Object Detection

Jannah, Ricky, Ken, Alakh, Michelle

Application: Self-Driving Cars (Video)

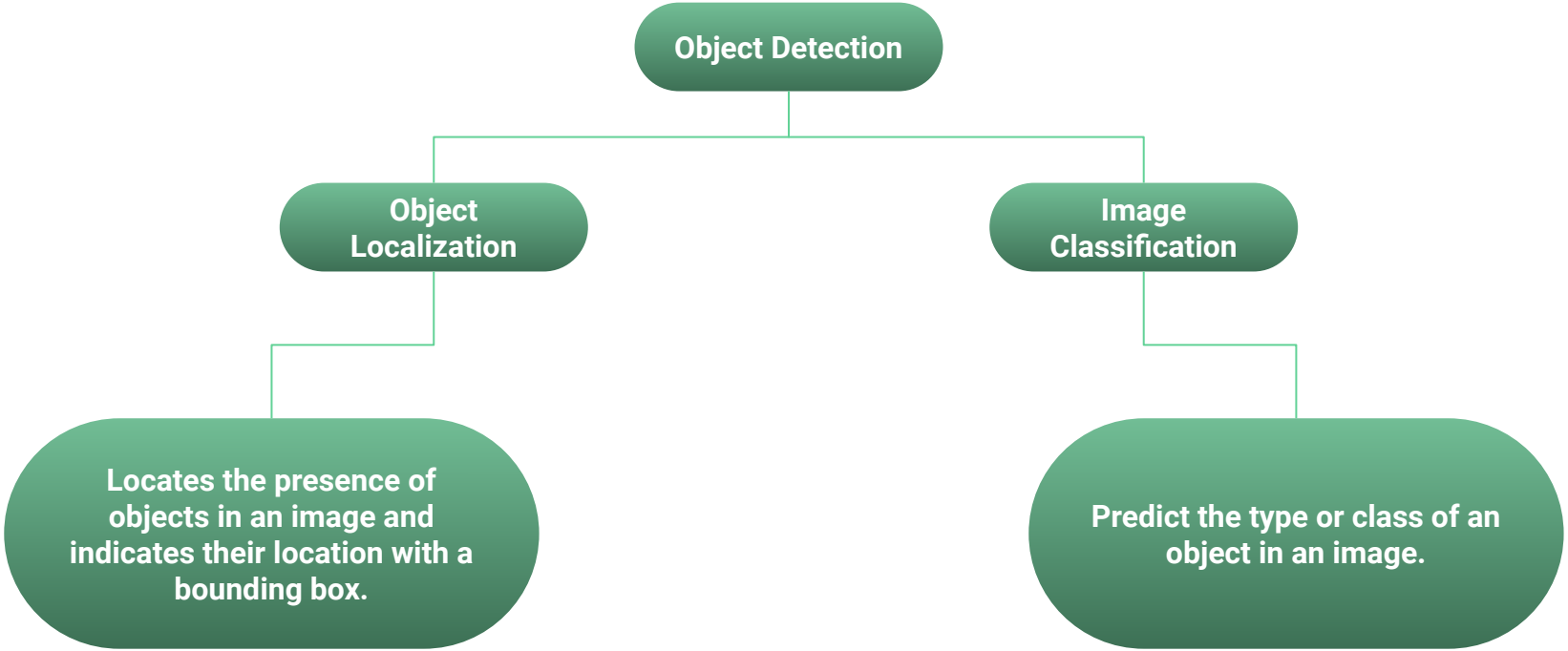
Why?

- Safety, efficiency, makes driving more accessible

How?

- Builds a map, plans a path, avoids obstacles, uses sensors, radar, camera...and **Convolutional Neural Networks!**

Problem Scope and Subproblems



Breaking Down the Task

Step 1 = detect objects while driving (including other cars, the road, traffic signs, pedestrians, etc.) excluding non-meaningful surroundings (trees, sky, roadkill, etc.)

