THE INTEGRATION OF THERMAL AND VISUAL IMAGES FOR HUMAN DETECTION USING DEEP LEARNING

GFM M.Sc. Final Defence

Qiao REN

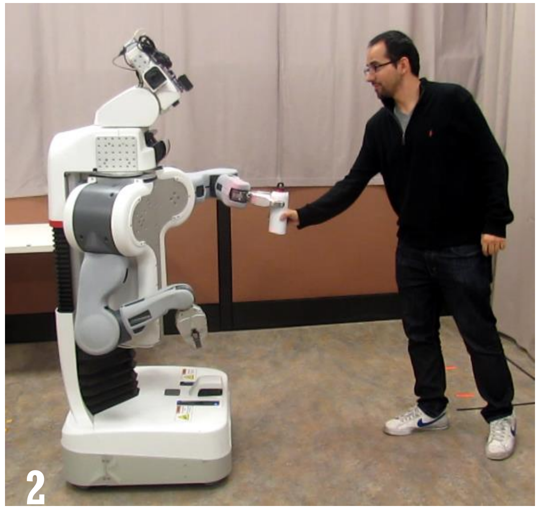
First Supervisor: Dr. Siavash Hosseinyalamdary

Second Supervisor: Dr. Francesco Nex

Mar 6, 2019

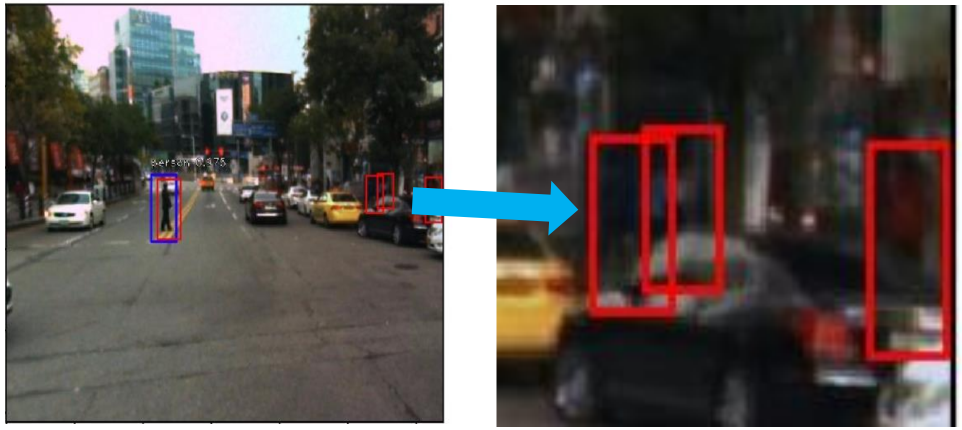
WHY HUMAN DETECTION IS IMPORTANT

- disasters management
- autonomous driving systems
- automated surveillance
- human-robotics interaction

