

# ML Workshop Day 1

Audun Mathias Øygard

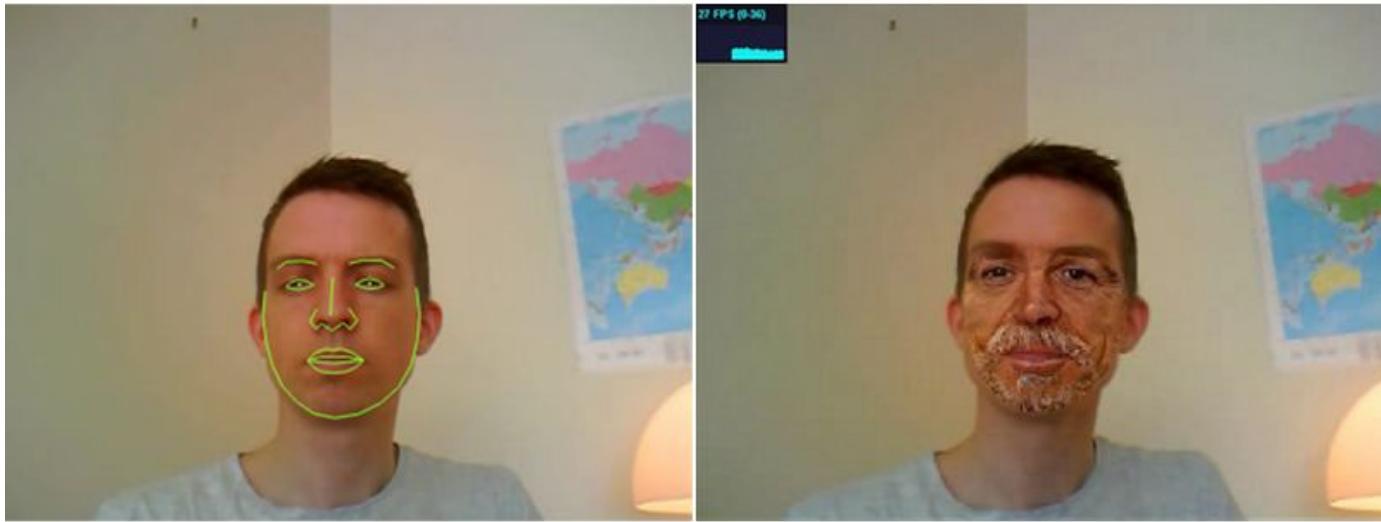
# Workshop outline

Day 1 (Monday):

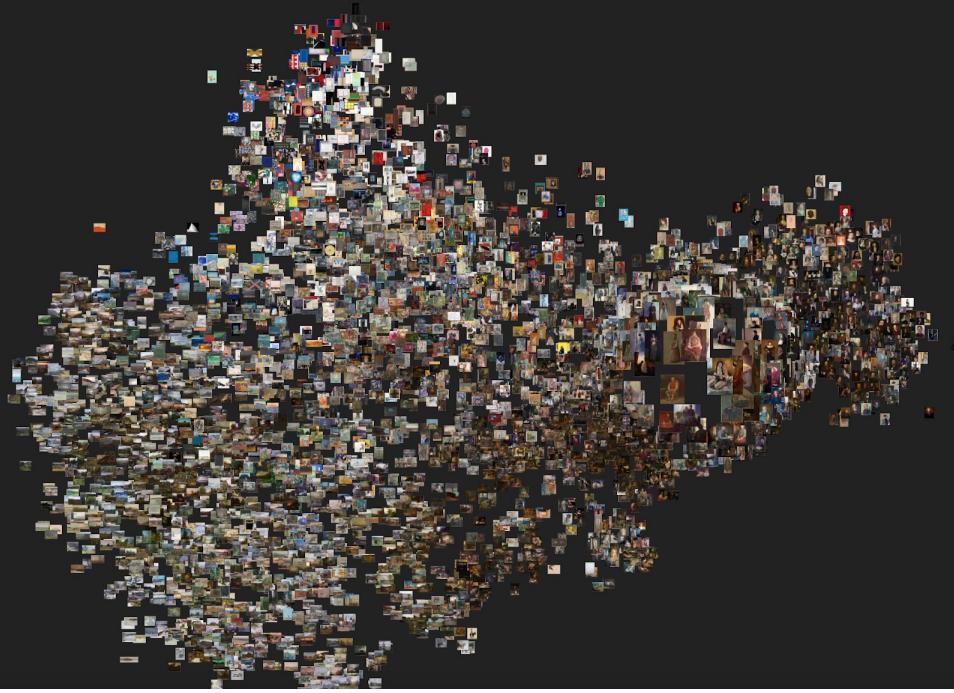
- Morning
  - Intro to ML/AI
  - Intro to p5.js
  - ML Basics
- Lunch
- Afternoon
  - Intro to ML5.js
  - Hands-on with pretrained models
  - Training models with Teachable Machine
  - Hands-on with trained models

Day 2 (Tuesday):

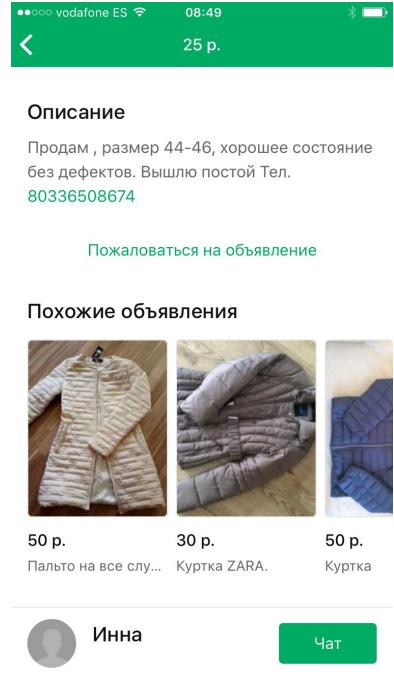
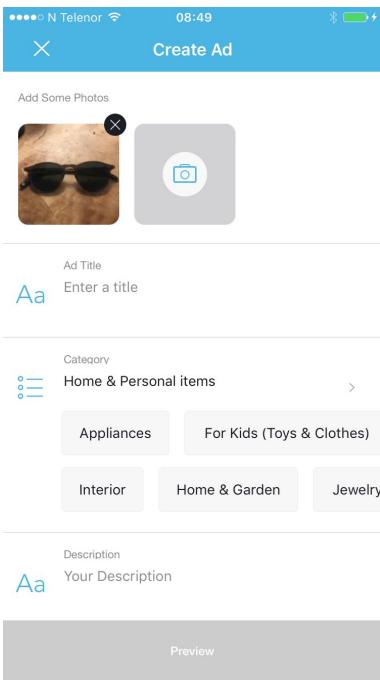
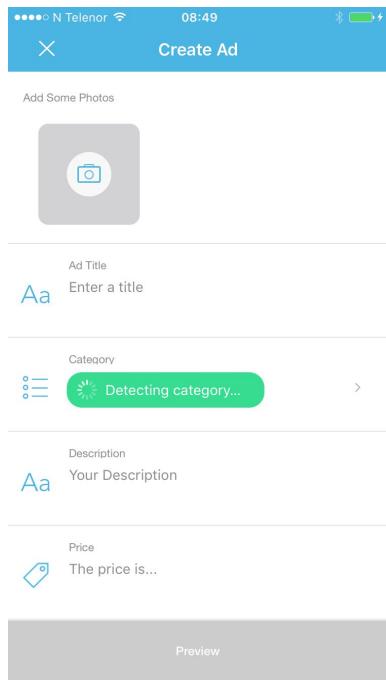
- Morning
  - Recap of day 1
  - Generative ML : sequential models and GANs
  - Hands-on session
- Lunch
- Afternoon
  - Hands-on session continues
  - Discussion and recap of the workshop ++



“clmtrackr”  
Realtime face tracking library in javascript  
<https://github.com/auduno/clmtrackr>

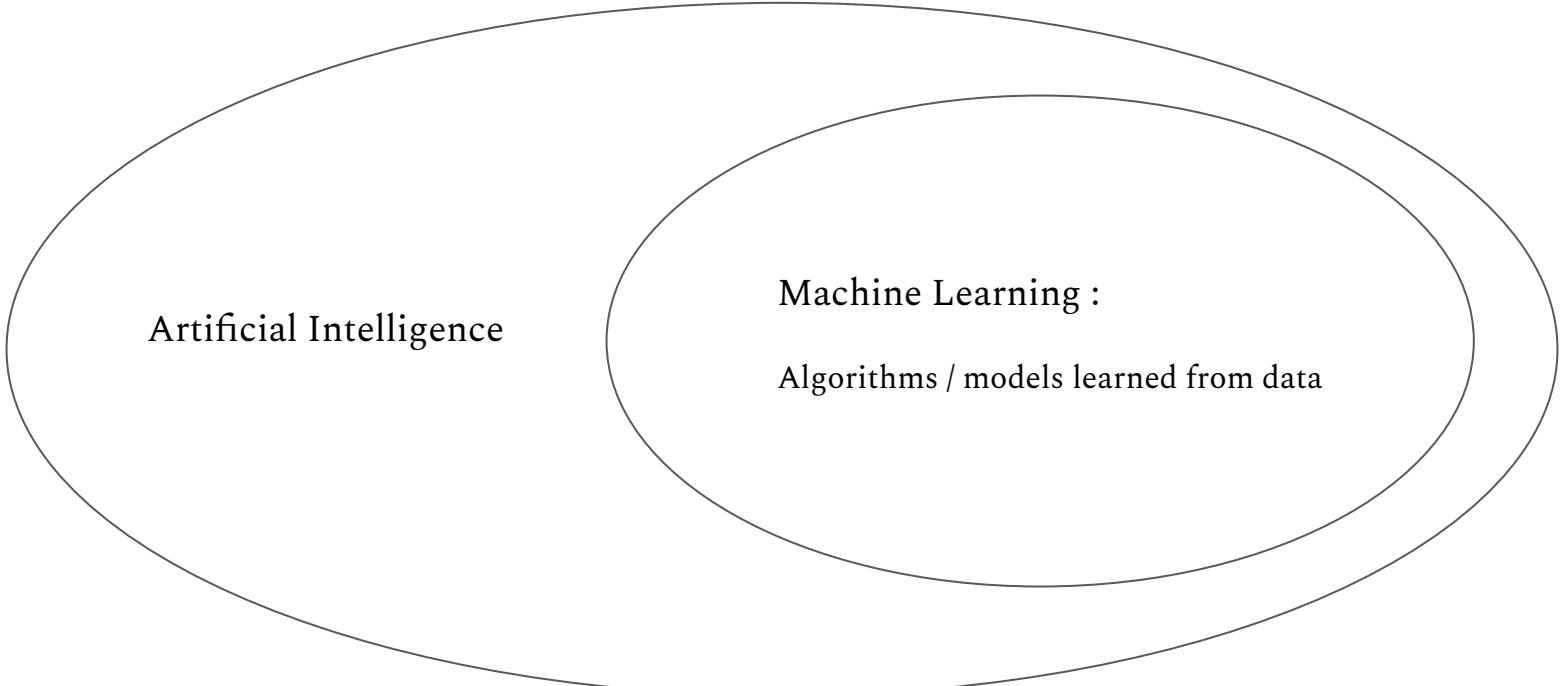


Visualizing the collection of Nasjonalmuseet  
(Collaboration with Bengler)  
<http://vy.nasjonalmuseet.no>



# Image recognition systems Schibsted marketplaces

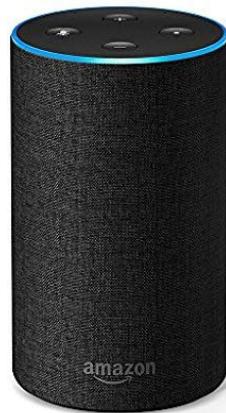
# What is Machine Learning?



Artificial Intelligence

Machine Learning :  
Algorithms / models learned from data

# Voice recognition

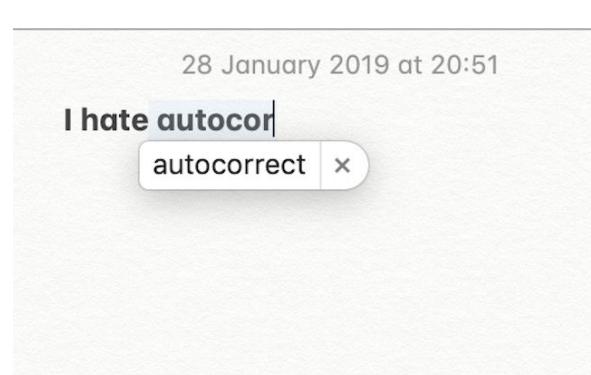
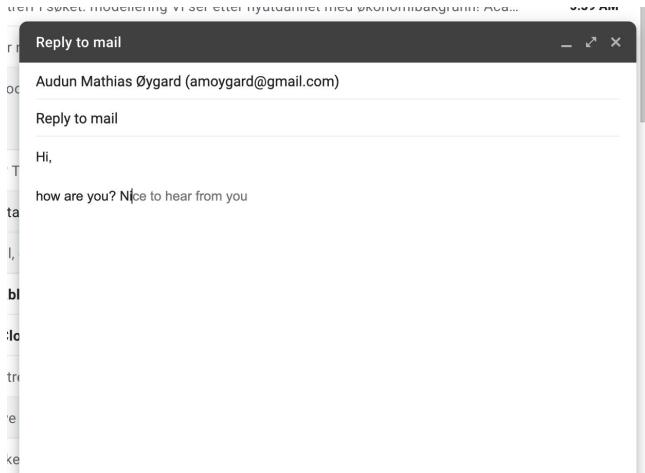
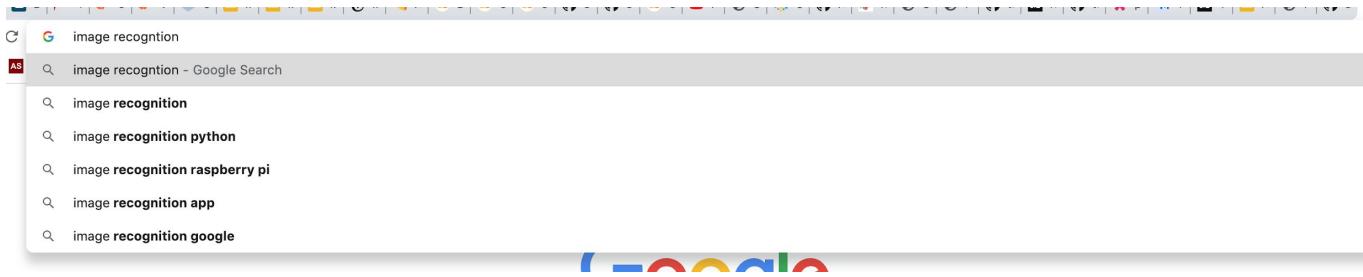


# Image search

A screenshot of a Mac OS X desktop showing a Google Photos search for "dogs". The search bar at the top says "dogs". Below it, a grid of dog photos is displayed, organized by date. The dates visible are Apr 16, 2016; Dec 13, 2015; Sep 13, 2015; Sep 5, 2015; May 16, 2015; Feb 12, 2015; and Dec 5, 2014. Each date group contains several photos of dogs in various settings, such as outdoors and indoors.

A screenshot of a Google search results page for the query "cat". The search bar shows "cat". Below it, the "Images" tab is selected. The results are a grid of cat-related images, including "Cute" cats, "And Kittens", "Clipart" (Cartoon cats), and "Drawing" (Line art cats). Below these categories are several larger, individual cat images.

# Text suggestion



# Translation services

≡ Google Translate

The screenshot shows the Google Translate interface. At the top, there are tabs for "Text" (selected) and "Documents". Below that is a row of language detection and selection buttons: ENGLISH - DETECTED, SPANISH, POLISH, ENGLISH, NORWEGIAN (selected), ENGLISH, POLISH. The main area displays a translation pair: "This is an example of a translation." in English and "Dette er et eksempel på en oversettelse." in Norwegian. There are icons for microphone, speaker, edit, and share below the text. A character counter shows 36/5000. At the bottom right, there are "Send feedback" and other sharing options.

ENGLISH - DETECTED SPANISH POLISH ENGLISH NORWEGIAN ENGLISH POLISH

This is an example of a translation. Dette er et eksempel på en oversettelse.

36/5000 Send feedback

# Recommendations

The screenshot shows a Spotify interface with a dark theme. On the left is a sidebar with navigation links: Browse, Radio, Your Library, Made For You, Recently Played, Songs, Albums, Artists, Stations, Local Files, Videos, and Podcasts. Below these are sections for Playlists like 'Women's March on ...', 'Jeep Stuff', 'Acoustic Covers', and 'Songs to Sing in the...'. The main content area features a 'Discover Weekly' mixtape for 'Kathleen Slattery Booth'. The mixtape cover features a photo of a woman and the text 'MADE FOR KATHLEEN' and 'Discover Weekly'. It describes it as a weekly mixtape of fresh music updated every Monday. Below the cover are buttons for 'PLAY', 'FOLLOW', and three dots. The mixtape has 0 followers. The tracklist table includes columns for Title, Artist, Album, and Date. The first few tracks are:

TITLE	ARTIST	ALBUM	
Watch Me (Whip / Nae Nae)	Silentó	Watch Me (Whip / ...)	2 days ago
Fight Song	Rachel Platten	Wildfire	2 days ago
Scream & Shout	will.i.am, Britney Spe... #willpower (Deluxe)	#willpower (Deluxe)	2 days ago
Gangnam Style - Radio Edit	Opa	Gangnam Style	2 days ago
Ooh La La (from "The Smurfs 2")	Britney Spears	Ooh La La (from "Th...")	2 days ago
Sax	Fleur East	Love, Sax & Flashba...	2 days ago

# Recommendations

NETFLIX Browse ▾ DVD

Search Joshua ▾

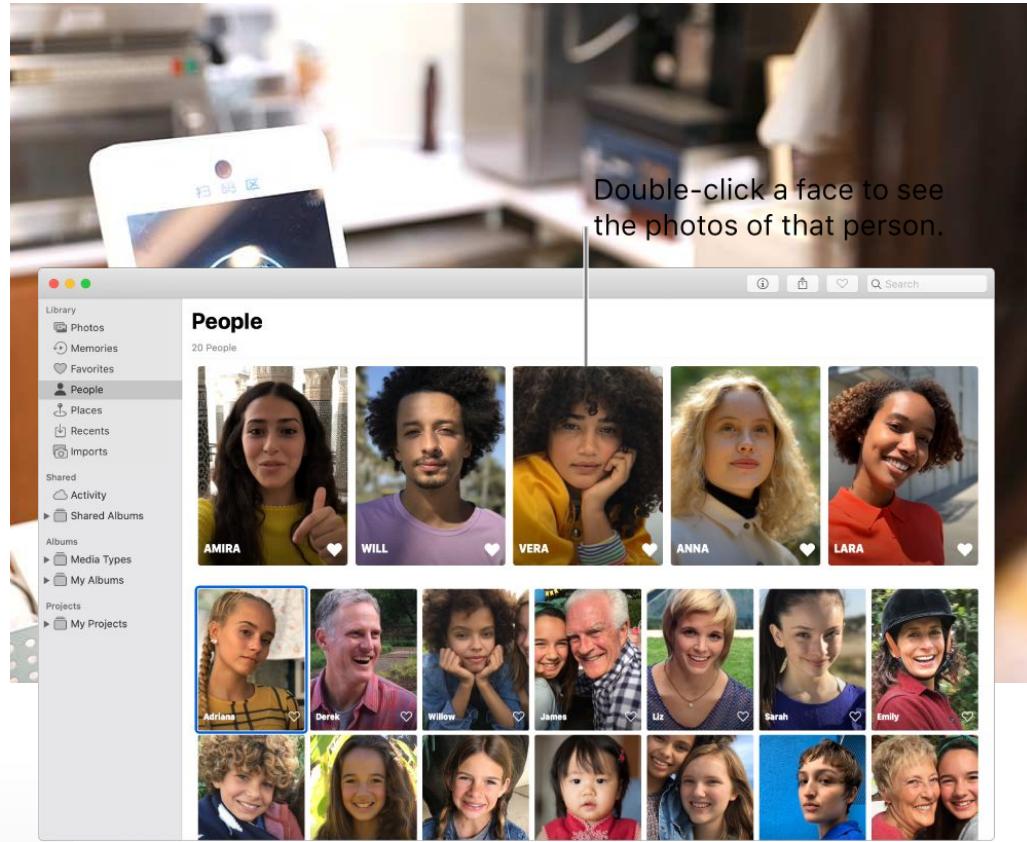
Top Picks for Joshua

Trending Now

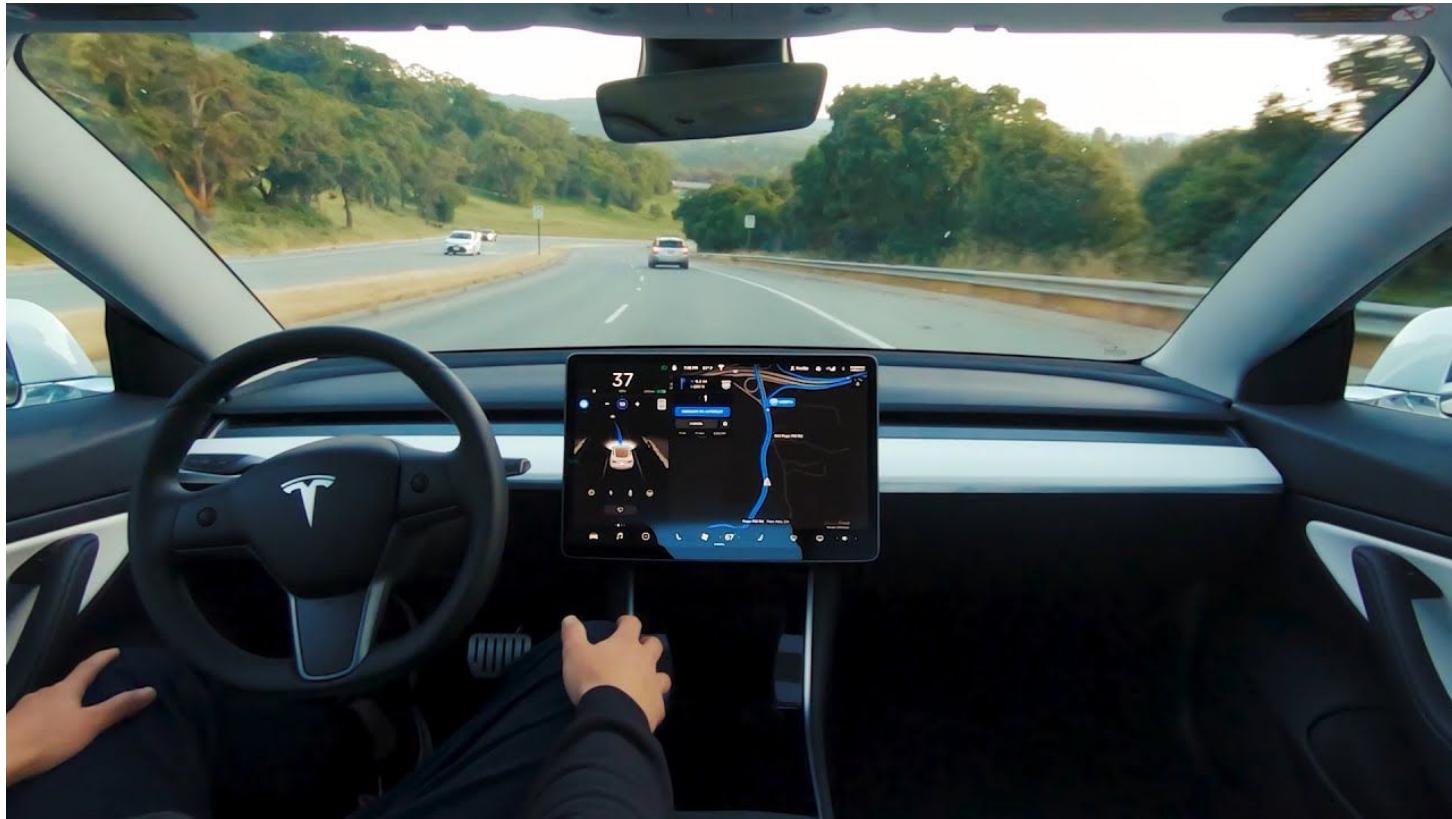
Because you watched Narcos

New Releases

# Face recognition

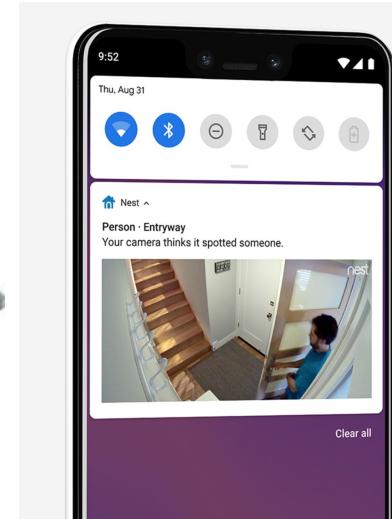


# “Self-driving” cars





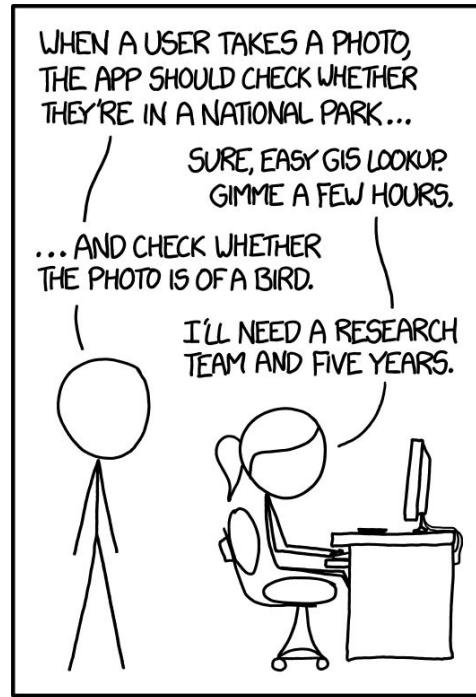
# Smart home devices



# Automated photo-editing

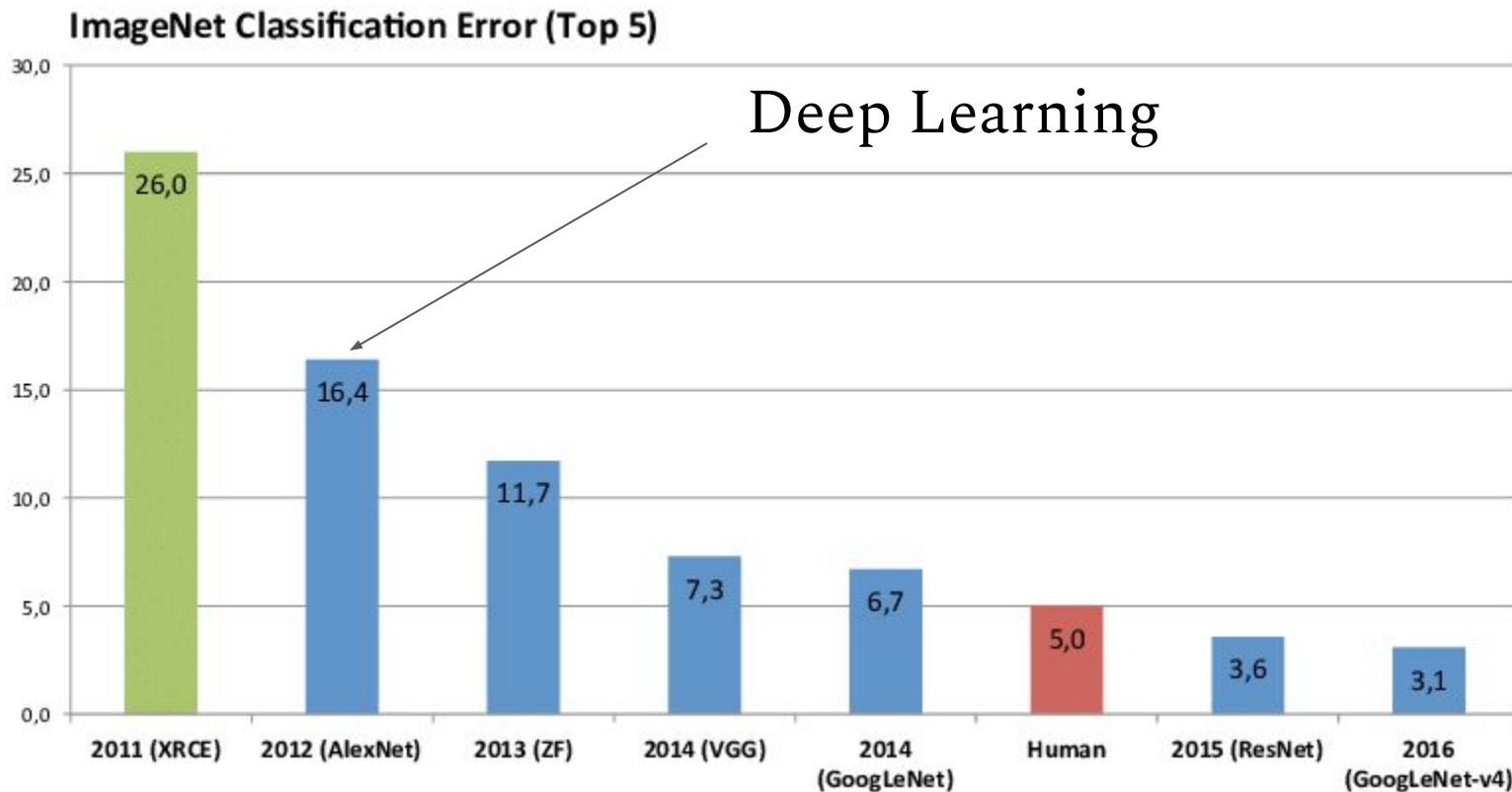


# State of image recognition (2014)



IN CS, IT CAN BE HARD TO EXPLAIN  
THE DIFFERENCE BETWEEN THE EASY  
AND THE VIRTUALLY IMPOSSIBLE.

# Improvements in Image Recognition



# AlphaGo (2016)



# Recent improvements in Machine Learning

- New and improved methods
- More processing power (GPUs)
- Lots of more labelled data available and standard datasets
- More mature tooling

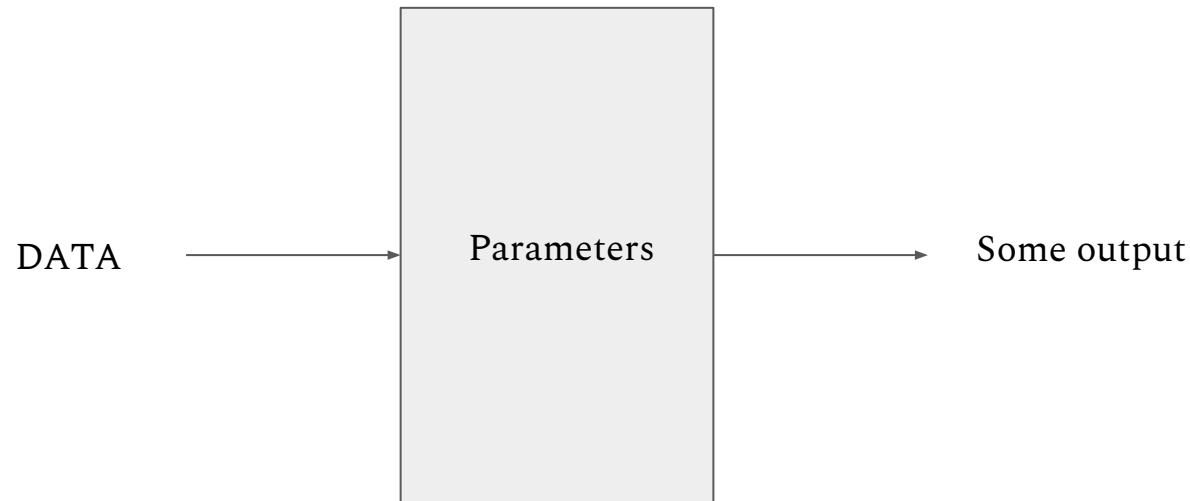
# p5.js

## Intro to p5.js

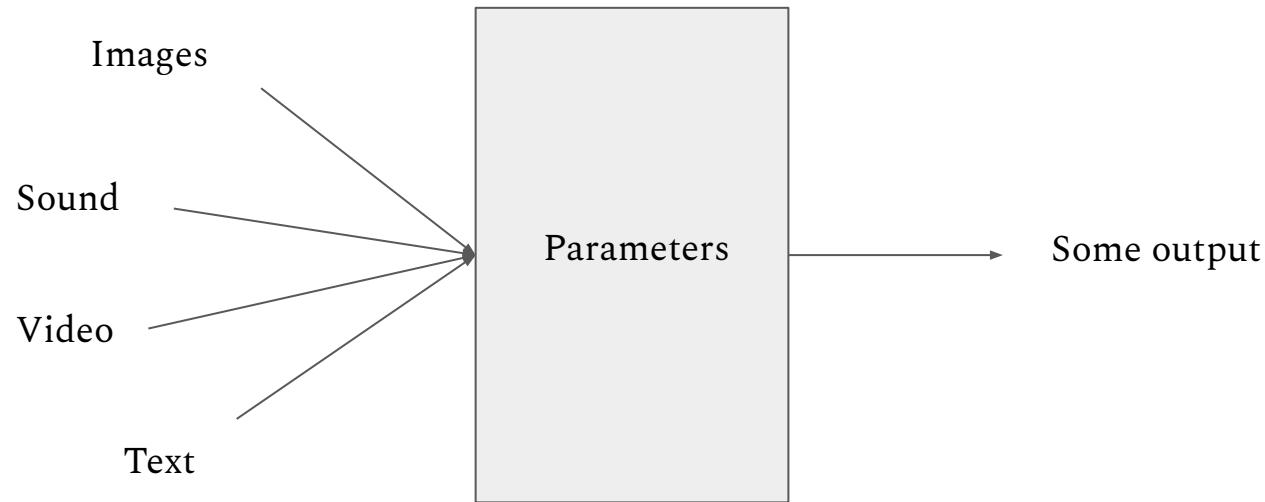
<https://p5js.org/>

# Intro to Machine Learning

# What is a model?



# What is a model?



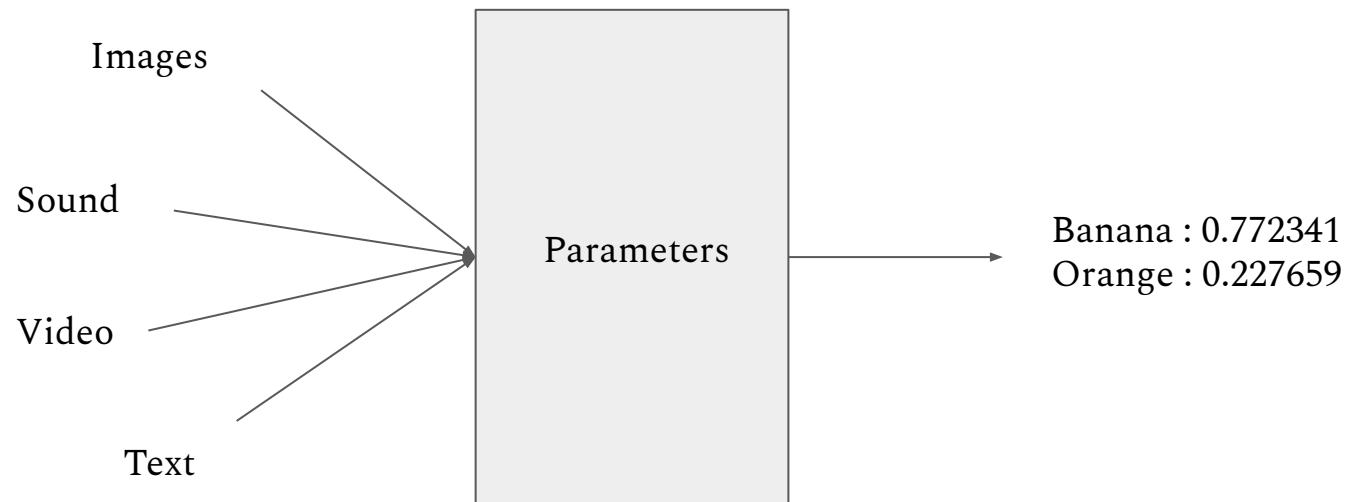
# What is classification?

- Predicting what *class* an object belongs to based on given input

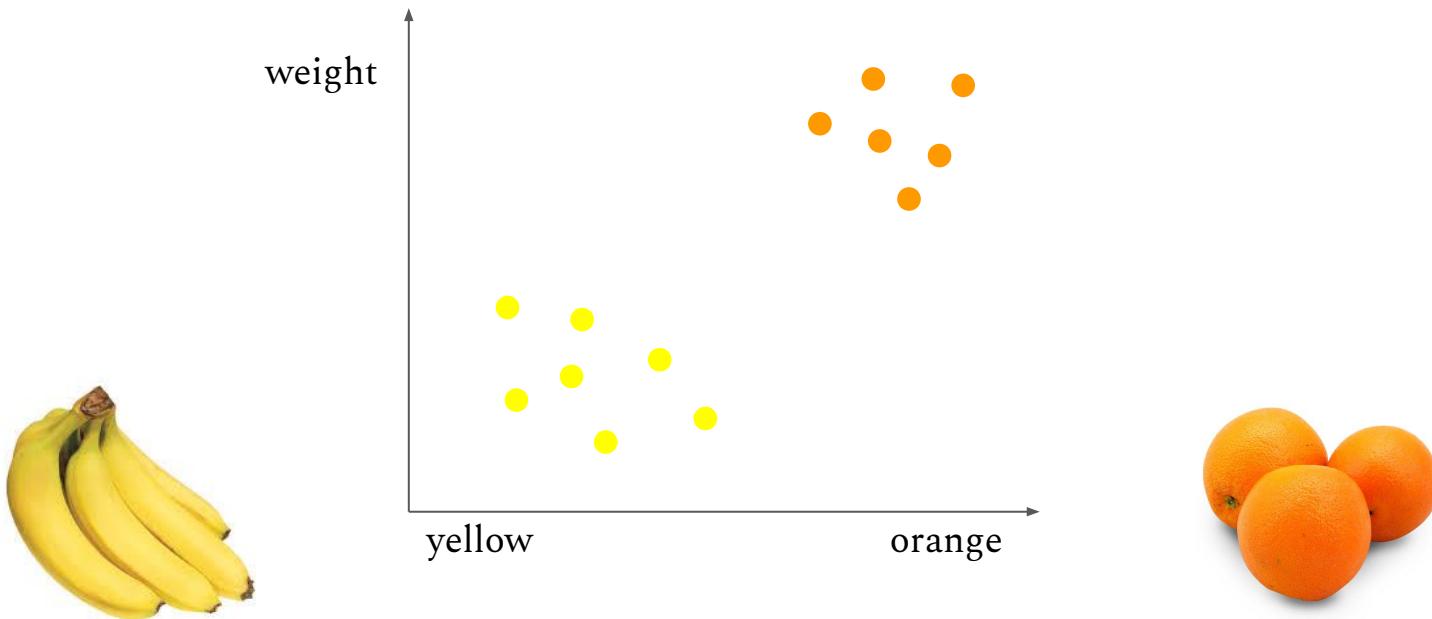


@teenybiscuit

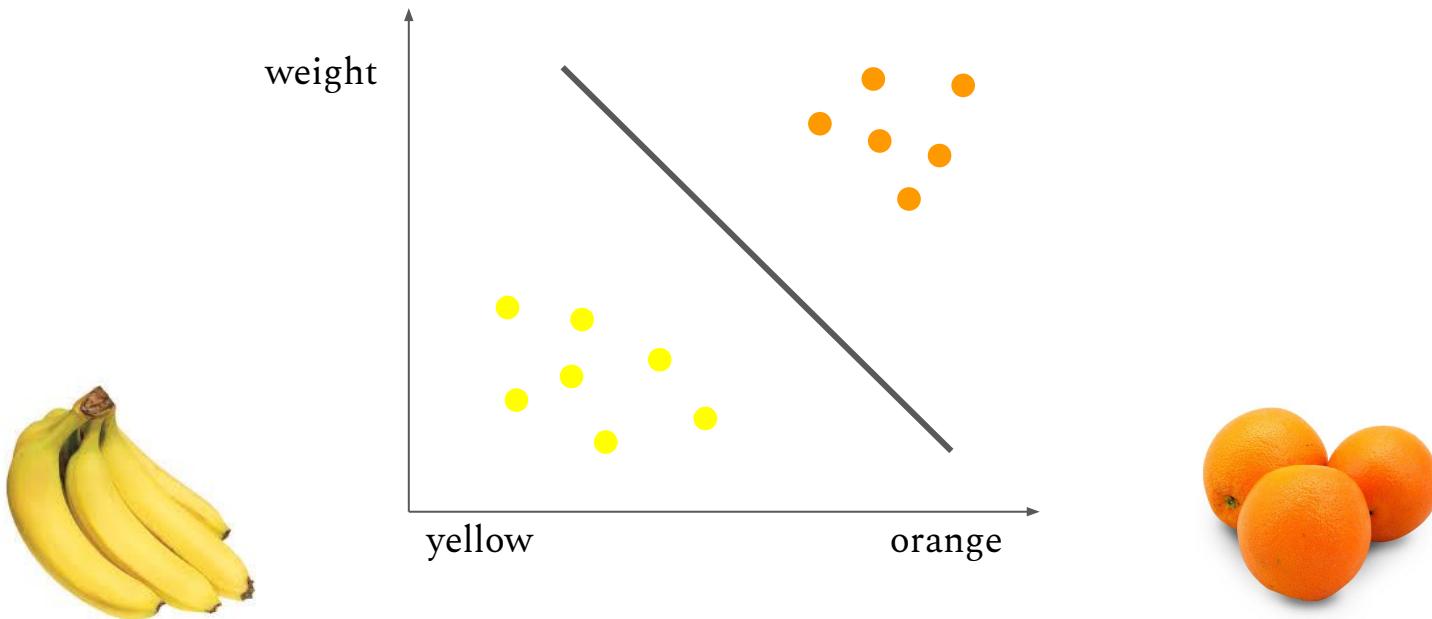
# Classification model



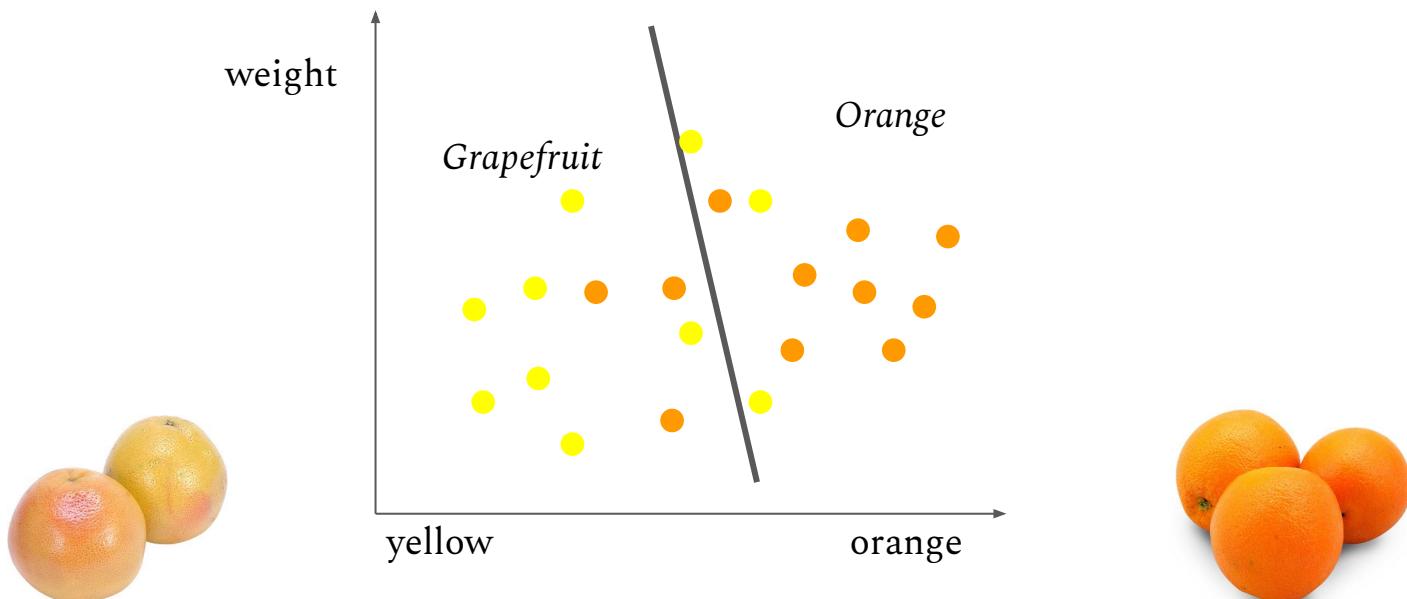
# Example of simple classification



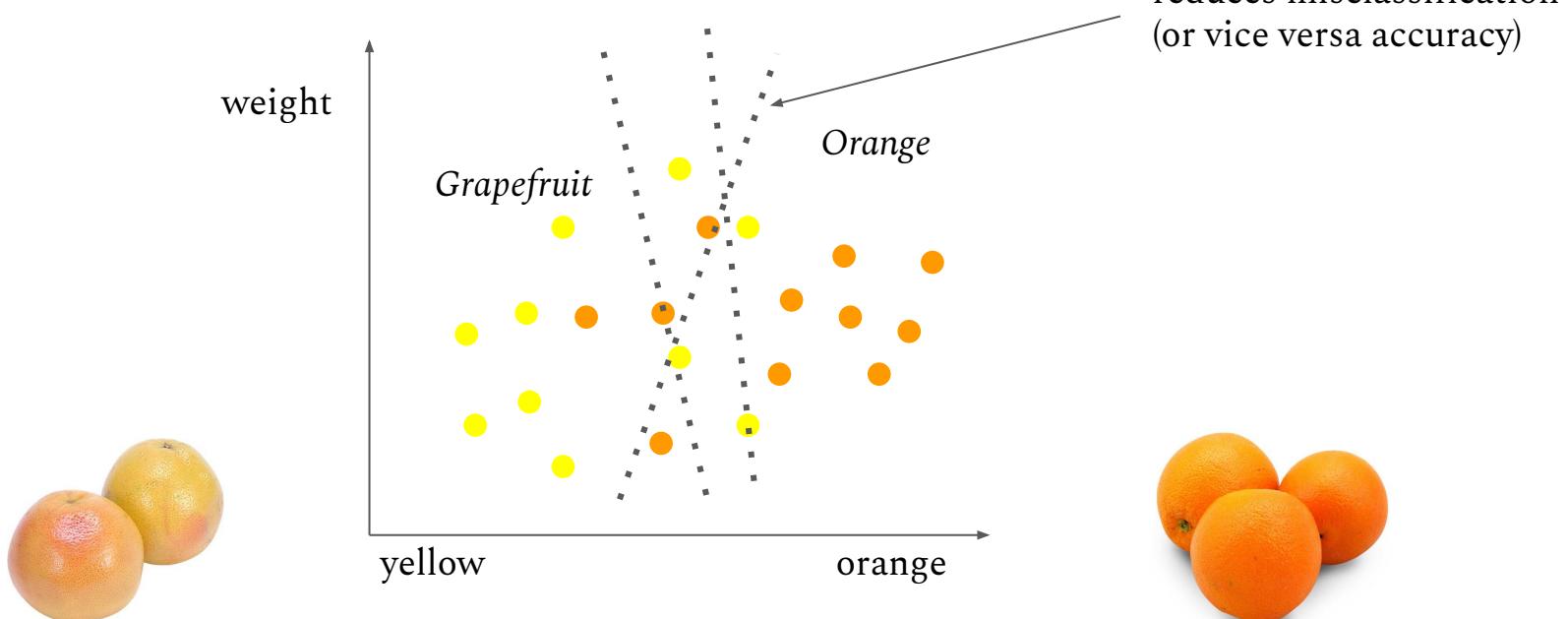
# Linear separation



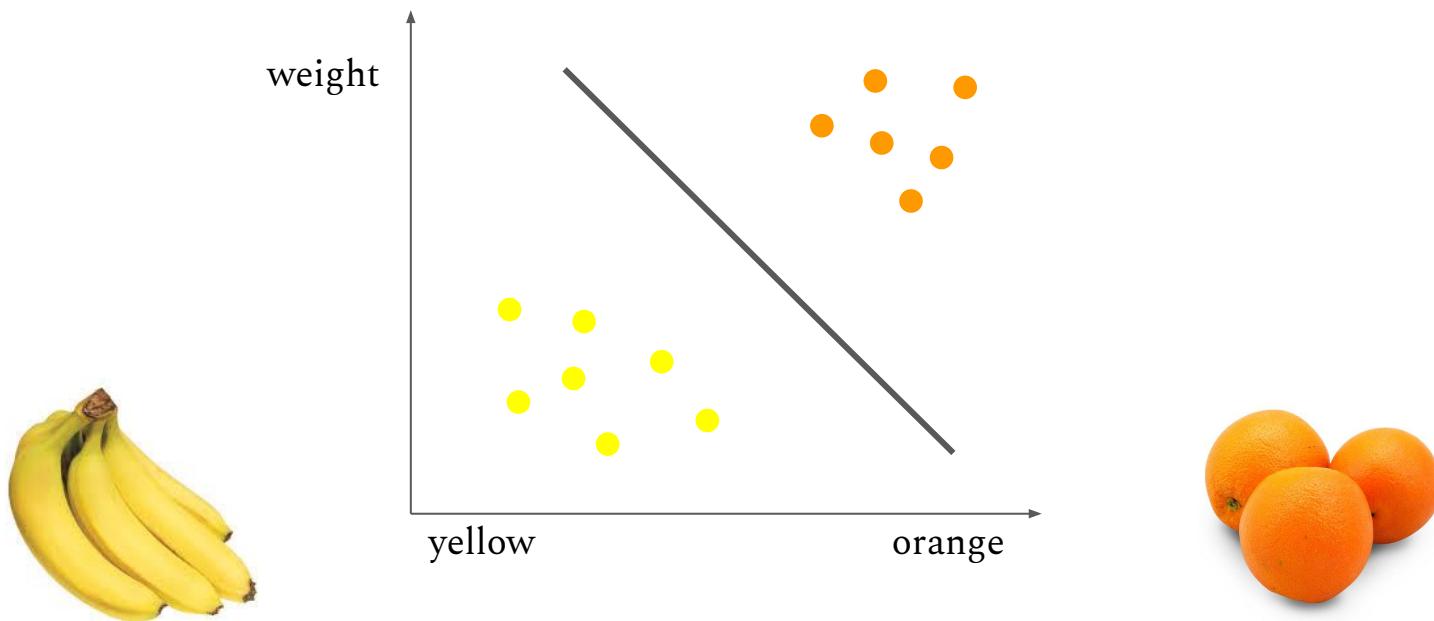
