

TIANCHI 天池

# 广东工业智造大数据创新大赛

数据引领 飞粤云端

都都都都都都

Defects Detection

主办单位

广东省人民政府  
阿里巴巴集团

承办单位

广东省工业和信息化厅  
佛山市人民政府  
阿里云计算有限公司

执行承办

佛山市南海区人民政府  
佛山市经济和信息化局



## CONTENTS

01

**Team Member**

02

**Problem Analysis**

03

**Innovations**

04

**Implementation Scheme**

05

**Extensibility & Conclusion**

## 01 Team Member

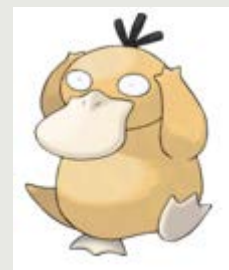
---



李浩宇  
华南理工大学



姚粤汉  
华南理工大学



朱焕升  
华南理工大学

博学慎思，明辨笃行



## 02 Problem Analysis

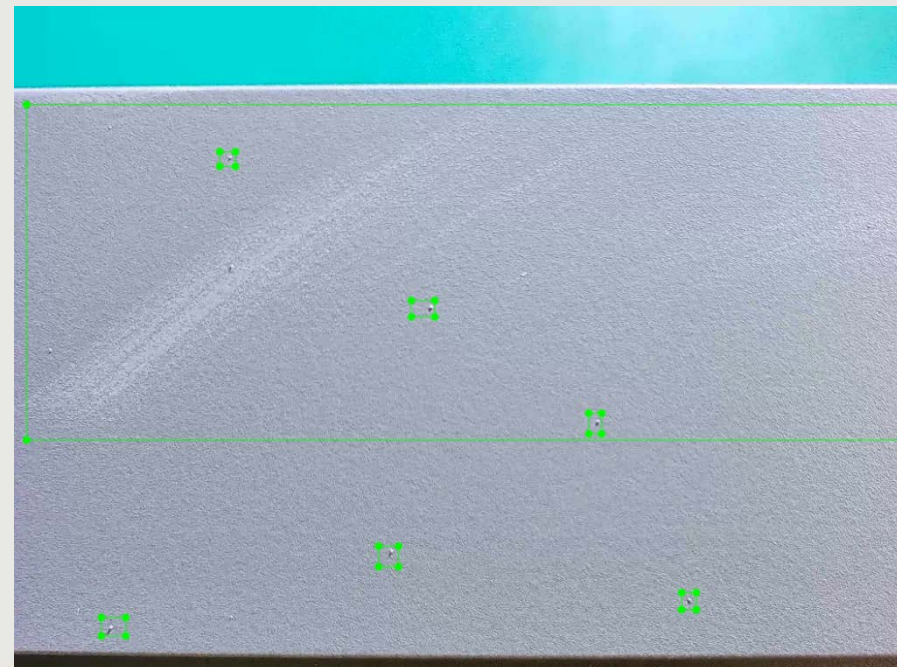
### ■ Summary

Given a picture of an aluminum sheet, detecting the defects of interest.

