(*) Residual connections / skip connections / highway connections (contd.)

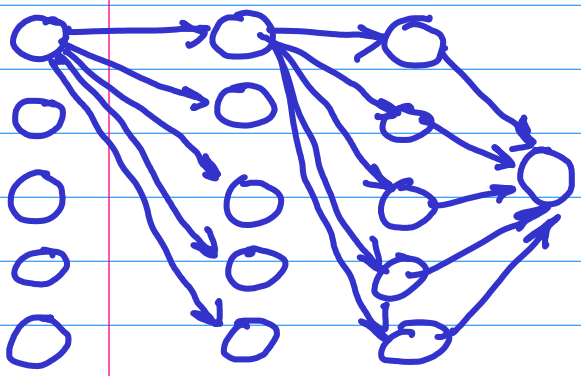
(*) The magnitude of the problem: →

- AlexNet (2012) had 5 convolutional layers
- 2014: VGG, GoogleNet
  - VGG → 19 layers
  - GoogleNet → 22 layers
- ResNet (34 → 50 layers deep architectures)
  deeper but had overall lower complexity.

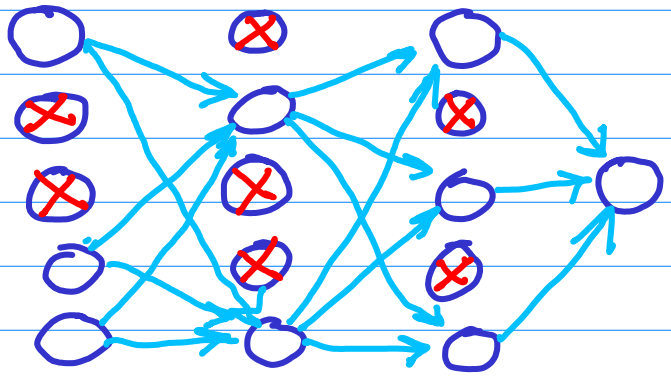
(*) 'DROPOUT' (AlexNet 2012)

Heuristic: applied at the training phase (FC layers)

Dropout is a kind of regularisation technique to reduce overfitting.



Usual FC scenario



Scenario with dropout

AlexNet : $p = 0.5$ at the first two fully connected layers

Neuron: has a probability p not to contribute to the feedforward phase & participate in the backpropagation. ⇒ Each neuron can have a larger chance to be trained, and not depend on some 'strong' neuron. No dropout at the test time

Second class of successful deep Architectures:

AlexNet (2012) and CaffeNet

↳ used 2 GPUs (two parallel paths)

↳ used 1 GPU

What & Why? ImageNet: 15-million labelled high-resolution images, with 22K categories (22,000)

2012: ImageNet Large-Scale Visual Recognition Competition

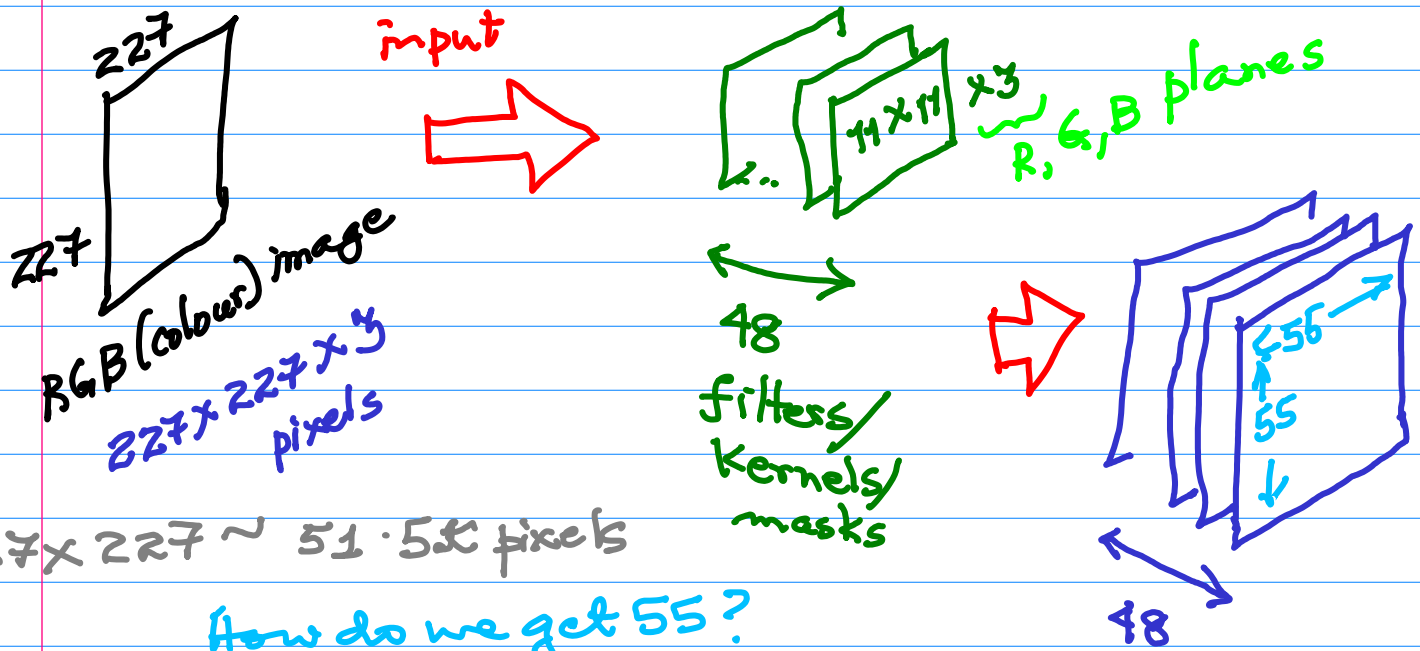
dataset: subset of ImageNet

* 1K images in each of the 1K categories/ classes.

