**Goal:**

Label all components in a webpage

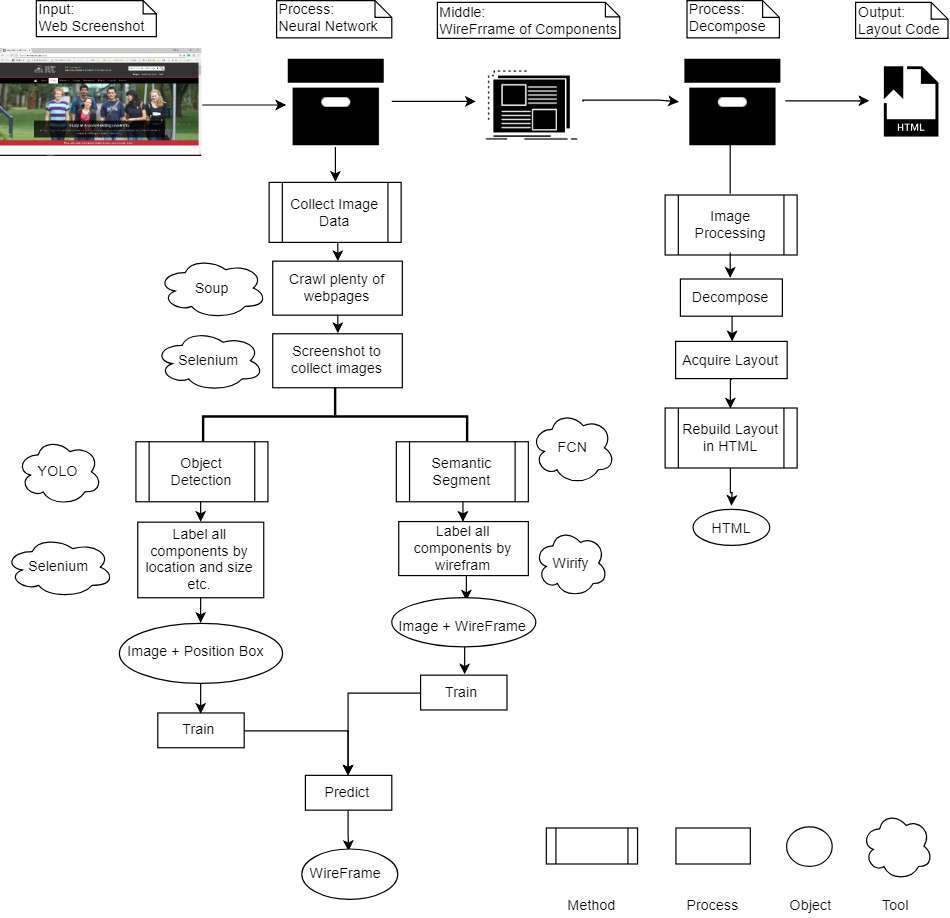
**Idea:**

Object Detection

Semantic label, semantic segment

FCN

COCO dataset



**Backlog:**

1. Label data in selenium
   1. Iterate all the elements √
   2. Get the position and size of each element √
   3. Export the info of each element into CSV √
   4. Screenshot the webpage √
   5. Debug opencv √
   6. Crawl several pages √
   7. Crawl pages with their screenshot √
   8. Adjust the result between the info and screenshot √
   9. Clean the labels √
      1. Hidden libels √
      2. Nonsense libels √
   10. Crawl the wireframe of each page √
   11. Tidy and debug the crawler
2. Build a YOLO darknet / FRCNN in attempt to train a model

**Backlog:**

1. Scrawl top websites
2. Try to get full image of the web page
3. Get limit info (label) according to the size of image
4. Build a simple NN to classify some obvious components
5. Build a simple NN to classify if a component is image