Driving competitive advantage with computer vision

Automated quality control for manual assembly

Overview

Jonas Conrad jonas@sentinus.ch +41 79 520 15 88
Felix Schnarrenberger felix@sentinus.ch +41 79 838 56 55

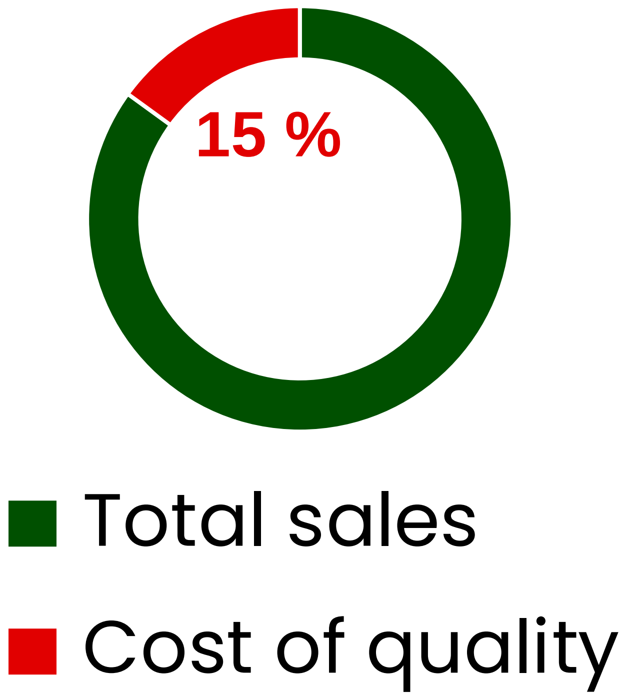
01 Sense of urgency

Manual manufacturing offers high optimization potential:

