File Name: T-ICML-I_M0_L1_introduction

Format: Presenter in Studio

Presenter: Evan Jones

# Image Classification Models

Evan Jones

# Specialization

End-to-End Lab on Structured Data

Production ML Systems

**Image Classification Models**

Sequence Models

Recommendation Systems

# Agenda

**Introduction**

Linear and DNN models

Convolutional neural networks

Dealing with Data Scarcity

Going Deeper Faster

Pre-built models

# Agenda

Introduction

**Linear and DNN models**

Convolutional neural networks

Dealing with Data Scarcity

Going Deeper Faster

Pre-built models

# Agenda

Introduction

Linear and DNN models

**Convolutional neural networks**

Dealing with Data Scarcity

Going Deeper Faster

Pre-built models

# Agenda

Introduction

Linear and DNN models

Convolutional neural networks

**Dealing with Data Scarcity**

Going Deeper Faster

Pre-built models

# Agenda

Introduction

Linear and DNN models

Convolutional neural networks

Dealing with Data Scarcity

**Going Deeper Faster**

Pre-built models

# Agenda

Introduction

Linear and DNN models

Convolutional neural networks

Dealing with Data Scarcity

Going Deeper Faster

**Pre-built models**

File Name: Qwiklabs-onboard-lab-M0-overview_V2

Format: Presenter in Studio

Presenter: Evan Jones

File Name: T-ICML-O_M0_l3_introduction

Format: Presenter in Studio

Presenter: Evan Jones

# Learn how to...

Recognize the applications for modern image classification models

Breakdown images as visual data (height, width, and depth)

Understand the limitations of comparing image data with traditional methods

# Advances in Computer Vision

# Sensors are now commonplace

# Visual data growing at an incredible rate

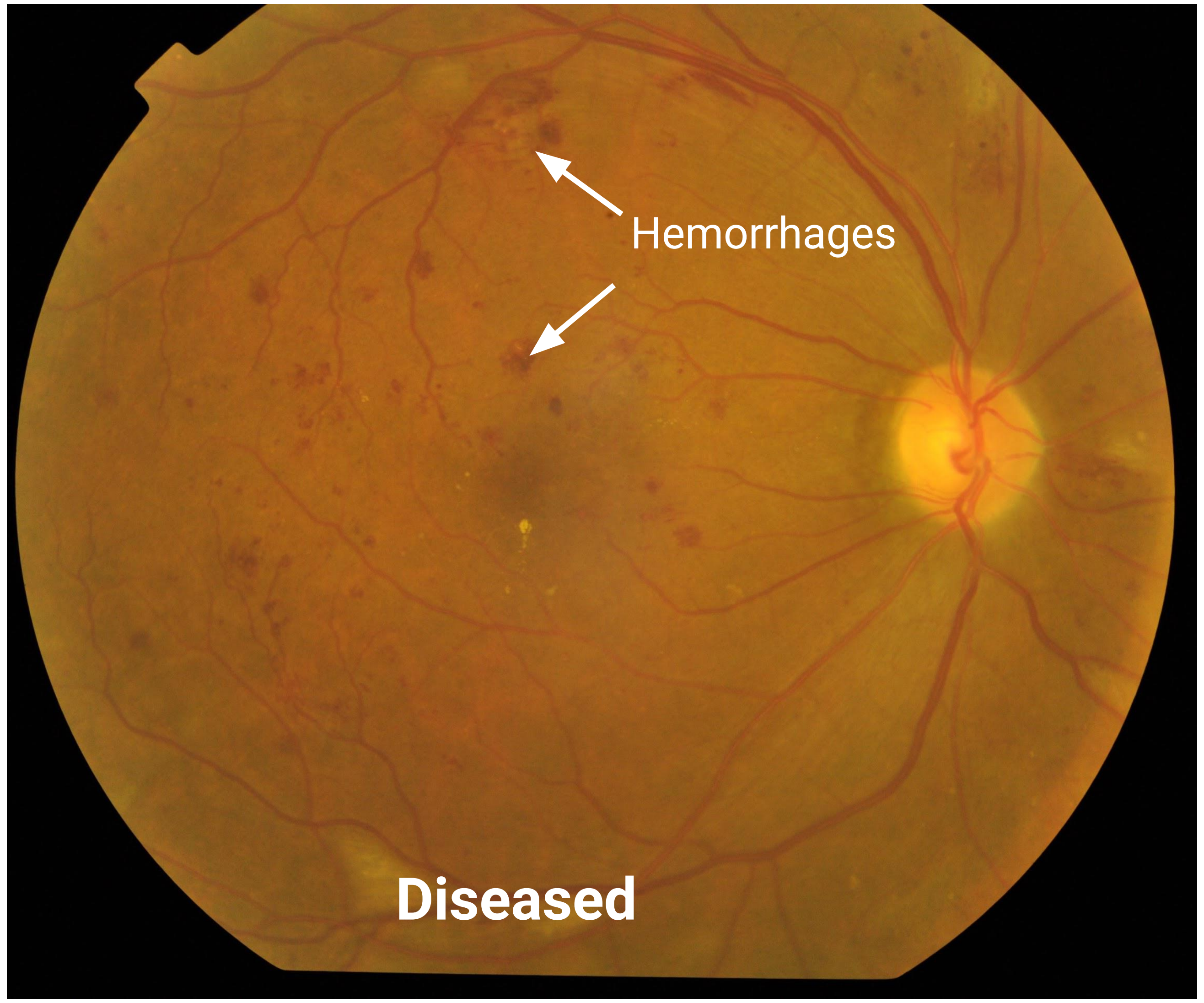
**High demand** for interpreting the meaning of images at scale

# Clouds or snow-capped mountains?

# Good or bad diced potatoes?

# Empty or Full?

# Diagnosing Diabetic Retinopathy

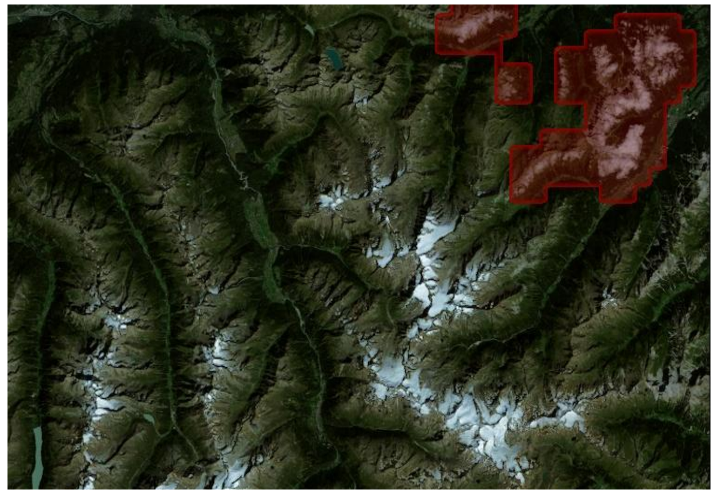
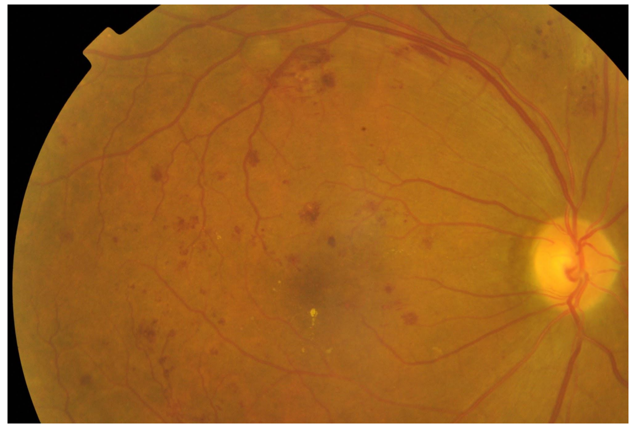
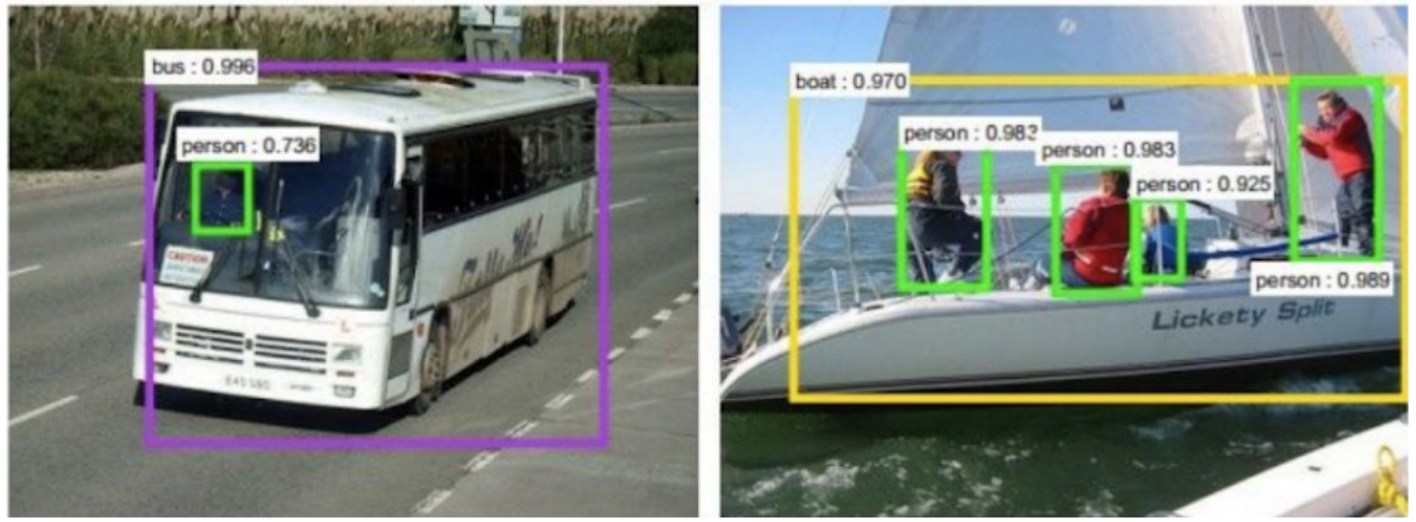


Image classification
**automates tasks** that are easy
(and not easy) for humans

# Labeling


# Object Detection




**Show and Tell: A Neural Image Caption Generator** Vinyals et al 2015:
https://arxiv.org/abs/1411.4555

# Pose Detection

g.co/movemirror

File Name: T-ICML-O_M0_I4_structured_vs_unstructured_data

Format: Presenter in Studio

Presenter: Evan Jones

## Review: Structured Data represented as a vector

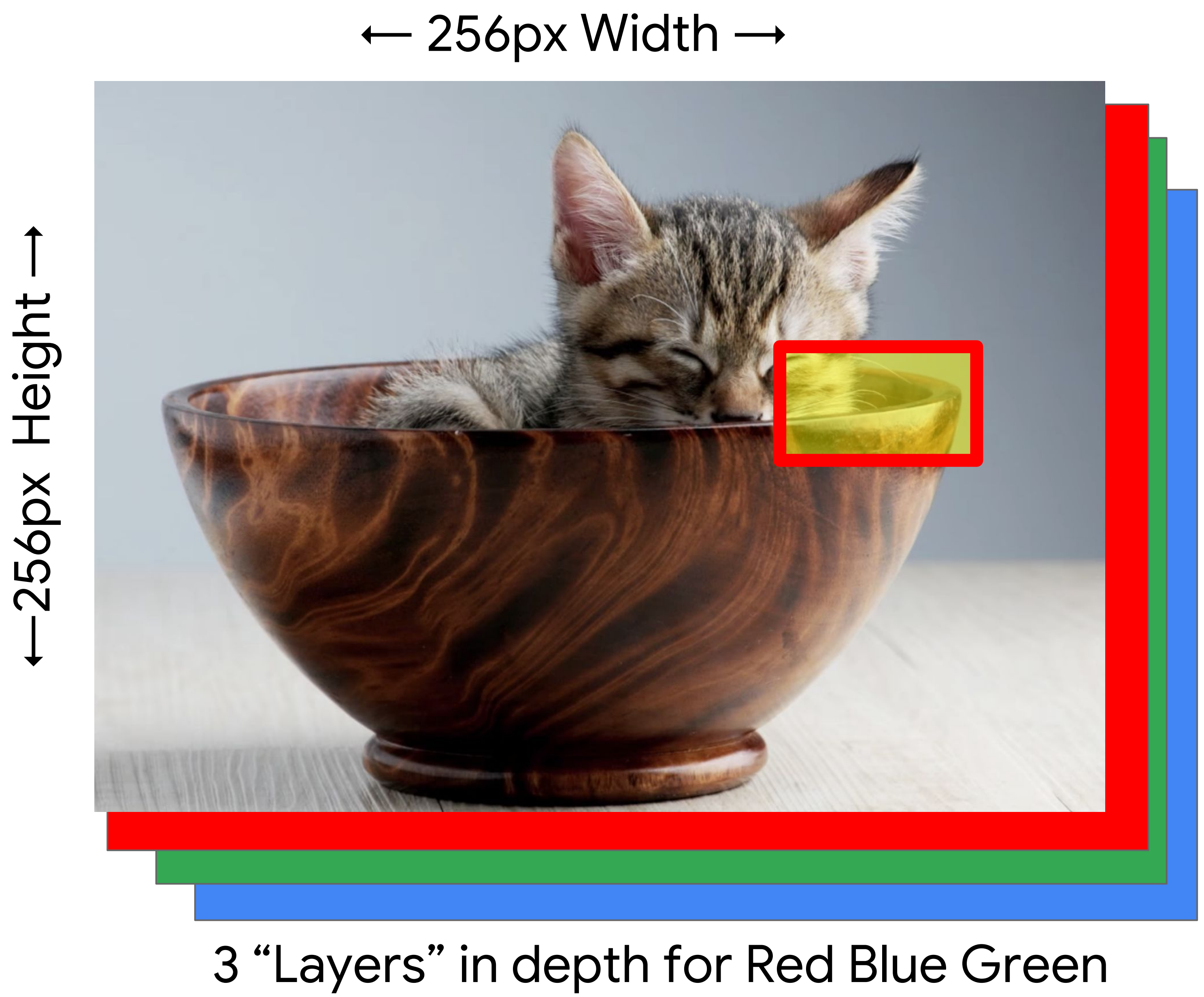
Cat's height and weight

[9.4, 7.9]

# Image Data is multi-dimensional

# Image Data has depth which represents color intensity

# Image Data has depth which represents color intensity




# A single high-res image can represent millions of weights to learn

← **3264px** Width →

←**2448px** Height →

3 “Layers” in depth for Red Blue Green

## 8 Megapixel resolution

3264 (w) x 2448 (h) x 3 (RGB) =

23,970,816 per image*

* ML training time impacted

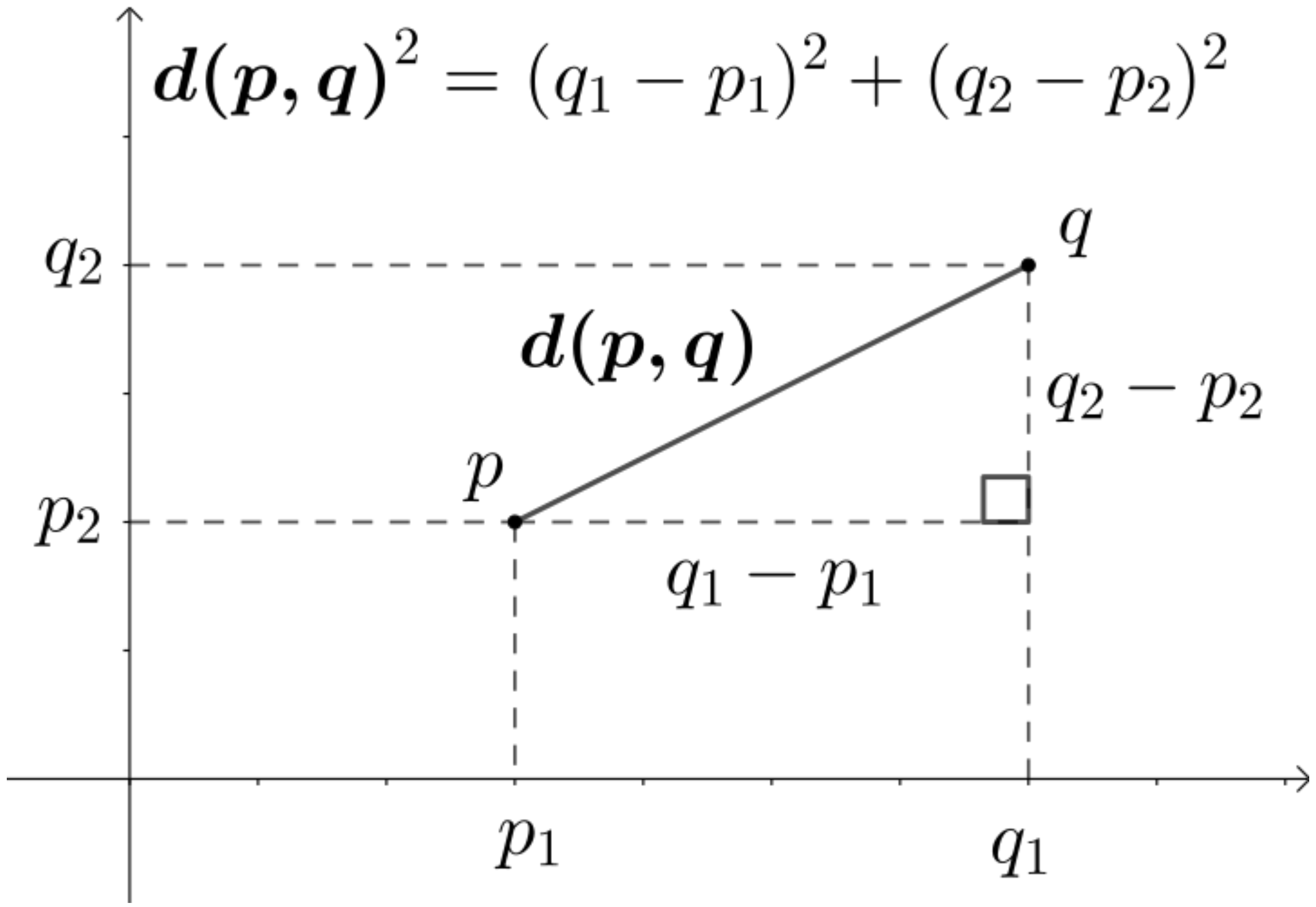
# Traditional ML methods do not handle translations well

