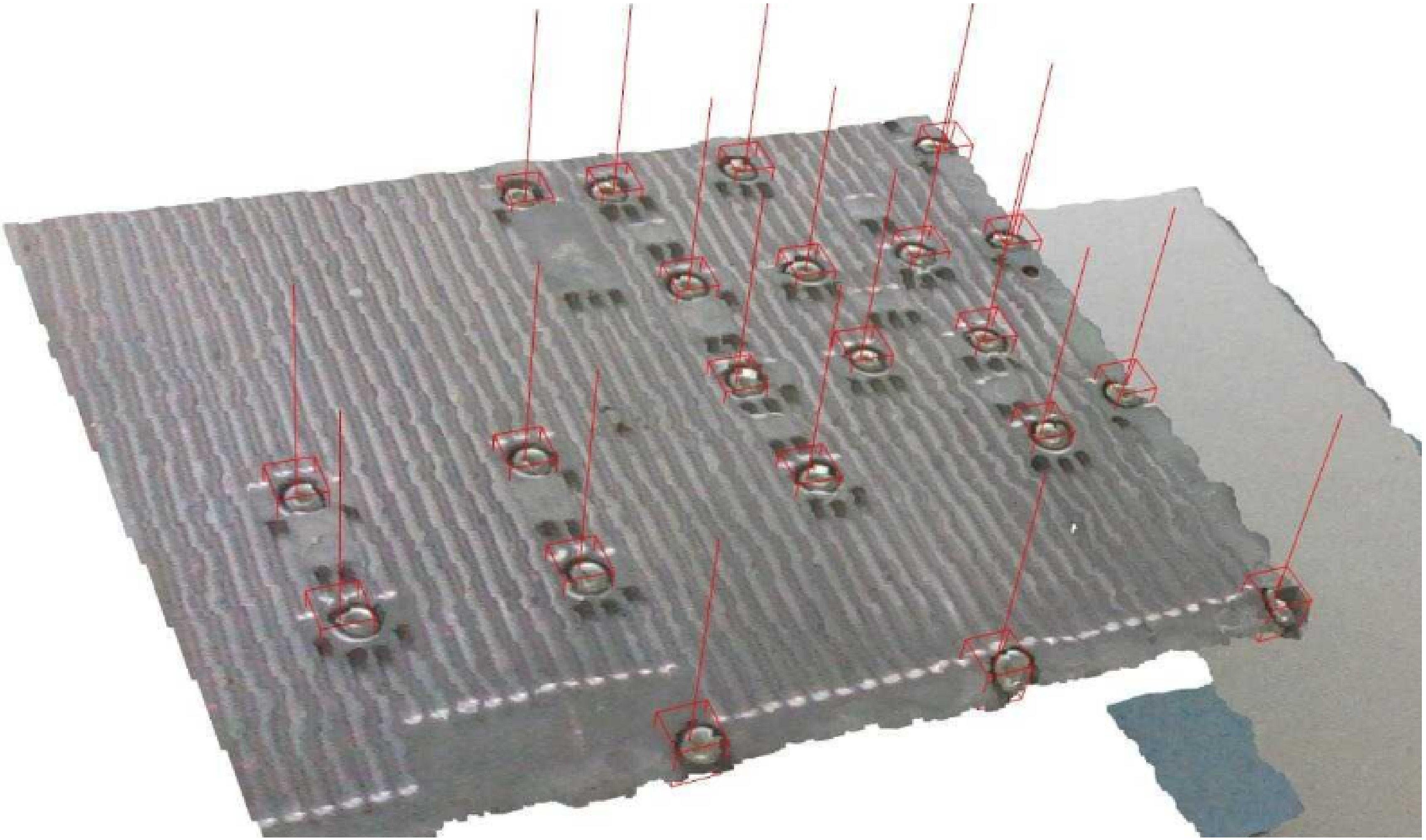
Learning from Synthetic Data

- Ideally, we learn novel objects from one showing
- In practice, need a model or many images
- Simulated data: paste model on random background



COPE and Object Surface for Unscrewing

- Shiny screws give bad object orientation → use context
- Surface patch around screw to estimate approach direction

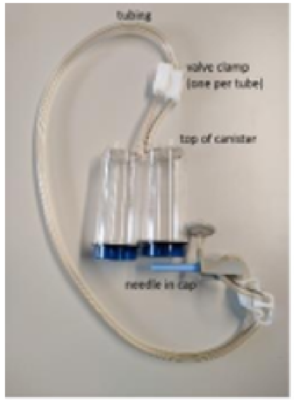


Towards Transparent Object Pose Estimation

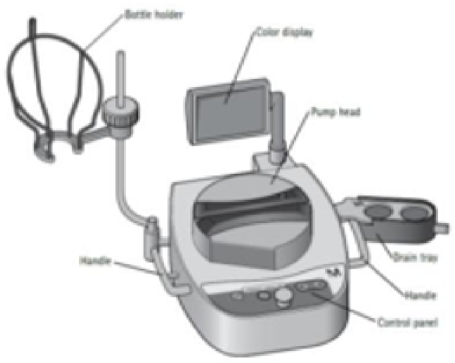
- CAD data for fast learning, robot datasets for evaluation
- Recognition of transparent objects

TRACEBOT

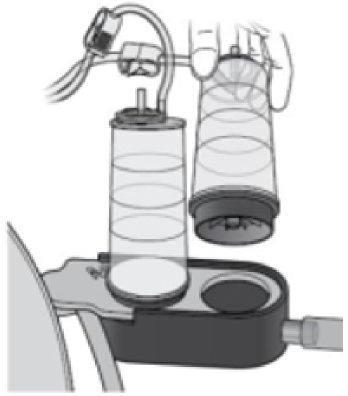
- Verification of every assembly steps
- Creation of an Audit trail



