Weston Robot

RSE2107A – Lecture 8

Deep Learning

Weston Robot

Agenda

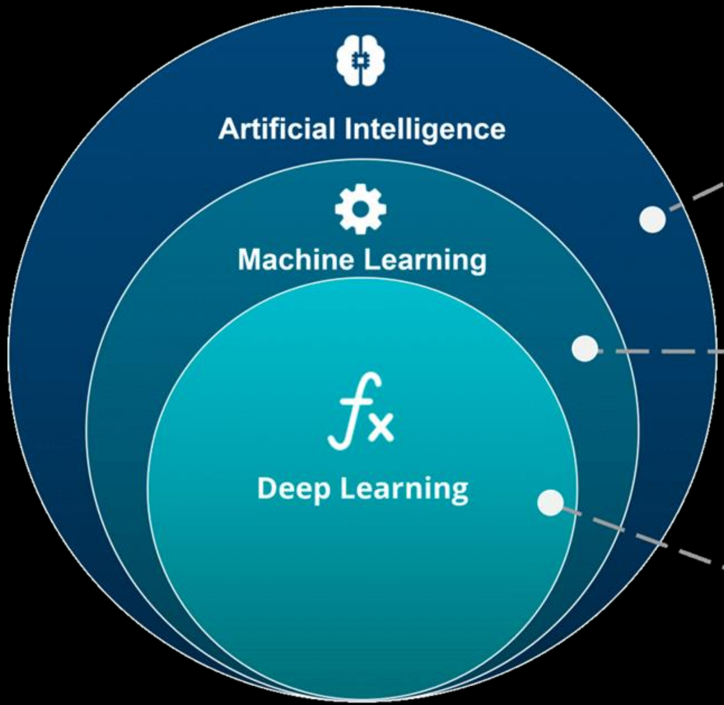
01
What is deep learning

02
Deep Learning in CV

03
Workflow for training & deployment of DL models

Deep Learning

Weston Robot



ARTIFICIAL INTELLIGENCE
A technique which enables machines to mimic human behaviour

MACHINE LEARNING
Subset of AI technique which use statistical methods to enable machines to improve with experience

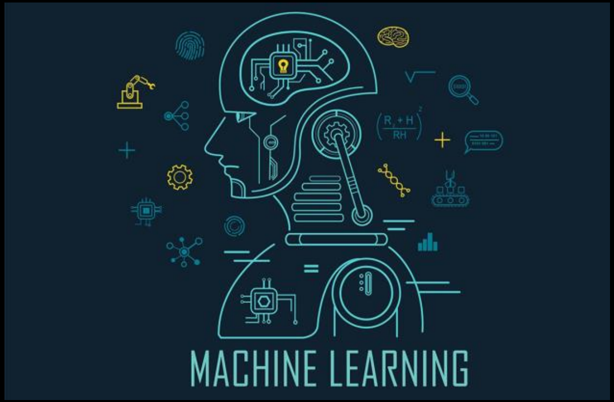
DEEP LEARNING
Subset of ML which make the computation of multi-layer neural network feasible

Artificial Intelligence

- Enables machines to think without any human intervention
- AI Systems falls into these 3 types
  - Artificial Narrow Intelligence (ANI)
    - Goal-oriented and programmed to perform a single task
  - Artificial General Intelligence (AGI)
    - Allows machines to learn, understand, and act in a way that is indistinguishable from humans in a given situation
  - Artificial Super Intelligence (ASI)
    - Capable of exhibiting intelligence that surpasses humans

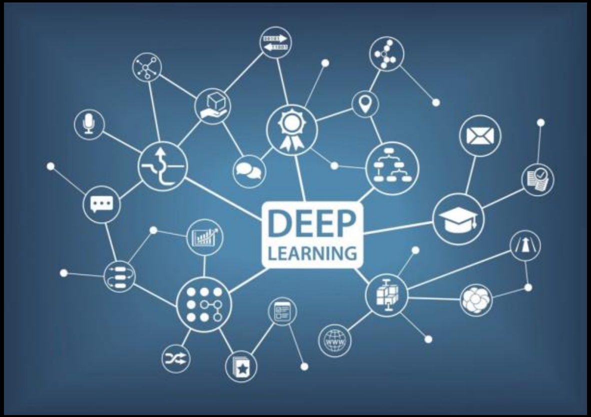
Machine Learning

- Uses statistical learning algorithm to build smart systems
- ML Systems can learn and improve automatically without explicitly being programmed
- Classified into 3 types
  - Supervised
  - Unsupervised
  - Reinforcement learning



