2017-18 Marketing & Business Analytics

↳ FMCG -

Instagram → impression - normal day

v/s

Impression with event

Images → $\boxed{\text{model}}$ → Y B

→ N B

TRANSFER LEARNING

3500 Images → 60-65%

↓

Beer company Europe → model → 75-80%

TRANSFER LEARNING

Transfer learning involves taking a pre-trained neural network and adapting the neural network to a new, different data set.

THE FOUR MAIN CASES WHEN USING TRANSFER LEARNING

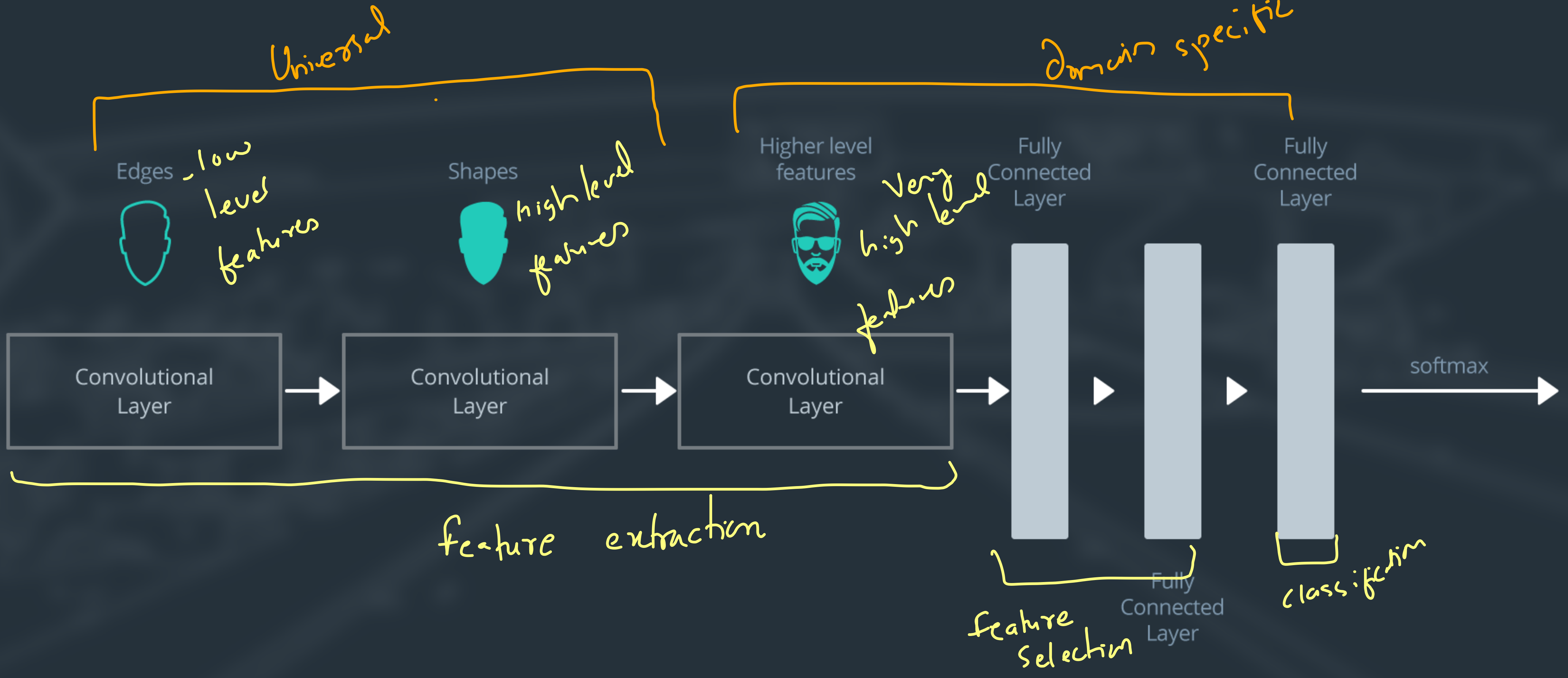
1. new data set is small, new data is similar to original training data
2. new data set is small, new data is different from original training data
3. new data set is large, new data is similar to original training data
4. new data set is large, new data is different from original training data

Guide for How to Use Transfer Learning

Size of Data Set	SIMILAR	DIFFERENT
LARGE	3 Fine-tune	4 Fine-tune or Retrain
SMALL	1 End of ConvNet	2 Start of ConvNet

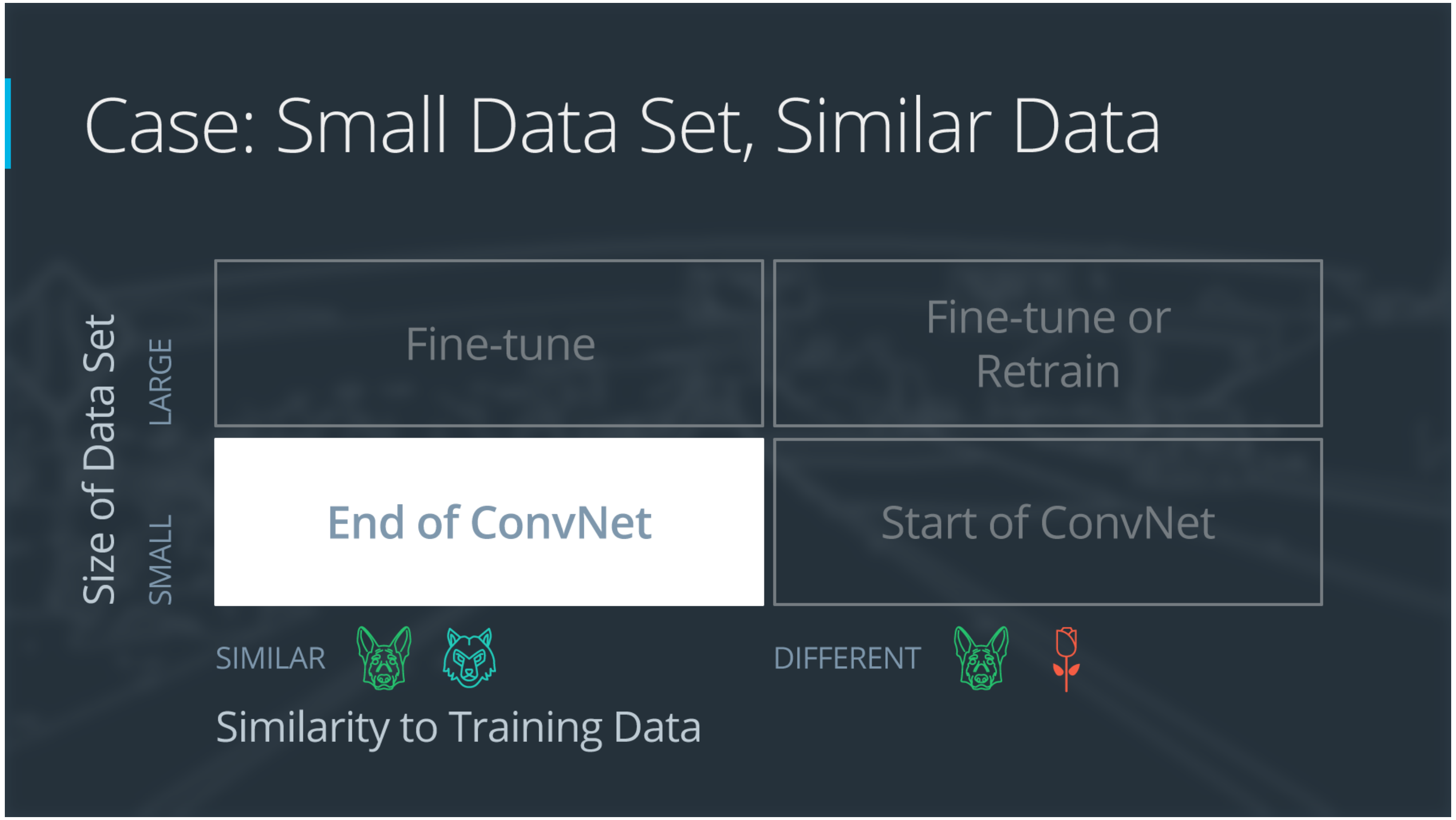
Similarity to Training Data

Pre-trained Convolutional Neural Network

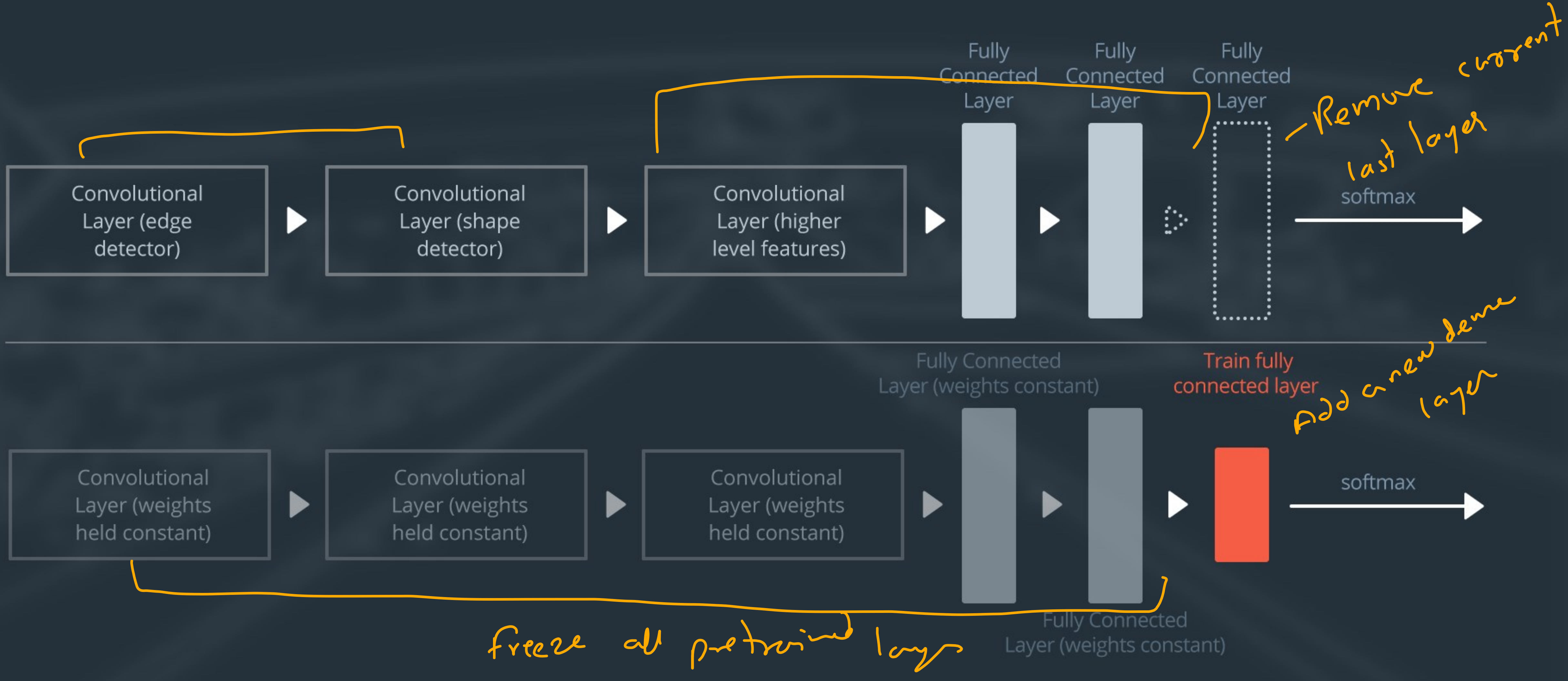


If the new data set is small and similar to the original training data:

- slice off the end of the neural network
- add a new fully connected layer that matches the number of classes in the new data set
- randomize the weights of the new fully connected layer; freeze all the weights from the pre-trained network
- train the network to update the weights of the new fully connected layer

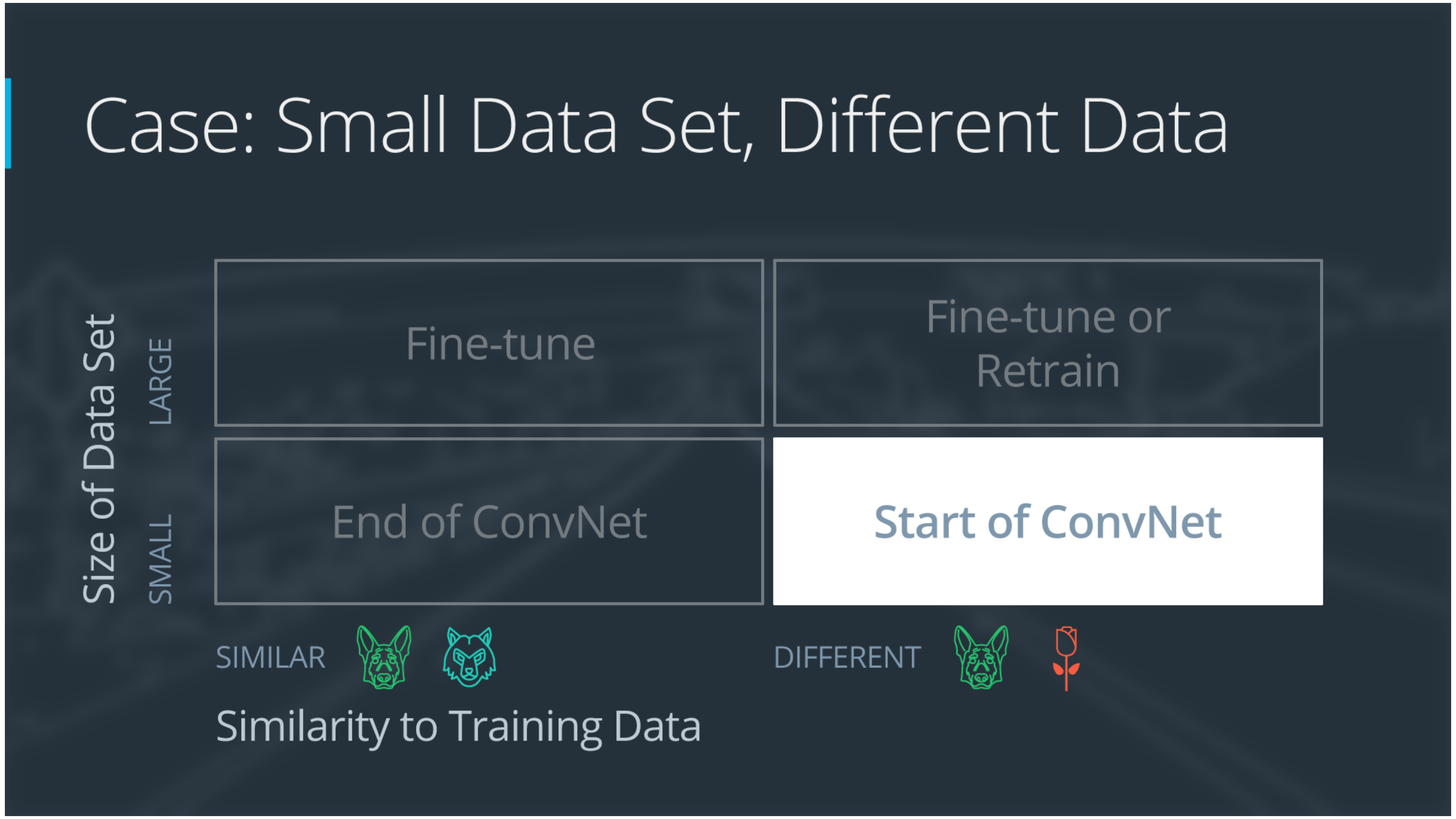


Case: Small Data Set, Similar Data

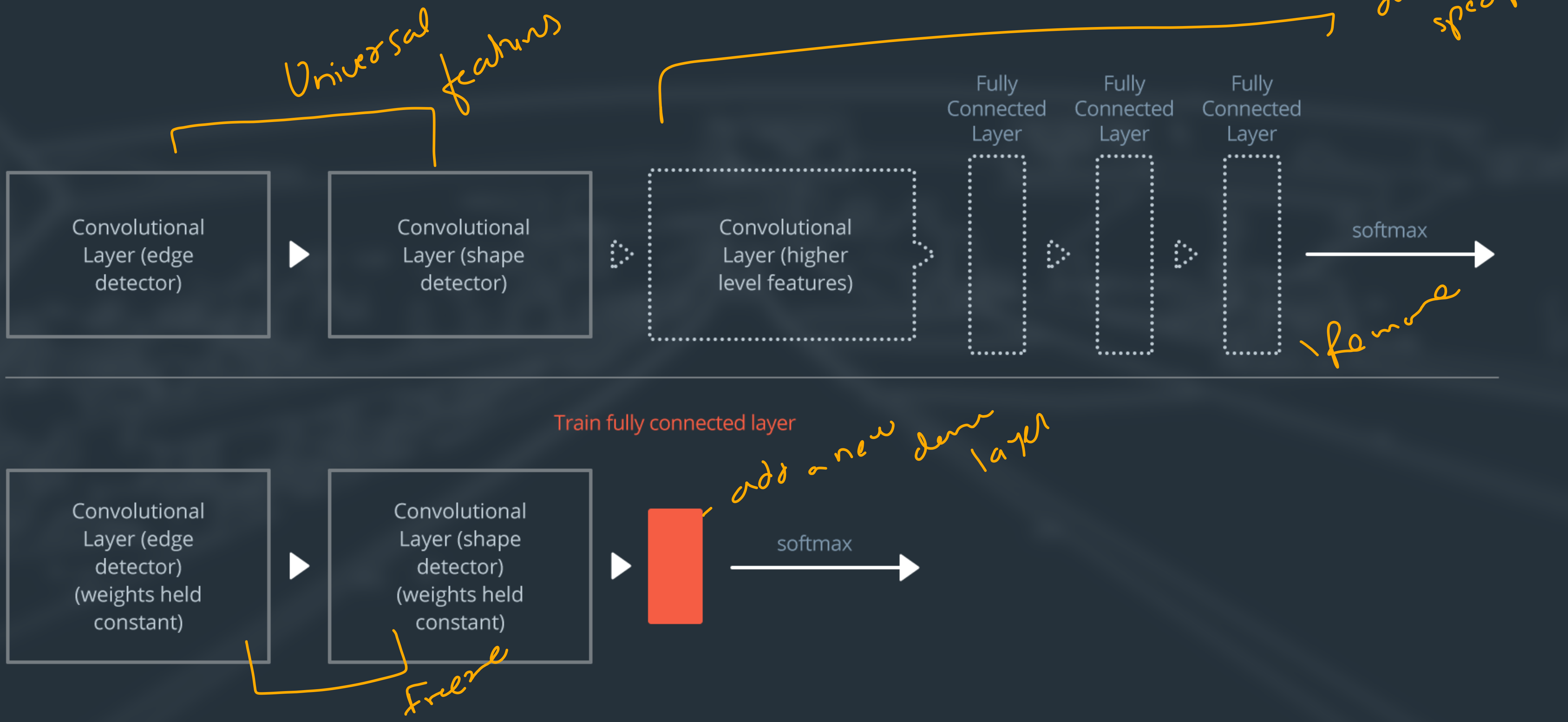


If the new data set is small and different from the original training data:

