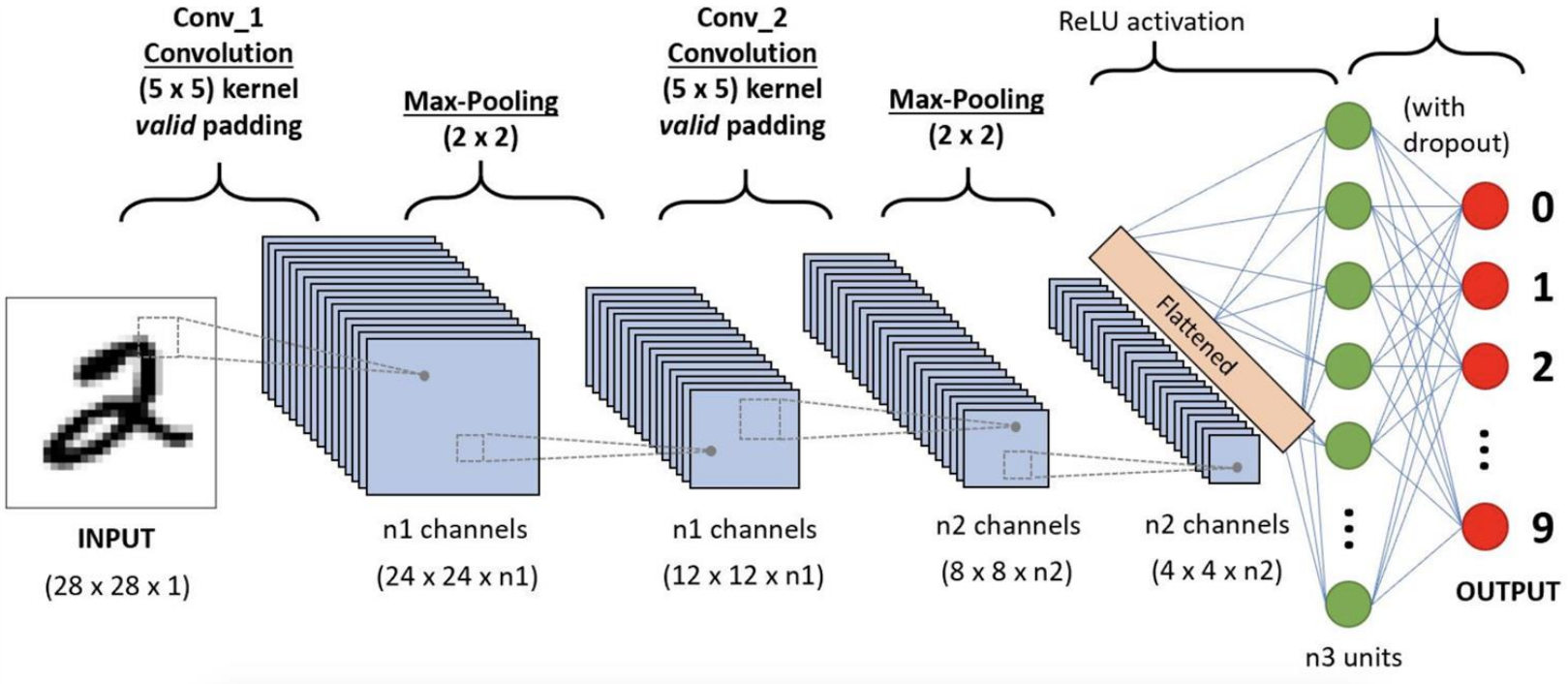
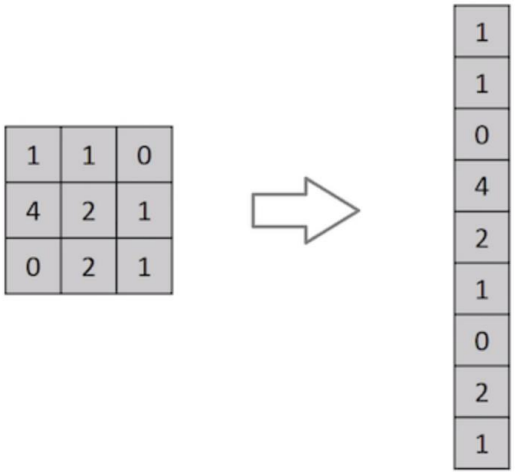
Convolutional Neural Networks

Let’s now deal with images!

