Image Classification with PyTorch

PRE-PROCESSING IMAGES TO USE IN MACHINE LEARNING MODELS

Janani Ravi
CO-FOUNDER, LOONYCORN
www.loonycorn.com

Overview

Image classification using machine learning

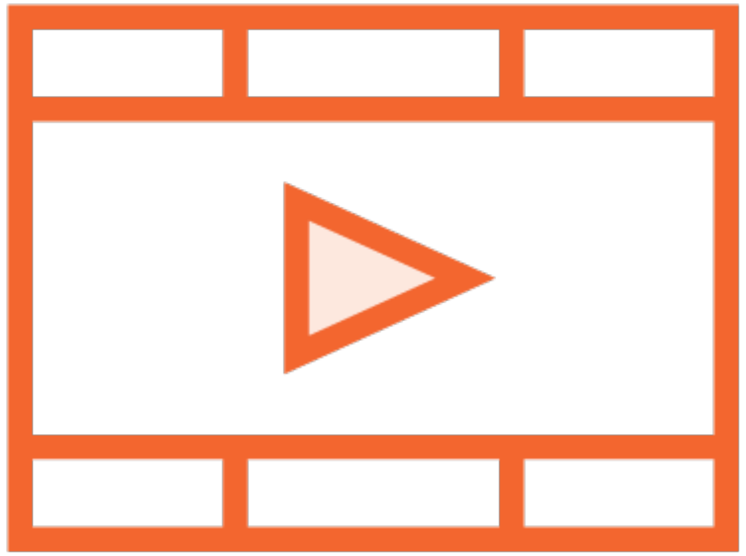
Representing images as tensors

Need for image pre-processing

Common image pre-processing techniques

Prerequisites and Course Outline

Prerequisites



Basic Python programming

Build and training machine learning models

Worked with PyTorch to build simple neural networks

Prerequisite Courses

Foundations of PyTorch

Building your first PyTorch solution

Course Outline

Images as features and pre-processing techniques

Drawbacks of Deep Neural Networks (DNNs) for image classification

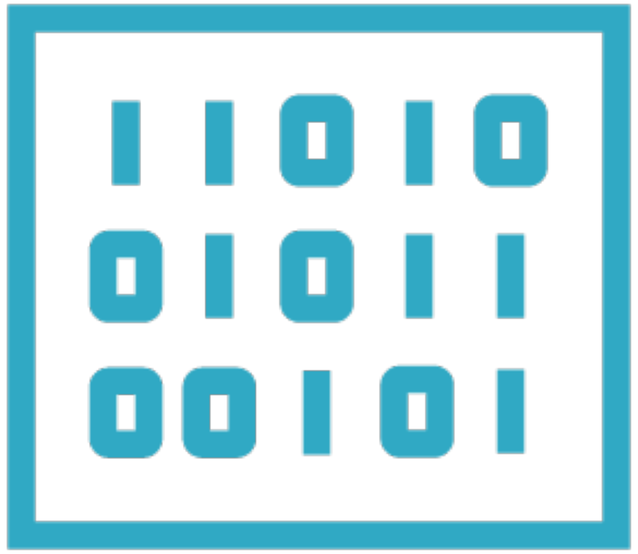
Introducing Convolutional Neural Networks (CNNs)

Hyperparameter tuning

Pre-trained models

Image Recognition

Image Recognition



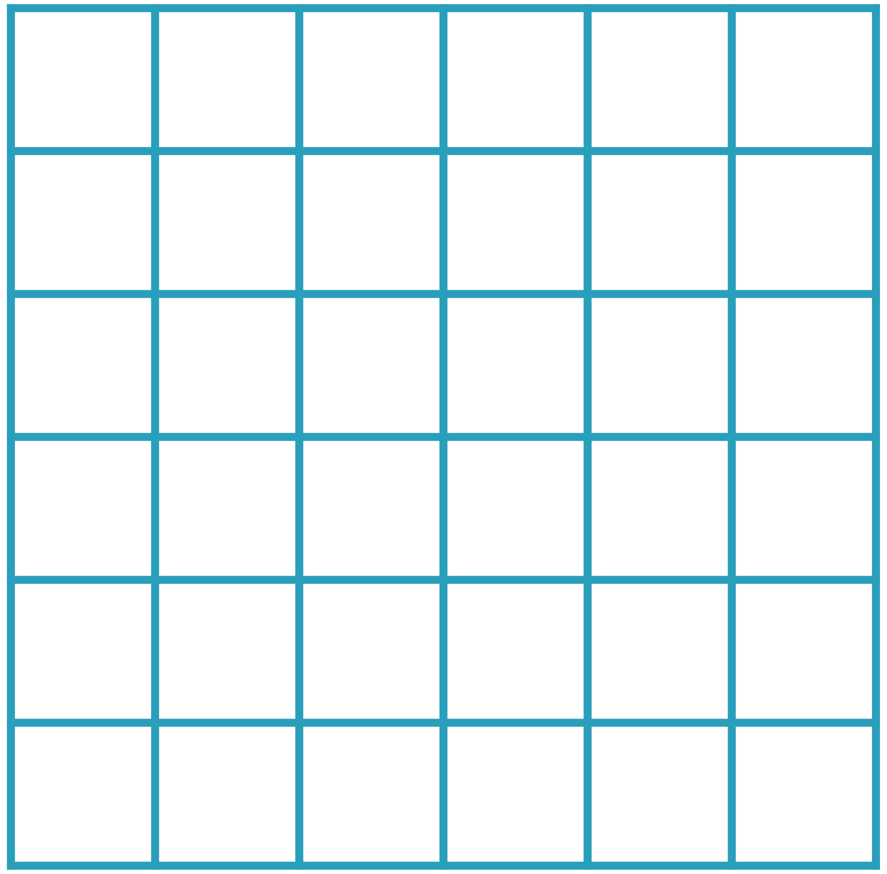
Images represented as pixels

Identify edges, colors, shapes

A photo of a horse

