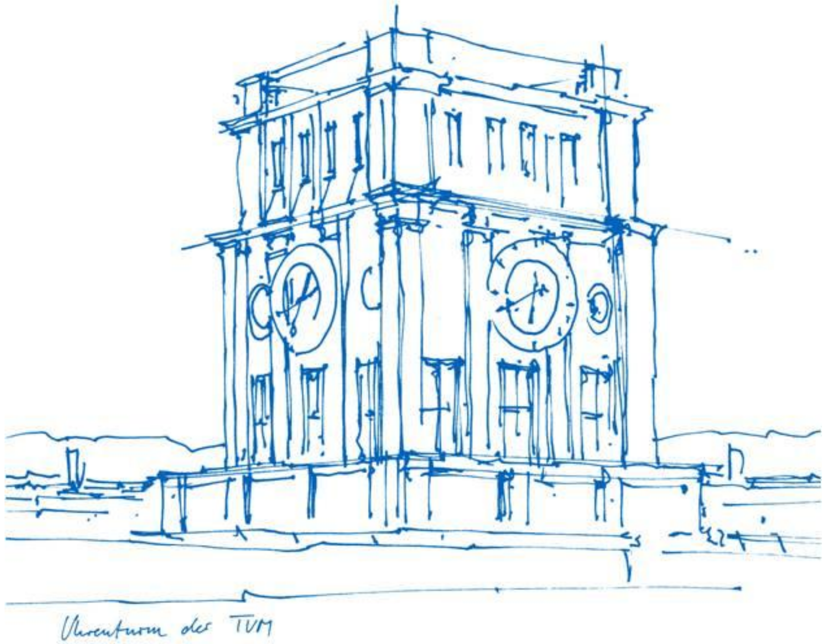
Advanced Deep Learning for Robotics - Final Presentation

Multimodal Sensor Fusion in Differentiable Bayesian Filters



Ali Rabeh, Tim Gerstewitz

10.08.2023

Too big? Push it!

Motivation

Taken from [1]

Pushing is hard though …

- Point of Contact?
- Pushing Angle?
- Material?
- Shape?
- …

Analytical Modelling is difficult

→ Use data-driven model instead

State Estimation with Learned Models

Problem Statement

Goal: estimate pose of pushed object

$$bel(\boldsymbol{x}_t) = p(\boldsymbol{x}_t | \boldsymbol{z}_{1:t}, \boldsymbol{u}_{1:t})$$

https://web.mit.edu/mcube//push-dataset/

Use Structure of Bayes' Filter:

$bel(\boldsymbol{x}_t)$

Learned Forward Model

$p(\boldsymbol{x}_t | \boldsymbol{x}_{t-1}, \boldsymbol{u}_t)$

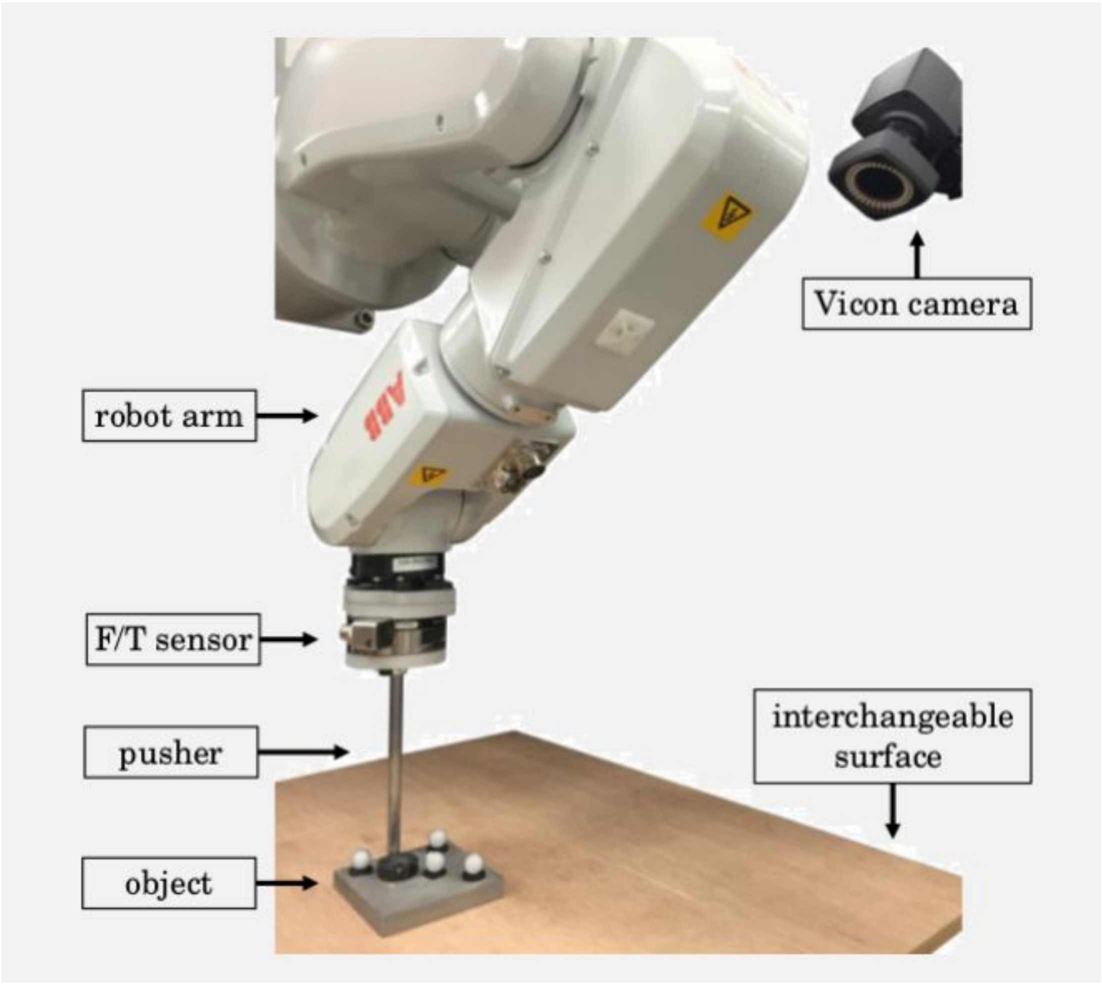
$\overline{bel}(\boldsymbol{x}_t)$

Learned Observation Model

$p(\boldsymbol{z}_t | \boldsymbol{x}_t)$

More than a million ways to be pushed

Experimental Setup



https://web.mit.edu/mcube//push-dataset/

MIT Push dataset:

- 11 objects
- 4 surface materials
- 250 Hz sampling rate

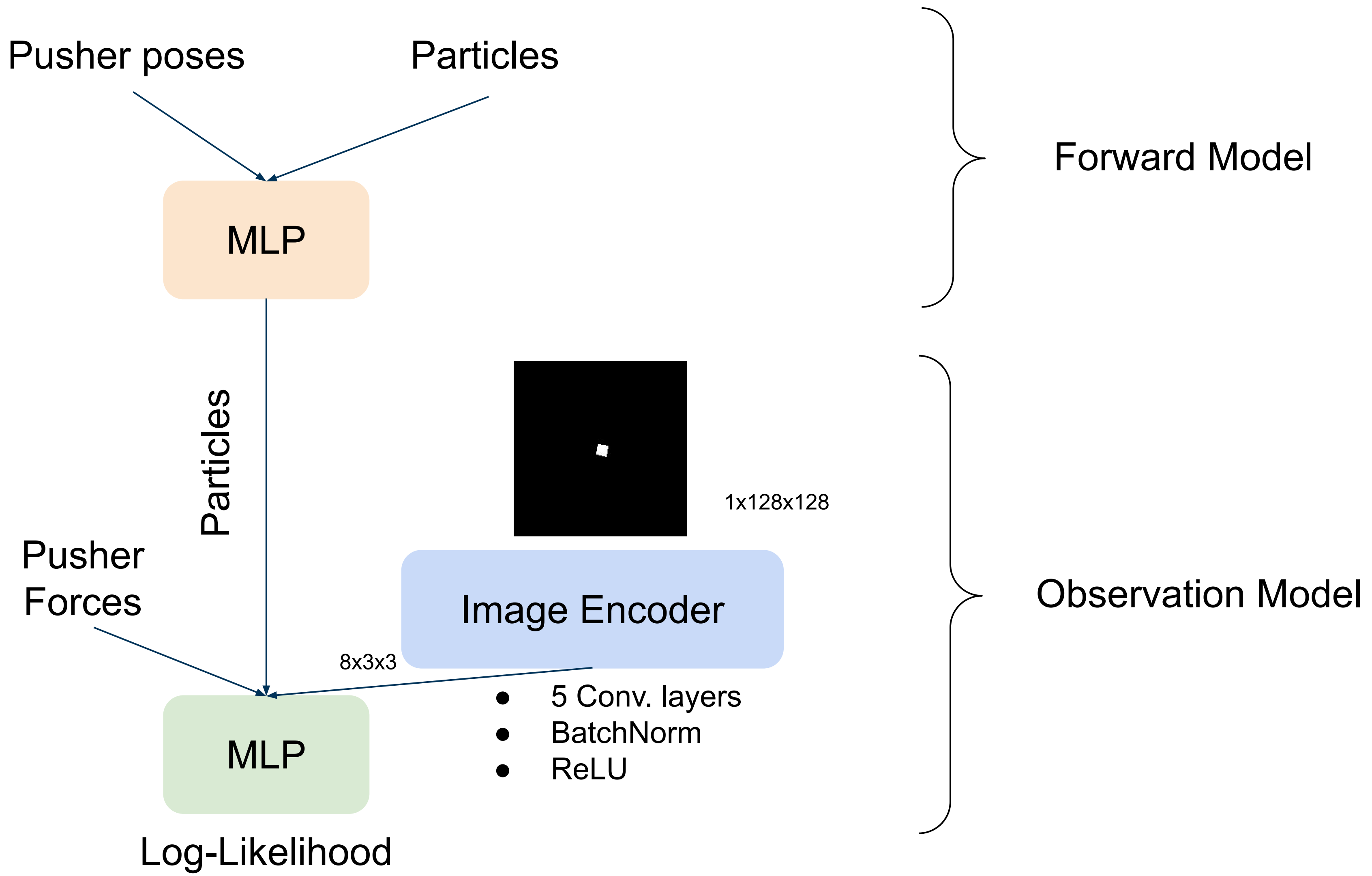
→ more than a million pushes

Use subset for training/testing:

- rectangular shape on plywood
- v = 100 mm/s, a = 0 mm/s^2
- downsample to 50 Hz

Architecture

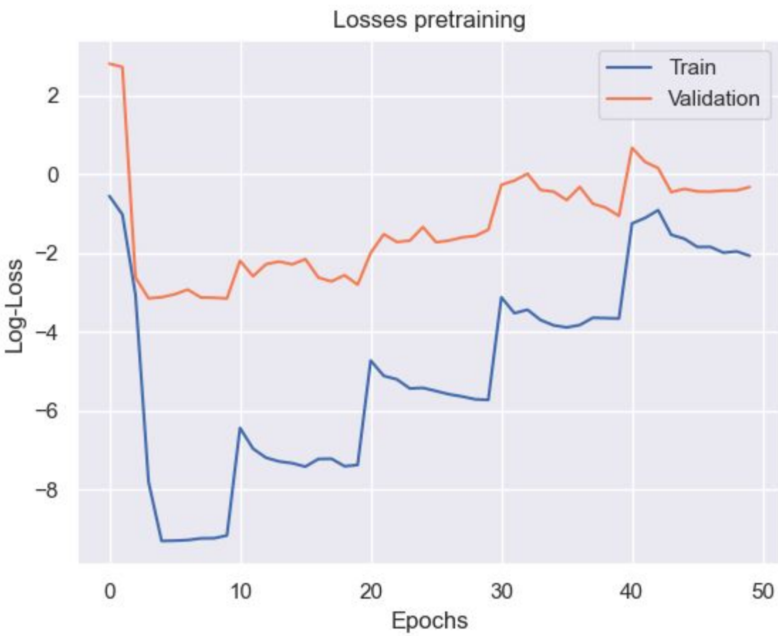
Methods



Training the models

Methods

- Pre-Train forward model
    - Minimize pose error over 1, 2, 4, 8 and 16 prediction steps into the future
    - L2-Loss
- Train forward model and observation model end-to-end
    - Use same sequence of steps
    - Serves to finetune forward model together with observation model

