Computer Vision

A brief, interactive introduction to the challenge and the methods

Review - what (extra) DA tasks are involved in working with these sources of data?

(files with tables of data)


(databases)


(spatial,lat/long)


(scraping)

(json, xml)

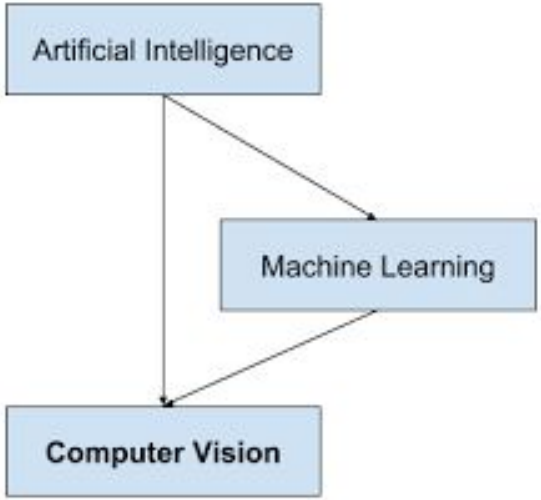

(text, descriptive)

What is the challenge of computer vision?

This is a task that humans find trivial

1. Describe a photo you have seen once
2. Summarise a video
3. Recognise a face

Even the youngest children can do this - they have been trained to since they first opened their eyes.



90%

13ms

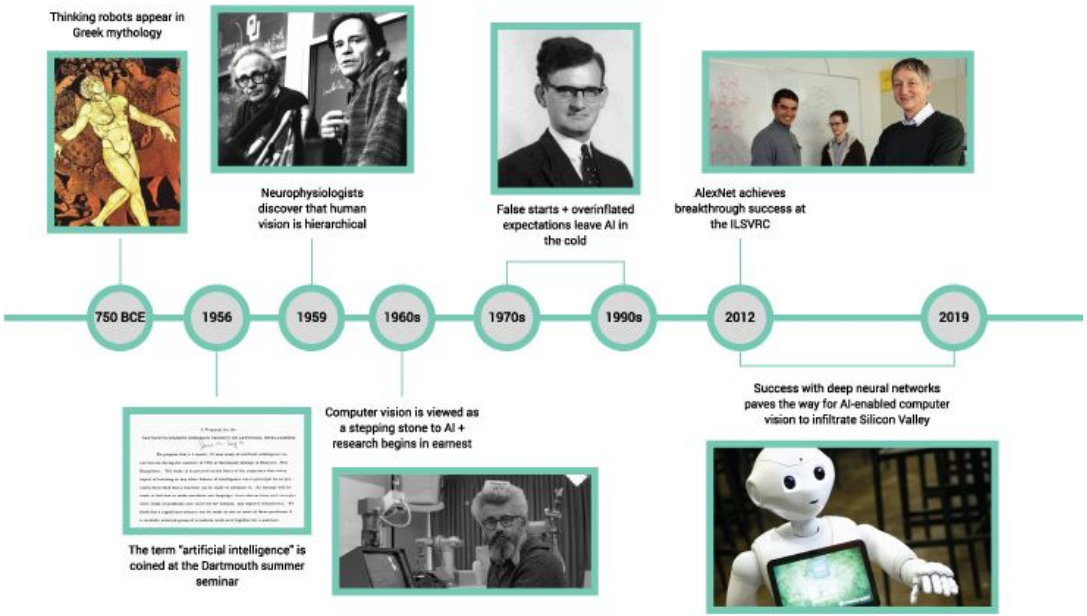
60,000 times

3 months old

Even the youngest children can do this...

Successes and failures

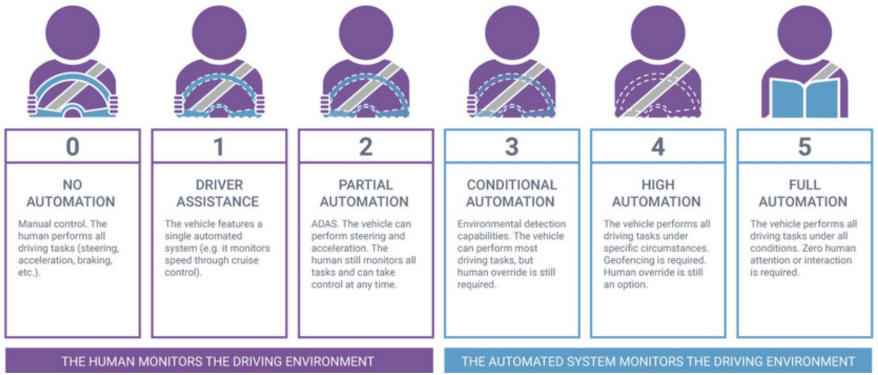
History of stunted progress in computer vision



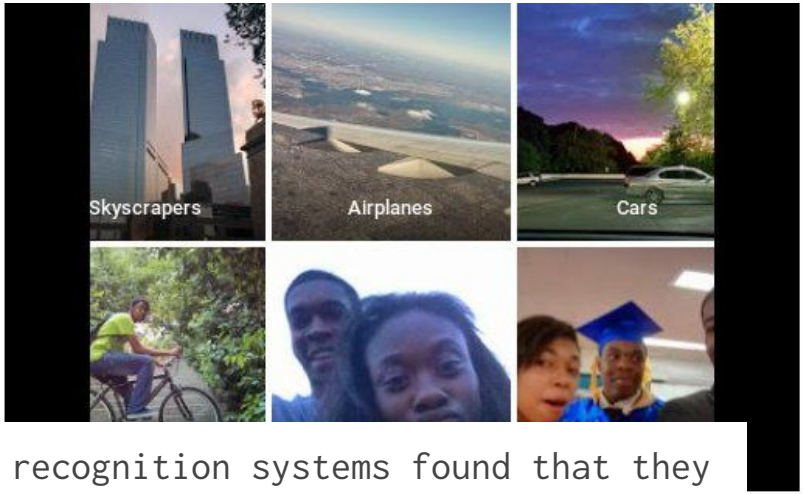
Case study / success stories of computer vision

- Automated checkouts - retail stores
- Medical imaging
- 3d model building
- Automotive safety
- Surveillance
- Fingerprint recognition / biometrics
- Machine inspection

The self driving car challenge



The race challenge in facial recognition



“In 2019, a federal study of over 100 facial recognition systems found that they falsely identified African-American and Asian faces 10 times to 100 times more than Caucasian faces.”

Why is it so difficult to interpret images?

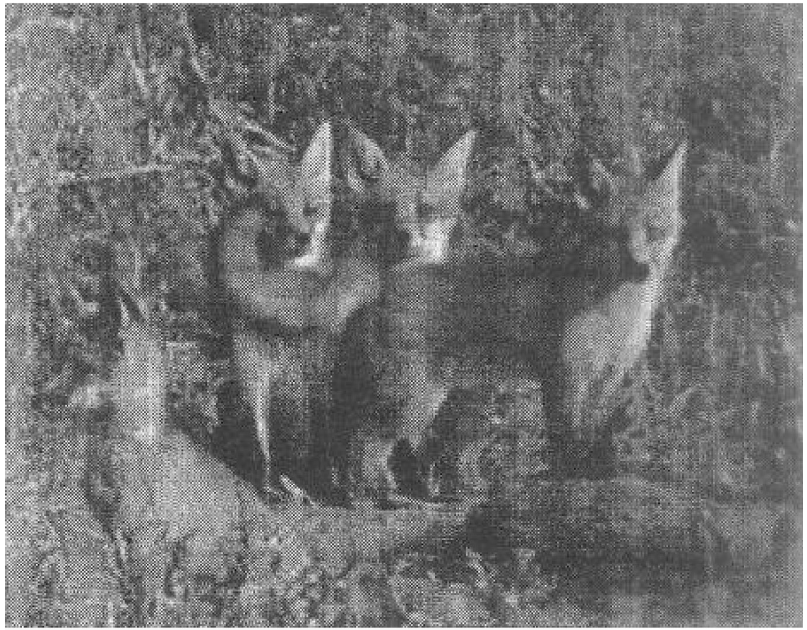
We don't understand how vision works in humans

The visual world is enormously complex

- Too many variables - orientation, lighting, occlusion
- An infinite number of scenes must be comprehended

Context is extremely important

Visual tasks are not purely data driven



Challenges this image presents to “bottom up” image analysis

- perform figure-ground segmentation of the scene (into its objects and background)
- infer the 3D arrangements of objects from their mutual occlusions
- infer surface properties (texture, colour) from the 2D image statistics
- infer volumetric object properties from their 2D image projections
- do all of this in “real time?” (if survival depends on it.)

When does the human brain struggle with images?

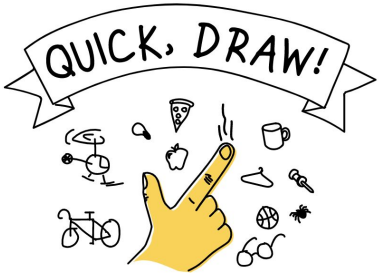
Describe what the image shows

Describe what happens in the video

Can you tell me who is in this image?

Applying the techniques

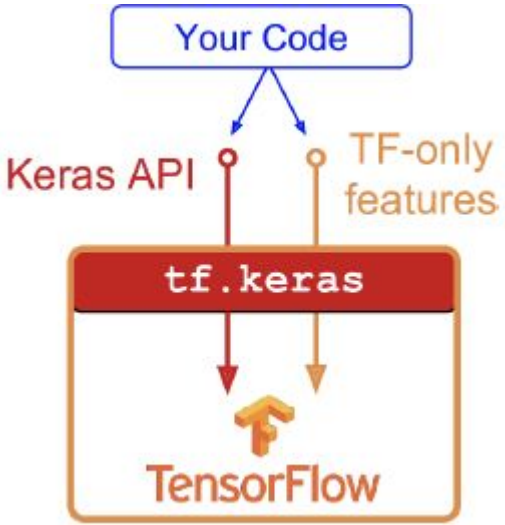
Tasks of computer vision



- Image classification with localisation - Labeling an x-ray as cancer or not and drawing a box around the cancerous region.
- Object detection - Drawing a bounding box and labeling each object in an indoor photograph.
- Object Segmentation - Segmenting all pixels in an image into different categories of object.
- Style Transfer - Applying the style of specific famous artworks (e.g. by Pablo Picasso or Vincent van Gogh) to new photographs.
- Image Colourisation - Datasets often involve using existing photo datasets and creating grayscale versions of photos that models must learn to colorize
- Object landmark detection - face recognition, pose recognition, emotion recognition
- Image Reconstruction - filling in missing or corrupt parts of an image.
- Image synthesis - Generating targeted modifications of existing images or entirely new images.
- Image captioning - Generating a textual description of an image
- Text to Image - Synthesizing an image based on a textual description - artistic style synthesis

“a giraffe standing on dirt ground near a tree.”

Lets try it- image classifier with Sequential (stack)



