Machine Learing: Convolutional Neural Networks

CPSC 501: Advanced Programming Techniques
Winter 2025

Jonathan Hudson, Ph.D
Assistant Professor (Teaching)
Department of Computer Science
University of Calgary

Friday, February 21, 2025

UNIVERSITY OF CALGARY

ImageNet

Deep Learning/ Convolutional Neural Networks

- Classify an image into 1000 possible classes:
  e.g. Abyssinian cat, Bulldog, French Terrier, Cormorant, Chickadee, red fox, banj o, barbell, hourglass, knot, maze, viaduct, etc.

→

cat, tabby cat (0.71)
Egyptian cat (0.22)
red fox (0.11)
.....

The Data: ILSVRC

- Imagenet Large Scale Visual Recognition Challenge (ILSVRC): Annual Competition (2009-published dataset)

1000 Categories

~1000 training images per Category

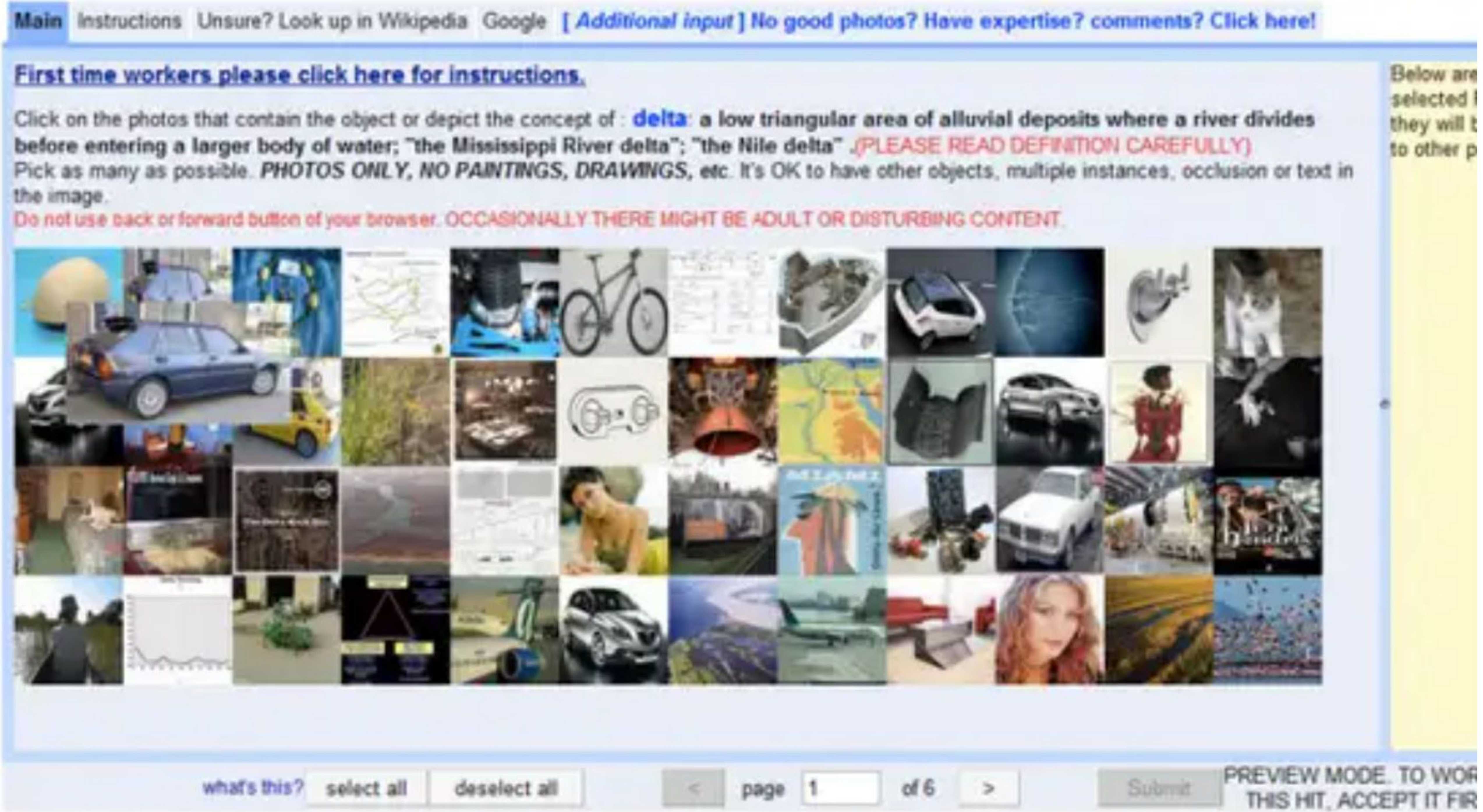
~1 million images in total for training

~50k images for validation

Only images released for the test set but no annotations,
evaluation is performed centrally by the organizers (max 2 per week)

https://arstechnica.com/ai/2024/11/how-a-stubborn-computer-scientist-accidentally-launched-the-deep-learning-boom/

Prof. Fei-Fei Li, Princeton

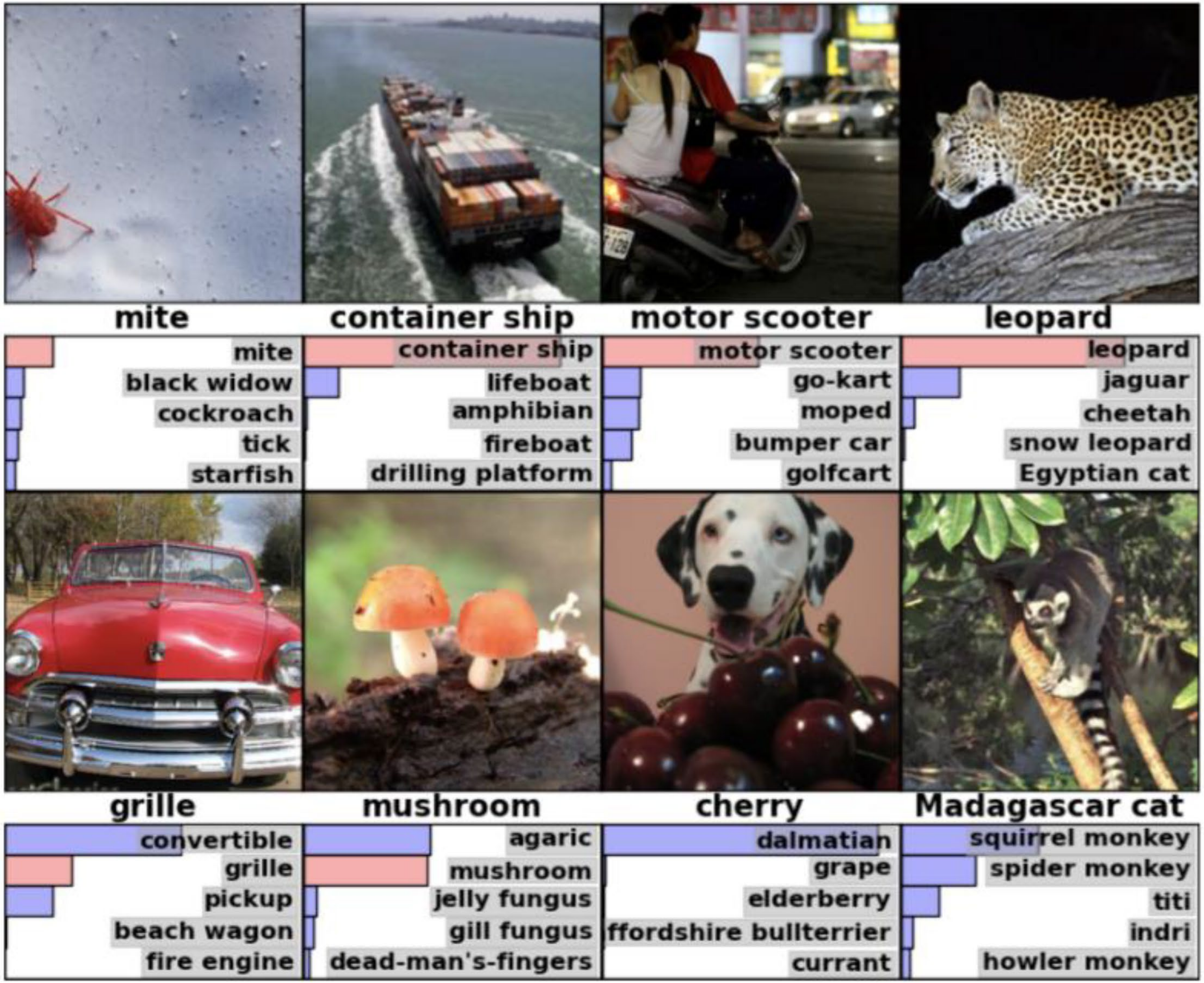

Main
Instructions
Unsure? Look up in Wikipedia
Google
[*Additional input*] No good photos? Have expertise? comments? Click here!

First time workers please click here for instructions.

Click on the photos that contain the object or depict the concept of : **delta**: **a low triangular area of alluvial deposits where a river divides before entering a larger body of water; "the Mississippi River delta"; "the Nile delta"** .(PLEASE READ DEFINITION CAREFULLY)
Pick as many as possible. ***PHOTOS ONLY, NO PAINTINGS, DRAWINGS, etc***. It's OK to have other objects, multiple instances, occlusion or text in the image.
Do not use back or forward button of your browser. OCCASIONALLY THERE MIGHT BE ADULT OR DISTURBING CONTENT.

Below are
selected
they will
to other p

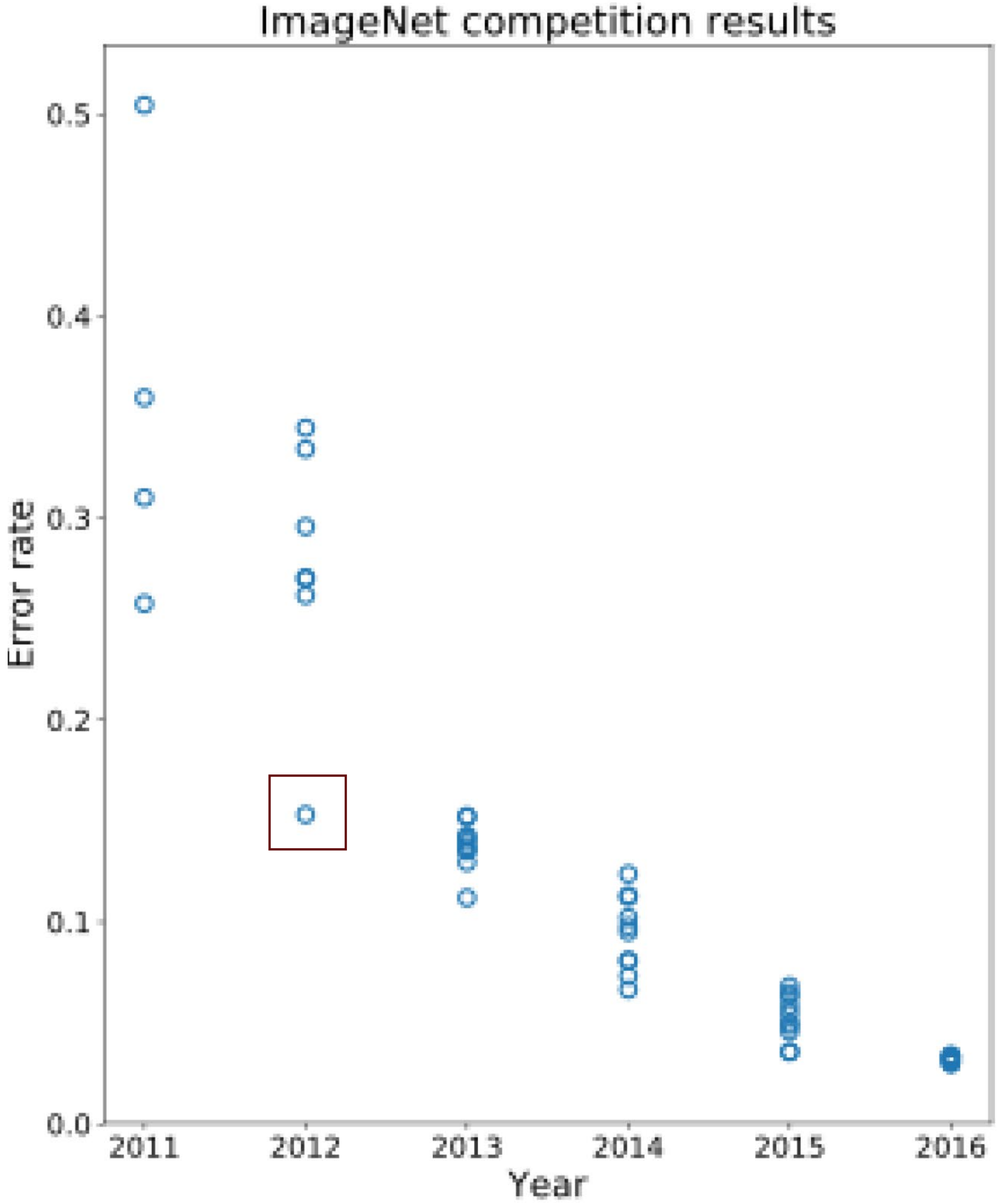
PREVIEW MODE. TO WOR
THIS HIT, ACCEPT IT FIR


mite
mite
black widow
cockroach
tick
starfish
container ship
container ship
lifeboat
amphibian
fireboat
drilling platform
motor scooter
motor scooter
go-kart
moped
bumper car
golfcart
leopard
leopard
jaguar
cheetah
snow leopard
Egyptian cat
grille
convertible
grille
pickup
beach wagon
fire engine
mushroom
agaric
mushroom
jelly fungus
gill fungus
dead-man's-fingers
cherry
dalmatian
grape
elderberry
ffordshire bullterrier
currant
Madagascar cat
squirrel monkey
spider monkey
titi
indri
howler monkey

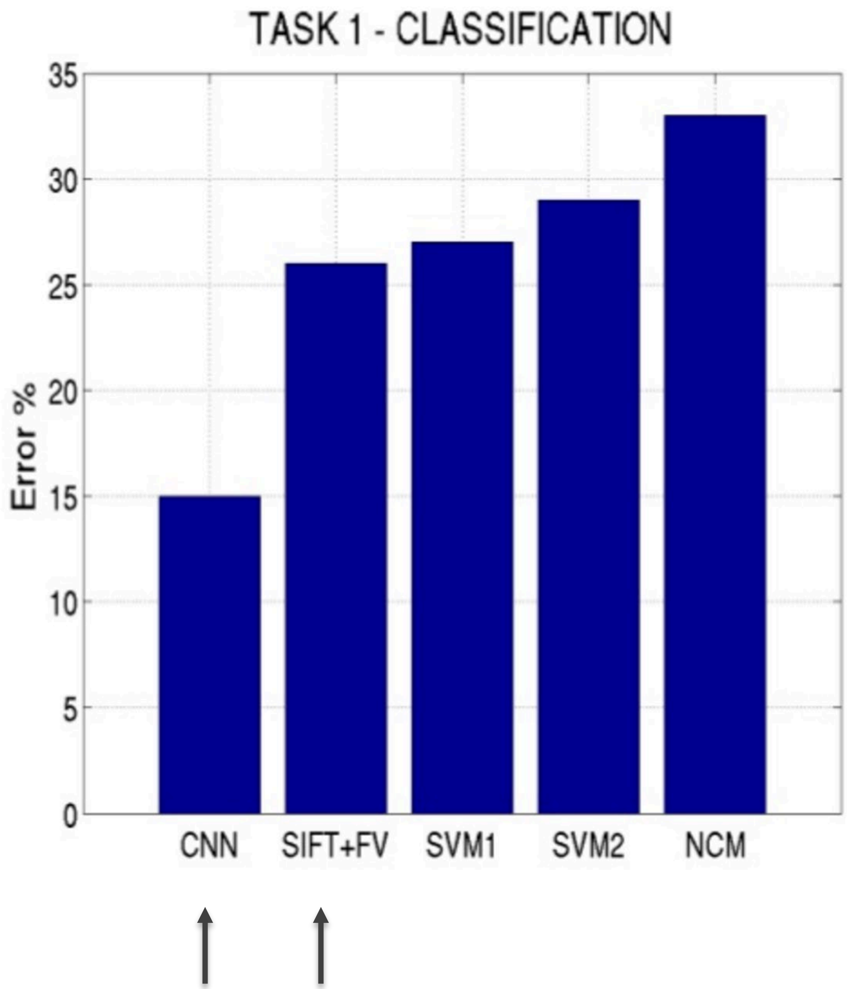
Results

- The ImageNet dataset was the basis for the ILSVRC, an annual AI competition that ran from 2010-2016.

- In 2012, a neural net calls **AlexNet** created by researchers at the University of Toronto won by a large margin.

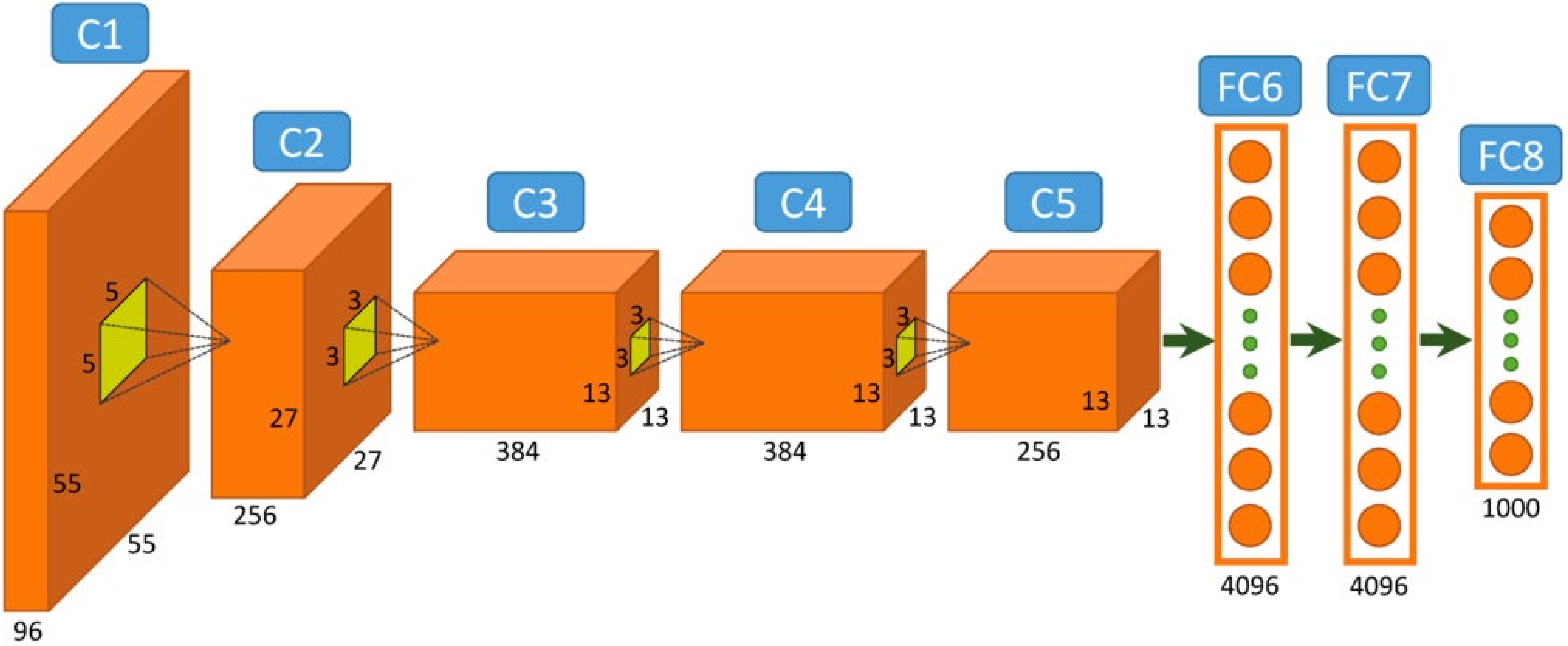


Top-5 error on this competition (2012 –AlexNet)



AlexNet

Alexnet (lead author Alex Krizhevsk, UofT)

