Keypoint-GraspNet: Keypoint-based 6-DoF Grasp Generation from the Monocular RGB-D input

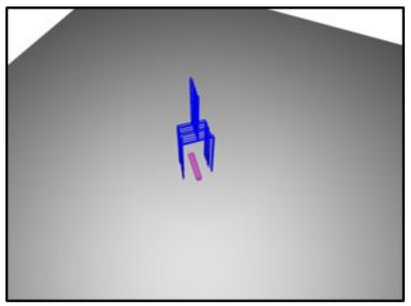
Yiye Chen, Yunzhi Lin, Ruinian Xu, Patricio Vela

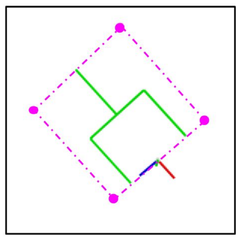
ICRA LONDON · 2023

Georgia Tech Robotics & Intelligent Machines

1. Motivation & Goal

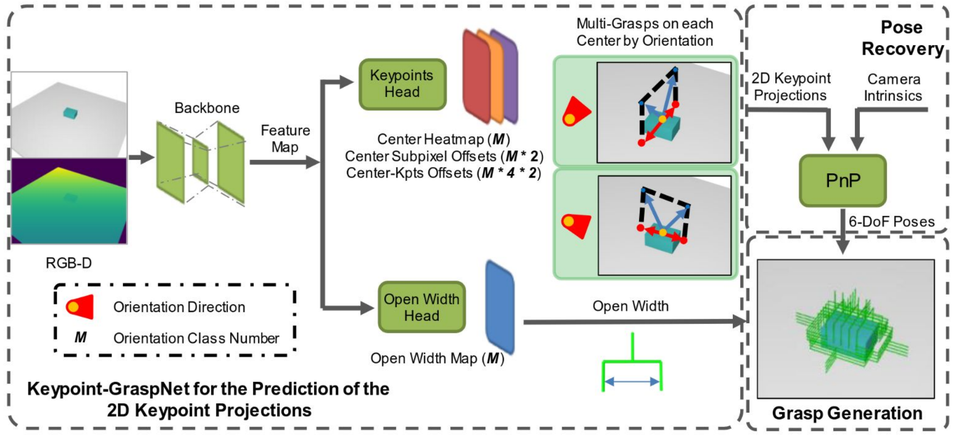
- Prior work in **3D Grasp synthesis** relies on point cloud input, which is time-consuming to process.
- **We aim to detect 3D grasps from the 2D image input.**





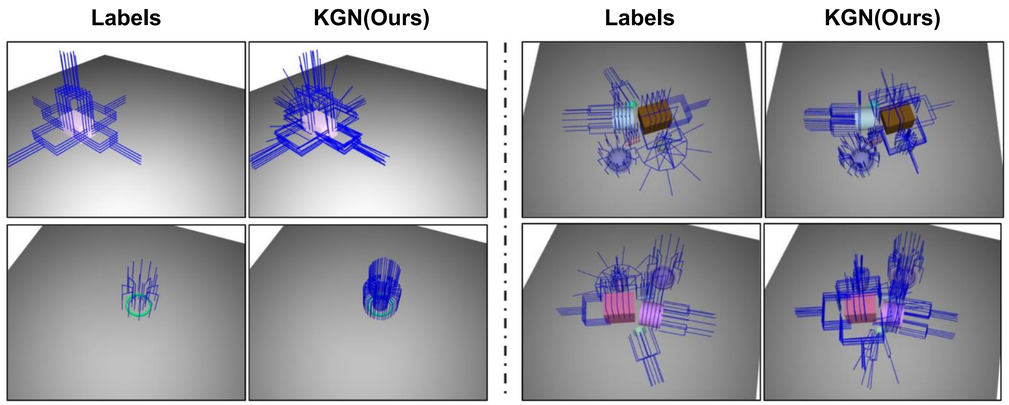
We adopt the keypoint representation to bridge the 2d-3d dimensionality gap

2. Method



- We detects keypoints **in the image space**, then recover 3D grasp poses using **PnP algorithm**.
- **Primitive shapes** (cuboid, sphere, etc) are used for training, which is **simple**, **free from artifacts of sampling**, and allows **arbitrary label density**.

