# VOT2019-RT Challenge

# SiamMargin

Guangqi Chen, Lei Chen, Guoxuan Li, Yanjie Chen, Shan You, Fei Wang, Chen Qian

SenseTime

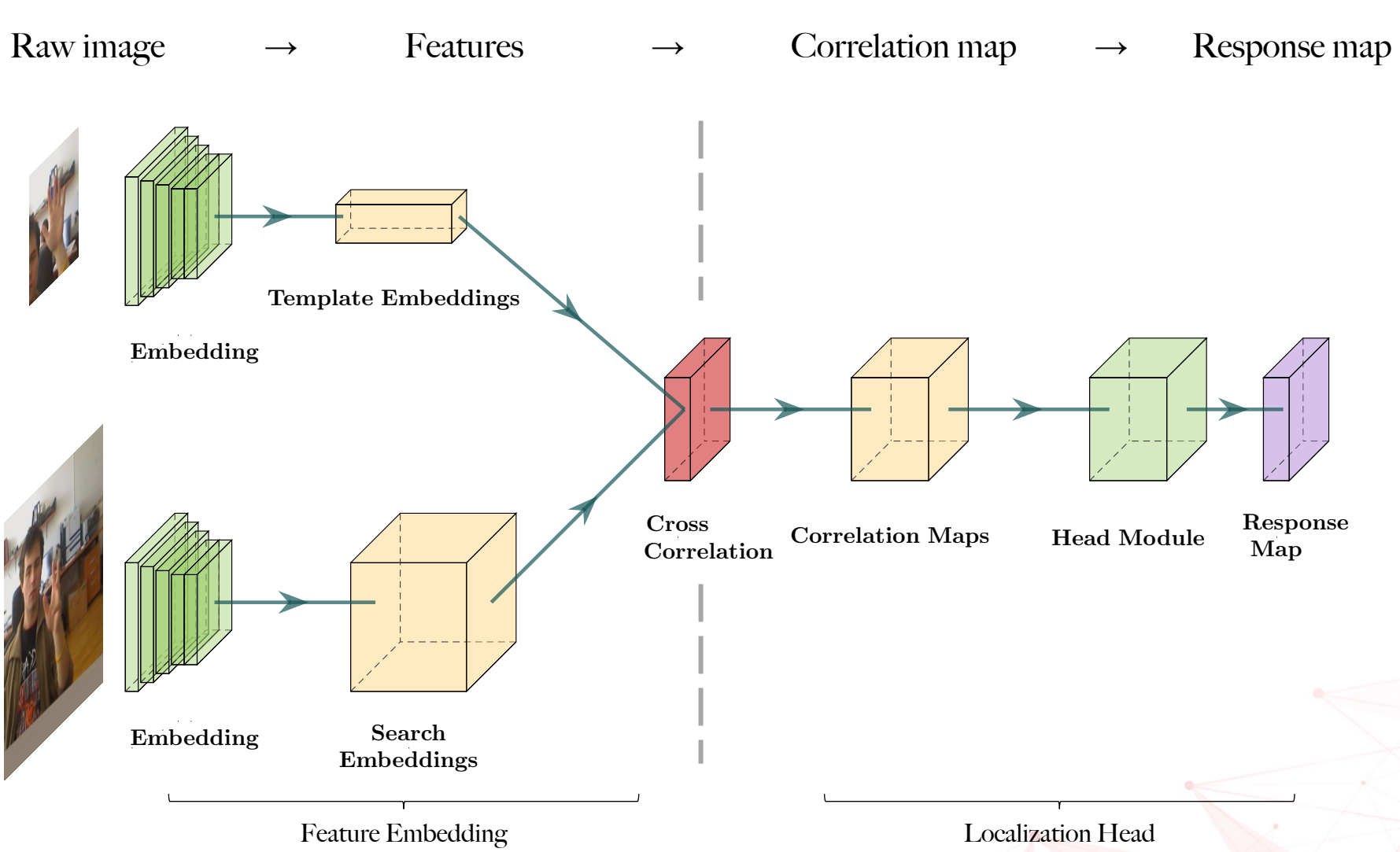
· 1 Rethinking the Architecture of Siamese Tracking

· 2 Learning Discriminative Feature Embeddings for Object Tracking

· 3 Ablation Study of Siamese Margin

# Rethinking the Architecture of Siamese Tracking

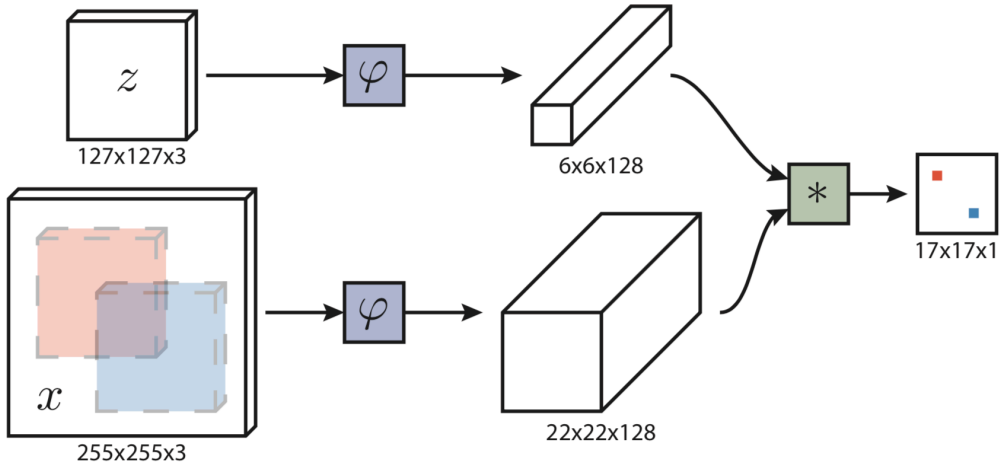
➢ **Summarizing Siamese tracking networks pipeline**



