Assignment 2: Weakly Supervised Object Localization

- Visual Learning and Recognition (16-824) Spring 2018

Name: Bhavan Jasani
Andrew ID: bjasani

Task 0

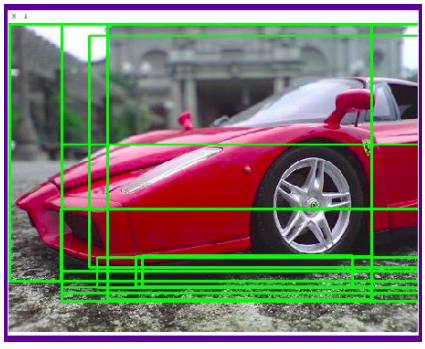
Q 0.1: WHAT CLASSES DOES THE IMAGE AT INDEX 2018 CONTAIN?

Class: car

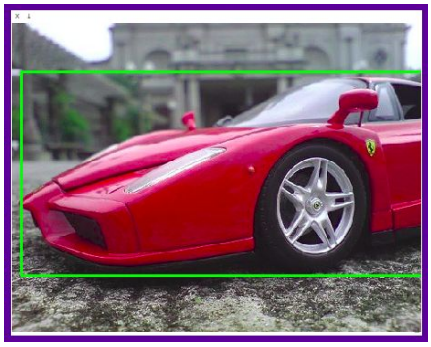
Q 0.2: WHAT IS THE FILENAME OF THE IMAGE AT INDEX 2018?

File name: 003998.jpg

Q 0.3 USE VISDOM+CV2 TO VISUALIZE THE TOP 10 BOUNDING BOX PROPOSALS FOR IMAGE AT INDEX 2018. YOU WOULD NEED TO FIRST PLOT THE IMAGE AND THEN PLOT THE RECTANGLES FOR EACH BOUNDING BOX PROPOSAL.



Q 0.4 USE VISDOM+CV2 TO VISUALIZE THE GROUND TRUTH BOXES FOR IMAGE AT INDEX 2018.



Task 1

Q1.1

Custom.py -> find_classes
This function creates a mapping from classes to index.

Custom.py -> Make_dataset
Reads images and and corresponding labels

Custom.py -> LocalizerAlexNet(nn.Module): __init__
Architecture definition for the Localizer AlexNet Model.

Custom.py -> LocalizerAlexNet(nn.Module): forward
Forward propagation for LocalizerAlexNet

Custom.py -> LocalizerAlexNetRobust(nn.Module): __init__
Architecture definition for the Localizer AlexNet Robust Model.

Custom.py -> LocalizerAlexNetRobust(nn.Module): forward
Forward propagation for LocalizerAlexNetRobust

Custom.py -> localizer_alexnet(pretrained=False, **kwargs):
Creates the Alex Net architecture and loads the pretrained weights

Custom.py -> localizer_alexnet_robust(pretrained=False, **kwargs):
Creates the Alex Net Robust architecture and loads the pretrained weights

Custom.py -> IMDBDataset
Getting data (images and the ground truth labels) from the class IMDB

Main.py ->
Defined loss function and optimizer
Created loggers for visdom and tboard
Added output from model
Global Max pooled output and also upscaled it to get the heatmaps
Computed gradient and do SGD step
Added Visualizations

Defined Metric 1
Defined Metric 2

Q1.2

Output resolution of the model is 20x29x29 for an input image of size 512x512, where 29x29 is the spatial dimension and 20 is the number of classes more generally it is batch_size x no_classes x spatial resolution

Q1.4

Initially we observe a very steep drop in lose because there is imbalance of number of labels/class of objects present in the images, as in most images have very few (less than 3) classes of objects present and so initially the network tries to optimize the loss by predicting everything as zero, that is no class is present in the images. This results in a very steep decrease in lose. But of course after that the network starts to learn the actual objects present in the images.

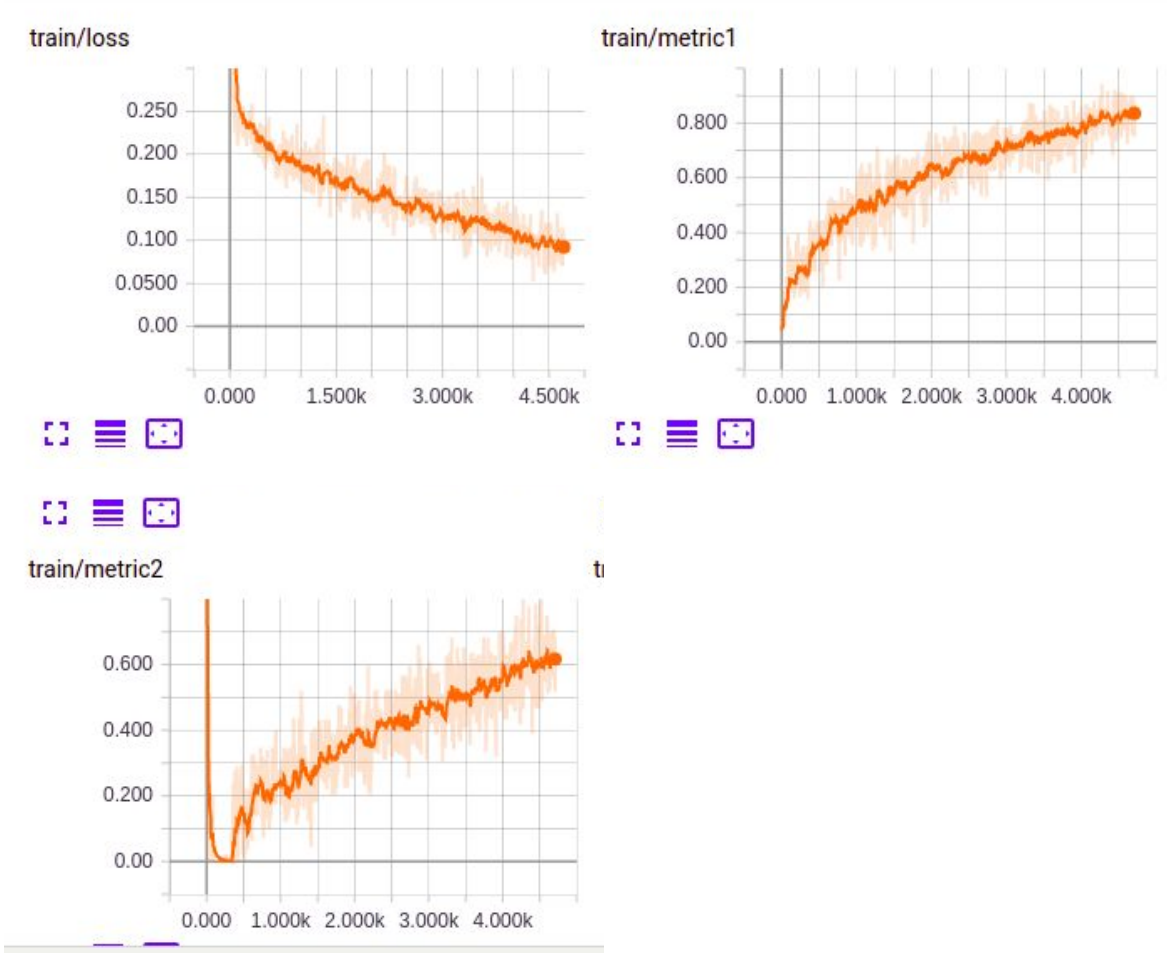
Q1.5

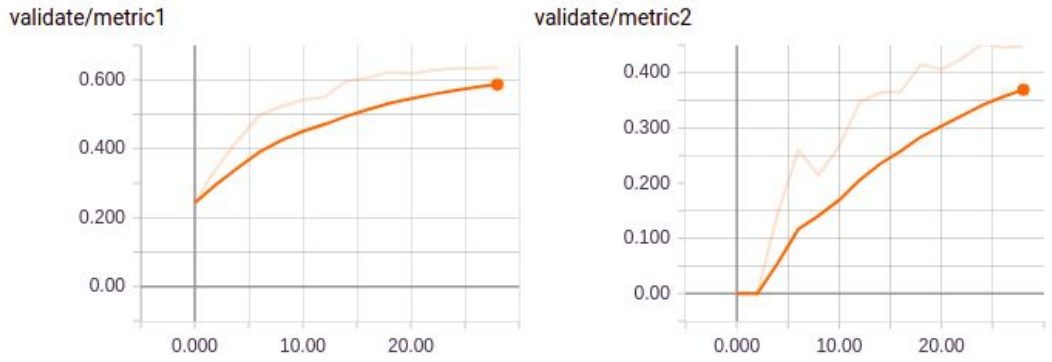
I am using the following:

Metric 1: mean average precision (mAP), as was used in HW1

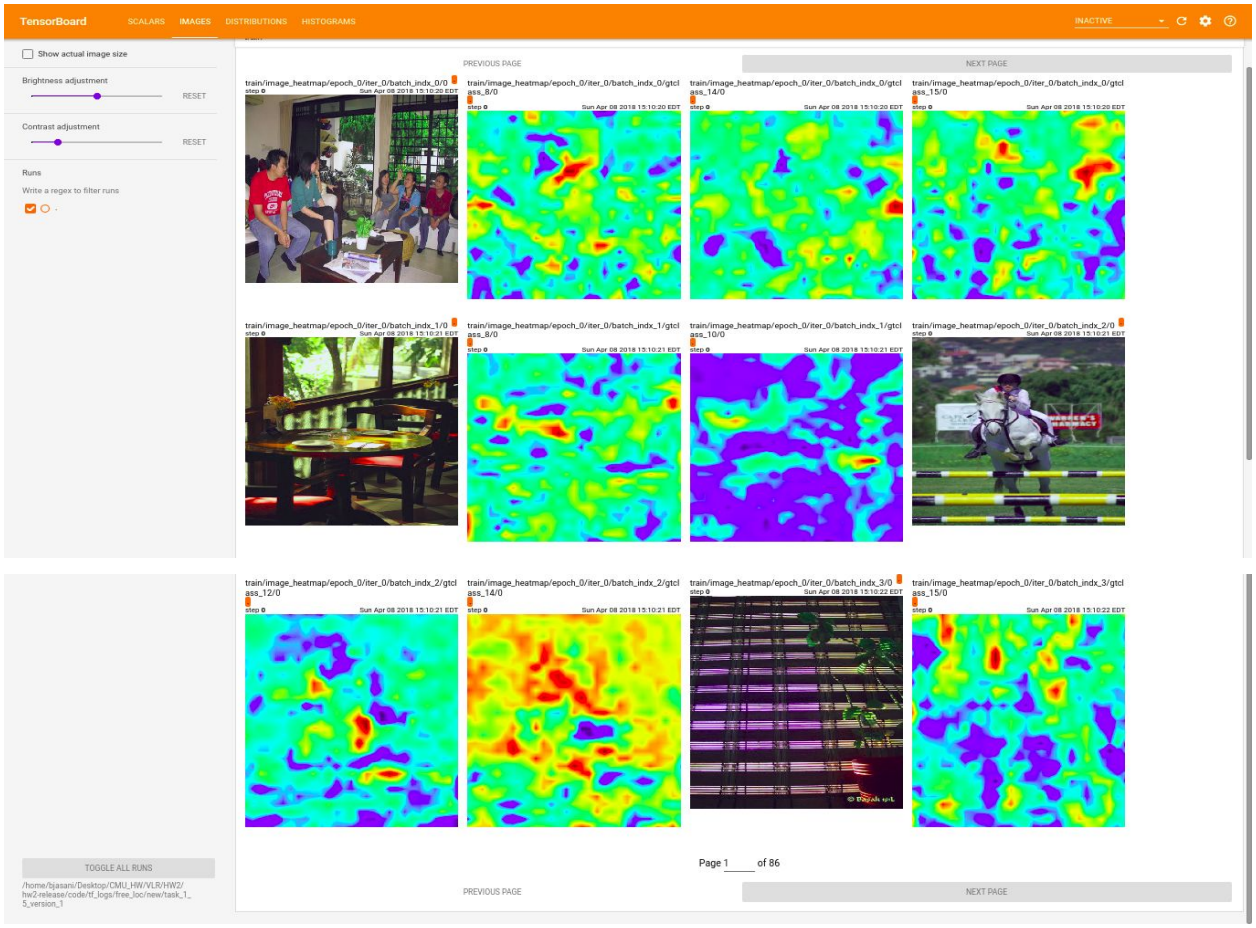
Metric 2: recall with a threshold of 0.5 (if the activation is above 0.5 then the class is considered to be present or else absent). Recall would only count the true positive for the classes which have ground truth labels present and so would be more robust.

