National University of Computer and Emerging Sciences

Data Science (8B)

Project Report

Submitted By

Ramin Rafi……..……L17-4284
Maryam Nizam…….L17-4311
Fatima Junaid……..…..L17-6301

Submitted To

Sir Irfan Younas

Object Detection Using Mask RCNN with Transfer Learning

Object Detection is a technique related to Computer Vision which is used to detect, locate and label objects in a picture or a video. There are various methods that can be used for the purpose of object detection but we are going to use YOLOV3 for detecting Name, identification number, Date of Birth, Country and Name in our dataset.

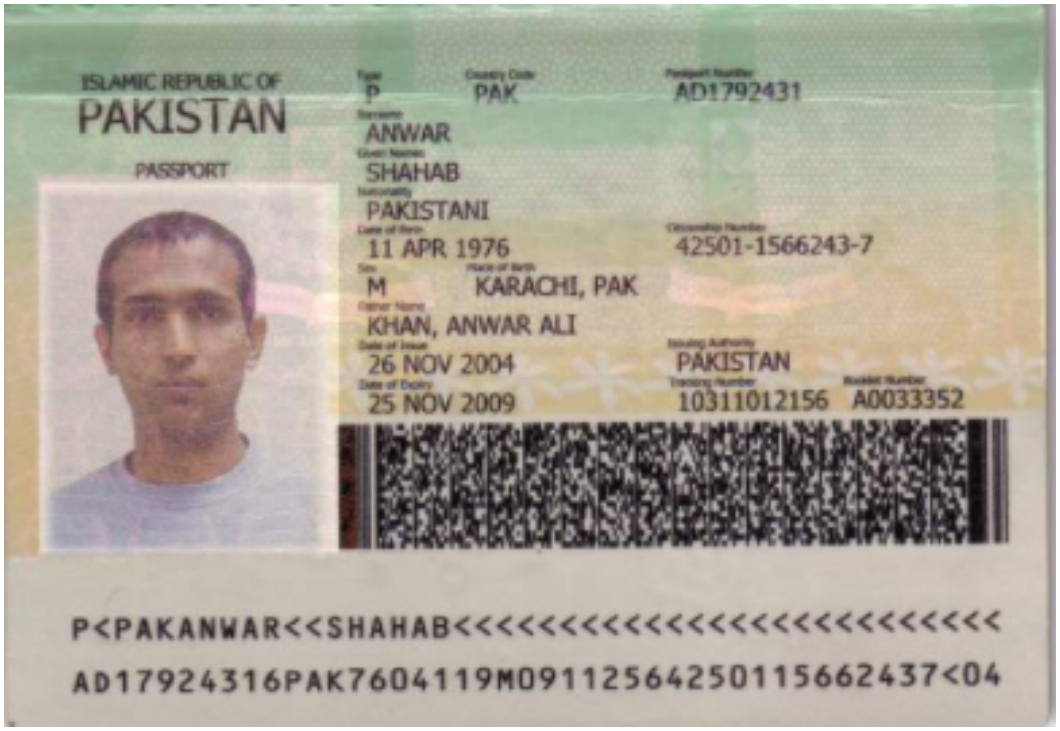
Dataset

For the purpose of object detection we gathered the dataset of UK, US and Pakistani passports. Since passports contain private information therefore it's difficult to find them lying around on the internet. Therefore collecting passport images as our dataset was one of the major challenges that we faced during the initial stage of our project. In the end we were just able to collect 25 images each of all UK, US and Pakistani passports.

UK Passport

US Passport

Pakistani Passport

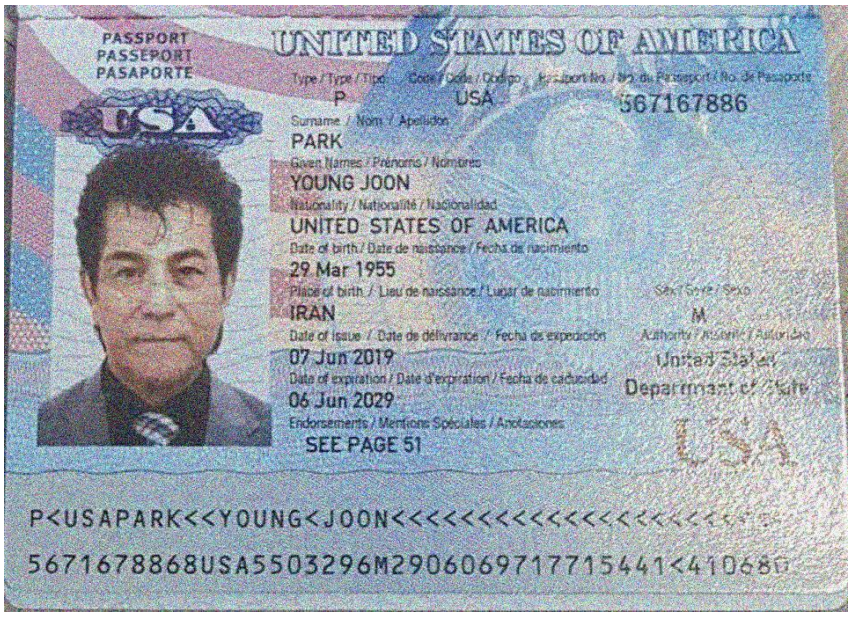


Data Augmentation

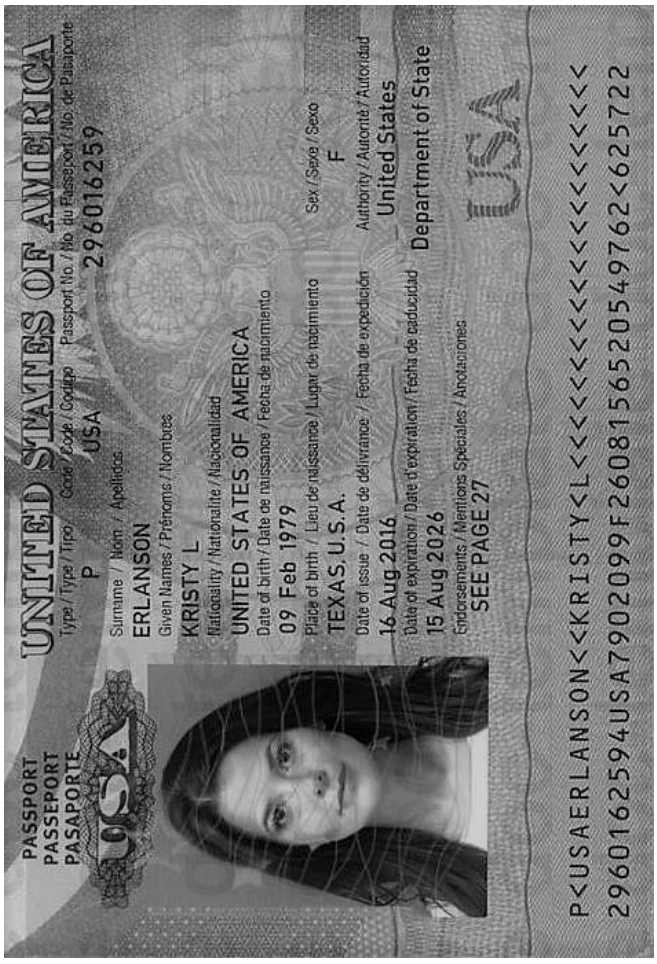
The accuracy of a neural network is very much dependent on the quantity and quality of the data collected. Therefore we need to have a sizable dataset because the efficiency of the neural network is based upon it. So in order to increase the size of the dataset we used several augmentation techniques such as addition of noise, rotation and cropping images, change in brightness etc.

Rotation

Noise Addition



Color Augmentation



Change in Brightness

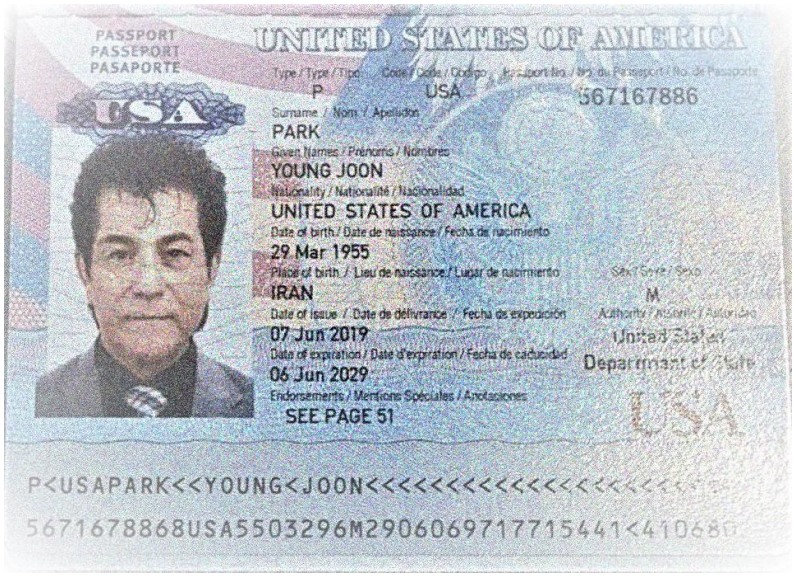


Image Annotation

After collecting the data we have performed annotations on each image and labeled them and converted them to the xml files. We used LabelImg for the annotation of the images in our dataset. LabelImg is a simple and standalone online annotator for image, video and audio. The name, identification number, date of birth and address of the person was marked and annotated in the Passport.

Number 1 indicates the name of the person, number 2 indicates date of birth, number 3 indicates address and number 4 indicates identification number of the person in the passport.

YOLO3 Algorithm

You Only Look Once, also known as YOLO is a very fast and accurate real time object detection technique used nowadays. YOLO is 4 times faster than the other object detection technique and we can simply tradeoff between accuracy and speed simply by changing the size of the model without needing to retrain the entire model. In order to use yolo3 we need to have darknet installed on our system. If you don't already have Darknet installed then you can use the following command to do that:

```
git clone https://github.com/AlexeyAB/darknet
```

Darknet prints out the objects it detected, its confidence, and how long it took to find them. We didn't compile Darknet with OpenCV so it can't display the detections directly. Instead, it saves them in *predictions.png*. You can open it to see the detected objects. Since we are using Darknet on the CPU it takes around 6-12 seconds per image. If we use the GPU version it would be much faster. After the training is complete we use the test images to test the output of the result. The new YOLOv3 uses independent logistic classifiers and binary cross-entropy loss for the class predictions during training. These edits make it possible to use complex datasets such as Microsoft's Open Images Dataset (OID) for YOLOv3 model training.

Here are some screenshots of our settings in .cfg, .data and .names files.

obj.data

```
classes =4
train = data/train.txt
valid = data/test.txt
names = data/obj.names
backup = /mydrive/yolov3/backup/
```

obj.names

```
Name
Address
ID
DoB
```

yolov3_custom.cfg

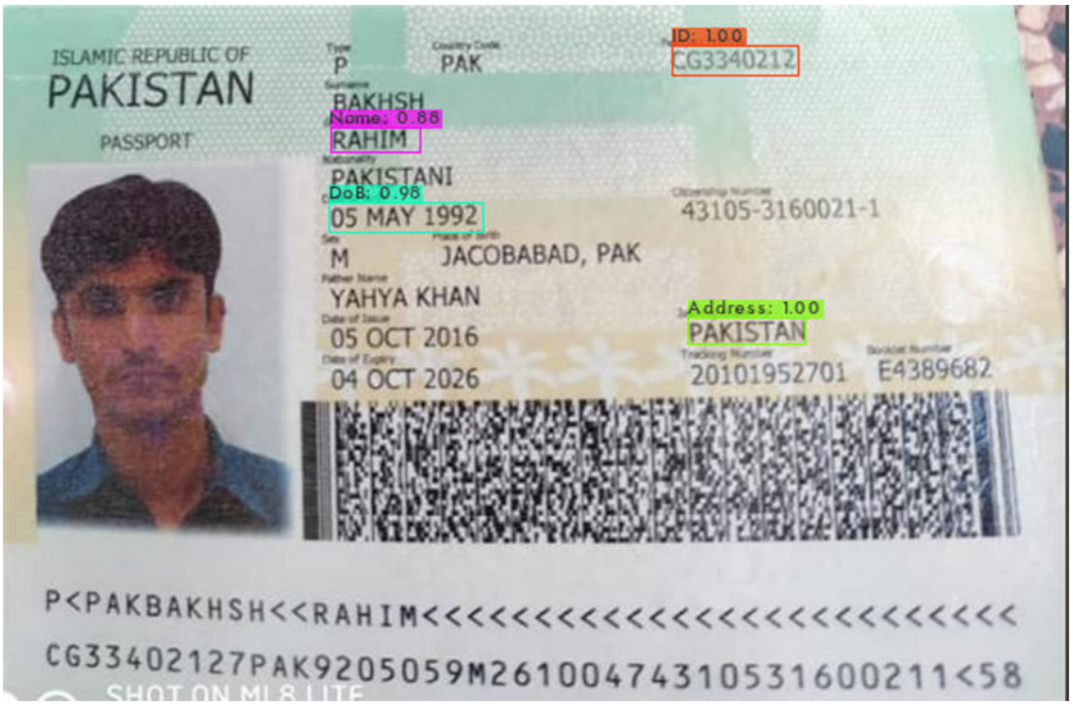
```
[net]
# Testing
#batch=1
#subdivisions=1
# Training
batch=64
subdivisions=16
```

```
learning_rate=0.001
burn_in=1000
max_batches = 8000
policy=steps
steps=6400,7200
scales=.1,.1


[convolutional]
size=1
stride=1
pad=1
filters=27
activation=linear


[yolo]
mask = 0,1,2
anchors = 10,13,  16,30,  33,23,  30,61,  62,45,  59,119,  116,90,  156,198,  373,326
classes=4
```

Object Detection in US Passport

Object Detection in UK Passport

Object Detection in Pakistani Passport



Questions Related to Project

1. How many classes have you considered to tackle this problem and how?

We gathered following classes to tackle this problem of Object Detection

- US Passport
- UK Passport
- Pakistani Passport

2. # of Images:

A. How many original images for US, UK and Pakistani Passport?

Initially we just had 25 images each of US, UK and Pakistani Passport

B. How many do you have in total after annotations?

After performing augmentation we increased our dataset to 175 images each for all UK, US and Pakistani passports. Making it a total of 525 images.

C. Training and test split (how many images for training and how many for test?)

We divided our dataset in 70/30 proportion. 70% of the dataset was used for training of our neural network while the other 30% of the dataset was used for testing the neural network.

3. What augmentation method/approach you used?

We used both Position and Color augmentation techniques for the lengthening of our dataset.

Position Augmentation

We rotated each passport in 4 directions and also added noise in the images using MATLAB.

Color Augmentation

After performing position augmentation we altered the color of the images by changing its brightness and applying filters.

4. Provide annotation files for all images (original+annotated). Also provide one example original image of Australia and PK and its annotation and similarly , one augmented image of both countries with annotation file.

//Already Given Above

5. Provide all code files (main code +augmentation code).

//Attached with the document

6. Also mention the piece of code for transfer learning: Mention in the report which and how many layers you have removed and on how much data you retrained?

Backbone:

Darknet Detector is used as backbone architecture to extract features from the images.

Pretrained weights:

Yolov3 pretrained weights are used for training the model which were downloaded from here: `wget https://pjreddie.com/media/files/yolov3.weights`

Layers removed:

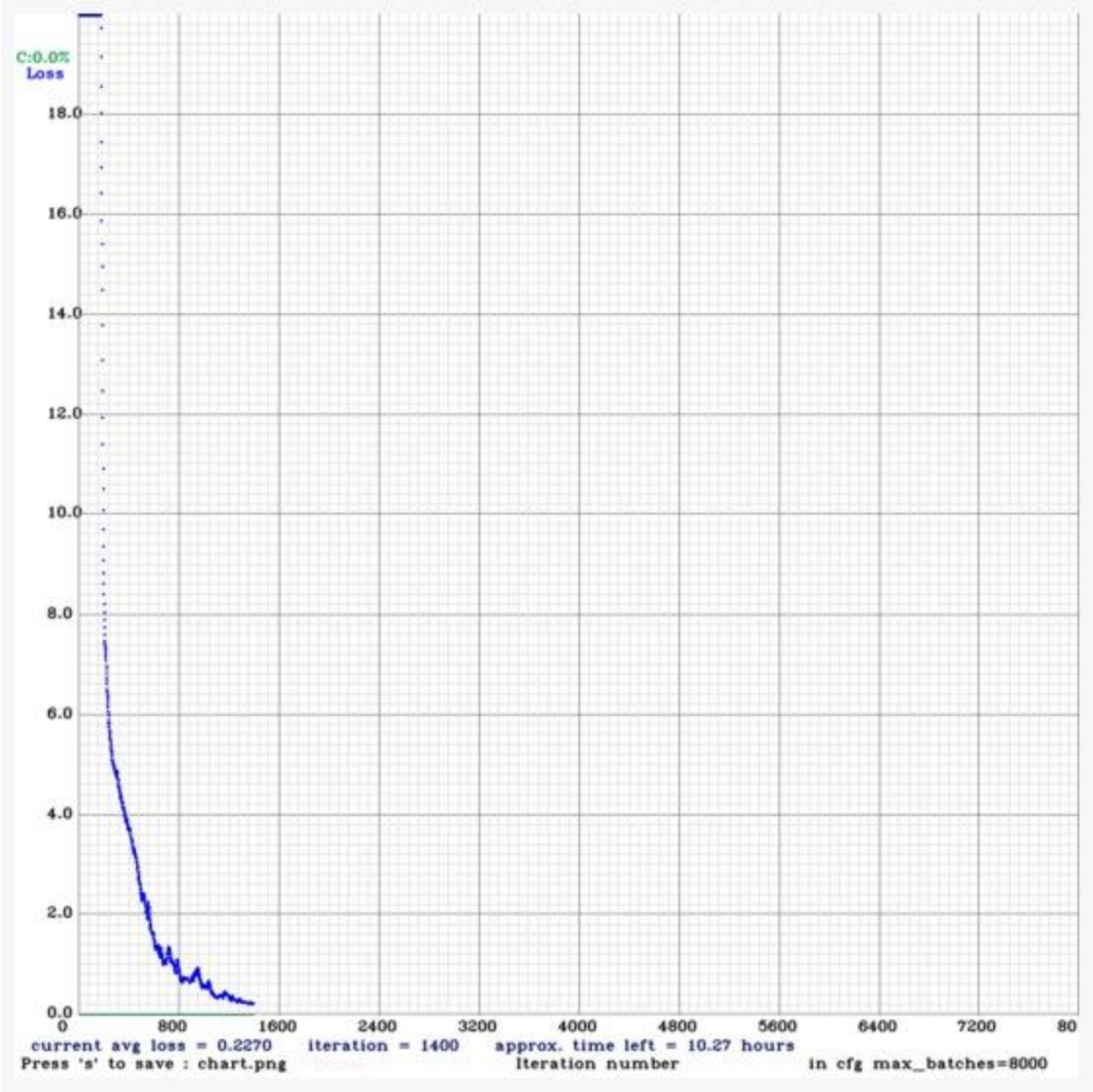
While training the model, the head layers have been removed only. The argument ‘layers’ is set as ‘heads’ in model.train(). This freezes all the layers except the head layers.

Hyperparameters:

By default, YOLO only displays objects detected with a confidence of .25 or higher which can change this by passing the -thresh <val> flag to the yolo command.

7. Provide detail on all the experiments with training +validation curves.

Given below is the graph for loss rate vs no. of iterations for training.

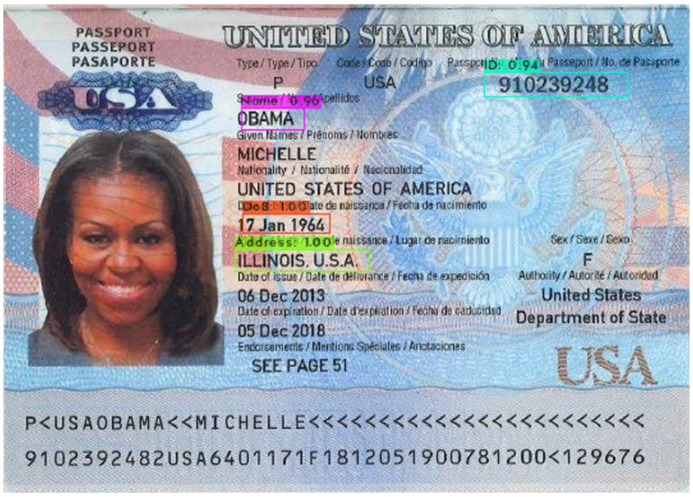


8. Report accuracy (How many cards were tested. How many labels were correctly identified and how many were missed for each class type?)

More than 65 images were tested for each class of our dataset. The lowest and highest accuracy for each label of all three classes are given below

USA Passport (Highest Accuracy for each label)				
	ID	Name	DoB	Address
Highest	1.00	0.98	1.00	1.00
Lowest	0.60	0.63	0.67	0.78
Average	0.80	0.805	0.835	0.88

Highest



Lowest

UK Passport (Highest Accuracy for each label)				
	ID	Name	DoB	Address
Highest	0.99	0.99	1.00	1.00
Lowest	0.72	0.64	0.67	0.63
Average	0.86	0.815	0.835	0.815

Highest

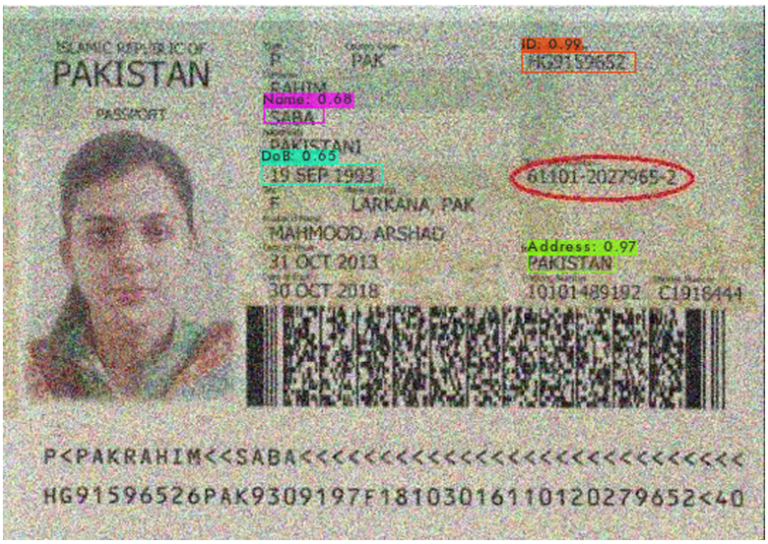
Lowest

Pakistani Passport (Highest Accuracy for each label)				
	ID	Name	DoB	Address
Highest	1.00	1.00	1.00	0.99
Lowest	0.70	0.64	0.64	0.67
Average	0.85	0.82	0.82	0.83

Highest


ISLAMIC REPUBLIC OF
PAKISTAN
PASSPORT
Type
P
Country Code
PAK
ID: 1.00
CX8109101
Name: 1.00
MUHAMMAD
DoB: 1.00
01 JAN 1993
Citizenship Number
36402-4338910-5
Sex
M
Place of Birth
BAHAWALNAGAR, PAK
Father Name
ALI, AMANAT
Address: 0.99
PAKISTAN
Date of Issue
07 JUL 2017
Date of Expiry
06 JUL 2022
Tracking Number
11631139177
E7825568
P<PAKJAHANZAIB<<MUHAMMAD<<<<<<<<<<<<<<<<<<<<
CX81091015PAK9301010M2207065364024338910 5<20

Lowest


PAKISTAN
PASSPORT
Type
P
Country Code
PAK
ID: 0.99
HG91596526
Name: 0.68
SABA
DoB: 0.65
19 SEP 1993
61101-2027965-2
Sex
F
Place of Birth
LARKANA, PAK
Father Name
MAHMOOD, ARSHAD
Address: 0.97
PAKISTAN
Date of Issue
31 OCT 2013
Date of Expiry
30 OCT 2018
10101489192
C1918444
P<PAKRAHIM<<SABA<<<<<<<<<<<<<<<<<<<<<<<<<<<<
HG91596526PAK9309197F1810301611012027965 2<40