50.039 – Theory and Practice of Deep Learning

Alex

Week 04: Data augmentation

Due: week5 Wednesday, 9pm

1 In class Problem 1 - Task: Test the performance of a pretrained net – simple crop.

Take the 2500 Imagenet val images, unpack them. the labels are in `ILSVRC2012_bbox_val_v3.tgz`. `getimagenetclasses.py` has example routines to get the label for one image.

- write a dataloader for this dataset. For mxnet you can transform the image label pairs into an ImageFolderDataset instance. For pytorch writing a dataloader directly should work, too.
  - Suggestion: do not load all images in the __init__ method of the dataset class. That does not scale well if you have 500000 images :) . load the filenames and the labels instead into a list or the like. load an image in `getitem` of your Dataset-derived class!
- rescale the images so that the smaller side is $s = 224$ and perform a center crop of 224×224
- initialize the weights so that you load weights from the so-called model zoo. pytorch model zoo or gluon model zoo.
- compare performance to the case when you do not subtract the mean and normalize the subpixels. If too slow, use only the first 250 or first 500 images.
- hint: use a net with little parameters, avoid VGG, Alexnet.

Learning goals: dataloader for a custom dataset, prediction with a pretrained neural network.
Another take away: prediction with a neural net runs fairly fast. it is not a must to use GPU for prediction.

2 In class Problem 2 - Task: Test the performance of a pretrained net – five crop (see below if you cannot use a five crop transform)

- Take the code from Problem 2, now rescale smaller side to 280 (or 256) and implement a five crop as shown above.
- For every of the five crops, compute the prediction.
- Compute for every image the average probability vector of the five predictions over the five crop. If too slow, use only the first 250 or first 500 images.
- Compare performance of the average over five crops to using just the centercrop.
- Consider the tencrop: this is five crop with vertical mirroring for five of the ten. For what datasets mirroring is a bad augmentation idea?

Leaning goal: You should observe improved performance, without training anything!

3 In class Problem 3 - Task: Different input size of the neural network.

Try to classify with two different pretrained neural network architectures with an input size of 330×330 .
If it does not work out of the box, then take the network, and create a derived class from it. Modify in the derived class the average/max pooling so that it works with an input size if 330×330 . Classify with 5-crop of size 330 as in the class problem 2.

4 Homework

Finish the above In class Problems 1,2,3 and submit them as homework. Submit the code and report prediction accuracies (simple classification error here does the job).