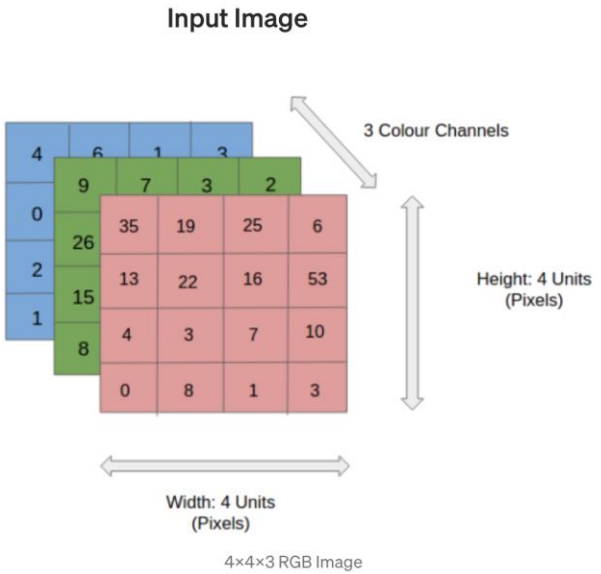


Images source: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

3-dimensional data

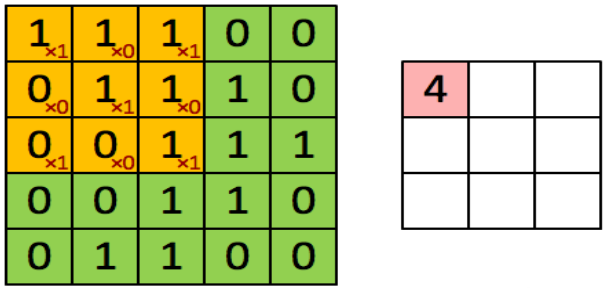


Flattening of a 3×3 image matrix into a 9×1 vector



Images source: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

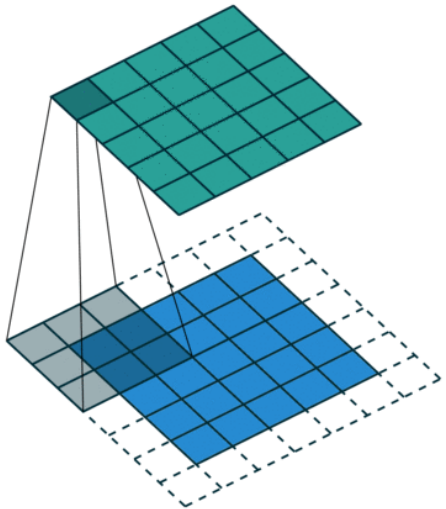
Convolution Operation



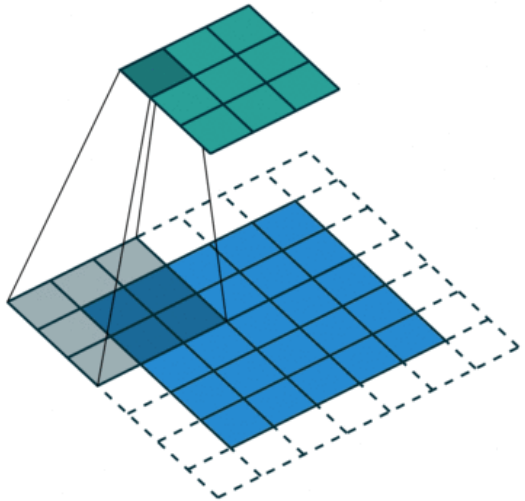
Images source: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

