# Product Recognition on Store Shelves

Introduction:

Object detection techniques based on computer vision can be deployed in super market scenarios for the creation of a system capable of recognizing products on store shelves.

Given the image of a store shelf, such a system should be able identify the different products present therein and may be deployed, e.g. to help visually impaired costumers or to automate some common store management tasks (e.g. detect low in stock or misplaced products).

Overall Task:

Develop a computer vision system that, given a reference image for each product, is able to identify boxes of cereals of different brands from one picture of a store shelf. For each type of product displayed in the shelf the system should report:

1. Number of instances.
2. Dimension of each instance (width and height of the bounding box that enclose them in pixel).
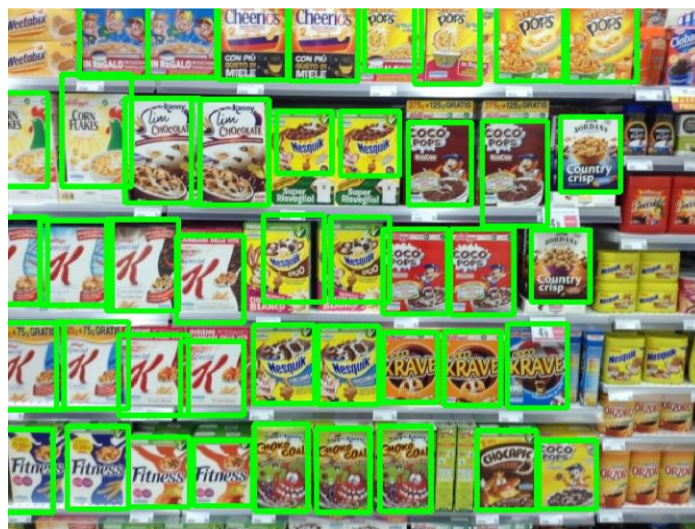
3. Position in the image reference system of each instance (center of the bounding box that enclose them in pixel).

For example, as output of the above image the system should print:

```
Product 0 - 2 instance found:

     Instance 1 {position: (256,328), width: 57px, height: 80px}

     Instance 2 {position: (311,328), width: 57px, height: 80px}

Product 1 – 1 instance found:

…
```

Material:

Students are provided with two folders of images:

- *Models:* contains one reference image for each product that the system should be able to identify.
- *Scenes:* contains 15 different shelve picture to test the algorithm developed in different scenarios.

Step A - Multiple Product Detection:

Test on scene image: {e1.png, e2.png, e3.png, e4.png, e5.png}

Use product images: {0.jpg, 1.jpg, 11.jpg, 19.jpg, 24.jpg, 26.jpg, 25.jpg}

Develop an object detection system to identify single instance of products given: one reference image for each item and a scene image. The system should be able to correctly identify all the product in the shelves image. One way to solve this task could be the use of local invariant feature as explained in lab session 5.

Step B - Multiple Instance Detection:

Test on scene image: {m1.png, m2.png, m3.png, m4.png, m5.png}

Use product images: {0.jpg, 1.jpg, 11.jpg, 19.jpg, 24.jpg, 26.jpg, 25.jpg}

In addition to what achieved at step A, the system should now be able to detect multiple instance of the same product. Purposely, students may deploy local invariant feature together with the GHT (Generalized Hough Transform). More precisely, rather than relying on the usual R-Table, the object model acquired at

training time should now consist in vectors joining all the features extracted in the model image to their barycenter; then, at run time all the image features matched with respect to the model would cast votes for the position of the barycenter by scaling appropriately the associated joining vectors (i.e. by the ratio of sizes between the matching features).

Step C (optional) - Whole shelve challenge:

Test on scene image: {h1.png, h2.png, h3.png, h4.png, h5.png}

Use product images: {from 0.jpg to 23.jpg}

Try to detect as much products as possible in this challenging scenario: more than 40 different product instances for each picture, distractor elements (e.g. price tags…) and low resolution image. You can use whatever technique to achieve the maximum number of correct detection without mistake.

Acknowledgements