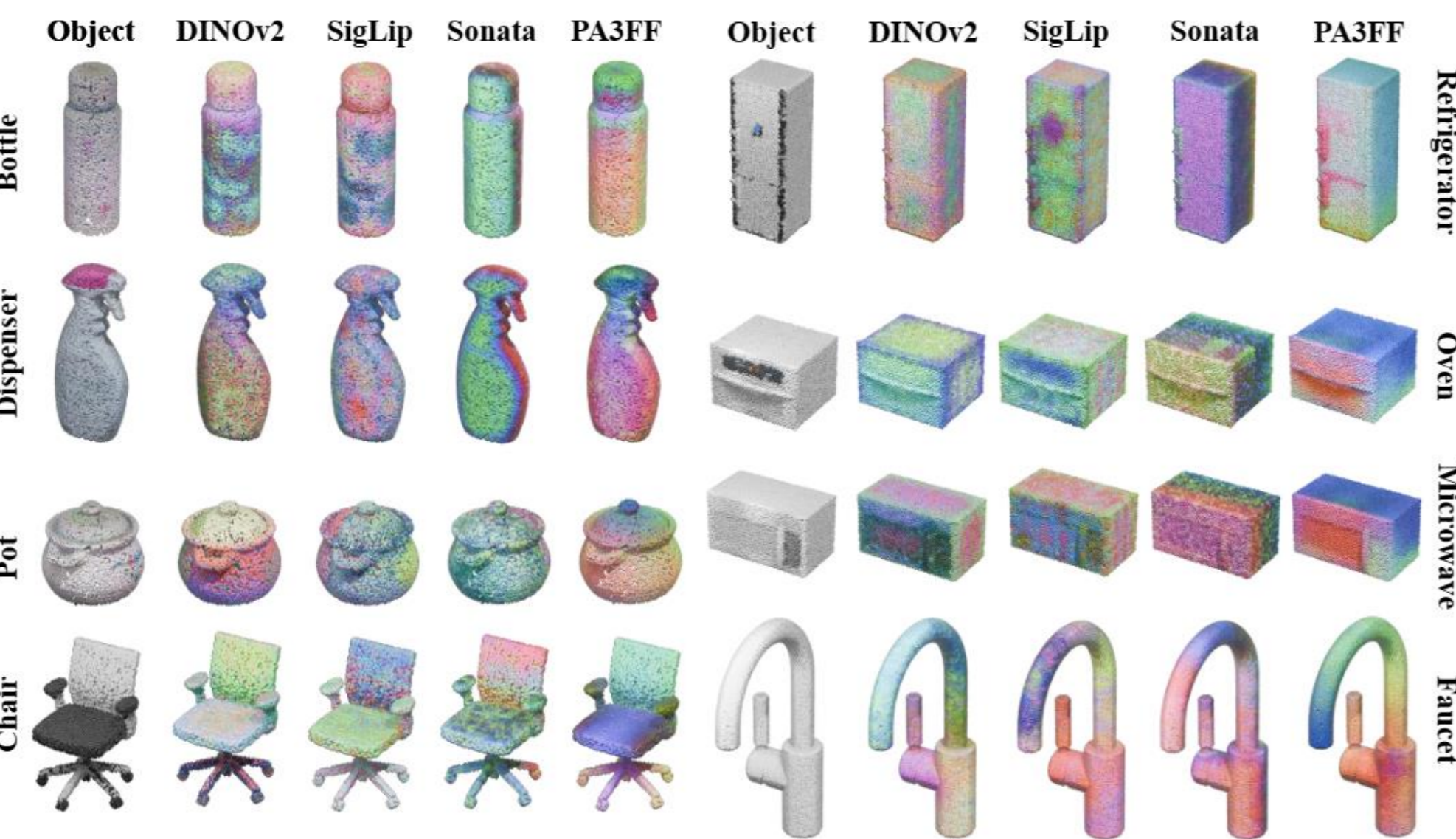
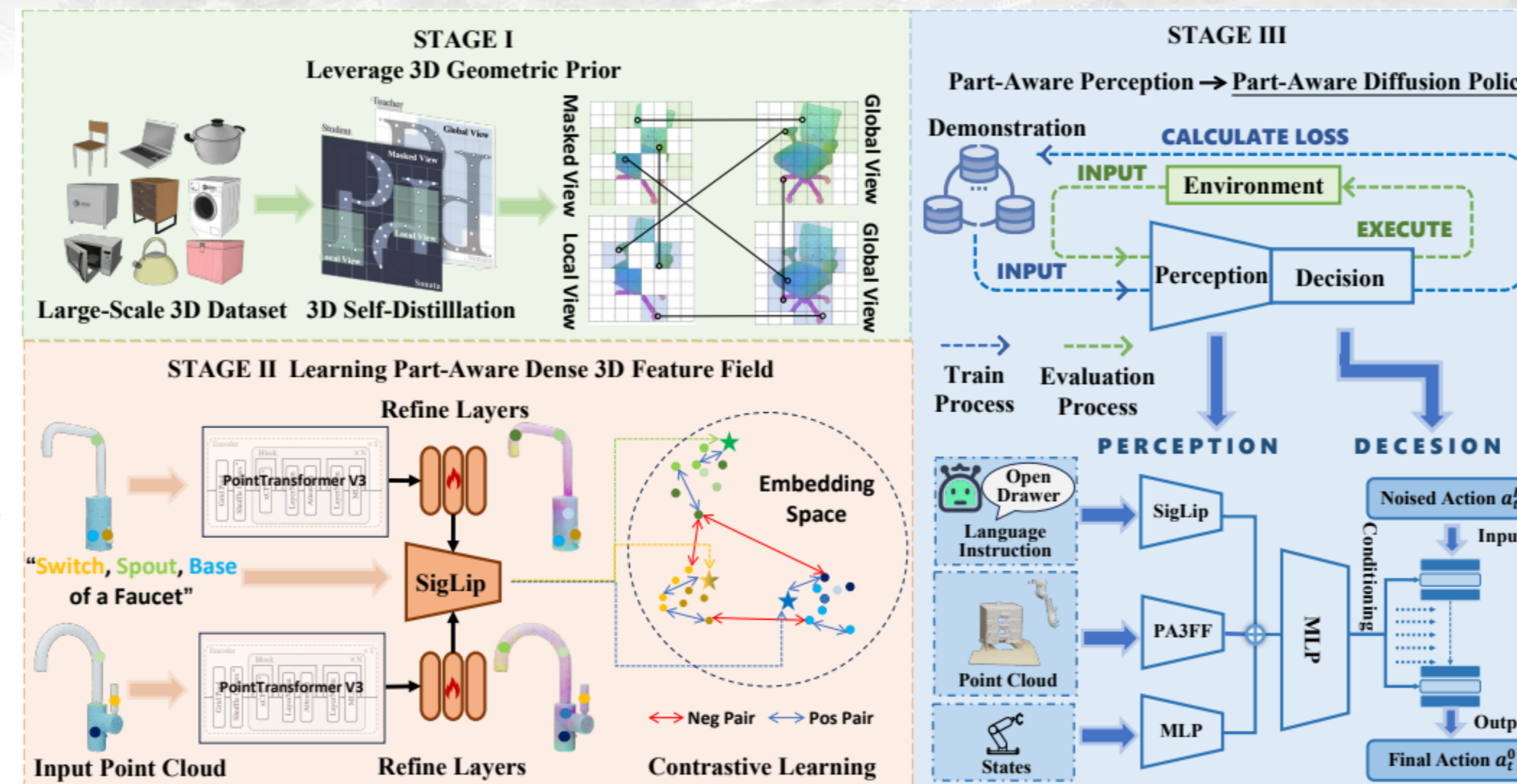


