The CasiaZhuasape is a robotic grasp dataset constructed by the Human-Computer Interaction Group, the Research Center for Brain-inspired Intelligence, Institute of Automation, Chinese Academy of Sciences. It contains 1498 images with 640x480 resolution ratio of 169 objects, and these objects were taken at various angles and scales for 15 shapes, including rectangle (140 images for 13 objects), dumbbell (100 images for 8 objects), fusiform (147 images for 11 objects), cross (108 images for 9 objects), trapezoid (86 images for 9 objects), strip (163 images for 15 objects), circle (75 images for 30 objects), loop (53 images for 11 objects), tstyle (181 images for 12 objects), ellipse (75 images for 10 objects), arc (70 images for 6, objects), irregular (103 images for 13 objects), star (76 images for 10 objects), triangular (61 images for 5 objects), square (50 images for 7 objects).

The directory structure of the “CasiaZhuasape” is as follow:

----CasiaZhuasape

--------Original

------------arc

------------circle

------------…..

--------BaseDetection

------------arc

----------------result\_txt

------------circle

----------------result\_txt

----------------…..

All files in the “Original” directory are the original images.

All images in the “BaseDetection” directory are the grasp visualization results obtained by the grasp prediction network (GP-Net) described in [Ribeiro2021] to generate the possible grasp anchors for the original images. The GP-Net network is trained on the mixed dataset of all Cornell images [Jiang2011] and 10,000 random Jacquard images [Depierre2018].

All files in the “result\_txt” directory give the detailed descriptions of the grasp detection results in the format as follows:

“Image\_name, confidence, x1, y1, x2, y2, x3, y3, x4, y4, ShortSide\_forGripper” or “Image\_name, confidence, x1, y1, x2, y2, x3, y3, x4, y4, LongSide\_forGripper”.

In the format of “Image\_name, confidence, x1, y1, x2, y2, x3, y3, x4, y4, ShortSide\_forGripper (LongSide\_forGripper)”, the item “Image\_name” means the image name of the detection image, “x1, y1, x2, y2, x3, y3, x4, y4” are the x, y coordinates of the four corner points of the detection grasp anchor, the item “ShortSide\_forGripper” means that in the four sides of the grasp anchor formed by the four points x1, y1, x2, y2, x3, y3, x4, y4, the two shorter sides of the grasp anchor in parallel are used as the location of the robotic gripper. And by the contrary, “LongSide\_forGripper” means that the two longer sides of the grasp anchor in parallel are used as the location of the robotic gripper.

Any questions, please contact [mhyang@nlpr.ia.ac.cn](mailto:mhyang@nlpr.ia.ac.cn). Thanks!

References:

[Ribeiro2021] Eduardo Godinho Ribeiro, Raul de Queiroz Mendes, Valdir Grassi Jr, Real-time deep learning approach to visual servo control and grasp detection for autonomous robotic manipulation, Robotics and Autonomous Systems, 139, 1037-1057, 2021.

[Jiang2011] Yun Jiang, Stephen Moseson, Ashutosh Saxena, Efficient grasping from RGBD images: Learning using a new rectangle representation,

2011 IEEE International Conference on Robotics and Automation (ICRA 2011), Shanghai, China

[Depierre2018] Amaury Depierre, Emmanuel Dellandréa, Liming Chen, Jacquard: A Large Scale Dataset for Robotic Grasp Detection, 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain