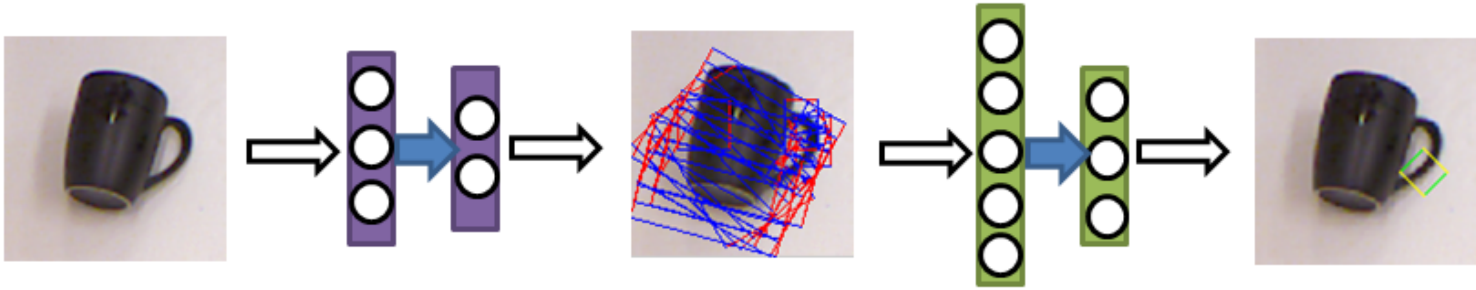
GAZEBO SIMULATED ENVIRONMENT FOR ROBOTICS MANUFACTURING

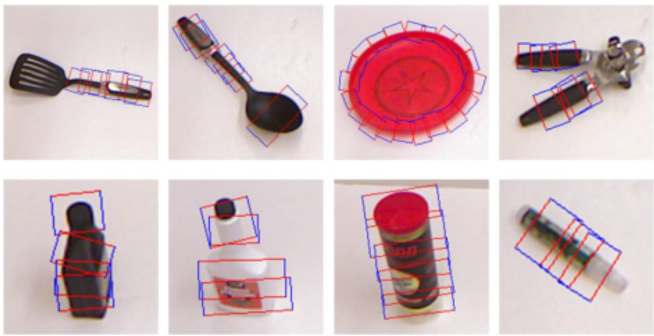
KONICA MINOLTA LABORATORY EUROPE

KONICA MINOLTA

CNN has achieved the state of arts solutions:

- Manually-labelled images: human labellers have determined the location of a suitable grasp point on an image of an object (not suitable for larger-scale training)

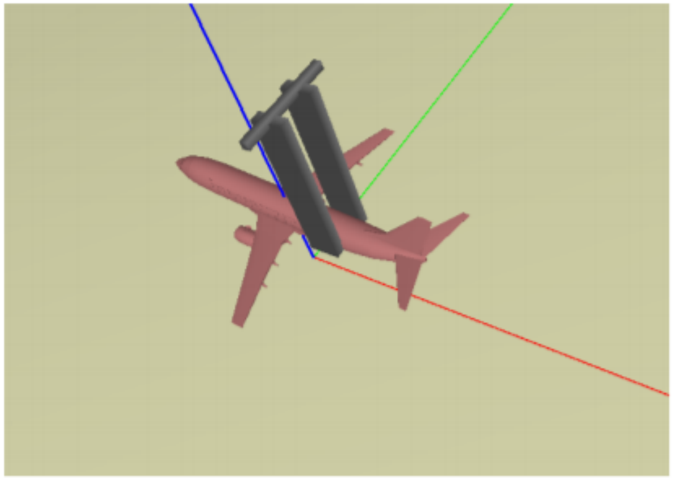




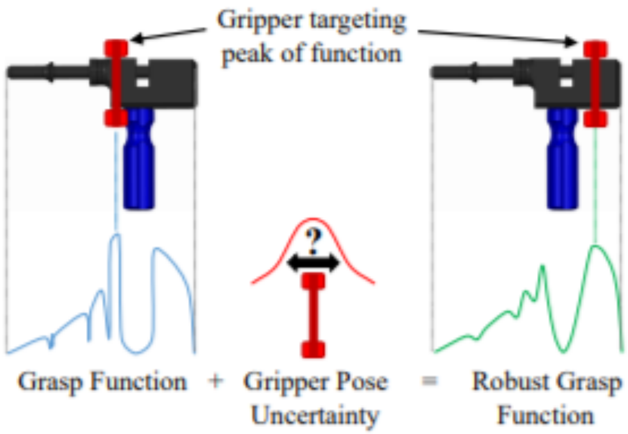
Cornell dataset: www.cs.cornell.edu/home/llee/data/

CNN has achieved the state of arts solutions:

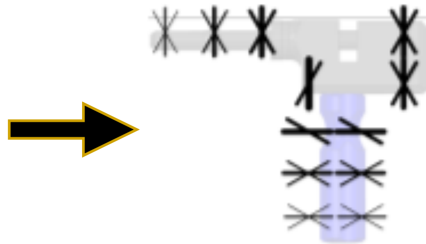
- Generate training data in simulation: Gazebo, GraspIt! & DART simulators for grasp of 3D meshes f objects



Simulated, virtual pick and place example



Estimation of suitable grasping points



Output scores

Introduction to Robot Grasping | Learning based approaches (3/3)

CNN has achieved the state of arts solutions:

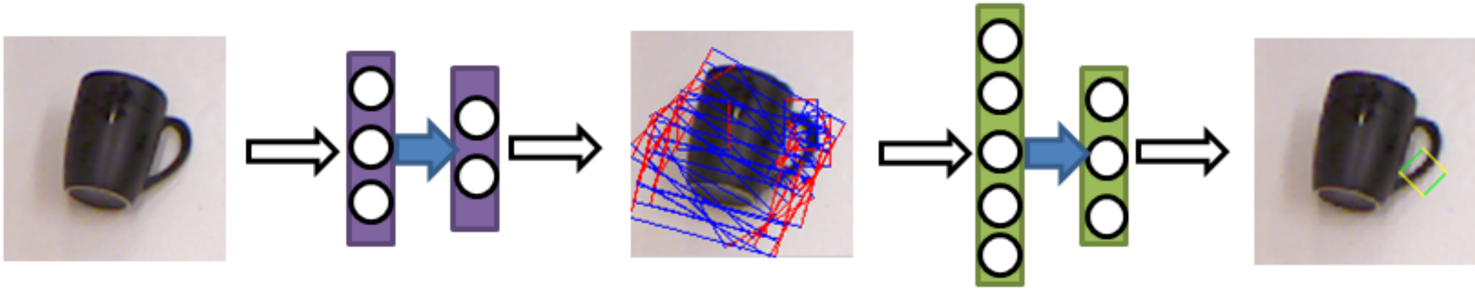
- Generate training data with real-world experiments: Time limitation during dataset acquisition (required several weeks/months)

By courtesy of Google

KMLE grasping solution

Learning based approach

- A custom CNN is trained in an end-to-end fashion, to infer a suitable grasping point by performing **regression** on the entire RGB-Depth image



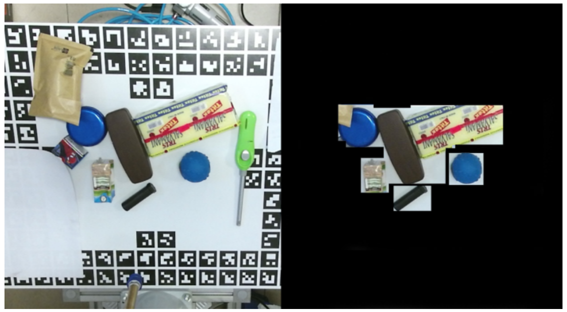
- Our Robotic cell grid is composed with Universal Robots UR3 equipped with: SCHUNK parallel Jaw Gripper; Realsense D435 RGBD camera

Custom Built Dataset | Acquisition Protocol

KONICA MINOLTA

- We collect training data by means of a try-and-test procedure

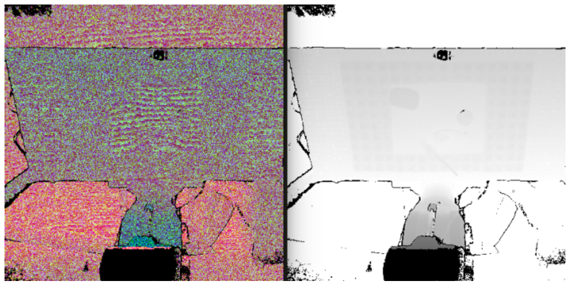
1) Region proposal using YOLO CNN



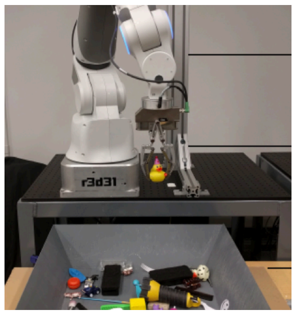
2) Random candidate pose of Grasping

- x,y random pixels coordinates
- Roll and pitch random angles

3) Check safe-depth for that point



4) Attempting to grasp



- Store results into dataset

KMLE Thesis proposal

Thesis Proposal

We want to generate synthetic dataset of grasping points starting from the Output of the KMLE CNN used for Grasping solutions.

- We exploit Gazebo simulator inside Ros Melodic environment to:
  - Replicate the same real grasping pipeline environment (Cell grid structure, camera sensor, robot manipulator arm)
  - Collect a new synthetic dataset by performing a series of simulated attempting of grasping in order to improve the overall accuracy of the network.
- At the end of the entire procedure we can exploit the new dataset to train our network to improve the actual overall accuracy. We validate the final output predictions in real robot context by exploiting sensors available (UR3 Cobot + Realsense RGB-D camera).

The student must create and develop the following tasks:

1) Prepare the environment of the UR3 robot inside Gazebo:

- Add the Transformations between the RGB-D camera with respect to the Robot baseline frame.
- Add the Gripper End-Effector, the API to control the jaws and physical components typical of real context of grasping.
- Add the correct light condition similar to our KMLE laboratory.

Thesis Proposal

2) Develop the Pipeline of Synthetic Grasping:

- Adapt the YOGO CNN w.r.t. the syntethic images provided by Gazebo
- Exploit the CNN output prediction to build a simulated attempting of grasping
- Store the dataset coming from successed attempting of grasping (a grasping will be consider as a success if and only if the object will be grasped at a fixed height with respect the table).

3) Train YOGO CNN with the new synthetic dataset:

- Adapt Images and labels to YOGO CNN format (Batch Normalization)
- Train the network (Try also data augmentation?)

4) Validate Synthetic CNN Outputs on Real Robot:

- Develop a module that understand if the real prediction for unknown objects is correct or not. The validation will be done in real time.
- Update the CNN with the validation obtained after a fixed number of tests.

Thesis Proposal

Step 1, 2 and 3 is now represented in the picture below: