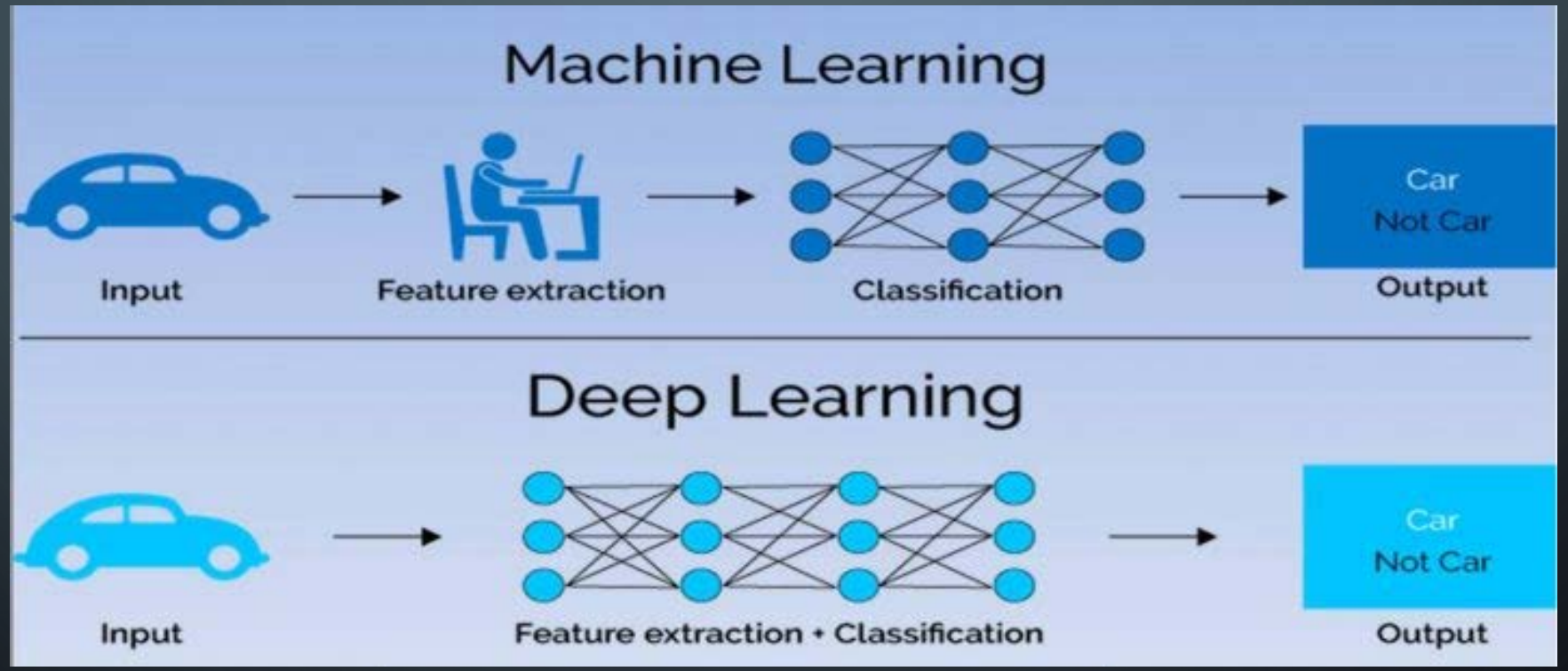
CONVOLUTIONAL NEURAL NETWORKS

MOHAMMAD GHODDOSI

MLP FOR IMAGES

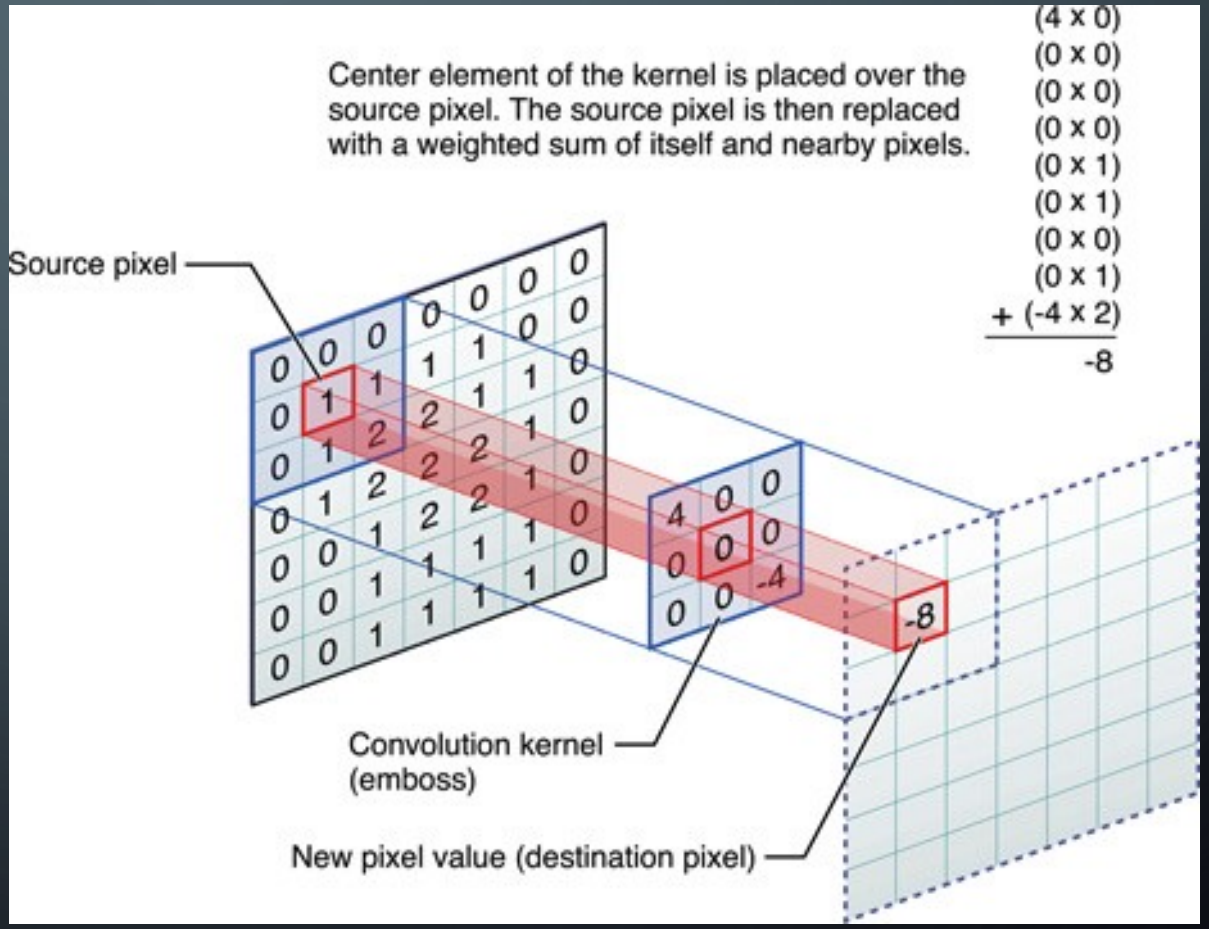
- MLP downsides for images
  - Local features
  - Position invariant
  - Big weight matrix ($n_{in} * n_{out}$)
  - Overfitting

DEEP LEARNING

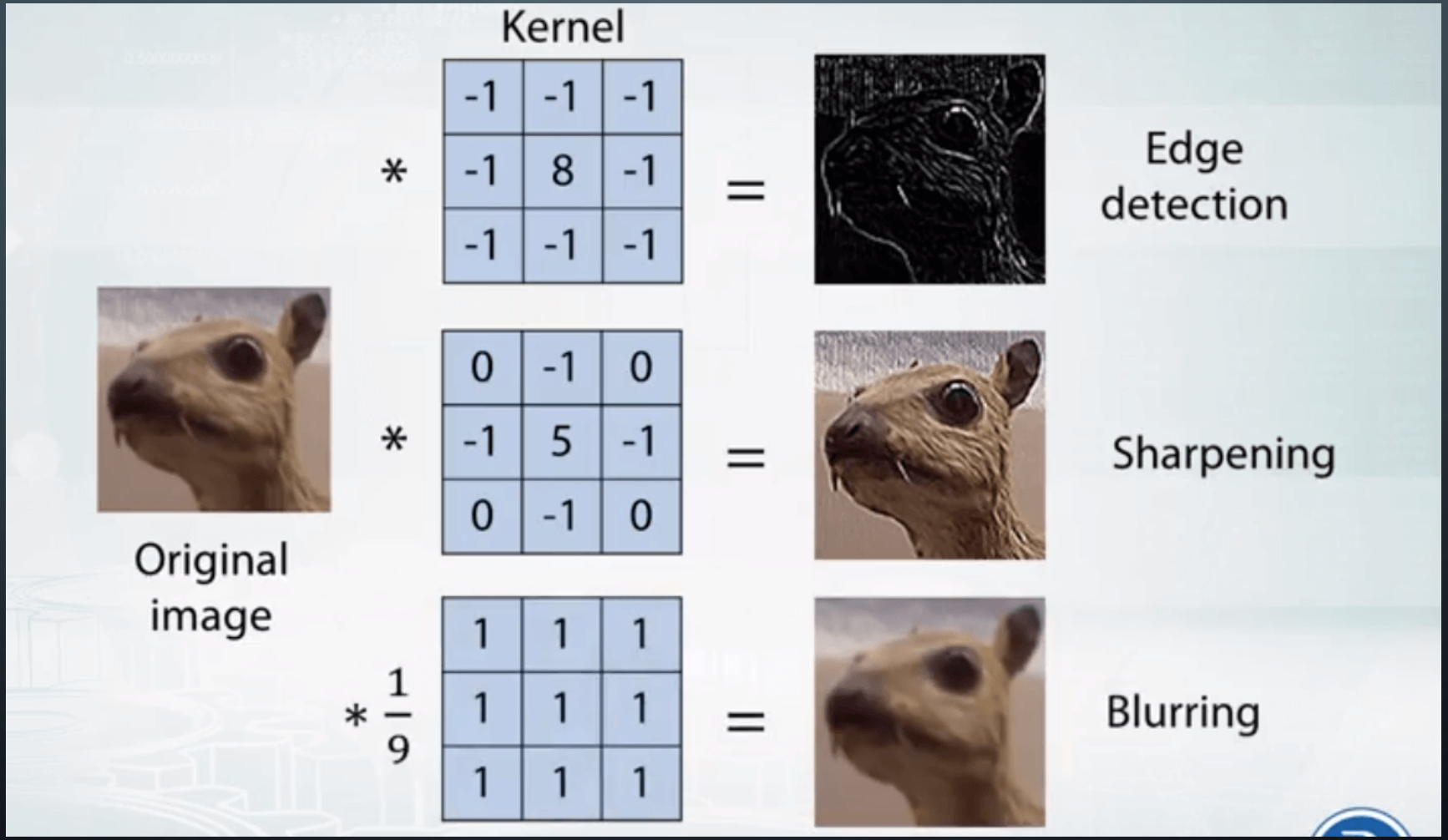


CONVOLUTION (2D)

- Cross-Correlation
- Output size
- Padding
- Stride
- $n_{out} = \left\lfloor \frac{n_{in}+2p-k}{s} \right\rfloor + 1$

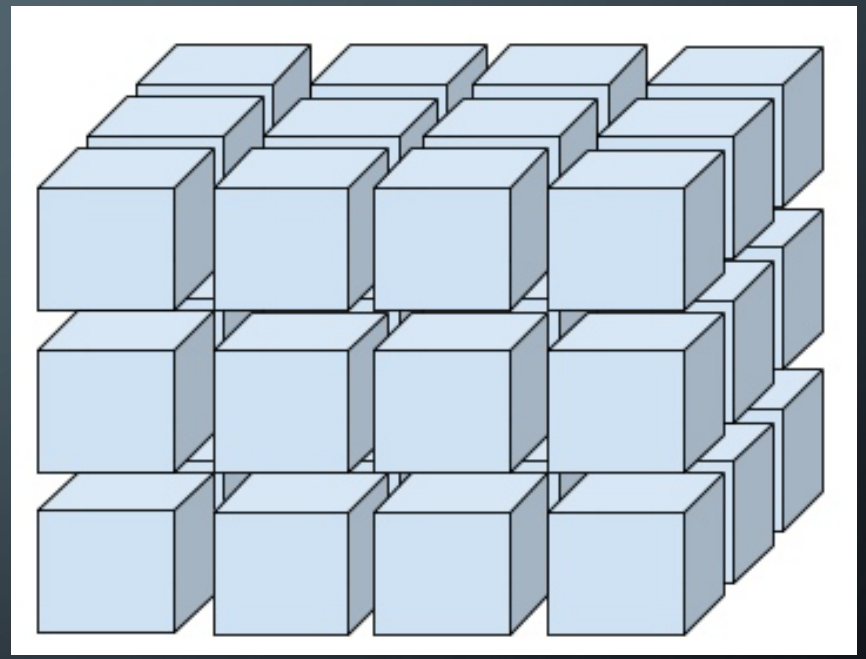


EXAMPLE (EDGE DETECTION)



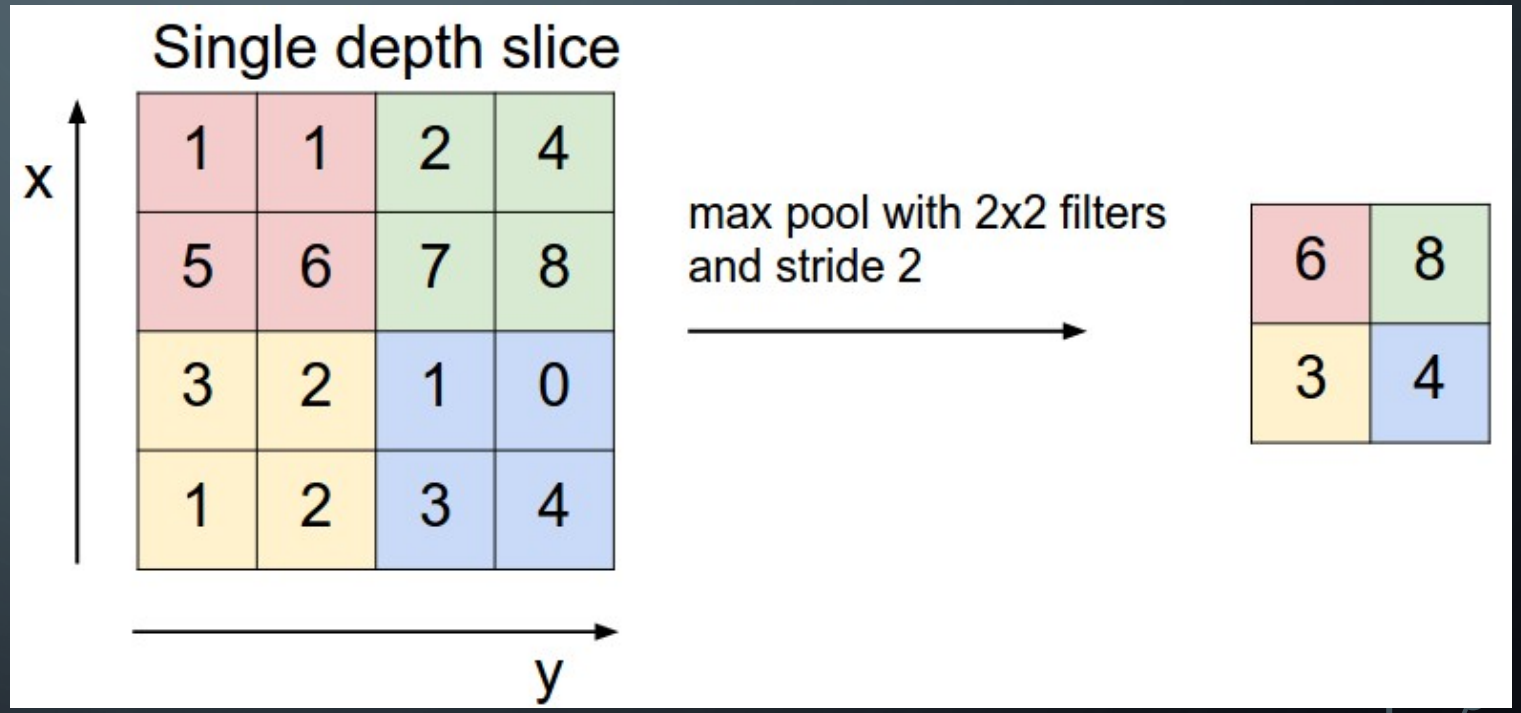
CONVOLUTION LAYER

- Convolves a filter
- Learns filters weight tensor
- Local processing
- Position invariant
- Small weight ($f * k * c_{in}$)
- Rank 3+ tensor
- Time consuming $O(WHk^2)$