- slice off most of the pre-trained layers near the beginning of the network
- add to the remaining pre-trained layers a new fully connected layer that matches the number of classes in the new data set
- randomize the weights of the new fully connected layer; freeze all the weights from the pre-trained network
- train the network to update the weights of the new fully connected layer

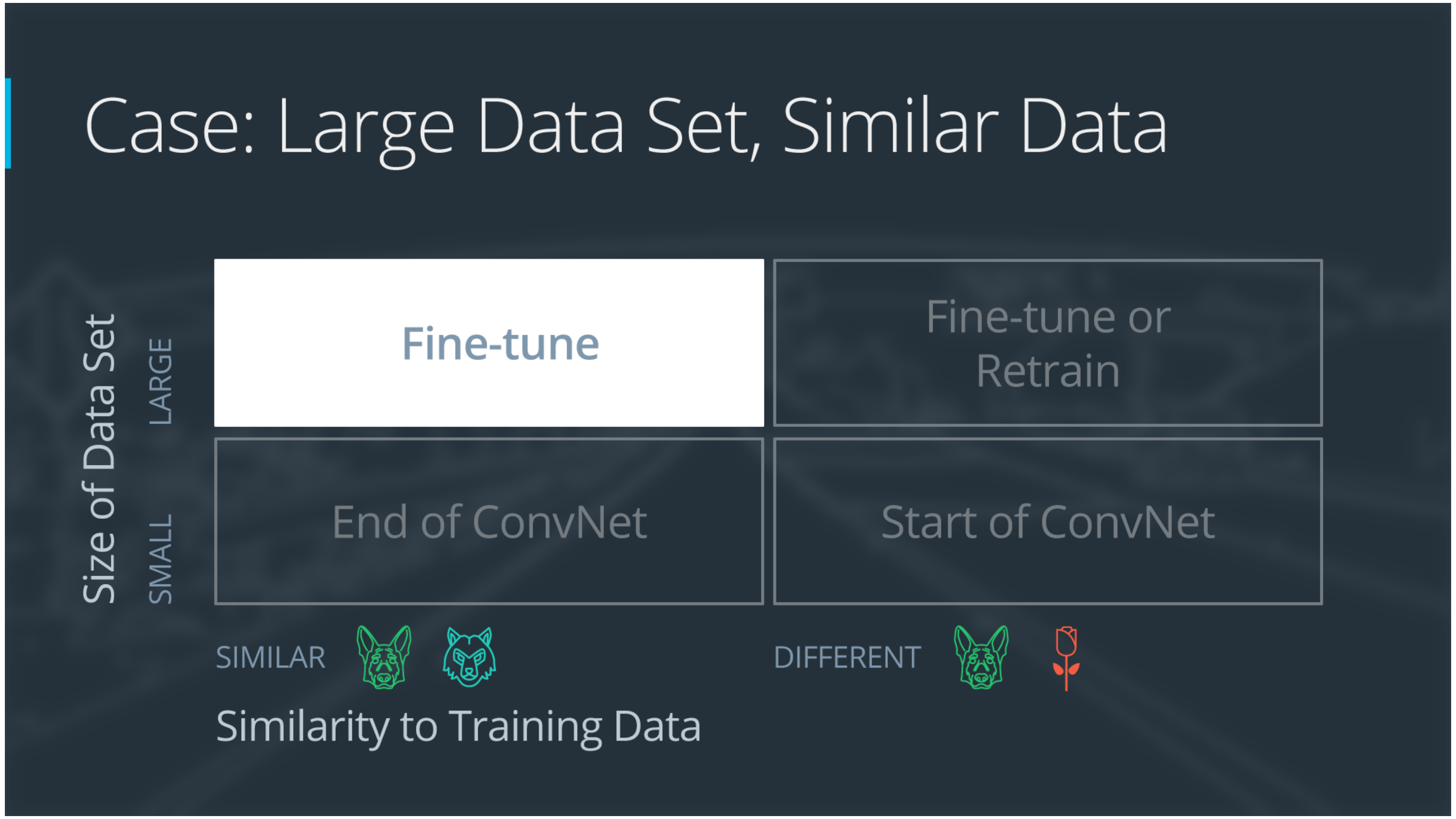


Case: Small Data Set, Different Data

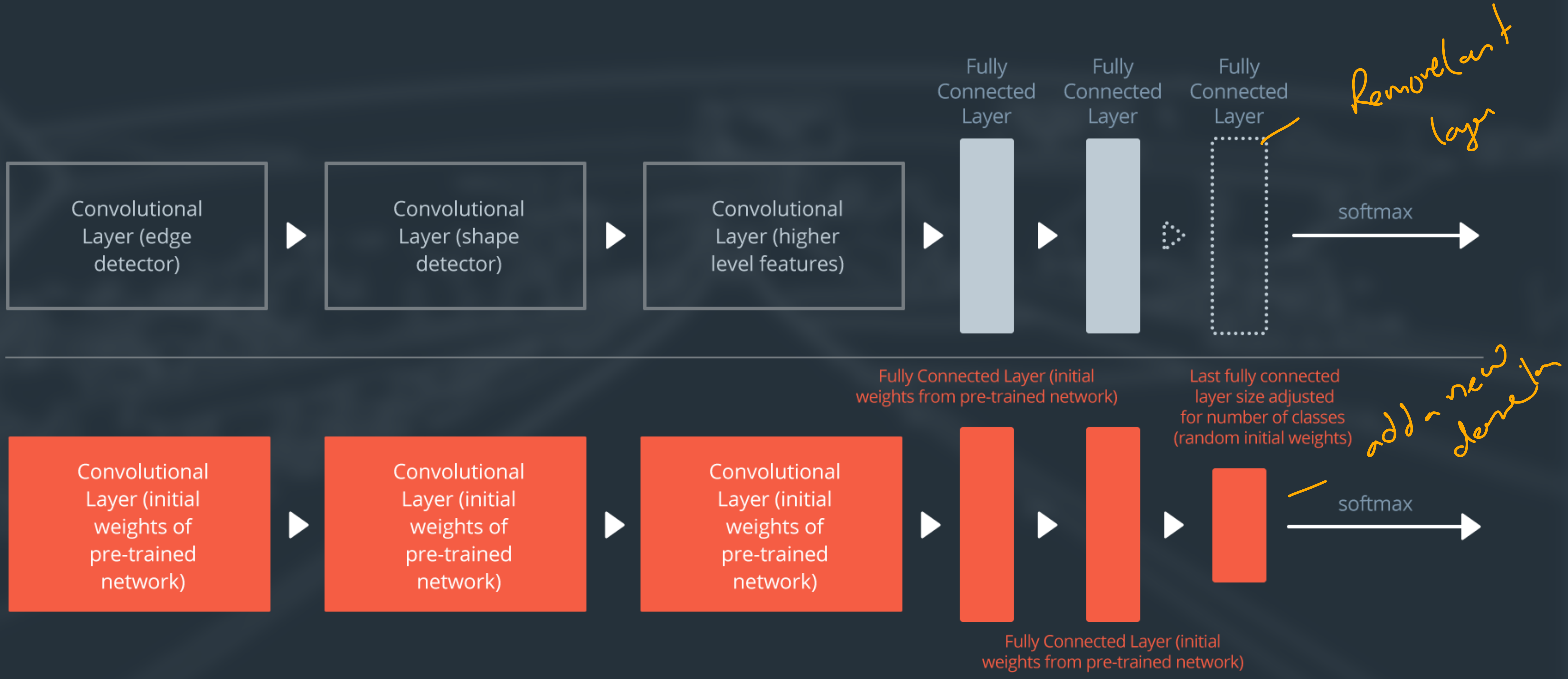


If the new data set is large and similar to the original training data:

- remove the last fully connected layer and replace with a layer matching the number of classes in the new data set
- randomly initialize the weights in the new fully connected layer
- initialize the rest of the weights using the pre-trained weights
- re-train the entire neural network