# Misclassifications



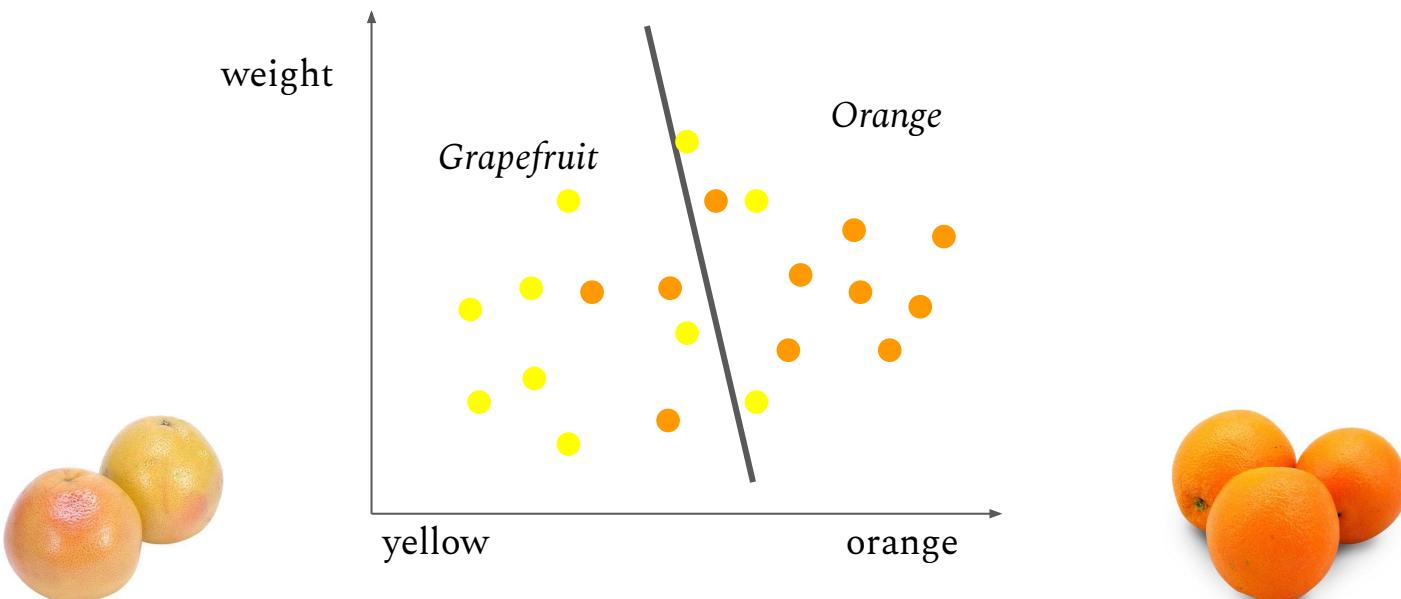
# “Training” the model



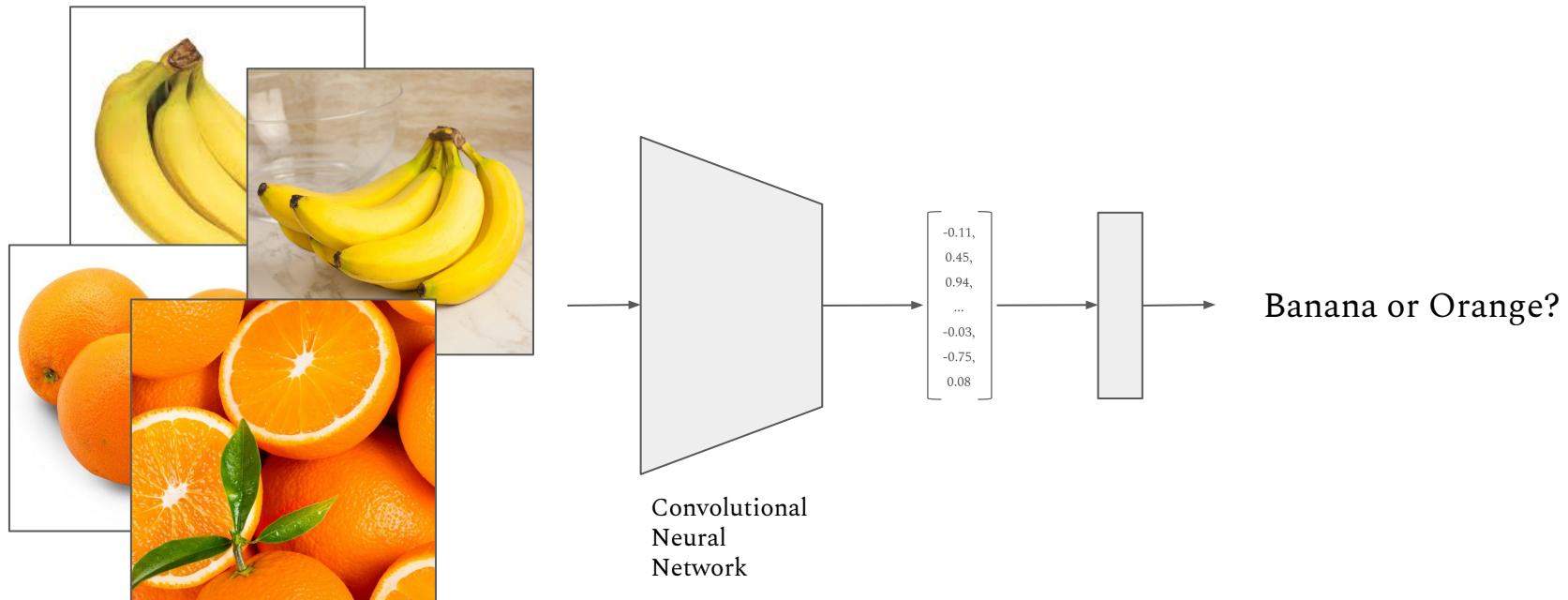
100% accuracy



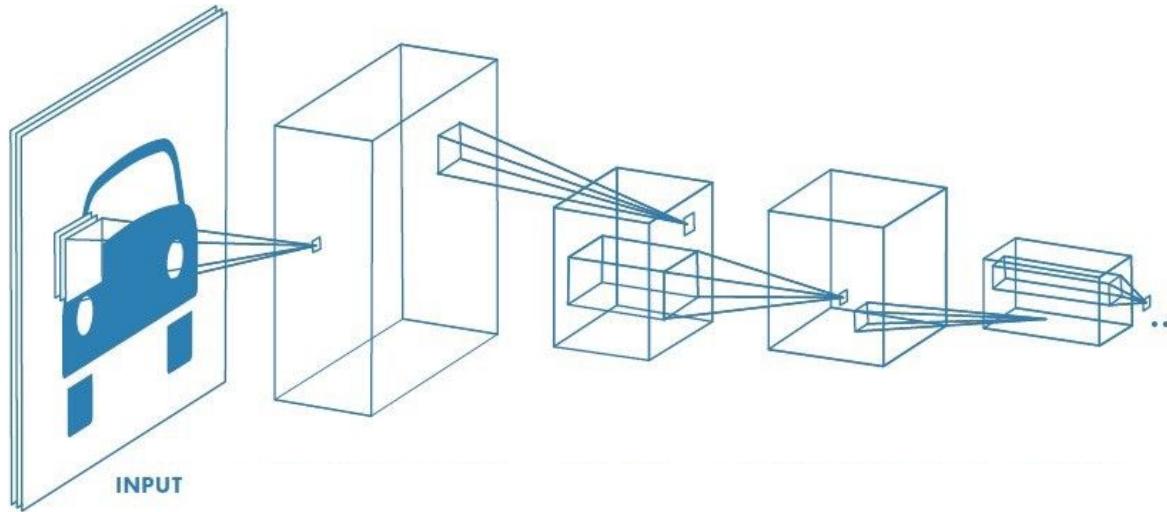
# 70% accuracy



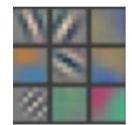
# Image classification



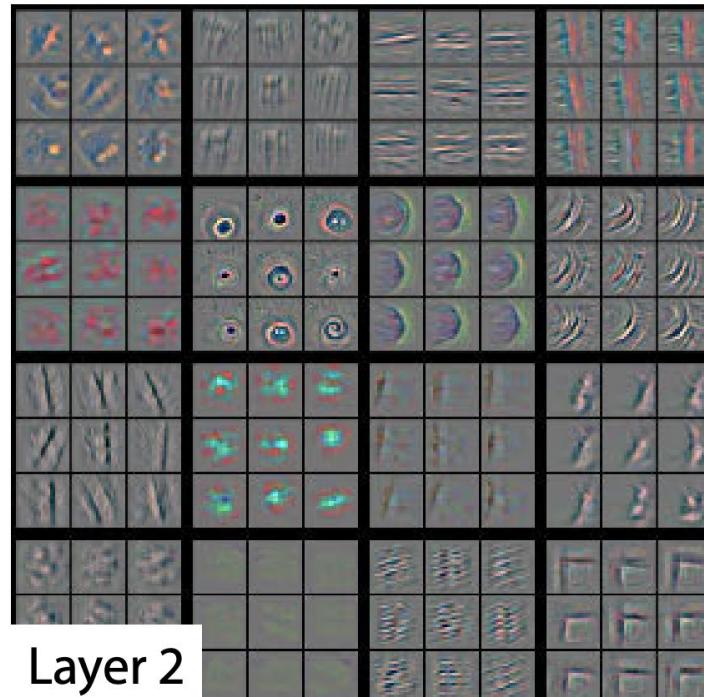
# Convolutional neural networks (images)