Start google colab

Import tensor flow & keras

Data set = fashion_mnist

Create validation set

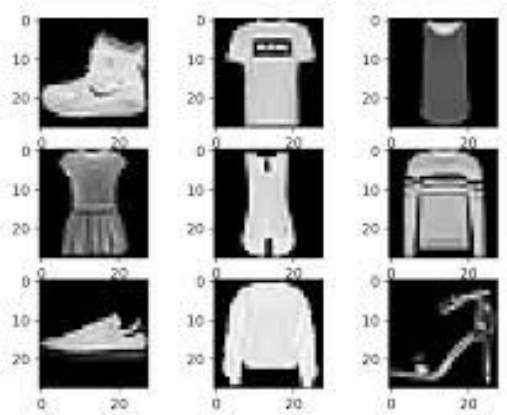
Scale data appropriate to algorithm

Define classes - use plot to validate y

Build neural network

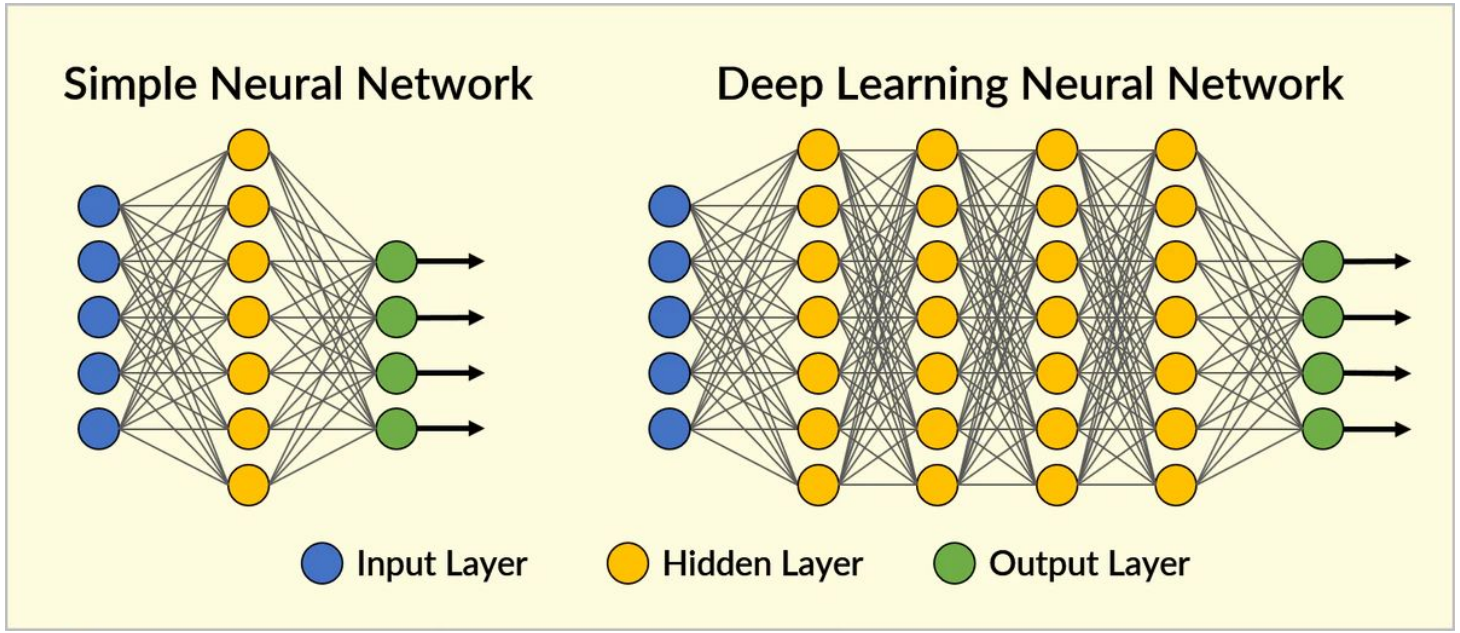
Compile model , Train & Evaluate

Predict using model (new images)

