跟踪

5.siamfc

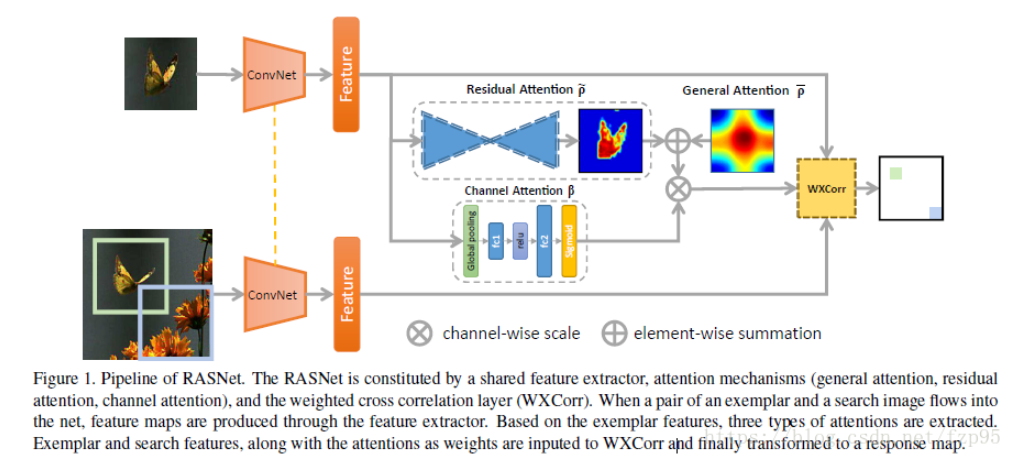
注意力：

11 Show, attend and tell: Neural image caption generation with visual attention **code**

Visual attention [11] is a sound candidate of suppressing background distractors in order to advance feature representation. It is to induce a mask map, i.e., feature weights, to dynamically highlight robust features while removing the information irrelevant to the target. Various attention modules such as soft attention [11] have been applied in visual tracking.

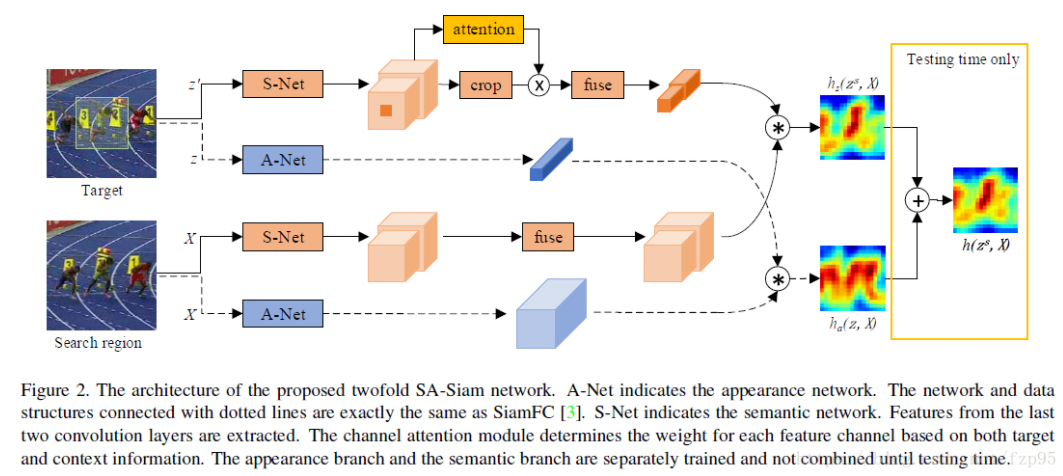
早期的encoder-decoder组成，使用lstm，注意力在特征提取之后

12 Learning attentions: Residual attentional siamese network for high performance online visual tracking



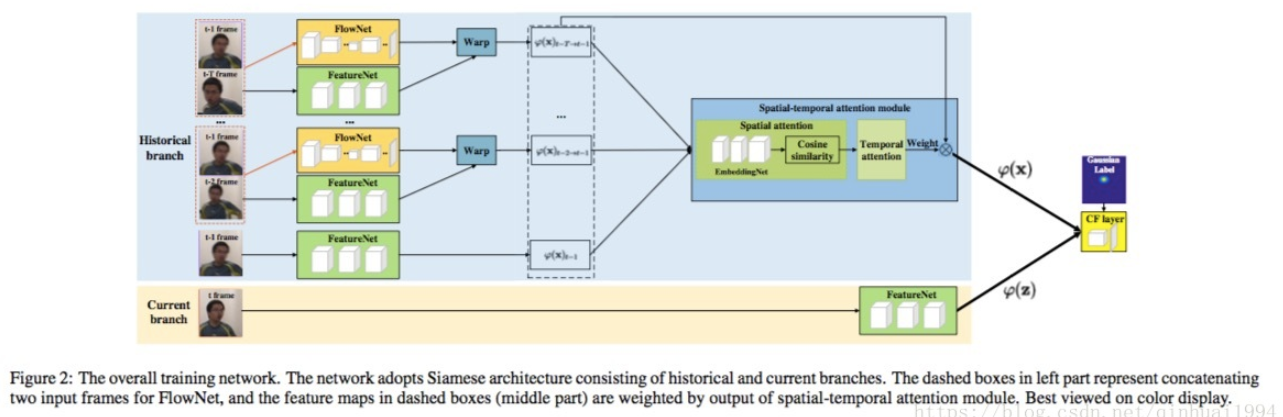
To name a few, [12] explores a multi-attention fusion strategy built on Residual Attentional Network.

13 A twofold siamese network for real-time object tracking  **code**



SA-Siam [13] tailors a channel attention for search branch to achieve the same goal.

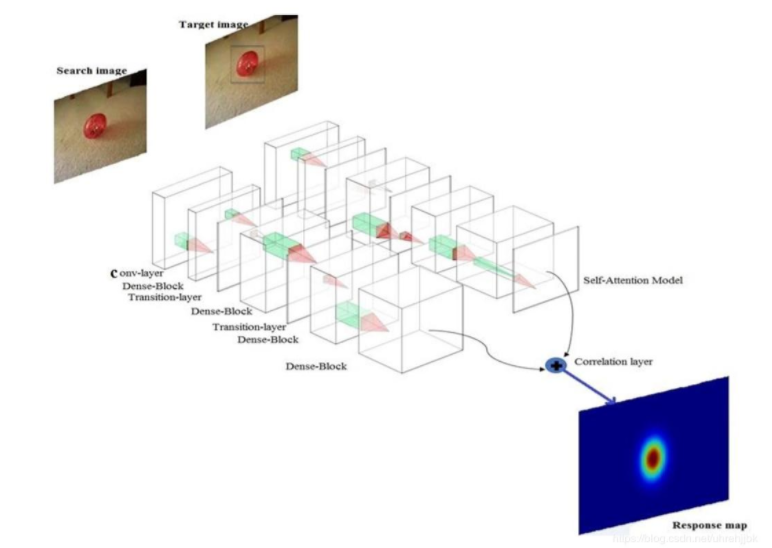
14 End-to-end flow correlation tracking with spatial-temporal attention **code**



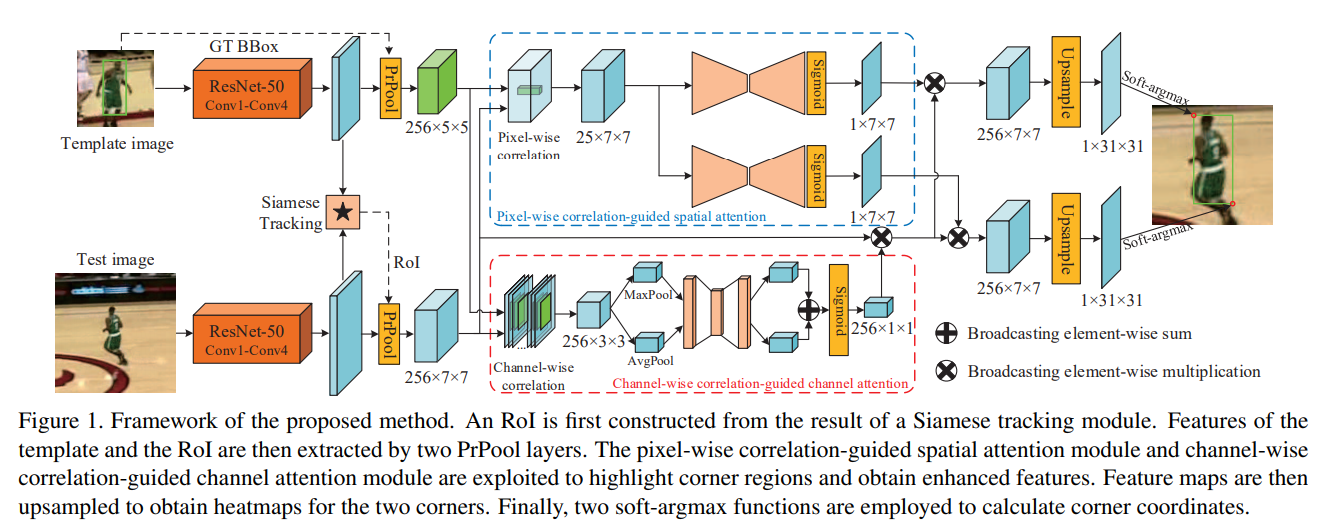
FlowTrack [14] devises a spatial-temporal attention mechanism by jointly learning optical flow and tracking model in a unified framework.