- 1.2 million training images
- 50K validation images
- 1.5L testing images

AlexNet: 8 layers (5 Conv layers)

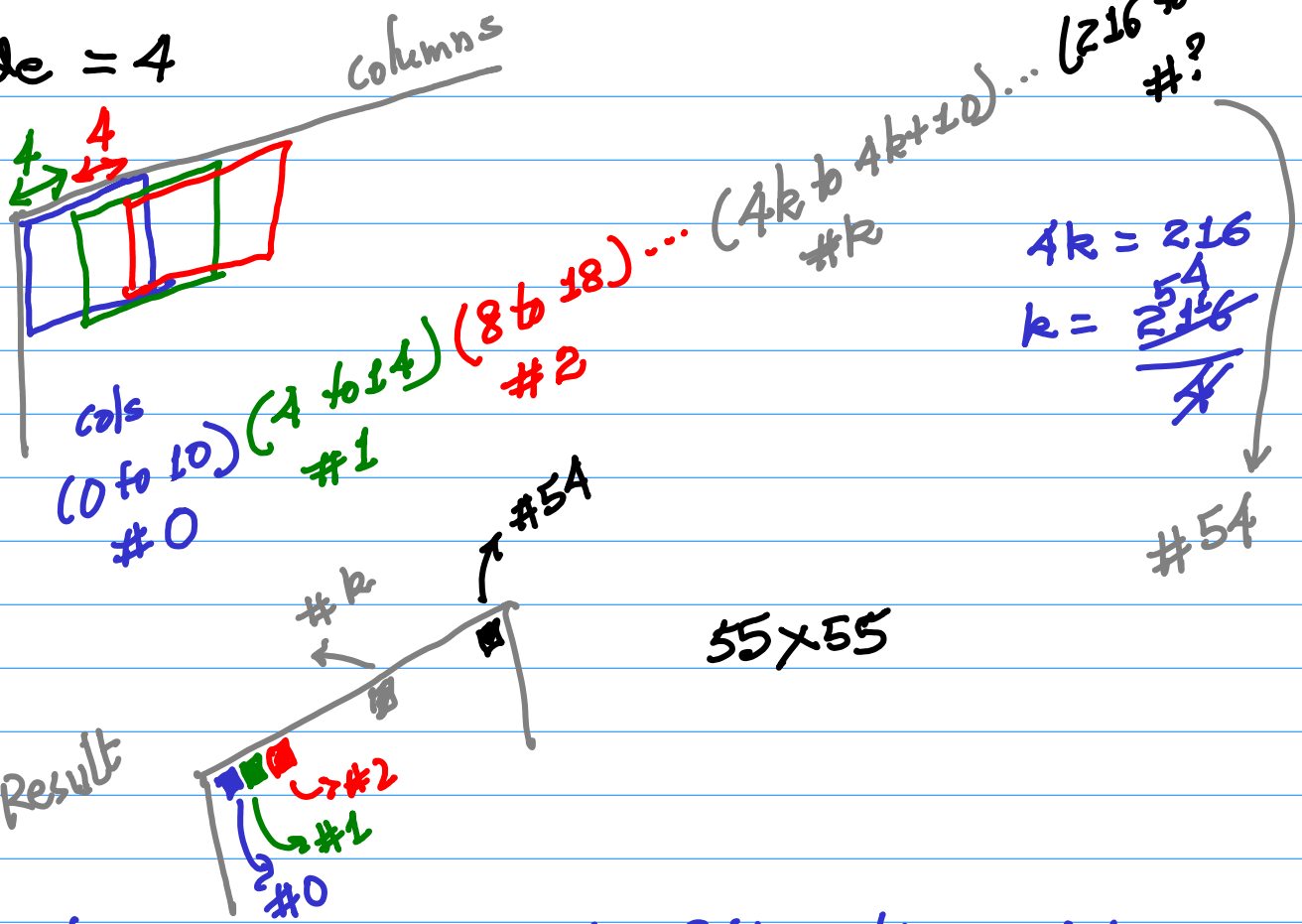
(1) 1st Convolutional layer



227 x 227 ~ 51.5K pixels

How do we get 55?

Stride = 4



AlexNet has 2 groups of 48 filters/kernels/masks each, each group given to one of the two GPUs used.

[CaffeNet is a 1 GPU version]

96 filters / kernels / masks

→ Result: 2 groups of 48 outputs, each of size 55x55

(AlexNet)