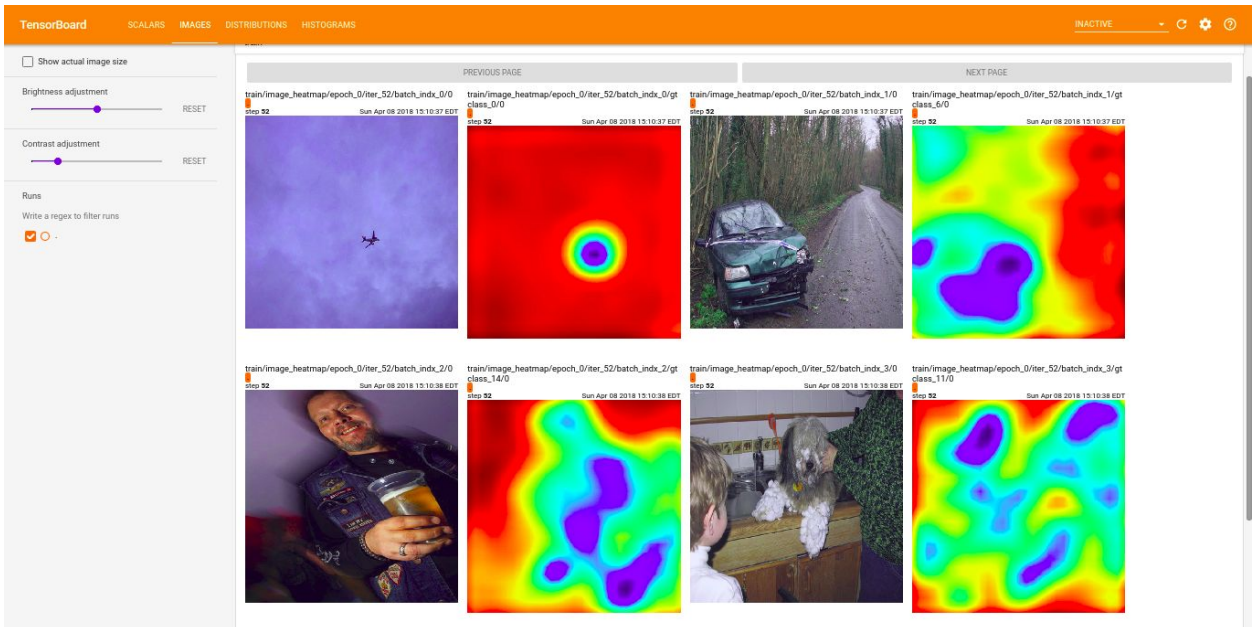
Both of these are implementates using sklearn in the TODO blocks

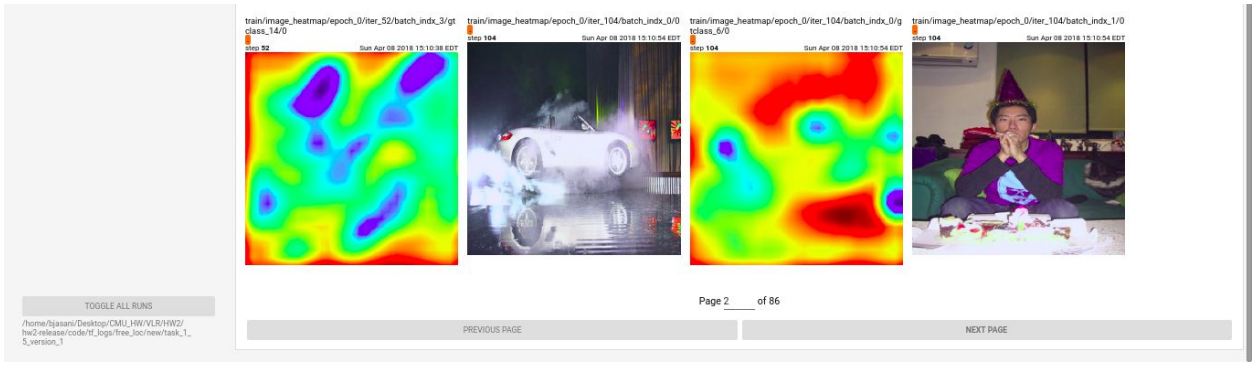


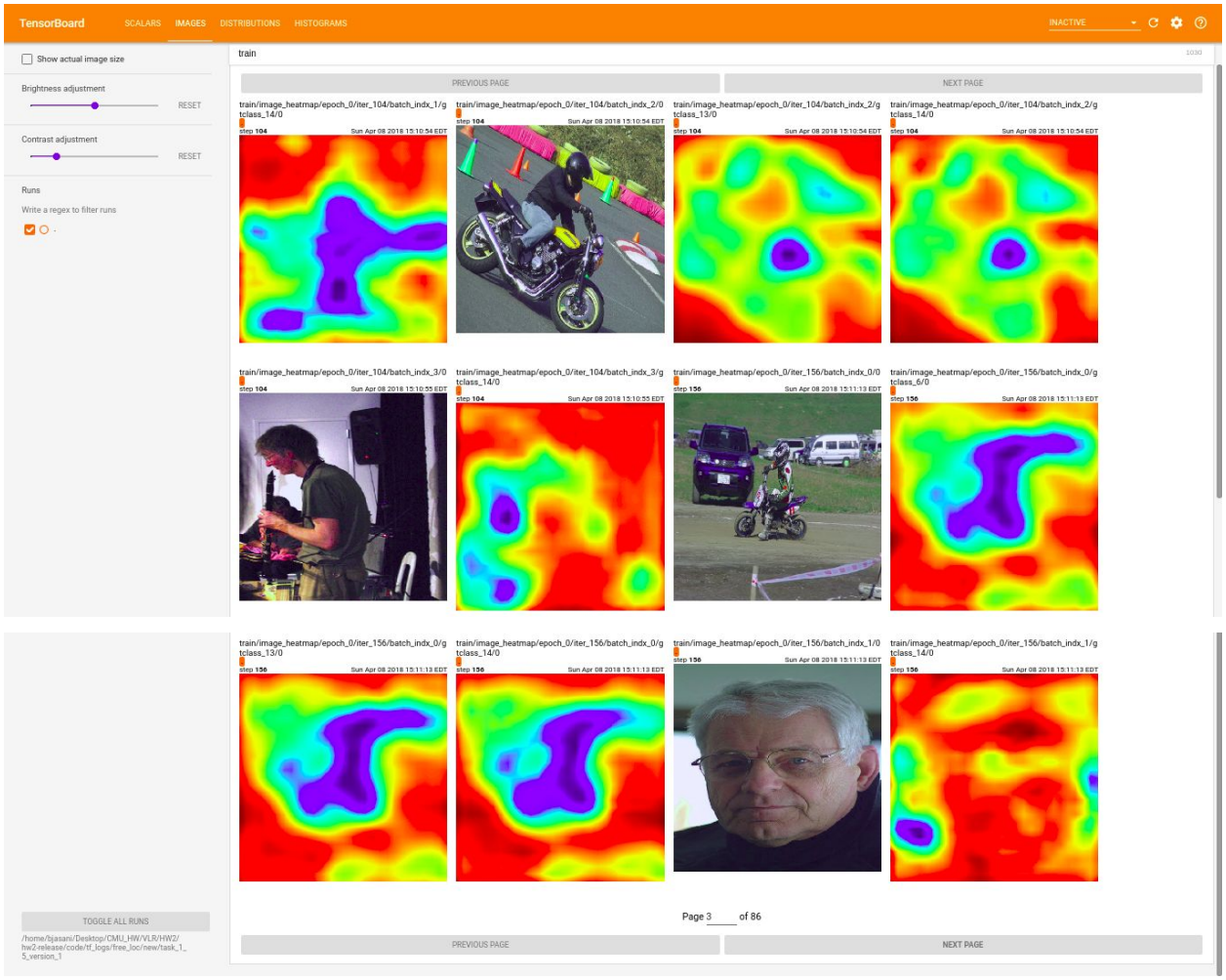


Tensorboard images and heatmap for first logged epoch


TensorBoard
SCALARS
IMAGES
DISTRIBUTIONS
HISTOGRAMS
INACTIVE
Show actual image size
Brightness adjustment
RESET
Contrast adjustment
RESET
Runs
Write a regex to filter runs
PREVIOUS PAGE
NEXT PAGE
train/image_heatmap/epoch_0/iter_0/batch_indx_0/0
train/image_heatmap/epoch_0/iter_0/batch_indx_0/gtclass_8/0
train/image_heatmap/epoch_0/iter_0/batch_indx_0/gtclass_14/0
train/image_heatmap/epoch_0/iter_0/batch_indx_0/gtclass_15/0
train/image_heatmap/epoch_0/iter_0/batch_indx_1/0
train/image_heatmap/epoch_0/iter_0/batch_indx_1/gtclass_8/0
train/image_heatmap/epoch_0/iter_0/batch_indx_1/gtclass_10/0
train/image_heatmap/epoch_0/iter_0/batch_indx_2/0
train/image_heatmap/epoch_0/iter_0/batch_indx_2/gtclass_12/0
train/image_heatmap/epoch_0/iter_0/batch_indx_2/gtclass_14/0
train/image_heatmap/epoch_0/iter_0/batch_indx_3/0
train/image_heatmap/epoch_0/iter_0/batch_indx_3/gtclass_15/0
TOGGLE ALL RUNS
/home/sjasani/Desktop/CMU_HW/VLR/HW2/hw2-release/code/tf_logs/free_loc/new/task_1_5_version_1
Page 1 of 86
PREVIOUS PAGE
NEXT PAGE

