Automated Benchmarks and Optimization of Perception Tasks

Maximilian Durner*, Zoltán-Csaba Márton*, Simon Kriegel, Manuel Brucker,
Sebastian Riedel, Dominik Meinzer and Rudolph Triebel

Deutsches Zentrum für Luft- und Raumfahrt
German Aerospace Center

Institute of Robotics and Mechatronics,
German Aerospace Center, Germany
*The authors contributed equally to this work

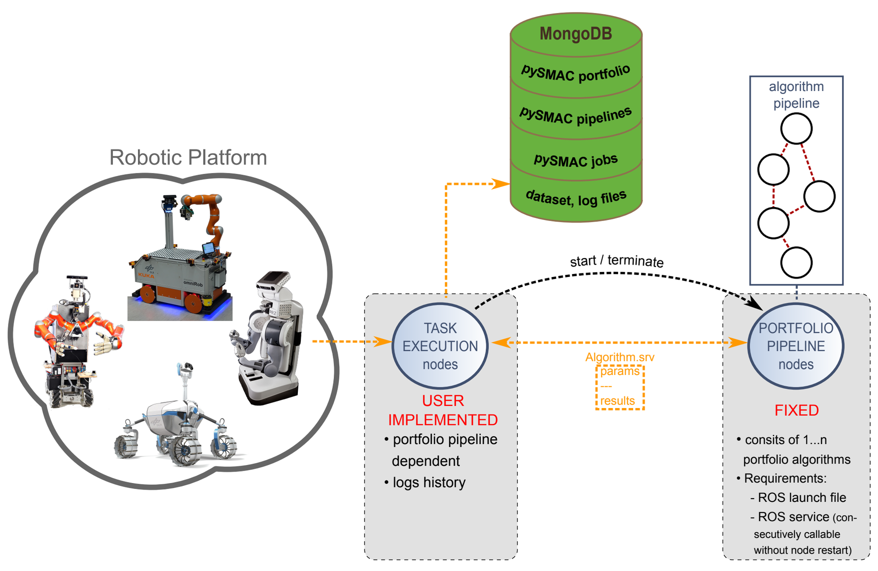
Problem Description

Optimizing task-execution performance of a robotic system is complex. A robotic expert is required to find poorly performed task steps, analyze reason for poor behavior and identify solutions to handle the identified problems.

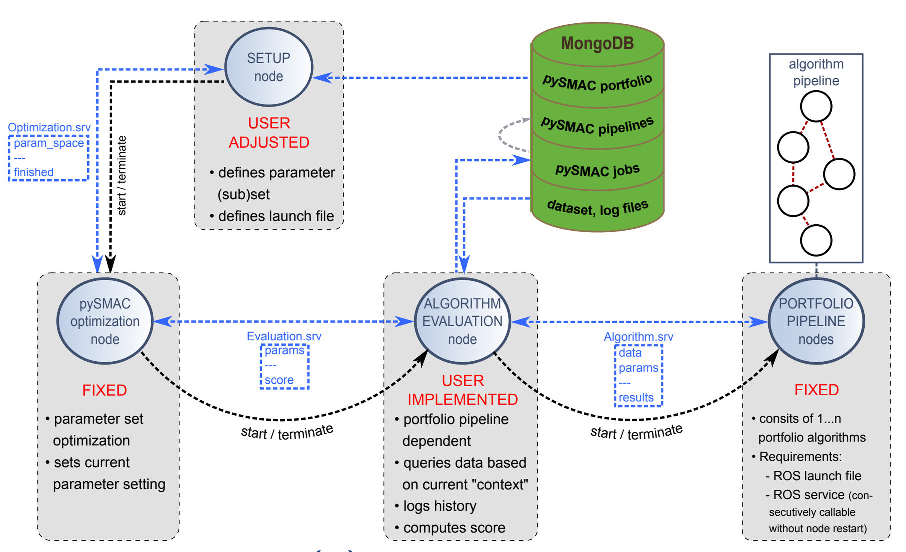
> *"In the context of manufacturing, the greatest potential is for functions that contribute to a reduction of programming and configuration requirements in deployed systems. There are clear benefits for small lot size systems in reducing the time and skill needed to reconfigure an adapt systems to new processes."*
>
> EU's Robotics 2020 Multi-Annual Roadmap [1]

We propose a Pipeline Optimization Framework (POF), which allows robots to improve its perception performance utilizing logged experiences and thus continuously improving their performance based on introspection.