Deep Learning

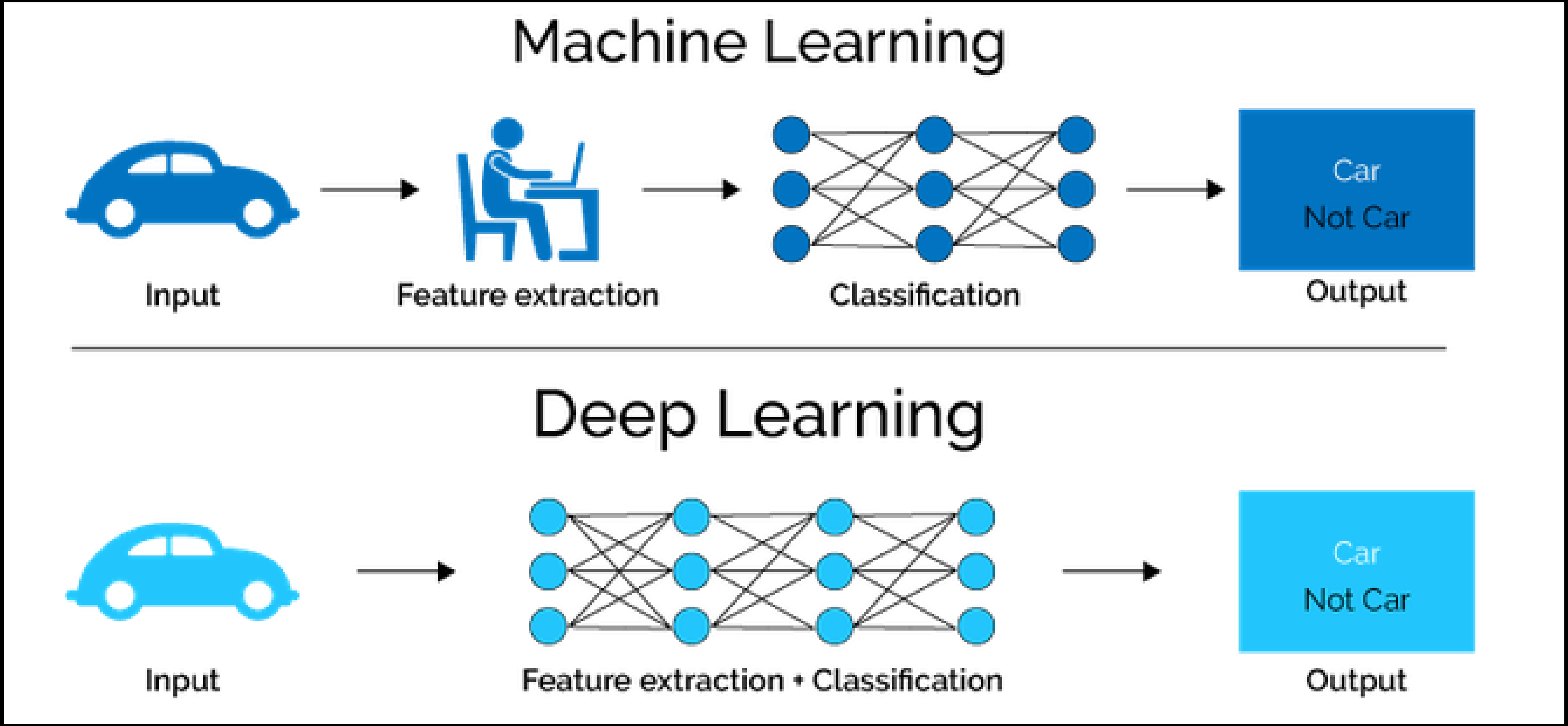
- Associated with learning from examples
- DL Systems help computer model to filter input data through layers to predict and classify information
- Similar to ML which requires large amount of data to learn and make informed decisions
- DL Network architectures
  - Convolutional Neural Networks
  - Recurrent Neural Networks
  - Recursive Neural Networks



Advantages of DL over ML

- Feature Extraction
  - ML algorithm requires pre-processing phase before being applied to data such as images, text, CSV
    - Example: To determine cat or dog using ML, we need to extract features like size, shape, etc and give these features to ML model to identify
  - DL, however, is capable of identifying these features without the need for features extraction, as it is capable of classifying it on its own

Advantages of DL over ML



Weston Robot

Advantages of DL over ML

- Big Data
  - Accuracy of ML will stop increasing at a certain point in time, even with an increasing amount of data
  - DL however, will increase its accuracy with increasing data
    - Like how brain works

Neural Network

Weston Robot

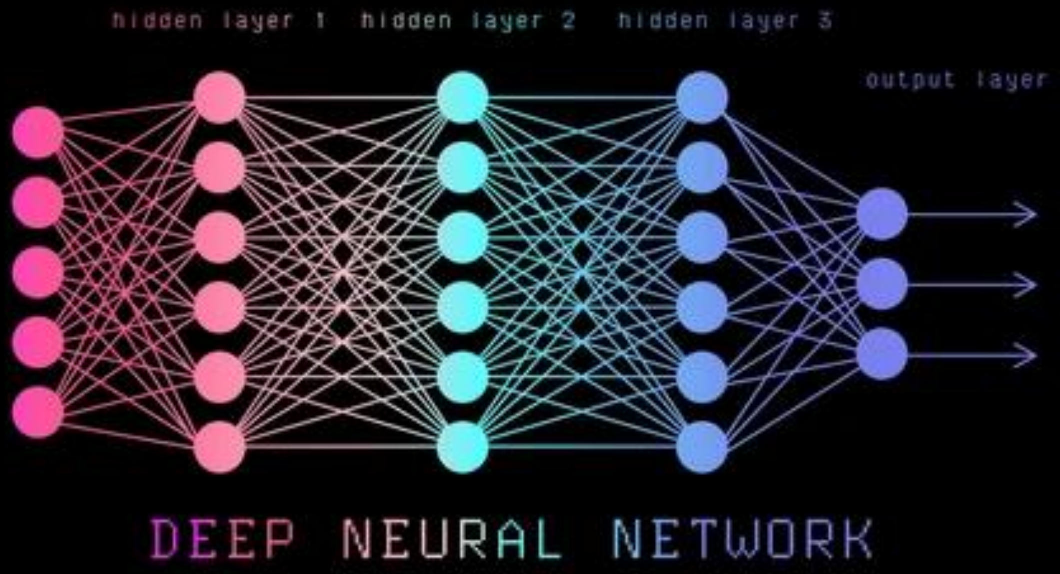
What are neural networks?

- Neural networks are a subset of machine learning and are at the heart of deep learning algorithms. Their name and structure are inspired by the human brain.
- Also known as artificial neural networks (ANNs), they comprise of an input layer, one or more hidden layers, and an output layer.
  - input layers take in raw inputs from a dataset
  - hidden layers take their inputs from the output of the previous layer, processes it and passes its own output onto the next layer.
  - output layers are the last layer of the ANN and will make a final prediction of the initial inputs might be.

Weston Robot

Deep learning neural networks

- A neural network that consists of more than three layers, which would be inclusive of the inputs and the output, can be considered a deep learning algorithm.