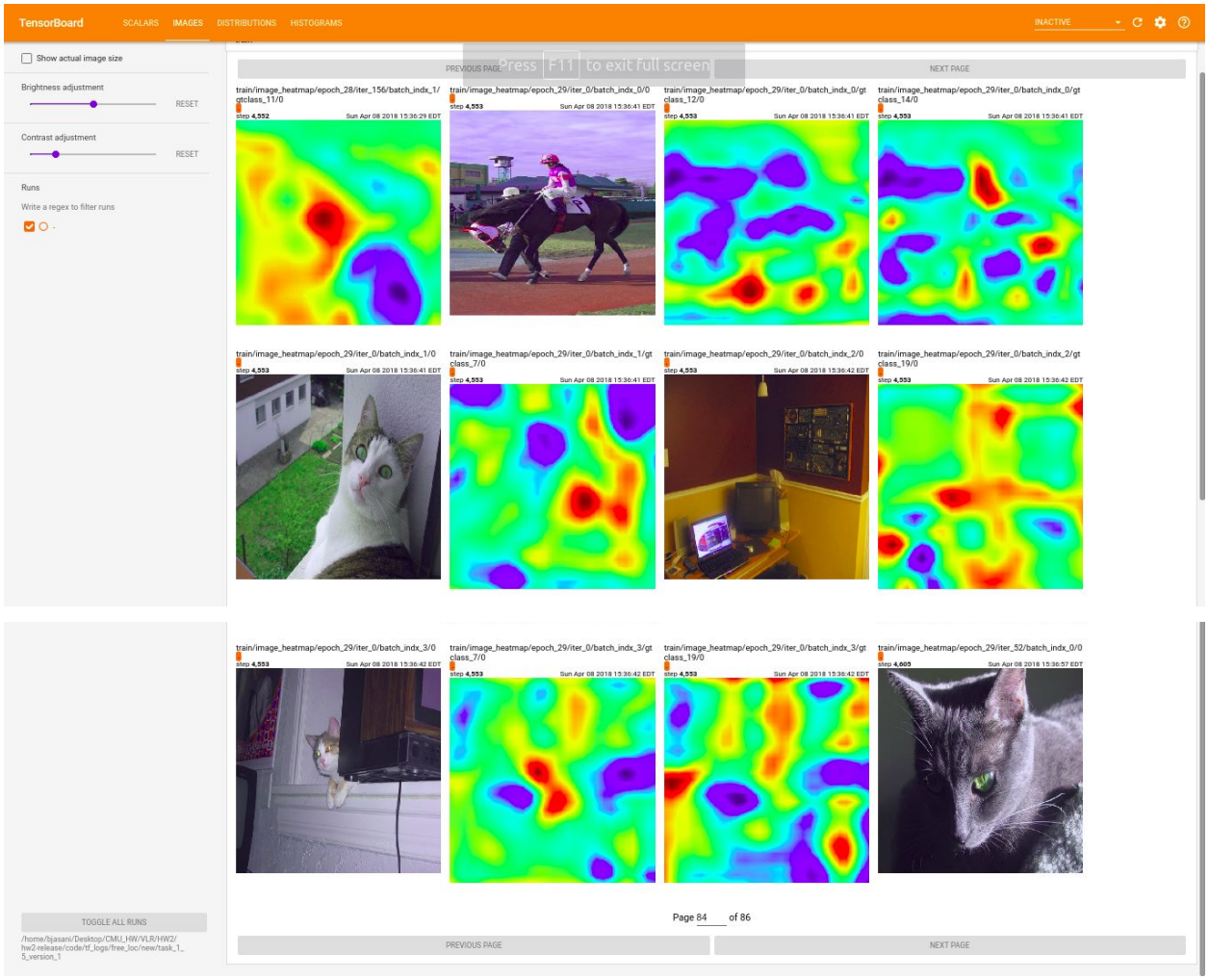

TensorBoard
SCALARS
IMAGES
DISTRIBUTIONS
HISTOGRAMS
INACTIVE
Show actual image size
Brightness adjustment
RESET
Contrast adjustment
RESET
Runs
Write a regex to filter runs
PREVIOUS PAGE
NEXT PAGE
train/image_heatmap/epoch_0/iter_52/batch_indx_0/0
train/image_heatmap/epoch_0/iter_52/batch_indx_0/gtclass_0/0
train/image_heatmap/epoch_0/iter_52/batch_indx_1/0
train/image_heatmap/epoch_0/iter_52/batch_indx_1/gtclass_6/0
train/image_heatmap/epoch_0/iter_52/batch_indx_2/0
train/image_heatmap/epoch_0/iter_52/batch_indx_2/gtclass_14/0
train/image_heatmap/epoch_0/iter_52/batch_indx_3/0
train/image_heatmap/epoch_0/iter_52/batch_indx_3/gtclass_11/0

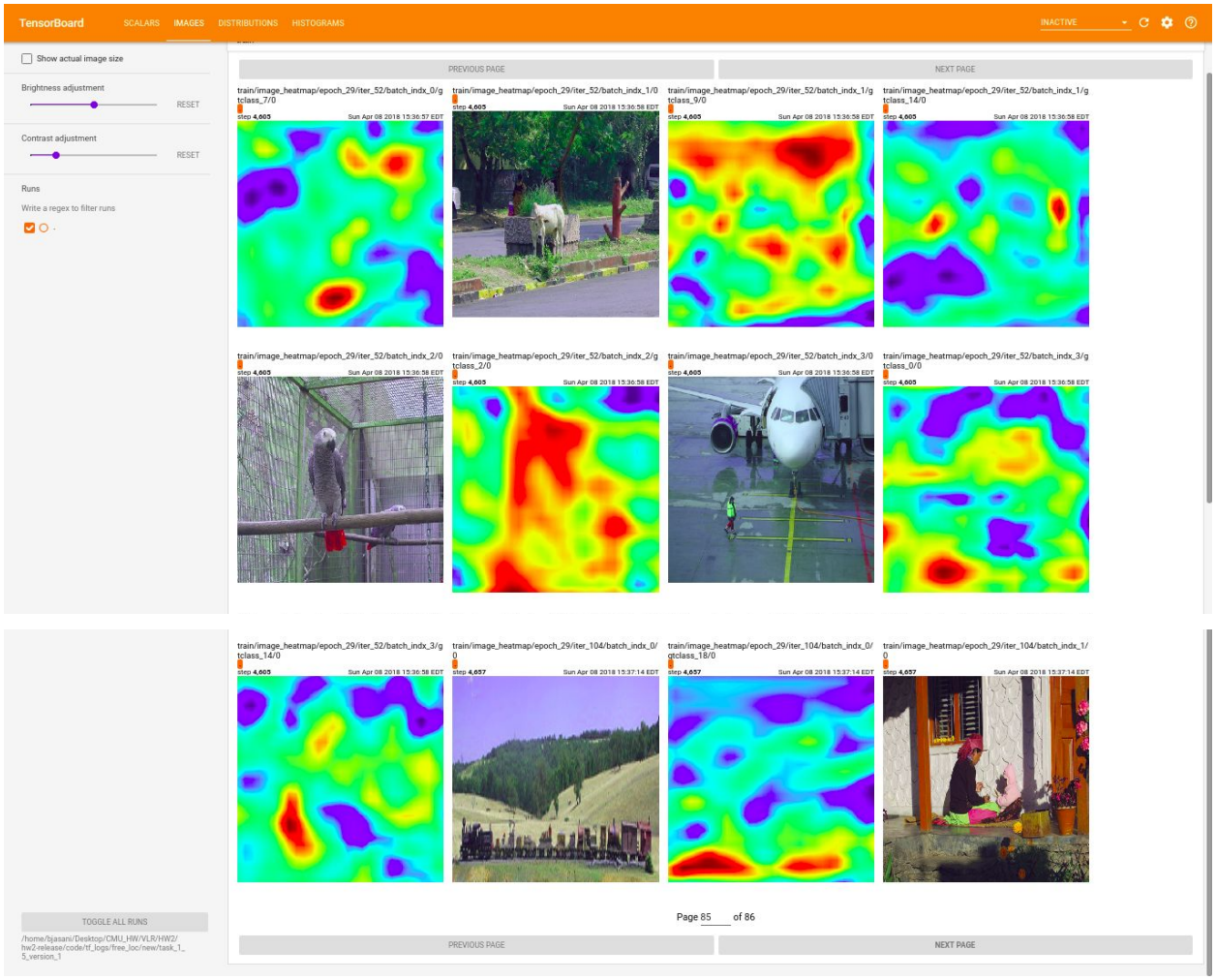

train/image_heatmap/epoch_0/iter_52/batch_indx_3/gt class_14/0
train/image_heatmap/epoch_0/iter_104/batch_indx_0/0
train/image_heatmap/epoch_0/iter_104/batch_indx_0/g tclass_6/0
train/image_heatmap/epoch_0/iter_104/batch_indx_1/0
TOGGLE ALL RUNS
/home/bjasani/Desktop/CMU_HW/VLR/HW2/hw2-release/code/tf_logs/free_loc/new/task_1_5_version_1
Page 2 of 86
PREVIOUS PAGE
NEXT PAGE