Parts of sampling kit

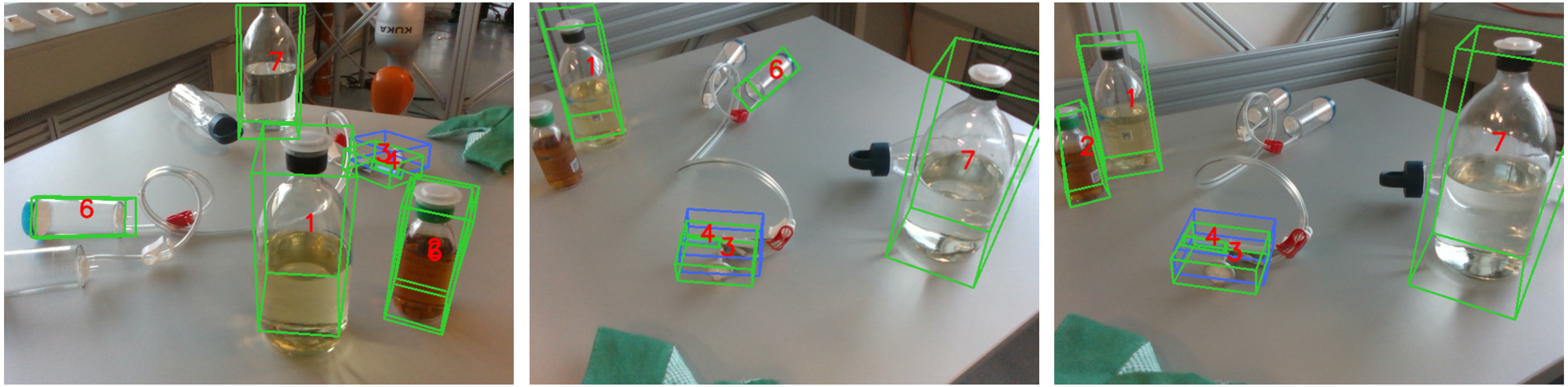


Part of pump tool



i) Fit canisters to drip tray

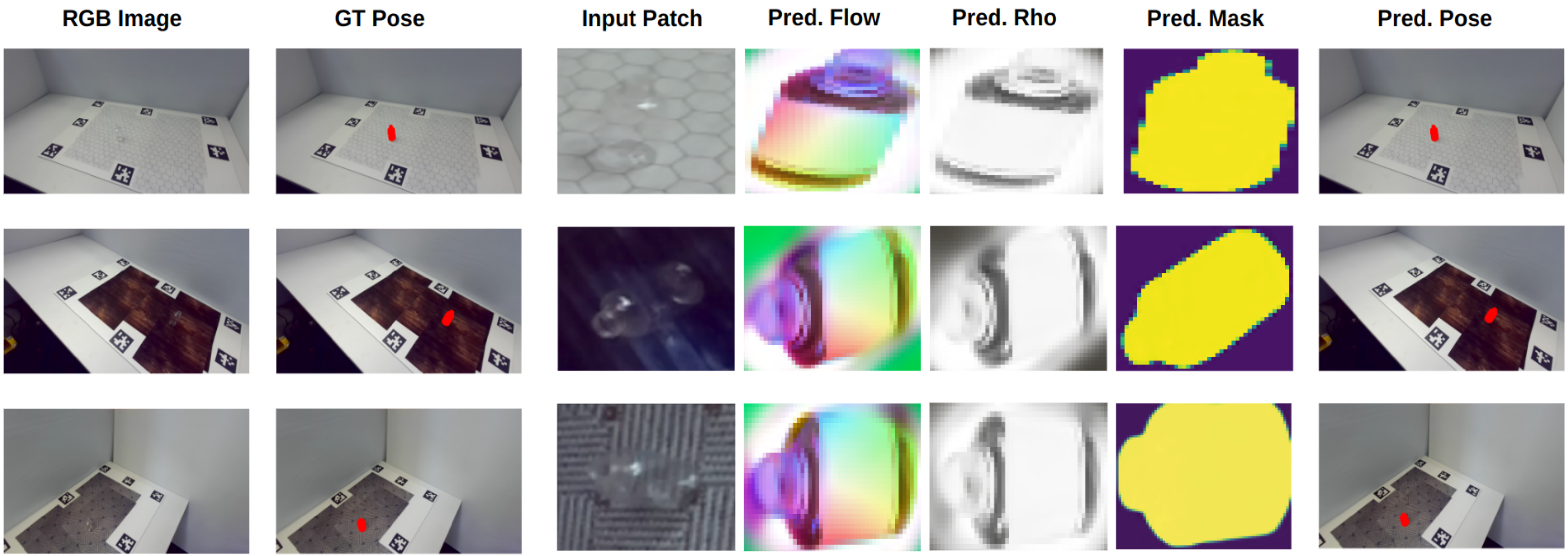
Partners: Tecnalia, invite, CEA, UBremen, astech, Biolago



[Thalhammer et al., COPE, WACV ’23]

Transparent Object Pose Estimation

- Learn how transparent materials look like =
- Learning refractive flow, attenuation (amount of light passing, rho), and mask

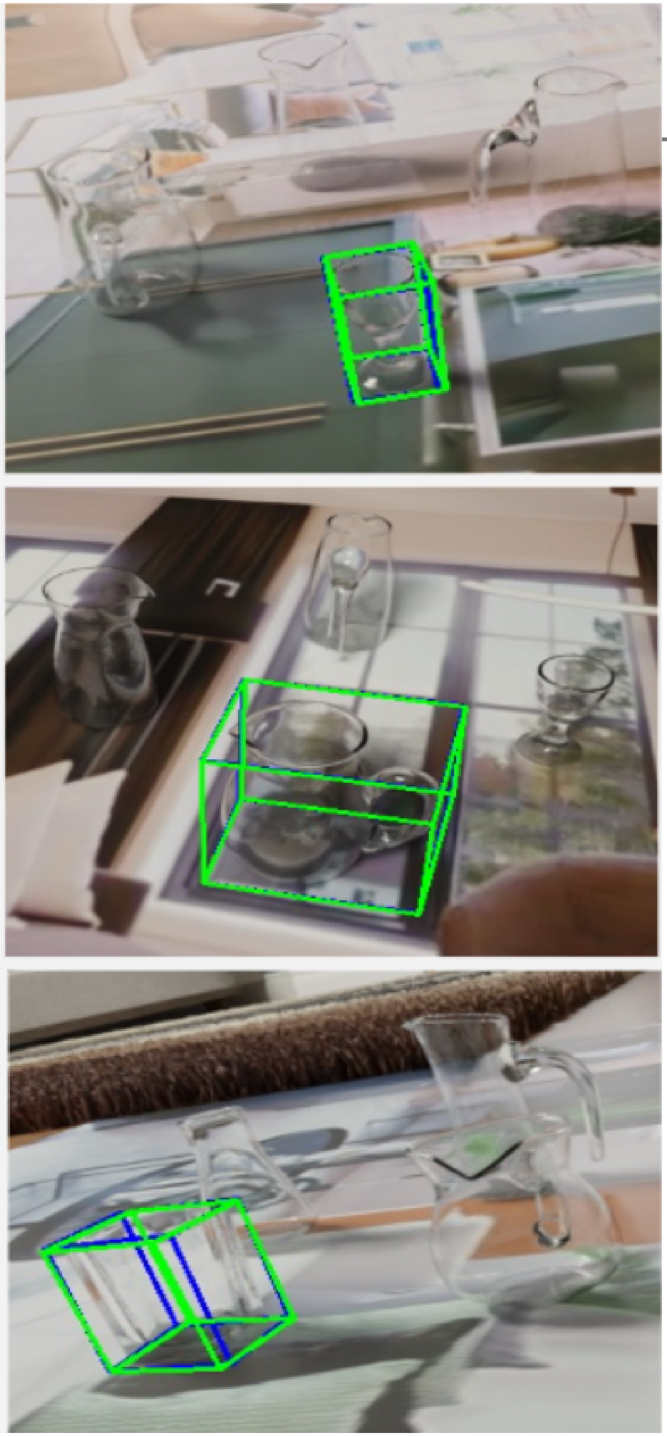


Transparent Object Pose Estimation

- Learning from samples on random background
- Examples of transparent object compositing from the TOD and Trans32K-6D datasets
- Random COCO background