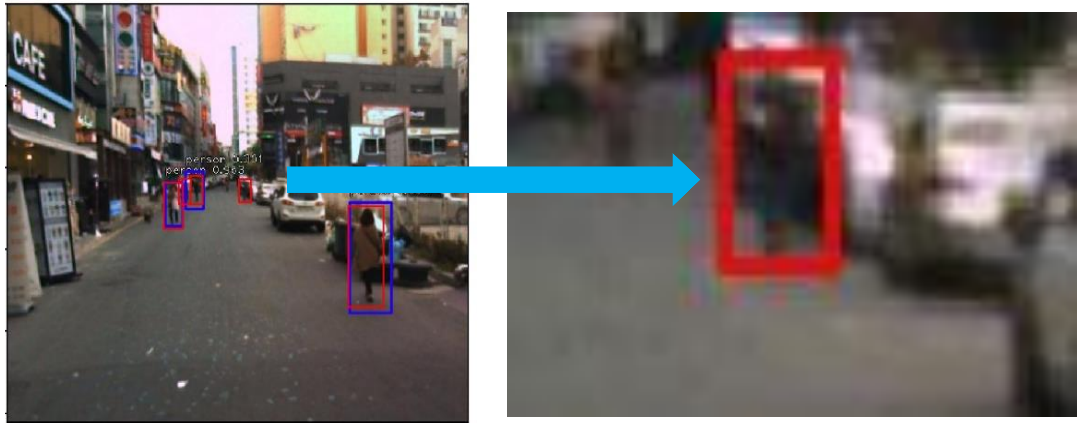


PROBLEM STATEMENT - CHALLENGES

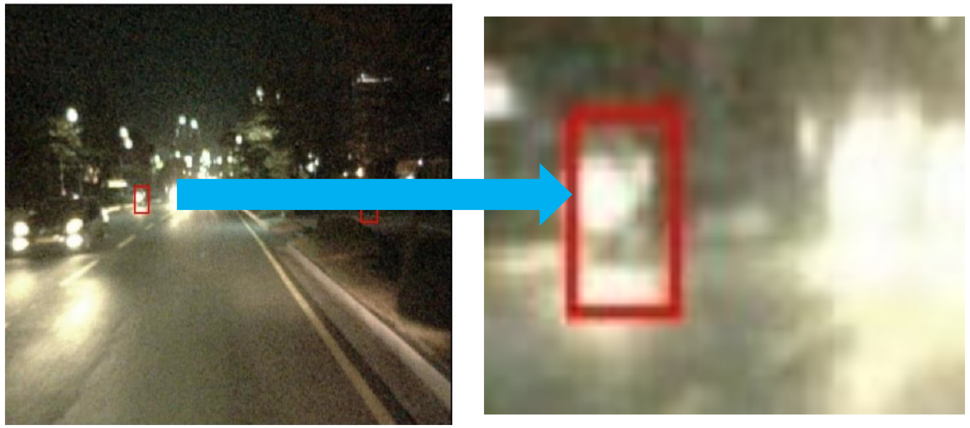
Under-exposure



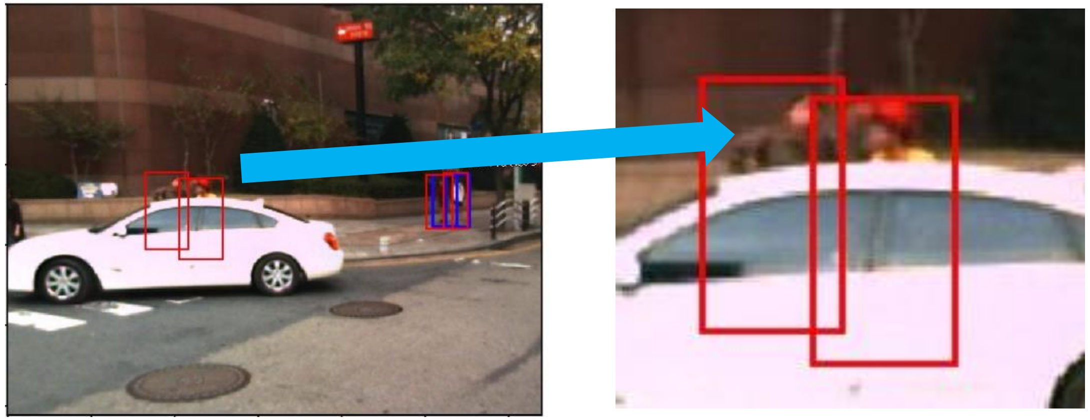
Persons are too far



Over-exposure

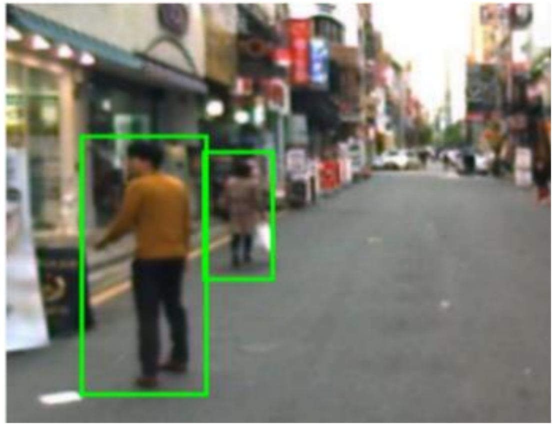


Occlusion



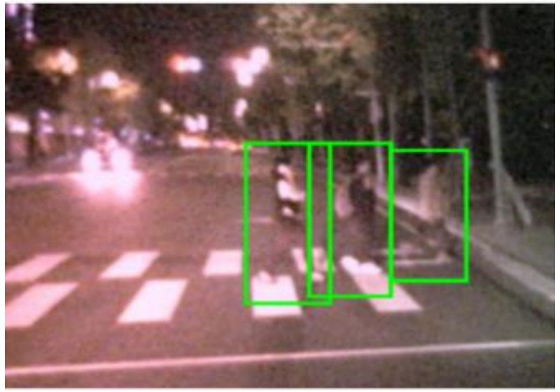
WHY TO INTEGRATE RGB AND THERMAL IMAGES

- RGB: sensitive to illumination
- T: temperature difference is small in summer, low resolution

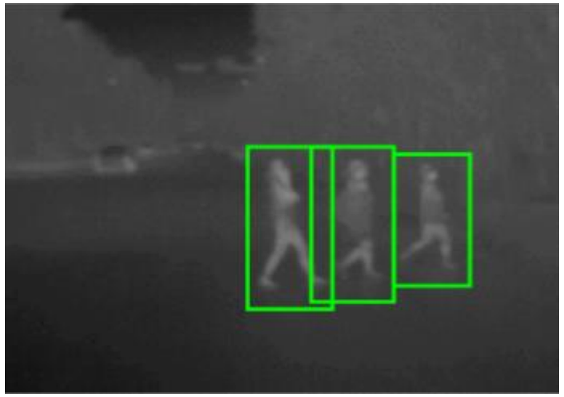


(a)

(b)



(c)



(d)

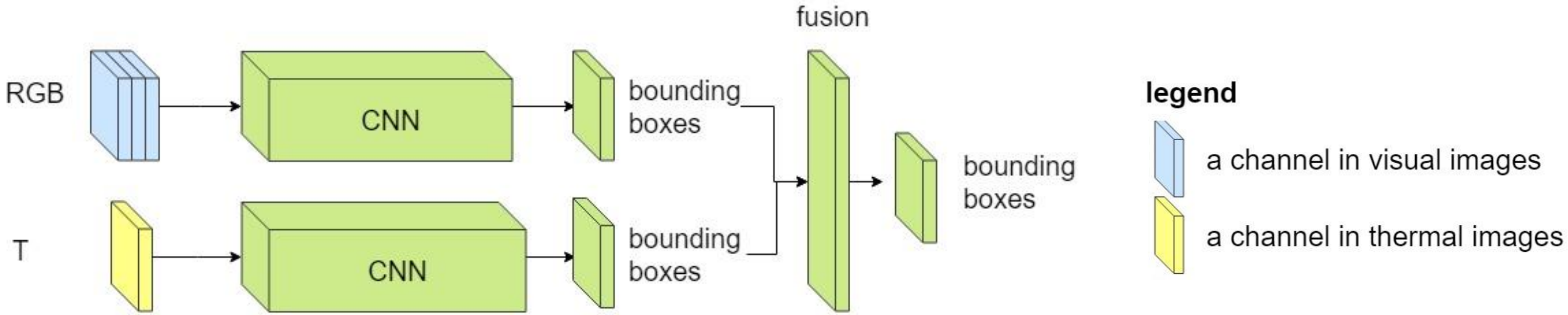
RESEARCH OBJECTIVE

- Test: different sensor fusion architectures
- Approach: convolutional neural network (CNN)
- Applied in: human detection
- Input: thermal and visual images
- Provide: the most accurate architecture

METHODOLOGY

a) Early Fusion (Pixel-Level Fusion)

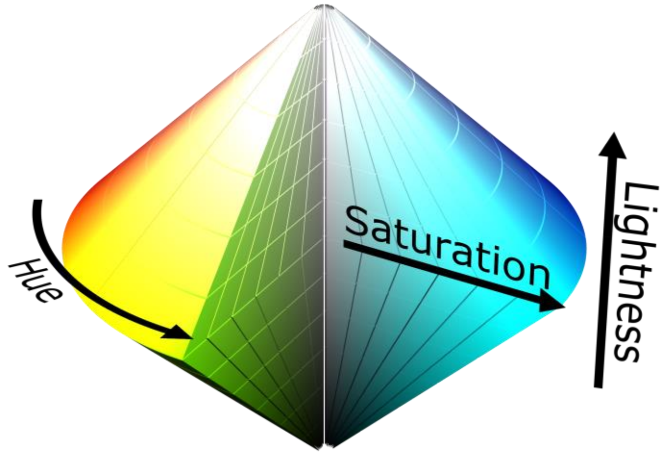
b) Late Fusion (Decision-Level Fusion)



c) kaistT
d) kaistRGB
e) cocokittikaistT
d) cocokittikaistRGB

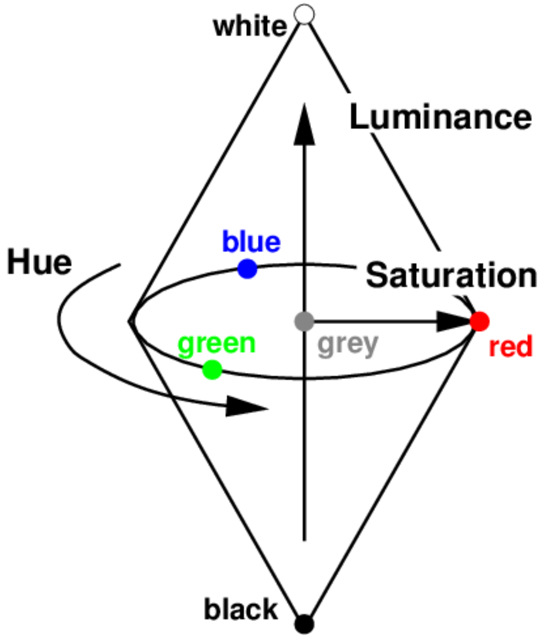
EARLY FUSION: WHY REPLACE SATURATION BY THERMAL

- definition
  - Hue: which color
  - Lightness or luminance: amount of reflection
  - Saturation: purity
- Replace S by T: because S is the least informative



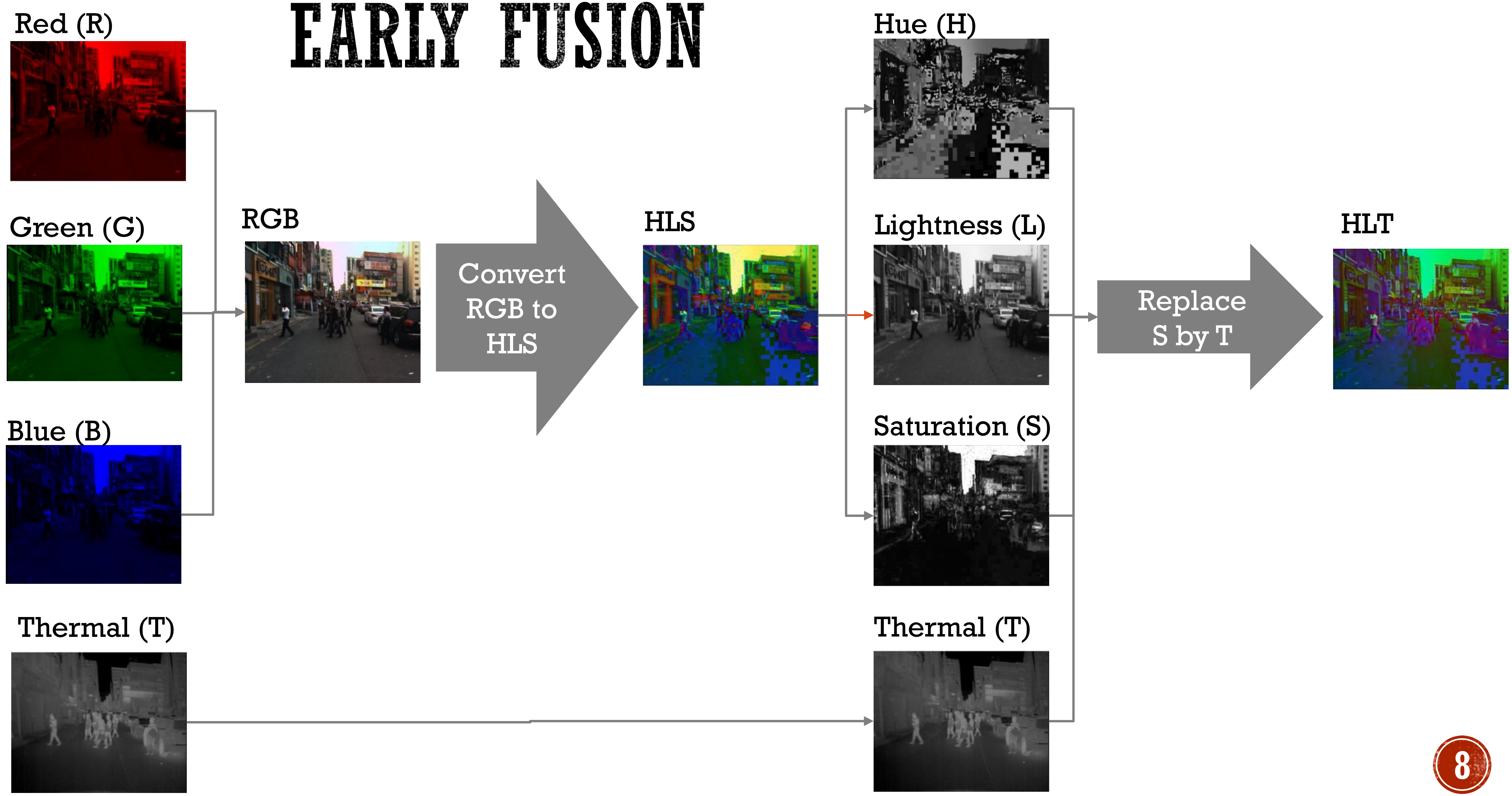
- Reasons why early is the worst:

1)Information loss: from RGB to HLS, replace S by T
2)underfitting



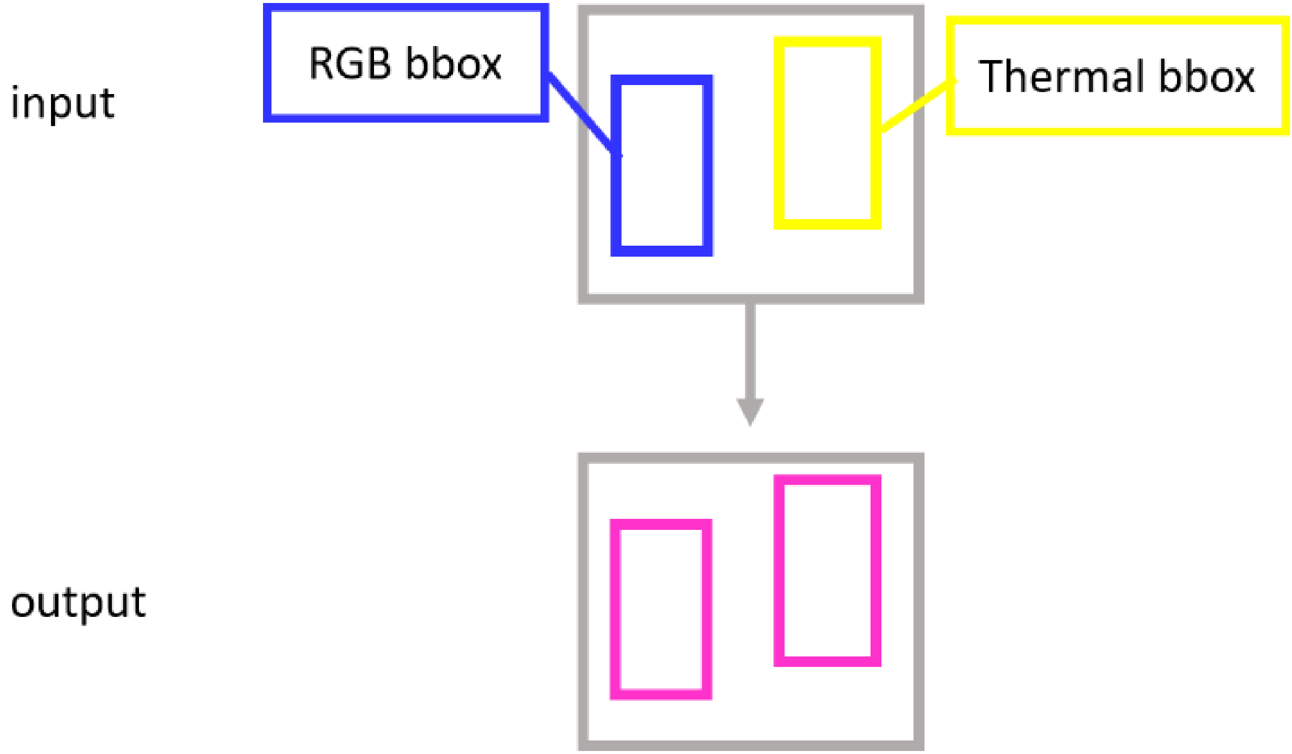
- Improvement:
  1) Change the code of RetinaNet to make it accept 4 bands
  2) Choose a model which accept 4 bands
  3) Remove a band from RGB, not from HSL

EARLY FUSION



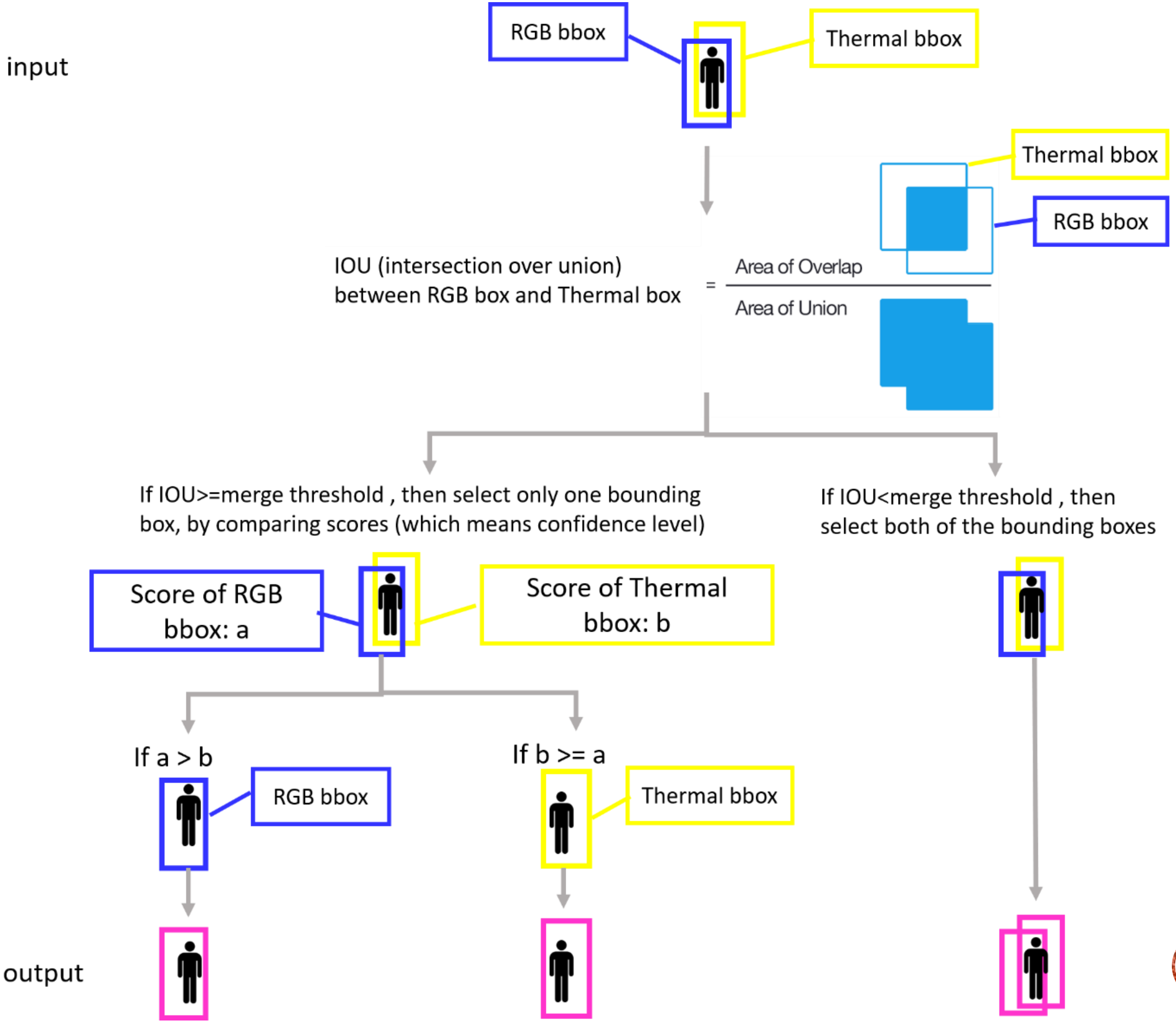
LATE FUSION

Case 1

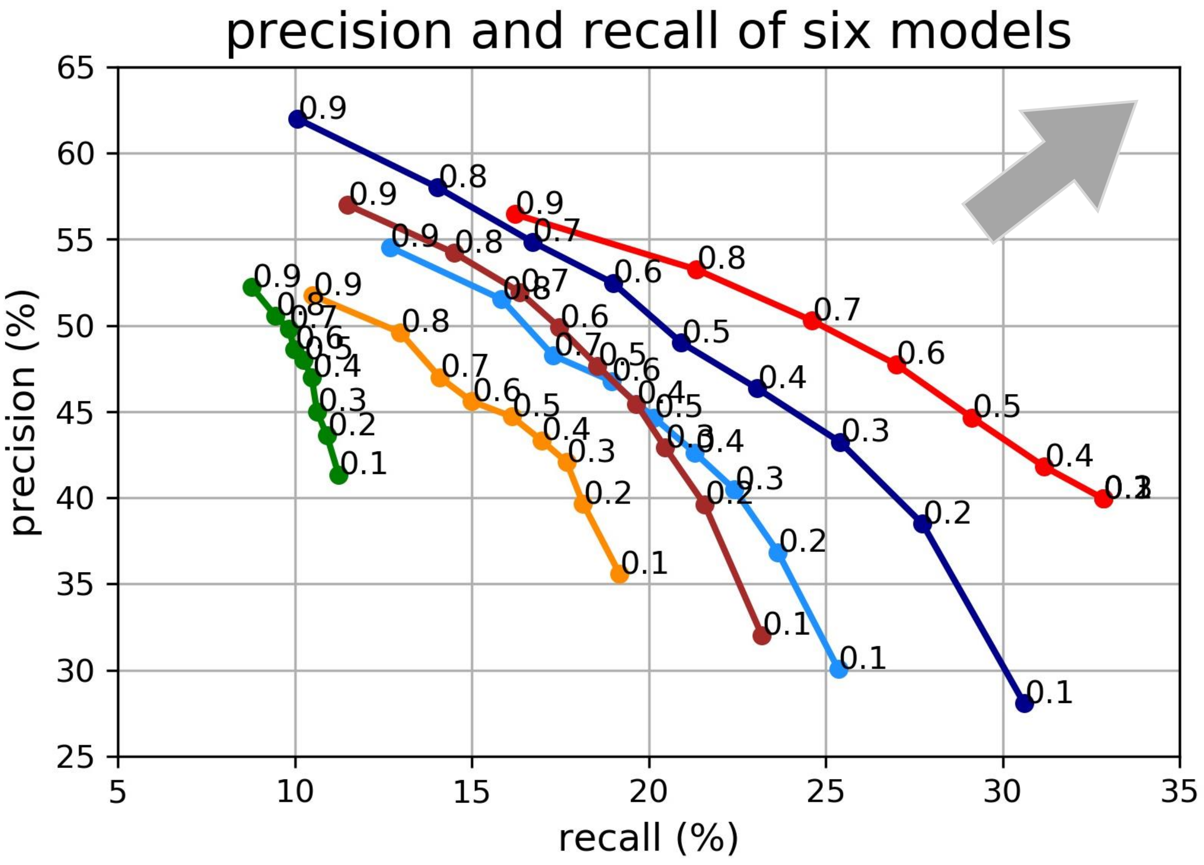


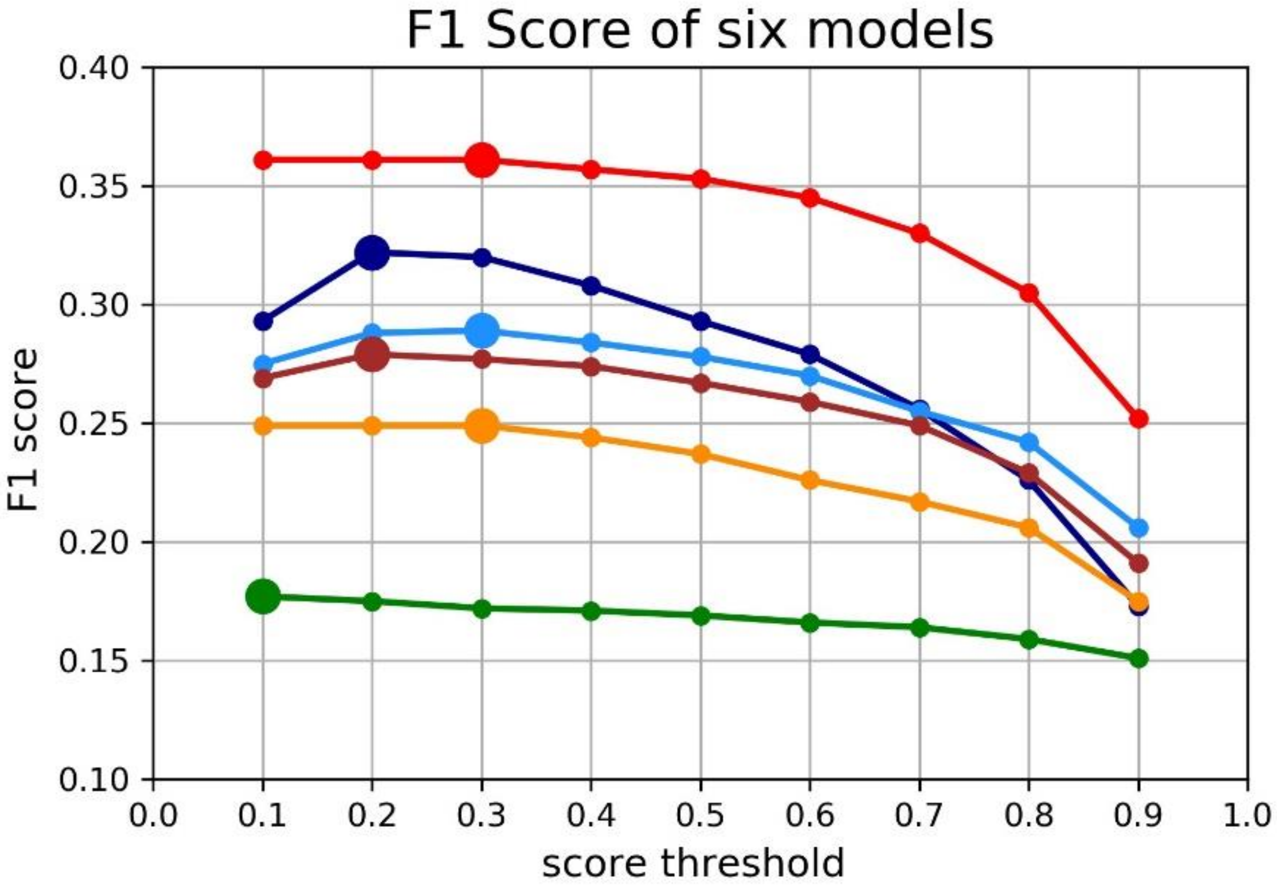
LATE FUSION

Case 2

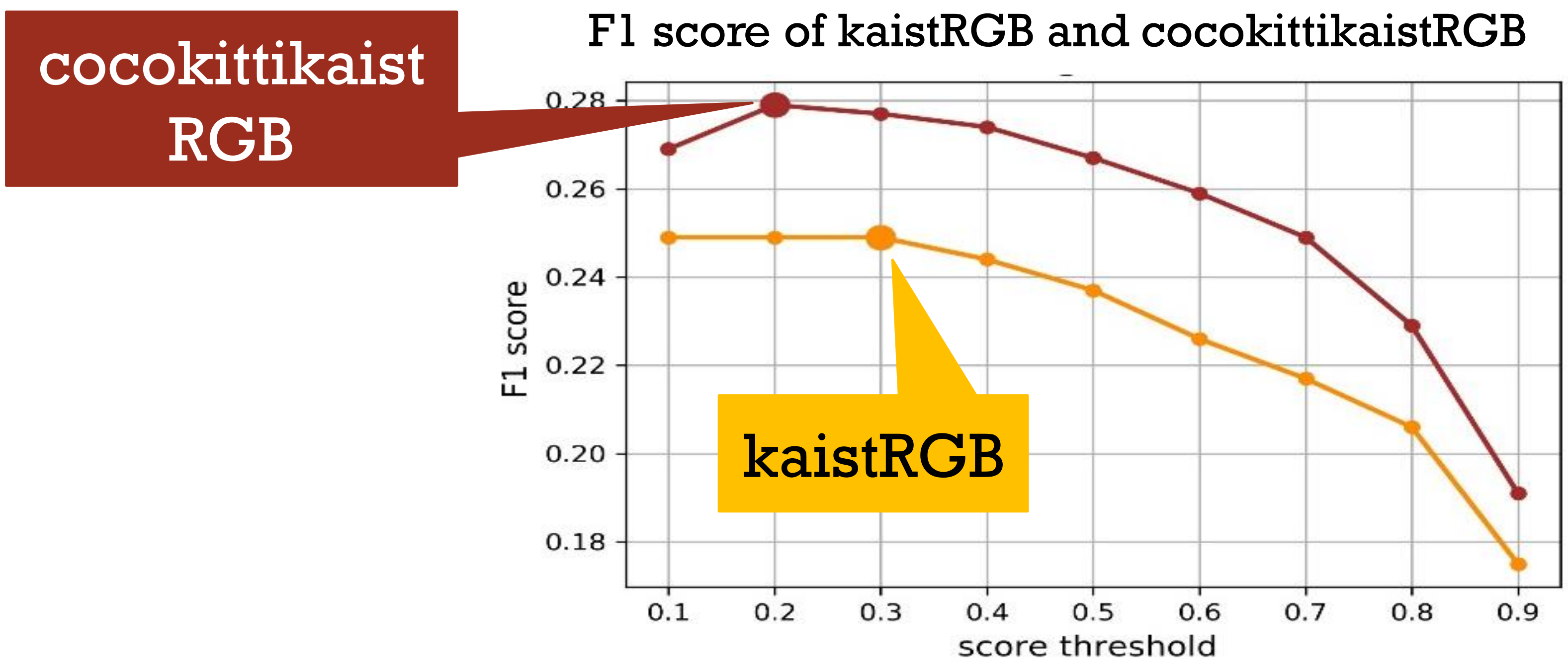


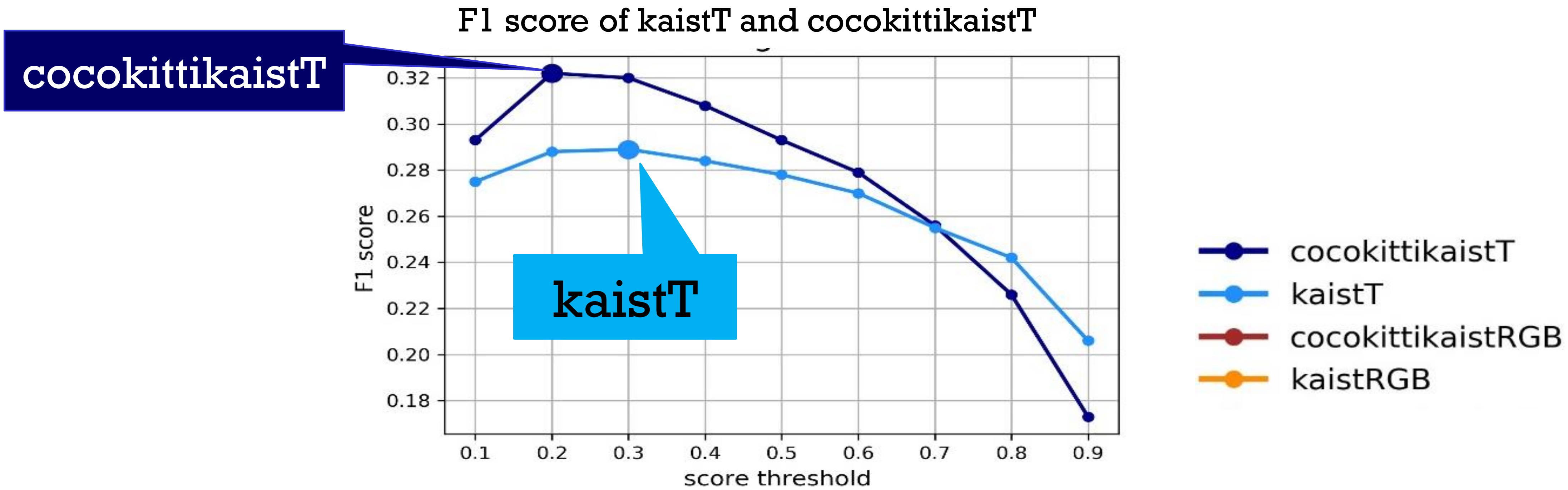
RESULT

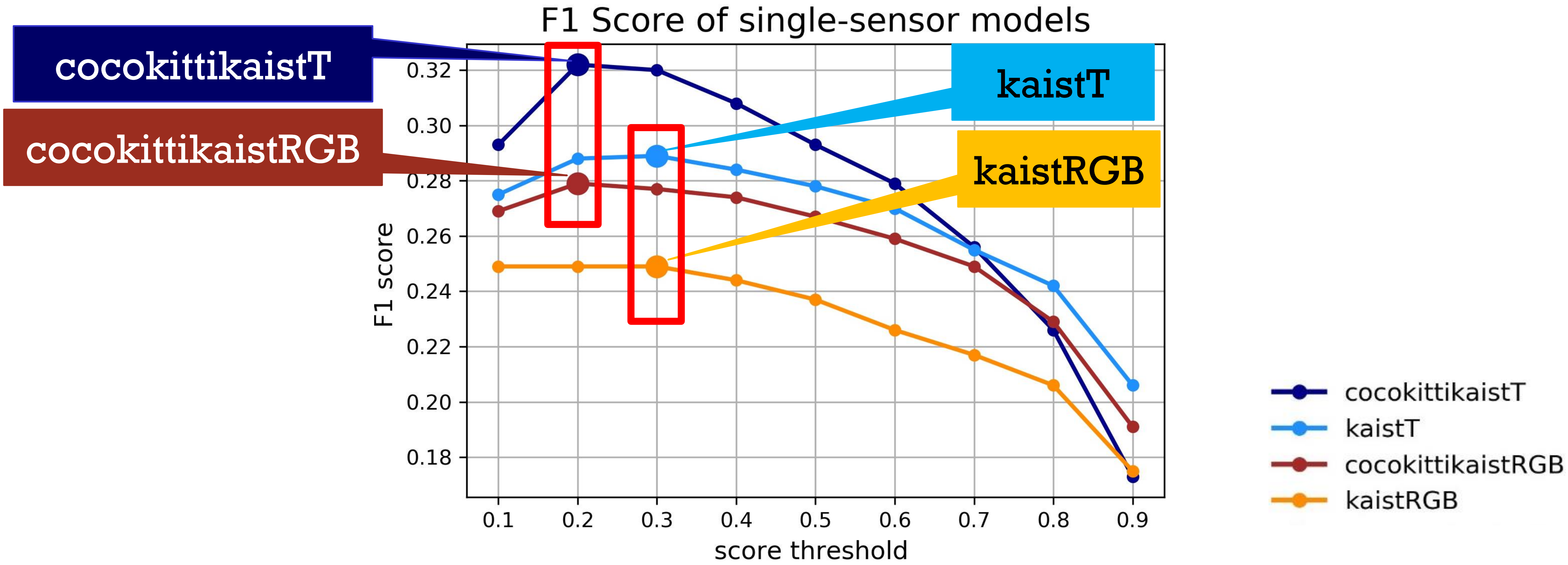

precision and recall of six models
precision (%)
recall (%)


F1 Score of six models
F1 score
score threshold

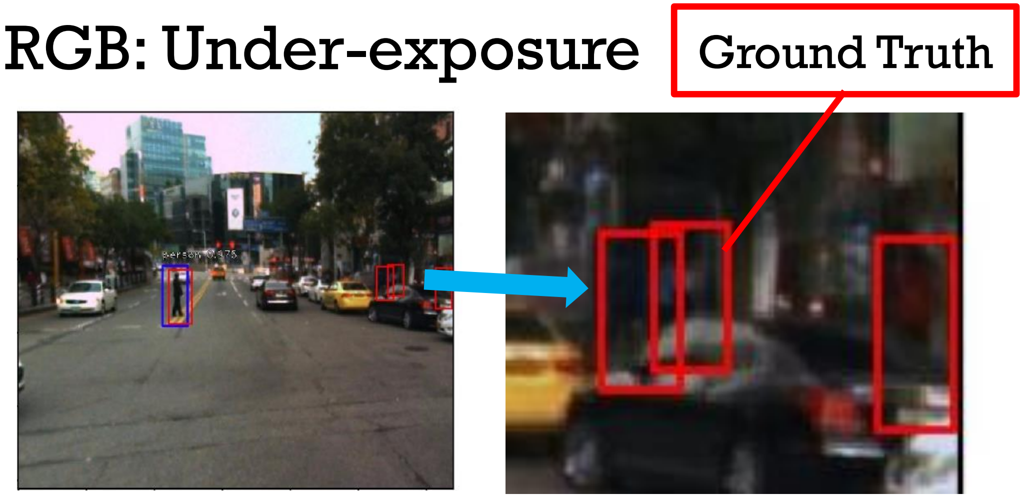

cocokittikaistT
kaistT
cocokittikaistRGB
kaistRGB
earlyFusionHLT
lateFusion


F1 score of kaistRGB and cocokittikaistRGB
cocokittikaist RGB
kaistRGB
0.28
0.26
0.24
0.22
0.20
0.18
F1 score
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
score threshold

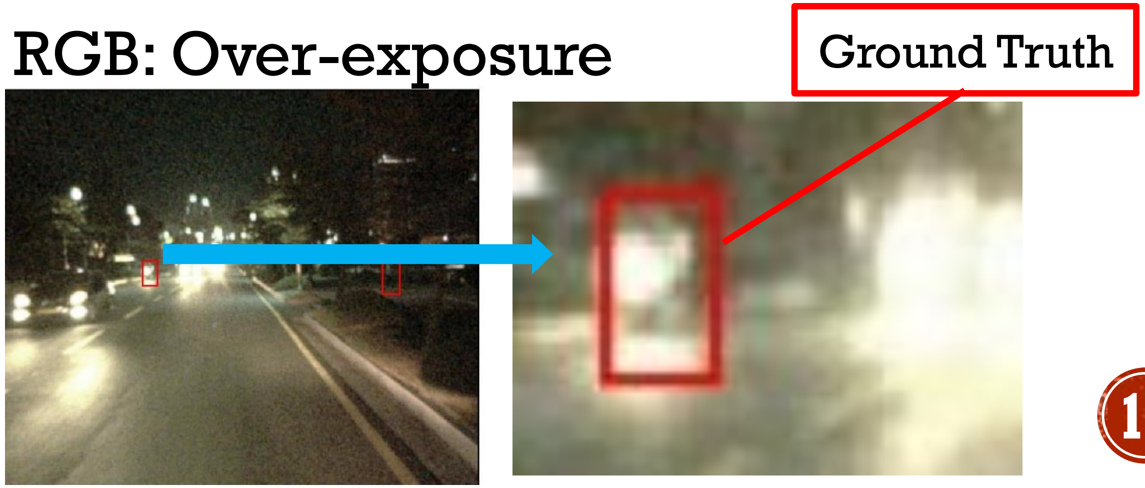

F1 score of kaistT and cocokittikaistT
cocokittikaistT
kaistT
0.32
0.30
0.28
0.26
0.24
0.22
0.20
0.18
F1 score
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
score threshold
cocokittikaistT
kaistT
cocokittikaistRGB
kaistRGB

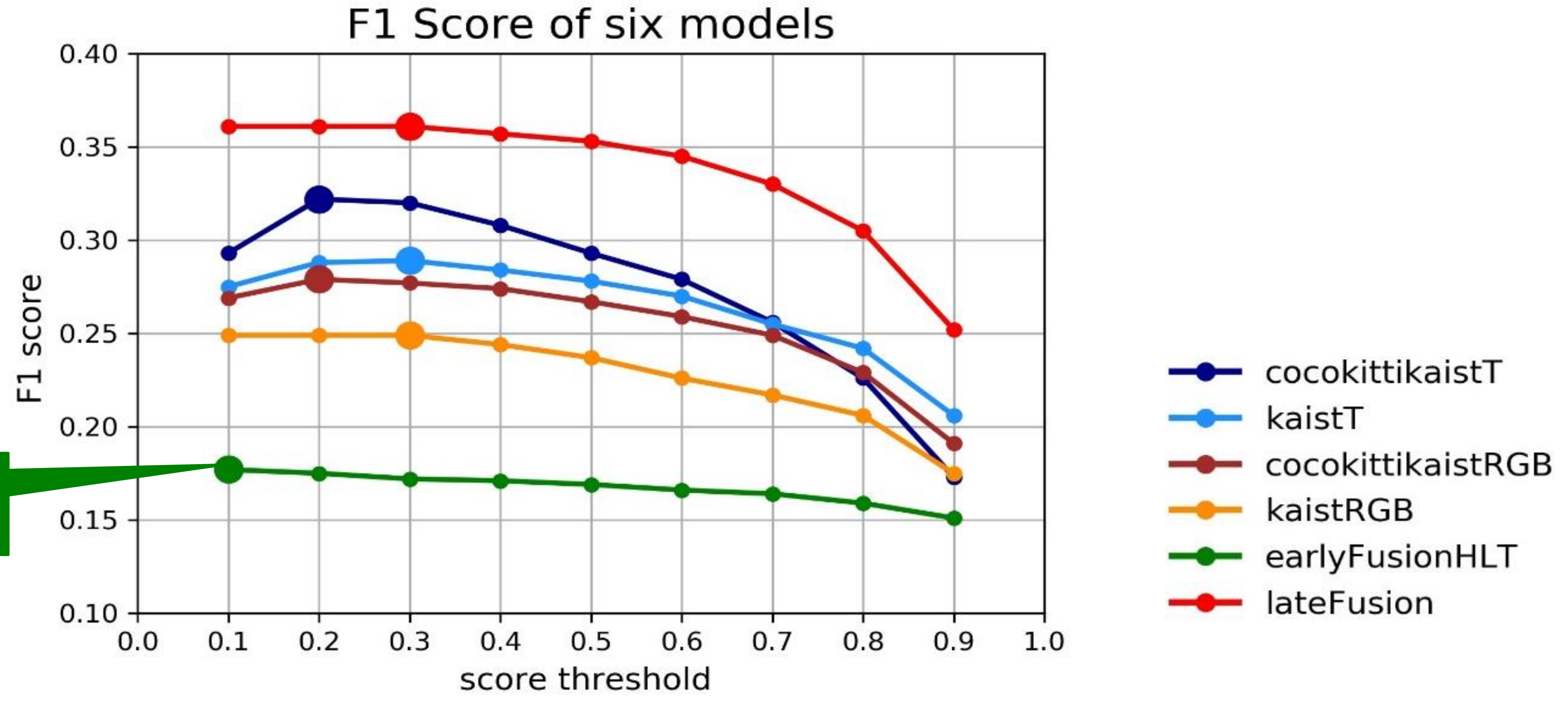


RGB: Under-exposure



RGB: Over-exposure





Early fusion

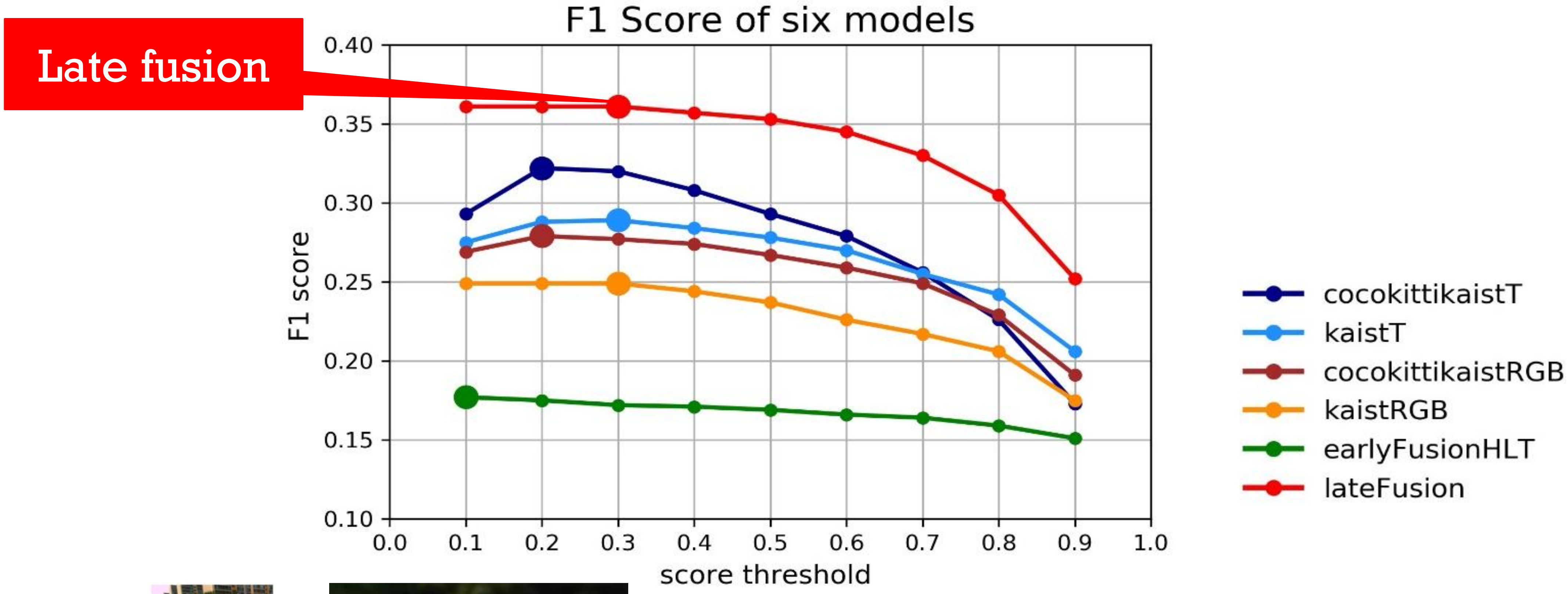
RGB

T

Replace saturation by thermal

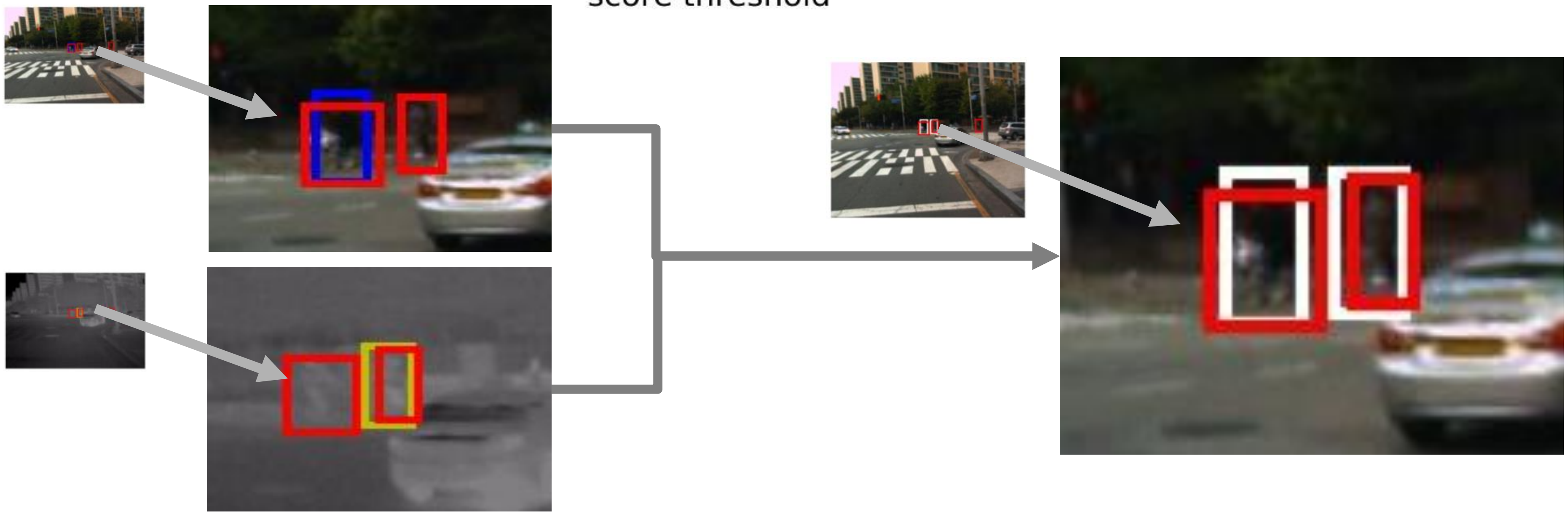
Early fusion: HLT

Ground Truth



cocokittikaistRGB: blue

cocokittikaistT: yellow

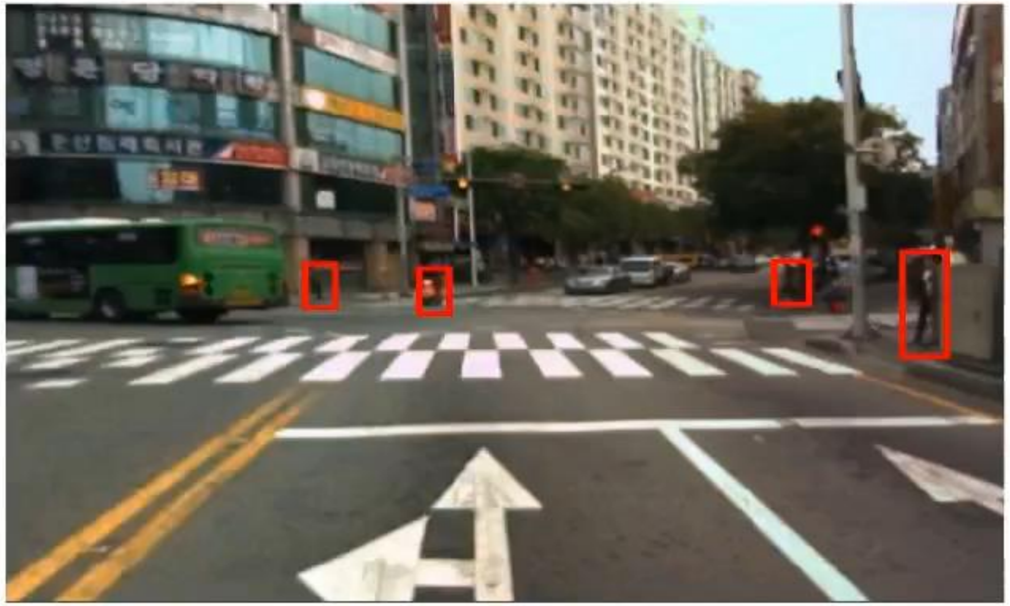


Late Fusion: white

CONCLUSION

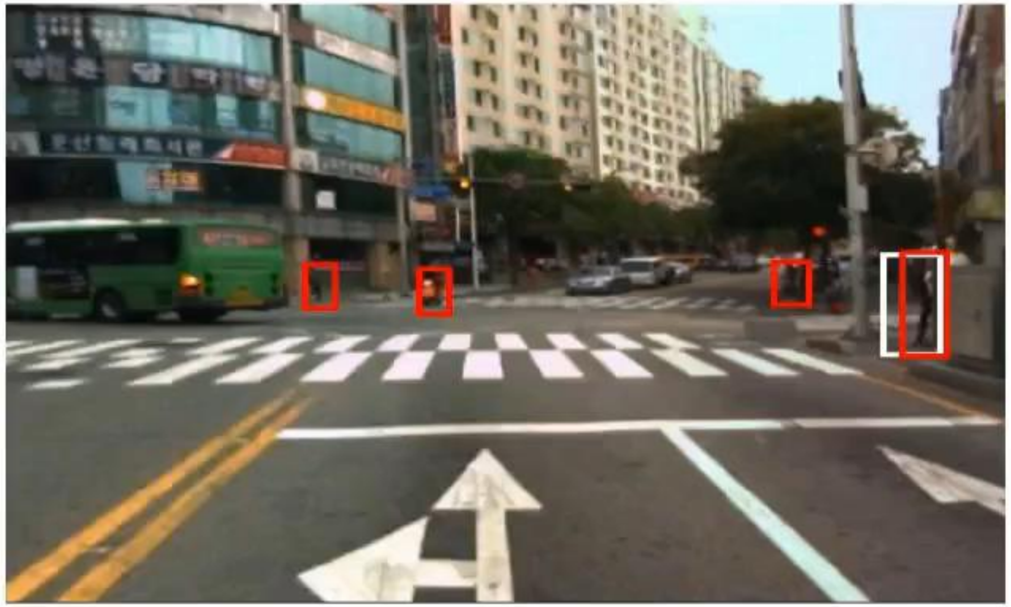
- **Late fusion has the best performance**:
  - 8.2% better than cocokittikaistRGB
  - 3.9% better than cocokittikaistT
  - 11.2% better than kaistRGB
  - 7.2% better than kaistT
- **Early fusion has the worst performance**
- **Fine-tuned models has better performance than non-fine-tuned models**
  - cocokittikaistRGB is 3.0% better than kaistRGB
  - cocokittikaistT is 3.3% better than kaistT
- In kaist dataset, **models trained on thermal images has better performance than trained on RGB images**
  - kaistT is 4.0% better than kaistRGB
  - cocokittikaistT is 4.3% better than cocokittikaistRGB

cocokittikaistRGB: Blue bounding boxes



cocokittikaistT: Yellow bounding boxes

Late fusion: White bounding boxes. Shown on RGB images



Late fusion: White bounding boxes. Shown on T images