3. Results

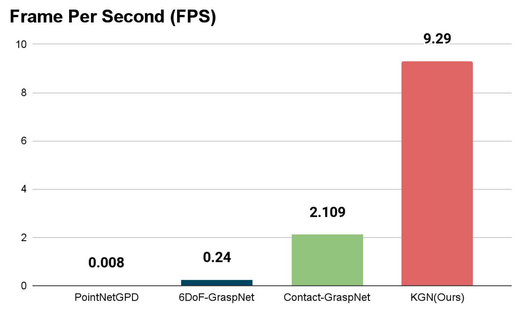


Methods	Modality	Single-Object Evaluation (GSR% / GCR% / OSR%)			Multi-Object Evaluation (GSR% / GCR% / OSR%)		
		1cm + 20°	2cm + 30°	3cm + 45°	1cm + 20°	2cm + 30°	3cm + 45°
PointNetGPD	PC	0.43 / 0.13 / 1.50	1.52 / 0.90 / 3.57	20.5 / 4.80 / 16.0	0.00 / 0.00 / 0.00	8.33 / 0.02 / 0.80	41.7 / 0.33 / 3.20
6DoF-GraspNet	PC	3.78 / 6.78 / 35.4	16.5 / 39.6 / 79.1	35.9 / **73.9** / 97.7	0.20 / 0.10 / 0.70	2.00 / 0.50 / 5.27	8.66 / 2.68 / 16.7
Contact-Graspnet[1]	PC	29.9 / 24.9 / 77.0	60.1 / 32.0 / 81.7	81.6 / 36.5 / 84.2	**22.1** / **15.5** / **44.1**	**54.2** / **28.5** / 51.4	**78.4** / **34.5** / 54.4
KGN	RGB-D	**55.5** / **42.9** / **97.0**	**78.5** / **63.3** / **99.6**	**86.9** / 73.2 / **99.9**	10.8 / 5.48 / 28.7	30.6 / 18.7 / **51.8**	49.6 / 33.8 / **62.4**

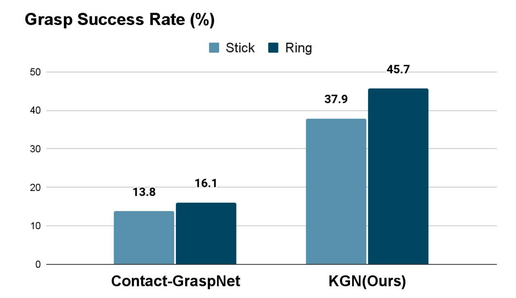
Our method **outperforms** baselines on **single-object** benchmark, and has **comparable performance** on **multi-object environments** with clutter methods **despite trained only with single-object data**.

4. Highlights

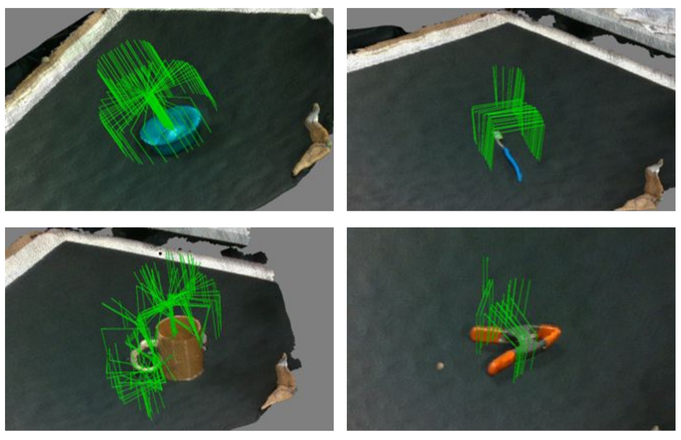
- **Faster**



- **Better at small objects**



5. Robotic Experiments



We perform physical experiments on objects with **diverse appearance and shapes**. Although **trained only with simple primitive shapes**, our method achieves **88% success rate** comparable with well-known 6DoF-GraspNet while being **20 times faster.**