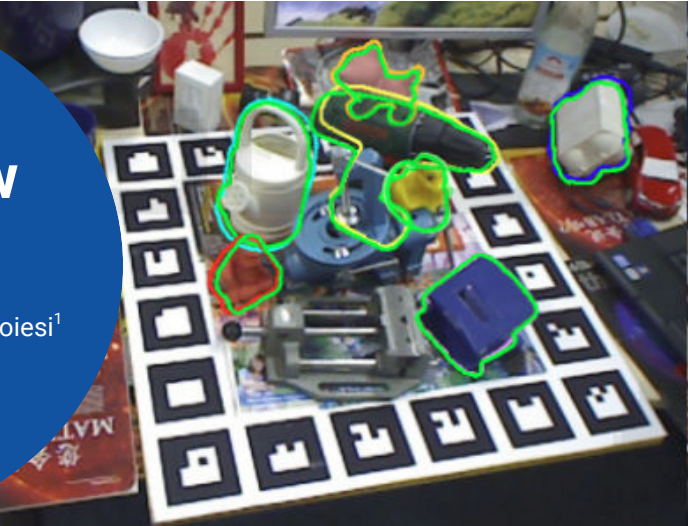


Flose : Generative 6D Pose Estimation via Conditional Flow Matching



Amir Hamza^{1,2}, Davide Boscaini¹, Weihang Li^{3,4}, Benjamin Busam^{3,4}, Fabio Poiesi¹

¹ Fondazione Bruno Kessler, ² University of Trento, ³ Technical University of Munich, ⁴ Munich Center for Machine Learning

Abstract & Motivation

- Goal: Robust 6D object pose estimation (3D translation & rotation) from single-view RGB-D data and CAD model.
- Existing 6D pose estimation methods fall into two camps, each with key limitations:
 - Direct methods regress pose in $SE(3)$ but struggle with object symmetries – multiple valid poses collapse to a single supervision signal.
 - Indirect methods match local features to recover correspondences but fail on objects lacking distinctive textures.

We formulate 6D pose estimation as conditional flow matching in \mathbb{R}^3 – learning a displacement field instead of explicit correspondences

Key Contributions

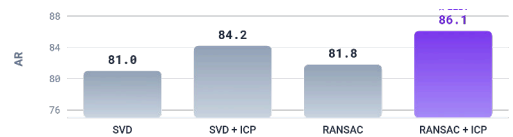
- First CFM formulation for instance-level object 6D pose estimation, learning dense displacement fields via generative denoising.
- Semantic feature injection from a frozen DINOv2 VFM to disambiguate symmetries and occlusions that purely geometric methods cannot resolve.
- RANSAC-based registration to robustly filter displacement outliers inherent to the denoising process, replacing fragile SVD global alignment.

Ablation Study

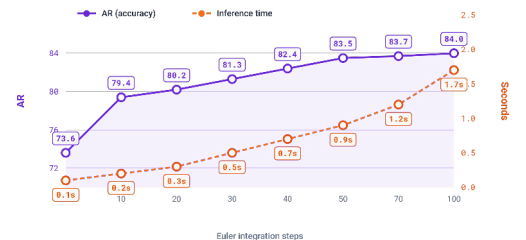
Average Recall (AR) and Inlier Ratio (IR) on LM-O under different feature conditioning



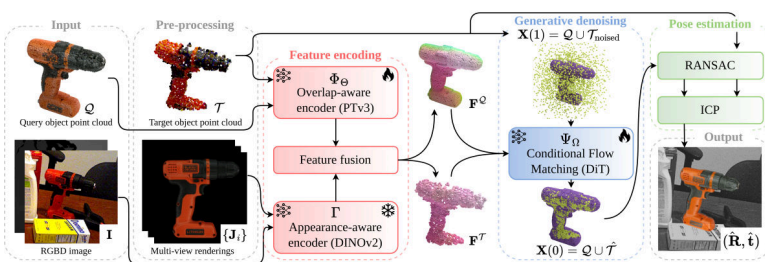
Comparison of Pose Estimation & Refinement Strategies



Effect of Denoising Iterations on AR & Inference Time



Proposed pipeline



Quantitative Results

Method	LM-O	T-LESS	TUD-L	IC-BIN	YCB-V	Avg
ZebraPose	75.2	72.7	94.8	65.2	86.6	78.9
GDRNPP (BOP23)	79.4	91.4	96.4	73.7	92.8	86.7
HccePose(BF)	80.5	87.9	94.4	72.4	91.1	85.3
PFA (S.M.)	79.7	85.0	96.0	67.6	88.8	83.4
SurfEmb (S.M.)	75.8	83.3	93.3	65.6	82.4	80.1
CIR (S.M.)	73.4	77.6	96.8	67.6	89.3	81.0
Flose (ours)	86.1	86.9	98.8	74.8	92.8	87.9

Qualitative Results