→Cost of quality

→Skilled worker shortage & nearshoring

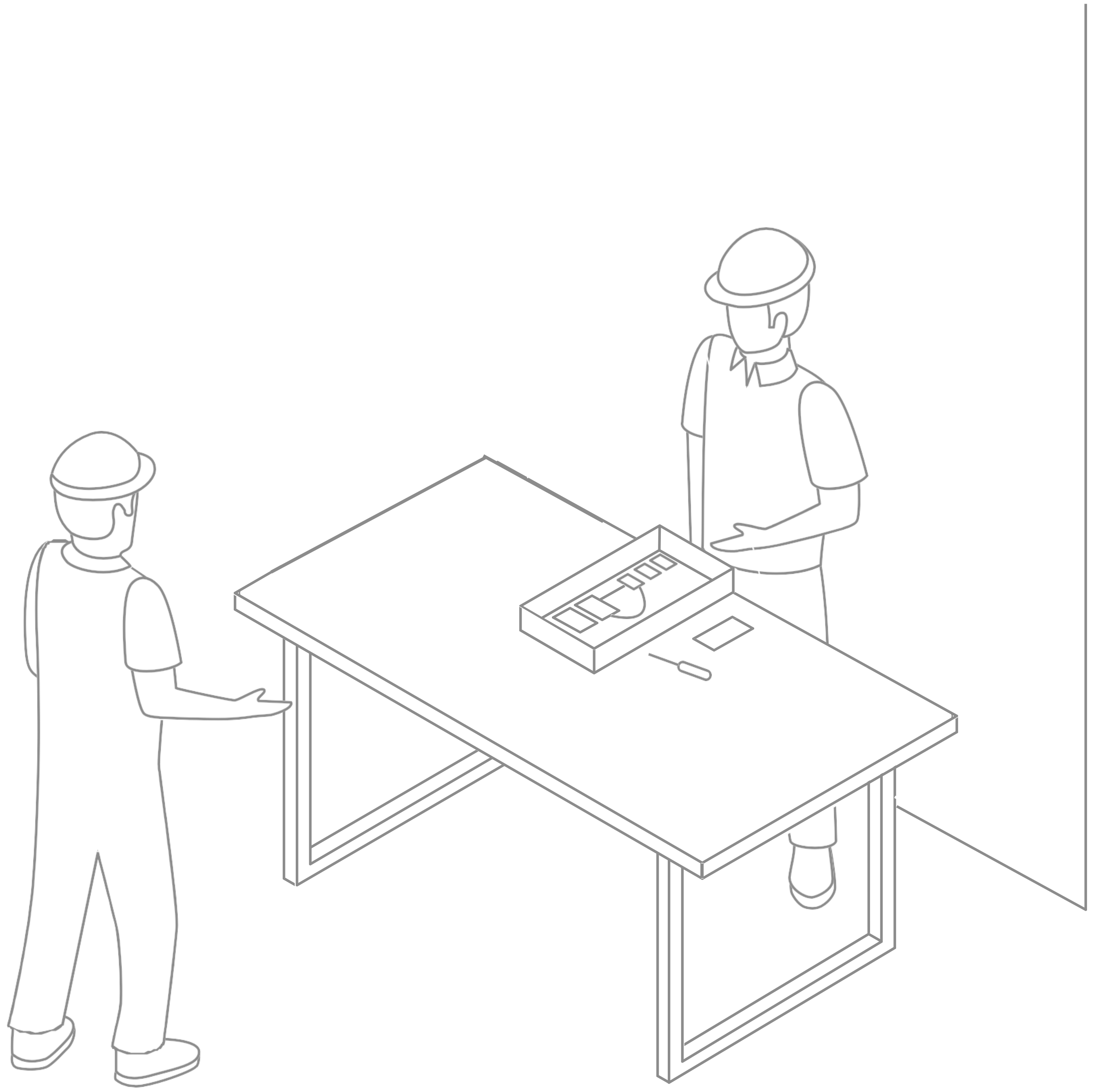
→Continuous process optimization



02 Solution

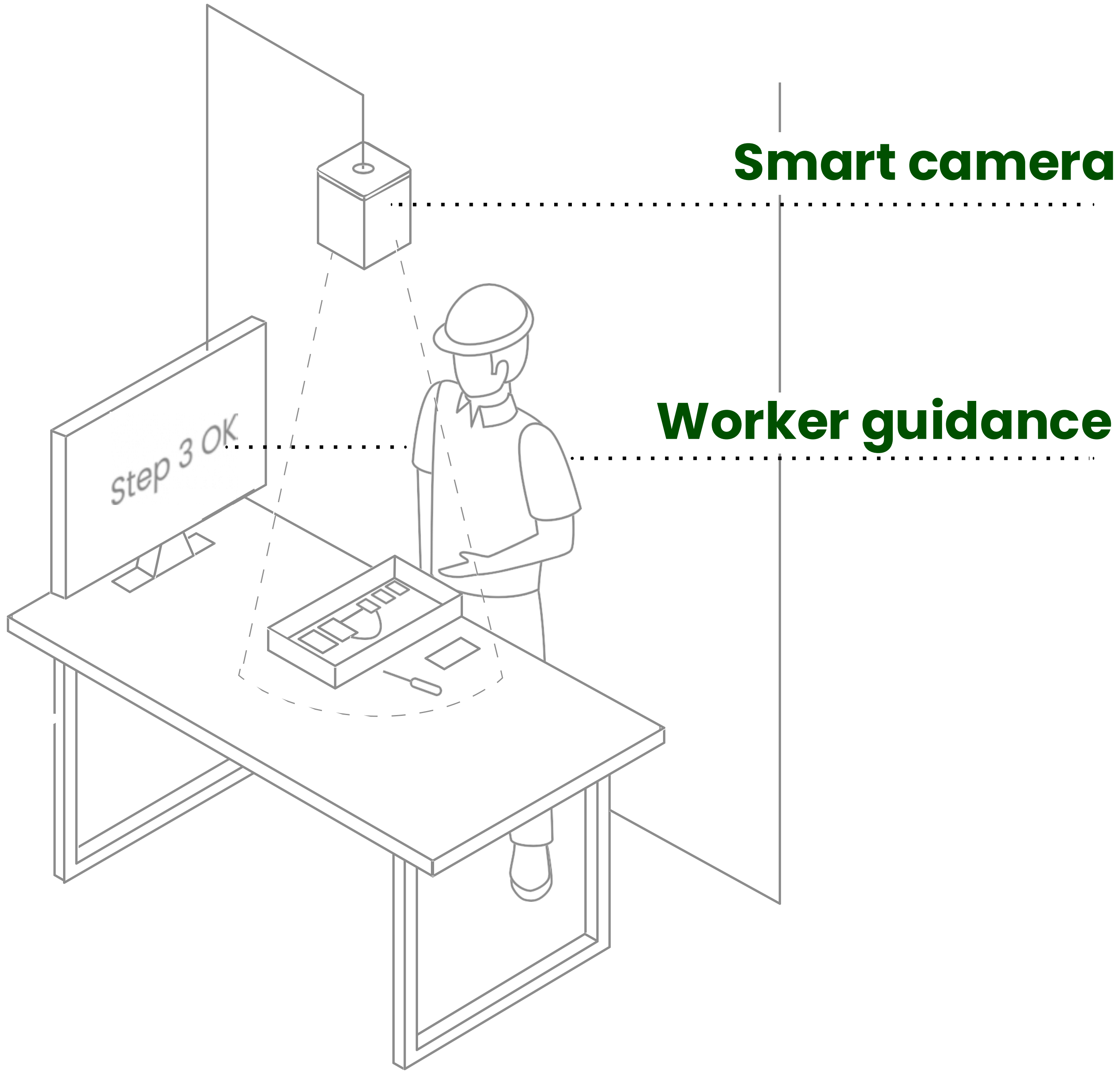
We digitize the four-eye-principle.

Four-eye-principle



Manual quality assurance

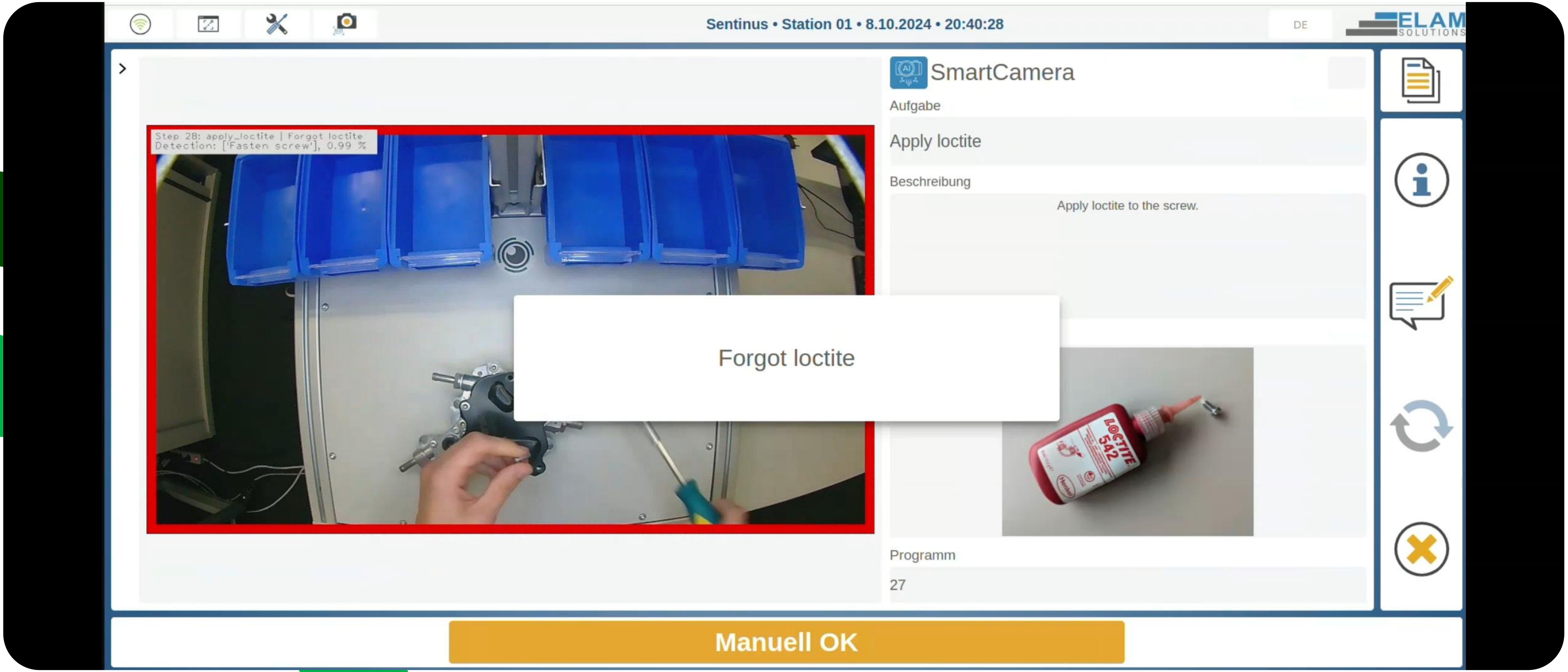
sentinus



Automated → **Evaluation**
→ **Feedback & instructions**
→ **Documentation**

04 Technology

Computer vision for assembly step recognition.



04 Technology

We understand the whole assembly process.

Other solutions:
→ Automation of single end-of-line checks

→ State confirmation only

→ Focus on automated / low-mix, high-volume production: **Lack of scalability**

→ **Fails in case of high variation**, e.g., variant task execution, lighting situation, …

State confirmation

Our solution:
→ Applicable for the whole assembly process

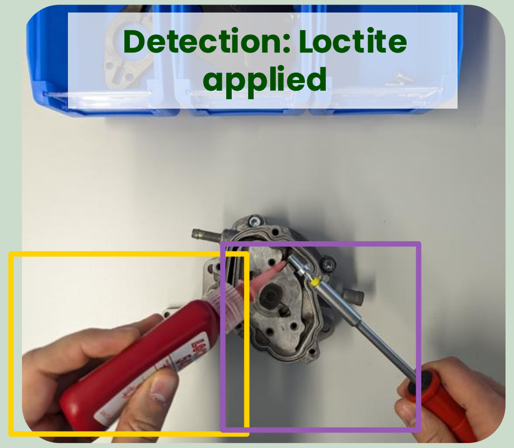
→ State confirmation + action recognition

→ Simple setup / adaption **manual for high-mix, low-volume production**

→ **Robust against variation**

State confirmation

Action recognition



05 Roadmap

We work towards 100% in-process control starting at lot size one.

2020: Applied research

ETH zürich inspire

2025: Automated quality control
→ Automated assembly step recognition
→ Live feedback, instructions, documentation