Convolution Operation

Padding: 1
Stride: 1



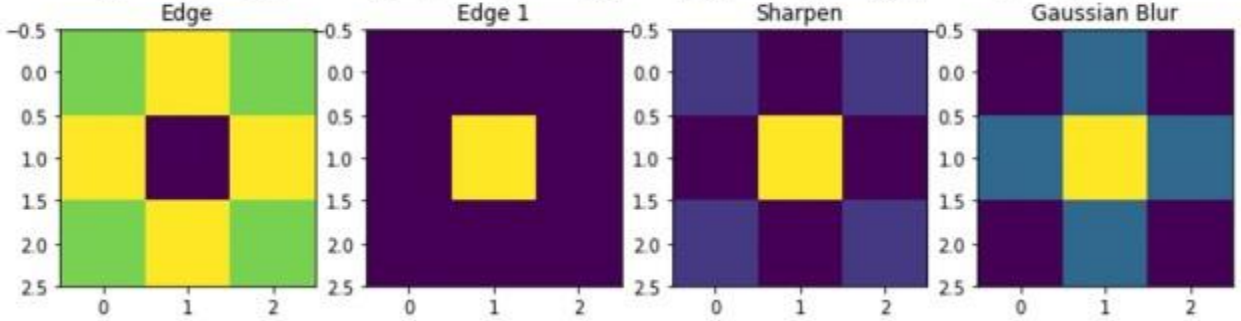
Padding: 1
Stride: 2



Images source: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

Different kernels for different feature extractions

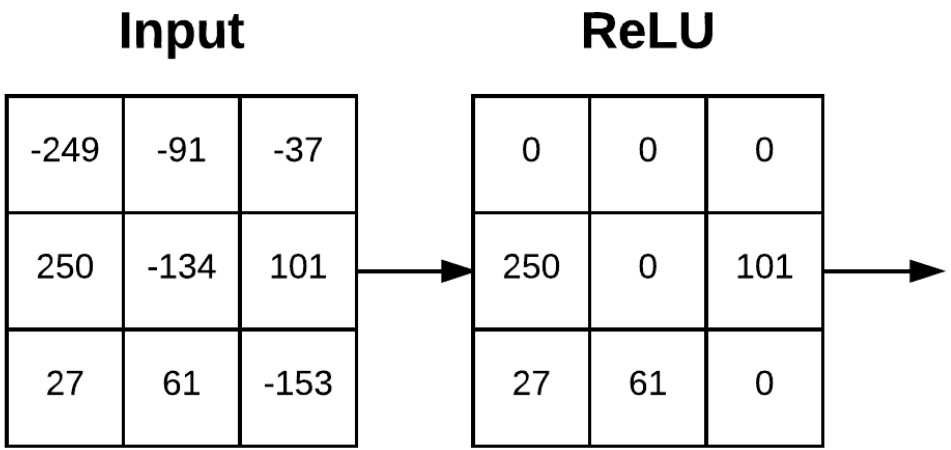
$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \times \frac{1}{16}$$



Images source: https://ucanalytics.com/blogs/convolutional-neural-networks-cnn-simplified-part-4/

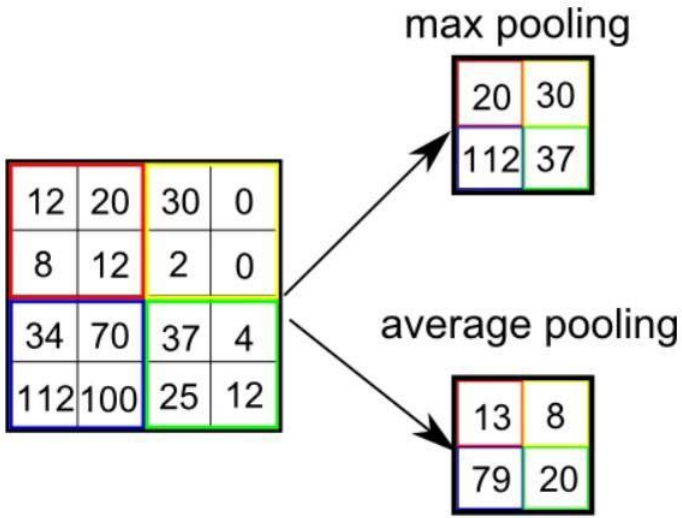
Activation layer

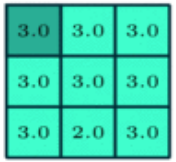
After each convolutional layer, we apply an activation function (usually ReLU)