Pipeline Optimization Framework



(a) Execution Phase



(b) Optimization Phase

Figure 1: Flow chart of the POF evaluation.

The general workflow surrounding the POF can be described as follows and and is illustrated in Fig. 1. During the *Execution Phase* (Fig. 1a), we create an execution log containing raw sensory data, intermediate processing results and contextual information. These are saved in a MongoDB object-orientated database [2]. After the Execution Phase, the logged data is reprocessed during the *Optimization Phase* (Fig. 1b), with the goal to improve the perception performance for further task executions. To enable the optimization, a ground truth of the object scenes needs to be given a priori.

Dataset

For our evaluation, we use the THR Dataset [3] which contains both individual objects and scenes with elements standing freely, and also mounted on a rail (Fig. 2).

Additionally high quality models of the objects using a hand-guided scanning system are acquired.

The dataset is publicly available:

`http://www.dlr.de/rmc/rm/thr-dataset`

Evaluation

We evaluate the POF on two of the most common visual perception problems, namely classification and pose estimation.

For the classification task two alternative classifiers are used: Linear SVM (linSVM), and Random Forests (RF). As input features we use deep-learned features (ResNet50 Network [4]) extracted from the RGB images.

For estimating the pose of known objects in a depth image, we use the method described in [5].

Two separated contexts are defined and evaluated separately: *unmounted object scenes* and *mounted object scenes*.

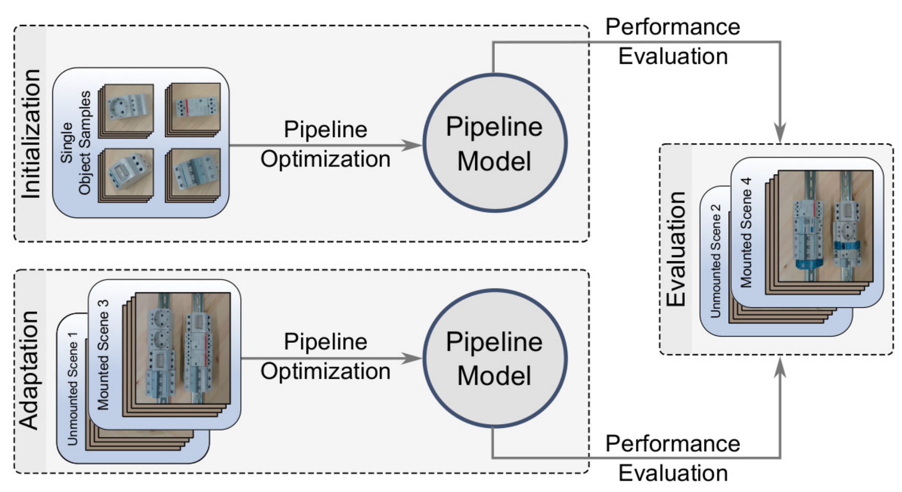


Figure 2: Flowchart of the POF evaluation

Data splitting:

- Initialization: Single Object Samples
- Adaptation: Scene with multiple instances
- Evaluation: Scene with single object occurrence

