

Pos3R: 6D Pose Estimation for Unseen Objects Made Easy

Weijian Deng¹ Dylan Campbell¹ Chunyi Sun¹ Jiahao Zhang¹
Shubham Kanitkar² Matthew E. Shaffer² Stephen Gould¹

¹Australian National University ²RIOS Intelligent Machines

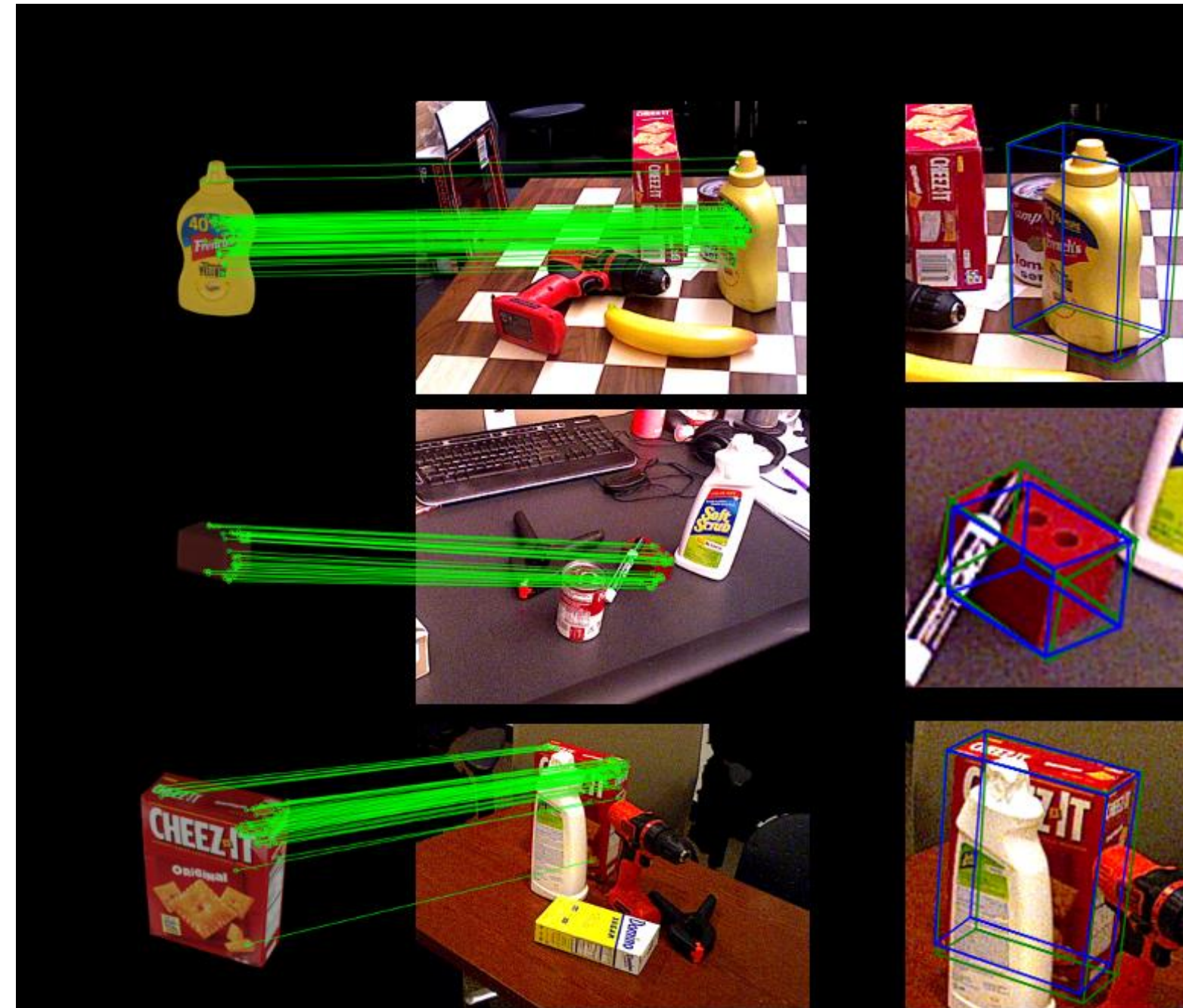


Australian
National
University



6 D OBJECT POSE ESTIMATION

□ Task Definition: Training Free, RGB-Only, and CAD Model

