CRISP: Object Pose and Shape Estimation with Test-Time Adaptation

Jingnan Shi, Rajat Talak, Harry Zhang, David Jin, Luca Carlone

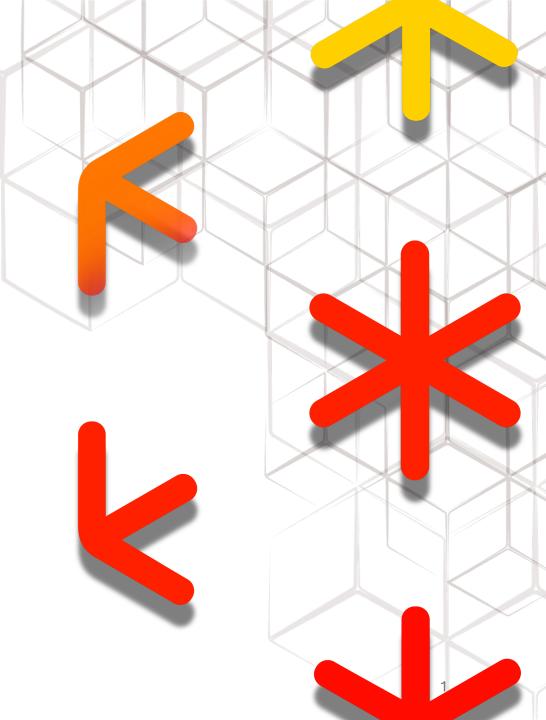
Massachusetts Institute of Technology











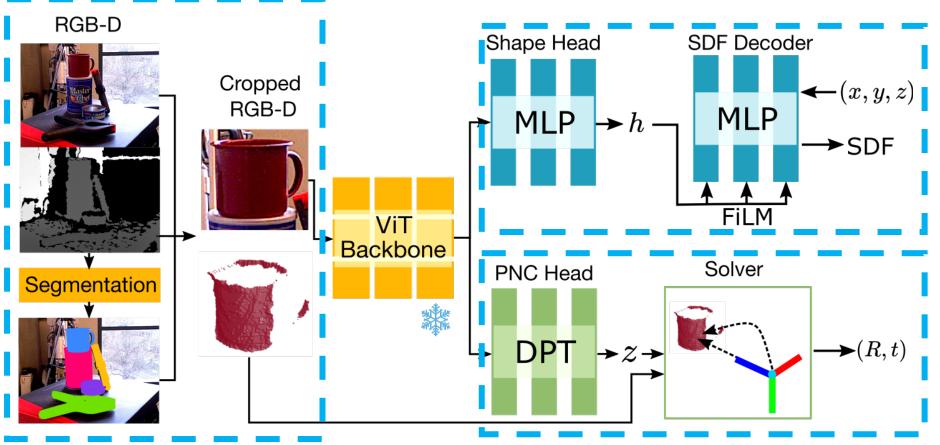
Problem Statement

How can we design an object pose and shape estimation pipeline that can self-train during test time?

Our Contributions

- 1. We introduce CRISP, an object pose and shape estimation pipeline.
- 2. We propose an optimization-based pose and shape corrector that can correct estimation errors.
- 3. We adopt a correct-and-certify approach to selftrain CRISP and adapt to new domain in a selfsupervised manner.

CRISP: Pose and Shape Estimation Network



We use FiLM conditioning + SIREN network for the SDF decoder

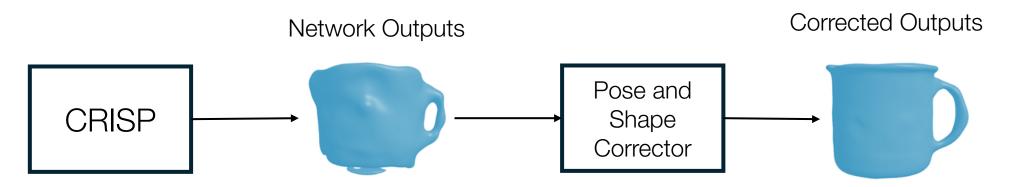
We adopt a Dense Prediction Transformer (DPT) network for predicting Pose Normalized Coordinates (PNC) for all pixels

We assume access to segmented RGB-D images

We use DINOv2 as the ViT backbone, but can potentially use any ViT models

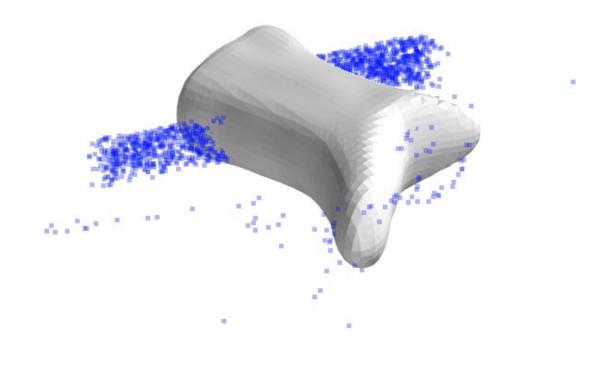
Pose and Shape Corrector

Motivation: Can we partially bridge the domain gap using optimization initialized with the network output, without object CAD models?



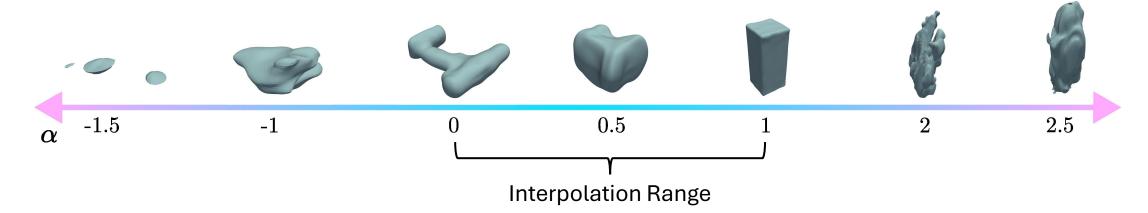
Pose and Shape Corrector

Gradient descent doesn't work well



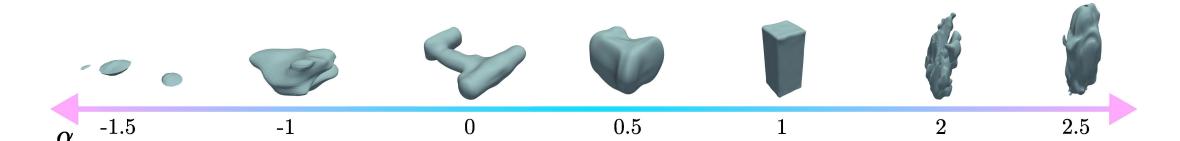
Pose and Shape Corrector

Observation: Interpolation within the simplex of learned latent codes produces much more reasonable shapes.

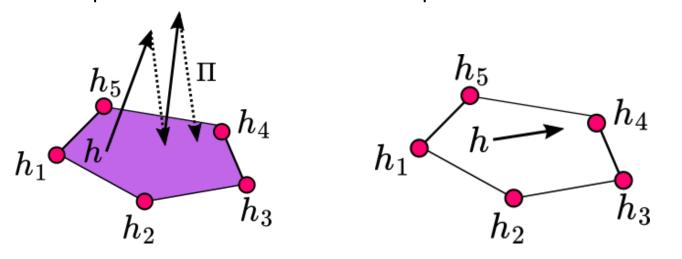


An active shape model for shape latent codes

Observation: Interpolation within the simplex of learned latent codes produces much more reasonable shapes.