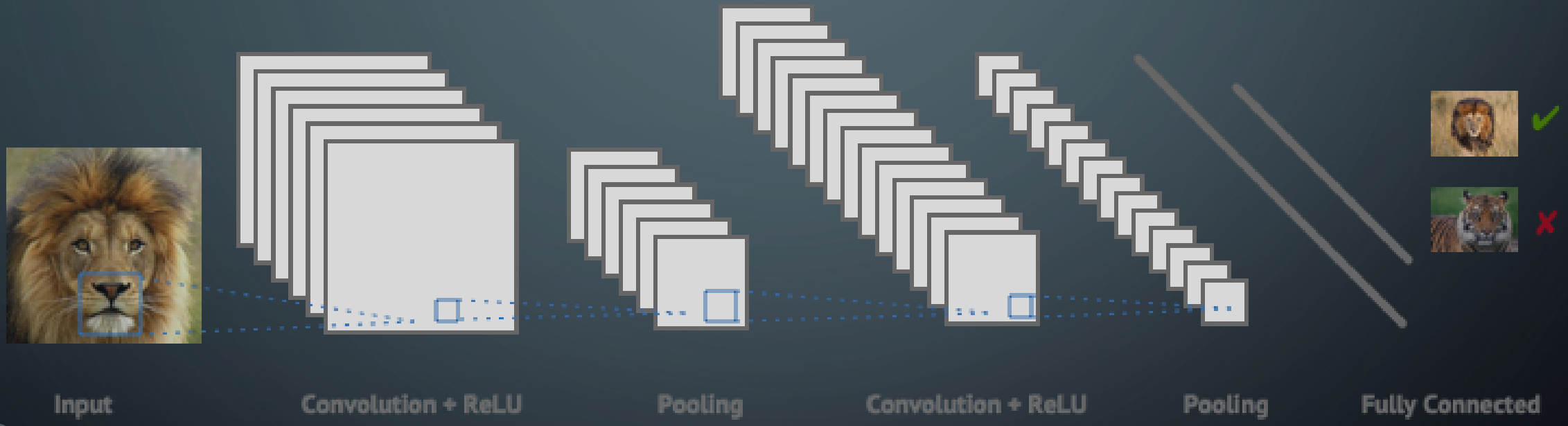


POOLING LAYER

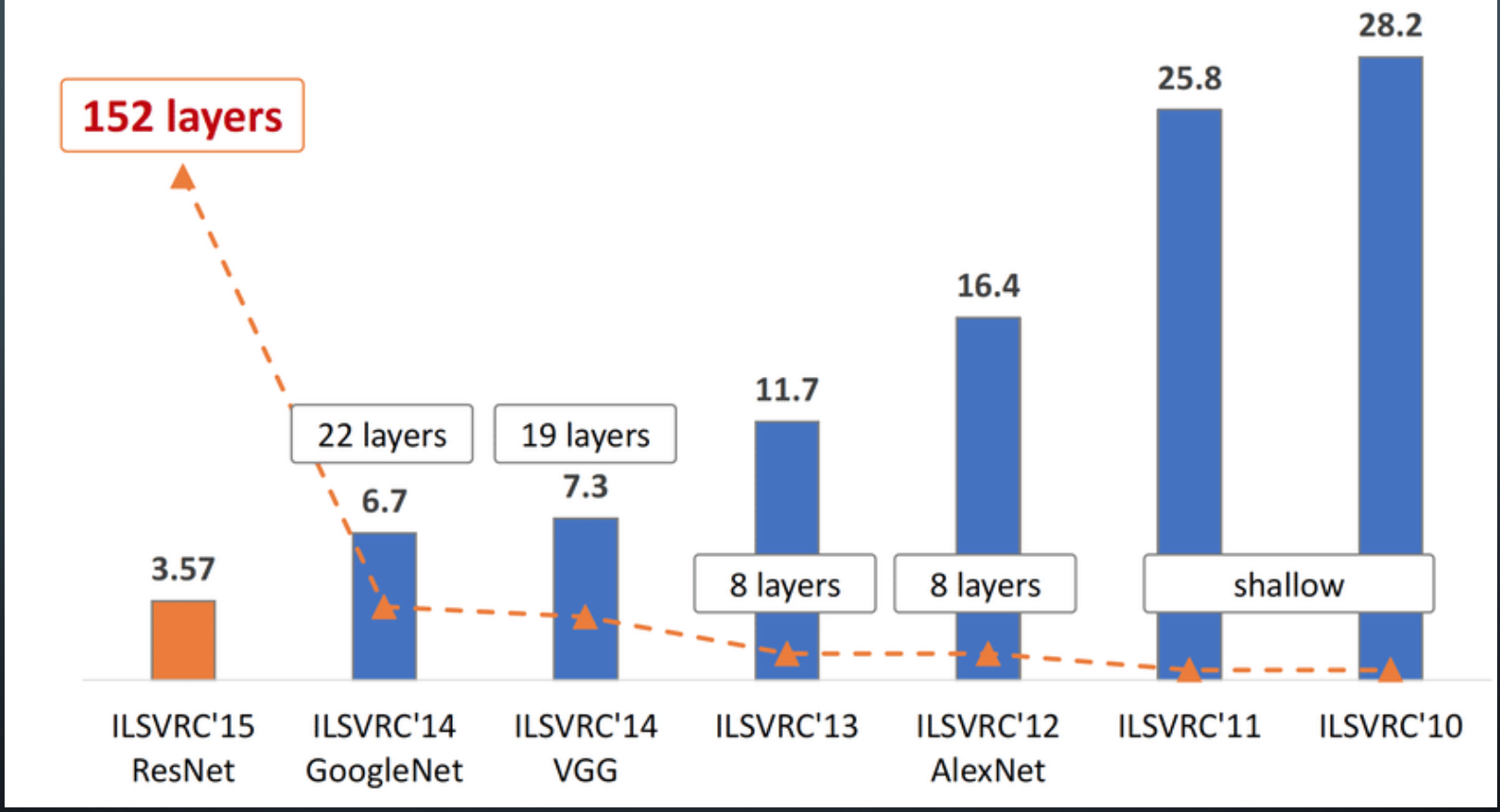
- Max / Average
- Not trainable
- Stride = kernel size
- Decrease image size
- Increase speed