Examples

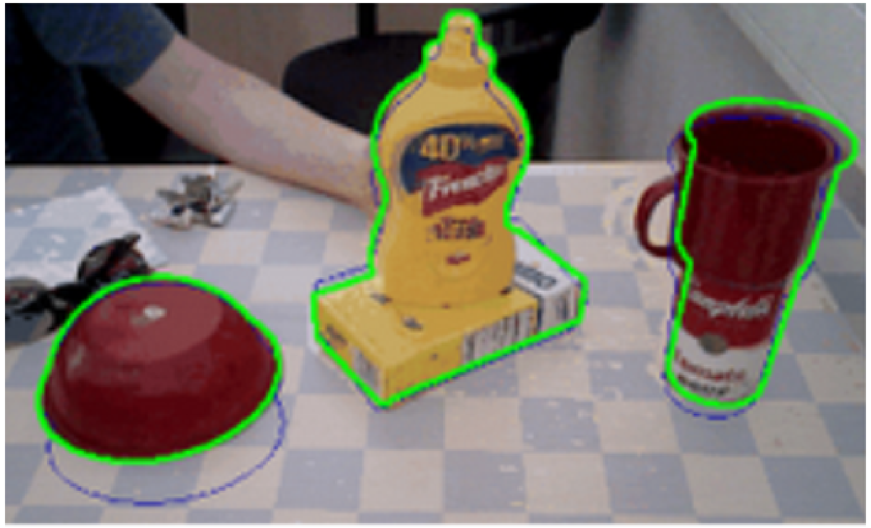
Original image

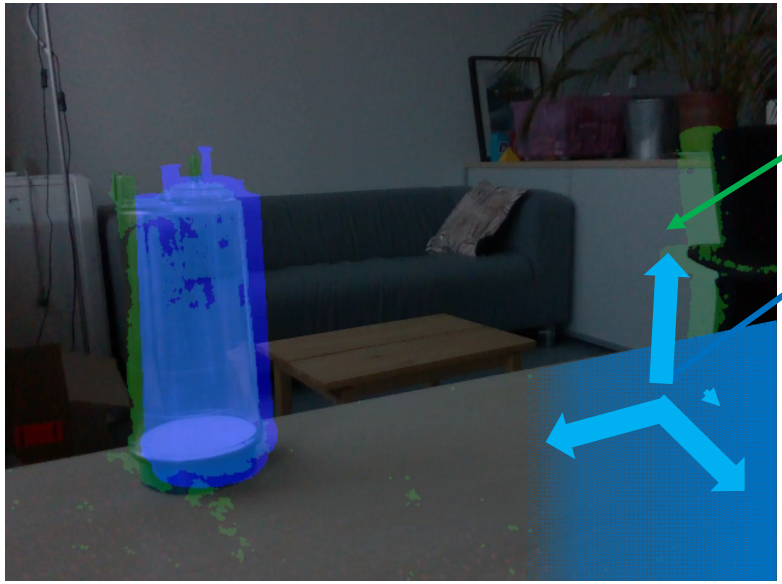


Target object, enlarged

Object Pose Verification

- Detectors create object and pose hypotheses
- Iterative physics simulation and pose refinement
  - Inverse rendering to align with visual evidence
  - Combine physical plausibility with observation: e.g. align to table plane





Observed mask (target)

Rendered mask (pose)

Plane coordinates (constrain to xy-translation and z-rotation)

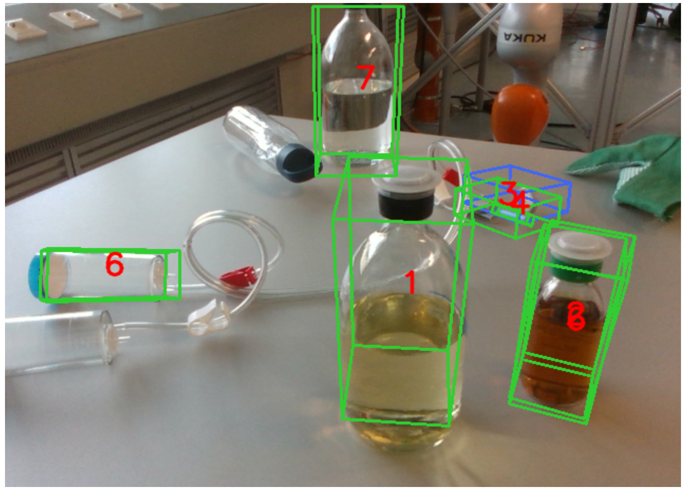
[Bauer et al.: VeREFINE, RA-L 2020; Bauer et al.: CVPR 2022, WACV 2022]

Grasping Transparent Objects

Lessons Learned

- Object detection: good methods for rigid, opaque objects
  - Make sure GT is trustworthy
  - E2E is „nearly“ as good as the data
  - Synthetic data: simulate sensor and expected scenes
- Hypothesise and verify - best of learning and simulation
  - Task sets the scope of what will be seen
  - Models help to verify what is seen
  - New, better learned method → better hypothesis → easier verification

Reliable robots need both: hypotheses & verification







TECHNISCHE UNIVERSITÄT WIEN

TRACEBOT

HORIZON 2020

SmartDis

FFG

BURG
HEAP
InDex

chist-era

FWF

FESTO

Ocado Technology: On Grid Robotic Pick (OGRP)

13-Mar-2024, ERF 2024

Dr Radhika Gudipati

Sr Research Coordinator (Robotics & AI)

Who we are

Ocado Technology - technology pioneers

We’re solving some of the **toughest technological challenges** of our age

We bring together some of the greatest minds in engineering, product, data science, robotics and UX

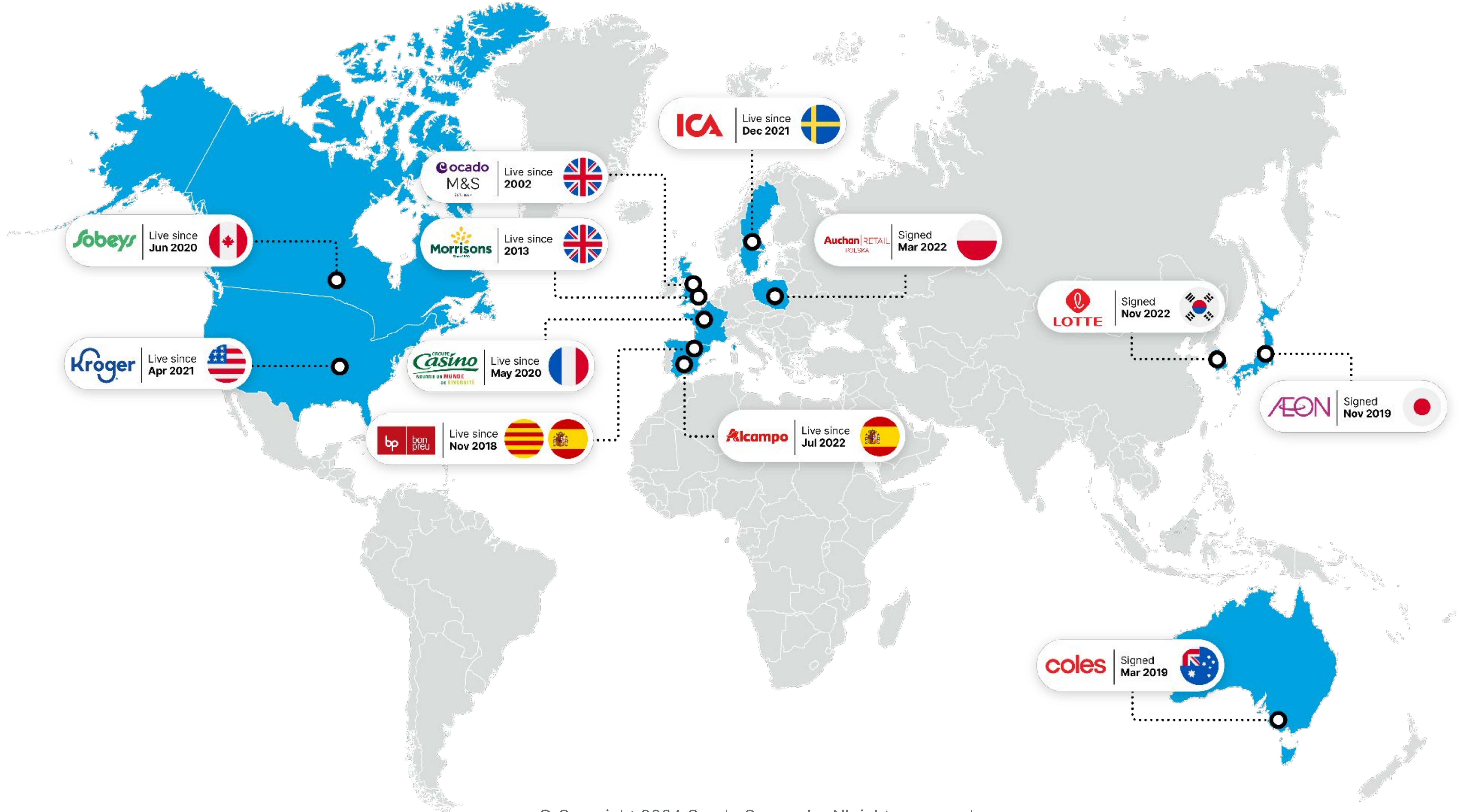
Twelve development centres out of **eight countries**

