# Vehicle detection with CNNs - Part 2

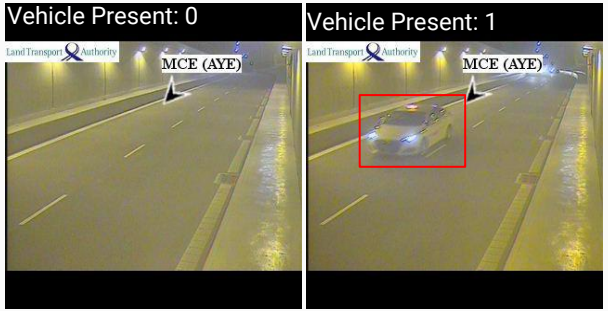
August 1, 2017

Marc Chan
marc.chan89@gmail.com

# A Recap..

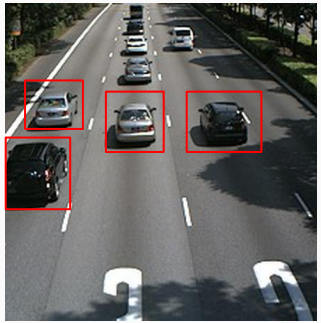
Part 1: Classification and Localization

- Trained a small CNN (LeNet-5) to do classification and localization



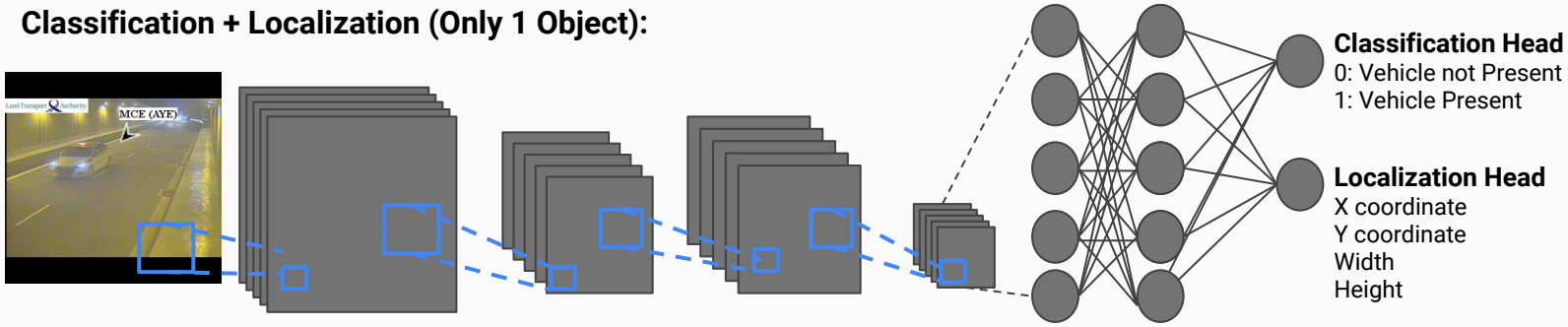

Part 2: Detection (Today!)

- Use a pre-trained CNN to detect multiple vehicles in the same image



# Challenges in detection: Sparsity

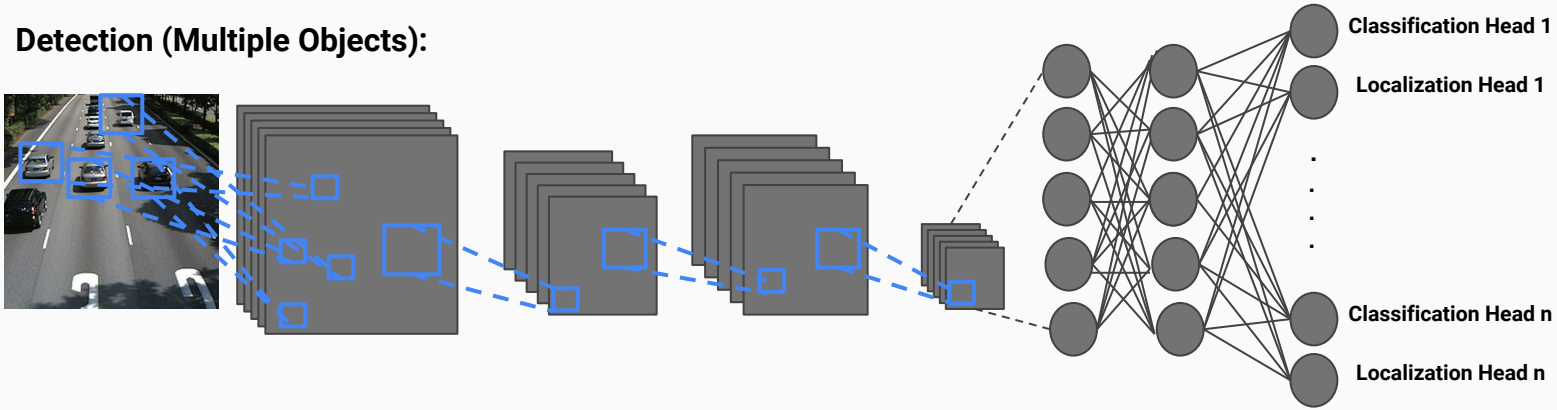
**Classification + Localization (Only 1 Object):**




