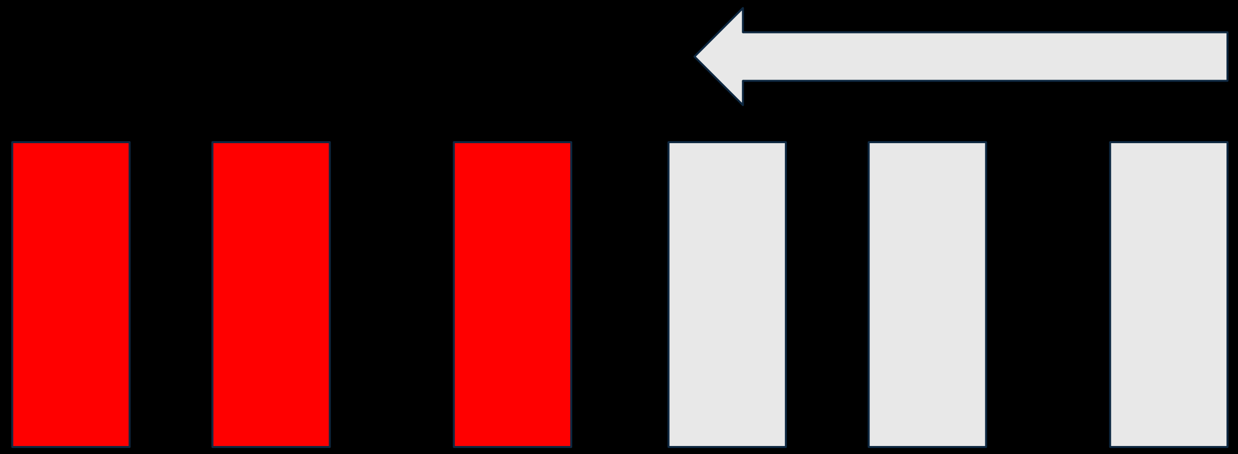
AIL 862

Lecture 4

Fine Tuning

Further train particular layers of the network

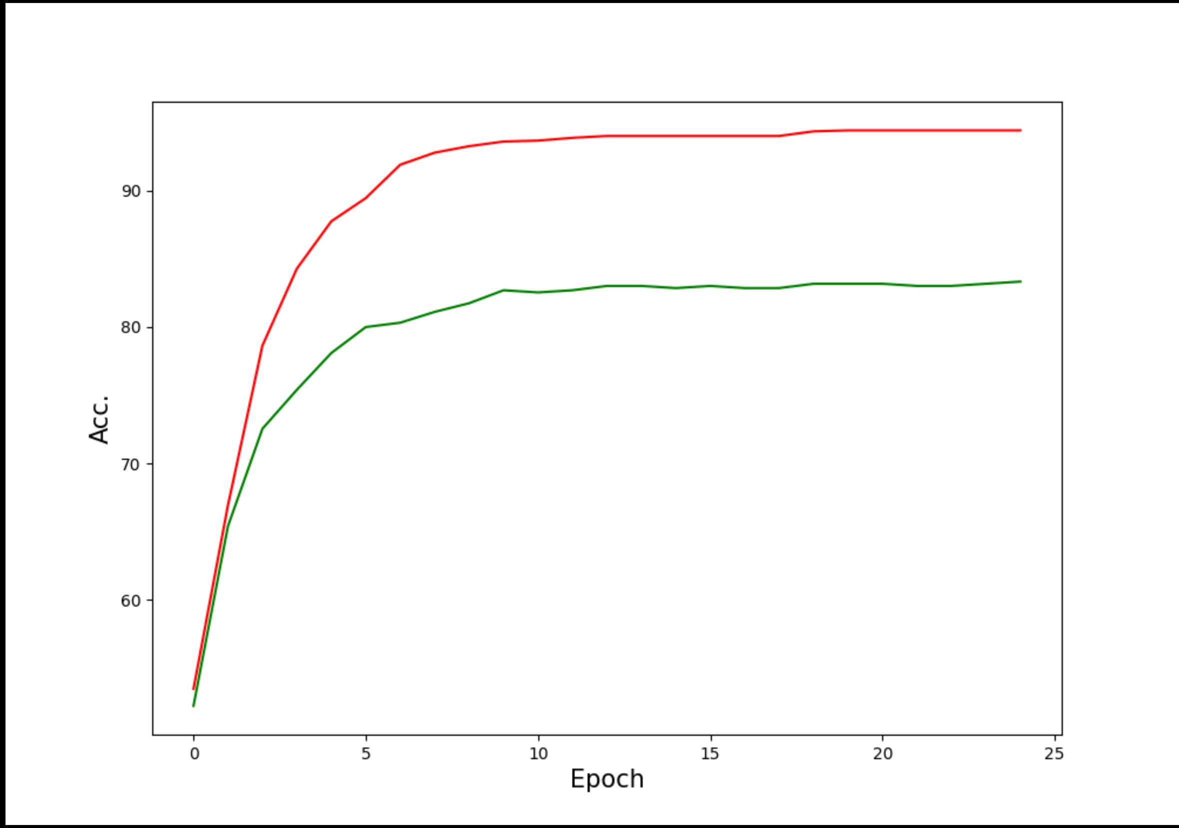
Red layers are frozen.



