The structure of the dataset is the following:

* nameOfObject1: a folder containing all the data for the object. All the files inside this folder are prefixed with \* = i\_nameOfObject1, where i is the index of the scene (between 0 and 4) for the current object.
  + \*\_grasps.txt: a text file with the grasps annotations. Each line in the file is one grasp written as x;y;theta in degrees;opening;jaws size. Please note that all values are in image coordinates, so they are expressed in pixels, y is toward the bottom of the image (and therefore the angle is horizontally mirrored). When the position is the same on multiple consecutive rows, the first one corresponds to the grasp with the default jaws size of 2 cm and the followings are just repetition of this grasp with different sizes.

A picture containing text, person, businesscard

Description automatically generated

* + \*\_mask.png: a png image with a binary mask (0 for background) of the object which can be used to perform background augmentation on the data for example.
  + \*\_perfect\_depth.tiff: a float32 tiff image with the perfect depth image.
  + \*\_RGB.png: a png image of the scene rendered with Blender.
  + \*\_stereo\_depth.tiff: a float32 tiff image with the depth obtained from an off-the-shelf stereovision algorithm. Pixels with an undetermined depth (because of occlusion or bad matching during the stereovision process) have a value of -1.
* nameOfObject2
  + \*\_grasps.txt
  + \*\_mask.png
  + \*\_perfect\_depth.tiff
  + \*\_RGB.png
  + \*\_stereo\_depth.tiff
* nameOfObject3
  + \*\_grasps.txt
  + \*\_mask.png
  + \*\_perfect\_depth.tiff
  + \*\_RGB.png
  + \*\_stereo\_depth.tiff