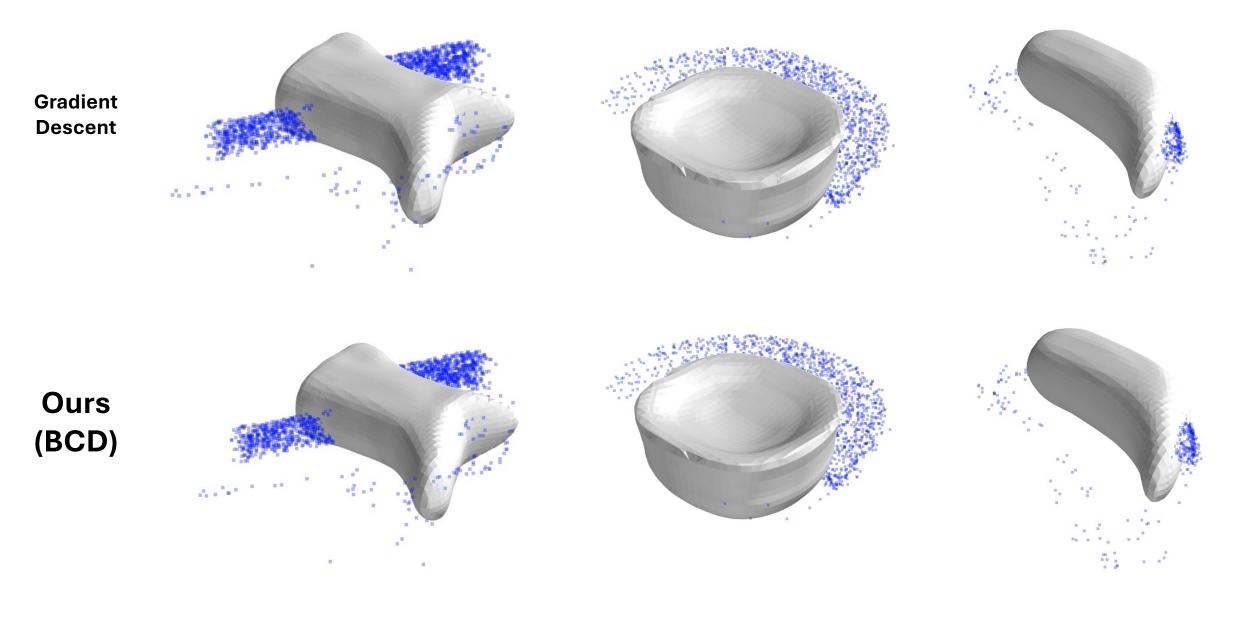


We develop two correctors that exploit this observation.

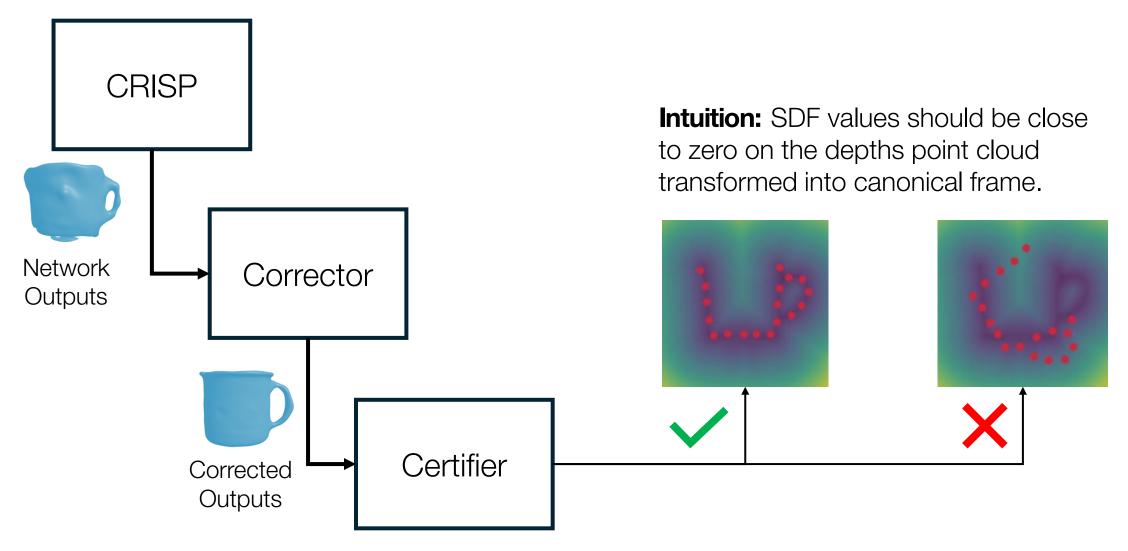


Block Coordinate Descent (BCD) Corrector

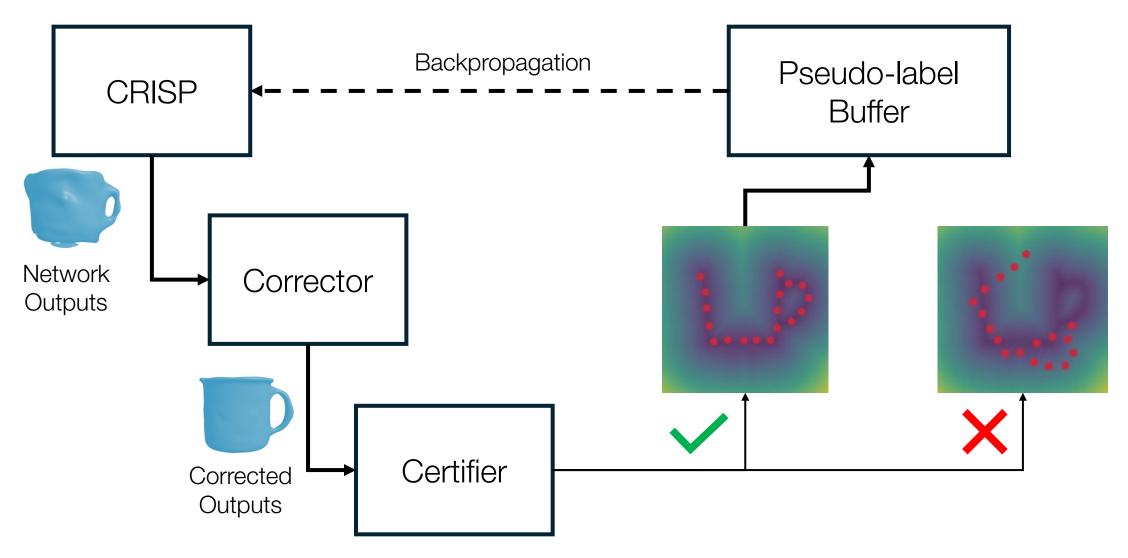
Least-squares (LSQ) Corrector



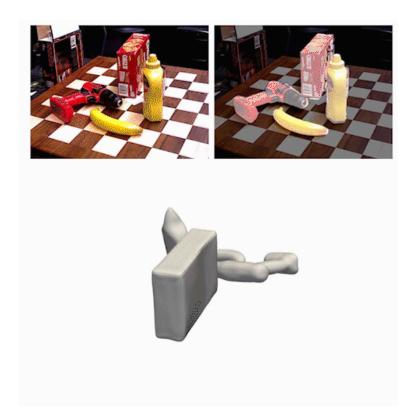
Self-training for Test-time Adaptation



Self-training for Test-time Adaptation



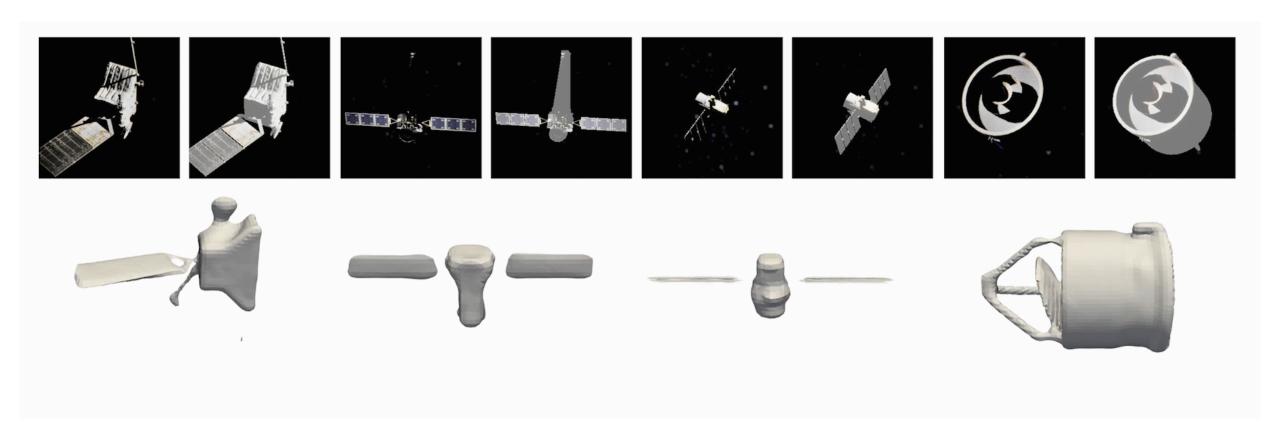
Experimental Results





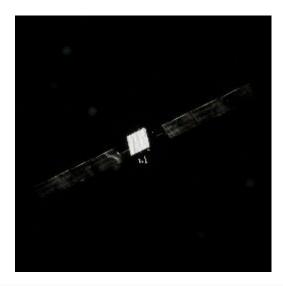


Experimental Results



Adaptation to Unseen Objects

Test Image





Network Output: Reconstructed Model Before Self-training





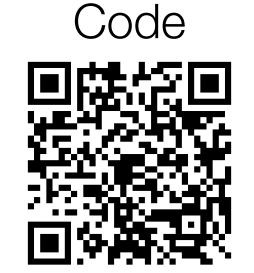
Similar Satellite in the Train Set

Reconstructed Model After Self-training



Thank you for your attention!





ExHall D Poster #96