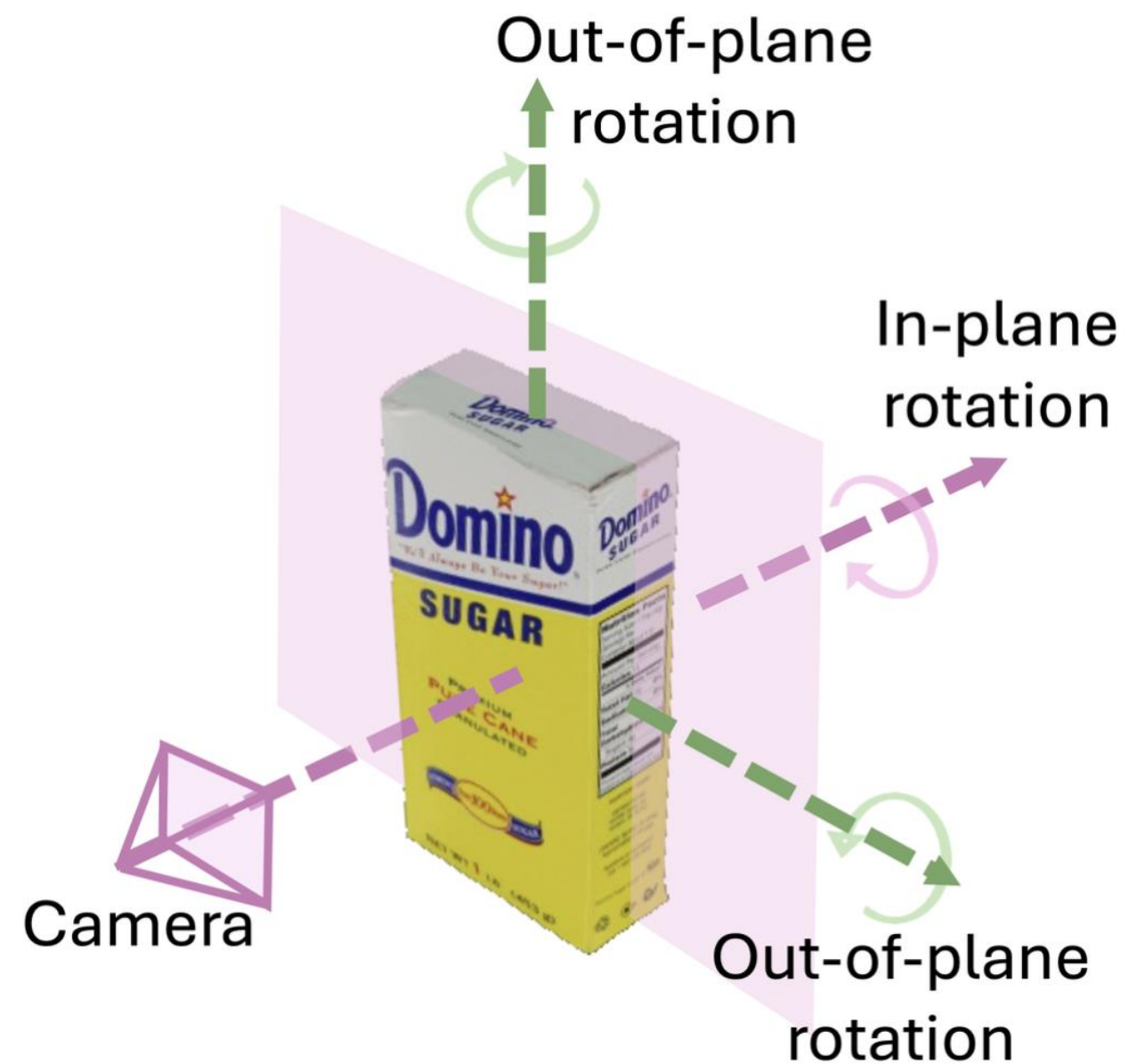


6 D O B J E C T P O S E E S T I M A T I O N

□ Task Definition: Training Free, RGB-Only, and CAD Model

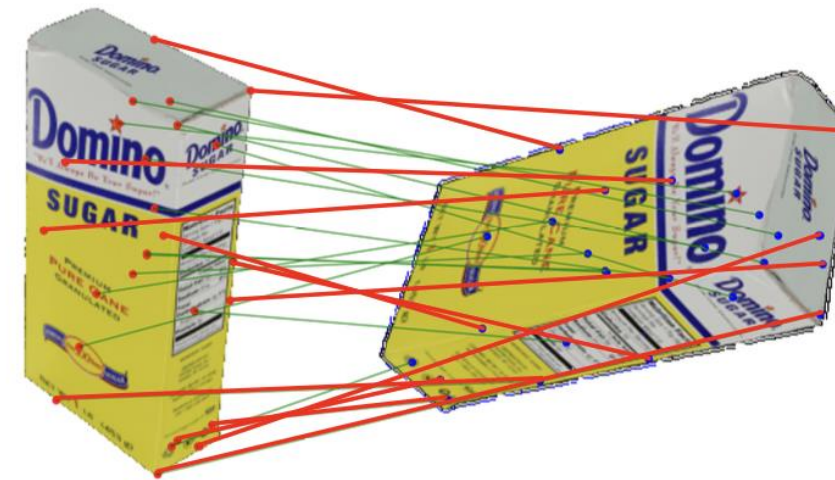
- *Training-free pipelines offer adaptability to unseen objects*
- *Model-based 6D localization estimates object pose from a 3D CAD model and an RGB image*

Why Use a 3D Foundation Model?

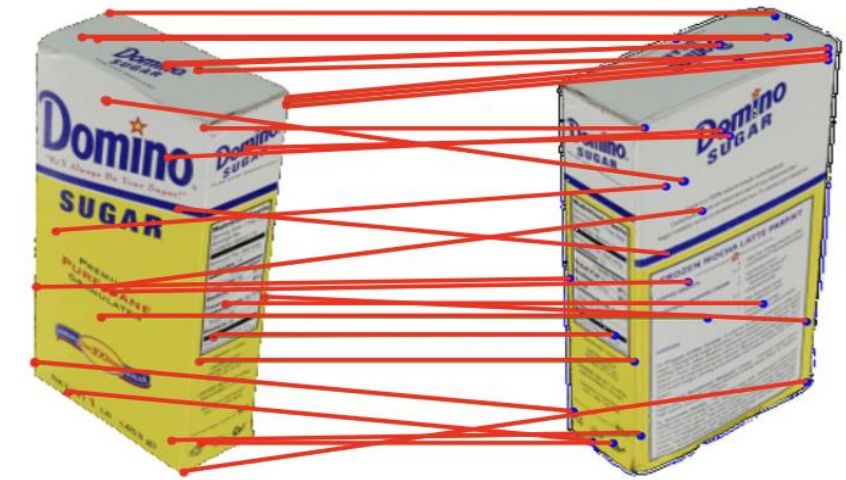


DINOv2

In-plane rotation

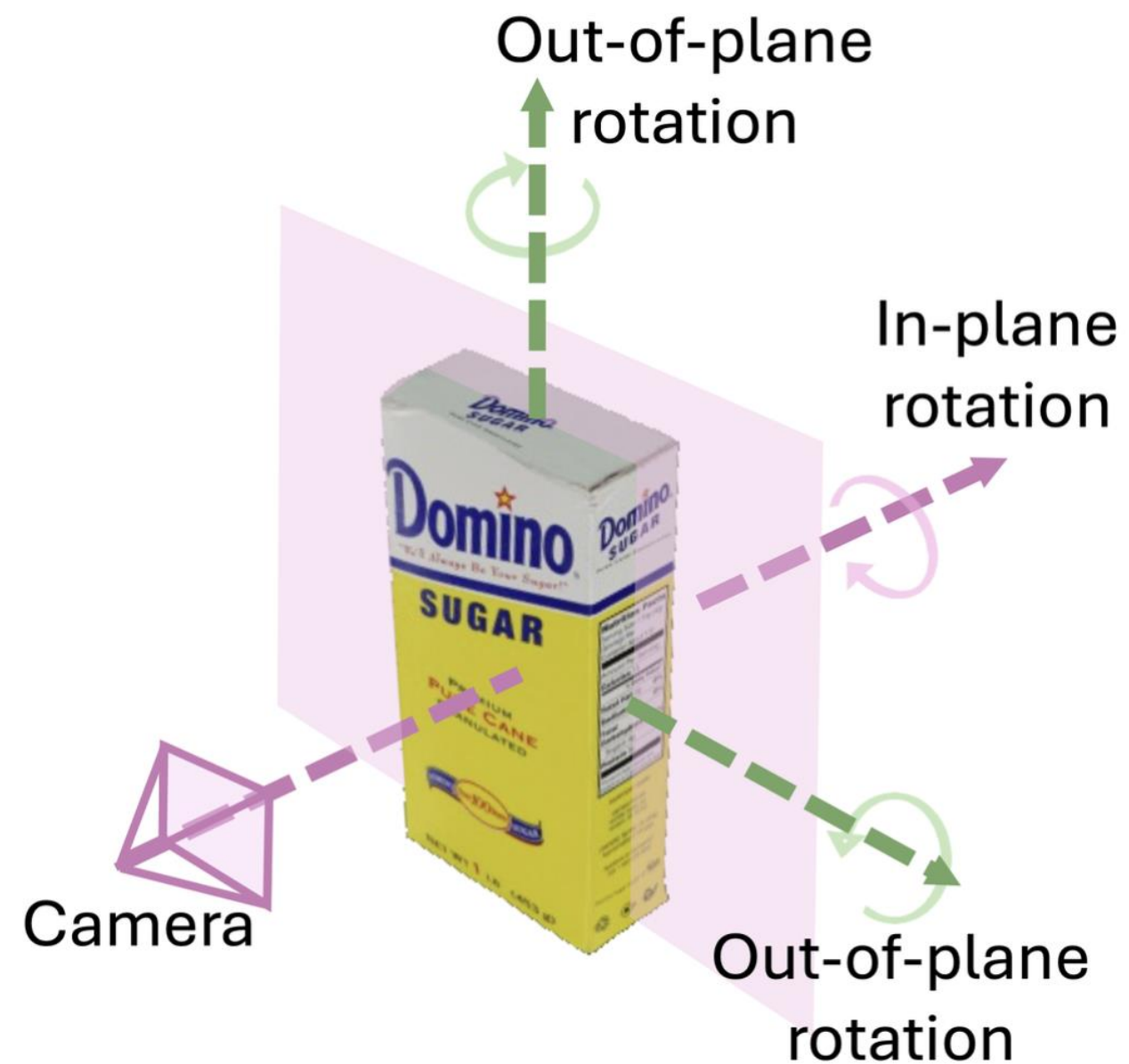


Out-of-plane rotation



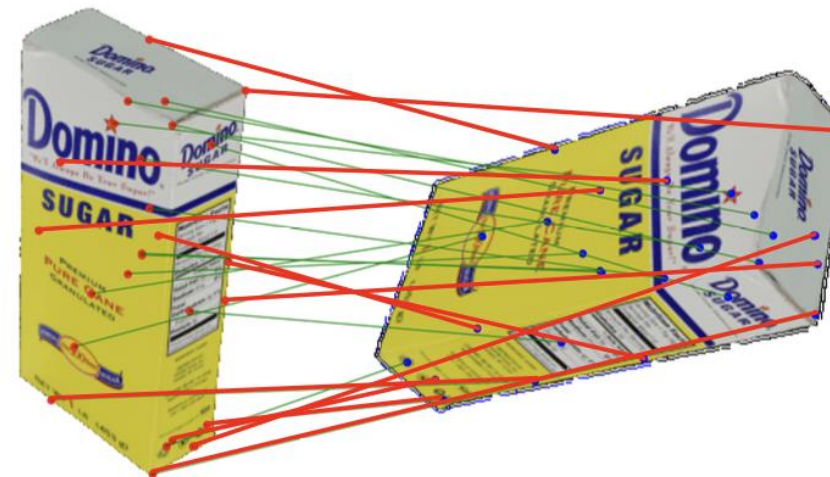
- *DINOv2 produces inconsistent matches under out-of-plane rotations since it was not trained under these transformations*

Why Use a 3D Foundation Model?