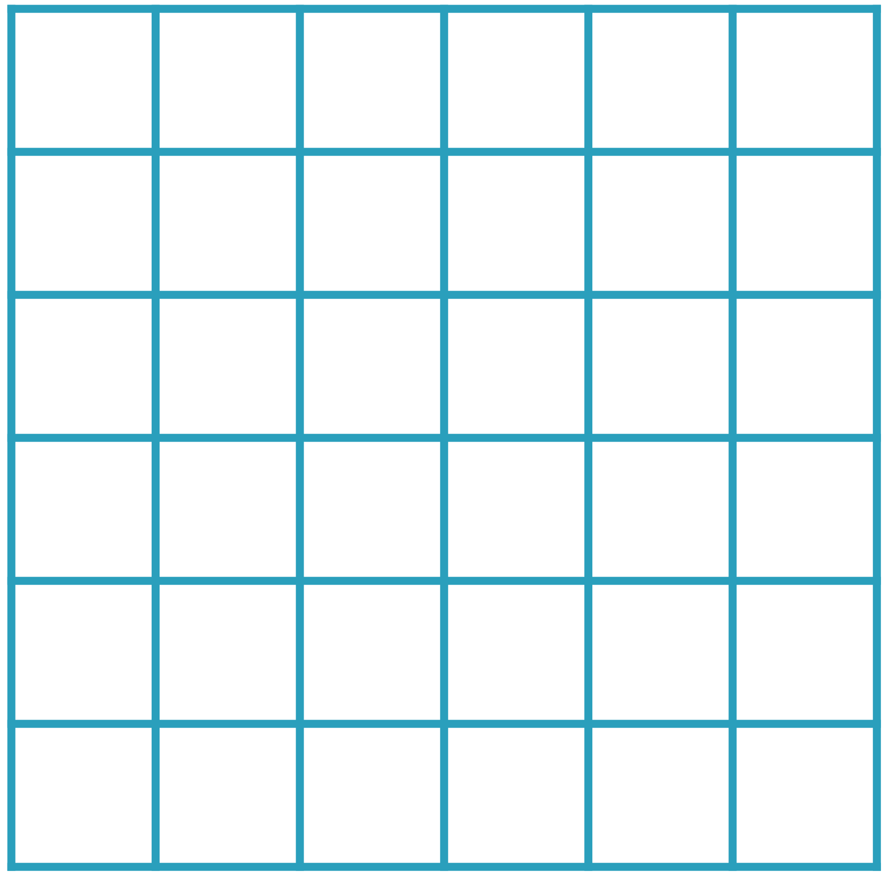
Images as Tensors

Images as Tensors



Each pixel holds a value based on the type of image

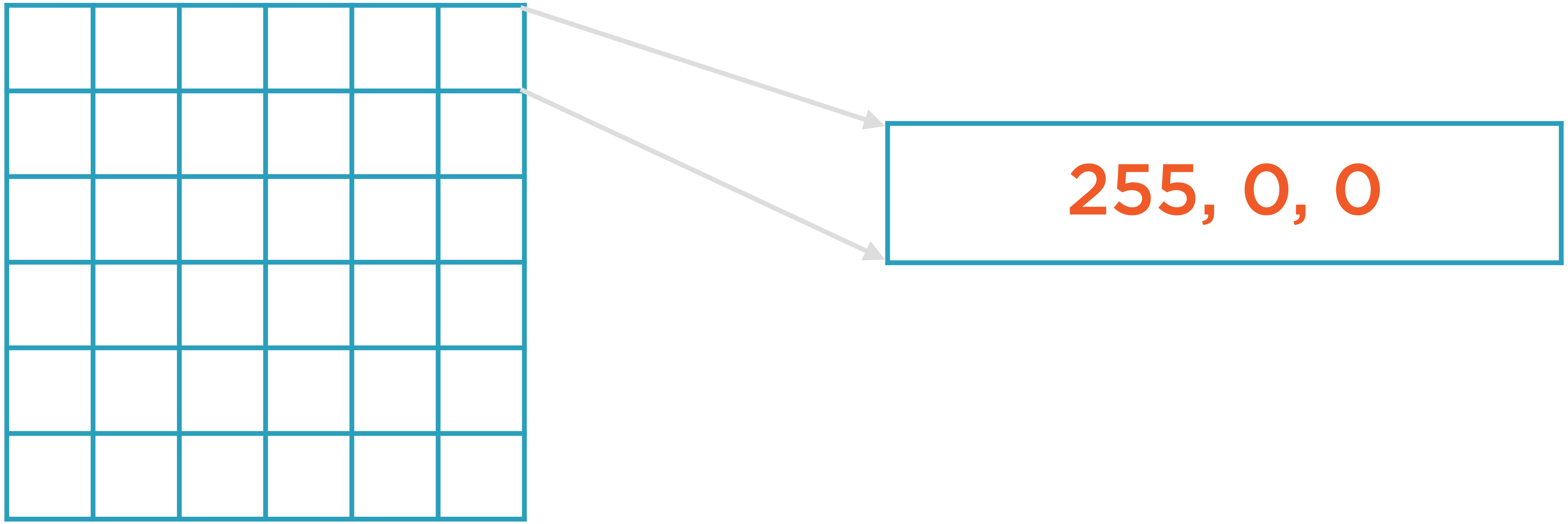
RGB Images



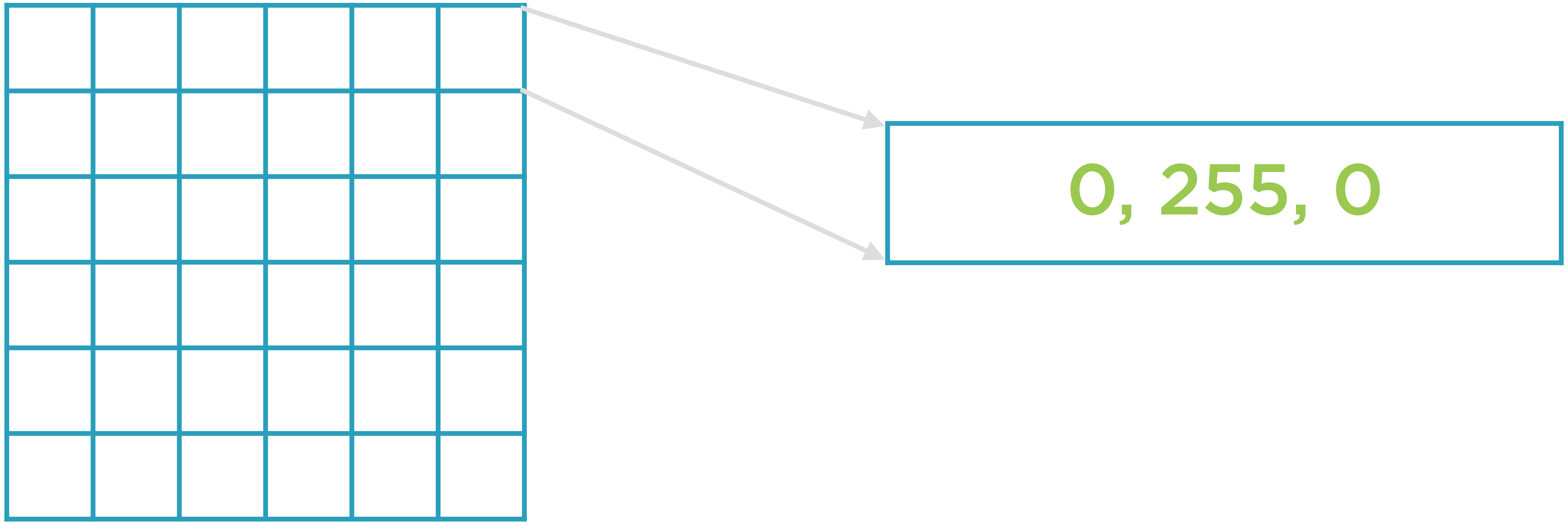
RGB values are for color images

R, G, B: 0-255

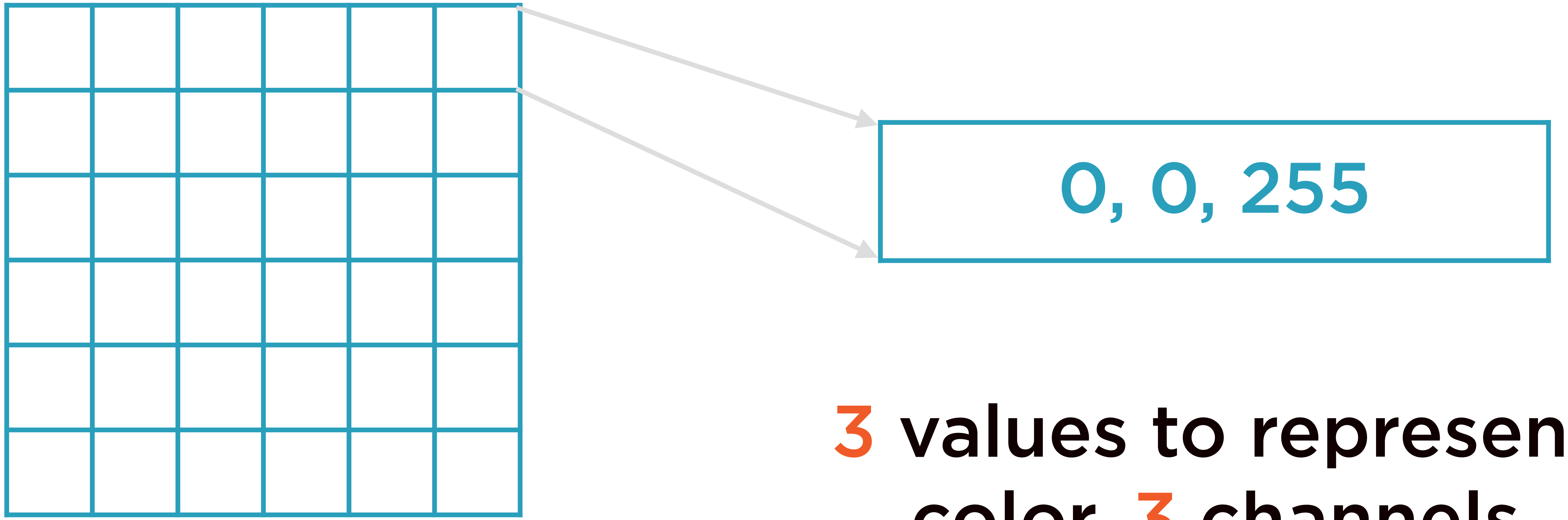
RGB Images



RGB Images

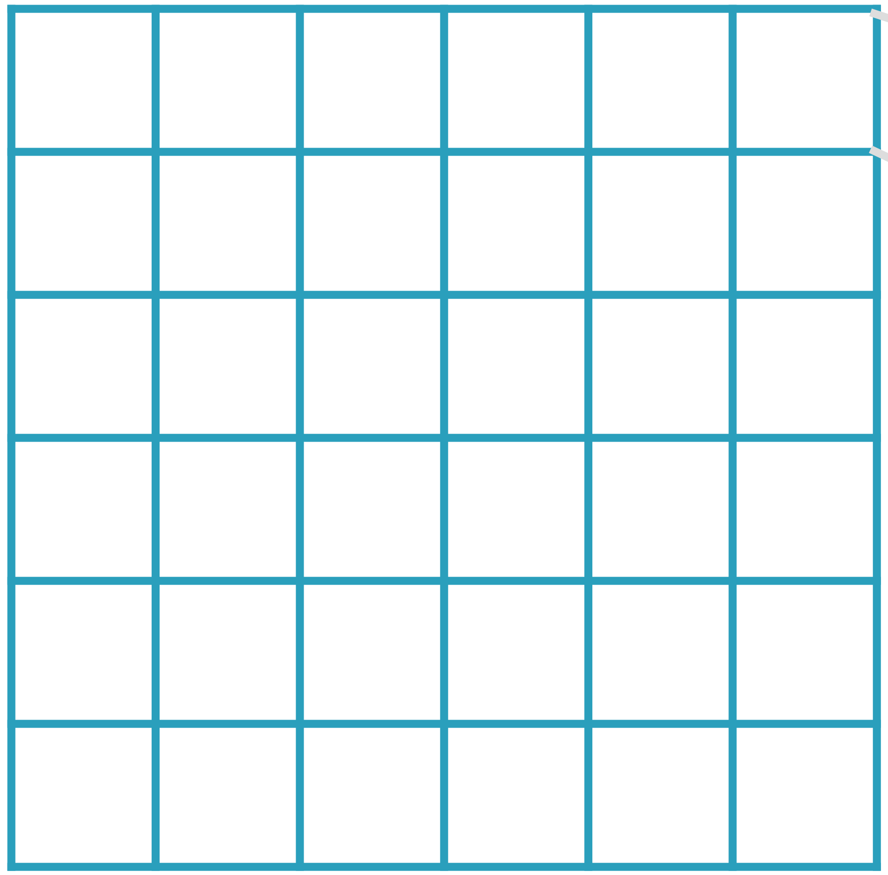


RGB Images



3 values to represent color, 3 channels

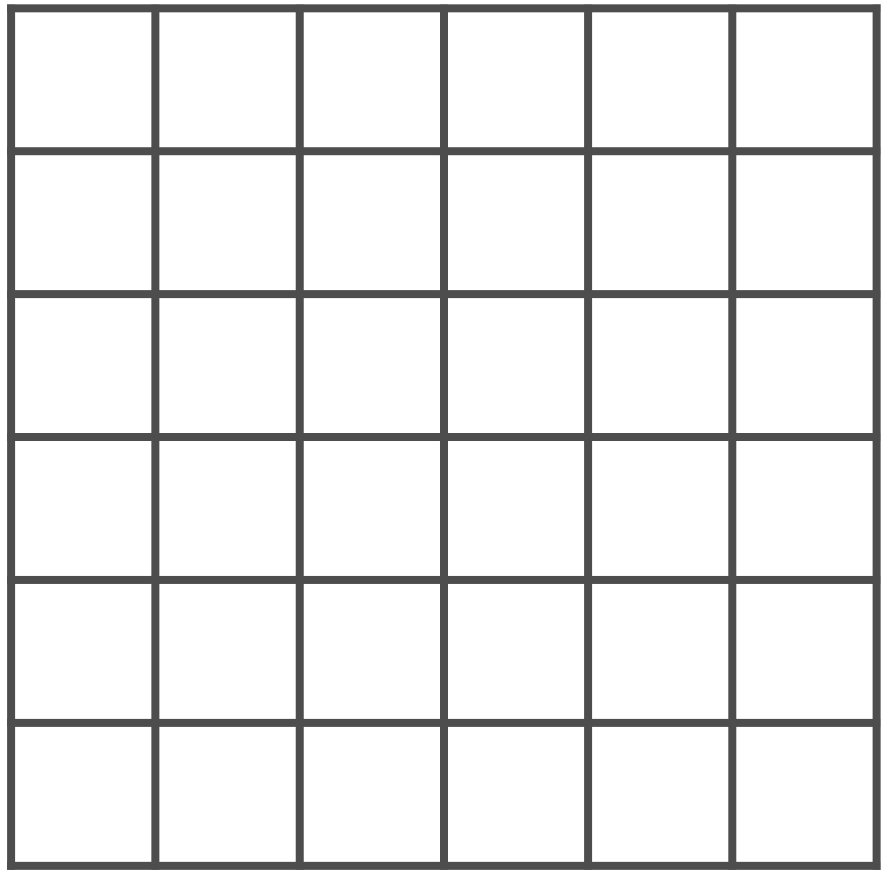
RGB Images



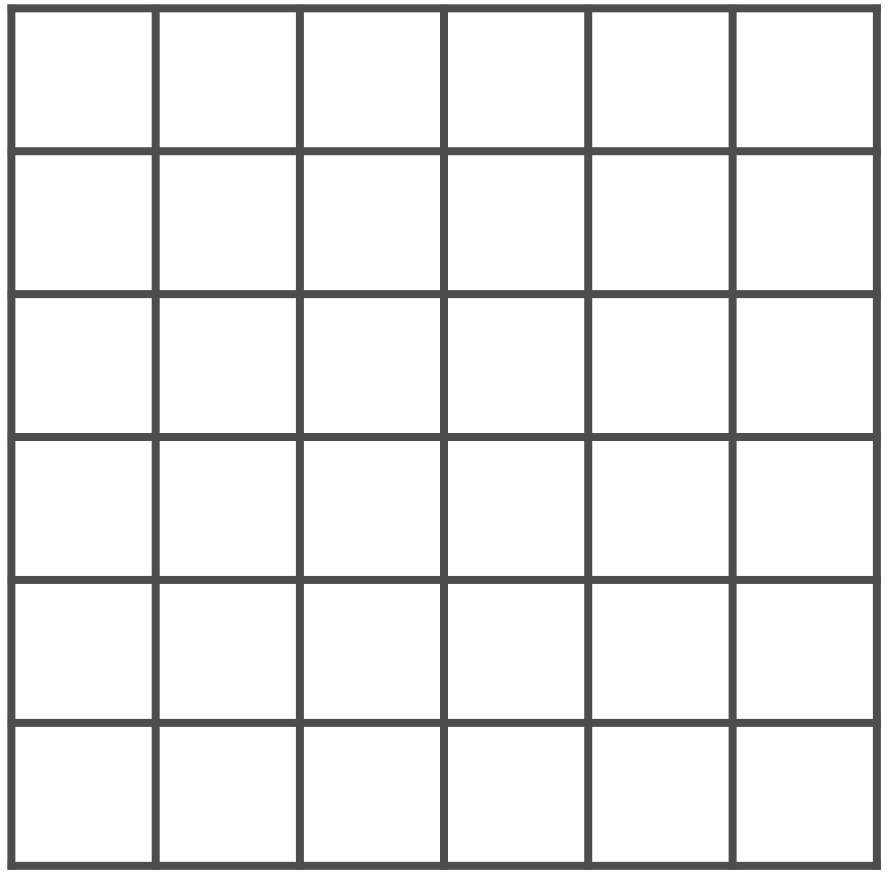
0, 0, 255

These are often scaled to be in the 0-1 range as neural networks work better with smaller numbers

