

Real-time Courier service provider detection system using YOLO framework

Phase- 2

DATA 606 - Capstone Project

Shrimanth Ajjamane Manohar

1. Data Collection

- Installed chrome driver for google chrome web browser
- installed selenium package in python
- executed python script to download google images for classes - amazon, fedex, ups, usps
- [python script](#)
- Looking at the images - [Colab notebook](#)
- Along with identifying the logos on the image, we will add another class with courier boxes to detect in the image as well.

Note: As Google has updated their policies on web scraping, I was not able to download more than 80 images from the Google Images link.

2. Data Augmentation

Neural networks, in general, are data-hungry that require large data to learn and perform better. Data Augmentation is the process of applying different transformation techniques like horizontal flip, cropping, shearing on the available data to synthesize new data.

For this, I have applied 3 kinds of transformations on the available data - [Colab Notebook](#)

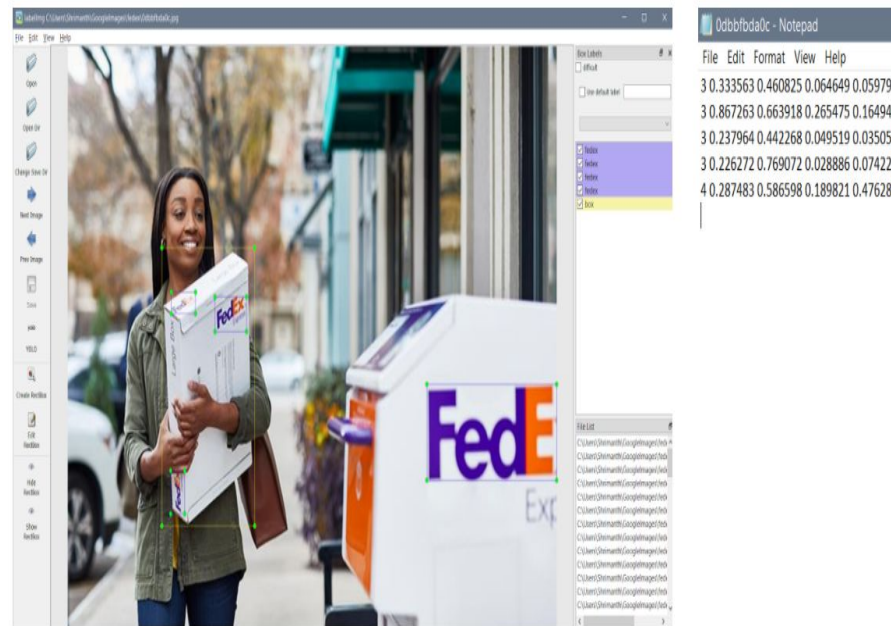
- a. horizontal flip (left to right)
- b. vertical flip (up-down)
- c. rotation by 45 degrees

After augmentation, we have 1280 images for five classes that can be split to train and validation.

3. Data Annotation

used [labellmg](#) open-source annotation tool

- Labellmg is a graphical image annotation tool.
- It is written in Python and uses Qt for its graphical interface.
- Annotations are saved as XML files in PASCAL VOC format, the format used by [ImageNet](#). Besides, it also supports the YOLO format
- The annotation file should be in .txt format
- (class) (X_CENTER_NORM) (Y_CENTER_NORM) (WIDTH_NORM) (HEIGHT_NORM)



4. Data configuration

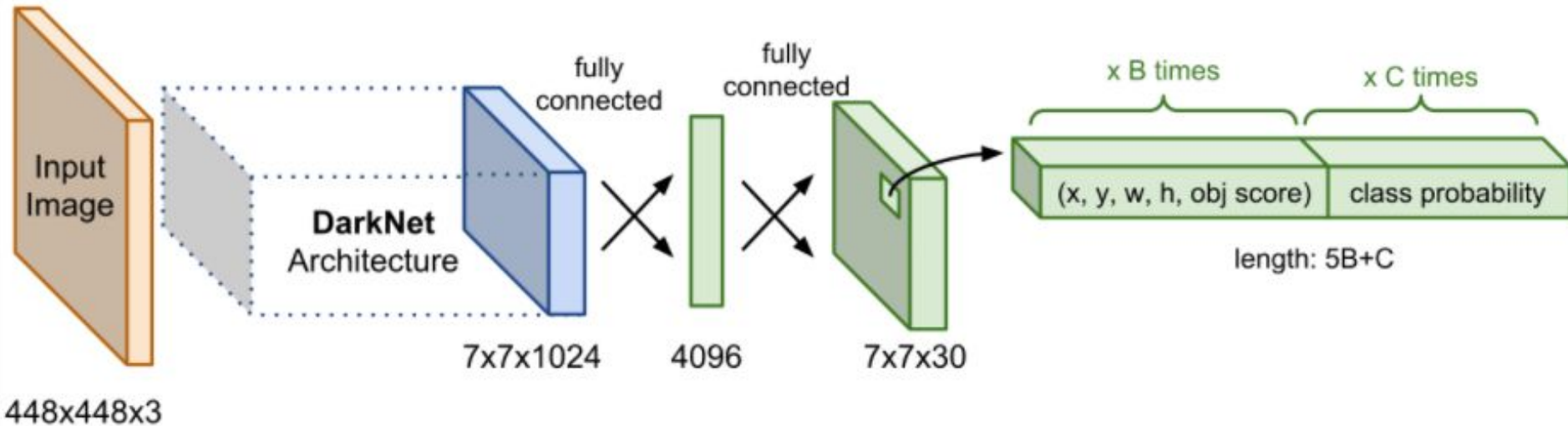
- Configuration files have to be changed as per the number of classes (5)
- Number of filters to be changed to is given by