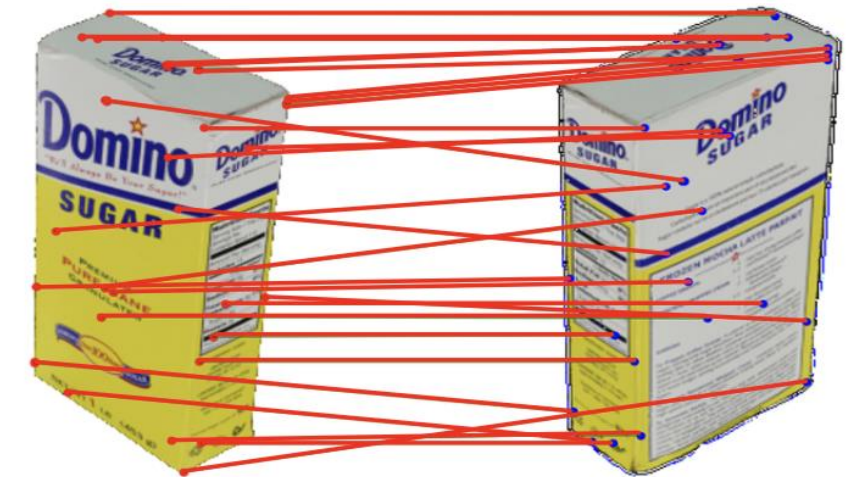


DINOv2

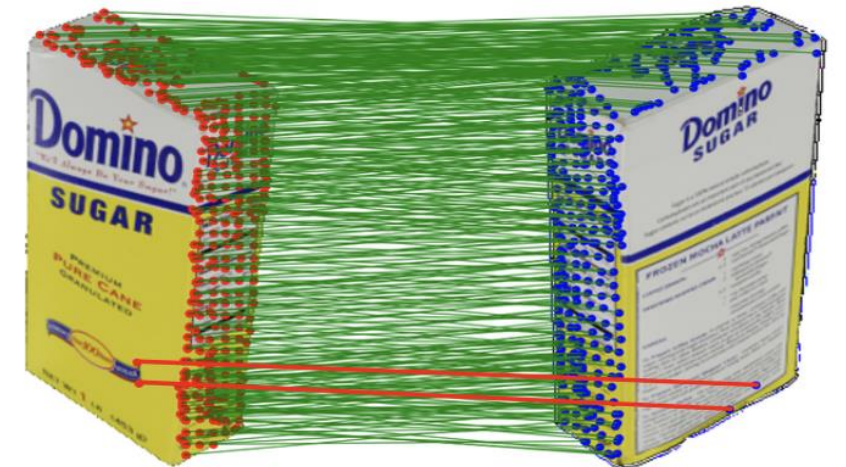
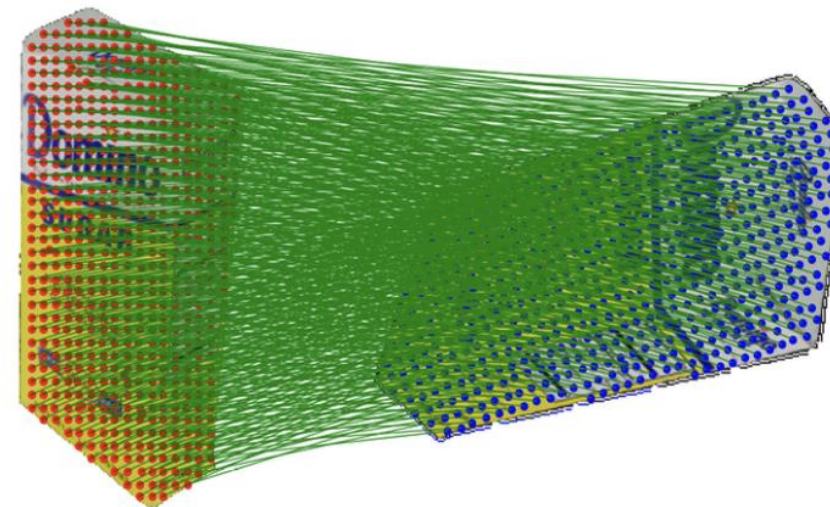
In-plane rotation