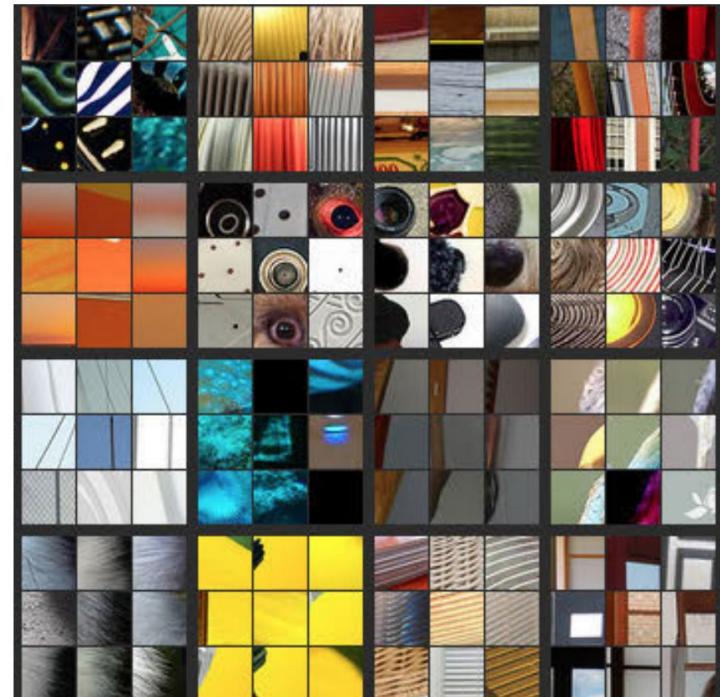
# Convolutional neural networks



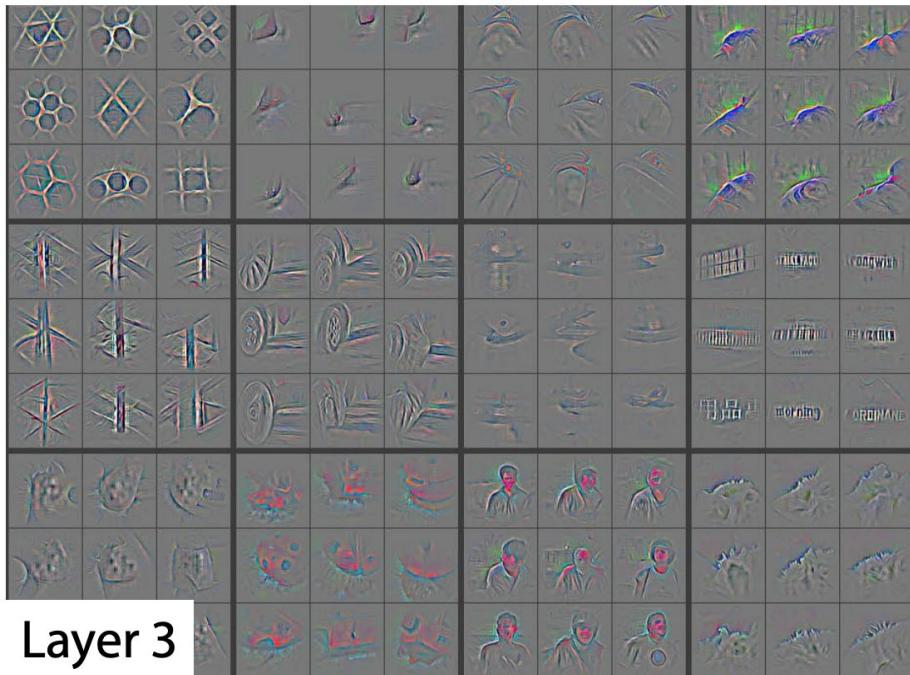
Layer 1



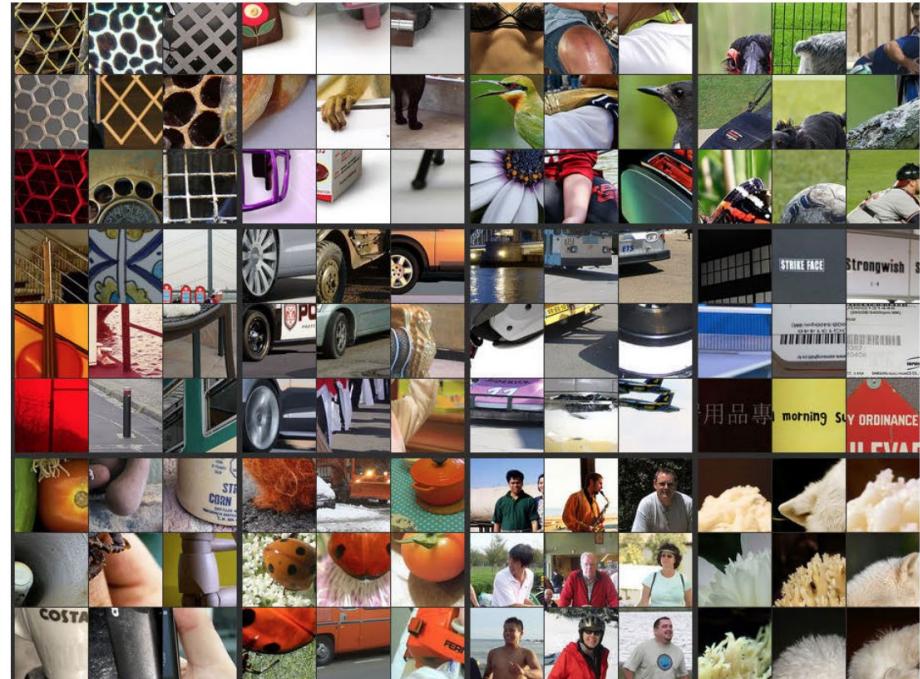
Layer 2



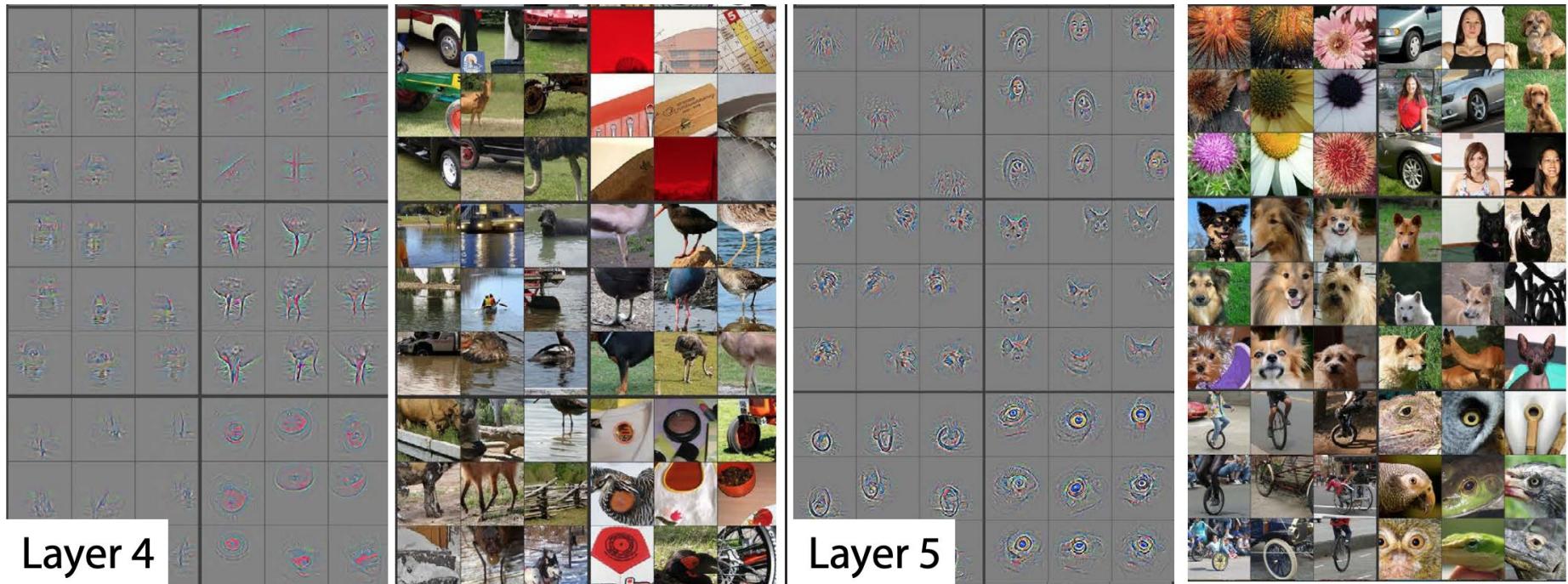
# Convolutional neural networks



Layer 3

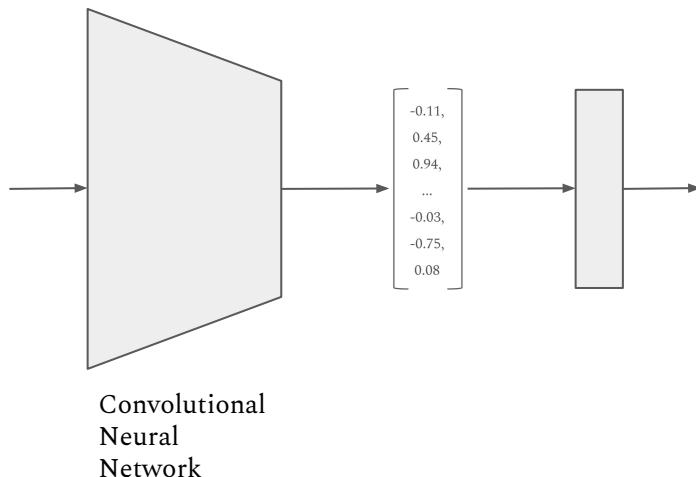
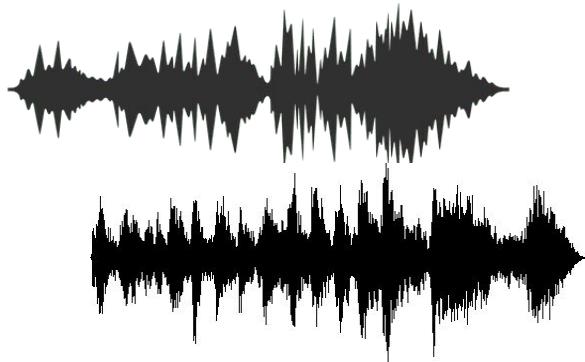


# Convolutional neural networks



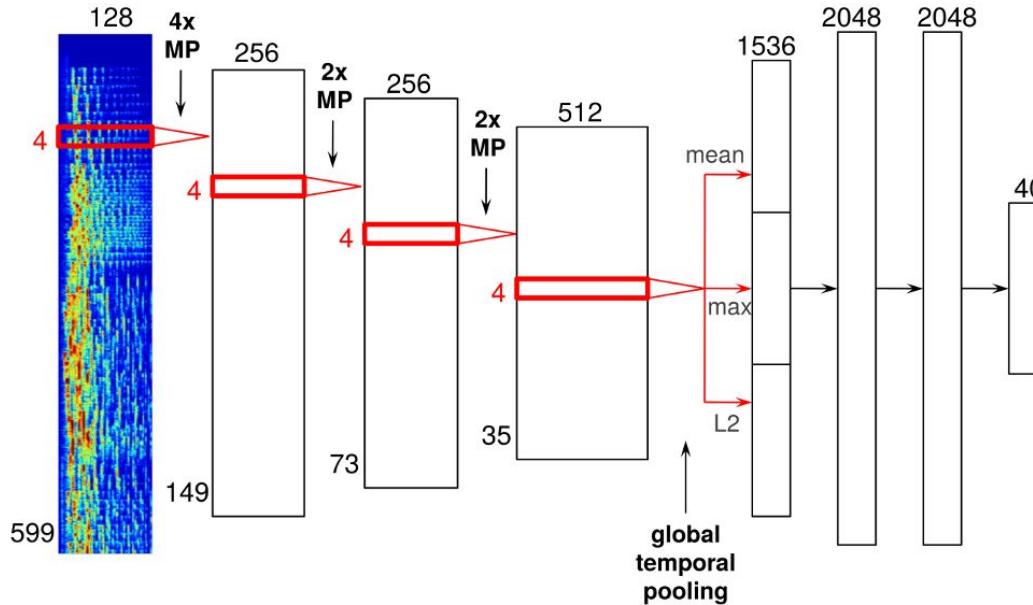
Source : “Visualizing and understanding Convolutional Networks”, Zeiler et al (2013)

# Sound classification



And also video etc.