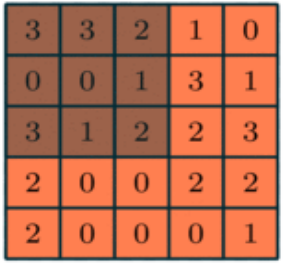


Images source: https://pyimagesearch.com/2021/05/14/convolutional-neural-networks-cnns-and-layer-types/

Pooling



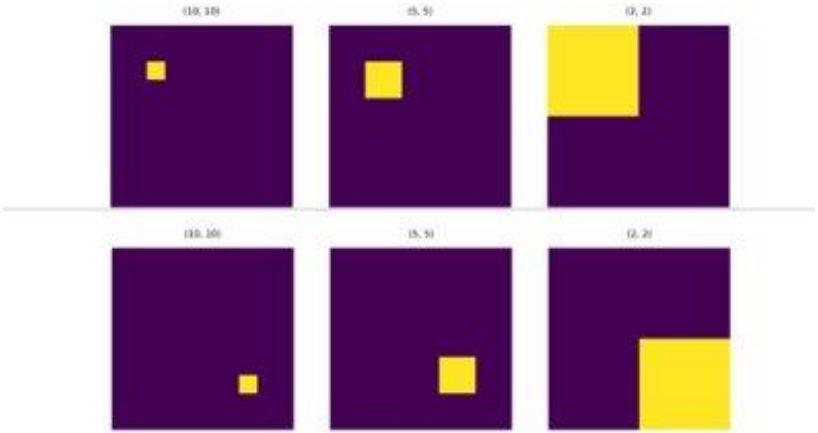


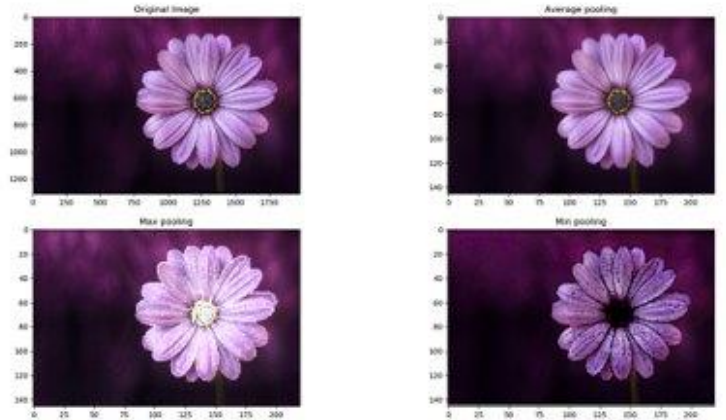


Images source: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

Pooling (comparison)

