GlobalTrack (AAAI 2020)

GlobalTrack: A Simple and Strong Baseline for Long-term Tracking

Huang, et al
Lianghua

① Problem Description: Long-term Tracking:

- ① target absences
- ② temporary tracking failures

Most existing trackers work under a strong temporal consistency assumption

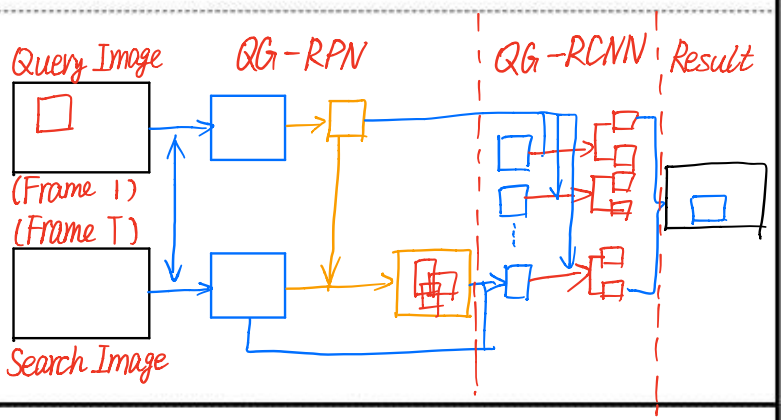
② Problem Solution:

using pure global instance search

- remove the locality assumption
- enable the tracker to search for the target at arbitrary positions and scales

thereby avoiding cumulative errors during tracking.

③ Conceptual Understanding



① QG-RPN: $\hat{x} = g_{qg\text{-}rpn}(z, x) = f_{out}(f_z(z) \otimes f_x(x))$

$L_{qg\_rpn}(z, x) = L_{rpn}(\hat{x}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{loc}} \sum_i p_i^* L_{loc}(s_i, s_i^*)$

② QG-RCNN: $\hat{x}_i = g_{qg\_rcnn}(z, x_i) = h_{out}(h_z(z) \odot h_x(x_i))$

$L_{qg\_rcnn}(z, x) = \frac{1}{N_{prop}} \sum_i [L_{cls}(p_i, p_i^*) + \lambda p_i^* L_{loc}(s_i, s_i^*)]$

Details of implementation

Implementation :

① Offline Training

sample frame pairs from training videos

- find the M instances
- construct M query-search image pairs

② Online Tracking

Initialization : specify query

Tracking : input { query, current image }

Result : top-1 prediction

③ Cross-query Loss

$$L_{cql} = \frac{1}{M}\sum_{k=1}^{M} L(\mathcal{Z}_k, \mathcal{X})$$

$$L(\mathcal{Z}_k, \mathcal{X}) = L_{cg\_rpn}(\mathcal{Z}_k, \mathcal{X}) + L_{cg\_rcnn}(\mathcal{Z}_k, \mathcal{X})$$

Architecture :

① Query-Guide RPN

generating query-specific proposals

use correlation to encode the query information in backbone features.

② Query-Guide RCNN

discriminating the proposals and producing the final predictions

- Feature modulation
- Traditional RCNN

③ Tracking Result

top-1 prediction as results.

Improvement :

① single object to multi-object.

② improve accuracy with post-processing.