Evaluation GUI Guide
(Detection and Classification)

General Information

Labels format: **JSON label only**

The evaluation Gui support both image classification and object detection

Section 1: Preparing the Dataset

Multiple Detection Case:

1. Folder with any name
    a. **Images** (Containing the images needed to be tested)
    b. **Labels** folder
        i. **Json** folder (Containing the labels of the images having the same name as the images)
2. In case of zip file the general format should be the same just zip the folder and upload that zip file

Image classification Case:

- Folder with any name
    - **Dataset_Class_1**
    - **Dataset_Class_2**
    - **Etc. …**

Section 2: Main Output of this evaluation GUI:

Case of object detection:

- General Evaluation for all the models containing the following:
  - Plots:
    - General Confusion matrix
    - Multi-class Confusion matrix
    - Histogram
    - Scatter in case we have bad IoU links

  - Excel file containing 2 sheets:
    - General section sheet:
      - General information about the inputs
      - Details of outputs and linkages types (TP, TN, FP, FN) with percentages
      - IoU details (with respect to thresholds given) and percentages
      - ED information (with respect to IoU thresholds)
      - Average IoU for the API
      - Average confidence for the API responses

    - Per Label sheet:
      - The GT, prediction and type of linkage
      - The IoU, ED and confidence
      - The type of coverage between the GT and Prediction + Which area outside the Intersection is Bigger
      - The coordinates of the Gt and prediction and their centroids.

- Specific Evaluation for each model
  - Plots:
    - Histogram and Scatter if available
- Draw the GT and predicted box on the detected images and store them.
- Error Folder containing corrupted images.

Case of image classification:

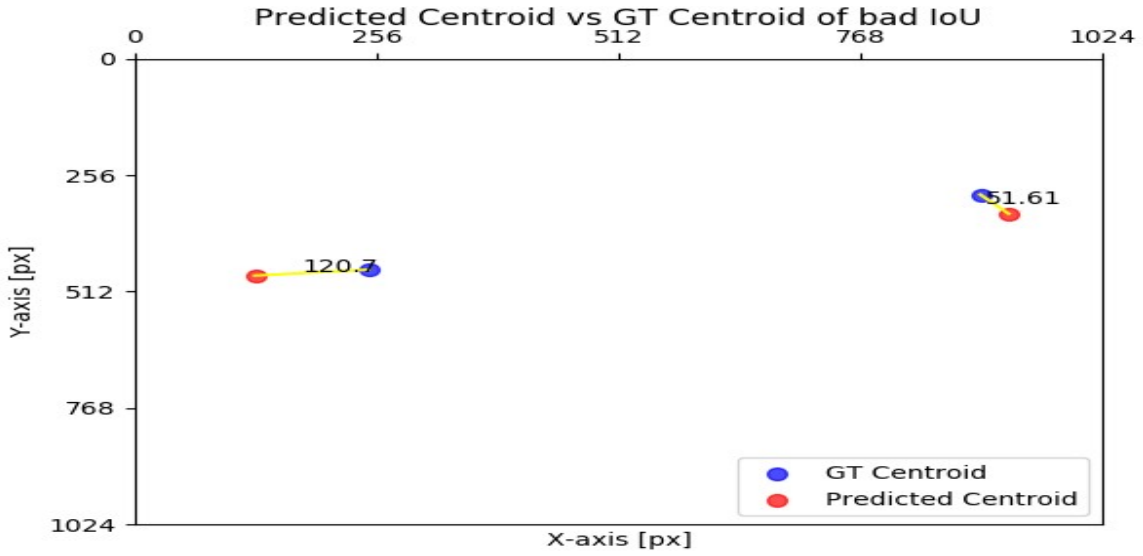
- General Evaluation
  - Plots:
    - General Confusion matrix
    - Multi-class Confusion matrix
  - Excel file containing 2 sheets:
    - General section sheet:
      - General information about the inputs
      - Details of outputs and linkages types (TP, TN, FP, FN) with percentages
      - Top-1 accuracy, accuracy, precision, recall, f-score
      - Average confidence for the API responses
    - Per Label sheet:
      - For each model and image name the prediction and the confidence
- Specific Evaluation for each class:
  - Partition of the images:
    - In the respective folder of the true label if classified correctly
    - In the Wrong Classified folder if wrong
    - In the Not classified folder if no classification resulted for the image

Section 3: Result Folders

Charts:

Scatter:

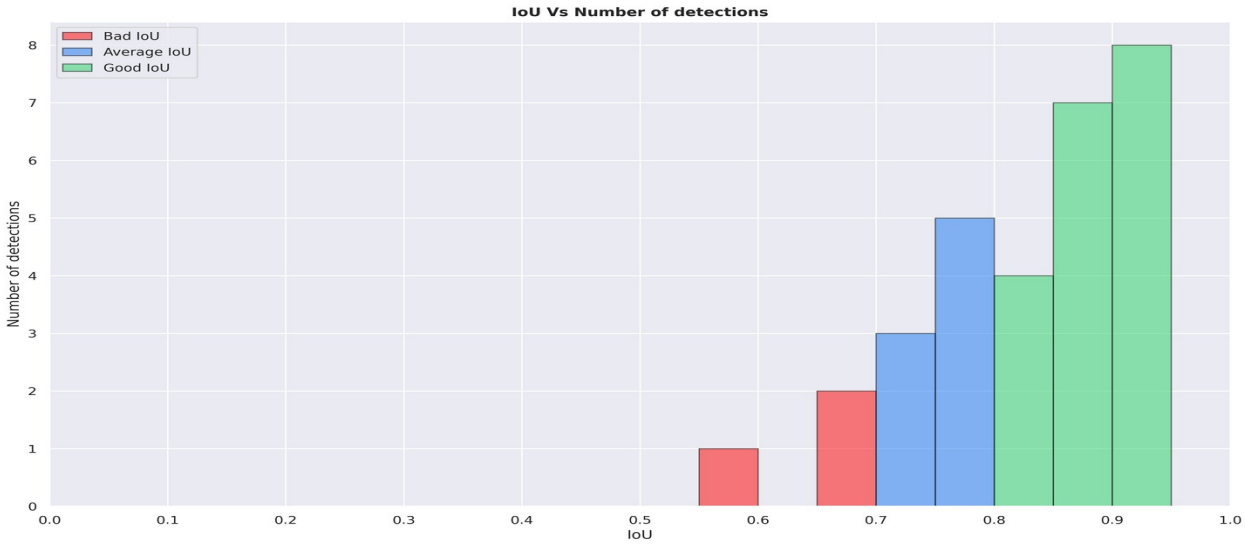
A scatter will be created, only covering the cases of bad IoU, containing the Euclidean distance between the centers of the GT and predicted box, since this parameter is important to know how far away the boxes are. This scatter has the boundaries of the largest image (in pixels) tested, with its origin on the top left (which is the same origin in the response of the detection given by the API). An example of a scatter with 2 bad IoUs is given in the following figure:



Histogram:

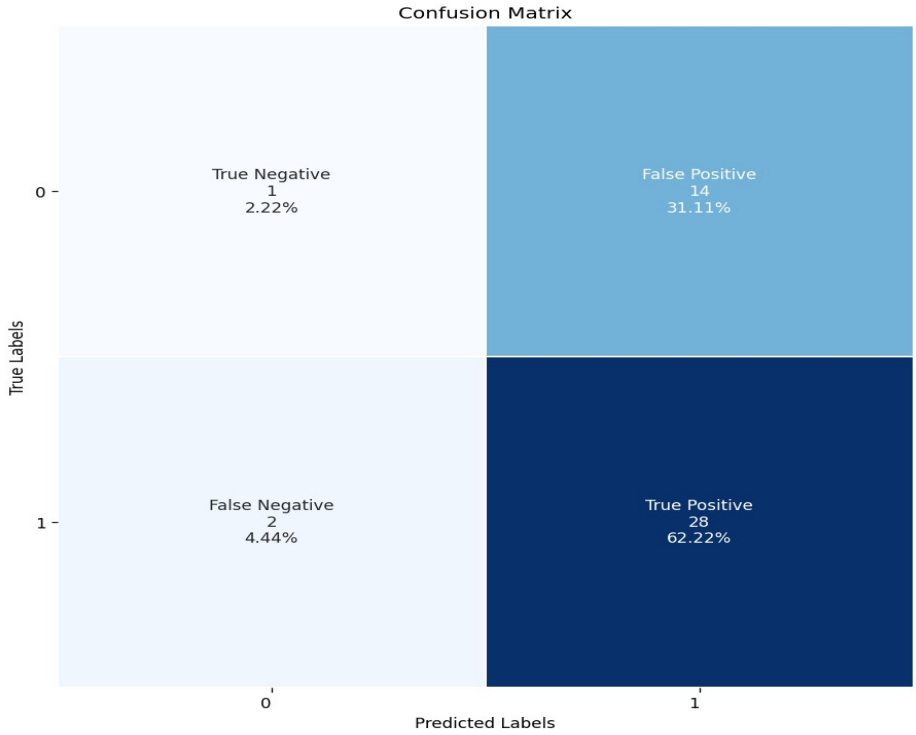
Histograms are also created showing for each model the IoU partition vs number of images and a general overview for all image's vs IoU. The