Fine tuning

- Pre-training + finetuning does not always work.

- Generally: initial layers are somewhat domain specific.



Training UC Merced (which is optical dataset) starting from a SAR-trained model

Using Models Trained in Supervised Fashion For Some Other Task

- ***Just use as feature extractor***

  Just use the features extracted from particular layer(s) without any tuning

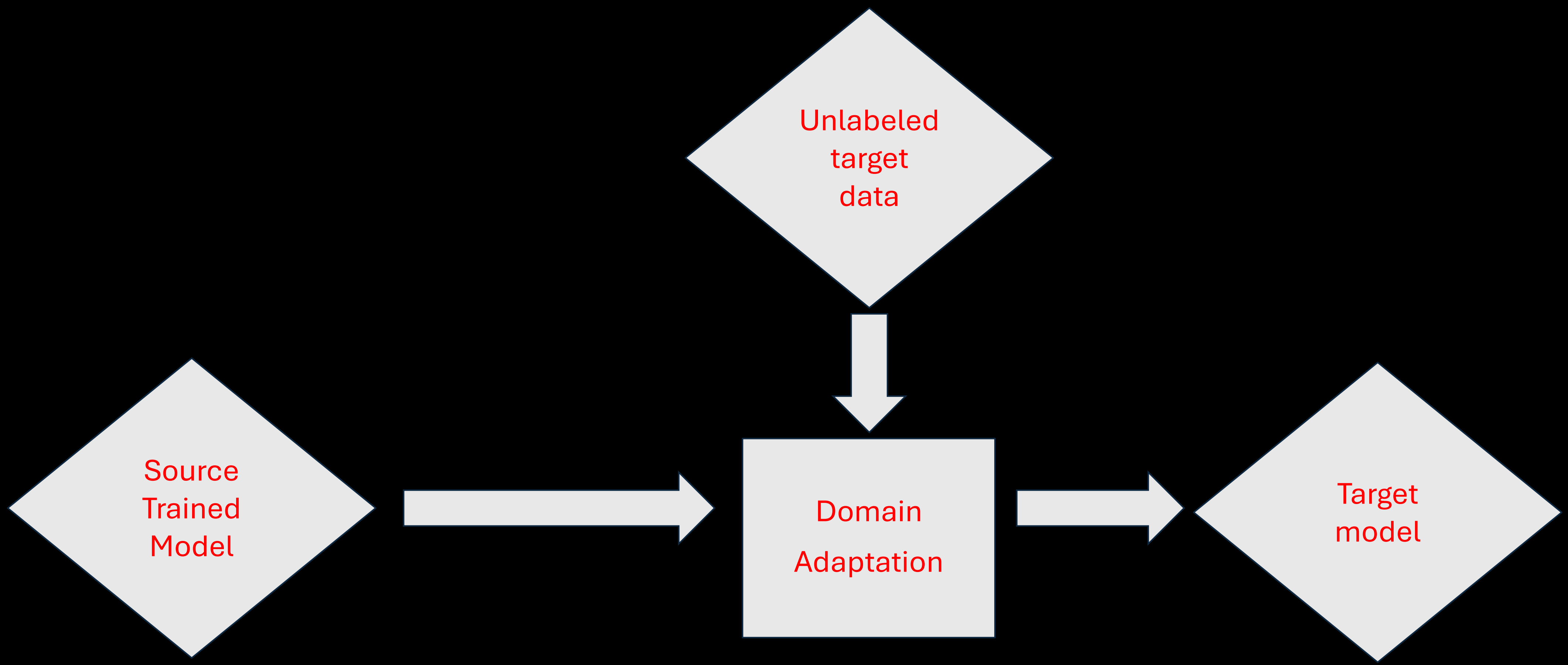
- ***Fine tuning on target data***

  Further train particular layers of the network

- ***Unsupervised domain adaptation***

