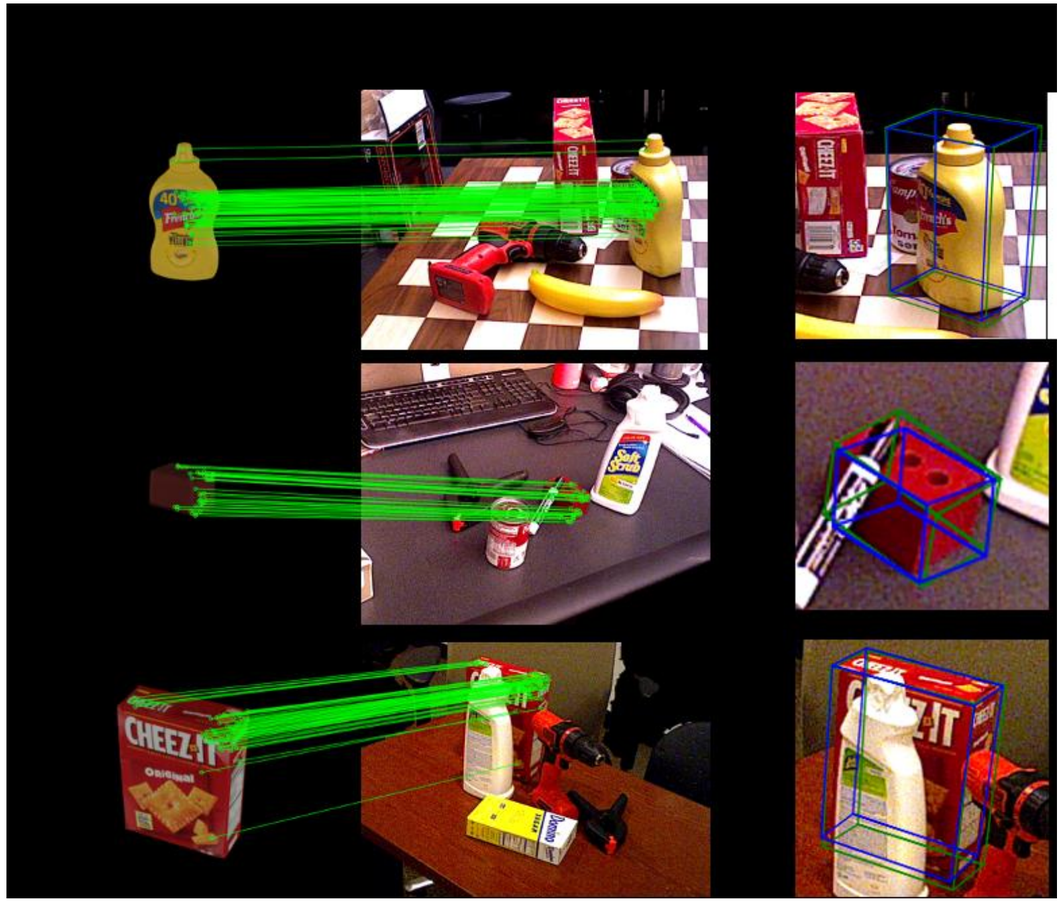
Pos3R: 6D Pose Estimation for Unseen Objects Made Easy

Weijian Deng[1] Dylan Campbell[1] Chunyi Sun[1] Jiahao Zhang[1]
Shubham Kanitkar[2] Matthew E. Shaffer[2] Stephen Gould[1]