CNN ARCHITECTURE

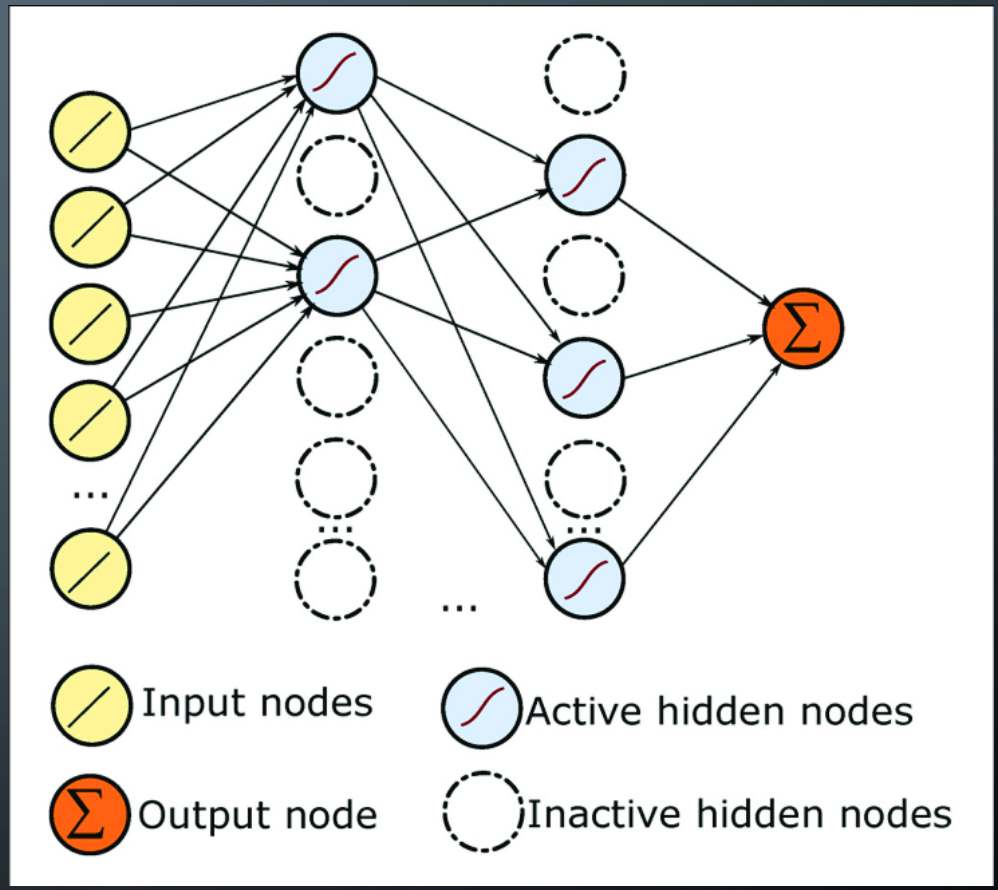


CNN VS SHALLOW



DROPOUT

- Prevent overfitting
- Weak model in train phase
- Powerful model in inference phase
- Randomness



OTHER CNN BASE LAYERS

- Conv1D
- Batch normalization
- Global max pooling
- Global average pooling

CNN

- Whole model is not totally position invariant
  - There is MLP layer
- Usually we use 3*3 conv with stride 1
- Usually number of features increase over layers
- Usually we use 2*2 pooling with stride 2