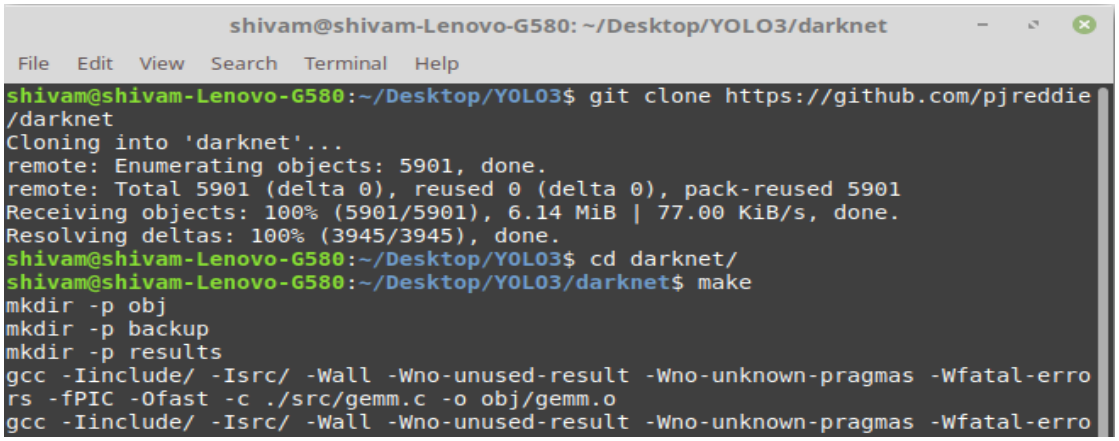
Project 3 – PII Detection Using YOLO - A Report

By: Shivam Pandit

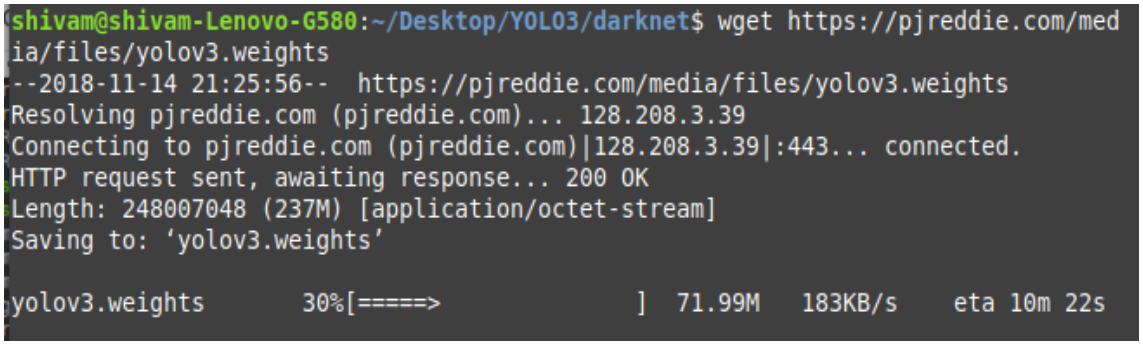
- Firstly, we start with setting up YOLO on our system. For that we first execute command

```
git clone https://github.com/pjreddie/darknet
cd darknet
make
```



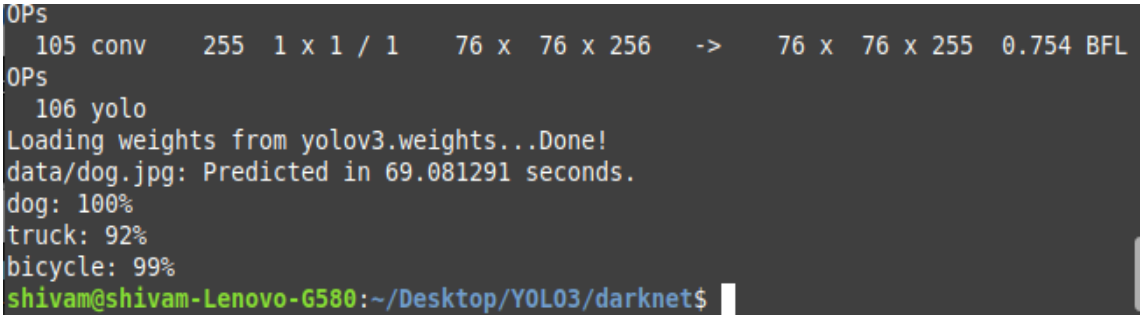
- Then, to use pretrained model weights we download pre-trained weight file using command

