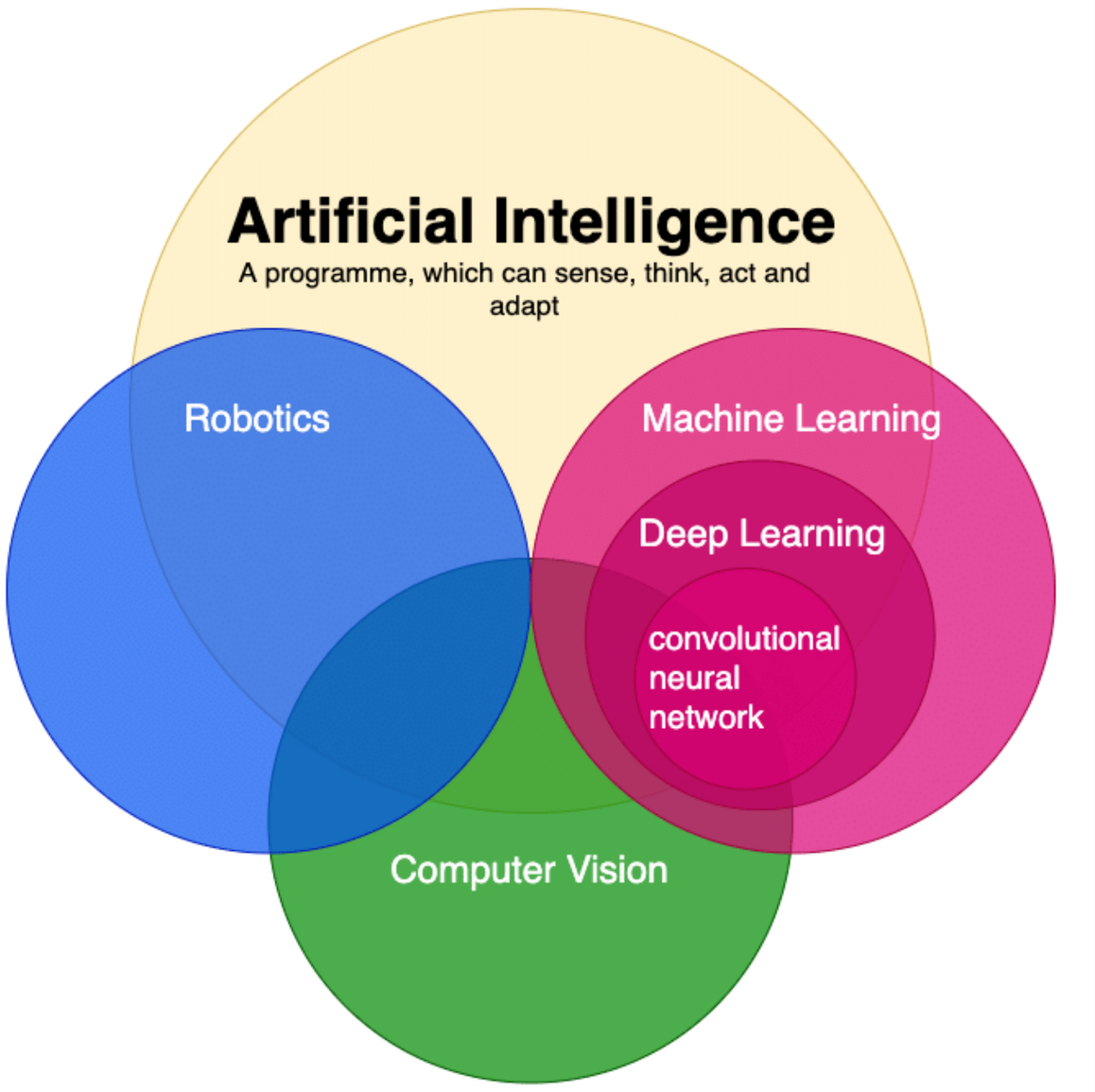
COMPUTER VISION WORKSHOP

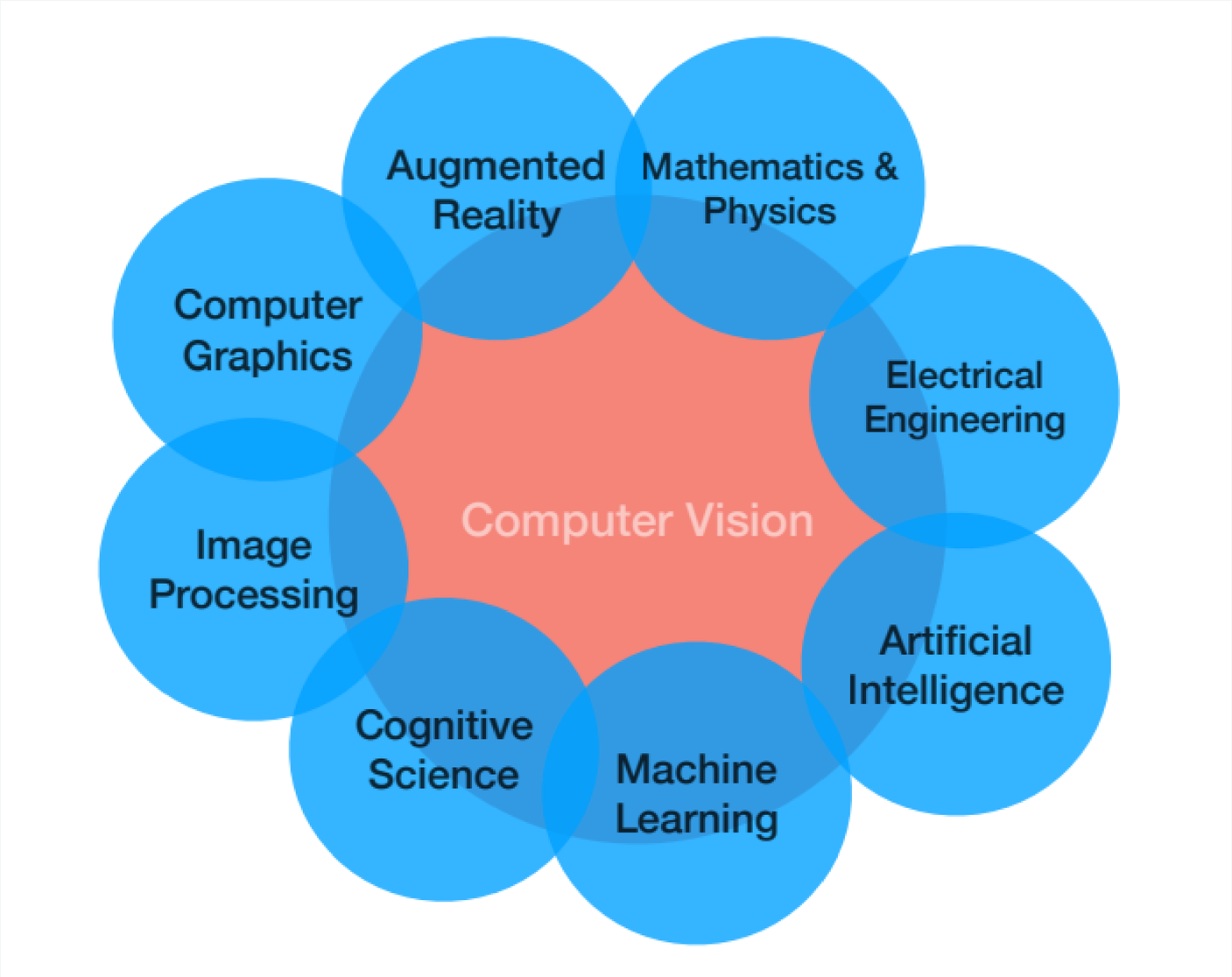
What exactly is computer vision?

" An interdisciplinary field that aims to enable computers to gain understanding of what is being seen in images and videos "

is computer vision artificial intelligence?



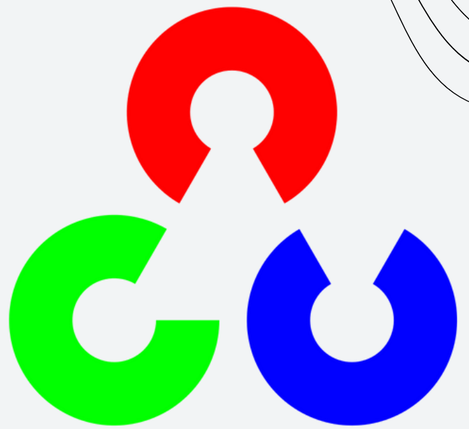
Computer Vision is an amalgamation of many fields



SOME CV APPLICATIONS

- Snapchat and Instagram filters
- Optical Character Recognition (OCR)
- Licence Plate Reading
- Self-driving cars
- Sporting Analysis
- Facial Recognition
- Object Detection
- Segmentation
- Image Similarity
- Deep Fakes

What makes Python so great for Computer Vision?



- Easy to learn and accessible
- Language of Artificial Intelligence
- The best and easiest to use Libraries, such as:
    - OpenCV for Classical Computer Vision
    - PyTorch (Meta) and TensorFlow with Keras (Google)

Classical Computer Vision

- What is meant by **Classical Computer Vision?**

- It encompasses Computer Vision algorithms that **do not** involve Machine Learning

- Before the advent of ML and DL, CV was a deeply explored field and many useful algorithms were developed for things like **feature extraction, OCR, Segmentation and simple transformations.**

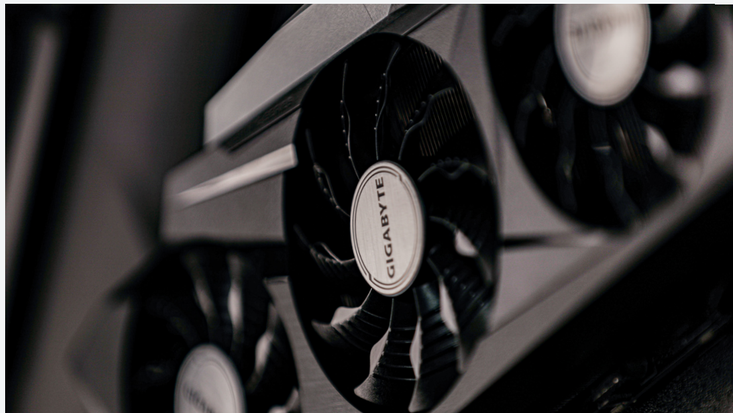
Deep Computer Vision

- Deep Learning was used in Computer Vision since the 1990s, however due to the computational requirements and intricate design, it remained on the sidelines for decades.

- **Mature Deep Learning libraries** (TensorFlow, Keras, Theano)

- **Accessible GPU processing** (NVIDIA’s CUDA)