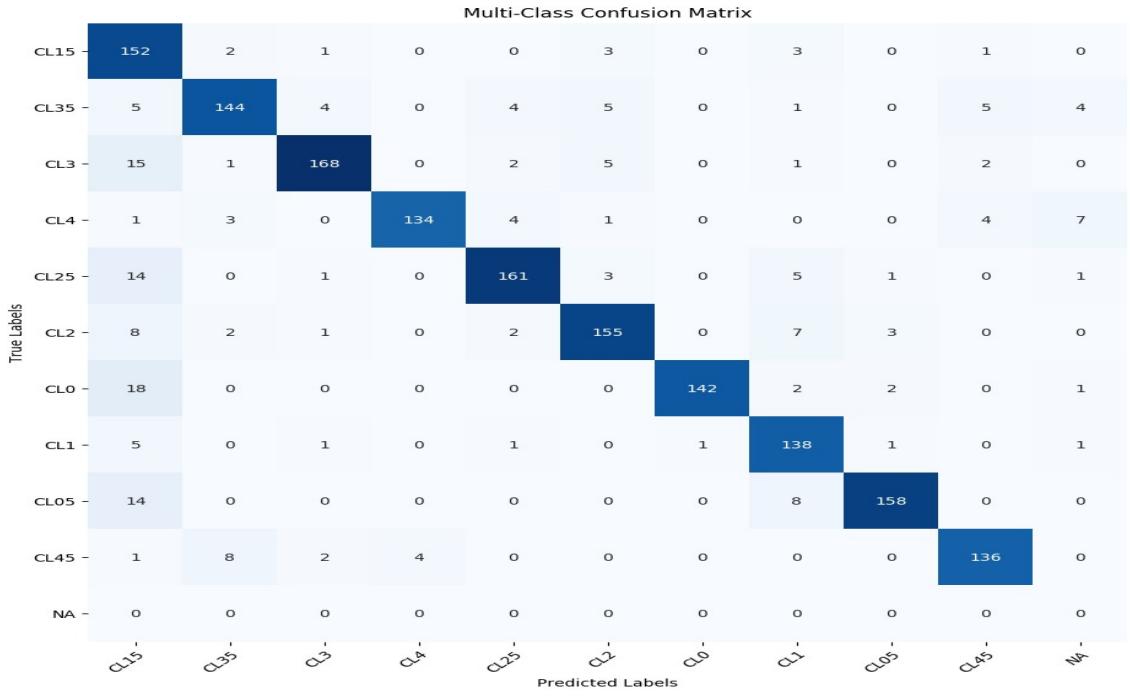
histogram is color coded, so good results are in green, average results are in blue and bad results are in red. An example of a histogram is given in the following figure:



Confusion Matrix :

2 confusions matrices are produced. The first one has a role of showing how many of the values are Correctly Predicted (True Positives), are wrongly predicted (False Positives), etc... and the second confusion matrix, the multi-class confusion matrix, will show the True classes Vs the Predicted classes. These confusion matrices are done with the aid of heat maps in order to help the user visualize the output. The best-case scenario is to have the darkest color in the True Positive case in the first chart, and on the diagonal of the second chart. An example of the confusion matrices is given in the following figures.


Confusion Matrix
True Negative
1
2.22%
False Positive
14
31.11%
False Negative
2
4.44%
True Positive
28
62.22%
0
1
True Labels
0
1
Predicted Labels


Multi-Class Confusion Matrix
CL15 152 2 1 0 0 3 0 3 0 1 0
CL35 5 144 4 0 4 5 0 1 0 5 4
CL3 15 1 168 0 2 5 0 1 0 2 0
CL4 1 3 0 134 4 1 0 0 0 4 7
CL25 14 0 1 0 161 3 0 5 1 0 1
CL2 8 2 1 0 2 155 0 7 3 0 0
CL0 18 0 0 0 0 0 142 2 2 0 1
CL1 5 0 1 0 1 0 1 138 1 0 1
CL05 14 0 0 0 0 0 0 8 158 0 0
CL45 1 8 2 4 0 0 0 0 0 136 0
NA 0 0 0 0 0 0 0 0 0 0 0
True Labels
CL15
CL35
CL3
CL4
CL25
CL2
CL0
CL1
CL05
CL45
NA
Predicted Labels

Bounding Boxes:

In the specific evaluation, or evaluation per model, we will have bounding boxes drawn on each image. The blue color will be for the GT, the red box for the predicted box, and the yellow line will link the centers of the boxes representing the Euclidean Distances.