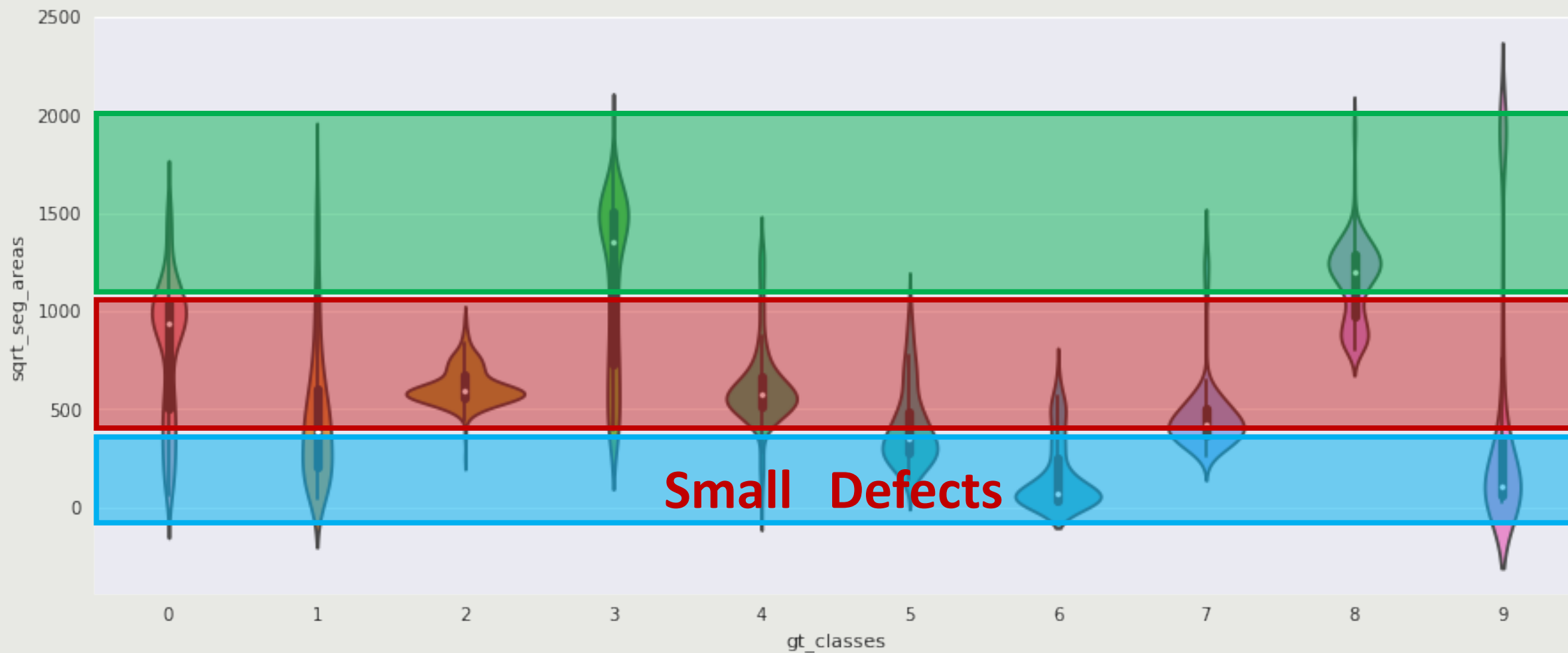
Detect = Locate + Classify

Metrics: mAP@IOU0.5

Difficulties: small defect、extreme aspect ratio

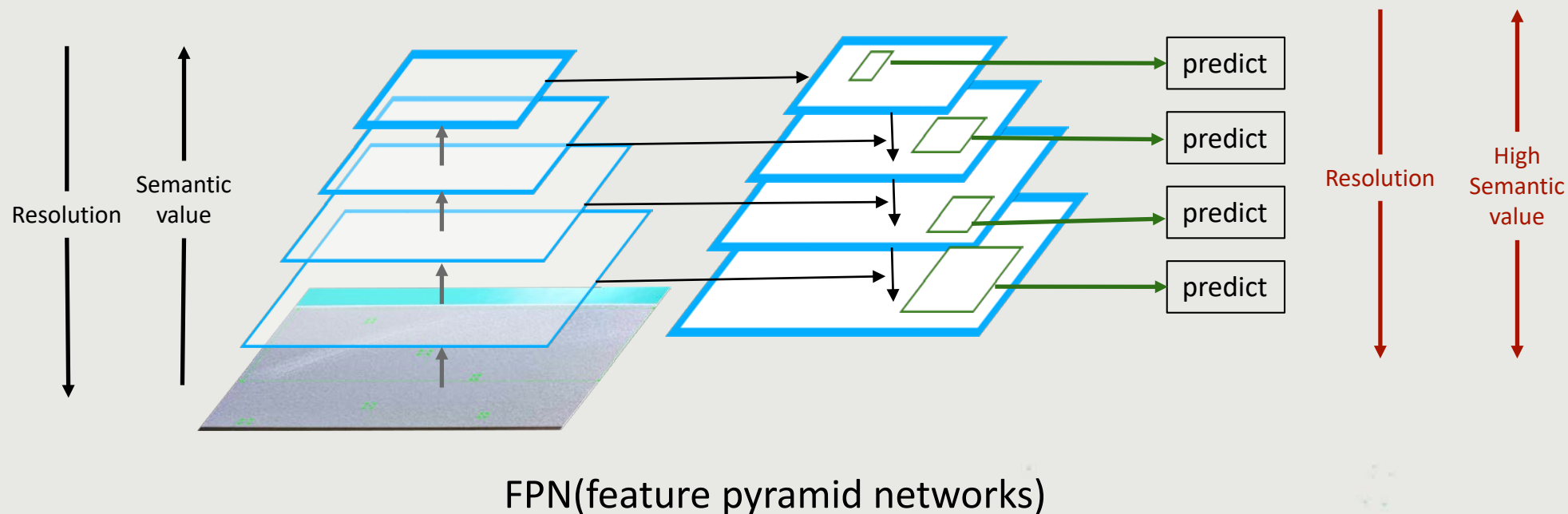


## 02 Problem Analysis



## 02 Problem Analysis

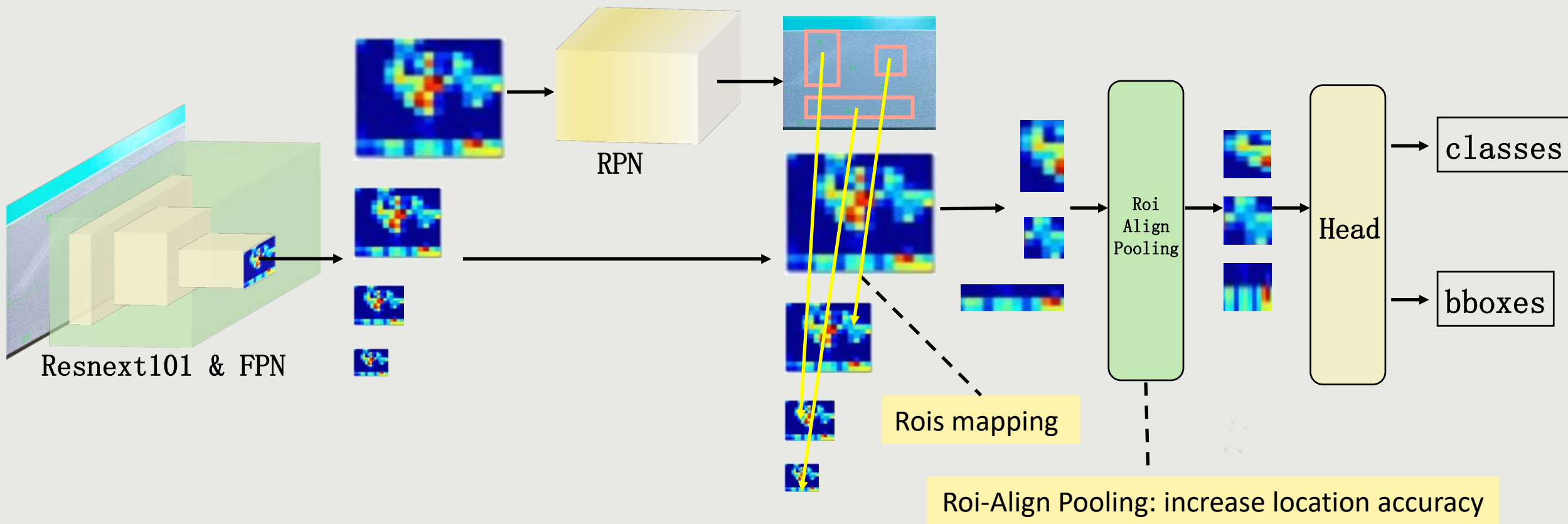
### ■ Feature Pyramid Networks for feature extraction



Lin T Y , Dollar P , Girshick R , et al. Feature Pyramid Networks for Object Detection[C]// CVPR. IEEE Computer Society, 2017.

## 02 Problem Analysis

### ■ Baseline: Faster R-CNN with FPN



Lin T Y , Dollar P , Girshick R , et al. Feature Pyramid Networks for Object Detection[C]// CVPR. IEEE Computer Society, 2017.

He K , Gkioxari G , Dollár, Piotr, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP(99):1-1.



## 02 Problem Analysis

### ■ Training Details

Data Augmentation:

Symmetric images → **Random horizontal/vertical flip**

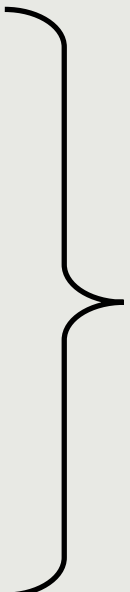
Different camera's parameters and lighting environment → **ColorJitter** Augmentation

Multi-scales Training:

More robust to defects' scale —— **scaling invariance**

Transfer Learning:

Pretrained on Coco dataset and then finetune on our dataset.



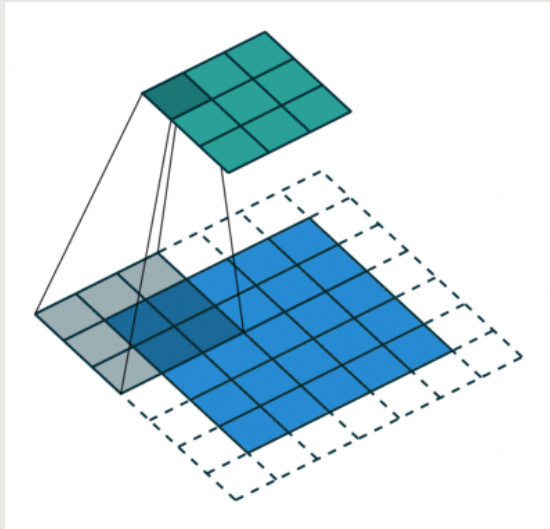
77.80  
↓ + 3.08  
80.88



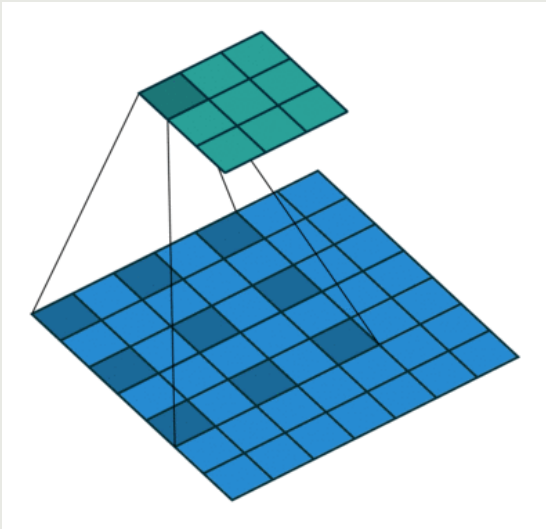
# 03 Innovations

■ **Problem:** Original receptive field (1027) is insufficient for high resolution input (800\*1066)