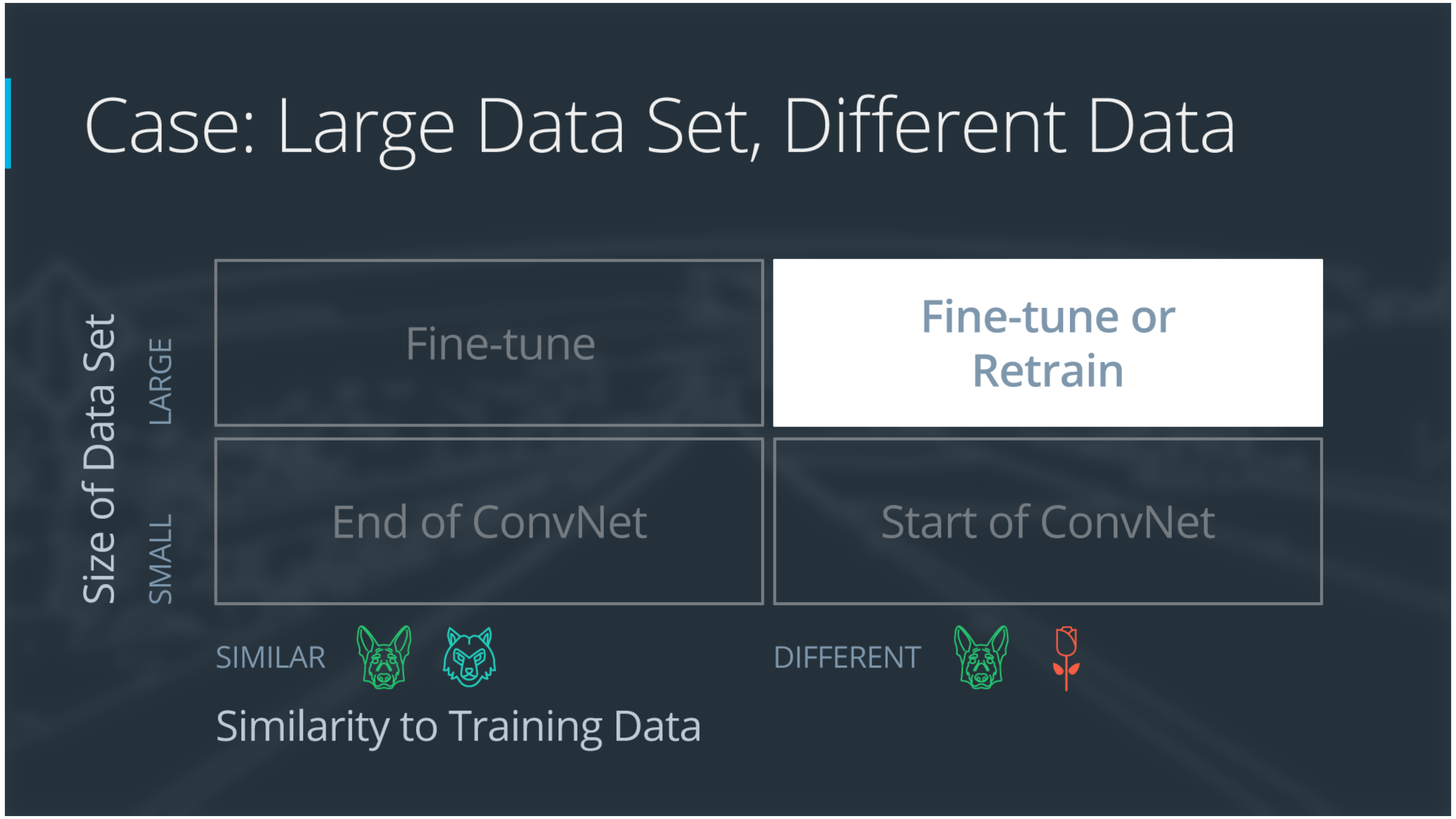


Case: Large Data Set, Similar Data

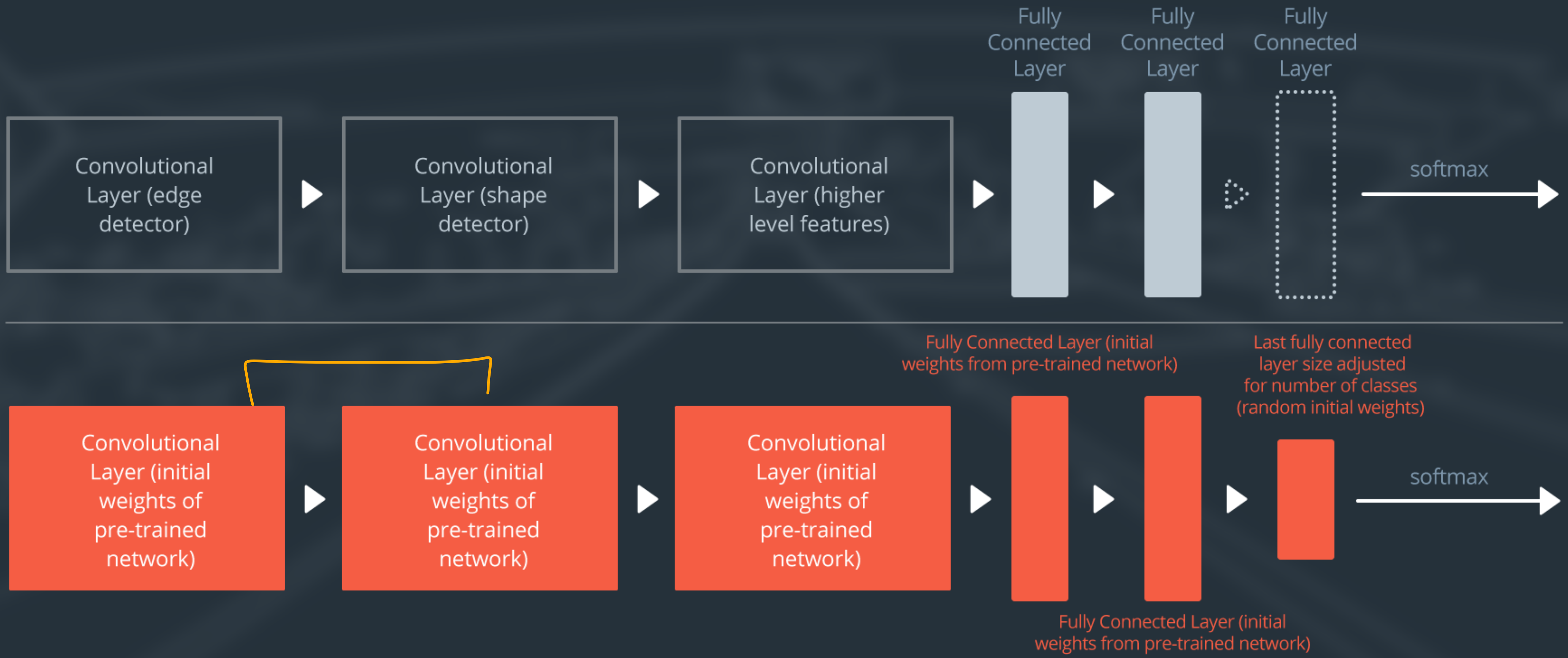


If the new data set is large and different from the original training data:

- remove the last fully connected layer and replace with a layer matching the number of classes in the new data set
- retrain the network from scratch with randomly initialized weights
- alternatively, you could just use the same strategy as the "large and similar" data case



Case: Large Data Set, Different Data



- https://keras.io/applications/

IMAGENET - WHERE HAVE WE BEEN?

- In 2006, Fei-Fei Li started ruminating on an idea.

- ***"We decided we wanted to do something completely unprecedented. We're going to map out the entire world of objects."***

- Resulting dataset was called ImageNet.

Fei-Fei Li

Chief Scientist of AI/ML at Google Cloud; Associate Professor at Stanford, Director of Stanford A.I. Lab

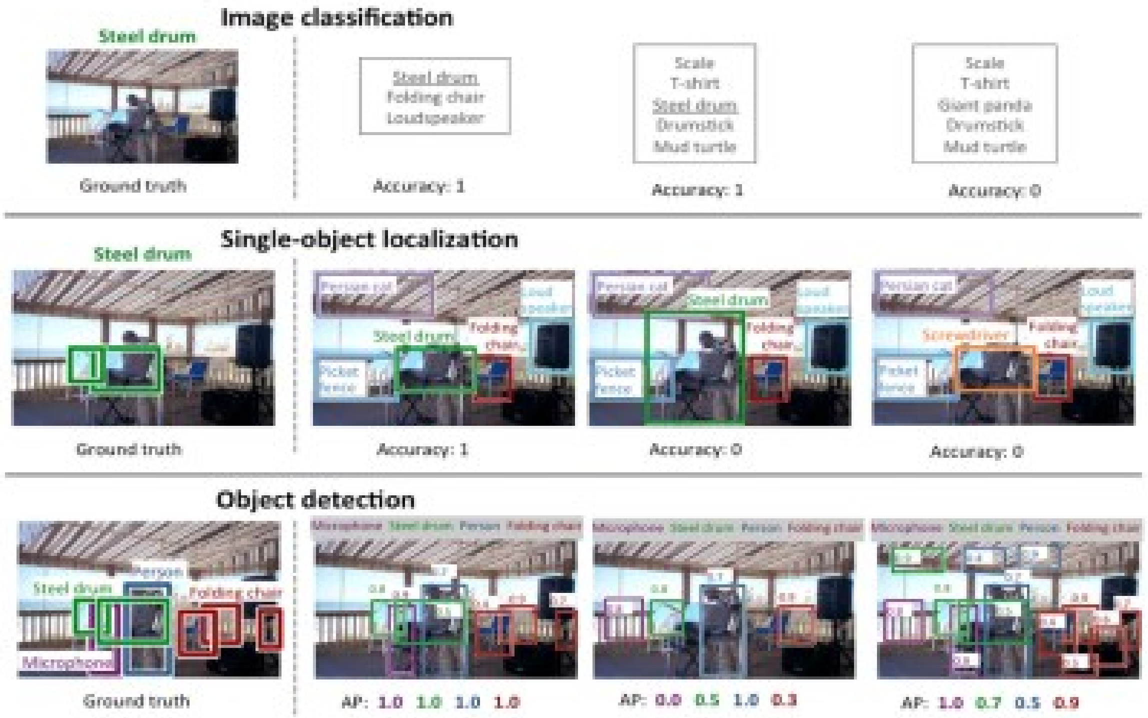
IMAGENET

- ImageNet project is a large visual database designed for use in visual object recognition software research.

- ImageNet database was presented for the first time as a poster at the 2009 Conference on Computer Vision and Pattern Recognition (CVPR) in Florida.