Thirteen global retailer partners

Over 2,400 patents granted and filed…

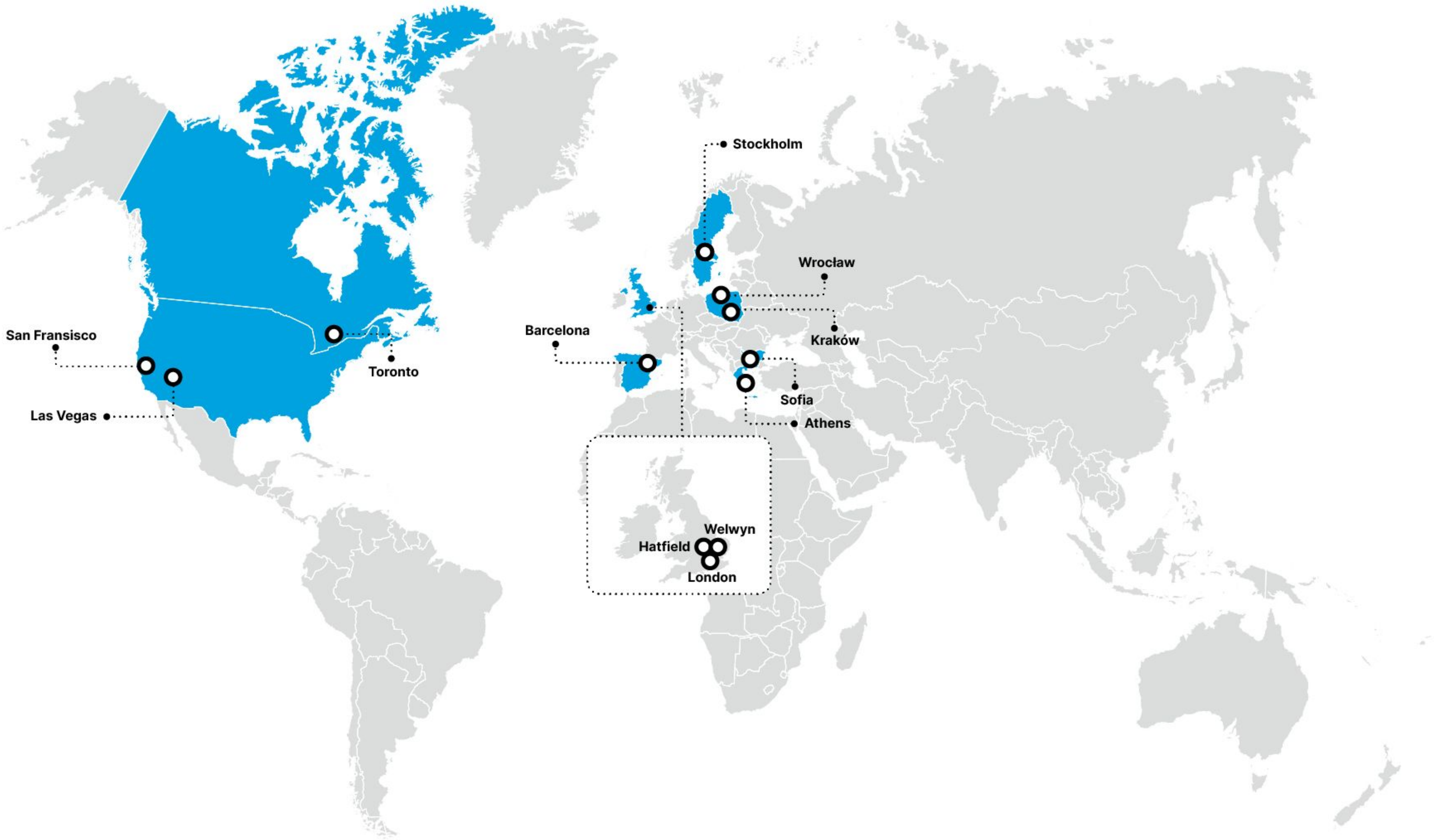
2,500+ technologists

Our retail partners



ocado TECHNOLOGY

Our development centres



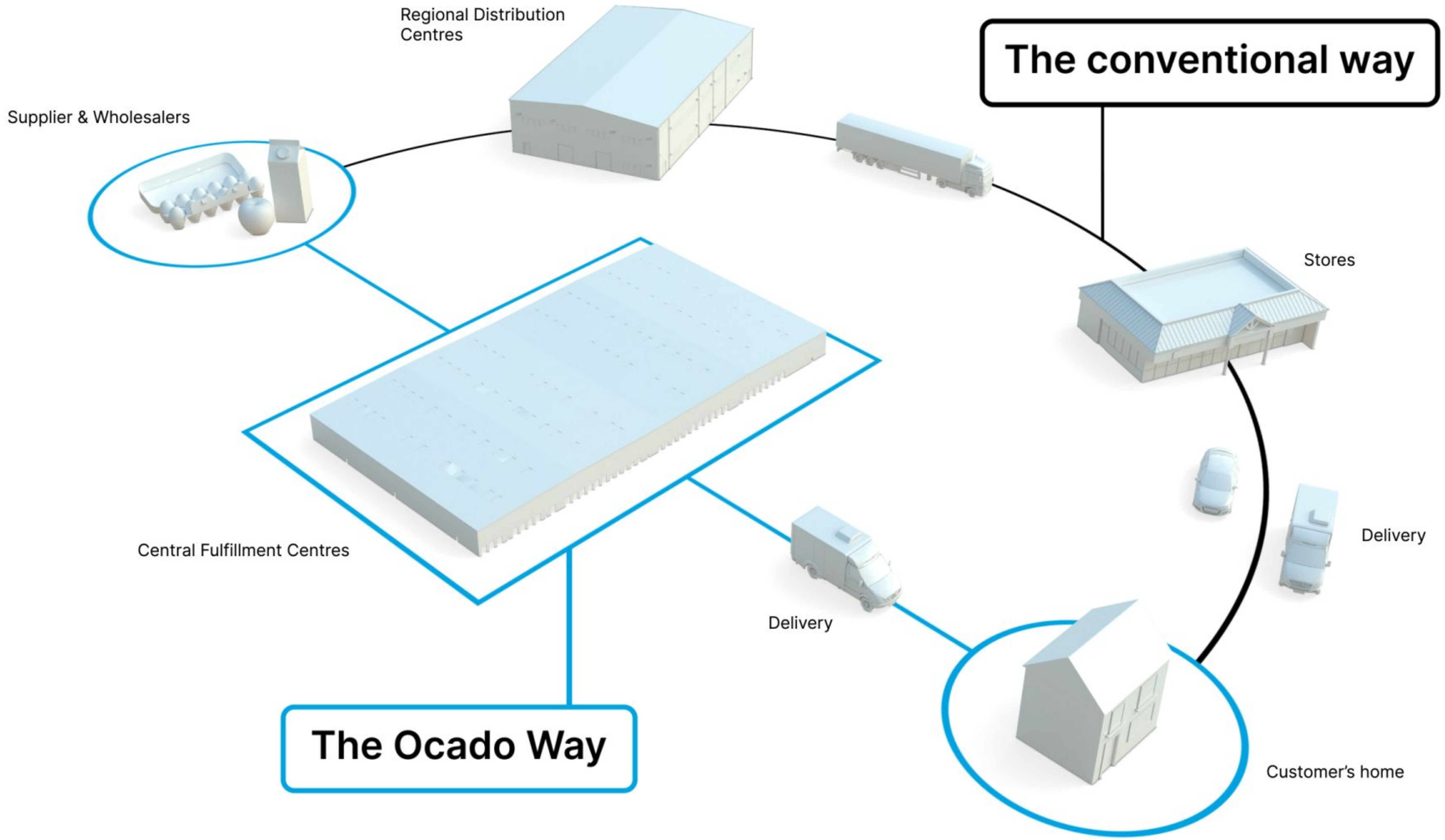
Our pioneering technology

The future of online grocery and beyond

We're pioneering the future of online grocery and beyond through cutting-edge technology and serial innovation