[1]Australian National University [2]RIOS Intelligent Machines

6D OBJECT POSE ESTIMATION

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



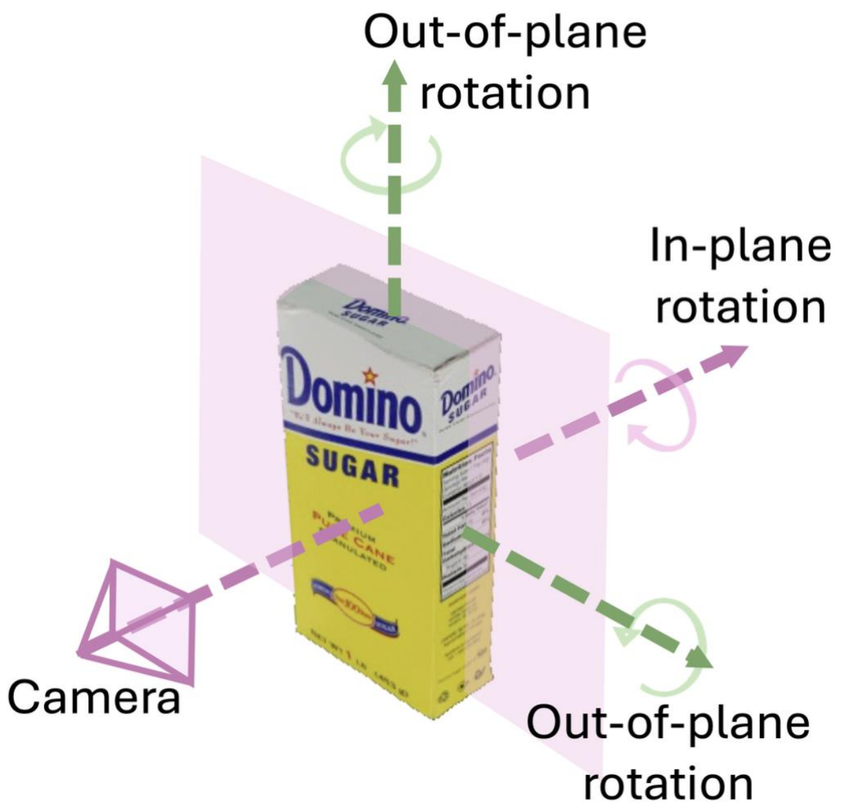
6D OBJECT POSE ESTIMATION

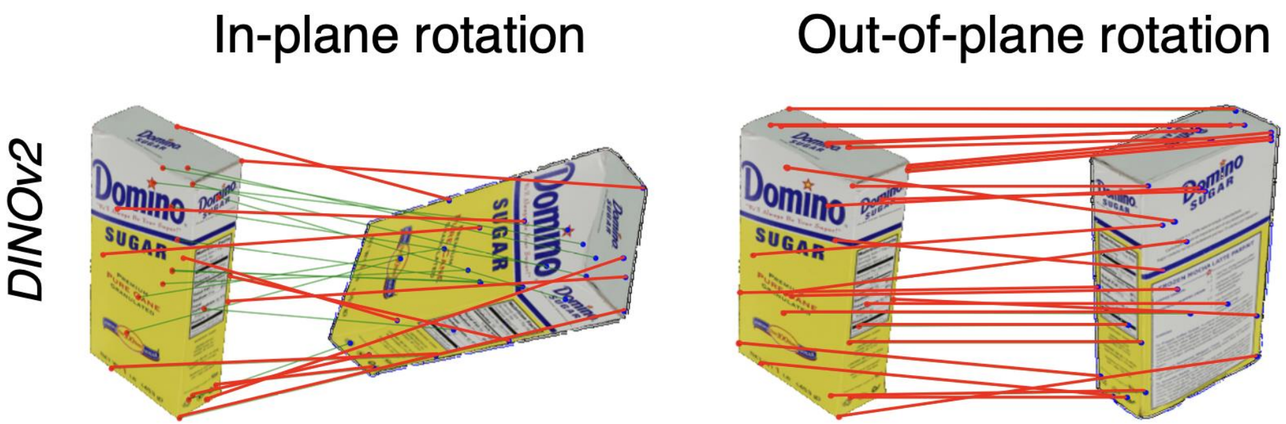
- Task Definition: Training Free, RGB-Only, and CAD Model

  - *Training-free pipelines offer adaptability to unseen objects*

  - *Model-based 6D localization estimates object pose from a 3D CAD model and an RGB image*

