PhD Literature

D3S – A Discriminative Single Shot Segmentation Tracker 1911.08862v1

2020-05-10 2:38:35 PM

Main difference from Siamese mask seems to be that this is single-stage and does end to end joint learning for both segmentation and localization

Two discriminative models are used – one allows broad geometric variations to generate precise masks while the other one only outputs a single location but is far more geometrically constrained so helps to disambiguate between multiple instance outputs by the former

Some sort of technique to infer a rotated rectangle from a segmentation mask is also proposed

2020-05-11 1:57:06 PM

The two models are respectively called geometrically invariant or GIM and geometrically constrained Euclidean or GEM And both take common features from the backbone network

GIM is used to produce 2 similarity maps – one for foreground and one for background – using some kind of top K averaging operator
spatially precise but do not distinguish between different instances of similar looking objects

GEM uses atom tracker and produces a single instance localization map

2020-05-13 12:45:00 PM

There is also a refinement or upscaling pathway to improve the mask resolution as common in segmentation

segmentation mask to rotated bounding box converter and using the segmentation network for tracking are mostly heuristics based

Only GEM is trained if only bounding box is available in the first frame while GIM is also trained if segmentation mask is available so there is proxy GT mask generation in the latter case using the network itself though its usefulness is dubious