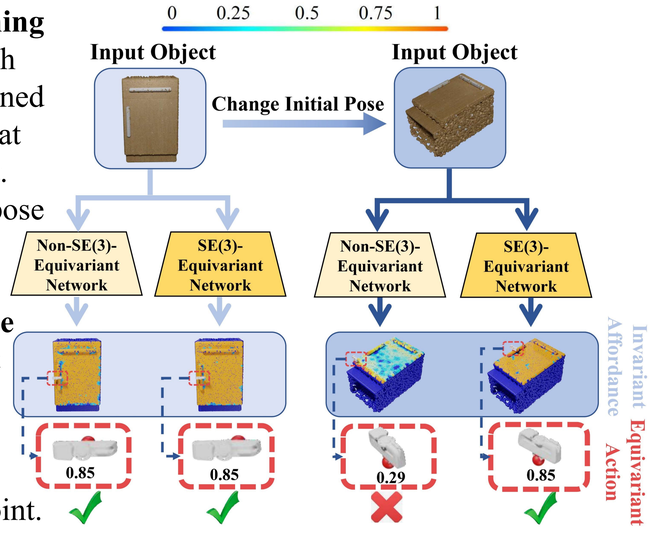
EqvAfford: SE(3) Equivariance for Point-Level Affordance Learning

YueChen*, Chenrui Tie*,Ruihai Wu*, Hao Dong (* denotes equal contribution)

Affordance Learning

Affordance Learning aims to interact with scenarios, and trained models to learn what and how to interact. Therefore, we propose a framework that leverages **SE(3) equivariance** to predict per-point affordance and downstream manipulation strategy on each point.



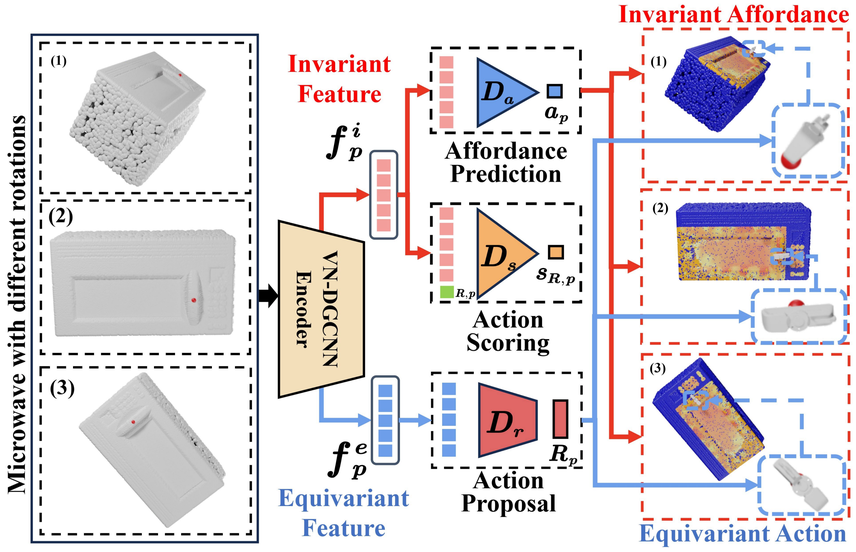
SE(3) equivariance & invariance

SE(3), which represents the Special Euclidean Group in three dimensions, characterizes the rigid body motion in 3D space, including 3D translation and rotation.

Given an input point cloud $\mathcal{P} \in \mathbb{R}^{n\times 3}$, for any rotation matrix $R \in \mathbb{R}^{3\times 3}$ and translation vector $T \in \mathbb{R}^3$.

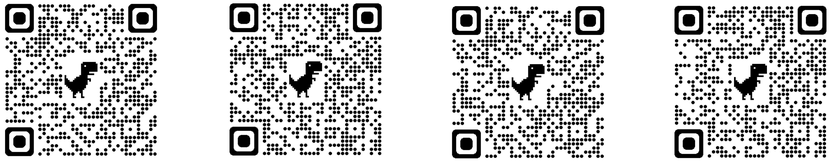
A SE(3) equivariant encoder $\mathcal{E}_{equiv}$ means $\mathcal{E}_{inv}(\mathcal{P}R+T) = \mathcal{E}_{inv}(\mathcal{P})$ and a SE(3) invariant encoder $\mathcal{E}_{inv}$ means $\mathcal{E}_{equiv}(\mathcal{P}R+T) = \mathcal{E}_{equiv}(\mathcal{P})R+T$

Method



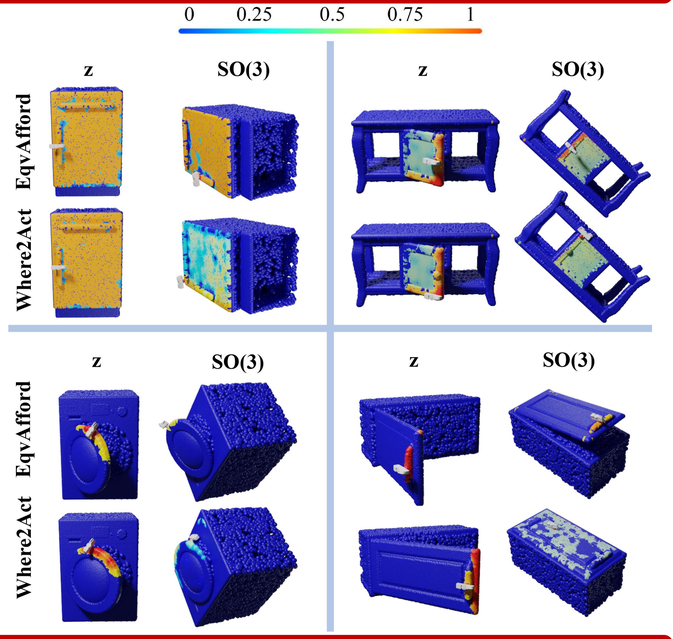
Taking as input a point cloud of the object, our framework first outputs a per-point SE(3) invariant feature f_p^i and SE(3) equivariant feature f_p^e. The invariant f_p^i results in the affordance map invariant to object rotations, while the equivariant feature f_p^e results in the manipulation actions equivariant to object rotations.

Links



Yue Chen | Chenrui Tie | Ruihai Wu | Hao Dong

Experiments



Conclusion

We propose a novel framework, **EqvAfford**, leveraging SE(3) equivariance for affordance learning and downstream robotic manipulation with novel designs to theoretically guarantee equivariance. Predicting per-point SE(3) **Invariant affordance** and **Equivariant interaction orientation**, our method generalizes well to diverse object poses. Experiments on affordance learning and robotic manipulation showcases our method qualitatively and quantitatively.