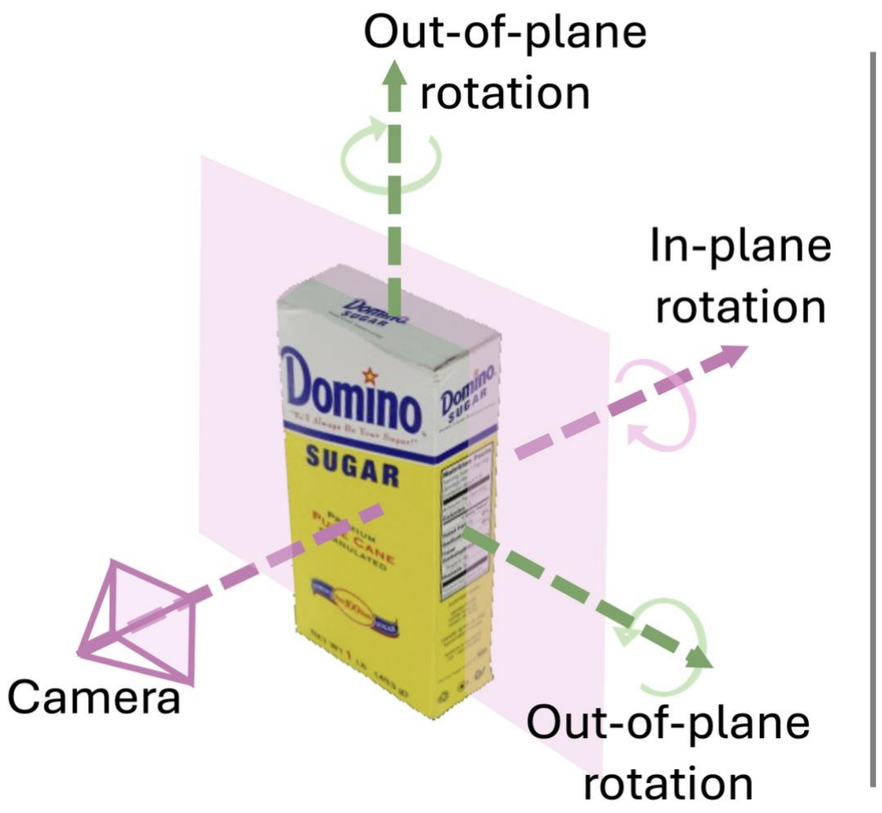
Why Use a 3D Foundation Model?

