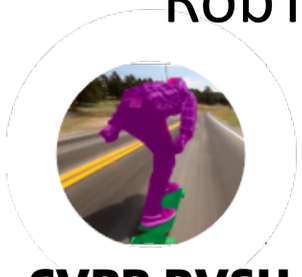
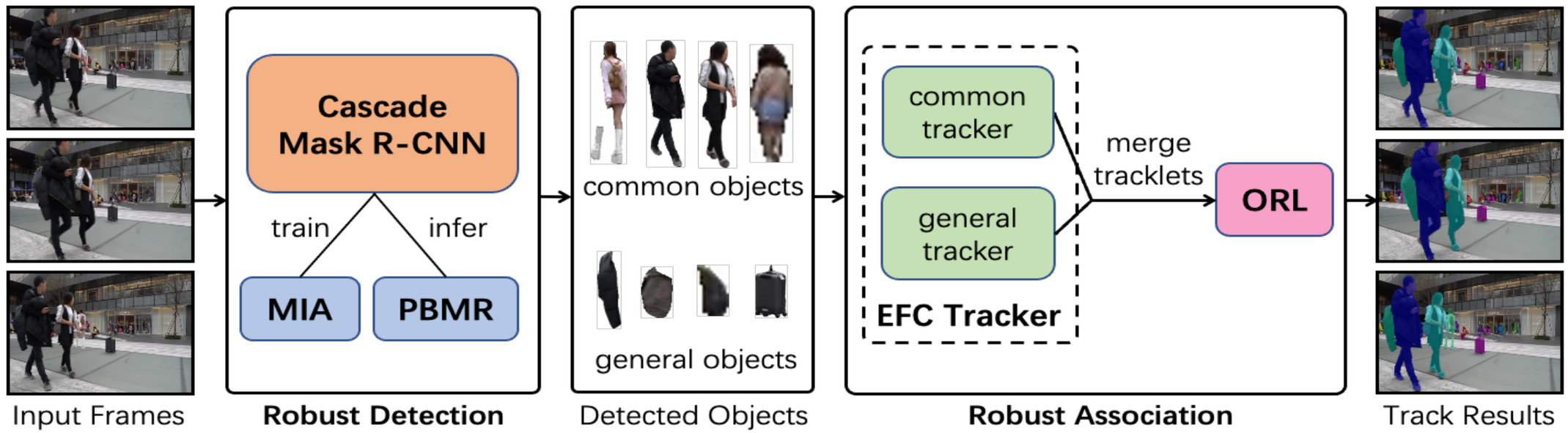
RobTrack: A Robust Tracker Baseline towards Real-World Robustness in Multi-Object Tracking and Segmentation



CVPR RVSU Workshop 2021

Dongxu Wei[1,2], Jiashen Hua[2], Hualiang Wang[1,2], Baisheng Lai[2], Kejie Huang[1], Chang Zhou[2], Jianqiang Huang[2], and Xiansheng Hua[2]

[1]Zhejiang University [2]Alibaba Group



Abstract: Multi-Object Tracking and Segmentation (MOTS) aims to detect objects' masks and track their trajectories throughout videos, which has a wide range of applications in autonomous driving, safety monitoring, video analytics, etc. Most existing MOTS methods focus on tuning and competing over one specific dataset at a time, which can only reflect their performance on a small subset of real-world scenarios. However, real-world tracking requires trackers to work robustly and consistently on any dataset without dataset-specific tuning. To achieve this, we propose RobTrack, a robust tracker baseline optimized from two aspects w.r.t. robust detection and association. For robust detection, a strong Cascade Mask R-CNN detector is utilized to first generate coarse object masks, incorporated with a Massive Instance Augmentation (MIA) method during training to improve its generalization ability in various scenarios. Then, a Patch-Based Mask Refinement (PBMR) module is applied to these coarse object masks for higher mask quality. For robust association, we propose an Ensemble Full-Category (EFC) tracker consisting of two expert trackers w.r.t. common (e.g., person, vehicle) and general (e.g., non-human and non-vehicle objects) categories, enabling optimal association for both category domains. To further improve the association accuracy, a well-designed Offline Re-Linking (ORL) strategy is applied to each video sequence to obtain the final tracking results. Extensive experiments are conducted on a very large-scale RobMOTS dataset, which consists of eight smaller MOTS datasets. The results show that our proposed method performs well in both detection and association for all of the eight datasets, and achieves the first place in the RobMOTS 2021 challenge, demonstrating its significance in robust real-world tracking.