# Convolutional neural networks (sound)



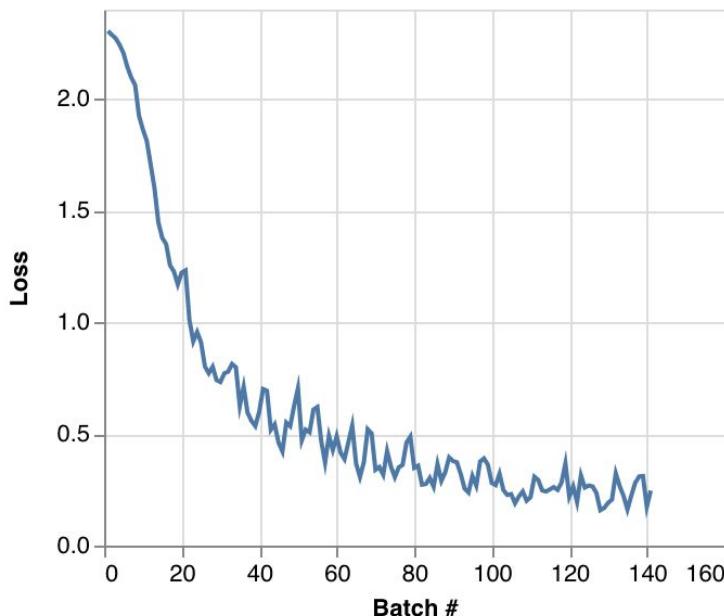
# IMAGENET Dataset

- Over 24 million labelled images
- Collected by Fei-Fei Li ++ at Stanford (2009)
- Many different classes
- 1000 classes are often used as reference

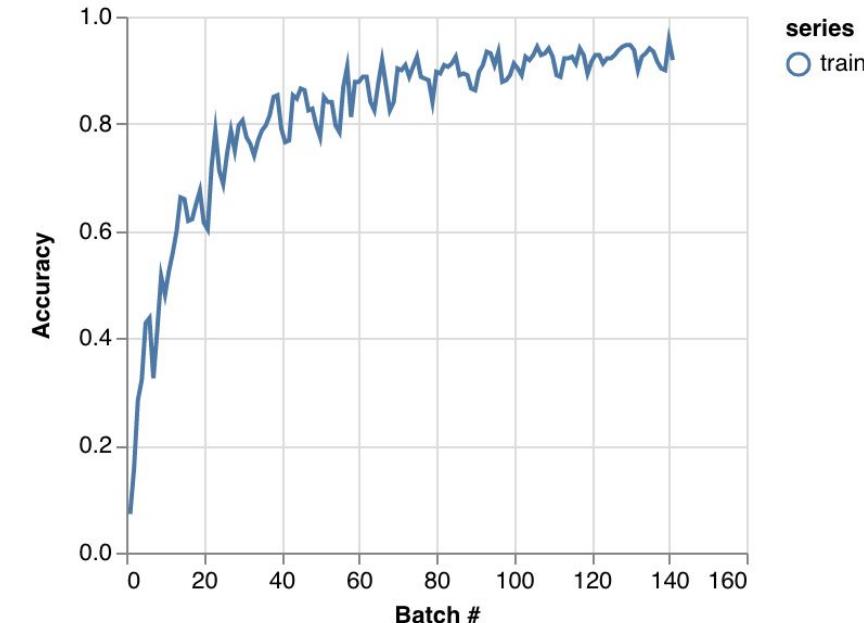


# Training on large datasets

last loss: 0.246



last accuracy: 91.9%



# Training hardware



CPU



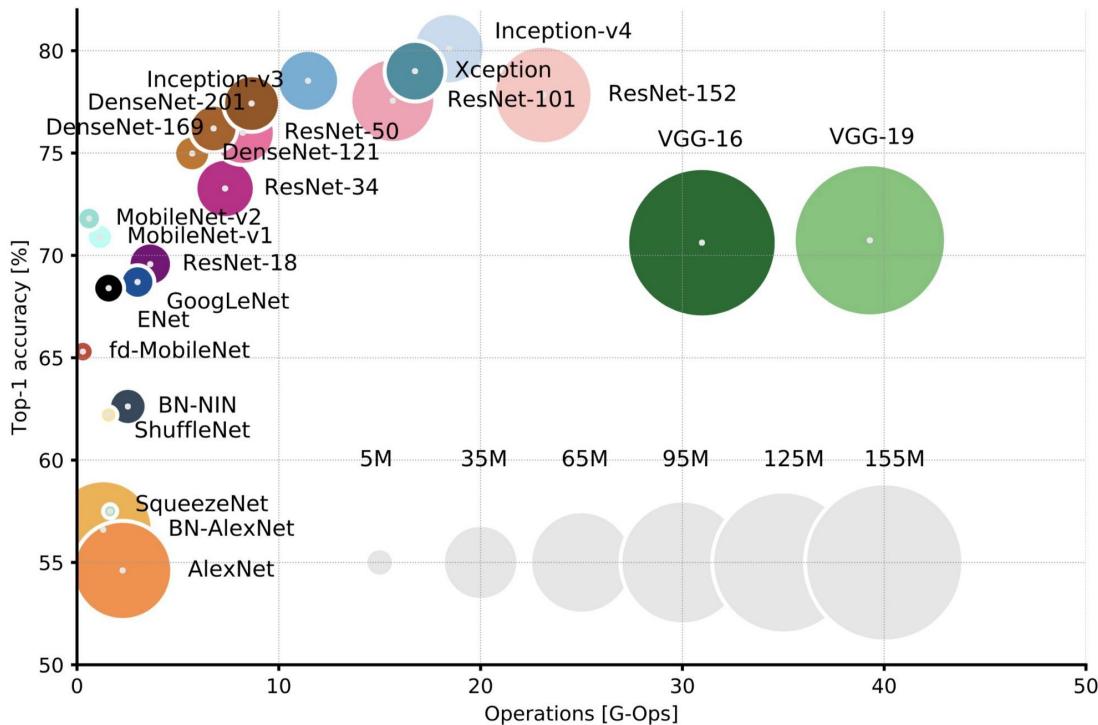
GPU



TPU  
(2016)

# Some common models

- Inception (2014)
- ResNet (2015)
- WaveNet (2016)
- MobileNet (2017)
- Transformer



Source : <https://towardsdatascience.com/neural-network-architectures-156e5bad51ba>

# Common frameworks for ML



(Facebook)  
Python, C++



(Google)  
Python, C++, Java, Javascript

+ many, many more



# Intro to ml5js

<https://ml5js.org/>

Lunch

# Image classifier

- “MobileNet”
- Trained on ImageNet dataset
- Classifies image into 1000 different classes

## Image classification example

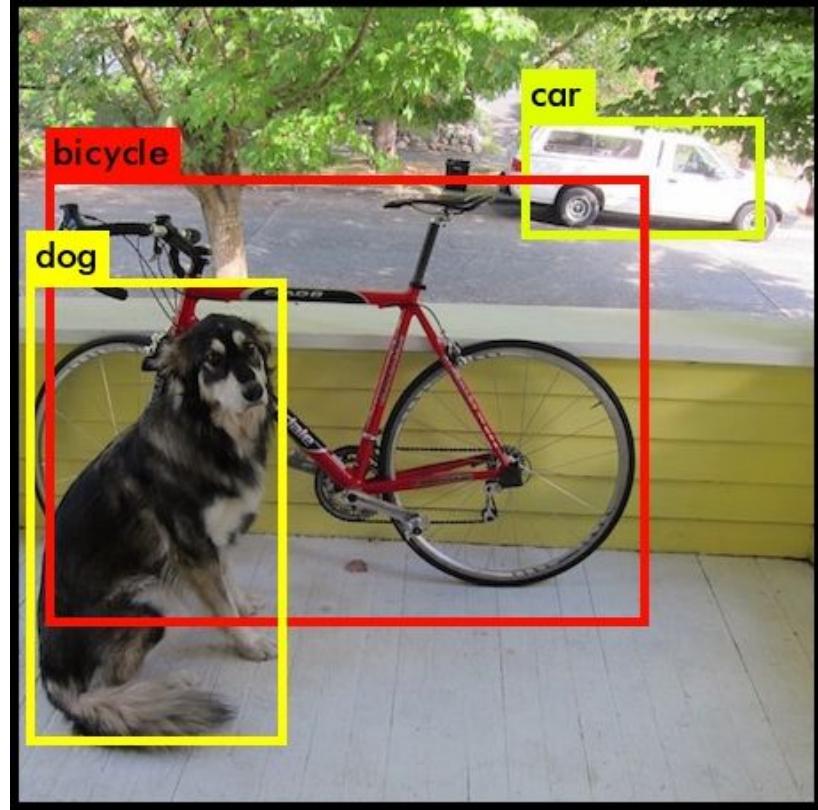


Choose file bananas.jpg

Label: banana, Confidence: 0.9408700466156006  
Label: orange, Confidence: 0.0037055329885333776  
Label: butternut squash, Confidence: 0.002862975699827075

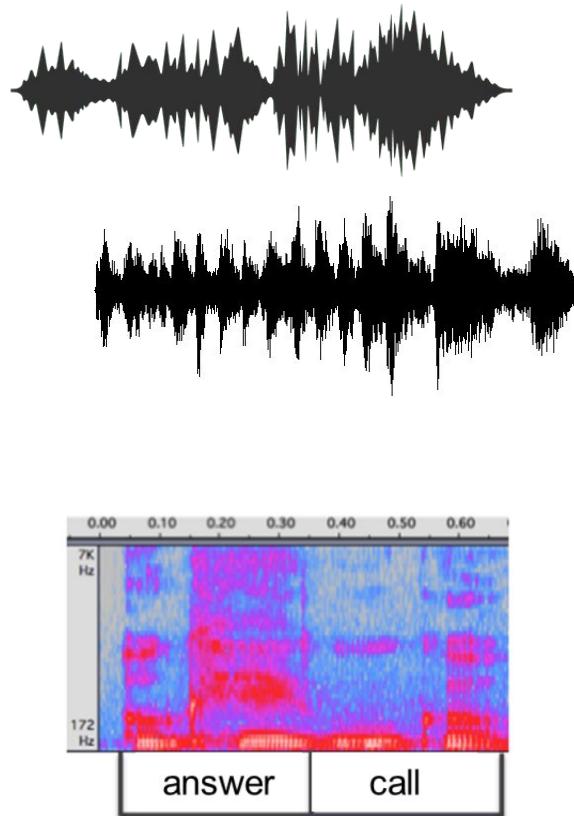
# Object detection

- “Coco-SSD” (Coco Single-shot multibox detector)
- Localizes and classifies objects
- Detects 91 different classes from the COCO dataset (“Common Objects in Context”)



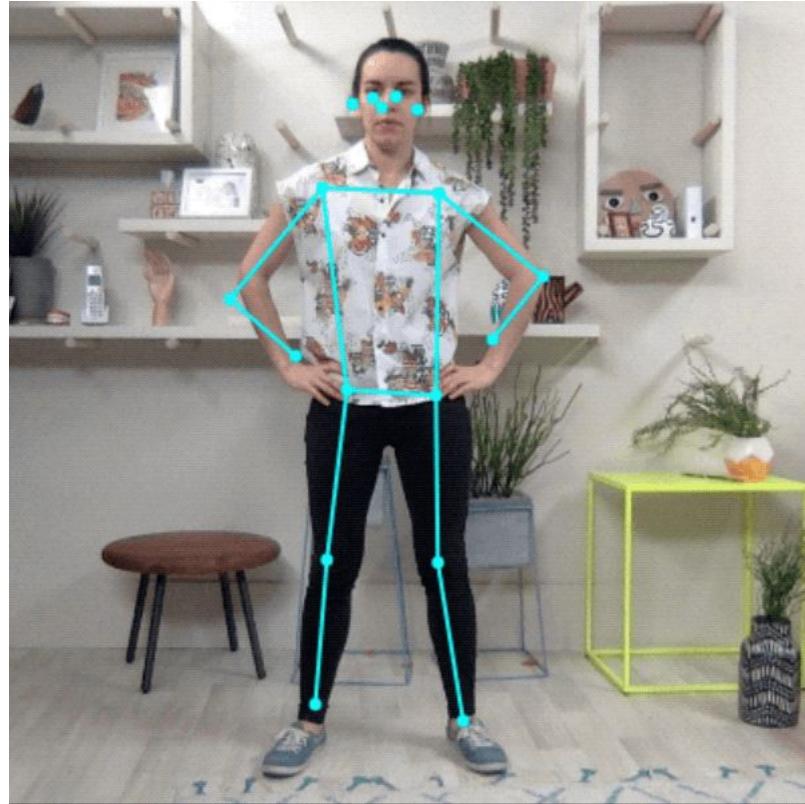
# Sound classifier

- Detects 1-second keyphrases
- 18 words: “zero”, “one”, “two”, “three”, “four”, “five”, “six”, “seven”, “eight”, “nine”, “up”, “down”, “left”, “right”, “go”, “stop”, “yes” and “no”
- Converts sound to spectrogram image and uses small convolutional network

