

1.

(a) Biases, filters, and size of filters

(b)  $(3(5*5*3))+3 = 228 \rightarrow 228*3 = 684$

2. To recognize the pattern shown in figure 1b, I would scan the image figure 2 in the convolution layer checking each given 6x6 for a pattern that heavily overlaps with both filters in figure 1a.



3.

In this image, the 2nd and 3rd images can be obtained by adding 50 or 100 pixels respectively to the first image.

4.

- (i) In the article *ImageNet Classification with Deep Convolutional Neural Networks*, 3 students at the university of Toronto train a model on the ImageNet dataset containing “over 15 million labeled high-resolution images belonging to roughly 22,000 categories.” Through the use of multiple CNNs and a method called “dropout” used to decrease overfitting, these students were able to achieve “top-1 and top-5 error rates of 37.5% and 17.0%” respectively.
- (ii) There may be a large fluctuation in loss because the training method used was stochastic gradient descent. Another reason could be because of the data augmentation method used. Dropout sets neurons to 0 with a probability of .5 meaning they don’t participate in forward or backward propagation. This means that the results of a specific epoch may not be accurate.