## Introduction

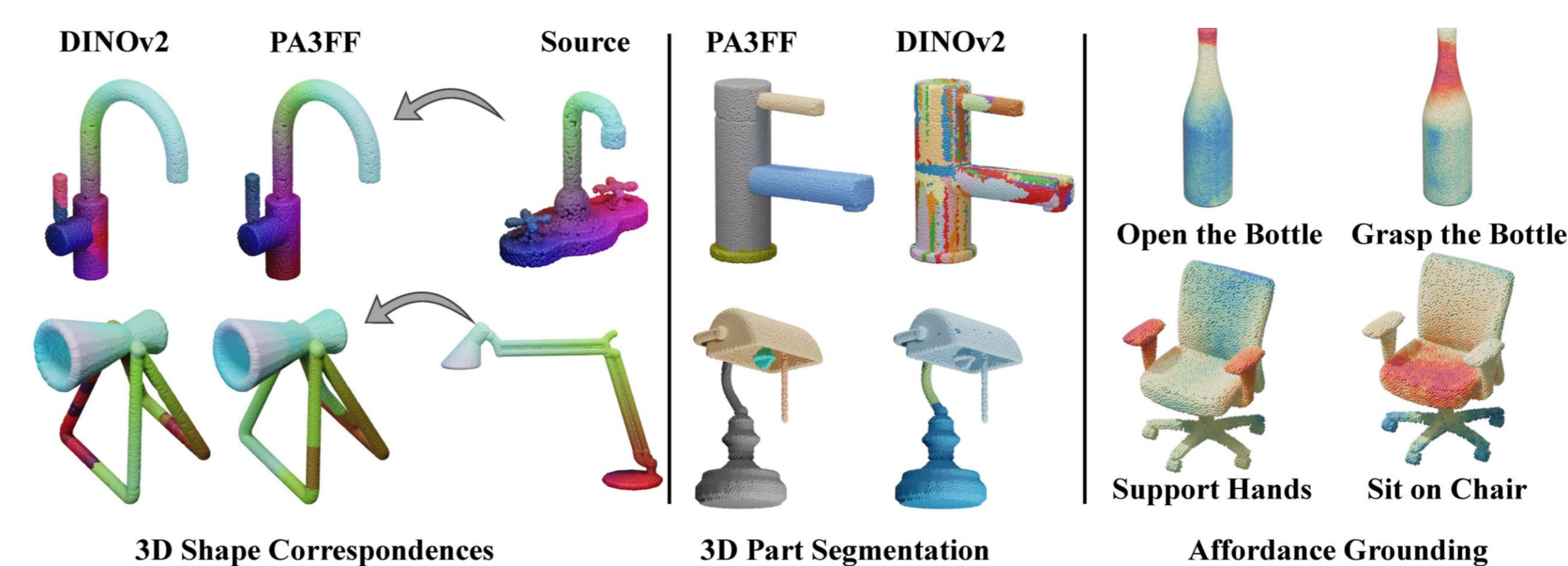
Generalizing robotic manipulation across diverse objects requires understanding *functional parts*. Prior approaches rely on lifted 2D features, leading to *long inference times*, *feature inconsistency*, and *limited spatial resolution*. We propose PA3FF, which encodes *dense*, *semantic*, and *functional part-aware* features from *point clouds*. PA3FF can further enable diverse downstream methods, including correspondence learning and segmentation, making it a versatile foundation for robotic manipulation.