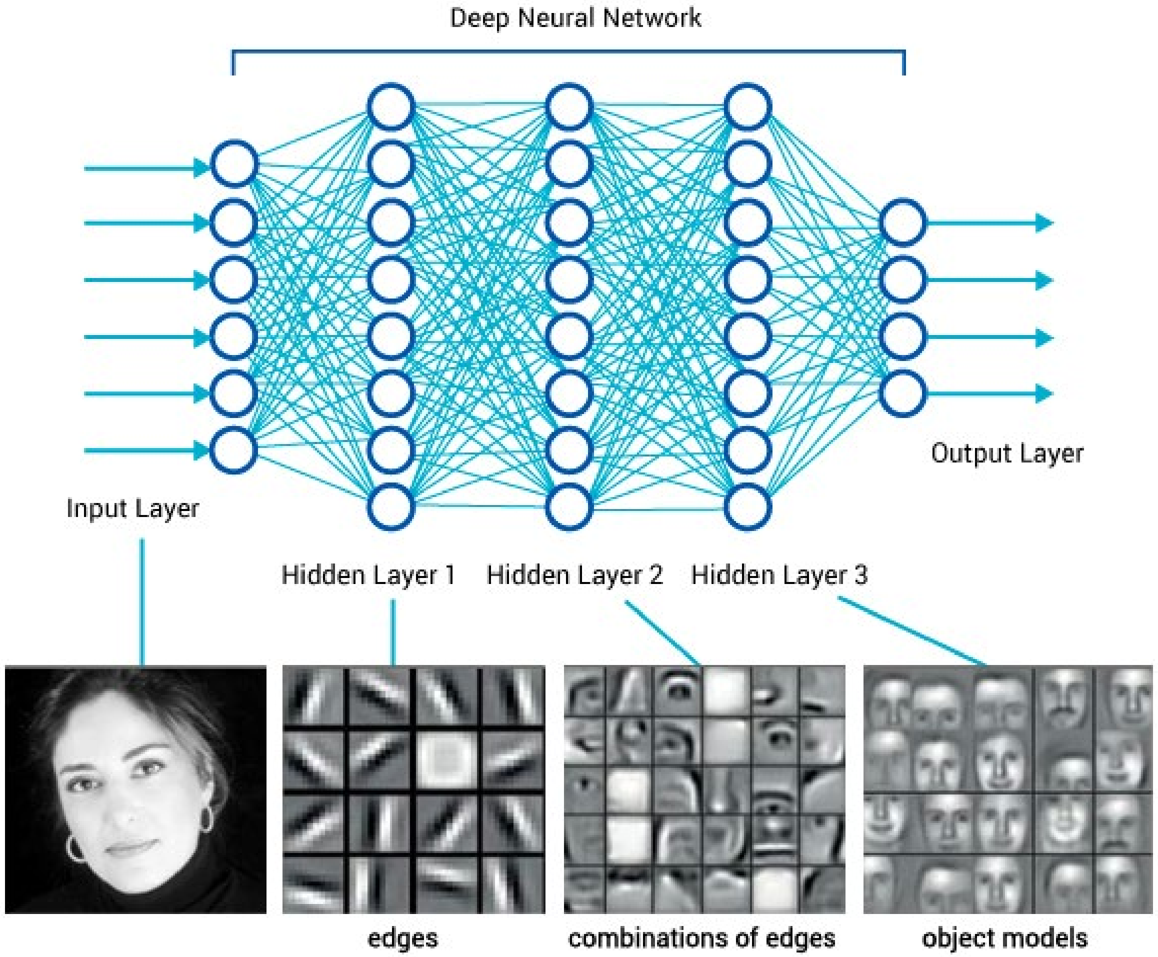


Using Nvidia CUDA (2006 announced)
Using pre-processing supp. (later)

https://www.saagie.com/fr/blog/object-detection-part1

UNIVERSITY OF CALGARY

What is happening?



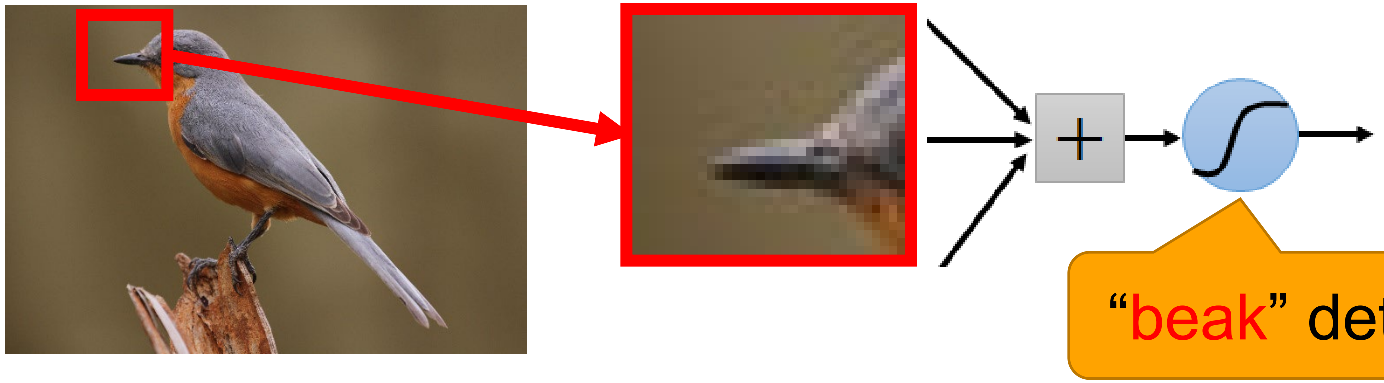
https://www.saagie.com/fr/blog/object-detection-part1

Convolution

Consider learning an image:

- Some patterns are much smaller than the whole image

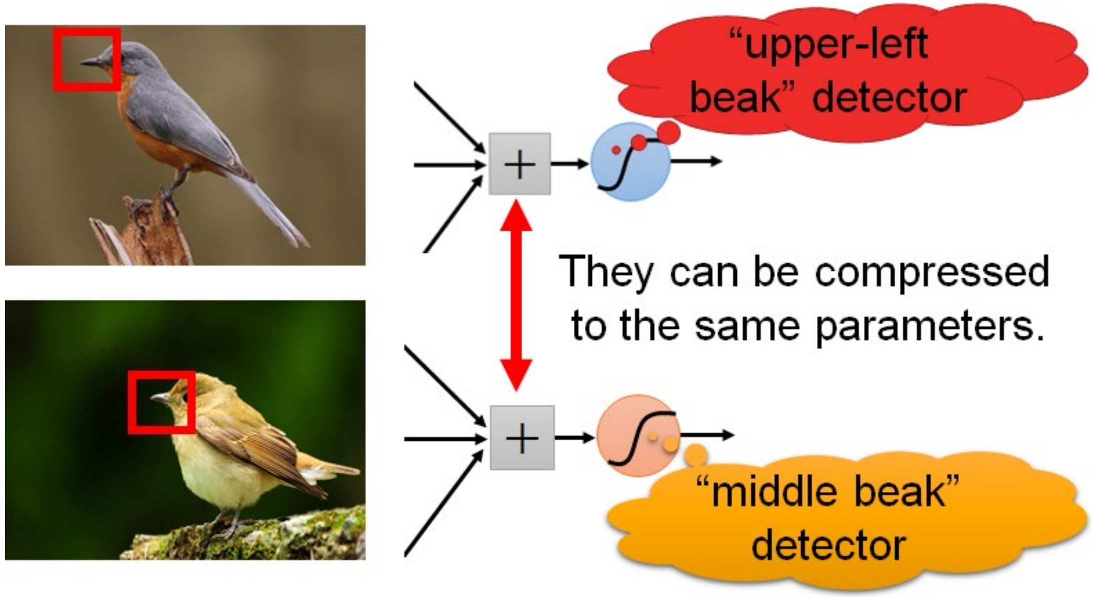
Can represent a small region with fewer parameters



“beak” detector

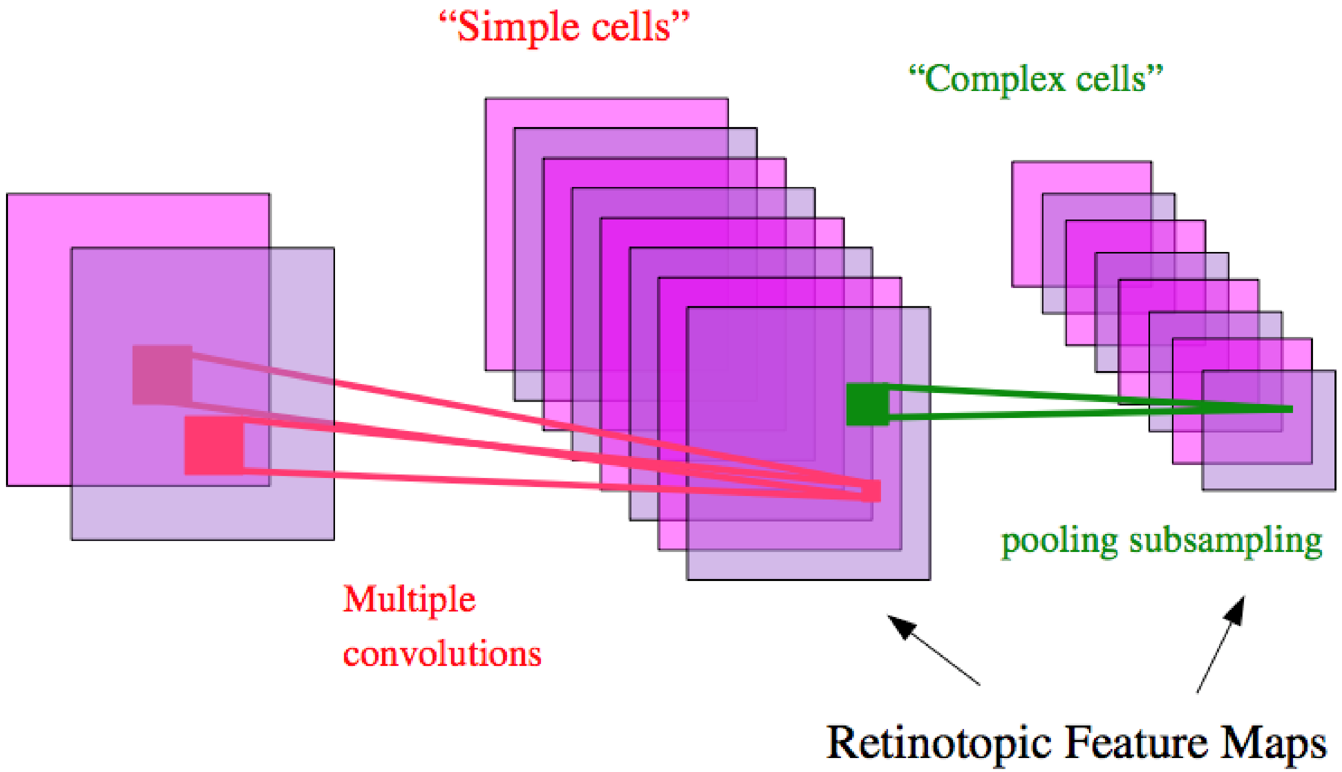
Detectors

- Same pattern appears in different places: They can be compressed! What about training a lot of such "small" detectors and each detector must "move around".


"upper-left beak" detector
They can be compressed to the same parameters.
"middle beak" detector

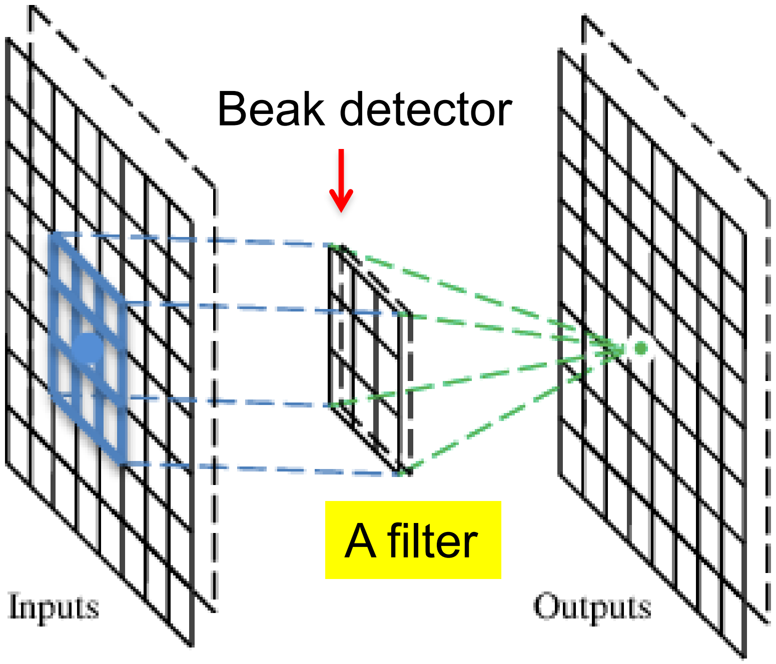
Model of vision in animals

- [Hubel & Wiesel 1962]:
  - simple cells detect local features
  - complex cells "pool" the outputs of simple cells within a retinotopic neighborhood.



