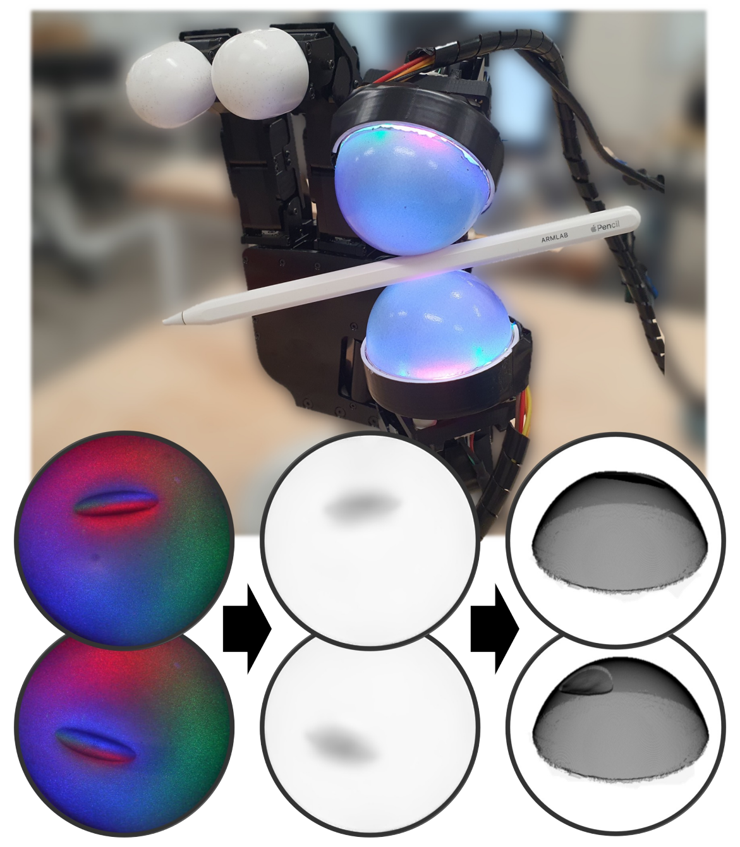
DenseTact: Optical Tactile Sensor for Dense Shape Reconstruction

PI: Monroe Kennedy III PhD, Assistant Professor of Mechanical Engineering, Stanford University
https://arm.stanford.edu/research/improving-robotic-assistant-dexterity

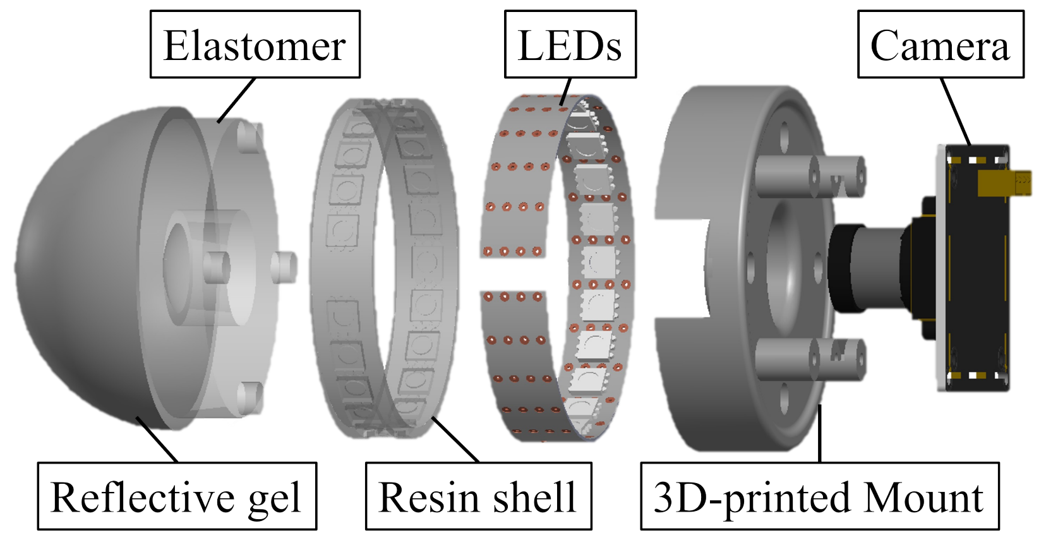
Motivation

Improving robotic dexterous manipulation is the catalyst to enabling ubiquitous robots capable of performing advanced collaborative tasks. Robots capable of manipulating small objects with the ability accurately assess stability and adapt or regrasp is necessary for systems to be effective in tasks ranging from rapid industrial assembly to assisted living tasks. DenseTact is an optical tactile sensor whose first generation provides *calibrated, high-resolution* shape reconstruction.



DenseTact Design

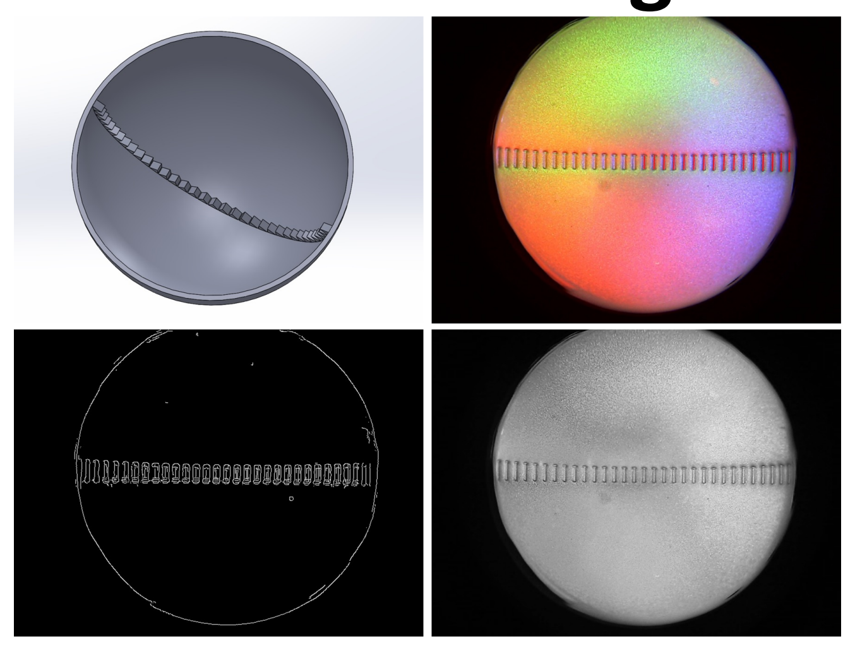
The optical tactile sensor consists of a fisheye camera, internal illumination, soft elastomer 'finger' with a reflective surface and chassis components.



Modeling and Calibration

Concept. The camera observes the LED illuminated interior boundary of the sensor whose deflection (surface normal) corresponds light incidence angle. The goal is to map the color image to a deflected boundary shape with high resolution and accuracy.

Fig 1.



Calibration. To account for distortion caused by the fisheye lens to points on the surface, we leverage a Gaussian Process with the calibration target shown in (Fig 1). To calibrate between interior images and surface deflection, we impress 3D printed parts (Fig 2) into the sensor and perform ray tracing and correspondence to construct an expected radial depth image (Fig 3).

Fig 2.

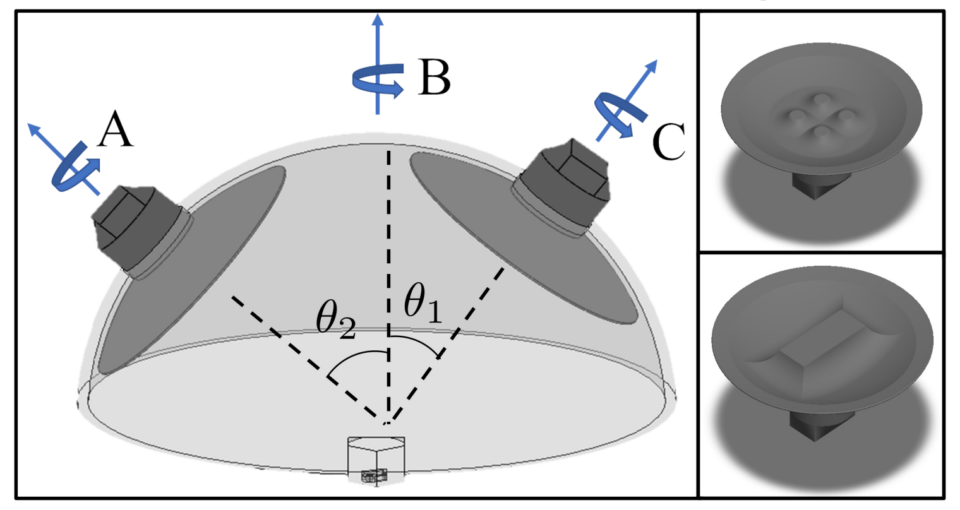
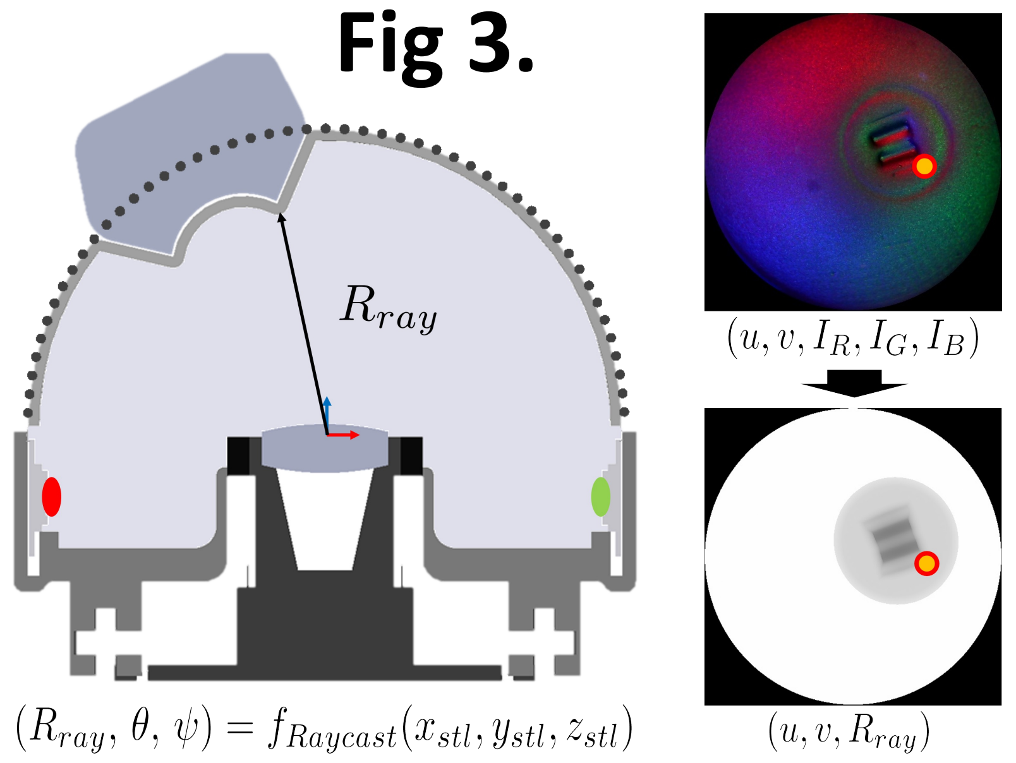


Fig 3.

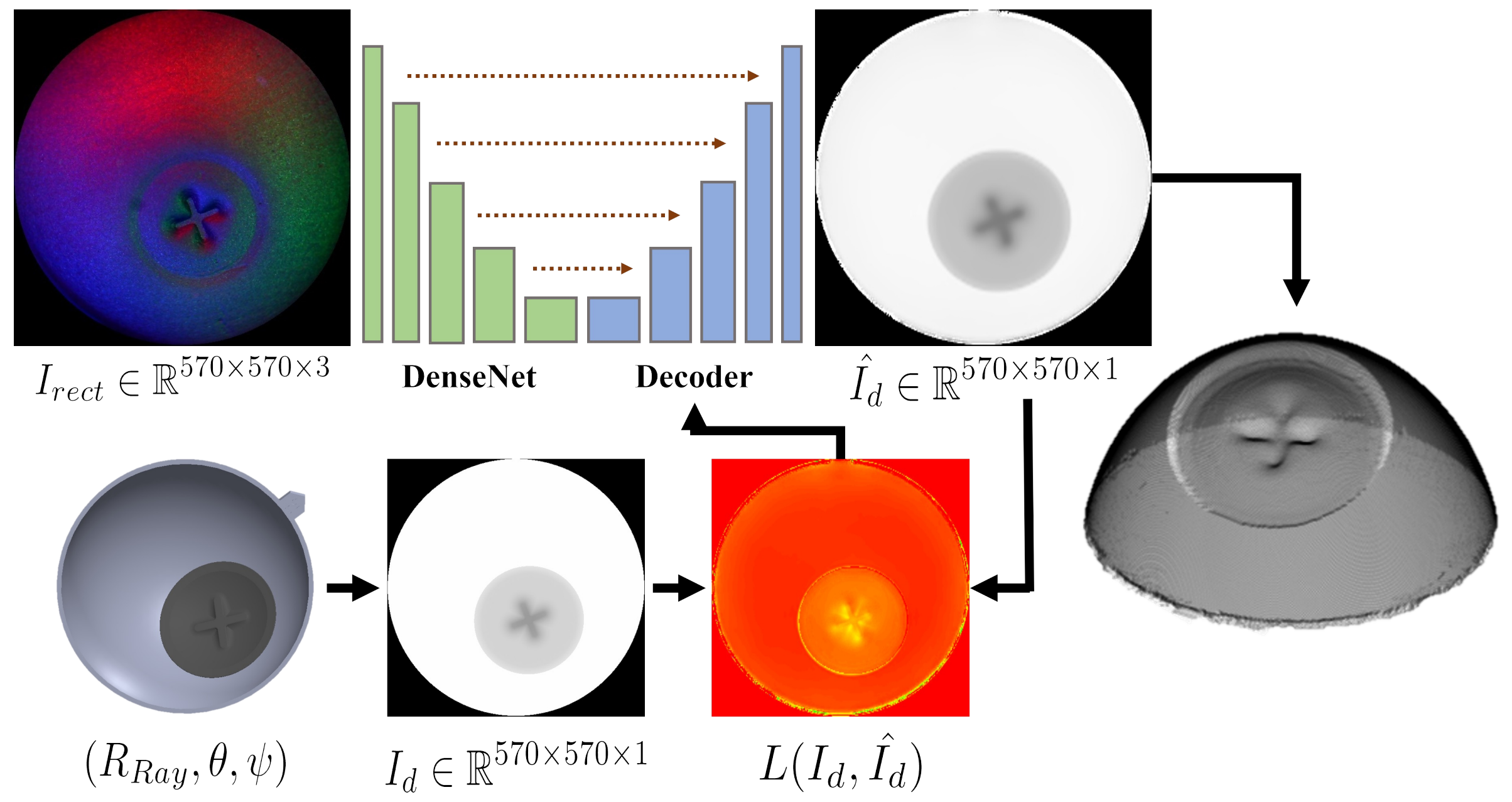


$(R_{ray}, \theta, \psi) = f_{Raycast}(x_{stl}, y_{stl}, z_{stl})$

(u, v, R_{ray})

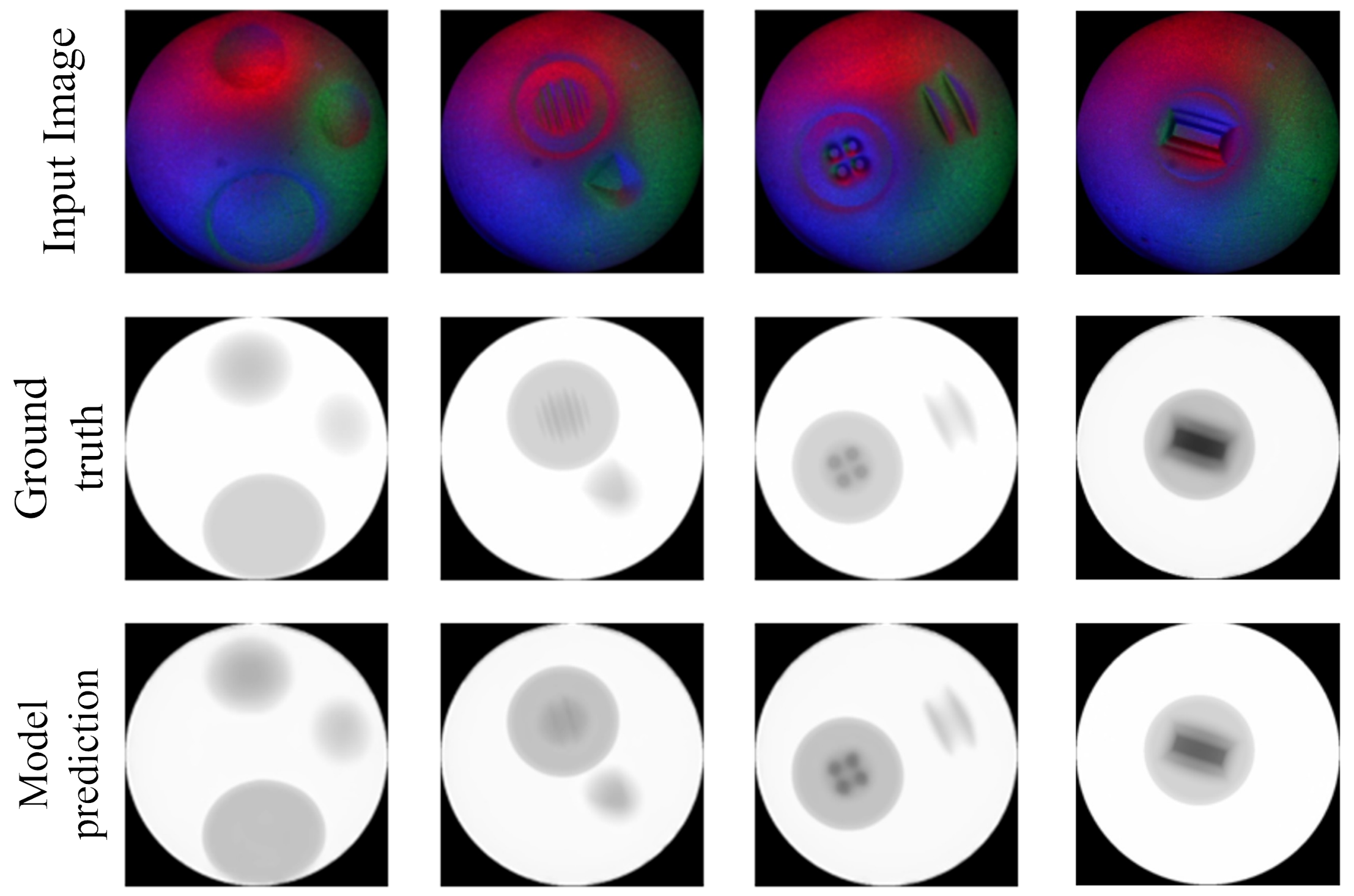
Modeling. A neural network with skip connections is used to map the input image to the radial depth image, trained with the ground truth from the STL files. A point cloud is then obtained.

Fig 4.

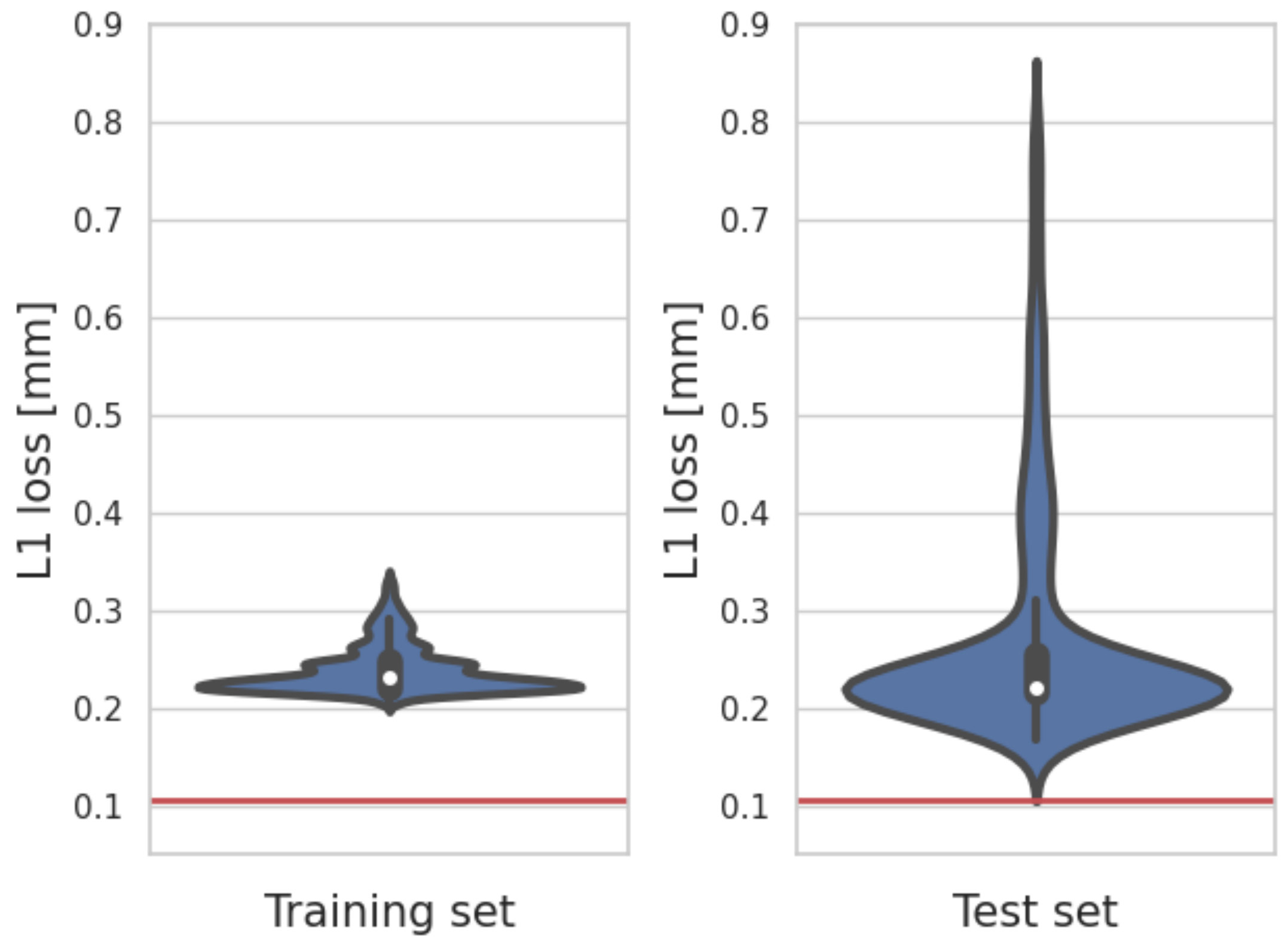


Results

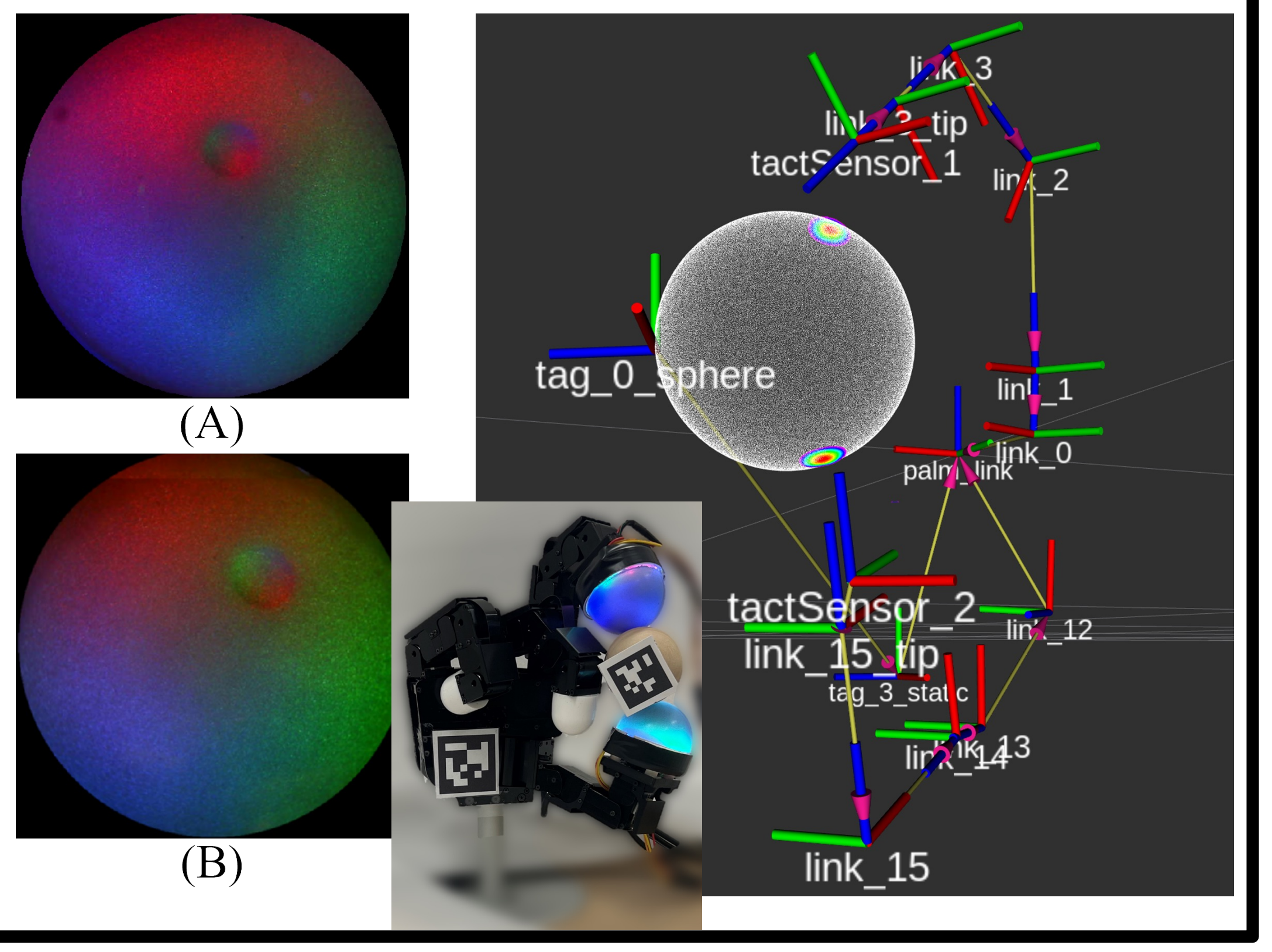
Shape Reconstruction. The images show the ability for our sensor and model to estimate radial depth. The violin plots demonstrate the quantitative performance with the L1 re-projection error for every pixel in an image (253,213 pixels per image), and that for 29,200 training images and 1000 test images. The mean L1 loss for training and test respectively is 0.2381mm and 0.2811 mm.



https://github.com/armlabstanford/DenseTact



In-hand Localization. The calibrated radial depth images were used with iterative closest point to localize an object within the grasp. With an average fitness score over 23 grasps of 0.597 with a variance of 0.238



2022 NRI & FRR Principal Investigators' Meeting
April 19-21, 2022

Assistive Robotics and Manipulation Laboratory

Award ID#: **2142773**