

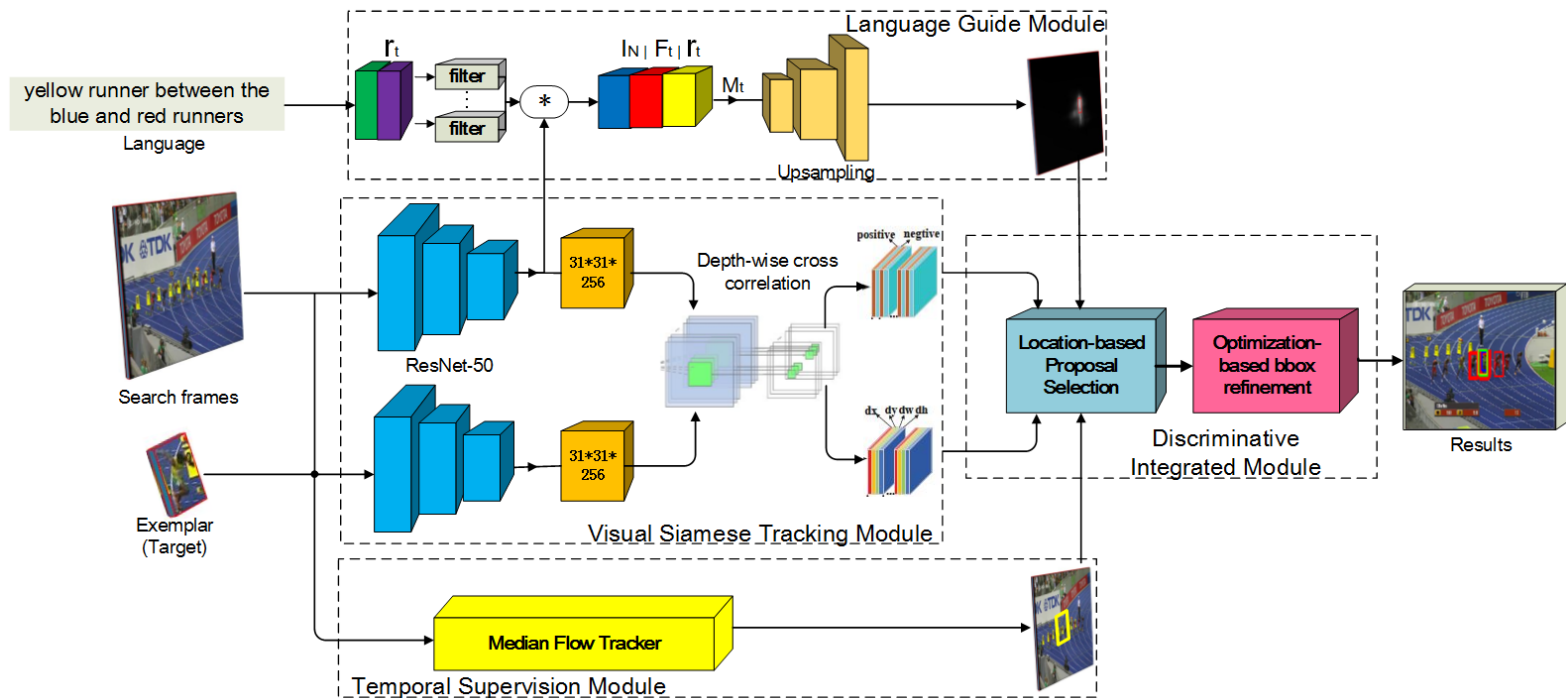
# Siamese single object tracking algorithm with natural language prior

**Qianli ZHOU, Rong WANG, Jinze LI, Naiqian TIAN ,  
Wenjin ZHANG**

Frontiers of Computer Science, DOI: [10.1007/s11704-020-0027-8](https://doi.org/10.1007/s11704-020-0027-8)

# Problems & Ideas

- Problems of performing unstably in complicated environment
  - Inherently have difficulty to differentiate intra-class instances.
  - Shortcomings of only employing classification score for localization.
  - Fail to utilize the inter-frame information of the video.
- Ideas: LangTrack to integrate the language prior, optical flow, visual feature and IoU-based refinement



Framework of LangTrack

# Main Contributions

- Comparison Study

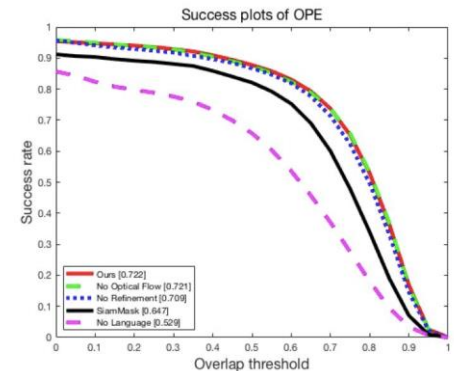
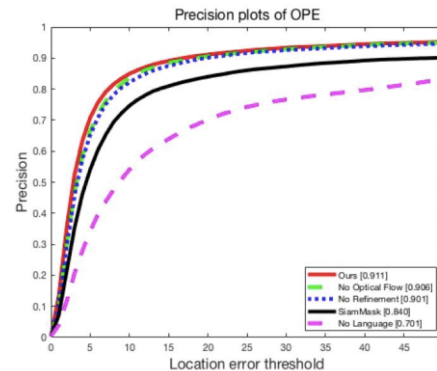
Algorithm	Staple	SINT	CREST	ADNet	SiamFC
Lingual OTB 2013	0.765/0.561	0.882/0.655	0.908/0.673	0.903/0.659	0.809/0.607
Lingual OTB 2015	0.755/0.538	0.789/0.592	0.838/0.623	0.880/0.646	0.771/0.582
Algorithm	SiamRPN	DaSiamRPN	SiamMask	MDNet	Ours
Lingual OTB 2013	0.884/0.658	0.888/0.656	0.841/0.643	0.948/0.708	<b>0.896/0.715</b>
Lingual OTB 2015	0.851/0.637	0.880/0.658	0.851/0.647	0.909/0.678	<b>0.911/0.722</b>

**Table 1** Comparison study on OTB 2013 and OTB 2015

- Ablation Study

Algorithm	Precision	Success
SiamMask	0.840	0.647
No Language	0.701	0.529
No Optical Flow	0.906	0.721
No IoU optimization	0.901	0.709
Ours	0.911	0.722

**Table 3** Ablation Study on lingual OTB2015



LangTrack achieves better performance than most of the popular trackers