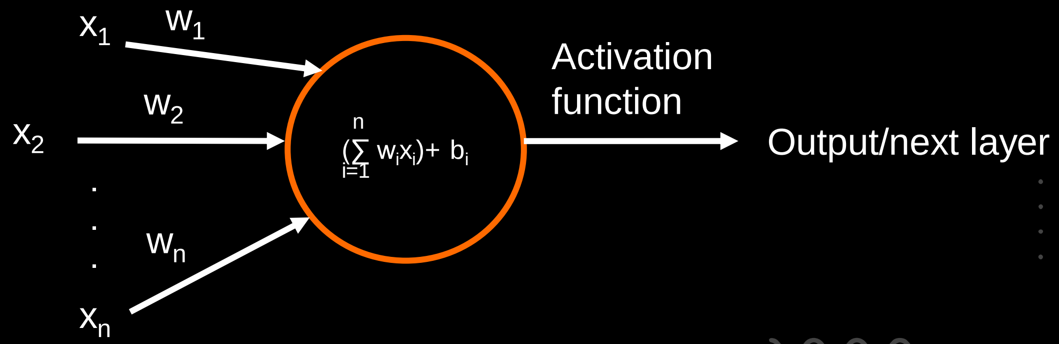
How do neural networks work? - Layers

- Each layer in a neural network are made up of multiple nodes (neurons), each connected to every node in the next layer.
- Each node represents an independent variable that affects the output that the neural network model is trying to predict.
    - Eg. A model that can predict the price of houses(output) will have input variables (nodes in the first input layer) such as age, house space, surrounding amenities, etc.

Weston Robot

How do neural networks work? - Nodes

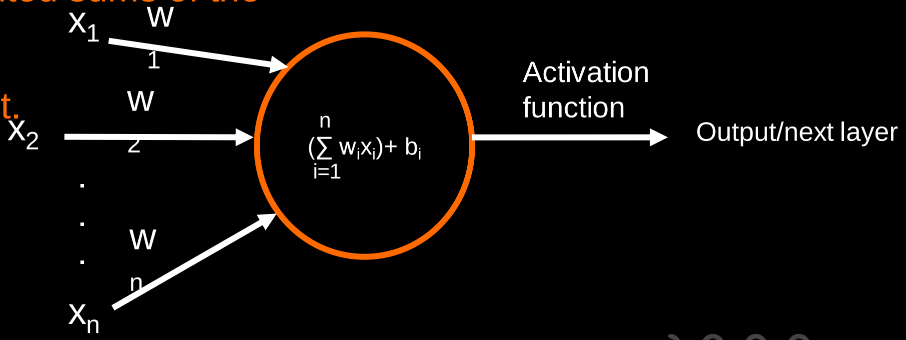
- Nodes are the smallest computational unit in a neural network, each doing a fixed mathematical formula (can differ across nodes).
- Each node calculates an output based on the weighted (W) inputs (X) from the previous layer's nodes' outputs and a given bias using its own formula (activation function).



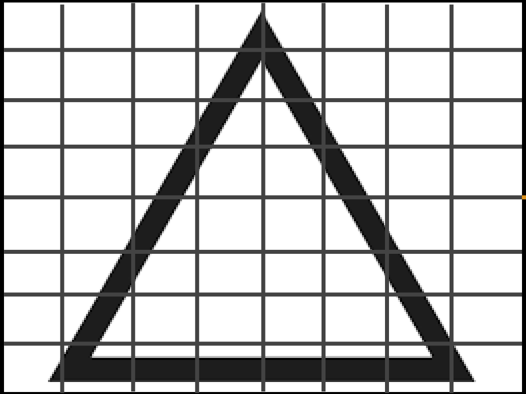
Weston Robot

How do neural networks work? - Nodes

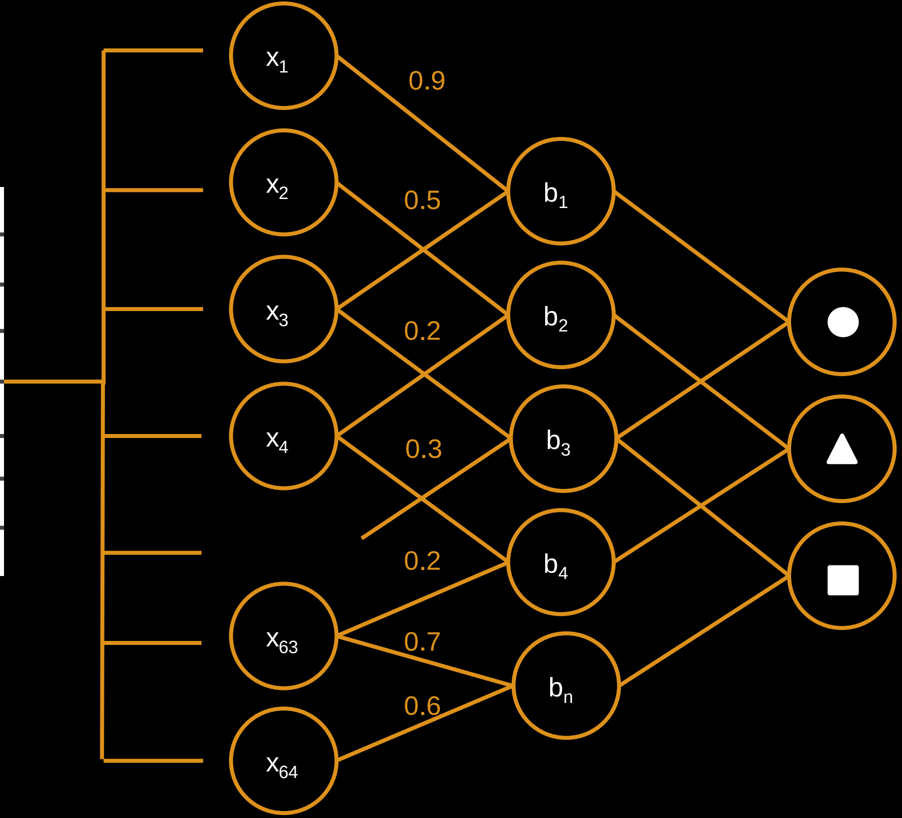
- Inputs (x)
  - Outputs from nodes in previous layer
- Weights (w)
  - Determines how much each input affects the output of the given node
  - Higher == Stronger
- Biases (b)
  - Used to independently adjust the weighted sums of the inputs.
  - Helps the network to best fit the dataset.
- Activation Function (f)
  - Computes the output of the node
- Output = $f(\Sigma_i(w_i \cdot x_i) + b)$



How do neural networks work?