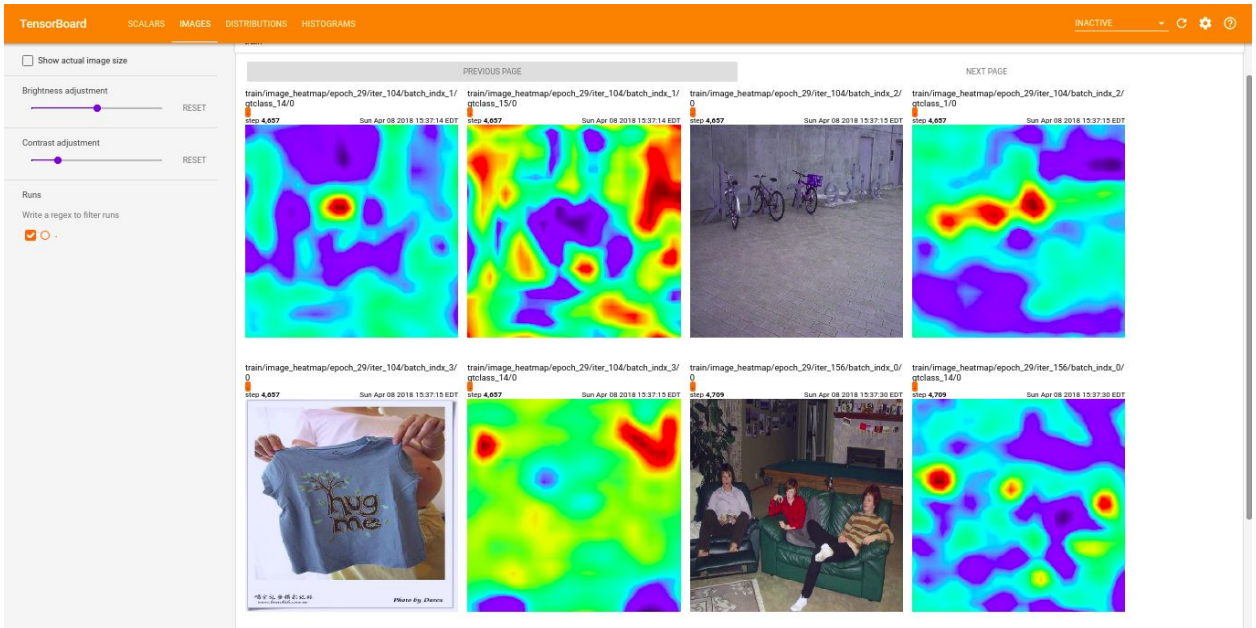

TensorBoard
SCALARS
IMAGES
DISTRIBUTIONS
HISTOGRAMS
INACTIVE
Show actual image size
Brightness adjustment
RESET
Contrast adjustment
RESET
Runs
Write a regex to filter runs
train
PREVIOUS PAGE
NEXT PAGE
train/image_heatmap/epoch_0/iter_104/batch_indx_1/g tclass_14/0
train/image_heatmap/epoch_0/iter_104/batch_indx_2/0
train/image_heatmap/epoch_0/iter_104/batch_indx_2/g tclass_13/0
train/image_heatmap/epoch_0/iter_104/batch_indx_2/g tclass_14/0
train/image_heatmap/epoch_0/iter_104/batch_indx_3/0
train/image_heatmap/epoch_0/iter_104/batch_indx_3/g tclass_14/0
train/image_heatmap/epoch_0/iter_156/batch_indx_0/0
train/image_heatmap/epoch_0/iter_156/batch_indx_0/g tclass_6/0

train/image_heatmap/epoch_0/iter_156/batch_indx_0/g tclass_13/0
train/image_heatmap/epoch_0/iter_156/batch_indx_0/g tclass_14/0
train/image_heatmap/epoch_0/iter_156/batch_indx_1/0
train/image_heatmap/epoch_0/iter_156/batch_indx_1/g tclass_14/0
TOGGLE ALL RUNS
/home/bjasani/Desktop/CMU_HW/VLR/HW2/hw2-release/code/tf_logs/free_loc/new/task_1_5_version_1
Page 3 of 86
PREVIOUS PAGE
NEXT PAGE

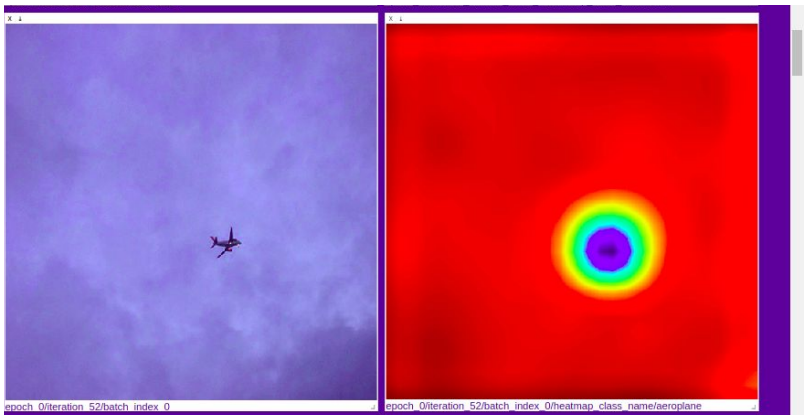
Tensorboard images and heatmap for last logged epoch

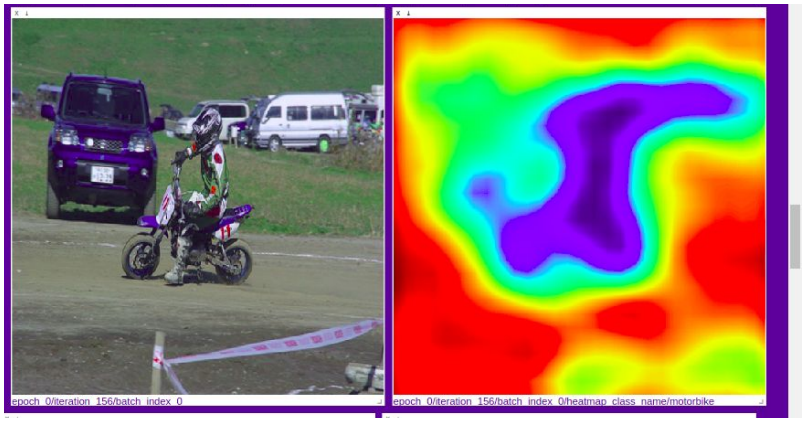


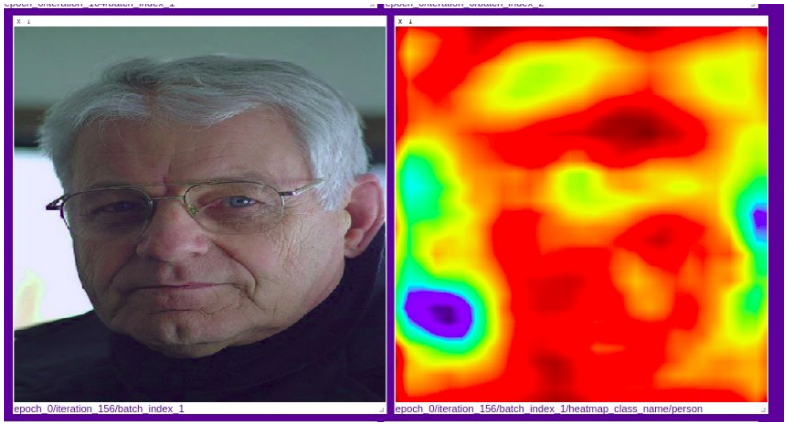

TensorBoard
SCALARS
IMAGES
DISTRIBUTIONS
HISTOGRAMS
INACTIVE
Show actual image size
Brightness adjustment
RESET
Contrast adjustment
RESET
Runs
Write a regex to filter runs
PREVIOUS PAGE
NEXT PAGE
train/image_heatmap/epoch_29/iter_52/batch_indx_0/gtclass_7/0
train/image_heatmap/epoch_29/iter_52/batch_indx_1/0
train/image_heatmap/epoch_29/iter_52/batch_indx_1/gtclass_9/0
train/image_heatmap/epoch_29/iter_52/batch_indx_1/gtclass_14/0
train/image_heatmap/epoch_29/iter_52/batch_indx_2/0
train/image_heatmap/epoch_29/iter_52/batch_indx_2/gtclass_2/0
train/image_heatmap/epoch_29/iter_52/batch_indx_3/0
train/image_heatmap/epoch_29/iter_52/batch_indx_3/gtclass_0/0
train/image_heatmap/epoch_29/iter_52/batch_indx_3/gtclass_14/0
train/image_heatmap/epoch_29/iter_104/batch_indx_0/0
train/image_heatmap/epoch_29/iter_104/batch_indx_0/gtclass_18/0
train/image_heatmap/epoch_29/iter_104/batch_indx_1/0
TOGGLE ALL RUNS
/home/sjasani/Desktop/CMU_HW/VLR/HW2/hw2-release/code/tf_logs/free_loc/new/task_1_5_version_1
Page 85 of 86
PREVIOUS PAGE
NEXT PAGE

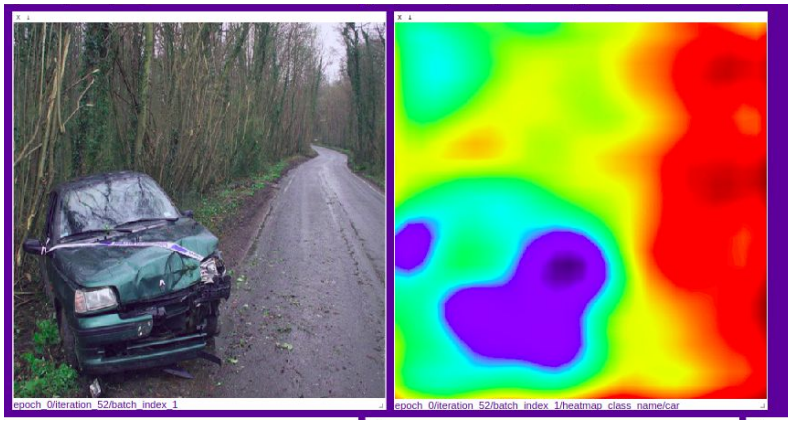

TensorBoard
SCALARS
IMAGES
DISTRIBUTIONS
HISTOGRAMS
INACTIVE
Show actual image size
Brightness adjustment
RESET
Contrast adjustment
RESET
Runs
Write a regex to filter runs
PREVIOUS PAGE
NEXT PAGE
train/image_heatmap/epoch_29/iter_104/batch_indx_1/gtclass_14/0
train/image_heatmap/epoch_29/iter_104/batch_indx_1/gtclass_15/0
train/image_heatmap/epoch_29/iter_104/batch_indx_2/0
train/image_heatmap/epoch_29/iter_104/batch_indx_2/gtclass_1/0
train/image_heatmap/epoch_29/iter_104/batch_indx_3/0
train/image_heatmap/epoch_29/iter_104/batch_indx_3/gtclass_14/0
train/image_heatmap/epoch_29/iter_156/batch_indx_0/0
train/image_heatmap/epoch_29/iter_156/batch_indx_0/gtclass_14/0

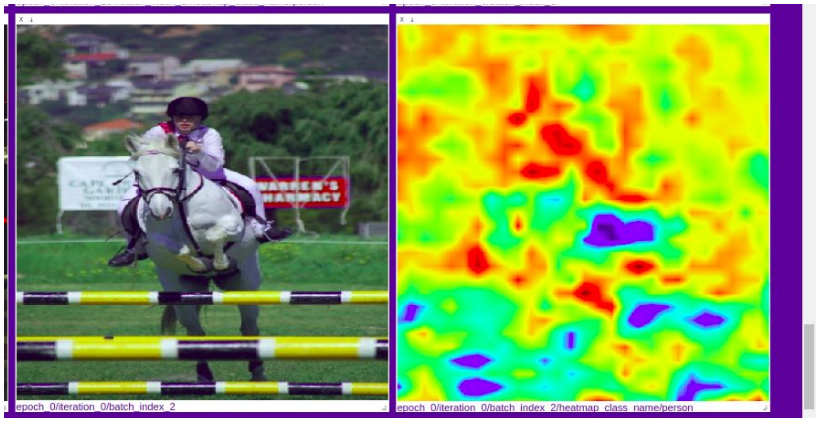
Visdom images and heatmap for first logged epoch

