

ATMF: Accurate Tracking by Multi-Modal Fusion

Chunhui Zhang¹, Haolin Liu¹, Tianyang Xu², Yong Wang³, Shiming Ge¹

¹Institute of Information Engineering, Chinese Academy of Sciences

²University of Surrey

³Sun Yat-Sen University

Outline

- Introduction
- Difficulties
- Tracking Pipeline
- Improvements
- Results

Introduction

- Anti-UAV tracking task requires to report the **UAV target position** and provide the **target invisible mark** from the **multi-modal dataset**.



Evaluation ranks are calculated according to the results on the **thermal infrared (IR) video**.

<https://anti-uav.github.io/dataset/>

Difficulties in Anti-UAV Challenge

➤ Fast Motion



Difficulties in Anti-UAV Challenge

➤ Target Frequent Disappearance



Difficulties in Anti-UAV Challenge

➤ Text Interference



Difficulties in Anti-UAV Challenge

► Weak UAV Target Signals



Difficulties in Anti-UAV Challenge

➤ Performance of SOTA trackers on Anti-UAV Test-Dev

	IR+RGB	IR
SiamFC	--	0.420
SiamDW-T	0.132	0.276
SiamDW-LT	--	0.475
ATOM	--	0.370
DiMP	--	0.397

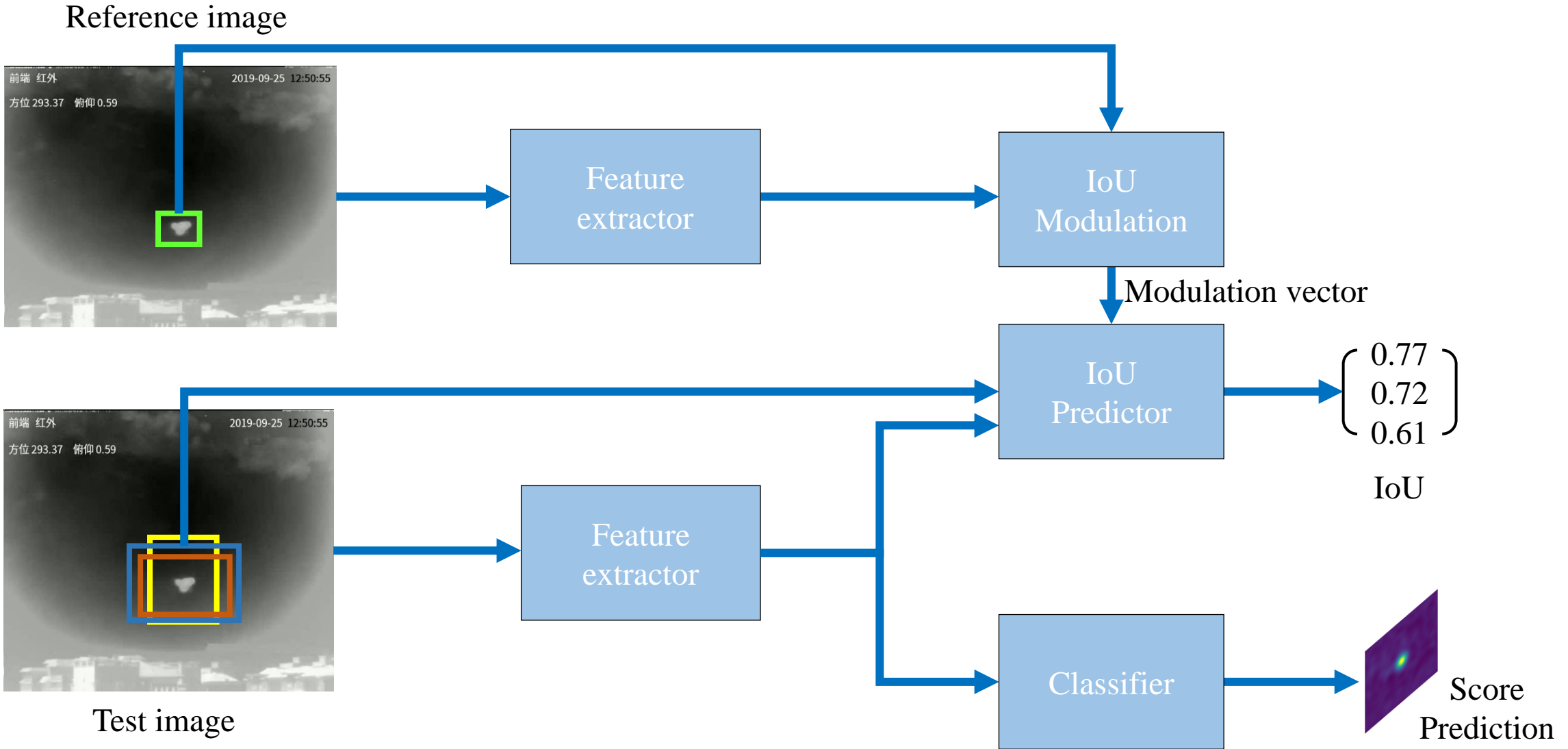
1. RGB-T tracker can not get good result.
2. The long-term tracking framework is better.
3. Existing SOTA trackers need to be improved.
4. Fuse IR and RGB videos?