FIXED:
1 Image,
1 Object,
6 Outputs

# Challenges in detection: Sparsity

**Detection (Multiple Objects):**




VARIABLE:
1 Image,
? Objects,
? Outputs?

# YOLO
## (YOU ONLY LOOK ONCE)

# YOLO Basics

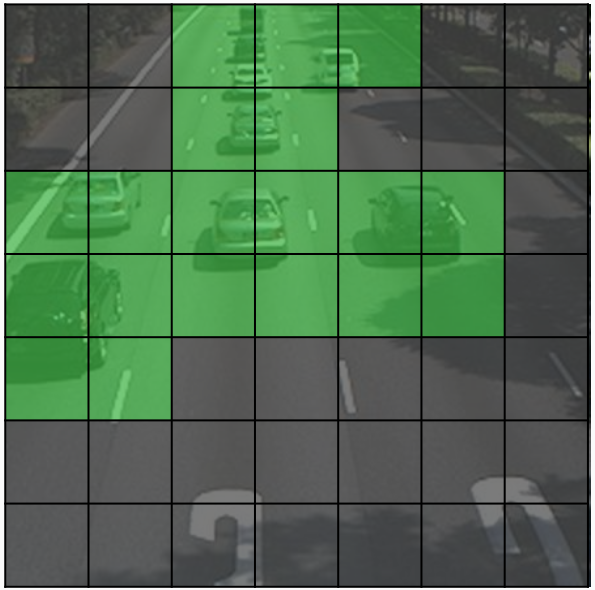
In principle, divides the image into a 7x7 grid. Each grid cell is responsible for predicting one object.

In theory, this limits the maximum number of detections to 49.

# YOLO Basics

For each grid cell, the model predicts the class of object it thinks the box contains.

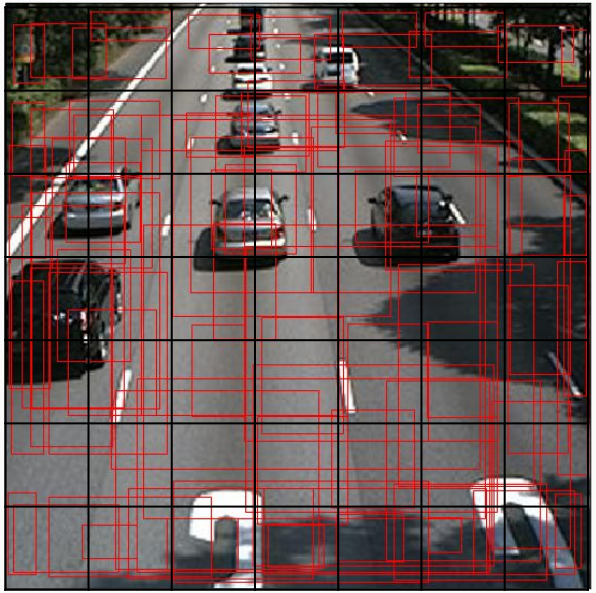
Implication is that YOLO fails to detect multiple small objects that are contained in a single grid cell.



# YOLO Basics

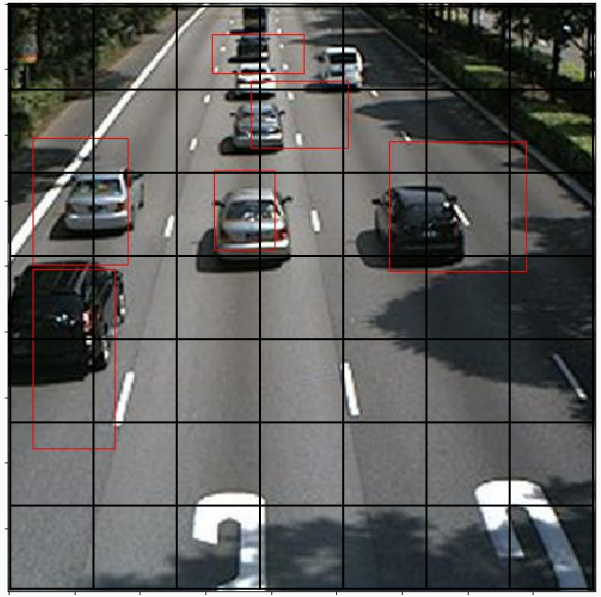
For each grid cell, the model outputs two sets of bounding box coordinates + confidence.

Predicting 2 sets of bounding boxes allows the model to better fit objects with differing aspect ratios.



# YOLO Basics

Based on the class + bounding box predictions, keep only predictions with high probability.

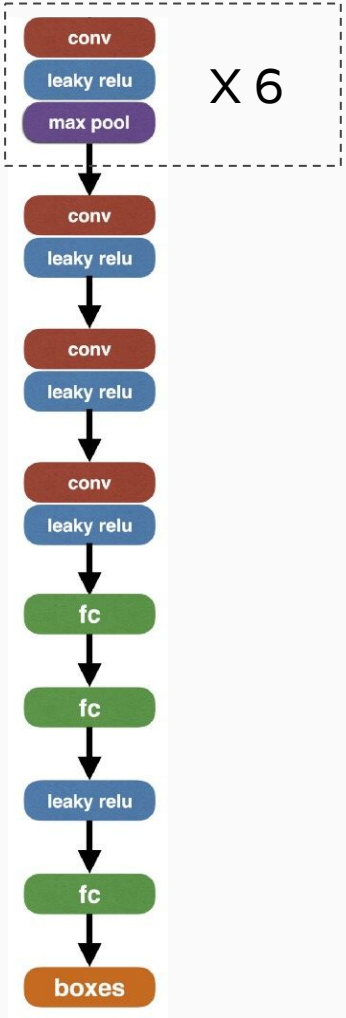


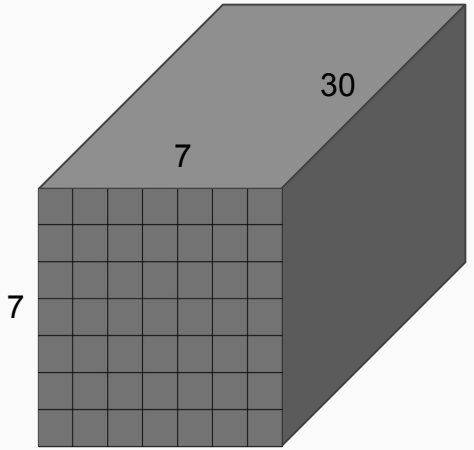
# Network Structure

We will be using tiny-YOLO v1 for this demo, since the network is relatively small and runs fast, even on CPU.

Very simple structure, stacking 3x3 convolutional, 2x2 max pooling and fully connected layers.

Output is a flattened 7 x 7 x 30 tensor. (20 predicted classes + 2x4 coordinates + 2 confidences)




Output: 7 x 7 x 30 Tensor

# iPython Demo!

# Acknowledgements


- Vehicle detection using YOLO in Keras, https://github.com/xslittlegrass/CarND-Vehicle-Detection
- J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, You Only Look Once: Unified, Real-Time Object Detection, arXiv:1506.02640 (2015).
- darkflow, https://github.com/thtrieu/darkflow
- Darknet.keras, https://github.com/sunshineatnoon/Darknet.keras/
- YAD2K, https://github.com/allanzelener/YAD2K
- Traffic Images API, https://developers.data.gov.sg/transport/traffic-images