Google Developer Student Clubs
SUPCOM



DEEP LEARNING CV VS CLASSICAL CV

Deep Learning Computer Vision

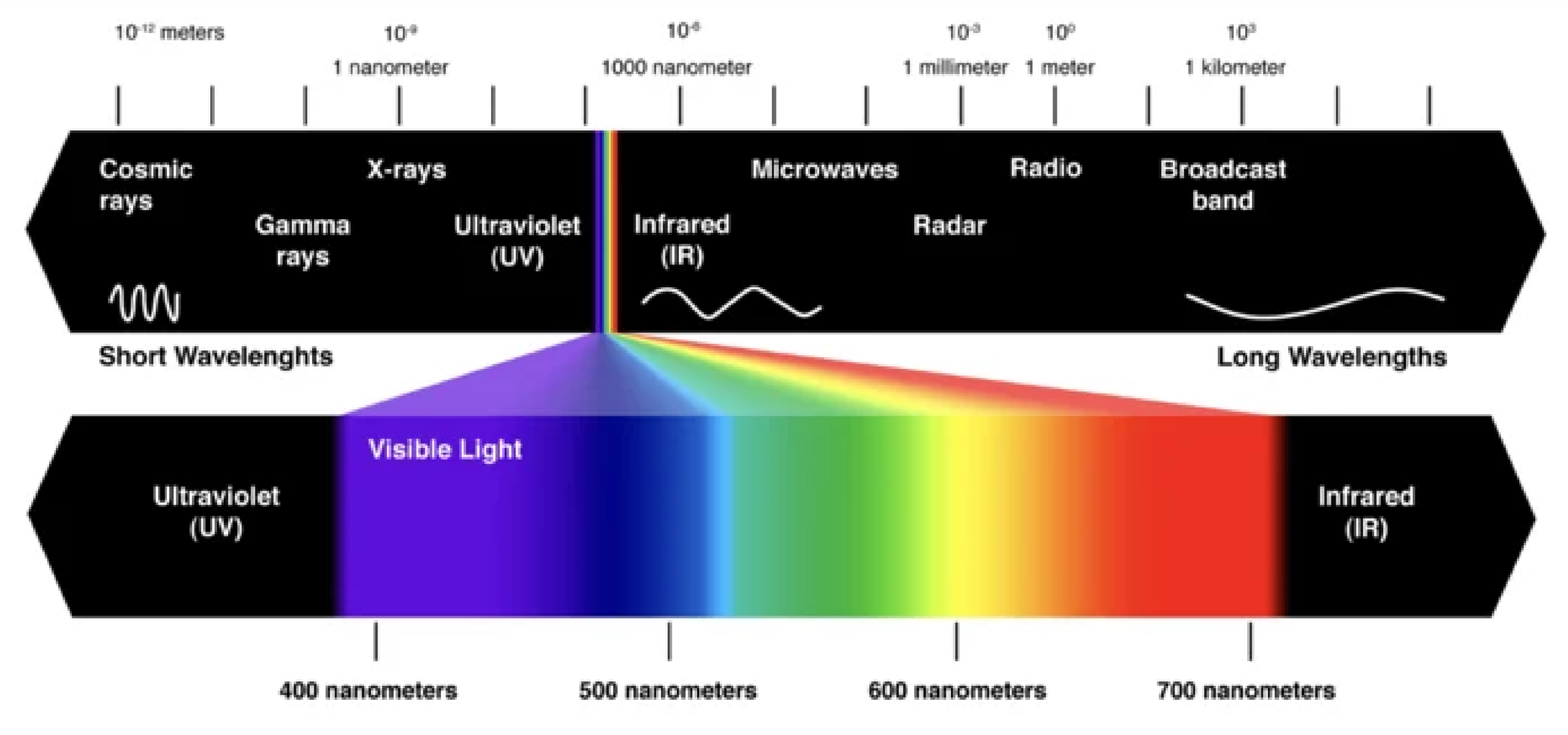
- Adapts to new images well (assuming it’s similar to the data it was trained on)
- Requires Models to be trained
- Requires GPU hardware (most times)

- Small changes can have big negative impacts
- Relies on hardcode features and parameters
- Can be run on CPU

