AI Deep Learning: Convolutional Neural Networks (I)

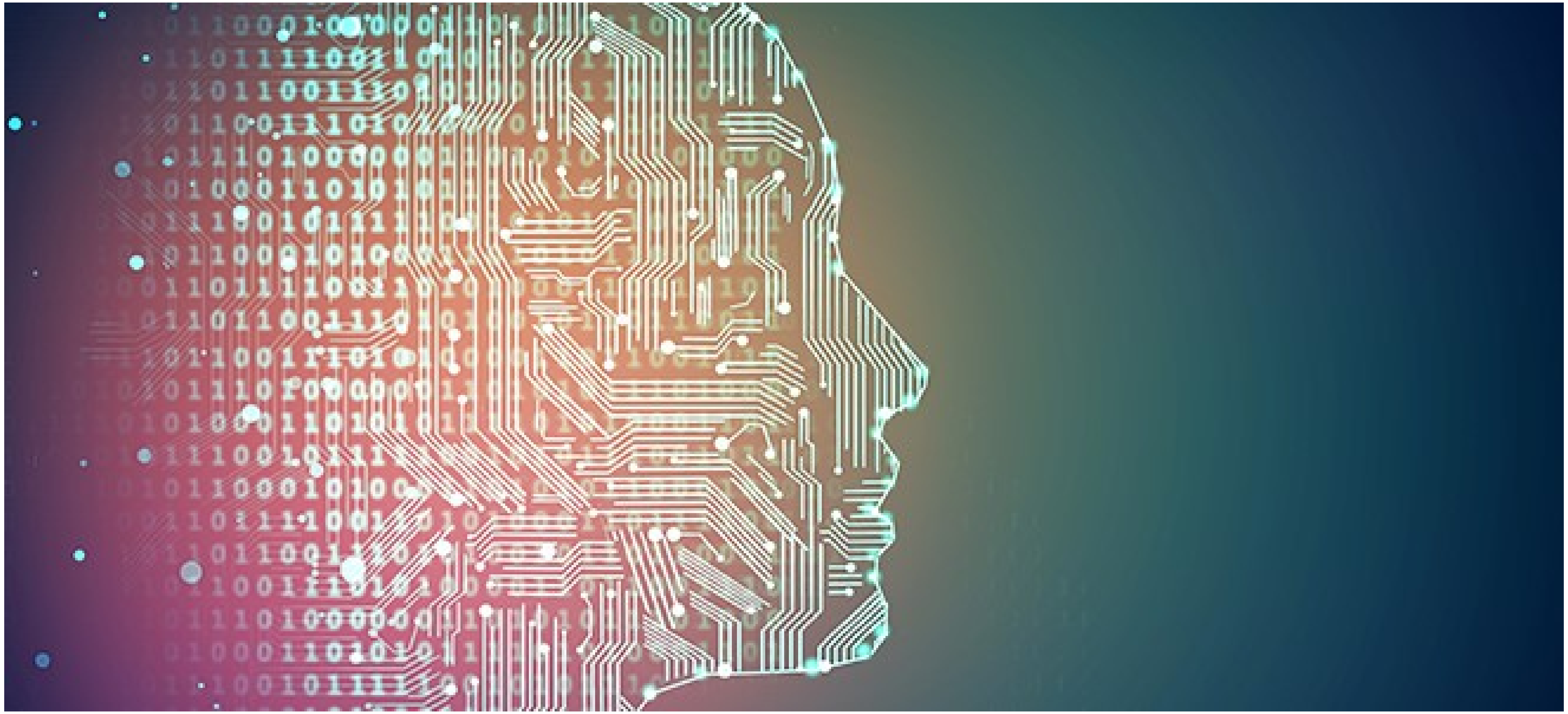
Deep Learning with Big Data

Thuan L Nguyen, PhD

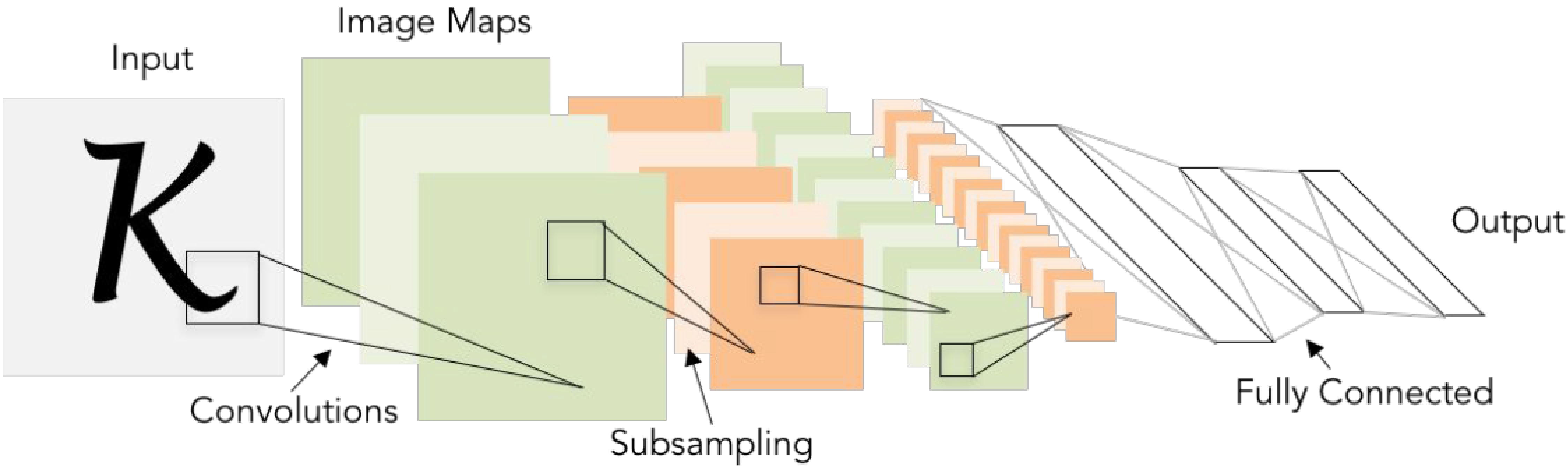
AI Deep Learning: Convolutional Neural Networks (CNN)

1. AI Deep Learning: Convolutional Neural Networks Overview
2. AI Deep Learning: Properties of Images
3. AI Deep Learning: Convolution Layers in Convolutional Neural Networks
4. AI Deep Learning: Pooling Layers in Convolutional Neural Networks
5. AI Deep Learning: Convolutional Neural Networks: Architectures: Overview
6. AI Deep Learning: Convolutional Neural Networks: Architectures: Layers
7. AI Deep Learning: Build and Train Convolutional Neural Networks

AI Deep Learning: Convolutional Neural Networks (CNN)



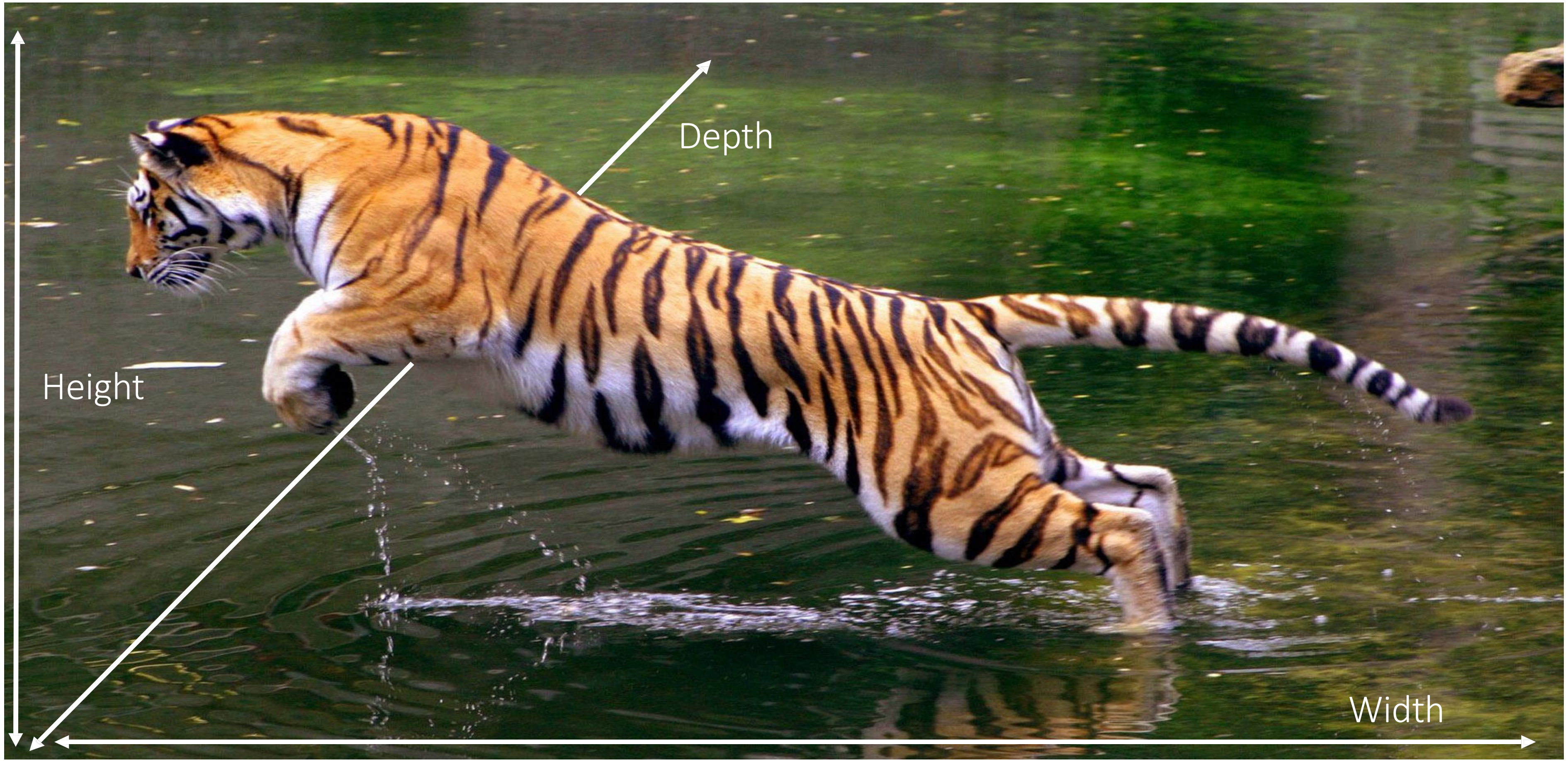
AI Deep Learning: Convolutional Neural Networks (CNN)



Sources: Stanford University

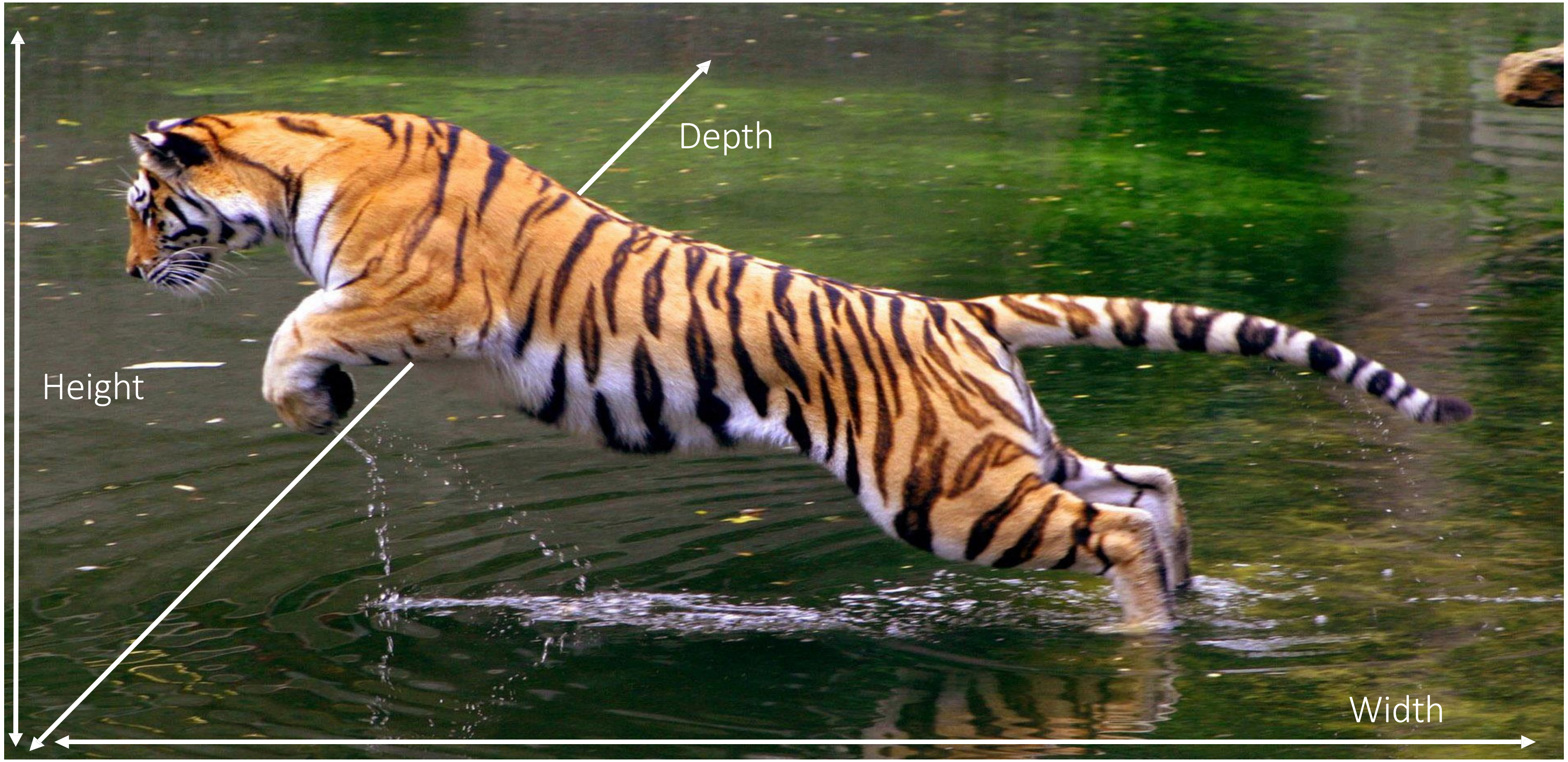
AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: Properties of Images



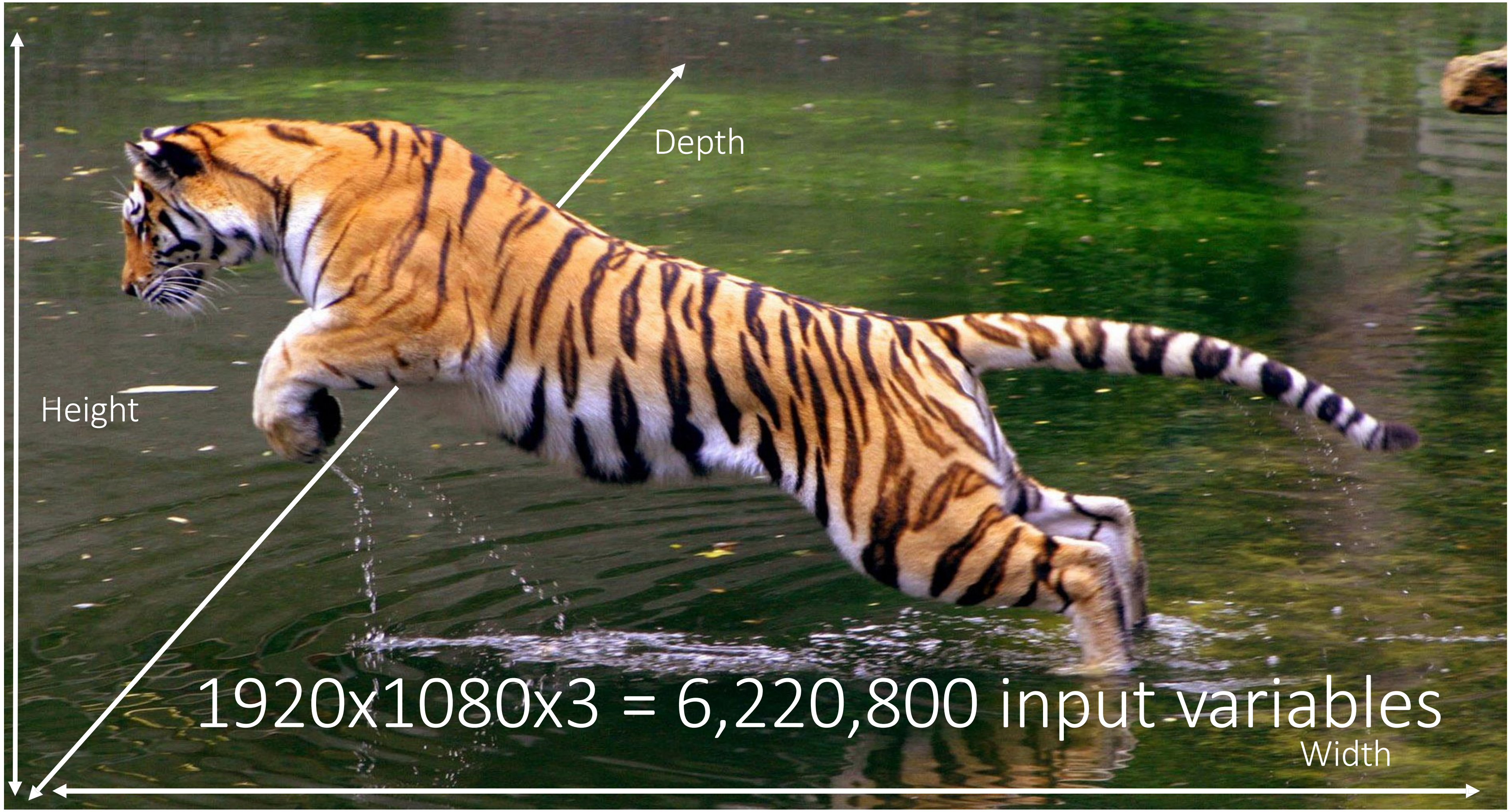
AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: Properties of Images



AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: Properties of Images



AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: Properties of Images

- An **image** is a collection of pixels.

For example:

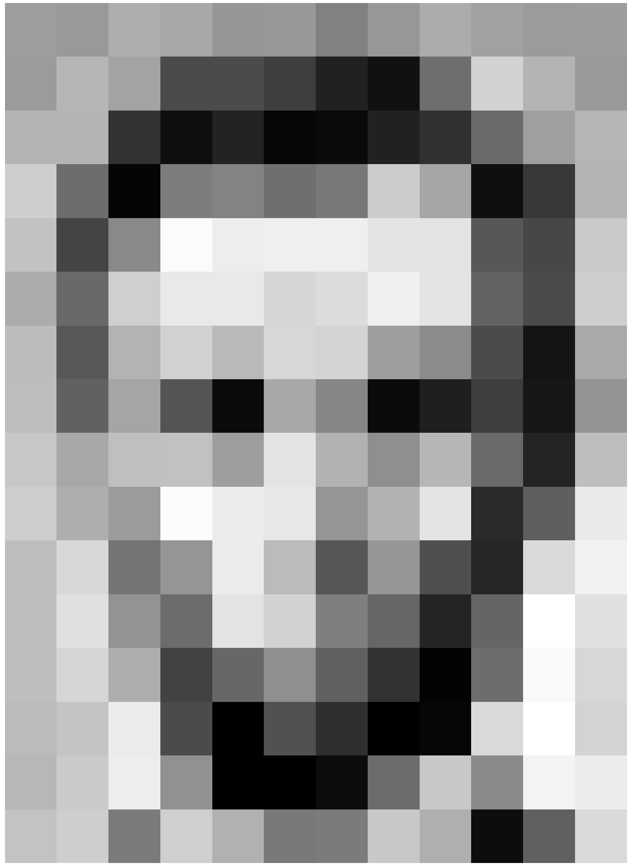
- A 32-by-32 image has 32×32=1024 pixels.

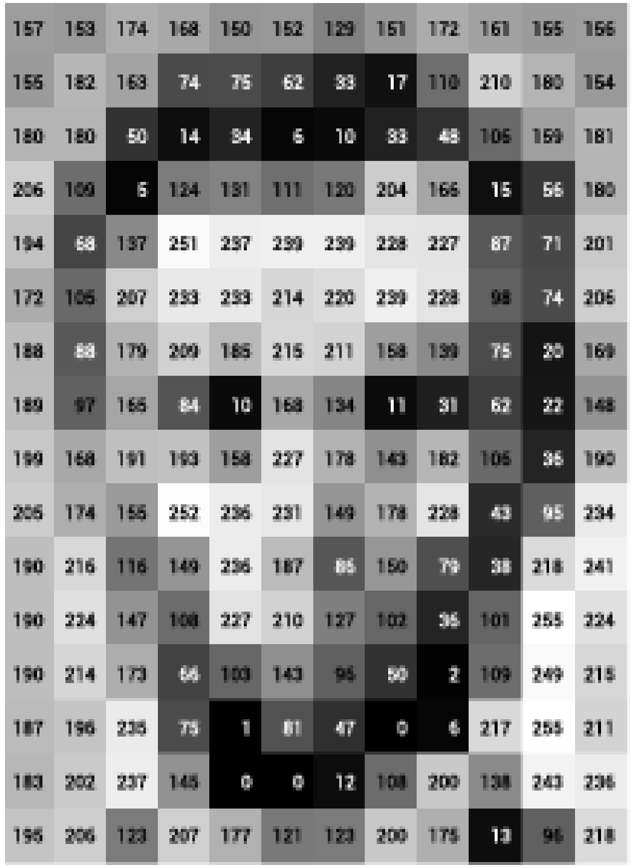
- Each **pixel** is represented by a number in the range [0,255].
  - 000 is black
  - 255 is white.