Results

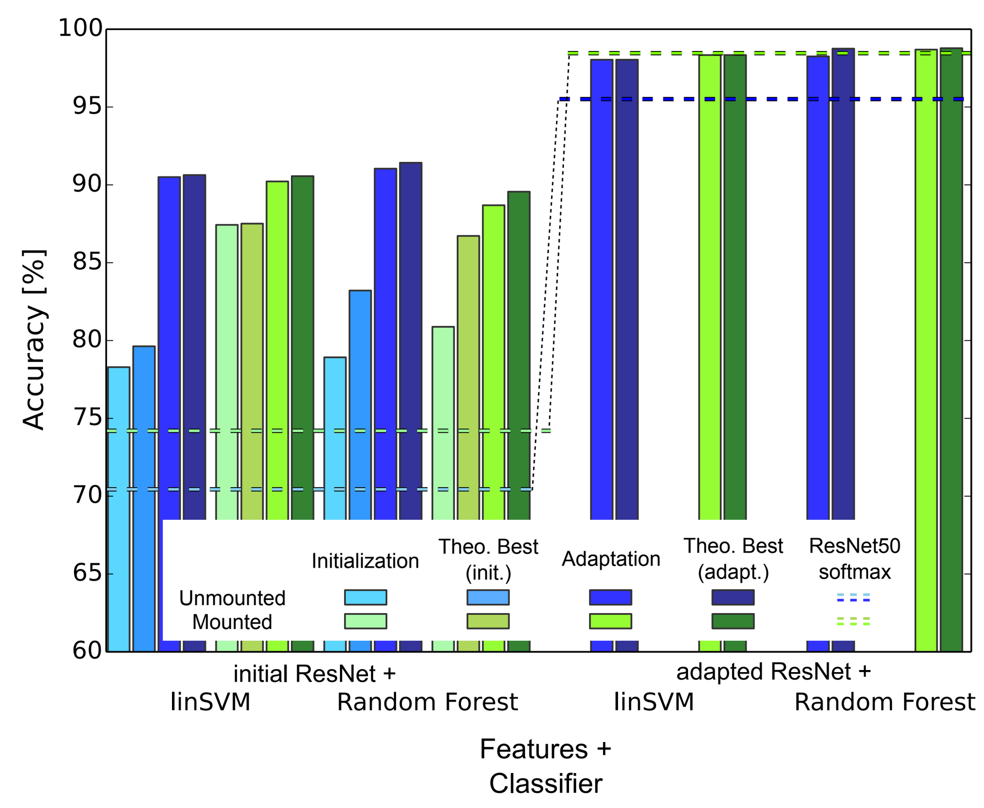


Figure 3: Results for classification task on all 9 objects

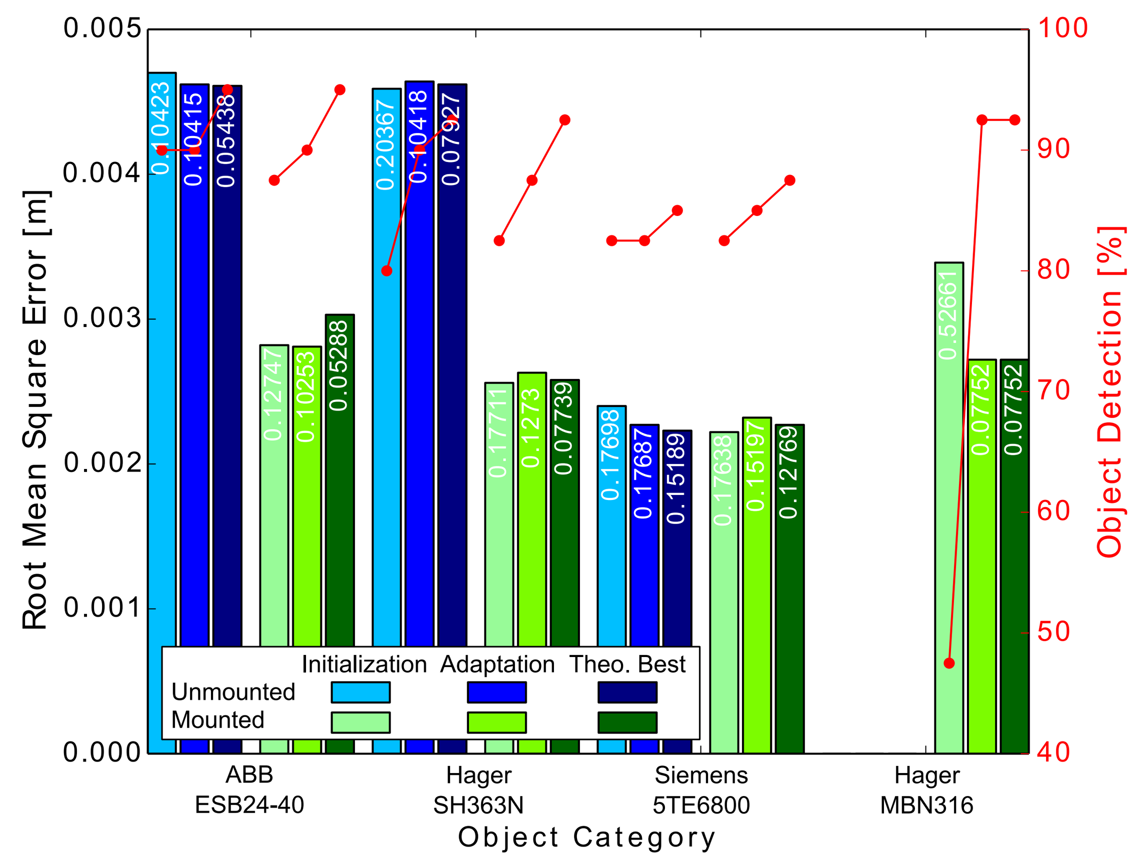


Figure 4: Results for pose estimation on 4 objects; **bar plot**: averaged root mean square error [mm] of *detected objects*; **white error values**: average root mean square error [mm] over *all* samples; **right axis**: object detection rate

Extension: Active Learning

In collaboration with the Institute for Artificial Intelligence, University of Bremen [6]: Using automatically labeled data for adaptation.