Out-of-plane rotation



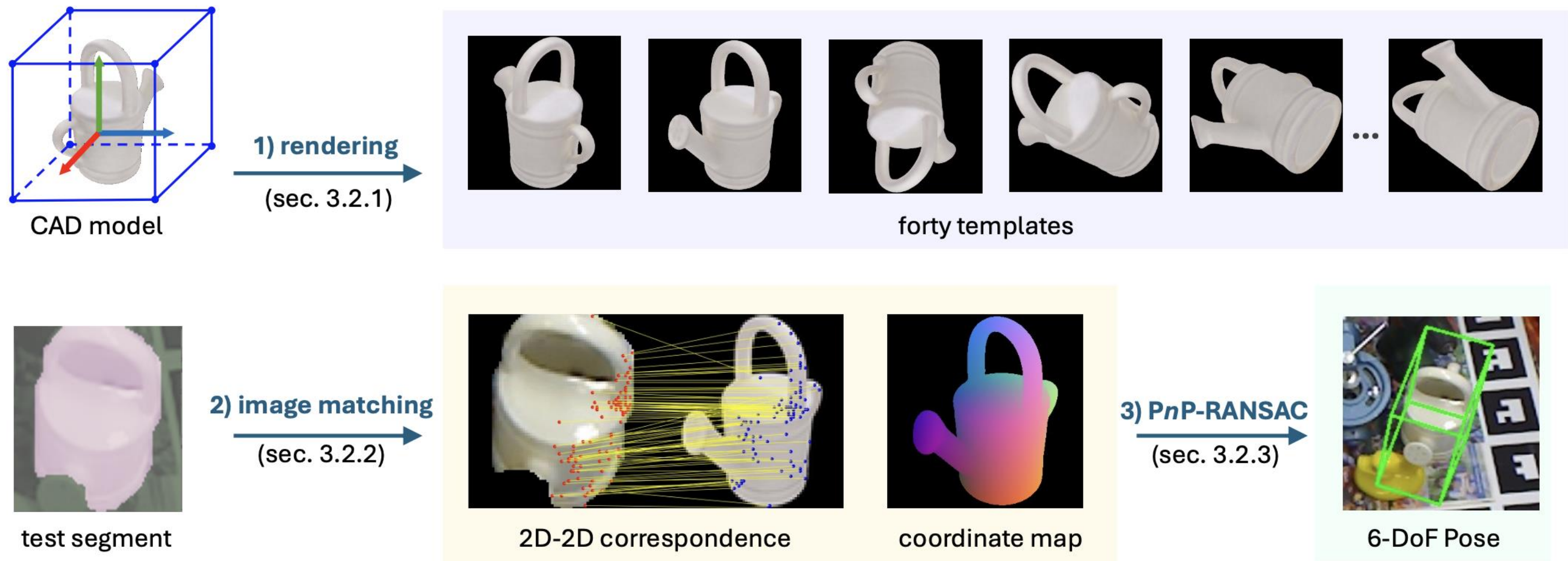
MASt3R



- *MASt3R provides dense and stable correspondences*

Pos3R

Pos3R: Training-Free and Fast — Render, Match, Fit



P o s 3 R

Step 1: Template Rendering



- *Eight base template, covering essential orientations*
- *For each, five in-plane (axial) rotations are generated around the camera's principal axis.*

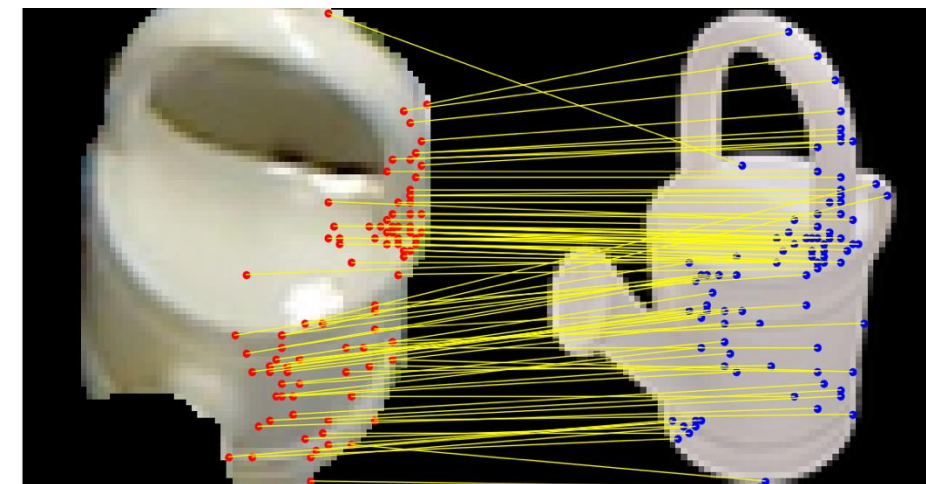
P o s 3 R

Step 2: Image Matching



test segment*

2) image matching →



2D-2D correspondence

** provided by CNOS (Cnos: A strong baseline for cad-based novel object segmentation)*

- *MASt3R produces dense 2D correspondences between test segment and every template*
- *Similarity is computed by summing feature similarities across correspondences*

