

# Computer Vision HW 4

YI NIAN

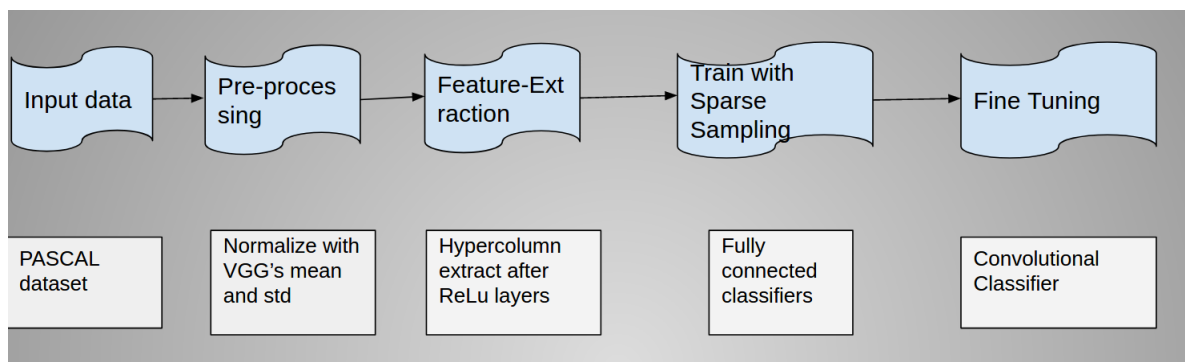
University of Chicago

3-10

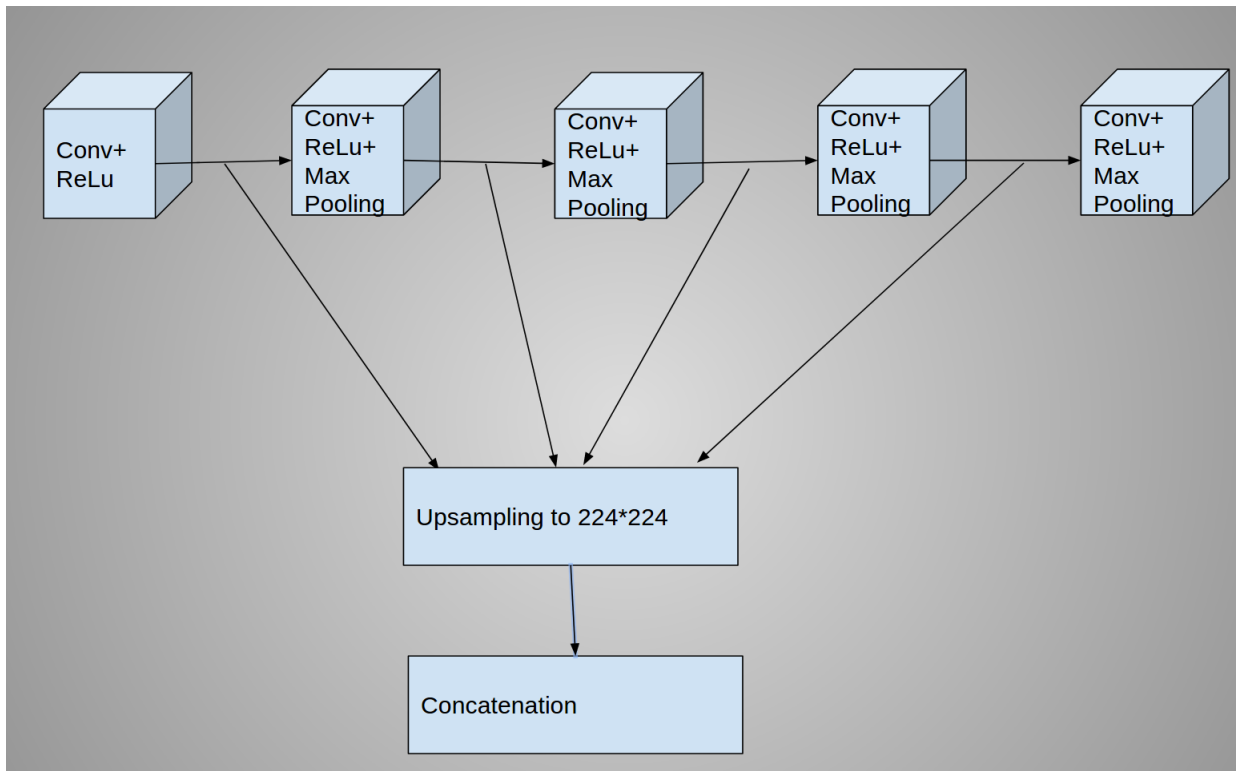
## 1 Submitted File

1. The writeup: hw4.pdf
2. a zip file including everything:
  - sampling.py: Sample 3 random pixel per class from each image
  - stats.py: To summarize the frequency of each class and use as weight
  - train\_cls.py: model for fully connect classifier
  - train\_seg.py: model for convolution(fine tune) model
  - zoomout.py: Extract zoomout features
  - utils.py: Some utility function for accuracy
  - Nets folder: includes the architecture for both classifier
  - losses folder: includes the loss function(cross entropy)
  - features folder: includes the extracted zoomout features and statistics from stats.py

## 2 Pipeline



### 3 Zoom-out-feature with pretrained VGG11

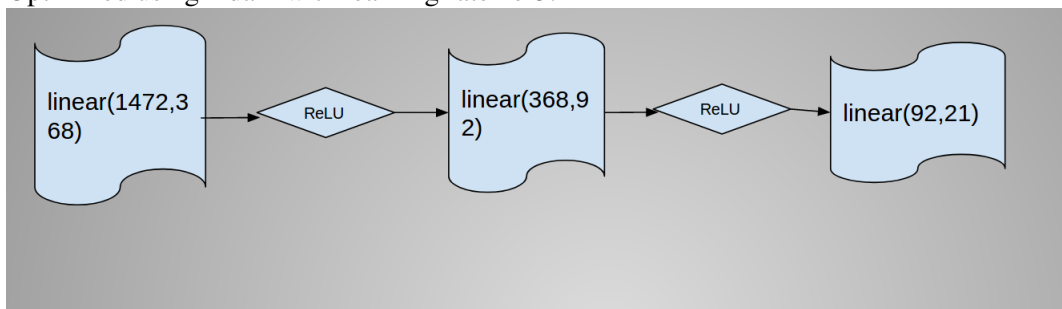


### 4 Sparse sampling

3 random pixels from each class per image are sampled and transformed to hypercolumns and send to fully connected classifier.

### 5 Simple Fully connected classifier

Optimized using Adam with learning rate  $1e-3$ .



## 6 Fine Tuning

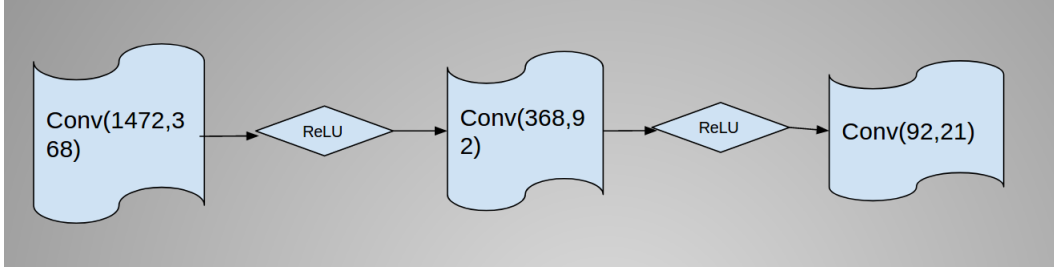
### 1. Loss Function

Cross Entropy Loss with weight =  $\frac{1}{\text{frequency of each class}}$

### 2. Optimizer

Adam with learning rate  $5 * 10^{-4}$  for first 5 epoch.  $10^{-4}$  for later 5 epoch.

Same Parameters from linear layers are used in 1 by 1 convection layers.



## 7 model accuracy

A test accuracy after 5 epoch of training :

```
UserWarning: Default upsampling behavior when mode=bilinear is changed to align_corners=False since
0.4.0. Please specify align_corners=True if the old behavior is desired. See the documentation of
nn.Upsample for details.
"See the documentation of nn.Upsample for details.".format(mode))
Accuracy: 43.62397294580358
Accuracy Class: 19.030923917496274
Mean IU: 7.05933208514151
FWAV Accuracy: 39.13983957968603
```

A test accuracy after 10 epoch of training :

```
nn.Upsample for details.
"See the documentation of nn.Upsample for details.".format(mode))
Accuracy: 64.65377000632033
Accuracy Class: 16.23438257349345
Mean IU: 8.660436837830353
FWAV Accuracy: 55.78138944057457
```

## 8 example output

1. It is more easy to classify an object if an image just includes a single object: for example, a cat or a person.
2. It is much harder to classify multiple object perfectly when they appear in an single image: for example, a person on a motor.