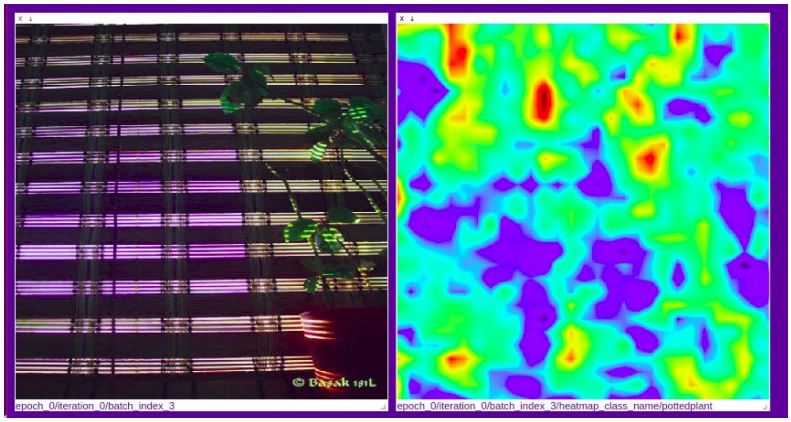


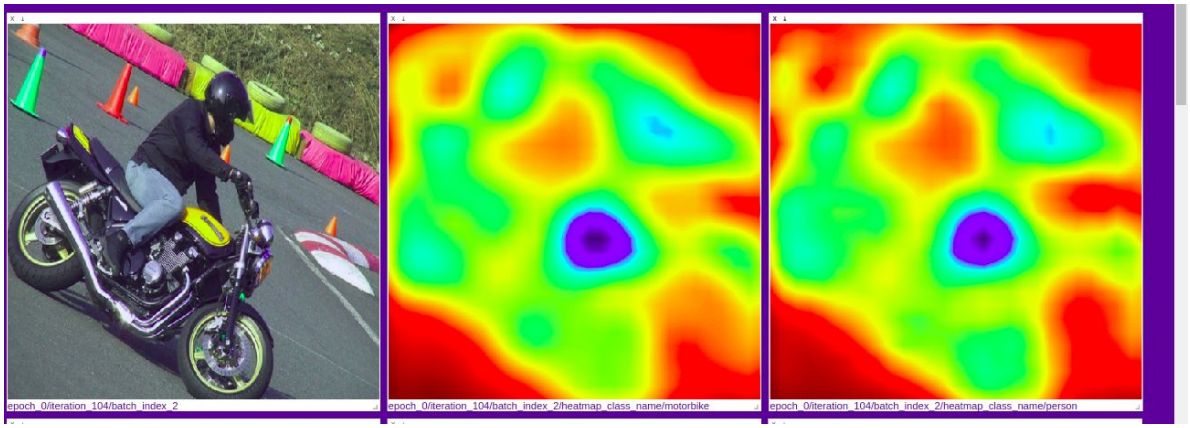


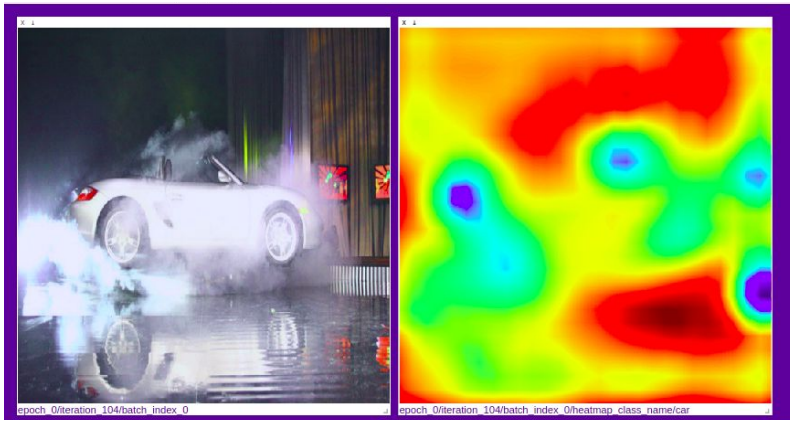



epoch_0/iteration_52/batch_index_1
epoch_0/iteration_52/batch_index_1/heatmap_class_name/car


epoch_0/iteration_0/batch_index_2
epoch_0/iteration_0/batch_index_2/heatmap_class_name/person

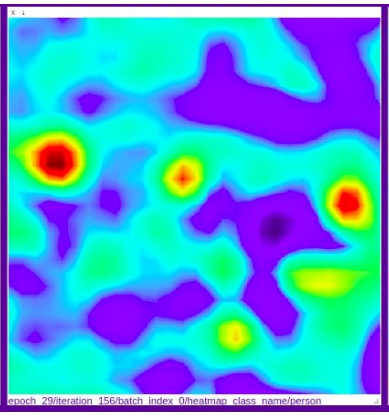


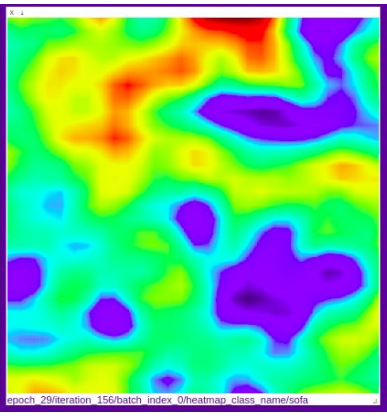




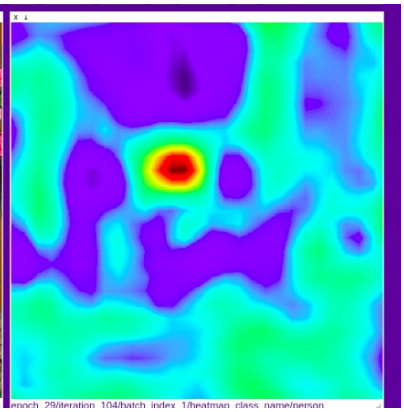
Visdom images and heatmap for last logged epoch