$(\text{Classes} + 5) * 3$ i.e. 30

- yolo.names - classes files
- yolo.data - configuration files details
- train.txt - train images path directory

```
-darknet
-cfg
    -yolo_train.cfg #train network architecture file
    -yolo_test.cfg #test network architecture file
-data
    -images # images folder
        -image1.jpg
        -image2.jpg
        -.....
    -labels # corresponding labels folder
        -image1.txt
        -image2.txt
        -.....
    -train.txt #train images path directory
    -val.txt #validation images path directory
    -yolo.data #details of all the configuration files
    -yolo.names # class names
-darknet53.conv.74 #weights file
```

5. YOLO model



5. Training

```
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.447401), count: 1, class_loss = 313.529663, iou_loss = 0.457703, total_loss = 313.987366
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.330744), count: 1, class_loss = 958.796448, iou_loss = 1.361572, total_loss = 960.158020
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.000000), count: 1, class_loss = 3911.594971, iou_loss = 0.000000, total_loss = 3911.594971
total_bbox = 8663, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.308299), count: 4, class_loss = 314.971436, iou_loss = 5.224274, total_loss = 320.195709
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.406053), count: 5, class_loss = 958.719788, iou_loss = 4.786316, total_loss = 963.506104
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.000000), count: 1, class_loss = 3894.319336, iou_loss = 0.000000, total_loss = 3894.319336
total_bbox = 8672, rewritten_bbox = 0.000000 %
Can't open label file. (This can be normal only if you use MSCOCO): data/obj/23120b5301.txt
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.366636), count: 2, class_loss = 311.716034, iou_loss = 2.683838, total_loss = 314.399872
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.000000), count: 1, class_loss = 954.713684, iou_loss = 0.000000, total_loss = 954.713684
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.000000), count: 1, class_loss = 3873.778809, iou_loss = 0.000000, total_loss = 3873.778809
total_bbox = 8674, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.231669), count: 1, class_loss = 311.526001, iou_loss = 2.292206, total_loss = 313.818207
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.302099), count: 3, class_loss = 960.059570, iou_loss = 6.256470, total_loss = 966.316040
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.236907), count: 2, class_loss = 3888.711426, iou_loss = 5.135254, total_loss = 3893.846680
total_bbox = 8680, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.298320), count: 2, class_loss = 313.809418, iou_loss = 1.757019, total_loss = 315.566437
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.360761), count: 1, class_loss = 954.117371, iou_loss = 1.363892, total_loss = 955.481262
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.000000), count: 1, class_loss = 3845.085449, iou_loss = 0.000000, total_loss = 3845.085449
total_bbox = 8683, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.210924), count: 2, class_loss = 315.193054, iou_loss = 3.836182, total_loss = 319.029236
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.000000), count: 1, class_loss = 960.083984, iou_loss = 0.000000, total_loss = 960.083984
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.000000), count: 1, class_loss = 3883.438721, iou_loss = 0.000000, total_loss = 3883.438721
total_bbox = 8685, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.442987), count: 1, class_loss = 313.242889, iou_loss = 0.938324, total_loss = 314.181213
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.346409), count: 1, class_loss = 953.477051, iou_loss = 0.954773, total_loss = 954.431824
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.000000), count: 1, class_loss = 3871.490479, iou_loss = 0.000000, total_loss = 3871.490479
total_bbox = 8687, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.535132), count: 3, class_loss = 315.087646, iou_loss = 1.359161, total_loss = 316.446808
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.184658), count: 1, class_loss = 958.268005, iou_loss = 2.455994, total_loss = 960.723999
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.000000), count: 1, class_loss = 3937.918945, iou_loss = 0.000000, total_loss = 3937.918945
total_bbox = 8691, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.000000), count: 1, class_loss = 312.338318, iou_loss = 0.000000, total_loss = 312.338318
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.115173), count: 2, class_loss = 961.652405, iou_loss = 5.483887, total_loss = 967.136292
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.000000), count: 1, class_loss = 3904.226807, iou_loss = 0.000000, total_loss = 3904.226807
```