P o s 3 R

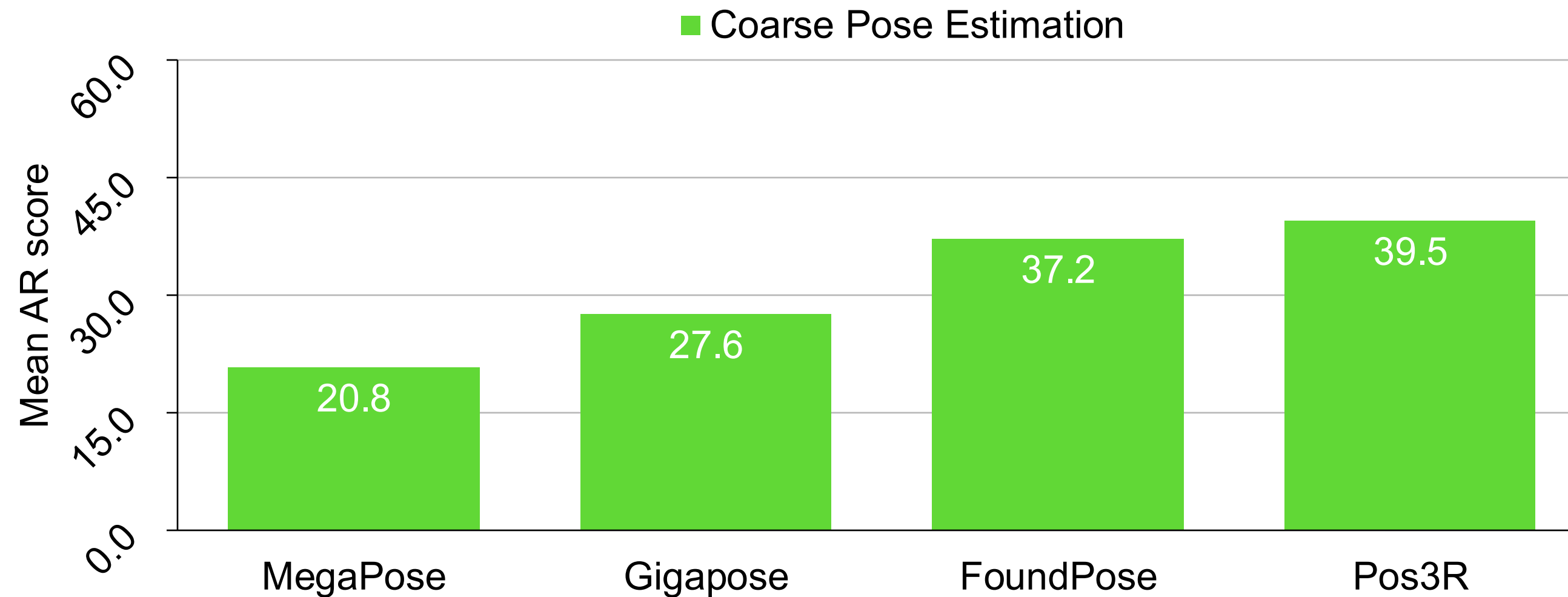
Step 3: Pose Fitting



- *The template with the highest score is selected*
- *3D coordinate map provides 2D-3D matches for PnP to estimate pose*

EXPERIMENTS

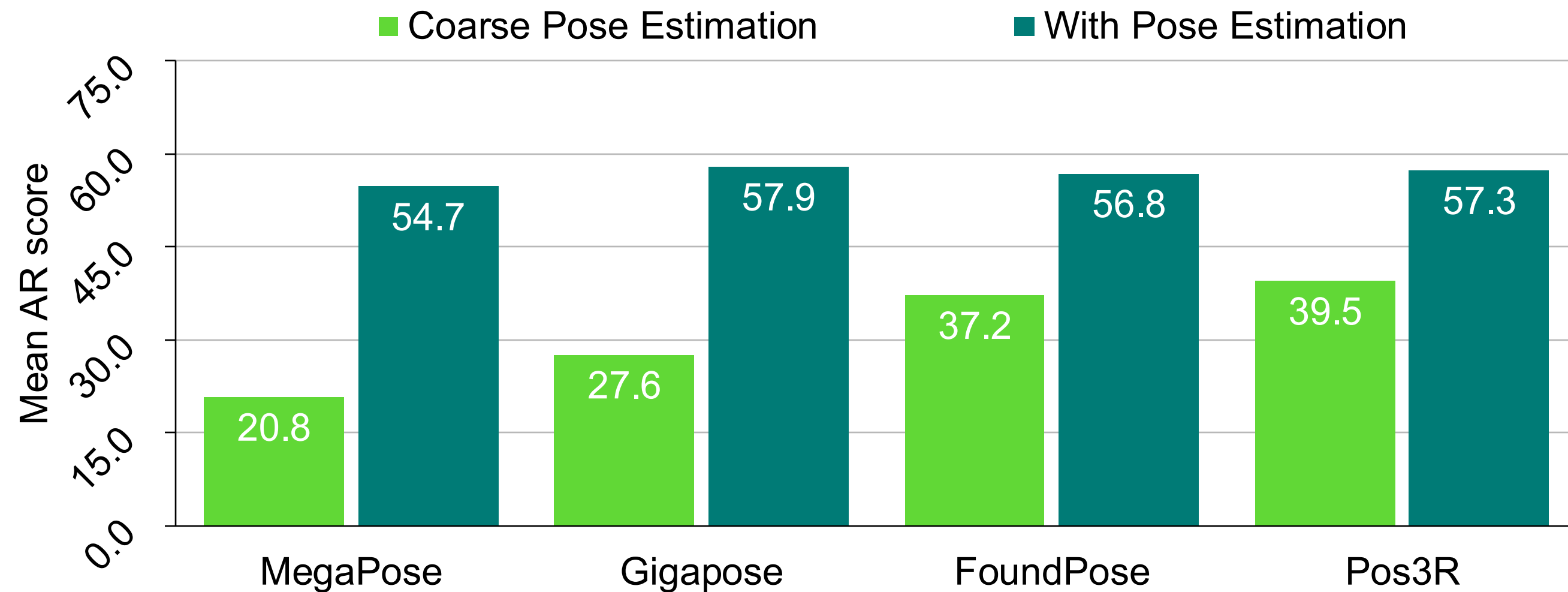
Performance Comparison on the BOP Challenge