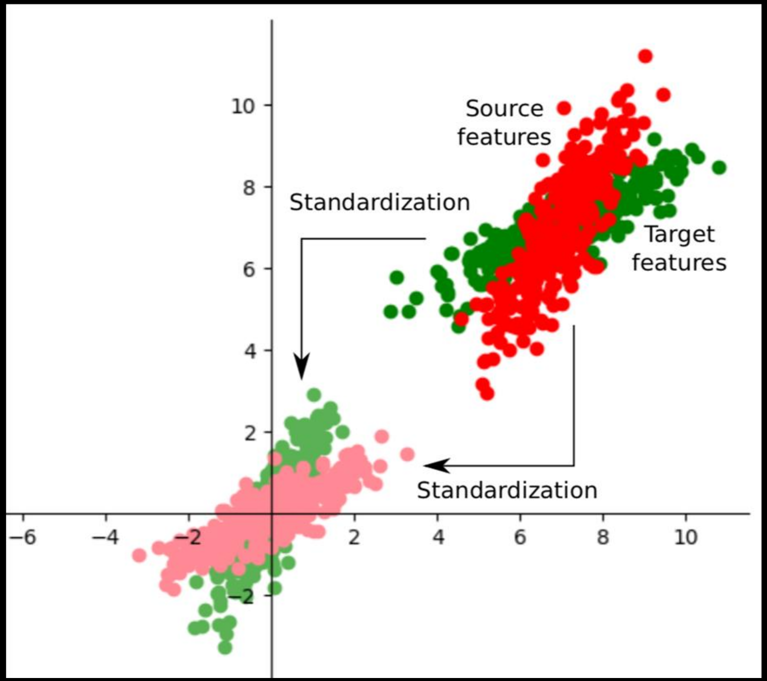
  Adapt the network with unlabeled target data

Unsupervised Domain Adaptation



Unsupervised Domain Adaptation

Batch Normalization based DA methods align feature distributions through feature standardization by setting mean of features to 0 and variance to 1.



Ack: Subhankar Roy, UniTn for this slide

Batch Normalization

- Generally, running mean and variance are estimated during training.

- However, BN (for domain adaptation) suggests to estimate the above from test time minibatches.