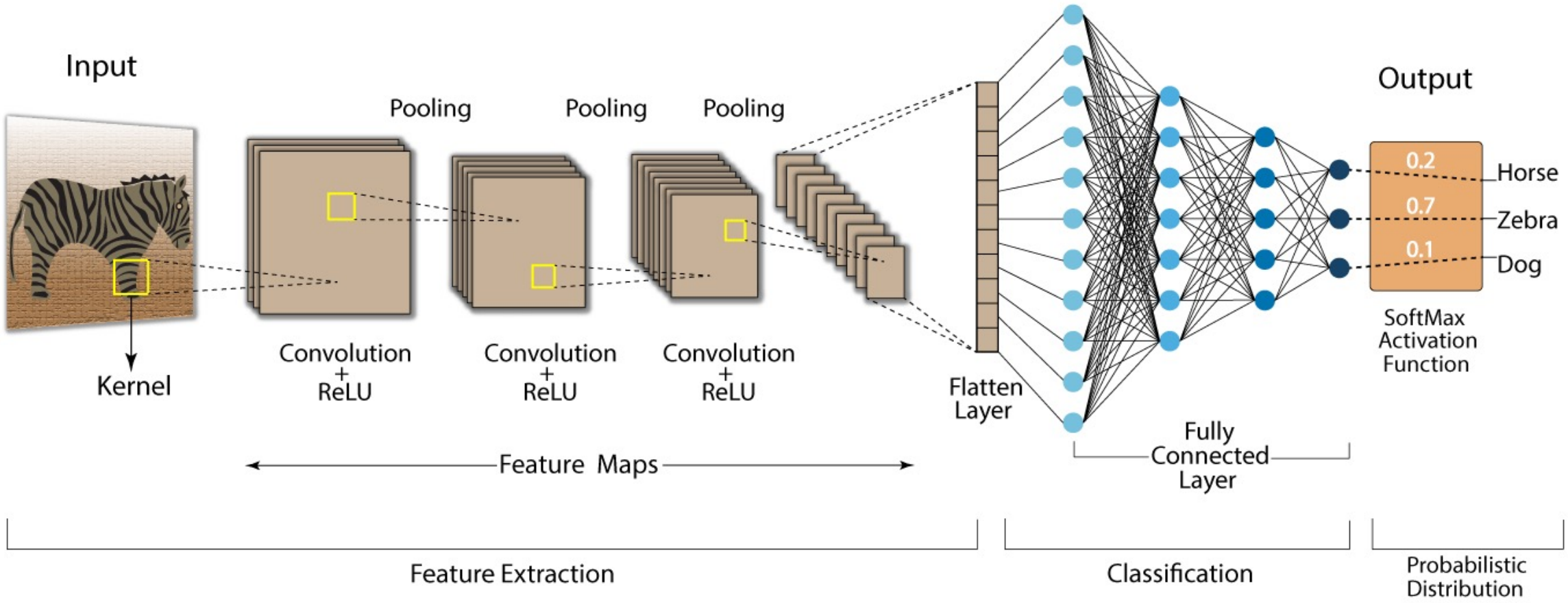
Max pooling





Images source:
https://blog.paperspace.com/pooling-and-translation-invariance-in-convolutional-neural-networks/
https://medium.com/@bdhuma/which-pooling-method-is-better-maxpooling-vs-minpooling-vs-average-pooling-95fb03f45a9

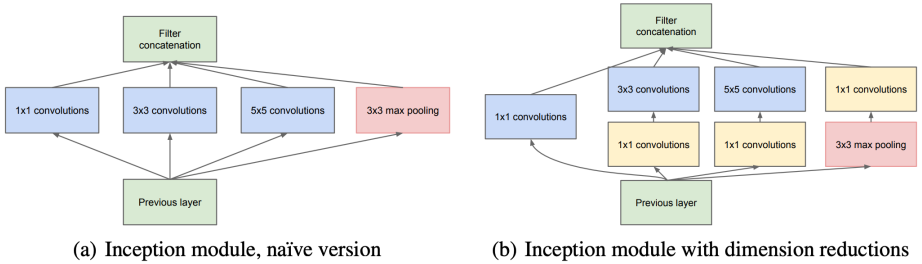
Overall Architecture of a CNN



Images source: https://www.analyticsvidhya.com/blog/2021/05/20-questions-to-test-your-skills-on-cnn-convolutional-neural-networks/