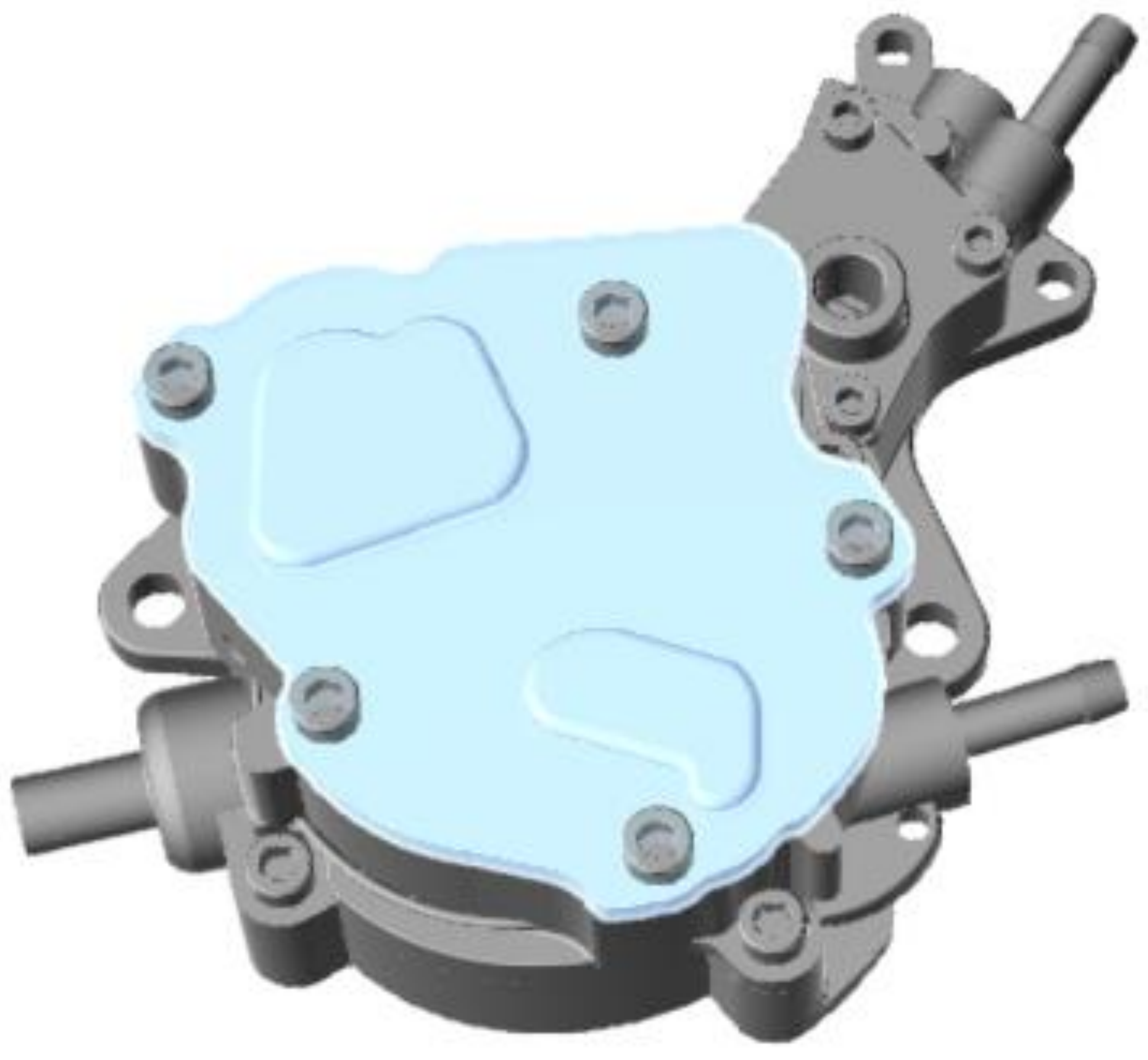
BOSSARD HUBER+SUHNER

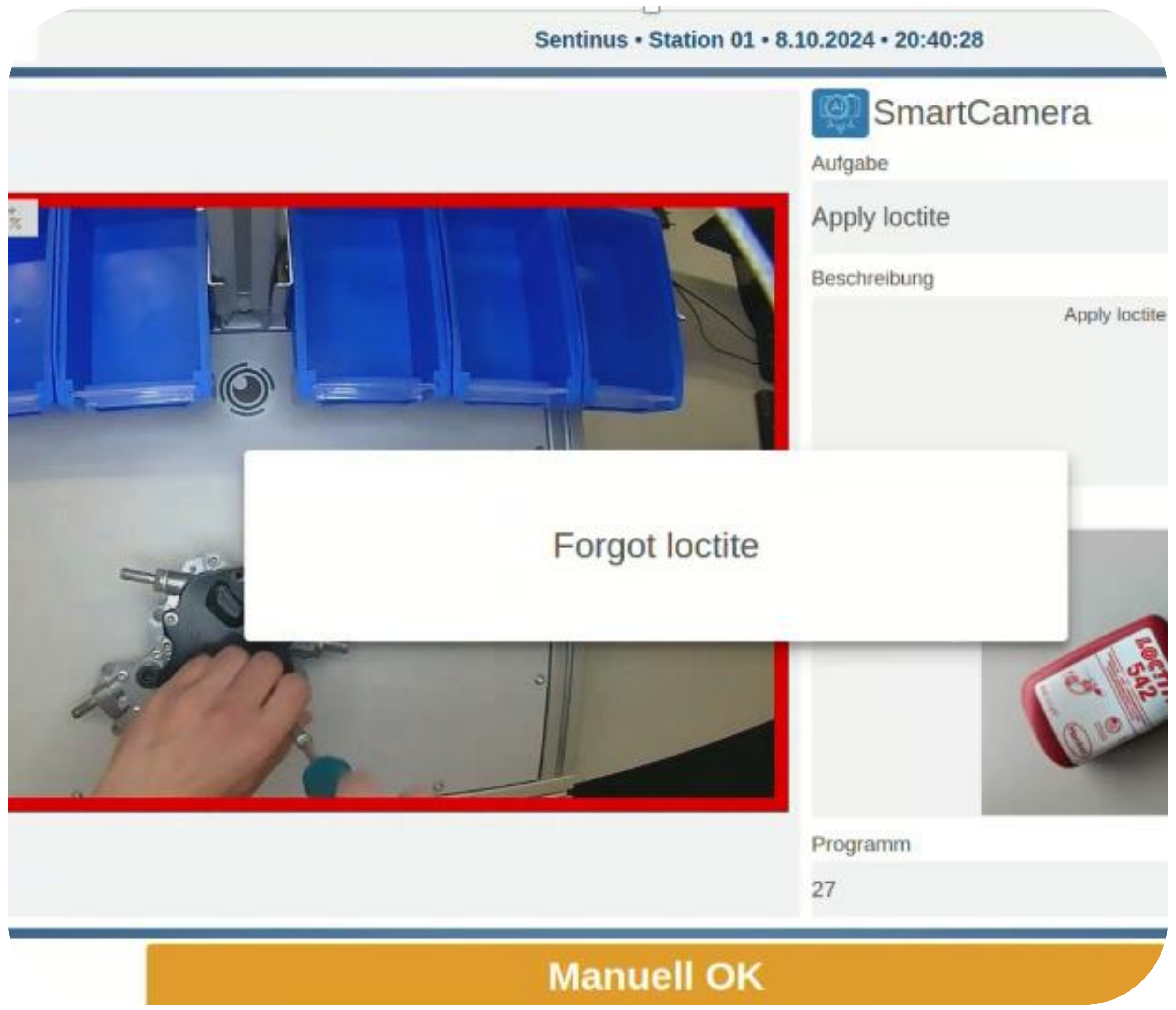
2027: Plug & play solution
→ Plug & play based on general dataset
→ Fully automated training pipeline using synthetic data

2029: 100% in-process control starting at lot size one

06 Technology vision

Tackling the hurdle of training data generation with synthetic training data.





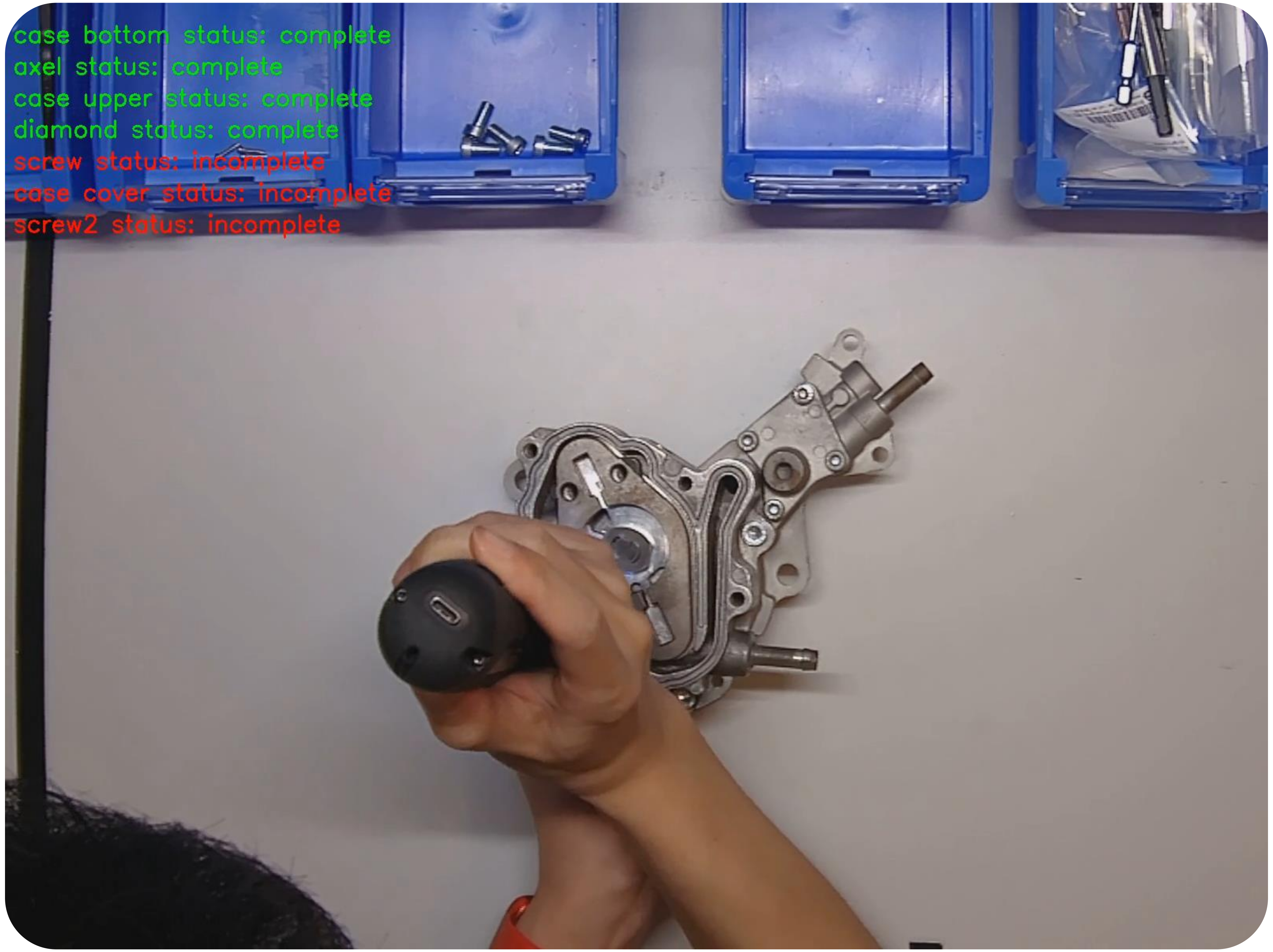
Digital part models

Generation of synthetic images & model training

Automated assembly step recognition

06 Technology vision

Tackling the hurdle of training data generation with synthetic training data.