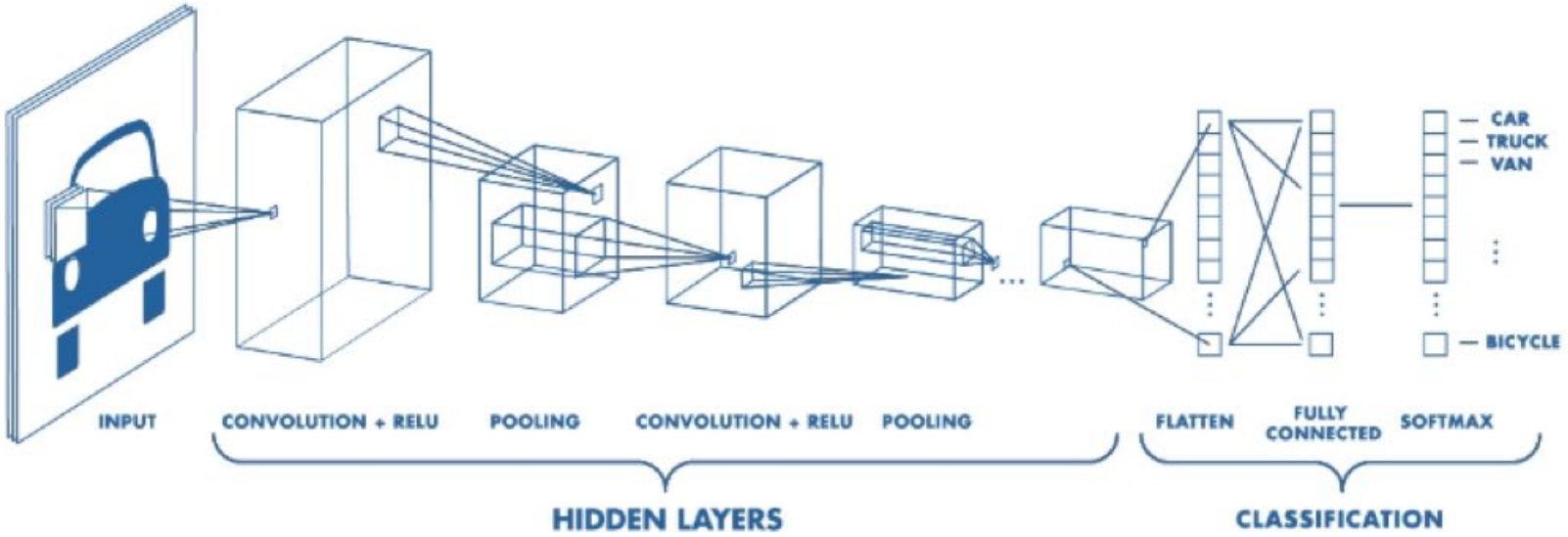


Label	Class
0	T-shirt/top
1	Trouser
2	Pullover
3	Dress
4	Coat
5	Sandal
6	Shirt
7	Sneaker
8	Bag
9	Ankle boot

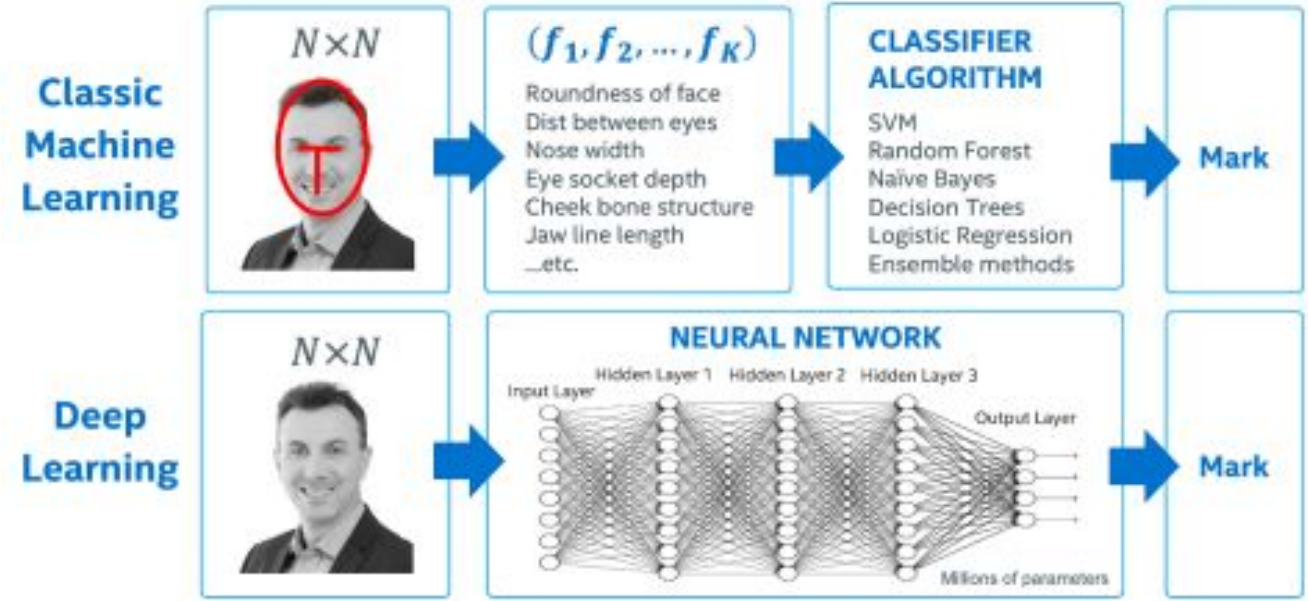
Neural network - artificially recreating the biological