Solution : **Dilated convolution**



standard convolution



dilated convolution

**Pros:**

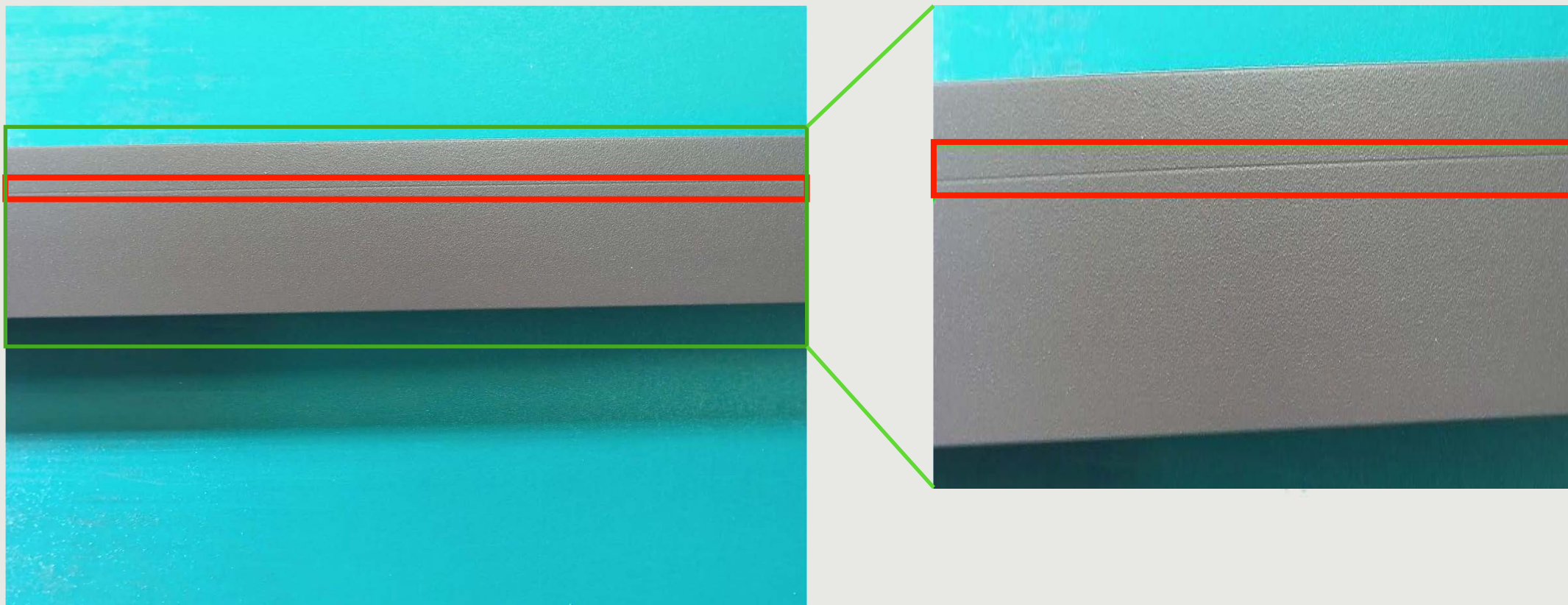
Cost-free (got larger receptive field without more compute )

Method	w.o. Dilated Conv	Dilated Conv	△
FPN	80.88	82.38	1.50



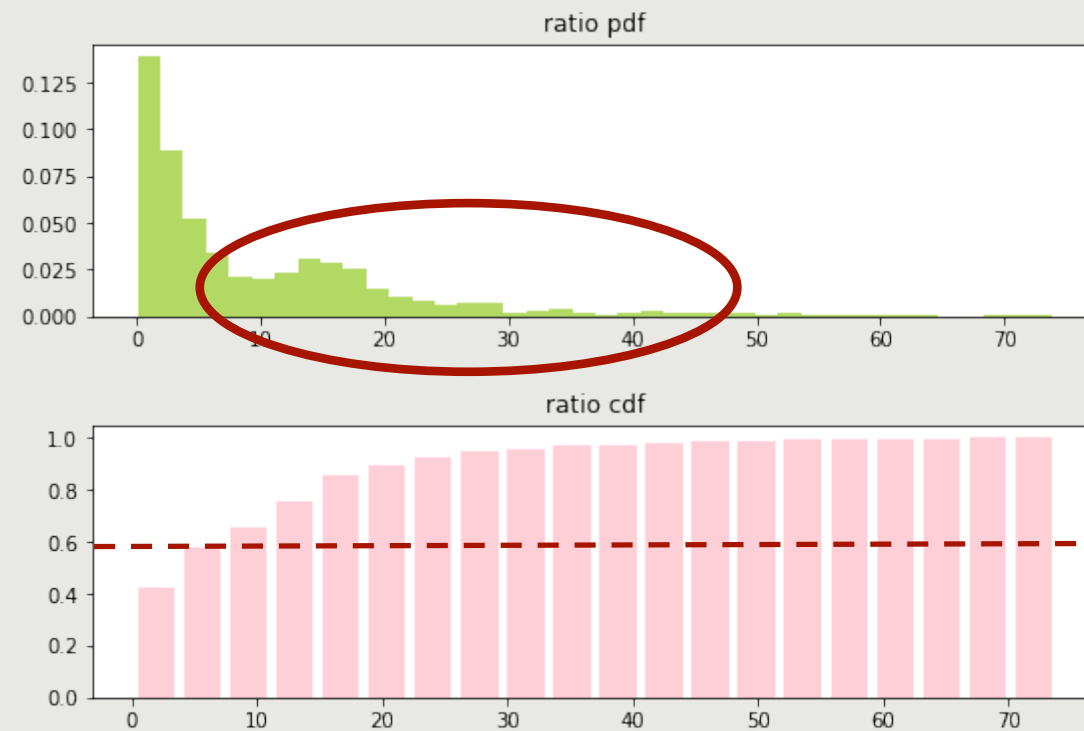
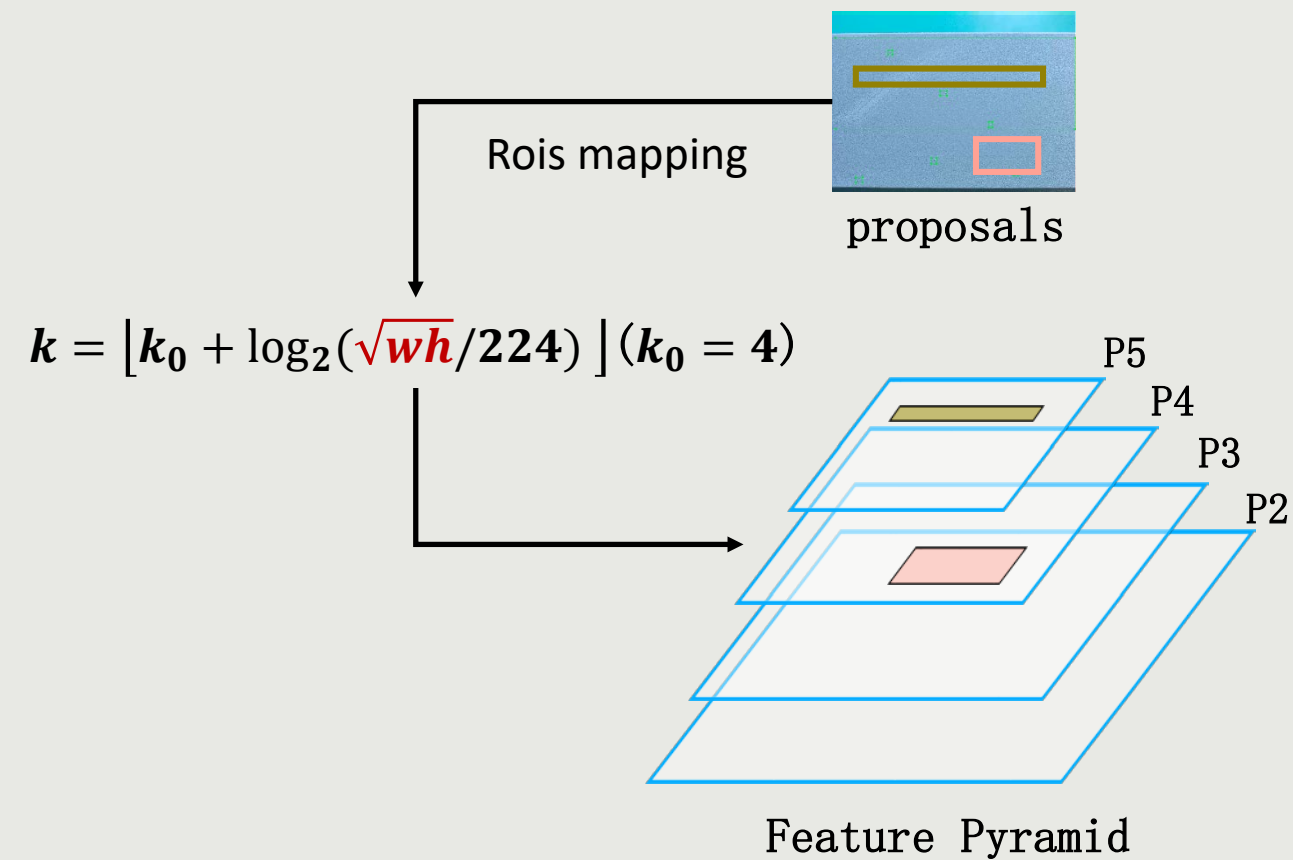
## 03 Innovations