Applied research

ETH zürich

inspire

06 Team

We are experts in computer vision applications and business strategy.

Jonas Conrad

→ MSc ETH Mechanical Engineering
→ PhD Deep learning in manual assembly (WIP)

RESPONSIBILITIES
→ **Executive lead**

Co-founder

Felix Schnarrenberger

→ MSc ETH Robotics, Systems and Control
→ Applied machine learning and computer vision

RESPONSIBILITIES
→ **Technical lead**

Co-founder

Constantin Herbst

→ MSc ETH Mechanical Engineering
→ Data Scientist CV

RESPONSIBILITIES
→ **Operational lead**

Co-founder

Christoph Conrad

→ Former Head Marketing "Solution & Service Portfolio" at Siemens AG

RESPONSIBILITIES
→ **Business development**

David Filliberti

→ MSc ETH Robotics, Systems and Control
→ Applied ML & CV

RESPONSIBILITIES
→ **Embedded systems**

Wsewolod Dubinski

→ MSc EBS Management
→ Strategy & Finances

RESPONSIBILITIES
→ **Strategy & Finance**

07 DSL Challenge

Automated anomaly detection in manual assembly based on synthetic data

→ Challenge Giver: Constantin Herbst, Jonas Conrad, Sentinus AG
→ Academic Coach:
→ Contact Details: constantin@sentinus.ch, jonas@sentinus.ch

Overview:

→ **Aim**: Perform robust anomaly detection in the manufacturing process trained on synthetic data
→ **Goal**: A pipeline that recognizes anomalies in single parts and wrongly assembled subassemblies based on CAD data of assembly group in real time.
→ **Tools**: python, blender, pytorch
→ **Focus**:
Robustness: needs to work on real data under realistic conditions
Data-Efficiency: no need to record data to prepare inspection for assembly
Semantic Scene Understanding: Correct segmentation of parts / subassemblies
→ **Support:**
Recording extra datasets for specific anomalies
Real-time testing with sentinus smart camera hardware
Guidance from Sentinus team
→ **Literature: Entry points to computer vision model training based on synthetic data**

https://www.mdpi.com/2076-3417/13/22/12316
https://www.sciencedirect.com/science/article/pii/S2212827124006668

07 DSL Challenge

Dataset:
→ Sources: MA-3 dataset
→ Content: 60 videos of manual assembly of different products; 20 operators. Focus on vacuum pump. The dataset can be enhanced by Sentinus.
→ Pickup SSD at Technopark

Challenge Components:

→ Build a CV pipeline that

1. Based on CAD data of each part / assembly stage creates a rendered dataset to train models for the following step
2. Segments all relevant assembly parts in a live video stream in real time
3. Performs anomaly detection on single parts and sub-assemblies
4. Anomaly detection on subassembly: wrong assembly!
5. Rejects frames in which no inspection can be performed reliably due to occlusion, motion blur etc,

