Improving Multispectral Pedestrian Detection by Addressing Modality Imbalance Problems

Computational Imaging Lab @ Nanjing University

http://cite.nju.edu.cn

Kailai Zhou, Linsen Chen, Xun Cao

{calayzhou, linsen}@smail.nju.edu.cn caoxun@nju.edu.cn

How to detect pedestrians at night?

traditional pedestrian detection dataset

Caltech, CityPersons, INRIA…

pedestrian detection dataset at night

NightOwls[2]

The performance of SDS RCNN[1] (Miss Rate)

Train \ Test	Caltech	NightOwls
Caltech	7.36%	63.99% ↑

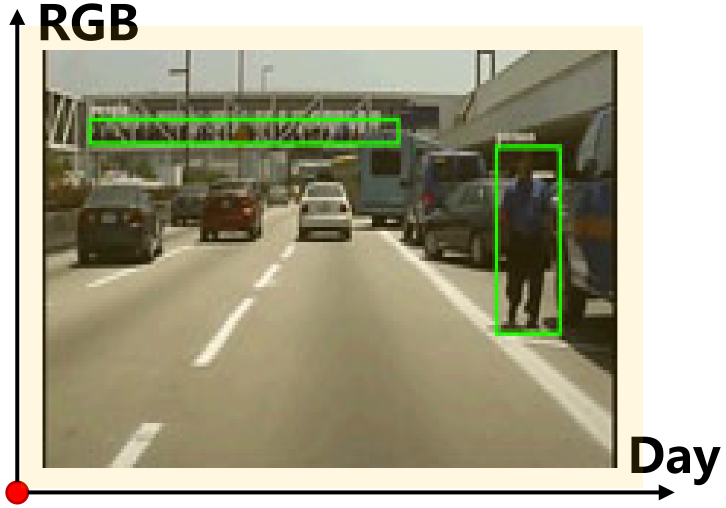
The performance of pedestrian detection model suffers at night!

SOLUTION: multispectral (RGB+thermal)

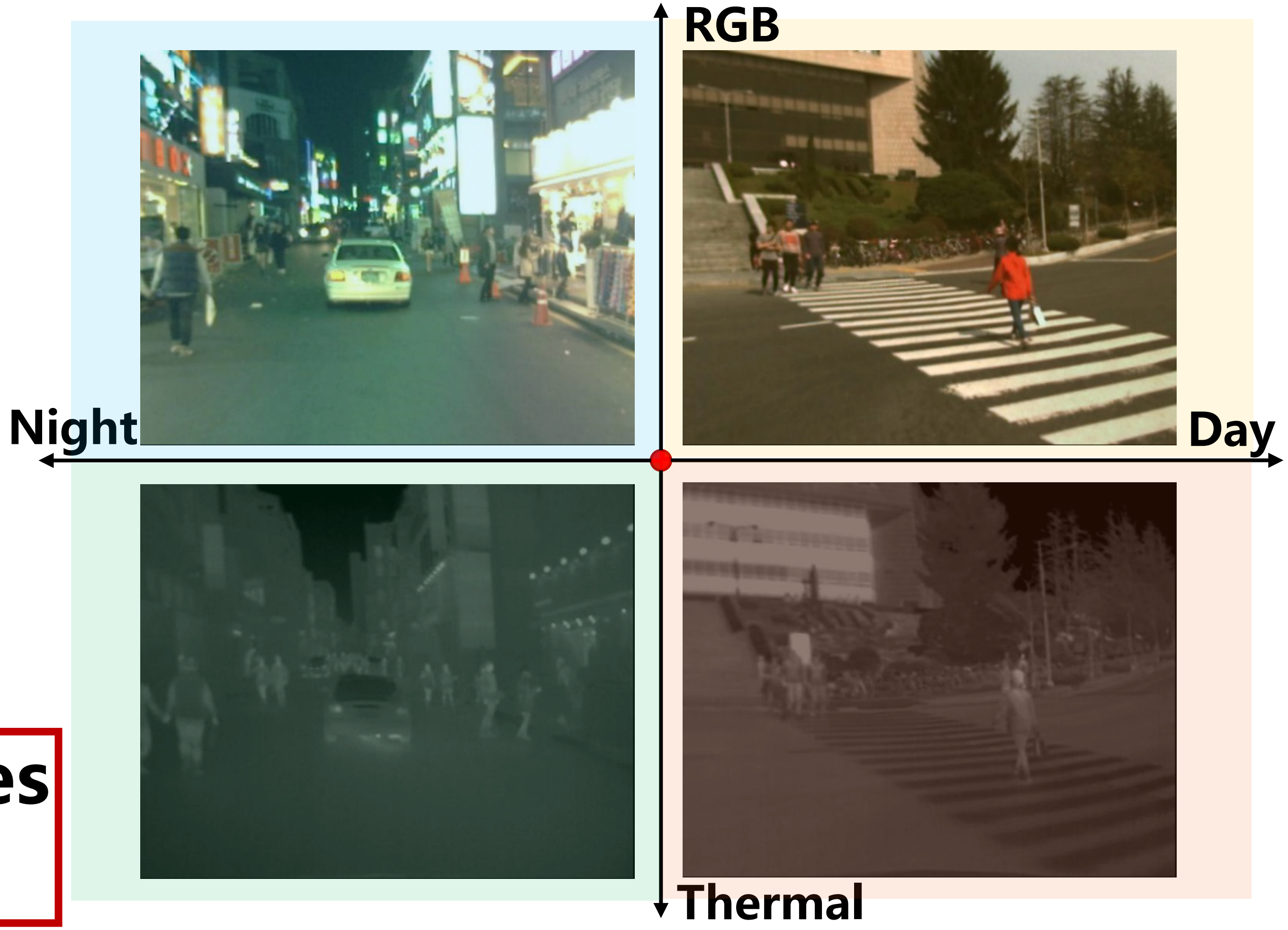
[1] Garrick Brazil et al. Illuminating Pedestrians via Simultaneous Detection & Segmentation. ICCV2017
[2] Lukáš Neumann et al. NightOwls: A Pedestrians at Night Dataset. ACCV2018