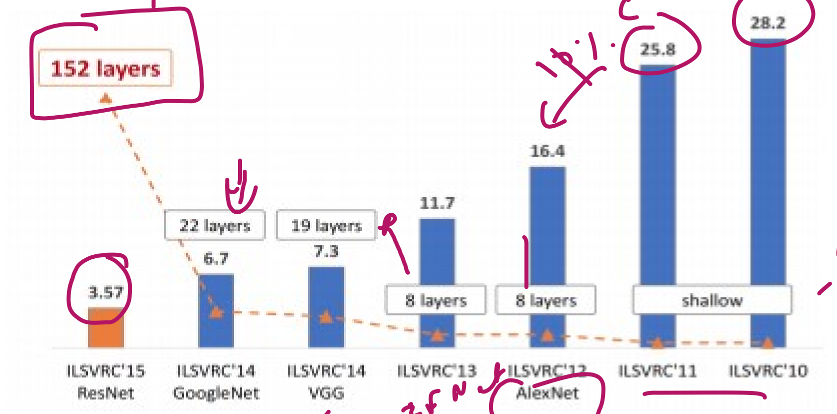
THE IMAGENET CHALLENGE - ILSVRC

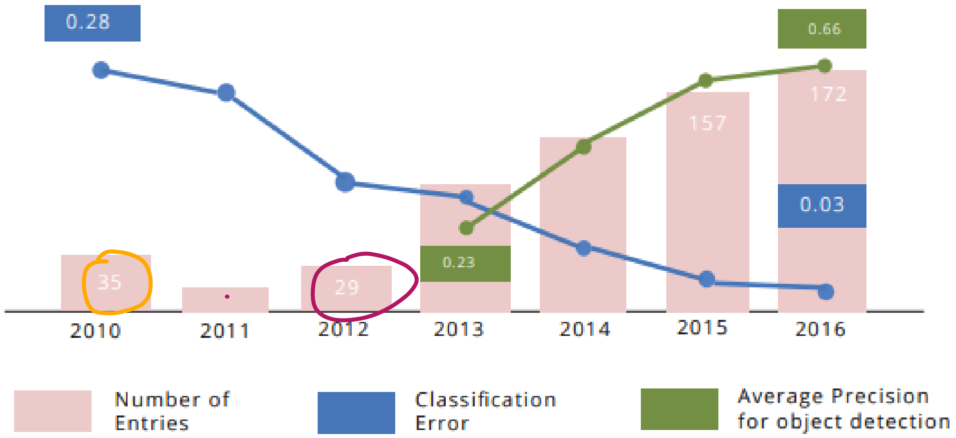
- Since 2010, the annual ImageNet Large Scale Visual Recognition Challenge (ILSVRC) is a competition where
- 1. Research teams evaluate their algorithms on given data set.
- 2. Compete to achieve higher accuracy on several visual recognition tasks



ILSVRC - EVOLUTION SINCE 2010

Andrej Karpathy — 5.8%



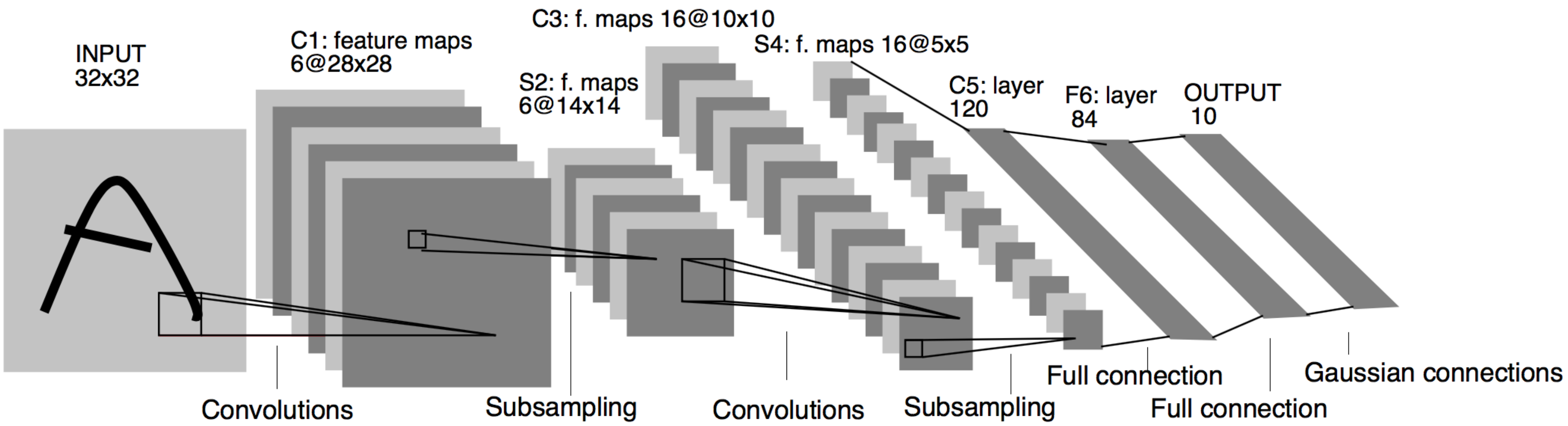


POPULAR CNN ARCHITECTURES

- LeNet
- AlexNet
- VGGNet
- GoogleNet
- Resnet

LENET

- Very first convolutional neural networks.
- Pioneered by Yann LeCun .
- Used mainly for character recognition tasks such as reading zip codes, digits, numbers on bank cheque etc



ADVANTAGES OF LENET

- LeNet architecture is simple.
- Good at detecting hand written digits & characters.

DISADVANTAGES OF LENET

- It’s not good for large data set.
- Doesn’t perform good on large number of classes.
- It doesn’t perform well on high resolution images.

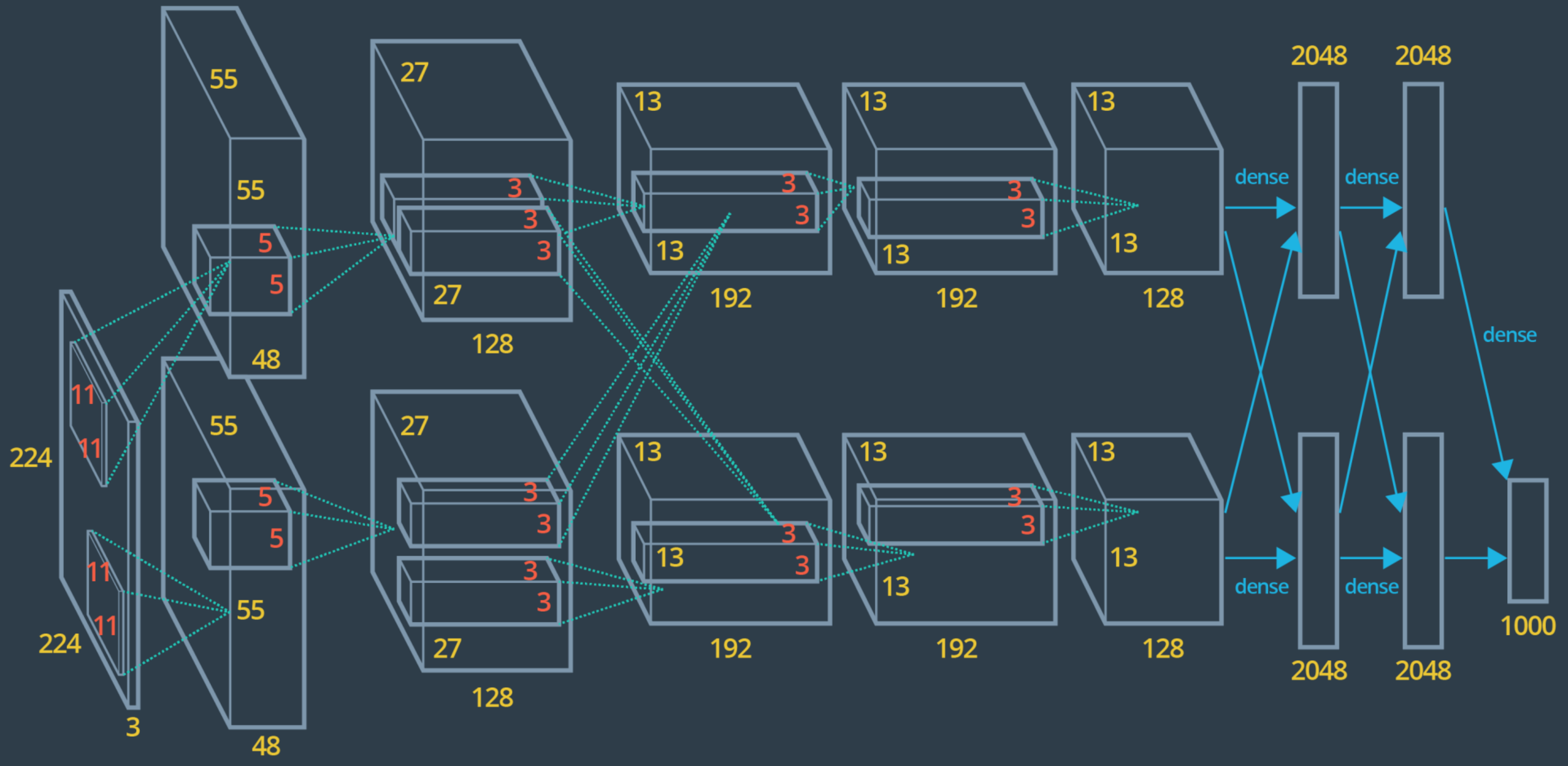
ALEXNET

- First Deep CNN.
- Winner of ImageNet LSVRC-2012

Geoffery Hinton

Pioneers of AlexNet: Alex Krizhevsky, Ilya Sutskever, Geoffrey Hinton

AlexNet



224
224
3
11
11
11
11
55
55
48
55
55
48
5
5
5
5
27
27
128
27
27
128
3
3
3
3
3
3
3
3
13
13
192
13
13
192
13
13
192
13
13
192
13
13
128
13
13
128
dense
dense
dense
dense
dense
2048
2048
2048
2048
1000

FEATURES OF ALEXNET

- 10 layers with 5 Conv,2 Maxpool and 3 FC layers.
- It uses 11x11,5x5 and 3x3 filter size.
- ReLU was used for the first time.
- Used dropout instead of ~~regularization~~ to deal with overfitting.
- Training time was doubled with the dropout rate of 0.5.
- Heavy data augmentation due to different image processing techniques.

