

Here is the structure of the dataset:

```
-TRAINING_DATASET
|
|_ IMAGES
|_ LABELS_TXT
|_ LABELS_XML
```

IMAGES : contains a subset of Mar_dataset + Kaggle_boat_dataset images (3220 in total). Since many images were not relevant, some of them were not considered in this dataset. **Note that some images (like those corresponding to numbers 1941, 1966, 2515, 2516, 2517) are absent:** they have been discarded because really hard to be labelled, or irrelevant. Consider this fact when you parse the images.

LABELS_TXT / LABELS_XML : contain the labels (one file for each image) in format .txt or in format .xml (**PASCAL VOC** typical structure: note that since the labels have been produced by many people, the `<path>---</path>` line is not relevant since it contains the absolute path to the image). For info on the labels, see (*)

```
-TEST_DATASET
|
|_ kaggle
|_ kaggle_labels_txt
|_ kaggle_labels_xml
|_ venice
|_ venice_labels_txt
|_ venice_labels_xml
```

kaggle / venice : contains the images of the kaggle dataset that we have to use to evaluate the performance of our approach. Note that here also very small and far boats have been labelled.

kaggle_labels_txt / kaggle_labels_xml / venice_labels_txt / venice_labels_xml : contain the labels (one file for each image) in format .txt or in format .xml of the aforementioned test images

(*) **The .xml file format** is easy to handle, since it is a common PASCAL VOC format. In particular, each .xml label specifies the dimensions and number of channels of the image (between `<size>---</size>`), then each boat is specified by the flag `<object>`:

`<name>---</name>` is the label that were assigned (**NOTE: the test dataset has only 'boat' labels. The training dataset instead labels a boat as 'boat' if it is clearly distinguishable and not hidden, while some other portions of boat are labelled as 'hiddenboat' or 'hidden_boat' -pay attention to the fact that there may be an underscore in the name- if there is a small portion of a boat.** Why this? Suppose you want only clear and distinct images of boat for instance, to train your CNN: then you can take only the bounding boxes indexed by 'boat' and discard all the 'hiddenboat'. Suppose instead you may want to include also some portions of a boat, for instance because you focus on features of boats: then you may include also 'hiddenboat' labels and consider them as pure boats. It's your choice. Note that too small portions of boats are not included).

`<bndbox>---</bndbox>` stores the coordinates of the bounding box that encloses the 'boat'/'hiddenboat'. There you may see 4 flags: xmin and ymin, that are the (x,y) coordinates of the top-left corner of the bounding box enclosing the boat, and xmax and ymax that instead encodes the bottom-right corner. x=horizontal direction, y=vertical direction

The **.txt files** are instead much more condensed: each line of the .txt file corresponds to one 'boat'/'hiddenboat', then the previously mentioned coordinates are saved in the current order:
 xmin;xmax;ymin;ymax;
 This means that coordinates (1,3) encode the top-left corner, while (2,4) the bottom-right.

NOTE: if you don't find a label corresponding to an image inside TRAINING_DATASET/IMAGES, it means that the image has not meaningful boats, or only a meaningless portion of a boat is framed (for instance just the side of a cruise ship, where only cabins can be distinguished). You may also decide to discard the images, if you think they may act like false negatives feeding them to a neural network.

SOME EXAMPLES:

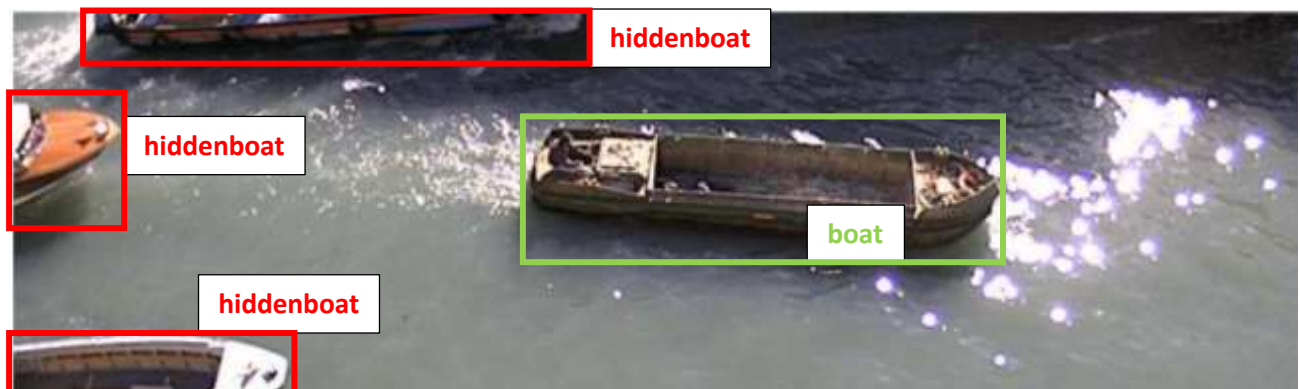


Image0763.png

```

<?xml version="1.0"?>
<image>
  <name>Image0763.png</name>
  <pose>Unspecified</pose>
  <truncated>1</truncated>
  <difficult>0</difficult>
  <object>
    <name>hiddenboat</name>
    <pose>Unspecified</pose>
    <truncated>1</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>1</xmin>
      <ymin>174</ymin>
      <xmax>197</xmax>
      <ymax>240</ymax>
    </bndbox>
  </object>
  <object>
    <name>hiddenboat</name>
    <pose>Unspecified</pose>
    <truncated>1</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>48</xmin>
      <ymin>354</ymin>
      <xmax>29</xmax>
    </bndbox>
  </object>
  <object>
    <name>hiddenboat</name>
    <pose>Unspecified</pose>
    <truncated>1</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>1</xmin>
      <ymin>76</ymin>
      <xmax>50</xmax>
      <ymax>129</ymax>
    </bndbox>
  </object>
  <object>
    <name>boat</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>316</xmin>
      <ymin>61</ymin>
      <xmax>618</xmax>
      <ymax>149</ymax>
    </bndbox>
  </object>
</image>
  
```

```

image0763.txt - Blocco note di Windows
File Modifica Formato Visualizza ?
hiddenboat:1;174;197;240;
hiddenboat:48;354;1;29;
hiddenboat:1;76;50;129;
boat:316;618;61;149;
  
```

image0763: screenshots of .xml and .txt files



Image2114.png: NO LABEL HAVE BEEN PRODUCED, since we are framing only the bow of a boat: there's not a full body of a boat