Classical Computer Vision

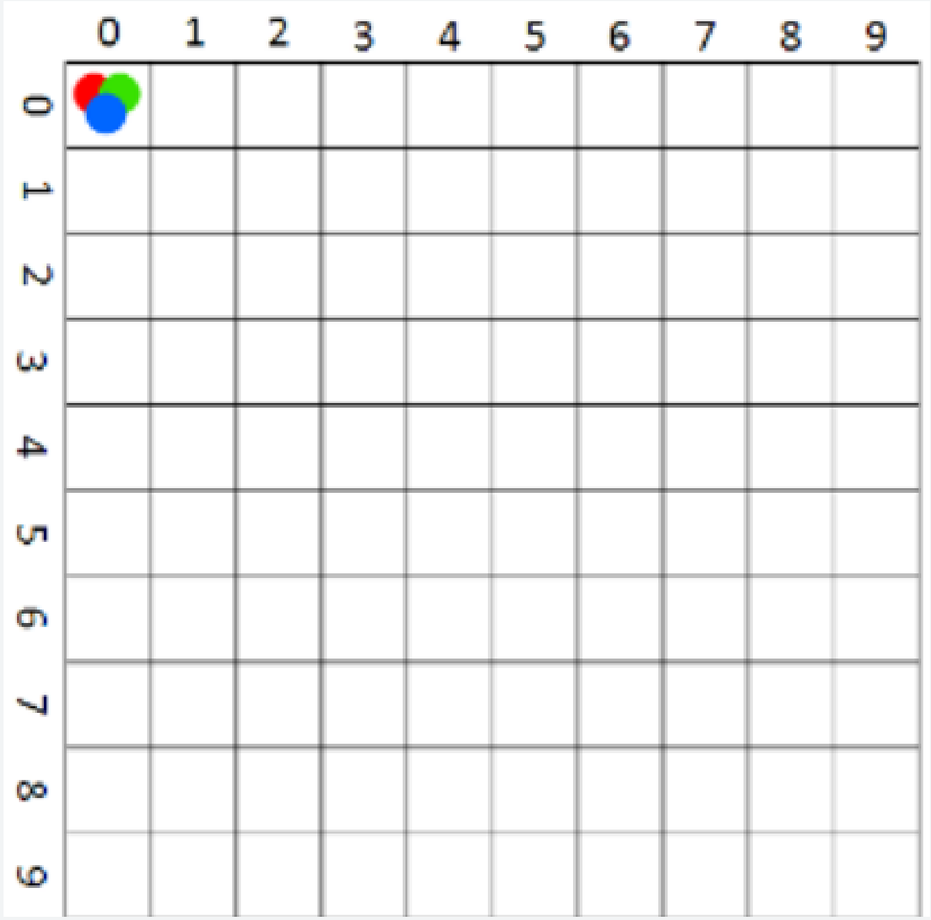
What are images?