- **Color images** have **three dimensions**: [Width, Height, Depth]
  - **Width x Height**: 2D size of the image
  - **Depth**: Layers of colors
    - Red, Green, Blue (RGB) values → Depth = 3
    - Gray scale → Depth = 1

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: Properties of Images





157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
206	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

Sources: ai.Stanford.edu

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: Properties of Images

- An image has spatial structure.
- An image has huge dimensionality:
  - A 256x256 RGB image amounts to ~200K input variables
  - Single-layered neural network with 1,000 neurons → 200 million parameters
- An image contains stationary signals that share similar features:
  - Removing variances between signals → the image is still meaningful
  - Small visual changes (often invisible to naked eye)
    - → big changes to input vector, i.e., numbers of variables, parameters

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: Properties

- Preserve spatial structure by convolutional filters
- Tackle huge input dimensionalities by local connectivity and parameter sharing
- Robust to local variances by spatial pooling

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: What is Convolution?

- In mathematics:
  - The **convolution of two functions** f and g is denoted by $*$ as the **integral of the product of the two functions** after one is reversed and shifted.

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau)g(t-\tau)\,d\tau = \int_{-\infty}^{\infty} f(t-\tau)g(\tau)\,d\tau$$

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: What is Convolution?

- A simple explanation with respect to AI deep learning :
  - An **image kernel** is a small matrix used to apply effects like the ones you might find in Photoshop, such as blurring or sharpening.
  - These small image kernel matrices are also used in AI deep learning for **feature extraction**, a technique for determining the most important portions of an image.
  - In the AI deep learning context, the process is referred to more generally as "**convolution**".

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: What is a Convolutional Neural Network?

- A convolutional neural network (CNN or ConvNet):
  - One of the most popular algorithms for deep learning, a type of machine learning in which a model learns to perform classification tasks directly from images, video, text, or sound.
- CNNs are particularly useful for finding patterns in images to recognize objects, faces, and scenes.
  - They learn directly from image data, using patterns to classify images and eliminating the need for manual feature extraction.
- Applications of object recognition and computer vision, such as self-driving vehicles and face-recognition applications, rely heavily on CNNs.
  - Depending on your application, you can build a CNN from scratch, or use a pretrained model with your dataset.

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: What Makes CNNs So Powerful and Popular?

What are the problems with normal multi-layer neural networks?
They do not scale well to handle full images.

- Neural Networks receive an input (a single vector), and transform it through a series of hidden layers.

- Each hidden layer is made up of a set of neurons where:
    - Each neuron is fully connected to all neurons in the previous layer
    - Neurons in a single layer function completely independently and do not share any connections.

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: What Makes CNNs So Powerful and Popular?

The problems with normal neural networks that do not scale well to handle full images:

- Let's consider one real scenario: In CIFAR-10 dataset:
  - Images are only of size 32x32x3 (32 wide, 32 high, 3 color channels): 3072 pixels/image
  - So, a single fully-connected neuron in a first hidden layer of a regular neural network would have 32*32*3 = 3072 weights.
    - Each pixel is an input that is associated with one weight.
  - But this fully-connected structure does not scale to larger images.
    - For example, an image of more respectable size, e.g. 200x200x3, would lead to neurons that have 200*200*3 = 120,000 weights.
  - Moreover:
    - We would almost certainly want to have several such neurons (for each layer), so the parameters would add up quickly!
    - Clearly, this full connectivity is wasteful and the huge number of parameters would quickly lead to overfitting.

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: The Solutions in Modern AI

- Using Convolutional Neural Networks (CNN) are everywhere now.
  - It is arguably the most popular deep learning architecture.
  - The recent surge of interest in deep learning is due to the immense popularity and effectiveness of convolutional neural networks.