- *Pos3R outperforms other methods in coarse pose estimations*

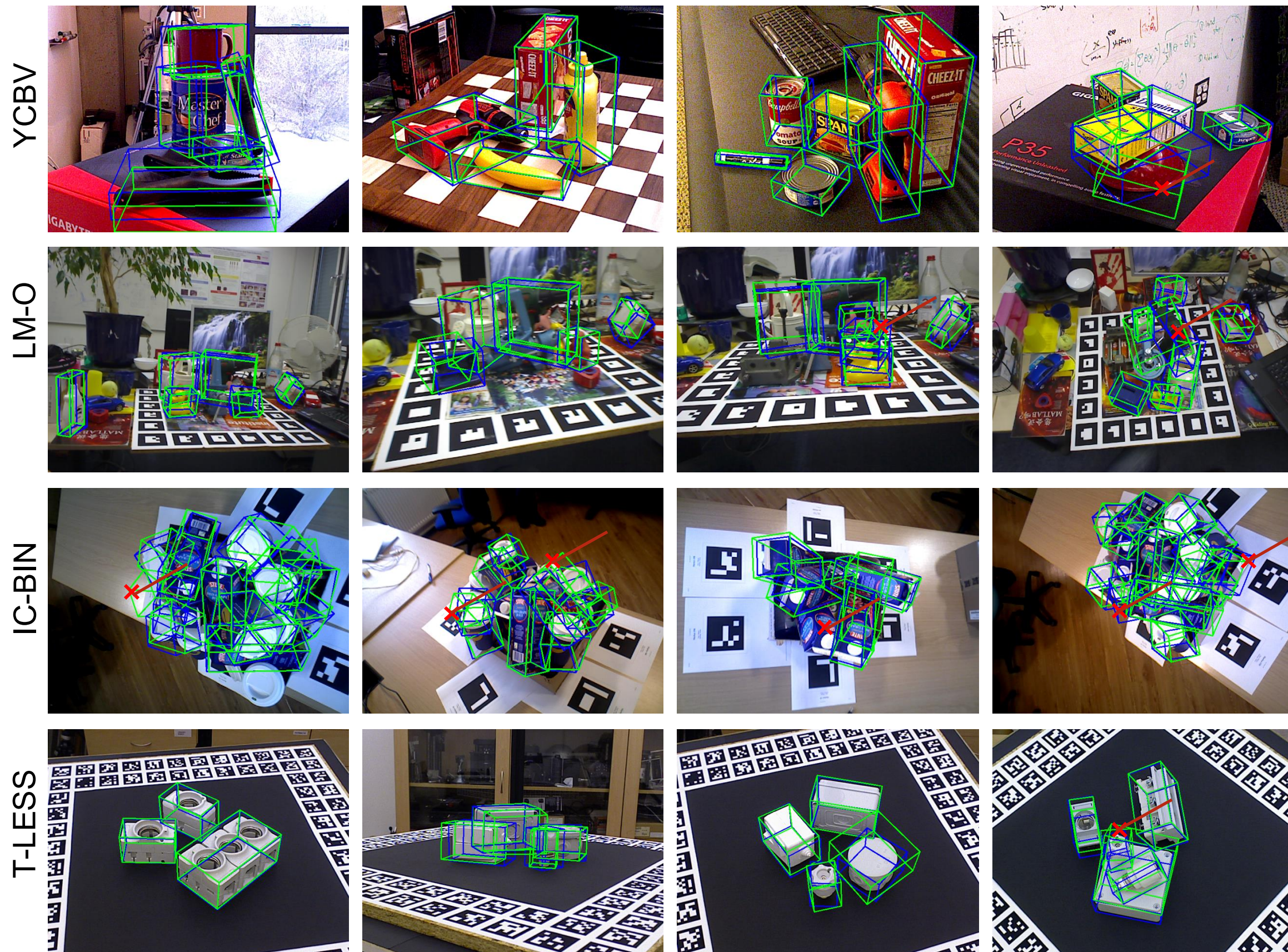
EXPERIMENTS

Performance Comparison on the BOP Challenge



- *Pos3R outperforms other methods in coarse pose estimations*
- *With pose refiner provided by MegaPose, Pos3R remains competitive*

EXPERIMENTS



- *Pos3R is robust to crowding, lighting changes, and texture-less objects*

- *Limitation: heavy occlusion (X) poses a challenge*

blue indicates ground truth; green indicates the estimate

PROJECT PAGE



Pos3R: 6D Pose Estimation for Unseen Objects Made Easy

Weijian Deng¹, Dylan Campbell¹, Chunyi Sun¹, Jiahao Zhang¹, Shubham Kanitkar², Matthew E. Shaffer², Stephen Gould¹

¹The Australian National University ²RIOS Intelligent Machines



I. Unseen Object Pose Estimation

Task Definition: Training Free, RGB, and CAD Model

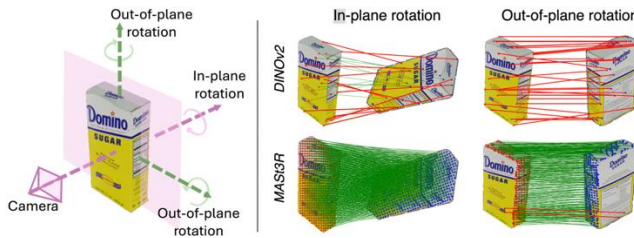
- Training-free pipelines offer adaptability to unseen objects
- Model-based 6D localization estimates object pose from a 3D CAD model and an RGB image

II. Why Use a 3D Foundation Model?

Motivation

- 2D foundation models (e.g., DINOv2) have been shown to be effective at training-free pose estimation, but are not consistent under significant 3D transformations
- 3D foundation models (e.g., MAST3R) predict 3D-consistent features, which we show to be useful for pose estimation

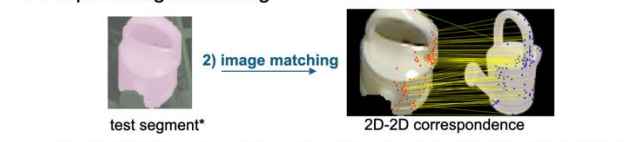
Correspondence Matching Quality



- DINOv2 produces inconsistent matches under out-of-plane rotations since it was not trained under these transformations
- MASt3R provides dense and stable correspondences