Grayscale Images



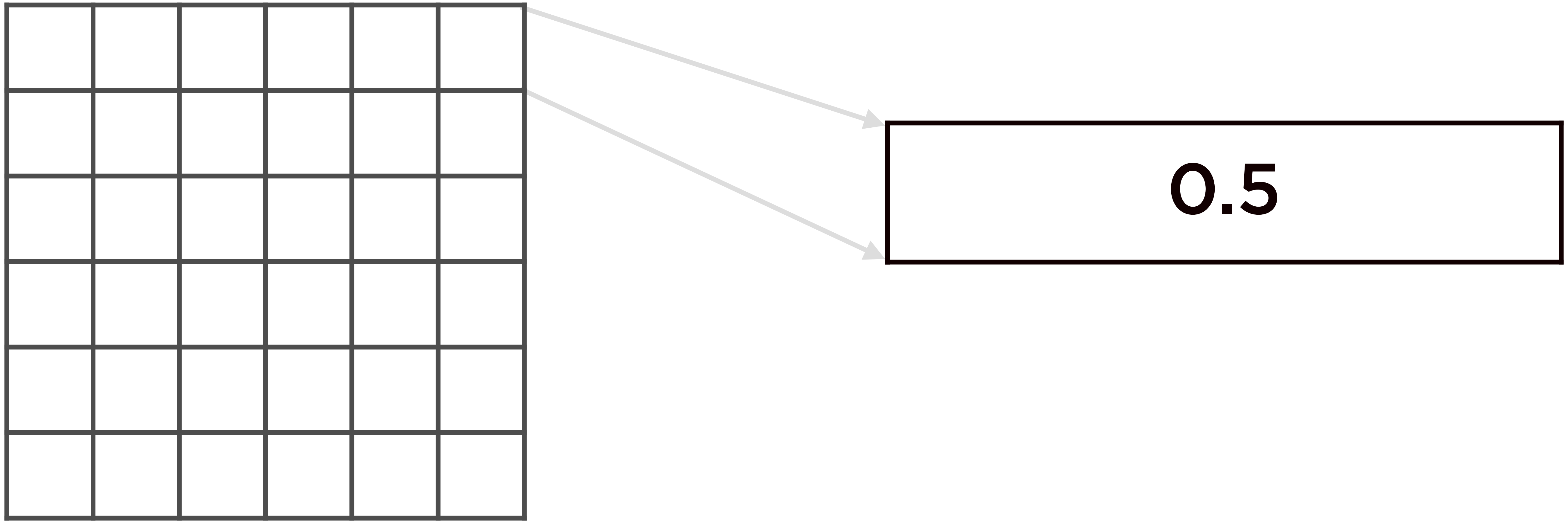
Grayscale Images



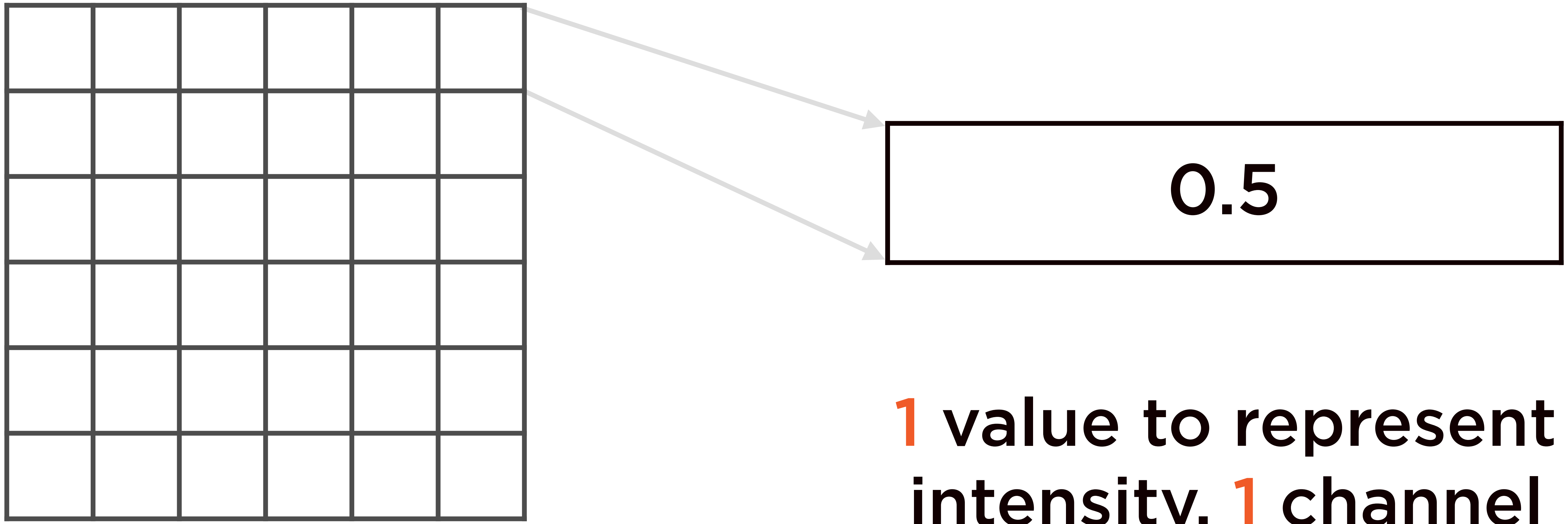
Each pixel represents only intensity information