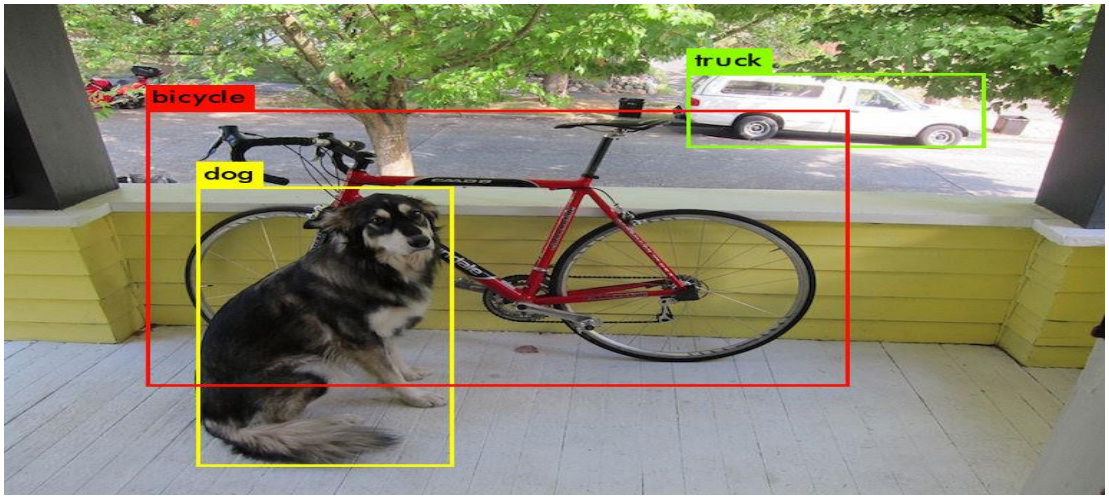
```
wget https://pjreddie.com/media/files/yolov3.weights
```