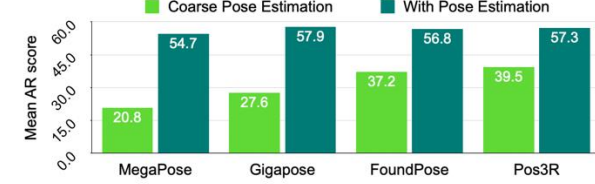
III. Pos3R

Pos3R: Training-Free and Fast — Render, Match, Fit

- Step 1: Template Rendering
 - Eight base template, covering essential orientations
 - For each, five in-plane (axial) rotations are generated around the camera's principal axis.
- Step 2: Image Matching
 - MASt3R produces dense 2D correspondences between test segment and every template
 - Similarity is computed by summing feature similarities across correspondences
- Step 3: Pose Fitting
 - The template with the highest score is selected
 - 3D coordinate map provides 2D-3D matches for PnP to estimate pose

IV. Experiments

Performance Comparison on the BOP Challenge

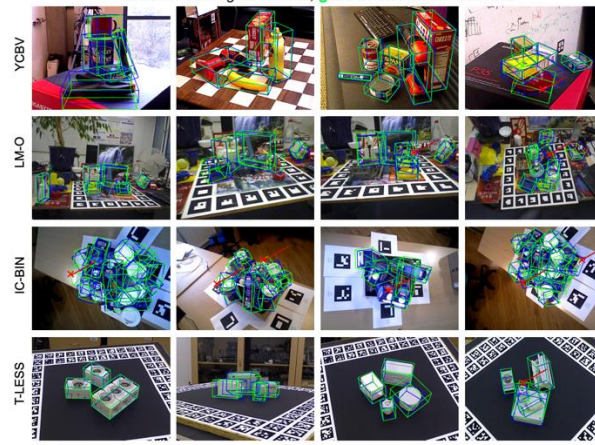


Method	Coarse Pose Estimation	With Pose Estimation
MegaPose	20.8	54.7
Gigapose	27.6	57.9
FoundPose	37.2	56.8
Pos3R	39.5	57.3

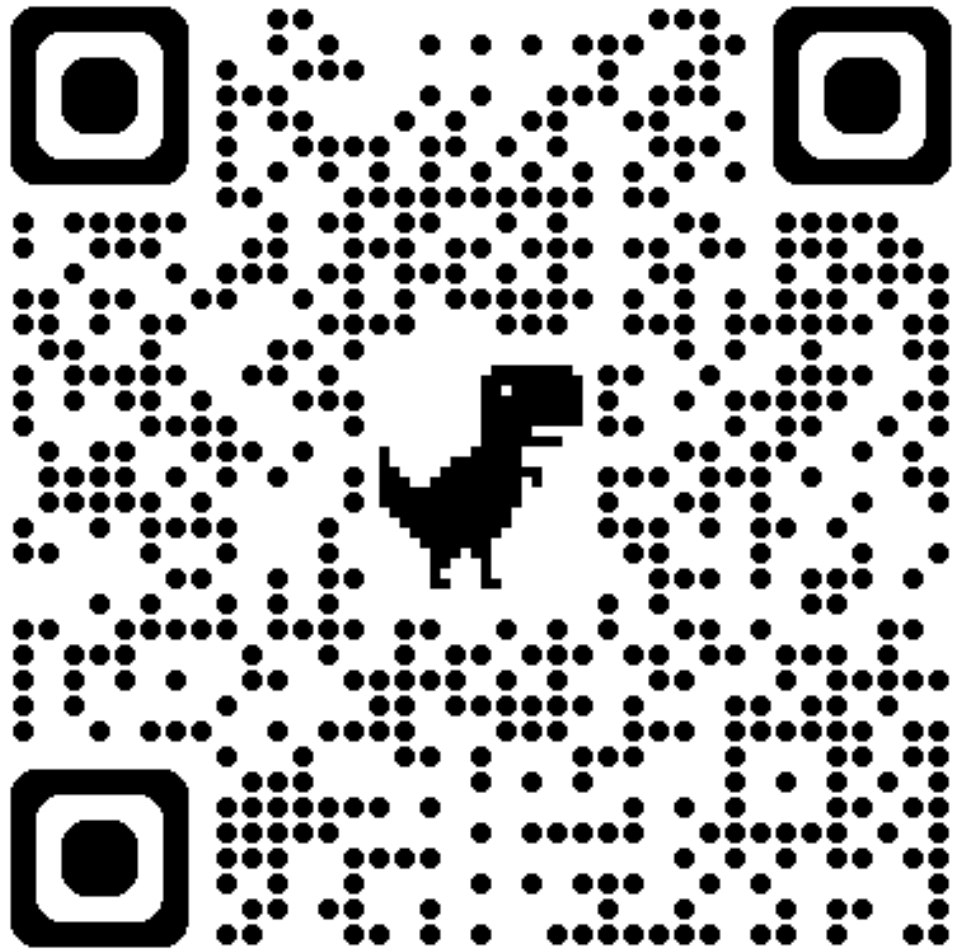
- Pos3R outperforms other methods in coarse pose estimations
- With pose refiner provided by MegaPose, Pos3R remains competitive

Qualitative Results of 6D Pose Estimates

blue indicates ground truth; green indicates the estimate



- Pos3R is robust to crowding, lighting changes, and texture-less objects
- Limitation: heavy occlusion (X) poses a challenge



Paper, code, and datasets

T H A N K Y O U !