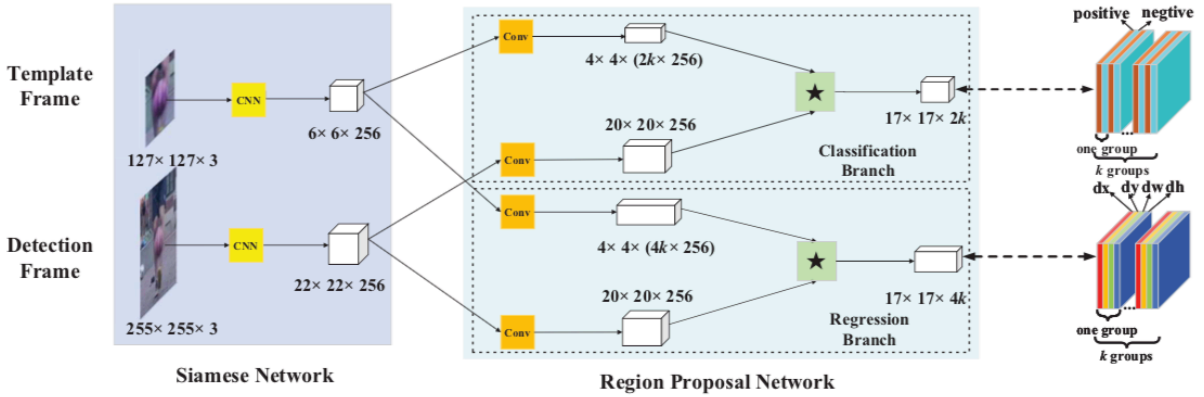
# Rethinking the Architecture of Siamese Tracking

## ➢ Siamese networks for tracking

1. SiamFC [1]




2. SiamRPN [2]





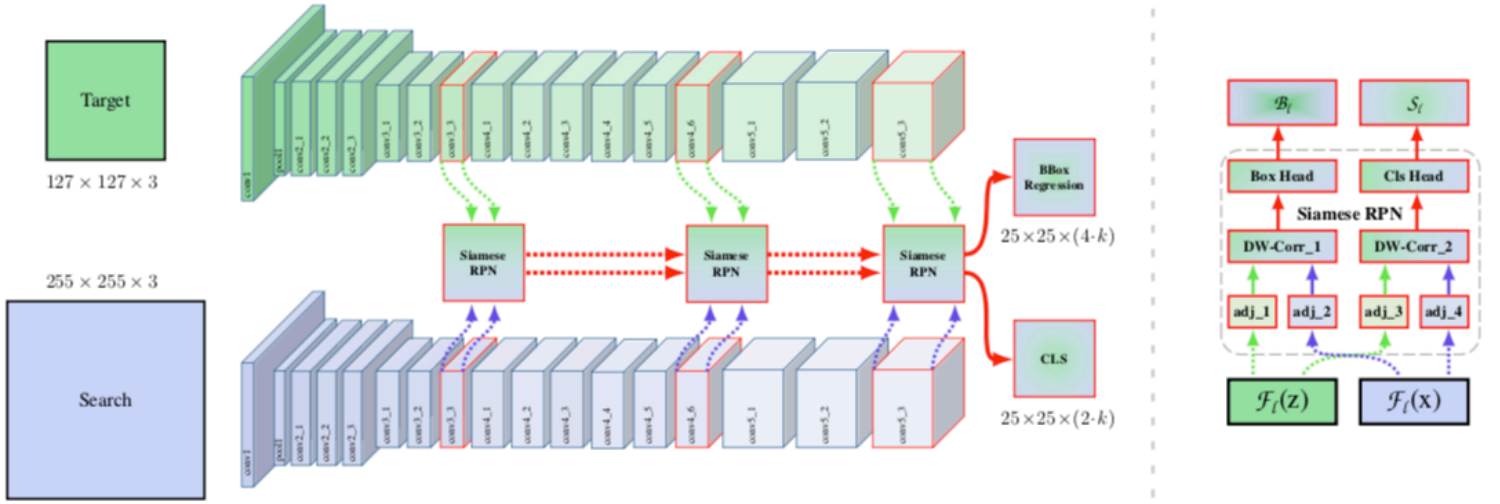
[1] L. Bertinetto, J. Valmadre, J. F. Henriques, A. Vedaldi, and P. H. Torr. Fully-convolutional siamese networks for object tracking. In *European conference on computer vision*, pages 850–865. Springer, 2016

[2] B. Li, J. Yan, W. Wu, Z. Zhu, and X. Hu. High performance visual tracking with siamese region proposal network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8971–8980, 2018.

