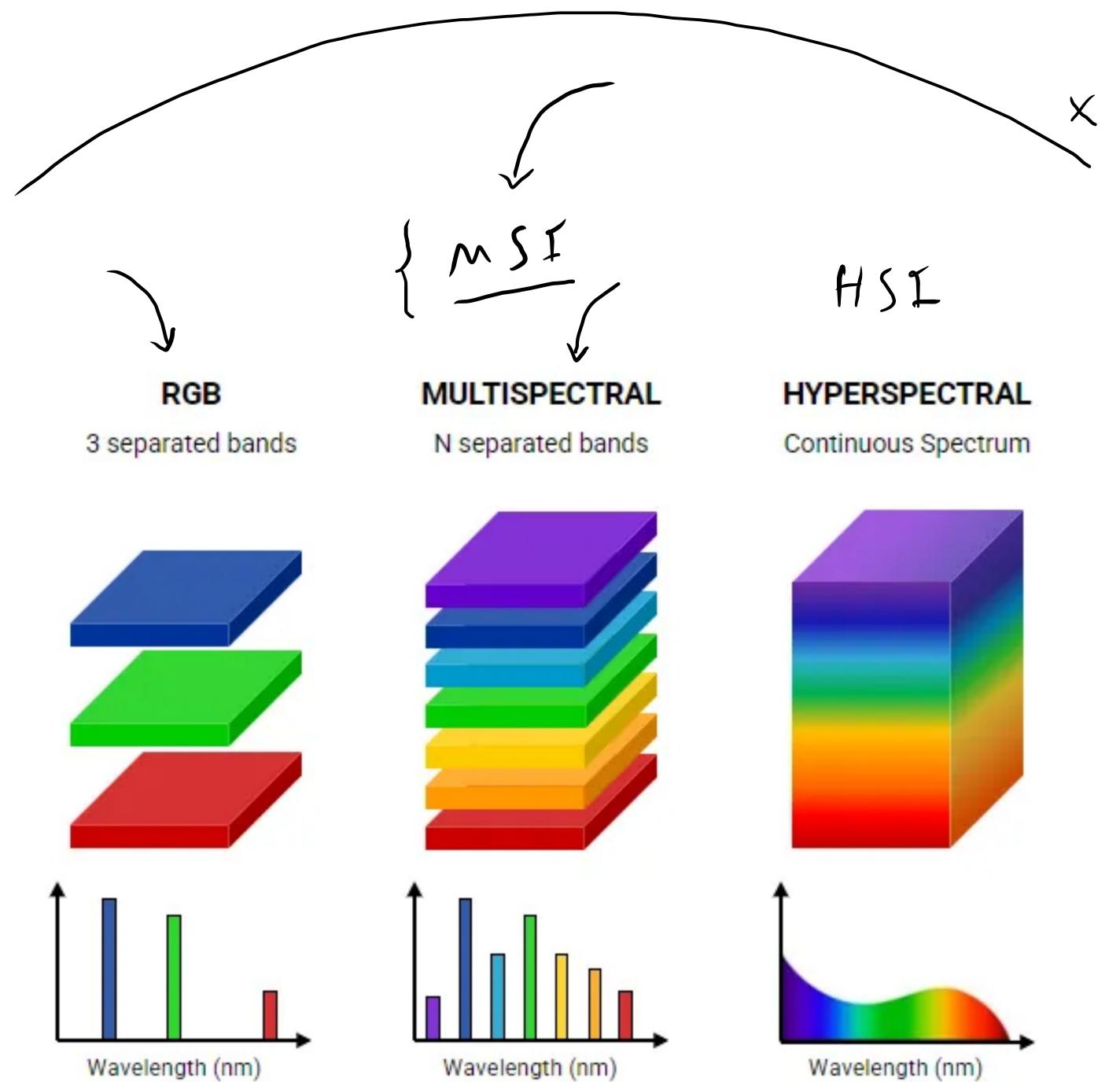


موضوع

تشخیص اشیا در شرایط چالشی  
(تصاویر ماهواره‌ای، ...)

×

## انواع تصاویر ماهواره‌ای

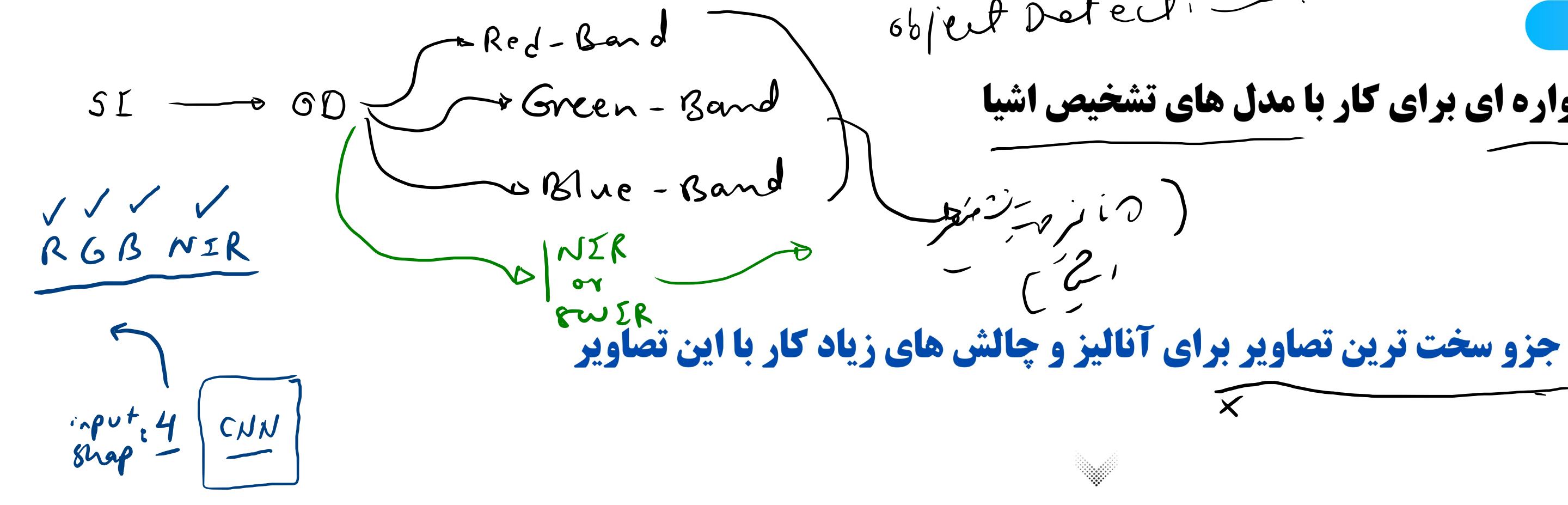


Thermal Image



SAR Image

## Object Detection



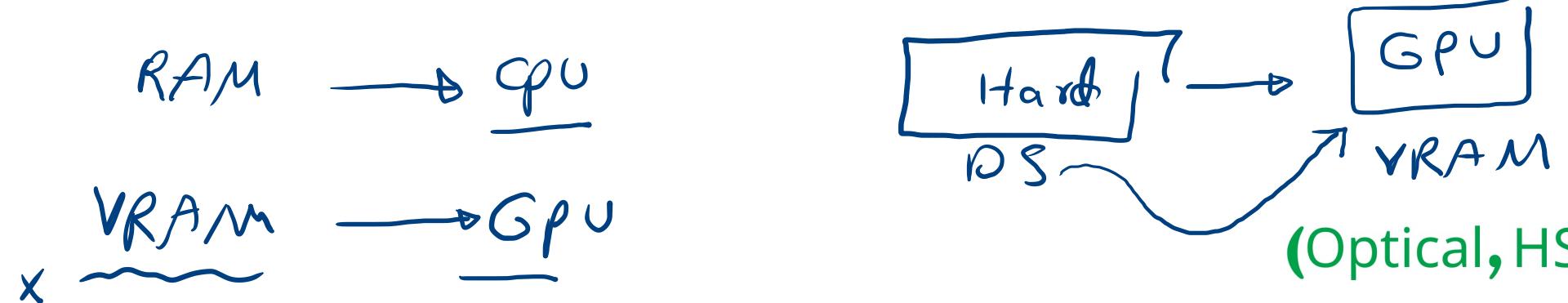
اگر بر اين چالش‌های غلبه کنيد، کار با تصاویر روتين ساده‌تر خواهد شد. (+) نتیجه معتبر را در ماهواره

نموده و نهاده

است. در عین حالا (+) نتیجه معتبر

نموده و نهاده

تصاویر ماهواره‌اي چه چالش‌هایی دارند؟



۱. ابعاد بسیار بزرگ تصویر

تصاویر ماهواره ای می توانند  
۱۰,۰۰۰ × ۱۰,۰۰۰ پیکسل  
یا حتی بزرگ تر باشند.

نمی توانند مستقیم روی GPU  
لود شوند.

## چالش های آنالیز ( تشخیص اشیا) تصاویر ماهواره ای (Optical, HSI, MSI)

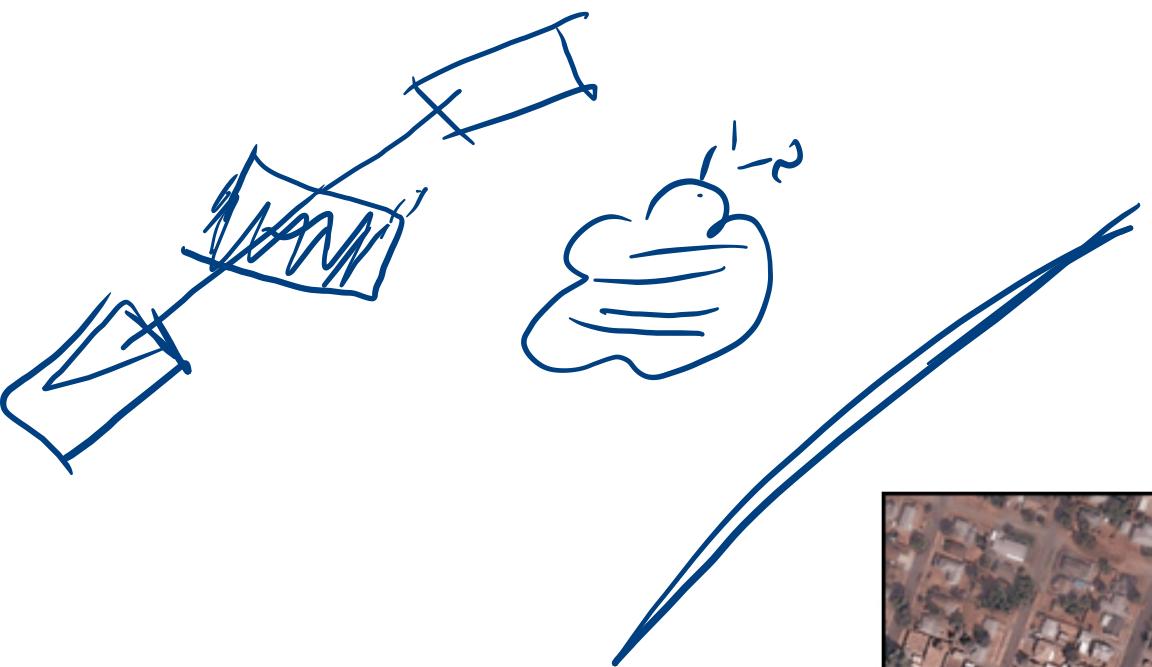


### ۲. اشیای بسیار کوچک

مثلا ماشین‌ها، قایق‌ها یا تیرهای  
چراغ برق فقط  $5 \times 5$  یا حتی  
 $3 \times 3$  پیکسل‌اند.

تشخیص یا سگمنت اشیا سخت  
می‌شود.