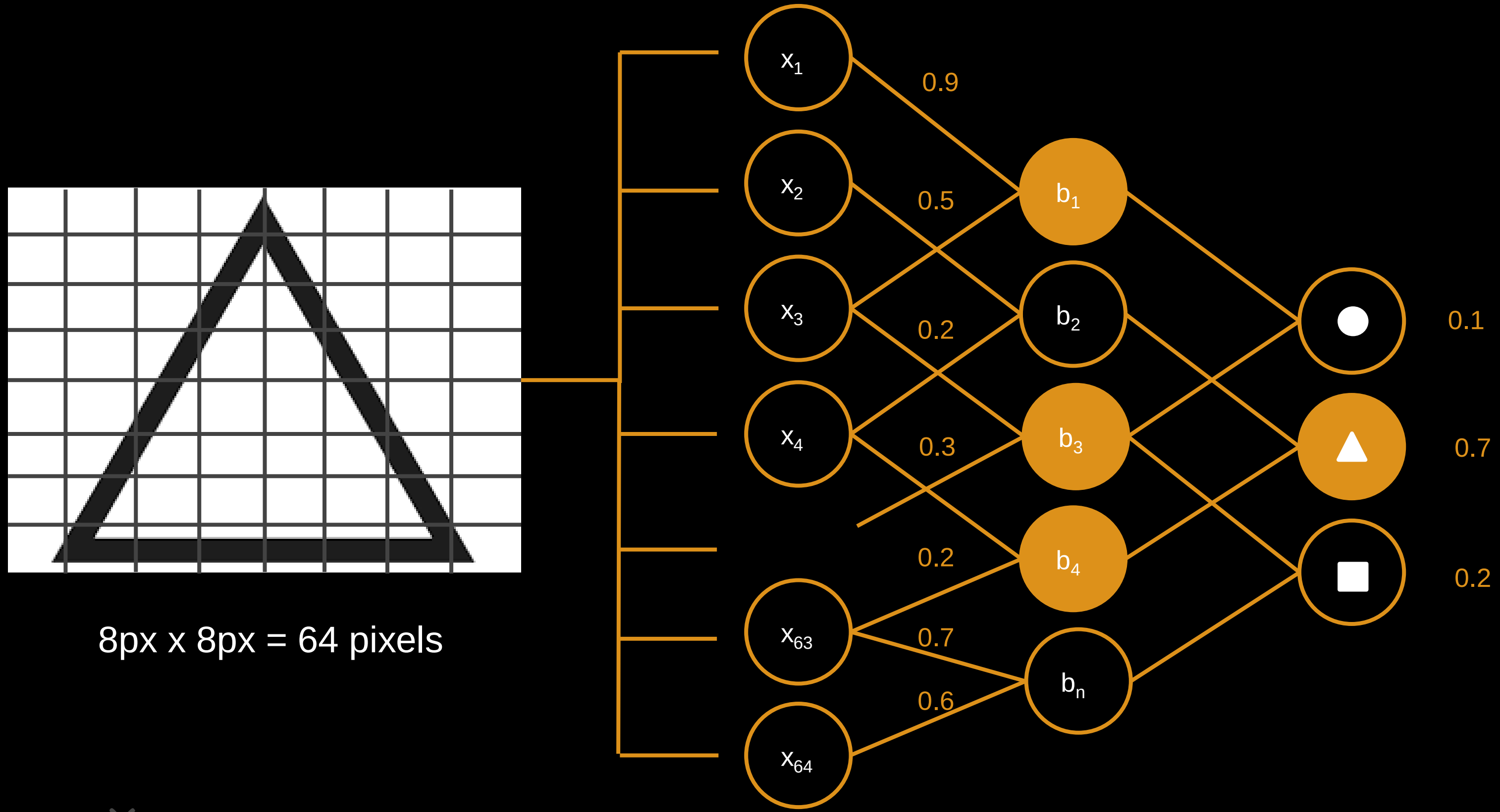


8px x 8px = 64 pixels



Weston Robot

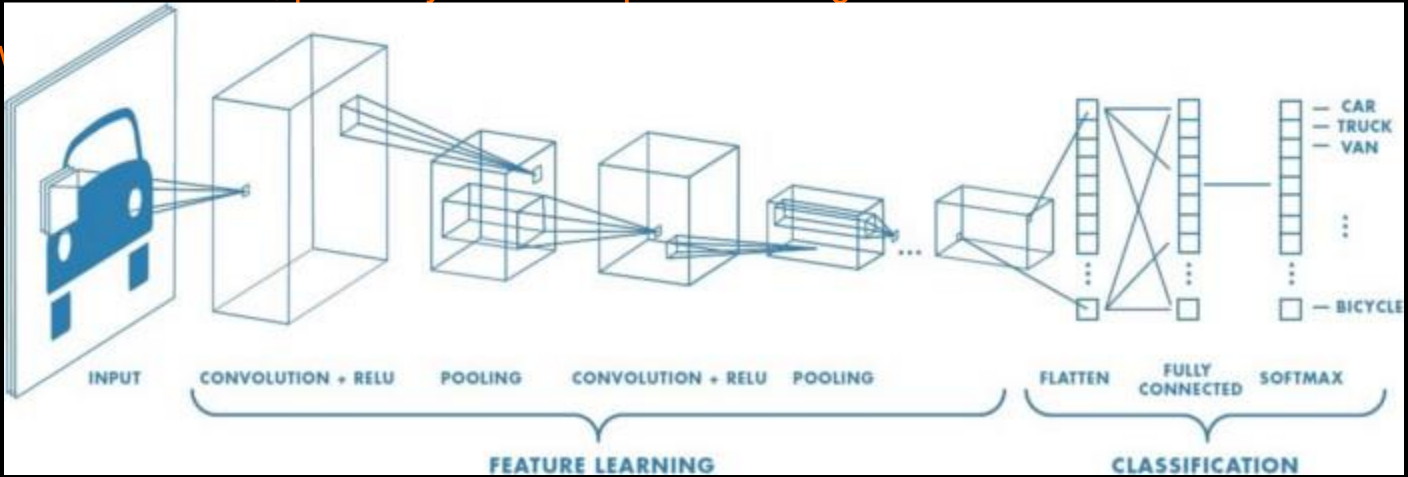
How do neural networks work?