Example CNN (convolutional neural network)



Deep learning NN v classic Machine learning

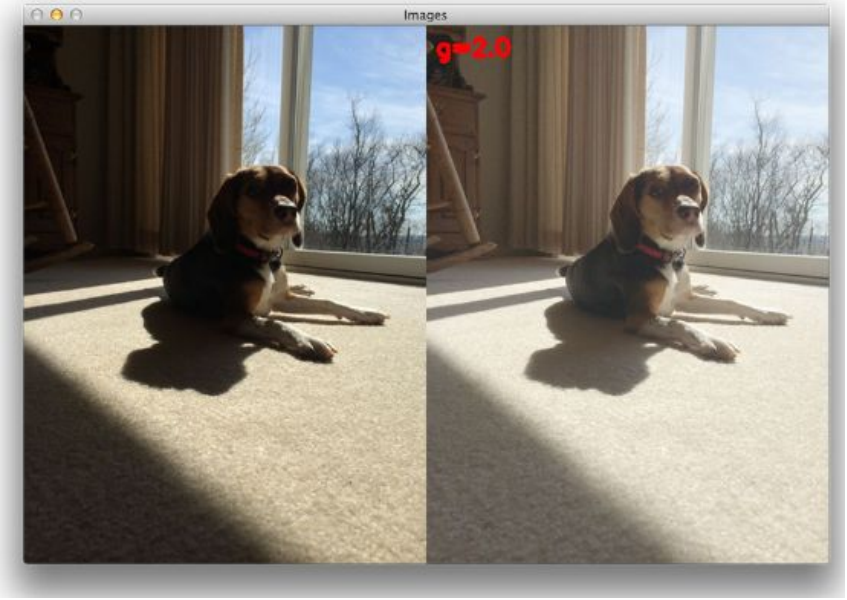


Required steps of any CV project

- Google colab (using google CPU, not your laptop)
- Keras + tensorflow ; sciPy *(eg. scipy.ndimage for 3d)*
- Images pre processing
- Data_augmentation processor
- Rescaling
- Dropout
- Construct a NN model appropriate for the scenario
- Train, Validation
- Inference + reusability
- Test and Evaluate