## چالش های آنالیز ( تشخیص اشیا) تصاویر ماهواره ای (Optical, HSI, MSI)

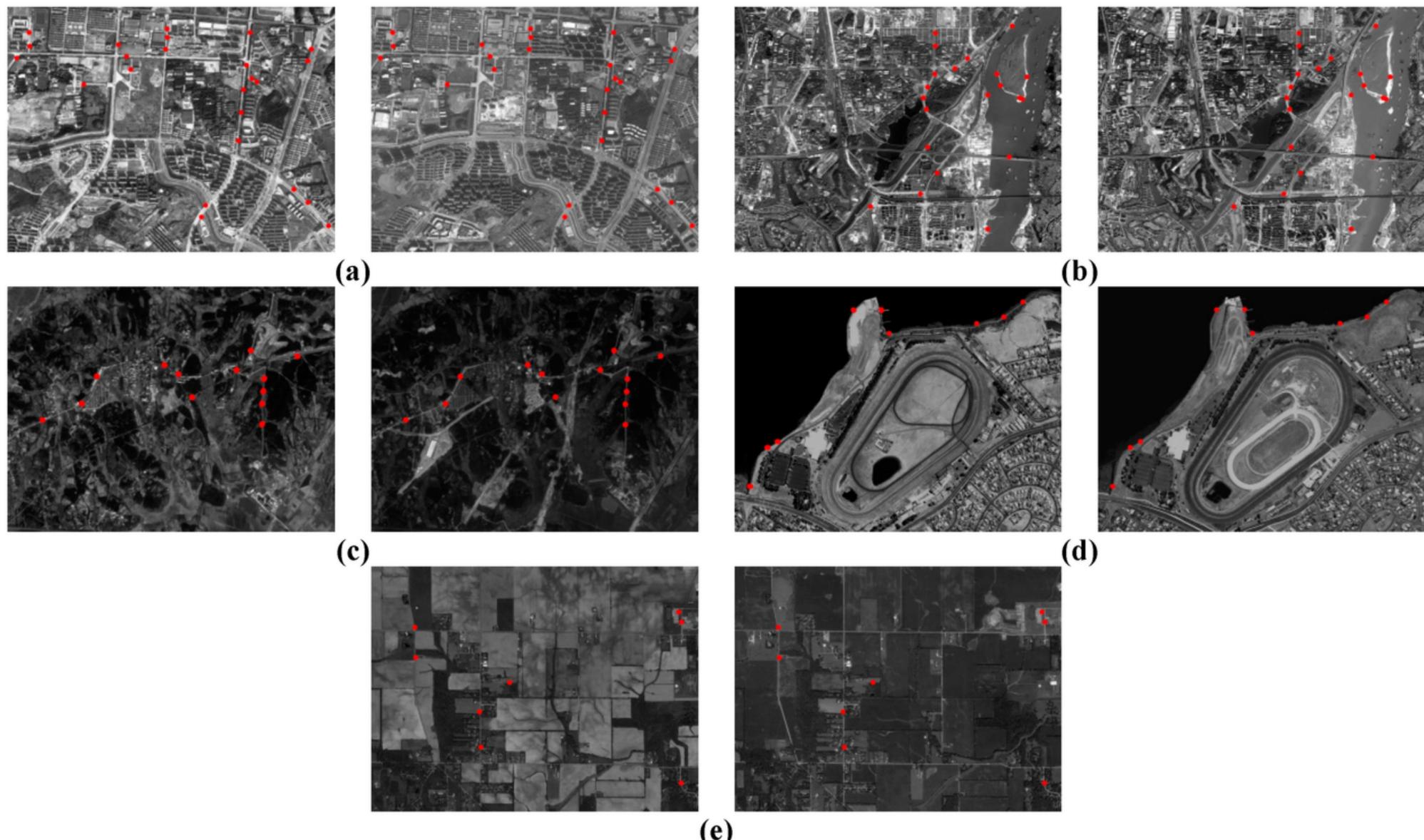
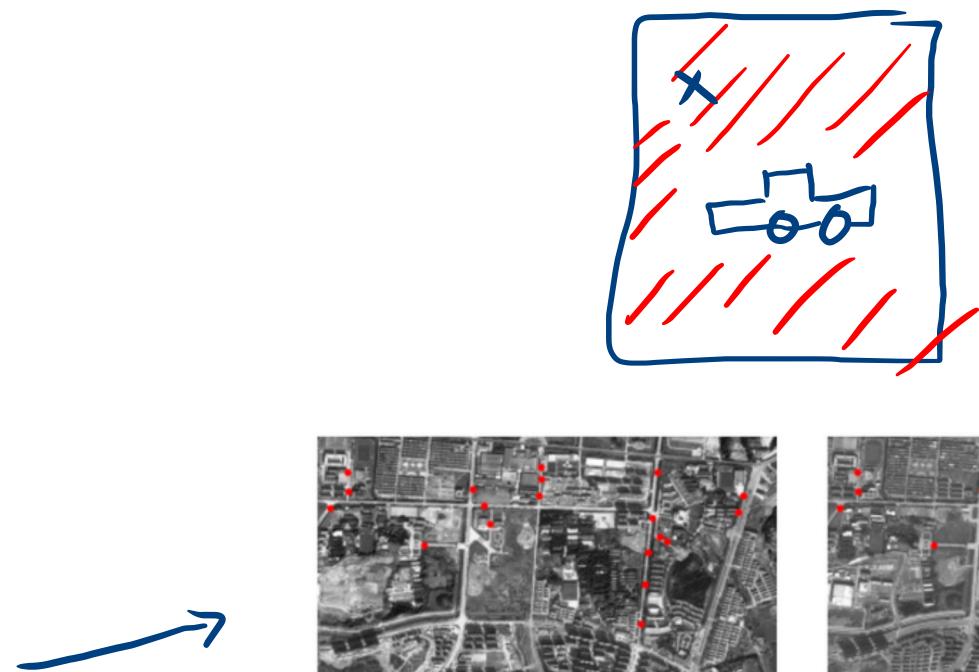


۳. وجود نویزهای اتمسفری (Cloud, Haze, Shadow)

ابرها، دودها و سایه های بزرگ باعث می شوند اشیا قابل تشخیص نباشند.



