Asymmetric Contextual Modulation for Infrared Small Target Detection

Yimian Dai[1] **Yiquan Wu**[1] **Fei Zhou**[1] **Kobus Barnard**[2]

[1] Nanjing University of Aeronautics and Astronautics

[2] University of Arizona

November 27, 2020

Detecting Infrared Small Target in a Single Image

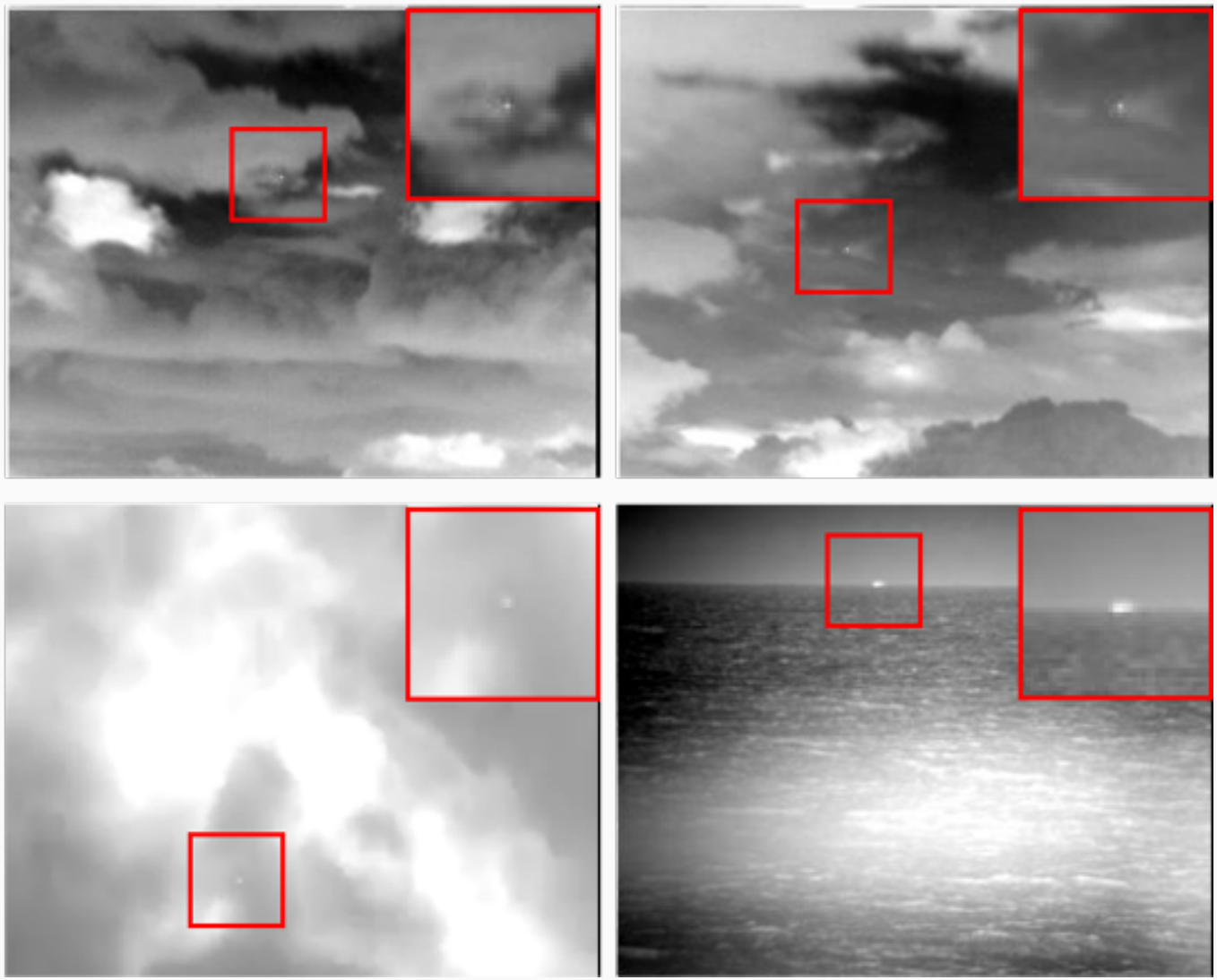


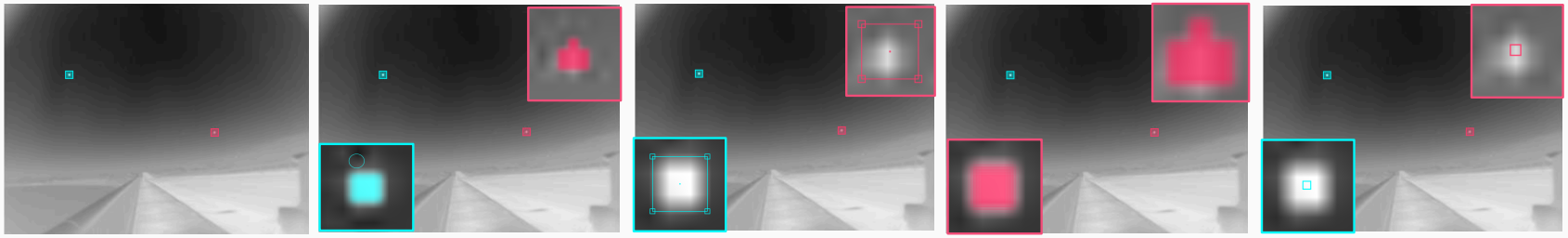
Figure 1: Representative Infrared Images

Bottlenecks in This Field

1. Lack of a public dataset so far.
    - Collecting Real-World Infrared Small Target Images Is Extremely Difficult.
    - Lacking a common benchmark for methods to be trained, tested, and compared.
2. Minimal intrinsic information
    - lacking texture or shape characteristics
    - Contradiction between feature map resolution and semantics

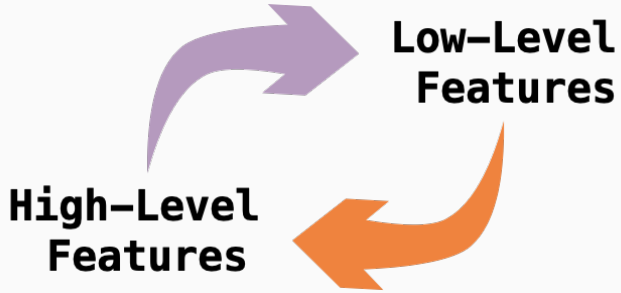
Contributions

- A First Open Dataset with High-Quality Annotations in This Field



(a) Category Labeling (b) Instance Segmentation (c) Bounding Box (d) Semantic Segmentation (e) Instance Spotting

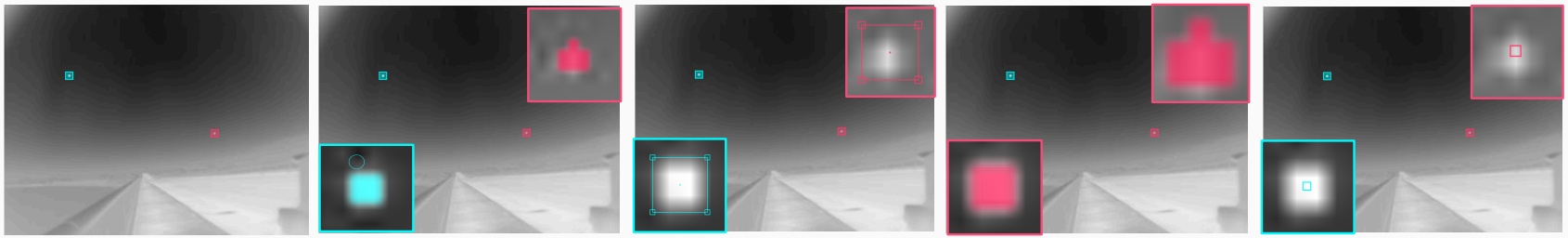
- An Specially Designed Asymmetric Contextual Modulation Module