← 256px Width →

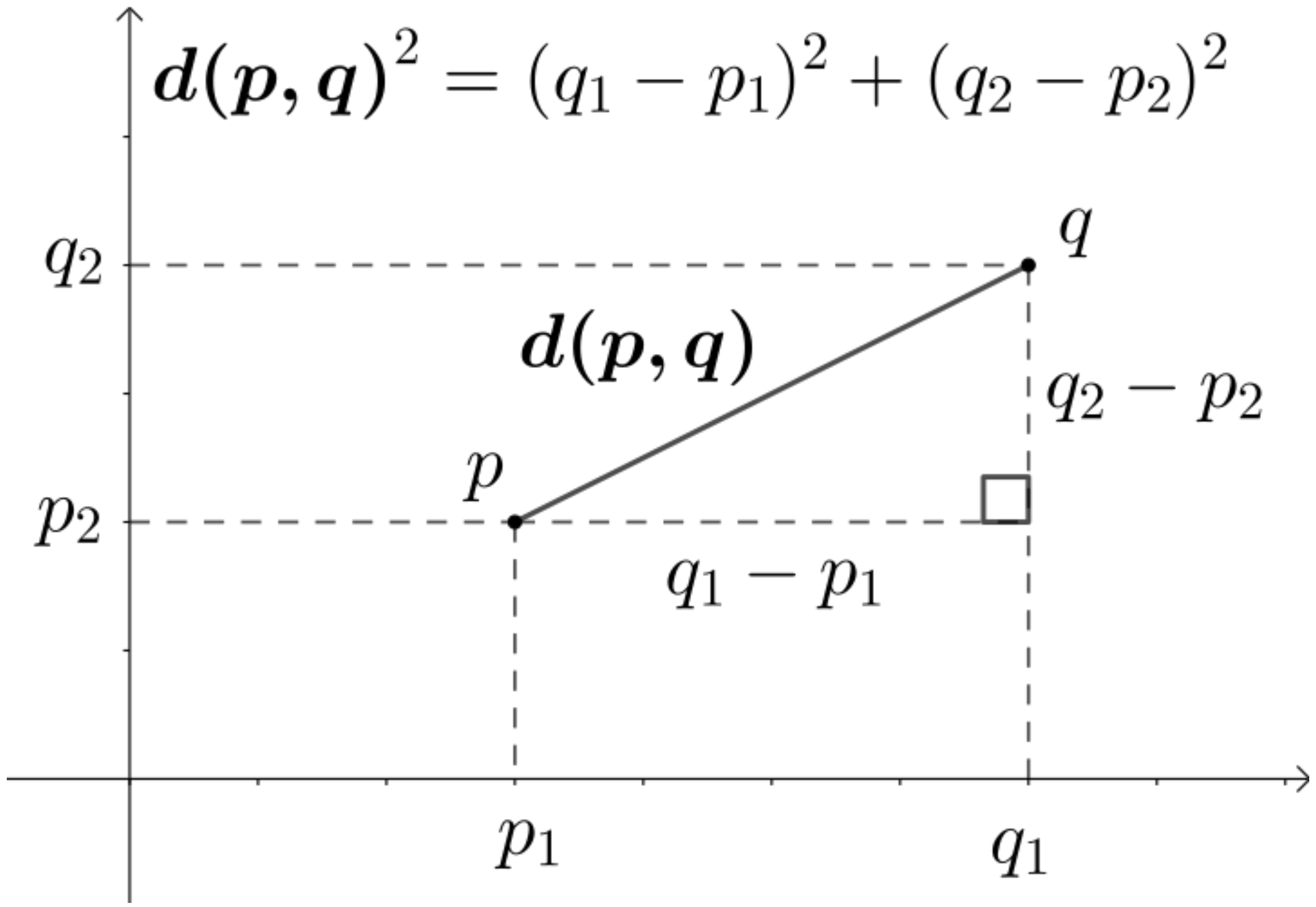
← 256px Width →

# Comparing pixel for pixel the image has significantly changed

# Comparing feature vectors with straight-line distance



# Comparing feature vectors with straight-line distance works well for structured data




| | Cat Height (in) | Cat Weight (lbs) |
|---|---|---|
| 1 | 9.4 | 7.9 |
| 2 | 9.7 | 9.9 |
| 3 | 42 | 320 |