Step 2 = classify detected objects (such as traffic lights, trucks, and all other meaningful data)

Object Localization = (step 1) locate the presence of objects in an image and indicate their location within a box

Image Classification / Object Recognition = (step 2) predict the type or class of an object (classify the object) in an image

Process

Sliding Window

Object Localization

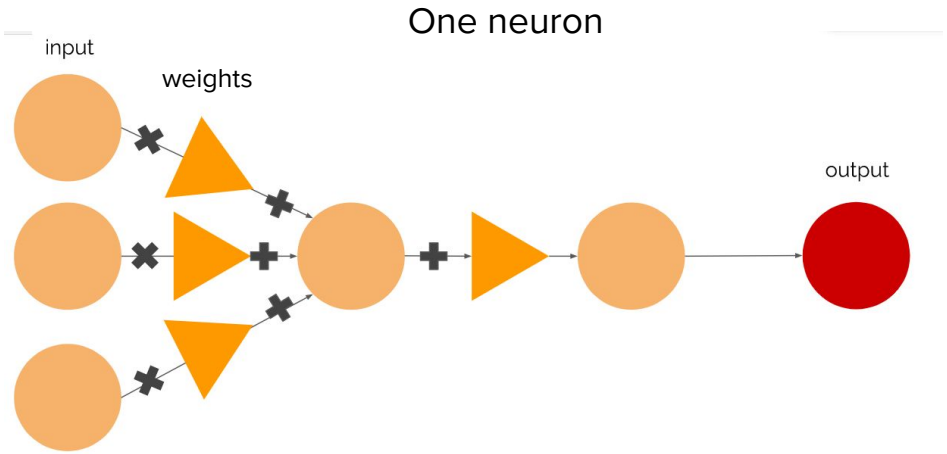
Image Classification

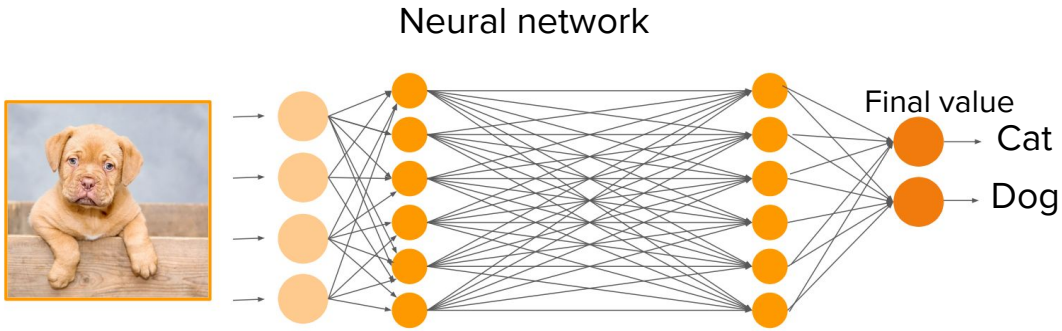
Sliding Windows

- Detects objects
- Crops images to prepare them for classification
- Reduces one big image into multiple smaller images
- Window slides through image continuously