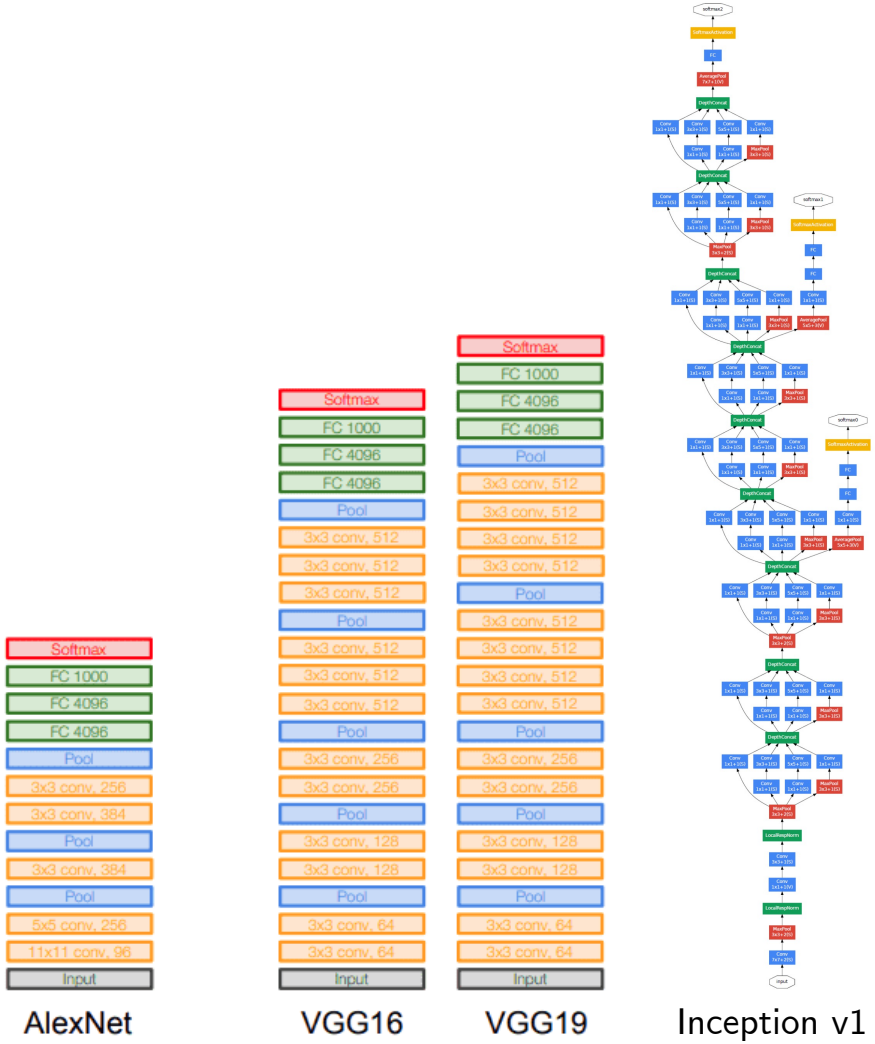
Historic CNN architectures

- Historic architectures:
- LeNet-5 (1989),
- AlexNet (2012),
- GoogleNet/Inception (2014),
- VGG (2014).