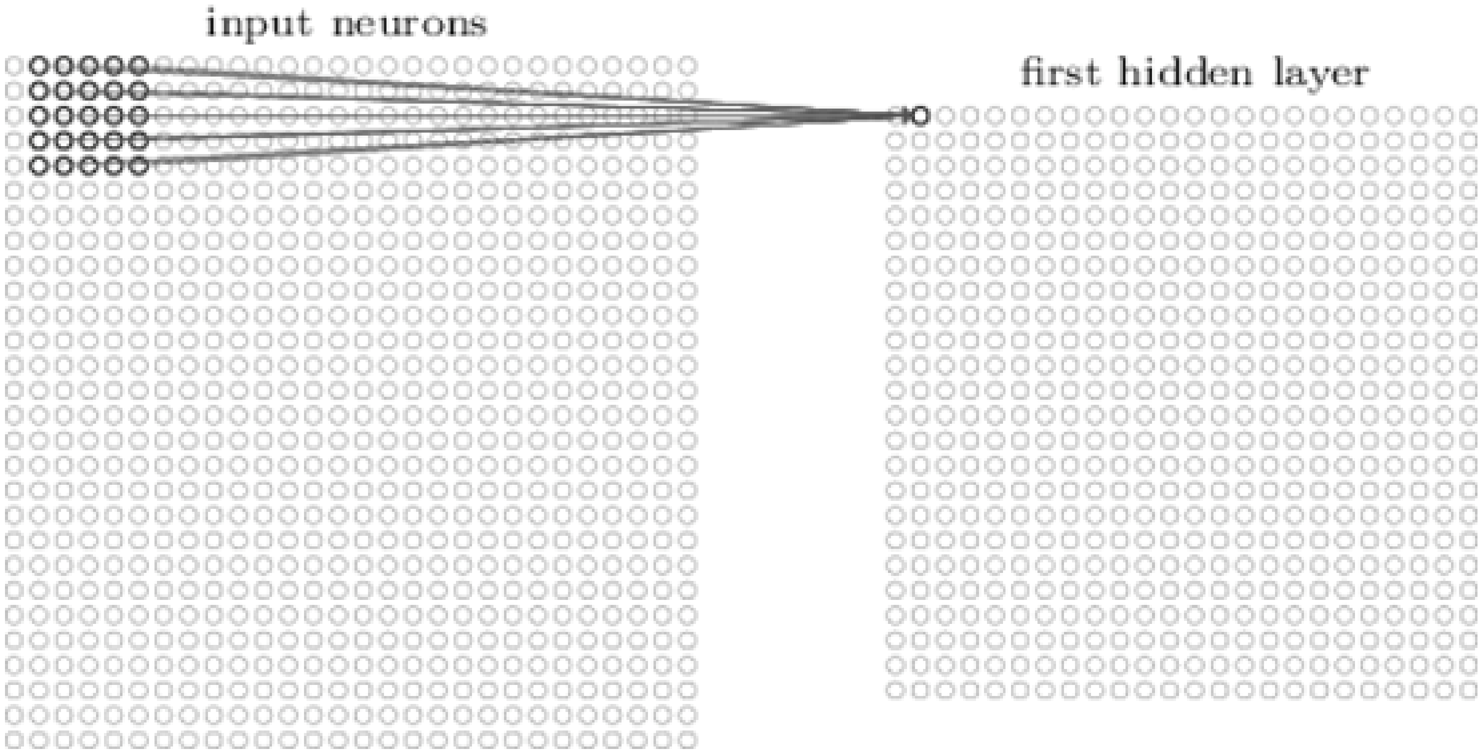
A convolutional layer

- A CNN is a neural network with some convolutional layers (and some other layers). A convolutional layer has a number of filters that does convolutional operation.

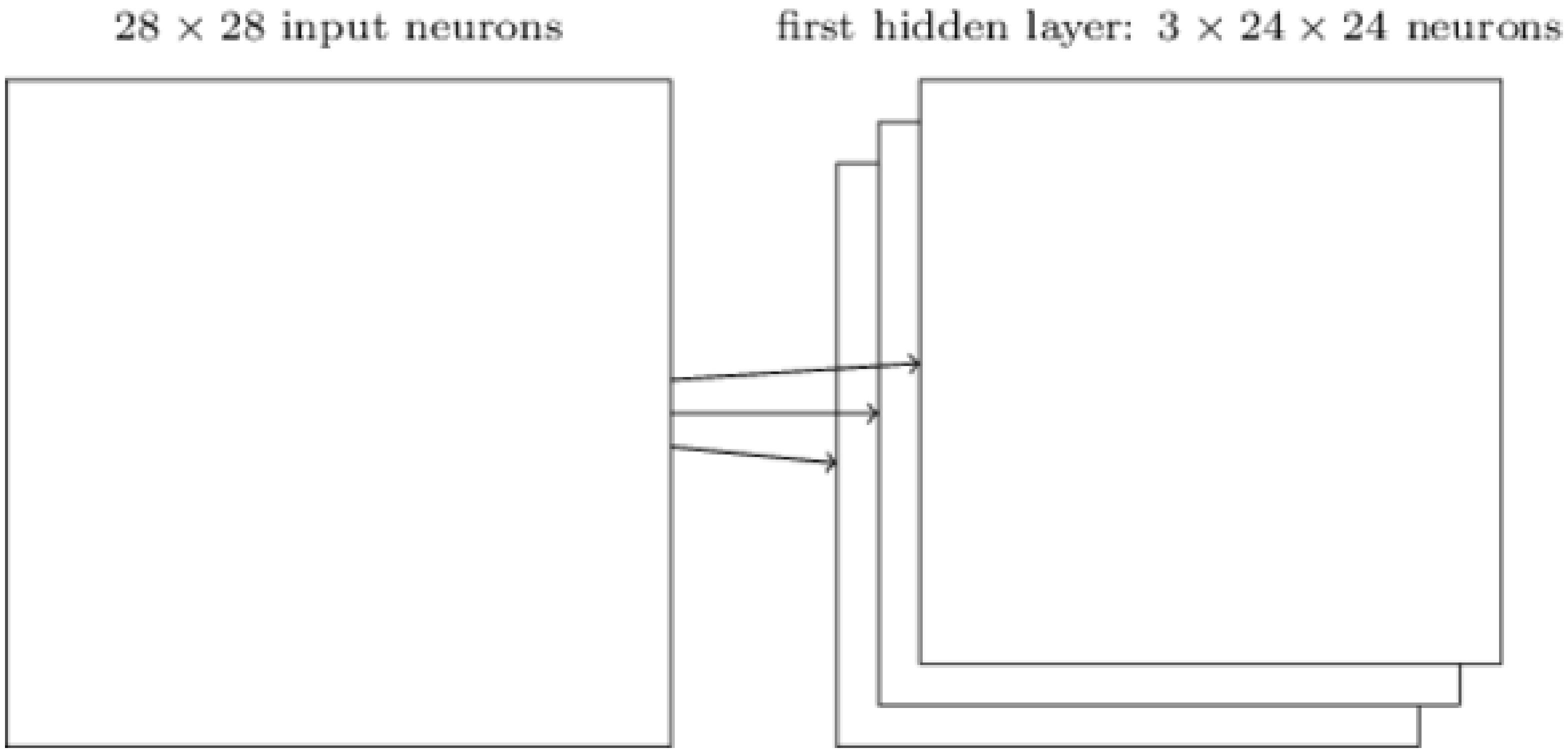


How do we convolve an image with an ANN?

Note that the parameters in the matrix defining the convolution are **tied** across all places that it is used



How do we do many convolutions of an image with an ANN?



