GOALS AND CONTRIBUTIONS

Context: Visually track and maintain identities of multiple objects in a videostream, as long as they remain in the frame. It relies on the multi-object tracking-by-detection strategy: (i) detect objects frame-by-frame, (ii) link detections over time and maintain consistent trajectories.

Constraints: Aim at further embedded applications (= CPU cost constraint).

Scientific contributions:
- all-in-one compact siamese CNN architecture to address both single object tracking and frame-by-frame reidentification within MOT context
- tracklet management strategy based on data association between current targets’ tracklet pool and input detections without any extra pre-processing cost
- comparative quantitative and qualitative state-of-the-art evaluation on MOT17 challenge with CPU cost reduction

METHODOLOGY

SINGLE OBJECT TRACKING AND REIDENTIFICATION FOR MOT
- Two branch siamese CNN dealing with position regression and reidentification at the same time
- Light-weight EfficientNet-B0 based feature extractor

MULTI-TASK LEARNING AND ALL-IN-ONE TRAINING PAIRS GENERATION

TRACKLET SET MANAGEMENT STRATEGY
- No thresholding done on the input detections
- Multicue score taken into account for data association (detector & tracker confidence, targets’ positions and appearance) and tracklet confidence update
- “Keep active” strategy based on tracklet confidence to compensate for missing detections
- Straight forward with one-stage management of active/inactive tracks

QUANTITATIVE RESULTS

Ablation study showing the impact of the tracklets multi-cue management strategy on the MOT17 SDP training set.

EXPERIMENTS & RESULTS

Comparison to the literature on the MOT17 test set with SDP detector inputs. We achieve state-of-the-art. Demo at https://bit.ly/2VNl9ks

CONCLUSION

- Online tracking-by-detection in a multi-object tracking context with an original architecture integrating a jointly trained all-in-one single light-weight siamese CNN for both tracklet position prediction and reidentification
- Comparison to the literature on the MOT17 benchmark, presenting state-of-the-art performance and high frame rates
- Work in progress / Perspectives: benefit from recent single object tracking techniques using intercorrelation to improve speed and performance; find an in-between strategy within appearance feature management.

ACKNOWLEDGEMENTS

This work is partially supported by the French National Association for Research and Technology (ANRT) within a CIFRE PhD agreement.