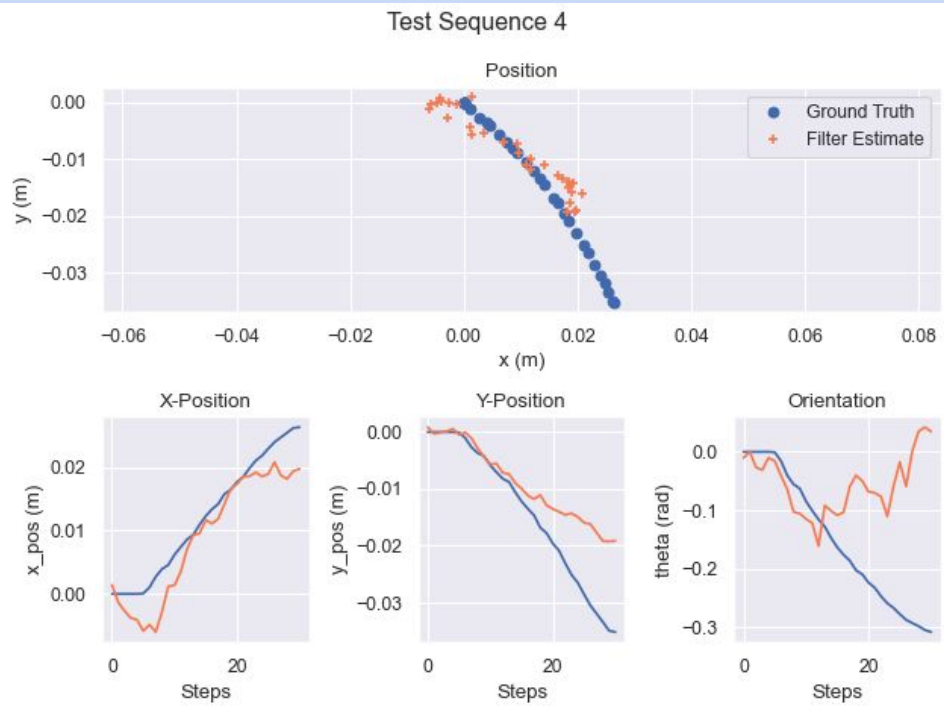


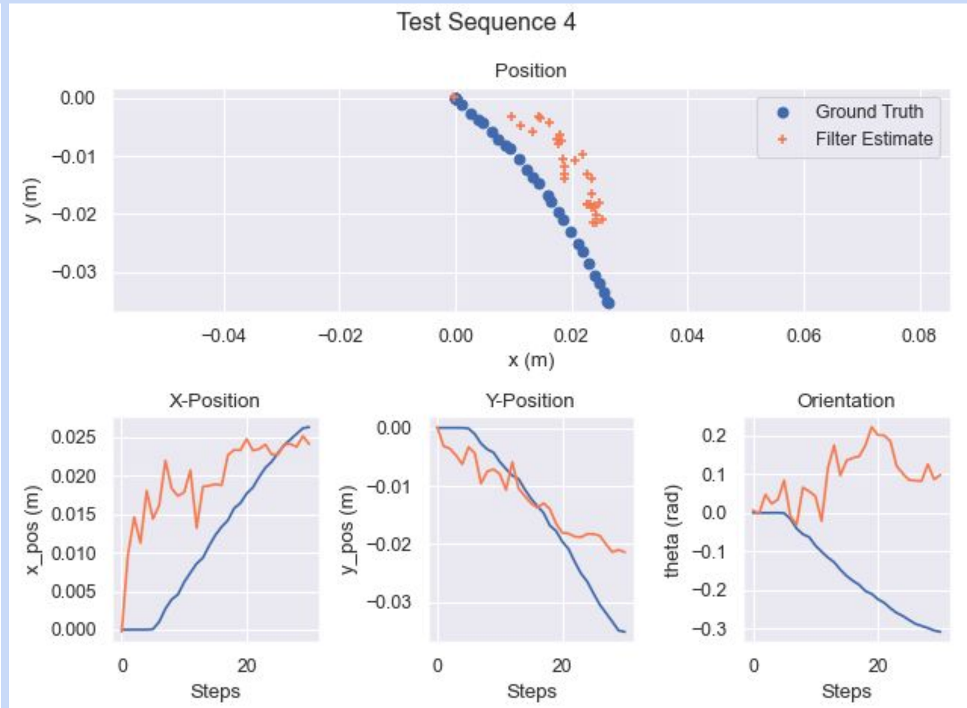
1	2	4	8	16

Prediction steps into the future

Unimodal Filter Models

Results

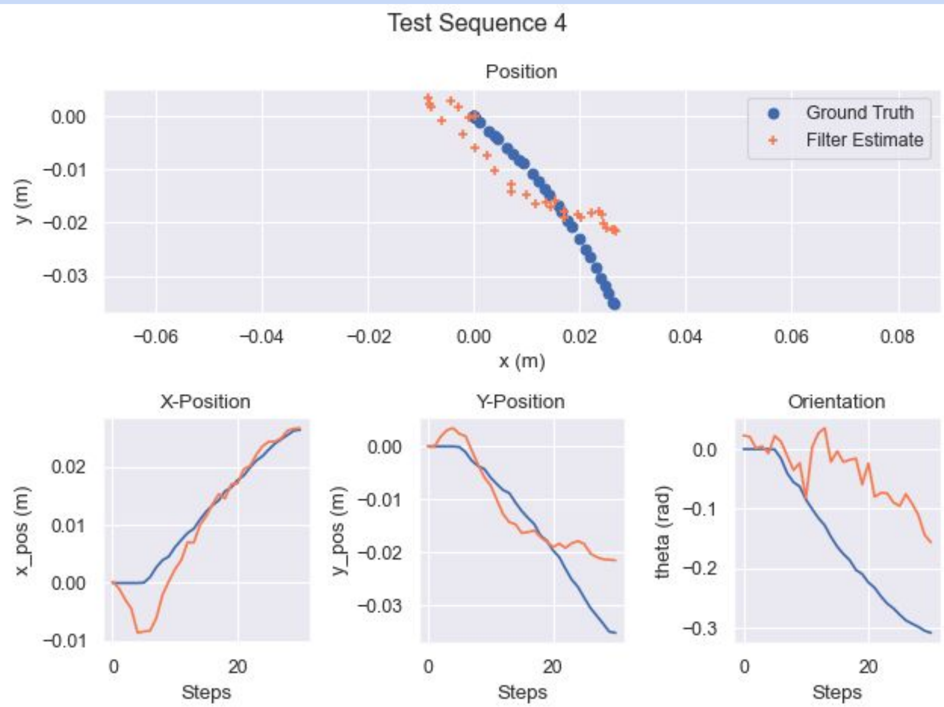


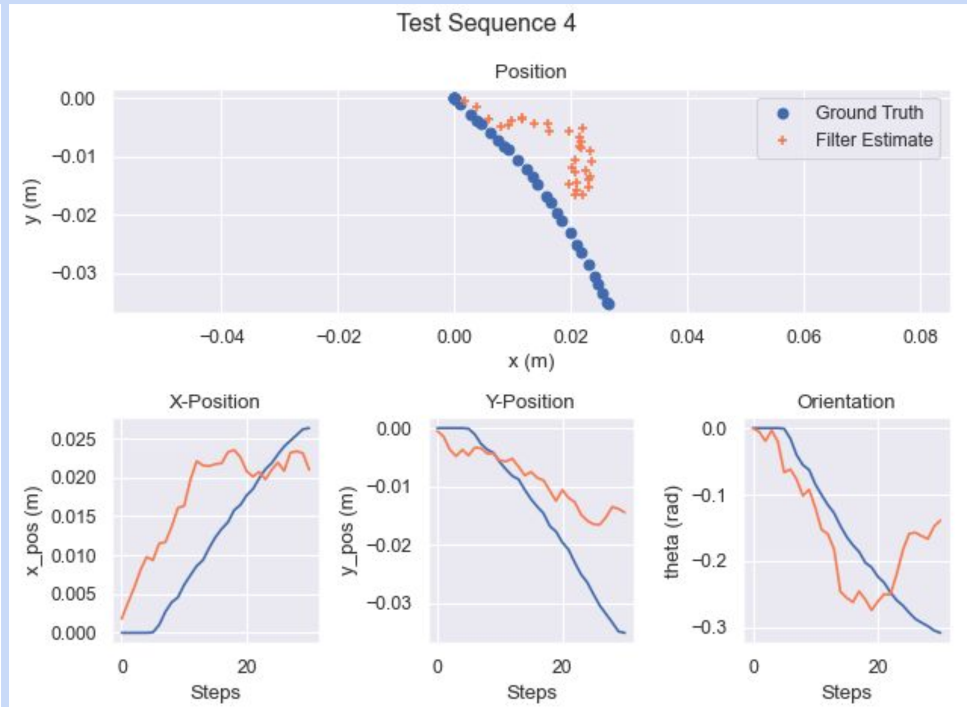


Forward Model	Pusher pose	Pusher pose
Observ. Model	Forces	Images

Multimodal Filter Models

Results



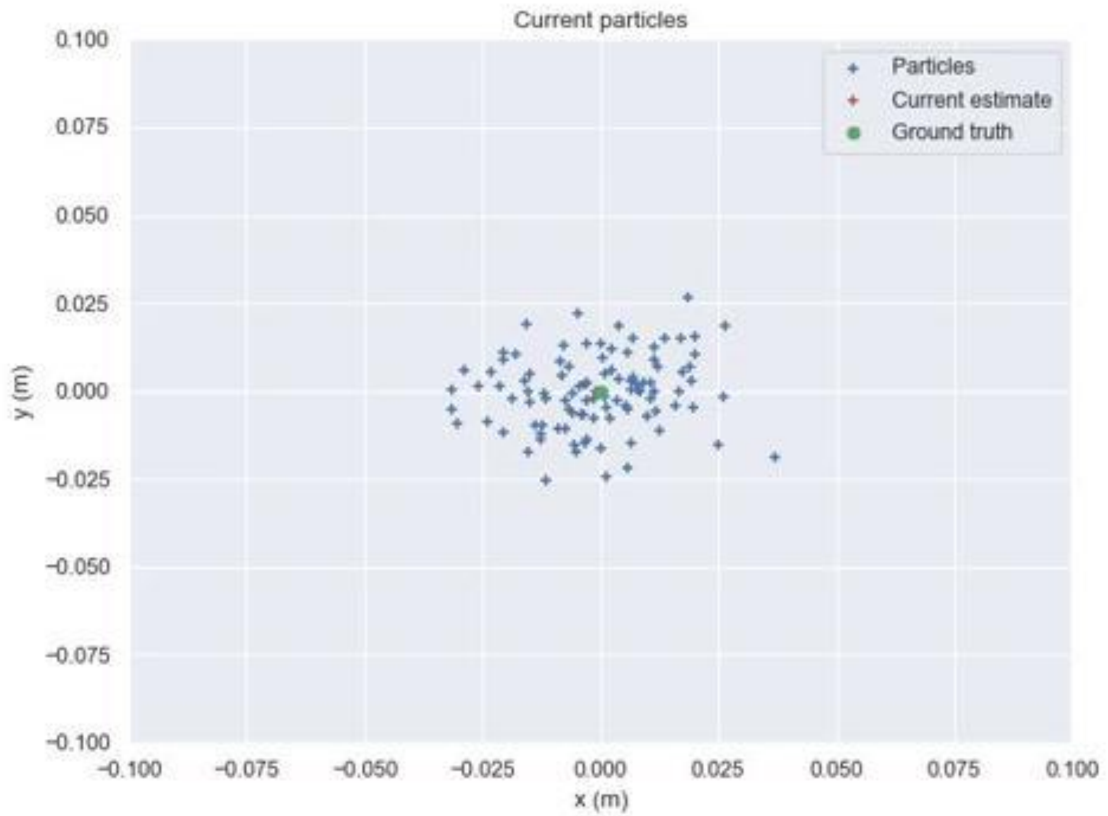


Forward Model	Pusher pose	Pusher pose & Images
Observ. Model	Forces & Images	Forces

→ **Informed proposal distribution**

How does the filtering look like?

Results



Why has the filter a hard time?

Discussion & Conclusion

- Difficult training
- Resampling makes difference during training

    → Observation model gets better

BUT: Filtering without known process or observation model possible!

→ combining modalities can lead to better results

“Classic” process model and observation model slightly better than a measurement-informed proposal distribution in this pushing scenario.

Thank you for your attention!

Questions?

References

[1] Mericli, Tekin & Veloso, Manuela & Akin, H. Levent, “Achievable push-manipulation for complex passive mobile objects using past experience”, in 12th International Conference on Autonomous Agents and Multiagent Systems 2013, AAMAS 2013, pp. 71-78.

[2] K. T. Yu, M. Bauza, N. Fazeli, and A. Rodriguez, “More than a Million Ways to Be Pushed: A High-Fidelity Experimental Data Set of Planar Pushing”, in 2016 IEEE/RSJ Internation Conference on Intelligent Robots and Systems (IROS), 2016, pp. 30–37.