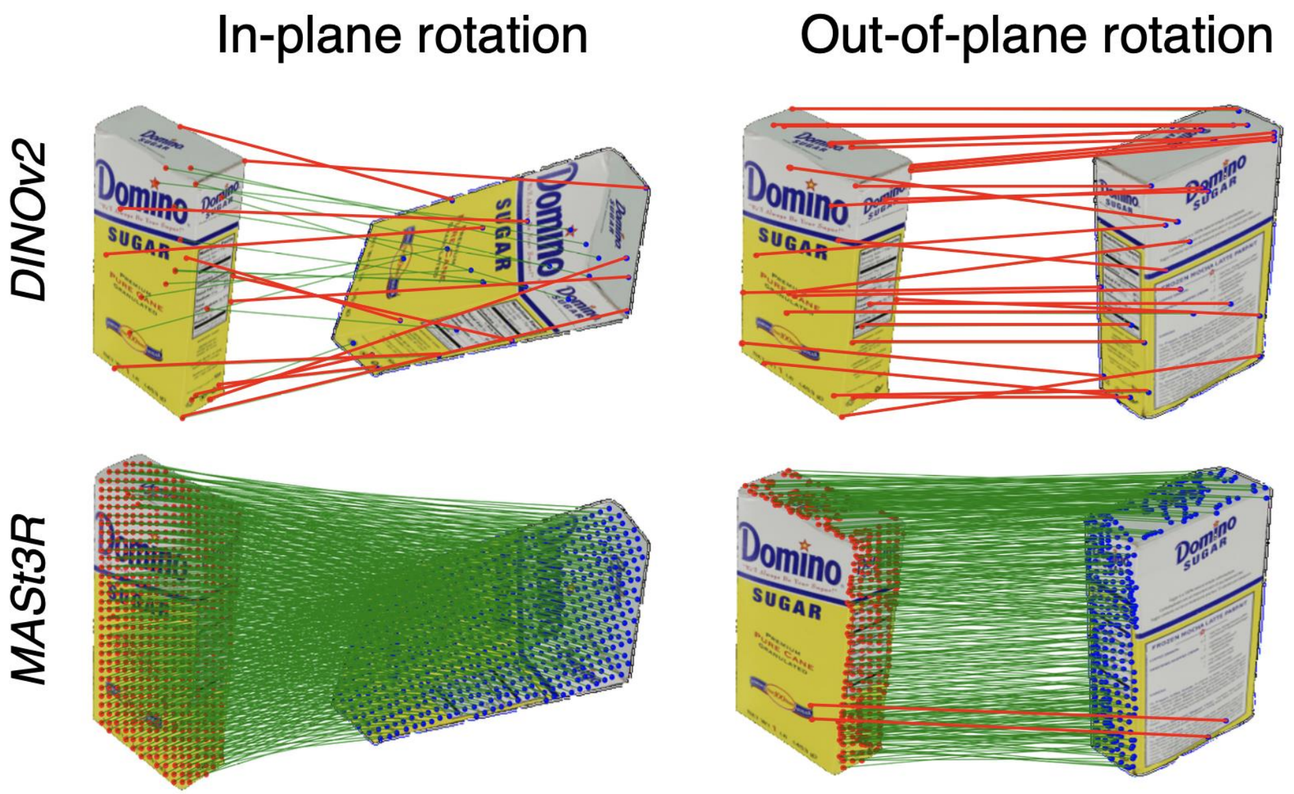




- *DINOv2 produces inconsistent matches under out-of-plane rotations since it was not trained under these transformations*

Why Use a 3D Foundation Model?

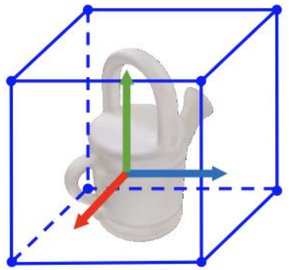




- *MASt3R provides dense and stable correspondences*

P o s 3 R

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



CAD model

1) rendering

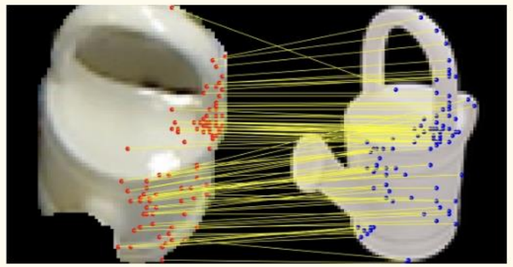
(sec. 3.2.1)

forty templates

test segment

2) image matching

(sec. 3.2.2)



2D-2D correspondence

coordinate map

3) P*n*P-RANSAC

(sec. 3.2.3)

6-DoF Pose

P o s 3 R

Step 1: Template Rendering

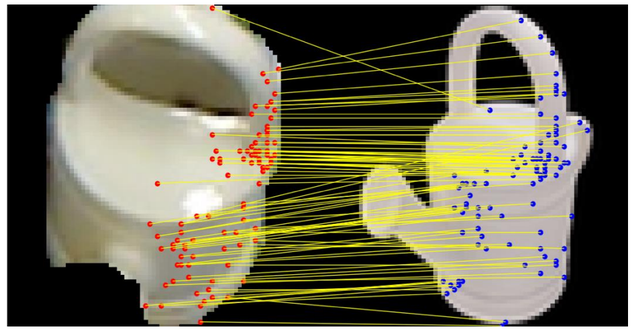
- *Eight base template, covering essential orientations*
- *For each, five in-plane (axial) rotations are generated around the camera's principal axis.*

