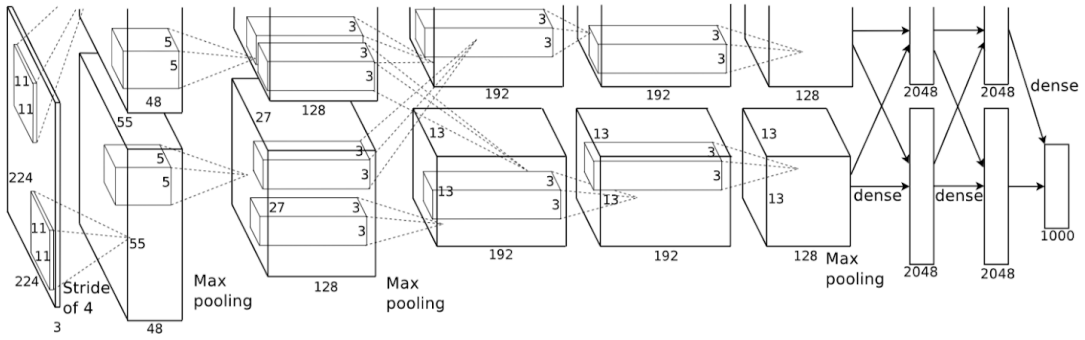
ImageNet Classification with Deep Convolutional Neural Networks, 2012

Bhavna Nagar
Parag Dudeja



Architecture of AlexNet

a) Is 224x224 input image size compatible architecture? Justify if yes, else provide the minimum correct image size for this architecture to work (3 marks)

$$\left(\frac{n+2p-f}{s}+1\right) * \left(\frac{n+2p-f}{s}+1\right)$$

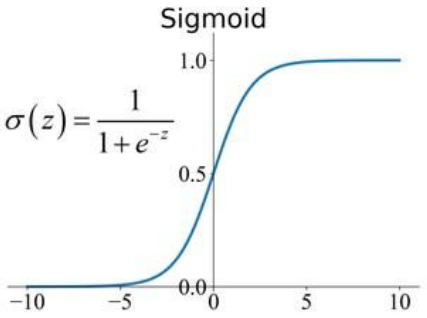
–> For n = 224, p = 0, f = 11, s=2 => 54.25 * 54.25

Wrong! Expected n = 55 in next layer

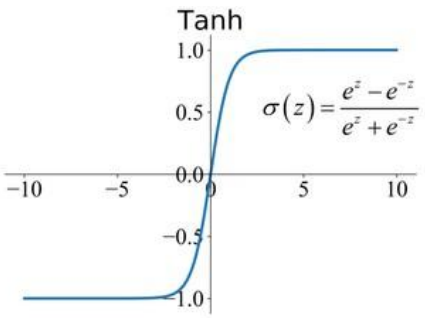
-> But For n = 227, p = 0, f = 11, s=2 => 55 * 55

Correct, as expected in next layer

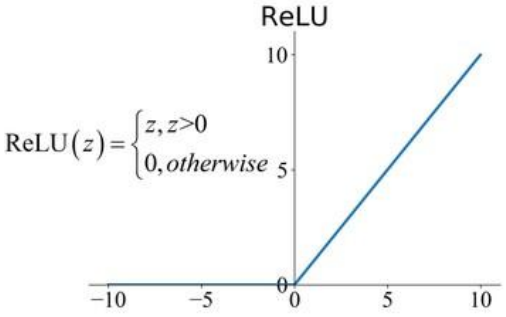
b) What difference does using non linear ReLU instead of using other standard nonlinear functions like sigmoid and tanh? (3 marks)



(a)



(b)



(c)

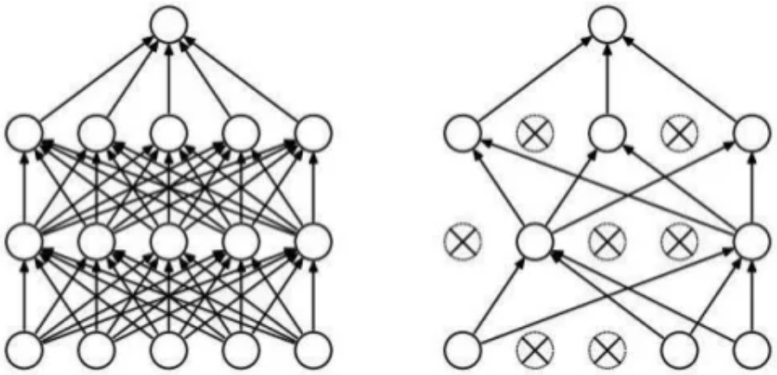
Sigmoid and tanh non linearities saturate whereas ReLU does not saturate on taking derivative at extreme values

Thus, ReLU makes it quick to train the model as compared to sigmoid and tanh

c) Explain the dropout technique used in AlexNet? (4 marks)

Dropout is a technique used to reduce overfitting and works in the following way

1. Zero the output of each hidden neuron with probability 0.5.
2. No longer contribute to forward pass and backward propagation
3. Neural network samples a different architecture every time
4. Reduce complex co-adaptations of neurons
5. Used in two fully-connected layers



(a) Standard Neural Net (b) After applying dropout.

d) How do you calculate the top-1 and top-5 accuracy? Explain with an example (4 marks)

Top-1 accuracy is the conventional accuracy: the model answer (the one with highest probability) must be exactly the expected answer.

Top-5 accuracy means that *any* of your model 5 highest probability answers must match the expected answer.

For instance, let's say you're applying machine learning to object recognition using a neural network. A picture of a cat is shown, and these are the outputs of your neural network:

- Tiger: 0.4
- Dog: 0.3
- Cat: 0.1
- Lynx: 0.09
- Lion: 0.08
- Bird: 0.02
- Bear: 0.01

Using top-1 accuracy, you count this output as **wrong**, because it predicted a tiger.

Using top-5 accuracy, you count this output as **correct**, because cat is among the top-5 guesses.