Pooling

Why Pooling

- Subsampling pixels will not change the object

bird

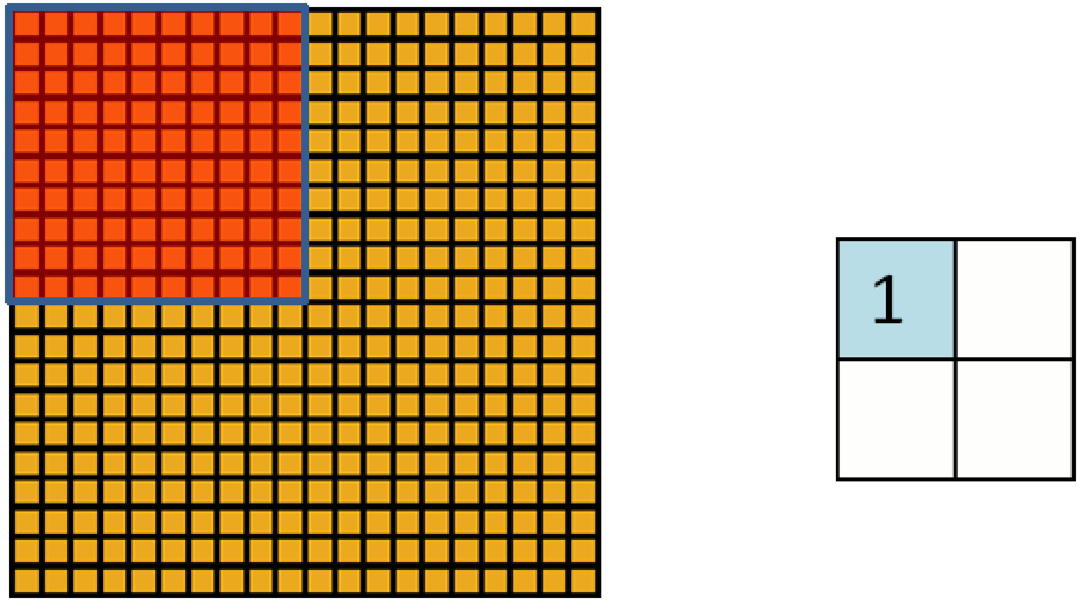
Subsampling

bird

We can subsample the pixels to make image smaller

→ fewer parameters to characterize the image

Pooling



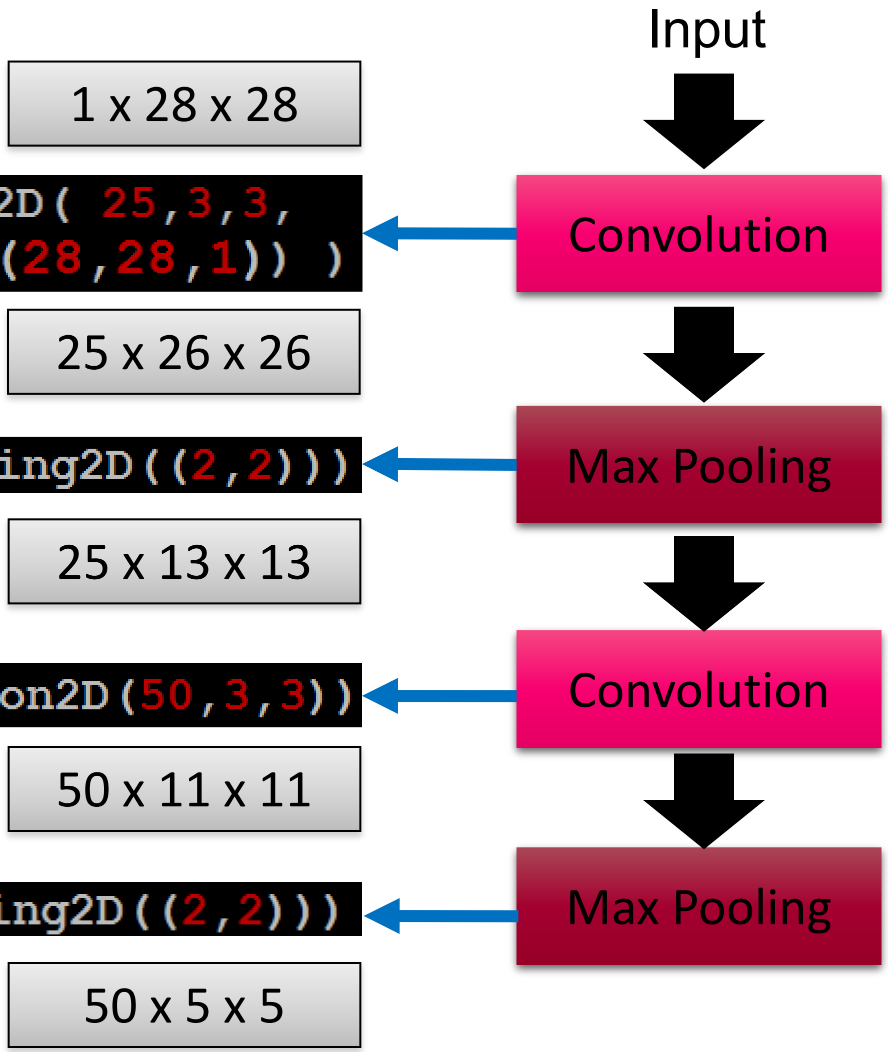
Convolved feature

Pooled feature

Full Convolution NN in Keras

CNN in Keras

Only modified the ***network structure*** and ***input format (vector -> 3-D array)***



```
model2.add( Convolution2D( 25,3,3,
            input_shape=(28,28,1)) )
```

How many parameters for each filter?

```
model2.add(MaxPooling2D((2,2)))
```

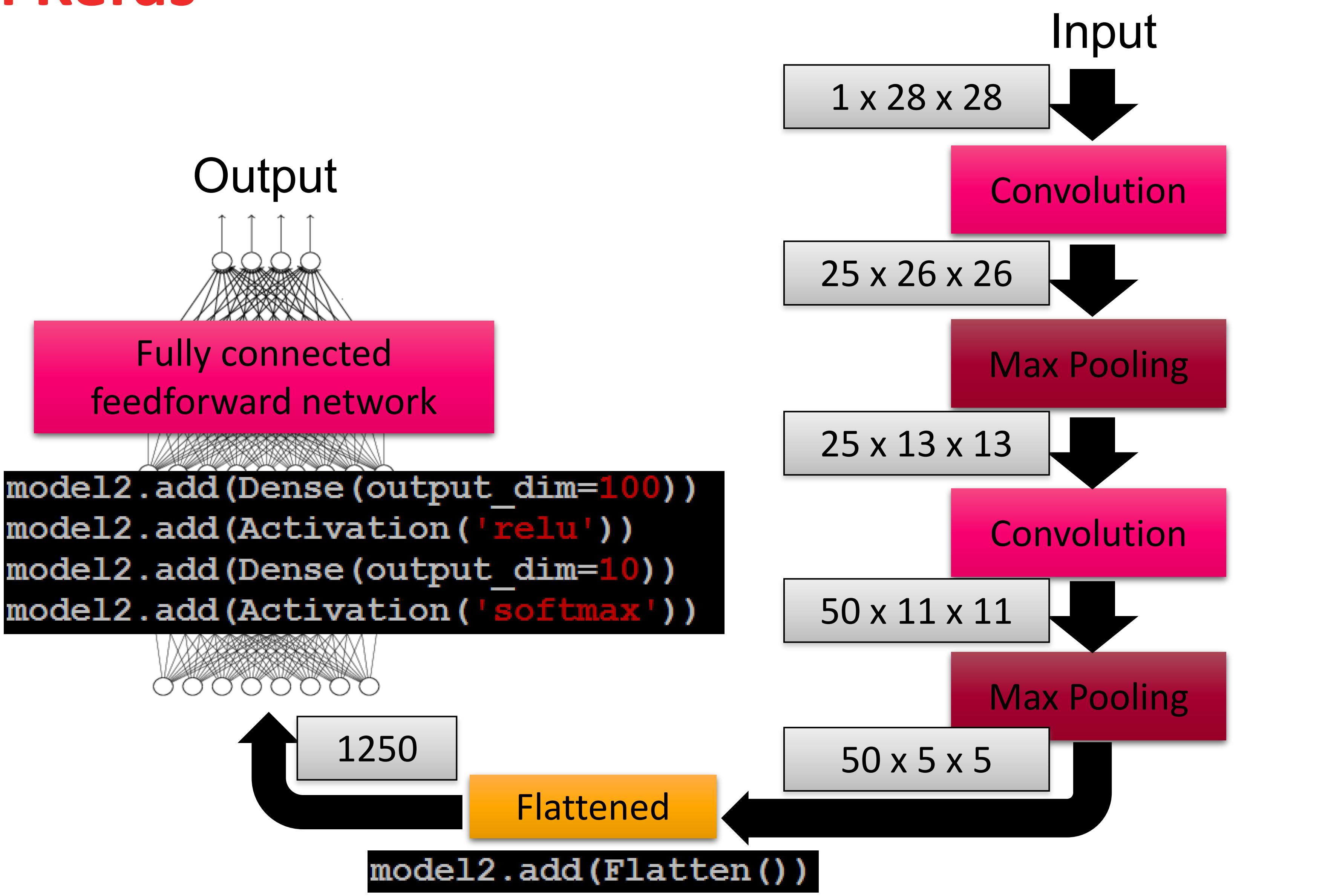
```
model2.add(Convolution2D(50,3,3))
```

How many parameters for each filter?

```
model2.add(MaxPooling2D((2,2)))
```

CNN in Keras

Only modified the ***network structure*** and ***input format (vector -> 3-D array)***



Input

1 x 28 x 28

Convolution

25 x 26 x 26

Max Pooling

25 x 13 x 13

Convolution

50 x 11 x 11

Max Pooling

50 x 5 x 5

Flattened

```
model2.add(Flatten())
```

1250

Fully connected feedforward network

```
model2.add(Dense(output_dim=100))
model2.add(Activation('relu'))
model2.add(Dense(output_dim=10))
model2.add(Activation('softmax'))
```

Output

Examples CNN

AlphaGo

19 x 19 matrix

Neural Network

Next move

(19 x 19 positions)

