

**PhD**

**Online Video Object Detection using Association LSTM iccv17**

2018/09/30 1:41 PM

- First run the SSD detector
- concatenate its bounding box predictions along with the corresponding features after passing the latter through an ROI pooling layer to get  $7 \times 7$  features;
- stacked together these concatenated features for several consecutive frames and pass these the stacked features two an LSTM;
- two different types of losses are used to lawn to perform regression and association simultaneously.
- Regression loss is just a slightly modified version of the SSD loss where a term is added to ensure smooth progression between frames
- association losses performed pairwise trained two consecutive frames though its details are not all that clear