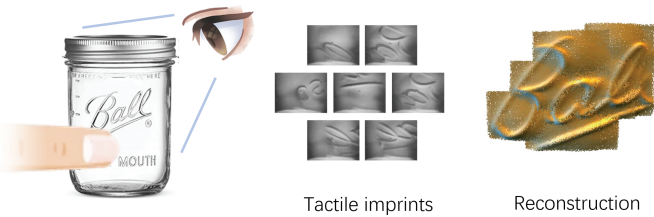
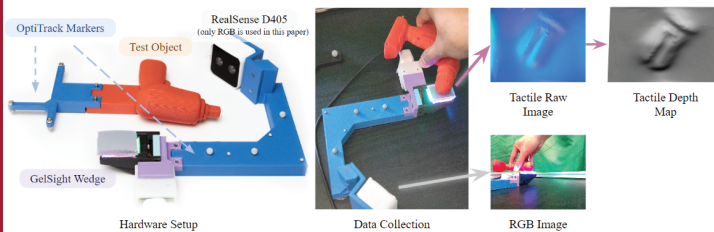


## In-hand SLAM: why vision + tactile?

Vision	Tactile	Vision + Tactile
Occlusion while in-hand	No occlusion requires contact	No occlusion
Affected by reflection/transparency	Not affected	Not affected
Global features	Local textures	Global + local



## Experiment setup

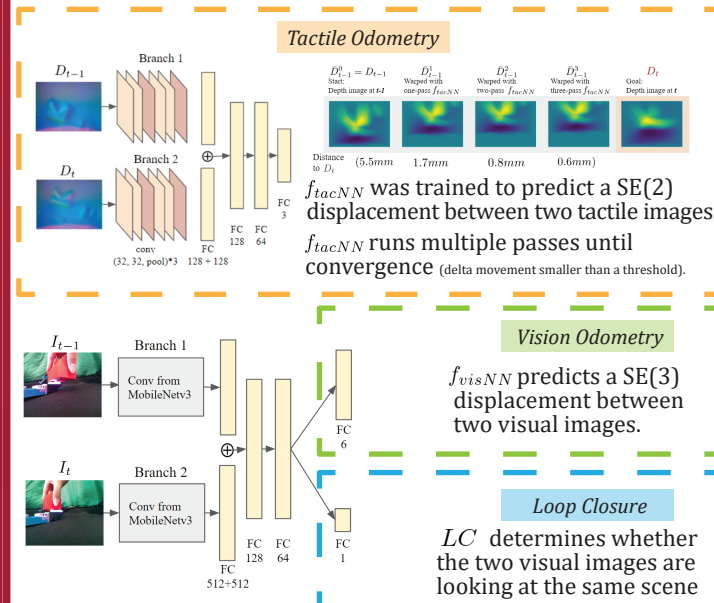


- Tactile sensor: GelSight Wedge<sup>[1]</sup> camera-based tactile sensor 200\*150px, depth reconstructed with photometric stereo
- Vision sensor: Intel RealSense D405 (only RGB is used)
- Pose tracking: OptiTrack Flex
- Test objects: YCB dataset<sup>[2]</sup>, ABC dataset<sup>[3]</sup>
- Dataset size: 9,500 raw datapoints, 40,000 after augmentation augmented by applying randomly sampled offset to tactile images

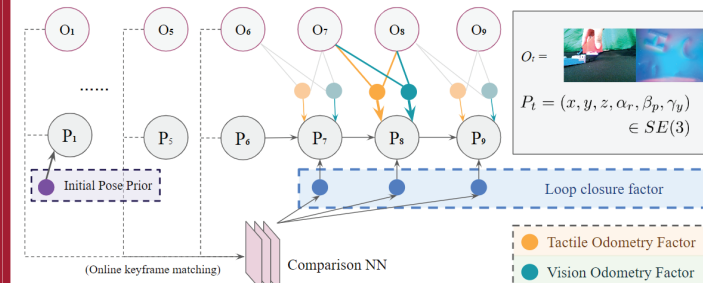
[1] Wang, Shaoxiang, et al. "Gelsight wedge..." ICRA, 2021.  
[2] Calli, Berk, et al. "Yale-CMU-Berkeley dataset for robotic manipulation research." IJRR, 2017  
[3] Koch, Sebastian, et al. "Abc: A big cad model dataset for geometric deep learning." CVPR, 2019.

## Pose prediction & 3D reconstruction

Pose at each time step is predicted with a **factor graph**, which consists of a **tactile odometry**, a **vision odometry**, and a **loop closure** factor.



## Factor Graph Optimization

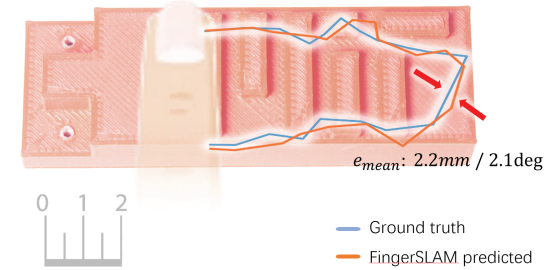


A factor graph is constructed to optimize for the best pose sequences. A covariance matrix which represents a noise model is assigned to each odometry. The factor graph was solved with GTSAM\*.

\* F. Dellaert, "Factor graphs and gtsam..." Georgia Institute of Technology, Tech. Rep., 2012







## Evaluation

### Trajectory prediction



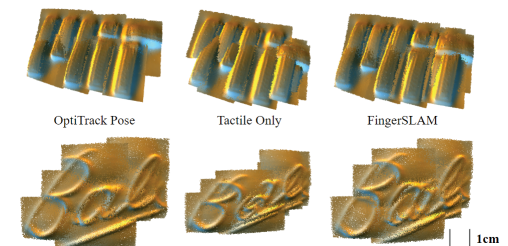
Low accumulated error observed for FingerSLAM.

### Ablation study

Name	Picture	Raw mm : deg	Tactile only mm : deg	Vision only mm : deg	Combined mm : deg	w/ Loop Closure mm : deg
Letter Cube		6.18 : 6.34	3.10 : 2.79	2.99 : 2.61	2.87 : 2.29	2.27 : 2.18
Graphic		5.93 : 5.59	2.45 : 2.61	2.86 : 2.11	2.28 : 2.13	2.13 : 2.07
Rubik's Cube		5.88 : 5.66	2.33 : 2.80	2.97 : 2.37	2.50 : 2.43	2.30 : 2.41
Mini Bleach		5.95 : 5.74	4.52 : 3.77	3.25 : 2.54	4.06 : 2.91	3.46 : 2.50
Painted Glass Jar		6.40 : 6.21	3.50 : 2.70	4.12 : 2.94	3.63 : 2.63	3.19 : 2.82
Glass Jar		6.46 : 6.02	3.58 : 2.47	4.79 : 4.21	3.79 : 2.27	3.38 : 2.27

FingerSLAM performs better than each individual components.

### Reconstruction



FingerSLAM produces realistic 3D reconstructions by stitching multiple tactile imprints.