## چالش های آنالیز (تشخیص اشیا) تصاویر ماهواره ای (Optical, HSI, MSI)

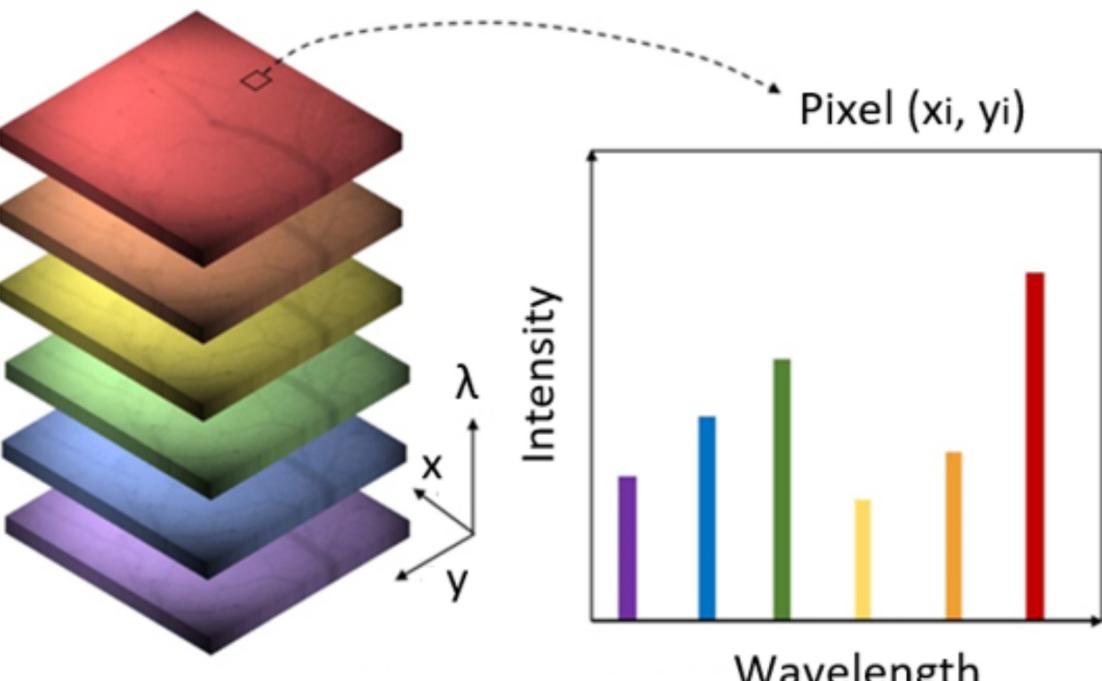
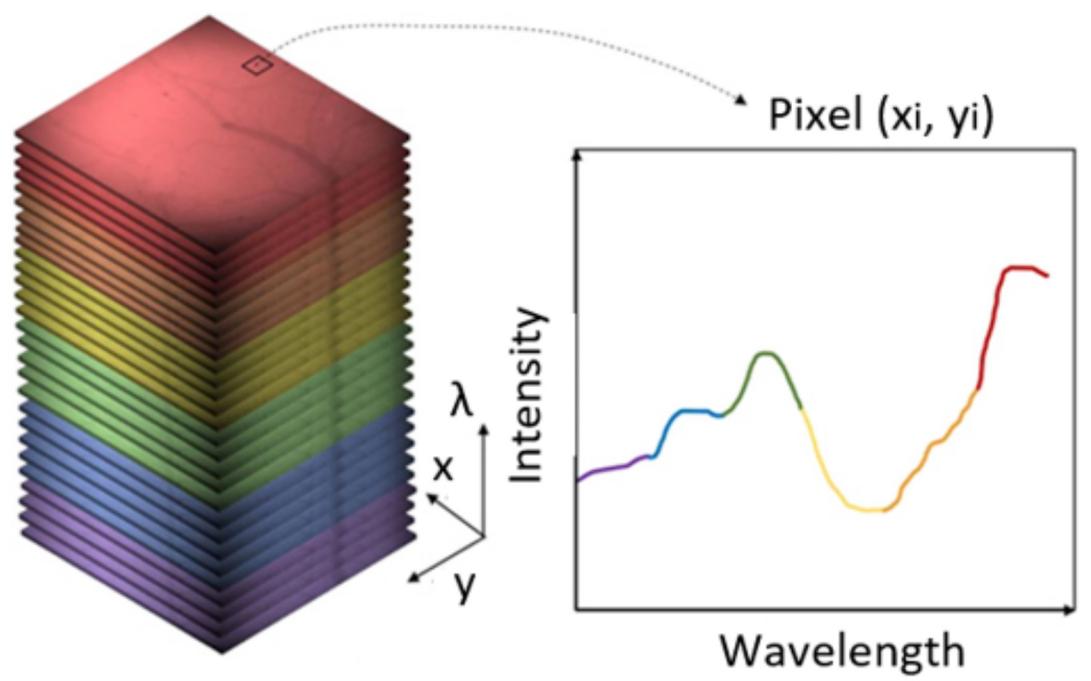


NIR

### ۴. پیچیدگی زمینه (Complex Backgrounds)

مثلًا خیابان‌ها و سطح جنگل، یا کشتی‌ها و سطح آب موافق؛ پس مدل باید بتواند foreground را از background به خوبی جدا کند.