0.0 - 1.0

Grayscale Images



Grayscale Images

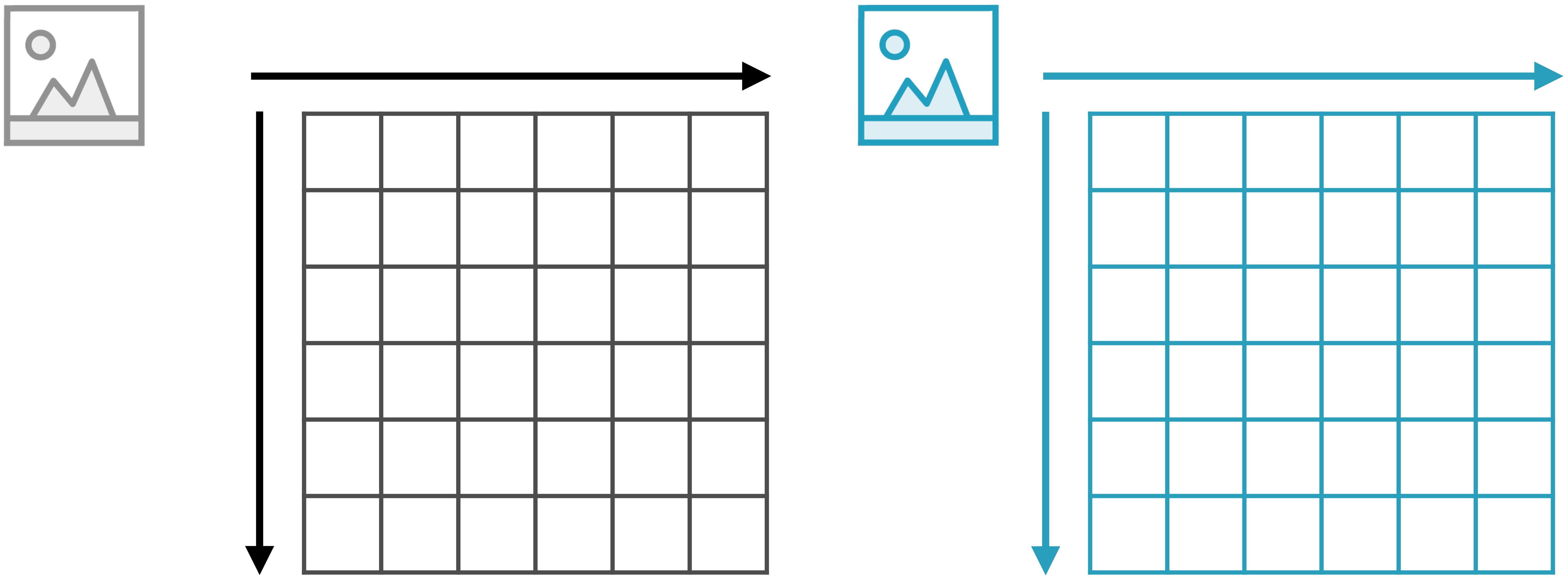


1 value to represent intensity, 1 channel

Images as Tensors

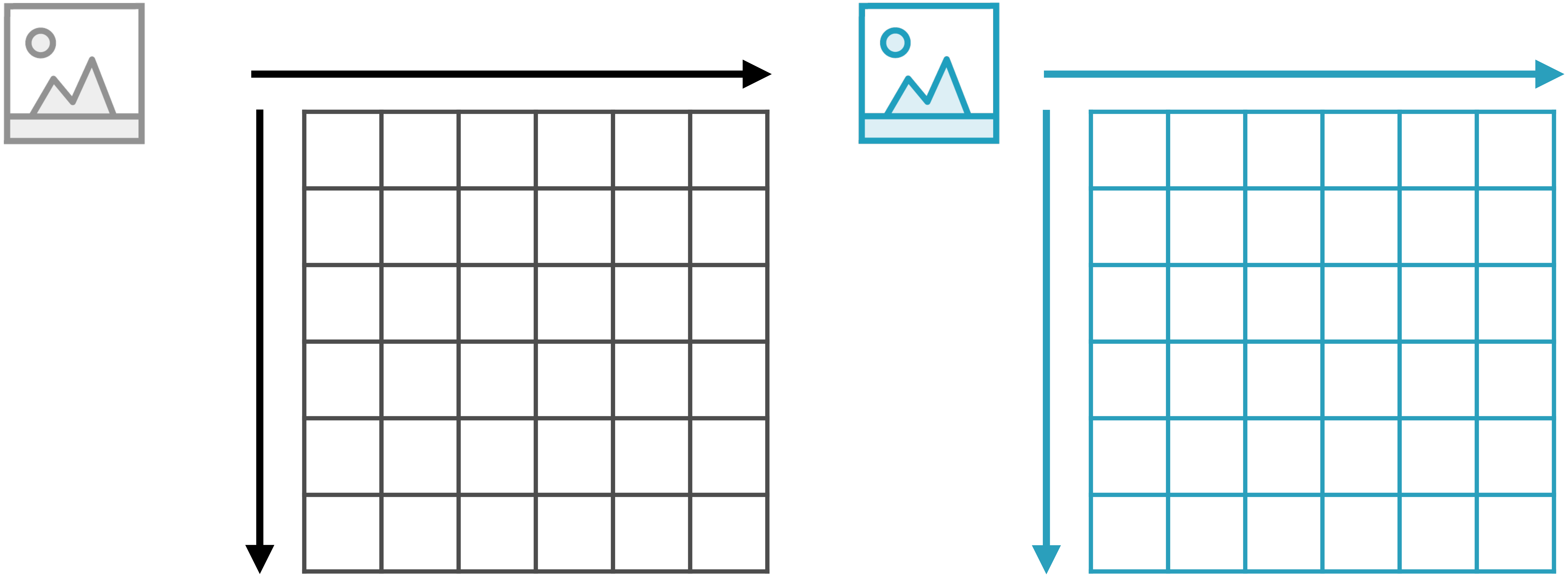
Single channel and multi-channel images

Images as Tensors



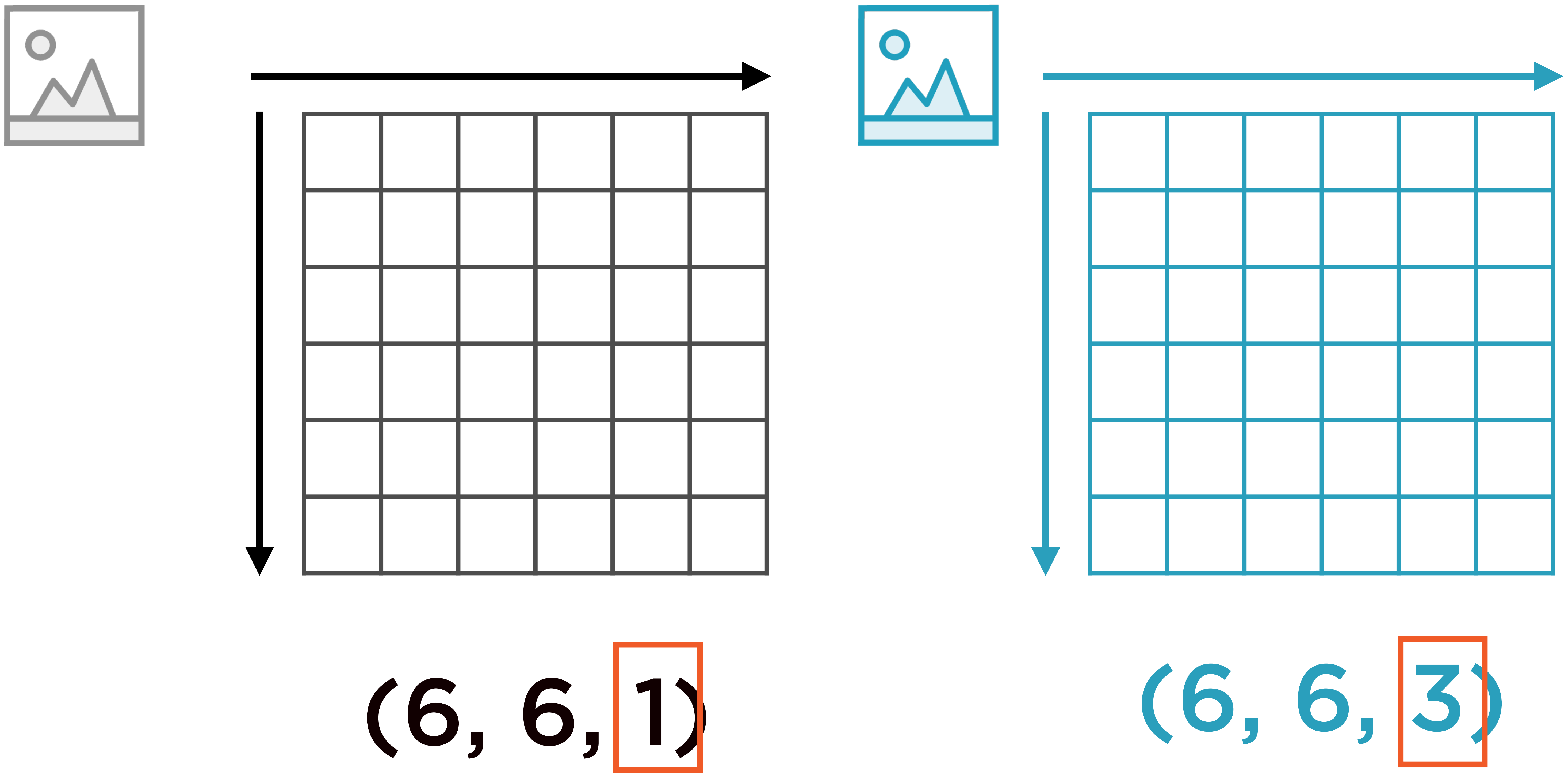
Images can be represented by a 3-D matrix