- **Problem:** Extreme aspect ratio



## 03 Innovations

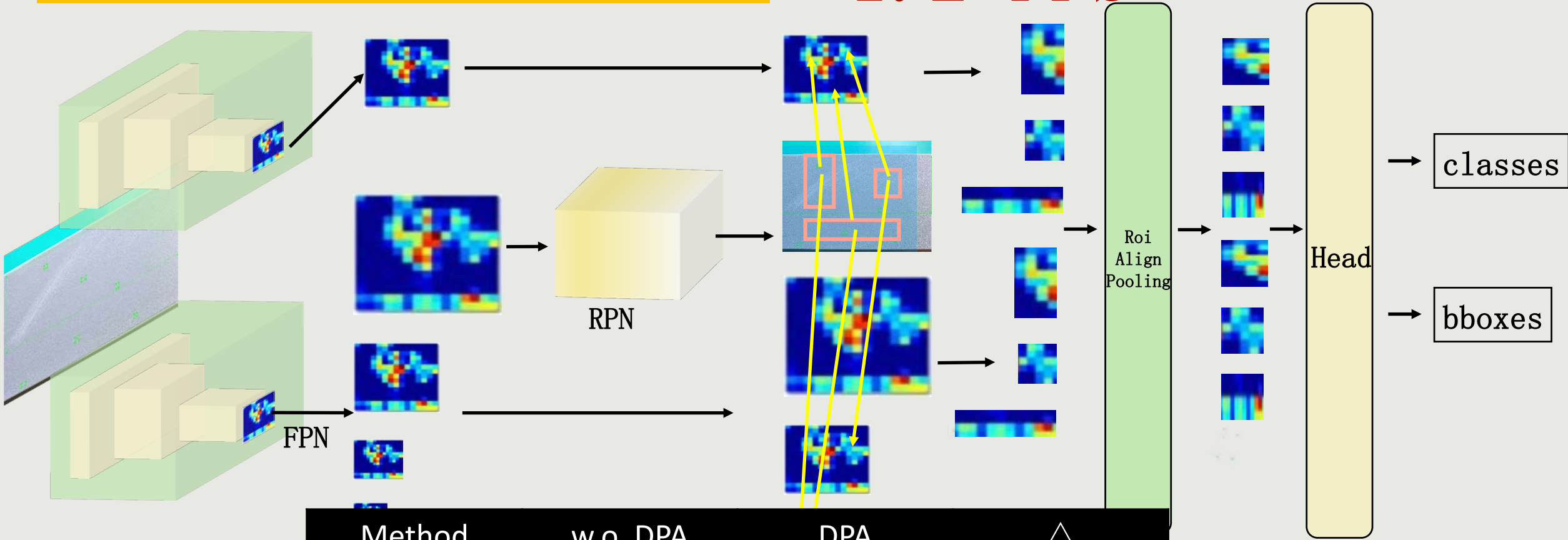
- **Problem:** Non-optimal Rois mapping due to extreme aspect ratio