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*

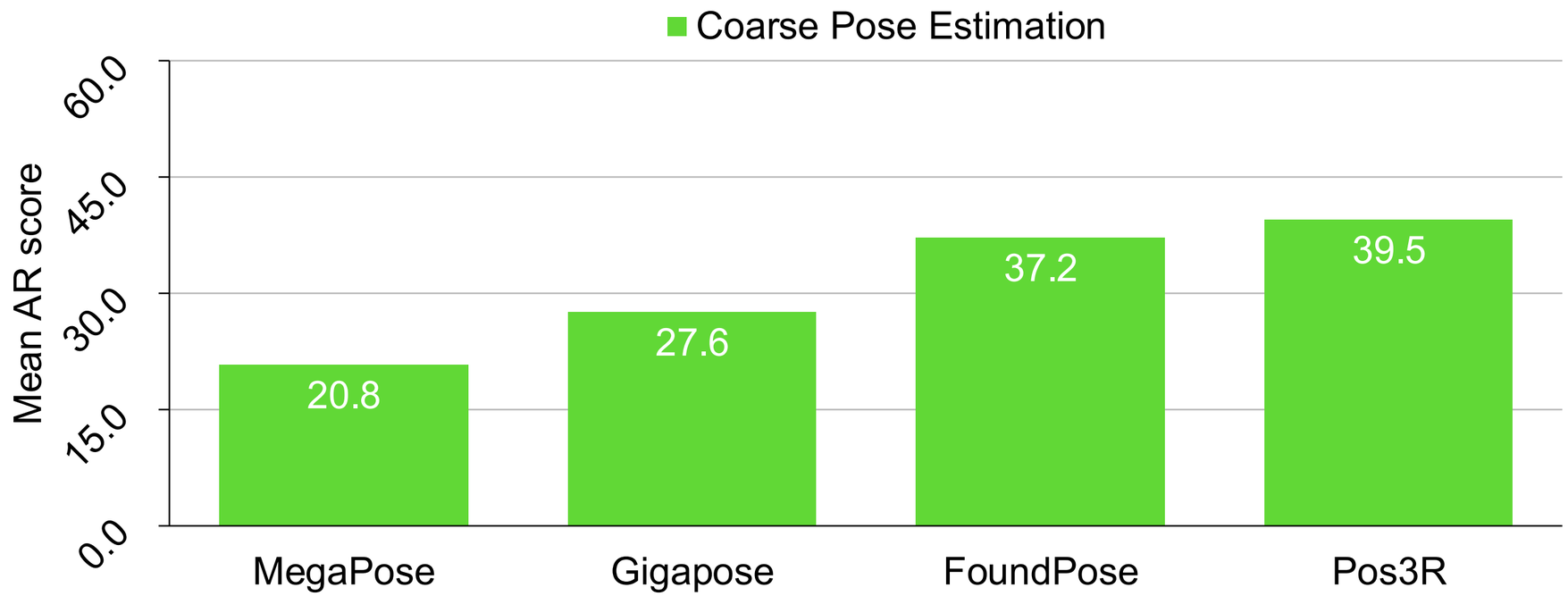
P o s 3 R

Step 3: Pose Fitting

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

E X P E R I M E N T S

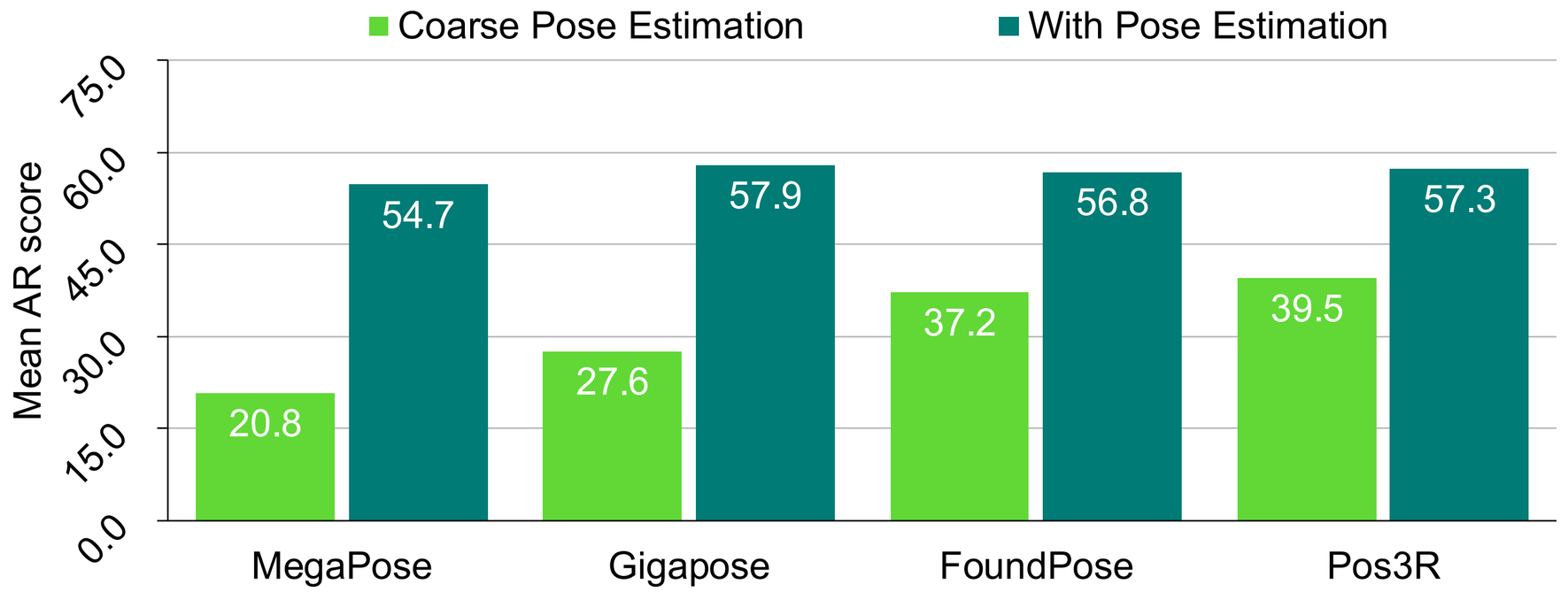
Performance Comparison on the BOP Challenge