## چالش های آنالیز (تشخیص اشیا) تصاویر ماهواره ای (Optical, HSI, MSI)



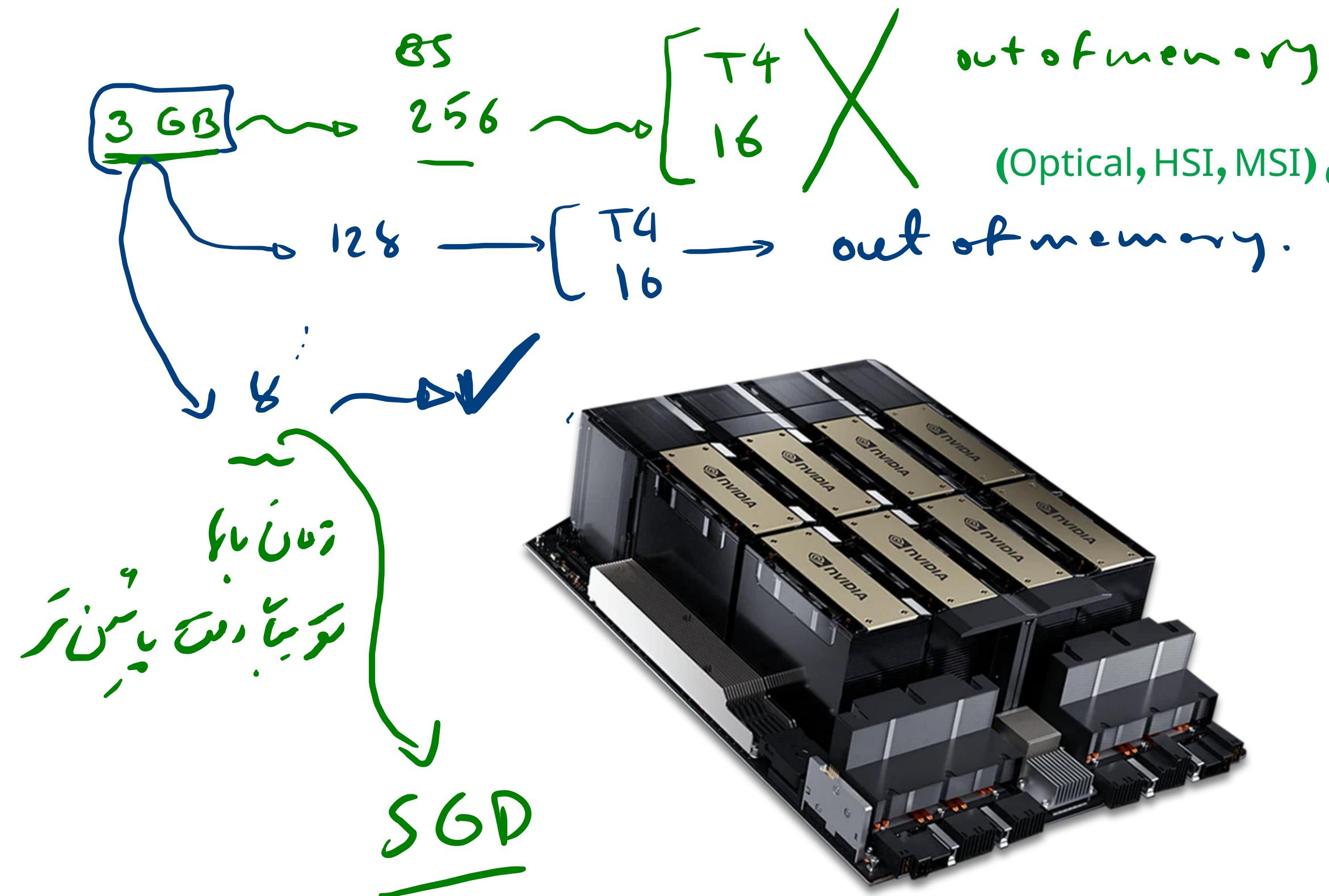
→ مرکز سرمه، مرل، سسرا

۵. چندطیفی بودن  
(MSI, HSI)

تصاویر می توانند چندین کanal  
غیرمرئی (مثل، NIR، SWIR،  
Thermal) داشته باشند

نیاز به مدل های پیچیده تر و

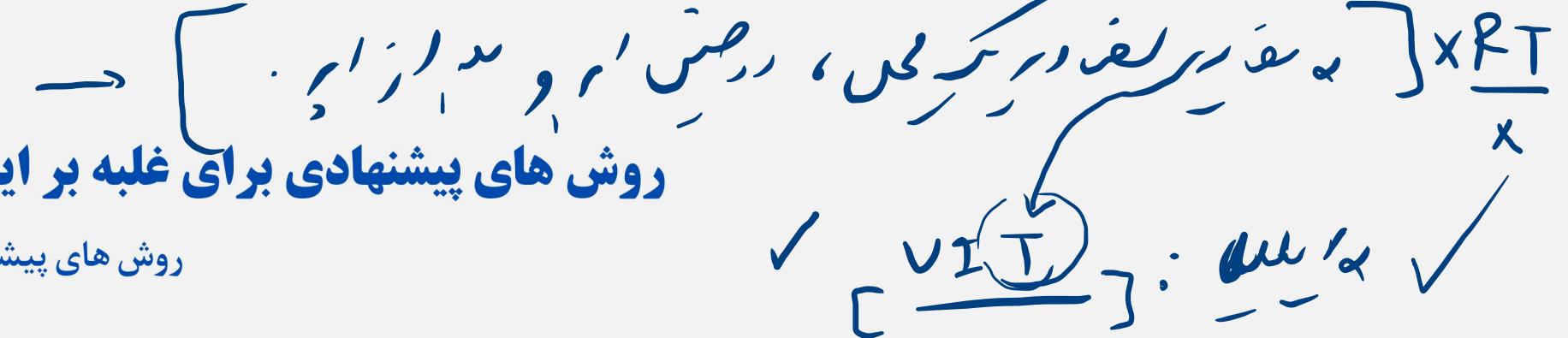
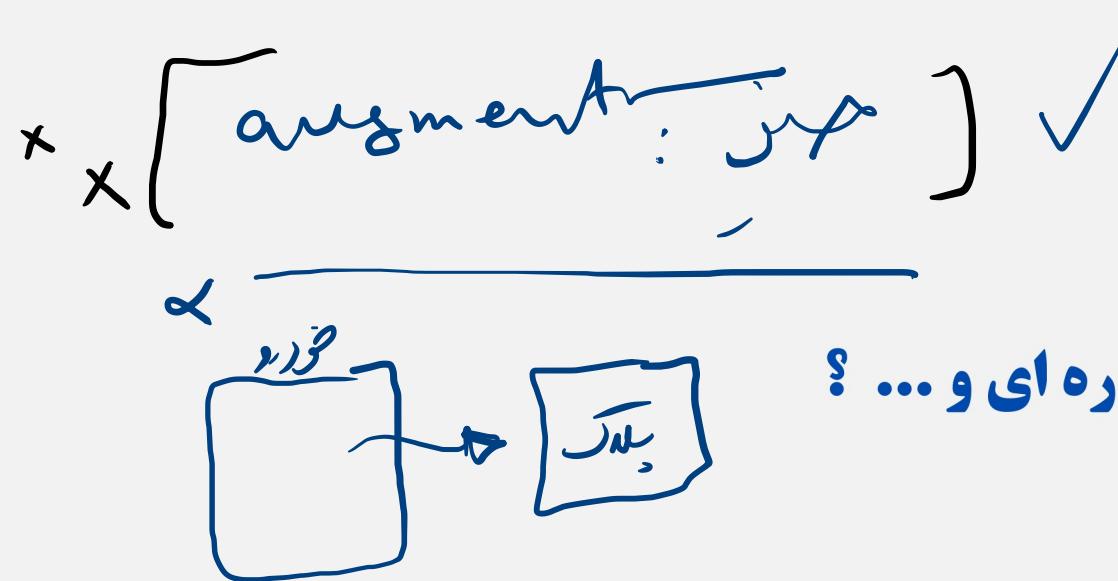
ریسورس بیشتر