ADVANTAGES OF ALEXNET

- Faster to train because it uses ReLU activation function.
- Uses Dropout which helps in overcoming over-fitting.

DISADVANTAGES OF ALEXNET

- Being a large network it produces 20 million trainable parameters which are computationally expensive.

$p=0$

$s=1$

VGGNET

$28\times28 \otimes 7\times7 \Rightarrow 22\times22 \rightarrow 50$

$28\times28 \otimes 5\times5 \Rightarrow 24\times24 \rightarrow 26$

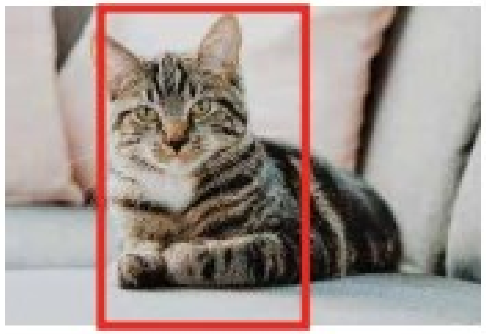
$28\times28 \otimes 3\times3 \rightarrow 26\times26 \otimes 3\times3 \rightarrow 24\times24 \otimes 3\times3 \rightarrow 22\times22$ — 30

- VGGNet is a neural network that performed very well in the Image Net Large Scale Visual Recognition Challenge (ILSVRC) in 2014.

$$FM = \frac{I - F + 2P}{S} + 1$$

CAT

1st Runner-Up (Image Classification)



CAT

Winner (Localization)

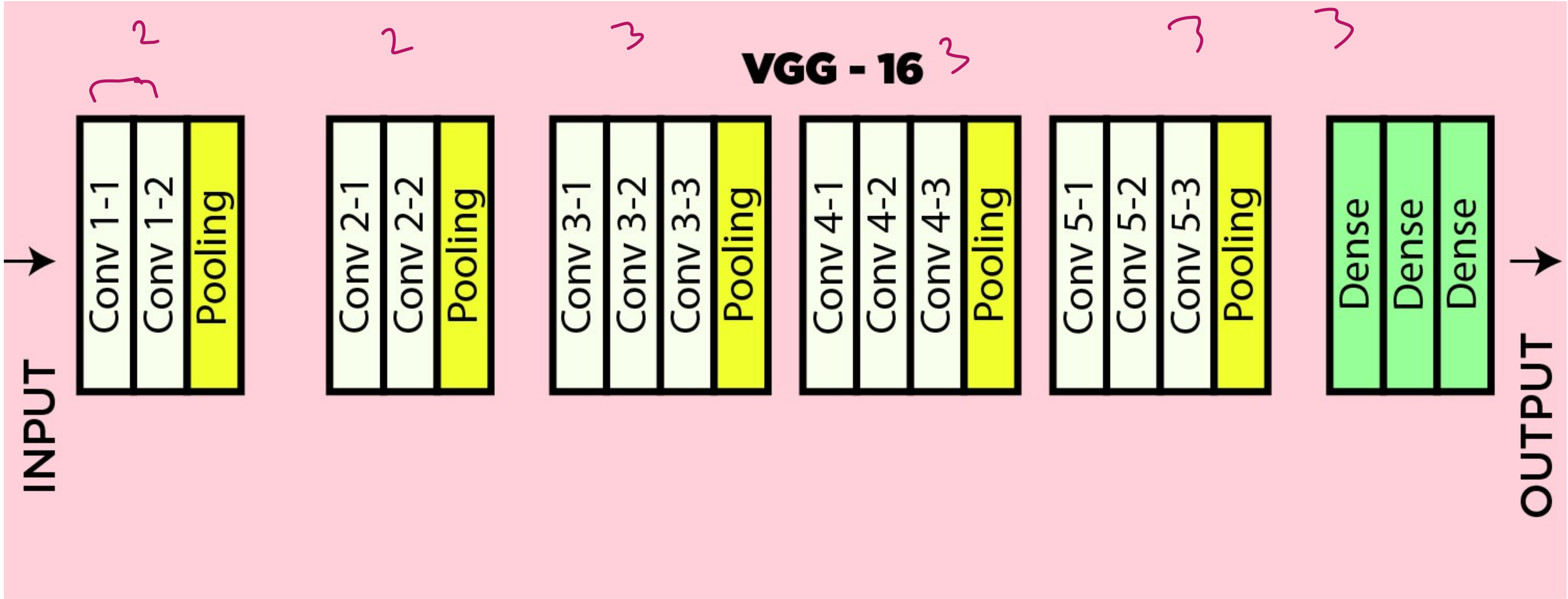
→ ① Cost - computationally expensive

② Vanishing Gradient

WHO DEVELOPED VGGNET

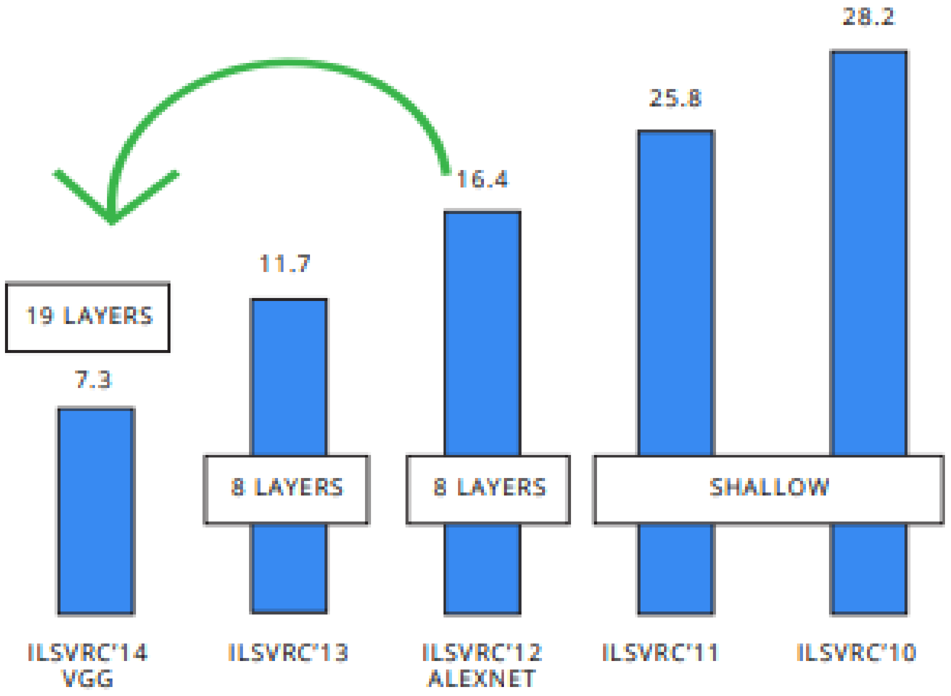
- VGGNet is invented by VGG (Visual Geometry Group),an academic group (Simonyan and Zisserman) focused on computer vision from University of Oxford.

VGG 16



VGGNet consists of 16 Convolutional layers and is very appealing because of its simplicity.

VGG - ERROR RATE IN ILSVRC 2014



Though VGGNet is the first runner-up, not the winner of the ILSVRC 2014 in the classification task, which has significant improvement over ZFNet (The winner in 2013) and AlexNet (The winner in 2012).

ADVANTAGE OF VGG

- Very good architecture for benchmarking image classification/ localization.
- Pre-trained networks for VGG are freely available on the internet, so it is commonly used for various applications.

DISADVANTAGE OF VGG

- Slow to train if trained from scratch. Even on a decent GPU, it takes more than a week to get it to work.

RESNET

$w_n = w - LR \boxed{\frac{\partial E}{\partial w}}$

0.01 → 0.002

Vanishing Gradient

$\boxed{\frac{\partial E}{\partial w} < 1}$

0.01

$w_n = w - 0.01 \times 0.002$

$\boxed{w_n = w - 0.00002}$

4

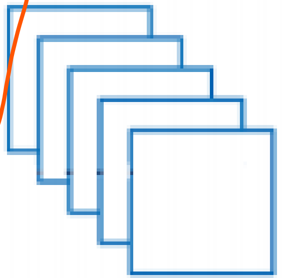
5000

200

1314

1314.000

Winner of all classification and detection competition in ILSVRC 2015 and COCO 2015



Ultra deep network with **152 layers**. Almost 7x more than GoogLe Net.

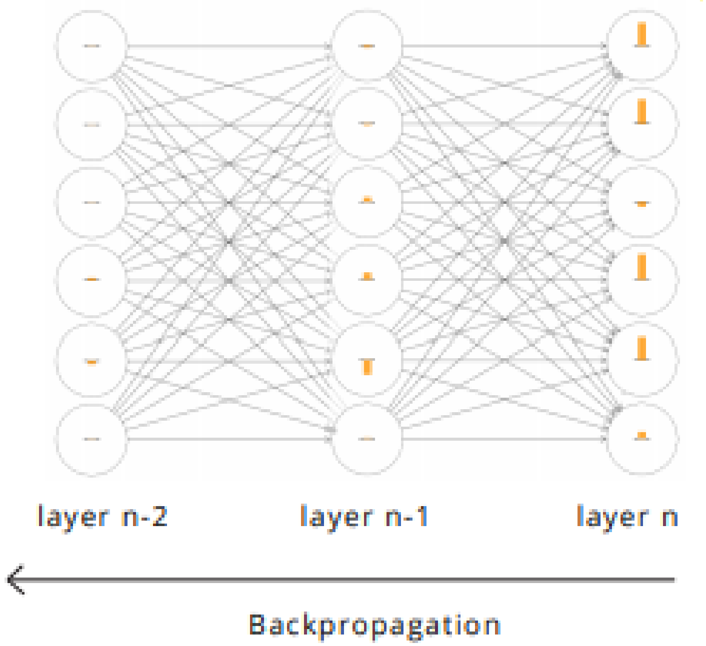
Best accuracy levels Better than "human performance"

0.04 0.5 3 7 12 21 34 62 78 100

E

ISSUES WITH DEEPER NETWORKS

- During the backpropagation process, the gradient update decreases as we go deeper in the network.

