2019-07-30 7:29:40 PM

Adapts the IOU net idea from detection to tracking by training a target specific models instead of class specific ones

Tracker is made up of 2 components – Discriminator or classifier that produces confidence map which seems quite similar to the Siamese thingy and State estimator that is based on IOU net and serves as a refinement step

estimator is trained off-line and the classifier is trained online

Object location from the last frame is used as input to the classifier and its output location is augmented by random sampling to get 10 locations that become input to the estimator

output of the IOU net is the overlap between the bounding box and an object in a given image where the image and the box are inputs
this is used for state estimation using gradient ascent type optimization on this network by treating it as a differentiable function that is to be maximized with respect to this IOU

Precise ROI pooling is used to allow training the IOU net by making the patch under a bounding box in of feature map differentiable with respect to the bounding box coordinates

Classification network is very shallow with only 2 convolutional layers that are trained using some sort of Gauss Newton/conjugate gradient approach