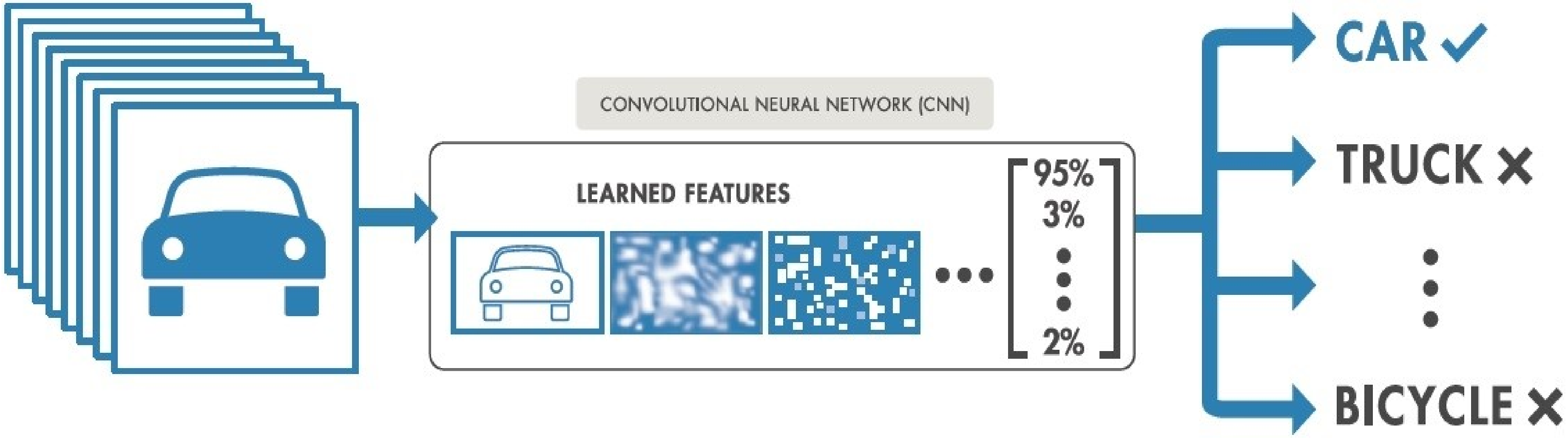
- CNN is now the model of choice on every image related problem.
  - In terms of accuracy, CNN is much better than alternative approaches.
  - CNN can also be applied to recommender systems, natural language processing and more.
  - The main advantage of CNN compared to its predecessors is that it automatically detects the important features without any human supervision.
  - For example, given many pictures of cats and dogs it learns distinctive features for each class by itself.

- CNN is also computationally efficient.
  - It uses special convolution and pooling operations and performs parameter sharing.
  - This enables CNN models to run on any device, making them universally attractive.

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: What Makes CNNs So Powerful and Popular?

- Using CNNs for deep learning has become increasingly popular due to three important factors:
  - CNNs eliminate the need for manual feature extraction—the features are learned directly by the CNN.
  - CNNs produce state-of-the-art recognition results.
  - CNNs can be retrained for new recognition tasks, enabling you to build on pre-existing networks.



Sources: http://www.mathworks.com

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: Convolutional Layers, Convolution Filters, and Feature Maps

- The **main building block** of Convolutional Neural Networks is the **convolutional layer**.
- **Convolution** is a mathematical operation to merge two sets of information, i.e., two functions.
- In AI deep learning, the convolution is applied on the input data, e.g., data sets of image pixels, using a **convolution(al) filter** to produce a **feature map**.

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

Input

1	0	1
0	1	0
1	0	1

Filter / Kernel

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: Examples of Convolution Filters

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

Input

1	0	1
0	1	0
1	0	1

Filter / Kernel

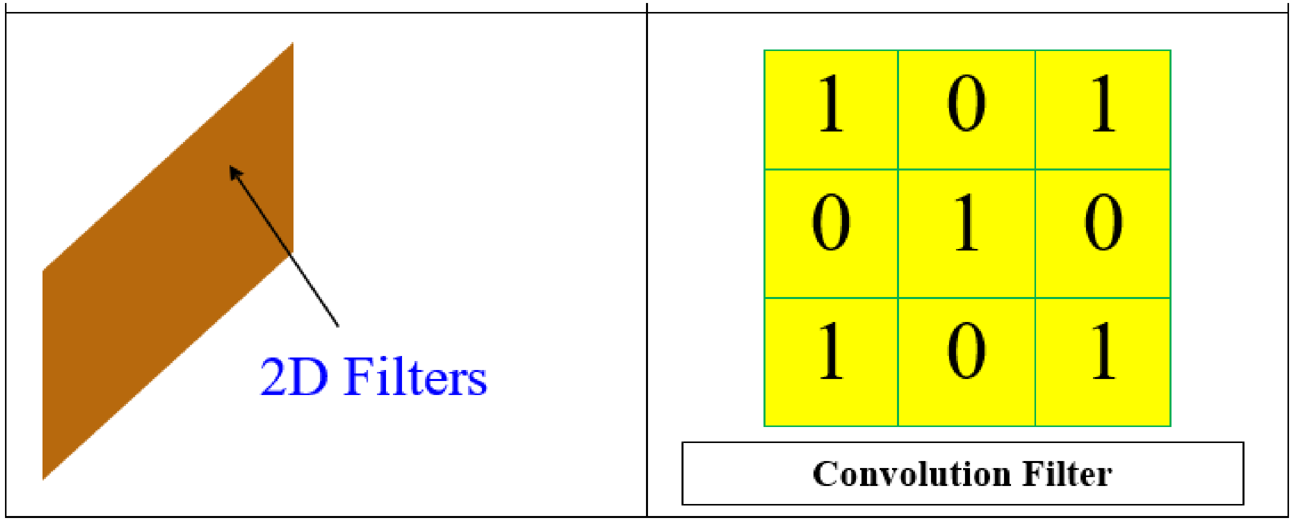
- In the image to the left is the **input** to the **convolution layer**, e.g., the input image.
- In the image to the right is the **convolution filter**, a.k.a. the kernel. This is called a 3x3 convolution due to the shape of the filter.