Black: 1

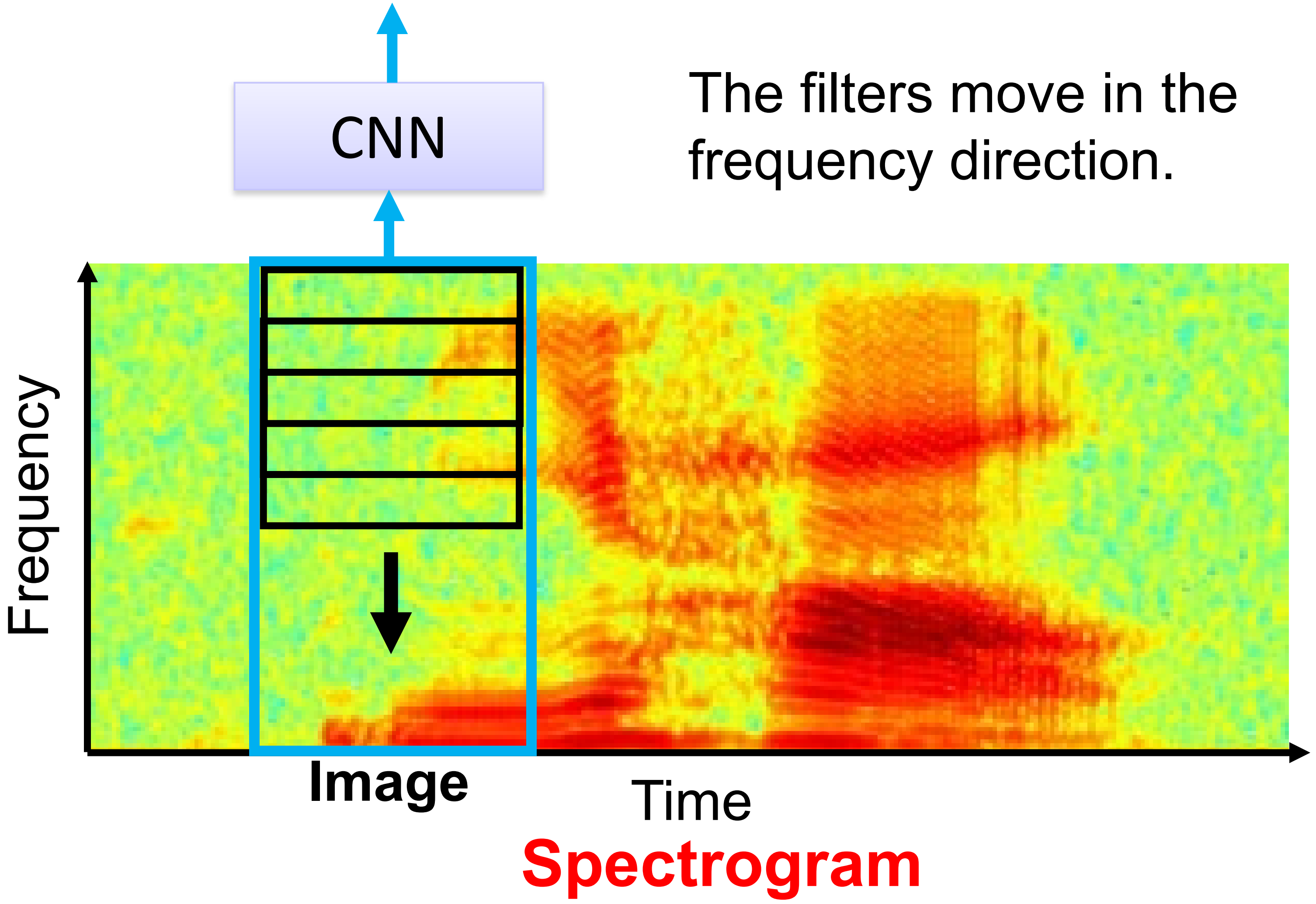
white: -1

none: 0

Fully-connected feedforward network can be used

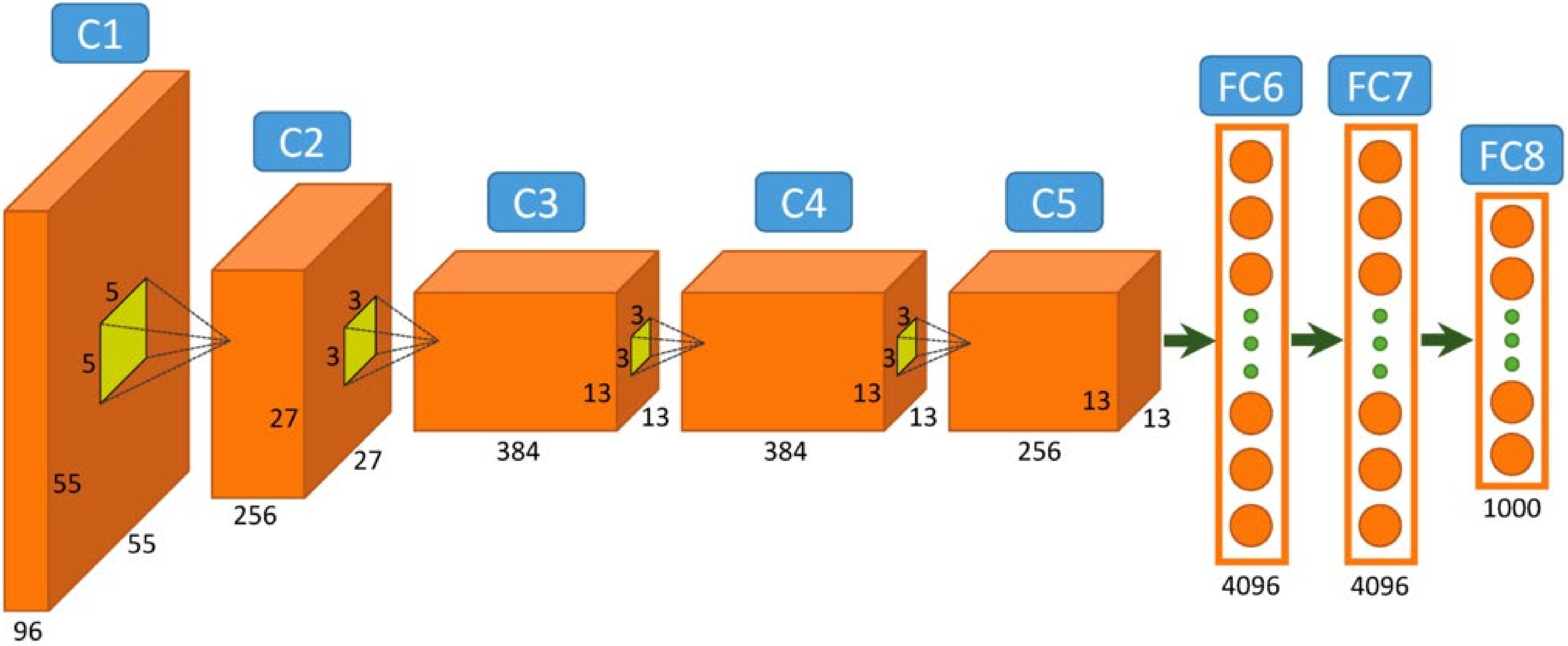
But CNN performs much better

CNN in speech recognition



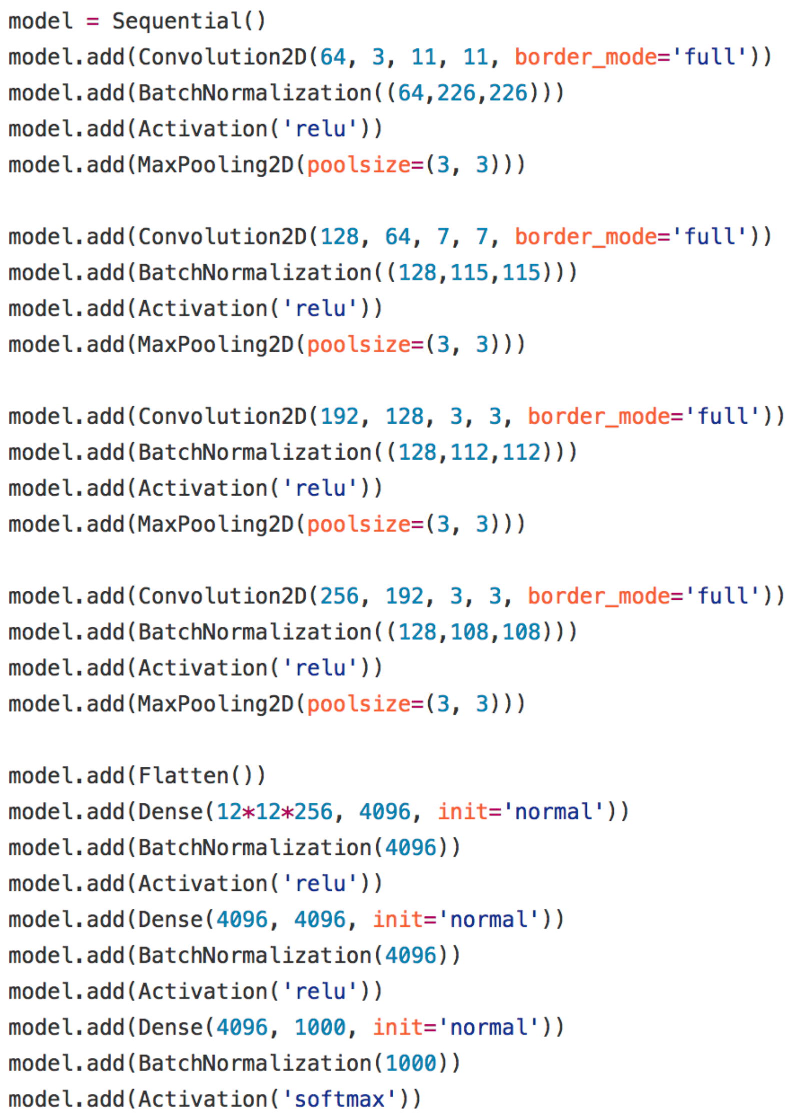
The filters move in the frequency direction.

Alexnet (lead author Alex Krizhevsk, UofT)



https://www.saagie.com/fr/blog/object-detection-part1

Alexnet in Keras

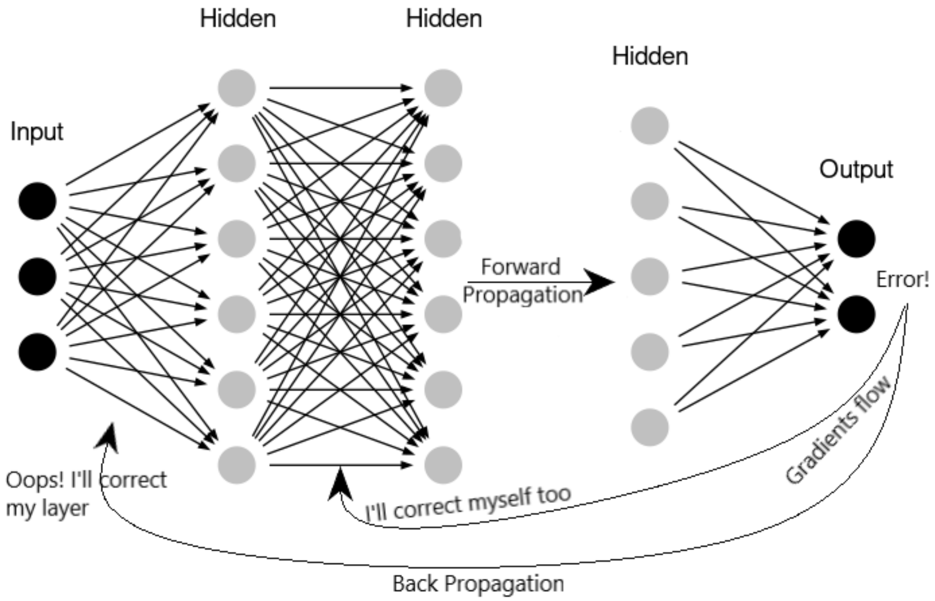


```
model = Sequential()
model.add(Convolution2D(64, 3, 11, 11, border_mode='full'))
model.add(BatchNormalization((64,226,226)))
model.add(Activation('relu'))
model.add(MaxPooling2D(poolsize=(3, 3)))

model.add(Convolution2D(128, 64, 7, 7, border_mode='full'))
model.add(BatchNormalization((128,115,115)))
model.add(Activation('relu'))
model.add(MaxPooling2D(poolsize=(3, 3)))

model.add(Convolution2D(192, 128, 3, 3, border_mode='full'))
model.add(BatchNormalization((128,112,112)))
model.add(Activation('relu'))
model.add(MaxPooling2D(poolsize=(3, 3)))

model.add(Convolution2D(256, 192, 3, 3, border_mode='full'))
model.add(BatchNormalization((128,108,108)))
model.add(Activation('relu'))
model.add(MaxPooling2D(poolsize=(3, 3)))

model.add(Flatten())
model.add(Dense(12*12*256, 4096, init='normal'))
model.add(BatchNormalization(4096))
model.add(Activation('relu'))
model.add(Dense(4096, 4096, init='normal'))
model.add(BatchNormalization(4096))
model.add(Activation('relu'))
model.add(Dense(4096, 1000, init='normal'))
model.add(BatchNormalization(1000))
model.add(Activation('softmax'))
```