Backup

Numerical Comparison

Results

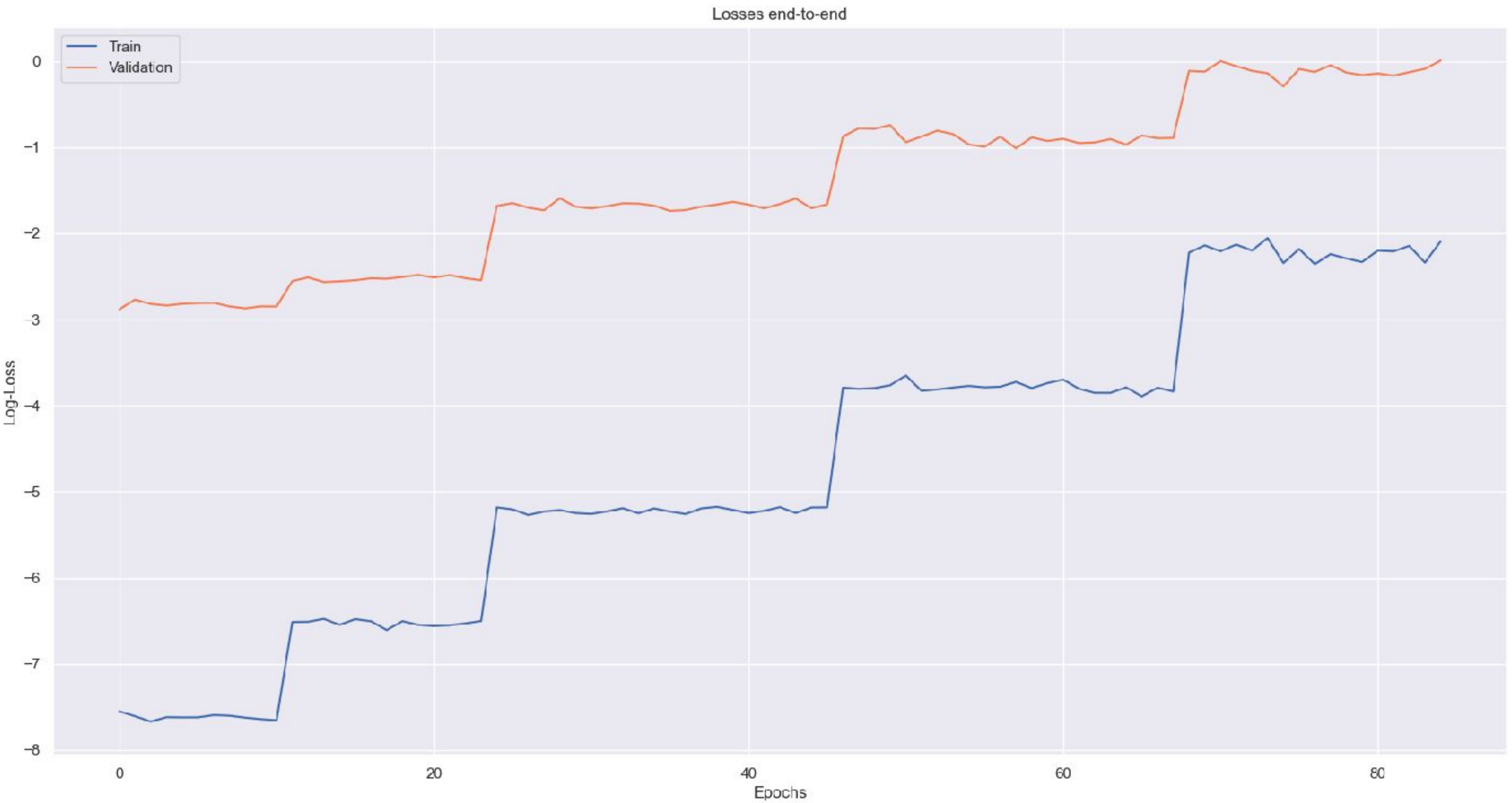
Models

RMSE in cm and rad		only Forces	only Images	Forces & Images	Proposal Images
	Pose x	0.92	1.01	0.79	0.87
	Pose y	0.79	0.77	0.79	1.15
	Pose theta	0.24	0.24	0.26	0.23
	Position	0.86	0.90	0.79	1.02

- Test dataset: 10 pushing trajectories
- Perform 3 runs across all test trajectories for each model