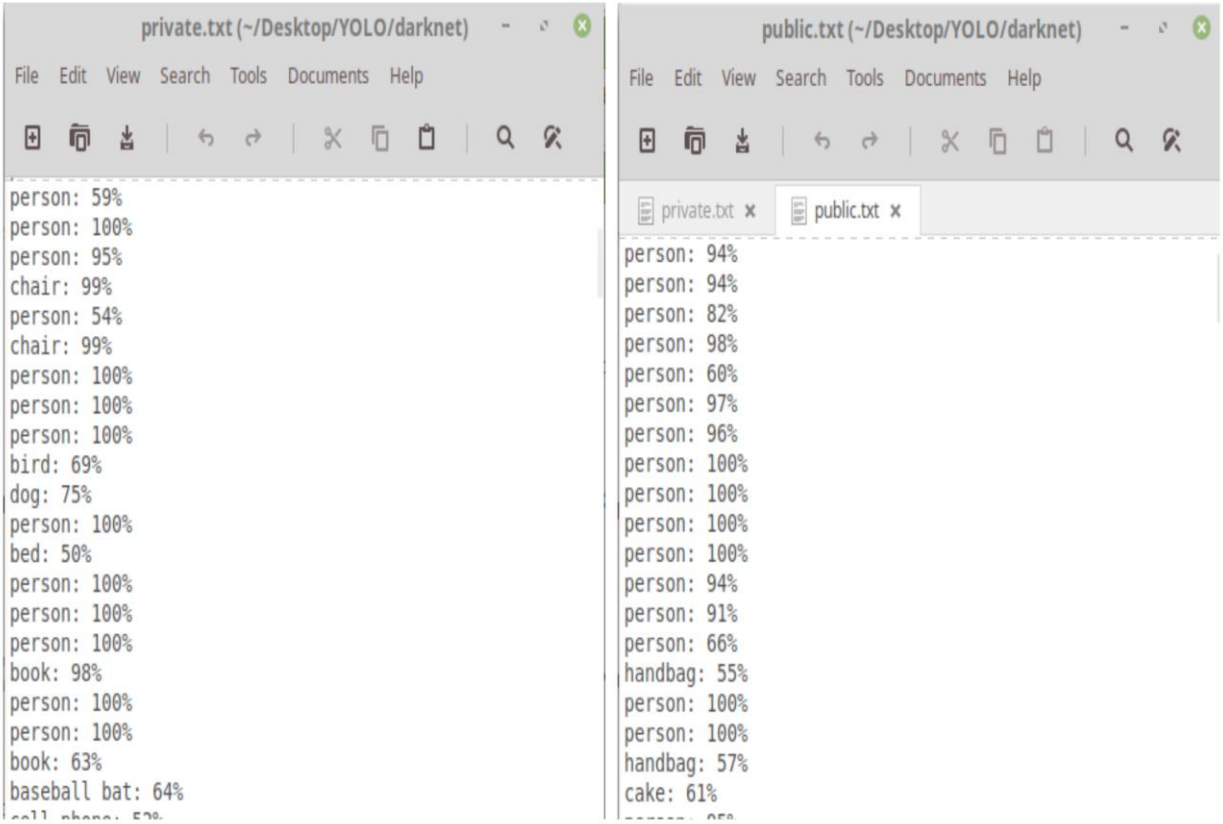
- Then we test test YOLO using command below which shows it is working as expected and we get prediction accuracy with labels.

```
./darknet detect cfg/yolov3.cfg yolov3.weights data/dog.jpg
```

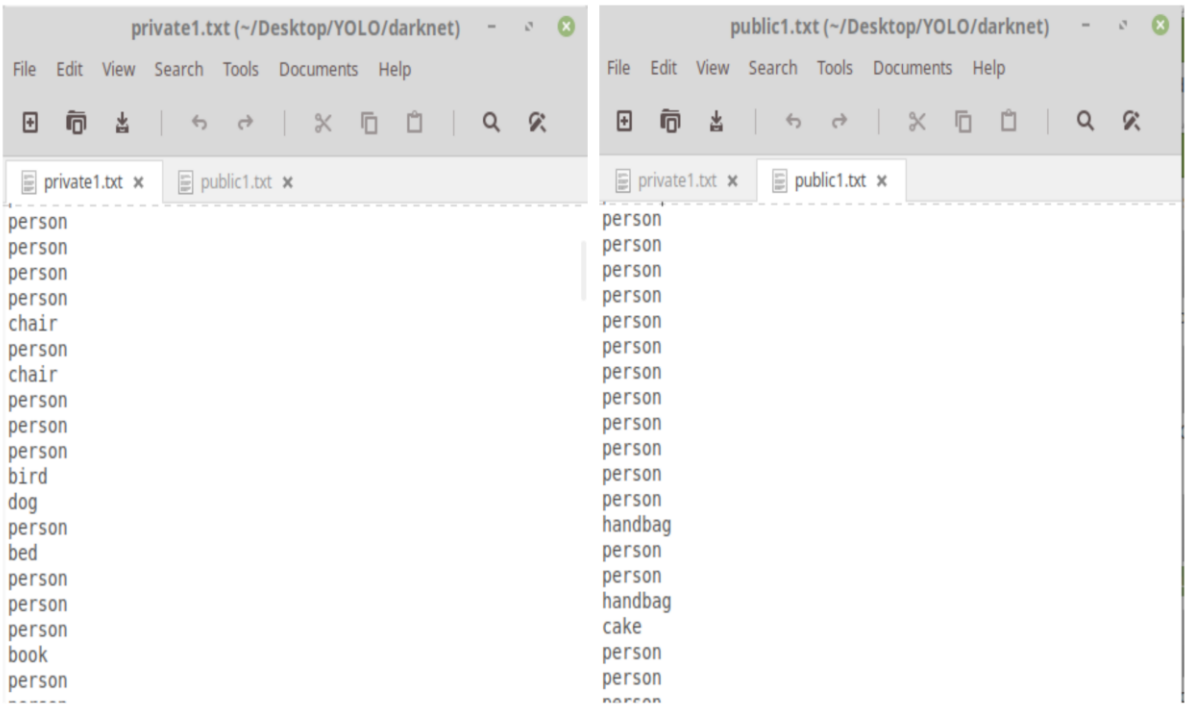




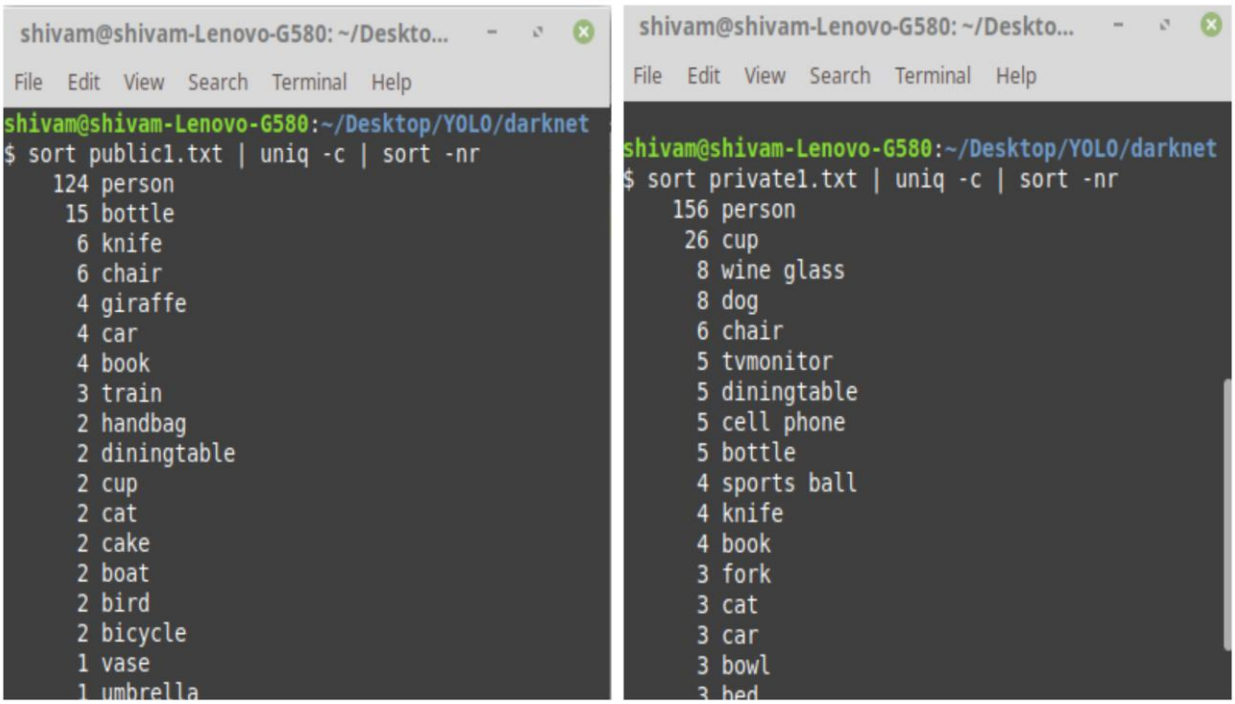
- Now, since YOLO is successfully compiled and tested we will now start to write a method to compute top 10 public and private objects. For that i wrote two scripts that runs a detector on public and private images to make a .txt files from the public and private data set with labels and prediction accuracy of the objects. After executing the script, we get public and private txt files with labels and prediction accuracy as shown below.



❑ Now, to get top 10 objects from this we have to first process the data to be able to count the occurrence of the labels in both public and private .txt files. For that, i wrote a script that strips every line after colon to get just labels and store in new files namely public1.txt and private1.txt as shown below.



❑ Then, to get count of unique labels in descending order, we use simple linux command uniq and sort on both public1 and private1 .txt files as shown below.



➔ Now, we have got the object labels with their corresponding count which we have denoted in tabular manner below for top 10 objects.

- For Private images

S.No.	Object Label	Total Count
1	person	156
2	cup	26
3	Wine glass	8
4	dog	8
5	chair	6
6	tvmonitor	5
7	diningtable	5
8	cellphone	5
9	bottle	5
10	Sports ball	4

- For Public images

S.No.	Object Label	Total Count
1	person	124
2	bottle	15
3	knife	6
4	chair	6
5	giraffe	4
6	car	4
7	book	3
8	train	2
9	handbag	2
10	diningtable	2

APPENDIX

- yoloprivate.py - Python file for running the detector and making private.txt file for private image from dataset with object labels and their accuracy predictions

```
import os
from subprocess import Popen, PIPE
file=open("private.txt", "a+")
images_dir = "/home/shivam/Desktop/YOLO/darknet/dataset1"
for img in os.listdir(os.getcwd()+"/dataset1/private"):
      cnt=0
      setting = "private"
      p    =    Popen(['./darknet',    'detect',    'cfg/yolov3.cfg',
'yolov3.weights',  images_dir  +  '/'  +  setting  +  '/'  +  img],cwd  =
'/home/shivam/Desktop/YOLO/darknet/', stdout = PIPE, stderr = PIPE)
      stdout, stderr = p.communicate()
      for line in stdout:
      cnt+=1
      psn=stdout.find('\n')
      stdout=stdout[psn+1:cnt-1]
      file.write(stdout)
file.close()
```

- yolopublic.py - Python file for running the detector and making private.txt file for public images from dataset with object labels and their accuracy predictions.

```
import os
from subprocess import Popen, PIPE
file=open("private.txt", "a+")
images_dir = "/home/shivam/Desktop/YOLO/darknet/dataset1"
for img in os.listdir(os.getcwd()+"/dataset1/public"):
      cnt=0
      setting = "public"
      p    =    Popen(['./darknet',    'detect',    'cfg/yolov3.cfg',
'yolov3.weights',  images_dir  +  '/'  +  setting  +  '/'  +  img],cwd  =
'/home/shivam/Desktop/YOLO/darknet/', stdout = PIPE, stderr = PIPE)
      stdout, stderr = p.communicate()
      for line in stdout:
      count+=1
      ps=stdout.find('\n')
      stdout=stdout[pos+1:count-1]
      file.write(stdout)
file.close()
```

- trimdata.py - Python file for removing object accuracy data from .txt files for counting top 10 objects

```
linecount=[line.rstrip('\n') for line in open('private.txt')]
file=open("private1.txt", "a+")
for i in linecount:
    psn=i.find(':')
```

```
    i=i[0:pos]
    file.write(i+'\n')
file.close()
linecount=[line.rstrip('\n') for line in open('public.txt')]
file=open("public1.txt", "a+")
for i in lines:
    pos=i.find(':')
    i=i[0:pos]
    file.write(i+'\n')
file.close()
```