Effect of Annotation Errors on Drone Detection with YOLOv3

Aybora Köksal, Kutalmış Gökalp İnce, Aydın Alatan

Middle East Technical University

Electrics and Electronical Engineering Department

Center For Image Analysis (OGAM)

METU

Outline

- Introduction
- Performance of YOLOv3 on Anti-UAV Dataset
- Performance of YOLOv3 with Simulated Annotation Errors
- Annotation Correction
- Conclusion

Introduction

The main aims of this study are,

- Understanding the effects of annotation errors in a training set for object detection
- Automatically detecting and correcting annotation errors in any tracking dataset with a ground truth

Introduction

- UAV detection by fine tuning YOLOv3[1]

- Bad results due to annotation errors

- What about correcting them?

[1]: Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. CoRR, abs/1804.02767, 2018

Performance of YOLOv3 on Anti-UAV Dataset

- Only thermal images of test-dev dataset is used

- 70 videos -> Training Set
- 30 videos -> Validation Set

Performance of YOLOv3 on Anti-UAV Dataset

- Metrics used to measure performance:
  - Hit: IoU between detection output and ground truth > 0.5
  - Hit Rate: $\frac{\#\ of\ hits}{\#\ of\ positive\ samples}$
  - False Alarm: IoU between detection output and ground truth = 0
  - False Alarm Per Minute: # of false alarms/minute through all samples
  - Tracking Accuracy: $\frac{1}{T}\sum_{t=1}^{T} IoU_t * v_t * p_t + (1-p_t)(1-v_t)$ $\quad v_t$: Visibility Flag $\quad p_t$: Prediction Flag
  - Modified Tracking Accuracy: $\frac{\sum_{t=1}^{T} IoU_t * v_t * p_t + (1-p_t) * (1-v_t)}{\sum_{t=1}^{T} max(v_t, p_t) + (1-p_t) * (1-v_t)}$

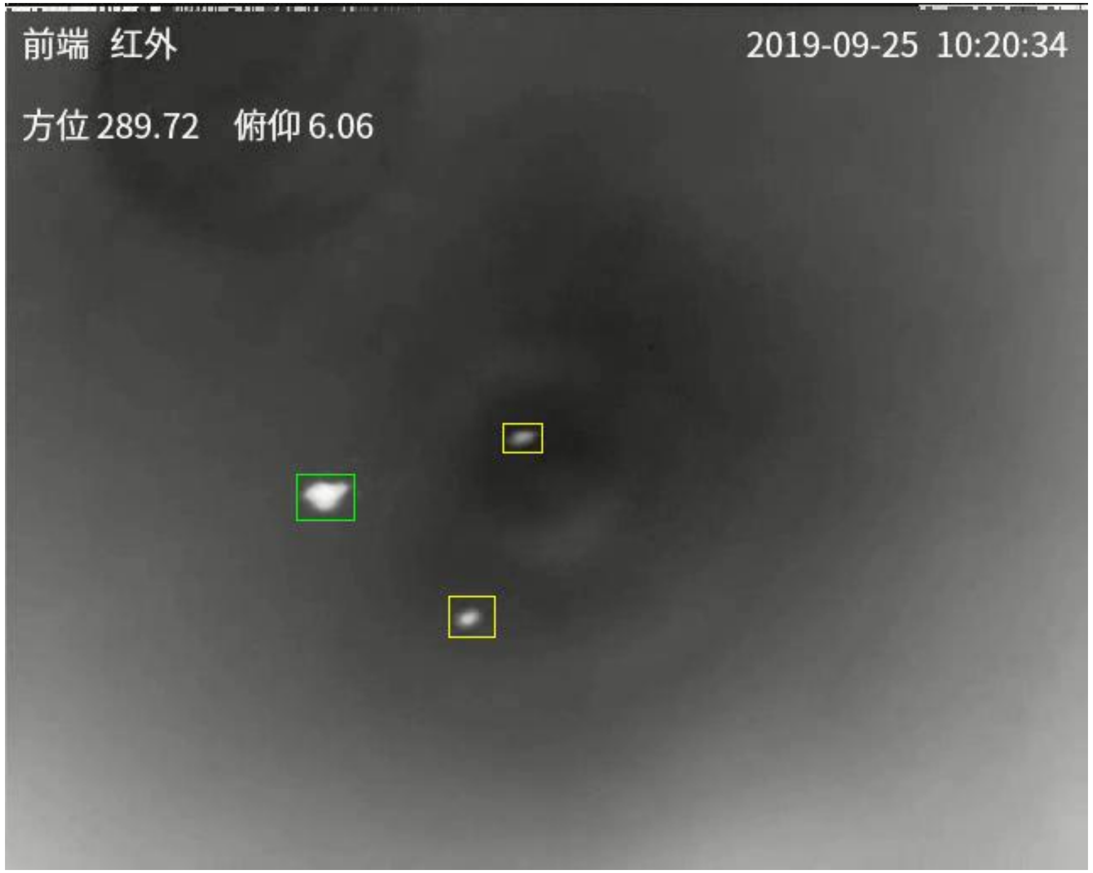
METU

Performance of YOLOv3 on Anti-UAV Dataset

- Modified Tracking Accuracy: $$\frac{\sum_{t=1}^{T} IoU_t * v_t * p_t + (1-p_t)*(1-v_t)}{\sum_{t=1}^{T} max(v_t, p_t) + (1-p_t)*(1-v_t)}$$

v_t : Visibility Flag

p_t : Prediction Flag



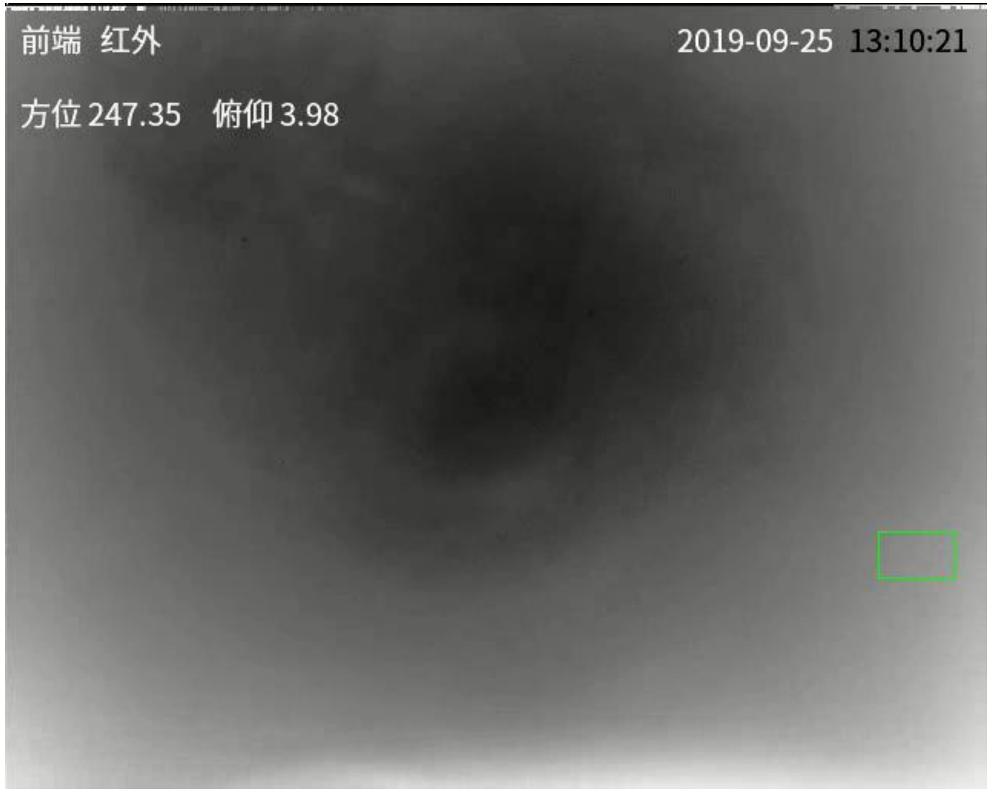
Performance of YOLOv3 on Anti-UAV Dataset

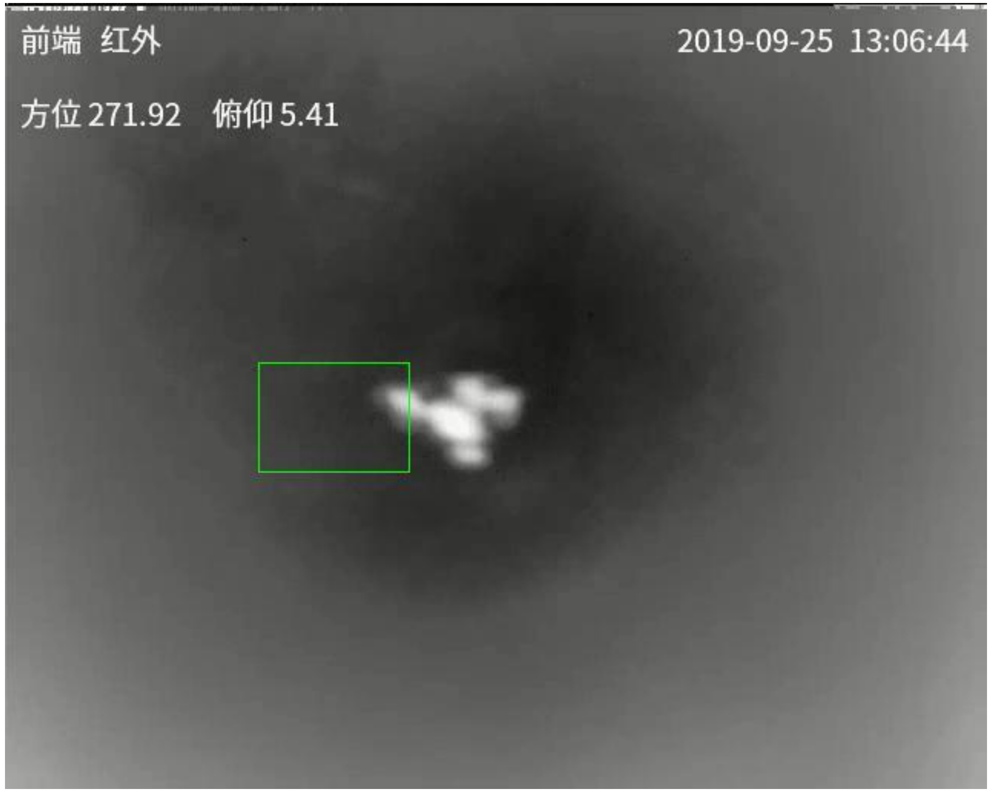
- Objectness Threshold = 0.5

- Hit Rate: 97.5%
- False Alarm Per Min: 2.4
- Tracking Accuracy: 73.6%
- Modified Tracking Accuracy: 73.5%

Performance of YOLOv3 on Anti-UAV Dataset

- Target detection and tracking results are quite acceptable, but not accurate due to existing annotation errors in the dataset





Performance of YOLOv3 with Simulated Annotation Errors

1) Additional Boxes: An extra box which does not contain any target

- Two types of additional boxes:
  - Temporarily Consistent
    - Random additional box created in one frame
    - This box tracked in consecutive frames
  - Completely Random
    - New random additional box created in each frame with a probability p

METU

Performance of YOLOv3 with Simulated Annotation Errors

2) Missing Boxes: Unavailability of the annotation of a true object

- Two types of missing boxes:
  - Temporarily Consistent
    - Boxes are missing in consecutive frames
  - Completely Random
    - Each box can be a missing box with a probability p

Performance of YOLOv3 with Simulated Annotation Errors

3) Shifted Boxes: Slightly translated version of the true object box

- Gaussian noise (0, σ) added to the center of the each box.

- Sizes of the boxes stay same.

METU

Performance of YOLOv3 with Simulated Annotation Errors

	FA	HR	TA	MTA
Original Annotations	2.4	97.5	73.6	73.5
Additional Boxes (25%)	9.7	97.8	74.9	74.3
Additional Boxes (50%)	18.8	95.6	69.4	68.6
Tmp. Cons. Add. Box. (25%)	5.6	96.5	72.7	72.5
Missing Boxes (25%)	0.3	94.1	71.3	71.3
Tmp. Cons. Mss. Box. (25%)	1.0	83.2	62.5	62.4
Tmp. Cons. Mss. Box. (50%)	0.9	34.7	27.2	27.2
Shifted Boxes (σ = 1.5)	2.2	90.8	68.8	68.8
Shifted Boxes (σ = 10%)	1.1	29.9	23.3	23.3
Combined	2.3	71.2	54.2	54.2

When Objectness Threshold is fixed to 0.5

Results are on validation set, which does not have any simulated errors

Performance of YOLOv3 with Simulated Annotation Errors

	Thrs	HR	TA	MTA
Original Annotations	0.5	97.5	73.6	73.5
Additional Boxes (25%)	0.72	94.1	72.1	72
Additional Boxes (50%)	0.68	92.2	67	66.9
Tmp. Cons. Add. Box. (25%)	0.58	95.6	72	72
Missing Boxes (25%)	0.4	97.2	73.2	73.1
Tmp. Cons. Mss. Box. (25%)	0.38	90.8	67.9	67.8
Tmp. Cons. Mss. Box. (50%)	0.3	56	43.2	43.1
Shifted Boxes (σ = 1.5)	0.48	92.4	70	69.9
Shifted Boxes (σ = 10%)	0.3	87.8	64.8	64.7
Combined	0.49	72.8	55.4	55.3

When FA/min is fixed to 2.4

Results are on validation set, which does not have any simulated errors

Annotation Correction

Proposed solution:

- Conventional template matching

Assumptions:

- No missing and additional boxes
- Appearance of the object does not change significantly between consecutive frames

If the assumptions hold:

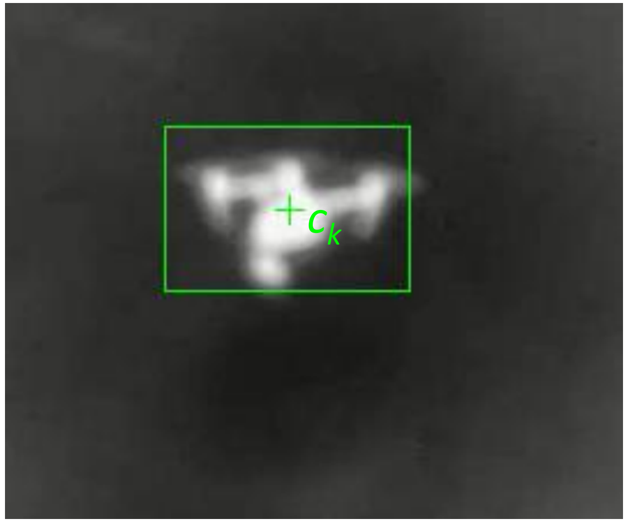
- Shifted box errors can be recovered using consecutive frames

METU

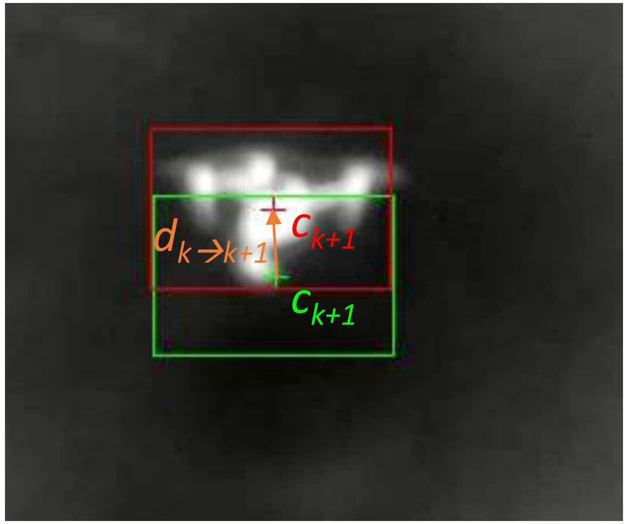
Annotation Correction

Let c_k be the object center at frame k

- Get the object template at frame k centered at c_k
- Search the template around c_{k+1} on frame $k+1$ using normalized cross correlation to find the displacement $d_{k \rightarrow k+1}$
- Find the displacement $d_{k+1 \rightarrow k}$
- If $d_{k \rightarrow k+1}$ and $d_{k+1 \rightarrow k}$ are consistent and correlation score is above selected threshold record the displacement, otherwise break the track
- If frame k+1 has no annotated objects, break the track



Frame k



Frame k+1

METU

Annotation Correction

For a track starting at frame k, corrected center at frame $k+n$

$$\hat{c}_{k+n} = \hat{c}_k + \sum_{t=k}^{k+n-1} d_{t \to t+1}$$

Assumption:

- Displacements should have zero mean

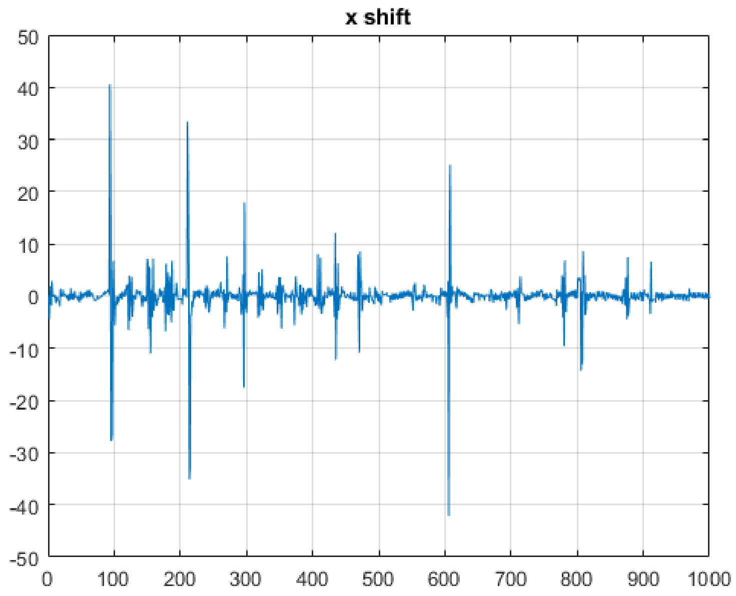
Problem:

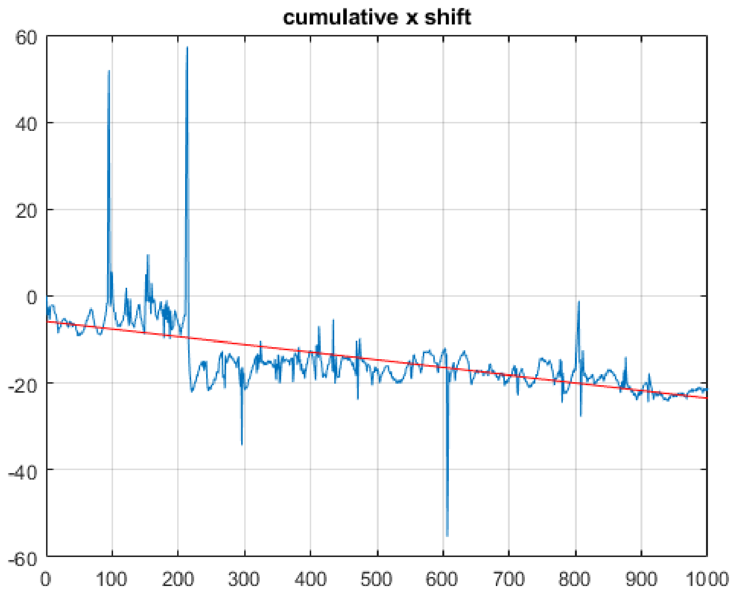
- Errors in $d_{k \to k+1}$ are accumulated

Solution:

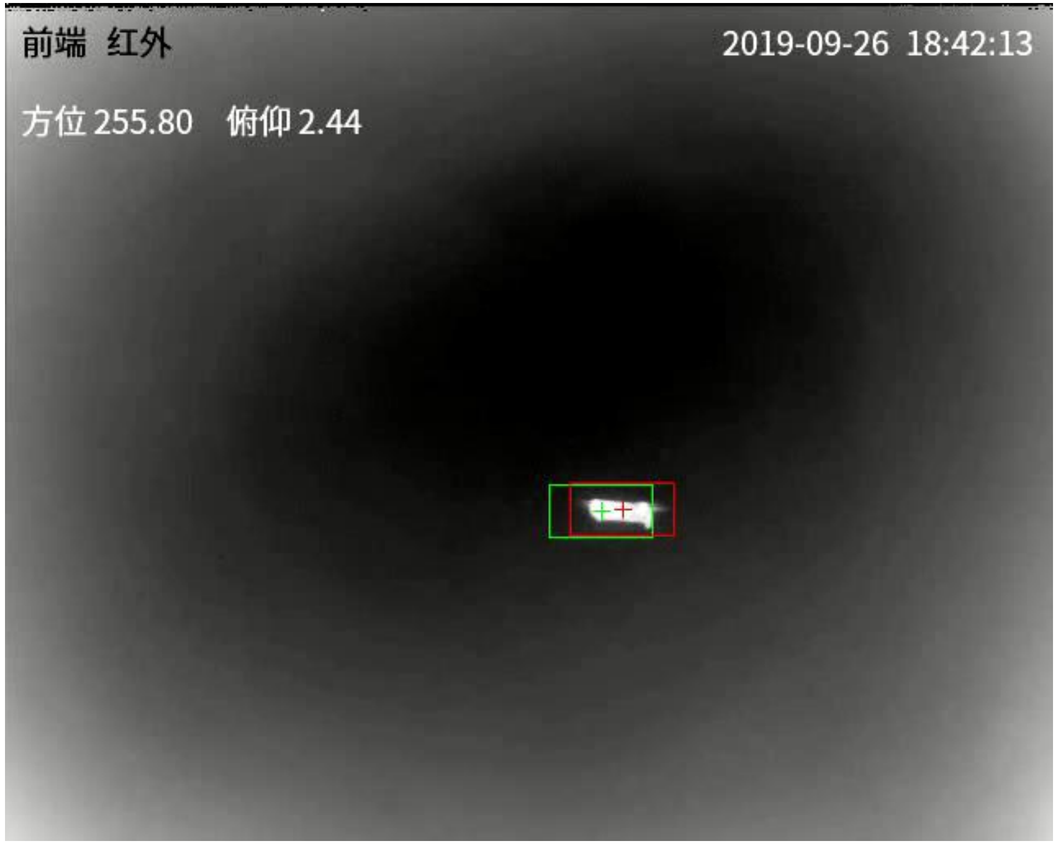
- Remove the trend of accumulative displacement

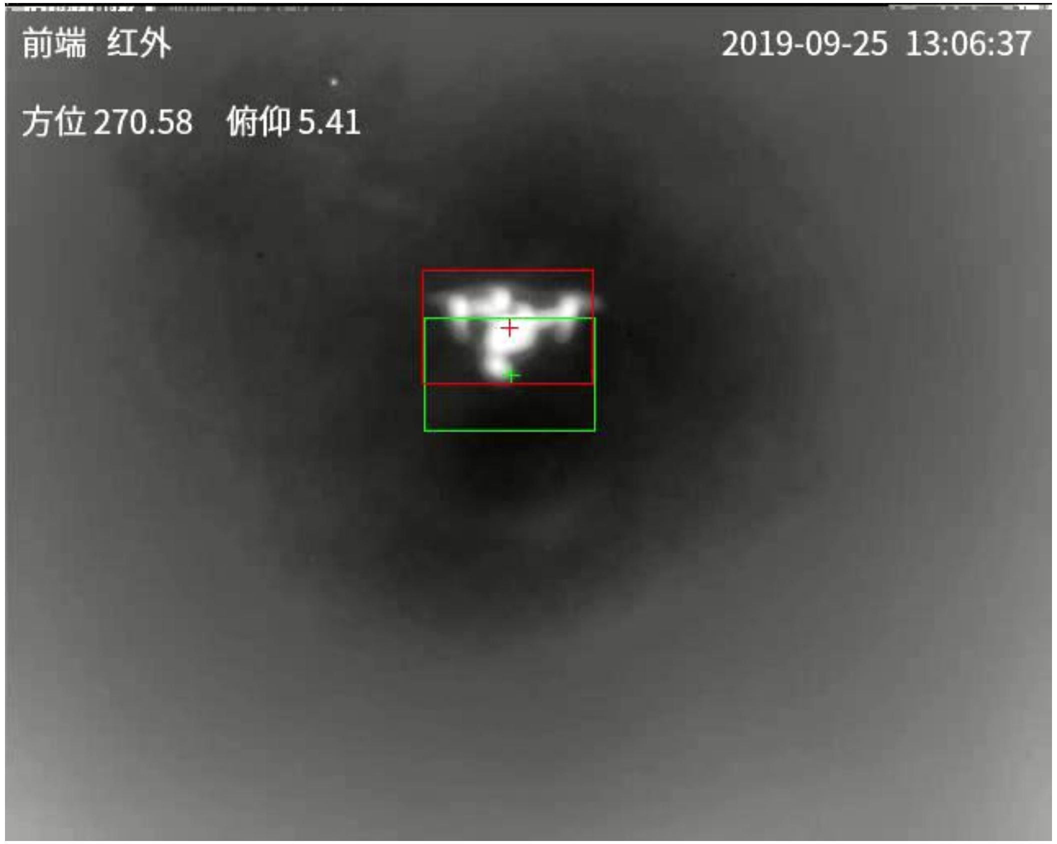
Annotation Correction





Annotation Correction





Annotation Correction

- Annotation correction applied 100 videos on Anti-UAV dataset
- Human subjects preferred corrected annotations on 66 videos
- 34 remaining videos are left with original annotations

	Thrs	HR	TA	MTA
Original Annotations	0.5	97.5	73.6	73.5
Corrected (Training Set)	0.55	98.0	74.8	74.7
Corrected (Training + Val Set)	0.55	98.8	76.3	76.2

When FA/min is fixed to 2.4

METU

Conclusion

- Performance of YOLOv3 on Anti-UAV dataset is relatively well.
  - Good in terms of hit and false alarm rates, not great in terms of tracking accuracy.
- Annotation errors degrade performance
- Shifted box errors are corrected with a semi-automated method
  - Results got better.
- Calculating IoU between corrected annotations and original annotations, we get 86.4% tracking accuracy
  - An algorithm which has a generalization ability can reach a maximum of 86.4% tracking accuracy (without memorizing)

Thank You!