# Rethinking the Architecture of Siamese Tracking

## ➢ Siamese networks for tracking

3. SiamRPN++ [3]

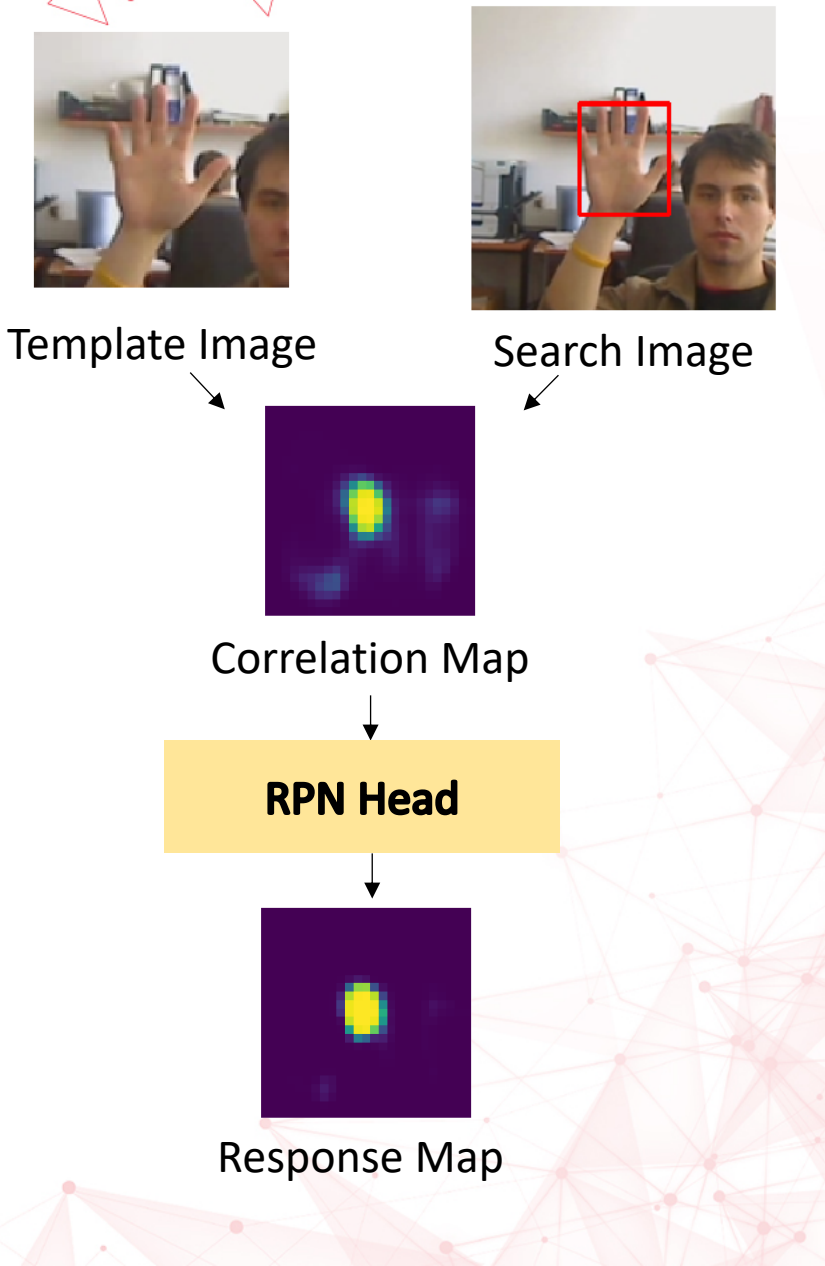



[3] B. Li, W. Wu, Q. Wang, F. Zhang, J. Xing, and J. Yan. Siamrpn++: Evolution of siamese visual tracking with very deep networks. *arXiv preprint arXiv:1812.11703*, 2018.

# Rethinking the Architecture of Siamese Tracking

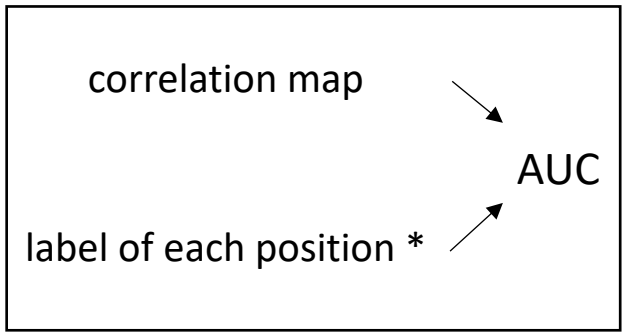
## ➢ The importance of Feature & Correlation map

- Correlation map: **similarity** between template and search local features
- Ideal correlation map ： Response peak at target location
  -- that's just how SiamFC[1] works.
- Region Proposal Networks (RPN) head can be viewed as a **refinement** network
- Therefore, RPN head performs better if correlation map is better. (more clear)

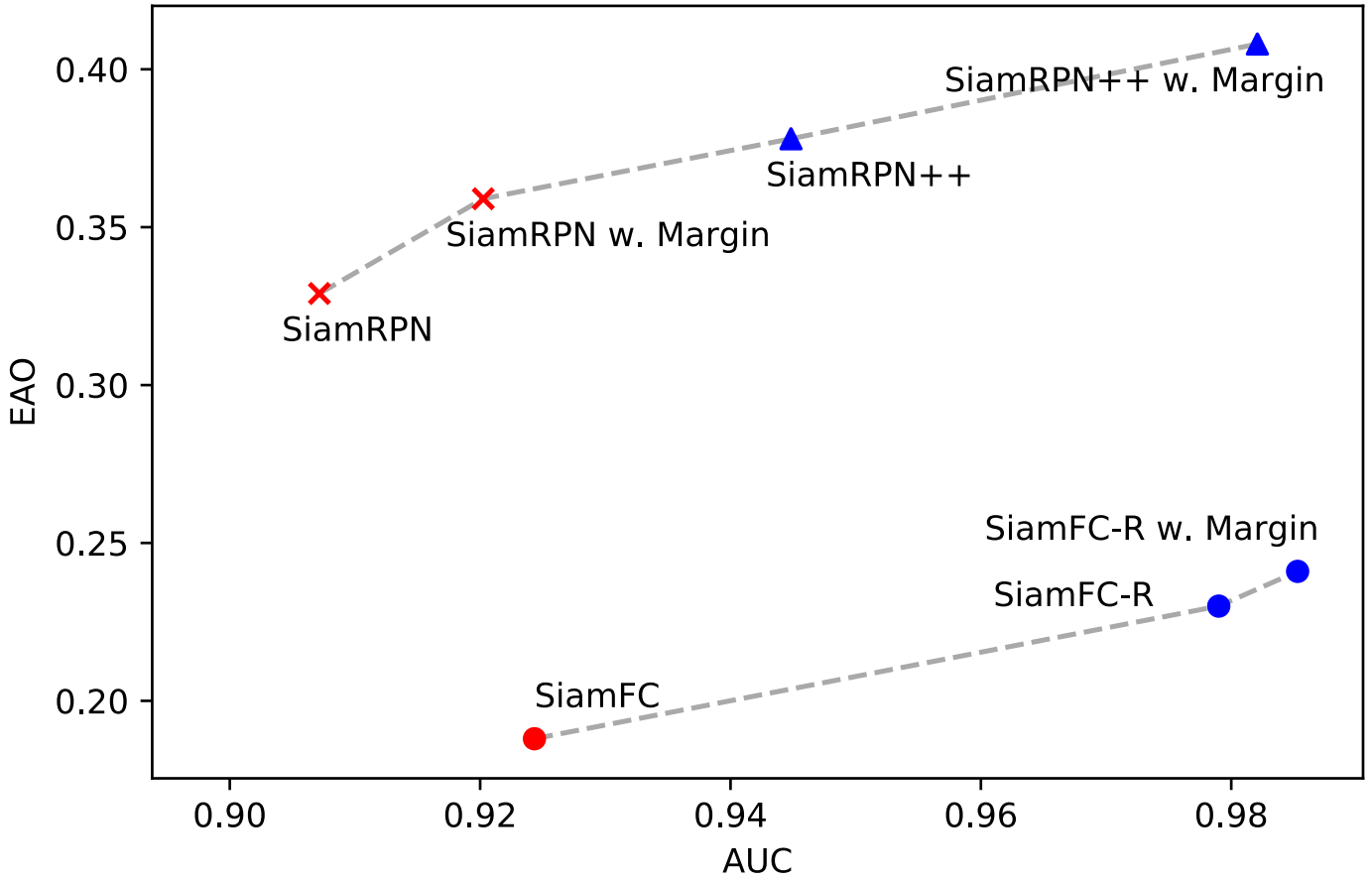


# Rethinking the Architecture of Siamese Tracking

➢ **The importance of Feature & Correlation map**



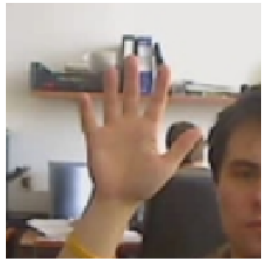

Good Feature -> Good correlation map -> Good Performance




*We follow the label definition of SiamFC[1] for AUC evaluation

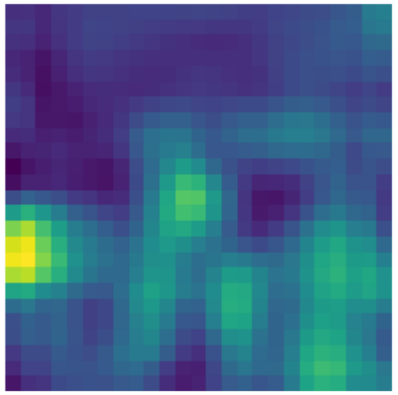
# Rethinking the Architecture of Siamese Tracking

## ➢ The distortion of Correlation map

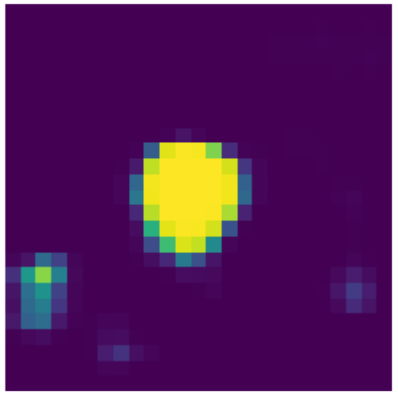


(a) Template image

(b) Search image



(c) **Distorted** correlation map of SiamRPN++



(d) RPN output of SiamRPN++

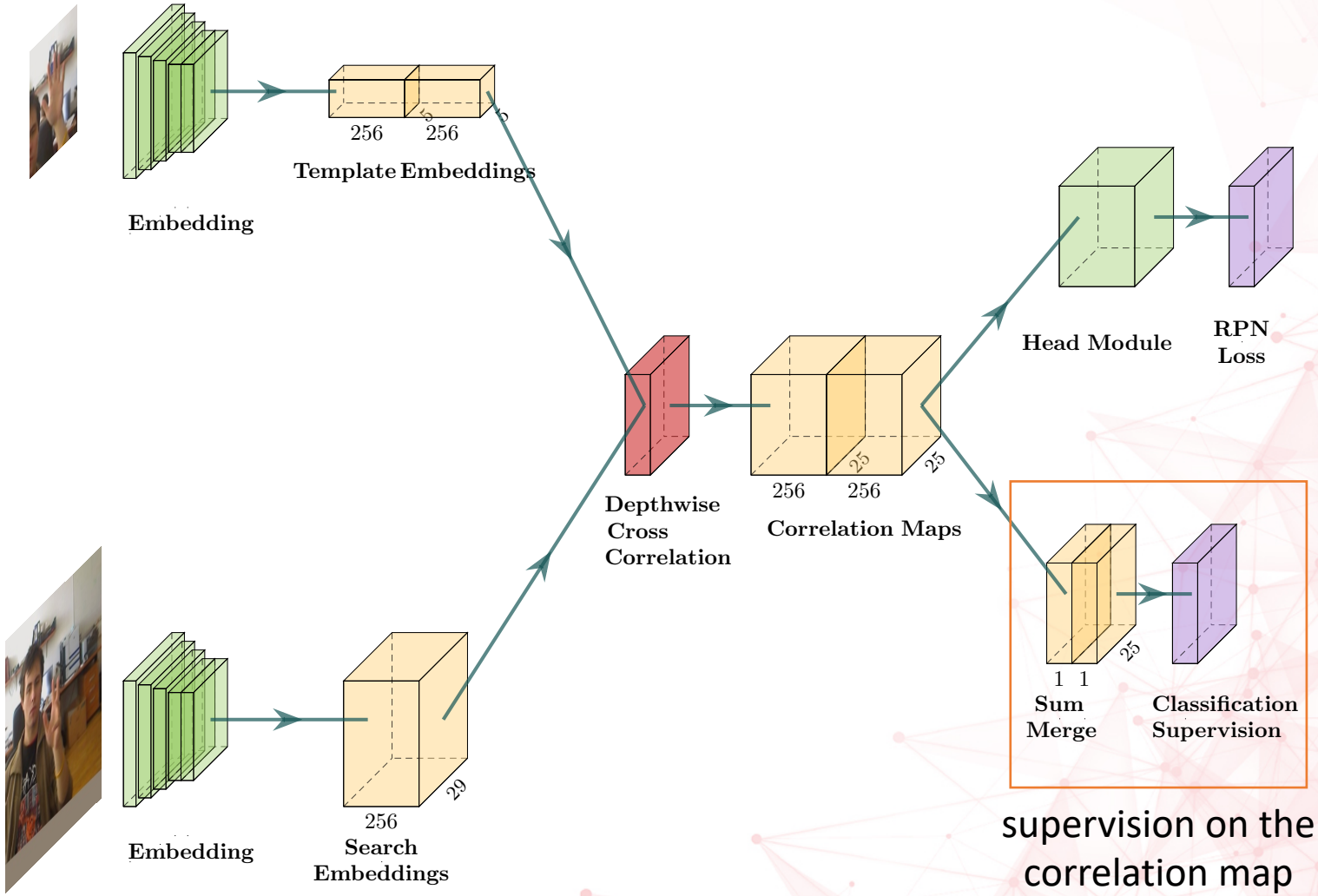
- The **distortion of** correlation map in SiamRPN++! (figure (c))
- Reason： RPN head can’t guarantee monotonically increasing of correlation similarity ！

# Learning Discriminative Feature Embeddings for Object Tracking

➢ **How to get ideal correlation map**
**naïve way : impose supervision on the correlation map**

- Conduct a binary classification supervision on the correlation map

- Label: shared from RPN
'1' for any anchor in the position is positive;
'0' for none of anchors in the position is positive.

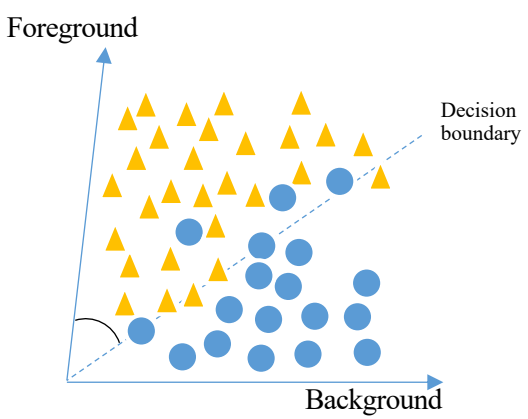



supervision on the correlation map

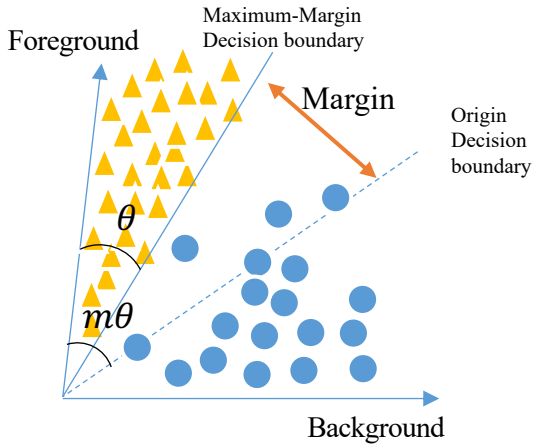
# Learning Discriminative Feature Embeddings for Object Tracking

➢ **Going further： SiamMargin**

Harsh supervision： **Large margin classification loss**




Regular Embedding



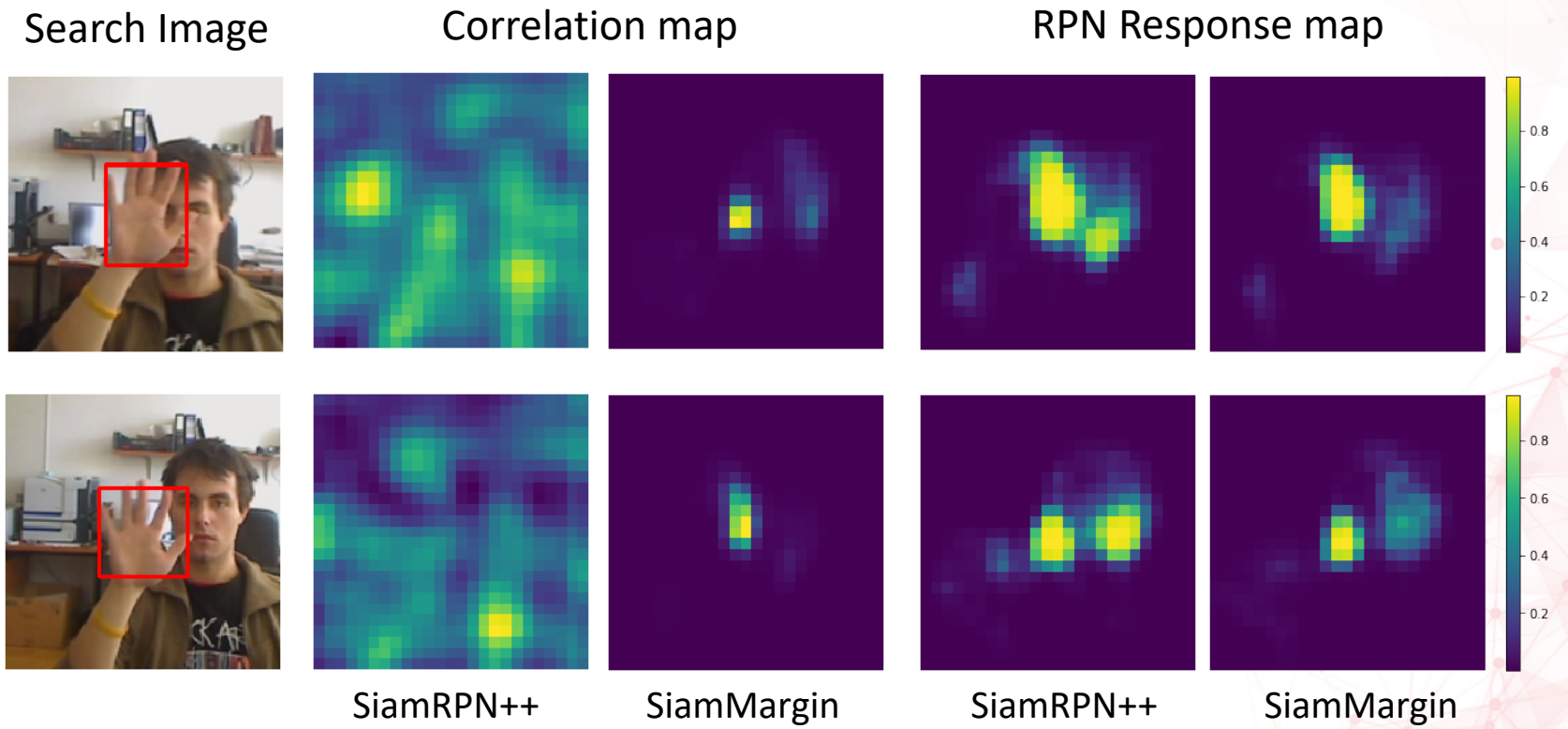

Discriminative Embedding

# Learning Discriminative Feature Embeddings for Object Tracking

## ➢ Correlation map visualization

- Semantically similar: e.g. ‘FACE’ and ‘HAND’ in the Search Image.

- Our SiamMargin model can suppress the response of hard negative samples while maintain true target response.

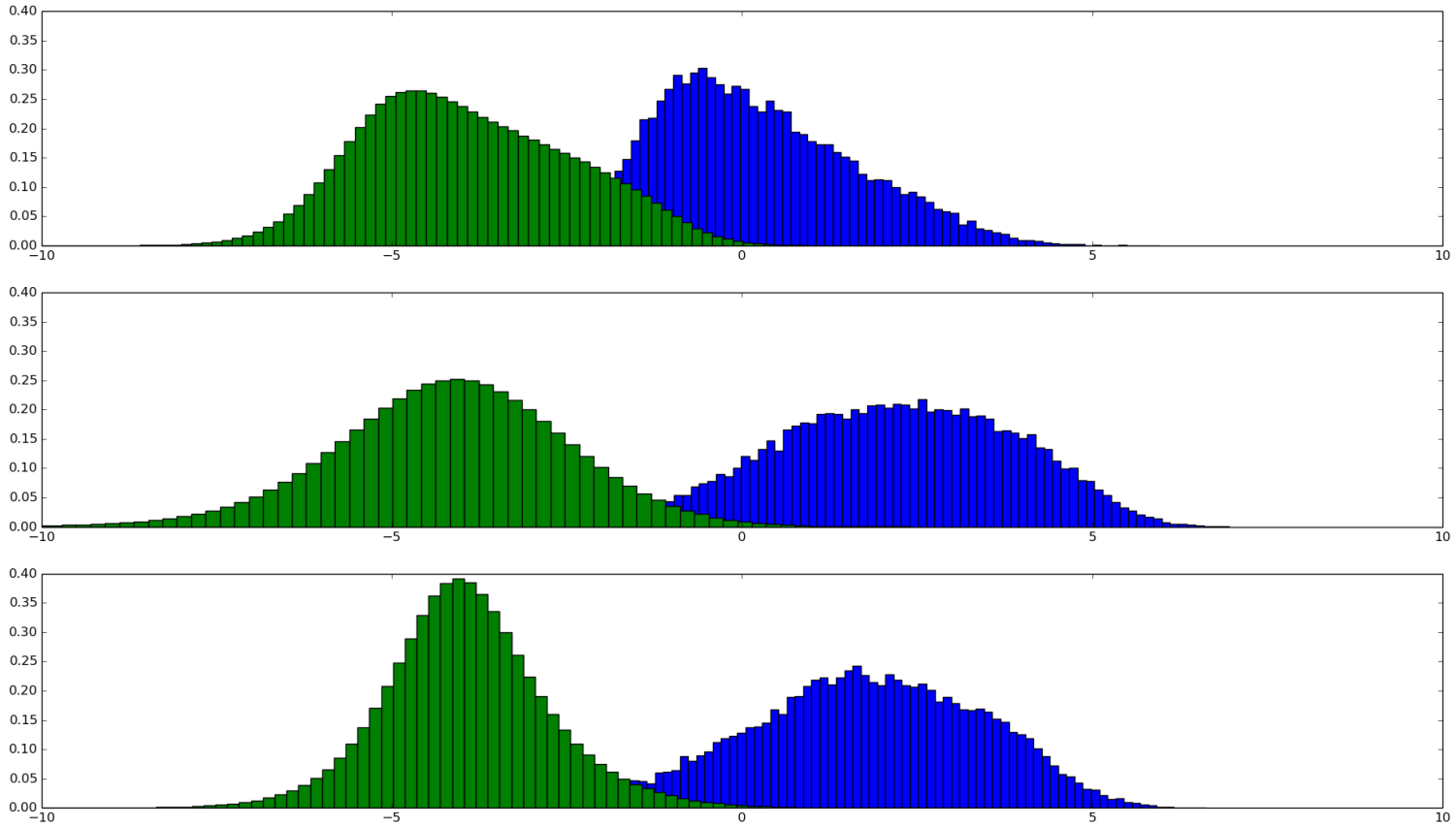


# Learning Discriminative Feature Embeddings for Object Tracking

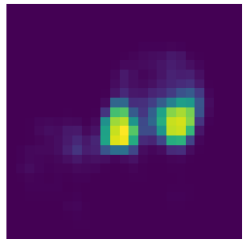
➢ **The RPN response distribution becomes more separable**

target

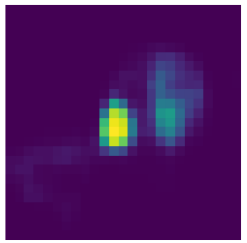
background



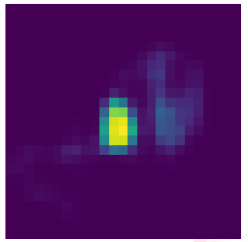

logits distribution of RPN classification response



Without supervision



Regular supervision



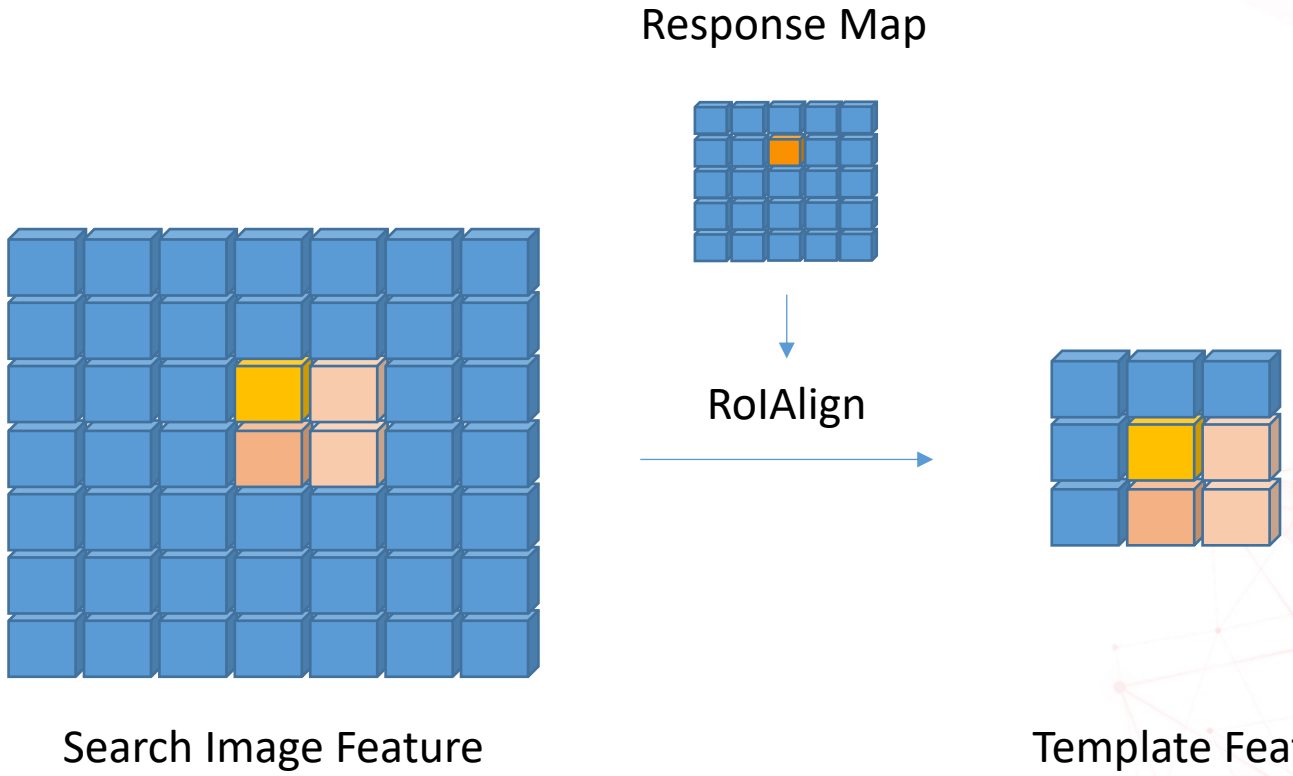
Large Margin supervision

RPN Response

# Learning Discriminative Feature Embeddings for Object Tracking

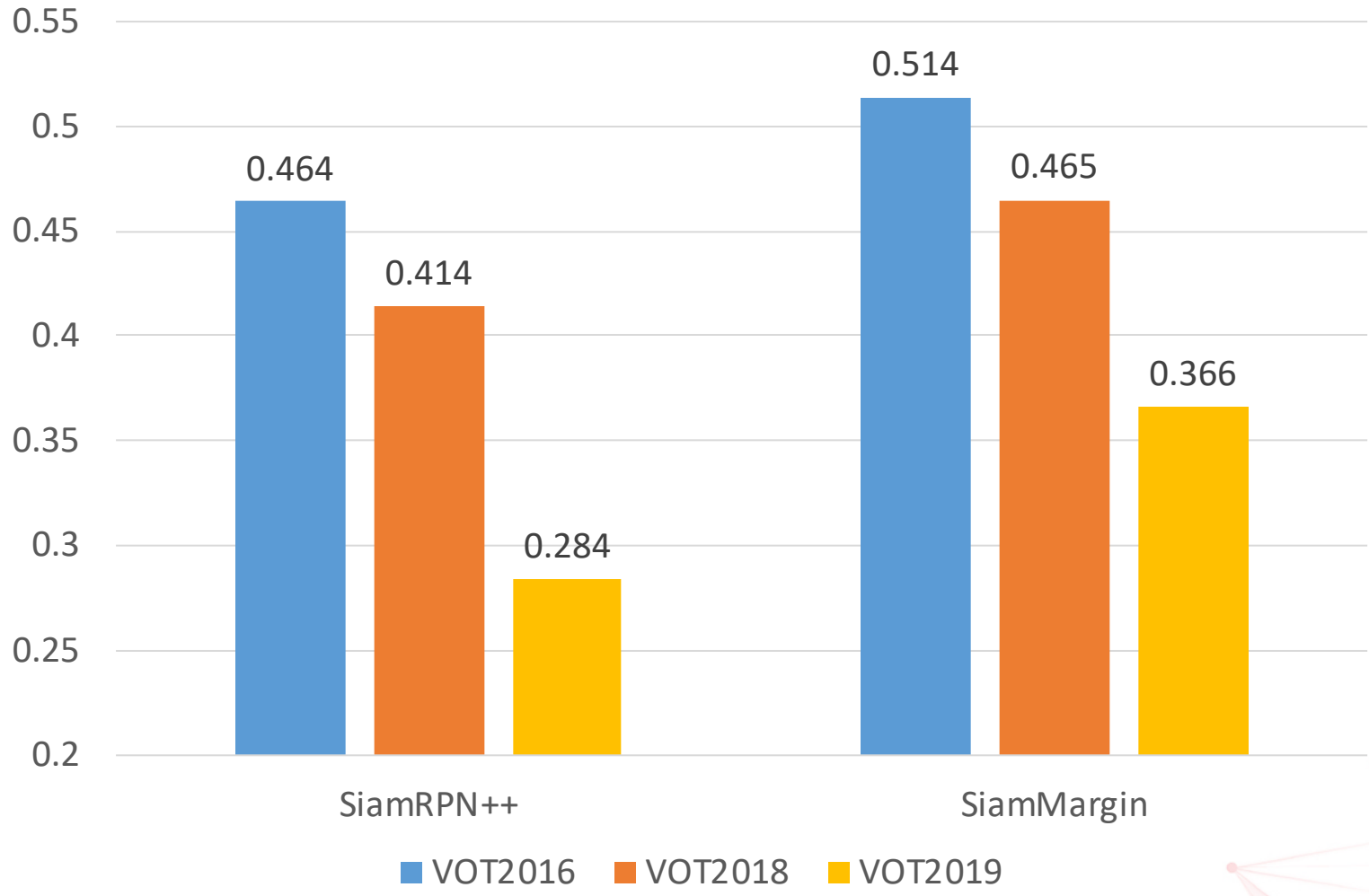
## ➢ Adaptive Online Update in Siamese Networks

- Previous Siamese networks lack of capacity to handle appearance change (lack of update)
- Discriminative embedding is crucial for reducing noise caused by template update
- Updating with a simple yet efficient moving average strategy
- Nearly cost free: only **1.86%** extra time cost!



# Ablation Study of Siamese Margin

- **SOTA Performance on the Several Benchmarks**



# Ablation Study of Siamese Margin

➢ **Results on VOT2019-RT**

| Large Margin | Online update（ROI-Align） | One more search | Adaptive Search Region | Accuracy | Failure | EAO |
|---|---|---|---|---|---|---|
| - | - | - | - | 0.585 | 91.3 | 0.287 |
| - | - | Y | Y | 0.587 | 87.3 | 0.297 |
| Y | - | Y | Y | 0.595 | 69.0 | 0.325 |
| Y | Y | Y | Y | 0.578 | 65.0 | 0.366 |

**More Robust!**

Failure from 91.3 times -> 65.0 times

28.8% relative improvement !

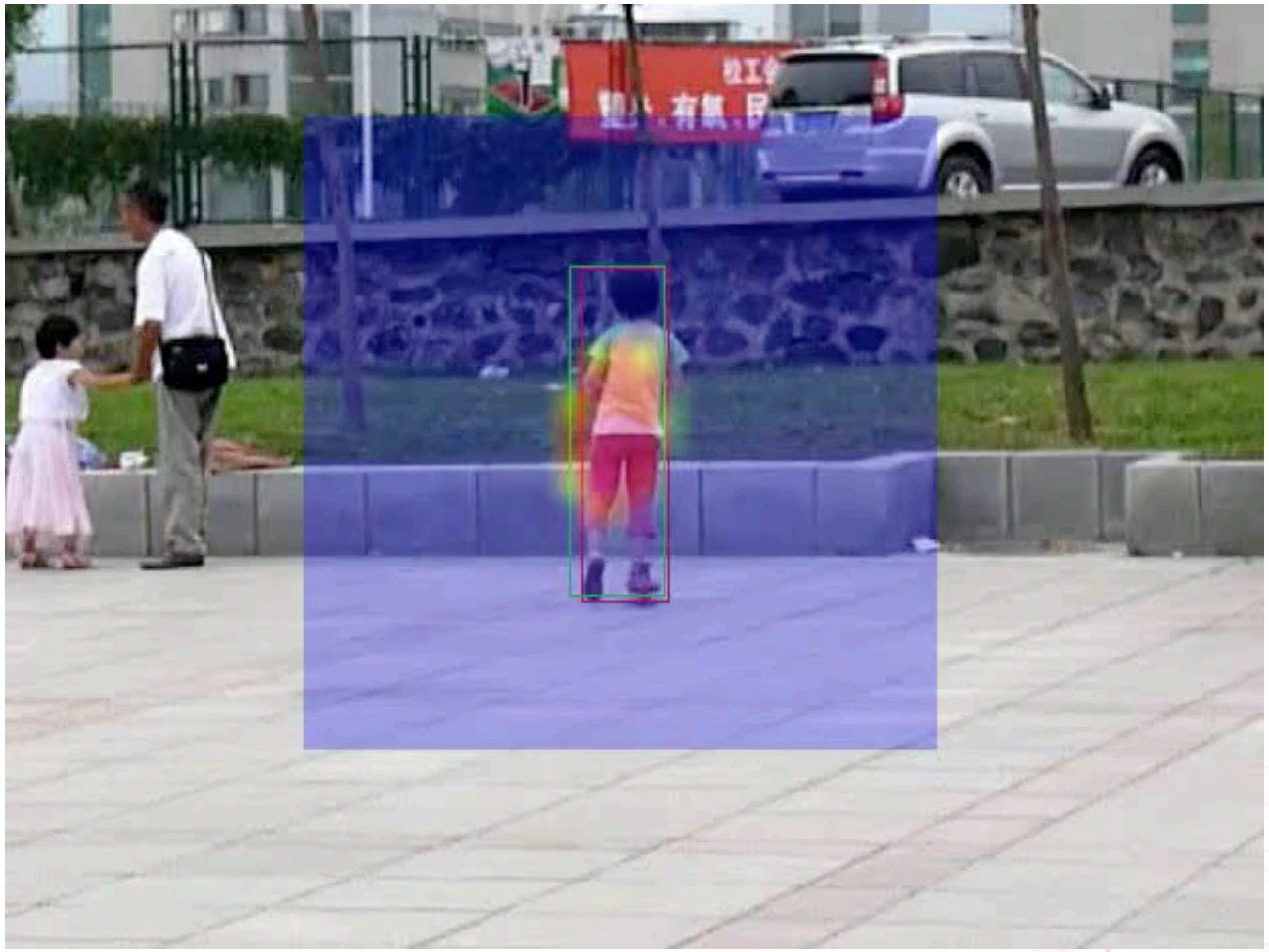
# Ablation Study of Siamese Margin

➢ **Video Demo**

| EAO | Accuracy | Failures | FPS |
|---|---|---|---|
| 0.366 | 0.578 | 65.0 | 45.618 |

GroundTruth

Prediction



# VOT2019-RT Challenge

Thanks!

Q & A

Guangqi Chen, Lei Chen, Guoxuan Li, Yanjie Chen, Shan You, Fei Wang, Chen Qian

SenseTime