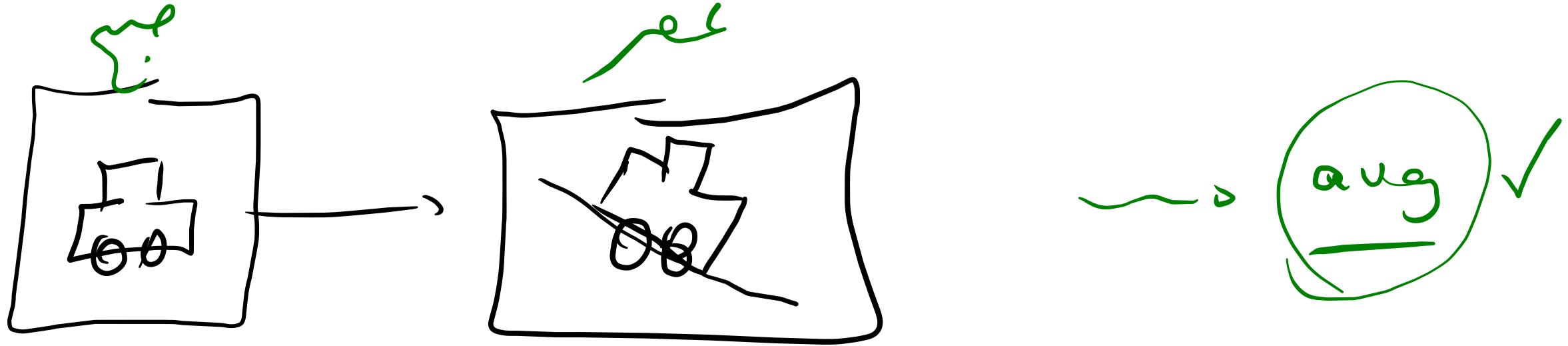
## ۶. محدودیت حافظه و پردازش

# آموزش تصاویر ۴ کناله‌ی $10k \times 10k$ با batch بزرگ

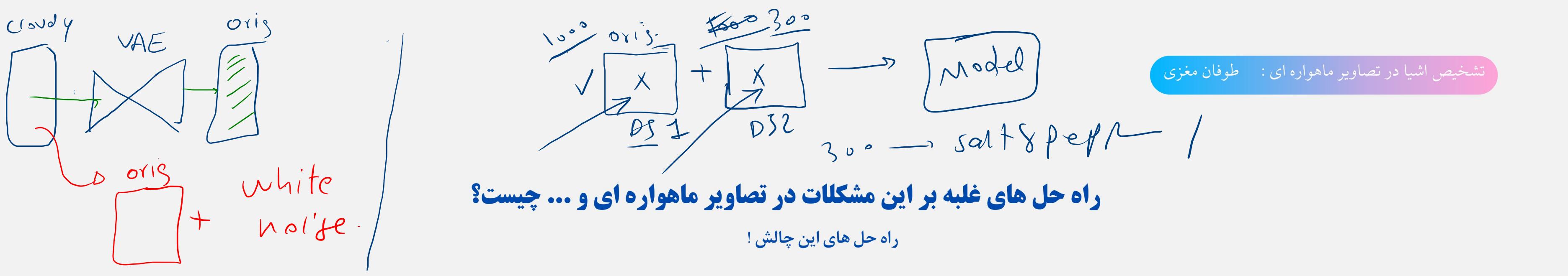
# نیازمند VRAM زیاد لے



- عمار، سفناوره ای مدل کرده: تصفیر کرده می‌شوند: تصفیر منع مسکن
- بیوی: مدل نظری نه کوچک: فزرگ
- عصرها: دسته نورهای متفاوت را با هم می‌کوچک کنند: موزائیک
- صفت مخفی: استفاده از آگوی مخفی برای تصفیر اسما، حرف سعنی ایجاد نمایش!
- عصرها: دسته نورهای متفاوت را با هم می‌کوچک کنند: Siamese
- بُرْج: صفت نهی: بلا نویمها را تصفیر کنند
- FPN: صفت



↙



تبدیل  $1024 \times 1024 \times 10000$  به  $1024 \times 1024$  تقسیم‌بندی تصویر (Tiling)

افزایش رزولوشن (High-Res Training) ورودی مدل باید بزرگ باشد (مثلاً  $1536 \times 1536$  یا  $1024 \times 1024 \times 10000$ )، چون اشیای کوچک در تصاویر کوچک گم می‌شوند.

آموزش بر روی چند کاناله (Multispectral Adaptation) علاوه بر ویژگی‌های لایه‌های دیگر همچون NIR می‌تواند به بهبود تشخیص کمک کند.

استفاده از Focal Loss یا Varifocal Loss به جای ساده‌ی BCE Loss قوی برای اشیای کوچک Loss Function



انتخاب مدل مناسب؟

در صدای ماهواره از Denoising

## مقایسه مدل های تشخیص اشیا

### DETRs Beat YOLOs on Real-time Object Detection

Yian Zhao<sup>1,2†</sup> Wenyu Lv<sup>1†‡</sup> Shangliang Xu<sup>1</sup> Jinman Wei<sup>1</sup> Guanzhong Wang<sup>1</sup>  
Qingqing Dang<sup>1</sup> Yi Liu<sup>1</sup> Jie Chen<sup>2✉</sup>

<sup>1</sup>Baidu Inc, Beijing, China <sup>2</sup>School of Electronic and Computer Engineering, Peking University, Shenzhen, China

[zhaoyian@stu.pku.edu.cn](mailto:zhaoyian@stu.pku.edu.cn) [lvwenyu01@baidu.com](mailto:lvwenyu01@baidu.com) [jiechen2019@pku.edu.cn](mailto:jiechen2019@pku.edu.cn)

[https://arxiv.org/pdf/2304.08069](https://arxiv.org/pdf/2304.08069.pdf)

