

Computational Vision

Sheet 1 — CSCI 4270

Homework 6: Discussion

1. **Neural Network Model:** For my neural network, I ran validation to select the final model by varying the learning rate and the number of epochs. When designing the network, I also had to make several decisions about the neural network which I didn't leave to validation. For example, when developing the loss function, the loss was returning a nan value (which I fixed by adjusting the batch size to be larger). In addition, during learning the program would often misclassify the car, bike, and motorcycle classes as nothing or people classes (since there are many more training points for the nothing and people class). To address this, I added class weights to the cross entropy loss function used for classification to address this (I adjusted the values based on observation, but generally speaking there is a higher weight assigned to classes with fewer data points). Finally, the model had a very low accuracy due to misplacement of the bounding boxes (even though classification accuracy was reasonably high). To address this, I increased lambda to 2 in order to double the weight placed on the bounding box (I tried higher lambdas, but at that point the classification was negatively impacted).

2. **Training Output:** Where the first confusion matrix is on the validation data and the second is on the training data.

Epoch 15

loss: 5.299726 [0/23649]

loss: 6.381811 [8000/23649]

loss: 9.730725 [16000/23649]

Validation Loss:

Accuracy: 63.2%, Avg loss: 6.592547

Confusion Matrix:

```
[[1599  145  339  147  458]
 [  16  155   0   12   2]
 [  12   3  504   7   11]
 [   4   13   6  138  12]
 [ 184  40   89  29 1789]]
```

Selected Model:

Epochs: 14

Learning Rate: 0.001

lambda: 2

Training Data Results:

Accuracy: 66.9%, Avg loss: 6.542037

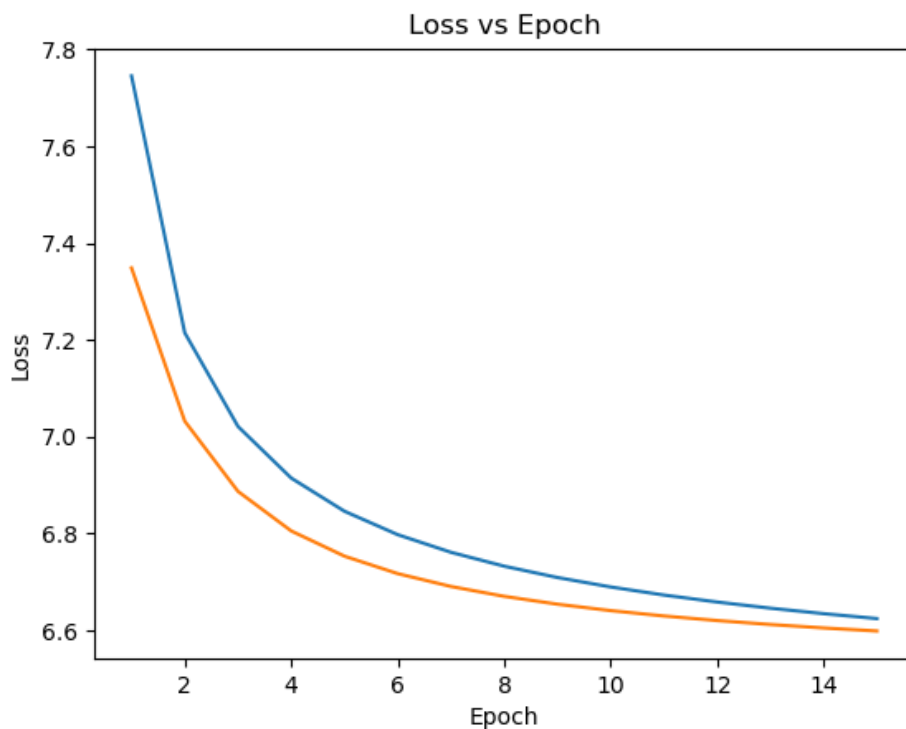
Confusion Matrix:

```
[[6261  476  776  446 2385]
 [  32  598   37   27   59]
 [  89   0 1839   25  111]
 [  18   33   48  584   41]
 [ 505  181  313  191 8574]]
```

3. Testing Output:

```
Image 1 AP is: 1.000
Image 16 AP is: 0.000
Image 28 AP is: 0.000
Image 30 AP is: 0.100
Image 39 AP is: 0.250
Image 75 AP is: 0.000
Image 78 AP is: 0.000
Image 88 AP is: 0.000
Image 93 AP is: 1.000
Image 146 AP is: 1.000
Average IOU: 0.67
Final mean average precision: 0.220
Success!!
```

See losses below (where blue is the validation loss and orange is the training loss):



4. **Summary:** Overall, I am happy with my results, although there are definitely some areas for improvement. To start off, I forgot to include output of the percentage of true detections that are true positive and percentage of ground truths found (I didn't notice until after my training and didn't have time to train again). Also, I do believe that with training over more epochs there could definitely be improved results. Looking at the bounding boxes and classifications from the selected images of the test dataset, there are several cases where the program correctly classifies the region, but it seems the bounding box simply has just under 50% overlap with the ground truth region. Consider the images below, where a bounding box was drawn around a human, but was marked as incorrect because the ground truth box was around another human. Also, the image of the car illustrates how the program was able to draw the bounding box close to the car (but not close enough).