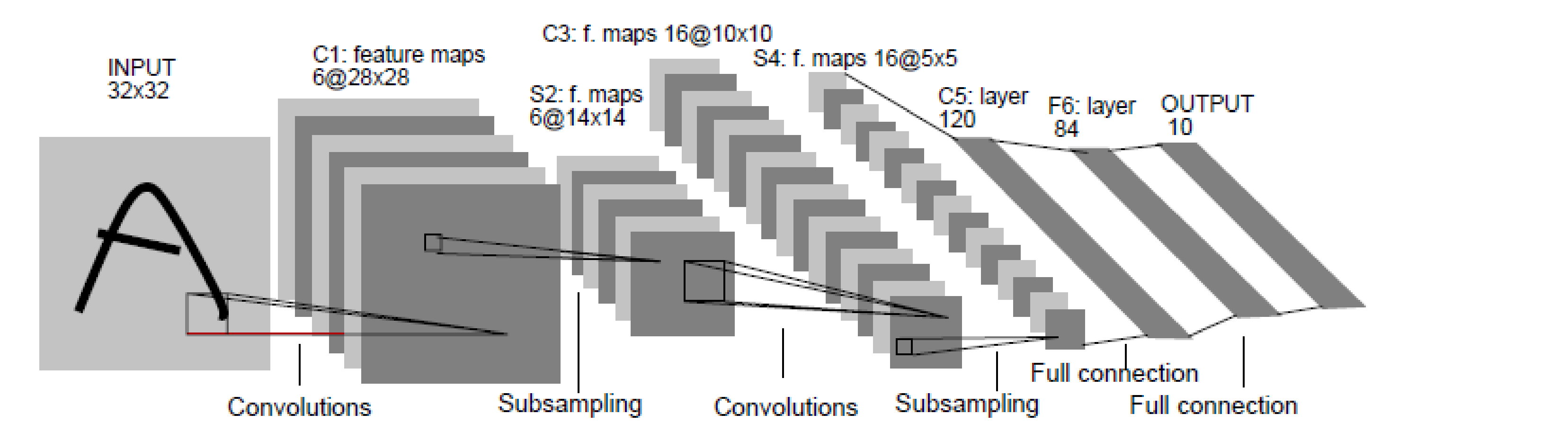
Unsupervised Domain Adaptation

- Domain Translation (GAN? However, why not used so much?)

Unsupervised Domain Adaptation

- Domain Confusion / domain adversarial training

LeNet


INPUT
32x32
C1: feature maps
6@28x28
S2: f. maps
6@14x14
C3: f. maps 16@10x10
S4: f. maps 16@5x5
C5: layer
120
F6: layer
84
OUTPUT
10
Convolutions
Subsampling
Convolutions
Subsampling
Full connection
Full connection

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	X				X	X	X			X	X	X	X		X	X
1	X	X				X	X	X			X	X	X	X		X
2	X	X	X				X	X	X			X		X	X	X
3		X	X	X			X	X	X	X			X		X	X
4			X	X	X			X	X	X	X		X	X		X
5				X	X	X			X	X	X	X		X	X	X

TABLE I

Each column indicates which feature map in S2 are combined by the units in a particular feature map of C3.