- *Pos3R outperforms other methods in coarse pose estimations*

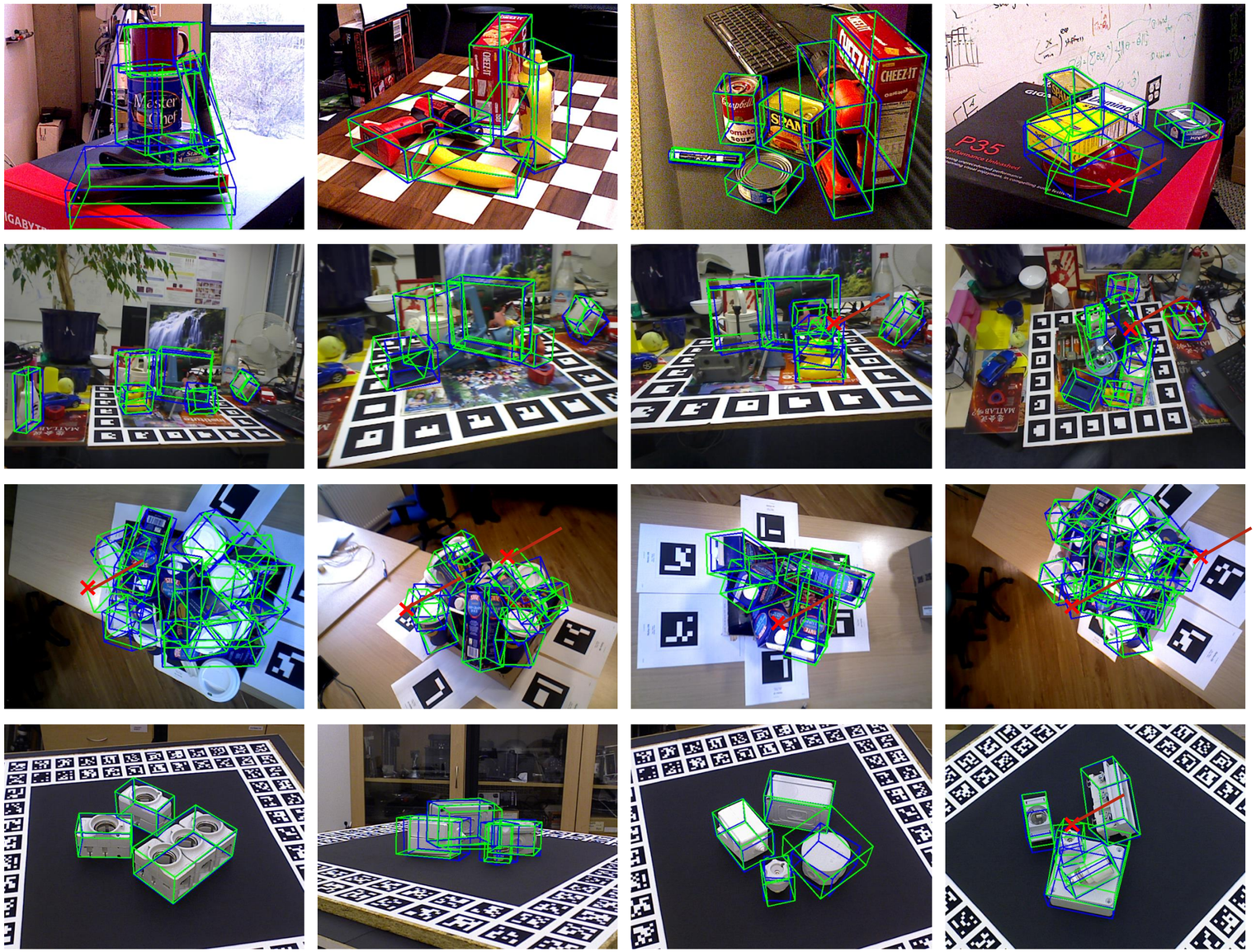
EXPERIMENTS

Performance Comparison on the BOP Challenge



