

Paper Review:

Online Multi-Object Tracking with Dual Matching Attention Networks



AutoML 연구실
석사과정 문아성
20.12.10

Multiple Object Tracking

MOT 17

Number of videos: 21 (train) + 21(test)

Number of annotations: 564,228

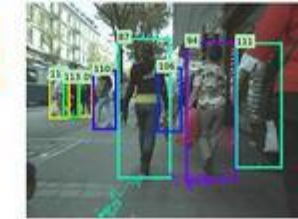
Year: 2017



MOT17-01



MOT17-03



MOT17-05



MOT17-07



MOT17-02



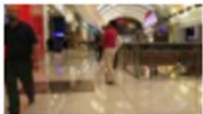
MOT17-04



MOT17-06



MOT17-08

Sample	Name	FPS	Resolution	Length	Tracks	Boxes	Density	Description
	MOT17-13-SDP	25	1920x1080	750 (00:30)	110	11642	15.5	Filmed from a bus on a busy intersection
	MOT17-11-SDP	30	1920x1080	900 (00:30)	75	9436	10.5	Forward moving camera in a busy shopping mall
	MOT17-10-SDP	30	1920x1080	654 (00:22)	57	12839	19.6	A pedestrian scene filmed at night by a moving camera
	MOT17-09-SDP	30	1920x1080	525 (00:18)	26	5325	10.1	A pedestrian street scene filmed from a low

Paper Review

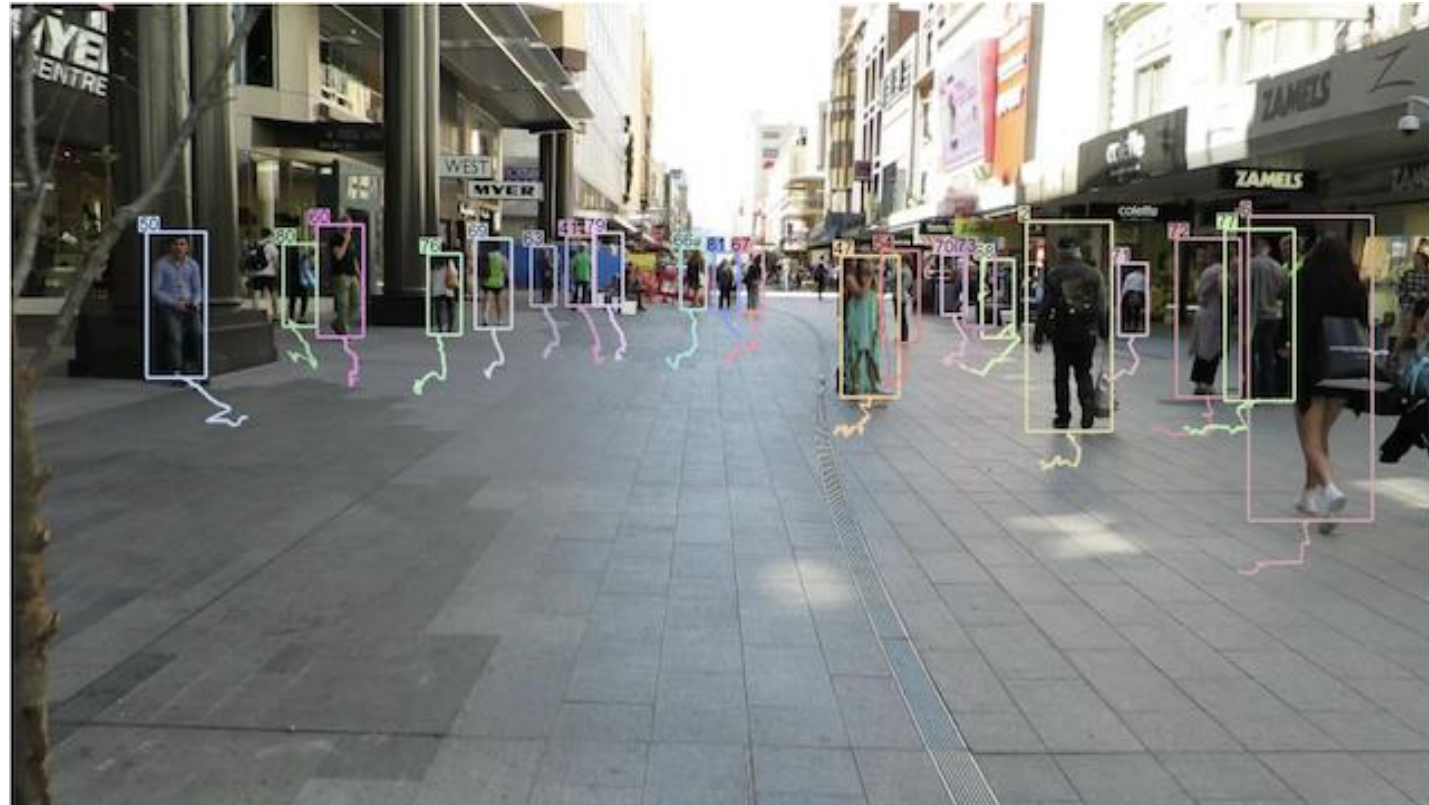
Online Multi-Object Tracking with Dual Matching Attention Networks

