8.7.2: Overfeat on cats and dogs feature layer

nathan.george **Aug 1**

In 8.7.2: Overfeat on cats and dogs
"We typically use either the -3 or -2 (with 4096 nodes) as our feature extraction layers, since higher level layers in the network learn more complex and discriminating features."

To be clear, we don't want to use a higher level layer because the feature dim would be too large? If the feature is more discriminating, isn't that a good thing?

Layer	1	2	3	4	5	6	7	Output 8
Stage	conv + max	conv + max	conv	conv	conv + max	full	full	full
# channels	96	256	512	1024	1024	3072	4096	1000
Filter size	11x11	5x5	3x3	3x3	3x3	-	-	-
Conv. stride	4x4	1x1	1x1	1x1	1x1	-	-	-
Pooling size	2x2	2x2	-	-	2x2	-	-	-
Pooling stride	2x2	2x2	-	-	2x2	-	-	-
Zero-Padding size	-	-	1x1x1x1	1x1x1x1	1x1x1x1	-	-	-
Spatial input size	231x231	24x24	12x12	12x12	12x12	6x6	1x1	1x1

Adrian Chief PyImageSearcher **Aug 2**

The feature dimensionality isn't too much of a big deal here. 4096-dim isn't *that* big in the grand scheme of things. 4096-dim is actually pretty small.

And yes, discriminating features *are* a good thing. But you can also run the risk of overfitting and your results not being able to generalize to images outside your training set.