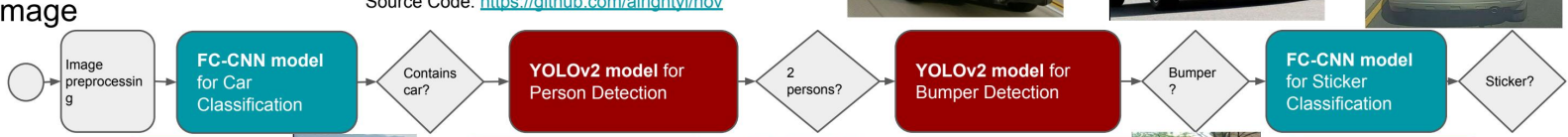


HOV Image Detection

Ivan Chit Hin Ho (ivanho1@stanford.edu) - CS230 Winter 2019

Presentation: <https://youtu.be/-PwwZRGPv0>
 Source Code: <https://github.com/alrightyi/hov>

Motivation: **Catch HOV-lane violators!!!**



FC-CNN Model
 Image size: 64x64x3
 Batch size: 32
 Epochs: 3
 Optimizer: Adam
 Loss: MSE
 Train/Val: 80/20
 Accuracy: 99.4%

(pre-trained model with dataset from ImageNet.)

Layer (Type)	Output Shape	Param #	Connected to
Input_1 (InputLayer)	(None, 608, 608, 3)	0	
lambda_1 (Lambda)	(None, 64, 64, 3)	0	
conv2d_1 (Conv2D)	(None, 64, 64, 16)	448	
batch_normalization_1 (BatchNormal)	(None, 64, 64, 16)	0	conv2d_1[0]
dropout_1 (Dropout)	(None, 64, 64, 16)	0	
conv2d_2 (Conv2D)	(None, 64, 64, 32)	896	
batch_normalization_2 (BatchNormal)	(None, 64, 64, 32)	0	conv2d_2[0]
dropout_2 (Dropout)	(None, 64, 64, 32)	0	
conv2d_3 (Conv2D)	(None, 64, 64, 64)	1792	
batch_normalization_3 (BatchNormal)	(None, 64, 64, 64)	0	conv2d_3[0]
dropout_3 (Dropout)	(None, 64, 64, 64)	0	
max_pooling2d_1 (MaxPooling2D)	(None, 8, 8, 64)	0	
conv2d_4 (Conv2D)	(None, 8, 8, 64)	0	
batch_normalization_4 (BatchNormal)	(None, 8, 8, 64)	0	conv2d_4[0]
dropout_4 (Dropout)	(None, 8, 8, 64)	0	



YOLOv2 Model
 Image size: 608x608x3
 Batch size: 32
 Epochs: 30
 Optimizer: Adam
 Loss: Classification & Coordinates Loss
 Non-Max Suppression, IOU boxes

(pre-trained model with dataset from ImageNet.)