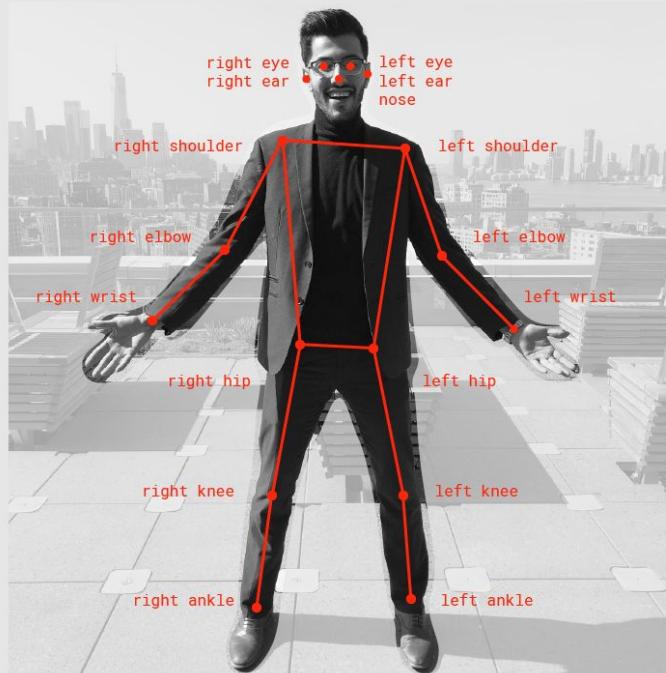


# PoseNet

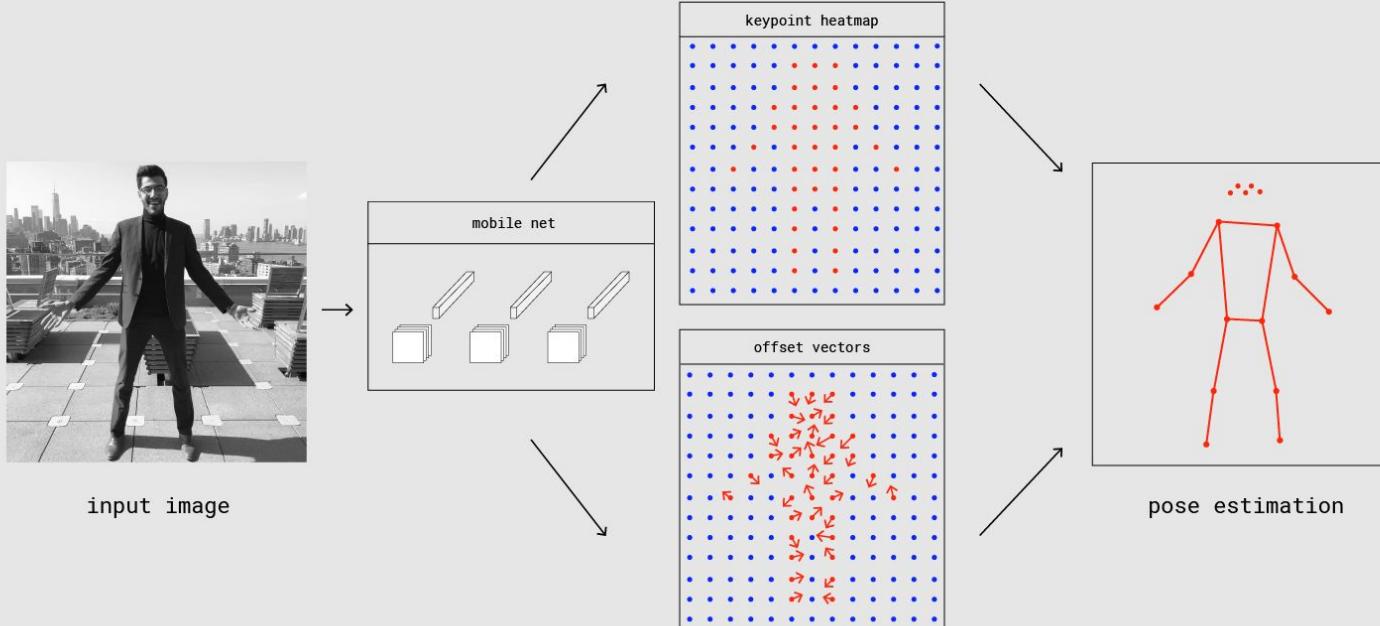
- Detects 17 body keypoints
- Is able to track both one person and multiple people



**17 Pose Keypoints  
Returned by PoseNet**

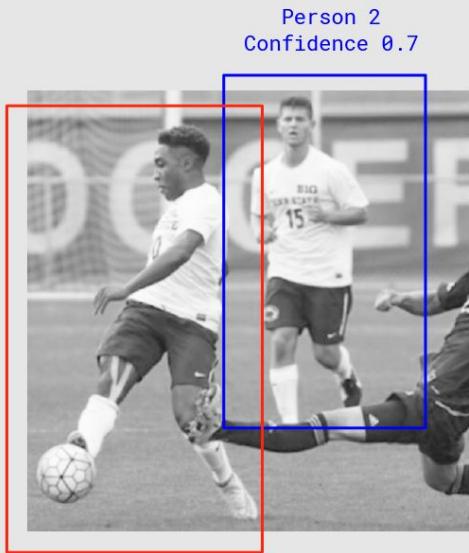


## Single-Pose Detection Algorithm



PoseNet model

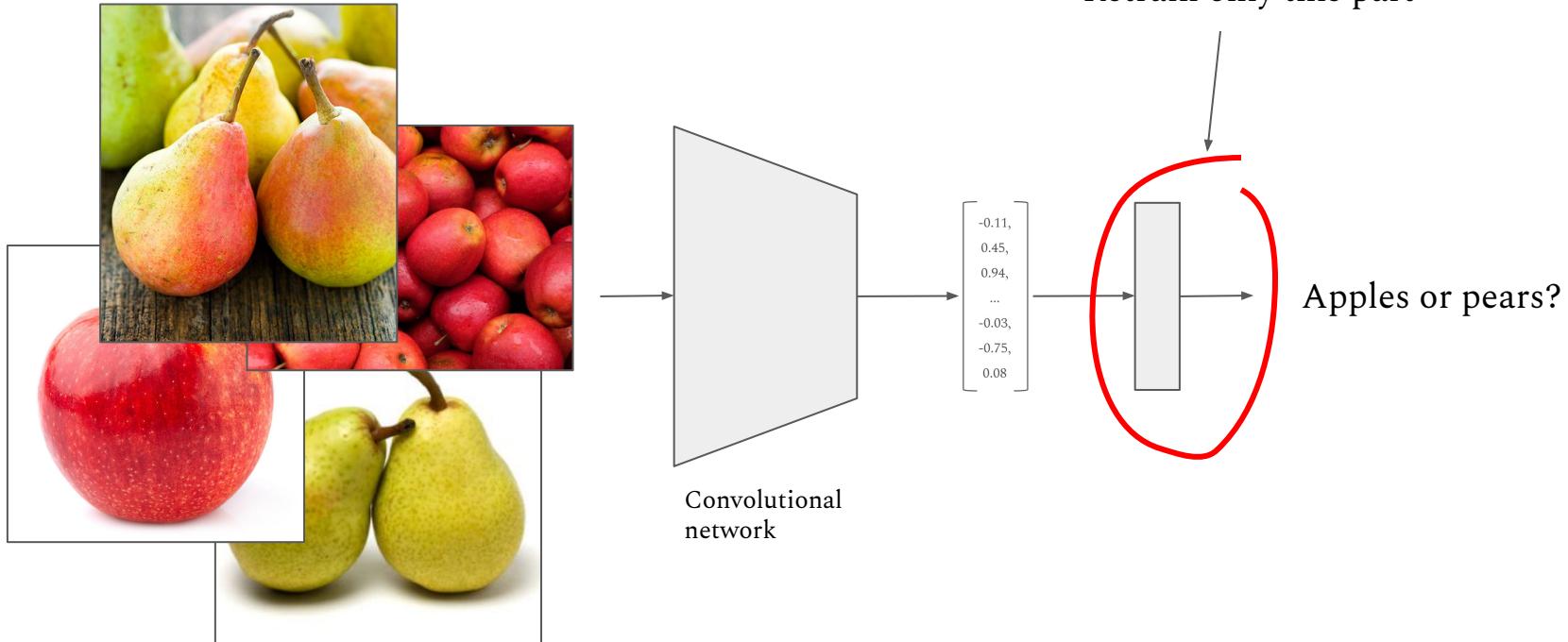
## Pose Confidence Scores vs. Keypoint Confidence Scores



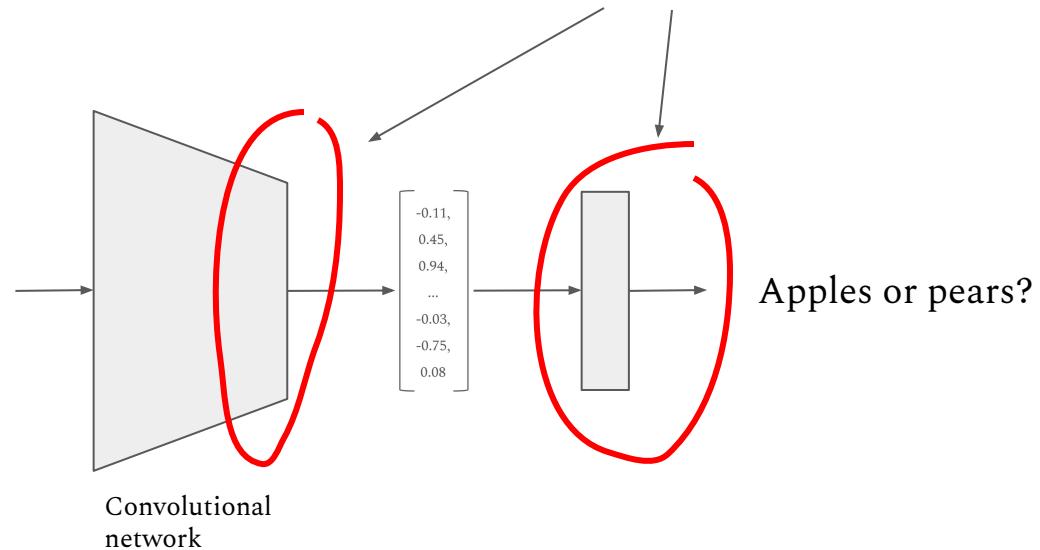
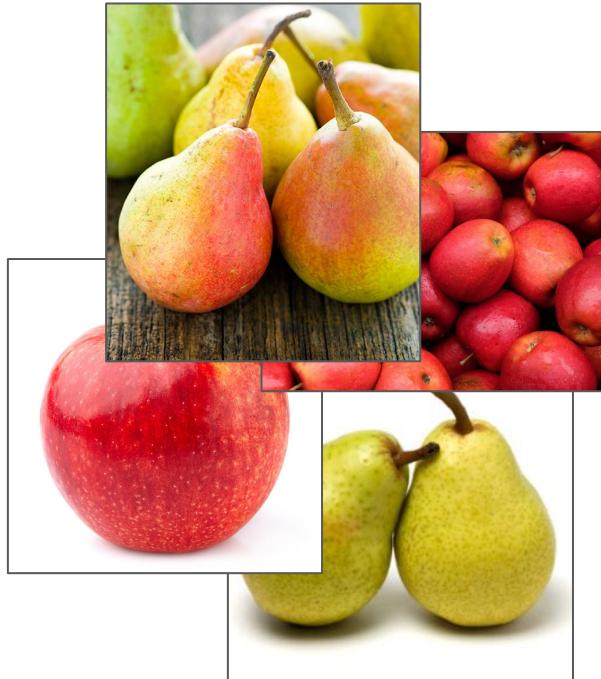
# Hands-on

# Training models

# Retraining a model



# Retraining a model



# Transfer learning

- Repurposing the model trained on one dataset for a different yet related task
- Advantages:
  - Faster training time
  - Less data needed
  - (Often) more accurate

# Teachable Machine

<https://teachablemachine.withgoogle.com/>

# Some tasks

- Try to figure out cases where your classification model fails
  - For images :
    - Change background / room
    - Find edge cases : try to only show parts of the object
- Try to train classification model on very similar objects
  - How much data do you need to train the model?

Try to build something!