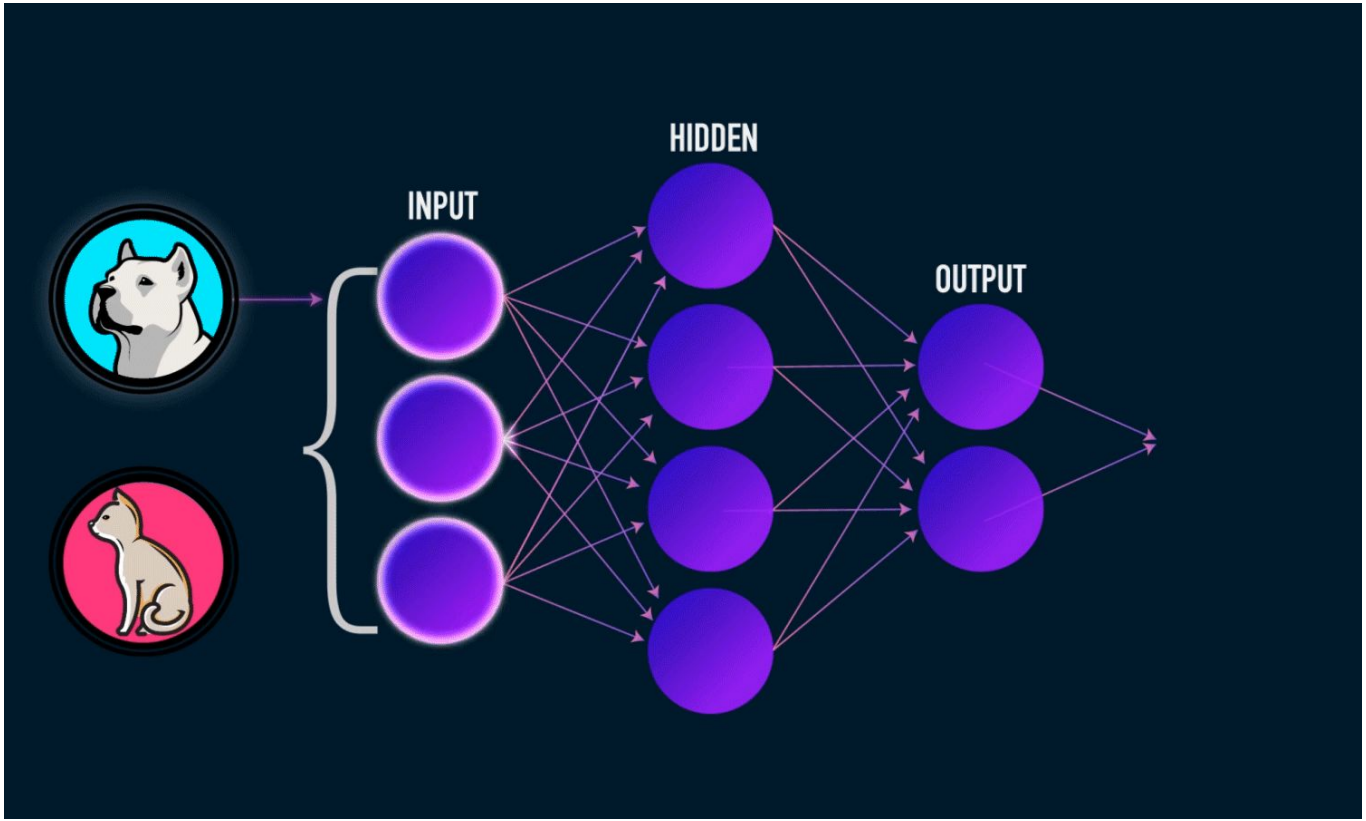
Neural Networks

- Imitates a brain's neuron
- Inside a "neuron" an input value goes through "weights" (conditions) to become an output value
- Neural network is a collection of these neurons
- Final output value will determine the label of image



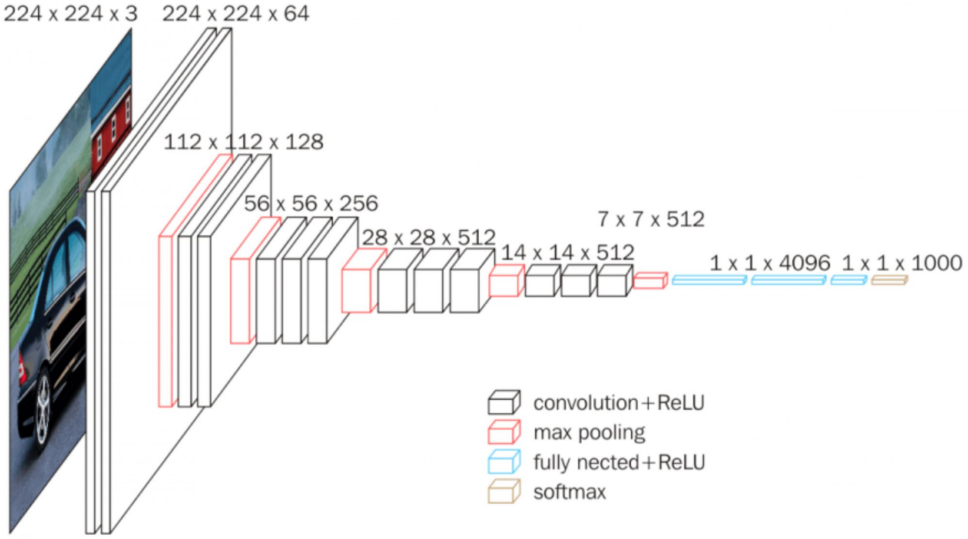


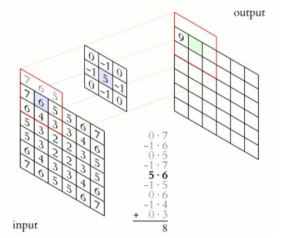
NN Prediction Process

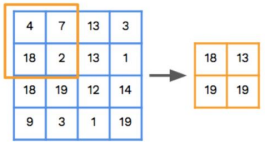


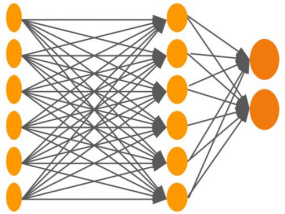
What is a CNN Model

1. Convolutional layer: apply kernel (remove background noise from image) to image
2. Pooling layer: simplify image further (ex: max pooling, greatest values kept)
3. Flattened layer: make image (3D bc RGB) 2D because CNN takes numerical data
4. CNN: neural network predicts the outcome









CNN Prediction Example

probabilities sum to one

0.15	airplane
0.01	dog
0.32	pizza
0.49	**hot dog**
0.00	sofa
0.03	bird

Transfer Learning

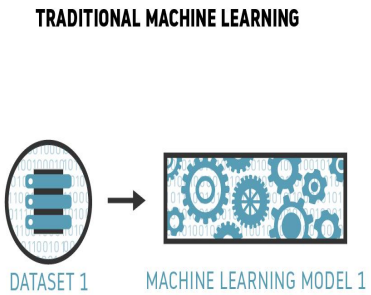
Uses more experienced machine learning models.

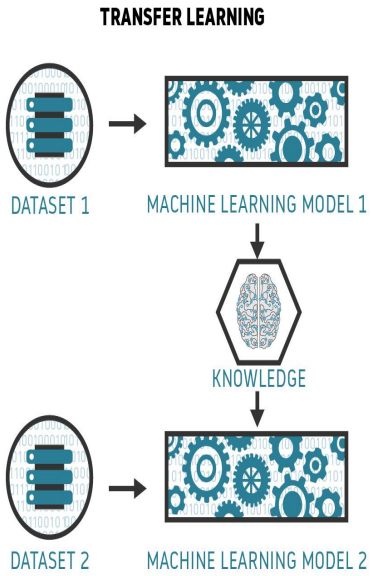
These are better than other newly built models as the have seen more of the “world”.

It focuses on storing knowledge and applying it to a different but related problem. For eg- knowledge learned while recognising cars can be used to recognise trucks etc.

The idea of using model trained on another task as a starting point for your model is transfer learning.

The model we used was VGG-16 which is a highly experienced CNN model.