Multispectral Pedestrian Detection: RGB + Thermal

traditional pedestrian detection dataset



KAIST [1]



one modality → four modalities

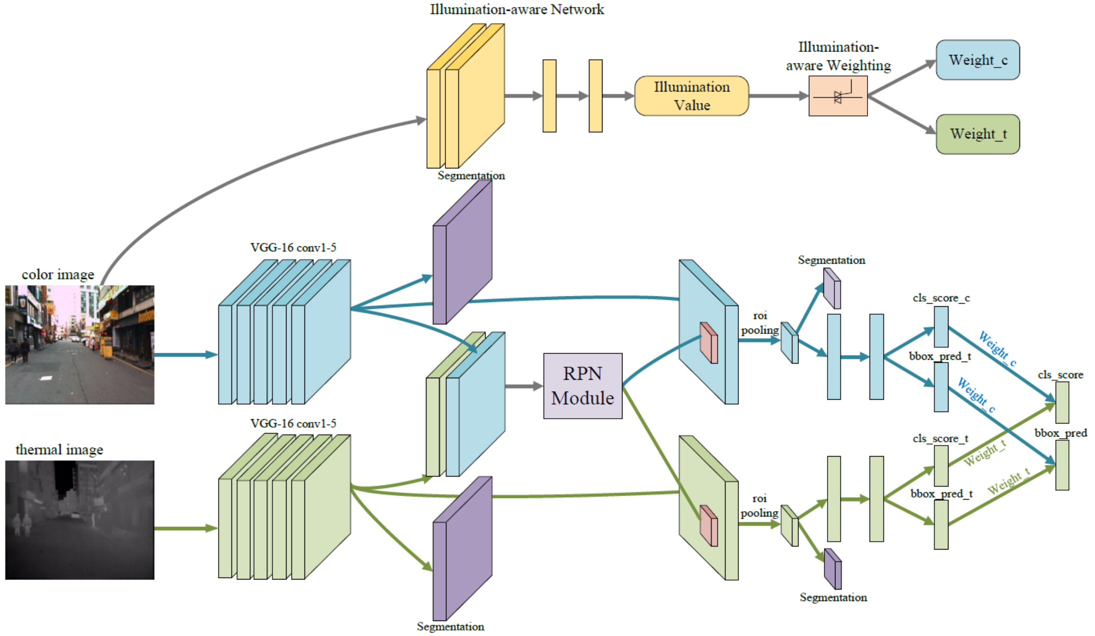
- RGB & Thermal
- Night & Day

[1] Soonmin Hwang et al. Multispectral Pedestrian Detection: Benchmark Dataset and Baseline. CVPR2015

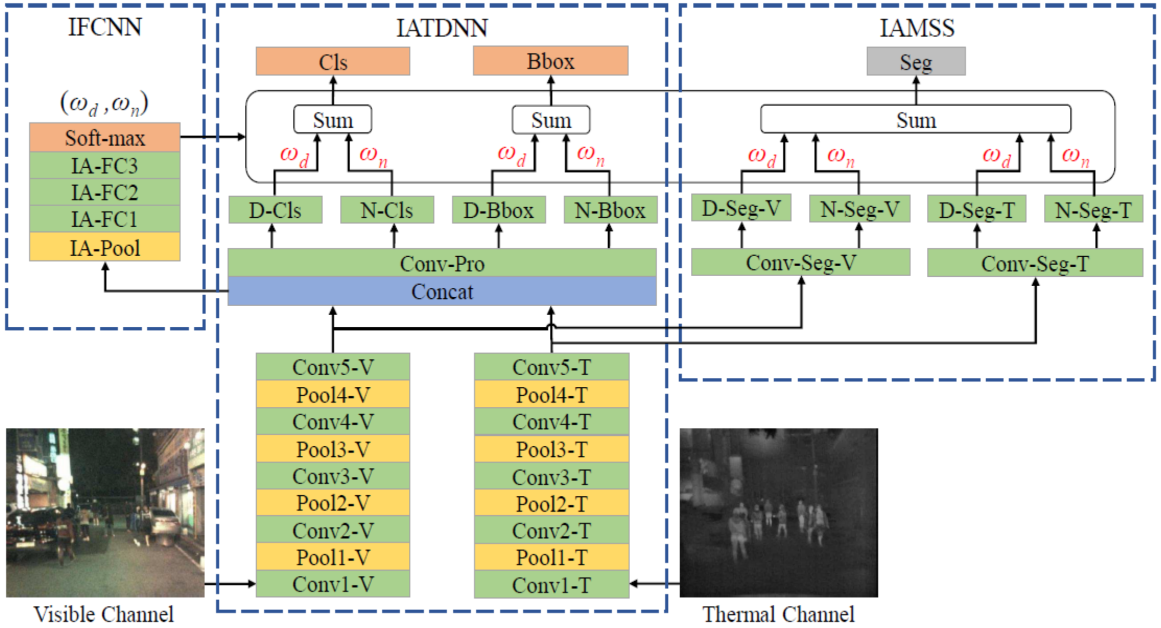
Review of Existing Works

- How to adapt to different illumination conditions
- How to fuse the RGB and thermal features

IAF−RCNN [1]



IATDNN−IAMSS [2]



Illumination − aware score fusion

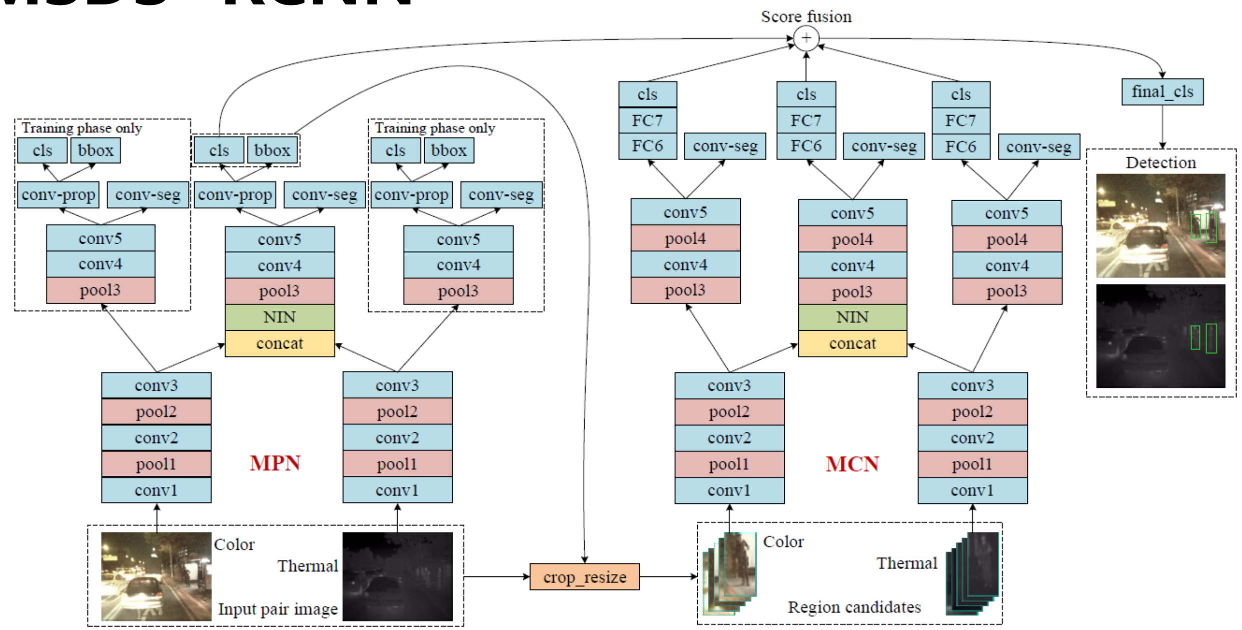
[1] Chengyang Li et al. Illumination-aware Faster R-CNN for Robust Multispectral Pedestrian. BMVC2018

[2] Dayan Guan et al. Fusion of Multispectral Data Through Illumination-aware Deep Neural Networks for Pedestrian Detection[J]. Information Fusion, 2019: 148-157.