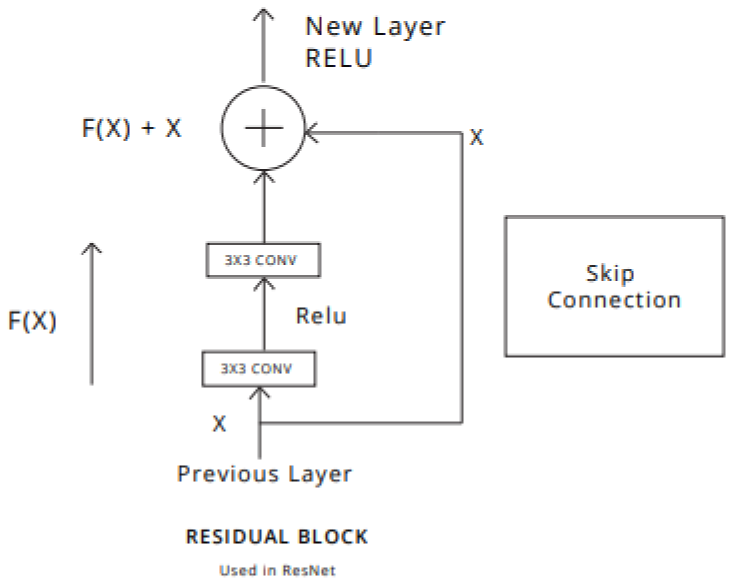


In deeper layers, the gradient update reduces to zero and the layers don't learn new weights.

This is why deeper models don't perform as good as shallow model.

HOW TO RESOLVE VANISHING GRADIENT PROBLEM

- Skip connections are shortcut connections which skip few convolutional layers.

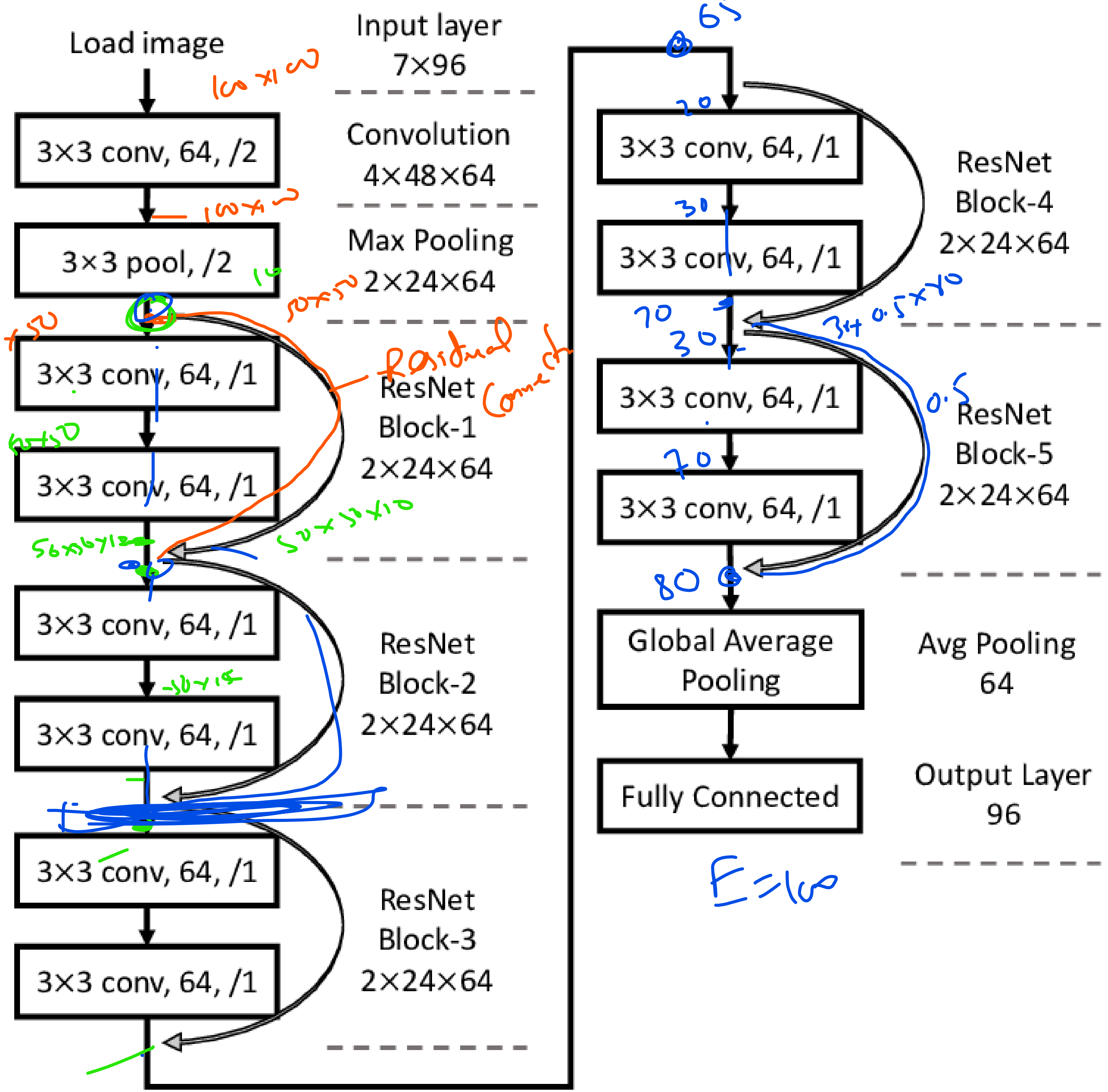


Residual Connections

RESNET

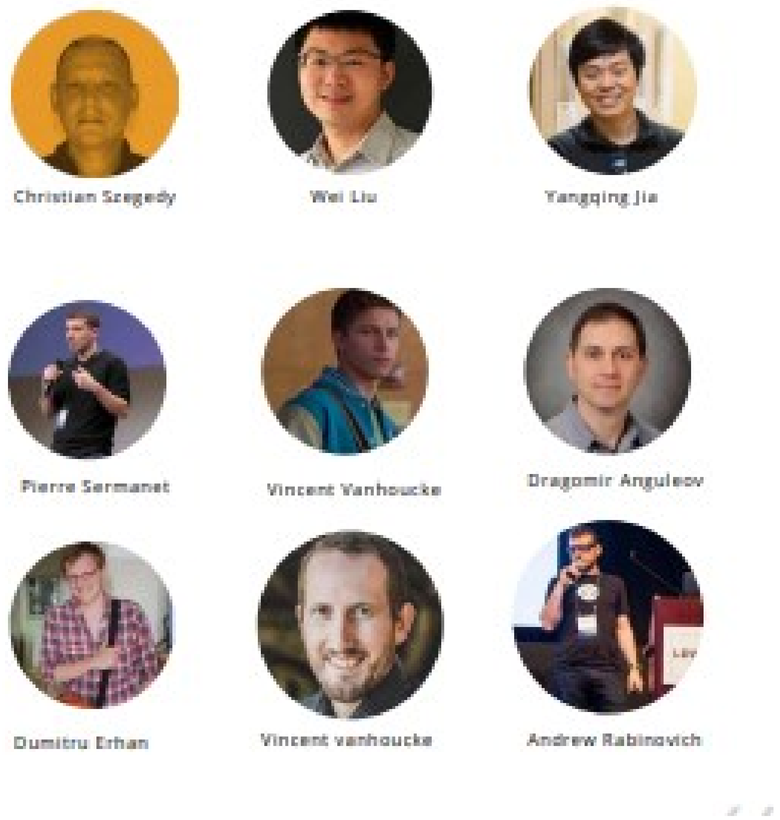
- ResNet architecture is a combination of Residual Blocks.
- You can stack more blocks to make it deeper.

Batch Normalization



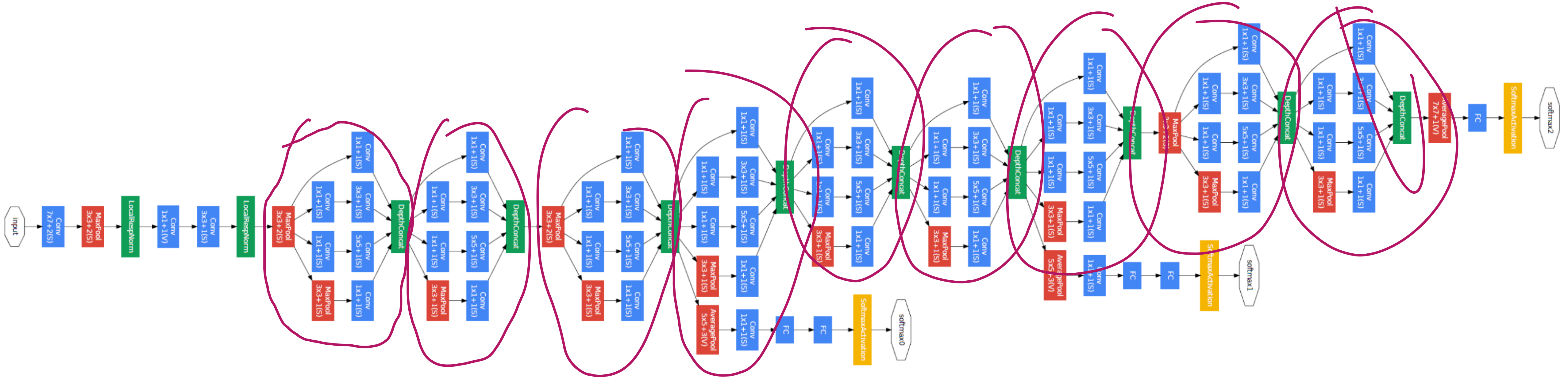
GOOGLENET

- GoogleNet was the winner of ImageNet 2014, where it proved to be a powerful model.
- GoogleNet (or Inception Network) is a class of architecture designed by researchers at Google.