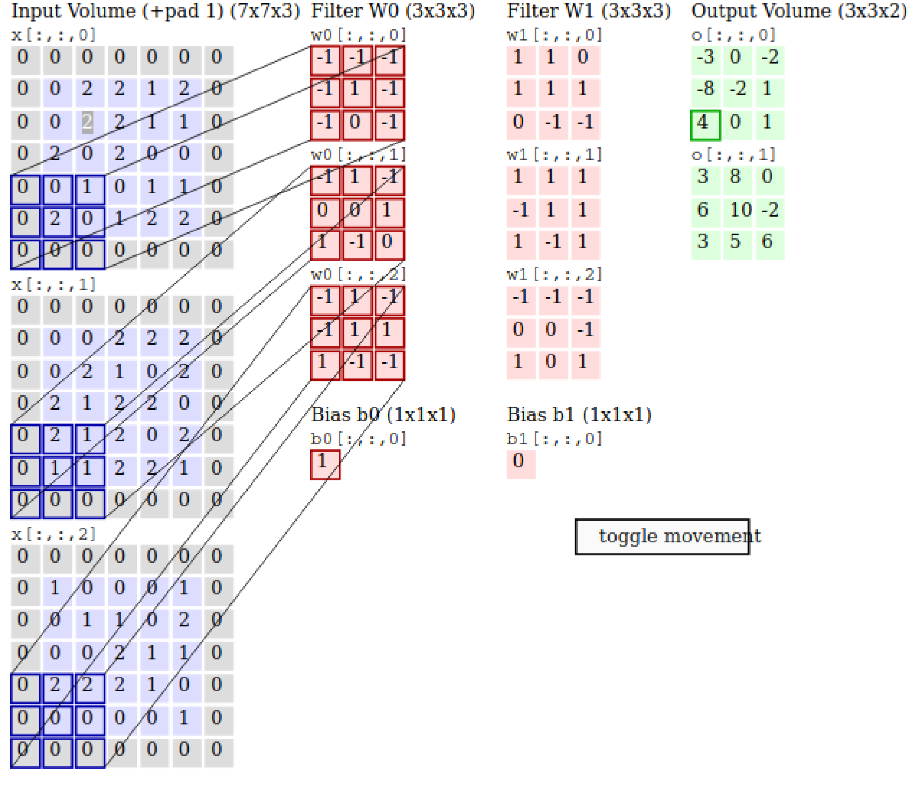
8px x 8px = 64 pixels

CNNs - Convolutional Neural Nets

- A convolutional neural network (CNN) is a type of artificial neural network used in image recognition and processing that is specifically designed to process pixel data.
- The advancements in Computer Vision with Deep Learning has been constructed and perfected with time, primarily over one particular algorithm — a Convolutional Neural Net

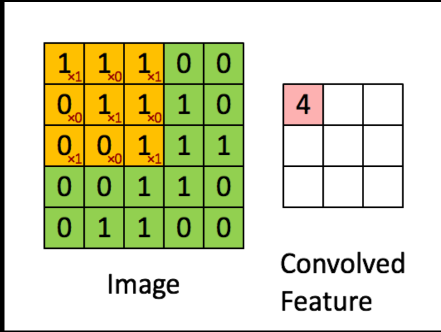


CNNs - Convolutional Neural Nets

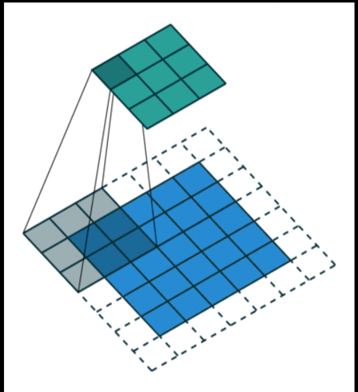


- The Convolution layer is the core building block of a Convolutional Network that does most of the computational heavy lifting.
- The role of the Convolution layer is to reduce the images into a form which is easier to process, without losing features which are critical for getting a good prediction.
- This processing is done by a kernel/filter.

CNNs - Convolutional Neural Nets



Convoluting a 5x5x1 image with a 3x3x1 kernel to get a 3x3x1 convolved feature



Convolution Operation with Stride Length = 2

- The objective of the Convolution Operation is to extract the high-level features such as edges and contours, from the input image.
- With added layers, the model adapts to the High-Level features as well, resulting in a network which has a better understanding of the images in the dataset.

