

## HW 8: Deep Learning / Computer Vision

Please remember the following policies:

- Submissions should be made electronically via the Canvas. Please ensure that your solutions for both the written and/or programming parts are present and zipped into a single file.
- Solutions may be handwritten or typeset. For the former, please ensure handwriting is legible.
- You are welcome to discuss the programming questions (but *not* the written questions) with other students in the class. However, you must understand and write all code yourself. Also, you must list all students (if any) with whom you discussed your solutions to the programming questions.

### V1. Calculating CNN kernel operation (6 points).

You are given the following kernel  $W$  of a convolutional layer:

$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

$W$

- (a) Apply the kernel  $W$  with a stride of two and one pixel of zero padding to the input  $X$  below. Fill in the missing values (blank box) of the output  $Z = X * W$ .

3	1	6	8	0
2	3	2	7	-6
-8	0	6	-5	-8
1	2	5	-2	-4
8	7	-2	-1	7

$X$

1		1
0		
	1	0

$Z$

- (b) Is this convolutional layer the same as a  $3 \times 3$  average pooling layer? Do they compute the same function? Do they behave similarly during train and test?

V2. Training from scratch (6 points).

hw8(1): In this question, you must create a neural network architecture that can be used to classify images in the CIFAR-10 dataset. Images are  $32 \times 32$  resolution and belong to one of six categories. There are 5000 images in each category in the training set.

(a) Fill in the layers array in Q1.m to create a neural network with the following structure:

Type	Kernel Size	Stride	Num channels	Padding	Num units
conv	5		16	2	
maxPool	3	2			
relu					
conv	5		32	2	
batchNorm					
avgPool	3	2			
relu					
conv	5		64	2	
batchNorm					
avgPool	3	2			
relu					
FC					64
relu					
FC					6

- (b) What accuracy do you get when training this network? Is it the same every time you train? Why/why not?
- (c) By how much does accuracy change if you remove the last convolutional layer from the network?
- (d) What is the effect of changing the average pooling layers to max pooling layers?

V3. Transfer learning (6 points).

hw8(2): A common strategy for improving image classification accuracy is to use pretraining/finetuning. In this question, you will copy the weights from a pretrained AlexNet and then finetune them to solve your CIFAR-10 problem. The code in Q2.m loads the AlexNet structure and pretrained weights and removes the last three layers of AlexNet. Your job is to append the five layers given below and illustrated in Figure1. When you run this case, it should train the whole neural net work with a focus on the last fully connected layers. You should end up with approximately 85% accuracy on the test set. Hint: if this code gets an error at ALEXNET, you may need to install the Deep Learning Toolbox Model for AlexNet Network: <https://www.mathworks.com/matlabcentral/fileexchange/59133-deep-learning-toolbox-model-for-alexnet-network>.

Type	Kernel Size	Stride	Num channels	Padding	Num units
FC					64
relu					
FC					4
softmaxLayer					
classificationLayer					

- (a) List three possible reasons why this network achieves higher accuracy than we attained in the last question.
- (b) Why does the softmax layer need to have four outputs?
- (c) What happens when you remove the first two layers that you added above (FC64 and ReLU)?

V4. Play with a pretrained model (2 points).

hw8(2): Let's have fun with the neural network. Run the code in Q3.m that uses a pretrained AlexNet to label objects shown through webcam. Notice that the caption sometimes swings between two or more categories?

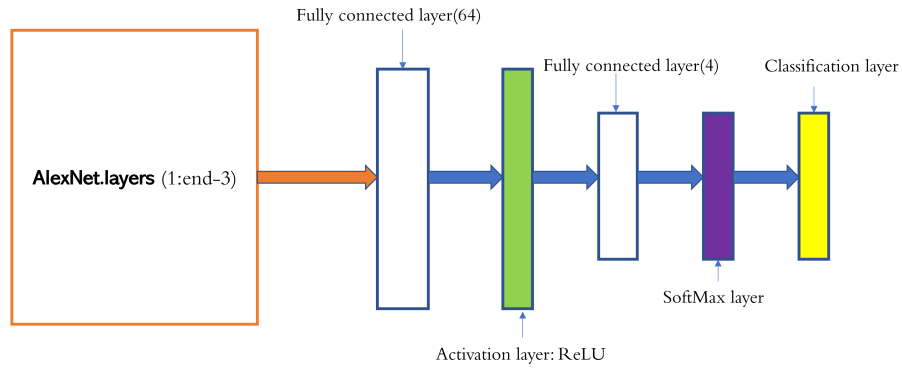


Figure 1: Model structure for the five layers that get added to AlexNet (V3).

Why would it happen? (Use Ctrl+C to stop running code and use "clear" command on command window to close camera)

Hint: if this code gets an error at CAMERA, you may need to install the MATLAB Support Package for USB Webcams: <https://www.mathworks.com/help/supportpkg/usbwebcams/ug/installing-the-webcams-support-package.html>.

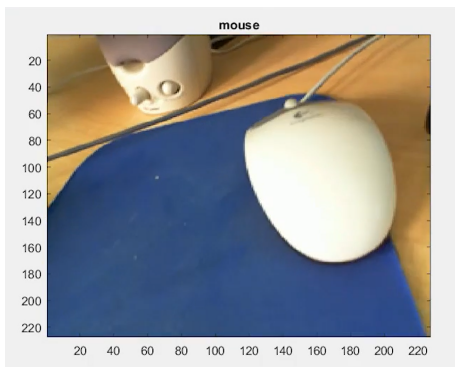


Figure 2: mouse

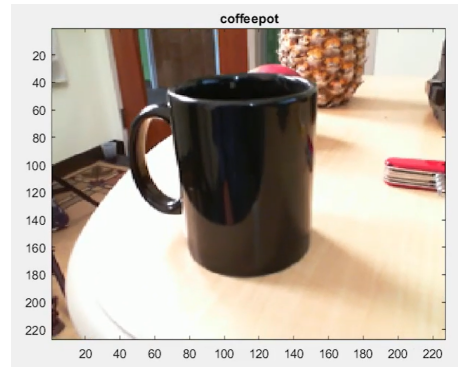


Figure 3: Coffee mug