# 03 Innovations

## ■ S1 : Dual Path Aggregation for Object Detection

1.2 FPS



Method	w.o. DPA	DPA	△
FPN	82.38	83.72	1.34

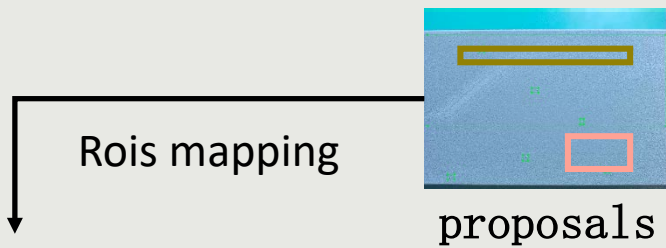


## 03 Innovations

## ■ S2 : Multi-levels Roi-Align Pooling (MLRP)

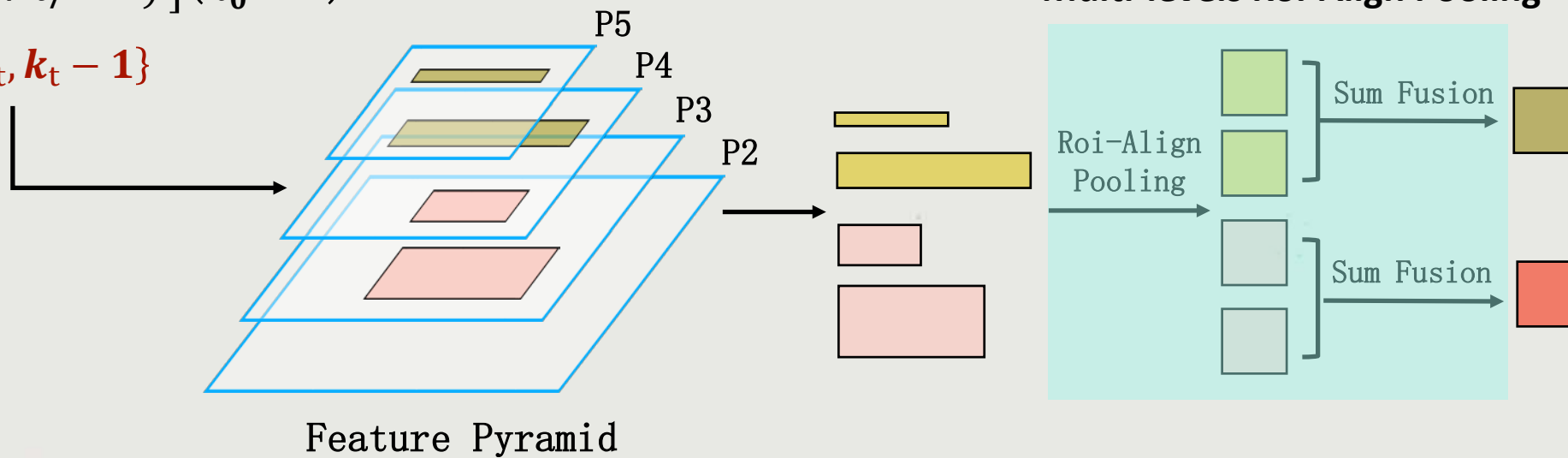
Motivation: unsuitable Rois mapping due to extreme aspect ratio

## Solution:



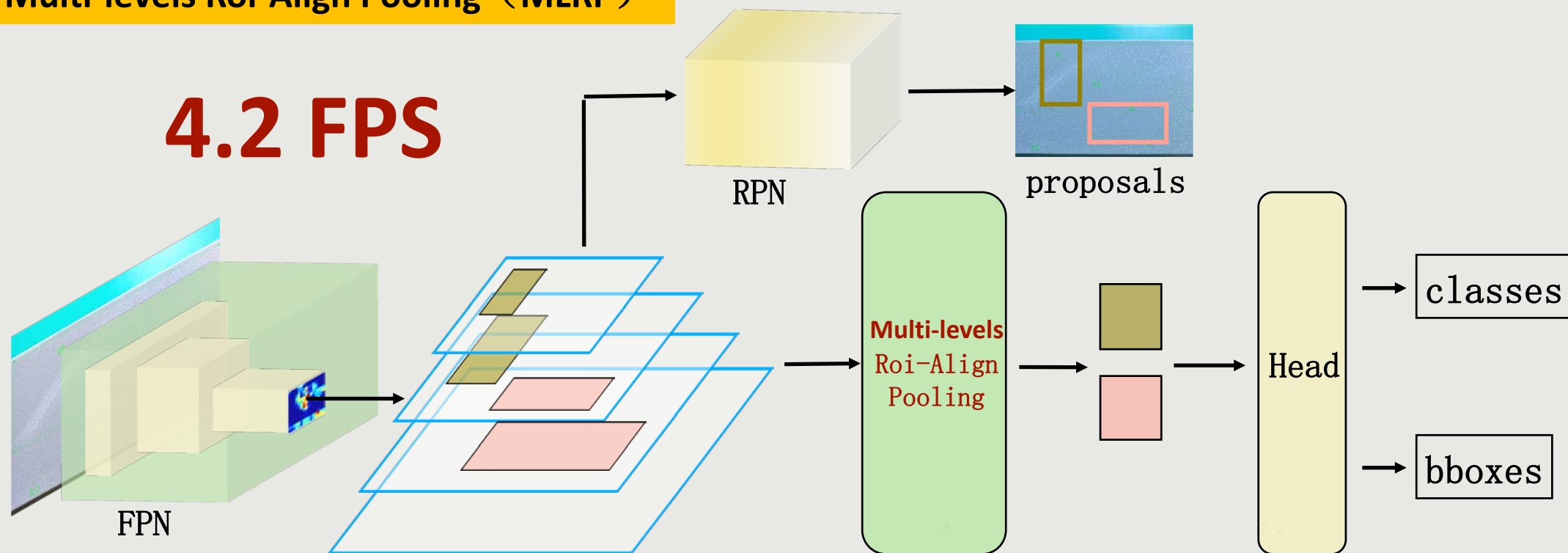
$$k_t = \lfloor k_0 + \log_2(\sqrt{wh}/224) \rfloor (k_0 = 4)$$

$$\mathbf{k} = \{\mathbf{k}_t, \mathbf{k}_t - \mathbf{1}\}$$



## 03 Innovations

### ■ S2 : Multi-levels Roi-Align Pooling (MLRP)



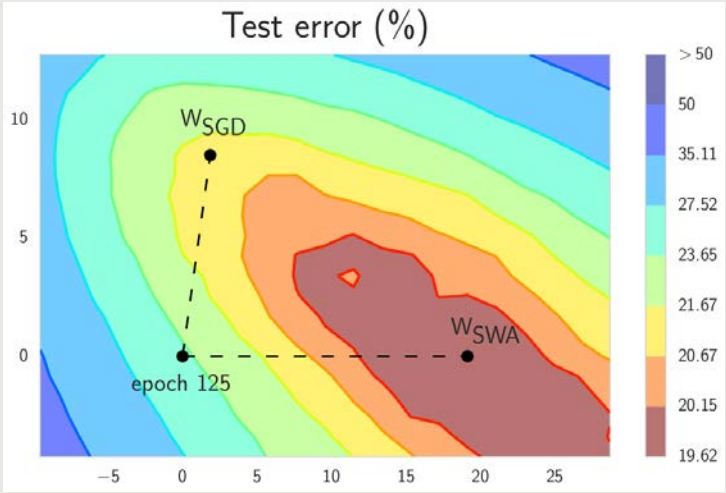
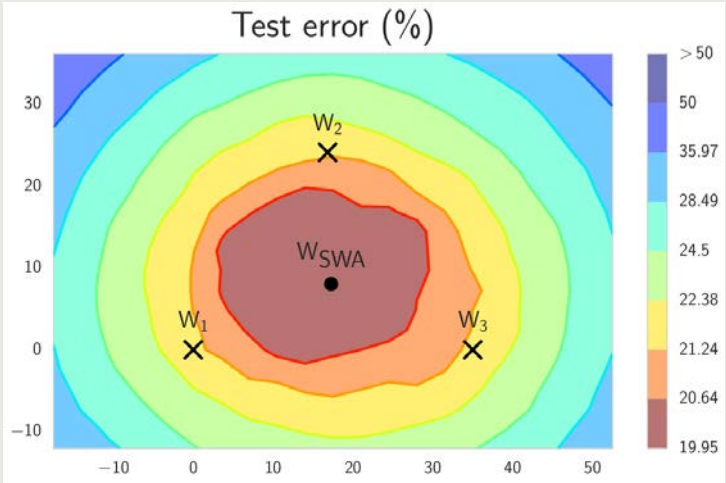
Method	w.o. MLRP	MLRP	$\Delta$
FPN	82.38	82.91	<b>0.53</b>

# 03 Innovations

## ■ Stochastic Weight Averaging (SWA)

- 1. Averaging Weights Leads to **Wider Optima and Better Generalization**
- 2. **Cost-free** Ensembling: Ensembles in the Weight Space, generating ensemble models **without any additional inference or training time.**

Method	w.o. SWA	SWA	$\Delta$
DPA	83.72	83.85	<b>0.13</b>
MLRP	82.91	83.02	<b>0.11</b>



Izmailov P, Podoprikin D, Garipov T, et al. Averaging Weights Leads to Wider Optima and Better Generalization[J]. 2018.

