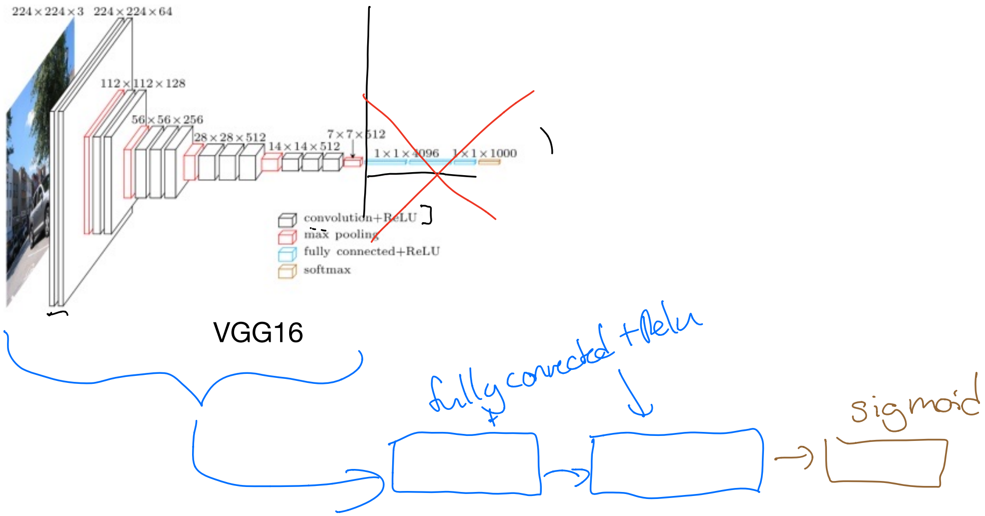
Reminder about transfer learning

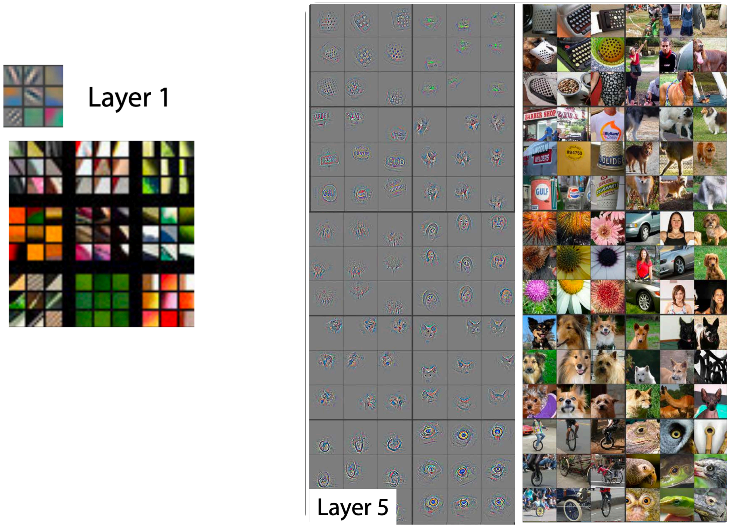
Cat vs. dog

Image from https://towardsdatascience.com/step-by-step-vgg16-implementation-in-keras-for-beginners-a833c686ae6c

