OBJECT DETECTION FOR SELF-DRIVING CARS

codingcashew.com

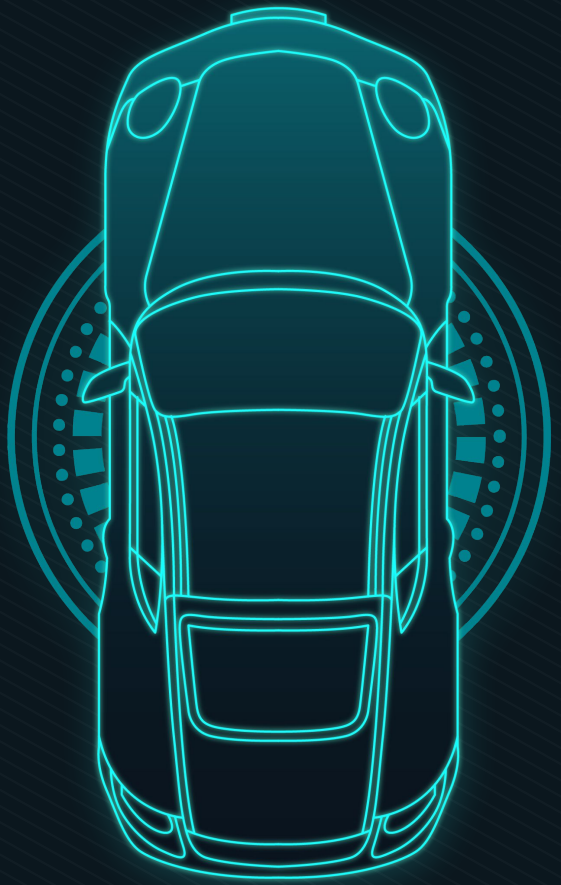
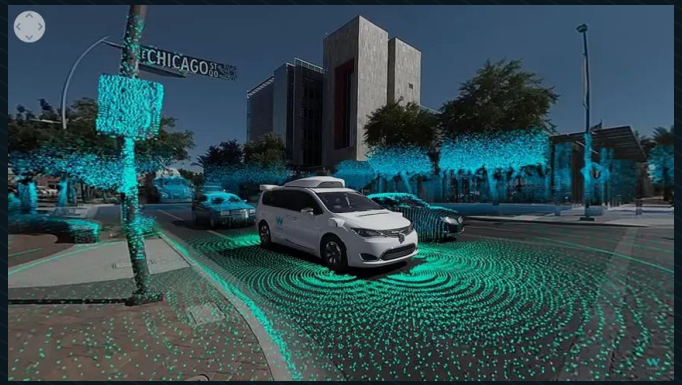


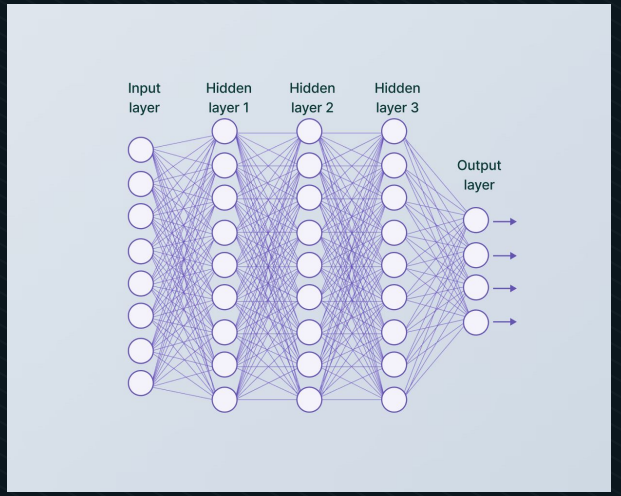
TABLE OF CONTENTS

1. Introduction



Object detection is key for self-driving cars. These systems must quickly recognize people, vehicles, and signs. We explored different detection methods to dictate which ones are the most efficient.

codingcashew.com

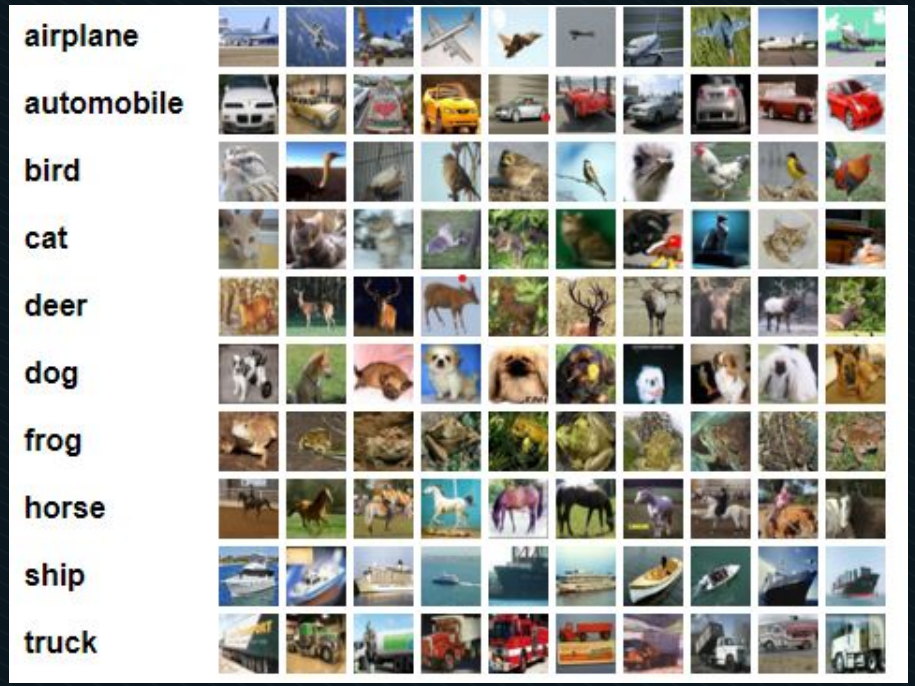


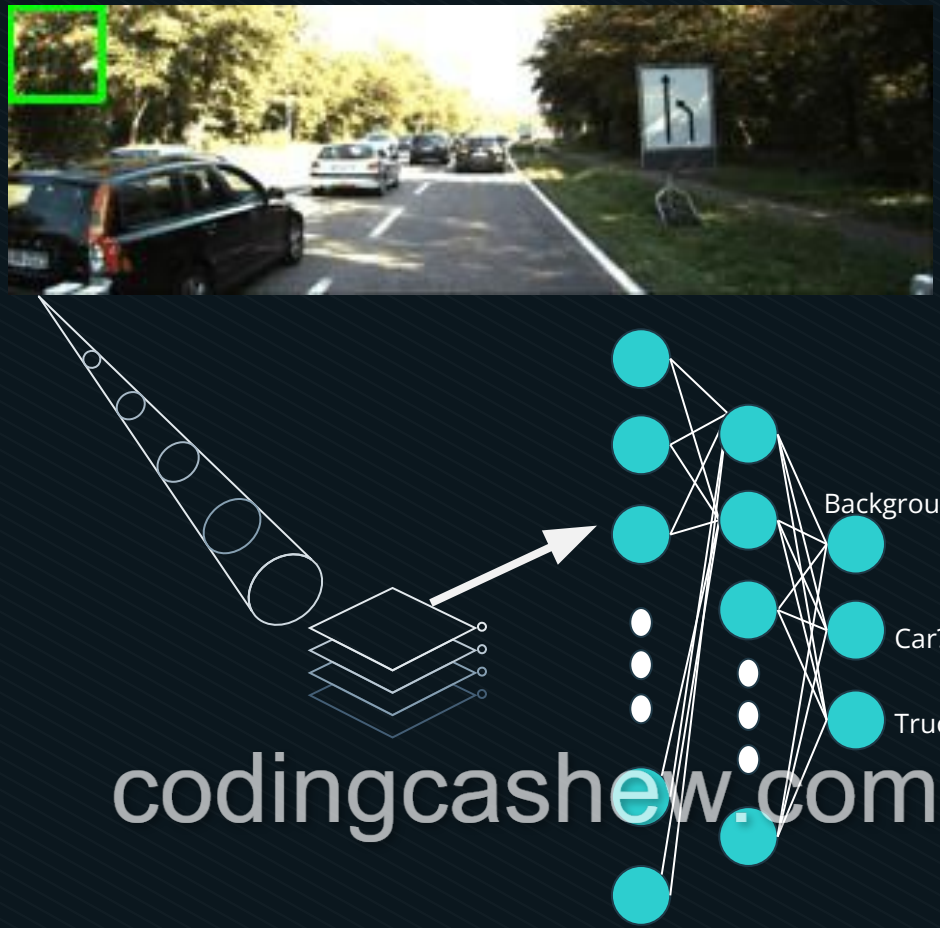
We integrated many methods such as perceptron, CNNs, and continued to build upon what we learned. In the end, we were able to create a model that was able to identify multiple objects at once.

2. Datasets Used

- CIFAR10 is the dataset we used for our image classification model. Classes include...
  - Airplanes
  - Cars
  - Birds
  - Cats
  - Deer
  - Dogs
  - Frogs
  - Horses
  - Ships
  - Trucks

- In total, this dataset consists of 60000 32x32 colored images.
  - There are 6000 per class, 5000 training and 1000 test images.
  - The test batch has 1000 randomly picked images from *each* class
  - Training batches do not always evenly represent classes





3. Sliding Windows

A fixed-size rectangular window that "slides" over the image.

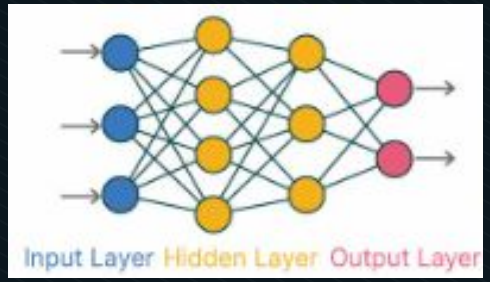
At each position, a patch of the image is extracted.

This patch is then classified.

4. MLP + CNN

What is MLP?

- Multilayer Perceptron, a basic neural network
- Flattens image into a 1D vector
- Loses spatial info (object shape/position)
- **Result:** Low accuracy



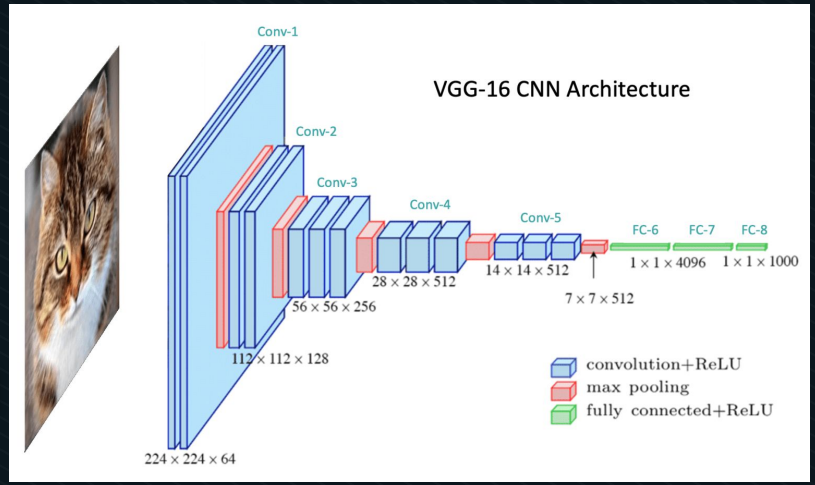
Results

71% MLP Accuracy

84% CNN Accuracy

What is CNN?

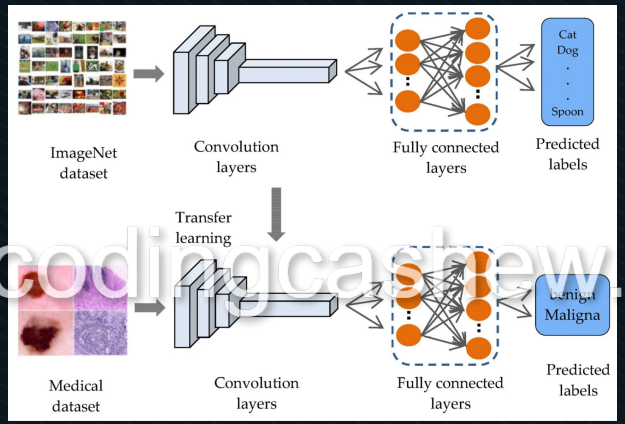
- Convolutional Neural Network
- Processes images in 2D
- Preserves spatial patterns – using filters to detect edges, curves, textures
- **Result:** Comparatively lot higher



5. Transfer Learning

What is it?

A model that is pre-trained for another task that can be used as a basis for models that we make. This allows us to spare the time of making a model from scratch.



Application

- **VGG 16**
- Very powerful and accurate!
- 14 million images–74 times
- Has many uses besides object detection

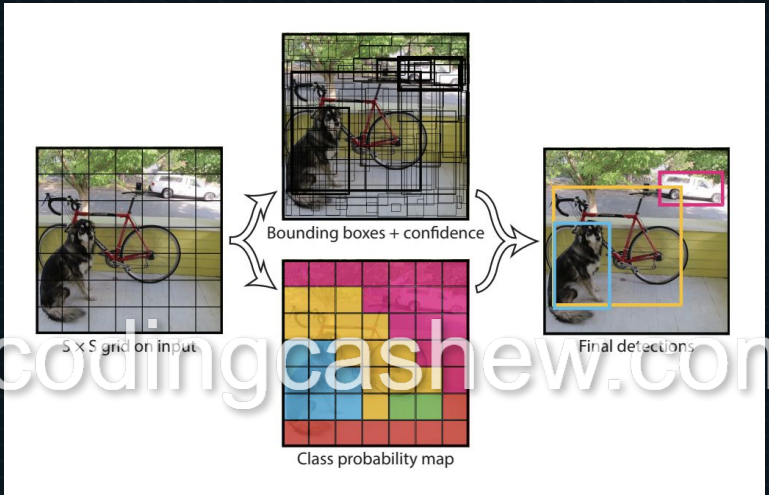
How to implement

- Choose a pre-trained model
- Modify it to closely match goals
- Train using dataset
- Evaluate its performance using test dataset or validation

6. YOLO

~~You Only Live Once~~
You Only Look Once

- Real-time object detection model
- Looks at the entire image once, splitting into grids, predicting objects presence



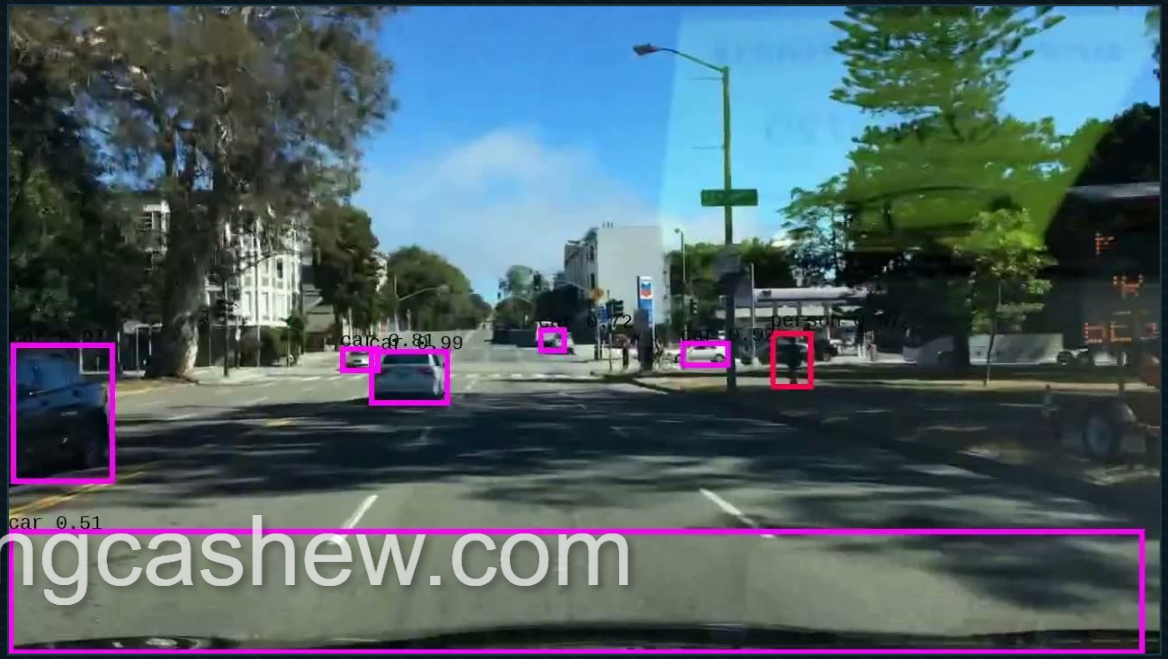
Why YOLO?