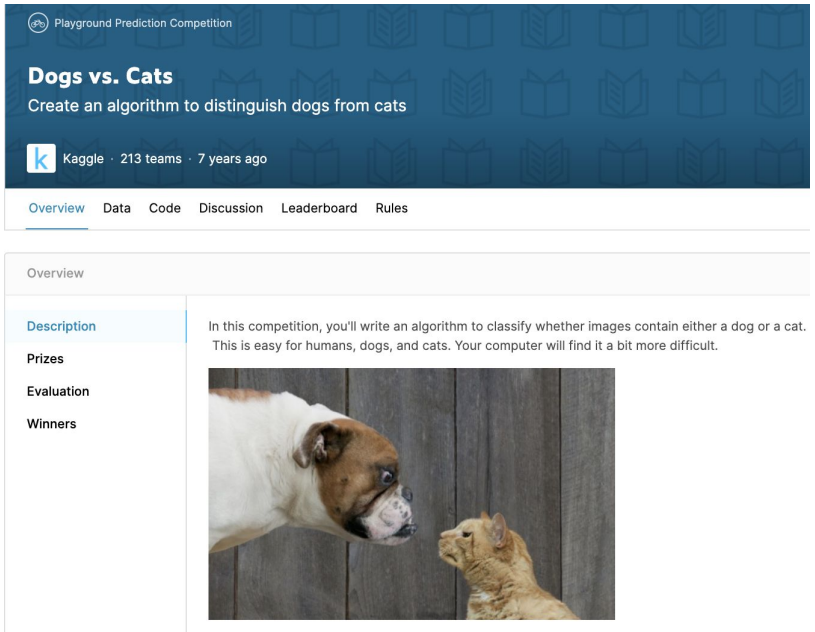
Image pre-processing

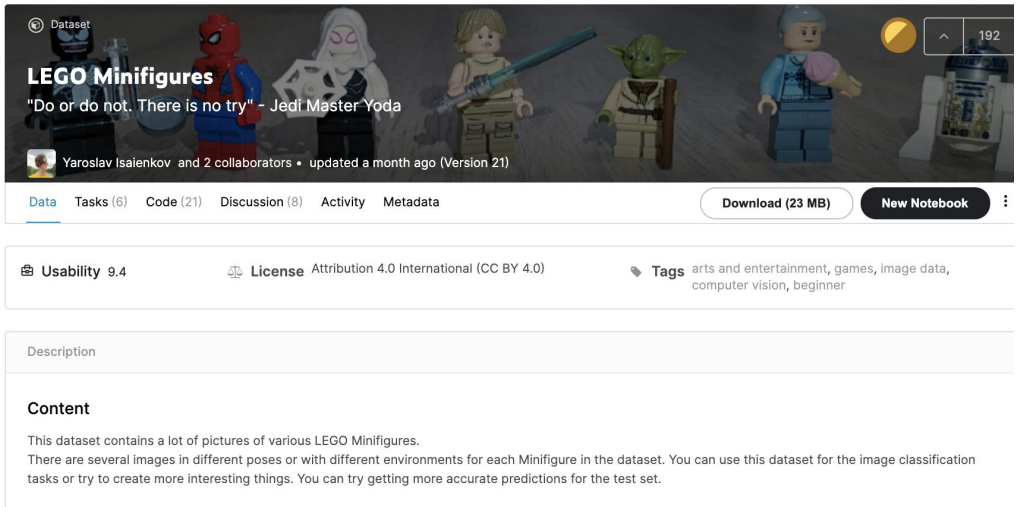
- Normalise brightness/ colour
- Crop image
- Geometric transformation
- Remove digital noise
- Filter and segment



(and by the way, you will want to deploy an enormous library of labelled training images)

Kaggle image resources





Google image resources

README.md

The Quick, Draw! Dataset

IMAGENET

14,197,122 images, 21841 synsets indexed

Download

- Download Image URLs
- Download Original Images
- Download Features
- Download Object Bounding Boxes
- Download Object Attributes
- API documentation
- FAQ