Ji Zhu, Hua Yang, Nian Liu, Minyoung Kim, Wenjun Zhang, Ming-Hsuan Yang

Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 366-382

Dual Matching Attention Networks

- Spatial Attention Network
- Temporal Attention Network



Online Multi-Object Tracking with Dual Matching Attention Networks

Dual Matching Attention Networks

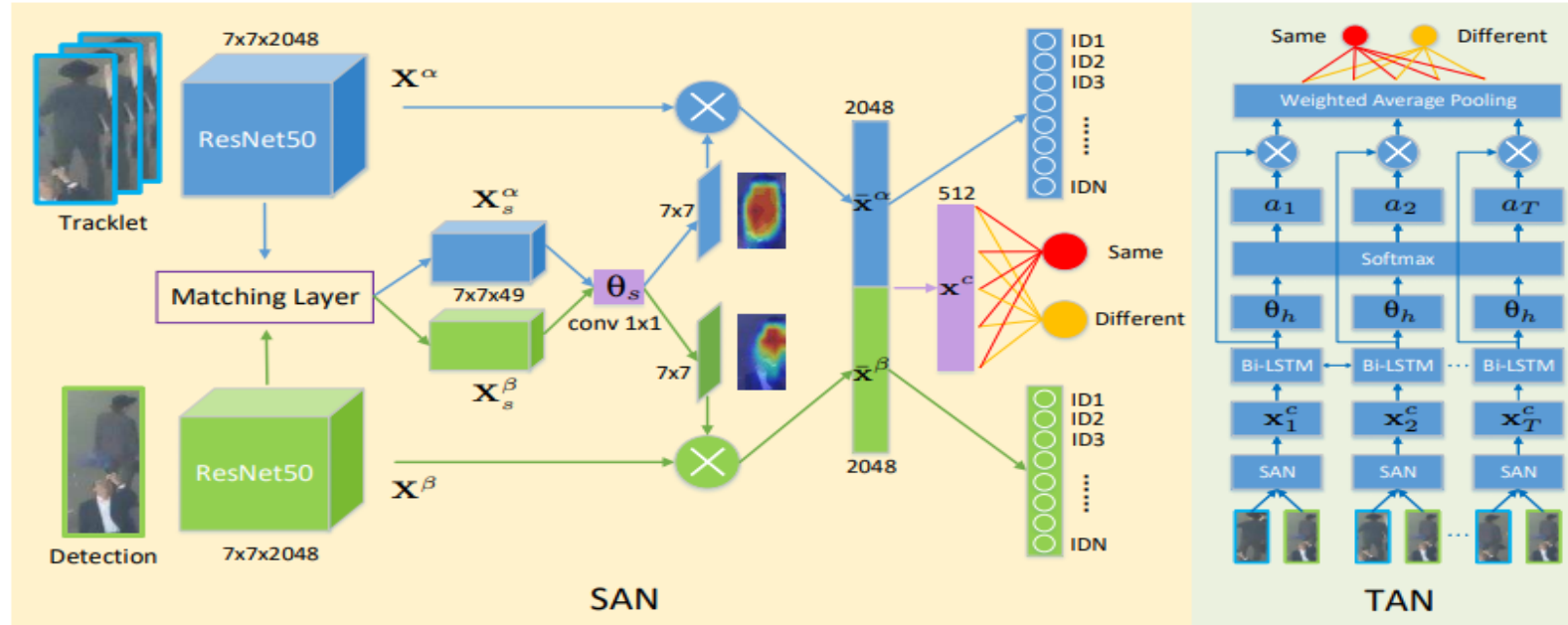


Fig. 4. Network architecture of the proposed DMAN. It consists of the Spatial Attention Network (SAN) and Temporal Attention Network (TAN). Given a candidate detection and a sequence of the target tracklet as inputs, the SAN repeatedly compares the detection with each sample in the tracklet to extract the combined features $\{\mathbf{x}^c\}_1^T$. Taking these features as inputs, the TAN integrates the information from the overall tracklet to infer whether the detection and the tracklet belong to the same target.

Online Multi-Object Tracking with Dual Matching Attention Networks

Multi Object Tracking Pipeline

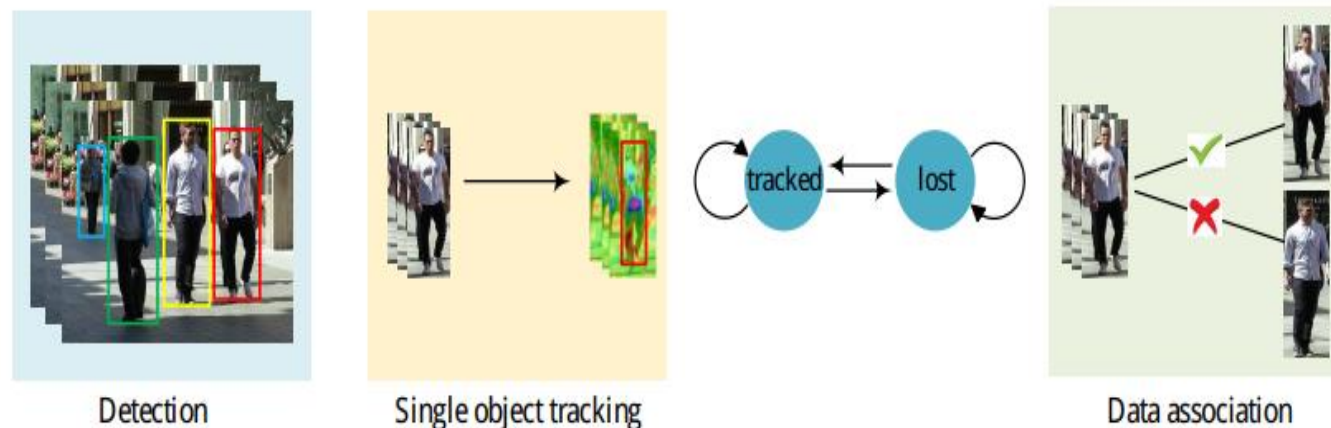
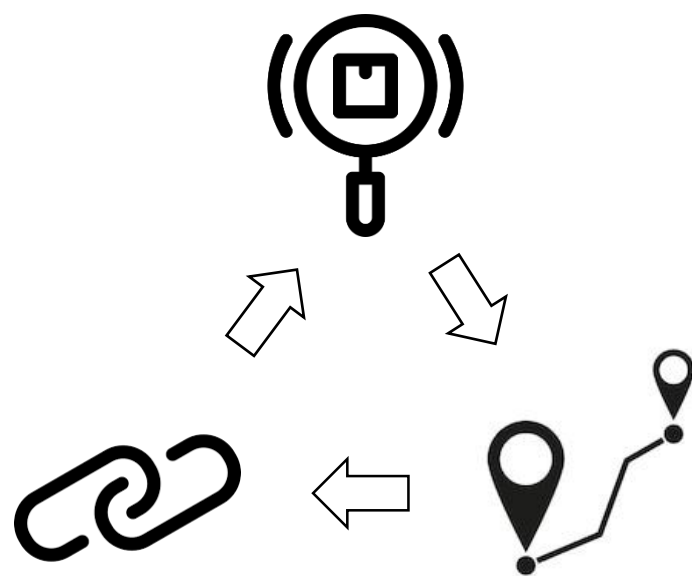


Fig. 2. Proposed online MOT pipeline. This pipeline mainly consists of three tasks: detection, single object tracking, and data association. The state of each target switches between tracked and lost depending on the tracking reliability. Single object tracking is applied to generate the tracklets for the tracked targets while data association compares the tracklets with candidate detections to make assignments for the lost targets.

Online Multi-Object Tracking with Dual Matching Attention Networks

Experiments Setting

DATA SET: MOT 16, MOT 17

- MOT 16: Train Set 7 (5,316 Images) / Test Set 7 (5,919 Images)
- MOT 17: Train Set 21 (15,948 Images) / Test Set 21 (17,757 Images)

Image Size: (224 x 224)

Name	FPS	Resolution	Length
MOT16-02	30	1920x1080	600 (00:20)
MOT16-04	30	1920x1080	1,050 (00:35)
MOT16-05	14	640x480	837 (01:00)
MOT16-09	30	1920x1080	525 (00:18)
MOT16-10	30	1920x1080	654 (00:22)
MOT16-11	30	1920x1080	900 (00:30)
MOT16-13	25	1920x1080	750 (00:30)
Total training			5,316 (03:35)

MOT16 Train Set

Name	FPS	Resolution	Length
MOT16-01	30	1920x1080	450 (00:15)
MOT16-03	30	1920x1080	1,500 (00:50)
MOT16-06	14	640x480	1,194 (01:25)
MOT16-07	30	1920x1080	500 (00:17)
MOT16-08	30	1920x1080	625 (00:21)
MOT16-12	30	1920x1080	900 (00:30)
MOT16-14	25	1920x1080	750 (00:30)
Total testing			5,919 (04:08)

MOT16 Test Set