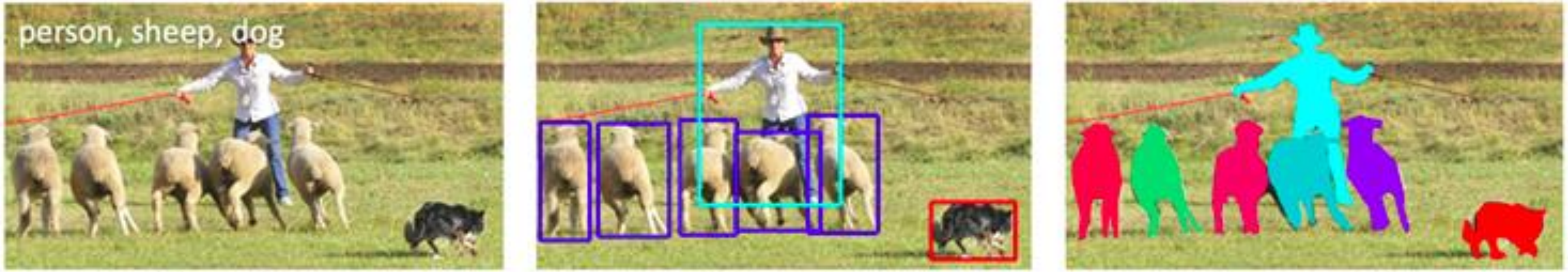
CNNs - Convolutional Neural Nets

- There are several other layers and functions that can be found in general CNN model architectures. These include
  - Pooling layers
  - Activation layers/functions
  - Fully Connected Layers
- On top of these there are Loss Functions.
- Finally we have our optimization algorithm. The most commonly used, is the Stochastic Gradient Descent (SGD).

Deep Learning in CV

Weston Robot

Types of Image Recognition

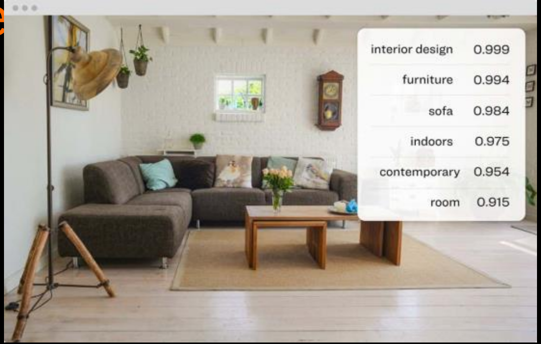


(a) classification (b) detection (c) segmentation

Weston Robot

Classification

- Type of labeling where image/video is assigned certain concepts/categories
  - Goal is to answer “What is there in this image/video”
  - Does not tell where the objects are located at nor how many there are, just that it exists in the given image



Segmentation

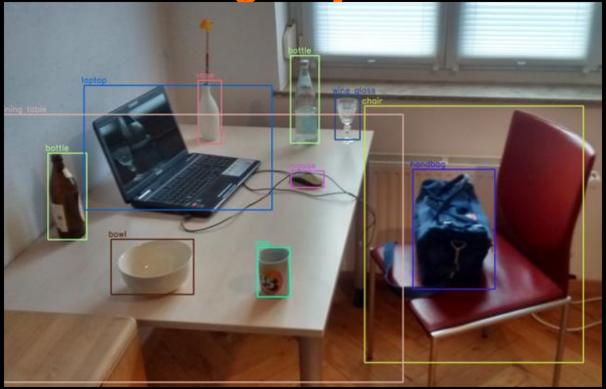
- Type of labeling where *each pixel* of an image is labeled with given concepts
  - simplifies an image or changing how it is presented to the model, making it easier to analyse
- Provides the exact outline of an object within an image
- Requires object classification and object detection to have took place before segmentation can be used
- Useful if there is a need to ignore the background of an image

Typical Object Detection Workflow

Weston Robot

What is Object Detection?

- Refers to detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos.
- Object detection has applications in many computer vision tasks such as face recognition, vehicle counting and tracking objects.

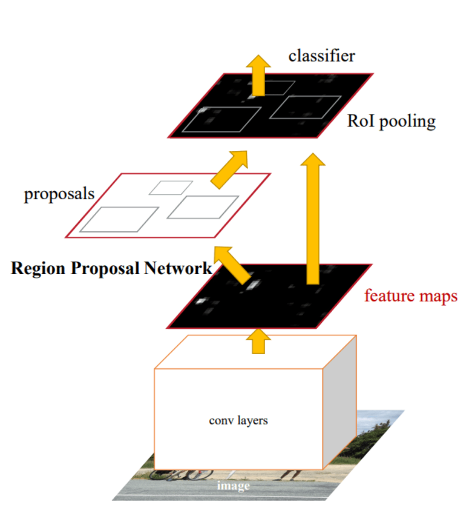


Weston Robot

How is Object Detection done?

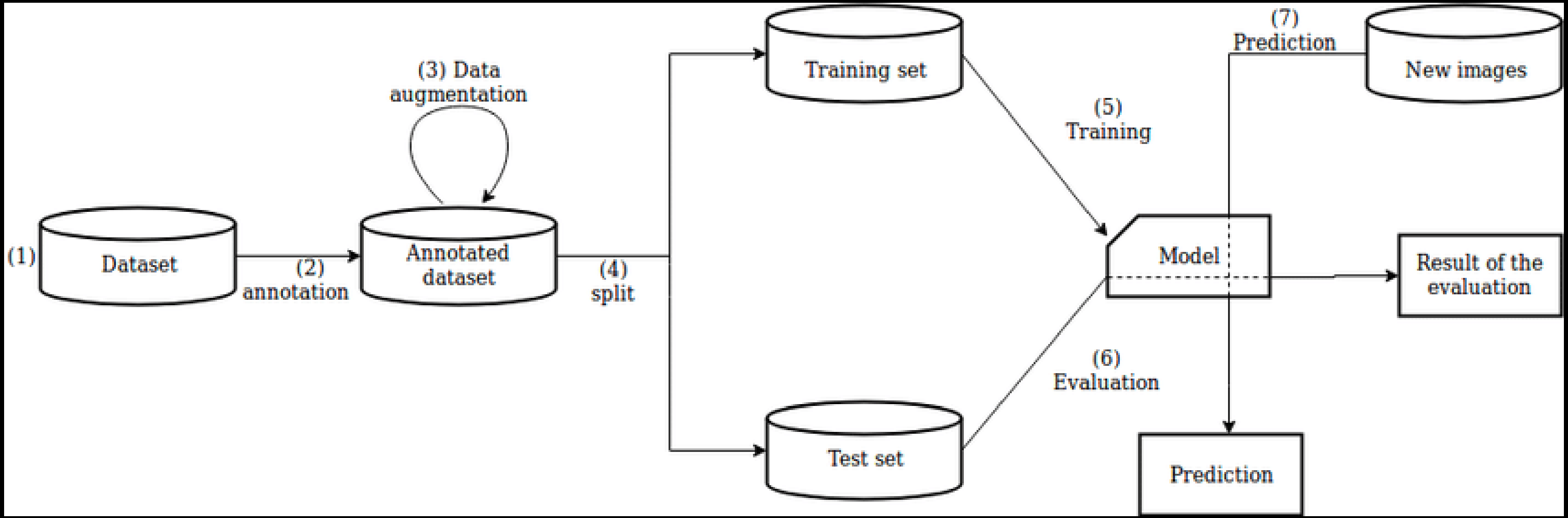
- Digital images are represented as array of integer values
- Image (defined as an array of values) is passed through a neural network
- Neural network makes predictions on the image & d

boxes to identify and locate objects with a certain pr



Weston Robot

Object Detection Workflow



Data Collection

1. Data collection - Methods

- Method 1: Taking from online datasets
  - Many open source datasets available online
  - Can come fully/partially annotated, or not annotated at all
  - Eg. Kaggle, COCO dataset, Github etc...
- Method 2: Self-source data
  - Taking Photos & Videos

Considerations

- Clearly define the object types to identify
  - General object types are hard for the model to distinguish
    - Eg. Trash can come in many forms, which makes it hard to distinguish
  - Use types where the objects have largely similar attributes
    - Eg. Splitting trash into plastic bottles, paper cups, drink cans etc
  - This will help in improving the accuracy of your model

Considerations

- How much data to collect?
  - Collecting more varied data is better for most cases
    - However, if your model is only required to identify objects in fixed scenarios/specialised environments, varied data can instead decrease the accuracy of the model.
  - Dependent on quantity of object instances in each image, rather than the total number of images
  - Best to maintain equal amounts of images for each class to prevent class imbalance
    - Otherwise, it can result in low accuracy for infrequent objects

Data Cleaning & Preparation

Annotation of images

- Object detection is a supervised learning problem, where our model has to be trained on annotated examples.
- Used to identify objects and tell the computer where the object is in the image
- Bounding boxes are drawn to identify & locate the objects in the images
- Use labelling tools for annotation of images
  - Roboflow
  - CAT
  - LabelImg

Weston Robot

Annotation of images

- Good practices
  - Tight-Fitting bounding boxes
  - Little/No overlap for bounding boxes
  - Cover full size of the image

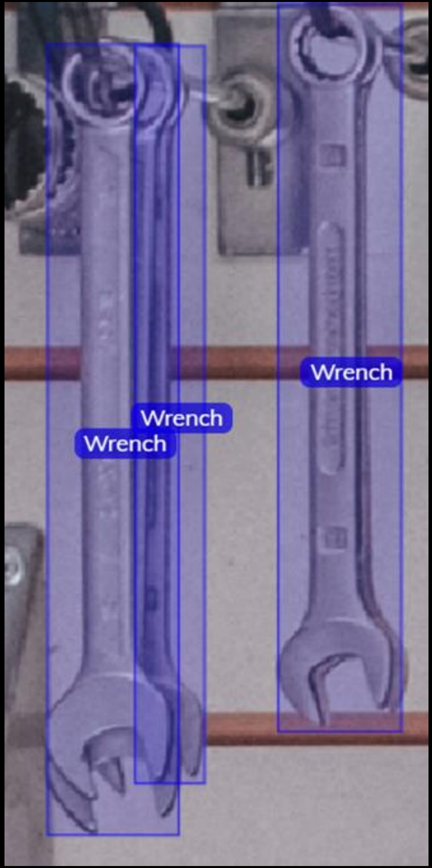
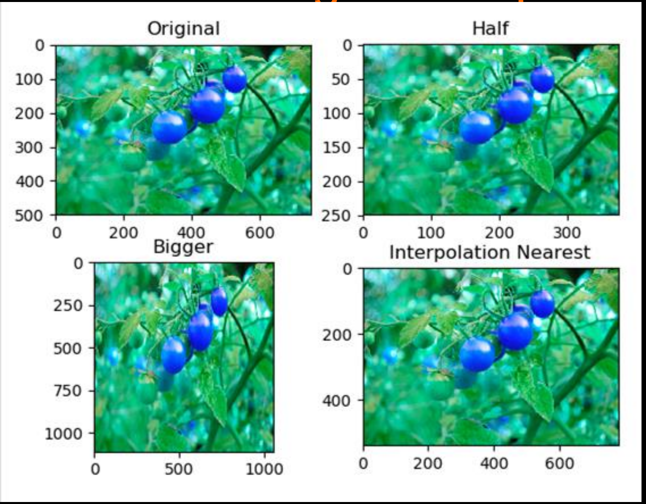


Image Pre-processing

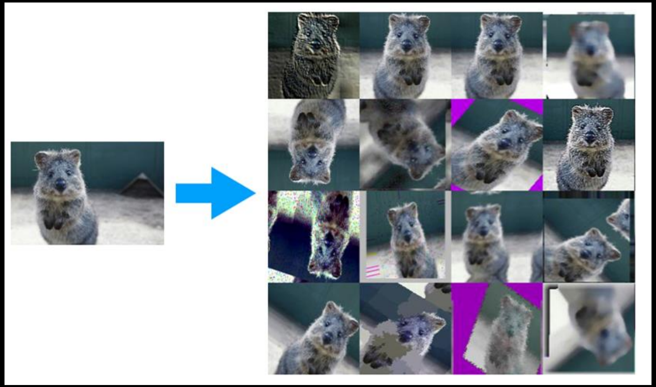
- Technique that suppresses undesired distortions or enhances some image features relevant for further processing and analysis task.
- Used to remove unwanted/noisy data from the images and possibly reduce training time
- Examples
  - Image Resizing
  - Cropping of images
  - Image Auto-Orientation