- We create and support the cutting-edge technology that powers the Ocado Smart Platform (OSP).
- OSP combines advanced capabilities in AI, Robotics, Digital Twins, IoT, Cloud and Big Data.
- Some of the world's leading grocery retailers have selected OSP to leapfrog their competition, offering online grocery with the best customer experiences and superior economic returns

The Ocado way



ocado TECHNOLOGY

The Hive

What is The Hive?

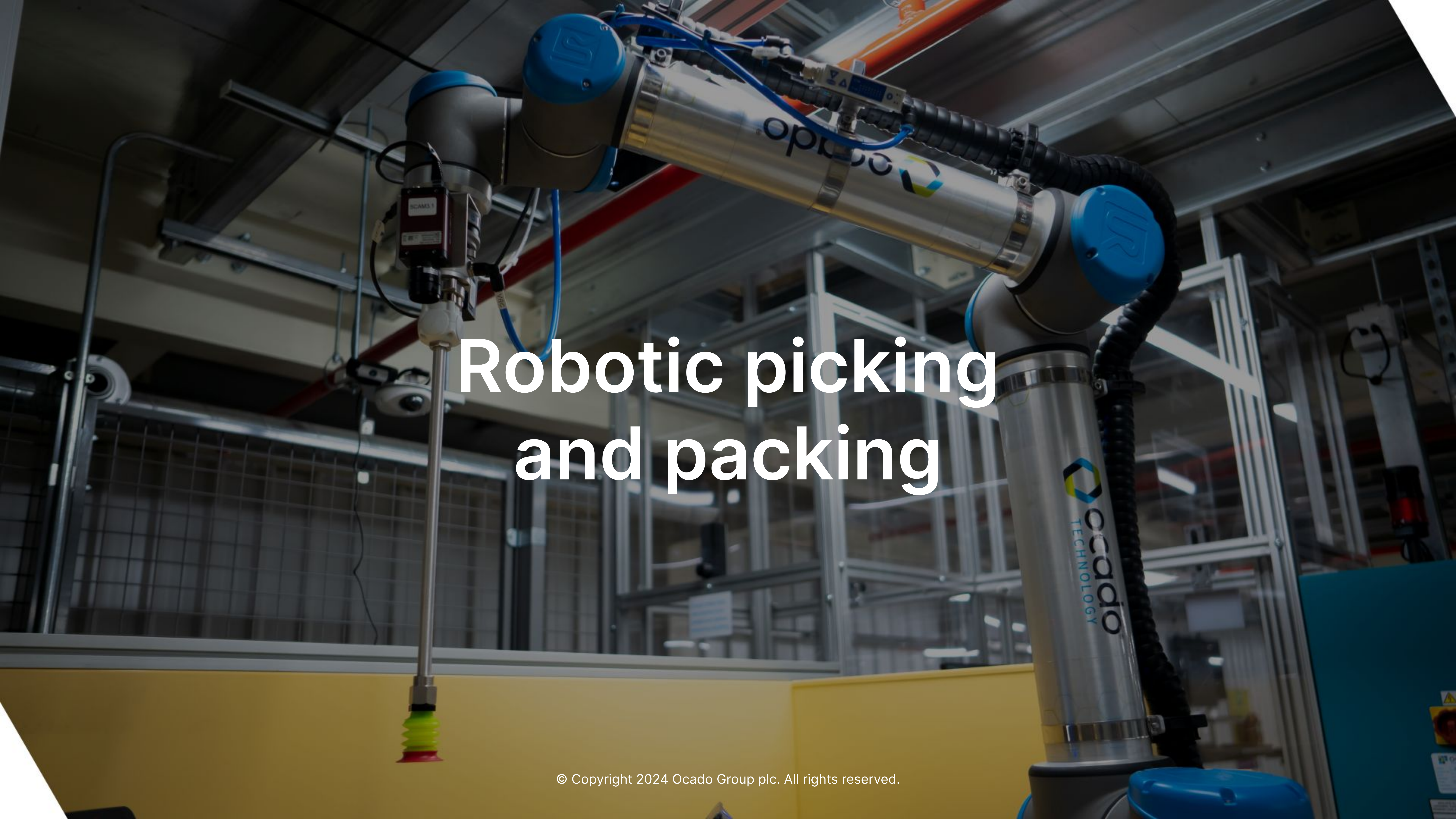
- The Hive consists of the grid and the bots which run on it
- Thousands of bots are orchestrated by AI, collaborating to pick 50-item customer orders in just five minutes, and up to 150 orders simultaneously
- Storage of items in the grid is constantly optimised for availability and efficiency
- The Hive enables fast, accurate picking for the best economics in grocery fulfilment.

ocado TECHNOLOGY

Our bots

How our bots work:

- Bots whizz around the grid at speeds of up to 4m per second, with just millimetres between each one
- Bots operate as a highly coordinated swarm, orchestrated by our AI ‘air traffic’ control
- Bots are modular and identical - any requiring preventative maintenance can instantly be replaced with no loss of throughput
- Each bot records 1 GB of data per day, or 4TB per day for an entire swarm*
- We use ML to analyse this vast data information for monitoring and oversight.



Robotic picking and packing

https://www.youtube.com/watch?v=UjczYsjO6fQ&t=17s

On-grid Robotic Pick

An innovation to automate the picking and packing of grocery orders

- Robotic arms need to be able to pick tens of thousands of products of varying shapes, sizes, weights, and fragility and pack them densely in bags with human precision and accuracy.
- On-Grid Robotic Pick combines cutting-edge machine vision, deep reinforcement learning and advanced sensing to pick and pack grocery items without any prior knowledge of what they are, making smart decisions on the fly.
- On-Grid Robotic Pick integrates with our 'Hive' solution and collaborates seamlessly with all Ocado bot fleets to fulfil customer orders without human touch, directly from the Grid.
- Being able to pick directly from the Grid opens up new opportunities to optimise warehouse design, enabling unmatched capital and site efficiency.

Challenges of robotic picking and packing

- Must be able to pick 50K+ different products without prior knowledge of what they're about to pick
- Need to be able to grasp and manipulate items of varying
  - Shapes
  - Sizes
  - Weights
  - Deformability
  - Fragility
- Must achieve speed and accuracy
- Avoid damage to delicate items
- Pack densely into bags

ocado TECHNOLOGY

Legal bits

Kroger is a registered trademark of The Kroger Co.

Morrisons is a registered trademark of Wm Morrison Supermarkets Plc.

Groupe Casino is a registered trademark of Casino Guichard-Perrachon S.A.

M&S is a registered trademark of Marks and Spencer Plc.

ICA is a registered trademark of ICA AB.

Lotte is a registered trademark of Lotte Corporation.

Bonpreu is a registered trademark of Bon Preu group.

Sobeys is a registered trademark of Sobeys Inc.

Coles is a registered trademark of Coles Supermarkets Australia Pty Ltd.

Dobbies is a registered trademark of Dobbies Garden Centres Limited.