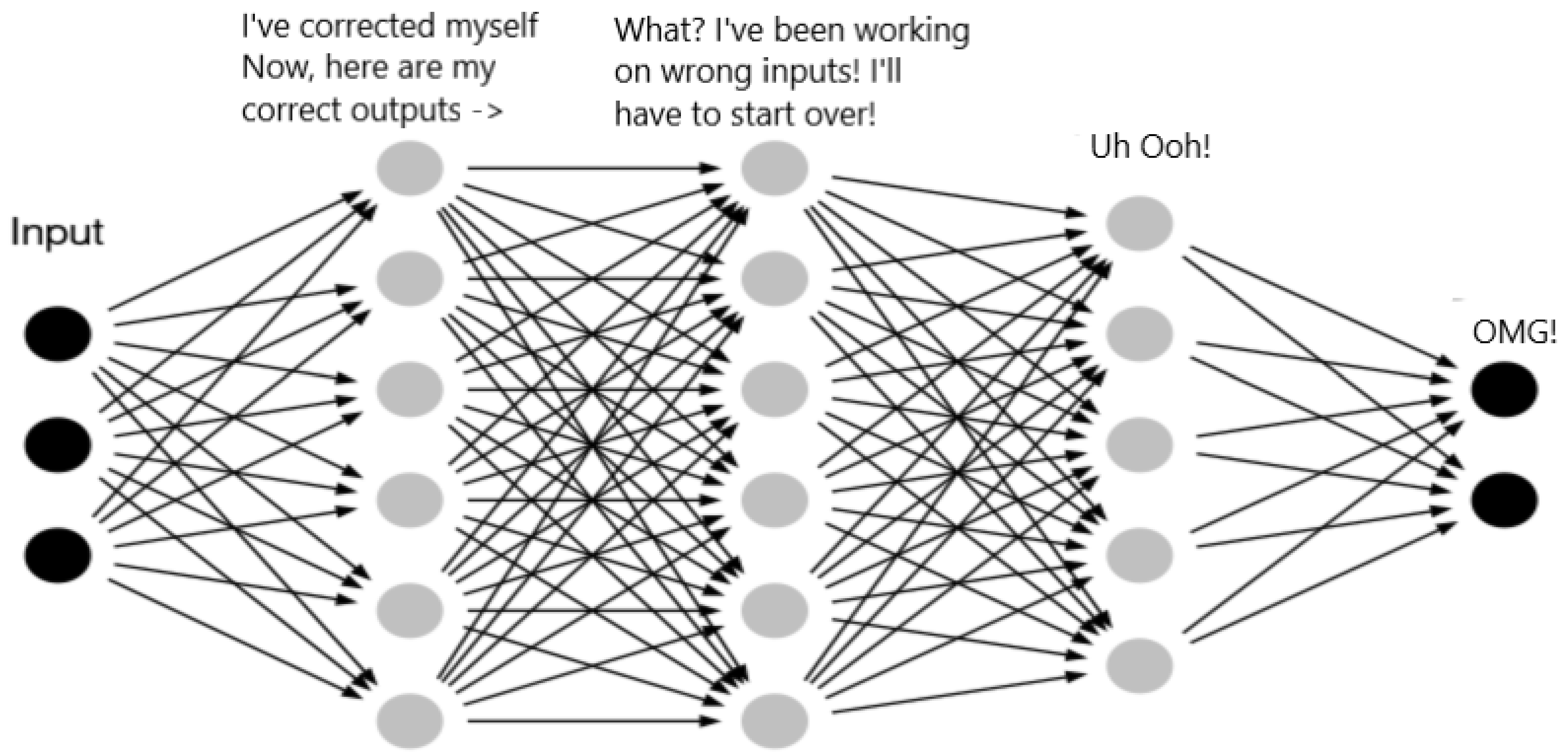
Batch Normalization

Batch Normalization?

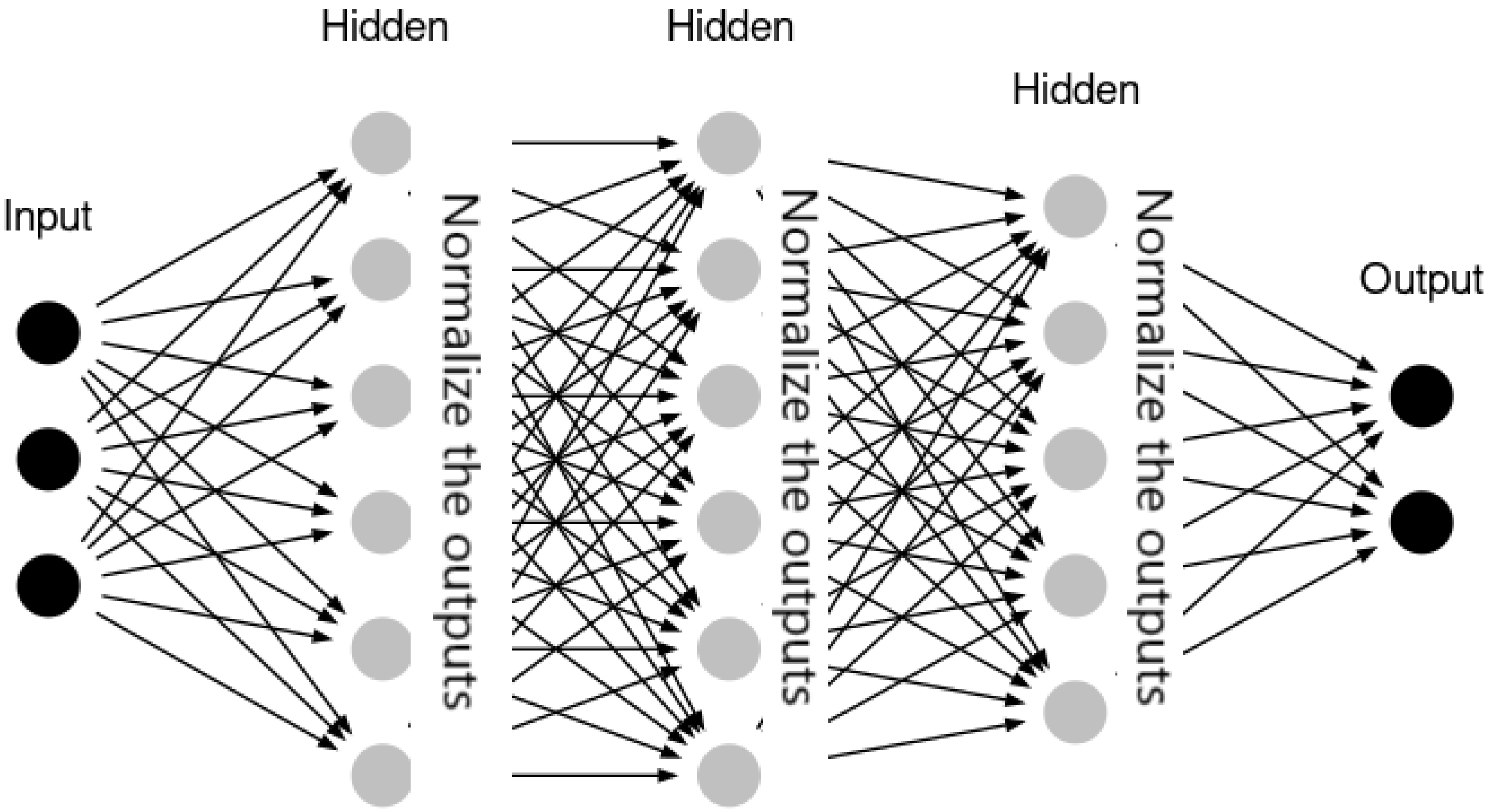
- Neural networks learn the problem using BackPropagation algorithm.
- BackPropagation involves computing gradients for each layer
- In deep networks this time explodes for training



Batch Normalization?



Batch Normalization?



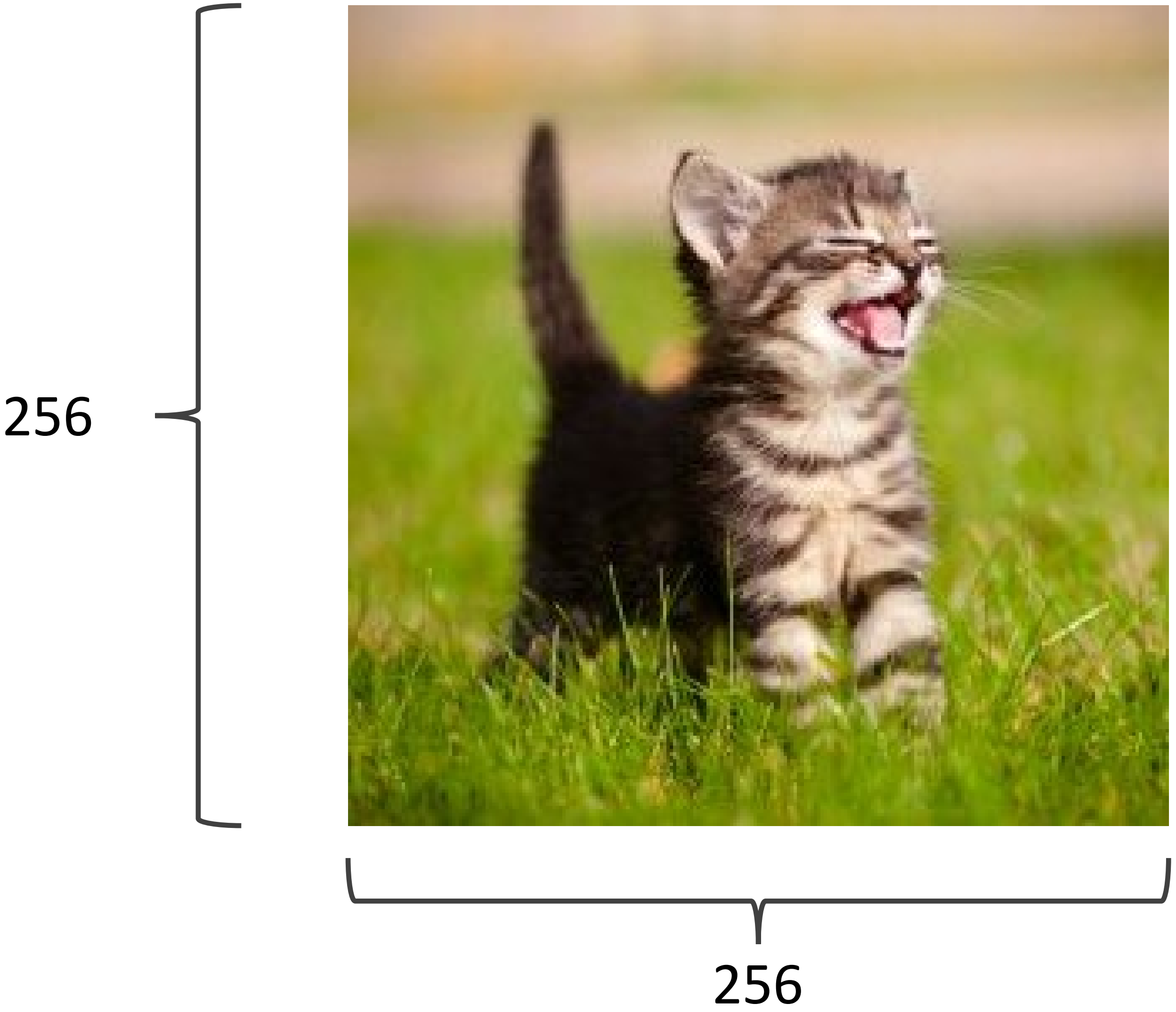
Batch Normalization?

- Normalization brings all the inputs centered around 0.
- This way, there is not much change in each layer input.
- So, layers in the network can learn from the back-propagation simultaneously, without waiting for the previous layer to learn.
- This speeds up the training of networks.