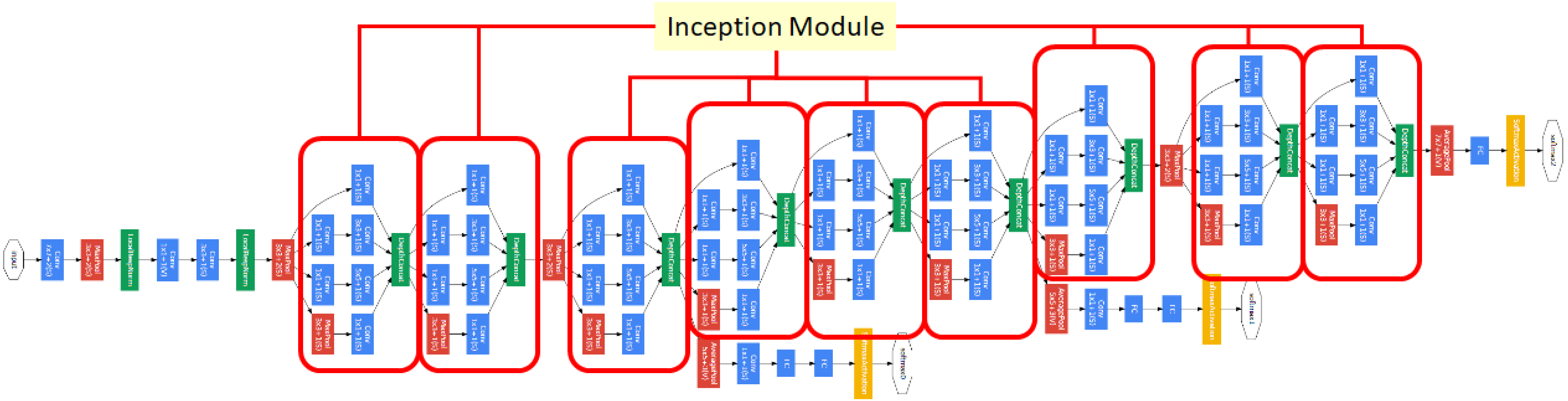


GOOGLENET

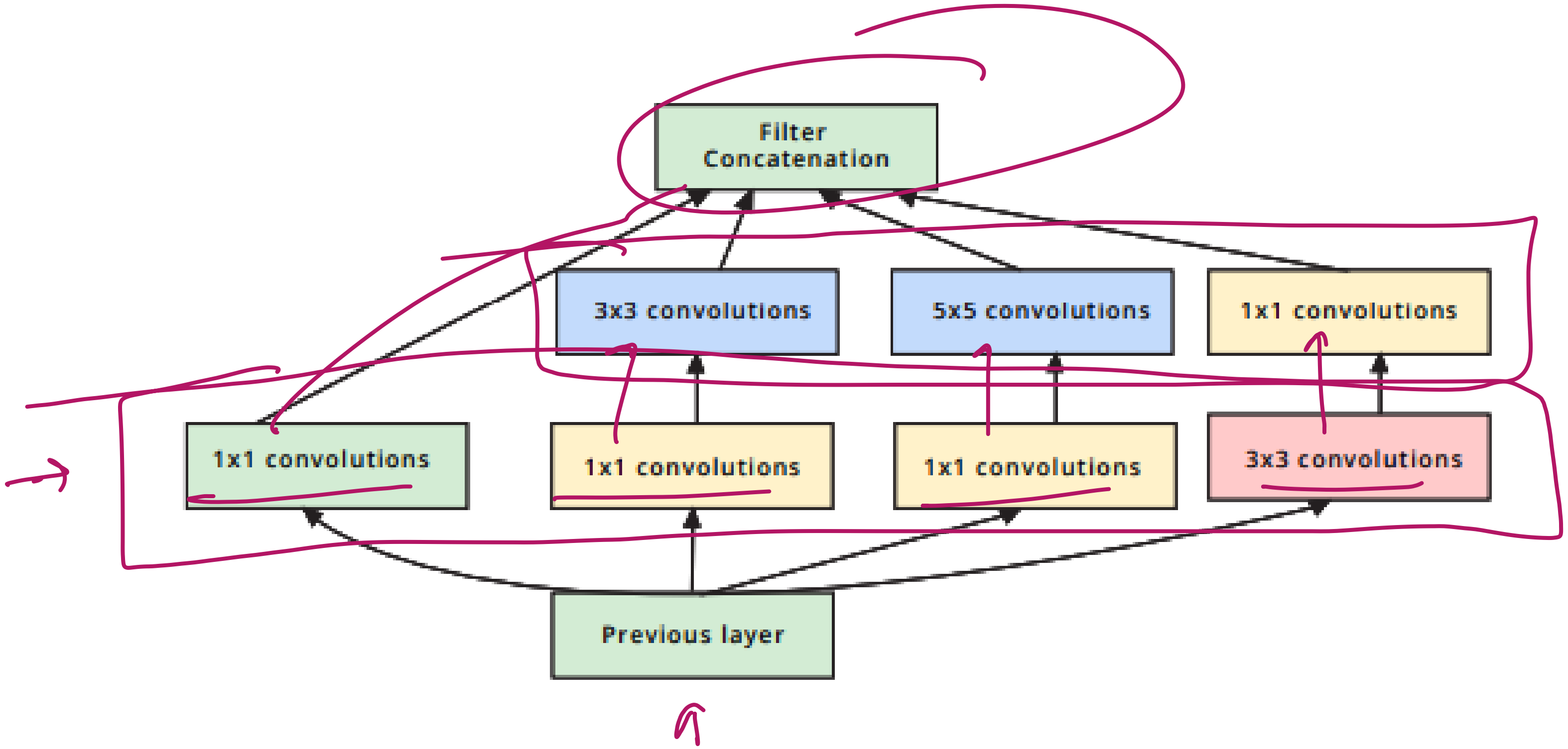
- Consists of a 22 layers deep CNN
- 12 times fewer parameters than AlexNet
- High Performance(top-5 error rate 6.77%)



INCEPTION

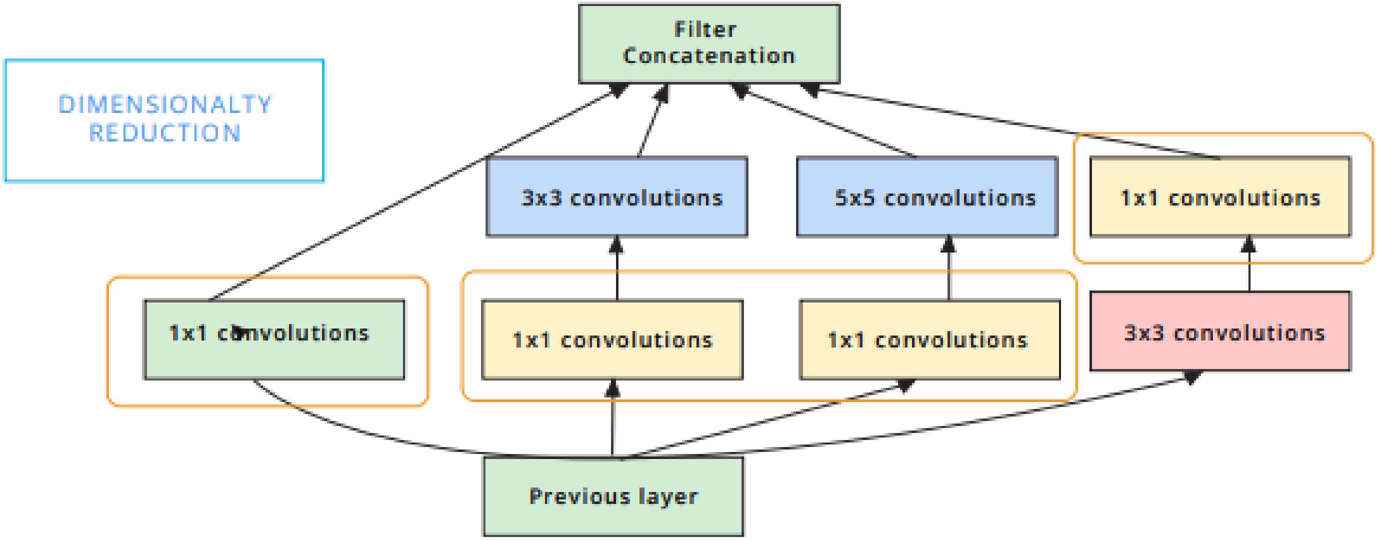


INCEPTION (NETWORK IN A NETWORK)



Image

GOOGLENET (INCEPTION)



MORE READS

- https://www.datacamp.com/community/tutorials/transfer-learning