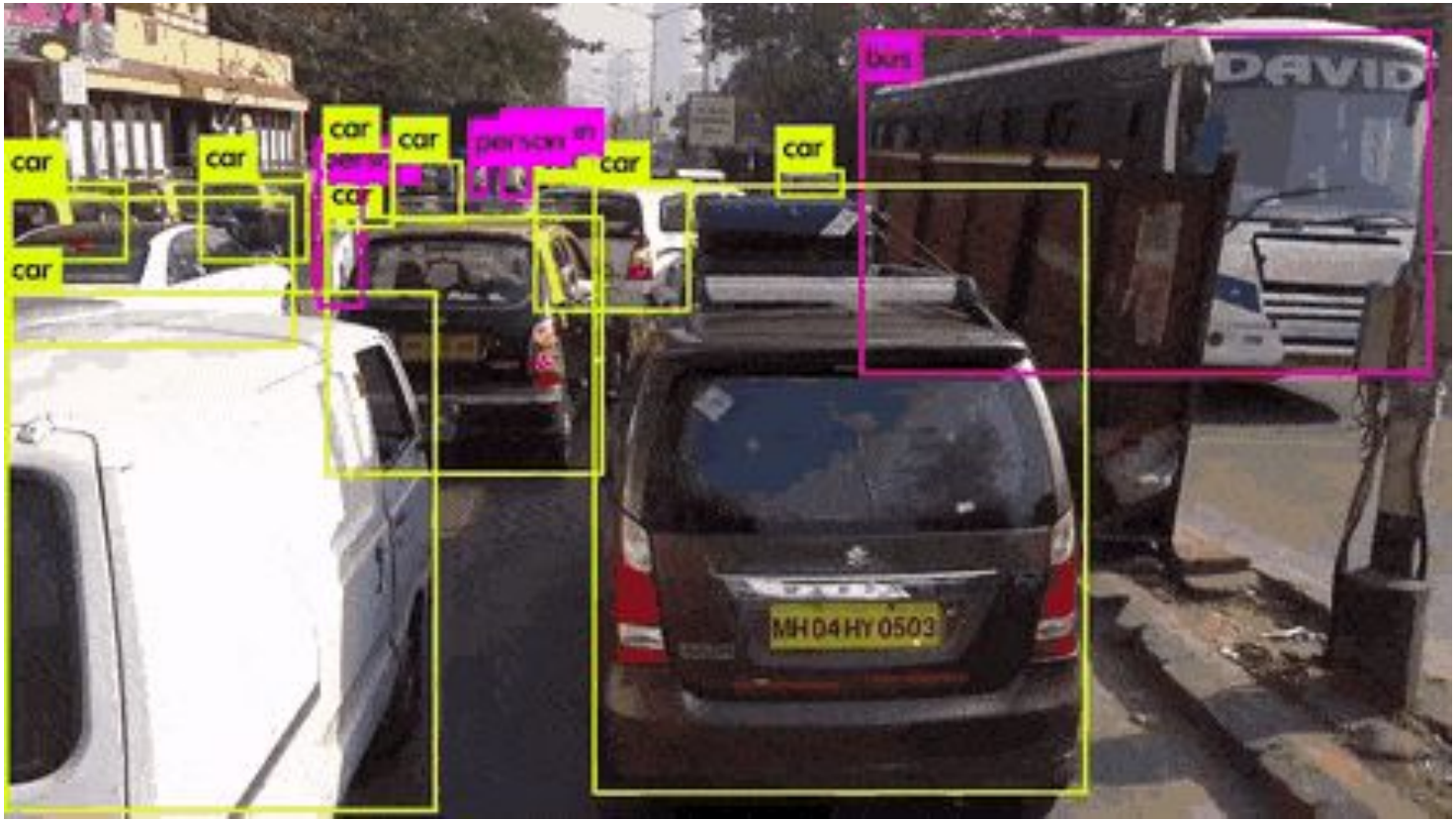




Our approach: Why we chose this method

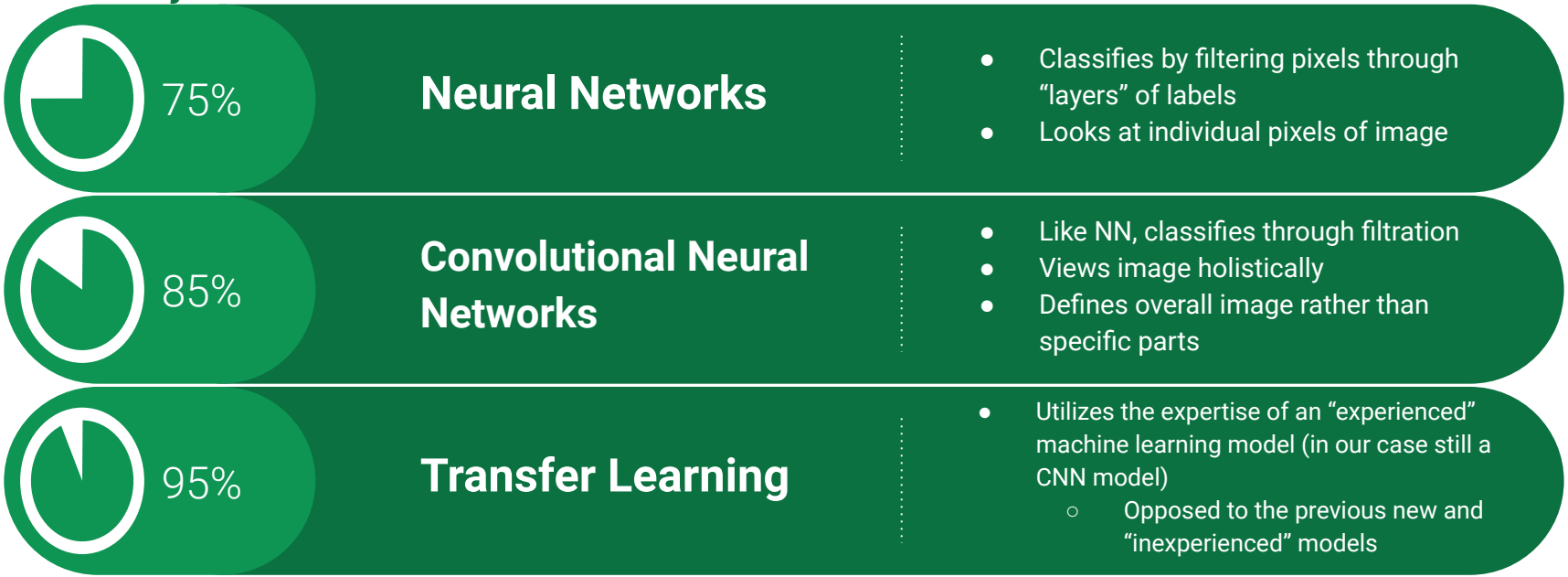
- Good at dealing with complicated data like images (data size of length * width * color channels)
- We want to get rid of the background/environment around the important images
  - Neural Networks (NN) and more specifically Convolutional Neural Networks (CNN) can classify different kinds of objects (cars, road, obstacles, humans, etc.)
- Trying new methods/algorithms to mark the objects in the image with rectangles in different colors (tag/classify the important images/sections of the greater image)