15 Denssiam: End-to-end densely-siamese network with self-attention model for object tracking **code**

DensSiam [15] imposes an attention mechanism over target branch to highlight its non-local features



16 Correlation-guided attention for corner detection based visual tracking



which are beneficial for generating a reliable response map.

CGACD [16] designs a channel and spatial attention, which is

derived from the correlation maps between the template and

the region of interest (ROI) for performance improvement. In

most existing attention based trackers, the attention modules

learn feature weights in a self-attentive way. That is, when

learning feature weights, no extra visual cues are accounted

for. This could weaken the true efficacy of the attention in

tracking systems, which a fortiori include Siamese trackers. In

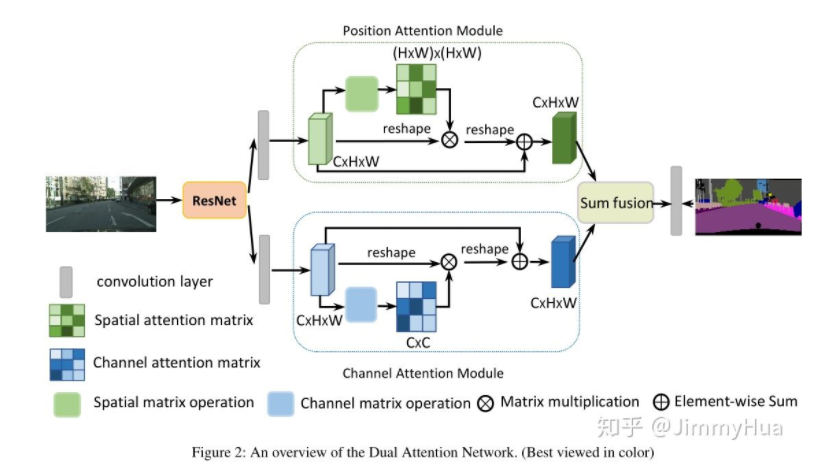
this paper, we will delve into how to integrate extra visual cues

into the non-local attention for visual tracking in a supervisedly

attentive way.

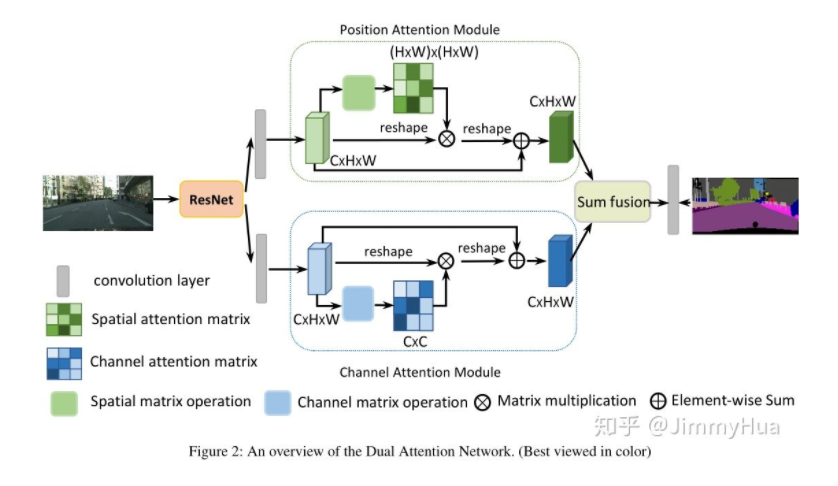
17 Non-local neural networks

18 Dual attention network for scene segmentation



更新：

4 Siamese Regression Tracking With Reinforced Template Updating



Inspired by recent works [17], [18], we design two nonlocal

blocks for Siamese tracker Ocean as our tracker (Nocal-

Siam) to jointly refine visual features and response maps