- Images are 2-D representations of the visible light spectrum



Digital images format

- Images are stored in Multi-Dimensional Arrays



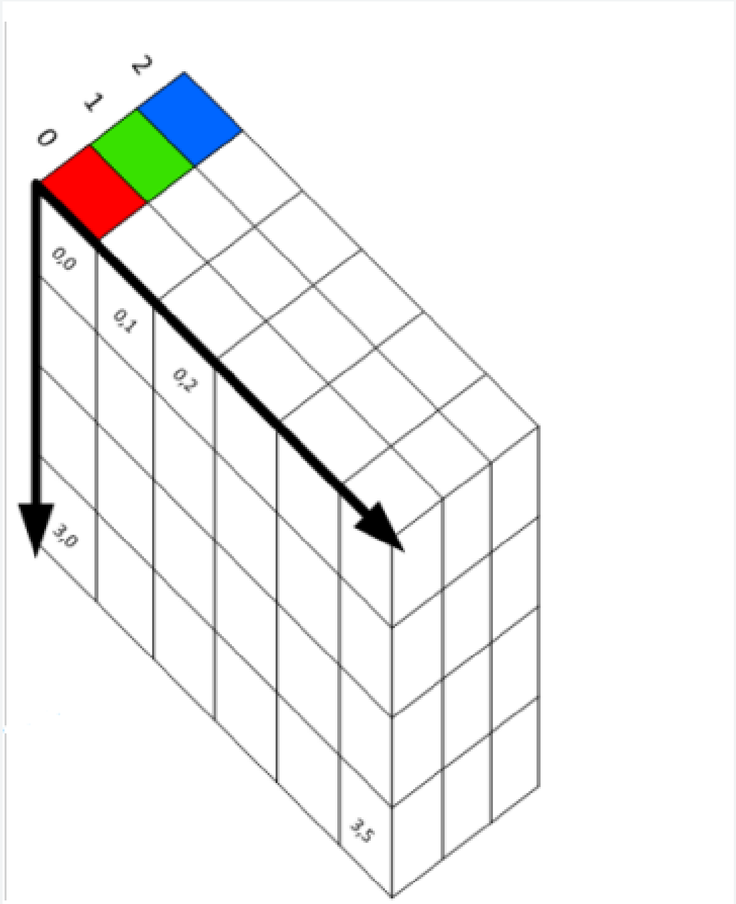


Image file format

Raster images

Pixel-based graphics
Resolution dependent
Photos & web graphics

JPG
Web & print photos and quick previews

GIF
Animation & transparency in limited colors

PNG
Transparency with millions of colors

TIFF
High quality print graphics and scans

RAW
Unprocessed data from digital cameras

PSD
Layered Adobe Photoshop design files

Vector images

Curve-based graphics
Resolution independent
Logos, icons, & type

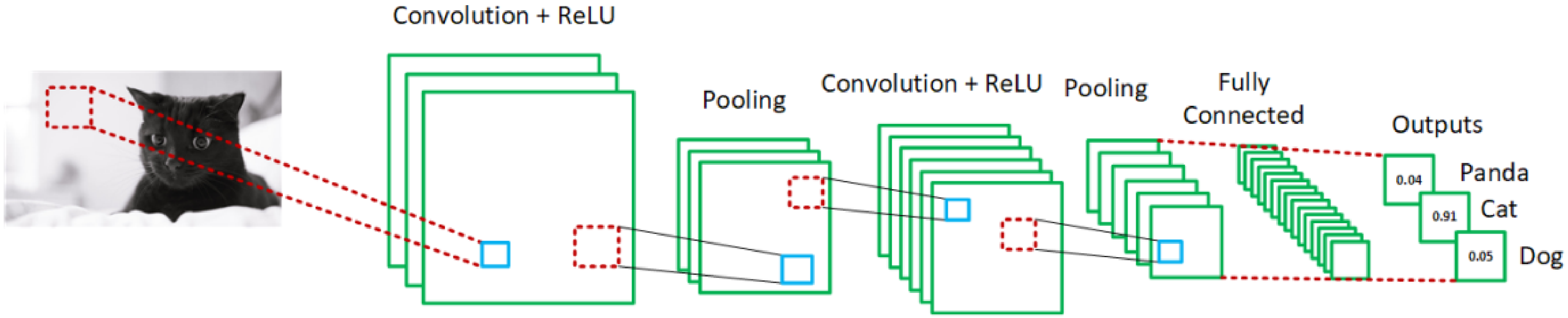
PDF
Print files and web-based documents

EPS
Individual vector design elements

AI
Original Adobe Illustrator design files

SVG
Vector files for web publishing

CONVOLUTIONAL NEURAL NETWORKS



The Convolution Operation

- Mathematical term to describe the process of combining two functions to produce a **third**
- In our situation, the output is called a Feature Map
- We use a matrix, called a Filter or **Kernel** that is applied to our Image
- So the first ‘Function’ is the image that is combined with the Kernel or Filter which produces a feature map