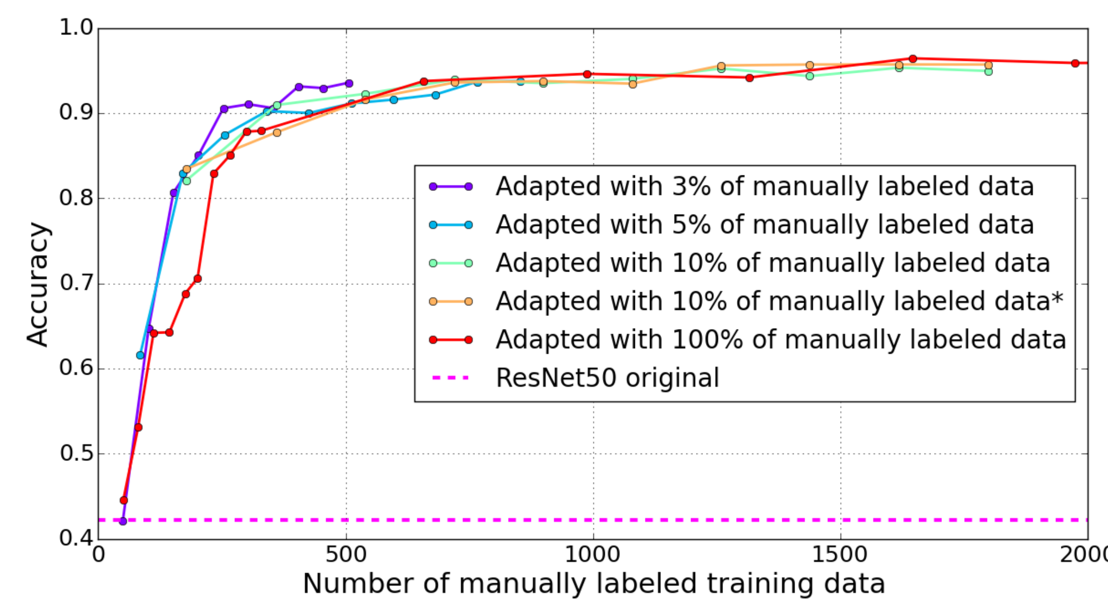


Figure 5: Accuracy of original classifier versus the adapted one, trained with different amounts of manually labeled training data. In the cases where some of the data was automatically labeled, only the most confident, thus correctly labeled, detections were used.

References

[1] European Union, "Robotics 2020 multi-annual roadmap," 2017. [Online]. Available: www.eu-robotics.net/cms/upload/downloads/ppp-documents/Multi-Annual_Roadmap2020_ICT-24_Rev_B_full.pdf

[2] F. Balint-Benczedi, Z.-C. Marton, M. Durner, and M. Beetz, "Storing and retrieving perceptual episodic memories for long-term manipulation tasks," in *International Conference on Advanced Robotics (ICAR)*. IEEE, 2017, Best Paper Finalist.

[3] M. Durner, S. Kriegel, S. Riedel, M. Brucker, Z.-C. Marton, F. Balint-Benczedi, and R. Triebel, "Experience-based Optimization of Robotic Perception," in *IEEE International Conference on Advanced Robotics (ICAR)*, Hong-Kong, China, July 2017, Best Paper Finalist.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.

[5] S. Kriegel, M. Brucker, Z.-C. Marton, T. Bodenmüller, and M. Suppa, "Combining object modeling and recognition for active scene exploration," in *IEEE/RSJ IROS*, Tokyo, Japan, Nov. 2013, pp. 2384–2391.

[6] F. Bálint-Benczédi, M. Durner, Z.-C. Márton, S. Kriegel, M. Brucker, T.-S. Wang, R. Triebel, and M. Beetz, "Uncertainty-based classifier adaptation for interactive life-long learning," *International Journal of Approximate Reasoning (Special Issue on Uncertainty Management in Machine Learning Applications)*, 2017, under review.

2$^{\text{nd}}$ Workshop on Machine Learning Methods for High-Level Cognitive Capabilities in Robotics
IEEE/RSJ International Conference on Intelligent Robots and Systems
Vancouver, BC, Canada, September 28$^{\text{th}}$, 2017