Mutually Exchanging High-Level Semantics and Low-Level Fine Details

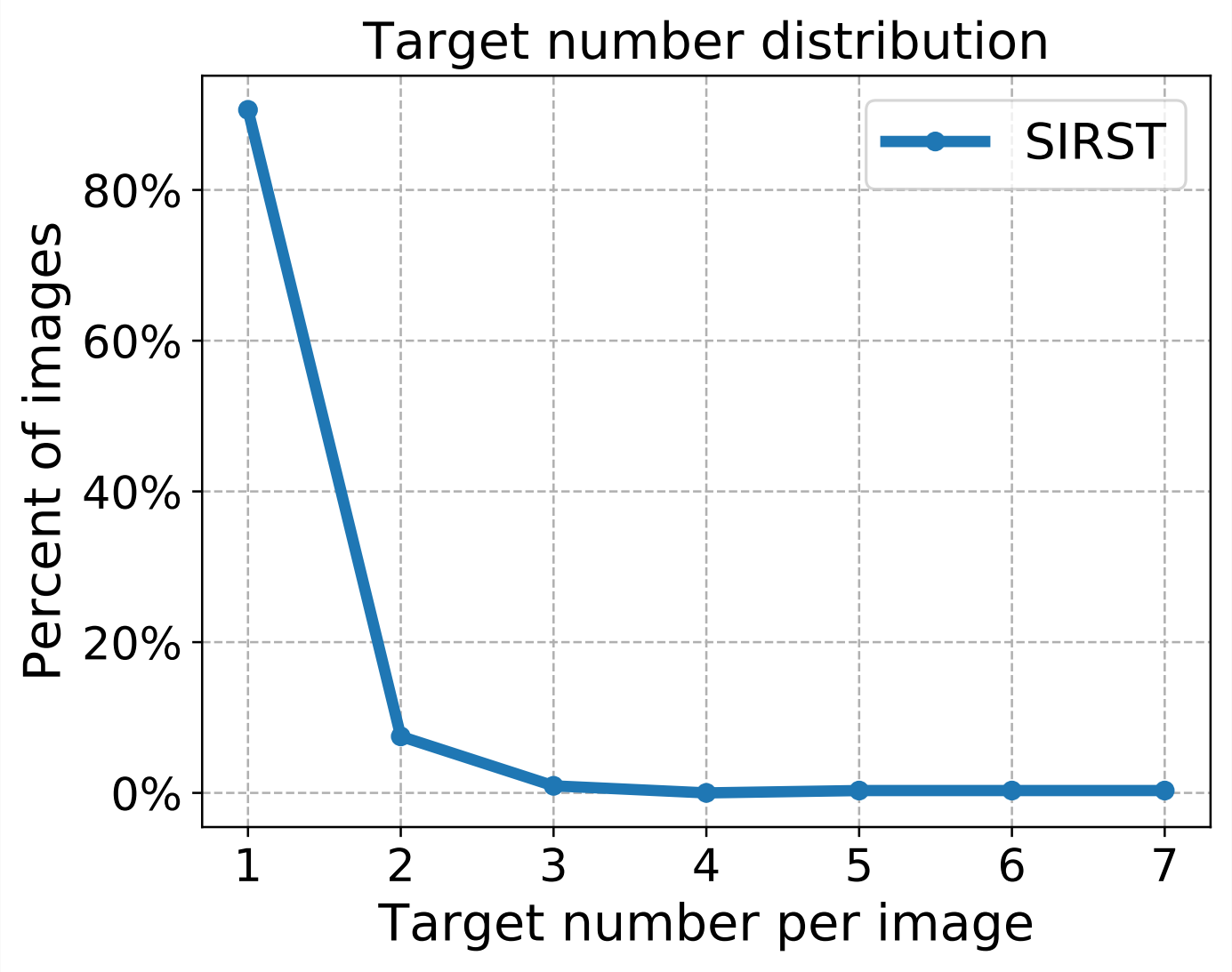
Image Collection and Annotation

- Our dataset contains 427 images including 480 instances
- Five kinds of annotations to support multiple tasks

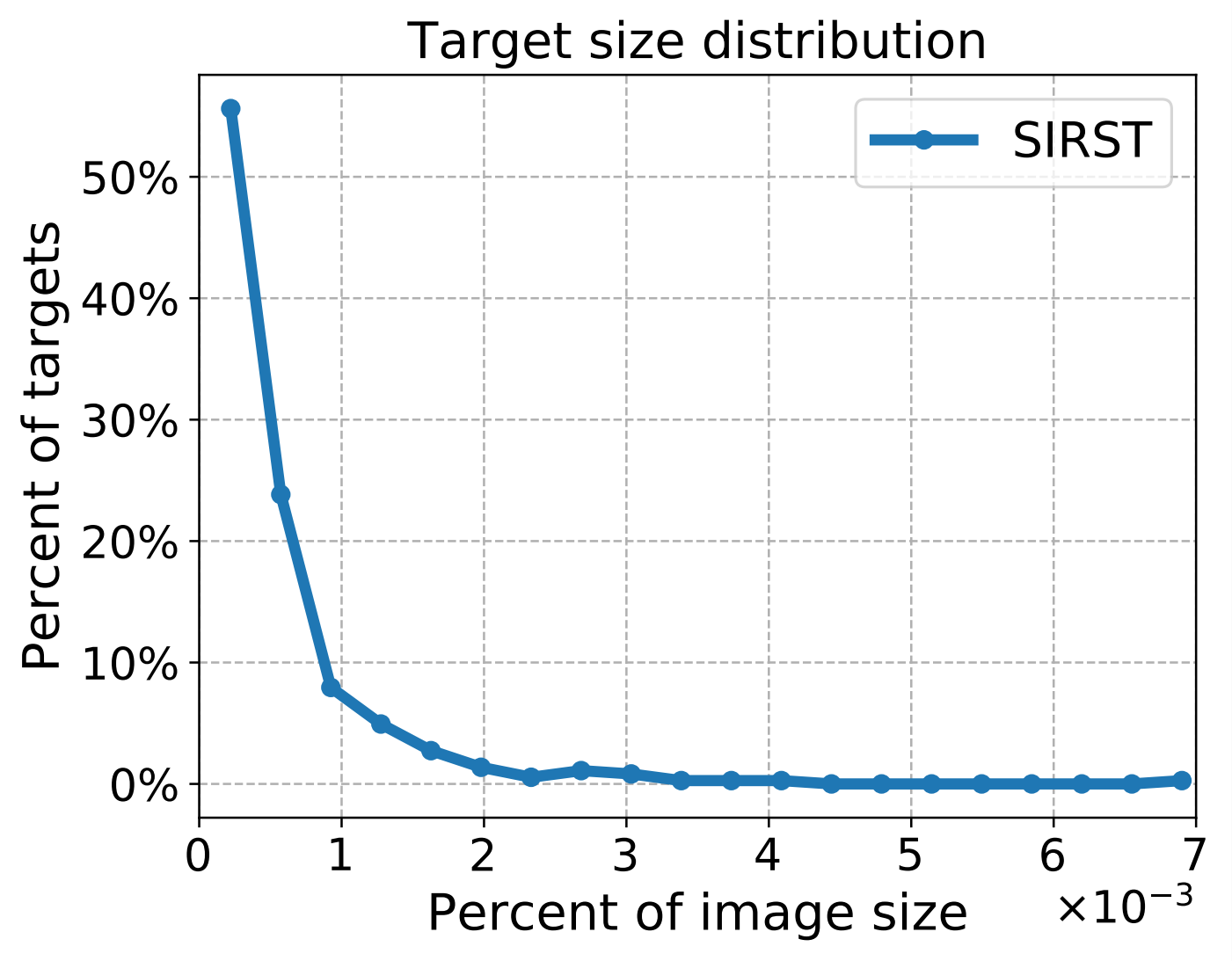


(a) Category Labeling (b) Instance Segmentation (c) Bounding Box (d) Semantic Segmentation (e) Instance Spotting