Aeon is a registered trademark of Aeon Co Ltd.

Alcampo is a registered trademark of Auchan Retail Group.

Thank you!

Questions?

radhika.gudipati@ocado.com

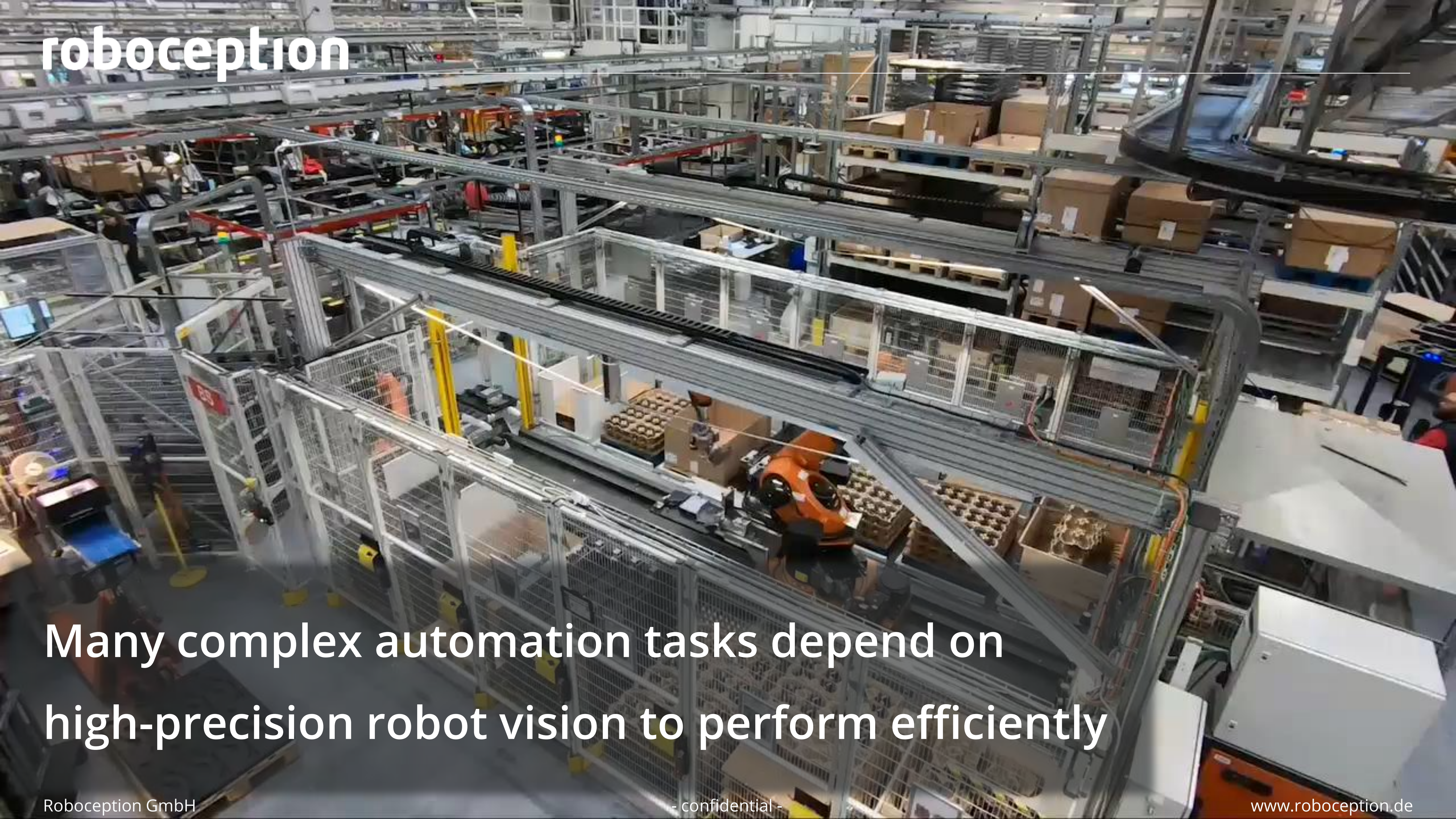
careers.ocadogroup.com

@OcadoTechnology

The Power of Synthetic Data in Agile Production

Dr. Michael Suppa

roboception



Many complex automation tasks depend on high-precision robot vision to perform efficiently

Challenges for Automation of Complex Tasks

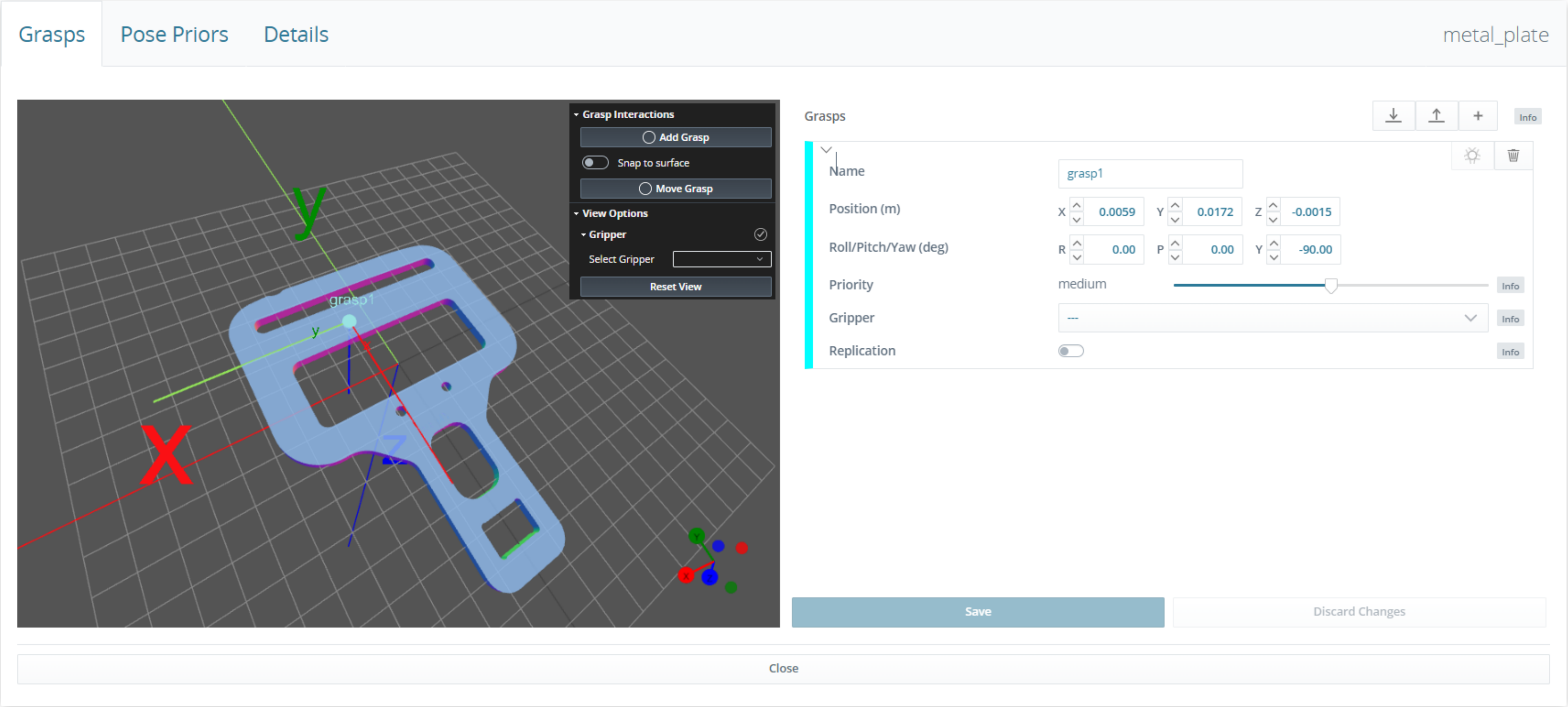
MULTIPLE VARIABLES ARE TO BE CONSIDERED

1. Variation in the environment
   - Illumination
   - Point of view
2. Variation of the object
   - Material ranges from shiny and transparent to translucent and black
   - Various sizes and distances
3. Variation in the application
   - Parameterization requires expertise
   - Test and commissioning time

Data must cover all variations, which is hard to achieve with real data or human labor.