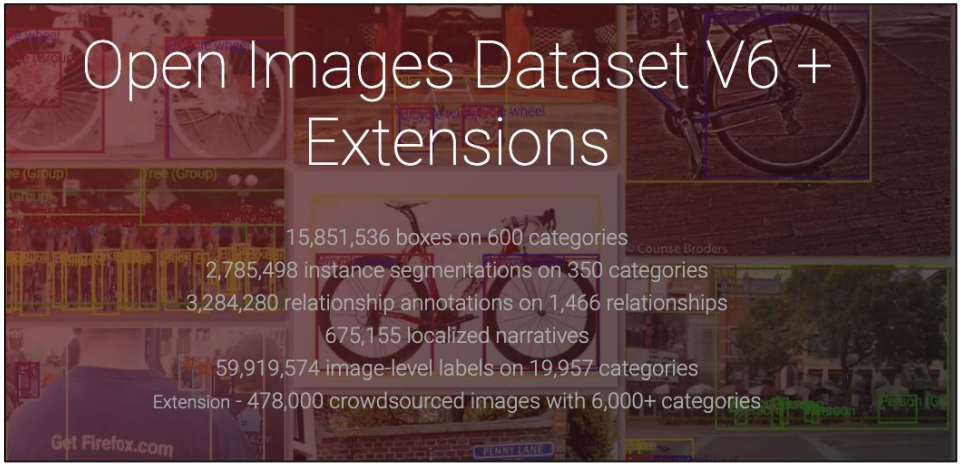
You can:

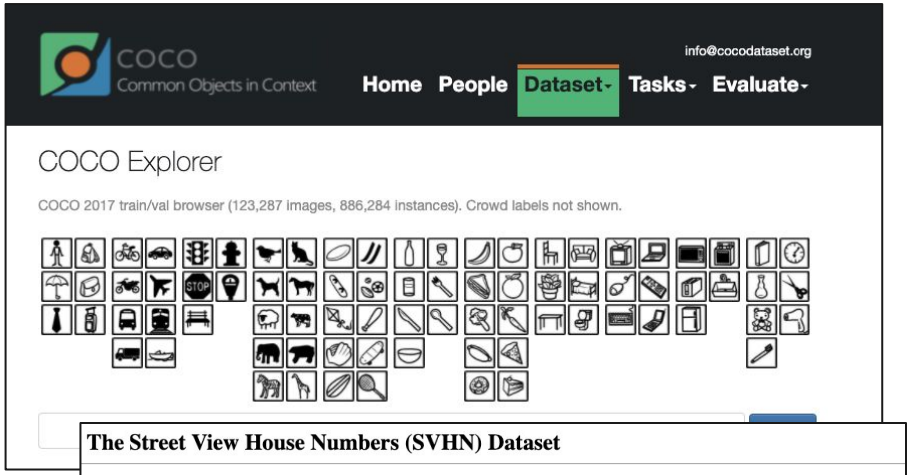
- Download Image URLs
- Download Original Images (for non-commercial r
- Download Features
- Download Object Bounding Boxes
- Download Object Attributes New!

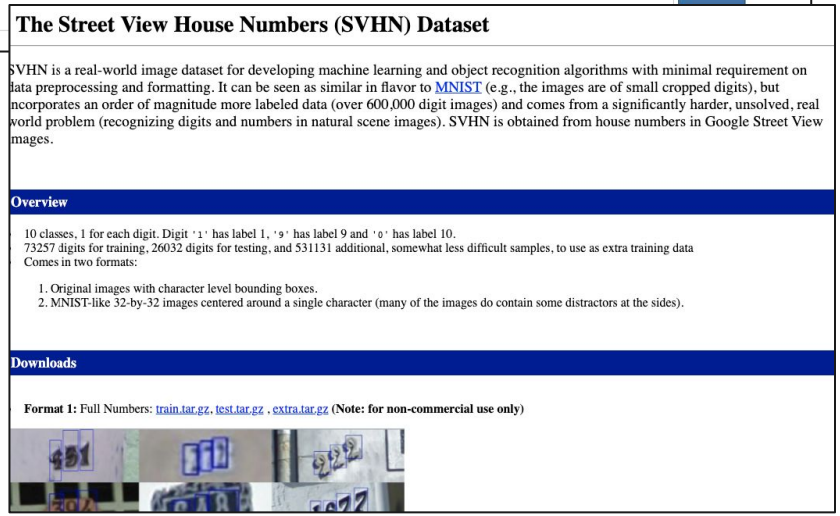
Next

To learn more about ImageNet API, Click Here.

Other image databases







questions?