Layer (type)	Output Shape	Param #	Connected to
Input_1 (InputLayer)	(None, 608, 608, 3)	0	
conv2d_1 (Conv2D)	(None, 608, 608, 32)	864	input_1[0]
batch_normalization_1 (BatchNormal)	(None, 608, 608, 32)	128	conv2d_1[0]
conv2d_2 (Conv2D)	(None, 608, 608, 32)	0	
batch_normalization_2 (BatchNormal)	(None, 608, 608, 32)	0	conv2d_2[0]
max_pooling2d_1 (MaxPooling2D)	(None, 304, 304, 32)	0	
conv2d_3 (Conv2D)	(None, 304, 304, 64)	18432	
batch_normalization_3 (BatchNormal)	(None, 304, 304, 64)	18432	conv2d_3[0]
conv2d_4 (Conv2D)	(None, 304, 304, 64)	18432	
batch_normalization_4 (BatchNormal)	(None, 304, 304, 64)	18432	conv2d_4[0]
conv2d_5 (Conv2D)	(None, 304, 304, 64)	18432	
batch_normalization_5 (BatchNormal)	(None, 304, 304, 64)	18432	conv2d_5[0]
conv2d_6 (Conv2D)	(None, 152, 152, 128)	36864	
batch_normalization_6 (BatchNormal)	(None, 152, 152, 128)	36864	conv2d_6[0]
conv2d_7 (Conv2D)	(None, 152, 152, 128)	36864	
batch_normalization_7 (BatchNormal)	(None, 152, 152, 128)	36864	conv2d_7[0]
conv2d_8 (Conv2D)	(None, 76, 76, 256)	73728	
batch_normalization_8 (BatchNormal)	(None, 76, 76, 256)	73728	conv2d_8[0]
conv2d_9 (Conv2D)	(None, 76, 76, 256)	73728	
batch_normalization_9 (BatchNormal)	(None, 76, 76, 256)	73728	conv2d_9[0]
conv2d_10 (Conv2D)	(None, 38, 38, 512)	147456	
batch_normalization_10 (BatchNormal)	(None, 38, 38, 512)	147456	conv2d_10[0]
conv2d_11 (Conv2D)	(None, 38, 38, 512)	147456	
batch_normalization_11 (BatchNormal)	(None, 38, 38, 512)	147456	conv2d_11[0]
conv2d_12 (Conv2D)	(None, 19, 19, 1024)	294912	
batch_normalization_12 (BatchNormal)	(None, 19, 19, 1024)	294912	conv2d_12[0]
conv2d_13 (Conv2D)	(None, 19, 19, 1024)	294912	
batch_normalization_13 (BatchNormal)	(None, 19, 19, 1024)	294912	conv2d_13[0]
conv2d_14 (Conv2D)	(None, 9, 9, 2048)	589824	
batch_normalization_14 (BatchNormal)	(None, 9, 9, 2048)	589824	conv2d_14[0]
conv2d_15 (Conv2D)	(None, 9, 9, 2048)	589824	
batch_normalization_15 (BatchNormal)	(None, 9, 9, 2048)	589824	conv2d_15[0]
conv2d_16 (Conv2D)	(None, 5, 5, 4096)	1179648	
batch_normalization_16 (BatchNormal)	(None, 5, 5, 4096)	1179648	conv2d_16[0]
conv2d_17 (Conv2D)	(None, 5, 5, 4096)	1179648	
batch_normalization_17 (BatchNormal)	(None, 5, 5, 4096)	1179648	conv2d_17[0]
conv2d_18 (Conv2D)	(None, 3, 3, 8192)	2359296	
batch_normalization_18 (BatchNormal)	(None, 3, 3, 8192)	2359296	conv2d_18[0]
conv2d_19 (Conv2D)	(None, 3, 3, 8192)	2359296	
batch_normalization_19 (BatchNormal)	(None, 3, 3, 8192)	2359296	conv2d_19[0]
conv2d_20 (Conv2D)	(None, 1, 1, 16384)	4718592	
batch_normalization_20 (BatchNormal)	(None, 1, 1, 16384)	4718592	conv2d_20[0]
conv2d_21 (Conv2D)	(None, 1, 1, 16384)	4718592	
batch_normalization_21 (BatchNormal)	(None, 1, 1, 16384)	4718592	conv2d_21[0]



YOLOv2 Model
 Image size: 416x416x3
 Batch size: 32
 Epochs: 30
 Optimizer: Adam
 Train/Val: 90/10
 Loss: Classification & Coordinates Loss
 Non-Max Suppression, IOU boxes, Early Stopping

Re-trained YOLOv2 model using YAD2K + modifications.
 Dataset from web search, hand-labeled using LabelBox (size: 680)