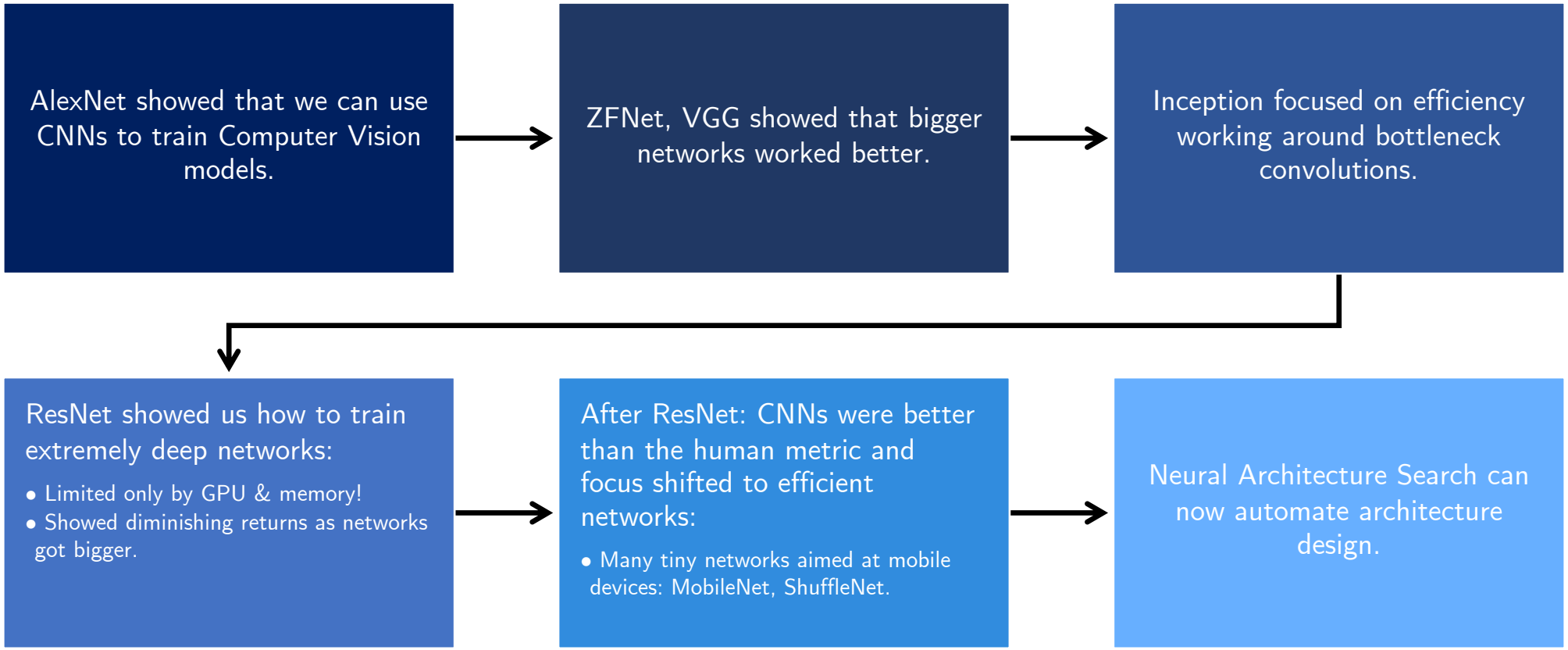


(a) Inception module, naïve version (b) Inception module with dimension reductions

Inception module from GoogleNet



CNN History: What to remember?

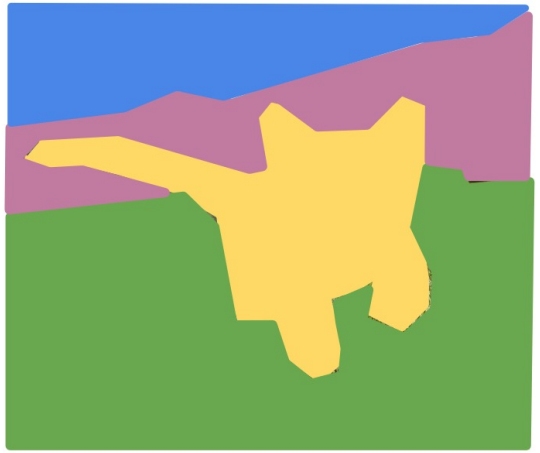


What kind of other tasks can we perform?

- Classification, Regression
- Segmentation
- Prediction of next image
- Content generation
- Feature extraction
- Descriptions
- ...

