My dataset at the moment is too large, and not the format that I require.

There are 8,476 images in just the training folder. Then 836 images in validation and 367 in test. Totalling half a GB, this is more data than I want to work with during the beginning stages of my capstone project.

I also need to update the format of the csv file to comply with Google Vision API's requested syntax for object detection.Current Syntax

[file_path], [bounding_box], [label]

Required Syntax

```
[ML_USE],GCS_FILE_PATH,[LABEL],[BOUNDING_BOX]*
```

The file path likely needs to be altered for use in the cloud. The label and bounding box must swap places. And the bounding box needs to be altered slightly.

- Empty columns must be added to make this:

1. Two vertices (two sets of x,y coordinates) that are diagonally opposite points of the rectangle:
   A. `X_MIN`, `Y_MIN`
   C. `X_MAX`, `Y_MAX`
   as shown in this example:
   A,,C,
   X_MIN, *Y_MIN*, , , *X_MAX*, *Y_MAX*, ,

- And those X_MIN, … values must be scaled to a float value between 0 and 1, as opposed to an integer value.