epoch_29/iteration_156/batch_index_0

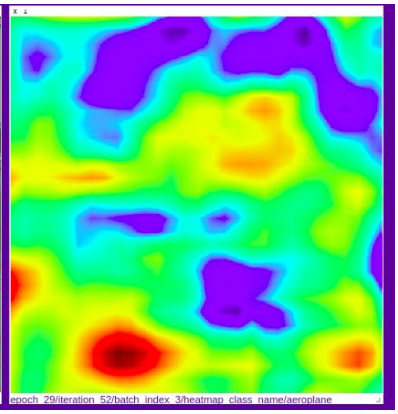

epoch_29/iteration_156/batch_index_0/heatmap_class_name/person


epoch_29/iteration_156/batch_index_0/heatmap_class_name/sofa

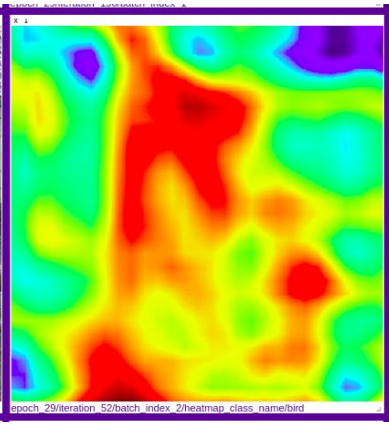

epoch_29/iteration_104/batch_index_1

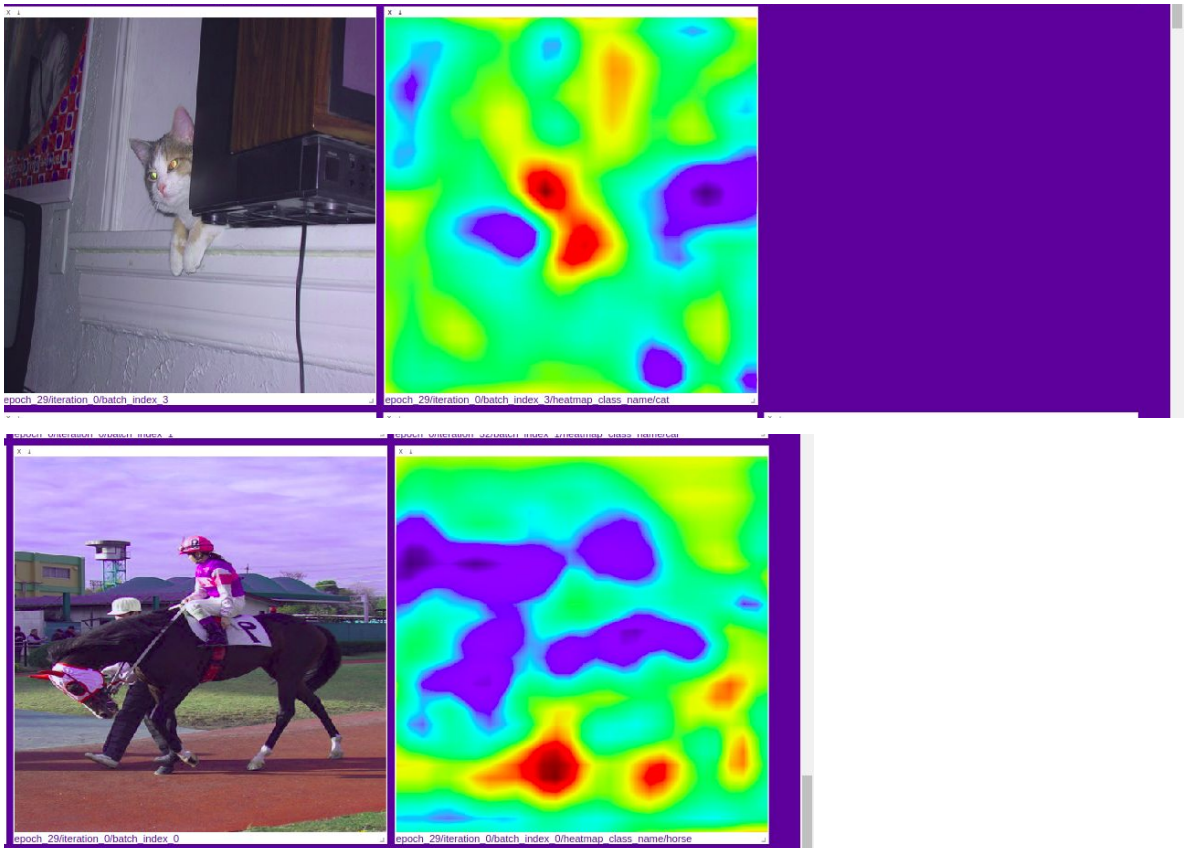

epoch_29/iteration_104/batch_index_1/heatmap_class_name/person


epoch_29/iteration_52/batch_index_3

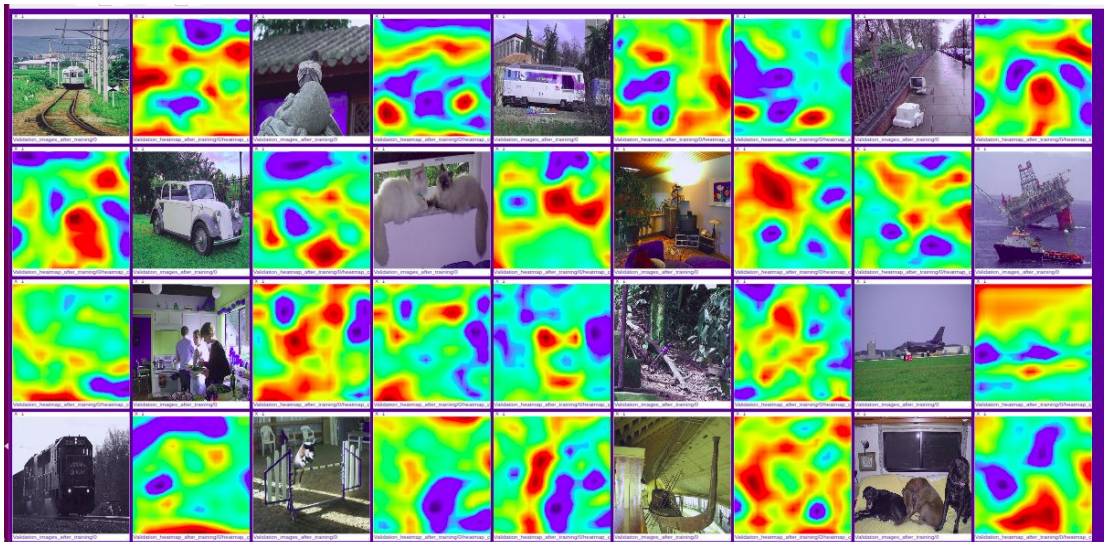

epoch_29/iteration_52/batch_index_3/heatmap_class_name/aeroplane

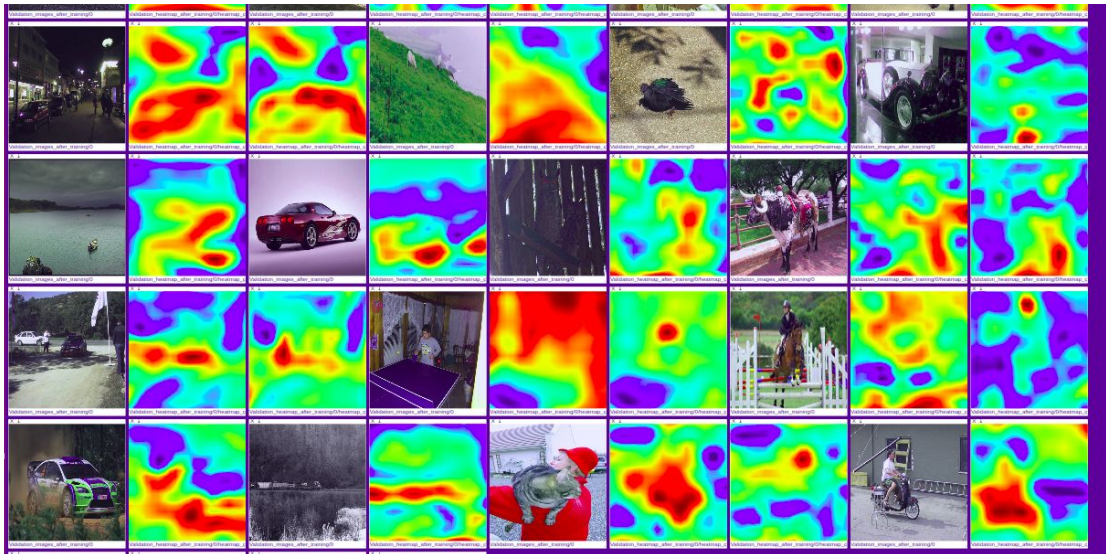

epoch_29/iteration_52/batch_index_2


epoch_29/iteration_52/batch_index_2/heatmap_class_name/bird



Visdom screenshot for 20 randomly chosen validation images and heat maps





At the end of training:

Training loss: 0.1104
Train metric 1: 0.7474
Train metric 2: 0.5185
Validation metric 1: 0.6345
Validation metric 2: 0.4486

Q 1.6 IN THE HEATMAP VISUALIZATIONS YOU OBSERVE THAT THERE ARE USUALLY PEAKS ON SALIENT FEATURES OF THE OBJECTS BUT NOT ON THE ENTIRE OBJECTS. HOW CAN YOU FIX THIS IN THE ARCHITECTURE OF THE MODEL?

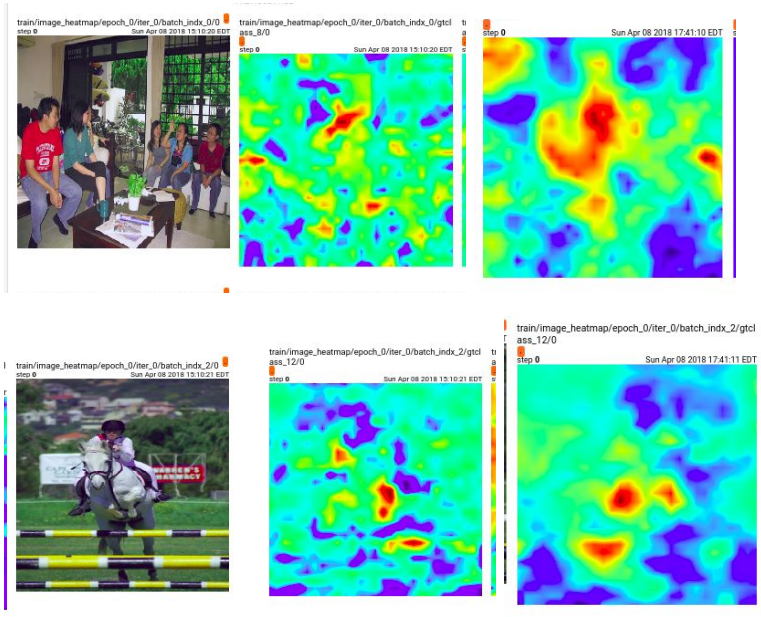
So I am doing two things for this

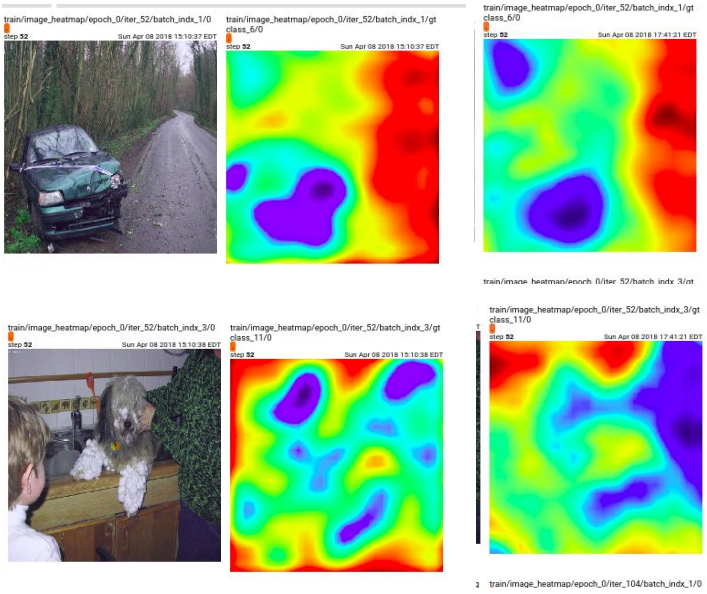
1) Apply dropout to randomly supress the activations just after the output of the classifier
2) Use a multiple average pooling after the dropout to get a more spreaded version of activations and then take an average of them

This is implemented as below:

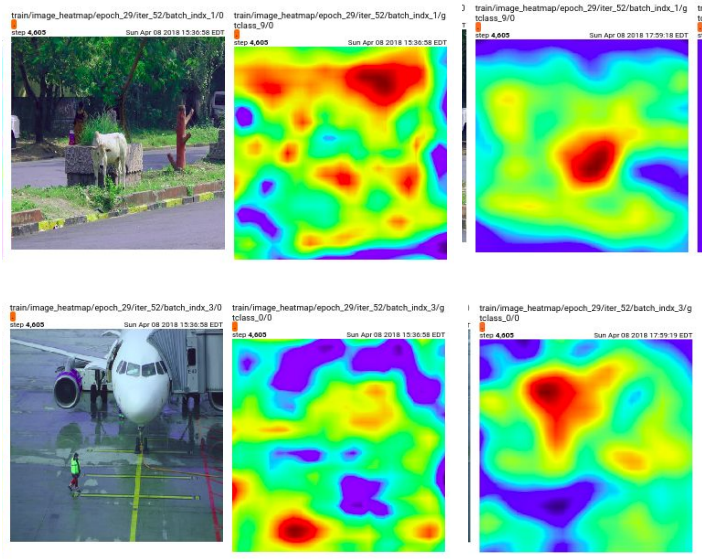
```
x = self.classifier(x)
x =  F.dropout2d(x, p=0.5)
x1 = F.avg_pool2d(x, kernel_size = 3 , stride  = 1, padding = 1 )
x2 = F.avg_pool2d(x, kernel_size = 5,  stride  = 1, padding = 2 )
x3 = F.avg_pool2d(x, kernel_size = 7,  stride  = 1, padding = 3 )
X   = (x + x1 + x2 + x3)/4.0
```

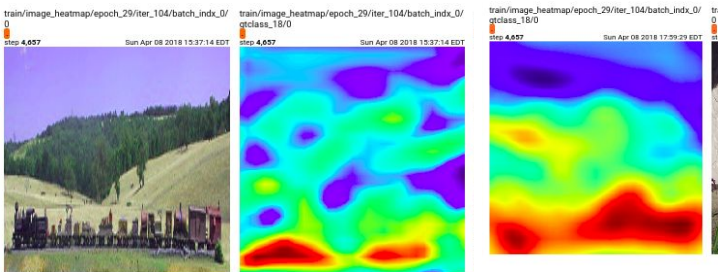
Screenshot of tensor board showing images and heat maps for the first logged epoch *for Q1.5 and Q1.6 side-by-side*.

