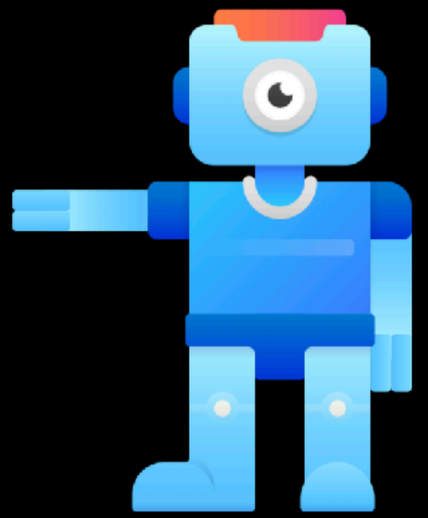


MODERN COMPUTER VISION

BY RAJEEV RATAN

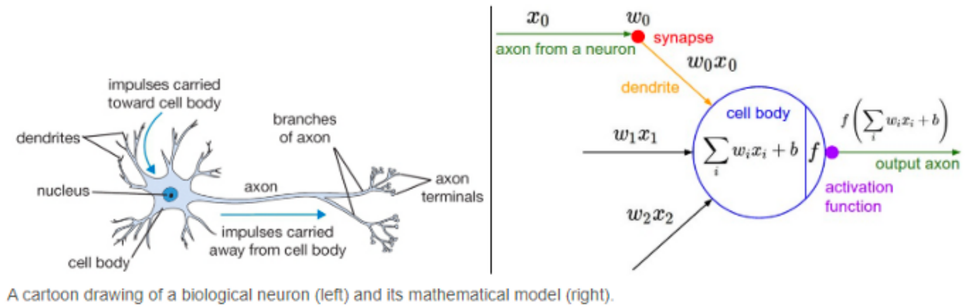
Activation Layer ReLU

What are Activation Functions and their importance

Purpose of Activation Functions

To enable the learning of **complex patterns** in our data

- Biological neurons fire (activate) on certain inputs, these are then fed into other neurons
- Introduces **non-linearity** to our network
- This allows a non-linear decision boundary via non-linear combinations of the weight and inputs



A cartoon drawing of a biological neuron (left) and its mathematical model (right).

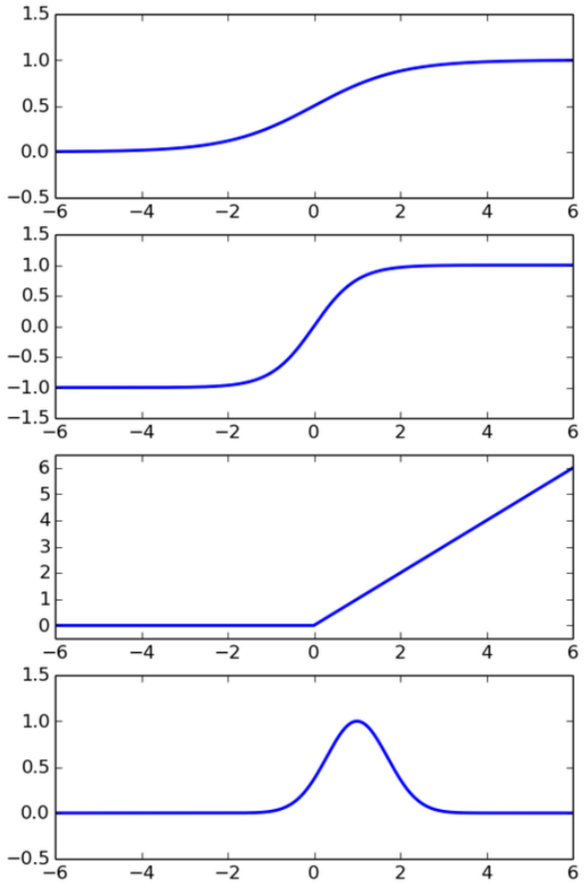
Stanford's CS231 Course

Types of Activation Functions

There are several activation functions we can use in our CNN. However, Rectified Linear Units (ReLU) have become the activation function of choice for CNNs.

ReLU is advantageous in CNN Training:

- Simple Computation (fast to train)
- Does not saturate



Sigmoid

$$\phi(z) = \frac{1}{1 + e^{-z}}$$

Hyperbolic Tangent

$$\phi(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}}$$

Rectified Linear

$$\phi(z) = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{if } z \geq 0 \end{cases}$$

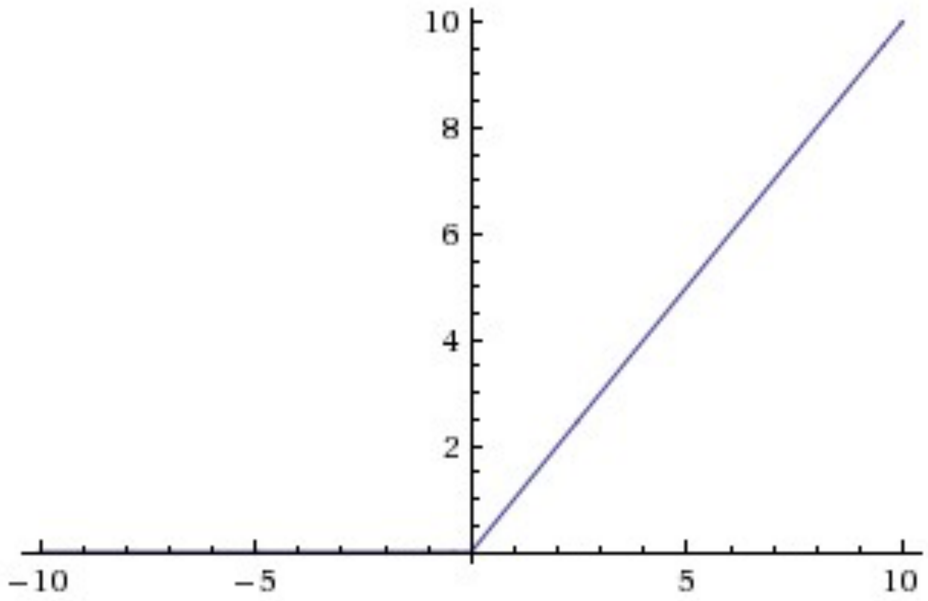
Radial Basis Function

$$\phi(z, c) = e^{-(\epsilon \|z - c\|)^2}$$

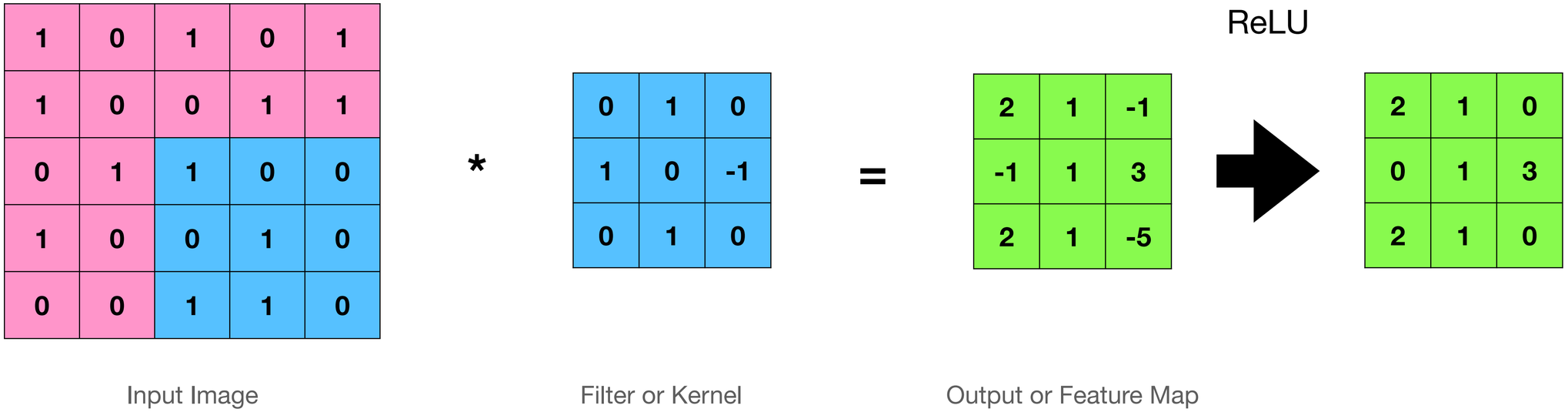
The ReLU Operation

- Change all negative values to 0
- Leave all positive Values alone

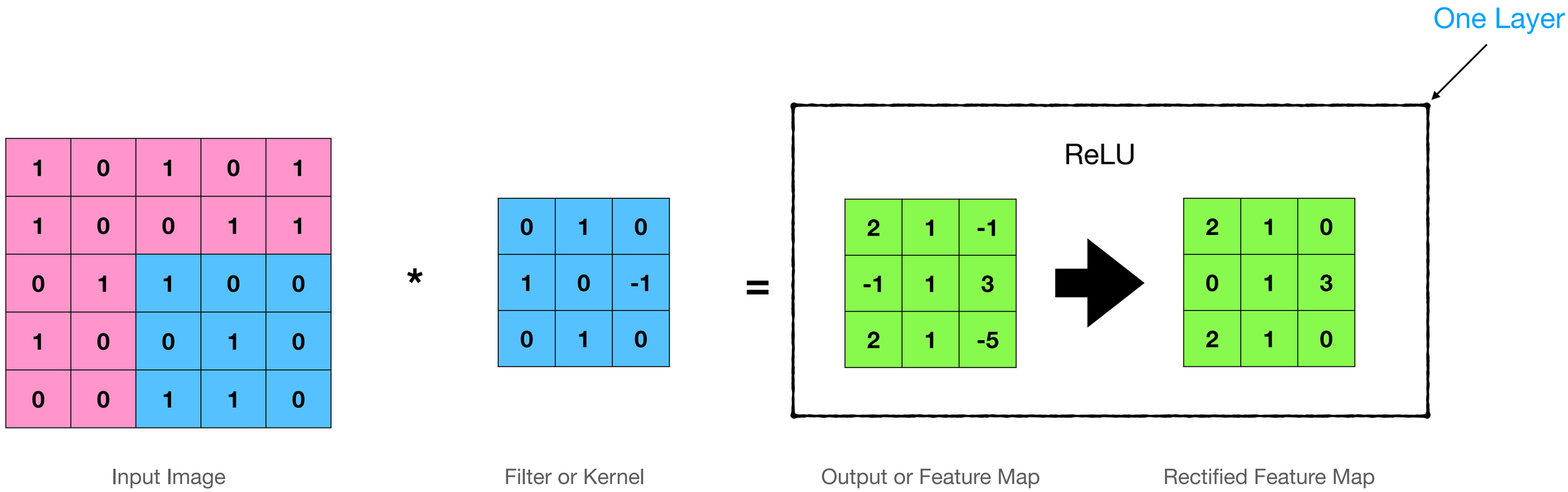
$$f(x) = max(0,x)$$