Pre-Processing

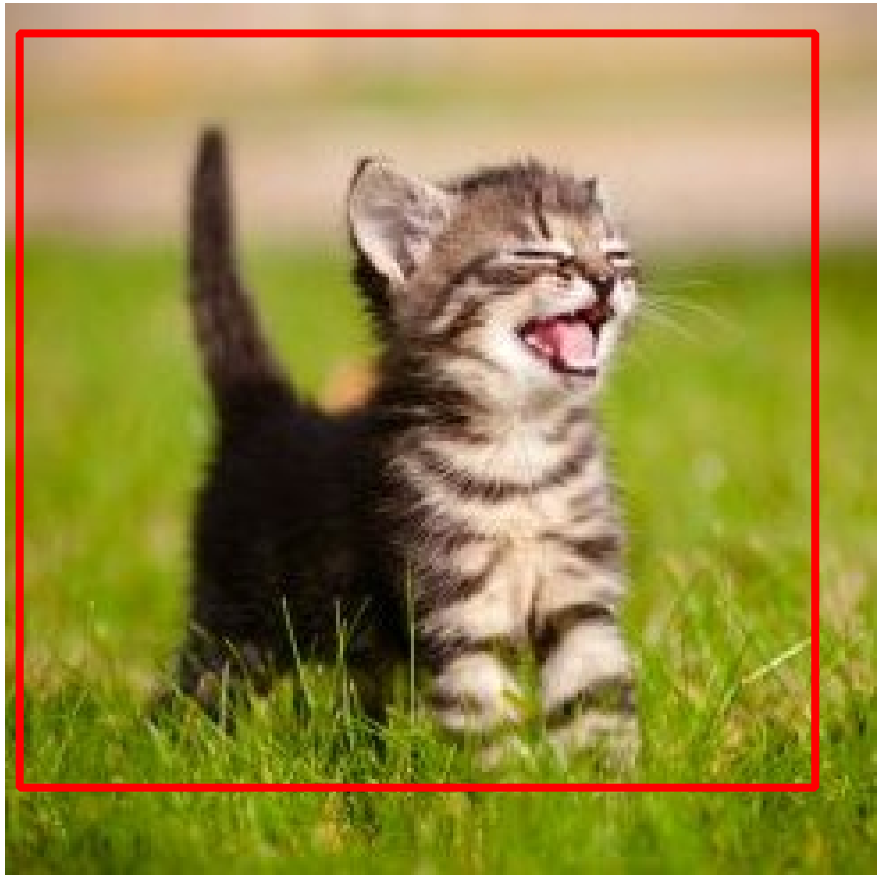
Preprocessing and Data Augmentation

Preprocessing and Data Augmentation



Preprocessing and Data Augmentation

224x224



Preprocessing and Data Augmentation

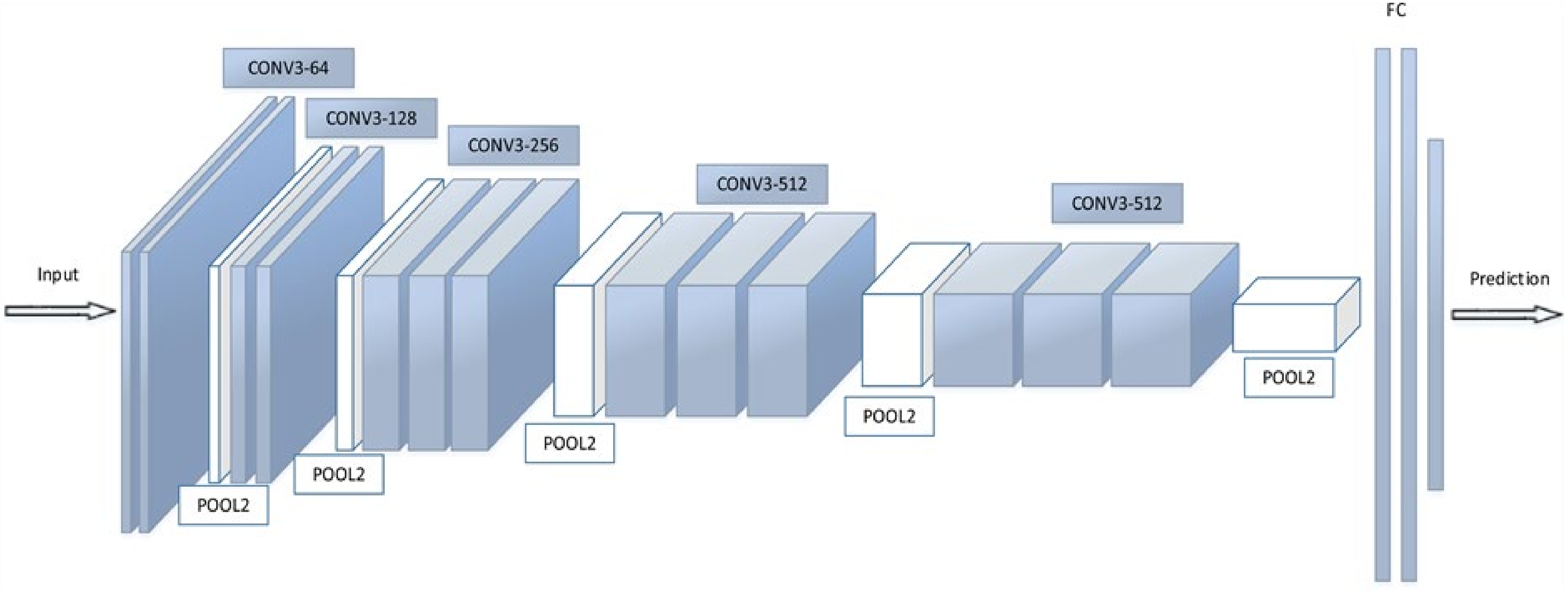
224x224

True label: Abyssinian cat

Other CNNs

VGG Network

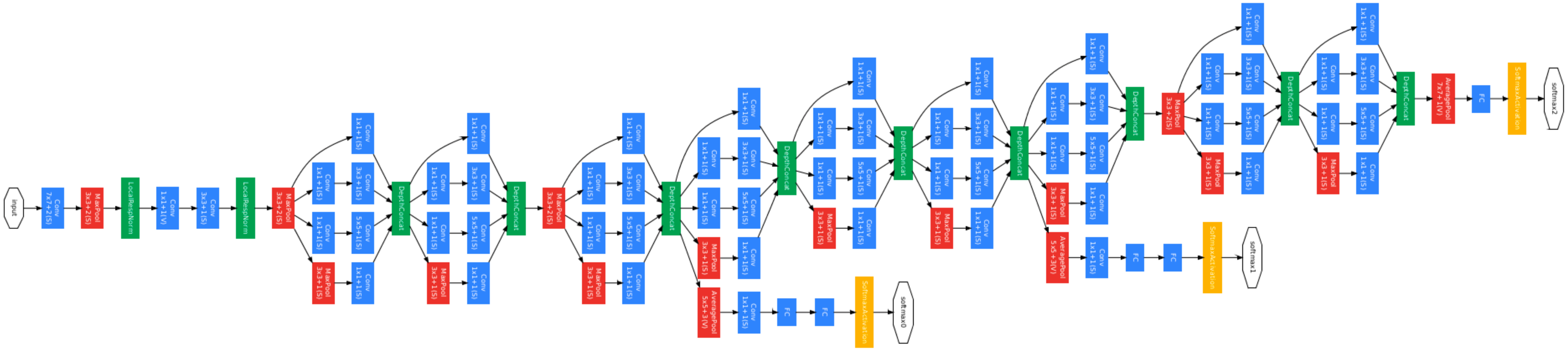
Top-5:



Keras: https://gist.github.com/baraldilorenzo/07d7802847aaad0a35d3

Simonyan and Zisserman, 2014.

GoogLeNet



Keras: https://gist.github.com/joelouismarino/a2ede9ab3928f999575423b9887abd14

Szegedy et al. 2014

ResNet

Sorry, does not fit in slide.

http://felixlaumon.github.io/assets/kaggle-right-whale/resnet.png

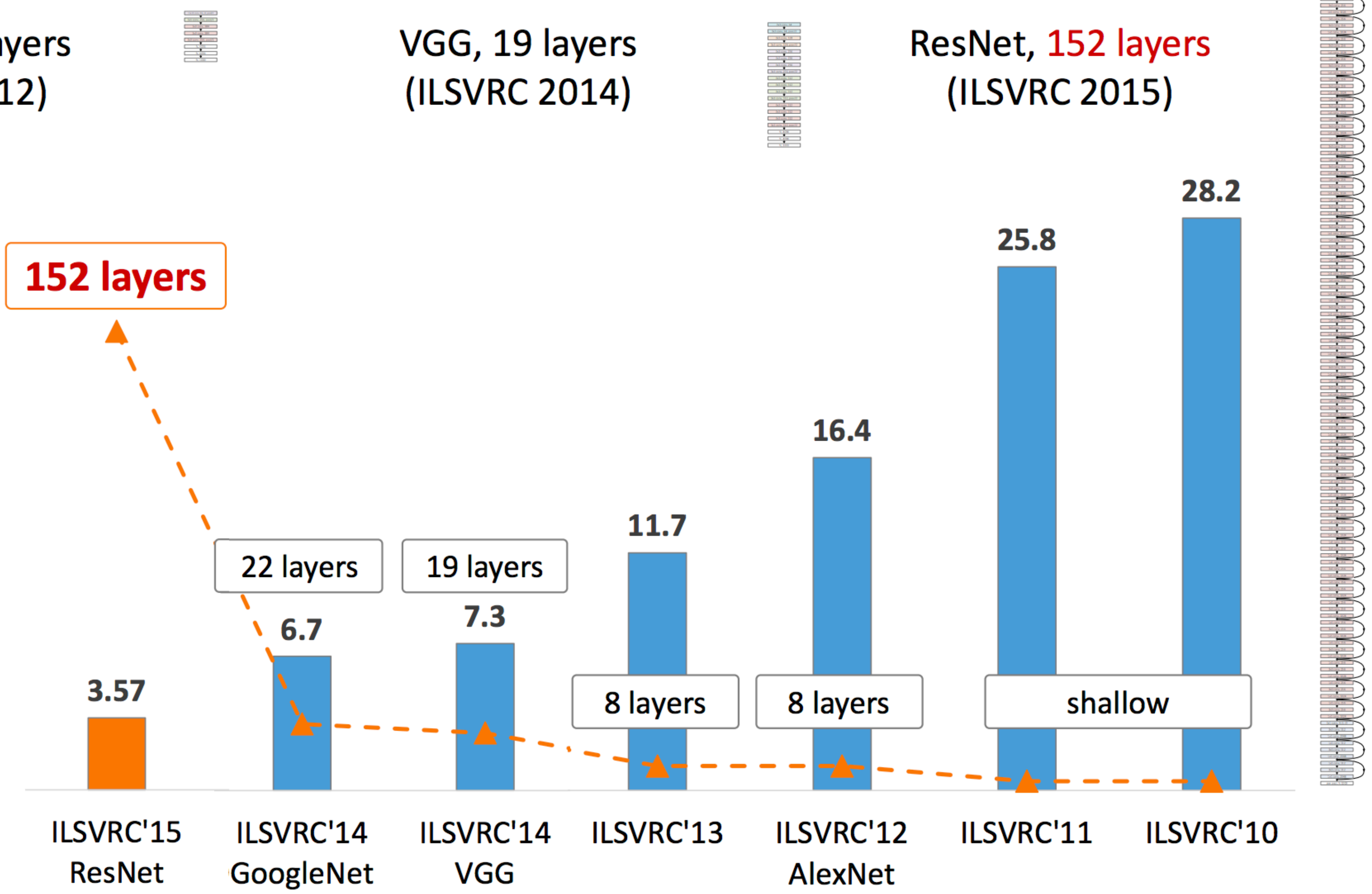
Keras: https://github.com/raghakot/keras-resnet/blob/master/resnet.py

Revolution of Depth

AlexNet, 8 layers
(ILSVRC 2012)

VGG, 19 layers
(ILSVRC 2014)

ResNet, 152 layers
(ILSVRC 2015)



UNIVERSITY OF CALGARY

Onward to … Deep Learning

Jonathan Hudson
jwhudson@ucalgary.ca
https://pages.cpsc.ucalgary.ca/~jwhudson/

UNIVERSITY OF CALGARY