## Method



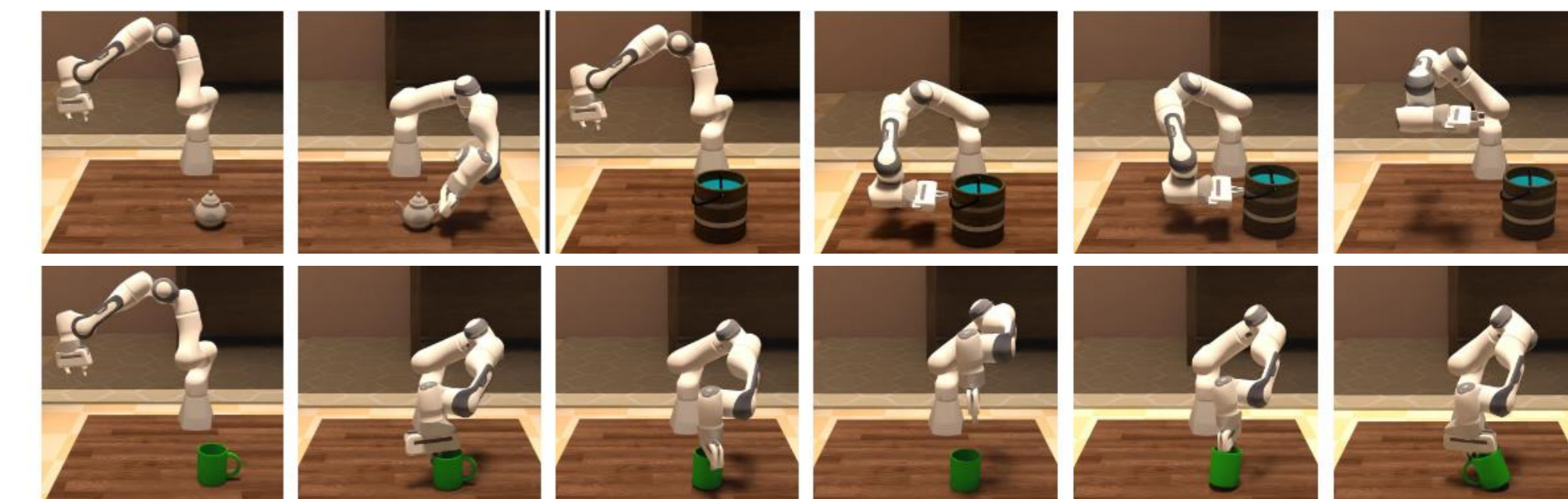
## Downstream Applications



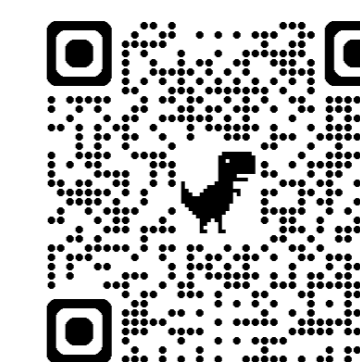
PA3FF can further enable diverse downstream methods, including correspondence learning, segmentation and affordance grounding.

## Experiments

We validate our approach on 16 PartInstruct and 8 real-world tasks, where it significantly outperforms prior 2D and 3D representations (CLIP, DINOv2, and Grounded-SAM), offering a 15% and 16.5% increase.



## Links



Website



Full Paper



WeChat