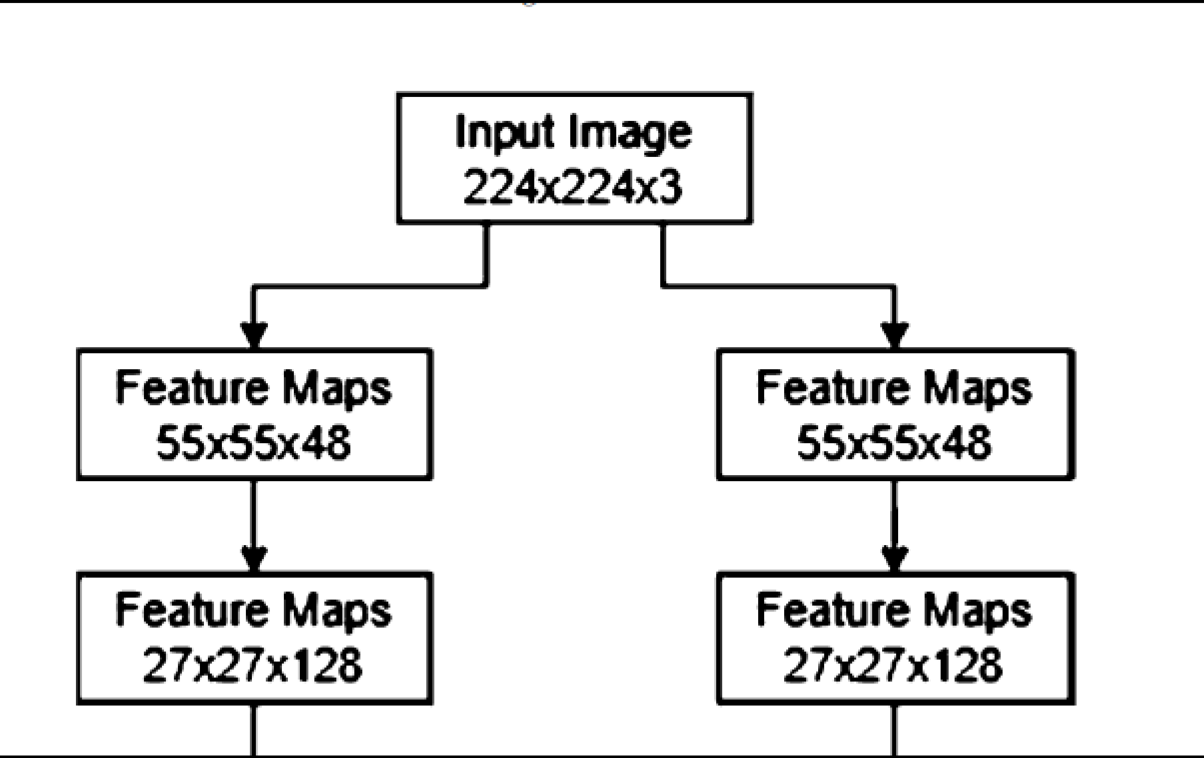
ImageNet Large Scale Visual Recognition Challenge 2010 (ILSVRC2010)

- 1000 object categories

- Descriptor Coding + SVM, 0.28 --- NEC-UIUC
- Fisher kernel + SVM, 0.34 --- XRCE
- LI2C, 0.58 --- NTU_WZX