Other Computer Vision Tasks

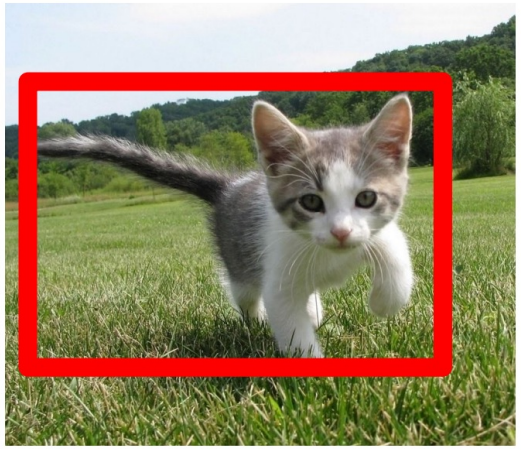
Semantic Segmentation



GRASS, CAT, TREE, SKY

No objects, just pixels

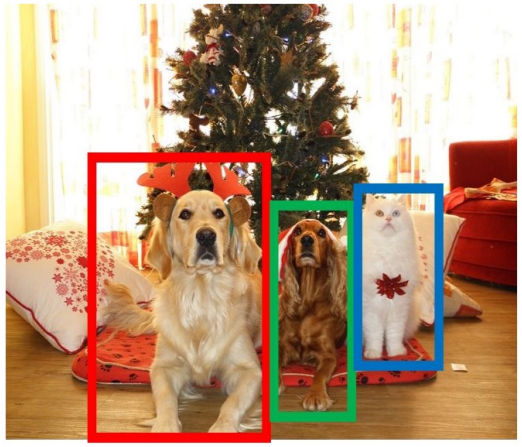
Classification + Localization



CAT

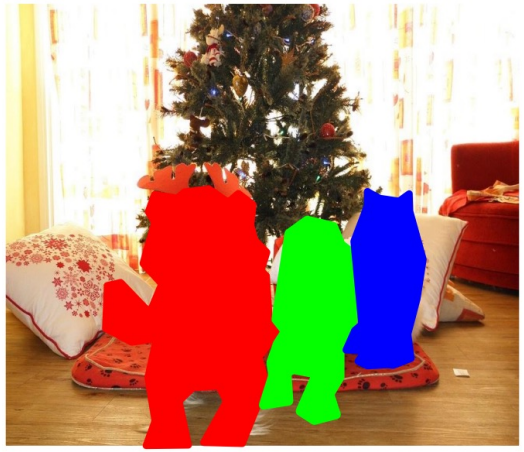
Single Object

Object Detection



DOG, DOG, CAT

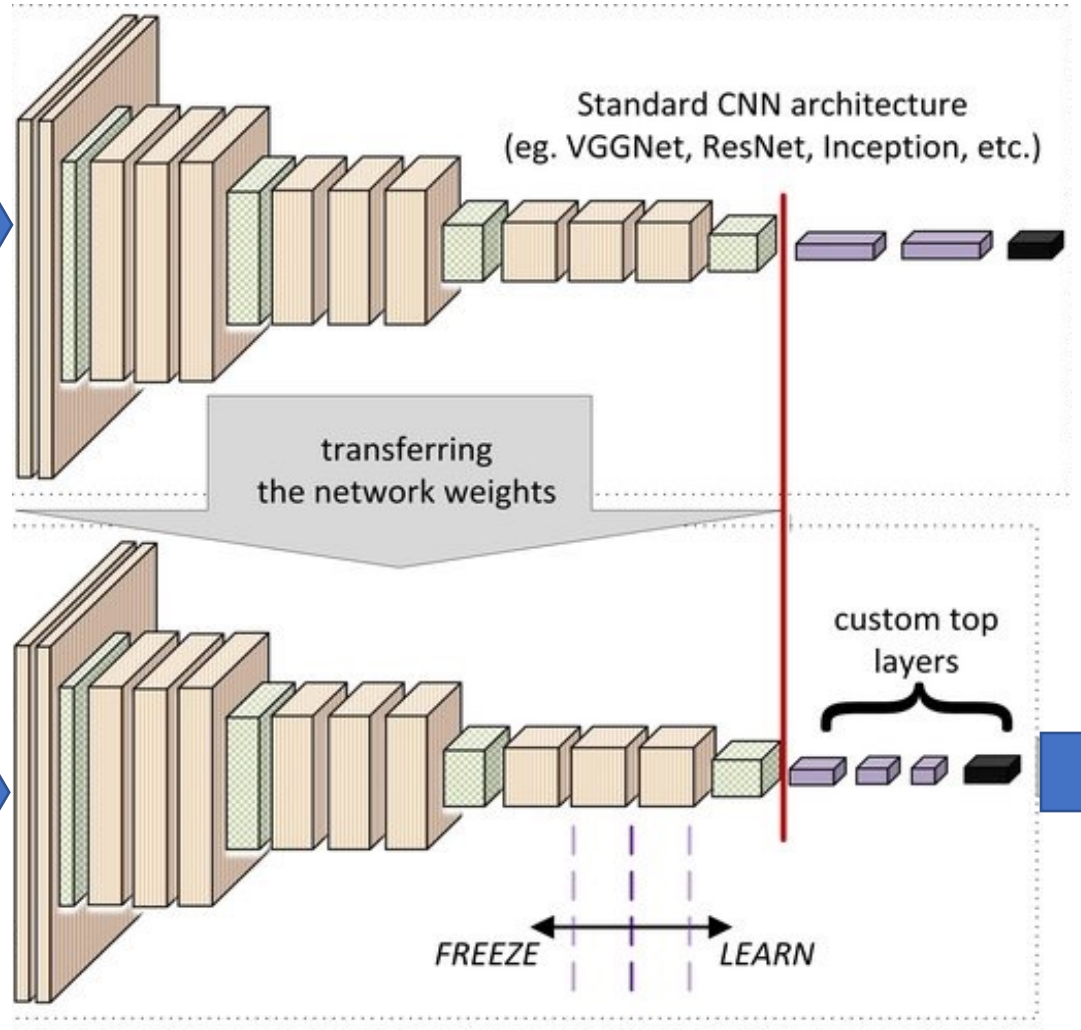
Instance Segmentation



DOG, DOG, CAT

Multiple Object

Transfer Learning



Prediction

Please ask any questions!

Thank you for your attention!