Tracking Pipeline

Danelljan, Martin, et al. "Atom: Accurate tracking by overlap maximization." CVPR 2019

ATOM

No re-detection module



Tracking Pipeline

H. Du, et al. "Online Deeper and Wider Siamese Networks for Long-Term Visual Tracking." 2019

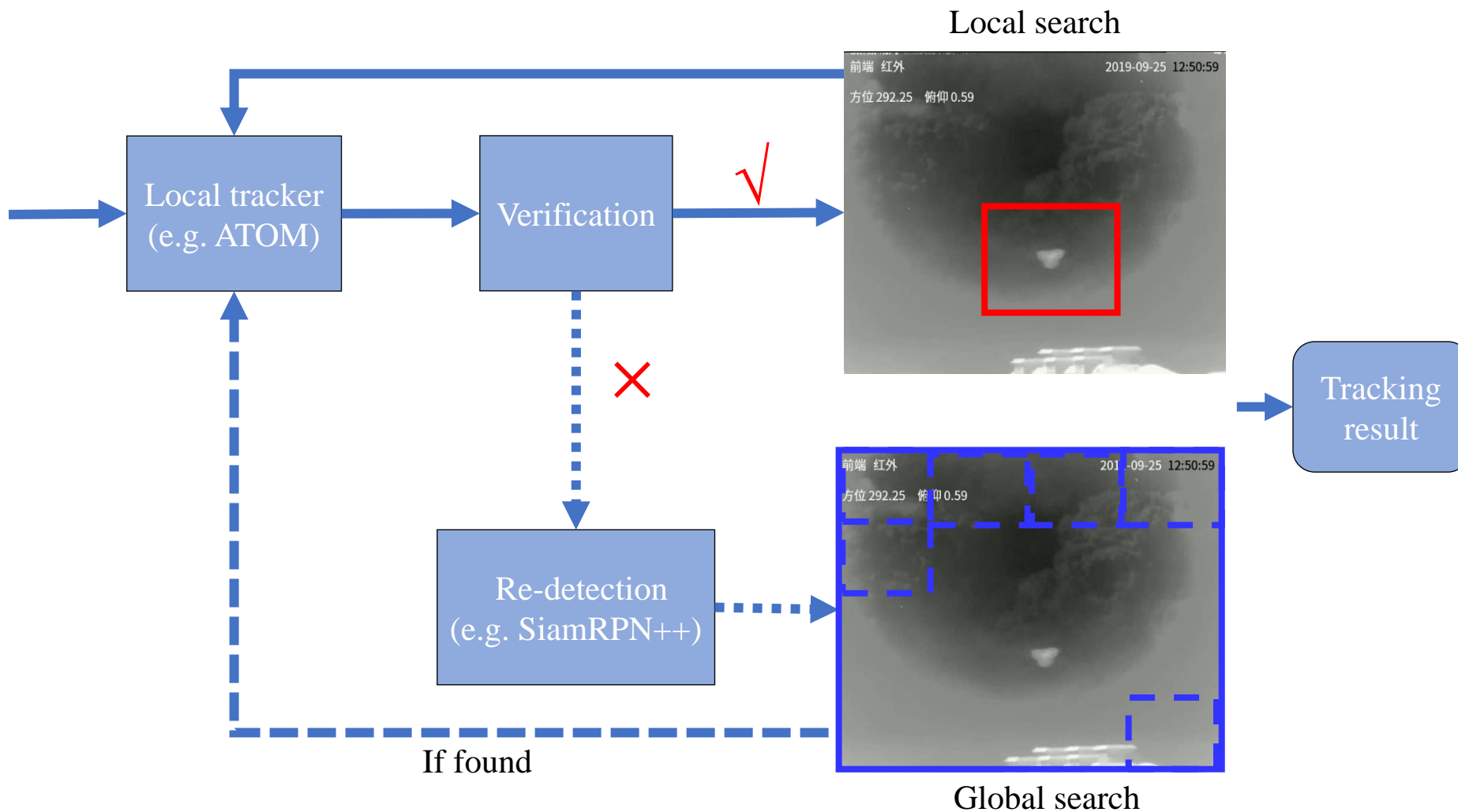
Reference image



Test image

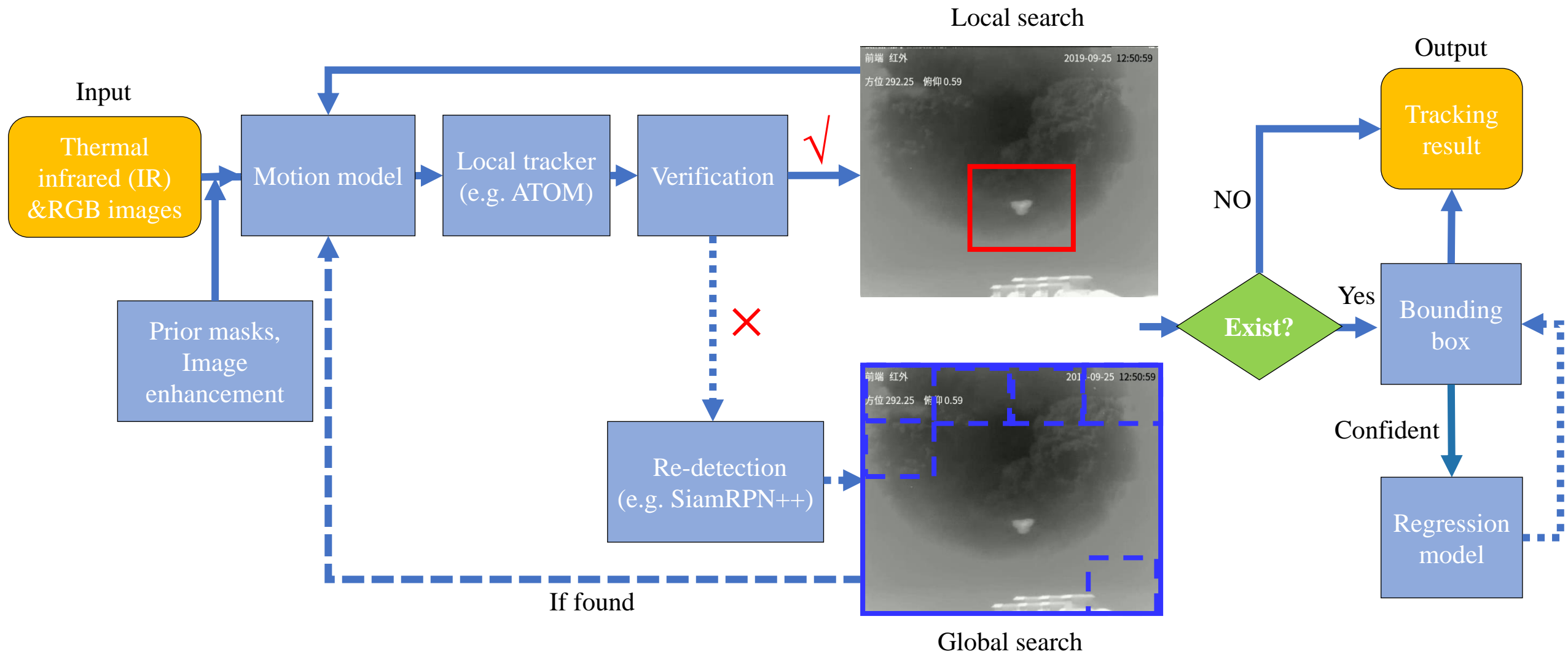
SiamDW-LT