Model-Based Detection For Pick-And-Place-Applications

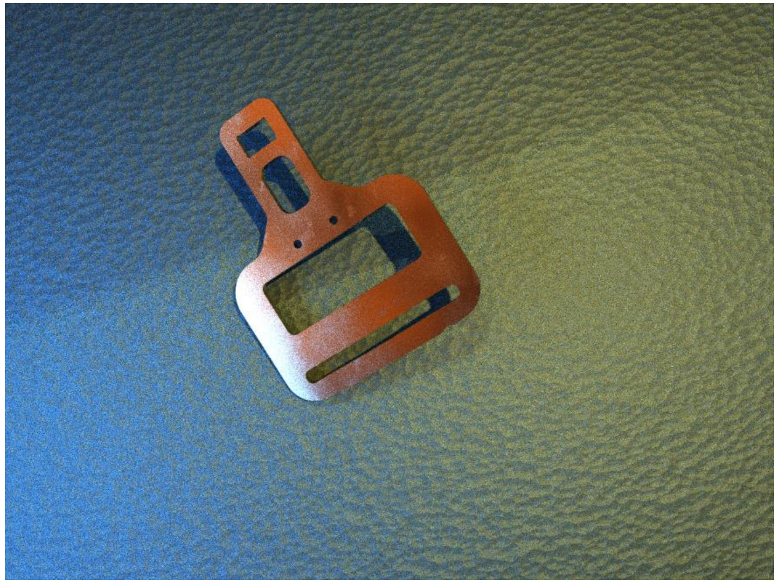
HIGH ACCURACY IS NEEDED FOR ORIENTED PLACEMENT

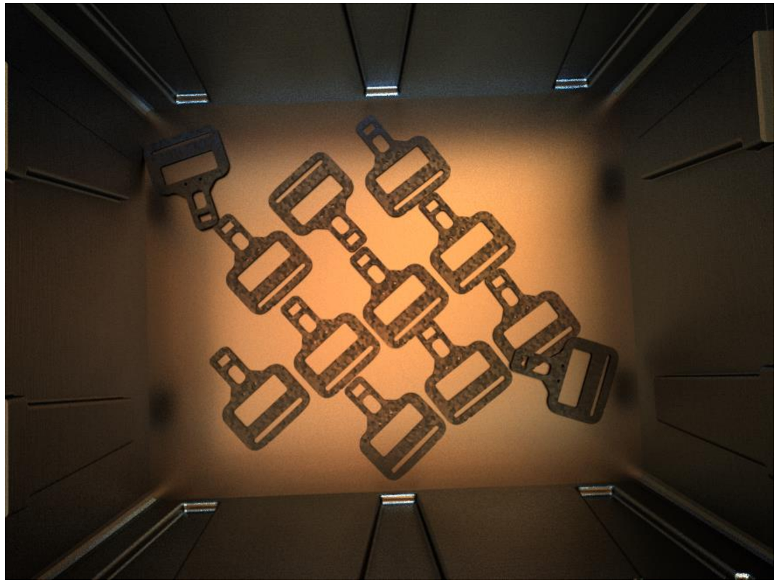


Template Generation Based on Synthetic Data

SIMULATION ENVIRONMENT

- Training images generated in a photorealistic simulation environment
- Large material library for robustness against color response and lightning conditions
- Requires no on-site data recording
- Support for different use cases and multi-material parts





roboception

AI-based simulation with synthetic data saves time and money by eliminating the need to collect data in the field.

Dealing with Variations

Issues:

- CAD model typically represents the final product stage
- Part may be seen at intermediate steps (e.g. casting, stamping, cutting)
- Part may include additional parts not in CAD (e.g. pipes, cables, plugs)
- Access to CAD may not granted, e.g. when packaging automotive parts

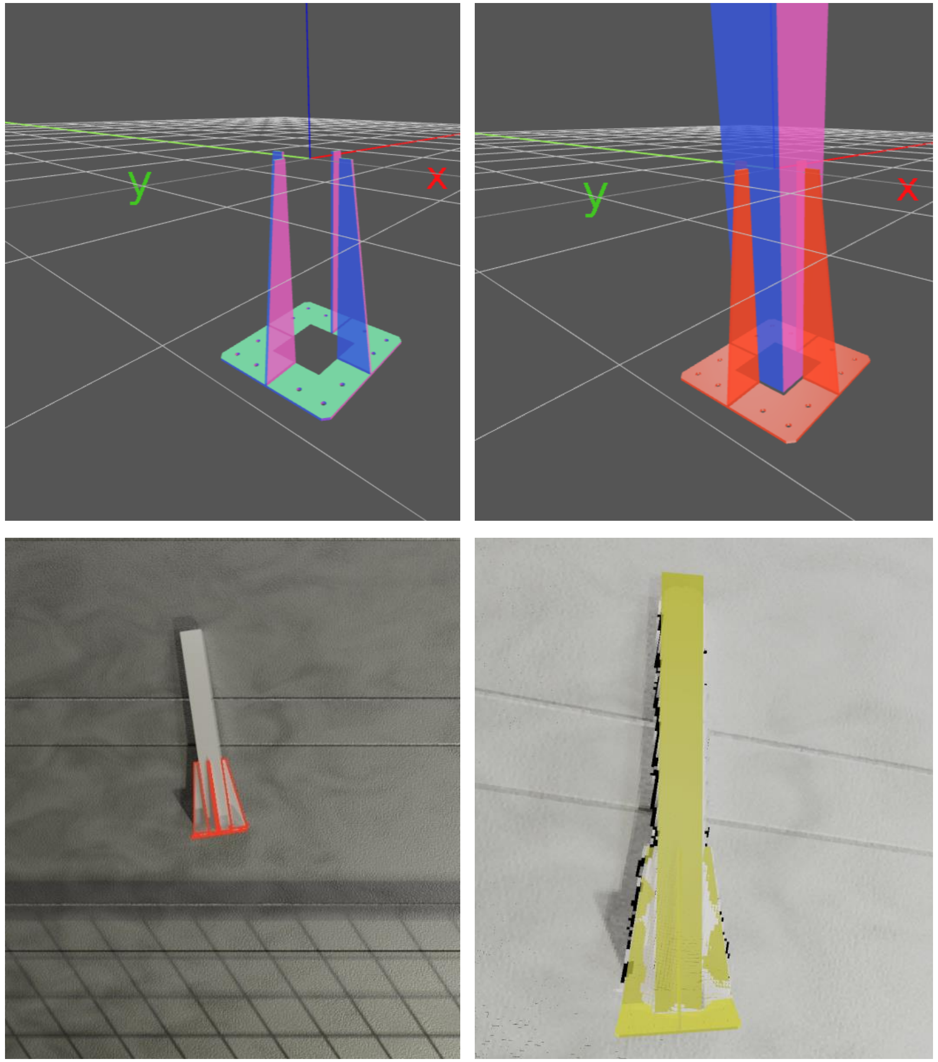
Possible solutions:

- Use only areas which are visible at the detection stage
- Generate model based on the data available