Images as Tensors



The number of channels specifies the number of elements in the 3rd dimension

Images as Tensors



List of Images

Deep learning frameworks usually deal with a list of images in one 4-D tensor

List of Images

The images should all be the same size

List of Images

$$(10, 6, 6, \boxed{3})$$

The number of channels

List of Images

$(10, \boxed{6, 6}, 3)$

The height and width of each image in the list

List of Images

($\boxed{10,}$ 6, 6, 3)

The number of images

Need for Image Pre-processing

Image Pre-processing Methods

Uniform Aspect Ratio	Uniform Image Size	Mean and Perturbed Images
Normalized Image Inputs	Dimensionality Reduction	Data Augmentation

Common techniques to improve CNN performance

Image Pre-processing Methods

Uniform Aspect Ratio	**Uniform Image Size**	**Mean and Perturbed Images**
Normalized Image Inputs	**Dimensionality Reduction**	**Data Augmentation**

Common techniques to improve CNN performance

Uniform Aspect Ratio

- Most models assume square shape
- Crop images to be square
- Usually, center of image most important
- Makes aspect ratio constant

Image Pre-processing Methods

Uniform Aspect Ratio	**Uniform Image Size**	Mean and Perturbed Images
Normalized Image Inputs	Dimensionality Reduction	Data Augmentation

Common techniques to improve CNN performance

Uniform Image Size

Fit image size to CNN feature maps

250 x 250 image to 100 x 100 image

Down-scaling factor of 0.4

Up-scaling and down-scaling

Image Pre-processing Methods

Uniform Aspect Ratio	Uniform Image Size	**Mean and Perturbed Images**
Normalized Image Inputs	Dimensionality Reduction	Data Augmentation

Common techniques to improve CNN performance

Mean and Perturbed Images

Mean image: average pixel across entire training dataset

Insights often emerge

E.g. faces usually in center of image

Mean and Perturbed Images

Perturbed image: intentionally distort pixels by varying them from mean image

E.g. to prevent CNN from only focusing on center

Image Pre-processing Methods

Uniform Aspect Ratio	Uniform Image Size	Mean and Perturbed Images
Normalized Image Inputs	Dimensionality Reduction	Data Augmentation

Common techniques to improve CNN performance

Normalized Image Inputs

"Normalize" each pixel

Subtract mean

Divide by standard deviation

Ensures each pixel has similar data distribution

Normalized Image Inputs

Converts pixels to N(0,1) distribution

Then scale to be in [0,1] or [0,255]

Helps neural networks converge faster

Image Pre-processing Methods

Uniform Aspect Ratio

Uniform Image Size

Mean and Perturbed Images

Normalized Image Inputs

Dimensionality Reduction

Data Augmentation

Common techniques to improve CNN performance

Dimensionality Reduction

RGB data has 3 channels

Can reduce to grayscale (just 1 channel)

Reduces dimensionality of all image tensors

Reduce the size of the problem so training completes faster

Image Pre-processing Methods

Uniform Aspect Ratio	Uniform Image Size	Mean and Perturbed Images
Normalized Image Inputs	Dimensionality Reduction	Data Augmentation

Common techniques to improve CNN performance

Data Augmentation

Perturbed images are a form of data augmentation

Scaling, rotation, affine transforms

Makes CNN training more robust

Reduces risk of overfitting

Demo

Set up a deep learning VM on a cloud platform

Demo

Explore common image pre-processing techniques

Demo

Implement image pre-processing using PyTorch

Summary

Image classification using machine learning

Representing images as tensors

Need for image pre-processing

Common image pre-processing techniques