

DATA.ML.300 Computer Vision

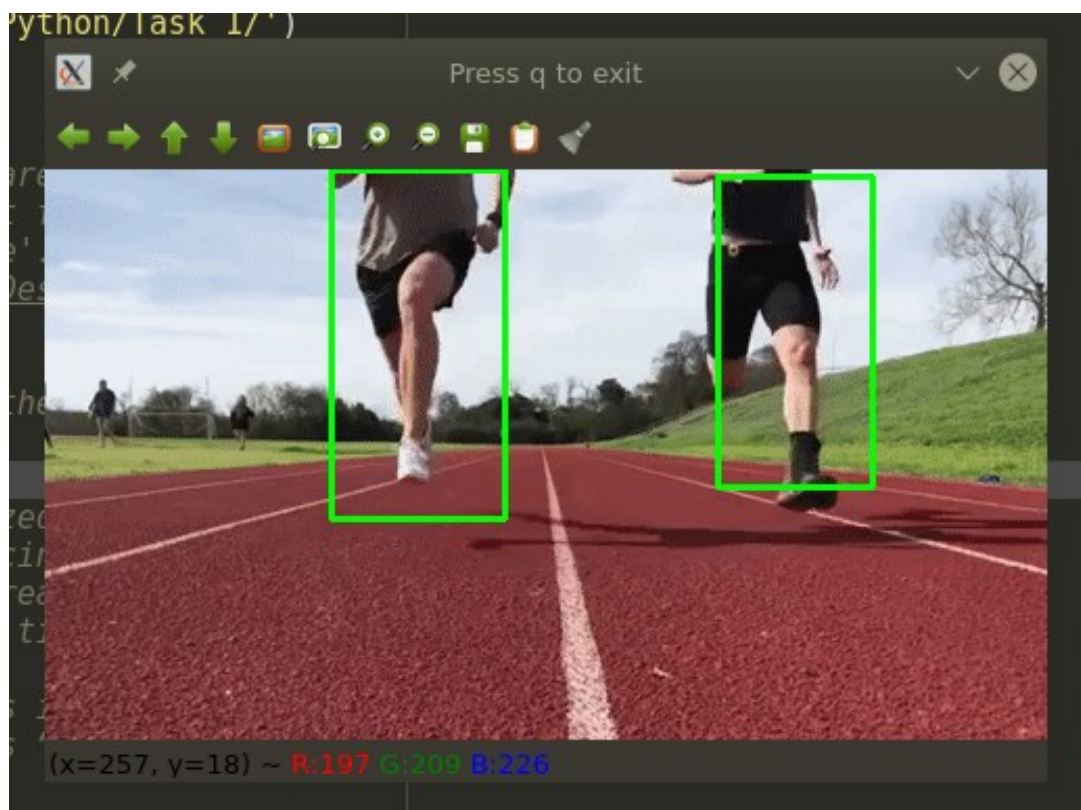
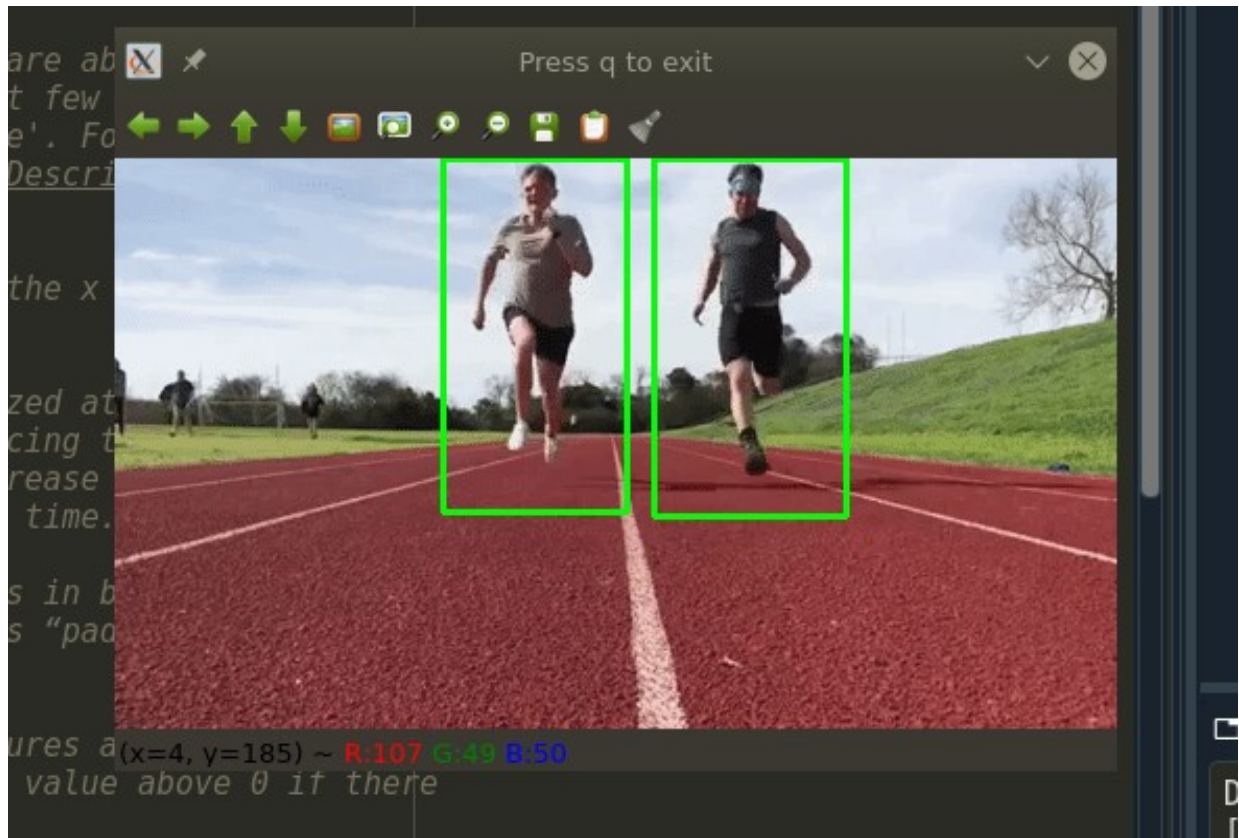
Exercise Round 3