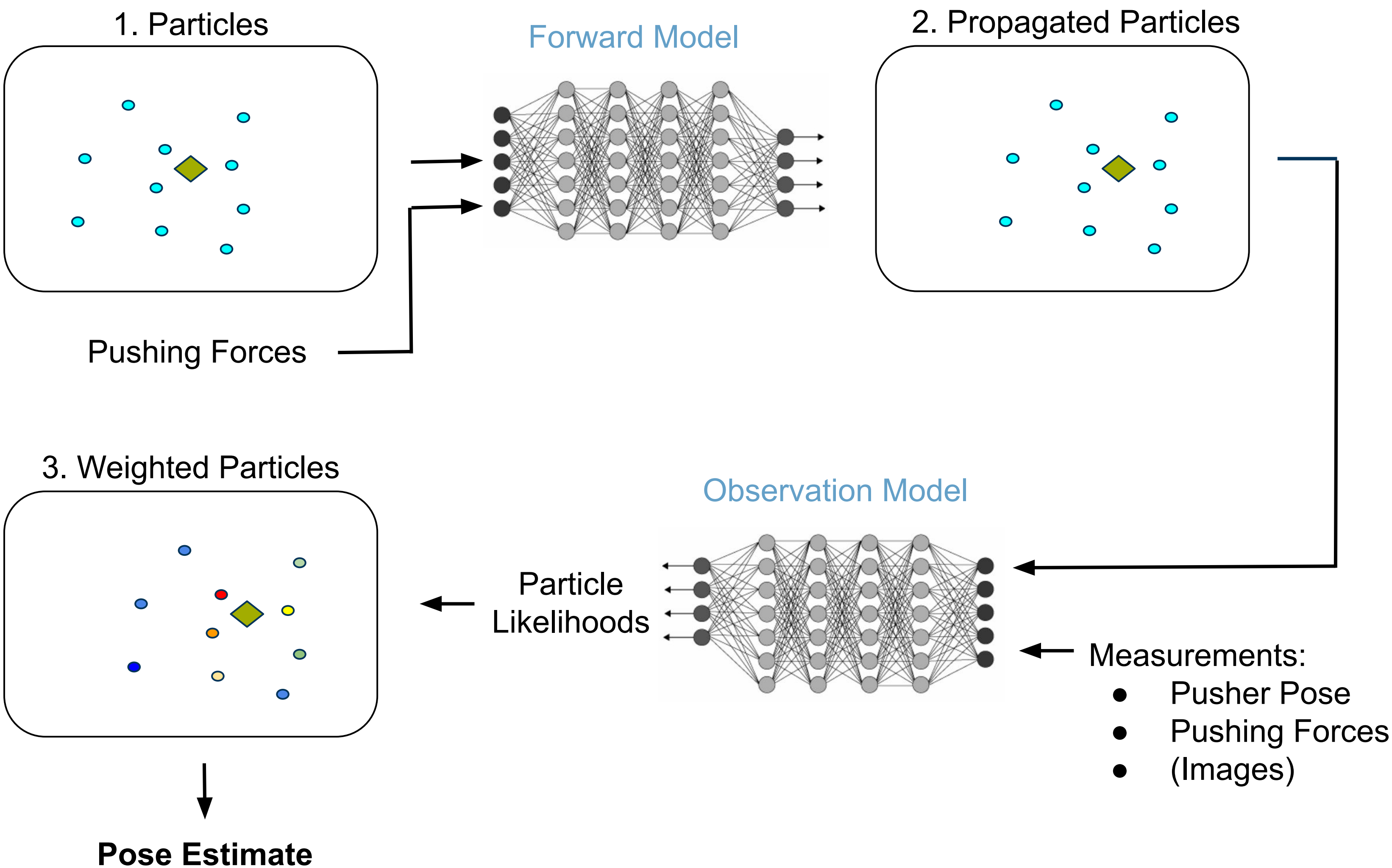
→ Average results

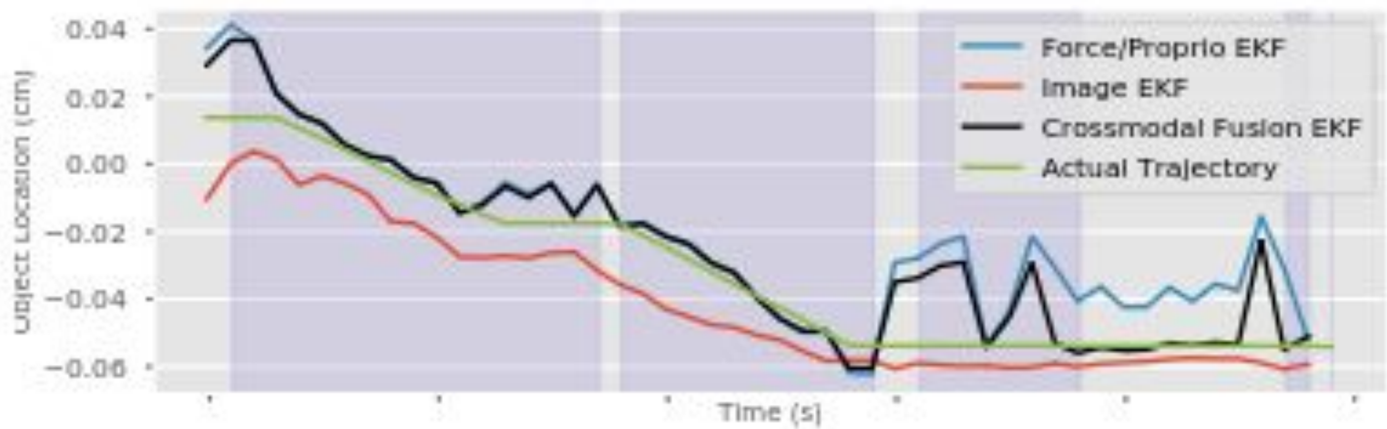
Loss end-to-end training



Differentiable Particle Filters

Methods





0.04
0.02
0.00
−0.02
−0.04
−0.06
Object Location (cm)
Time (s)
Force/Proprio EKF
Image EKF
Crossmodal Fusion EKF
Actual Trajectory