PhD

MOTS Multi-Object Tracking and Segmentation cvpr19

2019-12-14 5:08:48 PM

Code is available

new data sets for pixel level mask based multi object tracking by adding masks created by semiautomated processes to the MOT and Kitti data sets
Deeplabv3+ is used for annotating and is trained in an iterative manner and on a per object basis starting with 2 manually created segmentation masks for object

Baseline tracker called track–RCNN created by adding either a **3D convolutional** or a **convolutional LSTM layer** between the backbone and RPN of a mask RCNN to incorporate temporal context

Also adds an association head (analogous to an in parallel with the existing mask head) that outputs association vector for each predicted box which supposedly encodes its identity across frames

Apparently inspired by embedding vectors from human re-identification literature, these vectors are trained in such a way that vectors belonging to the same object are close to each other by some measure while vectors belonging to different objects are far apart – the distance measure itself is the simple Euclidean distance

Experiments are also done with optical flow based warping to carry out mask propagation as an alternative to the Association head

The actual association is done using the standard Hungarian algorithm and seems to be riddled with a mess of heuristical thresholds as usual

there does seem to be some sort of end to end Detection – tracking training where the detector is trained by the MOT ground truth but not clear exactly how effective it is

experiments seemed to indicate that the baseline tracker does seem to benefit from the extra information coming from the detailed segmentation masks when compared to existing MOT combined with mask RCNN segmentations