Only use the IR or RGB image
Can not address fast motion

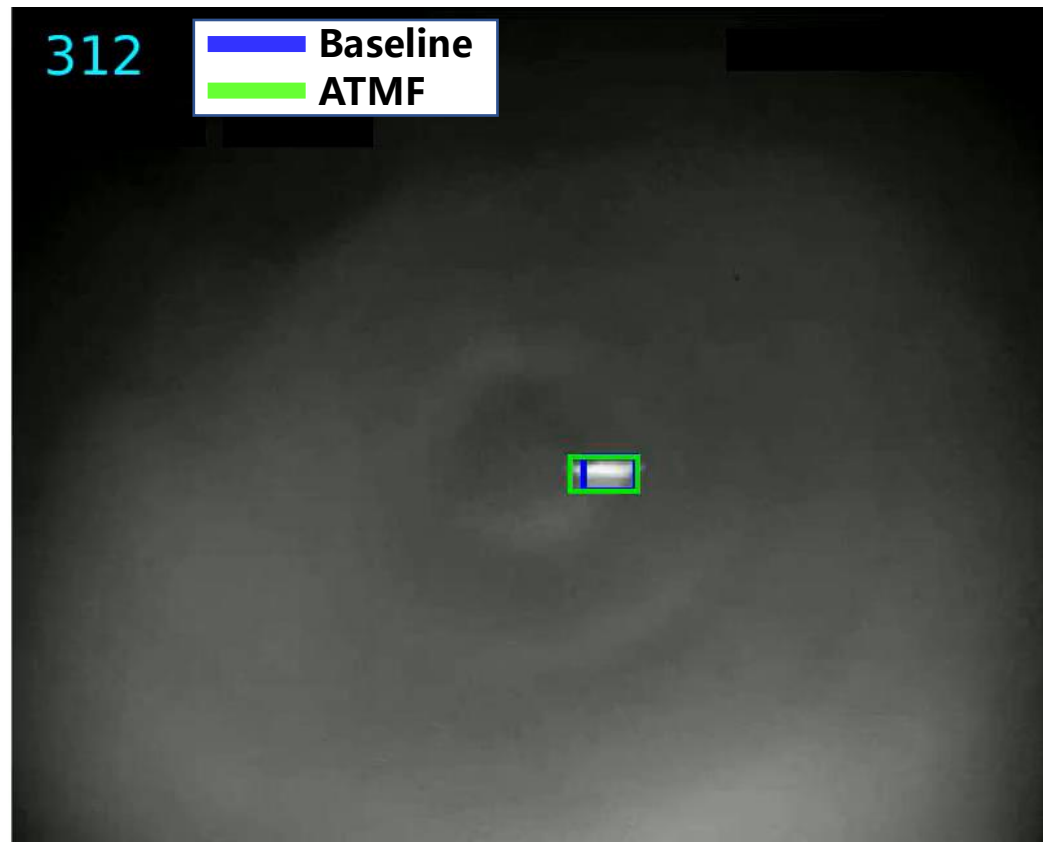
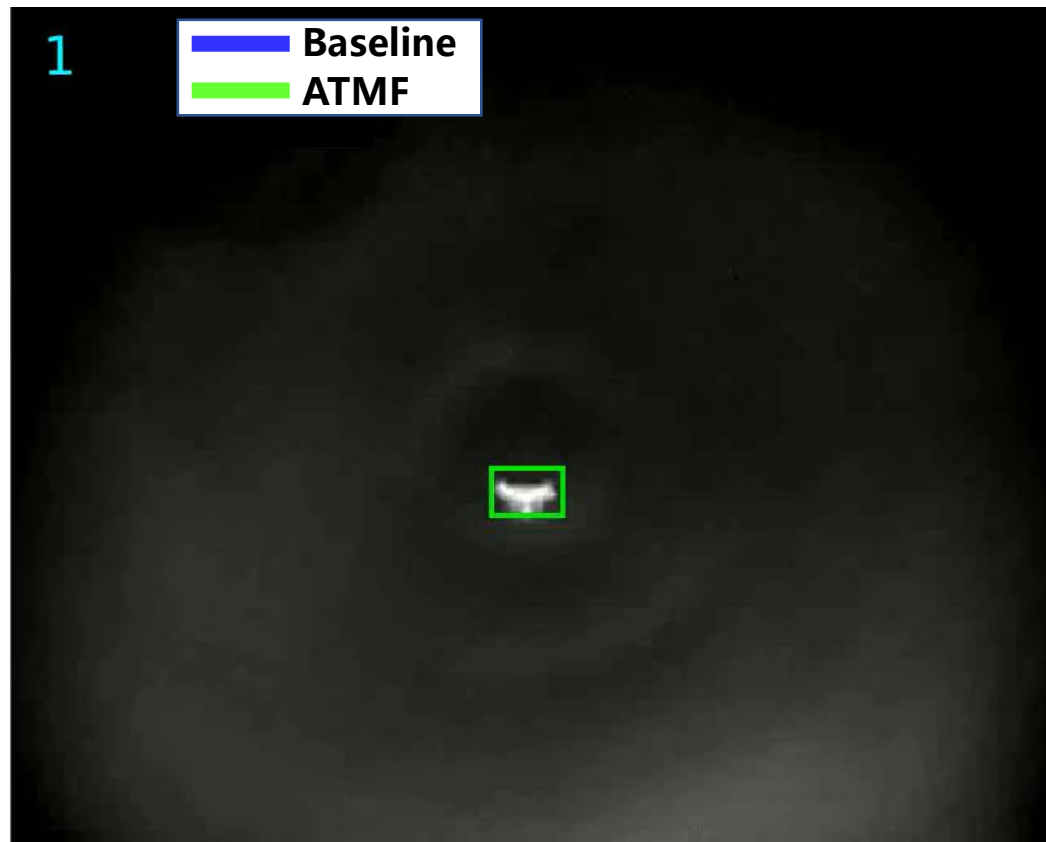


Tracking Pipeline

Our Architecture



Improvements 1-Re-detection & Model ensemble



1. When the short-term tracker is failed a **re-detection module** is activated to search the target in the images globally.
2. **Ensemble** ATOM, DiMP and SiamRPN++.
3. **Motion model**: adaptive search area based on target size and speed.

Improvements 2-Regression model



Baseline 



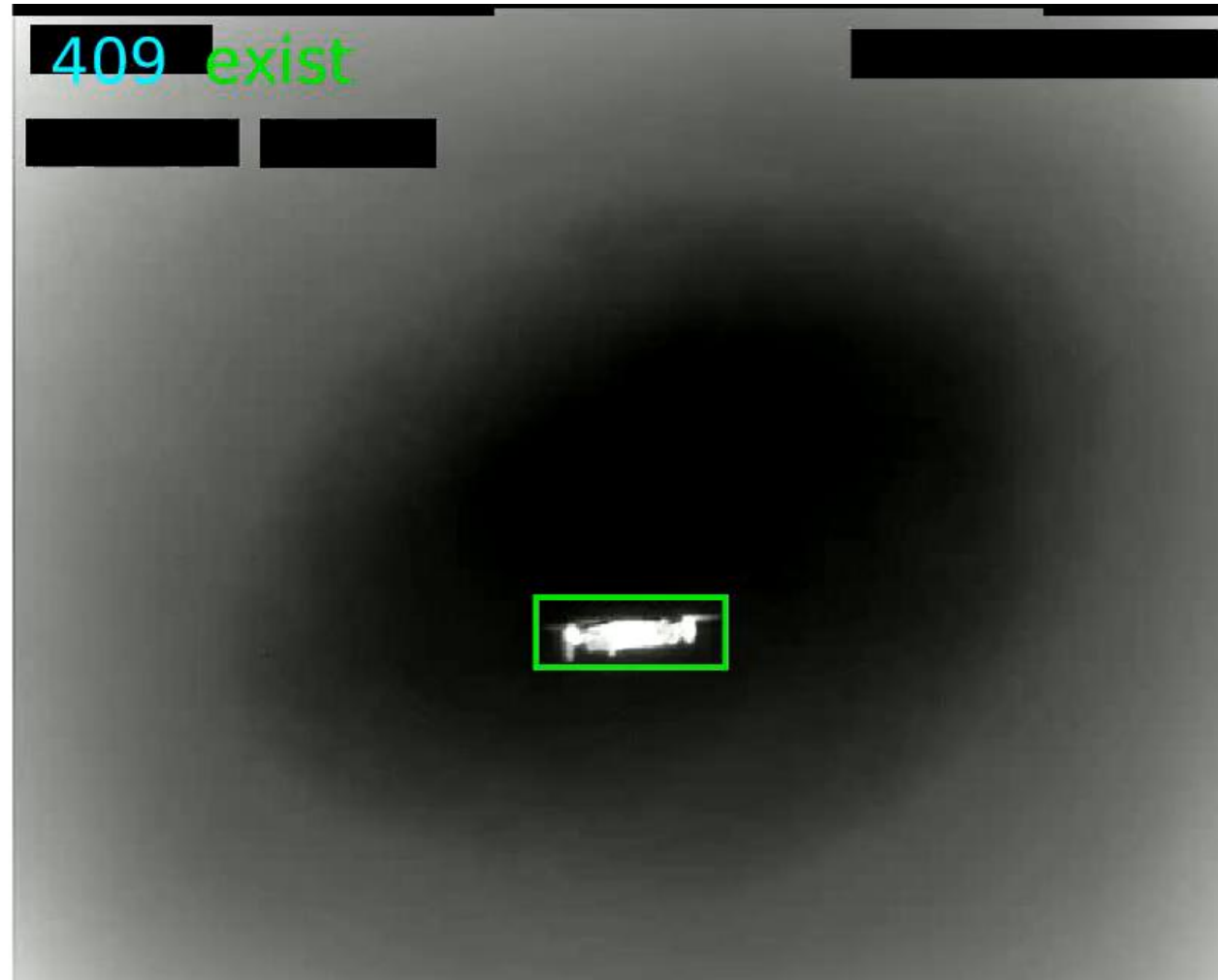
ATMF-RGB 



ATMF-IR 

1. Learn a regression network that converts tracking results from RGB images to IR images.

Improvements 3-Existence estimation



1. If the tracking score is below a given threshold, the target is judged not to exist.

Improvements 4-Other tricks

1. Use image enhancement to address low resolution.
2. Apply prior masks to address text interference.
3. Fine-tune the models on RGBT datasets (e.g. RGBT234, RGBT210).

Results

➤ Anti-UAV Test-Challenge

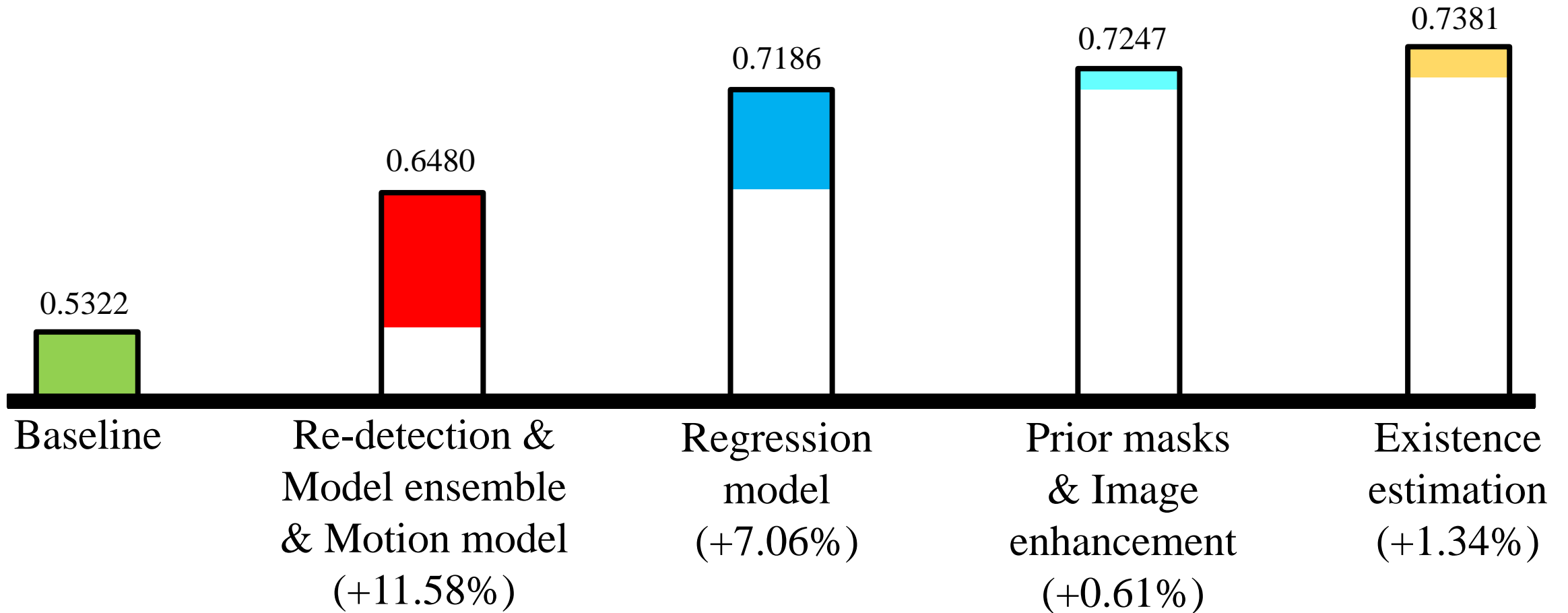
Tracker	ATOM	DiMP	SiamDW-LT	ATMF
acc	0.5322	0.5507	0.6379	0.7381

$$acc = \frac{1}{T} \sum_t \left(IoU_t \cdot \delta(V_t > 0) + P_t (1 - \delta(V_t > 0)) \right)$$

The IoU_t is Intersection over Union (IoU) between each corresponding ground truth and tracking boxes and the v are the visibility flags of the ground truth.

Results

➤ Anti-UAV Test-Challenge



THANKS

for listening



References

- [1] Z. Zhang, et al. “Deeper and Wider Siamese Networks for Real-Time Visual Tracking.” CVPR, 2019
- [2] H. Du, et al. “Online Deeper and Wider Siamese Networks for Long-Term Visual Tracking.” 2019
- [3] Danelljan, Martin, et al. “Atom: Accurate tracking by overlap maximization.” CVPR 2019
- [4] Goutam Bhat, et al. “Learning Discriminative Model Prediction for Tracking.” ICCV, 2019
- [5] Bo. Li, et al. “SiamRPN++: Evolution of Siamese Visual Tracking with Very Deep Networks.” CVPR, 2019
- [6] Bi, Li, et al. “ATP: Accurate Tracking by Progressively Refining.” 2019
- [7] C. Li, et al. “RGB-T Object Tracking: Benchmark and Baseline.” Pattern Recognition, 2019
- [8] C. Li, et al. “Weighted Sparse Representation Regularized Graph Learning for RGB-T Object Tracking.” ACM MM, 2017
- [9] Luca Bertinetto, et al. “Fully-Convolutional Siamese Networks for Object Tracking.” ECCVW, 2016