- **These reasons makes our approach efficient for object detection since unimportant parts of the image are left out and important parts are kept and then classified as their correct class/type**

Classification Boxes



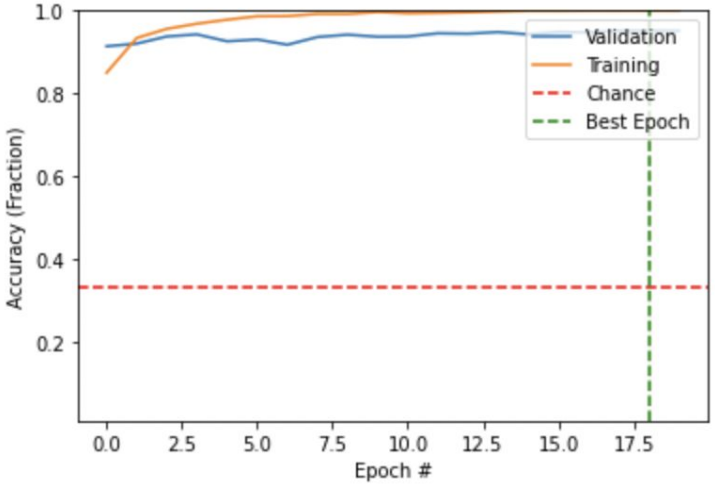
Results: Object Classification

Accuracy Score



Conclusion: Our Model

- Transfer learning was the most accurate method of object classification
  - Because it was trained on many examples before
  - Can see this because validation is close to accuracy on graph
  - Value of 1 = car, images on right show cars for predictions very close to 1
- Sliding windows was the used method of object localization
  - Simply and efficient way of breaking up the image into its component



Any questions?