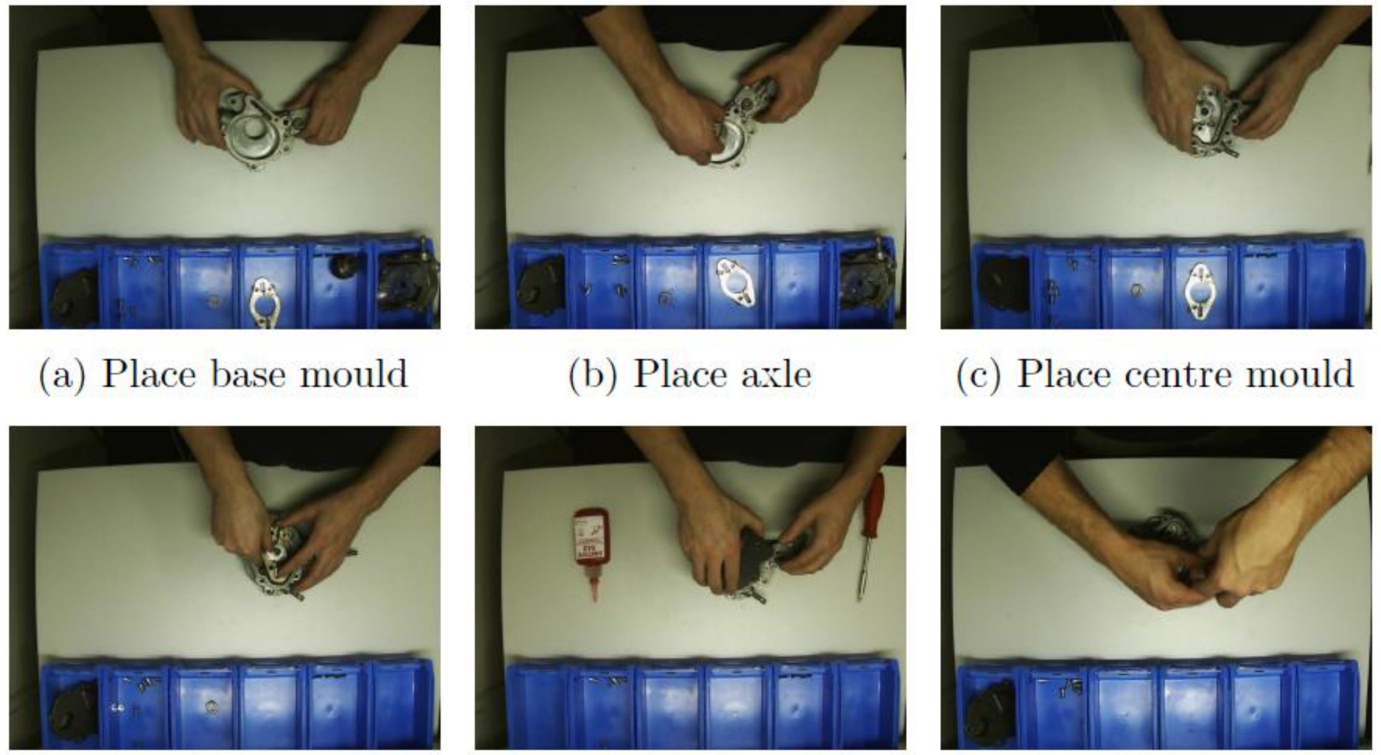
07 DSL Challenge

Dataset:
→ Sources: MA-3 dataset
→ Content: 60 videos of manual assembly of different products; 20 operators. Focus on vacuum pump. The dataset can be enhanced by Sentinus.
→ Pickup SSD at Technopark

Challenge Components:

→ Build a CV pipeline that

1. Based on CAD data of each part / assembly stage creates a rendered dataset to train models for the following step
2. Segments all relevant assembly parts in a live video stream in real time
3. Performs anomaly detection on single parts and sub-assemblies
4. Anomaly detection on subassembly: wrong assembly!
5. Rejects frames in which no inspection can be performed reliably due to occlusion, motion blur etc,



(a) Place base mould (b) Place axle (c) Place centre mould
(d) Place diamond plate (e) Place cover (f) Screw first location

07 DSL Challenge

Evaluation Metrics:

1. Accuracy of recognized assembly status
2. Accuracy of segmentation (mIOU)
3. Detection rate of anomalies (screw missing, flipped part) (AUROC/PRO)

Successful Project:

→ Can discern assembly status on real data

→ Rejects obvious disturbances / occlusions (e.g. hand covers assembly group

→ Can pinpoint anomalies robustly: rotation, position, illumination invariant, robust against slight variation in part appearance

→ Well-documented and reasonable code

SotA: Datasets exist, but do not accurately represent real application scenarios

Data set	Visual sensors	Assembly task	Reflective components	Cabling tasks	Symmetric base	Fine-grained labelling
Assembly101[15]	Eight RGB cameras (static mount), four monochrome cameras (head mount)	Toy cars	No	No	No	Yes
Wooden box[16]	Two RGB cameras (static mount)	Wooden box	No (screws)	No	Yes	No
MECANO [17,18]	RealSense300 (RGB + depth, head mount), Pupil Core (gaze signals, head mount)	Toy motorbike	No	No	No	Yes
IKEA ASM[19]	Three Azure Kinect V2 (RGB + depth, static mount)	Ikea furniture	No	No	Yes	Yes
IKEA FA[20]	One RGB camera (static mount)	Ikea furniture	No	No	Yes	No
IKEA Ego 3D[21]	HoloLens 2 (RGB + point cloud, head mount)	Ikea furniture	No	No	Yes	Yes
HA4M[22]	One Microsoft Azure Kinect (RGB + depth + IR + point cloud, static mount)	Generic assembly tasks	No	No	Yes	Yes
HA-VID[23]	Three Azure Kinect (RGB + depth, static mount)	Generic assembly tasks	No (screws, nuts)	No	Yes	Yes
MA-3 dataset (ours)	Two RGB cameras, 1 depth camera (all static mount)	Three manually assembled products	Yes	Yes	Yes	Yes

Contribution: Dataset including specific application hurdles for computer vision approaches in manual assembly + benchmark

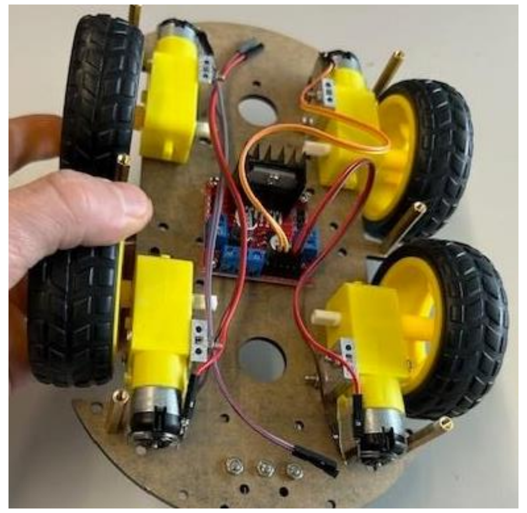
inspire

Use cases

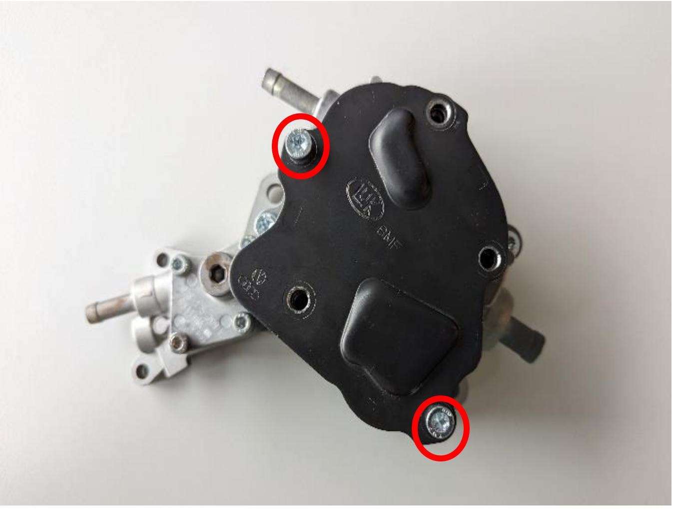
Challenges
- High operator-induced variance in task execution
- Reflective components
- Cabling tasks
- Symmetric parts (base plate)
- Fine-grained task distinction (e.g., screw coordinates)

Belt sander


- Task execution variance
- Fine-graínde task distinction

Robot car


- Task execution variance
- Fine-grainde task distinction
- Symmetric parts
- Cabling task