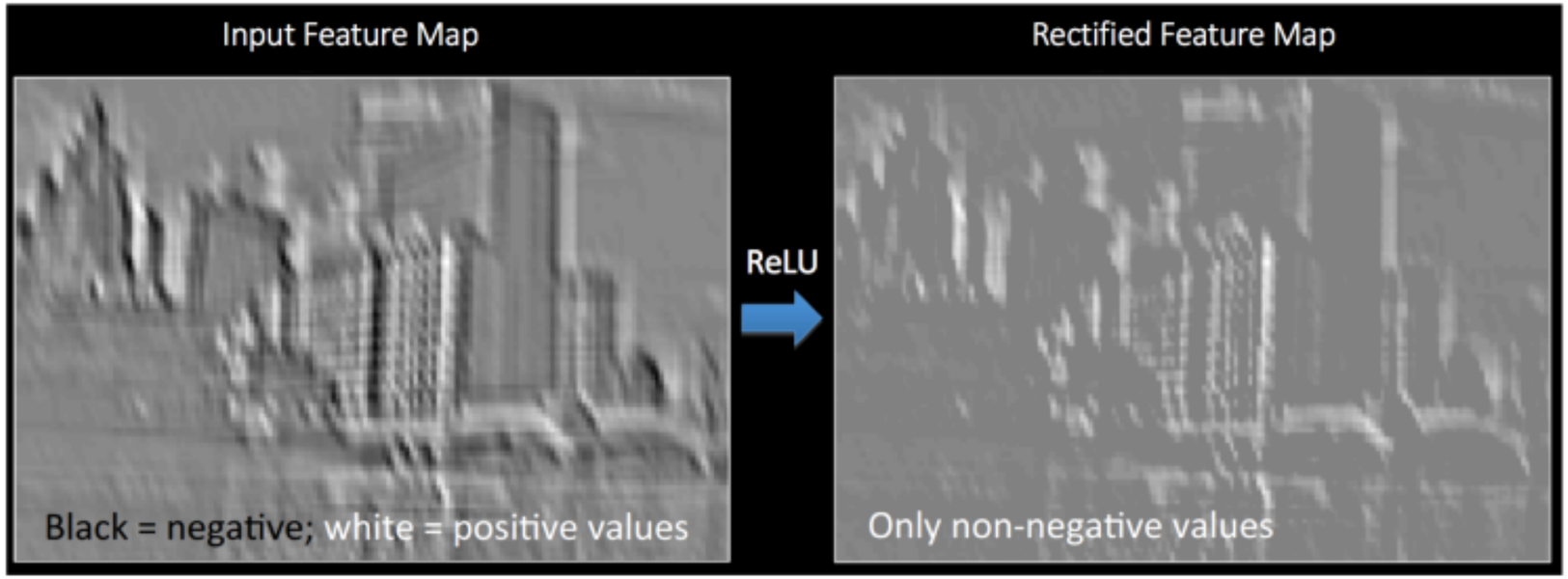
Applying the ReLU Activation



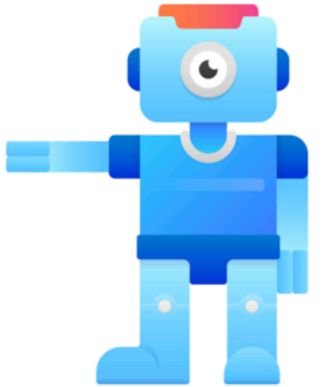
Applying the ReLU Activation



Example of a Rectified Linear Map



Source - http://mlss.tuebingen.mpg.de/2015/slides/fergus/Fergus_1.pdf



MODERN COMPUTER VISION

BY RAJEEV RATAN

Next...

Pooling