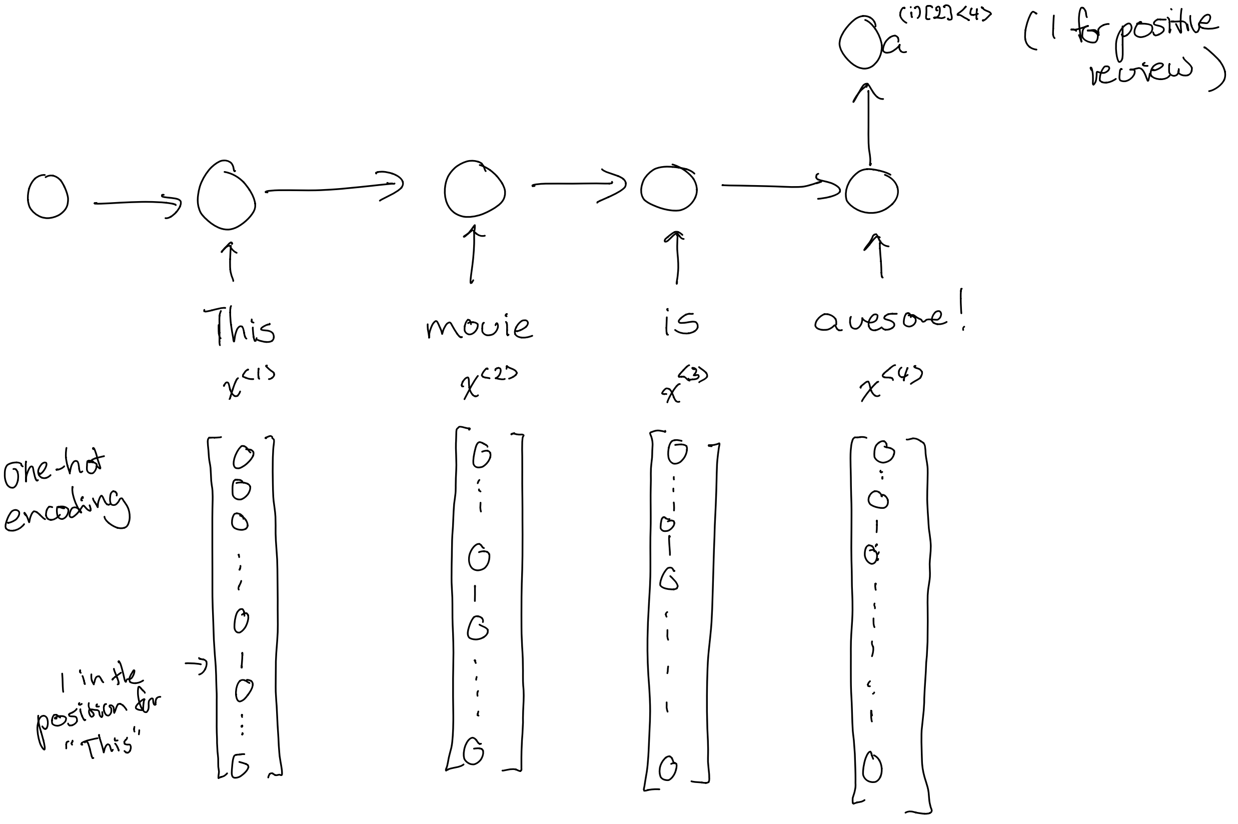


output of convolutional base:
high level representation of content of image

Images from 'Visualizing and understanding convolutional networks' by Zeiler and Fergus



Motivating Example: Sentiment classification for movie reviews



2 limitations to the one-hot encoding:

1) Input vectors very large (# of words in vocabulary)
(e.g. length 30,000 if 30,000 words in vocabulary
↳ lots of parameters to estimate.

2) Compare these two posssible sentences:

a. This movie is <u>awesome</u>

b. This movie is <u>great</u>

awesome and great are synonyms, but one-hot encoding doesn't take advantage of that.

Consider the inner product of their one-hot encodings:

$$O_{awesome} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad O_{great} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

so

$$O_{awesome} \cdot O_{great} = 0\cdot0 + 0\cdot0 + \cdots + 1\cdot0 + 0\cdot0 + \cdots + 0\cdot1 + 0\cdot0 + \cdots + 0\cdot0$$

$$= 0$$

The one-hot encodings of "awesome" and "great" are orthogonal.

↳ no similarity between these words according to one-hot encoding.

(since $\underbrace{v\cdot w}_{0} = \underbrace{\|v\|}_{1}\cdot\underbrace{\|w\|}_{1}\cdot\underbrace{\cos(\theta)}_{0}$)

Want: representation of these words $e_{awesome}$, e_{great} so that $e_{awesome}\cdot e_{great}$ is not 0.