5.1 Training loss for 100 epochs

```
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.000000), count: 1, class_loss = 697.872253, iou_loss = 0.000000, total_loss = 697.872253
  total_bbox = 13872, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.388086), count: 3, class_loss = 103.459915, iou_loss = 2.841896, total_loss = 106.301811
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.000000), count: 1, class_loss = 246.460800, iou_loss = 0.000000, total_loss = 246.460800
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.000000), count: 1, class_loss = 662.719177, iou_loss = 0.000000, total_loss = 662.719177
  total_bbox = 13875, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.378017), count: 2, class_loss = 101.834785, iou_loss = 1.186394, total_loss = 103.021179
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.000000), count: 1, class_loss = 249.422989, iou_loss = 0.000000, total_loss = 249.422989
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.000000), count: 1, class_loss = 730.541016, iou_loss = 0.000000, total_loss = 730.541016
  total_bbox = 13877, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.290716), count: 2, class_loss = 101.526329, iou_loss = 3.182159, total_loss = 104.708488
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.000000), count: 1, class_loss = 247.797241, iou_loss = 0.000000, total_loss = 247.797241
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.000000), count: 1, class_loss = 685.357178, iou_loss = 0.000000, total_loss = 685.357178
  total_bbox = 13879, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.312116), count: 3, class_loss = 102.337677, iou_loss = 2.871170, total_loss = 105.208847
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.369852), count: 1, class_loss = 248.866592, iou_loss = 0.672104, total_loss = 249.538696
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.000000), count: 1, class_loss = 670.343140, iou_loss = 0.000000, total_loss = 670.343140
  total_bbox = 13883, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.400528), count: 1, class_loss = 101.781265, iou_loss = 0.625145, total_loss = 102.406410
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.179916), count: 2, class_loss = 249.478882, iou_loss = 6.183289, total_loss = 255.662170
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.353176), count: 2, class_loss = 682.551514, iou_loss = 6.270325, total_loss = 688.821838
  total_bbox = 13888, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.000000), count: 1, class_loss = 99.218910, iou_loss = 0.000000, total_loss = 99.218910
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.338131), count: 5, class_loss = 251.118149, iou_loss = 14.522873, total_loss = 265.641022
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.230689), count: 7, class_loss = 675.643127, iou_loss = 15.998352, total_loss = 691.641479
  total_bbox = 13900, rewritten_bbox = 0.000000 %
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 82 Avg (IOU: 0.413937), count: 1, class_loss = 101.021568, iou_loss = 0.522636, total_loss = 101.544205
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 94 Avg (IOU: 0.000000), count: 1, class_loss = 251.187485, iou_loss = 0.000000, total_loss = 251.187485
v3 (mse loss, Normalizer: (iou: 0.75, obj: 1.00, cls: 1.00) Region 106 Avg (IOU: 0.222021), count: 1, class_loss = 703.448181, iou_loss = 1.806152, total_loss = 705.254333
  total_bbox = 13902, rewritten_bbox = 0.000000 %
```


5. Results

```
1 #/content/drive/MyDrive/Sharing_storage1/darknet/darknet/data/images/00d58c45ca.jpg
2 # Execute prediction on a single image
3 img = cv2.imread("/content/drive/MyDrive/Sharing_storage1/darknet/darknet/data/images/01c53361f1.jpg")
4 img = cv2.cvtColor(img,cv2.COLOR_BGR2RGB)
5 display_img(predict(img))
```

/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:5: MatplotlibDeprecationWarning: Adding an axes



6. Evaluation Metrics - mAP(mean average precision)

1. Popular metric in measuring the accuracy of object detectors
 - ❑ Precision: How precise the model is out of the predicted positives and how many of them are actual positives
$$\text{True positive} / [\text{Total predicted positives}]$$
 - ❑ Recall: Calculates how many actual positives did the mode capture
$$\text{True positive} / [\text{Total actual positives}]$$
2. computes the average precision - area under the precision-recall curve
3. mean of the average precision is calculated from the above average precision
4. Intersection over Union(IoU) - measures overlap between the two boundaries
 - there are predefined thresholds for specific datasets
 - A detection is a true positive if it has “intersection over union” (IoU) with a ground-truth box greater than some threshold (usually 0.5; if so, the metric is “mAP@0.5”)

Next Steps:

- Integrate with Pi Camera feed
- Execute the results on video feed
- Run it for more epochs and store the weights for each 100 iterations for obtaining test results using these weights file.