Review of Existing Works

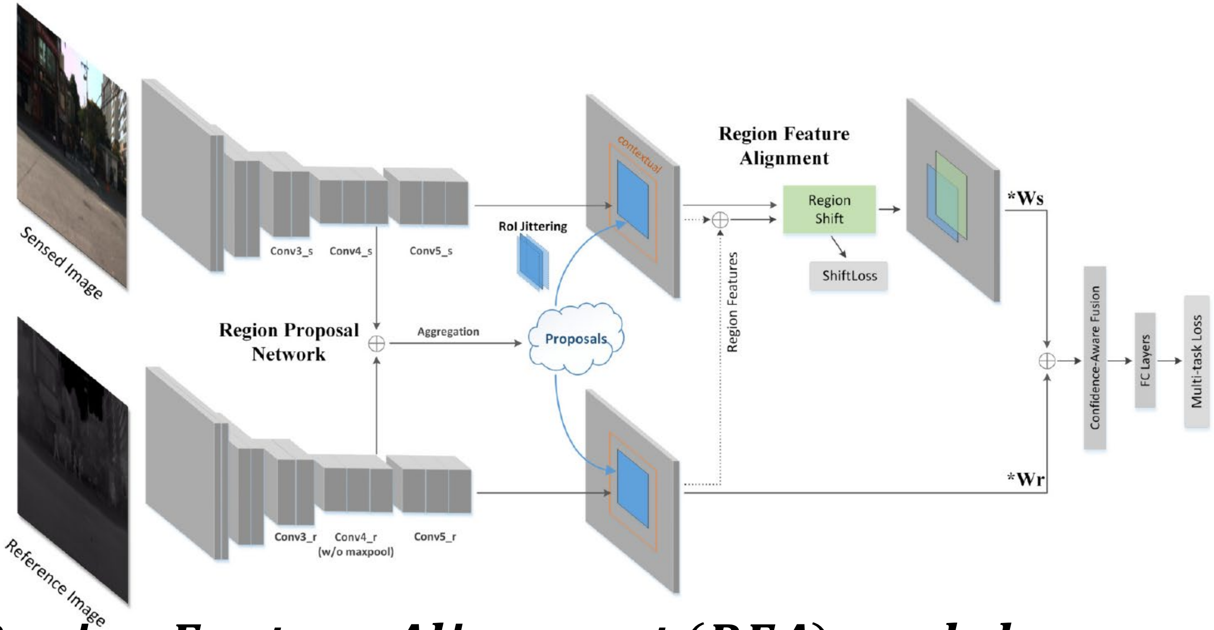
- How to adapt to different illumination conditions
- How to fuse the RGB and thermal features

MSDS−RCNN [1]



Concatenation

AR−CNN [2]

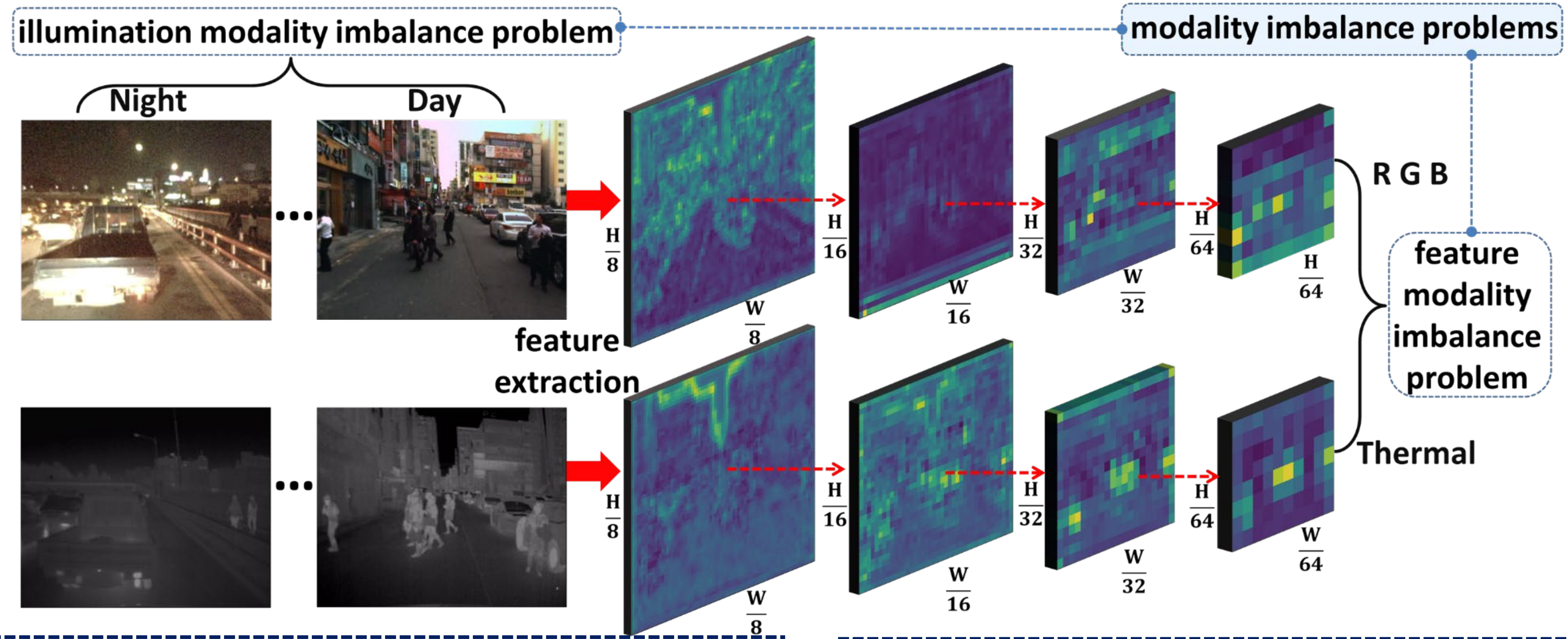


Region Feature Alignment (RFA) module

[1] Chengyang Li et al. Multispectral Pedestrian Detection via Simultaneous Detection and Segmentation. BMVC2018
[2] Lu Zhang et al et al. Weakly Aligned Cross-Modal Learning for Multispectral Pedestrian Detection. ICCV2019

Modality Imbalance Problems

- How to adapt to different illumination conditions
- How to fuse the RGB and thermal features



illumination modality imbalance

- Under different illumination conditions

Day&Night modalities contribute out-off-balance

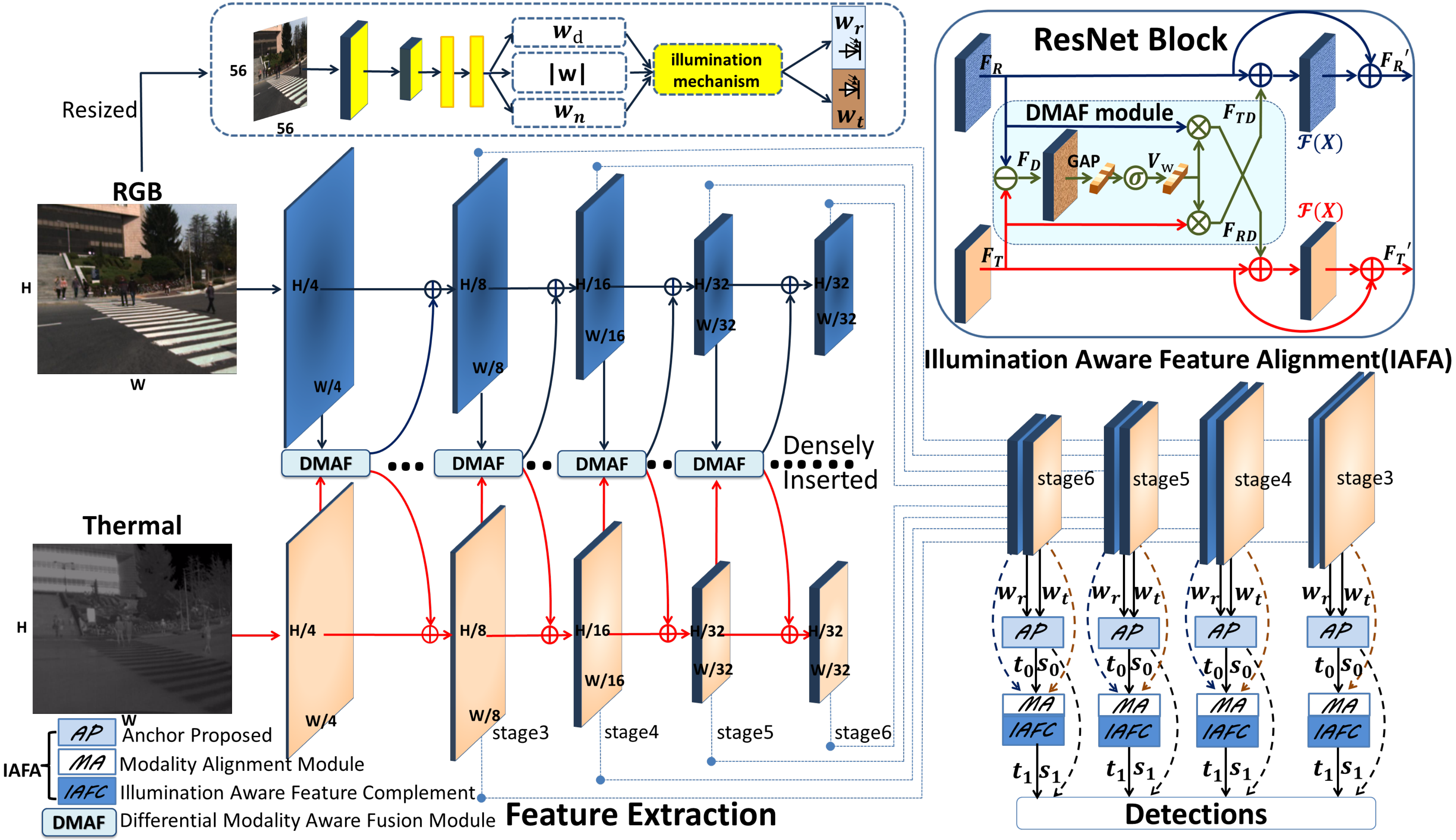
feature modality imbalance

- Feature Misalignment
- Inadequate Fusion

RGB&Thermal modalities contribute out-off-balance

Our Works

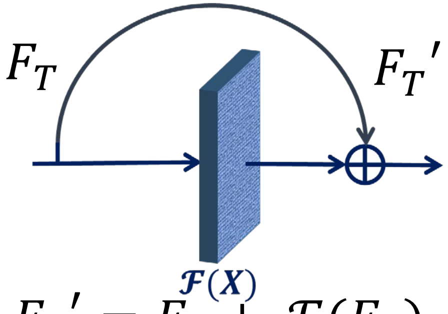
Modality Balance Network (MBNet)



Differential Modality Aware Fusion (DMAF) Module

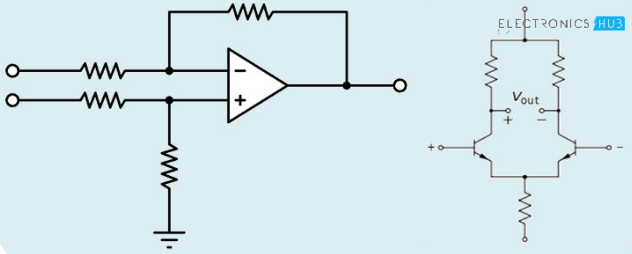
- Differential Amplifier: **amplify** differential signals & **suppress** common signals
- DMAF module: **compensate** differential features & **retain** original features

ResNet Block

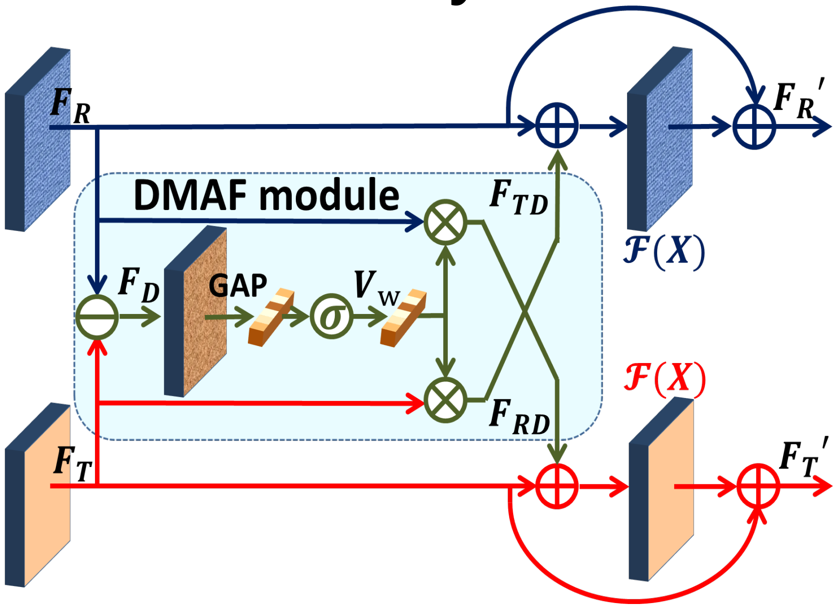


$$F_T' = F_T + \mathcal{F}(F_T)$$

Differential Amplifier



A multi-modality backbone

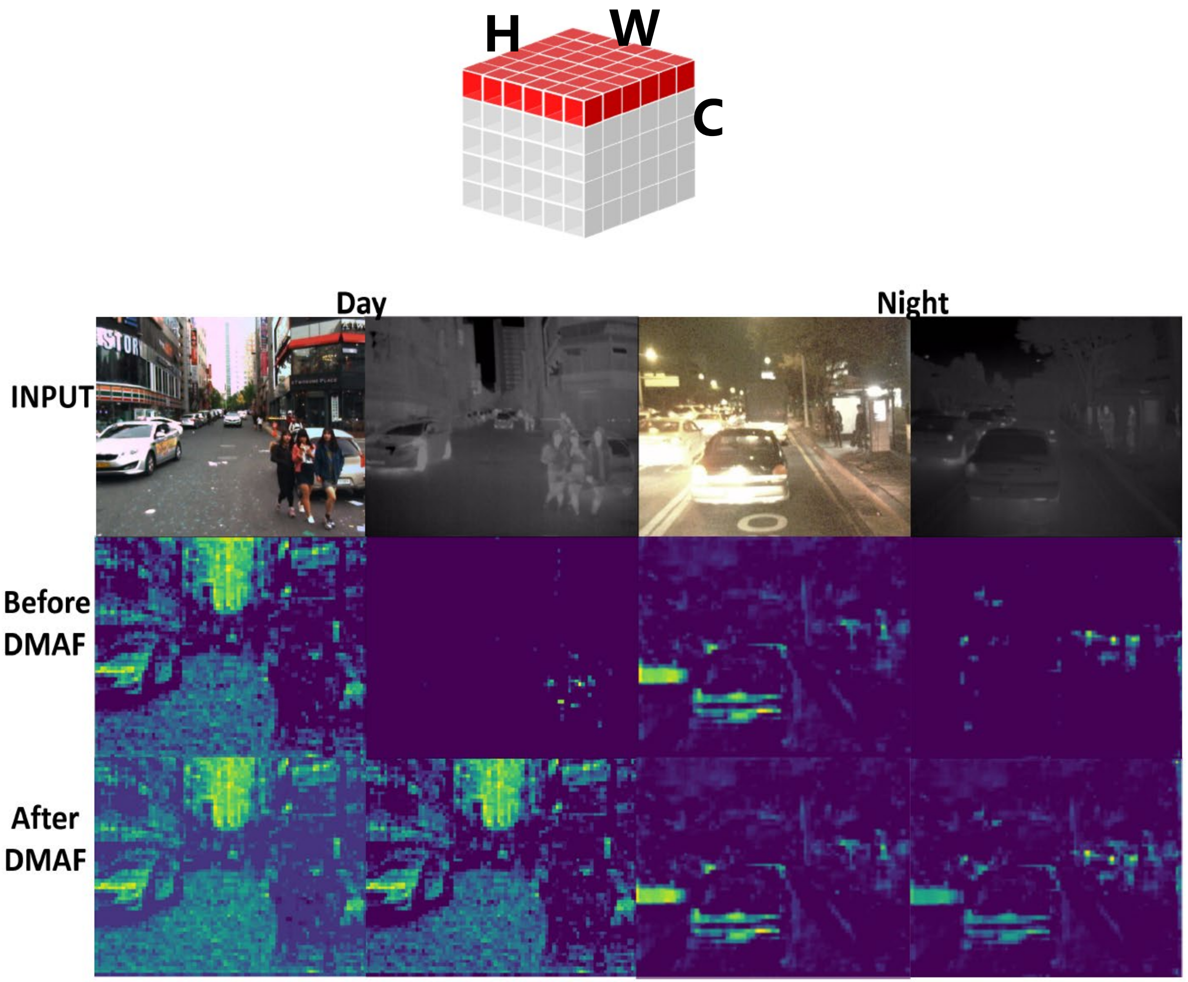


$$F_T' = F_T + \mathcal{F}(F_T \oplus F_{RD}) = F_T + \mathcal{F}(F_T \oplus \sigma(GAP(F_D) \odot F_R)$$
$$F_R' = F_R + \mathcal{F}(F_R \oplus F_{TD}) = F_R + \mathcal{F}(F_R \oplus \sigma(GAP(F_D) \odot F_T)$$

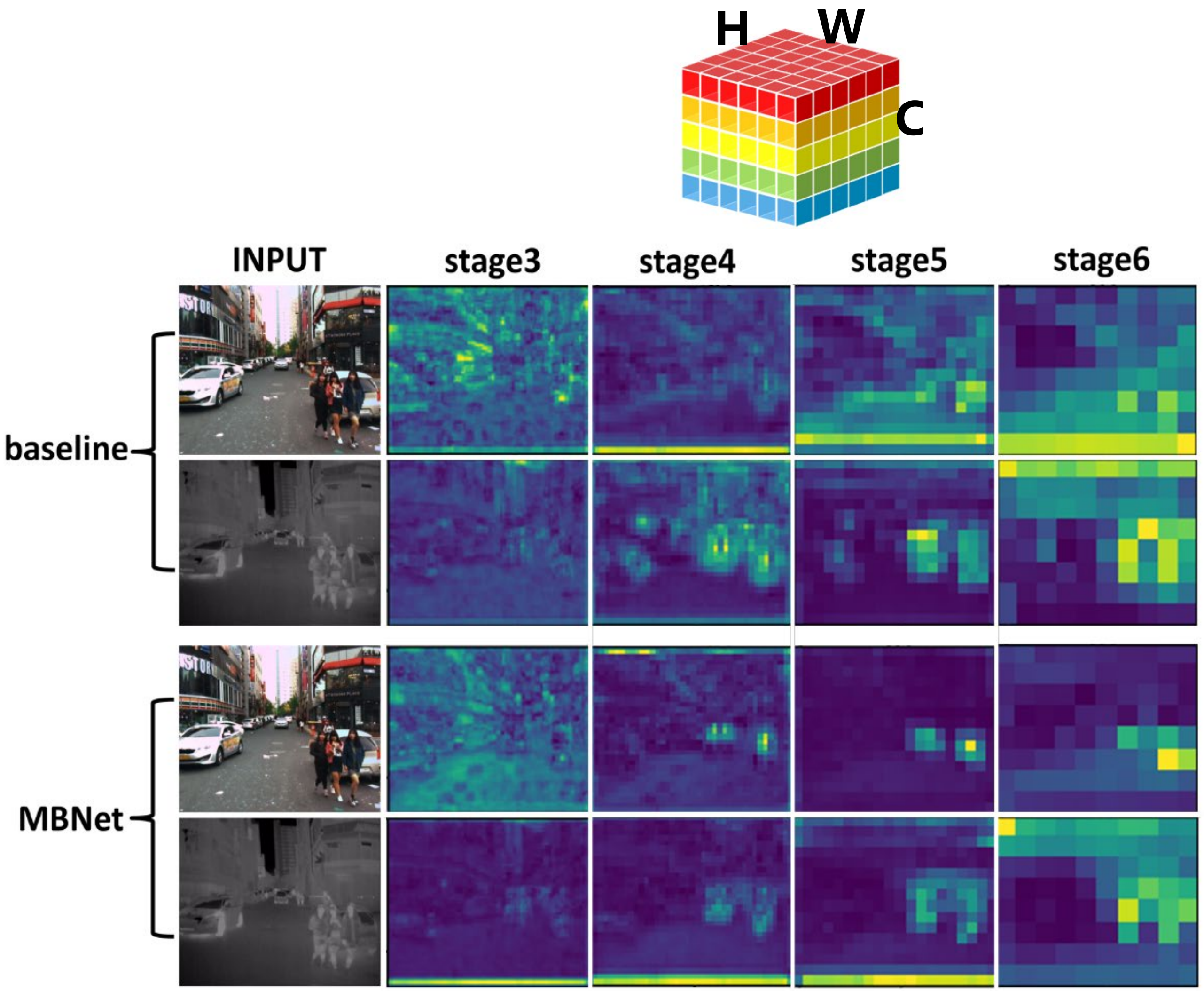
$\oplus$: element-wise sum $\odot$: element-wise multiplication GAP:Global Average Pooling

■ **sufficient fusion in the backbone with no extra parameters**

Visualization of the DMAF Module

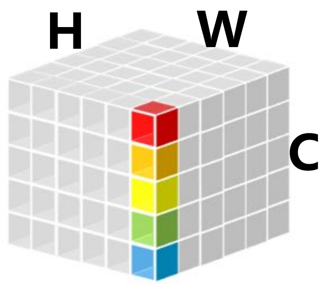


one channel

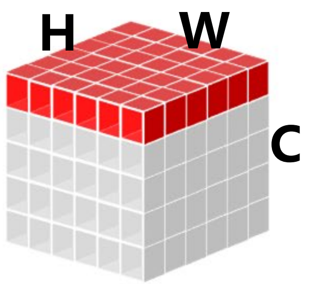


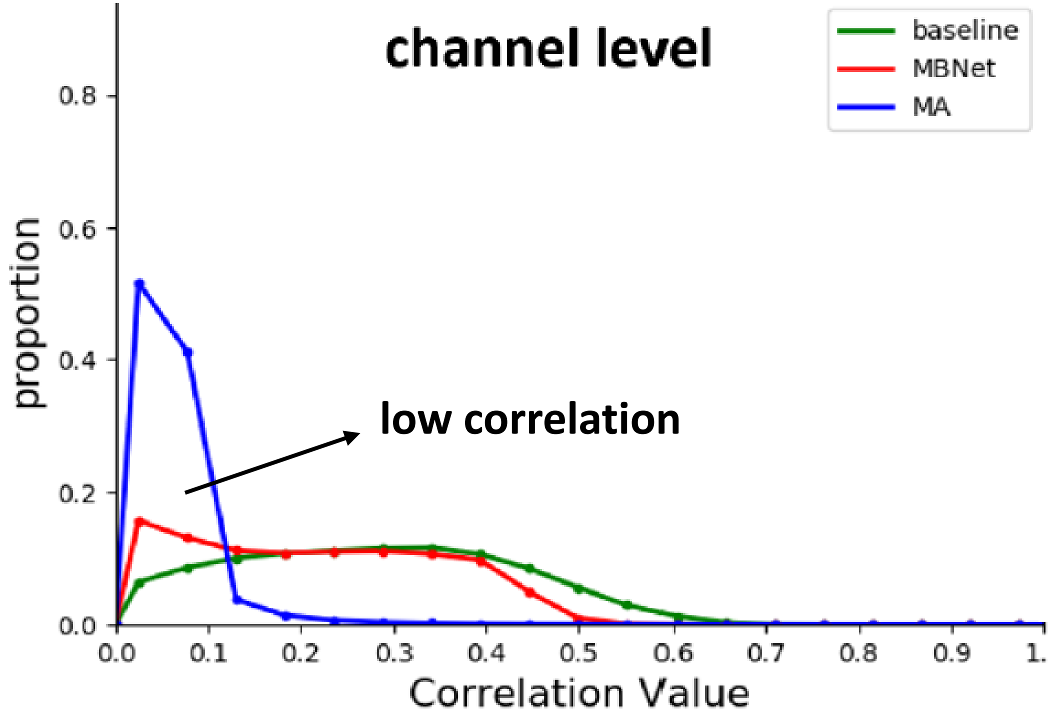
the whole feature map

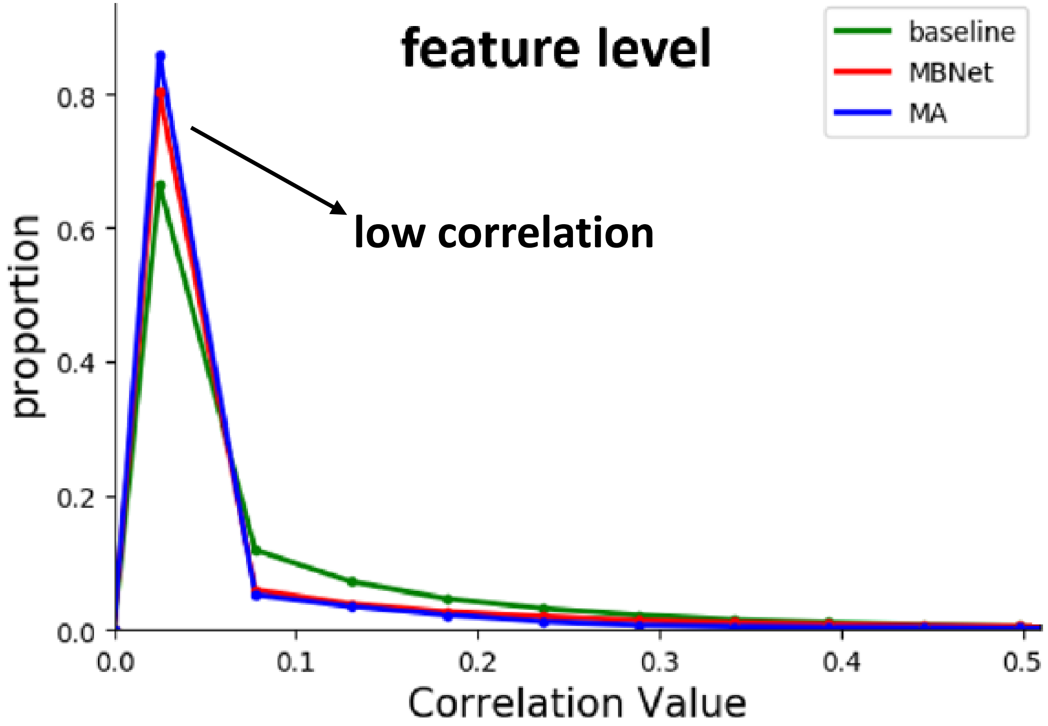
An explanatory perspective: Redundancy Analysis



Pearson product-moment correlation coefficient $|\rho|$ between two modalities

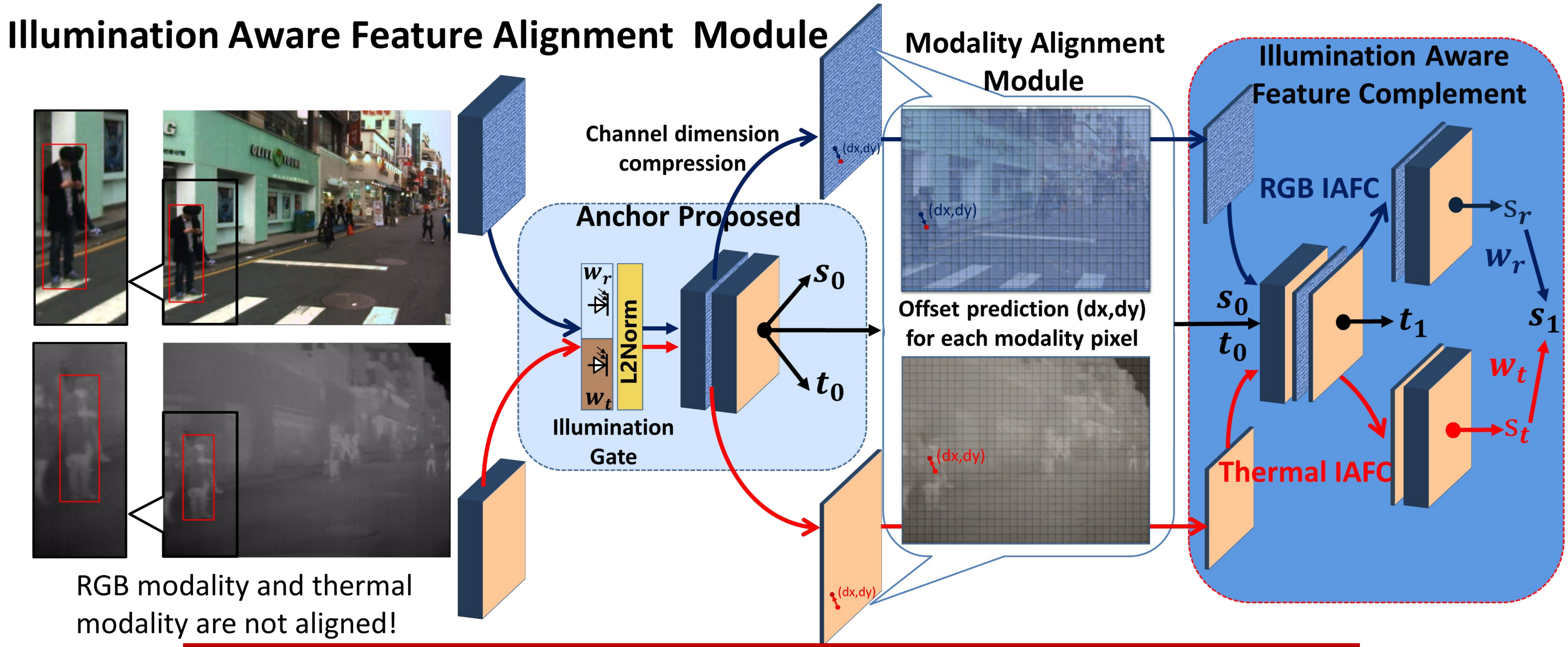






the DMAF module facilitates modality interaction in the backbone which reduces the learning of redundancy and conveys more information.

Illumination Aware Feature Alignment (IAFA) Module



- fix the misalignment between two modality features
- be adaptive to illumination changes

Results and Comparisons

Table 1. Comparisons with the state-of-the-art methods on the KAIST reasonable subset in terms of MR^{-2} [17] with different thresholds of IoU. In addition, Comparisons of running time are also provided.

Methods	MR^{-2} (IoU = 0.5)			MR^{-2}(IoU = 0.75)			Plateform	Speed(s)
	All	Day	Night	All	Day	Night		
ACF [17]	47.32	42.57	56.17	88.79	87.70	91.22	MATLAB	2.73
Halfway Fusion[27]	25.75	24.88	26.59	81.29	78.43	86.80	TITAN X	0.43
Fusion RPN+BF [21]	18.29	19.57	16.27	72.97	68.14	81.35	MATLAB	0.80
IAF R-CNN [23]	15.73	14.55	18.26	75.50	72.34	81.12	TITAN X	0.21
IATDNN + IASS[13]	14.95	14.67	15.72	76.69	76.46	77.05	TITAN X	0.25
RFA[42]	14.61	16.78	10.21	-	-	-	TITAN X	0.08
CIAN [43]	14.12	14.77	11.13	74.45	71.42	80.16	1080 Ti	**0.07**
MSDS-RCNN [22]	11.34	10.53	12.94	70.57	67.36	79.25	TITAN X	0.22
AR-CNN [44]	9.34	9.94	8.38	64.22	57.87	76.82	1080 Ti	0.12
MBNet(ours)	**8.13**	**8.28**	**7.86**	**60.12**	**54.90**	**68.34**	1080 Ti	**0.07**

Results and Comparisons

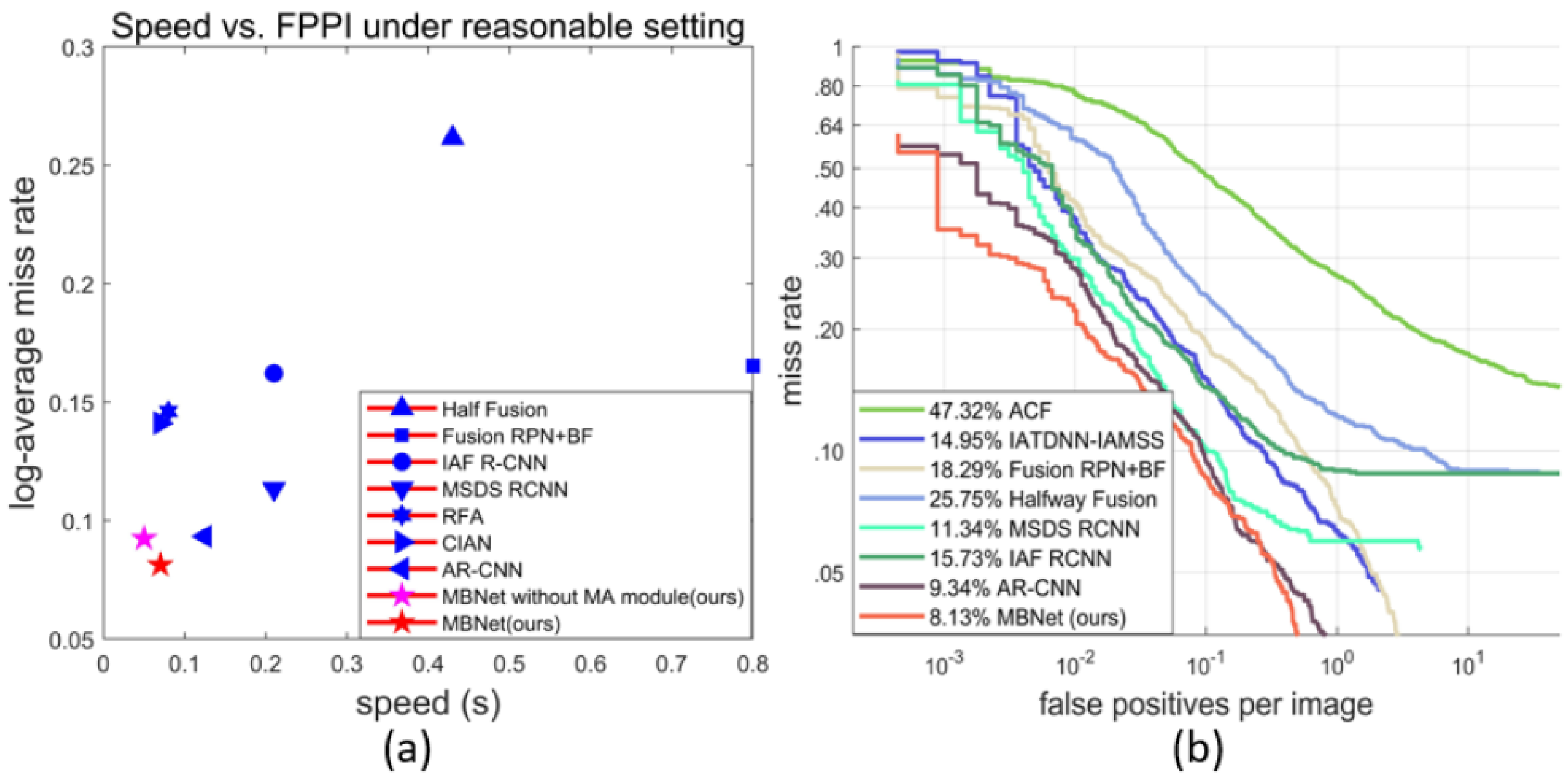


Fig. 5. (a) Log-average miss rate versus the running time of each detector. (b) Performance comparisons with the state-of-the-art methods on the KAIST dataset under reasonable subset.

Demo

Demo Link

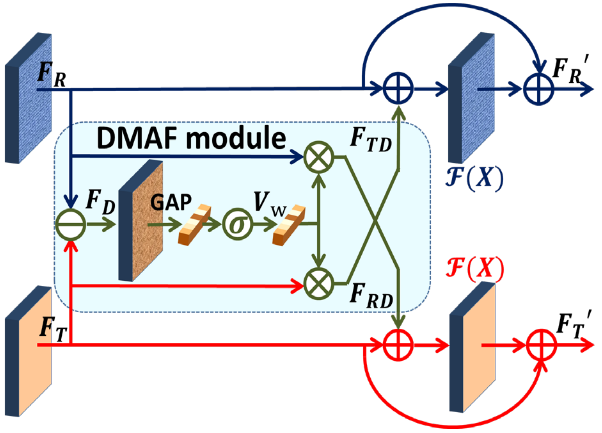
Full version: https://www.bilibili.com/video/BV1Hi4y137aS

Contribution

modality imbalance problems

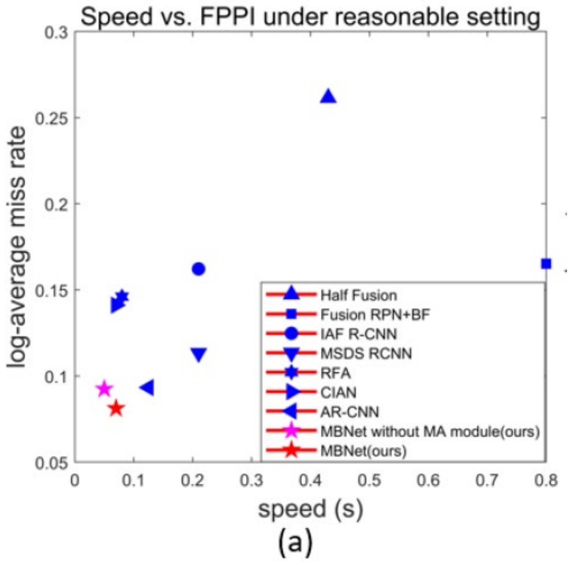
- RGB & Thermal
- Night & Day

the DMAF module



- Inspired by differential amplifier
- A multi-modality backbone
- No extra parameters
- Redundancy Analysis

performance



- Best results on KAIST and CVC14 dataset

ECCV'20 ONLINE

23-28 AUGUST 2020

16TH EUROPEAN CONFERENCE ON

COMPUTER VISION

WWW.ECCV2020.EU

Thanks for your attention

Computational Imaging Technology & Engineering Lab

CITE @ Nanjing University

NANJING UNIVERSITY 1902

Code Link

Demo Link