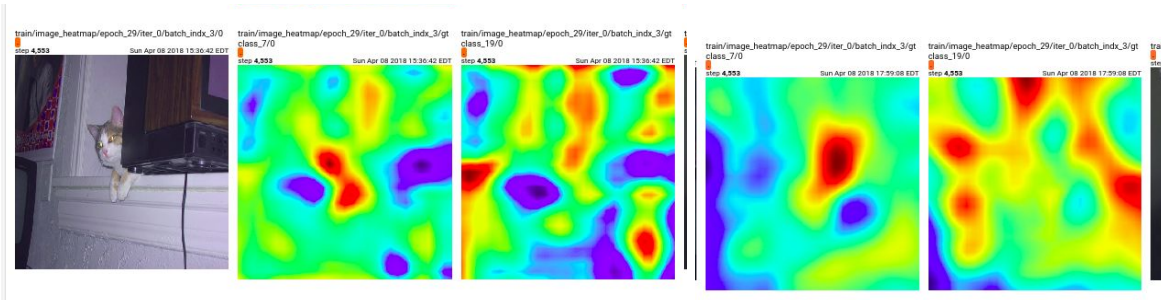




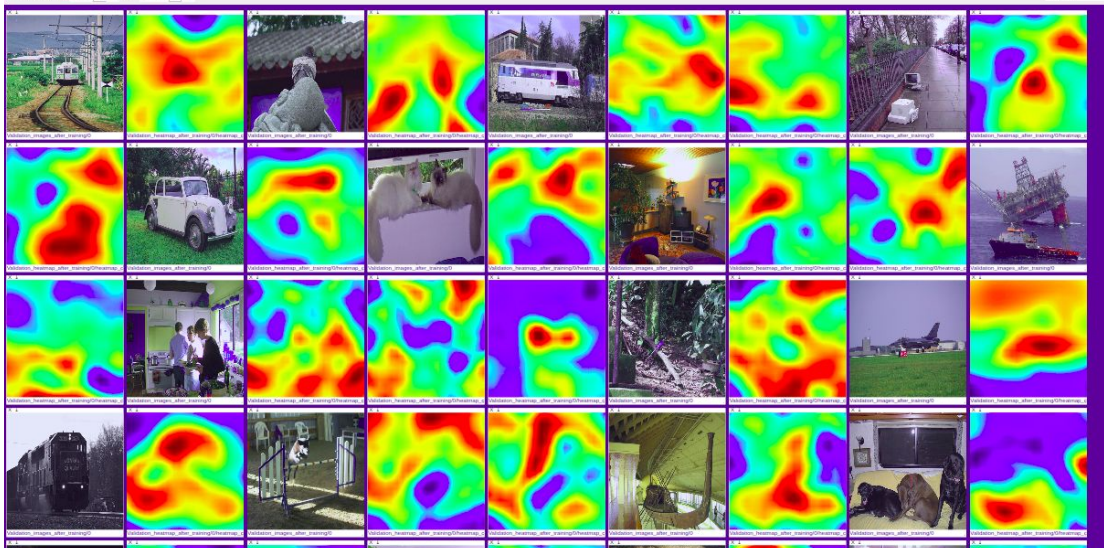
Screenshot of tensor board showing images and heat maps for the last logged epoch *for Q1.5 and Q1.6 side-by-side*.

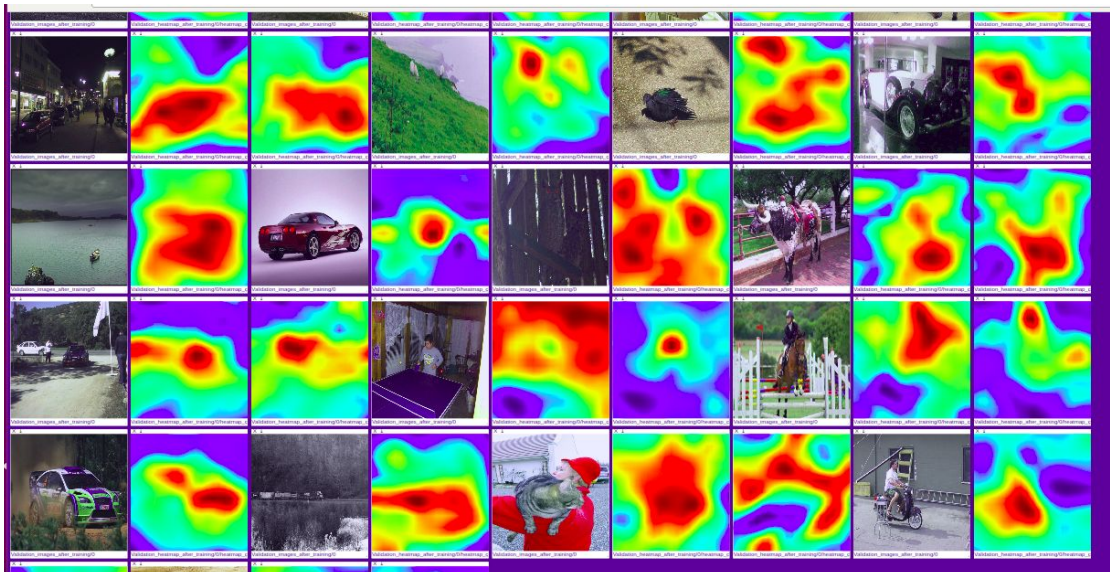






Visdom screenshot for 20 randomly chosen validation images (but same images as Q1.5) and heat maps





Report training loss, validation metric1, validation metric2 at the end of training

At the end of the training:
Training Loss : 0.3873
Validation metric 1 : 0.713
Validation metric 2 :0.505

Task 2

- visdom downloaded image of training loss vs iterations

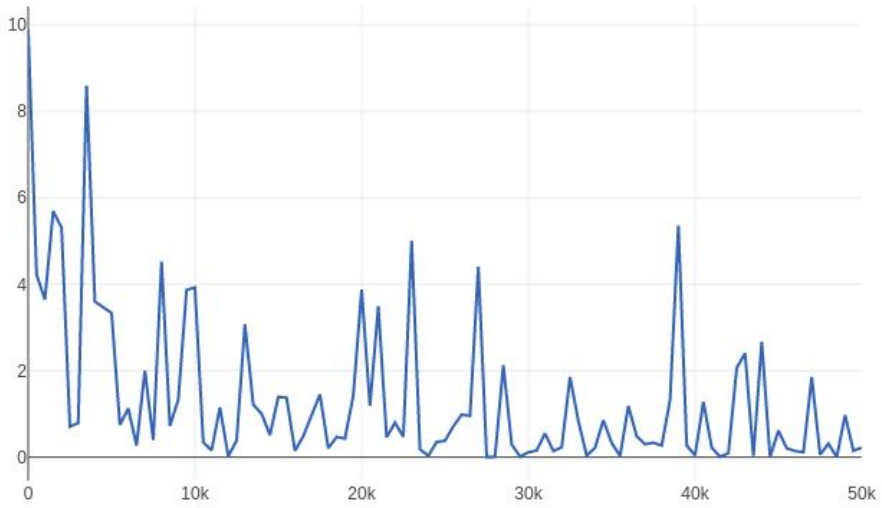
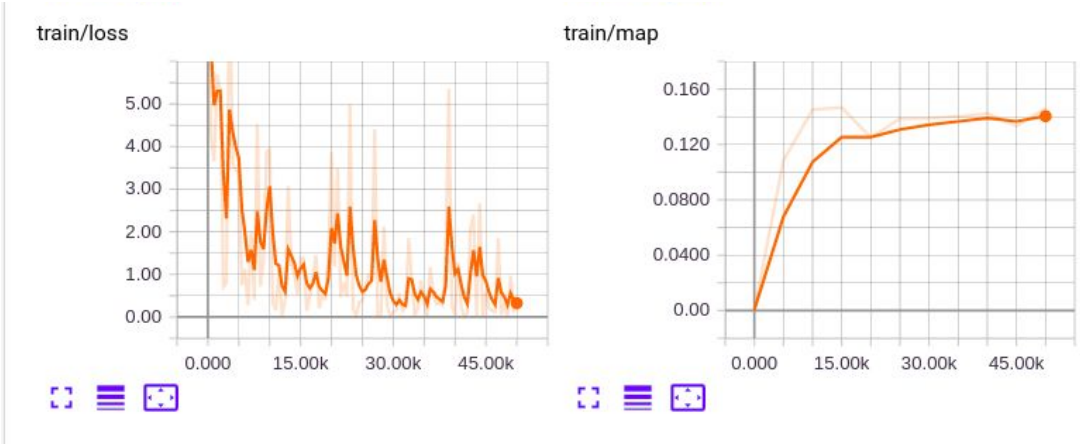


Figure: Training loss vs iterations

- tensor board screenshot of training loss vs iterations



- tensor board screenshot histogram of gradients of weights for conv1, conv2 and fc7