Feature Map Image × Kernel = Feature Map

The Convolution Operation

1	0	1	0	1
1	0	0	1	1
0	1	1	0	0
1	0	0	1	0
0	0	1	1	0

Input Image

*

0	1	0
1	0	-1
0	1	0

Filter or Kernel

=

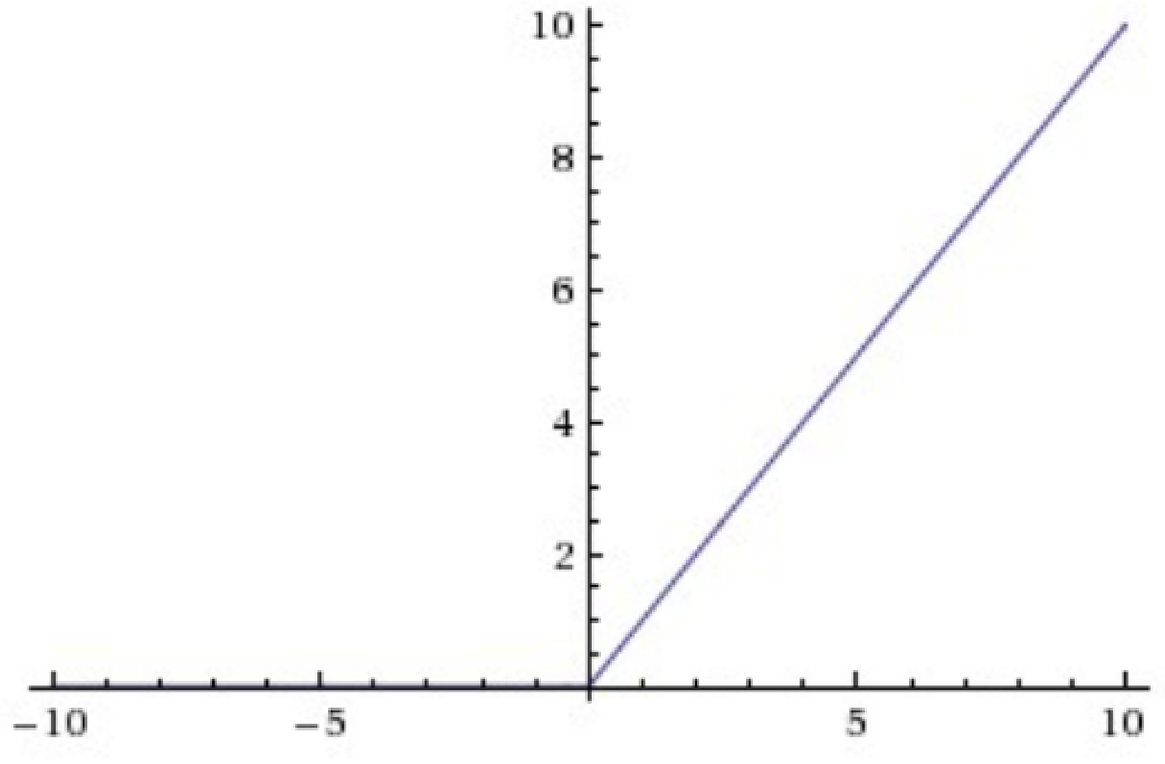
2	1	-1
-1	1	3
2	1	1

Output or Feature Map

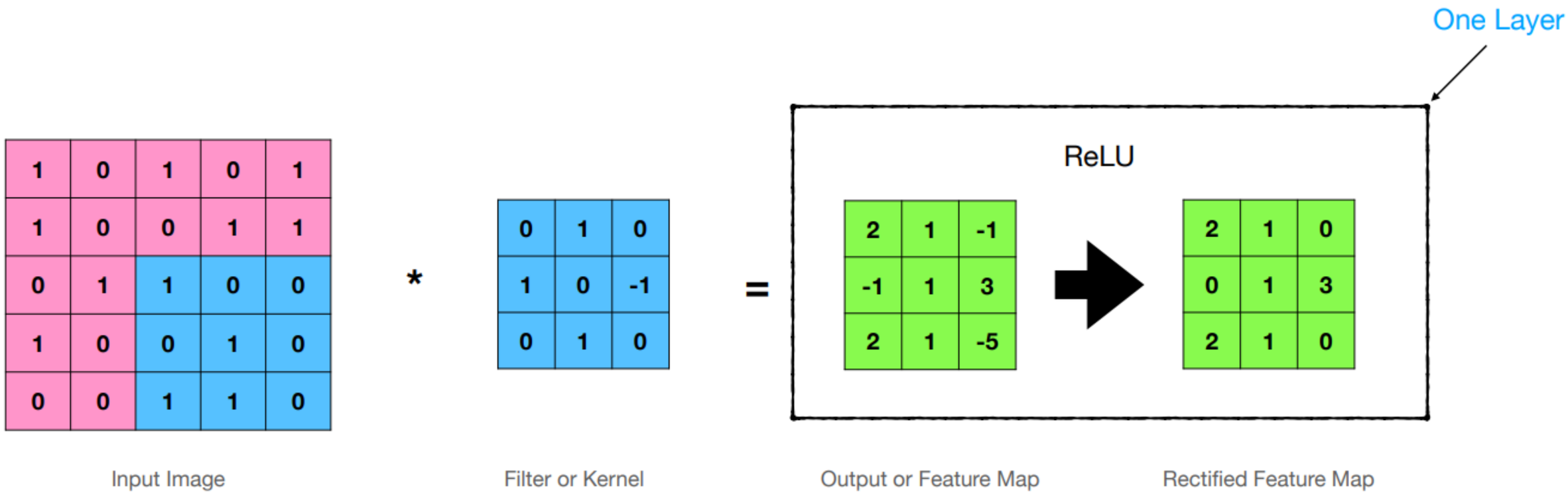
The ReLU Operation

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

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