# 03 Innovations

## ■ Stochastic Weight Averaging (SWA)

Method	w.o. SWA	SWA	$\Delta$	FPS
S1: DPA	83.72	83.85	0.13	1.2
S2: MLRP	82.91	83.02	0.11	4.2

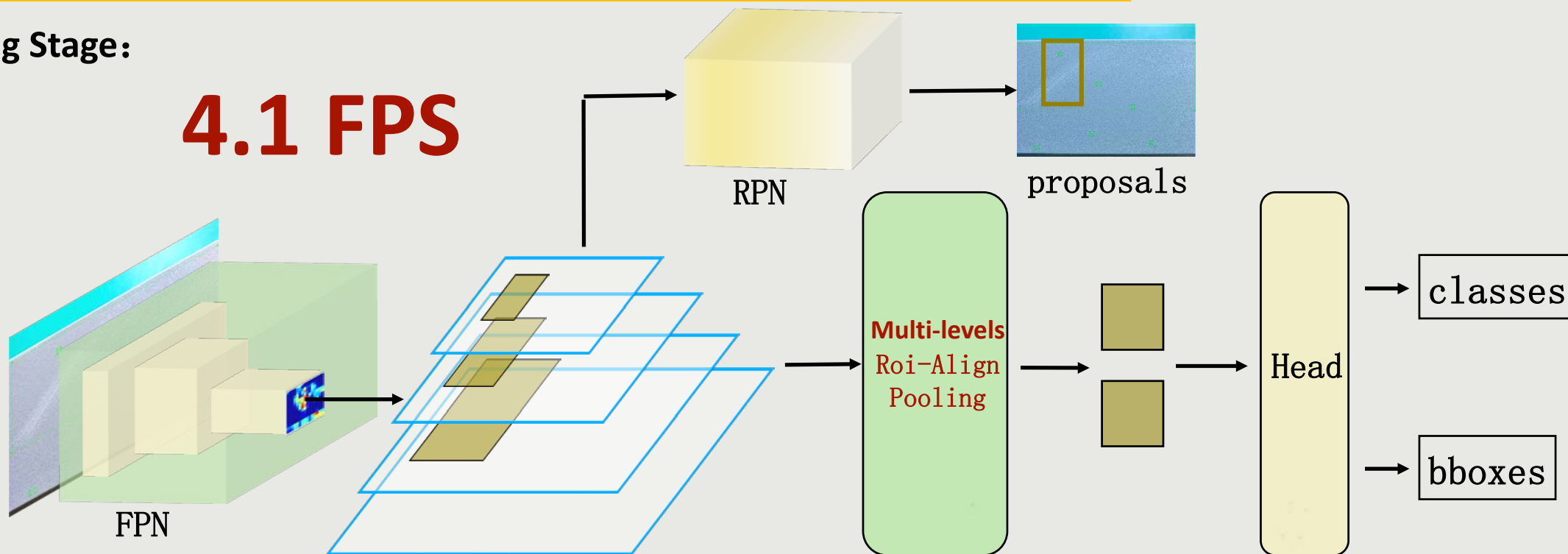
	B榜数据集	
团队名称	mAP分数	计算耗时（秒）
Are you OK?	0.8172	2838
BOOMBOOM	0.8055	1744
都都都都都都	0.7927	823
打怪升级	0.7897	7111
shuzhilian_ai	0.7897	6689
风不动	0.7886	13165



## 03 Innovations

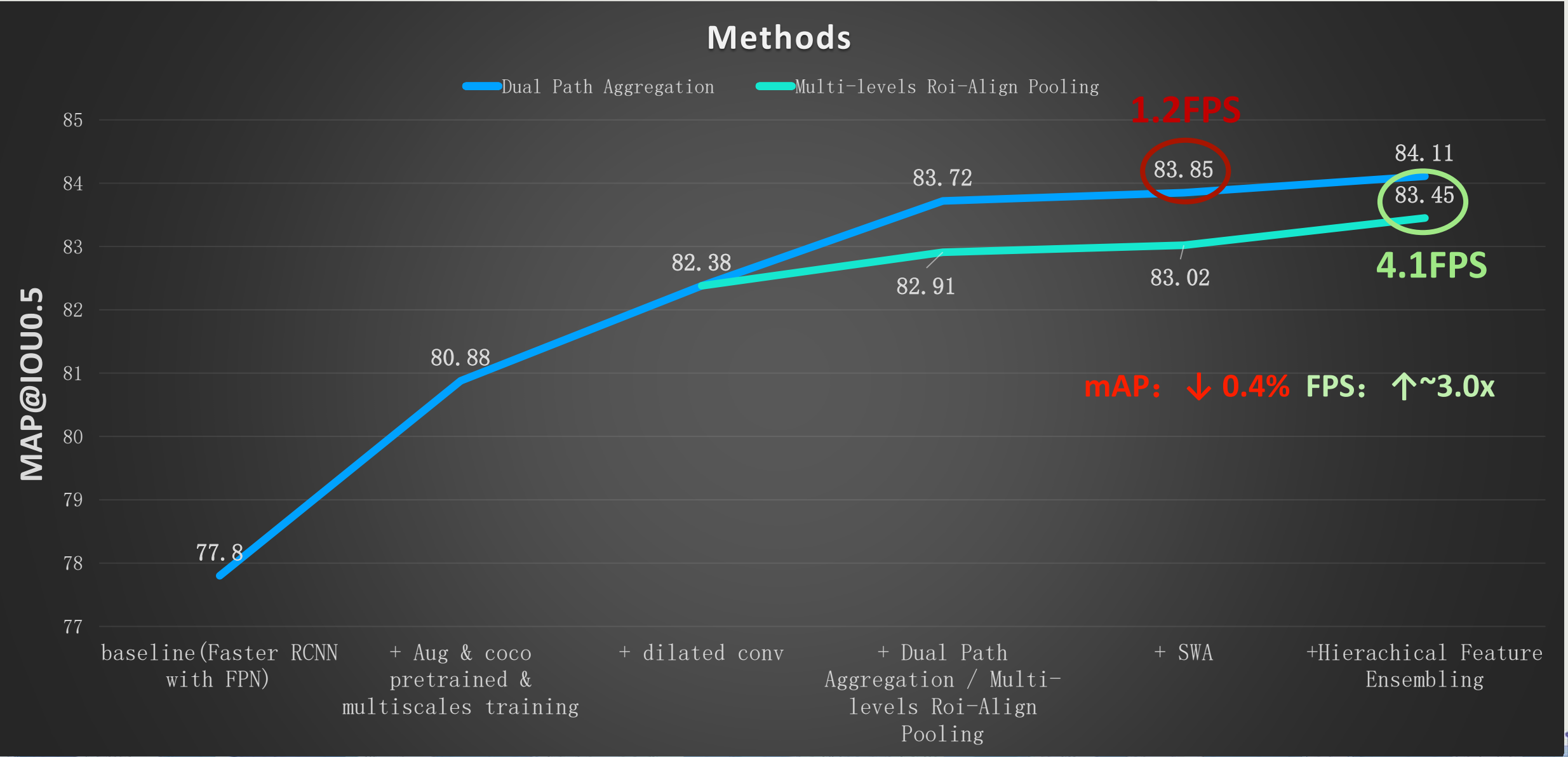
### ■ Hierarchical Feature Ensembling (HFE) : Infer 1, got better for free !