https://image-net.org/static_files/files/pascal_ilsvrc.pdf

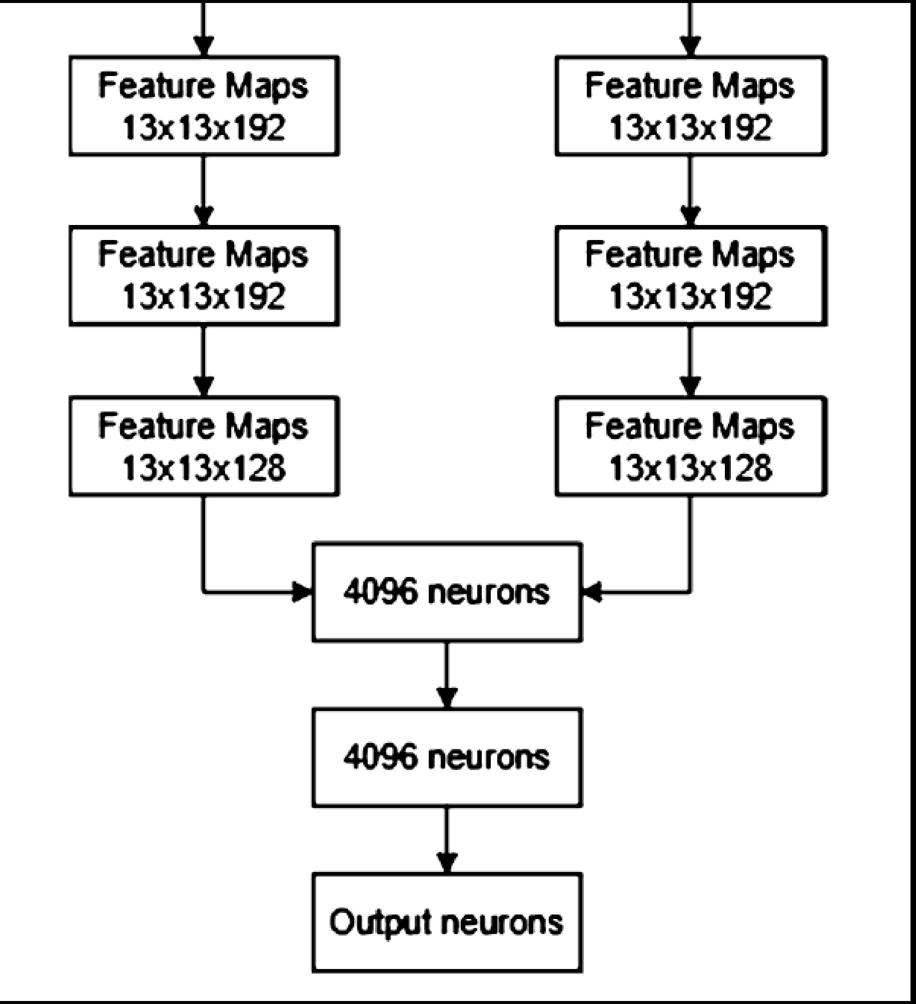
AlexNet


Input Image
224x224x3
Feature Maps
55x55x48
Feature Maps
55x55x48
Feature Maps
27x27x128
Feature Maps
27x27x128

AlexNet

The architecture of AlexNet is divided into two distinct parts, a design choice influenced by the limited memory capacity of GPUs. As a result, the network's architecture was split across two GPUs, with each GPU managing half of the neurons or feature maps in specific layers.

AlexNet


Feature Maps 13x13x192
Feature Maps 13x13x192
Feature Maps 13x13x192
Feature Maps 13x13x192
Feature Maps 13x13x128
Feature Maps 13x13x128
4096 neurons
4096 neurons
Output neurons

AlexNet

Five convolutional and three fully connected layers

The first convolutional layer uses 96 kernels of spatial size 11×11. These large filters allowed the network to capture complex and high-level features from images.

The subsequent convolutional layers use kernels of size 3×3.

