PhD Literature

Siam R-CNN Visual Tracking by Re-Detection 1911.12836 iccv19

2020-04-29 10:23:06 PM

Adapts faster R CNN using Siamese network to decide if RPN regions correspond to the target object and perform some corresponding regression

Some sort of New negative hard mining to learn distractor objects from very few samples

Checking table object candidates from previous frames and some form of tracklet formation to decide on the correct one

4.7 FPS for the standard version and 15 FPS for the speed optimized one so not really useful for practical purposes

2020-05-01 10:18:46 PM

main novelty seems to be that it uses two-stage faster R CNN instead of single a stage SSD type detectors that have apparently already been tried

ROI align is used to extract deep features from the template box as well as each one of the RPN boxes which are then concatenated after the usual 1x1 convolution thingy

the combined features and then passed to a 2–class re-detection head which is the only one that is trained while the backbone and the RPN weights are frozen

2020-05-02 1:13:45 PM

Hard example mining based on similar looking objects from other videos is used where and embedding method from a video object segmentation method is used to generate embedding vectors for a large set of objects followed by indexing and fast nearest neighbour search to find about 10,000 nearest objects for each training one

Tracklet generation and scoring process seems to be full of the usual heuristics based on the usual spatiotemporal cues