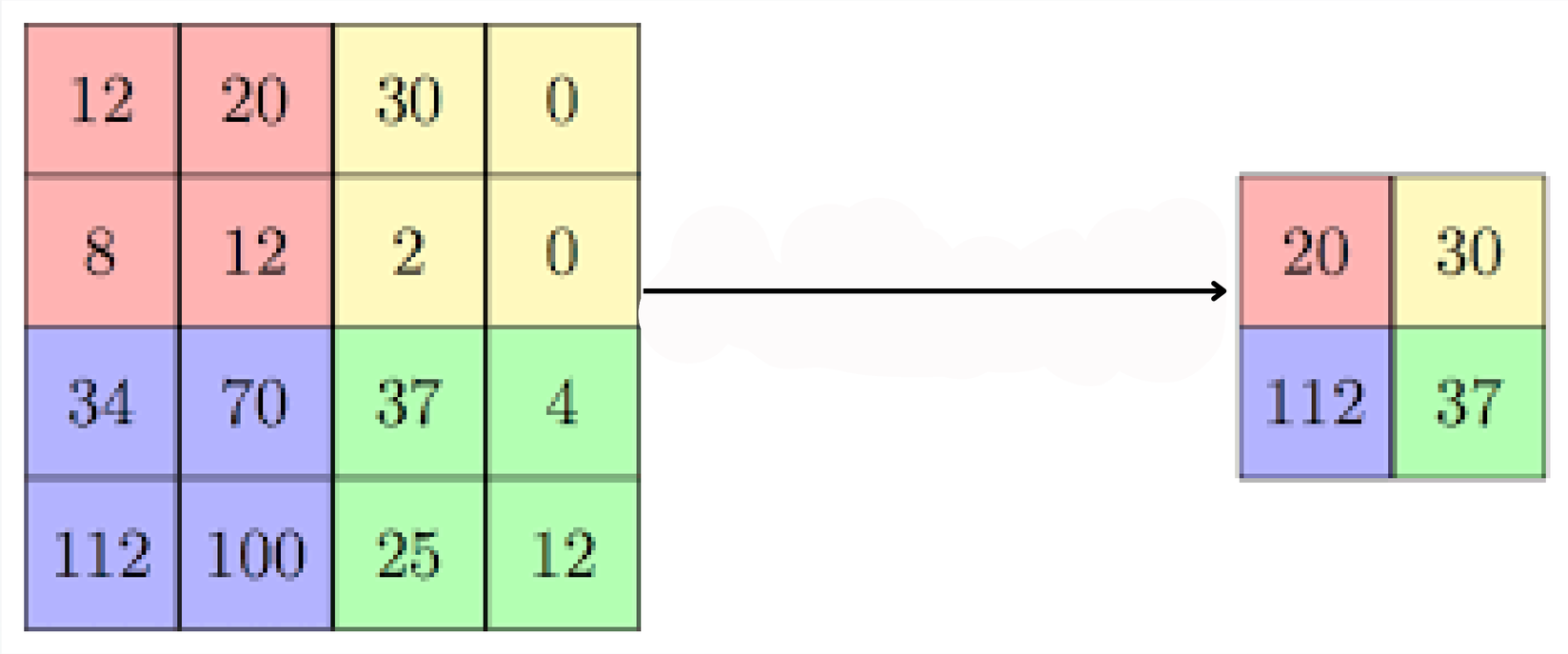
Applying the ReLU Activation



Pooling

• Pooling is the process whereby we reduce the size or dimensionality of the Feature Map

• This allows us to reduce the number of Parameters in our Network whilst retaining important features



USEFUL WEBSITES

- kaggle.com
- geeksforgeeks.org
- coursera.org
- udemy.net
- w3schools.com
- youtube.com

STRATEGY N°3

THANK'S FOR YOUR ATTENTION

Google Developer Student Clubs
SUPCOM