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

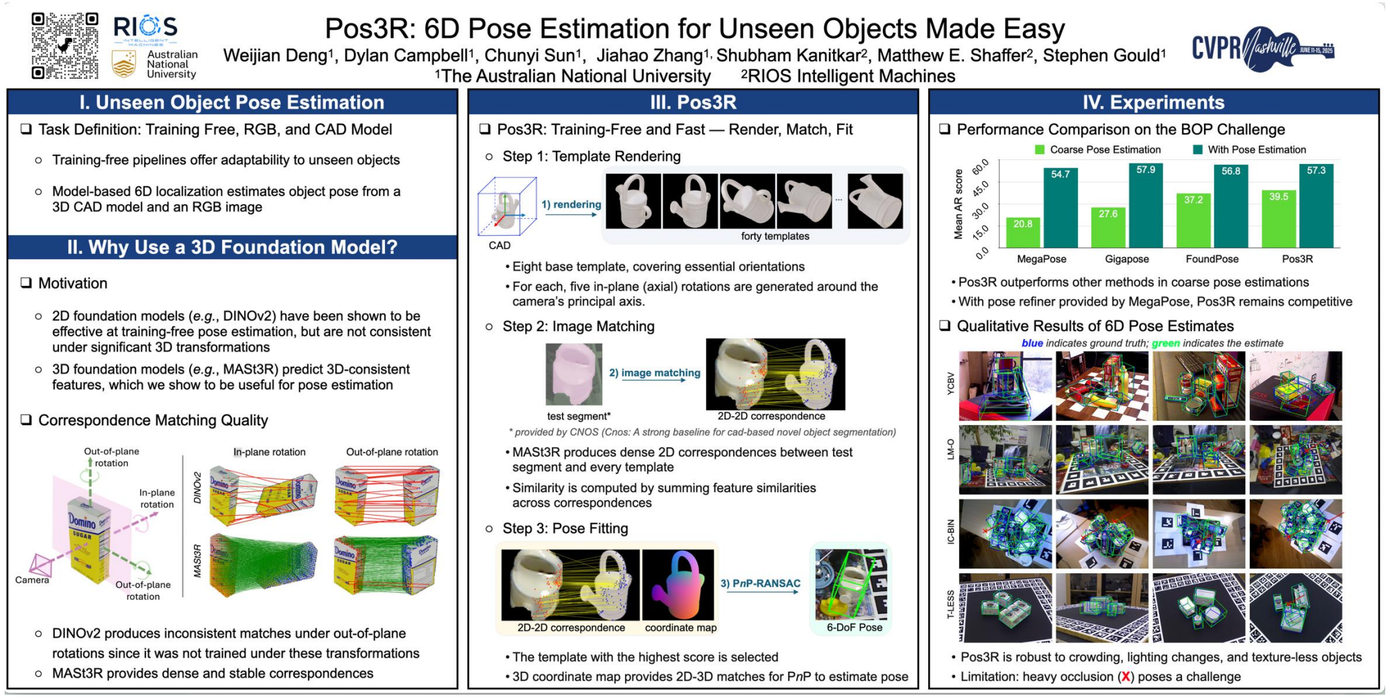
EXPERIMENTS



- *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



Paper, code, and datasets

T H A N K Y O U !