Report for Programming Assignment-2

Yogesh Tripathi
(CS16B044)

October 23, 2019

Checklist : The task you need to ensure before submission.

☑ I have read all the instruction carefully and followed them to my best ability.

☑ I have written the name, roll no in report.

☑ Run sanity_check.sh.

☑ I will be submitting only single submission on behalf of my team.

☑ I have not included unnecessary text, pages, logos in the assignment.

☑ I have not used any high level APIs(Keras, Estimators for e.g.).

☑ I have not copied anything for this assignment.

For consolidation of all codes, trained and saved models, please check the following drive link.

Answer 1

Implementation of Faster RCNN [1], RetinaNet [2] and YOLO-v3 [3] for DVQA-dataset is provided in the google drive link as `PA2 fasterrcnn.ipynb`, `PA retina.ipynb` and `PA yolo.ipynb` colab files. I have used detectron-this repository implementation for implementing Faster-RCNN and RetinaNet. Yolo-v3 is taken from darknet repository-here. The best performing parameters are:

Parameter	Best Value
Convolution Body	ResNet 50
Base Learning Rate	0.0025
Weight Decay	0.0001
Number of iterations	25000
IoU	0.5

Table 1: Best performing hyperparameters for Faster-RCNN

Parameter	Best Value
Convolution Body	ResNet 50
Base Learning Rate	0.001
Weight Decay	0.000125
Number of iterations	25000
IoU	0.5
Loss γ	2.0
Loss α	0.25

Table 2: Best performing hyperparameters for RetinaNet

Parameter	Best Value
Convolution Body	ResNet 50
Learning Rate	0.001
Number of iterations	25000
IoU	0.5
Batch	64
Subdivisions	16

Table 3: Best performing hyperparameters for YOLO-v3

Answer 2

The mAP for IoU=0.5 on public test split is reported below.

Class	mAP Value for IoU=0.5
bar	0.892
legend-heading	0.955
legend-label	0.929
title	0.984
xlabel	0.958
xticklabel	0.938
ylabel	0.985
yticklabel	0.957
Final mAP score	**0.9498**

Table 4: mAP value for IoU=0.5 for Faster-RCNN

Class	mAP Value for IoU=0.5
bar	0.656
legend-heading	0.917
legend-label	0.905
title	0.822
xlabel	0.907
xticklabel	0.925
ylabel	0.954
yticklabel	0.959
Final mAP score	**0.8805**

Table 5: mAP value for IoU=0.5 for RetinaNet

Class	mAP Value for IoU=0.5
`bar`	65.2
`legend-heading`	90.8
`legend-label`	89.7
`title`	81.4
`xlabel`	89.0
`xticklabel`	92.6
`ylabel`	93.2
`yticklabel`	94.3
Final mAP score	**0.8702**

Table 6: mAP value for IoU=0.5 for YOLO-v3

Answer 3

The mAP values for IoU = 0.75 is as summarized below:

Model	mAP Value for IoU=0.75
Faster-RCNN	0.9504
RetinaNet	0.8733
YOLO-v3	0.8491

Table 7: mAP value for IoU=0.75

The mAP values for IoU = 0.9 is as summarized below:

Model	mAP Value for IoU=0.9
Faster-RCNN	0.9374
RetinaNet	0.8733
YOLO-v3	0.8272

Table 8: mAP value for IoU=0.9

Answer 4

The given plot shows mAP scores on public test data for IoU = {0.5,0.75,0.9} for Faster-RCNN, RetinaNet and YOLO-v3. We observe that Faster-RCNN performes the best and in general, an increase in IoU threshold from 0.5 to 0.75 to 0.9, reduces the mAP score.

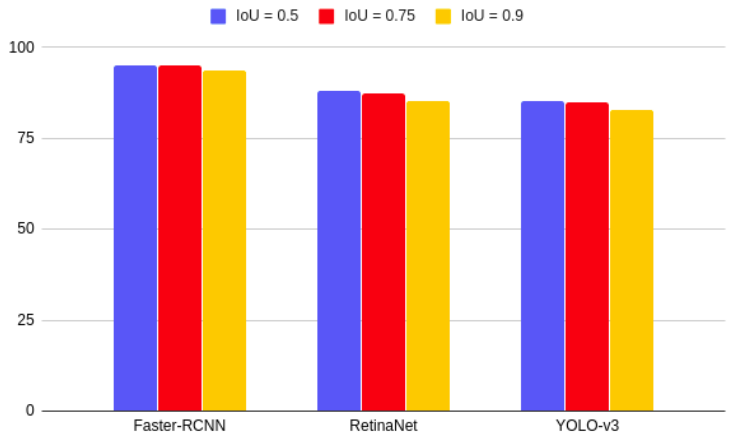

Figure 1: Plot of mAP scores for different IoU values

Answer 5

The code is consolidated in PA2-retina with Cross-Entropy loss configuration in `retinanet_R-50-FPN_1x.yaml`. The mAP score for IoU = 0.5 was obtained to be **0.8807**. The IoU for each class is summarized below. There is not much improvement observed ny replacing Cross-Entropy loss by Focal Loss in RetinaNet architecture, mostly due to easy examples in public test set.

Class	mAP Value for IoU=0.5
`bar`	0.655
`legend-heading`	0.939
`legend-label`	0.903
`title`	0.784
`xlabel`	0.935
`xticklabel`	0.923
`ylabel`	0.963
`yticklabel`	0.945
Final mAP score	**0.8807**

Table 9: mAP value for IoU=0.5 for RetinaNet(with Cross-Entropy loss)

Answer 6

Faster-RCNN performs the best among all the models for DVQA dataset. This is because this dataset has simple, easily identifiable rectangle proposals like bar, legend, etc. The training data is also quite small in size and therefore a very complex model would have high variance. Faster-RCNN is a simpler model and pre-trained ResNet-50 is also better fine-tuned due to lower variance in model. Hence, it outperforms RetinaNet and YOLO-v3.

The output files for private test is present in the submission in the specified format.

References

[1] Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun. Faster R-CNN: towards real-time object detection with region proposal networks. *CoRR*, abs/1506.01497, 2015.

[2] Tsung-Yi Lin, Priya Goyal, Ross B. Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. *CoRR*, abs/1708.02002, 2017.

[3] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv*, 2018.