## مقایسه مدل‌های تشخیص اشیا

Model	Backbone	#Epochs	#Params (M)	GFLOPs	FPS <sub>bs=1</sub>	AP <sup>val</sup>	AP <sup>val</sup> <sub>50</sub>	AP <sup>val</sup> <sub>75</sub>	AP <sup>val</sup> <sub>S</sub>	AP <sup>val</sup> <sub>M</sub>	AP <sup>val</sup> <sub>L</sub>
<i>Real-time Object Detectors</i>											
YOLOv5-L [11]	-	300	46	109	54	49.0	67.3	-	-	-	-
YOLOv5-X [11]	-	300	86	205	43	50.7	68.9	-	-	-	-
PPYOLOE-L [40]	-	300	52	110	94	51.4	68.9	55.6	31.4	55.3	66.1
PPYOLOE-X [40]	-	300	98	206	60	52.3	69.9	56.5	33.3	56.3	66.4
YOLOv6-L [16]	-	300	59	150	99	52.8	70.3	57.7	34.4	58.1	70.1
YOLOv7-L [38]	-	300	36	104	55	51.2	69.7	55.5	35.2	55.9	66.7
YOLOv7-X [38]	-	300	71	189	45	52.9	71.1	57.4	36.9	57.7	68.6
YOLOv8-L [12]	-	-	43	165	71	52.9	69.8	57.5	35.3	58.3	69.8
YOLOv8-X [12]	-	-	68	257	50	53.9	71.0	58.7	35.7	59.3	70.7
<i>End-to-end Object Detectors</i>											
DETR-DC5 [4]	R50	500	41	187	-	43.3	63.1	45.9	22.5	47.3	61.1
DETR-DC5 [4]	R101	500	60	253	-	44.9	64.7	47.7	23.7	49.5	62.3
Anchor-DETR-DC5 [39]	R50	50	39	172	-	44.2	64.7	47.5	24.7	48.2	60.6
Anchor-DETR-DC5 [39]	R101	50	-	-	-	45.1	65.7	48.8	25.8	49.4	61.6
Conditional-DETR-DC5 [27]	R50	108	44	195	-	45.1	65.4	48.5	25.3	49.0	62.2
Conditional-DETR-DC5 [27]	R101	108	63	262	-	45.9	66.8	49.5	27.2	50.3	63.3
Efficient-DETR [42]	R50	36	35	210	-	45.1	63.1	49.1	28.3	48.4	59.0
Efficient-DETR [42]	R101	36	54	289	-	45.7	64.1	49.5	28.2	49.1	60.2
SMCA-DETR [9]	R50	108	40	152	-	45.6	65.5	49.1	25.9	49.3	62.6
SMCA-DETR [9]	R101	108	58	218	-	46.3	66.6	50.2	27.2	50.5	63.2
Deformable-DETR [45]	R50	50	40	173	-	46.2	65.2	50.0	28.8	49.2	61.7
DAB-Deformable-DETR [23]	R50	50	48	195	-	46.9	66.0	50.8	30.1	50.4	62.5
DAB-Deformable-DETR++ [23]	R50	50	47	-	-	48.7	67.2	53.0	31.4	51.6	63.9
DN-Deformable-DETR [17]	R50	50	48	195	-	48.6	67.4	52.7	31.0	52.0	63.7
DN-Deformable-DETR++ [17]	R50	50	47	-	-	49.5	67.6	53.8	31.3	52.6	65.4
DINO-Deformable-DETR [44]	R50	36	47	279	5	50.9	69.0	55.3	34.6	54.1	64.6
<i>Real-time End-to-end Object Detector (ours)</i>											
RT-DETR	R50	72	42	136	108	53.1	71.3	57.7	34.8	58.0	70.0
RT-DETR	R101	72	76	259	74	54.3	72.7	58.6	36.0	58.8	72.1

Table 2. Comparison with SOTA (only L and X models of YOLO detectors, see Appendix for the comparison with S and M models). We do not test the speed of other DETRs, except for DINO-Deformable-DETR [44] for comparison, as they are not real-time detectors. Our RT-DETR outperforms the state-of-the-art YOLO detectors and DETRs in both speed and accuracy.