Vaccum pump


- Task execution variance
- Fine-grainde task distinction
- Reflective parts

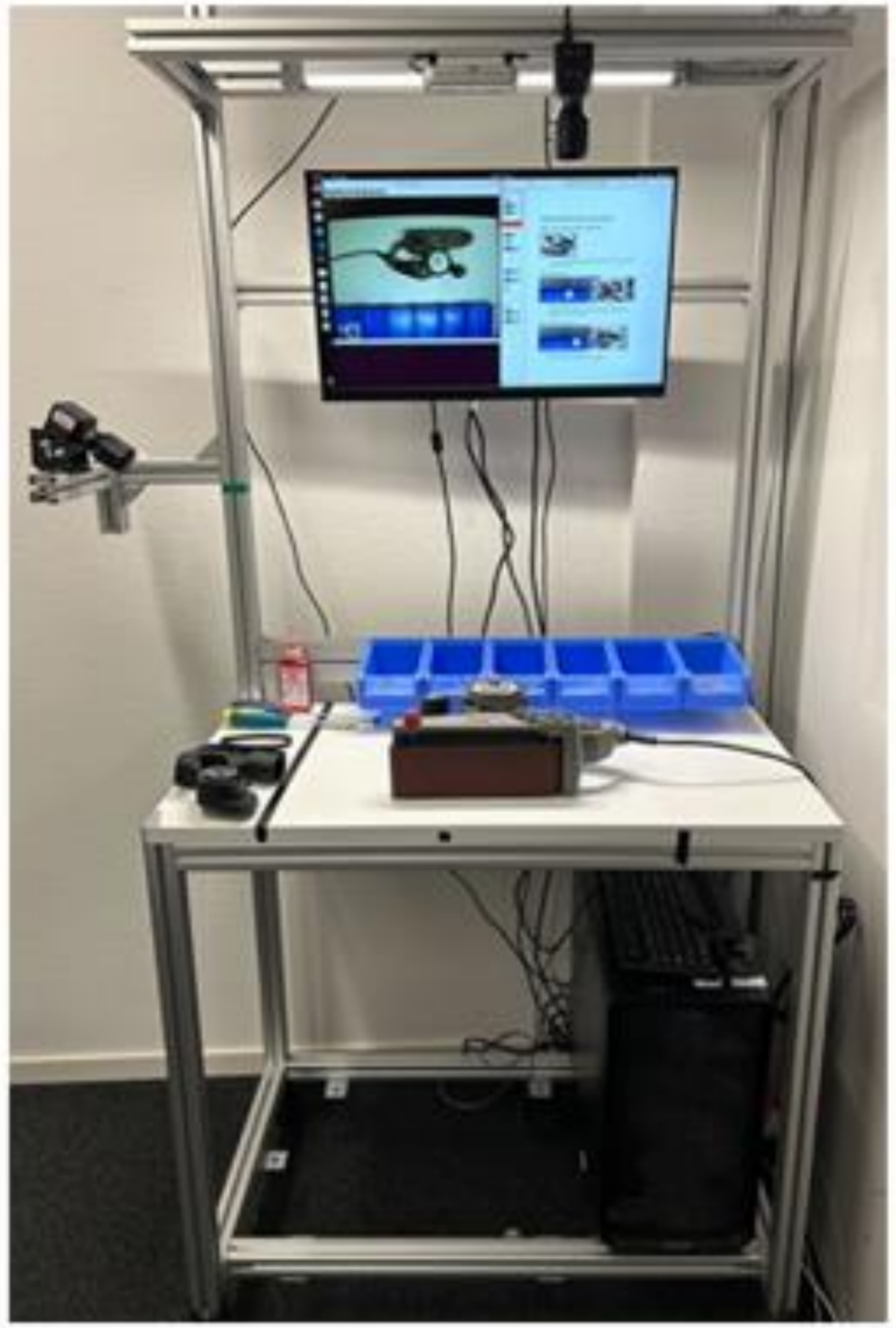
Methods: Dataset collection + benchmark

Data collection

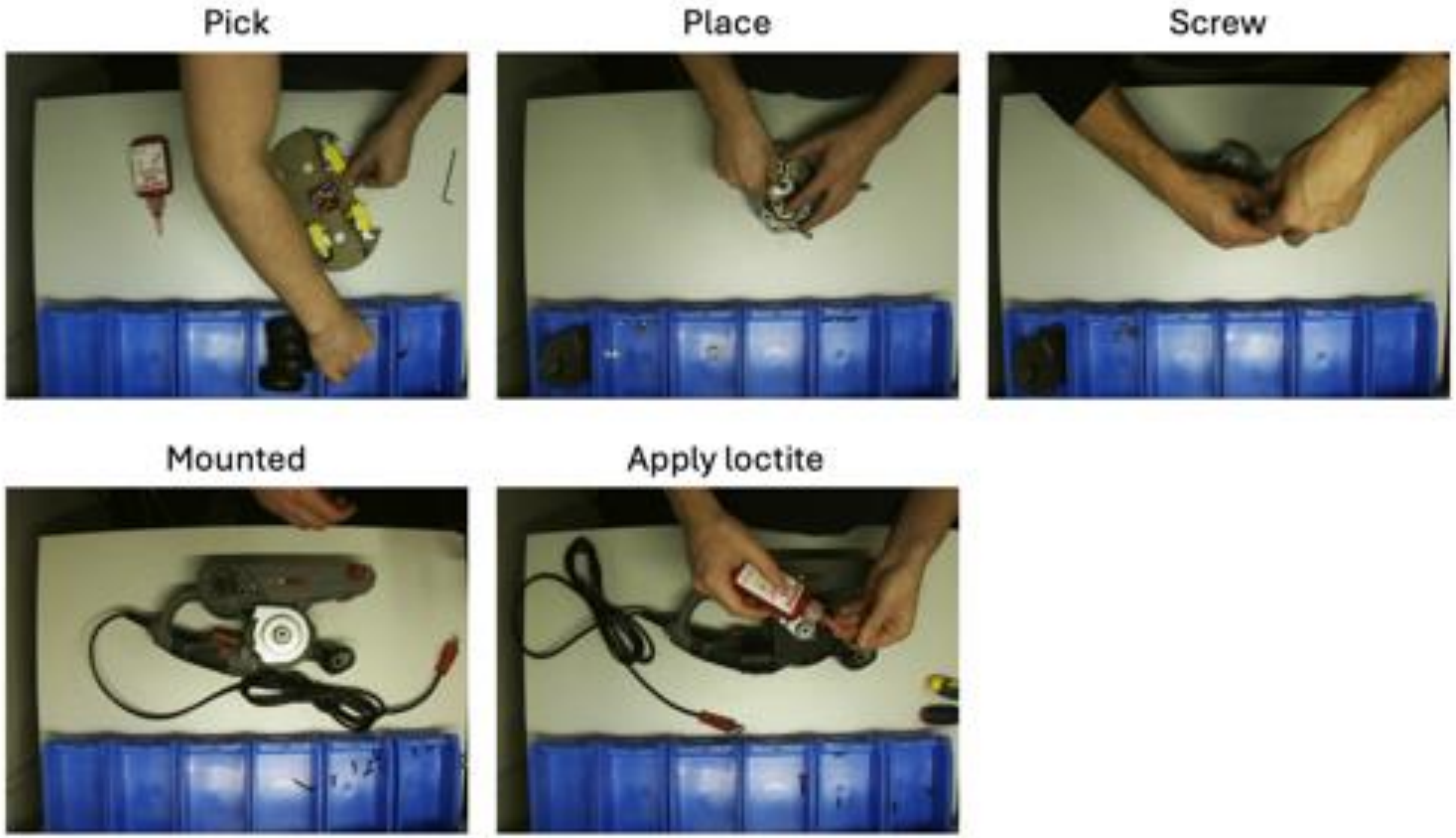
- 2 RGB cameras, 1 depth camera
- 20 participants, 3 recordings per use case
- Labelling with ground truth (79 distinct classes)
  - Action class + specifier, e.g., screw_screw1