Data Augmentation

Weston Robot

- Technique that applies variations to the images within the dataset and appends them to it
- Used to increase the size of our our training dataset
- This helps us to create a more robust and generalized model
- Examples:
  - Adjusting Brightness & Exposure levels
  - Blurring/Cropping parts of the image



Train-Test Split

- Refers to taking a dataset and splitting it into 2 subsets – Train & Test dataset
- Used to estimate the performance of the model on new data: data not used to train the model
- The train dataset is used to train the model while the test dataset is used to evaluate the model
- Minimizes overfitting as the model is evaluated on new data

Model Training

Weston Robot

3. Model Training

- The process which a Deep Learning Algorithm is being fed with sufficient training data to learn from, which helps it to identify and learn good values for all attributes involved
- Inputs from the training dataset are fed into the model, and training is done based on the deviation of the processed result from the documented result.
- Consistent training can significantly improve the prediction rate of the model
- 2 key methods:
    - Creating a Deep Learning Algorithm
    - Using a pre-trained Deep Learning Algorithm

Creating a Model

- Neural Network is usually loaded with random weights
- Upon each iteration of the training dataset running through the model, the weights will be re-calibrated and adjusted, based on the difference between the predicted and actual output.
- Not recommended for beginners as it takes a significantly longer amount of time to train a model from scratch

Using Pre-Trained Models

- Also known as Transfer Learning, we can use pre-trained models as the base for training our object detection model
- Pre-Trained Model examples include:
  - YOLOV5
  - Resnet50
  - EfficientNet
  - and more….
- Different pre-trained models are more suited for different use cases.
  - Example: Resnet50 is more accurate compared to the EfficientNet, however the Resnet50 is computationally more expensive compared to the EfficientNet.

How to train?

Weston Robot

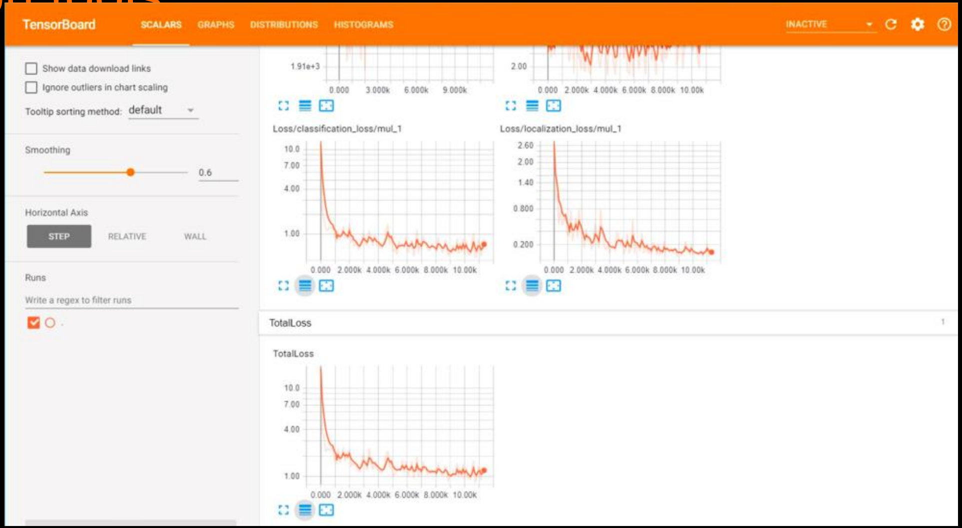
- Make use of Object Detection APIs
  - YOLOV5 API using PyTorch
  - TensorFlow Object Detection API
- Take note of configurable training parameters:
  - How many epochs to run?
    - Epochs refers to the number of runs the algorithm goes through the training set
    - Running more epochs usually results in better training results
    - However, running too many epochs might result in overfitting of the model
  - Image Size used for training
    - If too big, it will stretch the images and cause them to lose its aspect ratio
    - If too small, images might not have enough features for model training and learning

Model Evaluation

Weston Robot

4. Model Evaluation

- Use the model to run inferences of test images & videos
- Use training metrics to determine the accuracy of the model
- Training metrics visualization tools
    - Weights & Biases
    - Tensorboard



Mean Average President (mAP)

- A popular metric that tells you how precise your classifier is (how many instances it classifies correctly), as well as how robust it is (it does not miss a significant number of instances), where it computes the average precision value for recall value over 0 to 1.
- Precision refers to the number of correct positive results (True Positives) divided by the number of positive results predicted by the classifier (True & False Positives).
- Recall refers to the number of correct positive results (True Positives) divided by the number of **all** relevant samples (True Positives & False Negatives)

Model Deployment

Weston Robot

5. Model Deployment

- If user is unsatisfied with the performance of the model, it is possible to re-train on the same model once again
- Else, the model is then ready to be exported for deployment
- The newly trained model can be exported into numerous formats for different use cases, including:
  - Tensorflow-Lite
    - Running inference models on edge devices
  - Nvidia-DeepStream
    - Can use for running inference models for Nvidia controllers (ie. jetson nano etc)

Weston Robot

Appendix

- Popular Online Datasets
  - Kaggle Dataset https://www.kaggle.com/datasets
  - COCO Dataset https://cocodataset.org/#home
- Object Detection APIs
  - Tensorflow https://tensorflow-object-detection-api-tutorial.readthedocs.io/en/latest/
  - YOLOv5 https://github.com/ultralytics/yolov5/releases