AlexNet further used

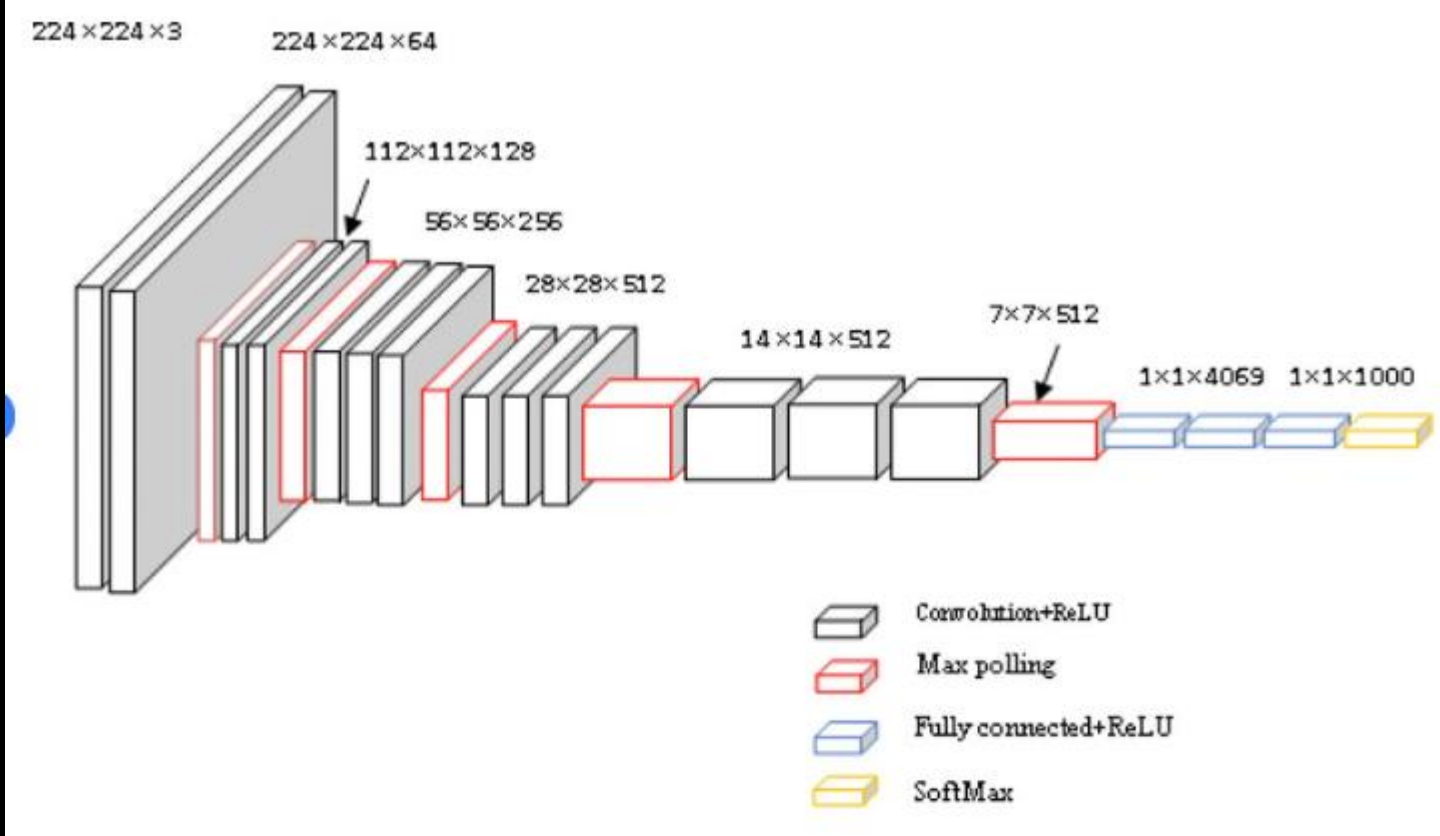
✓DropOut
✓ReLU
✓Max Pooling

- AlexNet advocated for extensive data augmentation as a way to reduce ovefitting, a practice that has since then become a norm.

VGGNet

- Unlike AlexNet, VGGNet promoted the use of 3×3 convolution filters. It suggested that maintaining a consistent convolution filter size across all layers and simply increasing the depth of the CNN could improve performance on visual recognition tasks.