## مقایسه مدل های تشخیص اشیا

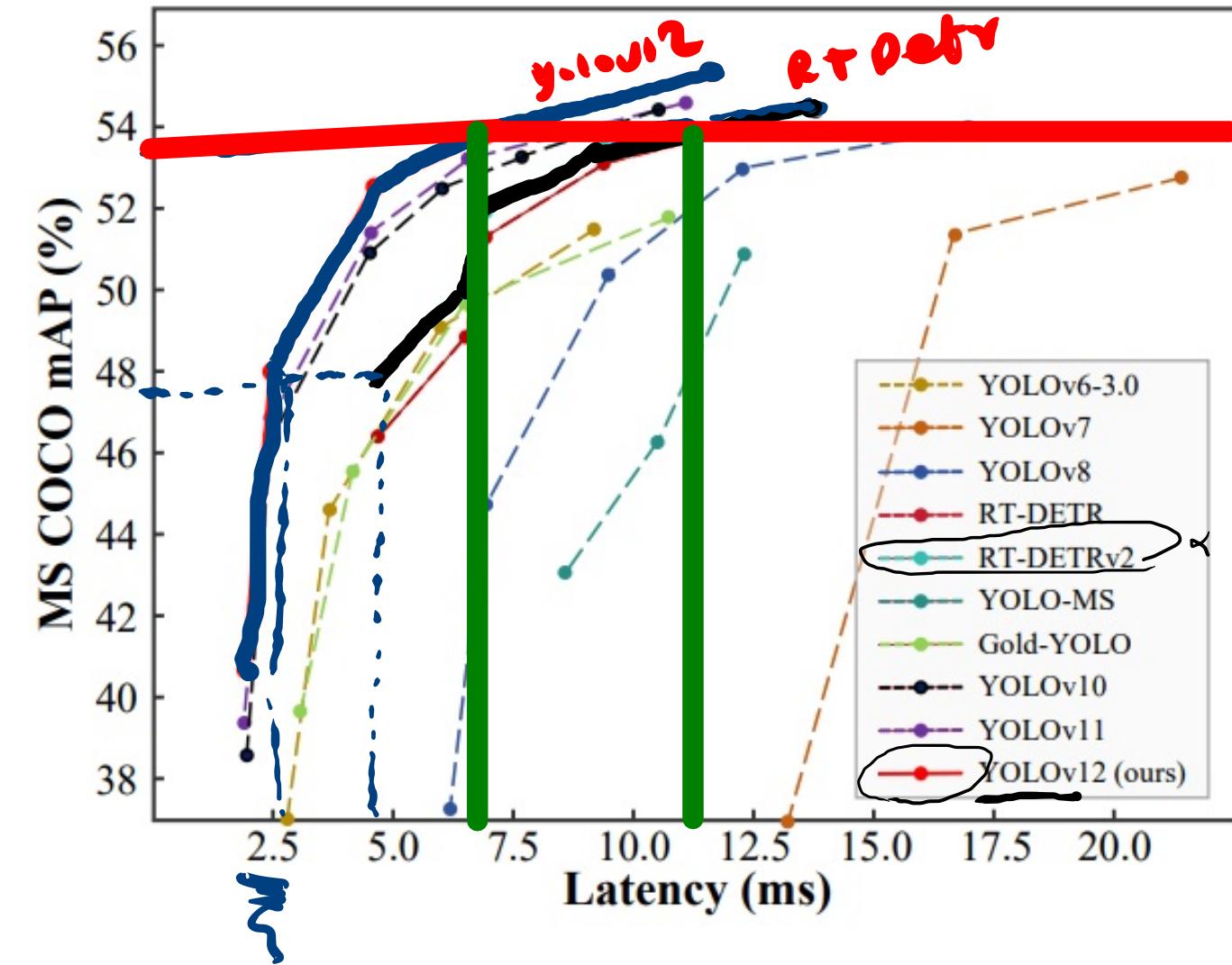
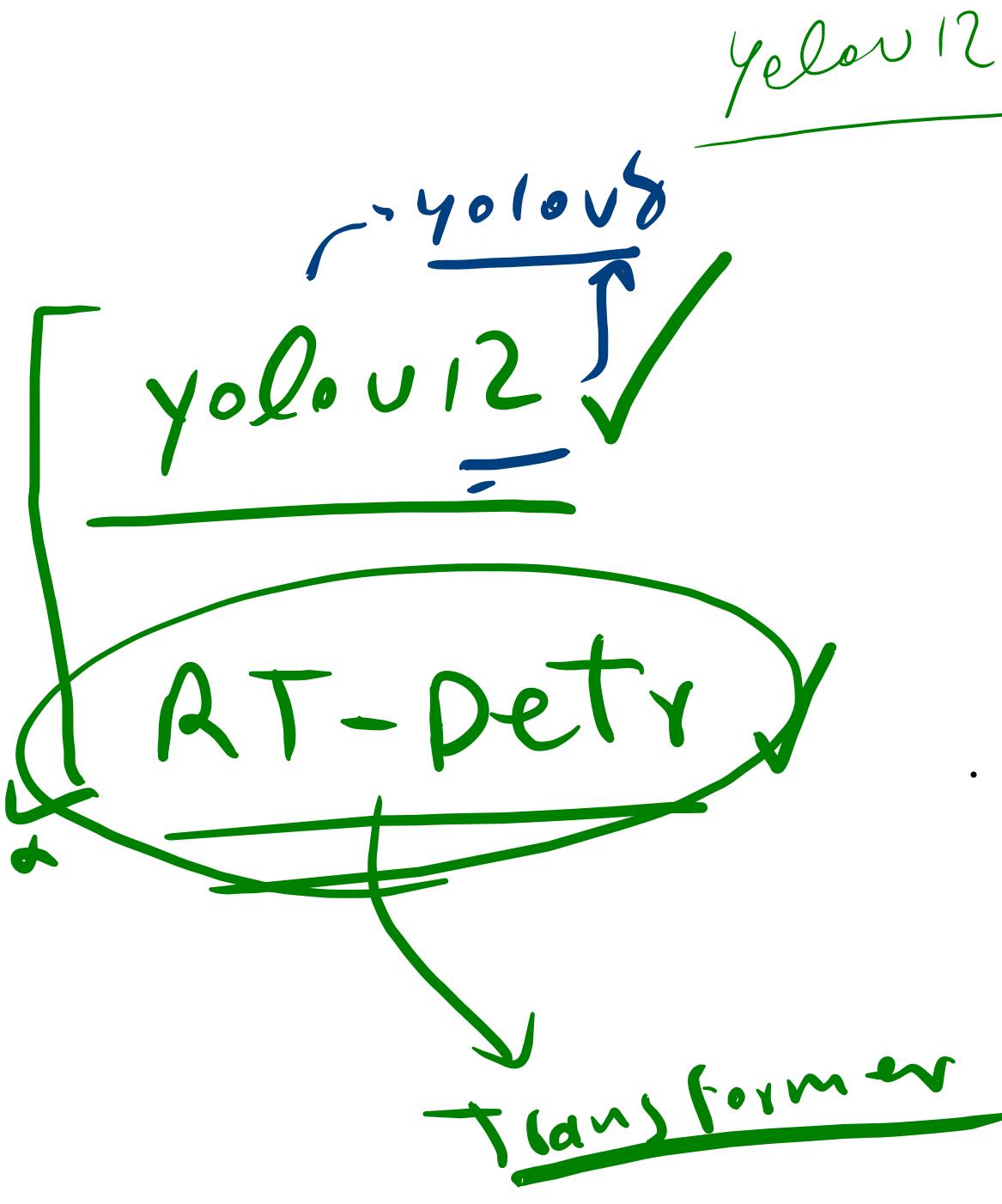
### YOLOv12: Attention-Centric Real-Time Object Detectors

Yunjie Tian  
*University at Buffalo*  
yunjeti@buffalo.edu

Qixiang Ye  
*University of Chinese Academy of Sciences*  
qxye@ucas.ac.cn

David Doermann  
*University at Buffalo*  
doermann@buffalo.edu

<https://arxiv.org/pdf/2502.12524>



## مدل های منتخب برای تشخیص اشیا

X

✓ 1. RT-DETR

✓ 2. YOLOv12

✓ 3. YOLOv12 + RT-DETR

Coarse detection

Fine Detection



5 4 6 0 7 0 2 ✓

## شروع کار با مدل Yolov12

جلسه بعد!