CADMatch Template Generation

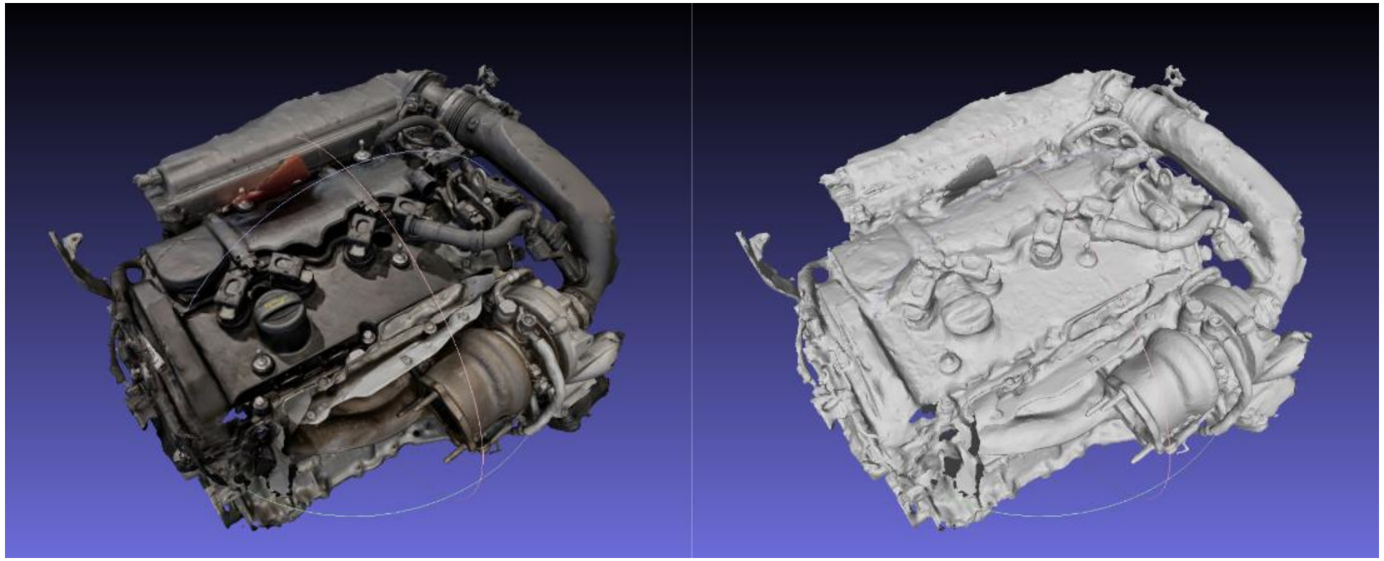
PARTIAL TEMPLATES

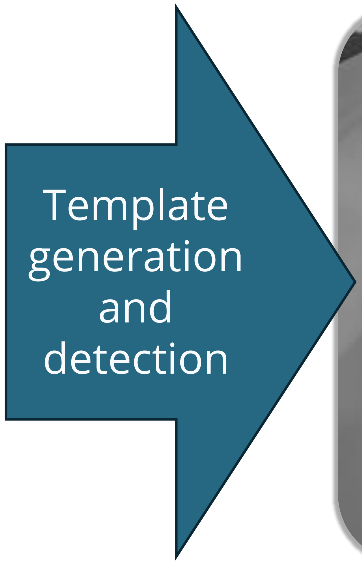
- Enables detection of portions of a complete CAD model (partial objects)
- Target use cases:
  - Large objects that cannot be entirely covered in one camera view
  - Objects that are largely occluded when placed in a bin (e.g. large stacks of flat parts)
  - Configurable objects (e.g. a switch that can change between two configurations)
  - Partially solid objects: object that have a partially soft or changing structure (e.g. brushes)

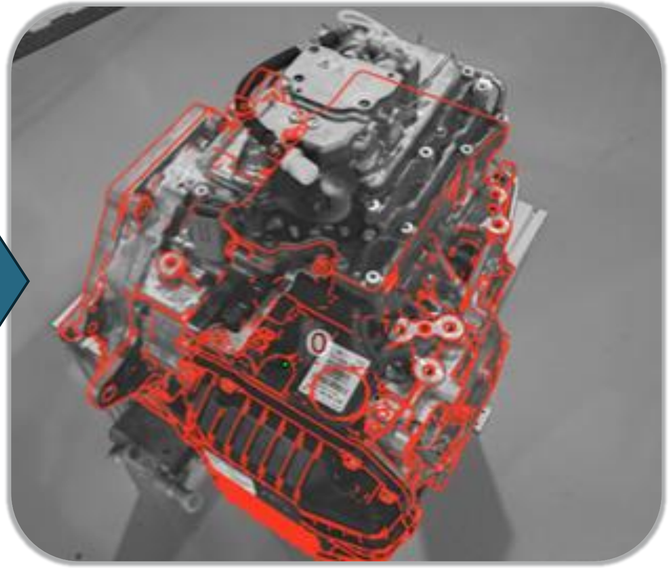


ODIN Project

- CAD Model did not contain the cables or plugs
- Detection error was large
- Motor was scanned, scanned model was used to generate synthetic data set
- Achieved good detection results
- Texture supports the detection results



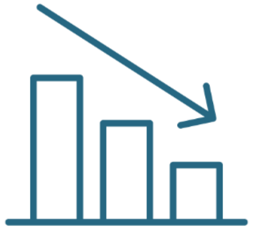




D2.4 ODIN Core Enabling technologies for perception enabled reconfigurable resources https://www.odin-h2020.eu/D2.4.pdf

Advantages of Using Synthetic Data

HOW SYNTHETIC DATA REDUCES COMPLEXITY OF A ROBOT VISION PROJECT



1. Robustness:
    - Data set can cover all variations (environment, object, application)
    - Validation against ground truth data
    - Lighting and material can be changed in simulation
    - Data can be adapted (partials and scanns)
2. Usability:
    - Only CAD model and some basic information required
    - No production downtime for real data recording or knowlegde for labelling data required
    - No expert knowledge needed to parametrize the vision solution
3. Effort and cost:
    - Feasibility study in simulation
    - Effortless commissioning on site
    - Enables efficient remote support

Contact

Dr. Michael Suppa

CEO & Co-Founder

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

Round Table Discussion

From Good Data to Ease of Use

#1

GOOD DATA INSTEAD OF BIG DATA REDUCES ONSITE TRAINING TIME

- Simulation helps create realistic training data using model-knowledge
- Ground truth can be used in the training
- Enrichment with real data images instead of complete data recording process
- Results in accuracy in mm and not detection rates in percent

#2

SCALABLE ML SOFTWARE PLATFORM FOR PLUG-AND-PRODUCE

- Share ressource by deployment concept
- Allow integrators and end customers to add modules on the same plattform
- Smart Sensors allow for distribution of computing ressources

#3

USING ML TO ENSURE EASE-OF-USE FOR NON-VISION EXPERTS

- ML reduces the parameter space for the customer
- Easy onsite optimization
- Web Interfaces with wizards allow for non-expert use

slido

Which level of expertise regarding 3D vision and machine learning is available in your area? (none, beginner, moderate, expert)

ⓘ Start presenting to display the poll results on this slide.

Model data of the final product is usually available in agile production. How do you deal with the potential lack of knowledge during the production process when object handling is required at intermediate steps?

ⓘ Start presenting to display the poll results on this slide.

slido

Synthetic training data generation requires model data for simulation. Do you have this data available, where does it come from and how do you asses the level of correctness?

ⓘ Start presenting to display the poll results on this slide.

Slides will be published on the website:
https://roboception.com/workshop-at-erf-2024/

Interest in participating in TG Perception:
michael.suppa@roboception.de

and/or registration at
https://www.robotics-portal.eu