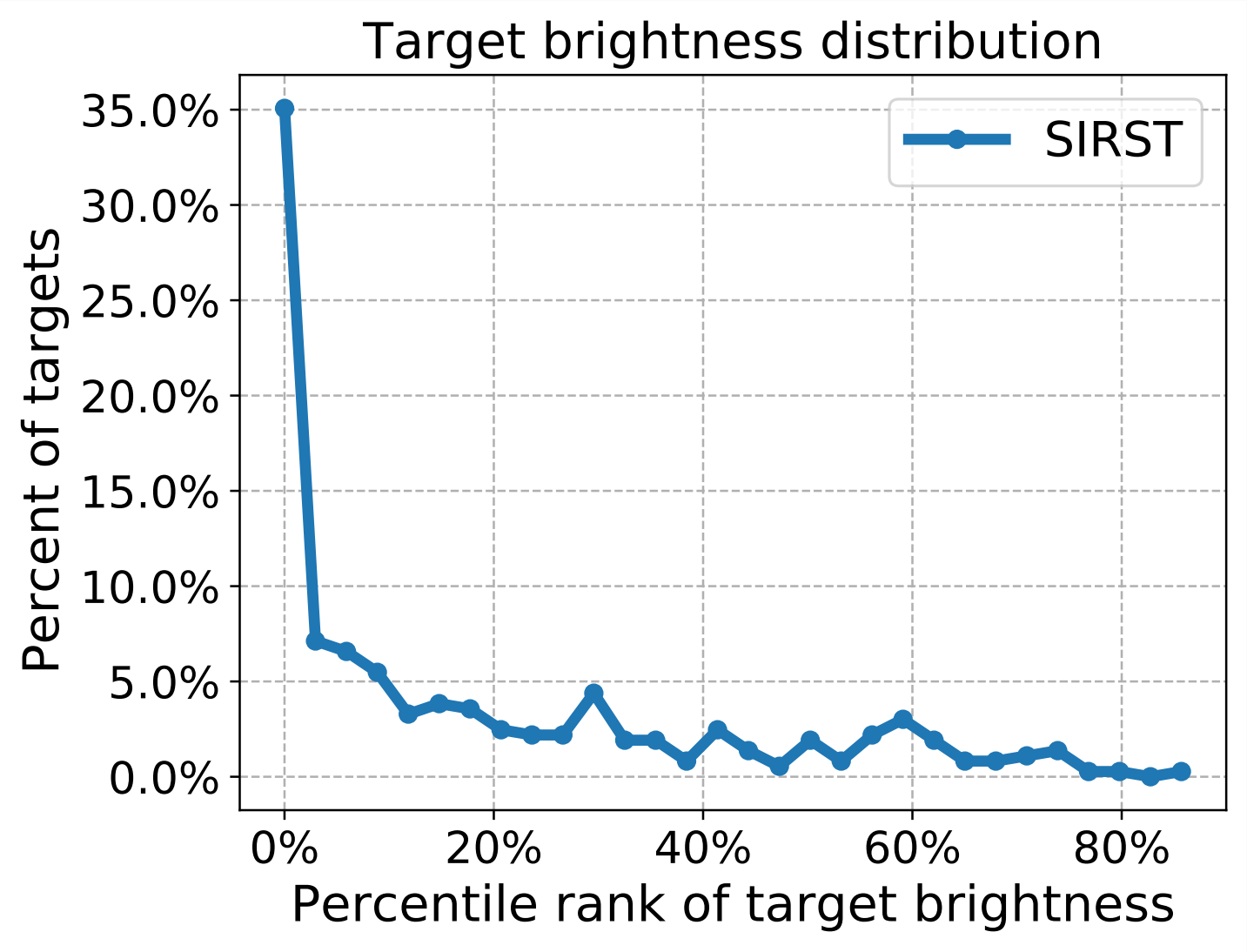
Dataset Statistics – Target Number Distribution



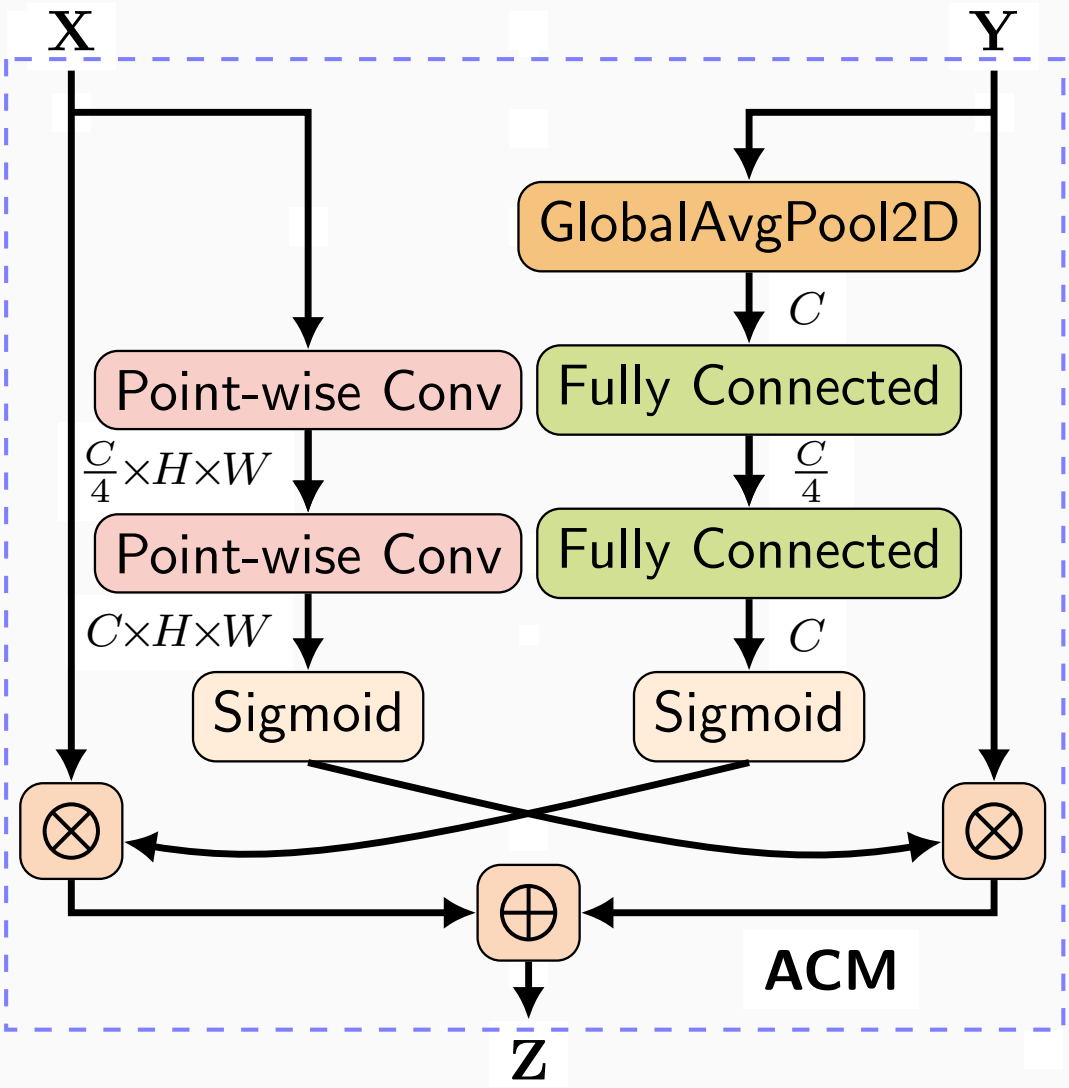
Dataset Statistics – Size Ratio Distribution



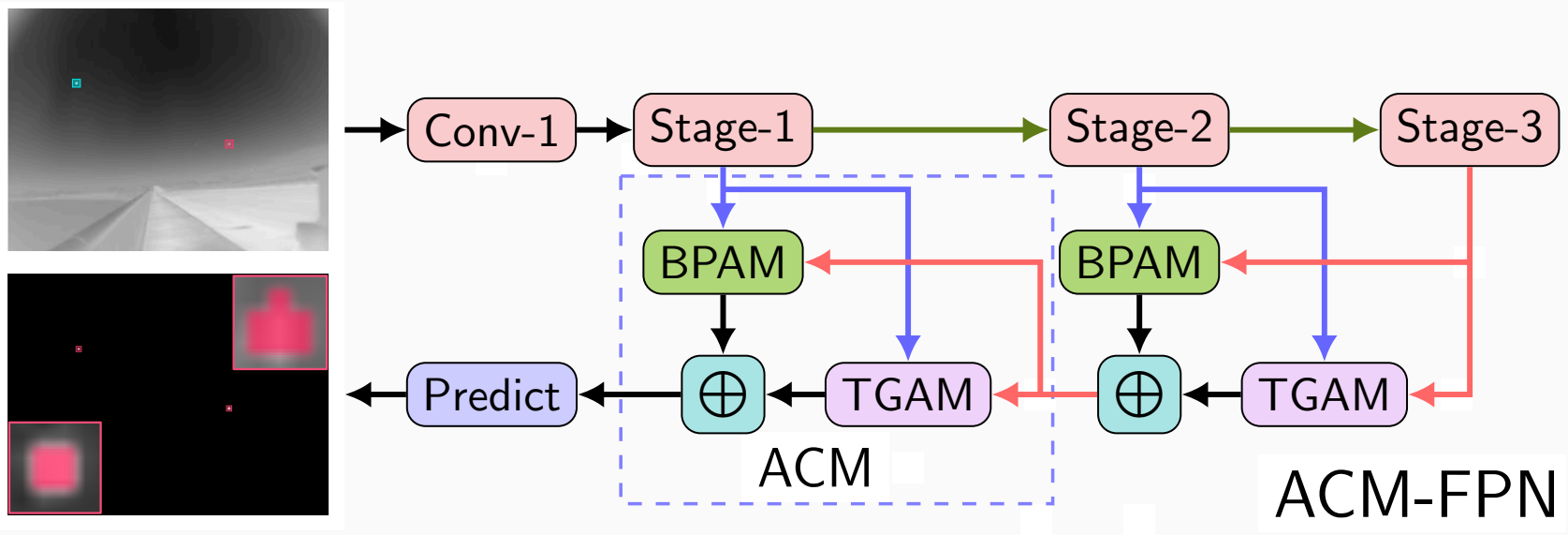
Dataset Statistics – Brightness Distribution

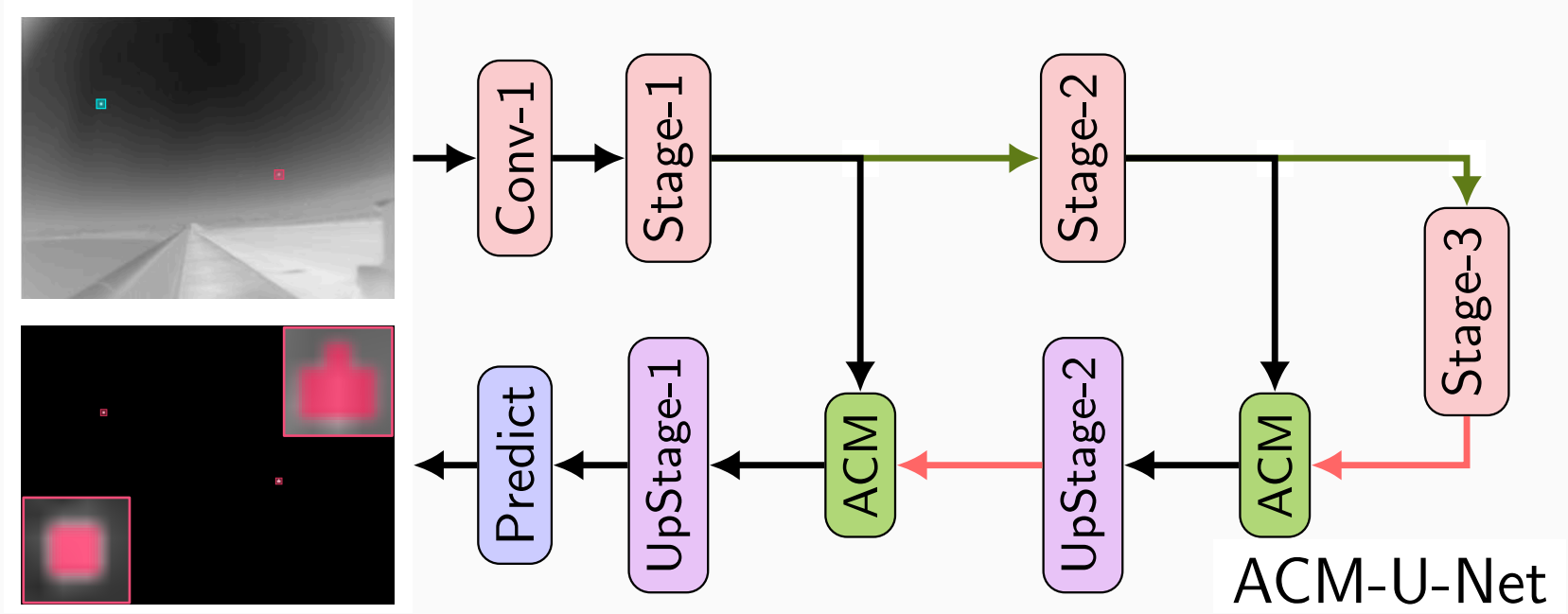


Asymmetric Contextual Modulation



Examples: FPN and U-Net





Backbone for ACM-FPN and ACM-U-Net

Stage	Output	Backbone
Conv-1	480×480	3×3 conv, 16
Stage-1 / UpStage-1	480×480	$\begin{bmatrix} 3 \times 3 \text{ conv}, 16 \\ 3 \times 3 \text{ conv}, 16 \end{bmatrix} \times b$
Stage-2 / UpStage-2	240×240	$\begin{bmatrix} 3 \times 3 \text{ conv}, 32 \\ 3 \times 3 \text{ conv}, 32 \end{bmatrix} \times b$
Stage-3	120×120	$\begin{bmatrix} 3 \times 3 \text{ conv}, 64 \\ 3 \times 3 \text{ conv}, 64 \end{bmatrix} \times b$

Results: Ablation Study on Modulation Scheme

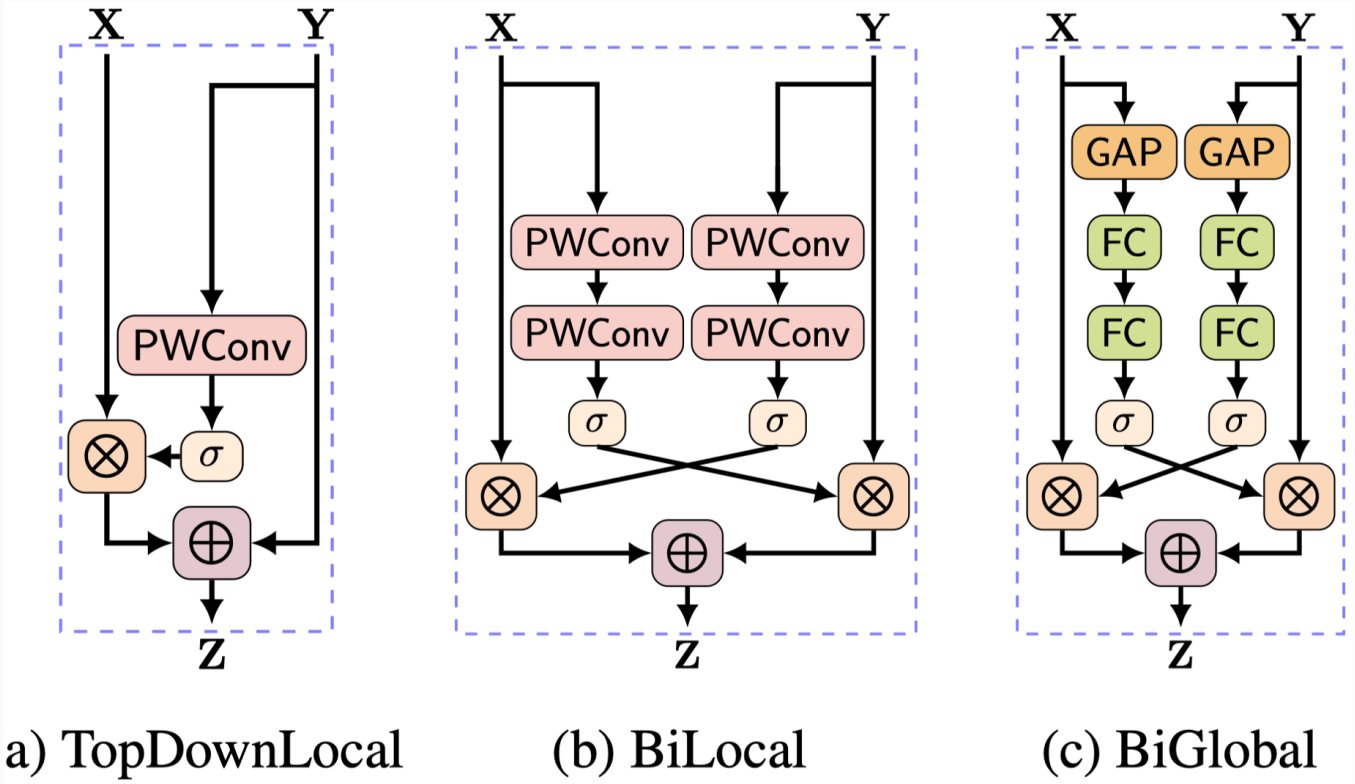


Figure 2: Architectures for the ablation study on modulation scheme

Results: Impact of Down-Sampling and Modulation Scheme

Modulation Scheme	FPN as Host Network							
	IoU				nIoU			
	$b = 1$	$b = 2$	$b = 3$	$b = 4$	$b = 1$	$b = 2$	$b = 3$	$b = 4$
TopDownLocal	0.595	0.648	0.693	0.713	0.635	0.662	0.688	0.703
BiGlobal	0.599	0.660	0.685	0.693	0.645	0.674	0.696	0.684
BiLocal	0.591	0.662	0.713	0.722	0.657	0.694	0.709	0.714
Regular-ACM	**0.683**	**0.703**	0.711	0.711	0.661	0.671	0.680	0.675
ACM	0.645	0.700	**0.714**	**0.731**	**0.684**	**0.702**	**0.713**	**0.721**

Results: Impact of the down-sampling scheme and modulation scheme

Modulation Scheme	U-Net as Host Network							
	IoU				nIoU			
	$b = 1$	$b = 2$	$b = 3$	$b = 4$	$b = 1$	$b = 2$	$b = 3$	$b = 4$
TopDownLocal	0.648	0.710	0.713	0.718	0.673	0.692	0.694	0.697
BiGlobal	0.682	0.716	0.723	0.730	0.688	0.708	0.707	0.719
BiLocal	0.670	0.715	0.718	0.742	0.680	0.710	0.713	0.720
Regular-ACM	0.684	0.700	0.692	0.692	0.637	0.650	0.646	0.643
ACM	**0.707**	**0.732**	**0.741**	**0.743**	**0.709**	**0.720**	**0.726**	**0.731**

Figure 8

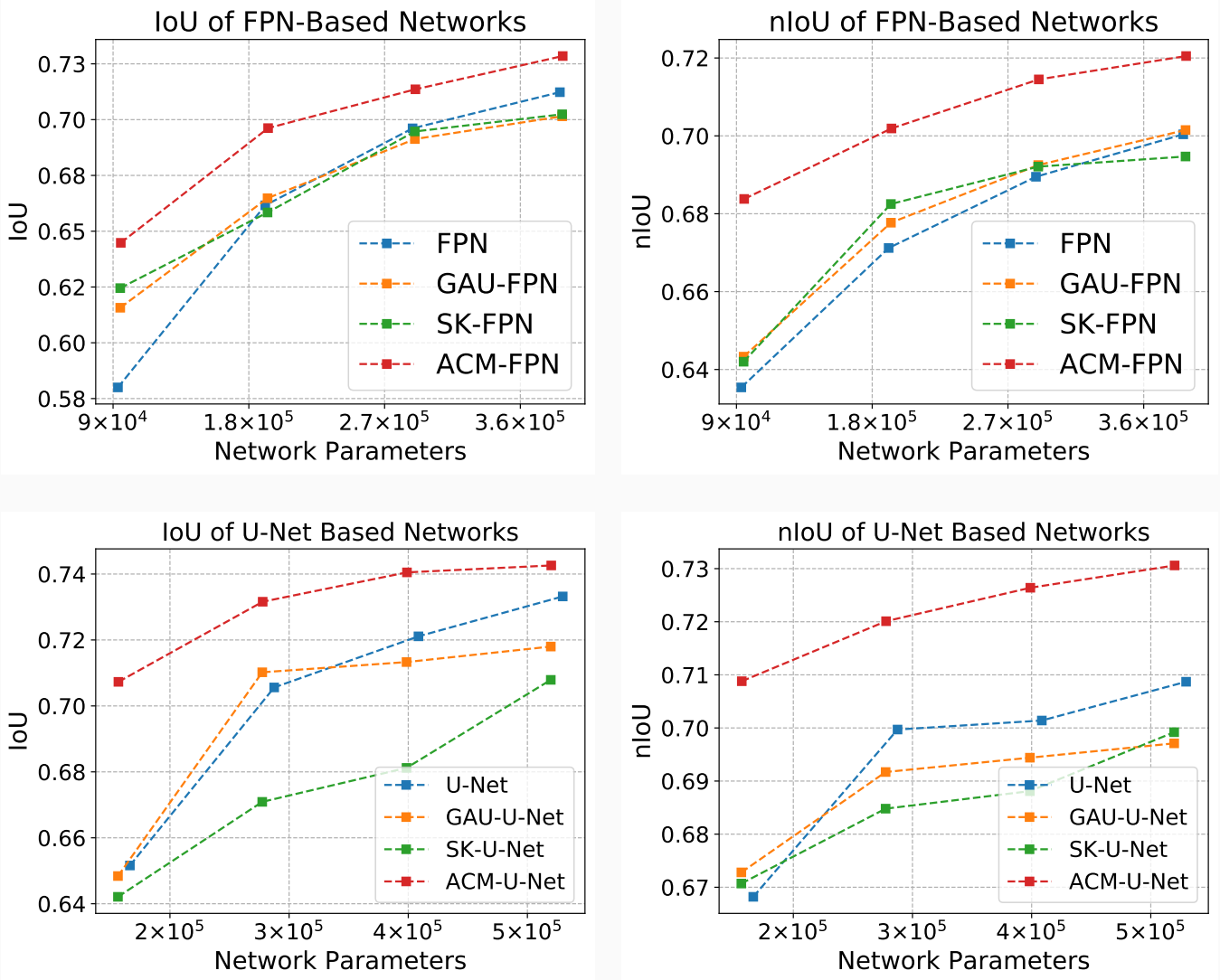


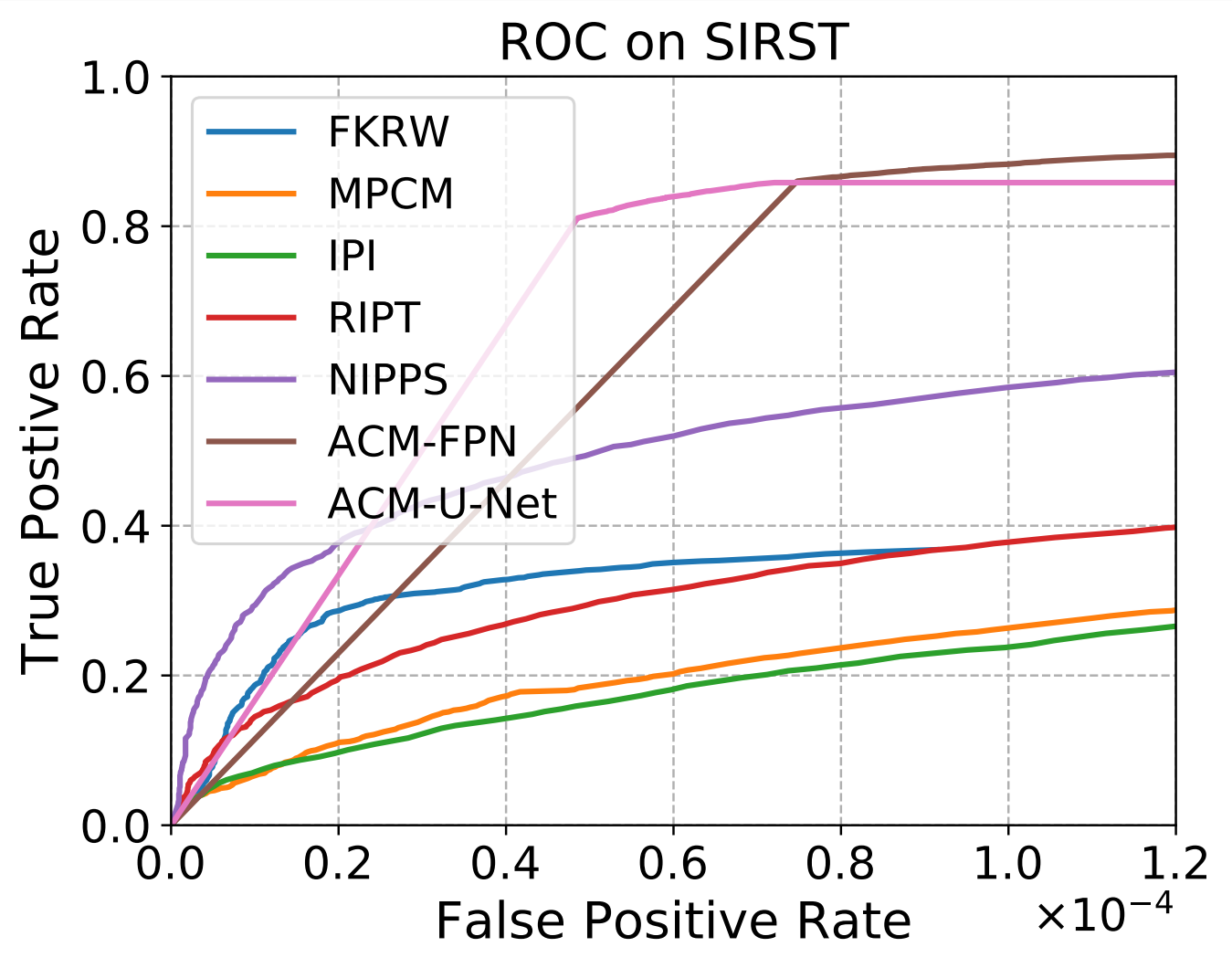
Figure 3: Comparison with other cross-layer modulation schemes

Table 3: IoU and nIoU comparison among 19 methods.

Methods	Model-Driven										
	Local Contrast Measurement						Local Rank + Sparse Decomposition				
	Tophat	LCM	ILCM	LSM	FKRW	MPCM	IPI	NIPPS	RIPT	PSTNN	NRAM
IoU	0.220	0.193	0.104	0.1864	0.268	0.357	0.466	0.473	0.146	0.605	0.294
nIoU	0.352	0.207	0.123	0.2598	0.339	0.445	0.607	0.602	0.245	0.504	0.424

Methods	Data-Driven							
	FPN Based				U-Net Based			
	FPN	SK	GAU	ACM	U-Net	SK	GAU	ACM
IoU	0.720	0.702	0.701	**0.731**	0.733	0.708	0.718	**0.743**
nIoU	0.700	0.695	0.701	**0.721**	0.709	0.699	0.697	**0.731**

Figure 9: ROC comparison of selected methods



Codes and Trained Models

- Our codes and trained models is available at `https://github.com/YimianDai/open-aff`
- Our dataset is available at `https://github.com/YimianDai/sirst`

Thank you!