Benchmark

- 4 Computer vision approaches
- F1-scores
- Reverse Leave-one-out



Action class overview



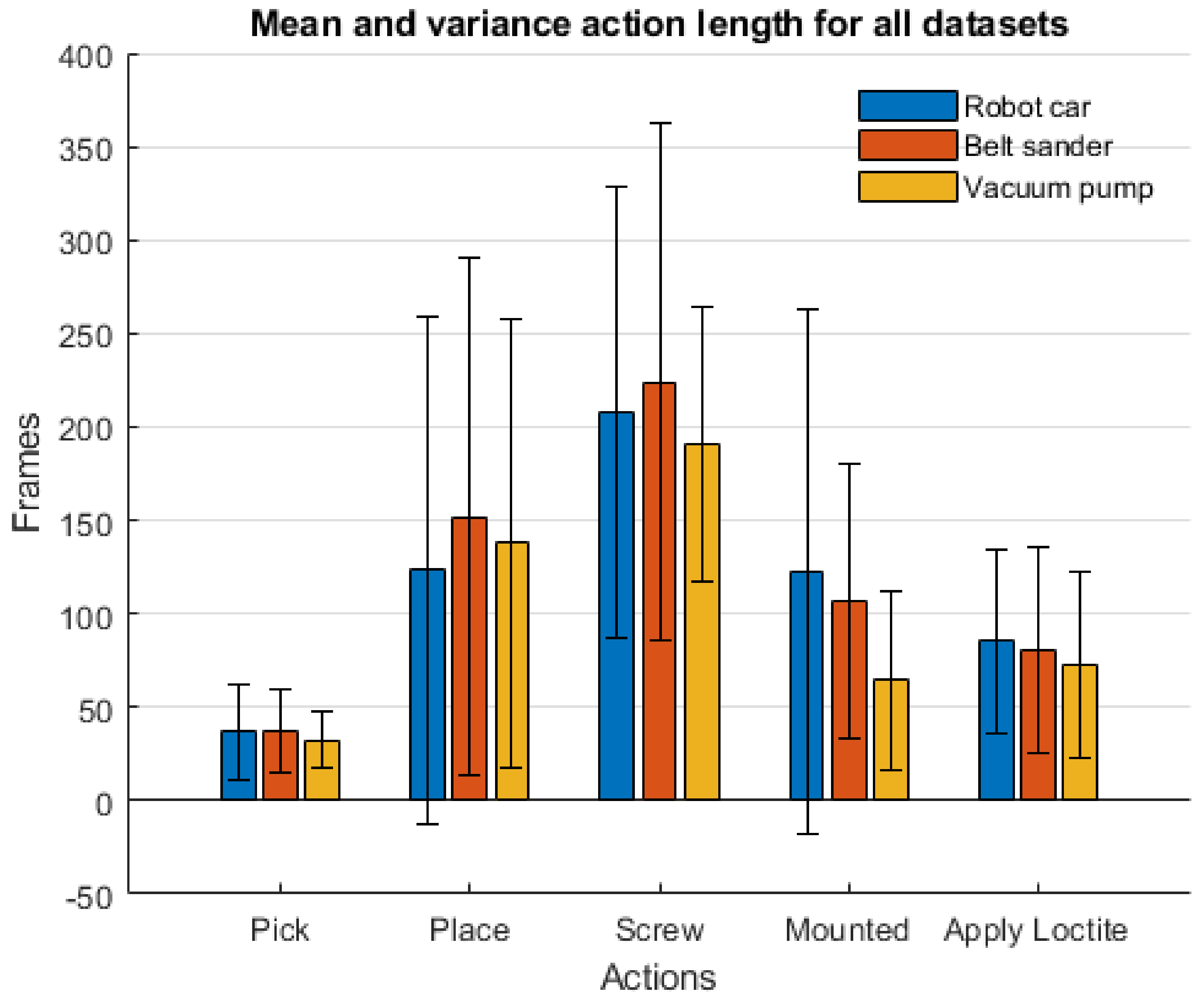
Technical validation

Dataset overview

- 18 hours, 9 minutes of assembly activity
- Recorded in 540 videos

Benchmark

- Overall performance
- Confusion matrix for one computer vision approach
- Results Reverse Leave-One-Out for one approach



Technical validation

Action	Instances	Mean Length (frames)	Variance (frames)	Min Length (frames)	Max Length (frames)
pick	1732	35.59	21.73	10	238
place	1286	134.43	133.01	10	1556
screw	961	208.65	116.6	12	904
mount	4015	100.26	104.15	9	1072
apply_loctite	801	79.15	51.78	10	339

Technical validation

F1-score of various models over all datasets				
Models	**Average**	**Robot car**	**Belt sander**	**Vacuum Pump**
YOLOv8m-cls	0.731	0.610	0.801	0.781
EfficientNetB4	0.796	0.677	0.885	0.825
UniFormerV2	0.795	0.656	0.905	0.824
CSN-r50	0.808	0.656	0.909	0.858
CSN-r152	0.790	0.625	0.888	0.858

08 Next steps

→ Get familia with dataset and theory
→ Define focus areas
→ Reach out if further inputs are needed (best via mail)

Thank you!

Please get in touch with us. Any day, any time!
+41 44 556 58 76

Reach us via mail to schedule a meeting:
jonas@sentinus.ch

Supported by: