HOV Image Detection

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

Presentation: https://youtu.be/-PwwZRGBpV0
Source Code: https://github.com/alrightyi/hov

Motivation: Catch HOV-lane violators!!!

Image

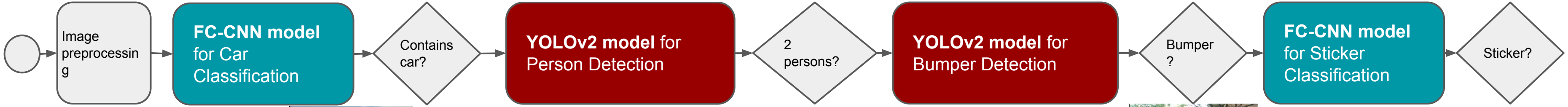
Image preprocessing

FC-CNN model for Car Classification

Contains car?

YOLOv2 model for Person Detection

2 persons?

YOLOv2 model for Bumper Detection

Bumper?

FC-CNN model for Sticker Classification

Sticker?

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 #
=================================================================
lambda_1 (Lambda)            (None, 64, 64, 3)         0
_________________________________________________________________
cv0 (Conv2D)                 (None, 64, 64, 16)
448
_________________________________________________________________
dropout_1 (Dropout)          (None, 64, 64, 16)        0
_________________________________________________________________
cv1 (Conv2D)                 (None, 64, 64, 32)
4640
_________________________________________________________________
dropout_2 (Dropout)          (None, 64, 64, 32)        0
_________________________________________________________________
cv2 (Conv2D)                 (None, 64, 64, 64)
18496
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 8, 8, 64)          0
_________________________________________________________________
dropout_3 (Dropout)          (None, 8, 8, 64)          0
_________________________________________________________________
fcn (Conv2D)                 (None, 1, 1, 1)
4097
=================================================================
Total params: 27,681
Trainable params: 27,681
Non-trainable params: 0
```

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

YOLOv2 Model

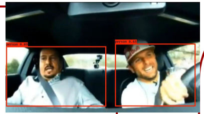

Image size: 608x608x3
Batch size: 32
Epochs: 30
Optimizer: Adam
Loss: Classification & Coordinates Loss
Non-Max Suppression, IOU boxes

(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][0]
__________________________________________________________________________________________________
batch_normalization_1 (BatchNor (None, 608, 608, 32) 128         conv2d_1[0][0]
__________________________________________________________________________________________________
leaky_re_lu_1 (LeakyReLU)       (None, 608, 608, 32) 0
batch_normalization_1[0][0]
__________________________________________________________________________________________________
max_pooling2d_1 (MaxPooling2D)  (None, 304, 304, 32) 0
leaky_re_lu_1[0][0]
__________________________________________________________________________________________________
conv2d_2 (Conv2D)               (None, 304, 304, 64) 18432
max_pooling2d_1[0][0]
__________________________________________________________________________________________________
…..
…..
…..
leaky_re_lu_21 (LeakyReLU)      (None, 38, 38, 64)   0
batch_normalization_21[0][0]
__________________________________________________________________________________________________
batch_normalization_20 (BatchNo (None, 19, 19, 1024) 4096        conv2d_20[0][0]
__________________________________________________________________________________________________
space_to_depth_x2 (Lambda)      (None, 19, 19, 256)  0
leaky_re_lu_21[0][0]
__________________________________________________________________________________________________
leaky_re_lu_20 (LeakyReLU)      (None, 19, 19, 1024) 0
batch_normalization_20[0][0]
__________________________________________________________________________________________________
concatenate_1 (Concatenate)     (None, 19, 19, 1280) 0
space_to_depth_x2[0][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 from 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][0]
__________________________________________________________________________________________________
batch_normalization_1 (BatchNor (None, 416, 416, 32) 128         conv2d_1[0][0]
__________________________________________________________________________________________________
leaky_re_lu_1 (LeakyReLU)       (None, 416, 416, 32) 0
batch_normalization_1[0][0]
__________________________________________________________________________________________________
max_pooling2d_2 (MaxPooling2D)  (None, 208, 208, 32) 0
leaky_re_lu_1[0][0]
__________________________________________________________________________________________________
…….
…..
…..
__________________________________________________________________________________________________
conv2d_22 (Conv2D)              (None, 13, 13, 1024) 11796480
concatenate_1[0][0]
__________________________________________________________________________________________________
batch_normalization_22 (BatchNo (None, 13, 13, 1024) 4096        conv2d_22[0][0]
__________________________________________________________________________________________________
leaky_re_lu_22 (LeakyReLU)      (None, 13, 13, 1024) 0
batch_normalization_22[0][0]
__________________________________________________________________________________________________
conv2d_24 (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 neg, 586 pos.)
Data Augment: flip, shift, rotate, shear, zoom, brightness

(pre-trained model with dataset from ImageNet.)

```
_________________________________________________________________
Layer (type)                 Output Shape
Param #
=================================================================
lambda_1 (Lambda)            (None, 64, 64, 3)         0
_________________________________________________________________
cv0 (Conv2D)                 (None, 64, 64, 16)
448
_________________________________________________________________
dropout_1 (Dropout)          (None, 64, 64, 16)        0
_________________________________________________________________
cv1 (Conv2D)                 (None, 64, 64, 32)
4640
_________________________________________________________________
dropout_2 (Dropout)          (None, 64, 64, 32)        0
_________________________________________________________________
cv2 (Conv2D)                 (None, 64, 64, 64)
18496
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 8, 8, 64)          0
_________________________________________________________________
dropout_3 (Dropout)          (None, 8, 8, 64)          0
_________________________________________________________________
```