Layer (type)	Output Shape	Param #	Connected to
Input_1 (InputLayer)	(None, 416, 416, 3)	0	
conv2d_1 (Conv2D)	(None, 416, 416, 32)	864	input_1[0]
batch_normalization_1 (BatchNormal)	(None, 416, 416, 32)	128	conv2d_1[0]
conv2d_2 (Conv2D)	(None, 416, 416, 32)	0	
batch_normalization_2 (BatchNormal)	(None, 416, 416, 32)	0	conv2d_2[0]
conv2d_3 (Conv2D)	(None, 416, 416, 32)	0	
batch_normalization_3 (BatchNormal)	(None, 416, 416, 32)	0	conv2d_3[0]
conv2d_4 (Conv2D)	(None, 416, 416, 32)	0	
batch_normalization_4 (BatchNormal)	(None, 416, 416, 32)	0	conv2d_4[0]
conv2d_5 (Conv2D)	(None, 208, 208, 32)	0	
batch_normalization_5 (BatchNormal)	(None, 208, 208, 32)	0	conv2d_5[0]
conv2d_6 (Conv2D)	(None, 208, 208, 32)	0	
batch_normalization_6 (BatchNormal)	(None, 208, 208, 32)	0	conv2d_6[0]
conv2d_7 (Conv2D)	(None, 104, 104, 32)	0	
batch_normalization_7 (BatchNormal)	(None, 104, 104, 32)	0	conv2d_7[0]
conv2d_8 (Conv2D)	(None, 104, 104, 32)	0	
batch_normalization_8 (BatchNormal)	(None, 104, 104, 32)	0	conv2d_8[0]
conv2d_9 (Conv2D)	(None, 52, 52, 32)	0	
batch_normalization_9 (BatchNormal)	(None, 52, 52, 32)	0	conv2d_9[0]
conv2d_10 (Conv2D)	(None, 52, 52, 32)	0	
batch_normalization_10 (BatchNormal)	(None, 52, 52, 32)	0	conv2d_10[0]
conv2d_11 (Conv2D)	(None, 26, 26, 32)	0	
batch_normalization_11 (BatchNormal)	(None, 26, 26, 32)	0	conv2d_11[0]
conv2d_12 (Conv2D)	(None, 26, 26, 32)	0	
batch_normalization_12 (BatchNormal)	(None, 26, 26, 32)	0	conv2d_12[0]
conv2d_13 (Conv2D)	(None, 13, 13, 32)	0	
batch_normalization_13 (BatchNormal)	(None, 13, 13, 32)	0	conv2d_13[0]
conv2d_14 (Conv2D)	(None, 13, 13, 32)	0	
batch_normalization_14 (BatchNormal)	(None, 13, 13, 32)	0	conv2d_14[0]
conv2d_15 (Conv2D)	(None, 13, 13, 30)	30750	



FC-CNN Model
 Image size: 64x64x3
 Batch size: 32
 Epochs: 3
 Optimizer: Adam
 Loss: MSE
 Train/Val: 80/20
 Accuracy: 67%
 Dataset size: 1663 (1077 pos, 586 neg)
 Data Augment: flip, shift, rotate, shear, zoom, brightness

(pre-trained model with dataset from ImageNet.)

Layer (Type)	Output Shape	Param #	Connected to
Input_1 (InputLayer)	(None, 64, 64, 3)	0	
lambda_1 (Lambda)	(None, 64, 64, 3)	0	
conv2d_1 (Conv2D)	(None, 64, 64, 16)	448	
batch_normalization_1 (BatchNormal)	(None, 64, 64, 16)	0	conv2d_1[0]
dropout_1 (Dropout)	(None, 64, 64, 16)	0	
conv2d_2 (Conv2D)	(None, 64, 64, 32)	896	
batch_normalization_2 (BatchNormal)	(None, 64, 64, 32)	0	conv2d_2[0]
dropout_2 (Dropout)	(None, 64, 64, 32)	0	
conv2d_3 (Conv2D)	(None, 64, 64, 64)	1792	
batch_normalization_3 (BatchNormal)	(None, 64, 64, 64)	1792	conv2d_3[0]
dropout_3 (Dropout)	(None, 64, 64, 64)	0	
max_pooling2d_1 (MaxPooling2D)	(None, 8, 8, 64)	0	
conv2d_4 (Conv2D)	(None, 8, 8, 64)	0	
batch_normalization_4 (BatchNormal)	(None, 8, 8, 64)	0	conv2d_4[0]
dropout_4 (Dropout)	(None, 8, 8, 64)	0	



pos images: 56, accuracy: 0.39285714285714285
 neg images: 50, accuracy: 0.98
 total images: 106, accuracy: 0.6698113207547169