Answered to all questions (in 1-3).

This file contains also all asked screenshots.

1)

Example screenshot



2)

1.

The targets values of the dataset are presented via the following attributes: *xmin*, *xmax*, *ymin*, *ymax* and *class_id*. The first four values present the limits of bounding box of the targets in the image, where (*xmin*,*ymin*)=top-left and (*xmax*,*ymax*)=top-right corners of the bounding box. The last variable is the id-number for the type of the target (eg. vehicle type).

In general sense, validation data holds previously unseen data (in the case of Udacity, bounding boxes for the targets that weren't marked on the training dataset) so that the final performance of the model can be fairly evaluated without the possibility to optimize the model specific to some dataset only.

2.

The first convolutional part starts at line 277 and ends at 309, where the whole structure consist of 7 convolutional blocks. Each convolutional block has specific number of filters (32/48/64) followed by batch normalization and finally an activation function which in this case is ELU (Exponential Linear Unit). After each block (except the last one), the output dimensionality is reduced by max pooling. Before feeding the data in network, there is also an implementation for lambda layer which performs a specific preprocessing task, such as normalization by mean or standard deviation.

3.

The overall loss function is a weighted sum of the two separate loss metrics: localization loss, which measures the mismatch between the predicted boundary box and the ground truth. This loss is defined by smooth L1-loss between the parameters of these two bounding boxes. Parameters of the bounding box are the ones presented in the first part of this example. Second loss metric is confidence loss, which measures the actual prediction error/mismatch regarding to object. Confidence loss is defined as a softmax (categorical crossentropy) loss over multiple classes confidences. For correct predictions and the prediction confidence, the loss is calculated according to the corresponding class confidence score (eg. 0.88 vs 1.0). In the case of missclassified object, the loss is calculated regarding to 'no object'-category's confidence score.

3)

Screenshots