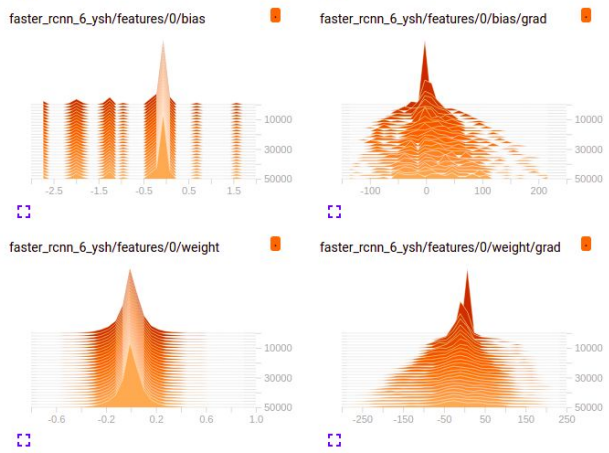


Figure: Conv1 plots

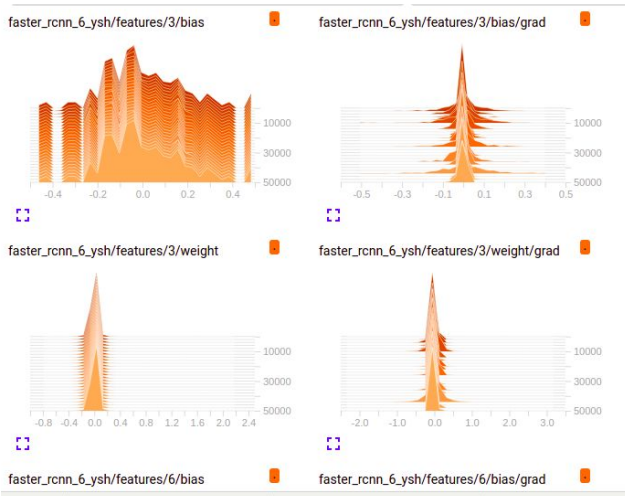


Figure: Conv2 plots

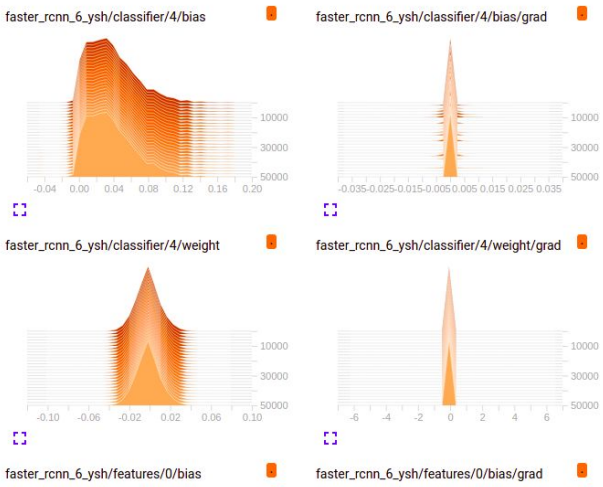


Figure: FC7 plots

- visdom downloaded image of test mAP vs iterations plot

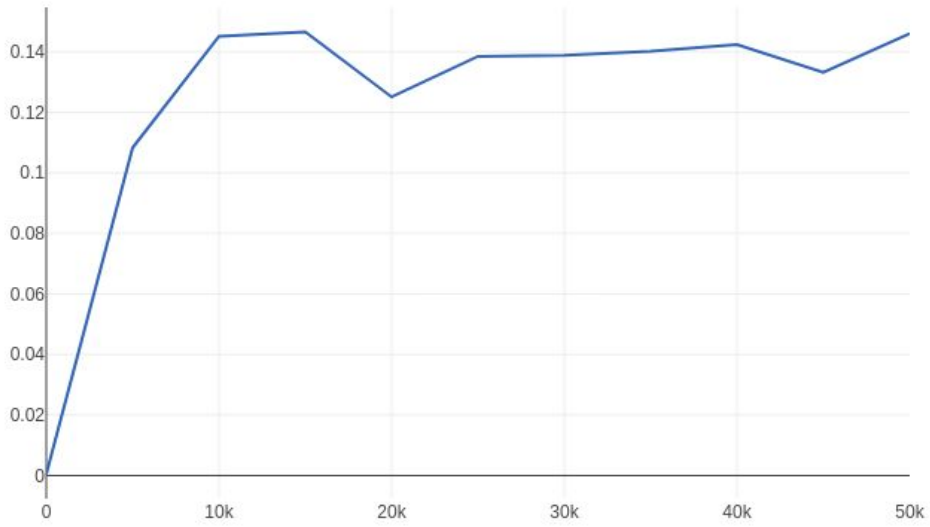
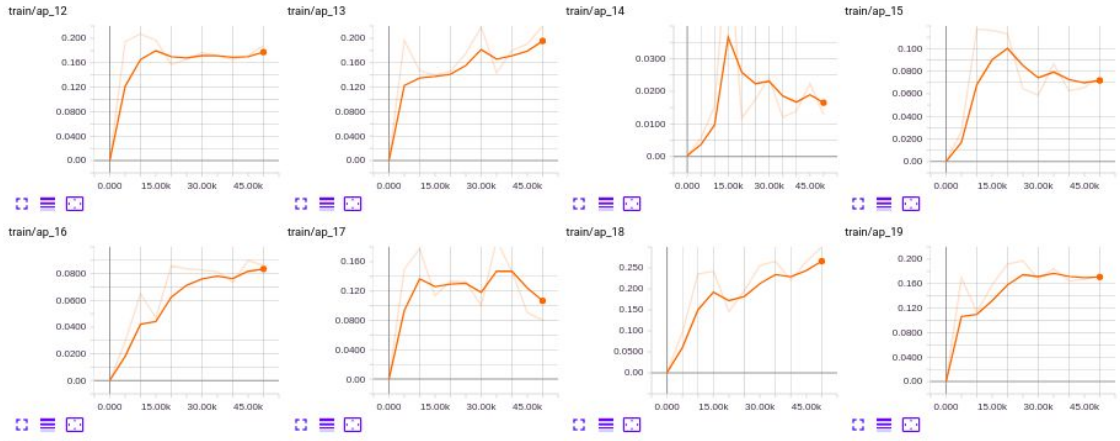


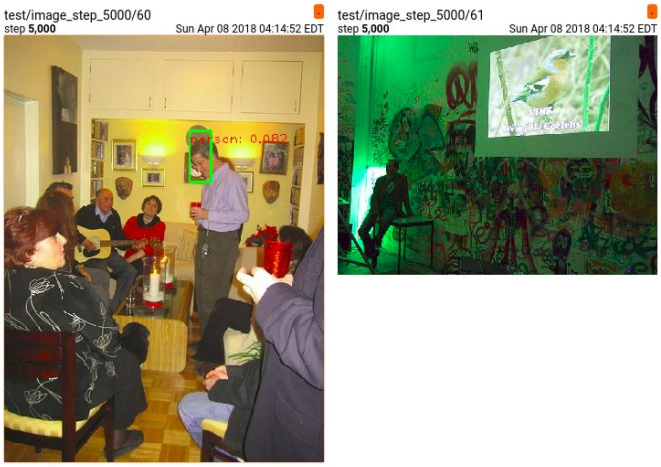
Figure: Test mAP vs iterations

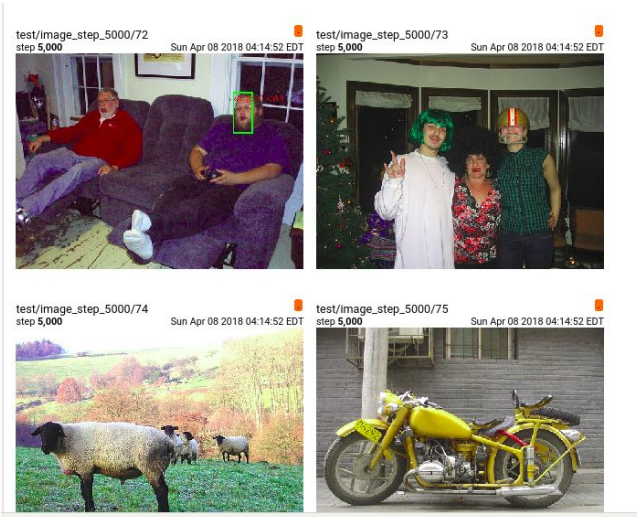
- tensorboard screenshot for class-wise APs vs iterations showing 3 or more classes


train/ap_0
train/ap_1
train/ap_2
train/ap_3
train/ap_4
train/ap_5
train/ap_6
train/ap_7
train/ap_8
train/ap_9
train/ap_10
train/ap_11

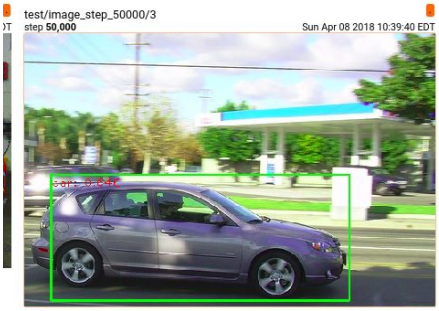

train/ap_12
train/ap_13
train/ap_14
train/ap_15
train/ap_16
train/ap_17
train/ap_18
train/ap_19

- tensor board screenshot of images with predicted boxes for the first logged iteration (5000)



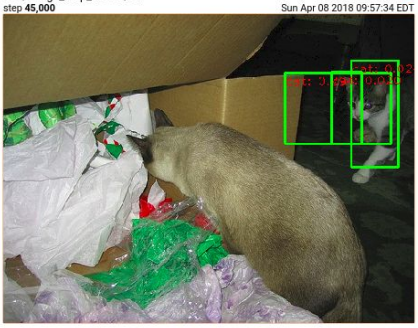


- tensor board screenshot of images with predicted boxes for the last logged iteration (50000 or 45000)

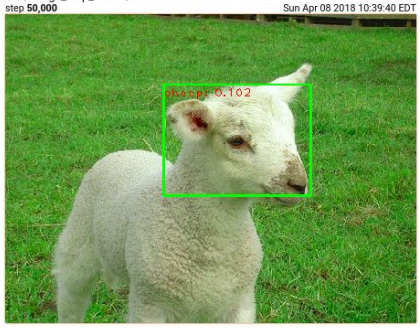


test/image_step_45000/90
step **45,000** Sun Apr 08 2018 09:57:34 EDT

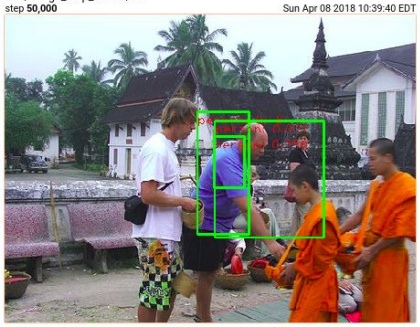
test/image_step_45000/91
step **45,000** Sun Apr 08 2018 09:57:34 EDT



test/image_step_50000/13
step **50,000** Sun Apr 08 2018 10:39:40 EDT



test/image_step_50000/24
step **50,000** Sun Apr 08 2018 10:39:40 EDT

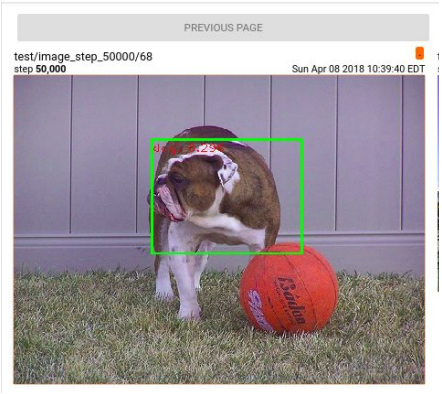


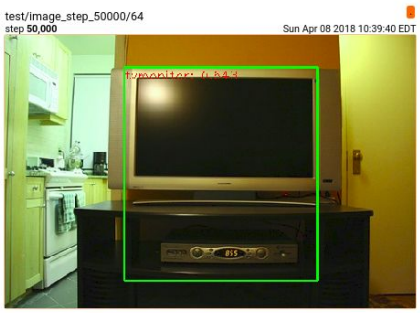
PREVIOUS PAGE

test/image_step_50000/32
step **50,000** Sun Apr 08 2018 10:39:40 EDT

test/image_step_50000/54
step **50,000** Sun Apr 08 2018 10:39:40 EDT

test/image_step_50000/55
step **50,000** Sun Apr 08 2018 10:39:40 EDT





- report final classwise APs on the test set and mAP on the test set

AP for aeroplane = 0.2485
AP for bicycle = 0.2866
AP for bird = 0.1879
AP for boat = 0.1146
AP for bottle = 0.0529
AP for bus = 0.2384
AP for car = 0.2324
AP for cat = 0.1684
AP for chair = 0.0195
AP for cow = 0.1182
AP for diningtable = 0.0223
AP for dog = 0.0962
AP for horse = 0.1883
AP for motorbike = 0.2202
AP for person = 0.0128
AP for pottedplant = 0.0749
AP for sheep = 0.0862
AP for sofa = 0.0802

AP for train = 0.3000
AP for tvmonitor = 0.1725
Mean AP = 0.1461