VGGNet

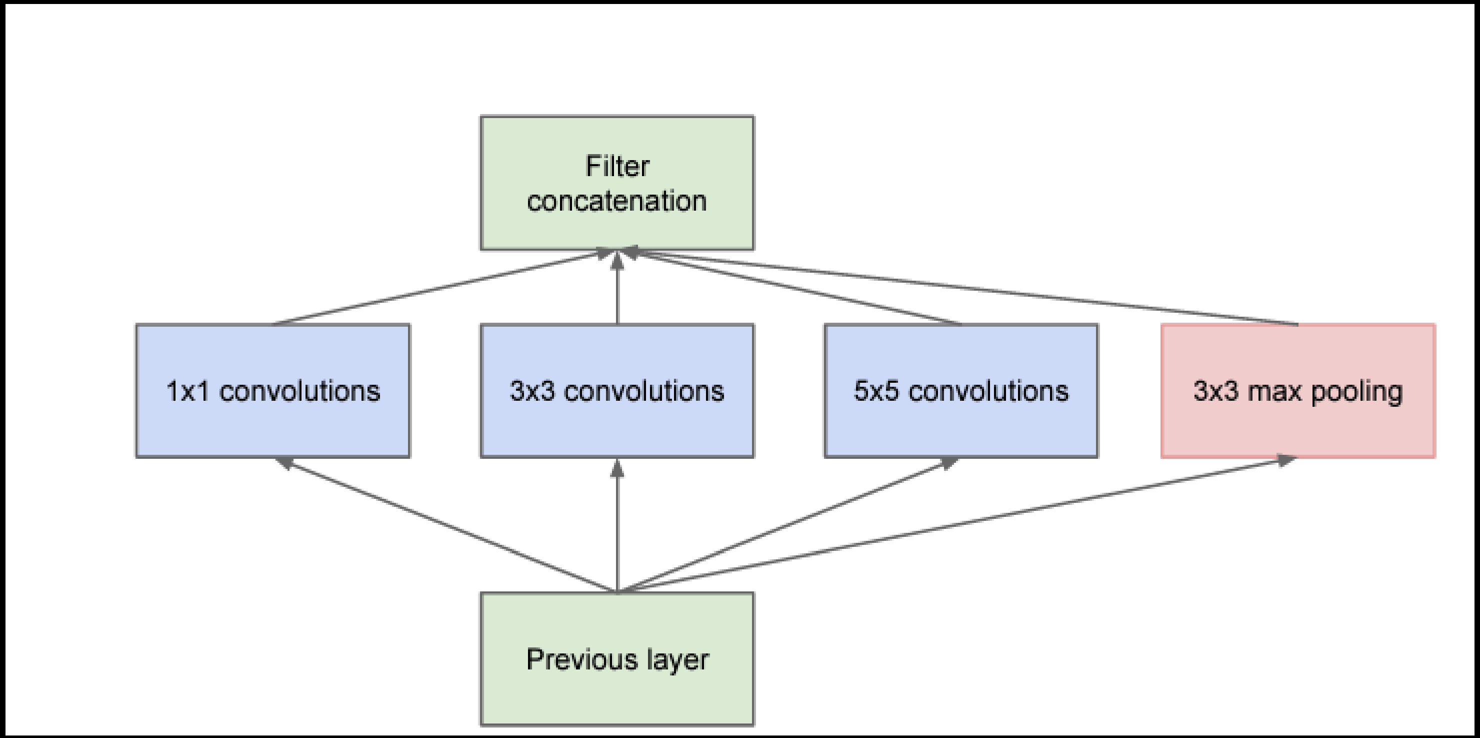


https://www.researchgate.net/figure/GG-Neural-Network-Architecture-Source_fig5_369178183

VGGNet

Their study suggests that, although VGGNet has more layers and a larger number of parameters compared to previous models, it requires fewer epochs for training. This can be partially attributed to the effect of smaller filter sizes, while pre-initialized weights also playing a significant role.

Inception Module

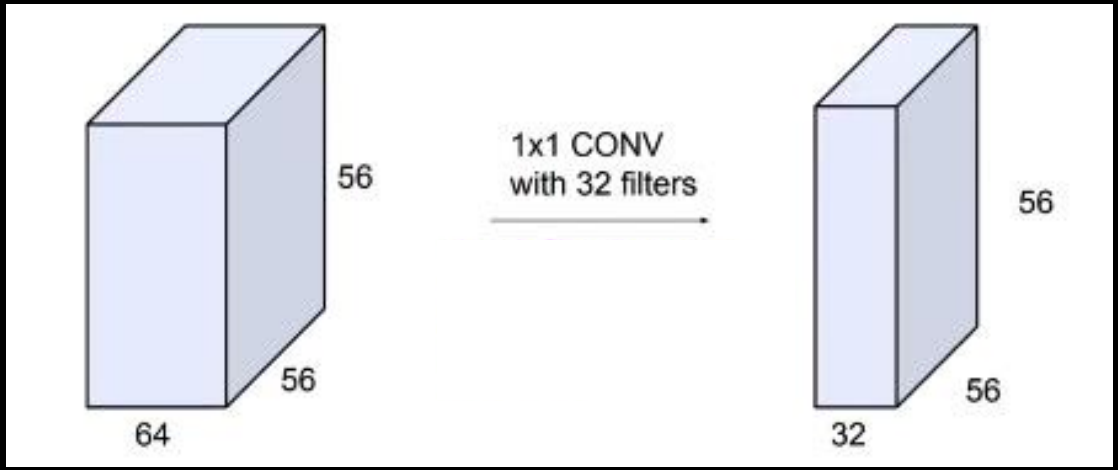


GoogleNet

- InceptionModule – authors call it inspired by "intuition of multi-scale processing"

- Uses inception module many times over the network

Project depth to lower dimension



Inception Module (Revised)