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: Convolution Filters: 2D Filters

- If images are 2-D, parameters should also be organized in 2-D:
  - → They can learn the local correlations between input variables.
  - → They can "exploit" the spatial nature of images.

Original Image

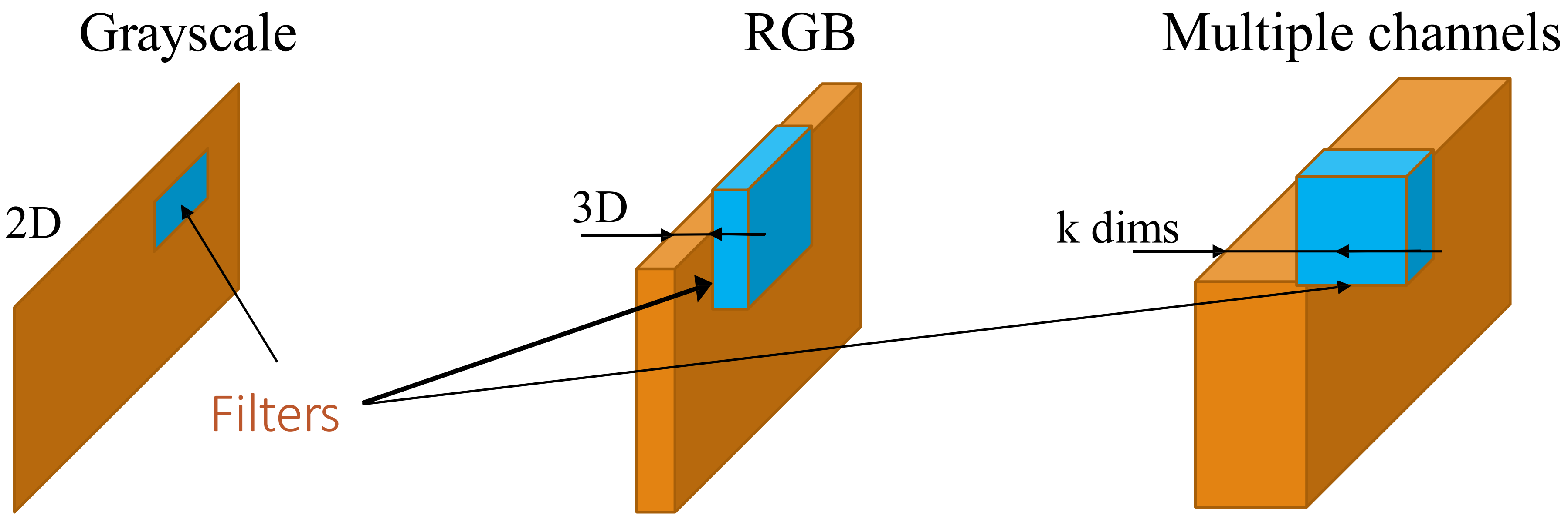


Convolution Filter

AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: Convolution Filters: k-D Filters

- Similarly, if images are k-D, parameters should also be k-D:
  - 2D filter: Height * Width (Color: Grayscale)
  - 3D filter: Height * Width * Depth (Colors: RGB: 3 channels: R, G, B)
  - k-D filter: Height * Width * Depth (Colors: RGB) * Light Exposure * … (more) → k dimensions



AI Deep Learning: Convolutional Neural Networks (CNN)

CNN: Convolutional Filters: k-D Filters

- Similarly, if images are k-D, parameters should also be k-D:
  - For example: A 3D filter

3-D Filter