- CNNs can classify, but not detect + locate
- YOLO does this in one go, and thus it is ideal for live use: traffic cameras, surveillance, and *object detection for self-driving cars*
- Doesn't lose spatial info (object shape/position)

- We used YOLOv3 (latest version: YOLOv13)
- • Input images → bounding boxes + labels + confidence scores (accuracy)

7. YOLO IN ACTION

Made on our own using OpenCV

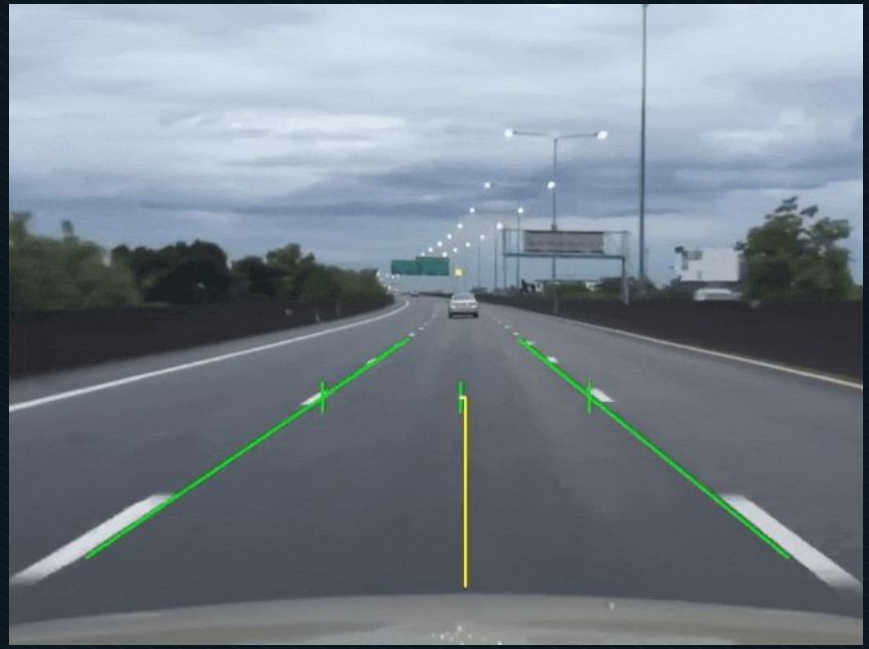


8. MODEL RESULTS

Ideas	Version	Comments	Accuracy
Image classifier with Perceptron	0	Unable to identify multiple objects in a single image.	**N/A for multi-object detection, 71% for image classification**
Sliding Windows + Perceptron	1	Detects objects one by one, but very slow and basic.	**76%**
Sliding Window + CNN	2	Detects objects better, but still basic, doesn't accommodate for large size differences in objects.	**84%**
Sliding Window + Transfer Learning (VGG16)	2.5	High accuracy with less training, but slower inference.	**95%**
YOLOv3	3	Fast, accurate, and adaptable	**28.2 - 33 on mAP**

CONCLUSION

- Live object detection with a very high accuracy is **essential** for self-driving cars
- *For example, if a system for detectings dogs and roads has a 95% accuracy, over a large amount of data, 1,000,000; it would mean 50,000 possible accidents*
- We explored MLP, CNN, Transfer Learning, and YOLO to compare speed and accuracy
- **YOLO** performed best, balancing real-time speed, object decision, mapping, and high accuracy
- This project suggested how computer vision + deep learning are transforming transportation and safety




THANKS
codingcashew.com