# Comparing feature vectors with straight-line distance works well for structured data

**Compare distance**

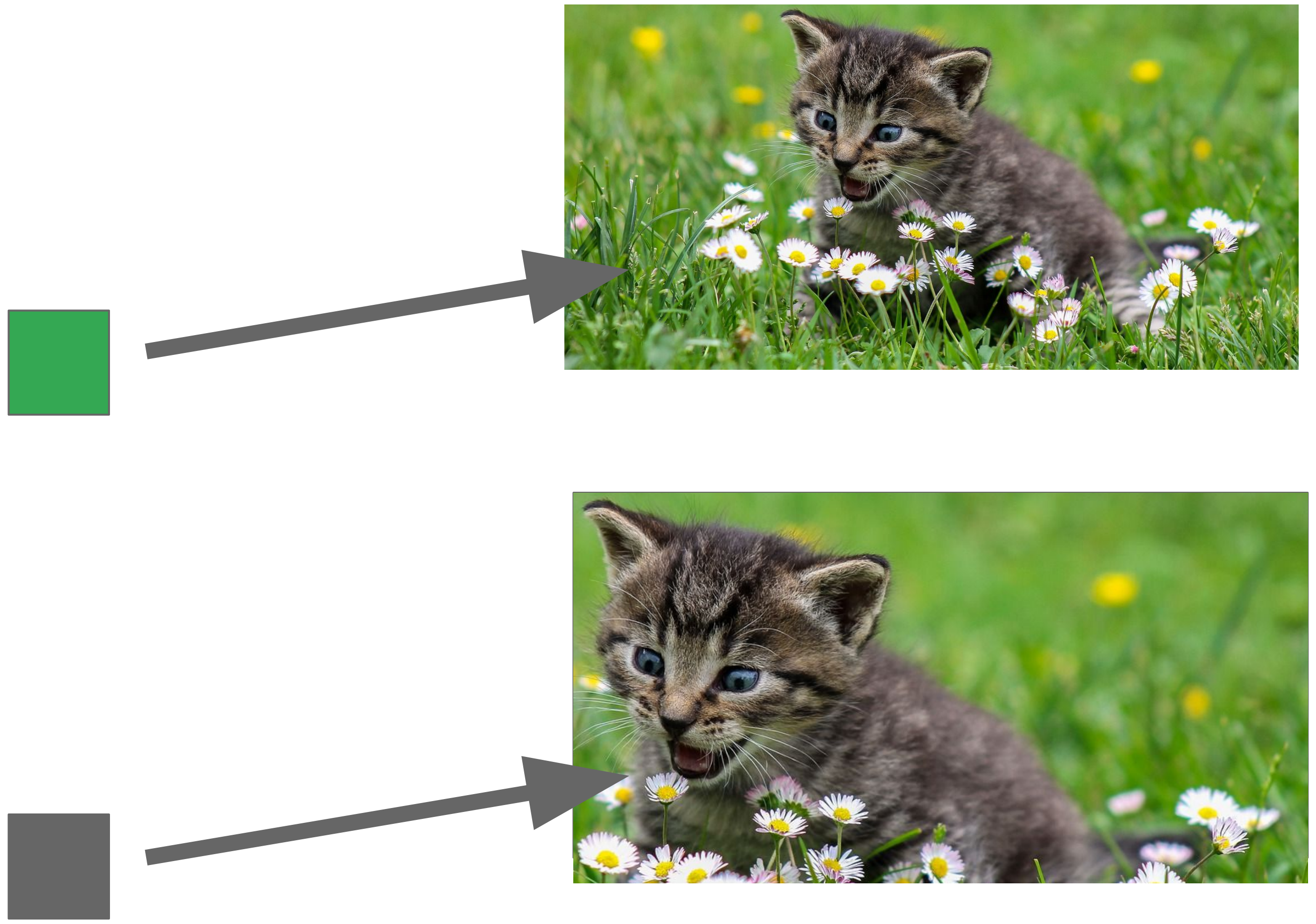
between two house cats
$(9.4 - 9.7)^2 + (7.9 - 9.9)^2 =$
$\sqrt{4.09} =$ **2.02**

between a cat and tiger
$(9.4 - 42)^2 + (7.9 - 320)^2 =$
$\sqrt{98,469} =$ **313.79**

| | **Cat Height (in)** | **Cat Weight (lbs)** |
|---|---|---|
| 1 | 9.4 | 7.9 |
| 2 | 9.7 | 9.9 |
| 3 | 42 | 320 |

# Comparing feature vectors with straight-line distance <u>does not work well</u> for unstructured data

# Comparing feature vectors with straight-line distance <u>does not work well</u> for unstructured data



# Humans can "see" the same image regardless of translation

## Humans can “see” the same image regardless of translation

## Humans can “see” the same image regardless of translation

# Humans can “see” the same image regardless of translation

# Humans can “see” the same image regardless of translation

The ideal decision boundary
of cat vs non-cat

is defined by the relationship
between the pixels

## How can you model the relationship between pixels?

??