Testing Stage:



Method	w.o. HFE	HFE	$\Delta$
S1: DPA	83.85	84.11	<b>0.26</b>
S2: MLRP	83.02	83.45	<b>0.43</b>

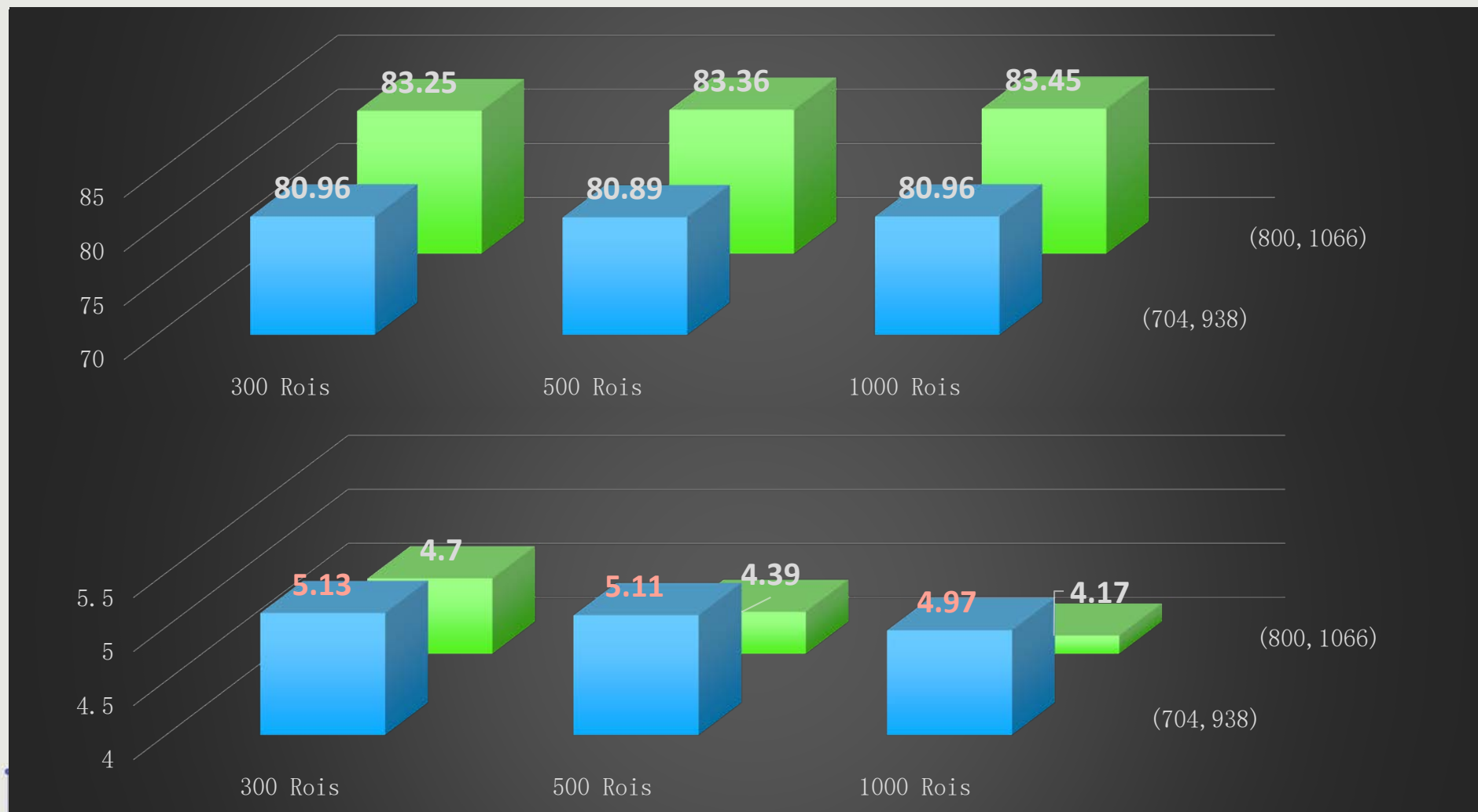
# 03 Innovations



## 04 Implementation Scheme

### ■ Scheme customization without retraining

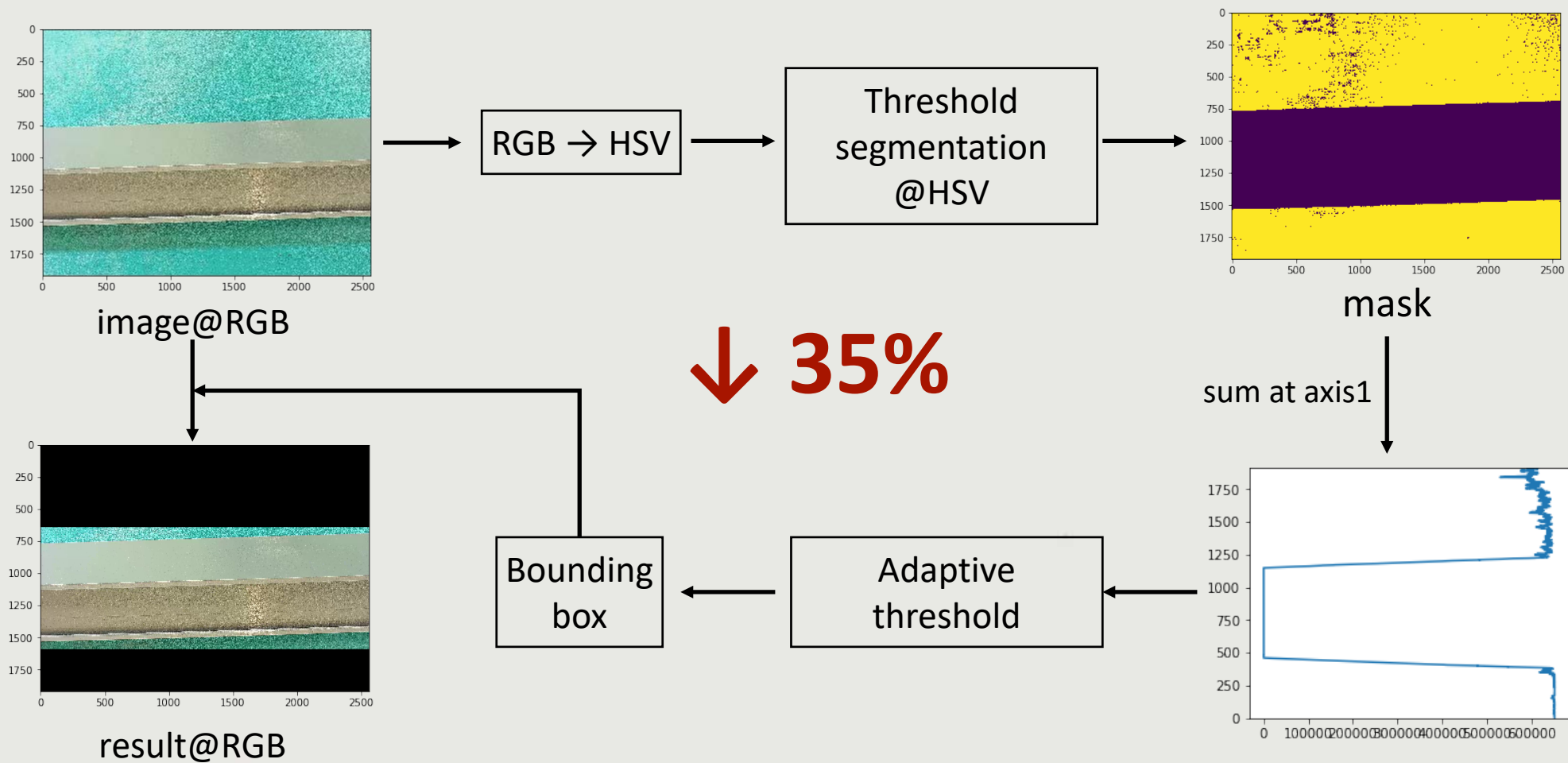
mAP



FPS

## 04 Implementation Scheme

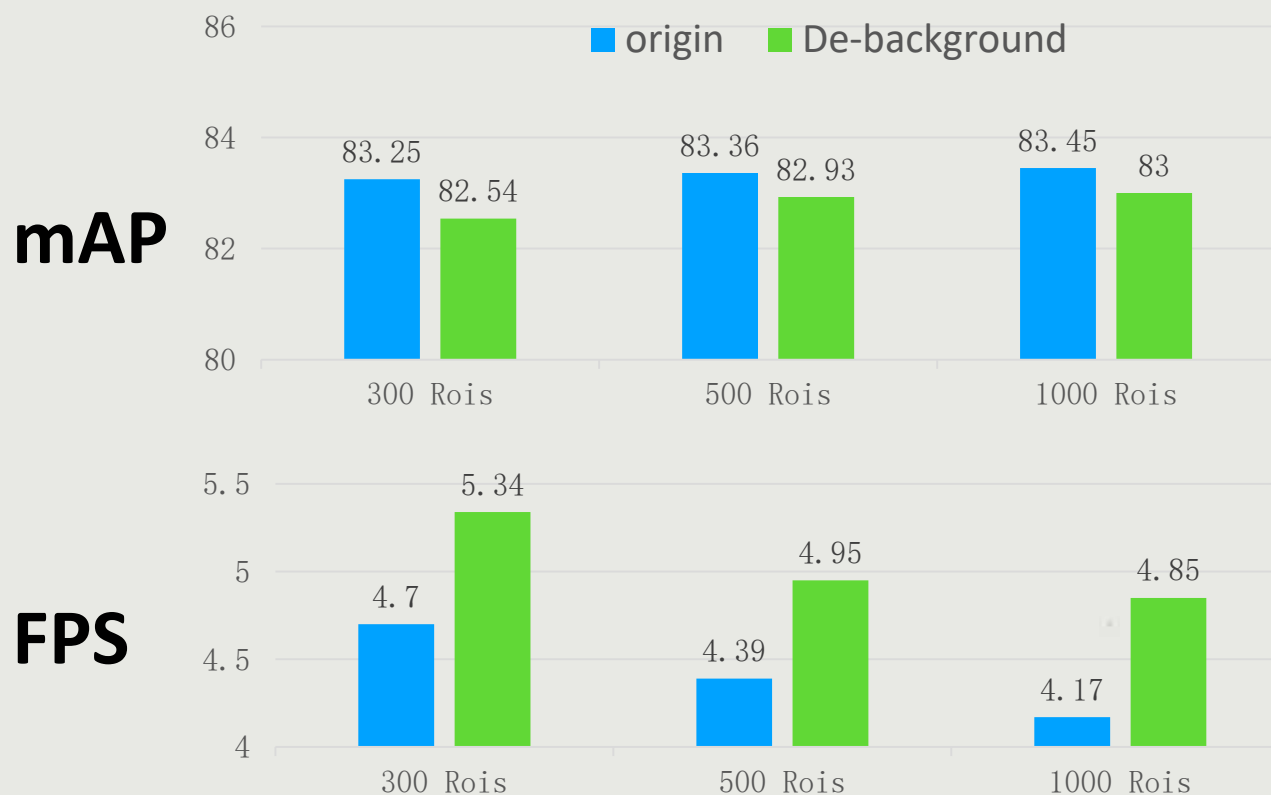
### ■ Run faster on this dataset —— De-background





## 04 Implementation Scheme

### ■ Run faster on this dataset —— Data preprocessing



FPN + MLRP + HEF + DB@(800, 1066)

## 05 Extensibility & Conclusion

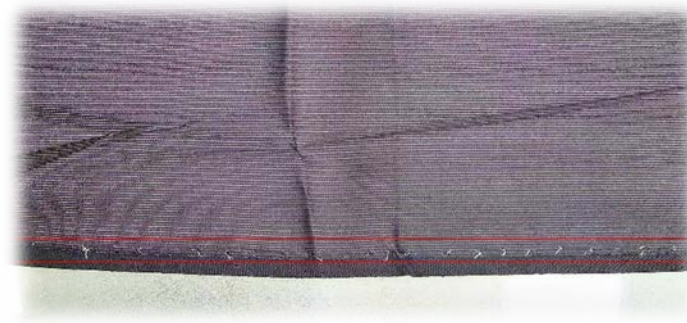
### Dual Path Aggregation for Object Detection

- ❑ Working well in complex scenario with small objects
- ❑ E.g. birds/traffic light/aircraft... detection



### Multi-levels Roi-Align Pooling

- ❑ Face to detecting objects with extreme aspect ratio
- ❑ E.g. Fabric flaw detection



### Hierarchical Feature Ensembling

- ❑ An **effective** method to improve FPN's performance.
- ❑ E.g. object detection / semantic segmentation tasks



## 05 Extensibility & Conclusion

---

- Our reported solution is at least **twice** as fast as other teams.
- **Without bells and whistles**, we got ranking 3<sup>rd</sup>.
- The proposed methods has **high extensibility** in object detection and semantic segmentation.
  - DPA: Dual Path Aggregation for Object Detection
  - MLRP: Multi-levels Roi-Align Pooling
  - HEF: Hierarchical Feature Ensembling
  - De-background
- **FPN + MLRP + HEF + DB : Faster(5 fps) & Cost-effective**